Wind Energ. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/wes-2021-16-RC1, 2021

Comment on wes-2021-16

Wim Munters (Referee)

Referee comment on "Modelling the Wind Turbine Inflow with a Reduced Order Modelbased on SpinnerLidar Measurements" by Anantha Padmanabhan Kidambi Sekar et al., Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2021-16-RC1, 2021

In their paper "Modelling the Wind Turbine Inflow with a Reduced Order Model based on SpinnerLidar Measurements", the authors present a POD-based representation of the turbine inflow from virtual lidar measurements. The paper is original and of interest to the community. The paper is well-written and easy to read. The conclusions are properly supported by the findings.

My main comment however is that it is not clear to me how the presented methodology could be used in practice, and that the presented work is focused on exploring the low-rank structure from SpinnerLidar measurements rather than resulting in a practical inflow reconstruction method. Furthermore, the paper is based on a single simulation setup, which is lacking some information for reproducibility. The paper would benefit from a clarification in these areas. This is further explained in my comments below.

Main comments
1. It is not clear to me how the model could be used to reconstruct turbine inflows in practice. Reconstruction capabilities of the POD model are shown solely for a data set which the modes were fitted on, so one could say the authors do not present a model but rather the presence of low-rank structure / compressibility of the lidar data as the result of a fitting/interpolation exercise. This is of course valuable information, but the true merit of such a ROM would lie in the application and performance for unseen data.
a) Could the authors clearly indicate, perhaps using a diagram with the flow of data and information, how the proposed method could be used in practice?
b) Can the authors elaborate on the three-parameter function $\hat{Y}(t)$, presented in Eq. 10
c) The merit of the proposed method seems to be mainly the fact that feeding the lidar data in a truncated POD basis, rather than feeding the entire signal, into these $\hat{Y}(t)$ functions allows the lidar data to be compressed and reduces the amount of data to be processed. Does this result in significant cost savings justifying the computation of the POD modes in the first place? Some indications on where the proposed method excels over using the full lidar data are necessary.
d) Could the POD modes be used outside of the dataset which they were fitted on? Or are the reconstruction capabilities of the model only available as a post-mortem processing? The authors mention in line 390 that converged POD modes could be re-obtained on the

fly relatively easily. However, reconstructing the wind field also requires the time-evolving coefficients in Eq. 7, for which the full signal $V'(x,t)$ is again necessary.

2. The authors build their argumentation based on a single simulation setup of an NREL 5MW turbine with a DTU SpinnerLidar in an unstable boundary layer.
a) Have the authors considered adding a second lidar to mitigate the cyclops dilemma? Please comment on this in the paper.
b) The details of the precursor simulaton could be improved. The paper would benefit from including information of the following:
- The authors use unstable stratification, what is the surface heat flux? Why did the authors opt for unstable stratification?
- How are the precursor simulations initialized? (temperature + velocity?)
- What is the boundary layer height after the spinup time? (Could you include snapshots of the precursor velocity field? (x,y) ; (x,z))
- p. 159, what do the 'default settings' of PALM imply? Please make the description of the setup self-contained.
- I'm assuming the precursor is periodic in x and y, please confirm.
c) I feel the paper would benefit from adding a second case in neutral or stable stratification, where turbulence structure will be significantly different from the large convective structures in the present case. If however the authors expect this would not influence their findings, they should at least discuss this in detail.

Minor / technical comments
- Figure 7, cumilative -> cumulative
- Figure 8, units
- p. 10, l 245: having has the largest... -> having the largest
- p. 10, l 245: cov -> $\rm{cov}$
- p. 11, l 262: "... start to converge around n = 3500 samples due to temporal correlations in the wind field." What do the authors mean by these temporal correlations causing convergence? Is this convergence caused by having snapshots that are separated by smaller time intervals when increasing the amount of samples, or are the n=1000 samples spaced equally as the n=3500 samples?
- p. 13, l 277: The authors mention that "energy is distributed over different scales and its representation might require an enormous amount of POD modes." This is an interesting comment. There is indeed a wide range of scales, but the energy cascade with decreased energy at small scales is effectively what allows us to do LES. On the other hand, it is true that further dimensionality reduction is notoriously difficult, perhaps even more so than widely considered in literature. A good reference here would be Bauweraerts, Pieter, and Johan Meyers. "Study of the energy convergence of the Karhunen-Loeve decomposition applied to the large-eddy simulation of a high-Reynolds-number pressure-driven boundary layer." Physical Review Fluids 5.11 (2020): 114603.
- p. 14, l 297: The authors claim that the M=1 reconstruction only allows to properly reconstruct the mean. Since the mean is not considered in the POD decomposition (Eq. 4), I suspect it is added afterwards to the POD reconstructions of the POD modes. Does this imply that the first mode offers no significant reconstruction information at all?
- p. 19, Figure 12 caption: u_projected refers to equation 1. But is not defined there. Please be more precise.