Please note that references to line numbers are for the revised manuscript.

Scientific questions:

1. Can you comment on how representative the specific case studies (stable in combination with mountain waves, unstable in combination with a recirculation zone) are for the wind conditions at Perdigão and for complex terrain in general? Do you think the conclusions in terms of wind-turbine wake behavior (terrain following in stable, deflecting upwards due to recirculation) will hold for all stable/unstable wind conditions at this site and for complex terrain in general?

Mountain waves occurred for almost 50% of the nights during the intensive observation period (Fernando et al. 2018). For recirculation, Menke et al. (2019) found that reverse flow with wind speeds greater than 0.5 m/s occurred over 50% of the time when the wind direction was perpendicular to the ridges. We have added these details into the manuscript at lines 130-132.

We expect that the conclusions in terms of the wind turbine wake behavior would hold for most convective and stable atmospheric conditions at the Perdigão site as long as the phenomena of interest (recirculation zones and mountain waves) are present. Other phenomena could be modeled using WRF-LES-GAD to examine wind turbine wake behavior in other cases. With regards to other sites, the wind turbine wake behavior would depend largely on the vegetation (surface roughness) and steepness of the terrain. If the previously mentioned parameters are similar to those in the examined case studies, similar conclusions for wake behavior in convective and stable conditions could be made. We have added this discussion to the manuscript in lines 514-518.

2. The inverse of the Froude number as defined in Equation 1 is often called the non-dimensional mountain height and it represents the ratio of the mountain height to the (vertical) wave length of the mountain wave. When discussing the case where the mountain wave length is shorter than the width of the mountain (line 133-134), wouldn't it be more appropriate to use a Froude number based on the width rather than the height of the mountain like for example in the book of Stull (1988, section 14.2.3 Flow over hills). Obviously, this also affects figure 3 and later calculations of the Froude number in section 4.1.

We thank the reviewer for bringing up this interesting point. Since we are dealing mostly with flow over the mountain rather than flow around the mountain, we believe the Froude number based on the mountain height is more appropriate. Additionally, the Froude number defined using the mountain width is for an isolated hill as opposed to an extended ridge as is the case for Perdigão.

The Froude number based on the ridge height has also been calculated for previous studies at Perdigão (Fernando et al. 2019, Palma et al. 2019).

The wavelength of a mountain wave is defined as lambda = 2*pi*U/N and is actually independent of both the mountain height and width. However, we can quickly calculate a Froude number based on width during the stable case. If the mountain width is 800 m, N is 0.035 /s, and U is 9.8 m/s, then Fr = 3.14*9.8/(0.035*800) = 1.1. Because of the proportion of the height and width of the ridge at Perdigão, calculations of the Froude number using either the mountain height or width are very similar.

The discussion comparing the wavelength to the mountain width has been removed to avoid any confusion. The sentence, lines 139-140, now reads: "For small Froude numbers (< 1), when wind speeds are low or the stability is very strong, the wavelength of the mountain wave is short resulting in weak mountain waves."

3. It is not entirely clear how you setup the semi-idealized simulations. What pressure gradient force do you impose, or alternatively what wind speed and direction do you enforce (and at what height)? Do you apply a negative surface heat flux in the stable case? What is the domain height of these simulations? Do you use any damping layers at the top? For how long did you simulate these idealized cases?

The semi-idealized simulations are set up using geostrophic forcing with a specified initial wind and potential temperature profile. The wind speed for the stable case comes from a sounding, while for the unstable case it is uniformly specified as 7 m/s. In both cases, the geostrophic wind direction is aligned with the x-axis, such that the forcing is entirely longitudinal. For the stable case, a cooling rate of -0.25 K/hr was used. However, the effect of this cooling rate is minimal as the simulation is only run for 10 minutes. The domain height is just above 1400 m and we use Rayleigh damping within 500 m of the domain top with a coefficient of 0.003 /s. These details have been incorporated into the semi-idealized modeling section in lines 198-209.

4. Line 182-183: What do you mean with the stratification is self-destructive? Do you mean the stable stratification turns into a constant temperature profile because of turbulent mixing? Does the simulation become unstable due to inertial oscillations? Something else?

Because the idealized simulation uses periodic boundary conditions, the flow will recirculate and the effects of the mountains and the GAD will induce mixing which will erode the stable stratification. We have edited the manuscript to replace the term “self-destructive” with “eroded by turbulent mixing” in line 207.

5. Figure 12: Is there any averaging of model results or measurements? Under higher turbulent conditions, does it make sense to compare instantaneous velocity fields with the point measurements of the met towers given the chaotic nature of turbulence?

There is spatial averaging of 30 m in each spanwise direction (60 m total) in figure 12 but no time averaging. While we agree that exact agreement between the model and observations is not expected under highly turbulent conditions, we believe that this figure shows important qualitative agreement. The four transects included aim to be representative of the hour-long simulation period, highlighting that the general dynamics of the flow are captured despite differences due to the chaotic nature of turbulence.

6. Line 361 “Differences in 80m mean wind speeds … are less than 1 m/s …”: Is this the difference averaged over the entire time horizon, or the root-mean square, or something else? When I look at figure 13b around 13:55, I obviously see differences of more than 1m/s. Similar for the discussion of the differences in the wind direction and also for the plots for stable conditions. You seem to be talking about a certain average difference being below a certain value, but the instantaneous differences are clearly higher than this value and it is not clear how the averaging is done. Please clarify.

The reviewer is correct in that instantaneous differences can be large and this is due to the previously noted chaotic nature of turbulence. This is why we have included both bias and RMSE as metrics comparing WRF-LES-GAD with observations. The RMSE captures the cumulative magnitude of the errors over the entire simulation period while bias captures the average behavior. To clarify the text, the discussion regarding mean metrics has been removed. In line 392, the variability quantified in terms of standard deviation now explicitly states that the standard deviation is calculated over the hour-long period. Additionally, errors displayed in Table 4 are discussed in the sentences that follow in the manuscript (lines 394-399).

7. Line 445: The blockage addressed in the papers by Meyer-Forsting, Bleeg and Segalini refers to the global blockage effect far upstream of large wind farms that arises due to the combination of many wind turbines. Here, you only have one turbine and you are only looking at

2D upstream, so I would say that turbine induction might be playing a role, not blockage. I don't think it is appropriate to refer to these wind-farm blockage studies here, so consider removing these references.

We have removed the blockage references and added two references that provide more detail on wind turbine induction zones.

Medici, D., Ivanell, S., Dahlberg, J.-A., and Alfredsson, P. H.: The upstream flow of a wind turbine: blockage effect, Wind Energy, 14, 691–697, https://doi.org/https://doi.org/10.1002/we.451, 2011.

Simley, E., Angelou, N., Mikkelsen, T., Sjöholm, M., Mann, J., and Pao, L. Y.: Characterization of wind velocities in the upstream induction zone of a wind turbine using scanning continuous-wave lidars, Journal of Renewable and Sustainable Energy, 8, 013 301, https://doi.org/10.1063/1.4940025, 2016.

Minor questions and technical comments:

Line 148: Please mention at least once how UTC relates to local time at this particular site. Maybe it is even more interesting to mention what time corresponds to sunrise.

The local time has been added in sections 2.2.1 and 2.2.2 at lines 167 and 177, respectively. In section 2.2.1, the local time corresponding to sunrise has also been added in lines 167-168.

Line 148-149: Please mention what value of the wind direction corresponds to a direction perpendicular to the ridge such that the statement can actually be assessed by the reader.

A southwesterly wind of 215 deg. corresponds to being perpendicular to the ridges. This has been added in line 169.

Line 149 reference to Fig 2a and 2b: I don't think this reference is correct. Figures 2a and b show wind speed in stable and convective conditions, while the sentence was talking about the wind direction.

The reference should be just Fig. 2c which refers to the wind direction during the stable case. This has been fixed.

Line 155-156 and figure 2d: Wouldn't it make more sense here to use an actual stability parameter like the Obukhov length L or z/L to assess the atmospheric stability? Especially when talking about moderately convective conditions, how can you assess that based solely on the potential temperature gradient?

Assessing atmospheric stability in complex terrain is an active research area. Using a gradient Richardson number is difficult because of the effect of terrain-induced flow-speedup over the ridge affecting the shear terms as mentioned by Menke et al. (2019). Additionally, Bodini et al. (2020) calculated the Obukhov length at Perdigão and found that it was not very powerful in predicting dissipation rate. However, we have calculated the Obukhov length and added it to Fig. 2 along with new discussion regarding its limited applicability for assessing atmospheric stability in complex terrain (lines 147-157). Additionally, we have removed "moderately" from this section since our primary intention is to say that the case is convective, rather than stable.

Line 158 reference to Fig 2d: Again this reference seems to be a bit misplaced as it refers to the wind direction plot, not the wind speed plot.

The reference should be Fig 2b and this has been fixed.

Fig2: Please mention hub height of 80m (?) and measurement points for the temperature gradient (10m and 100m?) in the caption for clarity.

This has been added in the caption of Fig. 2.

End of line 188: surface roughness length should be in units of meters.

Units have been added.

Line 291, 356, and many others : You often start a new paragraph without clearly indicating that you will be talking about a new figure. For instance on line 291, are you still referring to figure 7, or are you rather referring to figure 8 or table 2?

In lines 292, 318, 333, 357, 377, 387, and 412, we have added explicit references to figures to improve readability.

> Table 2 and 3: the results in these tables are not really discussed it seems. Can you somehow use these results when assessing/discussing the results of related figures?

For Table 2, we have added language talking about the RMSE of the temperature gradient in line 331. This is in addition to the specific numbers already mentioned and discussed in the paragraph from lines 324-332.

Specific numbers from Table 3 have been weaved into the paragraph from lines 338-356, where they can be related to figures.

> Line 312: How high above the actual surface is the nose of the jet which you use to calculate the Froude number? The Brunt-Vaisala frequency is based on the temperature gradient between 10m and 100m, but it is not clear whether the nose of the jet is at about the same height above the surface.

The nose of the jet is at 650 m a.s.l. in the model and 720 m a.s.l. in the measurements (which is 350 m and 420 m above the valley floor, respectively). This has been specified in the text at line 340 to avoid any confusion. The sentence has been rewritten as: "Using the velocity at the nose of the jet, which is at 650 m a.s.l. In the model and 720 m a.s.l in the observations, we can calculate a Froude number at the beginning of the period of interest."

> Line 320: How high is the inversion layer above the valley floor (600m asl but at what altitude is the valley floor)?

The valley floor is just under 300 m a.s.l. This has been added to the manuscript in lines 357-358:

"The small negative bias of -0.3 K for potential temperature is also apparent for the second sounding, but the height of temperature inversion is accurately captured close to 600 m a.s.l or 300 m above the valley floor."

> Line 323: When talking about the striation of lower wind speed close to 600 asl in fig 6b and d, in what ranges of x can this be seen?

There is a coherent striation from the beginning of the transect until the rotor in the lee of the second ridge at x = 2500 m. This has been added to the manuscript in lines 351-352:

"This decrease in wind speed could be the turbine wake or the striation of lower wind speed seen close to 600 m a.s.l in the multi-Doppler lidar scans from the beginning of the transect until the rotor in the lee of the second ridge at X = 2500 m (Fig. 6(b) and Fig. 6(d))."

> Line 357 "there is less variability in the potential temperature gradient for the model compared with observations": I'd say there is more variability in the model output compared with the observations (this is also mentioned later in the text). The statement seems to contradict the results shown in figure 13 g,h,i.

This should read "there is more variability…" which we have corrected in the manuscript at line 387.

> Line 427: The year of the publication by Churchfield and Sirnivas is missing (also in the list of references).

This has been fixed.

We thank the reviewer for the comments. We think that the revised manuscript and the responses above help to address most of these concerns.