Reviewer 2

The manuscript describes a case study of interactions between internal gravity waves (a bore initiated by nearby convection) and wind turbines in a wind farm, in Oklahoma. The case study uses a two-domain nested simulation, with the horizontal resolution down to 20m in the nested domain. Although the simulations present several difficulties (sensitivity of the convection and ensuing bore to the microphysics scheme, fine-scale processes, eg wakes, to include in the wind farm) the authors show convincing simulations, supported by the careful comparison to the observations. Publication after minor revisions is recommended.

We thank the reviewer for the time they spent reviewing our work and for their supportive suggestions and comments. Our replies are inserted below into the reviewer's comments.

Minor comments:

l34: there is a case study by Ralph, Neiman and Keller, 1999, worth mentioning with respect to fronts generating gravity waves near the surface

Thank you for bringing this study to our attention. It has now been included and cited in the introduction.

l39-49: repetition: '... in the US Great Plains. In the US Great Plains...'

The second "in the US Great Plains" has been removed.

l63: affects -> affect

Thank you, this has been corrected.

l124: it would be useful to indicate the local time already at ths stage of the manuscript; it is done later on of course, but the reader wonders when first encountering the time...

The local time is now stated at this point.

l146: the three numbers are odd, for the Obukhov length: the unit is not given, and two numbers are negative...

The text now includes the units, which are in meters. The negative Obukhov lengths represent unstable atmospheric conditions as the heat flux close to the surface is positive (upwards) in the following expression:

$$L = -\frac{u_*^3 \theta}{\kappa g \overline{w'\theta'}},$$

Where u_* is the friction velocity, θ is the potential temperature, κ is the von Karman constant, g is the gravitational constant, and w is the vertical velocity. $\overline{w'\theta'}$ represents the heat flux close to the surface. The fluxes are calculated using 30-minute averages to calculate the perturbations (line 117 in the manuscript).

l156-157: what is meant by 'or in the operational mode..'?

This sentence has been clarified to state that the operational mode is when turbines are operating with power-maximizing control:

> "Additionally, the hub-height wind speeds are relatively low, on the order of 4-6 m/s such that the power output of the farm is highly susceptible to wind speed fluctuations in this range (in region 2 of the power curve or in the operational power-maximizing control mode of the wind turbines)."

l198-199: the formulation is a bit surprising: 'the setup is unique (...) only two domains.' This is presented rather as a strength of the study, but using only two domains implies a strong ratio between the resolutions used in one and the other domain, which is generally considered unfavorable...

We agree that using a large ratio between the resolutions would traditionally be considered as non-standard. A major strength of using this setup is that it is computationally efficient, which has now been stated in line 203:

> "The setup is unique compared to other multi-scale WRF setups in that it uses only two domains, which makes it very computationally efficient."

The case study is unique in that there are two very clear dominant scales in the flow, which makes a two-domain setup feasible. The first are the scales related to the gravity waves themselves and the second are scales related to ambient small-scale turbulence in stable atmospheric conditions. Through sensitivity studies conducted both by us and Johnson et al. (2019), a horizontal grid spacing of 300 m resolves the relevant and important scales related to the gravity waves. The domain with 20 m grid spacing resolves the ambient small-scale turbulence (finer resolution would be desirable but we are limited by computational resources).

Domains between 300 m and 20 m grid spacings represent a gray zone in stable conditions where turbulence is either not resolved at all or not resolved accurately. It is possible that adding another domain could improve the results; however, this is not guaranteed due to the uncertainties of turbulence modeling in the gray zone (e.g., Chow et al. 2019). Given that our 20 m domain is large enough to include adequate fetch, while also including perturbations to encourage the development of turbulence, the benefit of an additional domain would likely be minimal, with a significant penalty in terms of computational cost.

Lastly, and most importantly, we achieved good results with this nesting configuration. The simulations results compared favorably with observations compared to other nesting configurations done in an initial sensitivity study.

l215: the formuation is ambiguous: "0 hour forecast (which includes the hour..." If data has been assimilated up to that point, it is odd to call this a forecast.. this should be reformulated

We agree that this formulation is unclear therefore it has been removed. The sentence at lines 218-219 in the revised manuscript now reads:

> "In this study, we use the HRRRv4 analysis product which includes assimilated NEXRAD data."

figure 4: it is somewhat disturbing to use the same colormap for the two panels, but with different ranges

The caption now explicitly states that the colormaps have different ranges due to the maximum and minimum elevations within each domain.

figure 12: plural: 'power output[s]'

This has been fixed.

l455: repetition of 'results'; vary the choice of words: simulations? flow variables?

This has been changed to "simulations".

l463: 'For the 20 m domain... ' formulation can be improved...

This beginning of this sentence now reads:

> "For the fine-scale domain with 20 m horizontal grid spacing and the wind turbines parameterized, …"

figure 14: the power output in the circles (white to red colors) is only indicated for the three central turbines... why?

This figure aims to show the coherent effect of the gravity waves. The three central turbines from each row are included because the aggregate effect of the waves in the power signal is still clear. When more turbines are included, the effect of the gravity waves are distorted because the waves are misaligned with the rows. A sentence has been added at line 407-409 in the revised manuscript for clarity:

> "Additionally, because the gravity wave propagation direction and turbine rows are slightly misaligned, only the three middle turbines from each row are considered to quantify the coherent effect of the gravity waves on individual turbines."

l476: 'by up to 56 %': it is good to give the extreme value, but other statistics (mean) could also be meaningful to include

The mean percent decrease in power for all four rows is now included in the conclusion (line 497 in the revised manuscript), which is 39%.

References:

Miguel Sanchez Gomez, Julie K. Lundquist, Jeffrey D. Mirocha, Robert S. Arthur, Domingo Muñoz-Esparza, Rachel Robey; Can lidars assess wind plant blockage in simple terrain? A WRF-LES study. *J. Renewable Sustainable Energy* 1 November 2022; 14 (6): 063303. https://doi.org/10.1063/5.0103668

Chow, F. K., Schär, C., Ban, N., Lundquist, K. A., Schlemmer, L., & Shi, X. (2019). Crossing Multiple Gray Zones in the Transition from Mesoscale to Microscale Simulation over Complex Terrain. *Atmosphere*, *10*(5), 274. https://doi.org/10.3390/atmos10050274

Johnson, A., and X. Wang, 2019: Multicase Assessment of the Impacts of Horizontal and Vertical Grid Spacing, and Turbulence Closure Model, on Subkilometer-Scale Simulations of Atmospheric Bores during PECAN. Mon. Wea. Rev., 147, 1533–1555, https://doi.org/10.1175/MWR-D-18-0322.1.