Summary

This manuscript provides the methodology and analysis behind a newly released publicly available offshore wind dataset called NOST (NOAAOffshoreWindProfiles-USA). The data set is six-hourly at 0.25-degree resolution and spans from 1987 to 2022. The authors implement a machine learning method (random forest regression) to extrapolate wind profiles from satellite-derived surface winds (NOAA's National Centers for Environmental Information's Blended Seawinds version 2.0 (NBSv2.0) product). They train their extrapolation model using lidar data and validate it at the same lidar sites as well as at two other locations whose data was not used for training. They compare their wind profiles with reference wind profiles (log-law, power law) against observations and find that their model outperforms reference wind profiles (as expected). They quantify their errors and also analyze subsets of the data to focus on low-level jets and high-shear conditions. Lastly, they use a triple collocation method to compare the NOSP product with ERA5 and the NOW23 wind dataset and find the lowest estimated errors for NOSP regardless of coastal region. The paper is very well-written, and I believe their dataset would be highly valued by wind farm developers. Overall, this is a great paper that can be improved with a few minor revisions as suggested below.

Specific Comments

1. I would suggest removing or streamlining the first three sentences in the abstract. Considering that this is a wind energy-focused journal, the statements are already widely understood by the journal's target audience.

2. I do not think that both the log-law and power law need to be used as references. I would recommend just choosing one as they are going to be similar and with similar limitations. Additionally, I could not find the surface roughness value used to generate the log-law profiles. Does the value vary spatially or temporally due to ocean conditions? The authors state that a lack of knowledge regarding u* and z_0 restrict them to only the neutral log-law, but you still need z_0 for the neutral log-law so I am confused why u* could not also be determined. I would also think that u* is also an output of ERA5 similar to the SST.

3. In general, there are a lot of figures with vertical profiles. I would recommend a few things to improve readability As in 2, I would suggest removing either the log-law or

the power law. RMSE and MAE are also similar metrics. I would recommend just showing one of the two.

4. In general, the figures need to take advantage of the width of the page to improve readability. They are too small as they currently exist. They also need to be centered. The captions for figures 8 and 9 need to be on the same page as the image.

5. Did the authors consider changing the shear exponent for the power law for the higher shear cases? I assume that the authors want a model that is independent of parameters that should be tuned; however, using a power law with alpha=0.1 is obviously going to underestimate wind speeds higher up during high-shear conditions.

Minor Comments

- It would improve readability of Tables 2 and 3 if the data subset (overall, normal, etc.) was a separate column
- Table 2: the entry on the second row for the MAE column is missing a '%'.