REVIEW: Vertical extrapolation of ASCAT ocean surface winds using machine learning techniques
Hatfield et al.

General comment:

The paper describes the application of machine learning algorithms to vertically extrapolate near-surface winds derived from satellites data. The focus is offshore, in Northern Europe.

The paper is generally well written (although some grammar improvements here and there would be recommended), and it represents a nice academic application of machine learning to the wind energy sector. However, I question the practical utility of this approach in real world applications. The main reason for this is that satellite-derived observations of wind speed at a given location are only available ~ twice a day, and always at around the same times. I am really struggling in finding a situation where someone would be interested in knowing hub-height wind speed only at these two hours of the day. While practical applicability is not a strict requirement for having a paper published, I still think this limit should be at the very minimum highly stressed in the paper, and language softened to reflect the limited applicability of the results described in the analysis.

Specific comments:
1. L.19: add "above the sea surface" or something similar.
2. L. 26: 275 m seems like a very specific threshold – can you provide a reference?
3. L.45: "predicting" instead of "predict" (or "to predict")
4. L. 66: "higher" instead of "greater"
5. L. 106: in the text you mention water temperature, while in Table 1 you mention sea surface temperature; please be consistent.
6. L.111: do you mean that all variables have exactly an availability of 85%, or greater than 85%?
7. L.123: what's the temporal resolution?
8. Consider moving all the details regarding data access to the 'Data availability' section towards the end.
9. Figure 1 caption: "shapes" instead of "rectangles".
10. It is hard to fully understand Table 3 (and some of the discussion in this section) without a clear explanation of the temporal frequencies considered here for the various data sources. Does "total data" refer to 30-min average time periods? And "Concurrent data with ASCAT" to 30-min average time periods in which an ASCAT data point was recorded? Please clarify in the paper.
11. Table 2: "FINO" is repeated twice in the left column.
12. Why using the cosine of wind direction only (and not the sine, too)?
13. Once again, being clear about the temporal resolution of the data used is key to understand whether the random split between train and test sets is a right choice, or autocorrelation effects might play a role in artificially enhancing the ML results.

14. Table 5: I would argue that the most relevant comparison is ML vs NEWA at each site, so I would suggest adding a horizontal line after each site, and highlight in bold the "winner" metrics at each site.
15. L.170: clarify that the error values refer to the test set.
16. L.173: "Note" instead of "not".
17. Section 3.2: please clarify how the mean wind profile from the RF was computed. Do you simply apply the RF to the whole period, and average results? Or to the test set only?
18. Section 3.3: why not including FINO2 as well?
19. Section 3.4: can you somewhat verify your hypothesis of horizontal homogeneity by looking at spatial variability of the meteorological variables from NEWA and the reanalysis product?
20. Figure 4: once again, more clarity is needed when explaining what is being plotted and described. What's the temporal extent of what is shown? Are NORA data taken only at time stamps at which ASCAT data are available? How about the ML-extrapolated winds? Are we really comparing apples to apples? Once that is clarified, please adjust text accordingly (it is misleading to state you are showing 2018-averaged winds, if you are only cherry picking time stamps).
21. Figure 4: why wasn't NEWA included?
22. Conversely, why wasn't NORA3 included in the earlier analysis (Table 5)? It is essential to know how well it compares to observed winds in order to use it as proxy for the truth here.
23. Figure 6: please change labels and instead list ALL the satellites available in each time period. Also, specify that this is a cumulative plot. What do you mean by "concurrent" in the caption?
24. Discussion of Figure 10 is key (see my main major comment above), and in my opinion should be moved way earlier in the paper.
25. Figure 10 caption: specify when referring to bars vs lines. Also, specify you are referring to local time and not UTC time (I believe).
26. L.332: decreases instead of increases?
27. Data availability: why not sharing the model algorithm scripts as well?
28. Please double check references and make sure each has a DOI.