Response to Referee #1

Referee comments appear in black and author responses appear in blue.

This is a WRF modeling study of the wakes of future offshore wind energy areas planned along the US east coast. It considers three scenarios of offshore wind development: Vineyard Wind only, lease areas in the Mid-Atlantic, and lease + call areas in the Mid-Atlantic. An impressive modeling effort was undertaken, with nested high-resolution WRF runs for one year repeated for the three scenarios above, plus the control case with no farms, plus runs with 0% and 100% TKE added. The team had also numerical noise issues, as seems to be the norm when the wind farm parameterization is turned on, and therefore had to deal with several additional runs to take care of it. The paper is definitely worth publishing as it includes interesting and valuable results. It is exceptionally well written. It was a pleasure to read such a good paper!

We thank the reviewer for reading thoroughly and providing thoughtful suggestions to improve this article.

Major Comments

1. **The study focuses too much on added TKE.**

The paper promises to study annual variability of wake impacts (see comment #4 below about how it is monthly and diurnal, not annual).

Thank you for pointing out our inaccurate use of word choice. Annual variability refers to multi-year studies while our research focuses on one year. You leave more title-specific suggestions, including this one, in comment #4, so we refer to all title changes in our response to comment #4.

Such a study should have a main run with certain fixed parameters, perhaps a few case studies of special interest or a validation effort, and a few case studies to assess the sensitivity to some of the parameters. Instead, in this study a lot of effort was put on the sensitivity.
The parameter of focus is the amount of added TKE, which is 25% in the default settings of the WRF model, and which was found here to have a relatively small impact on the power output (<5%). One would expect that the main run would be with 25% TKE and then a few cases (perhaps one week in each season) would be run with 0% and 50% and 100% TKE, in addition to 25% TKE, to assess sensitivity. The main results reported in the abstract and in the conclusions would be obtained with the default 25% TKE and a sentence or two would address sensitivity to TKE.

Thank you for the suggestion. In planning our simulations, we had extensive discussions about the merits of 0% vs 25% TKE, and finally decided to use 0% as a bottom limit rather than 25%. The 25% recommendation was based on only one study. Archer et al. (2020) recommended the use of 25% TKE based on idealized conditions, with neutral stratification, and for a one-wind-turbine setup. The 0% added TKE is more similar to the impact in the Volker et al. parameterization which has been used in several studies.

Instead, on one hand the study focuses excessively on the sensitivity to added TKE, because all the runs have been repeated entirely for 0% and 100% TKE, when a few weeks would have been sufficient. Of the 6 figures in the paper that describe modeling results (Figures 8-13), all of them are doubled to show 0% and 100% TKE. This would be understandable if TKE had a large impact on power output, but it did not (at most 5%).

Given the extensive discussion of the value of added TKE in the literature, we wanted to thoroughly document its variability. And, as pointed out above, 0% is more similar to Volker et al. and the 0-100% range therefore includes not just the Archer et al. suggestion but Volker et al. as well.

On the other hand, the team did not perform a year run with the recommended value of 25% TKE.

As noted above, the recommendation of 25% comes from one study, and we sought to provide more extensive bounds on the variability that could be introduced with a range of added TKE values. Archer et al. (2020) recommended the use of 25% TKE, and while extremely helpful to pioneer a suggestion for this issue, that recommendation is unfortunately limited in application for being run under idealized conditions, with neutral stratification, and for a one-wind-turbine setup. For this reason, Archer et al. reported that 25% TKE was the best choice for their setup, and further investigation is still required. There is uncertainty on what the "rule of thumb" TKE amount should be in regional wind plant modeling, and our results, because we explore the whole range of possibilities from 0% to 100%, provides a useful contribution by quantifying the (small) size of the impact of the TKE term.

A few weeks of simulation time may have been sufficient for future model development choices. However, the goal of our report was to provide the first year-long assessment of wake effects on power production, which is a highly sought-after dataset for industry partners and stakeholders.

Thus in principle every value that they report in the abstract should be a range, but it is not. Plus, no TKE results are shown (I would like to see the equivalent of Figure 8 but for added TKE). In summary, this study focuses excessively and at the same not enough on TKE. An obvious recommendation would be to ask the team to conduct a new one-year run with 25% TKE and rewrite the paper to focus on those results and reduce the sensitivity analysis. But I think that this would be an excessive request, plus there is already a lot of value in the current runs.

We appreciate that the reviewer recognizes that another set of 25% TKE simulations is computationally infeasible. Due to computational limitations, we cannot run an additional year-long simulation.

My first recommendation is therefore that the results in the abstract and conclusions, which do not report a range (i.e., the range of results with 0% and 100% TKE) but are presented without explanation as one value (see comment below), be modified by either reporting always the range, or by using an interpolation based on the few days of 25% TKE results that the authors have already run (Fig. A1).

We have modified the abstract as follows to incorporate the range of values:

> "Using a series of simulations with no wind plants, one wind plant, and complete build-out of lease areas, we calculate wake effects and distinguish the effect of wakes generated internally within one plant from those generated externally between plants. **We also vary the amount of added turbulence kinetic energy (TKE) between 0% and 100% to provide some uncertainty quantification.** The strongest wakes, propagating 55 km, occur in summertime stable stratification, just when New England's grid demand peaks in summer. The seasonal variability of wakes in this offshore region is much stronger than diurnal variability of wakes. Overall, year-long wake impacts reduce power output **by a range between 38.2% and 34.1% (for 0%-100% added TKE)**. Internal wakes cause greater year-long power losses, **from 29.2% to 25.7%,** compared to external wakes, **from 14.7% to 13.4%**. The overall impact is different from the linear sum of internal wakes and external wakes due to non-linear processes. Additional simulations quantify wake uncertainty by modifying the added amount of turbulent kinetic energy from wind turbines, introducing power output variability of 3.8%. Finally, we compare annual energy production to New England grid demand and find that the lease areas can supply **58.8% to 61.2%** of annual load."

Further, the conclusions are modified similarly:

- We now report "The average yearlong power deficits at Vineyard Wind considering internal wakes and external wakes from the LA range between 38.2% (TKE_0) and 34.1% (TKE_100)."
- Text is rewritten to include the range by "Yearly averaged wake losses induce power deficits at Vineyard Wind from 38.2% (TKE_0) to 34.1% (TKE_100)".

What I mean is that the team could obtain a relationship between average power output (or wind speed deficit or whatever the parameter of interest is) with 0%, 25%, and 100% TKE from the few simulated days. This relationship does not seem to be linear from Fig. A1. An example of this relationship might be something like: the power at 25% TKE is the mean of that at 0% and 100% TKE, on average. Then use that relationship to report one value (per parameter) in the abstract and conclusions, that "fitted" to 25% TKE.

As the reviewer has pointed out, the relationship between parameters and the amount of added TKE is a nonlinear relationship and so we have chosen to provide the range of values as above.

The second recommendation is that the authors add a figure and discussion on the TKE distribution in the wakes with 0% and 100% TKE, like Fig. 8.

We have added a section for the results and discussion of TKE at the hub height in new Appendix E, similar to Figure 8 (new Figure 11).

2. **The calculation of the losses from external wakes may be incorrect**

From the abstract, the effects of internal wakes are reported to be -27.4% and the effects of external wakes are -14.1%. The sum of the two is -41.5%. However, the combined effect is reported to be -35.9%. This is problematic. At first sight, this discrepancy may be the result of the non-linearity of the wake processes. If so, all the authors need to change is to reverse the order of two sentences and add a few words in the abstract to explain it: "Internal wakes alone cause greater year-long power losses (27.1%) compared to external wakes (14.1%). When both are present, however, the mean year-long wake impacts reduce power output by 35.9%, which is lower than the sum of the two due to non-linear processes."

Yes, we have noticed and discussed this nonlinear behavior, and have added a sentence to the abstract to explicitly note this behavior.

However, I suspect that there might be a design issue in the way the power losses are calculated in Eq. (9) and (10). Aside from the unclear notation (see comment below), the denominator of the two equations is not the same and that may be why the discrepancy arises. Eq. (10) is correct because there is no double counting: there are no losses in the denominator and the internal losses are only in the numerator. In Eq. (9), however, there are internal losses in both the numerator and the denominator, and they are not equal. The internal losses are not equal in the VW and CA cases because, as upstream conditions change due to external wakes in the CA case, the internal wakes change too and therefore the internal wake losses do not "cancel out", there is still some influence from the internal wake losses. As such, the ratio in Eq. (9) does not quantify just external losses because it still contains the effect of internal losses; it quantifies a mix of internal and external losses.
I suggest that the authors report Eq. 10 first (Loss_internal). Then, they should replace P_VW at the denominator with P_NWF in Eq. 9, to obtain the total effect from internal and external wakes due to the CA areas (call it Loss_total). We know this value: it should be -35.9% (from l. 384). The effect of the external wakes then is the difference between the total losses and the value from Loss_internal:
Loss_external = Loss_total – Loss_internal = -35.9% - (-27.4%) = -8.5% (Eq. 11)
This way the denominator is the same and the individual values for external and internal sum up to the correct total.

Thank you for this suggestion. There are several different methods for calculating the wake impact, and we have supplemented an additional method for calculating external losses as the difference between the total and internal losses, via a new equation (11).

Power losses from external, internal, and the total wake effects are calculated from:

$$Loss_{external} = 100 - \left(\frac{P_{LA,CA}}{P_{VW}}\right) * 100\% \quad (9)$$

$$Loss_{internal} = 100 - \left(\frac{P_{VW}}{P_{NWF}}\right) * 100\% \quad (10)$$

$$LOSS_{total} = 100 - \left(\frac{P_{LA,CA}}{P_{NWF}}\right) * 100\%\,, \quad (11)$$

where $P_{LA,CA}$ is the power production at Vineyard Wind grid cells in the presence of wakes by either the LA or the CA, P_{VW} is the power production in the presence of internal wakes from VW, and P_{NWF} is the power production from coupling hub-height wind speeds to the power curve. These methods are performed separately by added TKE amount. **We note that the upwind conditions change in a LA or CA scenario, due to external wakes, which can modify the internal losses in the numerator of Eq. 9. Thus, we provide an alternative method for calculating the external power losses as the difference between the total losses and the internal losses:**

$$\boldsymbol{LOSS_{external} = LOSS_{total} - LOSS_{internal}} \quad \mathbf{(12)}$$

3. **The stability classification is not adequate.**

The authors use a very simple classification for stability based on the value of L (Eq. 8). Neutral conditions are those with abs(L)>1000 m. This is inconsistent with the published literature, e.g., Gryning et al. (2007) and Sathe et al. (2011) used 500 m, Wharton and Lundquist (2012) used 600 m, Rajewski et al. (2013) used 400 m, Archer et al. (2016) used 500 m. In fact, too few neutral cases were found here, less than 2.5% of the time (p. 29 l. 595). I am unsure what to recommend here because there is not an "accepted" value of L for neutral conditions, but the authors need to assess the sensitivity of their results to a few values, at a minimum 500 m. This could possibly help with the previous inconsistencies in the areal extent and wake length, as days that were actually neutrally stratified may have been mixed in with days with other stabilities to obfuscate some of the relationships.

As the reviewer acknowledges, there is a wide range of thresholds that have been used to determine stability regimes and there is not an accepted value of L to demarcate the line between neutral and stable or unstable conditions. The threshold of 1000 m is consistent with the published literature as Muñoz-Esparza et al. (2012) use this cutoff for neutral conditions in the offshore environment. (Most of the references cited by the reviewer were for onshore conditions). Our finding that neutral stratification occurs 2.5% of the time is only for the CA simulations, which is a subset of August-September of 2019 and June-July of 2020. For the yearlong period, our original reported number is double this value, at 4.48% of the time.

Additionally, through discussion with other WRF modelers, we learned that the WRF-output Obukhov Length (which we were using in the original calculations) is not accurate because it is calculated in the timestep before the heat flux is calculated. We have recalculated the Obukhov length directly using model-output variables at the same location (all figures and calculations incorporating stratification have been updated). The new percentages of occurrence for unstable, stable, and neutral conditions using a 1000-m cutoff are 48.4%, 46.3%, and 5.2%, respectively (originally 53.6%, 41.9%, and 4.5%,). Using a 500 m threshold, these percentages change to 44.3%, 44.4%, and 11.2%. We choose to maintain the |L|=1000 threshold because that is consistent with offshore work (Muñoz-Esparza et al., 2012).

Minor Issues

4. **The title needs improvements**

The title suggest that the wake impacts "on" the wind farm development will be studied. This is somewhat inaccurate, as the study is about the wake impacts on offshore wind power production or output, not on the development. Development is choosing the number of turbines or their specs or their layout, which are all fixed in this study; or, development can be how the wind farm installations grow/change with time. Either way, the development here is given (3 scenarios), what changes is the power output.
Also, the title mentions the "Mid-Atlantic" as the focus area, but technically speaking the Mid-Atlantic stops as far north as New York state. From the U.S. perspective, the Vineyard Wind project is not in the Mid-Atlantic and neither are the northeastern lease areas of RI or MA. According to Wikipedia, the following states are included in the Mid-Atlantic: Delaware, Maryland, New Jersey, New York, Pennsylvania, Virginia, West Virginia, and Washington DC. To non-U.S. readers, "Mid-Atlantic" could be the Equatorial zone, as the Atlantic Ocean extends between the two Poles. I don't have a good recommendation for an alternative, but perhaps "U.S." should be added in the title because the study focuses on the U.S. offshore areas after all.
Last, "annual" variability suggests that many years were studied to understand how the production changes from one year to the next. Instead, only one year was simulated here. Thus the variability studied here is monthly/seasonal and diurnal, but not annual.

We keep the siting and characterizations of wind turbines constant in our work and agree that the main focus is on power production. However, our use of Mid-Atlantic is consistent with the Bureau of Ocean and Energy Management terminology. Thus, we will not change this nomenclature. We agree that annual variability implies studying multiple years, and that we should clarify the U.S. focus. We have changed the title to "Seasonal Variability of Wake Impacts on U.S. Mid-Atlantic Offshore Wind Plant Power Production".

5. **Simplify naming**

There is no need to add "_only" to the name of the run with only the Vineyard Wind farm. Just call it "VW."

All instances of "VW_only" have been changed to "VW".

6. **Unclear notation in Eqs. (9)-(10)**

These equations have already been discussed at comment #2, here I am focusing on the notation only. Eliminating "_only" will help (comment #5). P_WV_waked is not defined and uses a notation that differs from that of all other subscripts. All the other subscripts refer to a specific run, whereas "waked" refers to, I believe, a subset of grid points. But the same subset of grid points was used for all other denominators and numerators, thus the confusion. Plus the term P_VW_waked refers to run CA, I believe. I suggest something like (not including my recommendation from comment #2 above):

$$Loss_{ext} = \left(1 - \frac{P_{CA}}{P_{VW}}\right) * 100\%$$

In the text below the equation then you specify that this equation is obtained from the grid cells over Vineyard Wind.

We have changed the notation to reflect the simulation type such that $P_{LA,CA}$ refers to the power production at Vineyard Wind grid cells when exposed to external and internal wakes by either the lease or call areas, to reduce redundancy of writing the same equation twice. P_{VW} refers to power production at Vineyard Wind grid cells in the presence of internal wakes in a VW simulation. P_{NWF} represents power production at Vineyard Wind grid cells from coupling NWF wind speeds to the power curve.

7. **L. 318-320 ("While here ... schemes"): this discussion is irrelevant and unnecessary here.**

This sentence is also redundant and has been removed.

8. **P. 17: some of these results are rather counter-intuitive**

if the TKE_100 runs produce weaker deficits and smaller wake areas, then the wakes should be shorter, whereas the authors report 58 km for TKE_100 and 55 km for TKE_0. The explanation provided is vague and unsupported (l. 346: "larger reduction in momentum aloft"??). The authors do not report exactly how the wake length was obtained. I suspect the method was somewhat empirical and in fact it is giving counter-intuitive results. I suggest that either the authors develop an objective and automated method for calculating the wake length and, if the inconsistency persists, they document and explain it; or that they remove any discussion of the wake length.

We appreciate that, at first, this finding may seem counterintuitive. However, turbulence from the turbines enhances vertical momentum transport from aloft down to within the wake (Gupta and Baidya Roy 2021). The enhanced TKE in a TKE_100 simulation transports more momentum into the waked zone, leaving slower wind speeds above the wind plant (Fitch et al. 2012; Siedersleben et al. 2020). Reduced wind speeds above the turbines then offer a weaker reservoir of momentum available for wake recovery further downwind, leading to slightly longer wake propagation distances. References have been added to support this argument. We explain how the wake length was obtained in Section 2.8 "Wake Identification".

9. **Improve Fig. A1.**

Replace 10-m wind speed with 140-m wind speed in Fig. A1. In all panels (except c), add the results from NWF to appreciate the magnitude of the impacts.

The wind speeds shown in Fig. A1 are indeed hub height wind speed. An error in the y-label lead to a cascading effect in the text which has been fixed. NWF values have been added to the figure.

10. P. 26-27: the discussion is unclear, the authors report "reductions" in several sentences, but it is unclear what is changing and what the reference is: are they discussing changes from TKE_0 to TKE_100 or from TKE_100 to TKE_0 or from NWF?

- We now clarify "The wind speed reduction **during this time period** causes a corresponding decrease in turbulent transport of moisture."
- We now clarify "The reduction in heat flux **during this time period** causes 2 m temperatures to decrease and exhibit less variability by TKE amount, with a mean difference of 0.26 K between TKE_100 and TKE_0 (Fig. A1f)."

For the final occurrence of "reduction", it is obvious to the reader that we are referring to the same time period at this point, so to not add distracting information to a topic sentence, we leave this sentence as is: "The reduction in turbulent mixing lowers the PBL, regardless of TKE amount, to shallow heights between 30 to 80 m at 13:00 UTC (Fig. A1e)."

10. Fig. D1: need a legend for the colors. Also, are these wind speeds or wind speed deficits? The caption indicates wind speed.

We have added a color bar to Figure D1, included a color bar title delineating that the contours represent the wind speed deficit, and specified the "wake wind speed deficit" in the caption.

References

Archer, C. L., S. Wu, Y. Ma, and P. A. Jiménez, 2020: Two Corrections for Turbulent Kinetic Energy Generated by Wind Farms in the WRF Model. *Monthly Weather Review*, **148**, 4823–4835, https://doi.org/10.1175/MWR-D-20-0097.1.

Fitch, A. C., J. B. Olson, J. K. Lundquist, J. Dudhia, A. K. Gupta, J. Michalakes, and I. Barstad, 2012: Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model. *Monthly Weather Review*, **140**, 3017–3038, https://doi.org/10.1175/MWR-D-11-00352.1.

Gupta, T., and S. Baidya Roy, 2021: Recovery Processes in a Large Offshore Wind Farm. *Wind Energy Science Discussions*, 1–23, https://doi.org/10.5194/wes-2021-7.

Muñoz-Esparza, D., B. Cañadillas, T. Neumann, and J. van Beeck, 2012: Turbulent fluxes, stability and shear in the offshore environment: Mesoscale modelling and field observations at FINO1. *Journal of Renewable and Sustainable Energy*, **4**, 063136, https://doi.org/10.1063/1.4769201.

Siedersleben, S. K., and Coauthors, 2020: Turbulent kinetic energy over large offshore wind farms observed and simulated by the mesoscale model WRF (3.8.1). *Geoscientific Model Development*, **13**, 249–268, https://doi.org/10.5194/gmd-13-249-2020.