Wind Energ. Sci. Discuss.,
https://doi.org/10.5194/wes-2020-61-RC1, 2020

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Interactive comment on "Utilizing Physics-Based Input Features within a Machine Learning Model to Predict Wind Speed Forecasting Error" *by* Daniel Vassallo et al.

Javier Sanz Rodrigo (Referee)

jsrodrigo@cener.com

Received and published: 25 May 2020

Very interesting work investigating the potential use of atmospheric quantities measured at a met mast to improve very-short-term predictions of wind speed from 10 min to 3 hr ahead using machine learning. The relevance of different input feutures in isolation and combined is quantified showing significant error reductions with quantities that can be easily extracted from conventional mast measurement campaigns making the study particularly useful for implementation in operational tools. The testing methodology is convincing only requiring some clarifications. In particular, the selection of the input averaging interval of 5 min requires additional motivation. Have the authors

tested different input intervals to see the impact in the error reduction?

Comments:

Page 4, line 7: Can you elaborate further on how wind speed data is stationary? Is this tested at the prdiction timescales (10 min - 3hr)? I would also expect wind speed to be subject to seasonal and diurnal variability. Please clarify.

P5, L28: I'm curious why are you not using the Obukhov length (or z/L)? Isn't it a more commonly used parameter to characterize stability? You may want to motivate this selection even though from the results of Figure 6 it seems that stability parameters are not that important in the improvement of forecasts.

P7, L28: How are sonic measurements corrected for tilt? What is the interval used when deriving the fluxes? Is it equal, shorter or longer than the 5-min interval used in the moving average? This is just to know if 5-min is the actual filter in the data or if the data already came with a longer averaging time. This could also be relevant to understand the potential impact of this filter in the performance at 10 min prediction horizon (Figure 3).

P8, Figure 2: The map is difficult to read. It would be better to show a elevation contour plot where we can read the relative heights. I don't think it is necessary to provide an illustration of the mast levels if they are described in the text.

P8, L8: You end up using 5-min averaged data to build predictive models with prediction horizonts at 10 min, 1 hour and 3 hours. You previously mentioned that these are single-step forecasts. Wouldn't you have to use input data that is averaged at the same interval than the forecast step (e.g. use 3-hour moving averages to predict 3 hr ahead)? Or do you forecast {10min, 1hr, 3h} ahead based always based on 5-min data? If the latter is true, please clarify why not using a consistent interval between input and prediction data or, alternatively, how dependent are the results to the chosen interval in the time series.

P10. Figure 4: One may wonder how a Persistance-RF model would work. This might be a good result to include in the paper so that you can just isolate the impact of RF from that of the forecasting model to make the results more generally applicable. Maybe you get to the same conclusions with a simpler model.

P8. L11: How is the flux Richardson calculated between 100-20 m? Isn't it a local quantity derived from a sonic level? Is it the mean value between the two levels? Please clarify.

Interactive comment on Wind Energ. Sci. Discuss., https://doi.org/10.5194/wes-2020-61, 2020.