Wind Energ. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/wes-2021-76-RC2, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on wes-2021-76

Imad Abdallah (Referee)

Referee comment on "Data-driven farm-wide fatigue estimation on jacket-foundation OWTs for multiple SHM setups" by Francisco d N Santos et al., Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2021-76-RC2, 2021

Summary:

The article proposes a two-tier data-driven model to predict the 10-min fore-aft damage equivalent fatigue loads on the jacket foundation of offshore wind turbines. Data utilized in this research correspond to SCADA, estimated thrust load and a "low quality, i.e. 1Hz sampled" nacelle-based accelerations signals. The motivation behind this research is that strain gauge instrumentation for direct measurements of fatigue loads on OWT jackets is not economical, and a substitute surrogate data-driven model should be developed.

The topic is very relevant. The paper is well written.

Comments:

At the top of Figure 4, I read the expression "Measured thrust load". I would not agree that thrust is strictly measured, it is rather estimated from the strain measurements.

At the top of Figure 4, I read that you provide a validation set for the thrust load. This indicates that the thrust has been validated independently somehow. Might you please elaborate on this point? Were validation performed via an aeroelastic simulation? In order to get a good estimate of thrust, the strain gauges need to be calibrated for a know load

level. How was this done, in order for one to trust that the Thrust estimates are correct?

On page 9: I suggest that authors explain the basis by which they selected the features of the time series.

Section 2.2.2: another powerful feature selection approach is https://pypi.org/project/BorutaShap/

Page 12 line 289-290: "carefully selected as to be representative of all operating conditions". What are the criteria for such a selection? Turbine operating in partial load, full load and transition? Turbine operating at max Cp and rated power? Operating conditions covering all possible combination of pitch-tipSpeedRatio? Please explain what are all the representative operating conditions.

Page 12 line 292: why do you correct the thrust for the air density? Are you calculating the thrust coefficient?

Page 13 line 299: ANN as used in this article cannot extrapolate, i.e. they cannot make correct predictions when the input are not within the range of the training set. How do the authors ensure that the validation data (3 months of data outside of the training period) fall within the range of the training set?

Page 13 line 300: the cross-validation applied to a different turbine. Has the cross-validation set been chosen in such a way to reflect the conditions that occurred in the original training set? i.e. you cannot cross-validate on a set where the other turbine is known to be in a waked condition, right?

Table 1: you want to consider adding this to the mix of methods: https://pypi.org/project/BorutaShap/

Page 15: please note that random forest feature importance can be derived either based on mean decrease in impurity or based on feature permutation. Both methods could give slightly different results. Mind you that the results could also be affected by dependent input features. Please make sure you do not bias your feature selection based on the above. Kendall's Tau takes such dependence into account, and probably why it gives the lowest MAE in figure 7 at the cost of high number of features

Section 3.2.2: Table 2: once 1-second SCADA data is made available, by default then the

10min-SCADA can be calculated. Did you consider a scenario where both 10min and 1Hz SCADA are made available to the model?

Section 3.2.2: it is perfectly acceptable to propose several scenarios of various data sources, and check their effect on the model predictive error. However, wouldn't a more principled approach involving sensitivity analysis and sensors selection optimization with value of information be more adequate? Please discuss.

Page 20, line 472: please elaborate on the 80-20 train-test split. Do you respect the temporal evolution of the data or do you perform randomized split?

Page 21, line 489: "…for a different turbine". Please specify where this other turbine located with respect to the reference turbine where the model was trained.

Page 22, line 505-508: this might the case, but the question is whether the training set of the reference turbine included any data corresponding to wake? Indeed not all wakes/partial wakes/multiple wakes are created equal because of dependence on atmospheric stability, turbulence and shear. It is worth discussing this issue. The logical consequence of this is that your training set for the reference turbine should include a much larger amount of data in order to take into account the various wake effects in order to generalize to another turbine…

Page 22, line 505-508: 8-18 m/s according to the reference turbine where the model was trained or according to the different (second) turbine? If the second turbine were wake affected, then its wind speed should account for the velocity deficit.

Page 22, line 505-508: Avendaño-Valencia, L. D., Abdallah, I., and Chatzi, E.

Figure 16: what is mean by mean_DEL? DEL is a short term measure of fatigue conditional on wind speed. Is the mean calculated by weighting according to the pdf of the wind speed? Please elaborate how the mean is computed in this case.

Figure 18: same comment as above.

Figure 16: DEL across various wind turbines in the farm will be highly influenced by the mean wind speed (at each wind turbine), turbulence and shear. Maybe you would want to plot Figure 16 for various wind speed bins (e.g. 6-9m/s, 10-14m/s and >15m/s). You will notice that the DEL across wake-free and wake affected wind turbines will change quite a

bit.

General comment:

There are multiple grammatical and orthographic mistakes, and a proof-read is necessary.

It is not mandatory, but you might want to consider comparing the performance of your method to other state of the art methods dealing with a similar subject matter.

Methodological Suggestion:

A more direct approach avoiding the two-tier approach proposed in this article might be to use variational auto-encoders neural networks such as proposed here:

https://onlinelibrary.wiley.com/doi/full/10.1002/we.2621

https://link.springer.com/chapter/10.1007%2F978-3-030-12075-7_21