Wind Energ. Sci. Discuss.,
https://doi.org/10.5194/wes-2020-2-RC2, 2020

WIND ENERGY SCIENCE DISCUSSIONS

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

Raghavendra Krishnamurthy (Referee)

raghavendra.krishnamurthy@pnnl.gov

Received and published: 4 February 2020

The article "The importance of round-robin validation when assessing machine-learning-based vertical extrapolation of wind speeds" by Bodini & Optis details a round robin approach to vertical extrapolation from the 4 ARM SGP Doppler lidars using a random forest algorithm. The paper is well written and the reviewer agrees the necessity of a round-robin type approach to assess the accuracy of machine learning algorithms and make it more universal. Below are some comments/questions which probe into some of the details of the paper and would improve the paper if addressed in the next version.

1. Line 61: Extrapolation is not only generally done up to Hub-height but through the rotor swept area. So, I am not sure I follow the author's argument here, that if hub-height winds are available extrapolation is unnecessary. The approach to go above hub-height can also be treated as a "Gap filling" approach for met-masts (when Lidars are moved around from one location to the other for a short period). Please clarify.

2. For power law type extrapolations, measurements not only at the surface but at multiple heights is needed to estimate the dynamic power law exponent. So please define what you mean by near-surface in the paper? Is it within surface layer or also above surface layer?

3. The authors mention LLJs, frequently observed in the ARM site, how does this effect the ML output at higher heights?

4. The idea of round-robin is fair for machine-learning based extrapolation, but only if the training has been done accounting for all atmospheric conditions that would be representative of other sites. As you know, ML models can only learn what is in their training dataset. Therefore, the round-robin type approaches come with a caveat that the search space of the variables expands to many of the common conditions (including external forcings specific to each site) observed in the atmosphere and at all the evaluated sites. This comment needs to be addressed in the paper with supporting evidence.

5. For the SNR filtering, not only precipitation, but fog is also prevalent at SGP and it diminishes the range considerably at lower heights. Therefore, an upper limit on SNR could be important to filter out any abnormalities in radial velocity data.

6. Line 94: Maybe I am picky, but the poor data quality is because those measurements fall within the lidar blind zone? The Blind zone is generally 2 times the range-gate size, which fits the heights. If yes, please mention that for clarity.

7. Equation 2: The temperature used was from the sonic or from the cup anemometer

for the fluxes? Sonic anemometer temperature measurements have significant biases and are not considered very accurate (Berg et al., 2017). This would cause errors in classifying stability or L.

8. How is atmospheric stability defined? Based on Richardson number of MO length? Please provide the thresholds or a reference from which you picked the thresholds for classifying stability for the MO type extrapolation.

9. MO length is not known to be valid for complex terrain (Fernando et al., 2015), therefore these parameters would not fit well for all types of terrain/sites. Therefore, a note about applicability of the chosen parameters to different conditions/terrains would be needed to address the universality of these parameters for such an approach.

10. The effect of external forcings at different ARM sites are not considered, which is important in this context of machine learning (comment #4 above). The wakes from wind turbines have major impact on the hub-height winds at some of these sites. Sites E37 and E39 are far away from turbines or wind farms, while C1 and E41 are relatively closer and have considerable impact on the winds at hub-height in certain predominant wind directions. Please see attached the wind directions and distance from wind turbines at each of these sites and something similar must be included in your analysis. Therefore, I would recommend you can either discard the below sectors from your analysis or test the accuracy in waked conditions.

11. How much of these chosen parameters (TKE, L, WS4, WS65) explain the variance in the RF model? What is the unbiased predictor importance estimates of the chosen variables?

12. Figure 7: Maybe some additional explanation is required on how the dependence is calculated. Its not very clear if its just a correlation type analysis or something else. Please provide more details here. Also, the extrapolated wind speeds (Y-axis) for all plots are not same and it's not very clear why.

Very Minor comment: The language is a bit colloquial for a journal and would urge the authors to take that into consideration for their revised manuscript. For example: Line 225: Maybe you can but I am not sure if it's formal to end a sentence with "are": please rephrase. Similar sentence structuring needs to be considered throughout the document.

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

Lidar Location	Wind Direction Sectors	Approximate Distance of the nearest Turbine to the Lidar (m)	Common Turbine Height in that sector	Rotor Diameter	Type of Turbine	Built Year
C1	67 - 93	6700	90	116	GE 2.5 MW	2017
	112 - 196	3500	80	116	GE 2.3 MW	2017
	243 - 270	4600	80	82.5	GE 1.68 MW	2012
E32	45 - 60	11500	80	108	Siemens 2.3 MW	2016
E37	--	> 20000	--	--	--	--
E39	--	> 20000	--	--	--	--
E41	205-255	2500	87	126	Vestas V126-3.3	2016
	295 - 15	5000	80	108	Siemens 2.3 MW	2015

Fig. 1. SGP ARM Doppler Lidars Wind Farm Distance