Response to reviewer 1

Dear Anonymous Reviewer 1,

We appreciate your thoughtful feedback. It helped us to improve the manuscript. Below we comment on your suggestions in detail.
All reviewer comments appear in *italic text* below, while authors' responses appear in blue text. Line numbers referenced in the authors' responses refer to the revised document.

The authors present a statistical analysis of wind power of four wind turbines and wind direction shear, measured using a nearby Lidar. They discriminate between clockwise and counter-clockwise wind direction shear and show a correlation between under performance of the wind turbine and mean directional wind shear. This under performance is primarily visible for lower wind speeds. In general, the paper is very well written. Introduction, choice of data and statistical analysis are clear. The figures are suitable to present the data. My comments are therefore very minor asking for some clarification and some improvements of text and figures. I recommend that the paper is accepted after minor revisions.

General comments

The fact that turbine performance decreases by more than 15% for strong wind shear conditions is very relevant; I can clearly see that. I found the fact that is mostly high veering that is related to a drop in normalized power very interesting. However, I was missing any attempt for a physical explanation on why high veering has such a strong effect whereas high backing does not show the same signal. It seems that you are too defensive here, and a little bit of speculation could be appropriate in the discussion.

Thank you for giving relevancy to our findings. We agree that we did not provide a physical explanation on why high backing does not show a drop in normalized power. To provide a robust and defensible explanation, we would have liked to have a greater number of strong backing cases so that we could develop a rigorous hypothesis. Unfortunately, as we pointed out, an insufficient number of large backing cases occurred so we could not conclusively draw any conclusions regarding its effect on turbine performance. In the updated version of our manuscript, we examine in more detail the effects of small veering and small backing on normalized performance and found similar results for each scenario. We speculate that the variations in available power through the rotor disc and in turbine blades' efficiency are not significant for these subtle wind direction changes (see p. 21, lines 30 – 34; p. 22, lines 1 – 2).

Furthermore, I wonder if the unit deg m^{-1} is indeed the right one in this context. In my opinion, it depicts a certain generality that cannot be drawn from your dataset, given that you looked only at four wind turbines with the same rotor diameter. Given that directional shear is not linear, this may be relevant. Multiplying by 80 and using deg D^{-1} may help to communicate this important limitation.

We carefully considered this thoughtful suggestion, but literature regarding the effects of wind direction shear on turbine performance (e.g. Rareshide et al., 2009; Wagner et al., 2010; Walter et al., 2009) and characterization of the lower boundary layer for wind energy applications (e.g. Bodini et al., 2017) generally employs this unit. Further, comparisons made with other studies are easier to understand when maintaining the same units. Therefore, we deem appropriate using deg m^{-1} for characterizing shear in wind direction. We also consulted with colleagues using veer in their wind turbine control algorithms, and they specifically requested deg m^{-1}.

Third, I found that the directional wind shear is a function of height was overlooked in the analysis. This fact is clearly stated in the beginning but is ignored later on and only average directional shear is analysed. I wonder especially, what the effect of misalignment (with respect to the ideal 90 degree angle) would be if this misalignment is primarily below or above hub height. Given that wind speed generally increases with height, this seems important. You probably lack data on rotor orientation, but could you expand on this point in the discussion nonetheless? Splitting the analysis to the levels 40-80 and 80-120 may provide some insight.

Thank you for your comment. We considered your suggestion, but as is stated at the beginning of the manuscript, the main objective of this paper was to determine if direction wind shear across the whole rotor layer affects turbine performance to see if this phenomenon might have played an important role in conflicting results regarding the effects of atmospheric stability on turbine operation. If this overall veer did not show interesting results, then breaking it up into 40-80m and 80-120m segments likely would not have done so either. Hopefully our manuscript can motivate more detailed investigations using observations or simulations.

Finally, the rotation direction of the rotor is one asymmetry that may be of relevance given that you found such a distinct difference between veering and backing. Any thoughts?

Thank you for highlighting this as it most certainly plays an important role for large veering and backing scenarios. However, as stated above, we had insufficient large-backing cases to draw conclusions and make a more detailed comparison with turbine performance taking place during large-veering atmospheric conditions. Moreover, we did not expand on this topic as we found small veering and small backing to have similar effects on turbine operation (see p. 17, lines 9 – 18; p. 21, lines 30 – 34; p. 22, lines 1 – 2). The reviewer may be interested to see an investigation of the interactions of veering with wind turbine rotation direction and the effects on turbine wake structure, Englberger et al. (2019), at https://www.wind-energ-sci-discuss.net/wes-2019-45/.

Specific comments

1. Section 2.4: Do you really need a section of its own for two sentences? Furthermore, you mention that you average over 10 minutes in line 16, page 7 as well.

We agree that we did not present information in the best possible way. We have now clarified information in this subsection and expanded it. One of these expansions was to accommodate another reviewer's request for information on time synchronization between turbine and lidar data. Also, we consider it important to explain that lidar data is logged every 2 minutes (2-min averages), and turbine data is logged every 10 minutes (10-min averages) by their data-acquisition systems. Therefore, we match turbine-measured power operation to lidar-measured atmospheric conditions by averaging five 2-min lidar recordings. The text now reads as follows: "Turbine- and lidar-recorded data are averaged over different time intervals by their respective data-acquisition systems (2-min for lidar, and 10-min for turbine). Matching turbine performance with atmospheric conditions was performed by averaging 2-min lidar measurements for the corresponding 10-min turbine power production period. For example, turbine data for July 04, 2013 from 0500 to 0510 LT is matched with the average of five 2-min lidar data measurements corresponding to the same date and time period".

2. Page 15, line 13: Mentioning Turbine A-D comes a little bit out of the blue at this point because this is the very first time the text and the only other reference is Figure 1. You should refer to Figure 1 in this context.

Thank you for pointing this out, we did so in our updated version of our manuscript.

3. Figure 12: I found it confusing that there are four entries in the legend, but only three lines. You write in the text that bins with less than 30 data points are discarded, but I recommend removing the entry for high backing. Furthermore, it is rather strange that the error bars for the low veering case are not bold and black. It is a minor issue, but your readers should not play a guessing game. I have similar issues with Figs. 11 and 13, having four lines, but only 2 legend entries (specifically referring to the median).

Thank you for your comment. We agree that figures should be as clear as possible for the readers. We have included corresponding legends to all our new figures in the updated version of our manuscript.

4. Figure 2: The shading in this plot style leads to three colors in the graph, while there are only two in the legend. It is only a very, very minor remark, but wouldn't a simple line plot do a better job in communicating data availability?

Thank you for this suggestion. We considered a line plot for the updated version of our manuscript but preferred a bar graph. However, we have now separated data availability for turbines and lidar observations to avoid overlapping colors.

Typos

1. *Page 8, Equation 1: Full stop instead of comma after the equation.*

   We appreciate this comment and updated the manuscript accordingly.

2. *Page 15, line 7: 'was on average' instead of 'was in average'*

   We appreciate this comment and updated the manuscript accordingly.

3. *References For Rajewski et. al 2016 and Muñoz-Esparza et.al. 2017, it seems that there are some personal comments that have slipped from the .bib file into the references. The all-caps text is not part of the actual reference.*

   We appreciate this comment and updated the manuscript accordingly.

4. *Figures 11,12,13: You are using "Wind Speed in the x-label here, but "Wind speed everywhere else.*

   Thank you for catching this. We have updated axes labels in every figure and only left the first word of each label with an initial capital letter.