Review of "Onshore and Offshore Wind Resources and Operating Conditions in the Eastern U.S." by Rebecca Foody, Jacob Coburn, Jeanie A. Aird, Rebecca J. Barthelmie and Sara C. Pryor

This study describes the statistical analysis of a comprehensive lidar data set with focus on onshore and offshore wind speeds and power production in the U.S.

The data set described in this manuscript is interesting and of high relevance for wind energy research. The thematic itself is within the scope of the journal. However, a clear structure is missing, I could not identify clear objectives and/or hypotheses to be addressed and there is a lack of clear interpretations, discussions, and conclusions. The reader often needs either to accept statements without a clear proof or need to interpret the results by its own. Therefore, I recommend a major revision of the manuscript.

Specific comments

Introduction:
Entire Introduction: To my opinion, the big picture is missing here. What do you expect to find from the analyses. What do you want to explain or proof, what is your scientific question, the overall goal of your study?

Page 2, line 51: The authors mention that there are few previous studies without giving any references. I would also recommend to shortly summarize what has been done and found in those few studies.

Page 2, Point 1: I would expect a discussion/statement/explanation in the results part about the differences in the regions and what we can learn from it (e.g., an evaluation as to whether a region is more suitable as the others, beyond onshore-offshore differences)

Page 2, line 79: I could not find a guidance regarding optimal special scale. I don't feel guided with only a short statement that correlation is lower than 0.4 for distances >350 km. I recommend using either another formulation or provide a more detailed and profound discussion in the results part.

Page 3, paragraph 1: There is quite a harsh transition from the previous paragraph and topics to this one. The topic is completely new, and I miss a kind of introduction to why this is important in your study and what you want to show/discuss with results to this. What is your goal, what do you want to compare and what are possible consequences for your major hypothesis? In this paragraph, you mention something from structural loading and wakes but what does it mean for whatever you want to show and where is the discussion about it in the results/conclusion part?

Page 3, paragraph 2: Wouldn't it be better to paste a goal at the beginning of a paragraph? Otherwise, there is again a harsh transition from one topic to another from which the reader initially has no idea what the reason is, where to focus on.

Data Sources:
Page 4: The best year has a specific time frame for all positions, how about the analyses which are not based on the best years? Is the time span for all positions from January 2019 to December 2022 or are there some variabilities? If yes, how large are these and how would you expect them to influence your results?

Methods:
Page 7, chapter 3.3: Which question shall be answered by an analysis of the wind profile and why are you using shear and LLJ? In what sense are LLJ of relevance to wind energy applications?

Page 7, lines 223-225: Is this a commonly used method? Do you have references for this method which show that this can be done for wind energy or similar purposes? I wonder how suitable this method is considering the lower spatial and temporal resolution of ERA-5 compared to measurements, the

difference in time spans (comparing results from 44 years (ERA-5) to max 4 years (data) and the fact that ERA-5 has uncertainties on its own and additional uncertainties by the conversion from 10&100 m to 150 m. I would also suggest providing a kind of uncertainty or at least a discussion about this issue.

Results:
Page 8, lines 258-259: What does this mean? Why are low summer values a hint for a negative bias? Summer values are often lower, winter values often higher, so there are seasonal deviations from annual (and also long-term) mean values.

Page 8, lines 258-259: Definitely missing here is a detailed description and justification (preferably with reference) of how a long-term time series with relatively coarse resolution can lead to a meaningful error estimate of a point measurement, even more so when different time periods are used for this purpose. Why isn't it more likely here, that differences come from the interannual variability? What makes you believe that the data availability is responsible for these differences, in particular when you consider annual averages, and if so, wouldn't another way of calculating annual averages be the solution to avoid or at least minimize the influence? How do you calculate them that they have such a strong influence?

Page 11: Concerning the differences in Weibull scale parameters and AEP between best year and all data: What is the conclusion of this finding? Interannual variability? Data under-/overrepresentation? Any proofs for the one or the other?

Page 12, lines 290-293: A description/interpretation of figure 5 would be great. In general, the reader is a bit left alone with the interpretation of the figures. Either one understands it immediately on its own or not. Some help would be nice for all who didn't create the figures.

Page 13, lines 304-305: Could you explain this a bit more, please. There is an image, you could lead the reader through the image, just a bit, and let someone not being such deep into statistics see the same like you. Furthermore, do you have an explanation/expectation why the e-folding times at sea are larger than on land and why there is a slight difference in the onshore stations? E.g., any physical reasons for that?

Page 13, lines 308-309: Why does a large e-folding indicate the potential for more accurate power prediction? Are there any proofs? Did someone find this (citation?), or did you do any calculations?

Page 15: What is the conclusion from the analysis of shear conditions and LLJ?

Page 15/16, chapter demand matching: How did you calculate the normalized demand and site WPP and how do you relate it to the demand? What does it exactly mean, which conclusions can you draw from the findings? I guess, the couple of positions equipped with one turbine per position will not be able to cover the energy demand, but from the image it looks a bit like this. Means: further explanations are needed here to guide the reader into the right direction. And what does the comparison with ERA-5 reveal?

Concluding remarks:
At best, I see a summary here but no conclusion and no answer to a concrete scientific question or proof of a hypothesis. The reader is more or less left alone with the interpretation of this statistical analyses.

Page 16, lines 386-387: You compare the data to find out what? For what is it helpful, what does it aim for?

Page 17: 402-403: This is only a guess, I didn't see neither a proof nor a clear or understandable assessment of this (as has been stated in the introduction).

Page 17, line 410-411: To what extent does this follow from the Spearman correlation coefficients drop?

Page 17, line 414-420: The LCoE and its calculation was not mentioned in the results.

Figures:

Figure 3: For the comparison between onshore and offshore wind speeds I would suggest to create the same ranges for the y-axes. Also, the onshore figure is a bit crowded, maybe it would be a bit clearer to put the data availability into an own subfigure.

Figure 5: The figure is quite crowded, very small and differences are hard to interpret. I would also love to see a much better description in the text.

Figure 6: The lag time is in 10 Minute intervals, which needs an ad hoc recalculation to hours while reading the text, which in turn states the e-folding times in hours. I would recommend adapting either the text or, even better, the figures x-axis in a way that both becomes consistent.

Figure 8: Again very crowded, again, the image is not intuitively understandable without a more detailed description in the text.