Dear Julie Lundquist,

We believe that the required corrections enhance the quality of the work, and we appreciate your contribution. Here we address the three issues you have raised:

1) *On a scientific point, the definition of atmospheric stability bins requires some justification; please provide some references in at the end of Section 2 to justify your choice of |L| ~ 1000 m as the dividing line between stable/unstable and neutral (perhaps Muñoz-Esparza et al 2012 could be helpful here).*

   Thank you for your comment regarding the atmospheric stability bins.

   In literature, classifications based on the Obukhov length (L) vary slightly. For example, Hansen, 2012 defined very stable and stable conditions with ranges of $10<L<50$ and $50<L< 200$, respectively, while ranges of $-100<L<-50$, $-200<L<-100$ were used for very unstable and unstable conditions. However, many studies adopt $|L| <200$ for very stable/unstable and $200<|L|<1000$ for stable/unstable conditions (e.g., Motta, 2005; Watson, 2014).

   Since the primary purpose of our stability classification is model error characterization, we did not differentiate between very stable (unstable) and stable (unstable) conditions. Therefore, we chose $|L|=1000$ as a practical threshold for defining stable/unstable and neutral categories, as also used by Schneemann Schneemann, 2021. We have added the following to Section 2.6:

   "The choice was based on literature conventions, where ranges such as $0<L<200$ ($-200 < L < 0$) and $200<L<1000$ ($-1000 < L < -200$) are common for very stable (very unstable) and stable (unstable) conditions (Argyle and Watson, 2014; Motta et al., 2005). In this study, we do not distinguish between stable (unstable) and very stable (very unstable) conditions, as we use this classification primarily for error characterization."

2) *One reviewer pointed out that figure panels are not separately labelled, which makes it difficult for readers to understand exactly which panel is being discussed. I note that the author instructions state "Labels of panels must be included with brackets around letters being lower case (e.g. (a), (b), etc.).", so please revise to include panel labels, and you may also use the panel labels throughout your discussion to improve clarity. Specifically, Fig. 7, 11 need attention in this regard (Fig 10 is fine).*

   We updated the Fig. 7 and 11 in the new version and adjusted the text accordingly.

3) *Some of the figures are not friendly to color-blind viewers. As recommended at https://www.wind-energy-science.net/submission.html#figurestables, please run your figures through https://www.color-blindness.com/coblis-color-blindness-simulator/ and modify*

accordingly (I note that Fig 3 color loses meaning for Red-Blind/Protanopia, so you should check others as well.

We ran all the figures through https://www.color-blindness.com/coblis-color-blindness-simulator and updated the plots, making changes to the colors, when necessary.

References

Hansen, K. B. (2012). The impact of turbulence intensity and atmospheric stability on power deficits due to wind turbine wakes at Horns Rev wind farm. *Wind Energ.*, 183-196. doi:https://doi.org/10.1002/we.512

Motta, M. a. (2005). The influence of non-logarithmic wind speed profiles on potential power output at Danish offshore sites. *Wind Energy*, 219-236. doi:https://doi.org/10.1002/we.146

Schneemann, J. a. (2021). Offshore wind farm global blockage measured with scanning lidar. *Wind Energy Science*, 521-538. doi:10.5194/wes-6-521-2021

Watson, P. A. (2014). Assessing the dependence of surface layer atmospheric stability on measurement height at offshore locations. *Journal of Wind Engineering and Industrial Aerodynamics*, 88-99. doi:https://doi.org/10.1016/j.jweia.2014.06.002