Authors' response to Reviewer 1 comments

Overall, this manuscript is well-written and supported by an appropriate number of references, which demonstrates the authors' thorough engagement with the relevant literature. The use of a GPU-based WRF model is an excellent choice, and the configuration decisions—such as setting the TKE factor to 1 and employing a 1 km mesh resolution at the wind farm—are well-justified and technically sound.

I particularly appreciate the authors' decision to focus on power loss rather than velocity deficit maps, as this approach provides a more meaningful metric for assessing wind farm performance. The AEP loss map presented in Figure 7 is especially informative and visually compelling.

The analysis showing that large velocity deficits at wind speeds above the turbine's rated speed (11 m/s) have a limited impact on overall wind farm production is interesting. While this may not represent a groundbreaking finding, it is still valuable to publish, given the historical reliance on velocity deficit maps in WRF parameterization studies. This contribution helps clarify the limitations of traditional approaches and reinforces the importance of using power-based metrics.

We greatly appreciate the reviewer's positive feedback. We also carefully considered your comments regarding wind farm parameterization and sensitivity testing. Several of these suggestions would require conducting additional simulations. Unfortunately, given our remaining funding and computational resources, conducting all additional experiments is beyond our current capabilities but we did further examine the question regarding the stability classification raised by you and another reviewers by conducting an additional 1-year WRF simulation.

Overall, we still did our best in trying to address most of your comments and we believe this revise manuscript is stronger and more concise than the previous version.

Major comments

1- One of the most relevant findings in the manuscript relates to stability differences between the surface and the rotor area. This is an important aspect of wake modeling, and I encourage the authors to expand this analysis. Specifically, could you include a comparison with 1/L at the surface, which is a standard WRF output and widely used in the literature? Additionally, it would be helpful to investigate whether there is a larger mismatch between Ri at the surface and at rotor height near the shoreline. Please also clarify in the text how stability classes were defined using Ri values, as this is currently only indicated in Figure 4 and Table 3.

Thank you for your comment. Reviewer #3 also raised a similar question. Based on your comment, we have conducted additional simulations to extract the Obukhov length (L) to look into this as we did not output the Obukhov length in our initial study. Figure R1 compares h/L (with $h = 10$ m) against Ri calculated for both the near-surface layer (20–50 m) and the rotor layer (20–300 m) at locations within the wind lease region from the newly-conducted simulation. Over the simulated 1-year period, both stability metrics (L and Ri) agree most of the time, and their agreement improves when Ri is evaluated close to the surface (20–50 m). However, even in that case, approximately 14% of the time, L indicates unstable conditions while the near-surface Ri indicates stable conditions. This discrepancy increases to 28% when compared with rotor-layer Ri. Conversely, there are very few data points where L indicates stable conditions and Ri indicates unstable conditions. These results (in particular, the increase from 14% of points in disagreement to 28% of points in disagreement) demonstrate that surface-based stability (L) does not reliably represent stability in the rotor layer. Because detailed comparison of Richardson number and Obukhov length is not the foucs the scope of this paper, we did not include the corresponding figure in the revised manuscript. However, we have added a discussion on the differences between L and Ri to justify our choice of stability metric.

Lines 252-259: *To examine the difference in stability indication between Obukhov length, L, and Bulk Ri, similar analysis was conducted by comparing h/L (with $h = 10$ m) against Ri calculated for both the near-surface layer (20–50 m) and the rotor layer (20–300 m) over the wind lease region. The results suggested that approximately 14% of the time, h/L indicates unstable conditions while the near-surface Ri indicates stable conditions. The discrepancy increases to 28% when compared with rotor-layer Ri. These results further suggest that surface-based stability does not reliably represent stability in the rotor layer. Therefore, for the remainder of the paper, we will use the rotor-layer Ri for further analysis, as it is more appropriate for examining the effects of stability on the wind turbines. However, we note that defining stability in this manner has limitations, particularly when strong local inhomogeneity exists within the rotor layer (e.g., coastal low-level jets).*

Regarding your comments on the difference between Ri at the surface and at rotor height near the shoreline, we conducted the requested analysis and the results are shown in Figure R2. Eight points (colored in red; Figure R2a) are randomly selected to sample the nearshore region adjacent to the wind farm lease area. All points are approximately 5 km away from the shoreline. A comparison between Figure R2b) and Figure 5 in the manuscript shows clear differences: as we move closer to the shoreline, the agreement in stability classification between the surface layer and the rotor layer decreases from 87% to 70%. In addition, the occurrence of cases where the near-surface layer is unstable while the rotor layer remains stable approximately doubles, from 13% to 28%. These results further support our conclusion that surface-based stability metrics do not reliably represent stability in the rotor layer, and that the discrepancy becomes more pronounced nearshore. Although we did not include Figure R2 in the revised manuscript, we have added corresponding text in the discussion to address this point.

Lines 247-251: *Similar analysis is also conducted near the shoreline and the agreement in stability classification between the surface layer and the rotor layer decreases from 86% to 70%. In addition, the occurrence of cases where the near-surface layer is unstable while the rotor layer is stable increases from 13% to 28%. These results demonstrate, as also indicated in Rosencrans et al. (2024), that near-surface stability is not always representative of the deeper rotor layer and the discrepancy is more pronounced nearshore compared to offshore.*

In the revised manuscript, we have added specific text regarding how stability classes were defined using Ri values (Line 240: *In this study, only stable ($Ri > 0$) and unstable ($Ri < 0$) conditions are considered*).

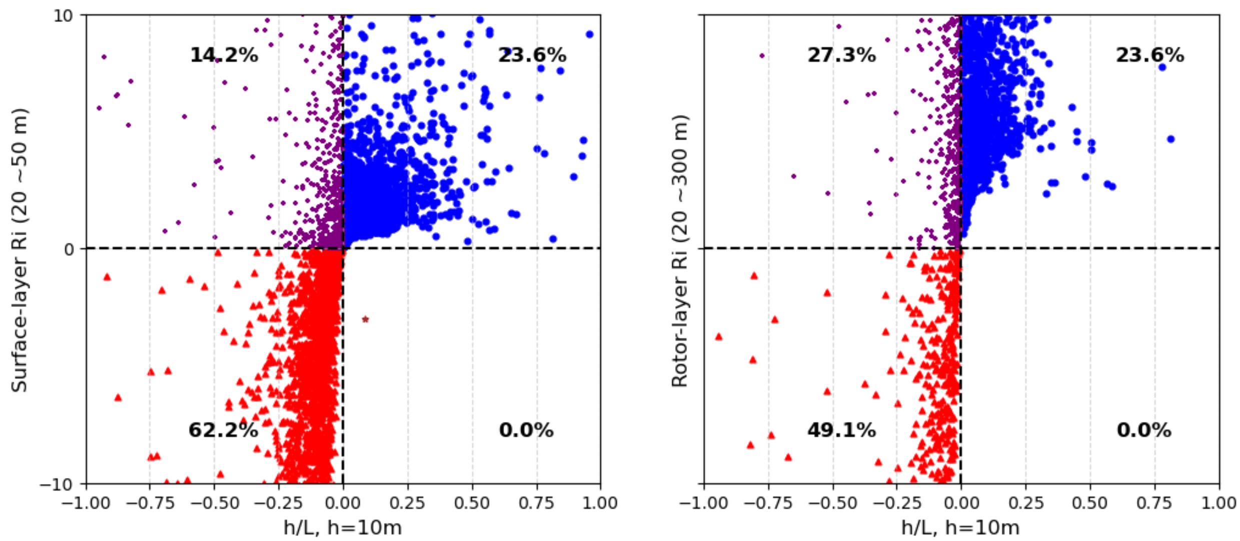

Figure R1: **Comparison of h/L (with h = 10 m) and Richardson number calculated for both the near-surface layer (20–50 m; left) and the rotor layer (20–300 m; right) over the wind lease region from the newly-conducted simulation.**

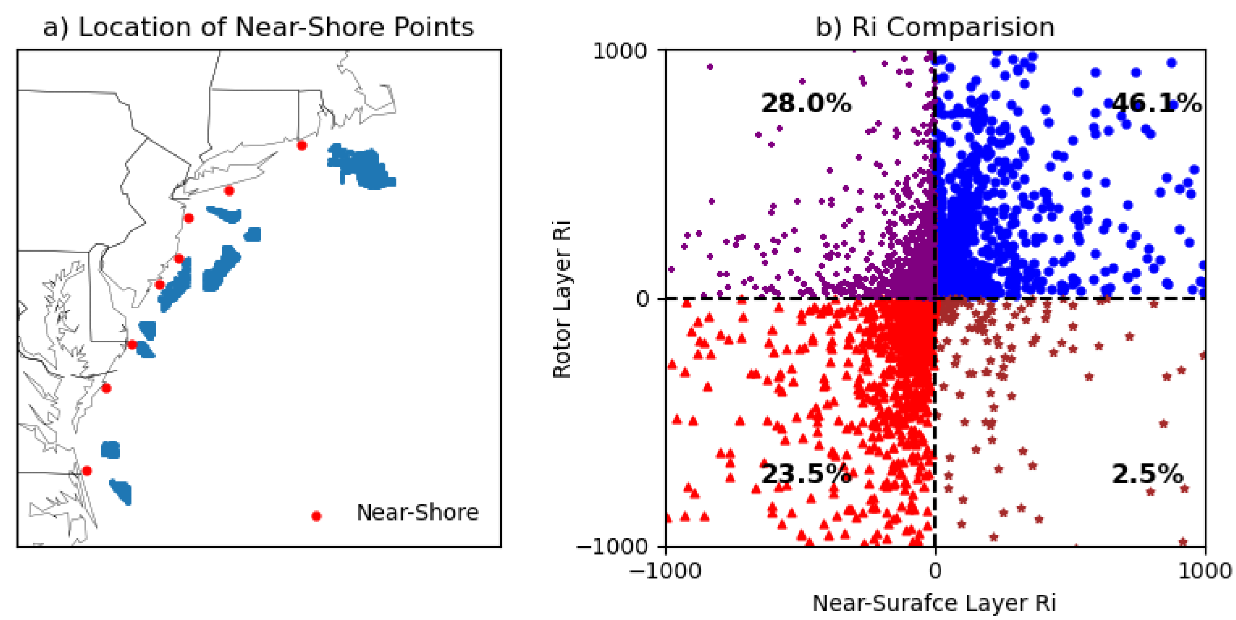

Figure R2: **Difference in near-shore Ri between the near-surface layer (20–50 m) and rotor layer (20–300 m)**

2- The configuration of WRF and the chosen parameterization is appropriate and well-executed. However, the study currently relies on a single model and configuration, which limits its robustness and generalizability. To strengthen the work, I recommend including a comparison with another widely used industry model, such as the Turbopark wake model (for at least one of the wind farms) or the Volker parameterization. This would provide additional context and enhance the credibility of the conclusions.

Thank you for your comment. We agree that adding comparisons with an industry model or an alternative parameterization would further strengthen the robustness of our conclusions. However, as mentioned in the overall response, such efforts are beyond our current resource limitations. Nevertheless, we did conduct an additional simulation and new analysis comparing the Obukhov length with the Richardson number, as this was a common request from both you and another reviewer. In addition, we have added a section on future work to address the uncertainties raised by the reviewer.

Lines 441-449: *There are several uncertainties associated with this study that motivate for future work. First, it would be valuable to apply the same analysis framework using an industry wake model (e.g., the TurboPark model) to assess whether the differences in wake map between wind speed deficit and energy loss approach remain consistent. Second, additional sensitivity studies related to different wind farm parameterization and value of added TKE are needed to quantify the robustness of our results. Finally, a comparison with recently published studies on offshore wind wake assessments demonstrates the complex interplay of multiple factors in determining the final wake area, including the (1) turbine design and layout, (2) simulation techniques, and (3) analysis methods and metrics employed. The variability in the results arising from these factors highlights the need for further research to standardize wake assessment methodologies and develop more robust, universally applicable wake characterization techniques.*

3- Regarding the contour maps presented in Figure 8, which are among the most impactful results for decision-making on wind farm siting: Is it necessary to simulate a full year to generate these maps? A sensitivity analysis using a shorter simulation period would be highly valuable. Furthermore, if a new wind farm were to be added, would it be necessary to rerun the entire cluster simulation, or could the new farm's results be combined with the existing data? Addressing these questions would significantly improve the practical applicability of the study.

Thank you for your comment. For studies involving wake-loss assessment, it is generally necessary to conduct long-duration simulations because accurate characterization of the regional wind statistics (e.g., the wind speed probability density function) is essential. While it is possible in some cases to use a shorter simulation period, that period must be carefully selected to be representative of the regional wind climate. A classic example is Pryor et al. (2021), who employed eleven five-day representative wind-flow scenarios. At the start of this project, we did look into the possibility of running a subset of days that could be representative of the year—however, we found that many days were needed for the probability distributions to converge, and the computational savings were limited. Due to this, we elected to simply run the full TMY.

Lines 110-112: *We considered running only a subset of days to minimize computation time, but found that we could not identify a subset that both suitably captured TMY conditions and significantly reduced the needed computation time, especially when using the GPU-based WRF, so we elected to run an entire year*

If a new wind farm were to be added, it would be necessary to rerun the full simulation for the entire wind-farm cluster. This is because wakes generated by neighboring wind farms can influence the target wind farm. It is also important to perform a simulation with only the target wind farm included. By comparing these two simulations, one can quantify wake losses associated with both internal and cluster wake effect for the new lease area.

Minor comments

1- The improved parameterization proposed by Vollmer et al. (2024) could potentially influence the conclusions of this work. Would it be feasible to conduct a limited test to confirm that the main findings remain consistent under this updated approach?

Thank your for your comment. While we agreed this additional work would strengthen the conclusions of this paper, such effort is beyond our current capabilities. However, this is a potential direction we would like to explore in future work.

Lines 443-444: *Second, additional sensitivity studies related to different wind farm parameterization and value of added TKE are needed to quantify the robustness of our results.*

2- Additionally, please elaborate on the rationale for selecting the three variables used in the TMY methodology. Would including additional variables such as TKE or stability indicators (Ri or 1/L) improve the representativeness of the dataset? Including a standard wind rose plot for the TMY dataset would also enhance the clarity of the analysis.

Thank you for your comment. The 100-m wind speed and direction were selected to provide an accurate representation of wind conditions at turbine-relevant heights from the TMY dataset, as this is critical for the wake loss analysis. The 2-m temperature was included to capture typical surface-level weather conditions. However, we do agree with reviewers that incorporating stability indicators would further improve the representative of the dataset, and correponding text has been added to address this limition. Regarding the wind rose, we did not include it in this revision because the probability distribution comparison between the TMY and ERA5 datasets serves a similar purpose. This comparison demonstrates that the long-term wind climate is well captured in our simulation.

Line129-131:*However, we acknowledge that the current TMY construction is not perfect. Its representation could be further improved by incorporating additional metrics such as ABL height and stability indicators (e.g., air-sea temperature difference, lapse rate)*

3- It would be useful to verify whether a single turbine in the domain reproduces the input power curve when using the parameterization, particularly under stable conditions where low-level jets affect the rotor area. This would help clarify whether some of the observed losses are due to wake effects or to the parameterization's response to vertical wind profiles.

Thank you for your comment. In the current implementation of the Fitch parameterization in WRF, the power output of each $n-$th turbine within the i, j grid cell is only a function of the hub-height wind speed and not the wind profile (Fitch et al., 2012; Skamarock et al., 2019). The power P_n of each turbine in WRF is estimated as follows:

$$P_n(i,j) = \frac{1}{2}\rho U_h^3 \pi (D/2)^2 C_P,$$

where U_h is the hub-height wind speed at the (i, j) grid cell, D is the turbine's rotor diameter, C_P is the power coefficient for U_h, and the air density remains fixed at $\rho = 1.23$ kg m^{-3}. Note that in WRF, the power curve is used to estimate the power coefficient as a function of wind speed, which is then interpolated based on U_h, and then the turbine's power is estimated following the equation above. Since the power curve is used as an input to the Fitch parameterization to translate wind speed into power, any turbine grid cell in the domain should reproduce the prescribed power curve, regardless of the conditions it is in. The power loss map presented in the manuscript corresponds to a single vertical level (150 m), and the associated power losses shown are due to the reduction in wind speed caused by the total wake effect—both internal and cluster—captured by the parameterization. Redfern et al. (2019) incorporated the rotor-equivalent wind speed into the Fitch parameterization, showing small improvements in power predictions compared to using only hub-height wind speed. Therefore, including the rotor-equivalent wind speed in our simulations is also expected to show margnial improvements.

4- In Figure 6, consider presenting energy loss aggregated at the wind farm level, in addition to the 1 km grid cell resolution. Furthermore, could you explain the energy loss observed far from the wind farms in this figure? Is this related to the well-known WRF parameterization numerical errors in stormy conditions?

Thank you for your comment. In the earlier submission, we did mention the energy loss at the wind-farm level in the manuscript (Line 276: "Under this scenario, power generation for wind farms like Community Wind, Attentive Energy, and Leading Light Wind could experience a greater than 30% reduction in power output."). However, because this information is sensitive and derived from model simulations, we only discussed it briefly. In the revised manuscript, to avoid confusion and at the request of Reviewer 2, we have removed this sentence. However, we have now added a section 3.3 comparing internal and cluster wake losses for one farm (Atlantic Shores South) that specifies the total energy losses for this farm in the presence of wakes.

Regarding the energy loss simulated far from the wind farms (Figure 6b), this feature is mostly due to numerical noise or errors from WRF. Similar behavior is not uncommon in WRF, and Rosencrans et al. (2024) discusses this in detail in Appendix F. In this revision, we have added a sentence to directly point the reader to that paper for further information.

Lines 330-333: *Regarding the energy loss simulated far from the wind farms (Figure 6b), this feature is mostly due to numerical noise or errors from WRF which has been explicitly discussed by Rosencrans et al. (2024) in Appendix F. We encourage the readers to refer to that study for more details.*

5- Finally, for context, are velocity deficit and power loss maps actually used by government agencies for planning purposes? If so, please provide a reference where this application is documented, as this would strengthen the practical relevance of the study.

Yes, velocity deficit and power loss maps are used by U.S. government agencies like the Bureau of Ocean Energy Management (BOEM) and the Department of Energy (DOE) for wind energy planning (Musial et al., 2013a,b). However, we note that the U.S marine spatial planning is a thorough process managed by BOEM, which is responsible for overseeing renewable energy development on the outer continental shelf. As a result, any assessment, including this study, may not fully reflect the actual availability of resources.

Lines 71-73: *We note that the U.S marine spatial planning is a thorough process managed by the Bureau of Ocean Energy Management (BOEM), which is responsible for overseeing renewable energy development on the outer continental shelf. As a result, any assessment, including this* *study, may not fully reflect the actual availability of resources (Musial et al., 2013a,b).*

References

Fitch, A. C., Olson, J. B., Lundquist, J. K., Dudhia, J., Gupta, A. K., Michalakes, J., and Barstad, I.: Local and mesoscale impacts of wind farms as parameterized in a mesoscale NWP model, Monthly Weather Review, 140, 3017–3038, 2012.

Musial, W., Elliott, D., Fields, J., Parker, Z., Scott, G., and Draxl, C.: Assessment of Offshore Wind Energy Leasing Areas for the BOEM Maryland Wind Energy Area, Tech. rep., National Renewable Energy Laboratory (NREL), Golden, CO (United States), 2013a.

Musial, W., Elliott, D., Fields, J., Parker, Z., Scott, G., and Draxl, C.: Assessment of offshore wind energy leasing areas for the BOEM New Jersey wind energy area, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2013b.

Pryor, S. C., Barthelmie, R. J., and Shepherd, T. J.: Wind power production from very large offshore wind farms, Joule, 5, 2663–2686, 2021.

Redfern, S., Olson, J. B., Lundquist, J. K., and Clack, C. T. M.: Incorporation of the Rotor-Equivalent Wind Speed into the Weather Research and Forecasting Model's Wind Farm Parameterization, Monthly Weather Review, 147, 1029–1046, https://doi.org/10.1175/MWR-D-18-0194.1, 2019.

Rosencrans, D., Lundquist, J. K., Optis, M., Rybchuk, A., Bodini, N., and Rossol, M.: Seasonal variability of wake impacts on US mid-Atlantic offshore wind plant power production, Wind Energy Science, 9, 555–583, 2024.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., et al.: A description of the advanced research WRF version 4, NCAR tech. note ncar/tn-556+ str, 145, 2019.