Authors' Response to Reviewer 2

General Comments. This paper presents a comparison of various numerical modeling tools while modeling wake steering operation for the IEA 22MW reference wind turbine. The paper is clearly written and presents useful results, and I believe the paper is publishable subject to a handful of comments.

Response: Thank you for your feedback.

We thank the reviewer for taking the time to read the paper and give detailed feedback. The feedback has helped improve the quality of the paper. We have adressed all the remarks point by point below.

Comment 1

My main comment is on the LES and whether it should be treated as the "truth" for comparison. I am not and expert on LES so I can't comment on the choice of simulation setup, although I have no reason to doubt that the authors have chosen a reasonable setup. However, I am interested in how closely the LES approximates true, atmospheric flows; and whether we should be treating the LES as the "truth" for comparison. For instance, my impression is that LES often contains lower wind direction variability than is usually observed in the atmosphere, and wind direction variability often strongly affects apparent wake location/expansion. Throughout the paper, the authors sometimes make statements comparing other models to LES, for instance " ... gives the closest results to LES", which I agree with. However, at other times, the comparison to LES is unstated/implied, such as "RANS underpredicts lateral deflections." My concern is that the implication is not strong enough, and some readers may believe that (in this case), RANS is "wrong", since it differs from LES. I think the authors could take a few steps to remedy this.
First, I would be interested for the authors to add a paragraph explaining how well the LES they used approximates observations of the atmosphere: what aspects of atmospheric flows is the LES capturing, and what aspects is it missing?
Second, that any "poor performance" from other models is only in comparison to LES should be made clear throughout (not necessarily in every sentence, but just so that there is no ambiguity for the readers).
Third, can the authors quantify the uncertainty in the LES so that readers know, when there is a small difference between the LES and another model, if that difference is meaningful?

Response: Thank you for this feedback. We have incorporated it in several places in the manuscript.

In lines 89-99 in the Introduction, we have further motivated the choice of reference results with the following text:

Out of the range of presented models, the large eddy simulations capture the most of the physics of the atmosphere and the turbine wakes. Hence, the LES results are chosen as a reference for the other model results. Further, the precursor simulations from the LES provide a series of inflow planes that can be used to clearly define the inflow for the other models, thus reducing uncertainty on the inflow such that differences in results between models can be traced back to differences in modeling choices not differences in the inflow. Thus making this a clean comparison between the different models. An alternative option would have been to use measurements as a reference, but they introduce uncertainties with respect to the inflow. While it is clear that the LES simulations do not capture all the physics of the atmosphere like the measurements, they do capture all the physics that the tested lower fidelity models are designed to capture, such that adding more physics of the atmosphere through measurements or meso-micro coupled simulations would be unlikely to materially change the conclusions of the present comparison [1].

In regard to the questions concerning the ability of LES to model atmospheric flows, comparison with measurement data and uncertainties, we have expanded the LES methodology section in lines 146-156 to address these further. Now included are references to validations/ verifications of the Ellipsys3D LES setup in modelling atmospheric flows, wind turbines and wakes. Like all CFD, LES results can be dependent on numerical choices, particularly if using an insufficient grid resolution. However, our setup is informed by previous verifications and cross-code comparisons to ensure grid convergence and hence minimise these uncertainties - this has been explicitly clarified in the text. The limitations of the setup - the constant geostrophic wind, and possible implications for wind direction variability - are also now included.

[1] P. Doubrawa, E. W. Quon, L. A. Martinez-Tossas, *et al.*, "Multimodel validation of single wakes in neutral and stratified atmospheric conditions," *Wind Energy*, vol. 23, no. 11, pp. 2027–2055, 2020. DOI: `https://doi.org/10.1002/we.2543`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.2543`.

Comment 2

Is the DWM driven by flow after filtering at the cutoff frequency f_C Or is the filter somehow built into the DWM model? Either way, how is the filtering applied to remove high-frequency content from the turbulence that drives wake meandering? What is the form of the filter?

Response: The filtering is built into the DWM model via spatial averaging of the background flow that is carried out at run time for the location of each particle. The filter is applied to the wake area with a wake diameter of $D_w(x)$ to obtain the particle advection velocity; this removes small-scale (high-frequency) turbulence components. Different options are available for the background flow, for this paper, the background flow included the mean flow, the wakes of upstream turbines and the large-scale turbulence. The wake of the turbine whose particle are being advected are not included, because the Hill Vortex model already accounts for that. We refer the reader to Equation 2 in the manusript and the text surrounding it.

Comment 3

Line 389: "It is clear that the LES and DWM signals are correlated". This is not really obvious to me looking at Figure 7, and besides is not very precise. Can the authors quantify how much correlation there is between the center positions (lateral and vertical) for DWM and LES?

Response: Thank you for this comment. We agree that the original statement was insufficiently precise. To quantify the correlation between the two models, we have added the following sentence to the manuscript: The correlation coefficients between the wake center location time series of the two models are $r_{\mathrm{Y,max}}(\Delta T = -39\mathrm{s}) = 0.80$ and $r_{\mathrm{Z,max}}(\Delta T = -35\mathrm{s}) = 0.69$, where ΔT denotes the time offset that maximizes the correlation between the two signals. In addition, Figure 7 and the associated discussion have been expanded in lines 447-465.

Comment 4

How is secondary steering defined? That is, how is the second turbine's wake isolated from the first turbine's wake?

Response: There is no strict quantitative definion for secondary steering. [2] defined it as "... the most striking observation is that in the SOWFA case, the second turbine's wake appears to be deflected, even though that turbine is aligned to the flow. In other words, the second wake appears also steered." A similar description is also given in the Introduction in lines 48-50. So for the definition of secondary steering, the wakes are not separated. However, in the cited paper, they did carry out an additional analysis where they just shifted the upstream turbine to see if it would also lead to the same amount of wake deflection as steering the upstream turbine and it did not.

We have added two more plots in Appendix B to better quantify the steering in the wake of the second turbine. Figure B1 shows the deflection of the wake as obtained from applying the waketracking algorithm to the full flow field (with the superposed wakes). Figure B2 shows comparison of wake center deflection and deficits between PyWake and LES for the aligned and the steered case. Additionally, it also shows the PyWake results of a cases where (i) the upper turbine remains unyawed, but is shifted in the direction of yaw deflection to overlap with the yawed wakes at the position of the second turbine, and (ii) a corrected steered case where the second turbine is also yawed by the average amount of yaw the second turbine sees as inflow in the LES data.

We have also modified some of the text in accordance with this in lines 503-526.

[2] P. Fleming, J. Annoni, M. Churchfield, *et al.*, "A simulation study demonstrating the importance of large-scale trailing vortices in wake steering," *Wind Energy Science*, vol. 3, no. 1, pp. 243–255, May 2018. DOI: 10.5194/wes-3-243-2018.

Comment 5

What computational resources (computational hardware and computer wall time) are needed for each of the models? A table comparing the models in terms of computational complexity would be interesting for readers.

Response: Thank you for this useful remark, we have added a table with the computation time at the end of the results section along with the following text in line 533.

Comment 6

The term "benchmark" not really defined. As I understand it after reading the paper, this is a "benchmark test" in the sense that multiple models are compared; however, it is not a "benchmark" that others can compare to, as I was originally expecting based on the title. Consider removing the term benchmark or clearly defining what you mean by "benchmark".

Response: Yes, that was not very precise in light of other papers using the term benchmark. We have changed the title to "A multi-fidelity model intercomparison for wake steering of a large turbine in a neutral ABL".

Comment 7

Am I understanding correctly that using LES for the background flows for the DWM is ok because it only needs to be run once (whereas the DWM model itself could be run many times to compare controllers)?

Response: Yes. We use the LES precursor inflow to enable a controlled, one-to-one comparison across models. The LES inflow only needs to be generated once; the DWM can then be run many times with different layouts or control settings using the same inflow. In principle, multiple precursor inflows could be generated to sample variability, but that is beyond the scope of this study.

Comment 8

Is the inflow WS, TI used for PyWake the average over the hour from the LES?

Response: Yes. For steady-state models (RANS, Fuga, PyWake), we use the one-hour LES time-averaged inflow. The description in the manuscript has been modified in the text in lines 317-318.

Comment 9

Why is the Hill Vortex deflection model used in DWM but Jimenez used in PyWake? Is that just to do with what is implemented?

Response: The differences between the two models are not that big, so we just wanted to show the results for both since both are available in both frameworks.

Comment 10

"Half a degree of wind direction offset corresponds to roughly three degrees of yaw steering" — what do the authors mean by this? Do they have analysis or citations to support this?

Response: This was the result of a simple back-of-the-envelope calculation based on the deflection curves in Fig. 5. For a 0.5° wind-direction offset (so not considering steering but assuming the turbine is aligned with the incoming wind), the downstream deflection is approximately $\Delta y(x/D = 15)_{WD}/D \approx \tan(0.5) * 15 = 0.1$. For LES, the change in deflection between $\gamma = +10°$ and $\gamma = -10°$ is about $\Delta y(x/D = 15)_{\Delta\gamma=\pm 10°} \approx 0.5$. Scaling linearly gives

$$0.1/0.5 \times 20° \approx 4°,$$

i.e., a half-degree wind-direction offset corresponds to roughly 4° of yaw steering in terms of downstream deflection. However, we have reformulated the text to make it less strong in lines 323-326.

Comment 11

Fig 2. Is the dramatic change in deficit in the near wake when steering in the DWM expected?

Response: Yes. The reduced near-wake deficit in the steered DWM case is an expected consequence of the reduced axial loading of the turbine under yaw misalignment. In the model, the effective thrust coefficient scales with the square of the cosine of the yaw angle, i.e.

$$\tilde{C}_T = C_T \cos^2 \gamma.$$

For $\gamma = -30°$, this gives $\tilde{C}_T \approx 0.75\, C_T$. Using the standard actuator-disk relation between thrust coefficient and axial induction,

$$C_T = 4a(1-a),$$

this corresponds to a reduction in axial induction of roughly 30%, which directly leads to a substantially weaker velocity deficit in the near wake. Please note that the actual inflow conditions for the DWM are a bit different since correction factors are introduced to give a good match in the far wake, see Equations (6a) and (6b) in the manuscript.

Comment 12

Fig. 6: Why do the colors in the legend not really match the bars? Also, there appear to be four "textures" defined, but in the figure I only see the "HAWC2 ghost turbine" and "velocity + power curve" cases in the plot. (or possibly the "HAWC2 turbine" does appear for DWM only at x/D = 0)?

Response: The colors are the same, for the ones with texture the opacity was decreased a bit. Also some of the textures are only there for the first turbine and since the errors there are small they are not very visible. We have modified the Figure caption to make it more clear.

Comment 13

397 "max out at" is a bit informal—consider revising.

Response: Indeed, that was a bit informal, it was changed to something more formal.