Second review of "The Sensitivity of the Fitch Wind Farm Parameterization to a Three-Dimensional Planetary Boundary Layer Scheme" by Rybchuk et al., submitted to Wind Energy Science Discussions

The authors have made concrete efforts to improve the paper and address the reviewers' concerns, including my own. I am especially glad that the authors checked the TKE advection issue and that they showed strong evidence that it was indeed active in their runs. That was my major reason of concern and it has been satisfactorily addressed.

As for the extremely long reply about the revisited motivation of the study to address uncertainty quantification (as opposed to validation), I could counter-argue that what the authors did is not exactly what Archer et al. (2014) recommended, which was basically to use ensembles. They meant several ensemble members, like 15-20, but here there are only 2 (MYNN and 3DPBL schemes). A 2-member ensemble is not sufficient to characterize uncertainty, but I will grant that the authors found a substantial difference in power, which is interesting and good to know, thus I will let this go.

Replies to previous comments

3. 101: I think I know what you are trying to say, but it needs to be defined better because an external wake cannot be defined as a "distance". Also, here you use 0.2 m/s as the threshold, but in the rest of the paper it seems to be 0.5 m/s (e.g., Figure 3 and 11, dashed blue line).

Following feedback from Reviewer 1, the literature review section has been removed. While the text has been removed, we want to address the challenge of characterizing external wakes here, as this feedback is also brought up later. As the recent WFP literature review paper by Fischereit et al. (2021) states
"One challenge that we identified is that from our review there is no standardized or common definition of a recovery length behind a farm. Studies used for instance the e-folding distance (Fitch et al. 2012), the location of 2% difference between a simulation with and without wind farms (Pryor et al. 2020) or the location where the wind speed has recovered to 95% of the freestream wind speed (Cañadillas et al. 2020). Due to this variety of different definitions, it is difficult to compare wake lengths across studies quantitatively."
Each of these studies selects a definition of an external wake that is reasonable for the question they are studying. In our analysis, we study wakes from large farms of large turbines. As such, we select three metrics: 1 m/s threshold, 0.5 m/s threshold, and the e-folding distance because we expect large wakes. This is imperfect, but we also acknowledge this challenge that the field faces.

I think that the authors misunderstood my comment. I did not criticize the use of a threshold at all, that was and is just fine. I just found it confusing that the authors used two different values (0.2 m/s and 0.5 m/s) in two parts of the same manuscript. I actually wonder if 0.2 was a typo perhaps? In fact, you no longer use 0.2 m/s in the revised manuscript. Not important anyway, but I thought I would clarify what I meant.

4. *Table 1: the same label here is used to indicate three different runs. Please use unique labels for each run, like "S-NWF" for stable, "U-NWF" for unstable etc.*

While it is entirely reasonable to label the idealized runs "S-NWF" and "U-NWF", it would not be reasonable to name the monthlong real run in a similar manner. Thus, for the sake of consistency between the ideal runs and the real runs, we retain the original label format.

I am sorry, but I disagree. You must not use the same label to indicate different runs, period. The point of a label is that it identifies uniquely the run you are talking about. However, the labels are not even used in the manuscript, so this discussion is moot. I would just suggest that the column "Label" be actually renamed "PBL scheme" and the prefix "NWF" be removed because this table is already for no-farm simulations.

8. *322: Why 0.5 m/s deficit if 0.2 m/s was stated earlier?*

As stated in response to Minor Comment 3, there is no standard for external wake characterization and 0.5 m/s threshold makes sense for the scale of problem we are studying.

Again, I was not criticizing the value, but noting the contradiction of using two different values in two different parts of the manuscript. All resolved now since there is no longer a 0.2 m/s threshold.

15. *476: Are these results with 0% TKE or 100% TKE? Why not 25% TKE as recommended?*

These results are with 100% TKE. Thank you for noting that omission, and we have updated the Methods section to include that detail. "All wind plant simulations are run with α=1. While validation of this parameter is limited, we note that Larsen and Fischereit (2021) saw more accurate results in an offshore wake study with that value (α=1) than the value of α=0.25 recommended by Archer et al. (2020)." (L232-234)

Actually, I read Larsen and Fischereit (2021) very carefully and they did not reach any such conclusion in their paper. They even explicitly stated that "It remains inconclusive … what are the correct C_TKE values to use; more measurements are needed for further investigation." The authors might be referring to Figure 14 in Larsen and Fischereit (2021), which shows very large TKE injection over the wind farm during one 2-hour flight, so large that none of the parameterizations, no matter which settings were used, could capture it. In particular, the fact that not even the results with advection turned off, which are well known to cause an overshoot of TKE in the grid cells of the wind farm because TKE has no way to go but continue accruing, could match the observed values indicate that perhaps something was off with the methods used to calculate TKE from the measurements. Even

Larsen and Fischereit (2021) themselves did not seem to trust the findings enough to make any recommendations about C_TKE based on their study. Anyway, I do not intend to conduct a review of Larsen and Fischereit (2021) here, but I think I provided enough evidence to suggest that the sentence at lines 232-234 should be removed.

17. *Figure 12: as in #14, not OK to have 4 colorbars.*

    Figures with multiple colorbars are employed within academic literature. For example, see Figure 3 in Pryor et al. (2020) and Figure 7 in Brugger et al. (2022).

    While other papers may have had good reasons to make such a choice, I really do not see the advantage of 4 colorbars here. It is extremely hard to compare results from one heatmap to the next. The only advantage is that one could discern patterns within the wind farm, but the authors do not actually discuss any such patterns in the text, so there is really no advantage.

New comments

1. P. 3, l. 79: the paraterization by Abkar and Porte-Agel does not have a name and calling it the "Abkar WFP" is presumptuous. The parameterization by Pan and Archer (2018) also is not called the Pan WFP, but it has a name, it is called the "hybrid" WFP.
2. P. 3, l. 89: your research question has changed to "How sensitive are modeled mesoscale wakes to the choice of PBL parameterization?", which is a much better fit to the study, but I think you should be a bit more humble and admit that you have only examined one additional PBL scheme, one that has not really been validated much because it has not even been released officially with WRF as far as I know. So I would recommend that you add a few sentences to explain your choices along the lines of: "Ideally, we should test all 13 PBL schemes with the Fitch WFP to fully characterize the uncertainty. Here we propose, as a first step, to compare two PBL schemes: the conventional NYMM and the newly-developed NCAR 3DPBL. We chose the latter because …. " and explain why you chose it over the other 13 schemes? A few sentences suffice, like the fact that it has a prognostic equation for TKE.
3. P. 6, l. 168: good idea to match the hub height wind speed rather than the geostrophic speed.
4. P. 9, l. 233-235: as explained above, this sentence should be removed because it is not consistent with the recommendations and findings reported by Larsen and Fischereit (2021) themselves.
5. What is "pp"? It is used only in the last sections of the paper (that is why I did not notice I earlier) but is not defined. Perhaps it means just "%", which was used in the abstract and other sections? Please use one convention only.