Response to Referee #1: Convergence and efficiency of global bases using proper orthogonal decomposition for capturing wind turbine wake aerodynamics

Overall Comments

Dear Authors,
I have reviewed your article and would like to provide my feedback on it.
The article investigates the performance of the Proper Orthogonal Decomposition (POD) method when applied to a dataset that is obtained combining data from different flow cases. A dataset of the flow inside a wind farm is used as the case study.
I think the topic of the article can be interesting for scientists working on fluid dynamics. However, there is a lack of significant developments that are specific to wind energy research, which is the topic of this journal. The objectives, most of the methodology and a large portion of the results are about the POD method itself and not about wind energy systems. So, I think this article is not entirely appropriate for this journal and, in its current state, is more suited to a journal on fluid dynamics.
I would also say that article is not very well written: many sentences, particularly in the Methodology section, are hard to follow and it's difficult to understand their content.
In the Discussion section, it is hard to understand what information is important. In the Conclusion there isn't any message specific to the wind energy field.
In reason of these comments, I am against publication of this article. My suggestion is to revise the text to improve clarity of presentation. Then, if Authors would like to discuss the POD in general, I suggest considering a journal of fluid dynamics, otherwise they should make it clear which is the impact and usefulness of this research for the wind energy community.

XXXXXXOnce again, thank you for the comments. This document contains written responses to each of the comments with links to changes in the paper, and we include our intermediate response previously posted online here in red. We have had particular focus on improving the Discussion and Conclusion.

Dear Reviewer,
Thanks for the comments. The specific comments mainly center around rephrasing for clarity, which we will address later. For now, we will just briefly address the overall comment of the review to provide additional background on the content and overall motivation of the article.
We fully acknowledge that the article is somewhere in between "applied wind energy" and "general fluid dynamics methodology". In fact, we have previously submitted the article to a more "fluid dynamic" journal, and we also believe that it has applications beyond wind energy as outlined in the Discussion. However, the article was rejected as it was viewed as too wind energy-specific. Hence, we have now resubmitted it to WES.
As the references show, other researchers have utilized global bases, but global bases have to our knowledge not been applied within wind energy before the development of PS-ROM (Andersen and Murcia Leon, 2022). Any newly developed model should in our opinion be continuously tested in

order to expand or verify the application, which is particularly important for data-driven models. This article extends the previous application beyond changes in CT to also encompass all turbines on a single basis, i.e. freestream, single waked and deep wake flows, which is a significant step toward generalizing the model. Wake aerodynamics are characterized by highly turbulent flows, and therefore we present new spectral metrics to quantify the efficiency across a range of conditions, particularly focusing on the optimality typically seen as a major advantage of POD although we here show it is less important than truncation. Finally, presenting a method to quantify the error related to stochastic realizations also has general implications for wind energy. Overall, a fast dynamic wake model delivering LES accuracy is in our view a significant contribution to the wind energy community.
Best regards,
Juan Felipe, Juan Pablo and Søren

Specific Comments

As said in the General comments, there are many sentences that are not clear and should be revised. These sentences are:

We have rephrased and clarified all sentences.

- First sentence of the abstract. This sentence, in the position it has (the first one of the article) is obscure. I suggest adding a few lines to introduce the article content.

We have rephrased the abstract and added more content to the first sentences.

- Line 35-36. These two sentences are not clear. You must explain what a single flow case is and the difference of considering multiple flow cases.

It was clarified what it means a single flow: *"the resulting bases are typically applied to data from a single flow case, which corresponds to a single point in a parameter space. A single flow case could for example be the flow around a cylinder at a specific Reynolds number or in the present case the inflow to a particular wind turbine in a wind farm operating at a single C_T value."*

- Line 46, "alongside ... parameter space". Not clear. This is the introduction. Try to give enough information so that any reader can understand.

We have rephrased the paragraph for clarity.

- Line 81, "which captures the wake dynamics". Maybe you mean the wake of the upstream turbine? This is not clear, because a turbine inside the line is affected by wakes of all upstream turbines.

"Wake dynamics" reference to the wake generated by the turbines upstream. The text has been clarifed.

- Line 101-103, "Ct of the wake-generating ... in the wake".

The description of the parameter space has been clarified.

- Line 104, "the operation of ... of the wake"

The text has been clarified.

- Line 106-107, "where the operating ... wake recovery".

Changed to *"where the flow has achieved a balance between extracted power and wake recovery"*

- Line 109-110. Isn't Fig.2 showing the Ct? Explain the relation between Ct and V.

We have clarified the explanation of the parameter space.

- Line 149-150, "Consequently, ... in the parameter space".

The sentences has been rephrased for clarity.

- Line 179-180, "these are not ... POD modes".

We have rephrased the sentence.

- Line 201-202, "which indicates ... different flows". Not clear, must be argued better.

A low error means that the flow is reconstructed efficiently despite the basis being derived from another dataset. We have rephrased for clarity.

- Line 211, "The basis error ... of modes".

Clarified.

- Line 232-234, "It is noted ... 0.4%".

Clarified.

- Line 82: "one radius upstream". The effect of induction is present also at a higher distance from the rotor. Maybe you should rephrase this sentence to explain what you mean with "reduce the influence of induction" and why you did not take the plane further upstream.

Yes, induction extends further upstream than $1R$, what we meant to say is that the induction will be approximately turbine agnostic further upstream. We have clarified with reference to Troldborg and Meyer Forsting, 2017.

- The Data availability section is missing. Code and data should be made available to the public to reproduce the results of the article.

We have added sections on code and data availability.

Technical Corrections

Line 68-69, "The grid ... outlet boundaries". If it is important, maybe it's worth adding a second panel in Fig. 1 with a sheme of the grid.

Thanks for the suggestion. We have elaborated on how the equidistant region is, but as it is a cartesian grid with simple stretching, we do not think a figure is not very informative nor necessary.

Line 90: "and U is the velocity", what velocity?

We have elaborated on this velocity.

Line 97: replace "atmospheric" with "undisturbed".

Fixed.

Line 274: of PS-ROM presented by Andersen and Murcia Leon (2022)

Fixed.

Line 299: you should use "" to make it (Andersen et al. (2015))

Fixed.

Response to Referee #2: Convergence and efficiency of global bases using proper orthogonal decomposition for capturing wind turbine wake aerodynamics

Overall Comments

The present manuscript tackles a potentially interesting topic for wind energy such as the development of a general basis for model reduction using POD. The subject is interesting and relevant for wind turbine's research, but there are some points that need to be clarified or developed.

Thank you for the comments to the paper. We have clarified aspects and provided responses to your comments below in blue.

First of all, the paper does not discuss in detail the contest of model reduction for wind turbines. In particular, the introduction lacks reference to the many previous works on POD analysis of wind turbine and wind farm flows (among others, VerHulst and Meneveau Physics of Fluids 2014, Bastine et al. Energies 2015, Hamilton et al., Wind Energy 2015 Physics of Fluids 2016, Phys. Rev. Fluids 2017, Wind Energy 2018; De Cillis et al. Wind Energy 2021, Renewable Energy 2022, Journal of Physics: Conference Series 2022). Also, DMD and ANN are mentioned but not put in the contest of wind turbine flows (see, for instance, Debnath et al. "Towards reduced order modelling for predicting the dynamics of coherent vorticity structures within wind turbine wakes" 2017; De Cillis et al. "Dynamic-mode-decomposition of the wake of the NREL-5MW wind turbine impinged by a laminar inflow" 2022, among many others).

We have added some references using POD for wind turbine wakes, but we have limited the references on DMD and ANN to general articles/reviews without providing a complete review as these methods are not used in the present paper, but merely mentioned as alternative approaches.

Moreover, I have some concerns regarding the convergence of the grid, which seems to me too coarse for the LES of 14 turbines, having only 20 points per radius in each direction. The grid convergence of the LES for the considered case needs to be shown in the manuscript, maybe in a dedicated appendix.

20 cells per radius is very highly resolved for simulations of wind farms with actuator discs. A full grid study is beyond the scope of this paper, but we have provided reference to Hodgson et al., 2023, who compared two state-of-the-art numerical solvers, including EllipSys3D used here. Hodgson et al. concluded grid requirements to be larger than 6.5 cells per radius, so we have 3x times the required minimum. Furthermore, typical design using the IEC-61400-1 standard corresponds to turbulent boxes with resolution of 32x32 covering the rotor, while these simulations has a resolution of 40x40 covering the rotor.

The same can be said about the choice of the time step set for the extraction of the snapshot from the LES, as well as the total time of the simulation. This quantities are known to affect considerably the convergence of the POD algorithm over one single dataset, so the convergence with respect to this parameters need to be discussed for some of the considered dataset.

The simulation time steps are chosen in accordance with standard requirements for actuator disc simulations to not conflict with CFL-conditions, where we maintain CFL less than 0.5. In principle, POD should be based on uncorrelated snapshots, i.e. separated by 2 integral length scales or more, but if the snapshots are not uncorrelated they will not add new information and eigenvalues will be 0 (machine precision). As stated the total simulation is 3.64 hours, which is very long for this type of simulation, and the simulations are statistically stationary. The insights into the convergence required by the reviewers are provided in Figure 7, Table 2, and lines 236-250 in the original manuscript, where the number of snapshots used for the local basis is varied and the number of independent snapshots is estimated to be 2048. We have clarified what independent snapshots mean. The alternative to changing the number of snapshots in each dataset would be to compare the convergence for a fixed number of snapshots, which was outlined the original discussion.

Moreover, I cannot see any information about the value of the tip speed ratio, and about the whether tower and nacelle are taken into account in the simulation, which are both very relevant information for the considered flow case . In case tower and nacelle are not taken into account, I suggest to discuss the relevance of the simulations to realistic conditions, and to reformulate using the work "rotor" instead of "turbine".

The actuator discs are fully coupled to Flex5, which gives dynamic estimations of loads and deflections. This also means that the turbines have a dynamic controller and therefore adjust the tip-speed-ratio relative to the local inflow according to the controller. We have added that nacelle and tower are not modelled, but we disagree with the reviewer that this exclusion should significantly impact the realism of the simulations with reference to the work by Zahle and Sørensen, who conclude that it has an influence of less than 2% based on blade resolved simulations. The influence on the resulting wakes would only decrease further into the row of wind turbines as the turbulence intensity increases.

Also, the discussion about the choice of the global POD basis is not sufficiently developed. I understand that the global basis is constructed in an iterative manner, but I am not sure what does this means exactly. For instance, the performance of a basis is evaluated only using the velocity error? Which is the condition for adding a dataset? And why the dataset with worst performance should be added? Since the construction of the global basis is a crucial point of the paper, it needs to be discussed in much more detail. The same can be said about the case study with stochasticity, which is not sufficiently clearly explained. For instance, the authors should explain in more detail the fact that "the actual projected spectrums are used.." etc.

We have clarified the iterative process for adding new datasets, where new datasets to be added are selected as the dateset with the largest error. It is an obvious choice as the global basis is intended to be efficient over the entire database and by adding a dataset with the highest error, the overall error is reduced the most and the global performance is improved. However, the Discussion also includes comments that the iterative process could be based on adding datasets based on alternative metrics, such as the velocity error, or arbitrary datasets.
The flows are projected into the global basis in the case study with stochasticity, and these spectrums are used. The alternative would be to predict the spectrums of unseen cases as done in Andersen and Murcia Leon, 2022. We have restructured the paragraph and clarified the text.

Finally, I think that the performance of the POD basis cannot be measured only using an integral quantity such as the velocity error, since even if the integral error is rather low, the flow field might have some important structural differences with the simulated one. The POD-reconstructed flow fields need to be shown and compared with the LES snapshots at given times, by showing a velocity error field for each velocity component.

Thanks for the good suggestion. We have added Figure 6, which show a very good reconstruction of all three velocity components.

Typos

page 2, line 59: Large Eddie –¿ Large Eddy

Fixed.

page 2, line 49: the acronym LES need to be defined

Fixed

page 13, line 275: spectrums –¿ spectra

Fixed