Referee Comments for wes-2021-25

Abstract

The manuscript presents novel application of collage method with an integration technique in an inverse scheme to identify the simplified drivetrain model and its torsional deformation for identification of wind turbine main shaft loads. The collage method is used to identify system parameters, in which the torsional displacement is initially required. As such, an inverse method for the time integration based on Tikhonov regularisation is applied to retrieve the torsional displacement of drivetrain form the measured angular velocity of rotor and generator. The novelty and impact of this work is centred on the use of SCADA data only and that no knowledge on the drivetrain model is required. Sufficient amount of results using both simulated and actual measurement data have been presented. The manuscript generally investigates an interesting research question, which is appealing to both academic and applied research areas in wind turbine science and technology.

General comments

The developed methodology and the way it has been implemented within the paper are well explained. However, some key questions on the theoretical basics and practical applicability of the presented method may arise. As such, further clarification, including explanation of the limitation of the method or adjustment of the paper's arguments seem to be required.

The first question concerns the simplified mathematical model of the drivetrain (Eqs 1-3). This model is a compact two-disk representation, in which the inertia of the main shaft and that of the gear box and high-speed shaft are collected into the equivalent generator inertia. Therefore for the sake of consistency of the model, the stiffness term (K) seems to stand for the collective stiffness of the whole drivetrain rather than solely that of the main shaft [1]. However, the manuscript is taking whole drivetrain apart from the main shaft perfectly rigid, so above stiffness is assigned to the main shaft and will then be identified and utilised it to infer the main shaft loads, which may not be right for drivetrains with multi-stage gearbox and shafts. There are several references available (e.g. [2]) that explain the equivalent drivetrain models as the combination of springs in series. The angular displacement of such models are basically the summation of the individual displacements of the drivetrain components. A comparison with simulation results cannot simply validate the above assumption, as the turbine simulation tools are also usually including a simplified drivetrain model, as they aim to capture the global motion and loads of system. In case that the contribution of the other components are known/found to be negligible, this has to be quantitatively demonstrated or at least be discussed adequately. It is not well described whether this work is identifying the main shaft load or the total equivalent torsional load of the drivetrain. Nevertheless, the manuscript's methodology and findings are still interesting. As such, it is recommended to either clarify this point or discuss the probable limitations.

The other point is regarding the manuscript's argument that it just needs SCADA data, which sounds compelling due to this data being readily available. The suitable sampling rate of the SCADA data that is required for this methodology has not been discussed within the paper and needs to be closely clarified. This point becomes more important when the manuscript argues that their findings are beneficial to the calculations for life extension of wind turbine system. However, many of the existing turbines that would require life extension are equipped with SCADA systems with very low frequency output data, normally averaged values of order of minutes. It seems that this drawback is going to be removed by artificially increasing the SCADA signal sampling rate without providing any information on the original and the resulted sampling rate. Therefore, it is a far question whether this methodology is really applicable to SCADA data or condition monitoring system's data is still required. As such, it is again recommended to either clarify this point or discuss the probable limitations.

Moreover, there are well-established integration techniques that are being used, even for online conversion of acceleration to velocity and displacement signals [3, 4]. In the absence of artificial noise in the simulated speed data, the type of drift shown in Figures 3 and 4 are apparently showing a linear trend due to the accumulation of the error from the initial value. This type of trend can be usually avoidable by the common digital filtering with restriction of frequency range within the pre-processing of the signal, particularly when the signal's mean value (static term) is going to be added later on separately, in which case the initial value of integration doesn't really matter. As such, the authors need to mention that the application of regularisation within a trapezoidal scheme is not the only available method to stably convert velocity into displacement.

Minor comments:

The derivatives for the inverse integration scheme is similarly given in Hong et al 2008, so please refer to this paper in the beginning of the corresponding section.

Paragraph 15: "and" missing after strain gauges.

Paragraph 75: This is questionable if older turbines really possess SCADA with a desirable sampling rate that suits the manuscript's methodology.

Paragraph 85: As discussed previously, the assumption of the fully rigid gearbox and higher stage shafts, particularly for a multi-stage systems, needs further clarifications.

Paragraph 190: computationally "more" expensive.

Paragraph 205: Please clarify those three yaw directions.

Paragraph 210: where is the noise coming from that needs to be damped. Please clarify.

Paragraph 220: "the" modes.

Paragraph 225: pds should be expanded.

Mixed use of symbols has to be avoided throughout the manuscript.

Eq 18, it seems that gearbox ratio is missing somewhere.

Fig 8, it seems that some constant or near constant offset exist (observable in both time and frequency plots), apparently theta_static is not appropriately identified. The plotted FFT is too noisy, one can use a better illustration, perhaps, a power spectrum with sufficient averaging to get more clear peaks, as at the moment it is barely possible to get a good picture of the main peaks.

Paragraph 20: what are the sampling rate details and how the SCADA data's frequency resolution was increased, is it practically legitimate? Please clarify.

The results using the actual data does not look to add substantial value to the manuscript, as it is not really supporting the validation of the methodology.

Summary

The manuscript deals with an interesting question, relatively well written and structured, and will be of interest to WES audience, both academic and industrial. At the moment, this methodology seems to give a first-pass clue on the drivetrain specifications, rather than being capable to identify the specific loads/parameters of main shaft. It was discussed above that further clarification on some fundamental assumptions and some less major aspects are required. The authors are encouraged to either amend these requested points or moderate their arguments. In summary, this work is intriguing and suitable for publication once the requested major revisions are addressed.

References:

[1] J. Berglind, R Wisniewski, and M Soltani, Fatigue load modeling and control for wind turbines based on hysteresis operators, in: 315 2015 American Control Conference (ACC), pp. 3721–3727, IEEE, 2015.
[2] P Girsang, J.S. Dhupia, E. Muljadi and M. Singh, L.Y. Pao, Gearbox and Drivetrain Models to Study Dynamic Effects of Modern Wind Turbines, NREL/CP-5500-58960
[3] A Brandta and R. Brincker, Integrating time signals in frequency domain – Comparison with time domain integration, Measurement 58 (2014) 511-519
[4] L Qihe, Integration of Vibration Acceleration Signal Based on LabVIEW, Journal of Physics: Conference Series1345 (2019) 042067, doi:10.1088/1742-6596/1345/4/042067