Wind Energ. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/wes-2021-107-RC2, 2021

Comment on wes-2021-107

Anonymous Referee #2

Referee comment on "Non-stationarity in correlation matrices for wind turbine SCADA-data and possible implications for failure detection" by Henrik M. Bette et al., Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2021-107-RC2, 2021

- General comments

The authors introduce an automated methodology to classify SCADA data in different operational regimes by defining wind speeds boundaries based on the observed correlations between variables. While the proposed methodology seems interesting, the benefit of using this method seems to be considered as a given. The paper also contains some statements that should be further investigated and clarified. These are discussed as specific comments.

- Specific comments

- Line 100 – The investigated set of variables is strange if the main goal of the paper is to identify the different operational regimes of the wind turbine. For example, a set of directly coupled variables (rotor and generator speed/line current and active power) are chosen, whereas the pitch angle of the blades is not. The latter is the most significant control variable of the turbine above nominal wind speeds, yet is not considered in the analysis.

- Line 263 – "In cluster 2 this changes and the observables RotorRPM and GeneratorRPM decouple from the others." This should be investigated, as a gearbox with a fixed transmission ratio should be the only component between the two tachometers measuring the rotational speeds. It therefore does not make sense from a mechanical point of view that these two variables can decouple from one another.

- Line 287 – "the response time of the turbine controller is finite." While this is true and while this can play a role during some events, the fact that 30 minute non-overlapping windows are used, surely plays a much larger role than this. Especially since windows are analyzed where no alarms or downtime was present.

- Figure 5 - The fact that no pitch angles are considered for the classification of the operational regimes, seems to result in a significant amount of misclassified points given no post-processing (section 5 is done). Is this resolved by extending the set of considered variables?

- Line 333 – The fact that the estimated rated wind speed seems to be 11.8% higher than the warranted one by the manufacturer, should be justified. This deviation is quite significant, and it does raise the question if these two wind speeds are directly comparable, or if the method intrinsically overestimates the threshold for classification. While the evolution of states over time is shown in multiple figures in the paper, these figures do not inherently convey any information without knowing the wind speed at a given time. In fact, the different control regimes of a wind turbine can normally rather easily be identified by making a three scatter plots, i.e. the active power, the rotor speed, and the pitch angle as a function of the wind speed. It would be good to see that the classification of the data represents what an analyst would manually identify on these plots.

- Figure 10/Line 350 – The direct use of this graph should be clarified when wanting to go towards a (general) classification framework. How can abnormal operating conditions (e.g. deration, curtailments) be taken into account? Do the authors aim to only classify based on wind speed, or should additional post-processing still occur to filter abnormal conditions? How are wind speeds dealt with where there is overlap between two control regimes?

- Line 367 - Our method enables the definition of multiple operational states for wind turbines, whereby an optimization for failure analysis and prediction could be undertaken.

  This seems like an overstatement. The same control regimes are described in the IEC standard, and are thus not 'defined' within this paper, but by the control design of the manufacturer.

- Line 369 – The fact that this pre-processing step might help the performance of "all techniques used for failure analysis" seems to not be a given. It would be good to briefly illustrate with an example that this is indeed the case, as the fact that different operational regimes exist for wind turbines is well known. The so-called boundary speeds v1, v2, and vnom are quite easily identifiable by an analyst (see question 5). This raises the question if their identification was indeed a bottleneck for the proposed analysis strategy, or if the benefit is simply not considerable. In any case, automated classification of control regimes is still interesting, but the paper clearly aims at this specific use case, which should therefore be illustrated.
- Regarding the "implications for failure detection" in the title of this paper: I agree that if the classification could be shown to have an added value for the failure detection, that this would be a novelty, but currently there is nothing in the paper that actually decisively shows that this classification improves failure detection. Without any proof that it actually helps the failure detection step, this sounds more like speculation than an implication. Can you provide proof, eg an example or reference, where the proposed methodology is used to augment the failure detection step?