Title: *Meso- to micro-scale modeling of the atmospheric stability effects in wind turbine wake behavior in complex terrain*

Authors: *Adam S Wise, James M T Neher, Robert S Arthur, Jeffrey D Mirocha, Julie K Lundquist, and Fotini K Chow*

Submitted to: QJRMS

Date: September 1, 2021

Reviewer general comment: *The current manuscript presents results of a suite of WRF-LES configurations around the complex topography of the Perdigão field site. Results represent the first time that WRF-LES is used to study the flow interaction in complex terrain and wind turbines. Results of the numerical simulations are qualitatively compared to the experimental measurements. According to the authors, the features of interest for the present work are: (a) mountain waves, (b) recirculation zones, and how these interact with a wind turbine and the corresponding wake as a function of atmospheric thermal stratification. The manuscript is very nicely written, clear, and with superb "candy-to-the-eye" type figures. Also the a priori aimed research question is of interest, and providing a thorough analysis of the problem posed would be of high interest to the community. Unfortunately the scientific analysis and content remains largely the same as the original version, mostly based on qualitative metrics and comparisons of numbers. The truth is that review of this manuscript has become quite a frustrating experience (hence in part the long delay in the review; my apologies for that), given that the authors instead of carefully considering the constructive suggestions and comments made, have instead taken a defensive approach. Review of scientific manuscripts is a volunteering act in support of the community. I don´t do it with any aim to go against anyone, or their work. I mostly do it to keep learning, and try to provide my opinion for others to consider and stir some additional thinking. As a critical scientist with myself, and with the community, I would reject publication of the manuscript as is. However, I also realize that this might be a controversial decision, hence I rather leave the community to judge it by themselves, and time will tell whether the work herein presented is worthy of publication. Beyond this rather important general comment, below I provide a few additional comments, that if the authors are willing to consider, could probably strengthen their manuscript.*

More Specific Technical Comments

1. **Reviewer comment:** In line 9-10; it might be appropriate to try to justify why the authors have selected to two study the mentioned periods. Are they relevant to something, or representative of something? Or is just a random selection. Providing this additional info my help motivate. This also relates to the text in Line 14-15. Are the authors expecting the dependence on the wake behavior to be continuous and smooth, or illustrating a sharp change in behavior at a certain thermal stratification?

2. **Reviewer comment:** In line 18-20; '*This study demonstrates the ability of the ...*'; I wonder has this been questioned in the literature? If the WRF model has been used and validated in a continuous manner through the years, and the GAD model has also been developed and

tested (as mentioned by the authors in line 61 to 63 in the text), where is the need for another comparison? Have the authors introduced any new element in the WRF platform that requires testing and verification? – This doesn't seem to be the case, but maybe I missed it again.

3. **Reviewer comment:** In line 55; '*and more realistic turbulence compared to*'; this is not necessarily true. Why would the turbulence initiation method based on the cell perturbation method provide more accurate turbulence than that on an idealized LES simulation? Maybe, and only maybe, under certain atmospheric conditions where very large scale forcing is of relevance, affecting the near surface turbulence, this could be true. However, this is not necessarily generalizable.

4. **Reviewer comment:** In line 74-75, the objective is nicely stated, which is great. It would be nice if the conclusions came around and provided an answer to this fundamental questions. Unfortunately I have the impression this is not the case, since the study only provides information for two independent stability cases. For example, the authors could also follow up on the results/hypotheses from previous works cited by the authors in line 80-81, and explore self similarity of the turbine wakes, in this case as a function of varying thermal stratification. This would add some interesting generalizable science, beyond the case specific observations.

5. **Reviewer comment:** In lines 87-88, why do the authors want to use idealized LES? – This hardly makes any sense for two reasons: 1. the authors mention from the beginning that they are interested in real conditions; 2. they run the simulations, show two instantaneous snapshots, and don't use that data any further. What is the point? I think the manuscript would be a lot simpler just using directly the realistic WRF-LES platform.

6. **Reviewer comment:** In line 86, '*based on atmospheric stability*'; this statement is a bit generous since the authors only consider 2 different stratifications. This is equivalent to that experimentalist that goes to develop a field experiment and only takes two data values to study a complex problem. One would expect at least three points, to discard the obvious linear fit, isn't it? Maybe saying something about the neutral stratification case to complement? If at the neutral stratification regime results are similar to those observed in convective conditions, what is the intensity on stable stratification needed for the wake to start changing behavior? – These are the kind of research question I would have expected to get resolved or studied in this manuscript as I read the introduction.

7. **Reviewer comment:** In line 161, note that the additional figure sent in the review seems to include periods of counter gradient fluxes (maybe my eyes are betraying me,...), but if this was the case a brief comment could be added, mentioning that those time periods have been discarded for example.

8. **Reviewer comment:** In line 206-207, when talking about the cooling rate, I interpret that being at the surface, is that correct? Mind you it could also be at the top boundary condition. Maybe it is worth clarifying this, is you decide to keep the text about the idealized LES.

9. **Reviewer comment:** In line 209-210, how is the turbulence initialized in the ideal LES cases? To that end, what is the turn over time of the flow around the periodic domain? Does the turbulent flow at least have time to properly develop before reaching the mountain? Does the turbulence reach some sort of convergence prior to reaching the region of interest, meaning the mountains?

10. **Reviewer comment:** In line 219, when discussing the surface roughness. If the authors have access to wind lidar data and three tall met towers, why don't they use the experimental data to extract an approximate z0 value? Regardless of how challenging the experimental data might be, at least the value will be of higher significance than that chosen by the authors (which doesn't seem to have much justification). Unless there is a stronger reason behind that decision that is not explained in the manuscript.

11. **Reviewer comment:** In line 224-225; '*Having demonstrated ..*'; once again, I wouldn't feel comfortable myself making that statement. The authors have only showed two beautiful colored figures showing reasonable results, but there is no strict 'demonstration'.

12. **Reviewer comment:** In line 245; instead of going around a rather lengthy justification of why the method to generate turbulence might or might not work, why don't the authors compare the turbulence spectra with that of the met towers? Wouldn't that be a simple, and really robust way to demonstrate there is no doubt on the approach? If the goal of the paper is to study the effect of thermal stratification in realistic conditions, then what better way to proof that the resultant flow is realistic than comparing turbulence spectra? – This seems to me a fundamental requirement.

13. **Reviewer comment:** In line 269-270, where exactly is the surface roughness being manually modified? It might be interesting to have a figure illustrating the surface roughness used in the different regions of the domain, so if someone else wants to redo these simulations for new purposes can actually do so.

14. **Reviewer comment:** In line 275, why does one need to compare the numerical data with the doppler data '*to understand*'? One can compare numerical data and experimental data to verify the accuracy of the numerical simulations, as it seems to be done in this section, or alternatively one can study in detail one of the datasets, or both at the same time to unravel the physics of the problem, which is not what is done in this section. It might be worth rephrasing this opening sentence to clarify the expectations for the reader.

15. **Reviewer comment:** In the paragraph contained between lines 300 to 305, it is discussed that comparison of averaged data in a non-stationary problem is complicated. This is indeed very much true. In fact, as the authors well indicate, given the random nature of the turbulent flow, comparison of instantaneous snapshots is not very significant. As an alternative, the authors provide some sort of 'cone averaging strategy (data within an easting and northing position)', as well as an additional spanwise averaging within 30m, if understood correctly. While one could discuss whether this approach is significant or not (which is not what I intend here), what at least the authors could do is better justify their decision. Why is this method robust? How sensitive are the results on this averaging strategy? – Given that this is more or less a hand-waved approach, I think it would be nice for the reader to have a better sense of the sensitivity of the outcomes on this analysis.

16. **Reviewer comment:** In lines 368-369; '*the height of the mountain wave does not extend as high in the model compared to the results*'; any thoughts on why this could be? – This kind of physical analysis and interpretation is what could make the manuscript a lot more interesting.

17. **Reviewer comment:** I would strongly encourage the authors to consider shortening the manuscript. While the main goal of the project, as stated in the introduction, is to understand the effect of thermal stratification on wind turbine wakes in complex terrain, the analysis

related to this point only expands between pages 25 to 30, among which only 1.5 pages would be of text and analysis, the rest are figures. So, said otherwise, the main focus and goal of the paper is discussed in 1.5 pages of text out of a manuscript of about 30 pages without counting references. The complementary information provided in the Results section (pages 13 to 25, more than 10 pages) relates to the comparison of the WRF obtained flow with data obtained from the met towers and wind lidars. I have the impression that this is quite out of balance, and maybe the paper should then be framed as a comparison between WRF data and experimental data, which once again I am not so sure on its value given that WRF is used daily around the world, and the manuscript does not introduce any new element or variation of elements in terms of WRF configuration. I think this final page-count also helps illustrates the point I was making in my opening remarks.