Reviewer 1 Response

**We thank the reviewer for their time and their responses to our manuscript. Our replies are inserted below into the reviewer's comments.**

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.
   **The first three sentences were removed.**

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.
   **While we agree that only using one model for comparison may suffice, we believe that comparing against both is more thorough and readers may wonder how our model compares to the law we drop. As such we would prefer to leave both in if the reviewer finds that acceptable.**

   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.
   **The surface roughness length used for the neutral log law is a constant z0 = 0.0001 (as in Optis et al., 2021). This has now been added into the paper (Line 180). As we used a constant z0, it would not be possible to calculate $u_*$ from z0 with any meaningful information as it would also be a constant.**

   **$u_*$ is also available from ERA5 but is unavailable at the exact location of lidar stations used to train the model. As ERA5 has lower spatial (0.25° vs. in situ) and temporal resolutions**

**(1 hour vs. 10 minute) than the lidar station wind profile data, using ERA5 $u_*$ could add more error into the log-law model when used for comparison at these stations. We use SSTs measured in situ at the lidar stations in the model training, not SSTs from ERA5. ERA5 SST and air temperature are only used in our analysis as RFR model inputs when extrapolating the Blended Seawinds dataset to full profiles as the two products have the same spatial resolution.**

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.
   **We agree that RMSE and MAE are similar metrics and have decided to remove RMSE from the paper. As for removing log and power law, please see our reply to comment #2 above.**

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.
   **Figures have been resized to fit the page better and centered. The captions for all images are now on the corresponding pages to the images.**

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.
   **The reviewer is correct that we wanted to compare against a model that is independent of tuning parameters. While alpha = 0.1 is obviously underestimating wind speeds during high shear conditions, a power law would have to calculate alpha in a way such that it knows when high shear conditions are occurring. This is not the case for most formulations we have seen, and the calculation we have seen for alpha that does take into account atmospheric stability and could potentially estimate wind during high shear cases also relies on stability correction functions that contain parameters unavailable from the lidar datasets.**

   **We also wanted to compare against models that could be used for the purpose of extrapolating winds at all grid points in our dataset, so calculating a different constant alpha for the purpose of high shear conditions specifically would not work as when applying the model across the whole region in our dataset as there would be no way to know beforehand which value to use without already knowing if the case is high shear or not.**

**Minor Comments**

- It would improve readability of Tables 2 and 3 if the data subset (overall, normal, etc.) was a separate column

**We agree with the reviewer and have added a new column for data subset "Profile Type" as a separate column.**

- Table 2: the entry on the second row for the MAE column is missing a '%'.

**Thank you for finding this. The missing ‘%’ is now added.**