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.

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 will update our introduction according to your recommendation by including some references to these works and stating how wind turbine models differ from other industrial applications.
- page 7, line 148: there is a <todo>.
    - This will be corrected. The reference is (Verdonck et al., 2023a)
- 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. We will add this figure to the paper and insert a discussion along the lines of this response on line 149.

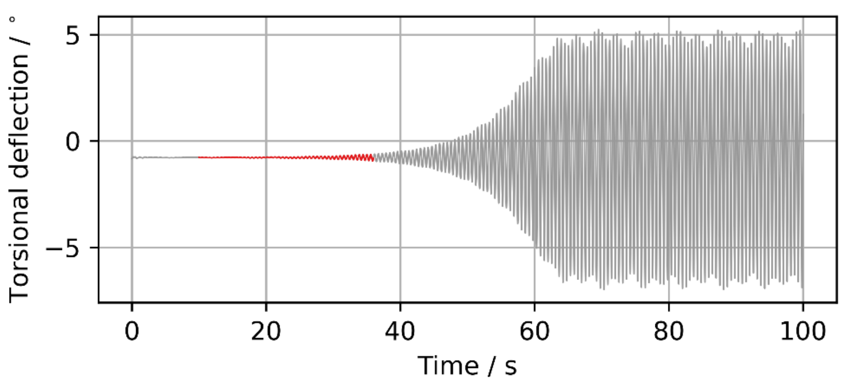

Figure 1: Example of a snapshot selection for the DMD postprocessing

- 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.
- 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 will add a statement how the Sobol indices are computed from the polynomials in line 257 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.
- 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.

- 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.
- 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.