

Wind Energ. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/wes-2020-119-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

The challenge of comparing ensemble-averaged mast data with canonical ABL simulations

Javier Sanz Rodrigo (Referee)

Referee comment on "Evaluation of idealized large-eddy simulations performed with the Weather Research and Forecasting model using turbulence measurements from a 250□m meteorological mast" by Alfredo Peña et al., Wind Energ. Sci. Discuss.,
<https://doi.org/10.5194/wes-2020-119-RC1>, 2021

Very complete assessment of the WRF-LES model in idealized conditions covering mean profiles and spectra at a 250-m boundary-layer mast. The main challenge in the study is in the inherent inconsistencies in the comparison of canonical ABL simulations with ensemble-averaged profiles. The study is based on the assumption that the observed profiles represent the horizontally homogeneous conditions that the LES model is based upon. This is not convincingly addressed or quantified so it deserves some additional discussion. The validation is quantified using normalized quantities based on u^* . A suggestion is made to use other set of normalized profiles to avoid the intrinsic bias in u^* in the computation of the RMSE.

P3. Section 2.1: I think you should justify why doing long-term statistics leads to mean profiles that can be considered good references for idealized ABL simulations. While bin averaging with long-term data will definitively help in the convergence of the mean profiles you don't guarantee that you are removing persistent heterogeneities from the mesoscale wind climate. To come up with reference profiles for canonical ABL you would want to filter out situations that deviate from quasi-steady and horizontally homogeneous conditions. Have the authors considering any filtering for such conditions? By filtering more you can decrease the value of the std-bars in the observed profiles to values that are closer to the ones from LES (e.g. Fig. 7 shows this discrepancy in the error bars). Even if you provide a reference about the measurements, please describe the case selection in more detail with regards to the binning process and number of samples in each case and discuss how you deal with these issues.

P4,83: Can you discuss if these stability functions have been found at the Østerild site and if they are also used in the surface layer model of WRF? Together with the roughness length, this may be relevant in the interpretation of surface fluxes.

P4.Setup: There is no mentioning of any sensitivity analysis on the grid settings that would support the selected grid configuration. While the setup looks reasonable it would be better to include some discussion on the adequacy of these settings for canonical ABL simulations. In particular, it would be interesting to compare with the parametric study of LES model settings done by Mirocha et al (2018, <https://www.wind-energy-sci.net/3/589/2018/wes-3-589-2018.html>) that also includes idealized WRF simulations.

P5.Figure 2: The x-axis is the vertical level height. For those not familiar with WRF grid, please indicate if this height represents the cell center or the cell height.

P5.113: Can you discuss how you come up with the values of geostrophic wind (14 m/s in neutral and stable conditions and 8 m/s in unstable conditions)? I would suggest using a table to collect all the input quantities that you use to define the three flow cases.

P6.144: "The choice of the time to extract LES statistics depends on the type of boundary layer." Can you elaborate on how you select this time? Based on the discussion later on, the selection seems a bit arbitrary although I understand that the profiles never reach a steady state. Maybe, at this point, you should discuss this challenge. My impression is that you end up choosing profiles that have sufficiently developed to a quasi-steady state close to the values you have from measurements and "have the right look" for canonical ABLs.

P7.Fig 3: Can you indicate (in the caption or in the legend) which time in the simulation corresponds to the blue line? A vertical line in the time series would also help.

Figs 3/4/5: The tke behaves similarly to the u^* so you might as well skip it. In neutral conditions you may consider replacing it with the wind speed of the jet nose since this is used to select the neutral profile. In unstable and stable conditions you may consider replacing u^* by the heat flux (as in Fig 5), or the Obukhov length (I'd use this one), to see how the profile evolves from the energy point of view.

P8.172: You should mention that the first model level is at around 5 m while the reference flux measurements are taken at 37 m. This difference may be significant in the value of the fluxes, specially in stable conditions. Why not using the closest level to 37 m (in all the plots) as the reference height for the time series? Why not showing the observed values of u^* and heat flux in Figs. 3 and 4? In principle, instead of using the u^* value produced by WRF at the surface, you could compute u^* from the high-frequency time series as if they were sonic measurements to try to represent the same quantity (as it is done in section 3.4 for the fluxes). To improve consistency, you could introduce filters in the measurements to remove mesoscale trends and mimic the behavior of the sgs model in LES (as the low-pass filter displayed in the spectra). I understand that this is beyond the scope of the paper at this point but I think it is worth discussing how "canonical" are the measurements when we compare with idealized LES simulations since these quantities integrate different turbulent scales in simulations and measurements.

P11.Fig. 7: Because of the differences in the reference heights of u^* (and the underlying filter in LES), you are probably introducing a bias in the normalized profiles. Although less formal, maybe a simpler but more robust way of normalizing the profiles would be using the wind speed at 37 m instead of the friction velocity.

P2.203: Please put the metrics of this section in a Table or put them together with those in Table 1.

Figures 9,10,11; Table 1: As discussed before, these plots and error metrics depend on quantities that are normalized with the u^* value, which may come with a bias due to using a different vertical level in measurements and observations. In addition, as pointed out by the authors, u^* in the simulations depend on the value of the roughness length that is used as input which is difficult to quantify in reality. For these reasons, I think it is better to choose a reference height and normalize each profile by the corresponding value of each quantity at that reference height, e.g. U/U_{37} , uw/uw_{37} , etc (and exclude the 37 m level in the computation of RMSE).

P18, 305: "The simulated means are always within the observed variability"... well, the variability is pretty large in the measurements, isn't it? Still, the agreement is pretty good considering the uncertainties in the definition of "canonical" flow cases.

P22.344: When running real-time simulations you can definitively extract mesoscale tendencies from WRF that you can use to filter out periods of strong heterogeneity. This can help you narrow down the case selection to obtain mean profiles that more closely match the idealized conditions of this study.