

# ***Interactive comment on “The importance of round-robin validation when assessing machine-learning-based vertical extrapolation of wind speeds” by Nicola Bodini and Mike Optis***

## **Anonymous Referee #3**

Received and published: 19 February 2020

The authors present a machine learning approach to vertical extrapolation of wind speeds compared to standard approaches. The research is quite robust, described in detail and well written. The conclusion that the machine learning approach, a random forest, can be extrapolated to other sites as shown by this round robin evaluation is an important scientific discovery.

There are a few areas that can be further described or clarified to make this an excellent paper.

1.) Page 5 line 108 - it is stated that precipitation periods were excluded from the analysis. Please explain why and what impact this has on the analysis.

[Printer-friendly version](#)

[Discussion paper](#)



2.) Page 5 Line 11 - it is stated that a 30-min average is used. Is there a reason why 30-min was chosen and would that averaging period affect the results? No further analysis is needed - just an explanation or including in the results discussion how the averaging period may impact the analysis.

3.) Page 7 Line 146 - please cite Pedregosa et al 2011 for Scikit-learn. <http://www.jmlr.org/papers/v12/pedregosa11a.html>

4.) Page 7 Lines 155-157 - the explanation of training, testing and cross-validation is not clear. Is the 5-fold cross validation performed on the 80% training data and the 20% testing data is held out for independent validation after the hyperparameters are chosen? Please describe so that it is clear the testing data was not used in the choosing of the hyperparameters.

5.) Figure 7 are the partial dependence plots for the random forest, which are an important aspect of the interpretability of the machine learning models. However, it is better to show both predictor importance and partial dependence plots so that the relative importance of each variable and its associated partial dependence is known. Recommend adding in predictor importance plots or list to add value to the interpretability.

---

Interactive comment on Wind Energ. Sci. Discuss., <https://doi.org/10.5194/wes-2020-2>, 2020.

Printer-friendly version

Discussion paper

