

Response to review

The reviewer's comments are given below in black font and our responses are in green font. A full tracked changes version of the manuscript is at the end of this document. Note also in responding to the reviewers comments (details below) we have also added 7 references.

Reviewer #1:

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.

Response: We regret that the reviewer did not find our structure clear. We felt that by writing in the paragraph starting on line 51 'We evaluate four aspects of the wind power generation potential on- and offshore:' and then listing them that we were setting out the objectives clearly. But in light of your concerns, we have completely restructured the introduction.

We have also changed the title to help the reader immediately know the purpose of the manuscript the new title is **Quantitative Comparison of Power Production and Power Quality Onshore and Offshore: A Case Study from the Eastern U.S.**

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?

We have re-written the introduction to read (quoting from the revised manuscript):

'Here we focus on the first of these reasons, and specifically seek to quantify the potential benefit of offshore wind turbine deployments using analyses of uniquely detailed wind profiles from an onshore LiDAR (Light Detection And Ranging) network and an offshore LiDAR network. We use these data sets to quantify and compare three aspects of the wind power generation potential on- and offshore:

- 1 Wind resource and power production. We present Weibull probability distribution parameters and derive energy density from the wind speed time series and compare and contrast the inferred wind resource at the onshore and offshore sites. We further compute and compare the Annual Energy Production (AEP) from the time series of wind speeds at each LiDAR site using a common wind turbine power curve.
- 2 Power quality. Intermittency is frequently cited as a barrier to increased wind power integration into the electrical grid (Bistline and Blanford, 2021). We quantify and compare the frequency of zero power production and intensity and probability of so-called ramp events (i.e., rapid changes in wind speed and power production) (DeMarco and Basu, 2018;Pichault et al., 2021) from each onshore and offshore site where LiDARs have been deployed.
- 3 Predictability and persistence of wind speeds and power production (Haghi et al., 2013;Haslett and Raftery, 1989). Within liberalized electricity markets, wind farm owner/operators bid in advance (e.g. 24 hours in advance) and are charged penalties for any imbalance between the bid

and actual production (Pinson et al., 2007). Hence, accurate forecasts of wind generation are important to reduce penalties and maximize revenue (Barthelmie et al., 2008). Persistence models where the power production at some future time is modeled as a function of power production in the recent past is often used as a benchmark forecast against which more sophisticated short-term power production models are compared (Kariniotakis et al., 2004). Also many statistical short-term forecast models are predicated in part on persistence (Zeng and Qiao, 2011) and thus are most skillful when the power production time series exhibits high temporal autocorrelation. We quantify the temporal autocorrelation of power production from each onshore and offshore site and compare the degree to which electrical power production from the onshore and offshore locations differ with respect to persistence and short-term predictability.

We further use these LiDAR measurements to quantify and compare a key driver of wind turbine loading at the on- and off-shore locations:

- 4 Extreme or anonymous wind shear across the rotor plane. Low-Level Jets (LLJ) are confined wind speed maxima within the lower atmospheric boundary layer (Stensrud, 1996) and are associated with enhanced vertical wind speed (and sometimes directional) shear relative to typical near-logarithmic profiles. LLJ within the wind turbine rotor plane are associated with higher aerodynamic and structural loading (Gutierrez et al., 2019;Gadde et al., 2021). Analyses of simulations with the Weather Research and Forecasting (WRF) model suggest that offshore coastal regions of the U.S. mid-Atlantic (including the locations of the buoys from which data are presented) generally exhibit a weakly sheared profile across the rotor plane and a relatively low frequency of LLJ (Aird et al., 2022). That analysis found LLJ in the lowest 500 m of the atmosphere are most frequent south of Massachusetts and during the summer (8% of all hours). They frequently occur at heights that intersect the wind turbine rotor plane, and at wind speeds within typical wind turbine operating ranges. Further, LLJ diagnosed from the WRF output were most intense and have lowest elevation under strong horizontal temperature gradients and lower planetary boundary layer heights. For comparative purposes, data from the NYSM LiDARs are used here to evaluate wind shear across the rotor plane and the occurrence, intensity, and height of LLJ at the onshore locations.

We also analyze the LiDAR data to quantify two other properties of relevance to wind energy integration into the electricity generation supply:

- 5 Co-variation of wind speeds and power production with varying distance separation (Pryor et al., 2014;Solbrekke et al., 2020). The electric power transmission network in the contiguous U.S. comprises three main interconnections (eastern, western, and Electric Reliability Council of Texas (ERCOT)) and 66 ‘balancing authorities’ that oversee regional operation of the electric grid and are referred to as Regional Transmission Operators (RTOs) or Independent System Operators (ISOs). New York (NY) state currently operates as a single state ISO. NY is both a net importer of electricity and the third most efficient state in terms of energy use per U.S. dollar of economic activity (<https://www.eia.gov/state/analysis.php?sid=NY>). Careful planning of wind farm locations on and offshore could ensure stable supply of wind-generated electricity into the grid and thus aid the transition from electricity imports and a current dependence on nuclear and natural gas (Eryilmaz et al., 2020). Here we quantify the spatial autocorrelation of power production from each onshore and offshore site where the LiDARs have been deployed to evaluate the decorrelation distance and hence provide guidance regarding optimal spatial scale of wind farm separation (on- and off-shore) for stability of wind power supply.
- 6 Seasonality and diurnal variability of wind power production (WPP) on- and off-shore for demand matching. Electricity demand varies with the level of economic activity and seasonal heating/cooling requirements which are a function of the regional climate (Castillo et al., 2022;Staffell and Pfenninger, 2018). Generally, electricity demand in the U.S. is minimized

between approximately 0400 and 0600 local time (LT), is high between 0800 and 1600 LT, and peaks between 1800 and 2100 LT (Burleyson et al., 2021). Diurnal variability of wind power generation is a function of location and land use but, for example, in ERCOT is highest at night (Kiviluoma et al., 2016), consistent with the expectation based on day-time variations in atmospheric stability caused by changes in net radiation and the surface energy balance. Because the oceans have higher specific heat capacity than land, this scale of variability is typically not present in the far offshore (> 20 km from the coast) (Barthelmie et al., 1996). At the seasonal scale, wind resources and power production in the midlatitudes and specifically the contiguous U.S. tend to peak in between October and April and are lowest in July or August due to pronounced shifts in the storm track and the frequency and intensity of mid-latitude cyclones (Pryor et al., 2020b). Recent research suggests WPP is highest in southeastern Canada and the northeastern U.S. during January and February (Coburn and Pryor, 2023). Thus, finally, we quantify whether electrical power from wind turbines deployed offshore exhibit higher or lower temporal matching with electricity demand in New York state at both the diurnal and seasonal scales.'

We feel this is a very clear statement of objectives, justification of the objectives and hope the reviewer concurs.

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.

We actually couldn't find any comparable studies but did not want to preclude the possibility that such studies exist, but since we could not find a comparable studies we have dropped this statement.

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)

We think the reviewer is referring to 'First, wind speeds tend to be higher and more consistent offshore due to both the lower surface roughness and lack of obstacles and topographic features that extract momentum and reduce both the wind speed and wind resource (Pryor and Barthelmie, 2002).

Accordingly, Capacity Factors (CF), which are the ratio of actual power generation divided by the theoretical maximum power generation, are typically higher offshore. Data from operating wind farms in Denmark indicate CF from four offshore wind farms with installed capacity (IC) of 160 to 400 MW of 41-53% while CF from smaller onshore wind farms (IC: 16-70 MW) have CF of 28-41% (Enevoldsen and Jacobson, 2021). Within the U.S., the mean CF for onshore wind farms built between 2014 and 2019 is approximately 41% (Wiser et al., 2021). Simulations using numerical models for offshore wind energy lease areas along the U.S. east coast indicate CF above 46% largely as a result of the higher wind speeds offshore (Pryor et al., 2021;Barthelmie et al., 2023). '

It certainly is true that some regions of the world exhibit higher wind resources but our focus here is listing the factors that are responsible for 'Enhanced deployment of wind turbines offshore offers great promise in terms of enhanced renewable energy penetration into the electricity generation portfolio for three primary reasons'... So, respectfully, we are unconvinced that adding a discussion of the relative wind resource in different regions of the world would be useful here.

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.

Our apologies – the concept of the spatial decay of correlation and the e-folding distance may, indeed, not be widely understood. We have elaborated on this matter a little in the Methods. We have modified this sentence to read; 'Temporal autocorrelation coefficients of the power production time series are used to derive e-folding time scales (i.e. the time delay at which the correlation coefficient drops to e^{-1} ,

i.e. to 0.37) which is used to represent the time scale at which the system ‘loses’ the memory of the initial state (Wilks, 2011).’

(i.e. to add 0.37)

And then have added this sentence:

‘The e-folding concept can also be applied in this context, to quantify the distance at the power production from two sites is no longer significantly correlated.’

So the bottom line is the separation distance at which the correlation coefficient for the power production time series from two sites two drops below about 0.37 (rounding to 0.4) is the distance at which the sites are no longer significantly correlated – or if you prefer if I wish to achieve a more stable electricity production through time I should place wind farms at sufficient separation that their individual generation is not significantly correlated. For this wind climate that distance is about 350 km. We hope our changes to the text help to clarify that point.

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?

We regret you found this to be a harsh transition. Hopefully with the re-structure introduction you find the flow is smoother.

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.

Again, we regret you found this to be a harsh transition. Hopefully with the re-structure introduction the flow is smoother. We do have an objective at the start of each bullet point.

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?

There is always a compromise to be made – use all the data that you have in order to increase the sample size for statistical testing versus use the data period that best represents the seasonal cycle. This is indeed a challenge. So, for statistical testing where sample size aids confidence we use all records (e.g. spatial autocorrelation) but we also present to the reader information regarding how estimated AEP varies as a function of the data sample.

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?

As we wrote:

‘The International Electrotechnical Commission (IEC) 61400-1 standard states the expected value of α over land is 0.2 and is typically in the range of 0.05 to 0.25 and uses a value of 0.2 in the normal wind profile model (IEC, 2019). The occurrence of α beyond this range implies shear across the rotor plane differs from this design expectation and hence may indicate higher mechanical loading.

Thus on the most fundamental level we are seeking to report how frequently the shear is outside this expectation. The LLJ analysis is really to examine if one source of anomalous wind shear profiles (i.e. the frequency of LLJ) is higher/lower on and offshore.’ We hope the modified introduction and the addition of an additional reference.

We have also added some text to the results section.: ‘The implication is that large wind turbines deployed in these locations may experience a relatively high frequency of large unbalanced rotor loads and reduced component lifetimes unless such loads can be appropriately compensated (Hur et al., 2017).’

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.

We believe the reviewer is referring to; ‘Hourly values from the 40-year $U_{150_{ERA5}}$ and $P_{150_{ERA5}}$ record are randomly resampled 1000 times with replacement using the number of hours from each month that the LiDAR data are available (Figure 3). For each of these 1000 bootstrapped samples the annual mean wind speed and AEP is calculated to provide an estimate of uncertainty due to the short time series from the LiDARs.’ Yes, bootstrap resampling is very frequently used to quantify confidence intervals around a metric (see for example the textbook by Dan Wilks or this text by Mudelsee; Climate Time Series Analysis: Classical Statistical and Bootstrap Methods (Atmospheric and Oceanographic Sciences Library, 51)). We have added a citation of the Wilks reference to the methods.

Perhaps the reviewer is speaking to our specific application. The mean power law coefficient 0.21 which for a height interval of 100 to 150 is equal to a correction of 0.06 (or if you prefer a multiplier of 1.09 on the 100-m wind speed) so it’s a small correction. We have now clarified this in the text. In terms of bootstrap resampling to derive uncertainties on wind speeds per se, it is a generalizable statistical method that can be applied to any geophysical property (see the book by Mudelsee)

We have clarified what the purpose of this analysis is by adding this statement; ‘This analysis explicitly acknowledges the presence of low-frequency variability (seasonal to multi-decadal) in mid-latitude wind speeds and wind resources (Pryor et al., 2020a) and is designed to quantify the uncertainty on mean wind speeds and power production computed from the relatively short LiDAR data time series.’

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.

We regret this statement wasn’t clearer. We have rewritten it to read:

‘Bootstrapping of ERA5 data indicates the mean annual wind speed computed from the LiDAR time series at the NYSM sites is likely underestimated by ~ 1.5-4.5% while AEP is underestimated by ~3-10% due to the high data availability in summer.’

We hope this clarifies.

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?

We regret any confusion – precisely we are examining inter-annual variability!. We hope we have removed any confusion by adding this statement in methods: ‘Although the LiDAR data sets that we analyze here are – to our knowledge – unique in terms of the duration and number of sites considered, we also contextualize the results and inferences drawn from these multi-year, but relatively short

duration, observations using the > 40 year duration ERA5 reanalysis product (Hersbach et al., 2020). This analysis explicitly acknowledges the presence of low-frequency variability (seasonal to multi-decadal) in mid-latitude wind speeds and wind resources (Pryor et al., 2020a) and is designed to quantify the uncertainty on mean wind speeds and power production computed from the relatively short LiDAR data time series. ‘

A minor note: If the cause is inter-annual variability due to differences in cyclone frequency/intensity that will be manifest in approximately equal magnitude in ‘point’ and spatially averaged values (unless the site is in complex terrain where directional channeling may be a factor).

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?

We did this analysis because it is important to acknowledge sources of uncertainty including incomplete time series. We note that in doing these analyses we also demonstrate that the on-shore off-shore differences are robust to data sampling issues. Accordingly, to avoid any confusion we have added the statement; ‘It is important to note that the differences in energy density computed from the on-shore and off-shore LiDAR data sets are robust to these sampling issues.’

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.

We regret any confusion regarding interpretation of this figure. The definitions of ramps are given in equation (3) and accompanying text in Methods. We have expanded that description this a little to read: ‘The probability of wind speed and power production ramp events are computed from the NYSERDA and NYSM LiDARs and in the case of wind speeds are normalized as follows:

$$\frac{\delta u(t)}{\sigma_{\delta u}} = \frac{u(t + \tau) - u(t)}{\sigma_{\delta u}} \quad (3)$$

where $u(t)$ is the wind speed at time t , $\delta u(t)$ is the wind speed increment from the prior time step, τ is the chosen time increment, and $\sigma_{\delta u}$ is the standard deviation of the wind speed increments (DeMarco and Basu, 2018). $\frac{\delta u(t)}{\sigma_{\delta u}} = 2$ indicates an increase in wind speed between two consecutive measurements (here $\tau = 10$ minutes) of a magnitude that is equal to two standard deviations of wind speed changes computed from the entire time series, and thus lies in highest 2.5% of values. Conversely, $\frac{\delta u(t)}{\sigma_{\delta u}} = -2$, has a similarly low probability but is associated with a large magnitude decline in wind speed between two consecutive measurements.’

We have also expanded the paragraph that links to Figure 5 to read:

‘The second component of power quality is the intermittency in terms of the probability and magnitude of ramp events – that is rapid changes in wind speed and/or power production. Wind speed time series at 150 m height from the NYSM and NYSERDA LiDARs indicate clear similarities in terms of ramp event magnitude and frequency to those derived using data from the FINO1 platform in the North Sea, Cabauw onshore in western portion of the Netherlands, Høvsøre in coastal Jutland, Denmark, and NWTC in the foothills of the Colorado Rocky Mountains (DeMarco and Basu, 2018) (Figure 5). Data from the NYSERDA buoys indicate a low probability of wind speed ramps of all magnitudes relative to the NYSM LiDARs (Figure 5), and all LiDAR time series indicate a substantially higher probability of a ramp-up (increase) than a ramp-down (decrease) of a given magnitude in wind speeds. Wind speed ramps in hourly ERA5 data exhibit a narrower distribution owing to spatial and temporal averaging, illustrating the need for in-situ data for capturing high resolution wind variability (Figure 5). Consistent with the lower probability of large-magnitude rapid changes in wind speed offshore, data from the NYSERDA buoys (Hudson North E05 and Hudson South E06) indicate probabilities of a wind power ramp with $> \pm 20\%$ change in power

are considerably lower than those from any of the onshore locations (Figure 5). Thus, the chance of experiencing an increase or decrease in electrical power production of 20% from one 10-minute period to the next is substantially lower for wind turbines deployed offshore. This indicates that wind turbines deployed offshore are likely to exhibit less intermittency in terms of electrical power production which is critical to efficient grid integration (Ayodele et al., 2012).'

We have also slightly modified the caption to Figure 5 to read:

'Figure 5. Left: probabilities of wind speed ramp events computed from the 10-minute data from the NYSERDA LiDAR buoys and the NYSM sites computed using Equation (3), and reported for four locations at or near operating wind turbines: FINO 1 is (offshore) in the North Sea, Cabauw is in the western portion of the Netherlands, Høvsøre is in Jutland, Denmark, and NWTC is in the foothills of the Colorado Rocky Mountains (data digitized from: (DeMarco and Basu, 2018)). Wind ramps computed from the hourly ERA5 output are shown by the gray polygon. Right: probabilities of wind power production ramp events at the locations of the NYSERDA buoys and the NYSM sites computed by applying the power curve for the IEA 15 MW reference wind turbine to the LiDAR wind speeds. The probabilities of no-change (i.e., power \pm 0%) are not shown to aid visibility.'

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?

The reviewer is referring to the following: 'The third aspect of power quality is predictability. The autocorrelation in power production at different time lags for the NYSM LiDARs exhibit clear diurnal oscillations and shorter e-folding time scales.'

We hope that adding text in Introduction

'Predictability and persistence of wind speeds and power production (Haghi et al., 2013; Haslett and Raftery, 1989). Within liberalized electricity markets, wind farm owner/operators bid in advance (e.g. 24 hours in advance) and are charged penalties for any imbalance between the bid and actual production (Pinson et al., 2007). Hence, accurate forecasts of wind generation are important to reduce penalties and maximizing revenue (Barthelmie et al., 2008). Persistence models where the power production at some future time is model as a function of power production in the recent past is often used as a benchmark forecast against which more sophisticated short-term power production models are compared (Kariniotakis et al., 2004). Many statistical short-term forecast models are predicated in part on persistence (Zeng and Qiao, 2011) and thus are most skillful when the power production time series exhibits high temporal autocorrelation. We quantify the temporal autocorrelation of power production from each onshore and offshore site and compare the degree to which electrical power production from the onshore and offshore locations differ with respect to persistence and short-term predictability.'

And in the Methods to this sentence:

'Temporal autocorrelation coefficients of the power production time series are used to derive e-folding time scales (i.e. the time delay at which the correlation coefficient drops to e^{-1} , i.e. to 0.37) which is used to represent the time scale at which the system 'loses' the memory of the initial state (Wilks, 2011).'

And this where we discuss Figure 5:

'These relatively large e-folding times for the buoy locations indicate a longer atmospheric 'memory' at these sites, indicating the potential for more accurate short-term power prediction forecasts because each time step is strongly dependent on previous time step(s).'

Clarifies this matter.

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? Please see the answer directly above.

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

We have added these two sentences that we believe makes the inference more concrete:

‘The implication is that large wind turbines deployed in these locations may experience a relatively high frequency of large unbalanced rotor loads and reduced component lifetimes unless such loads can be appropriately compensated (Hur et al., 2017).’

And

‘It is important to acknowledge that comparisons of LLJ climates derived from LiDAR measurements and WRF modelling should be done cautiously and that LLJ detection from the LiDAR wind speed profiles is critically dependent on unbiased data availability. Nevertheless, this analysis suggests LLJ within the rotor plane, as a source of large unbalanced rotor loads and reduced blade lifetimes, are less frequent at these onshore locations.’

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?

We calculate the demand as follows (see Methods):

‘Electrical demand (in MWh) for New York state are also presented and mean values are computed for each hour of the day and each month of the year based on hourly values for 2016-2022 as reported by the U.S. Energy Information Administration (EIA) hourly electric-grid monitor (https://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/US48/US48).’

And we normalize it as noted in the caption to Figure 9: ‘The data are normalized to a mean value of 1 and so that values of 0.9 or 1.1 in a given hour or month indicates WPP or demand that is 10% below or above the mean, respectively.’

We have expanded the paragraph that relates to Figure 9 to read as follows:

‘Electricity demand in New York state tends to peak in the afternoon (~ 1700 eastern standard time, EST) and in summer (highest values in July), though a secondary maximum occurs in January (Figure 9). Wind power production calculated from the NYSM/NYSERDA LiDARs and ERA5 grid-cell data ($P150_{ERA5}$) show highest values at night (0100 to 0500 EST) and during winter to spring (December-April), with the lowest production during the day (1300 to 1600 EST) and during the summer (Figure 9). Wind power production estimated based on LiDAR data from the NYSEDA buoy locations exhibits markedly lower diurnal and seasonal variability than is estimated at the NYSM sites, varying by $\pm 10\%$ around the mean versus $\pm 25\%$ at NYSM. This results in a reduction in mean absolute error (MAE) between time series of normalized WPP from the offshore LiDAR and electricity demand on both diurnal and seasonal timescales. The MAE computed from the mean hourly offshore WPP and demand is 0.19 when computed over the 24 hours of the day (Figure 9a) and 0.13 when computed from the time series of monthly mean values (Figure 9b). Both are smaller than MAE computed from WPP from the onshore LiDARs and demand on these time scales which are 0.25 and 0.20, respectively. This implies there will be better matching to electricity demand for power production from wind turbines deployed offshore.’

We hope this clarifies.

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.

We regret the reviewer found this section challenging to interpret. We have supplemented materials in the conclusions to help the reader. The conclusions now read:

‘Comparative analyses of wind resources and projected power production quantity and quality at onshore and offshore locations have been hampered by the lack of high-quality hub-height wind speed

observations. Here we use uniquely detailed LiDAR measurements from an onshore profiler network and offshore campaign to compare projections of potential power generation quantity and quality from offshore and onshore locations in New York State (Figure 1). Returning to the study objectives articulated in section 1, the study results indicate there are significant benefits to offshore deployments of wind turbines:

- Wind resources at locations in the New York Bight (coastal offshore areas southeast of New York state, Figure 1) greatly exceed those of all onshore locations within New York state. The mean wind speeds at 150 m (\bar{U}) offshore are above 10 ms^{-1} , while \bar{U} is below 8 ms^{-1} at all onshore sites. Weibull distribution fits to the 10-minute wind speed time series indicate scale parameters that are higher by 2 ms^{-1} than all onshore locations (Figure 4) except EHAM which is on Long Island and is within 1 km of the coastline (Figure 1). Accordingly, energy densities are 40% higher offshore and power production estimated offshore using the power curve of the IEA 15 MW wind turbine (Figure 2) yield over twice the AEP estimated for all onshore sites except EHAM (Table 2). Power generation estimated from wind speed time series offshore also exhibits lower variability on diurnal and seasonal time scales (Figure 6) and improved matching to current electricity demand in New York State (Figure 9). This implies that not only is the offshore resource considerably larger offshore, but the ability to meet electricity demand is better for wind turbines deployed offshore.
- Analyses presented herein also suggest that power generation intermittency is lower for the offshore sites. The probability of wind speeds below cut-in or above cut-out for the IEA reference wind turbine is lower offshore, as is the probability of large magnitude wind speed and power ramps (Figure 5). For example, the probabilities of wind power ramps with $> \pm 20\%$ change in power over a 10-minute period are less than half as probable offshore as onshore. The higher temporal autocorrelation of wind power production offshore (Figure 6 and Table 2) may also aid the accuracy of short-term wind power forecasting for wind turbines deployed offshore, yielding economic benefits to wind farm owner/operators and enabling grid integration.

Conversely, the frequency of anomalous wind speed shear and LLJ close to, or within, the rotor plane computed from the NYSM LiDAR wind speed profiles are slightly higher than those previously reported for the offshore areas from numerical simulations (Aird et al., 2022) but LLJ also exhibit higher elevations of the jet cores (Figure 8) and thus may be of less concern to wind turbine loading.

An analysis of the distance dependence of the co-variability of power production derived from measured 10-minute mean wind speed time series at the onshore and offshore sites indicates that the non-parametric Spearman correlation coefficient drops below 0.4 at distances of about 350 km (Figure 7). This indicates that in order to ensure consistency of electrical power production from wind farms in New York state, major developments should be separated by more than 350 km. This information could be used to guide judicious selection of wind farm locations to minimize the probability of concurrent low generation from onshore and offshore sites.

Thus, in accord with a priori expectations, analyses presented herein indicate there are advantages to the emerging trend towards offshore wind energy deployments in terms of the wind resource and the expected power quality and predictability (reduced ramp events, higher probability of rated power, etc.). Despite higher project AEP for the offshore locations, the additional costs involved in installing and operating offshore wind farms results in higher LCoE estimates for the offshore sites (Table 2). LCoE estimates derived using AEP at the NYSM sites are 26 to 49 \$/MWh while estimates for the NYSEDA buoy locations are 62-64 \$/MWh. Nevertheless, projected LCoE from wind energy for all of the sites investigated here in NY are competitive with all other electricity generation sources, with the possible exception of utility-scale PV, and much less expensive than traditional sources such as coal and nuclear that, according to a recent analysis, have an unsubsidized LCoE of 65-152 \$/MWh and 131-204 \$/MWh, respectively (Lazard, 2023).'

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

We hope the rewrite of the introduction and conclusion helps – we are seeking to quantify the relative benefits of deploying wind turbines offshore.

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).

No its not a guess the statement “higher temporal autocorrelation of wind power production offshore (Figure 6 and Table 2) may also aid the accuracy of short-term wind power forecasting” can be readily made because virtually all statistical short-term forecasting methods employ directly or indirectly methods that are predicated on red noise characteristics of the atmosphere. See additional text that explains that matter.

Page 17, line 410-411: To what extent does this follow from the Spearman correlation coefficients drop? Exponential decay of atmospheric properties in time and space is well documented. We hope that explanatory text we have added about this matter helps.

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

We have moved some discussion of LCoE into section 4.1

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.

We do not concur with the reviewer. We think there is importance to having the reader readily be able to note the data availability with the monthly mean wind speeds.

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.

We have elaborated the text as requested.

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.

Text changes enacted as requested.

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

We regret the reviewer struggled to understand this figure. We have added this text to the caption:

'A LLJ frequency of 5% calculated from the LiDAR deployed at QUEE for the calendar month of May indicates that LLJ were indicated in 5% of all 10-minute periods during this month. Data in panels (a) and (b) indicate that at that site in the month of May the associated LLJ mean core wind speed is 11 ms^{-1} and the mean core height above ground is 330 m.'

Reviewer #2:

General comments:

The paper presents data from a set of onshore and offshore lidars. The paper mostly presents statistics based on these data and the results of the analysis are as expected, as the wind blows more and more steady offshore compared to onshore. There is no new methods, concepts or ideas introduced in the manuscript, so I have rated the scientific significance as low. Nonetheless, the paper could be useful for somebody that is looking specifically for information about the wind climate in this region.

We regret the reviewer did not find the scientific significance to be higher. The lack of measurements at

wind turbine hub-heights and above has long been a source of concern for the wind energy community so we believe that the availability of network based LIDAR data will be of general interest. Further, we have found no previous research that used these types of data sets to quantitatively compare; the resource, power quality, demand matching etc on – and offshore in the same climate zone. And yet these properties are of critical importance in charting the future of wind energy deployments.

My main comment on the analysis itself is about the low data recovery percentage of the lidar data. It is not demonstrated that there is no correlation of when data recovery is low and what the wind climate is. For example, one would expect that the lidars return 'not available' when a measurement cannot be obtained. Most of these data will be during low wind speed conditions when there is not enough aerosols to measure the wind. You have your long-term measurement time series from ERA5, so you could correct for this. Also in general I miss some discussion of the type of lidars you are using, because they are not the same offshore (zephyr) and onshore (windcube). What kind of filtering was done (precipitation? CNR?).

It is indeed regrettable that the NYSM network does not have higher data recovery. We asked this question of the operators of the NYSM but they are not able to provide further information beyond what is available via the readme documentation (available at: http://www.nysmesonet.org/networks/profiler#stid=prof_alba) which simply states: 'Sensor and/or system failures are not uncommon as the Profiler equipment are sensitive to a variety of environmental factors. Data gaps may be due to sensor failures; calibration errors; power failures; and/or communication failures. ... Only manufacturer-developed QA/QC procedures are applied to the data and there might still be some undetected errors.'

More information is available regarding the NYSERDA buoy deployments – e.g. via technical reports available for download from:

<https://oswbuoysny.resourcepanorama.dnv.com/download/f67d14ad-07ab-4652-16d2-08d71f257da1>

We have added a note to that effect to the Data availability statement: 'Reports documenting LiDAR performance verification are also available for download from:

<https://oswbuoysny.resourcepanorama.dnv.com/download/f67d14ad-07ab-4652-16d2-08d71f257da1>.'

Regarding the comment 'It is not demonstrated that there is no correlation of when data recovery is low and what the wind climate is.' We think there are two concepts of importance here:

- 1) We do document at the monthly scale data availability versus wind speed.
- 2) There is no documented evidence from NYSM that the LiDAR scans failed to report wind speeds due to low aerosol concentrations but given they state 'manufacturers QA/QC procedures were employed it is highly likely a CNR screen was employed'. That having been said air quality measurements in New York state DO NOT imply a high frequency of sufficiently low aerosol concentrations to render these LiDARs likely to be unable to operate (see for example; Squizzato, S., Masiol, M., Rich, D. Q., & Hopke, P. K. (2018). A long-term source apportionment of PM_{2.5} in New York State during 2005–2016. *Atmospheric Environment*, 192, 35-47.)

We have added this text to section 4.1; 'Documentation associated with the data set notes the causes as; 'calibration errors; power failures; and/or communication failures.' And further notes 'Only manufacturer-developed QA/QC procedures are applied to the data and there might still be some undetected errors.' (readme accessible from

http://www.nysmesonet.org/networks/profiler#stid=prof_alba).

Technical comments:

l13: factor -> parameter

Done.

l166: conventional usage would be capital gamma

changed as requested.

I211-215: since the lidar signal depends on aerosol concentration this method will likely miss many low-level jets as the lidar will simply not return a signal above the jet. Would be good to discuss this.

We have added this cautionary text:

'It is important to acknowledge that comparisons of LLJ climates derived from LiDAR measurements and WRF modelling should be done cautiously and that LLJ detection from the LiDAR wind speed profiles is critically dependent on unbiased data availability. Nevertheless, this analysis suggests LLJ within the rotor plane, as a source of large unbalanced rotor loads and reduced blade lifetimes, are less frequent at these onshore locations.'

In addition: I am not quite sure how to interpret the comment about the comparability with the 500 m height: did you use data up to 500 m? It would be good to show what the recovery percentage is at this height, related to the remark above.

We have calculated the data recovery rates at each LiDAR location at each height < 500 m and now report them, see text that's reads; 'Analyses of wind speed data from the NYSM LiDARs at all measurement heights from 100 to 500 m indicates that averaged across all stations the data availability as a function of height varies only by +/-2.5%.'

I232: move bracket from before Barthelmie to before 2023.

Done.