

Wind Energ. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/wes-2021-25-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.



Comment on wes-2021-25

Edward Hart (Referee)

Referee comment on "Identification of wind turbine main-shaft torsional loads from high-frequency SCADA (supervisory control and data acquisition) measurements using an inverse-problem approach" by W. Dheelibun Remigius and Anand Natarajan, Wind Energ. Sci. Discuss., <https://doi.org/10.5194/wes-2021-25-RC1>, 2021

General comments

The paper presents a novel method by which a coupled-stiffness value related to wind turbine main-shaft dynamics, along with torsional displacements of the main-shaft, may be estimated from measured data extracted from an operating wind turbine. A benefit of the method is that all analysis is carried out using measured data, without the need for proprietary information regarding turbine design or structural dynamics. Outputs of the methodology allow main-shaft damage equivalent loads to be estimated from relatively standard turbine outputs. A validation of the method is undertaken which includes simulated and real world data. The results seem to indicate good performance of the developed methods, with errors for estimated quantities all lying between 4 and 12%. The work undertaken herein therefore appears to provide a valuable new capability and, as such, is appropriate for publication in WES. However, the manuscript itself requires significant revision before it can be considered suitable for publication. A brief overview of the main concerns will be outlined here, with more specific details listed under "Specific comments", below.

Structure, clarity and readability – in its current form the paper is haphazard and difficult to follow. The introduction jumps between background information and descriptions of the methodology in the current work. The structure and flow of the paper overall is not clear, with the reader often left wondering where a given section is headed, or having to try and work out whether a certain concept is being newly introduced or if the author is referring to something already discussed. The paper needs to be reorganised such that introduction, background, methodology and results are all clear and distinct from one another – with the overall structure and flow of the paper made clearer from the start, and better signposted on the way through.

Modelling and analysis presentation and associated discussions – Descriptions of the modelling and analysis undertaken in this paper is currently lacking in clarity, completeness and context. The lumped parameter model is introduced without discussion of the approximations being made and the potential impacts this could have on model accuracy. Inputs to the model are not described in sufficient detail, e.g. frequency requirements for input SCADA data applied throughout this work, and the effect data frequency may have on the presented method, are not properly discussed. Note, when SCADA data is mentioned, one normally assumes 10 min mean values are being referred to, whereas high frequency SCADA (50Hz) is used here for at least part of the analysis. This is therefore a possible source of confusion for readers of the current paper. Also, the stiffness value being approximated is stated to contain contributions from components other than the main-shaft, again the effect of this (good/bad/neutral) is not discussed, the fact is merely stated. The descriptions of the Collage method and Tikhonov regularisation, in their current form, are not appropriate. Much of the mathematics is simply stated without proper definitions for the various terms, or suitable descriptions to ensure the reader can follow what is being said. Notation also needs to be revised, with the same symbols currently representing both scalar and vector quantities. While I applaud a proper treatment of the mathematics, this is only useful if the formulations can be followed and reproduced by the reader. This requires careful definitions for all terms and better descriptions of how the resulting algorithms are implemented. I suggest that the authors consider providing a shorter description of each methodology, including why it is necessary and roughly how it works (e.g. provide an analogy), along with the resulting equation to be solved (e.g. Equation 6 for the collage method) and the solution method used. Then, the full mathematical details can be referred to an appendix where they can be fully and carefully described, without derailing the flow of the overall paper.

Context, motivations and claims – While undoubtedly useful to the wind energy field, context regarding the wind industry is somewhat lacking from the paper in its current form. For example: is main-shaft fatigue currently of concern, are there available failure rates for this component, has it been much researched in the literature to date? As mentioned above this work seems to assume high frequency SCADA data for its analysis, but that isn't necessarily standard, especially for older turbines. The paper claims that the method is useful for older turbines, but what if they only have 10-min SCADA available? The paper shows certain levels of error for damage equivalent loads estimates, what does that accuracy level mean in terms of usefulness in the field or an ability to predict remaining useful life? Claims are made regarding remain useful life, but the link to that from results in the paper is non-trivial given variability of conditions and potentially incomplete histories of operating data – statistical analysis would likely be required. Note the above points don't necessarily all have to be bottomed out fully, but they should be discussed at some level in this work.

Specific comments

Page 2, lines 30-44 - why list all of these methods if you're not going to discuss them? Beyond stating that they exist its necessary to tell the reader something about them, e.g. pros, cons, what's relevant to the current study, why you used the method you did instead of these.

Page 2, lines 42-43 – “However, to the best of the authors’ knowledge, there is no such study available on estimating the structural response of a wind turbine component from SCADA measurements.” You mentioned previously that many studies exist which consider OMA for wind turbines, do these not use SCADA data at all? Maybe describe them in more detail or list what they do use to make this claim easier to interpret.

Page 2, lines 45-47 - We're now into methodology and haven't even reached the problem formulation yet. This is why splitting into intro, background, methodology seems to make sense. You can then give the summary overview of the types of models to be used in the intro, and then just go straight to the literature in background, that makes the flow feel better.

Page 3, line 60 - Again, this should be in Background - well in the summary of the methods at least, wherever that ends up being.

Page 3, lines 73-74 – “Hence the same site-specific torsional load can also be used to quantify the RUL of gearbox and other drive train components as well”. These are very big claims that need to be demonstrated if they are being made in the paper. Also in your drivetrain model don’t you assume the gearbox is completely stiff? Does that not impact your ability to make conclusions regarding the gearbox? These are some of the reasons why I feel a better discussion of the applied model and its underlying assumptions is necessary.

Page 3, lines 74-76 – Older turbines will likely not have a full service history of 1Hz SCADA data, how is the method applied in that case? Are you proposing that new data can be collected and used to estimate yearly damage for those turbines? That would also need some linking to historical weather records I’d imagine. This is not a trivial claim to make so more discussion is necessary.

Page 3, line 86 – “gearbox is perfectly stiff”. This approximation, among others, needs discussion and perhaps references, is it reasonable to assume this? What are the likely errors from it? Do any other papers deal with this type of approximation and show its effect?

Page 3, line 88 – “The main shaft is modelled by an inertia free viscously damped torsional spring”. Same as for gearbox, the validity and effect of these approximations is necessary to allow the reader to understand how this model relates to the real world.

Page 4, line 98 – “the high frequency gearbox dynamics do not play a significant role in it.” Do you have a reference for this claim? It needs to be demonstrated, referenced or at least discussed.

Page 4, line 99 – “Hence, the two-mass model is sufficient enough to model the shaft torsional dynamics for the wind turbine normal operations as it includes both the low frequency modes” This is a claim, can you back it up with evidence?

Page 4, line 100 – “Also, given the system parameters and rotor and generator torques, the two-mass model is capable of predicting the shaft torsional displacement as close as that of the full-fledged aeroelastic simulation as shown in Fig. 2.” Another un-evidenced claim. Please discuss.

Page 4, line 102 – HAWC2 needs better description, what does this include (BEM model + multibody dynamics? etc.). Not all readers will be familiar with this code.

Page 4, line 104 – “In forward problem approach” You are using lots of terminology that has not been defined, don't assume everyone knows what this means.

Page 4, line 109 - How do you differentiate the theta values? Finite difference can go wrong for 'noisy measurements', do you apply a filter?

Page 5 – see comments in 'General' concerning presentation of mathematical methods in the paper. Too high a burden is placed on the reader here, you need to decide what to keep, what to add and what to move to an Appendix. Currently readers would struggle to recreate these methods from the descriptions provided and the overall paper gets bogged down in technical details which are important but probably best in an Appendix and with improved definitions and descriptions.

Page 6, Equation 6 – is there an issue here? Should the 'u' in the integral operator now be 'theta'?

Page 6, line 137 – “hence it is important to check” This sentence does not necessarily hold. Just because something hasn't been tried it doesn't mean it should. There should be a better justification than that alone.

Page 6, line 150 – “As explained earlier,” by this point the earlier descriptions have all merged with other details, another reason to have a clear, concise description of the methodology somewhere which is easily remembered and referred back to. Remember readers might not be familiar with these types of method so much will be new to them, some of these descriptions feel like they assume quite a high level of familiarity. Also I'm not sure 'time integration' was specifically mentioned earlier in the paper.

Page 7, Figure 3 – It is not clear to the reader why the numerically integrated estimate of theta diverges here. A 'lack of initial conditions plus measurements noise' is given as a reasons, but that doesn't help explain why. Is it because noise leads to greater 'areas under the curve' when integrating? But I'd expect that effect to give both higher and lower values. Please give a more detailed description, or analogy even, that gives the reader a sense for why this happens.

Page 7, Figure 3 – There seems to be a lack of clear definition for what theta is. I assumed it was displacement of the shaft, normally taken in-line with one of the blade roots. But that definition would see theta move from 0 to 2π repeatedly, which is not what is happening for the 'Actual' theta in Figure 3. Now it seems that theta is instead the torsional deflection in a moving frame which rotates with the shaft. The definition needs to be made clear when the model is introduced as otherwise this sort of confusion can arise. Maybe add to Figure 1 to show theta more explicitly there?

Page 7 and 8 – Mathematical presentation again. Please see previous comments. Additionally, in Equation 8 it is not clear what we are minimising over, in Equation 9 theta now denotes a vector (please use new notation when going from scalar to vector and make all definitions explicitly). These same developments include mention of 'fictitious nodes' with no description. Overall the reader is left struggling to understand what is happening and why, with little hope of being able to reproduce what is done here.

Page 8, line 180 – "Since ten-minute SCADA measurements with a sampling frequency of 50 Hz are considered for theta(t) estimation, n becomes 30000." Now SCADA of much higher frequency is mentioned. This is even less common than 1s data. Assumptions and requirements concerning data from the wind turbine needs much clearer description and discussion throughout the paper. Also, if 50Hz is used here, will the method suffer if that's not available? Did you test to see what you can get away with at lower frequency?

Page 8, line 180 – Now that n is becoming large it seems that a discussion of required computing power and computational times is necessary. Does this take a long time to run? How about if you were assessing a lot of data for remaining useful life analysis across a whole wind farm? I don't believe any such discussions are undertaken in the manuscript currently.

Page 9, line 209 – "the inputs". Again, what frequency is this data at?

Page 10, Figure 5 – The flow chart is blurry and appears squashed. Please remake as a clearer vector image. Also I'd include this in a 'methodology' section earlier in the paper to help make the process clearer earlier in the paper.

Page 10, line 216 – "At this point, it is important to realize that the static component of

the displacement does not... and hence only the dynamic component of the torsional displacement" Neither of the 'static' or 'dynamic' component of torsional displacement is defined in the paper and this is the first mention that either of them get.

Page 11, line 219 - "By removing the static displacement from the simulated displacement, the resulting dynamic component can be compared with...". Is that useful? Is static component important for damage in the shaft? Does it relate to the mean about which fluctuations are occurring? None of this is clear and so the highlighted line of text is also unclear regarding what is being compared and how the comparison is to be interpreted.

Page 11, line 228 - "12.06%". This sounds good but how good is it in the context of the problem at hand? How else might it be approximated and if there are other methods how does this compare?

Page 11, line 228 - "At this point, it is important to remember that the estimated K value is a combination of the main shaft and blade edgewise stiffness." What does this imply for the reader? How does this change our interpretation of findings? If other components are present, how large are they relative to shaft K. Can you perform a simple analysis by taking blade edge stiffness values from the HAWC2 model and comparing with a known value of shaft torque to demonstrate that the dominant contribution to K is from the shaft?

Page 11, line 231 - Again, static and dynamic theta values need to be introduced much earlier in the paper, and described without assuming familiarity. Also the method by which static theta is estimated should form part of a methodology section, rather than being introduced in amongst results.

Page 11, line 233 - What about gearbox efficiency? If generator efficiency is taken into account won't gearbox efficiency also have an effect? If not then please outline why, e.g. maybe the efficiency of one is much higher than the other? Otherwise both should be accounted for.

Page 12 - Damage equivalent load calculations: The Wöhler exponent takes on different values, which is used here? Does hollowness of a shaft have an effect to these calculations?

Page 13, line 254 - "This could be due to fact that the pitch control will be active beyond the rated wind speed and hence the instantaneous changes in the pitch angle affects the rotor speed." Isn't this a bit of a gap in the modelling? Wind turbines spend a fair amount of time operating while pitching so this is an important factor. This all ties to the approximations being made when the chosen model is applied. A better discussion of the

model in relation to a real world turbine is needed when the model is presented, and this must include effects such as pitch that are accounted for, along with some indication of the possible impact that will have. E.g. "The proposed model does not account for shaft speed changes related to pitch activity, however, in above rated conditions the controller is acting to maintain a constant rotational speed, and so such changes (and associated errors) are expected to be small..." or something like that.

Page 13 (and onwards) – You start referring to "1Hz DELs" without defining what these are.

Page 13, line 258 – "SCADA measurements of the Vestas V52-850 kW research turbine..."
What frequency of data?

Page 14, lines 278-282 – You make big claims about remaining useful life, but only actually calculate damage equivalent loads. Unless you demonstrate how this link is made (which is non-trivial) you maybe need to make the claims less strong. E.g. 'the proposed methodology may assist in decision making regarding maintenance and remaining useful life' kind of thing. Or, reference to material that shows how damage equivalent load predictions are then made into remaining useful life assessments. But either way, claims need to be tentative unless backed up by references or evidence.

Page 15, line 281 – "It is also feasible to reconstruct the torsional moment time series, which may be used an input to gearbox design tools to predict the loading within the gearbox." This is another big claim, especially considering that you have assumed a perfectly stiff gearbox with 100% efficiency, if included it needs to be demonstrated more fully that this is indeed the case. It is also important to consider what you have shown, which is that damage equivalent loads are well predicted by your method, that is not necessarily the same as saying that the time-series loads are reconstructed to a similar level of accuracy. They may be, but that needs discussion and consideration and doesn't follow immediately from the damage estimate results.

Page 15, lines 297-298 – "As the proposed methodology does not require any design basis, it can be used for any older turbines for the estimation of the main shaft torsional RUL." This is not necessarily true. As mentioned, older turbines will at best tend to have 10min averaged SCADA data available. What does this mean for your methodology if that's the case? If high-frequency SCADA is always required then that needs to be made much clearer throughout the paper, and probably indicates that the term 'high-frequency SCADA' should be used in the paper title, since most people assumed SCADA = 10min averages. The methodology is still potentially relevant to older turbines since new data can be collected to link fatigue damage to its operating conditions, but life estimates then require some sort of hindcasting or statistical analysis that goes beyond what is covered here. Again, please consider which claims are reasonable to make based on what is demonstrated in the paper, if reaching beyond what is presented here then please add more discussion and references which evidence the claims being made.