.
Response to the comments from referee 2

Thank you very much for the comments on our manuscript. Here our response to each of your comments. The response is given in blue color.

Best regards,
The authors

General comment

The paper presents a validation of WRF LES model on three canonical flow cases: neutral, stable, and unstable. The motivation, structure, content, graphical presentation, and the referenes, are all excellent.

The high-quality data for the validation is collected at Østerild test centre, from 5 heights between 7 and 240 metres.

The three flow cases were defined using the Monin-Obukhov stability at 37 meters above the ground. The surfaces heat fluxes and/or surface temperature tendency, and the roughness length, resulting in the observed variables, were used to set-up the WRF model. This approach is adequate, but the assumption of the homogeneous surface should be validated to ensure that the flow characteristics are not too much affected by the fetch distance for different heights and wind speeds. It is possible that the apparent roughness length would be different if other heights than 37 metres would be used to derive the necessary parameters.

As the reviewer points out, the roughness length is different when computed from different heights. This was illustrated in Peña (2019). In the same study, it was also shown that the behavior of the dimensionless wind shear, $\phi_m = (\kappa z/u_*)(dU/dz)$, with stability, z/L, follows closely surface-layer scaling over a homogeneous, flat surface when looking at the 37-m height instead of 7-m. We now extend Sect. 2.1 (where we first describe the measurements at Østerild) with this information

Given that a lot of effort has been invested into estimations of the surface roughness using satellite and lidar data, especially around Østerild, one could ask why you haven't used the available roughness data for these WRF simulations. Are you suggesting that the roughness length is a function of the flow, and not the other way around?

The surface roughness length is a function of the flow, at least in the light of the logarithmic law of the wall (it is actually an integration constant). This is not very helpful for modeling as we need some inputs and boundary conditions to bound the flow, the roughness length being one of them. As we want the simulations to represent the observed conditions, we need to extract the roughness information from the observed flow behavior. This is the approach used here. But, looking at the bigger picture, the idea is that satellite/lidar-derived roughness maps do also reflect roughnesses consistent with those one can extract from the flow observations so that we can use those products for flow modeling elsewhere. In this work, however, our focus is not on the goodness or accuracy of those products.

The model's capability to accurately simulate the three flow cases is nicely analyzed in terms of the mean properties, and all stress components. The agreement between the LES model and real data is quite good. In particular the velocity spectra are fantastic, in the resolved part of the spectrum of course.

All three cases are shear-driven which, especially for the stable case, somewhat limits the applicability of the results. You have proposed a few directions for further work and we can look forward to it being performed and presented.

Thank you very much for your general comment about our analysis. We will hopefully provide the community with further contributions on this topic soon.

Specific comments

1. What was the period of the data collection, i.e. how much data is actually used in the aggregation for U, N, and S? From Figure 1 it seems that about 6 months of data is used, however from L274 it follows that there is

3000 seconds of data. Please clarify.

Following the recommendations of the reviewer, we now also extend Section 2.1 (with the description of the selection of the flow cases from the measurements) so that information with regards to the binning process, extension of the analyzed periods, and the amount of cases per stability condition is provided.

2. Figures 3-5: would it be possible to estimate the Ri number for these cases? It could be useful for comparisons when more of the flow cases will be constructed in the time to come.

The issue with the Ri number is that it is very sensitive to the levels used for its calculation. As shown in the new version of the manuscript, both simulated fluxes (in the forms of friction velocity and surface heat flux) approach rather well the observed values, and so the cases can be readily classified based on the surface fluxes, i.e., on the dimensionless stability z/L (Fig. 1-left).

3. Figure 6: it would be easier to compare the instantaneous flow representations if the color scales were the same.

As suggested by the reviewer we tried that but setting the color scales within the same range either distorts the flow structures or disappears some of them, as the wind speeds are not the same for the three conditions

4. P12L203: Here the RMSE is introduced, but it is not immediately clear if this is the statistics derived from 5 values (5 heights, and using the whole-period-average values), or? Please clarify.

As suggested by the reviewer, at the end of the first paragraph in Sect. 3 (where we first presented RMSEs), we add a sentence clarifying that the RMSEs are based on the mean of the values of the examined quantity. In the previous sentences in that paragraph, it is stated that RMSEs are computed between simulations and observations across the observed heights. We now also add 'five' between 'the observed' to further clarify this.

5. P12L214: Please elaborate/comment on the possible reasons why the neutral simulation differs most from the observations, i.e. the stable and unstable simulations match the observations better.

As suggested, we now extend the discussion within that paragraph: "The overprediction under neutral conditions of the simulated dimensionless wind shear within the first tens of meters from the surface is the result of an overprediction of the simulated vertical shear of the u-component. Thus, the contribution of the v-component to the turning of the wind diminishes, which results in low values for the relative direction."

6. Figures 9-11 (also elsewhere but perhaps easiest to discuss here): the differences between N and S cases are very small, it requires quite an effort to find any significant differences. Have you considered finding "more stable" flow cases, or would that be difficult at this location?

More stable cases are much less frequent at Østerild as observed from the frequency of stability conditions in Fig. 1-left. We do not fully agree with the reviewer in that the neutral and stable cases are similar: as already noticed from the observations of normalized winds in Fig. 1-right, the wind shear in stable conditions is much higher than that in neutral conditions. The observed friction velocity under neutral conditions is nearly doubled that of stable conditions and the neutral TKEs are more than twice the values in stable conditions.

7. P16L262-267, Figure 12: Please expand the discussion of the TKE profiles. For example, the neutral simulation seems to "compress" the TKE profile towards the surface, evident by the "nose" of the profile being quite lower than 37 metres, which is the height of the maximum observed value

As suggested by the reviewer, we now expand the discussion in that paragraph: "In the three ABL regimes, there is a 'local maximum' in the total turbulent kinetic energy profile close to the surface. This shows the limitation of the LES to resolve turbulence below the local maximum; from this level down to the surface, the SGS contribution increases substantially."

8. P16L272: diving $->$ dividing.

Changed as suggested.

9. P21L339-340: there is some confusion regarding the inversion strength, and the adiabatic lapse rate for dry atmosphere. The capping inversion and the adiabatic lapse rate should have the opposite signs, and are

unrelated in this context. What was the purpose of the statement in L339-340?

As suggested, we now delete the comment on the adiabatic lapse rate as it was confusing.

References

Peña, A.: Østerild: a natural laboratory for atmospheric turbulence, J. Renew. Sustain. Energ., 11, 063 302, 2019.