Green = our response to questions, blue = indication how manuscript has been changed or statement why no changes were necessary
All line numbers apply to the revised preprint manuscript

Review Ozan Gozcu

The authors present a framework to analyze the effects of uncertain blade properties on aeroelastic stability of a turbine. The study is very relevant to wind turbine community since turbine stability is one of the hot topics as turbine sizes grow. The work includes results from different aeroelastic tools, surrogate model generation, damping and mode determination from time signals and uncertainty quantification test cases. This wide selection of tools and complex steps make it hard to understand the details for the reader.
We understand that the process is complex. To improve the clarity of the process and therefore the readability of the paper, we introduced a flowchart at the end of the introduction which visualizes the process of the study and the corresponding structure of the paper.

I have some comments that I hope can improve the paper:

- The references for the aeroelastic tools especially for their theory and capabilities are not presented in the study. Although I am familiar with some tools, I spend some time to understand the theory behind the tools I am not familiar. It is very time consuming and not always a successful process. I think, authors should give references for each tool's theory and capabilities, so readers can find the correct source for the tools.
    - We added citations to the theory manuals or equivalent papers in the simulation overview table (Table 1) and we cited these papers in the text of Sect 2.2 at the applicable statements. We tried to give references to 1) the main theory description document, 2) structural model description, and 3) aerodynamic model description. Unfortunately, this information is not always fully publicly available (e.g. alaska/Wind and Simpack).
- HAWC2 and HAWCStab2 are presented as same tool but in fact they are separate tools which use different formulations for beam solvers. In table 1. HAWCStab2 is presented as the linearization tool for HAWC2 results but in fact, it can compute equilibrium point and linearize its own solution.
    - We modified Table 1 by giving HAWCStab2 and Bladed (lin.) individual entries, indicating that these can be seen as separate tools (which is less applicable for Bladed (lin.), but gives a better overview of all models this way). We also corrected the properties of HAWCStab2 in the table and in the text (l. 103-104, 108, 110).
- Setting the operational points in each tool is not explained well. HAWC2 can be run in a constant rpm, constant pitch point but I don't think you can do the same for all tools. So, these differences also need to be explained.
    - We do indeed use an open-loop configuration for the linearized models. We tried to replicate this in the non-linear time domain simulations. In HAWC2 and Simpack it is possible to do time domain simulations with an open-loop configuration, i.e. it is possible to fix the rotor speed and pitch angle to a constant value, without controller intervention. In Bladed (v4.9) and alaska/Wind it was not directly possible to run time simulations with a fixed rotor speed. We therefore used a controller which aims to maintain the rotational speed as constant as possible. The rotational speed is almost constant, but the variation is minimal (especially because the inflow is uniform and constant).
    We added a discussion of this topic on l.112-116 in section 2.2
- The beam property conversion might be another source of error when different tools are used and it might help explaining the differences given in Figure 9. The shear center location is actually hidden in coupling terms for the tools which uses 6x6 stiffness terms directly (e.g. HAWC2), on the other hand SC location is generally a direct input for other tools. Besides, different tools use stiffness and inertia values which are computed at different locations. Although I don't expect to see all these details in the paper, I would like to see available references related to this.
    - All beam property modifications are done on the reference 6x6 matrices (BECAS output). This includes the stiffness reductions to obtain the critical reference condition and the uncertain parameter modifications for the case studies. The initial comparison between the models (section 2.3) serves as verification that this derivation was done correctly in all tools. This was described on line 270-274.
    - Note that we initially indeed modified all parameters in the tools individually. This resulted in significant differences. This was fixed by applying all modifications on the reference matrices and deriving the inputs accordingly.

- The modifications of the 6x6 matrices and the derivation of the beam properties is based on the work by Hodges (2006). The implementation can be found in the python modules preprocessor.py and libCrossSection.py in the software repository (Verdonck, 2023b)
- This topic is discussed on l.291-295 in section 3.2. The reference to Hodges is given for theoretical background how the modifications have been done and the reference to the implementation in the software repository has been added for readers that want to see all details.

- Although there is a blade frequency comparison given in Figure 1, mode shapes are not mentioned anywhere. I wonder if all tools give similar torsion and flapwise motions for the 1st edgewise frequency.
  - In preceding studies, mode shapes were satisfactorily compared, but this was not repeated for the modified model in this study. With previous experience of having a good match in mode shapes in addition, we concluded to reduce the verification to the blade eigenfrequencies, static blade deformation, and static aeroelastic test cases only.
  - No changes to the manuscript were required.
- The complex mode shapes and phase differences are not mentioned in stability analysis part. It would be interesting to see if different tools give similar phase differences from their complex eigenvalue analysis.
  - We agree that a mode shape analysis for the stability analysis might reveal additional insight on the instability mechanisms. The detailed analysis of these mechanisms was not the focus of this study, but rather the comparison of the sensitivities between the tools. A detailed analysis of the stability can be subject of further work.
  - This remark was added to l. 379-382 in the conclusion.

I have more specific comments below:

- Page 3, line 70 : “This effect can not be eliminated, but causes only a negligible periodic excitation” I think this is eliminated in HAWCStab2. Tower is an important element for system eigenvalues but can be assumed rigid for steady-state analysis and then real tower stiffness values can be included to eigenvalue and stability analysis. Of course this requires a lot of work for other tools but possible.
  - The proposed strategy is worth trying. However, as the reviewer already stated, it is a lot of work and can thus not be included in this paper. It may be subject of future investigations. We were not aware that this effect was eliminated in HAWCStab2.
  - Our statement on l.74-76 was corrected.
- Page 3, line 85 : “A stiffness reduction of 70% in flapwise direction, 30% in edgewise direction, and 70% in torsional direction was required to accomplish the desired instability behavior.” You could also alter the geometry such as prebend in flapwise direction, the swept in edgewise direction, aerodynamic center offset etc.
  - Yes, and it might have been better to include other parameters as well. However, as mentioned on lines 84-86, we established the critical reference condition in the beginning of this work, at a point where we did not have the automatic processes to modify all parameters. To keep the parameter space limited and manageable, we chose to limit ourselves to these three parameters. With the framework as it is implemented now, this could have been done better and faster. Nevertheless, we deem the critical reference condition to be suitable for the presented uncertainty quantification.
  - We did not deem it necessary to elaborate our discussion on l. 84-86.
- Table 1: Although, HAWC2 uses MB, HAWCStab2 doesn't use multibody approach. It uses corotational formulation for blades. Interesting, Simpack-Aeordyn has 6x6 stiffness definition for tower but not for blades. Can you give some references? Also see my comments about references for tools theory/capabilities above.
  - We corrected the properties of HAWCStab2 in Table 1 and on l. 103-104, 108, 110.
  - For Simpack-Aerodyn we use two different preprocessing tools for the tower and the blades. For the tower we use Ansys as preprocessor and import the tower as a single linear elastic body. Ansys uses the 6x6 properties directly. For the blades we use the Simpack internal beam description (SIMBEAM) in a multi-body implementation. At the moment of the work, SIMBEAM uses the engineering beam properties as input (6x6 input for SIMBEAM models is in development).

Green = our response to questions, blue = indication how manuscript has been changed or statement why no changes were necessary
All line numbers apply to the revised preprint manuscript

- Page 5, line 122: “only only” the same word typed twice.
  - This was corrected.
- Page 6, Figure 2. Have you talked with FAST developers (e.g. Jason Jonkman) about tip deflection results? I haven’t seen it in other studies such as IEA15MW turbine. You can check ORCAFlex IEA15MW report where OpenFAST is used for comparison.
  - This problem was raised as an issue and was worked on by Jason Jonkman and Andrew Platt. This issue is referenced in the paper (NREL, 2019).
  - Some commits have been added to the issue, but it is still open as of today (18.09.2023). The presented model was established before the latest commits by the OpenFAST developers.
  - No changes to the manuscript were required.
- Page 7, line 148: “<todo>” missing reference.
  - This was corrected
- Page 8, line 165: The difference for 1st EW BW seems the largest.
  - This was corrected
- Page 9, line 192-194: I expect to see Bladed time domain and linearized results match much better. Is it related to DMD or Bladed time domain results?
  - This question concerns the comparison between the Bladed linearization and the Bladed DMD-postprocessed time domain simulation in figure 4c. This figure shows a good match in frequency and damping between the 2nd edgewise BW mode within its unstable range. The 1st edgewise BW mode matches well in frequency and shows a similar trend in damping over the wind speeds, but the magnitude of the damping of the time domain simulation is significantly higher for most of the domain. Note that the DMD markers for the 1st edgewise BW are really small, indicating the small participation of this mode in the analyzed signals. We assume that the DMD postprocessing is sufficiently accurate in this case, and the main difference originates indeed from inherent differences between the linearization and time domain simulation. The results at operating points 9 and 14 m/s give the main reason for this opinion. The 1st edgewise BW mode is negatively damped in the linearization, but the time domain simulation is stable, i.e. there are no diverging signals. This is correctly identified by the DMD. The reason for this additional damping on the 1st edgewise mode in the time domain simulation is unknown.
  - We added this response in brief to l. 209-212. Additionally, the sentence on l. 206-208 missed a section, this was corrected: “The 1st edgewise BW modal component is also identified, but its participation in the time signal **is significantly smaller** and damping ratio is higher compared to the linearization results and the other time domain simulations.”
- Page 9, line 200-201: Any root cause of low damping values of SimPack? It is not particularly away from other tools in steady state analysis results. Any difference in unsteady aerodynamic part?
  - The root cause for the low damping values and the deviating frequencies in Simpack still has to be investigated. Simpack-AeroDyn uses a Beddoes-Leishman-like unsteady aerodynamic model, similar to the other tools. Unlike the other tools, we did not enable the dynamic wake model in Simpack. We verified that this is not the reason for the damping differences with a Bladed simulation without dynamic wake model.
  - For clarity and transparency, we added the statement that the root cause for the difference is unknown to l. 218.
- Page 10, Figure 4.: Is there any given small disturbance in time domain simulation to excite the modes, so that you can observe them clearly in the signals?
  - For the present work we were focused on the identification of the lowly damped modes in unstable operating points. For this use case we did not have to disturb the simulation, the interesting modes could be identified successfully with the DMD process. If the same approach is to be applied for stable time domain simulations, it might indeed be necessary to specifically disturb modes for a clear identification. This was not attempted in this work. A comment on the suitability of the DMD process is given on lines 372-374.
  - We did not deem it necessary to extend the comment on l. 372-374.
- Page 12, Figure 250-257: I think you explained the Sobol indices very well. Can you just elaborate the interactions? How should I interpret them?

  - The uncertainty quantification is based on a Sobol decomposition of the output quantity of interest, i.e. the function for the quantity of interest which depends on a certain number of uncertain parameters can be decomposed in a sum with terms depending on 1) none of the uncertain parameters, 2) only one of the uncertain parameters, or 3) combinations of two or more uncertain parameters. The variance of the quantity of interest can be decomposed by looking how each of these terms contributes to the total variance.
  - In the case of a variance-based uncertainty analysis with a PCE surrogate model, this becomes a lot easier to grasp. Assume we want to approximate a Quantity of Interest (QoI) of an unknown model with two uncertain parameters (X1, X2). Take following three basis functions for the polynomial expansion (X1, X2, X1*X2). This is just exemplary, depending on the uncertain parameters, different basis functions should be used. The PCE model will look like this: QoI(X1, X2) = C1*X1 + C2*X2 + C3*X1*X2. The unknown coefficients (C1, C2, C3) are determined by a least-squares regression (or something equivalent) based on the samples of the model. The variance of a polynomial is the sum of its squared coefficients. The total variance of the PCE model would therefore be (C1**2 + C2**2 + C3**2). To compute the first order Sobol index of uncertain parameter X1, we gather the polynomials which **solely** depend on X1 and compute the ratio of their variance with the total variance. This would be computed as C1**2 / (C1**2 + C2**2 + C3**2). Similarly, the first order Sobol index of uncertain parameter X2 would be C2**2 / (C1**2 + C2**2 + C3**2). For the total Sobol index of uncertain parameter X1 we gather all polynomials which depend on X1 and compute the ratio of their variance with the total variance. This would in this case be (C1**2 + C3**2) / (C1**2 + C2**2 + C3**2). Similarly, for X2 the total Sobol index would be (C2**2 + C3**2) / (C1**2 + C2**2 + C3**2).
  - No changes to the manuscript were required. A detailed description of the uncertainty quantification and surrogate modeling theory was out of the scope of the article and can be found in the referenced papers, e.g., Sudret, 2008; Le Gratiet et al., 2017.
- Page 15, Figure 6: HAWC2 total is more than 1 for total Sobol indices. Is it correct?
  - Yes, the sum of the first order Sobol indices can be maximum 1, the sum of the total Sobol indices will be at least 1.
  - Continuing the example from the last question. The sum of the first order Sobol indices would be (C1**2 + C2**2) / (C1**2 + C2**2 + C3**2). This will always be smaller than 1, or exactly 1, if there is no interaction between the uncertain parameters (C3 = 0). The sum of the total Sobol indices would be (C1**2 + C2**2 + **2 *** C3**2) / (C1**2 + C2**2 + C3**2). This will in the very least be equal to 1. The variance contribution of polynomial term X1*X2 counts double, as it contributes to the variance contribution of both X1 and X2.
  - No changes to the manuscript were required.
- Page 17, figure 9: The SC location is very critical for aeroelastic stability and Damage Equivalent loads in flapwise direction. I expect, it should also have substantial effect on EW direction stability. I don't understand why you don't observe it in the tools other than HAWC2 and HAWCStab2. Can it be related to my comment above about stiffness conversion? I might be wrong, but it would be great if you can add some physical explanations about the differences and observed results?
  - As described above, the stiffness modification and the uncertain parameter modifications are done directly on the 6x6 reference matrices. We believe that our verification proves sufficiently that the derivation of the input for the different tools is implemented correctly. We therefore believe that any difference in the uncertainty quantification study originate from the fully aeroelastically coupled simulation models themselves and are not caused by differences in the input generation.
  - No changes to the manuscript were required.
- Page 18, figure 10: I expect HAWC2 and HAWCStab2 damping curves with opposite slope, so that the damping decreases as the SC moves towards TE. Again, I might be wrong, but it would be great if you can add some physical explanations about the differences and observed results?
  - The focus of this work was the identification of differences in the sensitivity between different tools. Investigating the underlying physical explanations for these differences was beyond the study's scope. Nevertheless, it would be great to answer this question in combination with a study on the detailed instability mechanism in future studies.
  - No changes to the manuscript were required.

Green = our response to questions, blue = indication how manuscript has been changed or statement why no changes were necessary
All line numbers apply to the revised preprint manuscript

Review Anonymous Referee

The authors present a methodology to assess the uncertainty related to the blade properties on the aeroelastic stability of a wind turbine. The authors propose to construct a surrogate model of the wind turbine with the PCE approximation to reduce the computational cost and compute the Sobol indices. Different solvers are compared for the computation of the aeroelastic damping of the wind turbine and the effect of the uncertain parameters are compared. The work is interesting as it addresses an important topic for wind turbine design. It covers a large band of methodologies and tools, sometimes making understanding the general workflow difficult.
To improve the understanding of the general workflow, we introduced a flowchart at the end of the introduction which visualizes the process of the study and the corresponding structure of the paper.

Here are some general comments to improve the quality and understanding of the paper:

- if applying surrogate modelling to wind turbines is a recent topic, many works have already been done to propagate uncertainties on other mechanical systems with instabilities (squeal, flutter etc). The authors should include references to some of these works and emphasize how wind turbine models differ from other industrial applications.
    - Surrogate modeling has been widely applied to wind turbine models to quantify uncertainties in load analysis. There are very few applications of surrogate model uncertainty quantification to wind turbine stability in the literature. There is indeed literature available for other mechanical systems with instabilities. The instability mechanisms for wind turbines can be very different from those of e.g., an airplane wing or propeller whirl, so the uncertainties are also likely to be different.
    - We updated the introduction according to your recommendation by including some references to these works and a statement why wind turbine stability analysis differ from these other industrial applications (l. 34-37).
- page 7, line 148: there is a <todo>.
    - This was corrected.
- Table 2 gives the different parameters for the DMD methods. It is mentioned that the snapshot must be placed at the beginning of the instability. How do you set this up? How long is the selected time signal? I guess, if the signal is too long, the hypothesis of linearity loses its validity. Maybe an illustration of the different time signals of one dof could help in the understanding?
    - Indeed, if the signal is too long, the assumption of linearity breaks down and the result of the DMD becomes erroneous. The setup was a process of trial and error. Different setups had to be used for the different tools because the instability characteristics, such as the damping ratio or the time at which the signal ceased to be linear, were very different. Throughout, our assumption was that we wanted the snapshot to be as early as possible, but long enough for the DMD to give accurate and robust results. The window length was typically in the range between 10-40 seconds. Figure 1 shows exemplary how this looks like for a single signal. The figure shows the torsional deflection at 50 m blade radius for one of the Bladed simulations. The grey line shows the full signal, the red section is the snapshot which is selected for the DMD postprocessing.
    - This figure was added to the paper (Fig. 5)

Green = our response to questions, blue = indication how manuscript has been changed or statement why no changes were necessary
All line numbers apply to the revised preprint manuscript

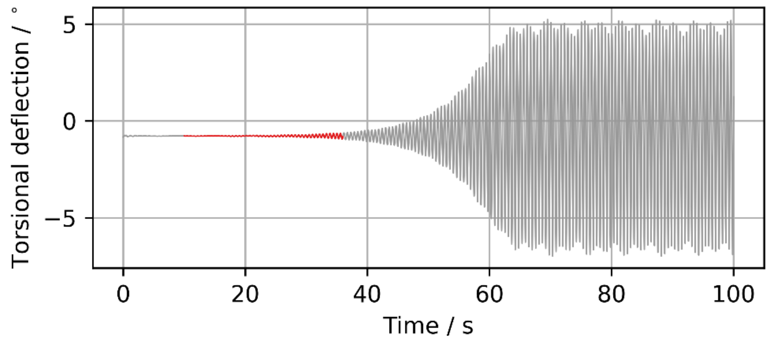

Figure 1: Example of a snapshot selection for the DMD postprocessing

- why consider only uncertainty on the mechanical properties and not include uncertainty on the wind speed? In some simulations, are there sometimes several unstable modes? If yes, how do you deal with it? If not, how would you generalise your methodology? Similarly, if you extend the variation range of your parameters, you may have a case where your instability is not always present, how would you deal with this case?
  - A whole range of other uncertain parameters could be considered. The selection of parameters shown in this paper is only a small subset of all uncertain parameters of an aeroelastic wind turbine model. We chose the structural beam properties, because they are known to affect the aeroelastic properties and because they are relevant design parameters.
  - Yes, in some simulations there are multiple unstable modes. This can be seen in the Campbell diagrams in figure 4. There are multiple modes with negative damping. In the case studies, we focused only on the most negatively damped mode, which was the second edgewise BW mode in all tools. For the given case studies with relatively small uncertainty distributions, this mode remained the most critical in all tools.
  - The presented methodology could be generalized and extended as follows:
    1. Instead of only one quantity of interest, all modes could be considered as quantities of interest. In addition to damping, the frequency or mode shape could also be considered. For each quantity of interest an individual PCE model should be fitted. This does not increase the computational effort significantly, since the same samples can be used. Only the least squares fitting of the PCE model needs to be repeated for each quantity of interest.
    2. Positive damping values are in essence no problem. The presented uncertainty quantification procedure could also be applied to stable systems. The only method that needs to be improved for this is the damping determination of time signals. The presented DMD methodology was tuned for unstable signals. This is reflected by the poor DMD results in the Campbell diagram at stable operating points (e.g. 8 m/s).
    3. The general procedure can also be extended to multiple operating points. This would require an accurate and robust mode tracking, which is often a difficult task.
  - The questions are valid and interesting, but we deemed this discussion out of the scope of the article. A statement on the need for an update of the presented process for other operating conditions or other uncertain parameters is included on l. 372-374 in the conclusion.
- how do you compute the Sobol indices? Are they directly deduced from the PCE coefficients, or do you use some sampling technics?
  - There is no sampling needed. The Sobol indices can be computed directly from the PCE coefficients. The Sobol indices are therefore also mathematically exact for the given polynomial. The variance of a polynomial is the sum of the squares of the coefficients (except for the coefficient of the zero-order term). The variance contribution of the different uncertain parameters (= Sobol indices) can therefore be calculated by computing the ratio between the variance of the terms dependent on a parameter and the variance of

the entire polynomial. We have given a more in-depth example in our answer to question 17 and 18 by referee Ozan Gozcu. A good description of the theory and application of PCE models for UQ is given by Sudret (Sudret, 2008).
- We added a brief statement how the Sobol indices are computed from the polynomials on l. 277-278 of the paper.

- for the PCE, what is the size of the expansion? Did you use some truncation technics to reduce the size of the basis?
  - For both case studies, a fourth order polynomial was used. For case study 1, with 3 uncertain parameters, this resulted in an expansion with 35 terms. For case study 2, with 4 uncertain parameters, this resulted in an expansion with 70 terms. This setup was not optimized for either the case studies or the tools. We could have optimized this setup and achieved similar surrogate model accuracy with fewer samples. However, as the verification of the PCE models for both case studies shows, the obtained PCE models are an accurate representation of the true model for all tools, which was the principal goal of this work.
  - The basic setup of the PCE is given on l. 244-248. We did not deem it necessary to expand this discussion.
- could you give some details on the simulation time associated with the different solvers? This would help to emphasize the interest in using surrogate models.
  - Within this work, the models in the different tools were not optimized for computational time. A direct comparison of the computational time could give a false impression of the performance of the presented tools. In unfavorable conditions, we observed wall clock times in the order of magnitude of 10 times the simulated time. We performed simulations of 100s simulated time for each sample point (in retrospect, this could have been shorter, since we only use a snapshot in the beginning of the time signal, as discussed in our response to your third question). This resulted in wall clock times of >10 minutes per sample. The linearization computations were also in the range of a few minutes per operating point.
  - A comparison of different numerical approaches for the sensitivity analysis was beyond the scope of the article. Crestaux et al. (Polynomial chaos expansion for sensitivity analysis - ScienceDirect) have shown typical differences of necessary model evaluations associated for meta-models, especially for PCE surrogates, in comparison to Monte Carlo simulations. They show that PCE surrogate models are especially well suited for low-dimensional problems with a maximum of 10-20 uncertain parameters. If the desired number of uncertain parameters exceeds this number, the classical direct Monte-Carlo uncertainty quantification should be used. Alternatively, a hierarchical approach, which first identifies the sensitive parameters with simple screening methods, followed by a detailed variance-based uncertainty quantification on the subset of most sensitive parameters as shown by Hübler (Hübler, 2017) could be used.
  - In our opinion, a quantitative statement on the computational times could be misleading, so we did not include this in the manuscript.
- For the second test case, there are strong differences between the Sobol indices depending on the solver. Could this be explained by differences in the solvers, the initial modelling and/or the uncertain parameter considered? What good practice would you give to engineers in this context?
  - Within the scope of this work, we did not perform a detailed analysis of the instability mechanism itself. We can therefore not answer the question why some variables are more sensitive in one tool than in another. We did verify that:
    1. The aeroelastic models are as similar as possible. This was shown by the model verification in section 2.3.
    2. The uncertain parameter modifications that we applied in the case studies were applied on the common reference dataset with beam properties described with 6x6 mass and stiffness matrices. This makes sure that our modifications represent the same modification in each tool. A further verification of this parameter modification was done, but not shown in the paper.
    3. The surrogate models were verified by the leave-one-out tests, such that we can be confident that the Sobol indices are not significantly affected by inaccuracies in the surrogate modeling.

  - The main message we want to convey here is that although the basic aeroelastic properties of different models and the comparison of Campbell diagrams may look very similar, the parameters influencing said instability could still be significantly different for different tools, as the second test case shows.
  - We added the last point of our response to the conclusion (l. 388-390)
- Only Sobol indices are compared. However, are the damping distributions impacted in similar ways? What about the resonance frequencies and mode shapes?
  - The resulting PCE models can be resampled in a computationally efficient way to provide detailed insight in the uncertainty propagation. In this way, a detailed damping distribution can be generated. This can be done for all uncertain parameters together, but also for a single parameter or combination of parameters. The damping follows a normal distribution. We decided not to include these plots in the paper, because they do not provide any additional information. The total spread of the damping and a detailed insight into the isolated influence of each parameter can be seen in figures 7 and 10, and the contribution of each parameter to the total uncertainty is shown by the Sobol indices.
  - In this study, we did not look into resonance. All models were symmetric, such that there could not be any periodic excitation. The modal frequencies are also a lot less sensitive to the given uncertain parameters. Nevertheless, this could be an interesting topic for further studies.
  - A detailed analysis of the stability mechanisms was out of the scope. This would have required an in-depth analysis of the complex aeroelastic mode shapes, which would be an interesting study in itself. This is something we will look into in the future. We therefore also did not analyze the influence of the uncertain parameters on the mode shapes.
  - The last point of our response was added as outlook in the conclusion on l. 377-382. No further changes to the manuscript were required.

Review Leonardo Bergami

I am not a reviewer, would just like to share a couple of thoughts/comments from reading through the pre-print, in case they could be useful:

- It is not clear from the text how the turbine controller is setup in linear and non-linear simulations. Typically, by default the linear stability analysis tools assume an open-loop configuration, ie. they model a system where the steady equilibrium state is assumed to be maintained without any intervention from a controller (either in torque/speed or pitch). A closed-loop linear analysis requires a linearized version of the controller to be included in the analysis (aero-servo-elastic analysis in HawcStab2 terminology).
  - We do indeed use an open-loop configuration for the linearized models. We tried to replicate this in the non-linear time domain simulations. In HAWC2 and Simpack it is possible to do time domain simulations with an open-loop configuration, i.e. it is possible to fix the rotor speed and pitch angle to a constant value, without controller intervention. In Bladed (v4.9) and alaska/Wind it was not directly possible to run time simulations with a fixed rotor speed. We therefore used a controller which aims to maintain the rotational speed as constant as possible. The rotational speed is almost constant, but the variation is minimal (especially because the inflow is uniform and constant).
  - We added a discussion of this topic on l.112-116 in section 2.2

The open loop configuration (no interaction with a controller) could be tricky to reproduce in non-linear aeroelastic simulations where a specific steady operational point should be kept.

The manuscript could benefit from better explaining whether a open-loop or a closed-loop configuration is reproduced in both the linearized and non-linear simulations, and if a closed-loop configuration is used also in the linearized tools, how the controller linearization is performed.

A mismatch in between open and closed loop, or in the linearization of the controller behavior could possibly (partly?) explain some of the mismatch observed between linear tools and DMD estimations especially around rated wind speed (where the controller response is typically less linear).

- In the uncertainty quantification case studies (section 3) the operating parameters of pitch and rpm are kept constant. Although practical this could lead to a variation in the steady state and in cases steady state that would not happen in real operation. Variations of eg. torsional stiffness and/or shear center would modify the angle of attack distribution along the blade, and thus change the power output when not compensated for by changes in pitch (whereas in more realistic operation nominal power output would be kept above rated). In other words, keeping a steady state condition more "similar" to the baseline one in terms of eg. power output, loading of the blades and aoa distribution would require instead a variation also of the pitch angle. Without changing the pitch, large variations of torsion or shear would actually bring the blade to be loaded in a completely different way from baseline, and thus make it hard to distinguish whether the observed changes in damping come from the completely changed loading distribution, or from changes in aeroelastic behavior per se.
    - Yes, you are correct that we shift the operating point when we change e.g. the torsional stiffness. It has not been our intention to realize a constant torque or constant power model. Our results would indeed likely differ, if this was done. Rather, our intention was to increase the reproducibility and simplify the comparison between the different tools, such that differences in the stability analysis and differences in the sensitivity of the uncertain parameters on the stability are most likely the result of differences in the structural dynamic and aerodynamic modelling in the tools. Note that the uncertain parameter variations in the uncertainty quantification studies are also relatively small. We will include this explanation of the consequences resulting from our simplification to maintain constant operating settings.
    - This clarification was added to the first paragraph of section 3 (l.225-229) and to the disclaimer on line 393 of the conclusion.
- Page 8. L.165. Isn't it the other way around? HS2 and BladedLin have better agreement on 2nd modes than 1st?
    - You are correct. The 2nd edge modes are in better agreement than the 1st.
    - This was an editing error and was corrected (l.178-179).
- Small "appearance" comment, please consider whether color sequences a bit more friendly towards color blindness could be used in the plots.
    - The color sequences in the figures were updated to improve the clarity for readers with color blindness. Figures 1, 2, 3, 6, 8, 9, 11 show comparisons between the tools. The same color scheme is used for all these plots. This color scale was tested with Coblis – Color Blindness Simulator. Where applicable different linestyles and markers were used. Figures 5, 7, 10 use one of the color maps suggested by Crameri et al. (2020). These figures were also tested with the color blindness simulator.

Green = our response to questions, blue = indication how manuscript has been changed or statement why no changes were necessary
All line numbers apply to the revised preprint manuscript

List of manuscript modifications in order of appearance in paper

(readability or spelling improvements are not listed)

Location	Improvement	Referee
Authors	Name change due to marriage Otto Braun -> Otto Schramm	/
l. 4-5	To put less emphasis on the computational benefit of PCE, the statement on the increased numerical efficiency was removed from the abstract. This benefit of PCE was not in detail investigated in this work	Anon
l. 34-37	Added references to application of PCE on stability analysis for other mechanical systems	Anon
Fig. 1	Paper overview workflow added to improve clarity of the procedure. This figure is referenced on lines 51-63	OG / Anon
l. 74-76	Corrected statement on periodic effects due to tower deflection	OG
l. 76	Removed statement of fixed operating conditions and rigid clamping at the root. This was mentioned elsewhere.	/
l. 92-99	Improved introduction of the simulation models	OG
l. 102-112	Added references to the theory manuals or scientific literature if applicable and added information for HAWCStab2	OG
l. 112-116	Added discussion on modeling of the control systems in the different models	LB / OG
l. 127	Added footnote stating that OpenFAST was only used for the verification studies and not for the stability analysis and UQ	/
Table 1	Modified format + references added + properties HAWCStab2 corrected	OG
l. 134	Corrected typo	OG
Fig. 2	Modified colormap	LB
Fig. 3-4	Modified colormap, introduced different linestyles	LB
l. 162	Added missing reference	OG / Anon
Fig. 5	Figure with exemplary snapshot for DMD analysis added. Reference to this figure added in Table 2.	Anon
l. 178-179	Corrected statement (differences 2nd edgewise < differences 1st edgewise)	LB / OG
l. 209-212	Added statement on likely cause for difference between Bladed time domain and Bladed linearization Campbell diagram results	Anon
l. 218	Added statement that root cause for differences between Simpack and other tools are unknown	Anon
Fig. 6a	Modified colormap	LB
l. 225-229	Added statement on the implication of the fixed operating condition assumption	LB / OG
l. 277-278	Added statement how Sobol indices are computed from PCE model and repeated reference to Sudret for further information	OG / Anon
l. 292-293	Added reference to the software repository for the python implementation of the structural beam property modifications	OG
Fig. 7	Modified colormap + introduced different marker types	LB
Fig. 8, 11	Modified colormap	LB
Fig. 9, 12	Modified colormap + introduced different linestyles	LB
Fig. 10	Modified colormap + introduced different marker types + corrected erroneous cropping of original figure (Simpack results were not visible)	LB
l. 377-382	Elaborated statement why detailed instability mechanism investigation is interesting to pursue in further studies and why it is out of scope here	OG / Anon
l. 387-390	Added main take-away for readers concerning the differences observed in case study 2 of the UQ	Anon
l. 393	Added disclaimer that assumption to use fixed operating conditions will have an influence on the result and should be taken into consideration when generalizing the results	LB