Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-16-RC2, 2020

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Can machine learning improve the model representation of TKE dissipation rate in the boundary layer for complex terrain?" *by* Nicola Bodini et al.

Anonymous Referee #2

Received and published: 24 May 2020

The manuscript "Can machine learning improve the model representation of TKE dissipation rate in the boundary layer for complex terrain?" by Bodini et al. provides an interesting look at using machine learning techniques to generate estimates of TKE dissipation rate and comparing those results to the approach used in the MYNN parameterization. This work should be of interest to the community and provides a useful road map for scientists wanting to apply a similar approach to other data sets. Overall, I think the manuscript will be acceptable for publication in Geoscientific Model Development after relatively minor revisions. The text is generally clearly written and straight forewarned to follow. I wonder, given that focus on data analysis rather than atmospheric model development, if the manuscript is a better fit for Atmosphere Chemistry

and Physics or Atmospheric Measurement Techniques. I leave that, however, up to the editor.

General comments: âĂć Machine learning techniques generally do not increase our physical understanding. The authors try to address this in Section 5.1 and 5.2 where additional analysis is provided. Section 5.2, however, is very brief and should be developed more to provide additional insight into the results. âĂć In section 3, the authors show that the MYNN approach does a reasonable job in unstable conditions, but much worse when the boundary layer is statically stable. I was surprised that the authors didn't carry this analysis into the subsequent sections. It would seem natural to examine the model behavior with stability in Section 5.

Specific comments: 1. Figure 1. I appreciate the histogram shown in Figure 2, but could you also differentiate the points in Figure 1 to indicate measurement heights? Maybe that doesn't work well if the measurements made at a single location are at several heights? 2. Section 2.1: Can you say anything more about how the sonics are distributed on the towers? For example, how many were deployed on the 100 m tower? 3. Lines 78-80: Double check this sentence, the wording seems odd. 4. Line 101: Is the mean potential temperature computed from the sonic data or does it come from a different source? 5. Lines 104-109: Can you point the reader to the terrain data set that was used? What was the resolution of that data set? Does that have any impact on the results? 6. Line 138: I agree that the length scale assumption is the best you can do given the data set that you have, but I think some additional discussion is warranted to help defend that selection. Can you argue that Ls is likely dominate near the surface? 7. Figure 6: You show the mean bias in Figure 6, could bars be added to indicate the standard deviation of the bias? This would help show how significant the biases are. In addition, the figures shows a decrease with height. Is this significant, or could it (at least partially) be related to the horizontal distribution of the measurements taken at different heights? 8. Section 4: It would be helpful if you could include a brief discussion of why you selected these particular algorithms for this application. 9.

Section 5.2: Is there a better header for this section to help the reader understand the importance of the analysis that is presented? 10. Section 5.2: This section seems to end abruptly. Can you guide the reader to anything important? What additional insight is gained from the analysis? What does it tell us about what is controlling the dissipation rate at large values of wind speed and/or TKE?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-16, 2020.