Review of the manuscript wes-2016-38, entitled "Stochastic Wake Modeling Based on POD Analysis", by D. Bastine, L. Vollmer, M. Wachter and J. Peinke

This manuscript deals with a POD analysis of LES data of a single wind turbine wake, which extends results from a previous paper by the authors, Bastine et al. 2015. After a classical POD analysis and truncated POD reconstruction, two stochastic models and a spectral model for predictions of unsteady flows and loads connected with wake flows are proposed. Furthermore, an empirical technique is proposed to include small-scale turbulence in the prediction of dynamic loads.

The manuscript is very lengthy. Sometimes discussions are redundant, quite trite, or providing superfluous bridges between different sections. In contrast, key points of the work, such as description of the models, fitting procedures, are not described in detail. As reported in my detail comments, several figures and paragraphs can be completely removed.

Besides, writing and presentation, I have major concerns from the scientific and technical standpoints as well, which I am going to list in the following:

1. From Figs. 5 and 6, I guess POD analysis has not achieved a statistical convergence and several inaccurate conclusions might have been drawn. See my detail comments 16 and 17.
2. Performance of the uncorrelated model are extremely poor. Therefore, I recommend to remove this model from the manuscript.
3. Nonetheless, From Figs. 14, 15, 17 show that predictions obtained with the spectral and OU-based model are very poor as well. Even if mean and standard deviation of the original signal are predicted with a good accuracy, these predictions are completely out of phase. This makes me thinking that applications of these models to real flows with a varying atmospheric stability or operative conditions of the wind turbines will lead to very poor predictions.
4. In my opinion, the method proposed to include predictions of small-scale turbulence is quite rudimental and without any theoretical background. I am concerned that these models might fail for real atmospheric flows. Actually, we have already quite robust models, such as these cited in my detail comment 5, to reproduce synthetic turbulence or, if needed, CFD tools.

Therefore, according to my comments, I cannot recommend this manuscript for publication this time. I hope that the comments reported below might be useful for the authors.

Detail comments:

1. P1L5: "…load static characteristics"; if I am not mistaken, the proposed model can only predict load fluctuations, is that right? In that case please revise your abstract.
2. P1L2 and throughout the paper: "which" typically goes after a comma.
3. P2L32: "… differential equations can be obtained by projecting…" I guess you mean performing Galerkin projection
4. P3L3: there is a typo, Kalman.

5. P3L16: "it is principally possible to capture the small-scale properties of the flow by adding a homogeneous turbulent field to the wake structure modeled by the POD-based approach". In my opinion this is theoretically incorrect and, thus, it lacks of generality for the model. The model can be satisfactory from a statistical standpoint because approaching smaller and smaller scales turbulence becomes more isotropic. However, turbulence theory clearly indicates that there are specific relations between correlations and energy content at different scales, which vary for different characteristics of the specific turbulent flow. A good example to produce a synthetic turbulent signal is the Mann's model (J. Mann, The spatial structure of neutral atmospheric surface-layer turbulence, JFM, 273, 141-168, 1994), or the modified version for stably stratified flows proposed in A. Segalini et al., A spectral model for stably stratified turbulence, JFM, 781, 330-352, 2015.

6. Fig. 1: the mean velocity field looks skewed in the vertical direction. Some comments are reported later in the paper. Please provide your justifications here.

7. P4L21: "Snapshots of this plane are shown in Fig. 3 revealing a variety of shapes of the wake structure". This information is trite. I suggest removing text and related figure.

8. P5:4: Revise Data in data.

9. Fig. 4: You filter out data with deficit lower than 40% of the maximum deficit. The maximum deficit is about 4, thus any value lower than 1.6 should be removed. How is it possible you still have negative values?

10. P7L24-30: Please rephrase this paragraph. It is quite cumbersome.

11. Sect. 3.3: The stochastic methods are described too quickly and it is difficult to get the main differences among them. I suggest dividing this section is sub-sections for each model.

12. P11L9: explicit to which models belong to u or $\tilde{u}$.

13. P11L9: Remove "This discussion will enable us to gain a deeper understanding of the results presented in the next sections 4-6." That's obvious, and as it should be indeed. Please remove this sentence.

14. P11L12: " flow structures in the rotor plane change in time due to the hydrodynamics of the flow field". What do you mean for hydrodynamics of the flow field?

15. P11L19-28: I suggest to remove it. It is a quite obvious discussion.

16. Fig. 5a: Can you show the convergence of the POD eigenvalues and POD modes of interest for different numbers of snapshots and different sampling time?

17. Fig. 6. POD modes typically capture flow dynamics as couple of two POD modes with about same energy content (POD eigenvalues), spectral content, but they are orthogonal. Your first POD mode is clearly isolated and decoupled from the other modes. Therefore, it should not be associated with flow dynamics. In contrast, this might be a sign of not-achieved convergence of the POD analysis. If you try to reduce the number of snapshots, then energy of this mode should increase. Can you please verify my speculation?

18. Fig. 6: Showing the POD modes does not provide any essential information. I would save space by removing this figure.

19. P18L18: Explain more in detail this fitting procedure.

20. P18L19: "S_0 is systematically underestimated due to the logarithmic function". Why a fitting with a log function always underestimates?

21. P18L18-P19L4: You present 2 figures (6 panels) is 6 lines. If these plots are not crucial,

then just remove them.

22. P19L6-L19: Since here and in the following you will show that the uncorrelated model is highly inaccurate (see Fig. 14, 15c, 17 etc.). Then, why do you present this model? In my opinion, a scientific paper should present the main information for the community in a concise way.

23. Fig. 14. In my opinion, these models do not provide a satisfactory prediction. Are you sure it is worth to document these results?

24. Fig. 17. "For the *OU-based-* and 10 *spectral model*, the time series resemble the loads of the truncated POD but drawing further conclusions from a single short time window is difficult", In my opinion, the model predictions are completely out of phase. Why we should learn about these models?

25. P27L10-12: "We use a three-dimensional spectral surrogate of this region, as introduced in Sect. 3.4, to build a homogeneous turbulent field with similar structures. This surrogate is shown in Fig. 21c." This small-scale turbulence is already included in your POD modes. Why don't you try to recover this information from your POD results?

26. P27L16: "Outside the structure, we use the atmospheric boundary layer flow from the LES which is uninfluenced by the turbine" Do you add the mean flow or the instantaneous turbulent flow? In the second case, in my opinion this procedure is theoretically incorrect. You can find a large number of papers describing interaction between wakes and boundary layer flows.

27. Fig. 24: Is this a satisfactory prediction?