Interactive comment on “Analysis of Different Gray Zone Treatments in WRF-LES Real Case Simulations” by Paula Doubrava et al.

Anonymous Referee #1

Received and published: 13 February 2018

In this article, the authors present a sensitivity analysis of the turbulence closure approach utilized in an intermediate domain with grid spacing within the gray zone (GZ) when performing coupled mesoscale-LES simulations with the WRF model. This is an interesting aspect of multiscale simulations that requires quantification, and the authors present systematic analysis to that end, including comparison to field observations from the Prince Edward Island Wind Energy Experiment. Overall the manuscript is well written and logically organized. However, there are major aspects that must be addressed before the manuscript can be considered for publication in Wind Energy Science. The two main issues are:

1) Figure 9 clearly exhibits that the model results are not able to capture the basic diurnal evolution of the atmospheric boundary layer (ABL), with the near-surface sensible heat flux being nearly constant in time throughout the day in the model (\(\sim 0.005 \text{ K m s}^{-1}\)). This is strong evidence that there is something wrong with the model setup. The model does not capture the diurnal cycle and therefore, any attempt to see the impact of the GZ on the nested LES results through comparison to observations is hopeless. While a multi-day period as the authors present here is desirable, it does not add much if the results are not properly representing the meteorological conditions of interest to a minimum degree. The authors should look for cases within this 16-day time window (if any) where mesoscale results reasonably compare to the observations, at least in terms of basic features, and exclude the rest of the days from the analysis. If none is found, the authors need to explore the use of different re-analysis datasets and/or physics options to try to find a better model performance at the mesoscale before running the GZ and nested LES simulations.

2) While the grid resolution employed at the innermost LES domain may be suitable for daytime conditions; it is definitely not the case for the stable boundary layer. A horizontal grid spacing of 111 m is too large to resolve the characteristic turbulent scales that dominate stable boundary layers, typically of the order of few meters to several tens of meters. This implies that resolutions of at most 10 m are required, as it is often used in both ideal and real world LES modeling of stable ABLs. With that consideration in mind, the authors should either change their setup consistently with this requirement or to remove stable ABL instances from the analysis.

Here a list of other comments, following the order of appearance in the manuscript:

- Page 1, line 2: “momentum”. Is not only momentum balance, also energy equation is involved. Please modify to: “produced by the model”.

- Page 1, line 7: It would be more correct to say an LES parameterization. Or, a "large-eddy simulation (LES) parameterization where the most energetic turbulent motions are resolved and only the effects of the sub-grid scales on the resolved flow are parameterized".

Any explanation for the -3 slope?

“Navier-Stokes equations”. You should employ a better term. NWP models are "geophysical fluid dynamics" models. You could perhaps use instead "engineering computational fluid dynamics (CFD)" models, which is more commonly used in that context.

This literature review is incomplete. There are other works in the literature including these by Munoz-Esparza et al. JAMES2017 and Rai et al. BLM2017 that need to be acknowledged here.

"LES mode". Please add some references here.

"single code". Please add some references here.

"GZ is avoided". Better would be to say: "can potentially be avoided". However, the problem remains, since one likely wants to use high-resolution mesoscale forcing to drive the stand alone LES calculations. It could also be avoided within the same model by skipping GZ resolutions.

"Mirocha et al. (2014)"

This speculation may not be needed here, since this is what you will be analyzing in detail in the remaining of the manuscript.

"Study Domain". What do you mean? NWP model domain? This is confusing, please change.

"indicates".

Is it neutral the rest of the time? What is the threshold in L used to determine the classification? I guess it is described in the reference, but it would be a good help to the reader if it is mentioned here as well.

How can that be possible? What is the top pressure? A CBL can easily be several km deep, in which case very few grid points remain for the rest of the atmosphere. Could you instead provide the \Delta z of the first and last levels within the ABL.

Please include the rest of model physics options used in the study.

What about spinup time? Did you perform multiple initializations or not? Please describe. If a single run, is the SST updated? Also, make sure the landuse is such that the gird cell where you are outputting the results is considered as "land". If that is not the case, such mismatching would explain the issues with the diurnal cycle...

"including the ABLP budget". What does that mean? Which budget?

It would be interesting to see the sub-grid scale contribution (and total) for the GZ domain as well.

Results in this section (Figs. 3, 4, 5, 7, 8, 9) should be separated at least into two stability classes (convective and stable). The model errors can cancel out and its nature is expected to be stability-dependent. Otherwise, it is difficult to identify any clear trends in the analysis. In particular, stable results are highly questionable (as
mentioned earlier), and therefore should be separated from the rest not to corrupt the rest of the analysis.

- Page 10, line 2-4: It is somewhat puzzling that although the authors claim the LES_LES is the least performing, the PDF in the GZ LES is more similar in structure to all the microscale LES (although the peak is over-predicted). So the micro LES departures from the GZ ABLP solutions. Could the authors elaborate on this?

- Page 10, line 5: There is no budget presented, just TKE. Remove "budget".

- Page 11, line 1: Figure 5. Why do you use U to correlate with TKE? It would be more appropriate to plot time of the day, since that is likely a better estimate of stability than U.

- Page 11, line 3-5: This is not correct. The PSD you are seeing at low frequencies is that of mesoscale motions, which are essentially quasi-two-dimensional in nature. The ABLPs act as diffusion terms, and will therefore extract energy from the model. Also, how can there be more variability in the GZ LES TKE from Fig. 5 and less energy in the PSD? "Act to generate turbulence energy at length scales higher than the grid size". How is that possible? What is the mechanism? Please explain.

- Page 11, line 10-11: WRF has been shown to produce mesoscale -5/3 kinetic energy spectra (e.g., Skamarock 2004). Please plot the spectrum from the 3 mesoscale domains and see whether that pattern is induced by the parent domains or not.

- Page 11, line 14-16: In the paragraph just above you claim usage of the 4 Hz output in the previous section...

- Page 13, Figure 7: Any explanation of why the wind speed bias is very similar between all the GZ-driven simulations while there were large differences in between the GZ solutions?

- Page 13, line 10-13: The differences here are smaller than for other wind directions. That is a plausible explanation but you need to support that with additional analysis.

C5

- Page 14, line 7: The micro LES distribution of the LES_LES case is quite similar in structure to the ABLP-driven cases, clearly departing from the 333 m distribution. This points to some issue/s in the modeling. Could the authors please comment on this?

- Page 14, line 14: This is hard to understand. If errors were larger for low wind speeds, what is causing the best TKE agreement?

- Page 16, Sect. 4.3: Given that there is no diurnal-cycle forcing in the simulations, and such is the extent of the frequency range considered in this section, this analysis is totally inconclusive (it compares to very different things, model and observations). Needs to be repeated once a better model setup is found.