Correlations of power output fluctuations in an offshore wind farm using high-resolution SCADA data

Authors: Janna K. Seifert, Martin Kraft, Martin Kühn, and Laura J. Lukassen
DOI: 10.5194/wes-2020-96

Authors response to referee comments

Dear referees, we appreciate the time you again spend helping to improve our manuscript. Your constructive comments help us to improve our paper further. As recommended, we had our manuscript professionally proofread by a native speaker. Below we will discuss your comments in detail.

Authors response to comments from referee #1

RC1 *Overall I believe this line of work is important and relevant for the community and based on this I believe this work is suitable for publication in Wine Energy Science. The authors have replied in detail to the comments of the referees and the corresponding changes have improved the manuscript. Obviously, the analysis of the field data is a challenging task as articulated by the authors. I believe that in various places, as indicated below, the presentation in the manuscript can be further clarified.*

AC Thank you for your positive feedback. We appreciate your effort and will incorporate your remarks as follows.

Comments on previous review

Referee #1

RC2 *RC4: So to clarify; the data on U is provided to you (and is based on Power signal) and you did not calculate it yourself.*

AC Yes, that is correct. U is not calculated by us but provided as is without detailed information.

RC3 *RC6: The slight asymmetry is mentioned in the conclusions (line 340), but I am not sure it is mentioned clearly in the body of the text, please double check.*

AC We double checked the text and we mention the asymmetry first here l. 83f: "They are installed in a grid-like, slightly asymmetric pattern with a triangular shape towards the south (see Fig. 1)." And again here l.252f: "Figure 4 shows the average power output fluctuation correlation around 90° and 270° in detail as cuts through Fig. 3. The absolute peaks are at 90° and 270°. For wind directions where the wind turbines in a pair are less streamwise aligned, the peak decreases, and the correlation curve flattens. In contrast to 90°, the correlations for 270° are more defined and show slightly larger peak values. This may be due to the asymmetric wind farm layout (cf. Fig. 1).

RC4 *The information provided as response to RC11 is very useful, but I could not find this information in the manuscript. I think the answer can be summarized shortly in the manuscript, and the authors can refer to the graph provided in the response document, which is already available online. So the graph itself does not need to be incorporated in the manuscript. But it is useful to inform the reader of the main manuscript that it is available.*

AC We agree that this information is helpful for the reader and decided to include it in the appendix of the paper to enable the reader to check on the information directly. In the text we added:

p.6, ll. 154f: Depending on the data availability, the next interval of 600 consecutive seconds could go from $[t_j + 1\,\mathrm{s}, t_j + 1\,\mathrm{s} + 599\,\mathrm{s}]$, and thus overlap the previous one up to 599 s. This does not result in significantly different findings compared to non-overlapping intervals as shown in App. B.

RC5 *RC20: the provided answer does not really answer my question. The provided answer states the general benefit of the approach, but I am wondering whether the authors can clarify the specifically new insights obtained for the case considered here.*

AC To clarify further clarify this, the other cluster approaches could help to find more defined correlation curves (higher and more narrow peaks) or to find further clusters. We revised the text as follows:

p. 18, ll. 409f: In addition, it is worth considering alternative clustering methods like k-medoids (Kaufman and Rousseeuw, 2008), which is less sensitive to outliers or Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996) which is also less sensitive to outliers and has no fixed cluster shapes and fixed number of clusters. Using these algorithms could improve the clustering of the intervals and more defined correlation curves or could identify further clusters.

Referee #2

RC6 *RC19: The authors state "The 20° interval is the result of the 10-degree tolerance in the wind direction measurements of the wind turbines". I think this information is very important and it should be incorporated in the main text.*

AC We agree. Thank you for pointing this out. We revised the text as follows:

p. 8, ll. 209f: In the following, we average correlations over a wind direction interval of 20° and all available time intervals of the considered wind turbines (either all wind turbines or a selection of wind turbines). We consider 20° intervals due to a 10° tolerance in the wind direction measurements of the wind turbines.

Specific comments

RC7 *I also read the manuscript again. In various places the authors could be more precise on the formulation, i.e. making clear to what case, data, parameter is exactly referred. Or the English grammar should be removed. I provided a list of some of the main instances below, but please check the entire manuscript, especially the abstract, on this.*

AC We agree and revised the text accordingly. Please see the track of changes as well as the answers to your specific comments below.

RC8 *"Space-time correlations of wind turbine pairs" ==> the power output is correlated; not the turbine pairs*

AC Thank you for pointing this out. Revised as follows:

p.1, ll. 1f: Space-time correlations of power output fluctuations of wind turbine pairs provide information on the flow conditions within a wind farm and the interactions of wind turbines.

RC9 *"Such information plays an important role for the control" ==> Can provide important insights for controls, i.e. information obtained from the analysis you presented is not yet used in wind farm control*

AC Thank you for noting this. The text was revised as follows:

p.1, ll. 2f: Such information can play an essential role in controlling wind turbines and short-term load or power forecasting.

RC10 *which overcomes the challenge of highly variable flow conditions within the wind farm ==> make this a new sentence. New it refers to the SCADA data. However, this statement is on your analysis approach.*

AC We agree, the text was revised as follows:

p.1, ll. 4f: Here, we present an approach to investigate space-time correlations of power output fluctuations of streamwise-aligned wind turbine pairs based on high-resolution SCADA data. The proposed approach overcomes the challenge of spatially variable and temporally variable flow conditions within the wind farm. We analyse the influences of the different statistics of the power output of wind turbines on the correlations of power output fluctuations based on eight months of measurements from an offshore wind farm with 80 wind turbines.

RC11 *Wind direction investigations show ==> More effect of the wind direction.*

AC That is correct. The text was revised as follows:

p.1, ll. 8f: First, we asses the effect of the wind direction on the correlations of power output fluctuations of wind turbine pairs. We show that the correlations are highest for the streamwise-aligned wind turbine pairs and decrease when the mean wind direction changes its angle to be more perpendicular to the pair.

RC12 *line 21 ==> average size of [new]? Wind farms*

AC We clarified this as follows:

p.1, ll. 23f: Due to the increased size of the newly installed wind farms, the average size of offshore wind farms rose to 621 MW (Ramírez et al., 2020).

RC13 *line 44: "Furthermore, the variance of the wind velocity and the mean velocity turned out to be important parameters in the modelling set up." ==> Please clarify to which modeling setup you are referring.*

AC Thank you for this note. We clarified the text as follows:

p. 2, ll. 47f: In an LES study by Lukassen et al. (2018), space-time correlations of velocity fluctuations within a wind farm with periodic boundary conditions (modelling a periodic array of wind turbines) were analysed and modelled analytically. Velocity fluctuations, which are directly related to power output fluctuations, showed pronounced space-time correlations. Furthermore, the variance of the wind velocity and the mean velocity turned out to be important parameters in the space-time correlation model.

RC14 *line 58: "includes unstable inflow conditions," ==> please be more precise. LES and wind tunnel data also include unsteady effects. However, typically LES and wind tunnel have constant wind directions, while wind direction varies in field data, etc.*

AC We agree and revised the text as follows:

> p. 3, ll. 61f: In contrast to the wind tunnel measurements by Bossuyt et al. (2017) and the LES analysis by Lukassen et al. (2018) mentioned above, the data set processed here includes unstable inflow conditions (varying wind speeds and wind directions), dynamically operating wind turbines as well as changing flow conditions within the wind farm.

RC15 *line 70: "relevant wind turbine statistics" ==> statistics on the power production of the turbines not statistics of the turbines themselves.*

AC Revised as follows:

> p.3, ll. 76f: With this and the results from the LES analysis by Lukassen et al. (2018), we identify relevant wind turbine power output statistics that influence the correlation.

RC16 *line 95: Please be more precise on what wind direction fluctuations you refer to. Fluctuations around some mean value? The average wind direction does namely matter.*

AC Yes, that is correct. Here, wind direction fluctuations around a mean wind direction are considered. We added this in the text as follows:

> p.2, ll. 54f: Dai et al. (2017) analysed 1 Hz wind farm SCADA data concerning the influence of wind speed fluctuations around a mean velocity and wind direction fluctuations around a mean wind direction on the wind turbine power output fluctuations of single wind turbines.

RC17 *line 109-110: "All these factors multiply to an order of unpredictable variability" ==> please clarify your formulation.*

AC We revised the text as follows:

> p. 5, ll. 115f: The combination of all these factors causes highly dynamic flow conditions and thus, an unpredictable variability.

RC18 *line 135: As mentioned before, these effects have a limited ==> sentence is not complete.*

AC Thank you for noting this. We completed the sentence as follows:

> p. 6, ll. 142f: As mentioned before, these effects have a limited effect on the power output fluctuations of the wind turbines.

RC19 *line 186: why do you have the reference of 13ms^{-1}?*

AC The value was empirically chosen based on the power curve of the wind turbines considered here. Full load is reached at 12.5 ms^{-1}; thus, 13 ms^{-1} (including tolerance) was chosen as we only consider partial load. We clarified this in the text as follows:

> p. 7, ll. 187f: Next, the correlation curves with the normalised lag $\tau_{norm,intv}$ are discretised using a histogram with a reference time lag of
>
> $$\tau_{norm} = \tau \cdot \frac{U_{max}}{x_{AB,mean}} \quad (1)$$

where τ is the time lag (0 s to 300 s), U_{max} is an artificially introduced velocity that has to be at least equal to the maximum possible wind speed to fit all normalised curves ($U_{max} = 13\ \mathrm{ms}^{-1}$ for this case). This value is based on the wind turbine power curve characteristics, including a tolerance as the wind turbines considered here reach their rated power at $12.5\ \mathrm{ms}^{-1}$.

RC20 *Figure 2 /3: For certain wind directions you have less measurements. The gap seems to be at slightly different wind directions in figure 2 and 3 (left / right of the 0/360 degree line). I guess it has to do with the way you defined your measurement intervals, but please make sure this is clarified to the reader.*

AC Thank you for pointing this out. We revised the text as follows:

p. 8, ll. 230f: Due to the varying data availability per wind direction and the applied data filtering (see sec. **??**), the average correlation curve per bin is based on a different number of data. It turns out that after filtering, no data is available for the bins around 330° to 10°.

RC21 *line 224: "are ideal compared to free field measurements" ==> I agree. Can you clarify to the reader what specific aspects you believe are the main reason for the observed differences between field data and LES and wind tunnel data?*

AC We believe the main reasons for this are the wind direction and wind speed changes and the individual operation of the wind turbines (yawing, limited power, shut off). To further clarify this, we added the following text:

p. 9, ll. 237f: These dynamics are most likely caused by the wind direction and wind speed changes and the individual operation of the wind turbines (yawing, limited power, shut off). Even though the correlation curves were adapted to the average wind speed per interval, the wind speeds were just an assessment and could change during the interval. Also, the wind direction is averaged over the whole wind farm, which means certain intervals could include data from wind turbine pairs facing a slightly different direction. Further, we only consider the intervals of wind turbine pairs that fit the data filter; however, other wind turbines could be yawing at the same time or start pitching. Thus, the flow within the wind farm could still be influenced by these wind turbines.

RC22 *line 337: "more defined averaged correlations" ==> please clarify*

AC Here we would like to point out that the peaks of the average correlation curves for 270° are more narrow, thus more defined. We revised the text as follows:

p. 16, ll. 364f: In general, we found that the averaged correlation curves of power output fluctuations for 270° with a maximum correlation coefficient of 0.21 have more defined (narrower) peaks compared to those of the averaged correlation curves for 90° with a maximum correlation coefficient of 0.16.

RC23 *line 339: "introduced parameters" ==> please clarify what parameters you refer to.*

AC We revised the text as follows:

p. 17, ll. 366f: Further, the standard deviation of the power output fluctuations of the wind turbines in a pair was larger for 270° than for 90°.

RC24 *line 343: "more defined" ==> please clarify what you mean*

AC Similar to **RC22**, we revised the text as follows:

> p. 17, ll. 372f: We found different correlation curves for the rows of the wind farm, becoming more defined (more narrow peaks) towards the back of the wind farm.

RC25 *line 349: "turbine pairs in the same row are affected by the same flow conditions," ==> please clarify*

AC The flow conditions within the wind farm are dynamically changing due to the individual control and operation of the wind turbines. We clarified this in the text as follows:

> p. 17, ll. 379f: As mentioned before, the flow throughout the wind farm is highly variable due to the individual control and operation of the wind turbines. This means that not all wind turbine pairs in the same row are affected by the same flow conditions, as upstream wind turbines could be turned off, could be yawing or could be pitching. This further means that they show different correlation curves and should be sorted into different correlation states.

RC26 *line 354: "clearly distinguishable parameters" ==> What parameters do you mean?*

AC We revised the text as follows to clarify our focus:

> p. 18, ll. 385f: The clustering showed similar results for the wind directions 90° and 270°. The clusters had distinguishable values in the standard deviation of the power output fluctuations and in the normalised power output differences of the wind turbines, which were directly related to the average correlation curve per cluster.

RC27 *line 361: "The remaining clusters were not as significant as the other but also showed" ==> In this sentence it is not so clear to which clusters you refer.*

AC Thank you for noting this, we clarified this in the text as follows:

> p. 18, ll. 395f: Clusters 2, 3 and 4 were not as significant as Clusters 1 and 5 and showed distinguishable correlation curves with their peaks ranging from 0.14 to 0.22 for 90° and from 0.2 to 0.31 for 270°.

Authors response to comments from referee #2 - Mark Kelly

RC1 *The draft has been improved with revision; however, some issues still remain. These are pointed out below, with author comments (AC) addressed first, and later point-wise comments with line numbers referring to the file that showed revision/changes ("...ATC1.pdf").*

There are also numerous linguistic errors which need to be corrected (most of them new). As in the previous review, I again suggest asking someone with native-level English proficiency to proof-read the (updated) draft; they are too numerous to note individually here.

E.g., on l.6: "in free field" is not proper English; on l.8 the word 'of' is missing after "eight".

Before getting to specific points line-by-line, I'll respond to the author comments (ACs) that addressed my previous reviewer comments (RCs), and then I add some general comments.

AC Thank you again for your feedback. We hope the professional proofreading improved our manuscript. Please see the track of changes for the overall revision. In the following, we will address your comments specifically.

Comments on previous review

RC3 *AC11: It is an improvement to mention the "reconstructed" aspect of wind speed, but to be open/clear, why not include your AC statement "details on the reconstruction are not available" in the text? This and/or the unknown transfer function should be mentioned; e.g. the latter can affect the direction as well as the speed.*

AC Thank your for noting this, we added this information to the text as follows:

> p. 3, ll. 90f: The reconstructed wind speed U is not directly measured but provided as a variable that results from the measured power and control variables of the wind turbine (details on the reconstruction of U are not available).

RC4 *AC12: Yaw misalignment isn't only due to "false...measurements, calibrations, or sensor installation"–especially if there is a transfer function used for nacelle-mounted 2d-anemometers. Your addition on l.106-8 helps to allay this issue.*

AC Yes, we agree and revised the text as follows:

> p. 3, ll. 91f: Due to this, U is considered as an approximated and idealised value that does not include the wind speed independent power reduction, e.g. by a yaw misalignment of the wind turbine.

RC5 *AC15: Your response about yaw misalignment threshold and rate per 10-minutes is reasonable, but you have not included this in the revised text.*

AC We clarified our description of the selection of the wind direction as follows:

p. 6, ll. 138f: Since this data filter only applies to the average wind direction Φ_{av}, individual wind turbines might have a slightly deviating relative wind direction for this specific time interval. This deviation could be caused by a false wind direction measurement, a yawing process that has taken place asynchronously to the majority of other wind turbines or a wind direction deviation due to local changes over the area of the wind farm. This means there is no threshold for yaw misalignment within the 600 s intervals.

RC6 *AC17: How is your approach "similar to" Taylor's, but not actually simply assuming it? Again, it appears you've assumed it over the entire range of τ (and τnorm); this can become problematic for small enough U (large lags).*

AC We agree that Taylor's frozen eddy hypothesis could be problematic. However, we do not assume frozen eddies here. Turbulence structures travel downstream, but they are subject to change when doing so. We just use the mean velocity as an advection velocity for turbulent structures, which is similar to Taylor's hypothesis. In our case, small velocities (causing large lags) are not a problem as the power output fluctuations would not be correlated anymore in this case. To clarify this, we revised the text as follows:

p. 7, ll. 175f: Similar to Taylor's hypothesis (Taylor, 1938), we assume that the wind structures responsible for the power output fluctuations measured at an upstream wind turbine A that travel a certain distance to the downstream wind turbine B with an advection speed that is similar to the average wind speed over that distance. But in contrast to Taylor's hypothesis, we do not assume frozen eddies but expect wind structures to change and thus decorrelate while travelling downstream. Further, as we have no access to the average wind speed over the distance between wind turbines A and B, we use the average wind speed measured at wind turbine B as a reference.

RC7 *AC18 and l.189-205: you have used τnorm before it is defined in (4); this can be quite confusing for the reader, particularly if they have not read this before. Also, why is the second 'normalization' (4) done? If Umax = 13m/s always, and xAB; mean is also a simple constant for all cases, then why normalize again? If you have done this to force the peaks closer to 1, then this should be stated. Also, how was Umax chosen–doesn't this just arbitrarily squeeze/stretch the correlation curves (as you wrote for τnorm;intv)? There appears to be no physical justification for (3) and (4) together, unless perhaps you could explain what is meant by "at least equal to the maximum possible wind speed to fit all normalised curves".*

AC Thank you for pointing this out. We adjusted the order of the text and further clarified the usage of τ_{norm} and $\tau_{norm,intv}$. While $\tau_{norm,intv}$ is used to shrink and stretch the correlation curves based on the average wind speed within a time interval, τ_{norm} is a reference time lag which is only created to bin $\tau_{norm,intv}$ in a histogram. τ_{norm} is thus the reference time lag used for representing all processed results. U_{max} was empirically chosen based on the power curve of the here considered wind turbines. Full load is reached at 12.5 ms^{-1}, thus 13 ms^{-1} (including tolerance) was chosen as we only consider partial load. The text in the manuscript was revised as follows:

p. 7, ll. 180f: Hence, to compare the correlations calculated for intervals with different average wind speeds and different wind turbine distances, the time lag τ is normalised for each time interval starting at t_j individually

$$\tau_{norm,intv} = \tau \cdot \frac{\langle U_B(t+\tau)\rangle_{\Delta t_{300}}}{x_{AB}} \tag{2}$$

where $\tau_{norm,intv}$ is the normalised time lag, $\langle U_B(t+\tau)\rangle_{\Delta t_{300}}$ is the average reconstructed wind speed from a certain (downstream) wind turbine B for a time interval $\Delta t_{300} = 300$ s for t in the discretised interval $[t_j, t_j + 299\ \mathrm{s}]$ and a certain lag τ. This means for a certain τ, the averaging interval of $\langle U_B(t+\tau)\rangle_{\Delta t_{300}}$ is $[t_j+\tau, t_j+\tau+299\ \mathrm{s}]$. x_{AB} is the distance between wind turbine A and wind turbine B.

Next, the correlation curves with the normalised lag $\tau_{norm,intv}$ are discretised using a histogram with a reference time lag of

$$\tau_{norm} = \tau \cdot \frac{U_{max}}{x_{AB,mean}} \tag{3}$$

where τ is the time lag (0 s to 300 s), U_{max} is an artificially introduced velocity that has to be at least equal to the maximum possible wind speed to fit all normalised curves ($U_{max} = 13\ \mathrm{ms}^{-1}$ for this case). This value is based on the wind turbine power curve characteristics, including a tolerance as the wind turbines considered here reach their rated power at 12.5 ms^{-1}. $x_{AB,mean}$ is the average distance between wind turbine A and wind turbine B of the considered wind turbine pairs. Note that $\tau_{norm,intv}$ is used for stretching and shrinking of the correlation curves. τ_{norm} is only a reference time lag that is only created for binning of the stretched or shrunk correlations and does not change the correlation curves.

RC8 *AC20: The response statement "the temporal autocorrelation of a wind turbine decorrelates in the considered time intervals of 300 s" does not make sense. Do you perhaps mean that the the correlation decreases to effectively 0 as lag (τ) approaches 300s?*

AC Yes, exactly. We meant that the correlation decreases to 0 within the considered time lag of 300 s.

RC9 *AC23: If you are to insist on using the term "filtering" in place of 'data selection' or similar—knowing that WES is not a data science journal, but a wind energy journal where you are also mentioning turbulence—then you should at least include the word 'data' before it. Further, I strongly recommend section 2.1 to be renamed "Data selection and filtering" or similar—again, spectral filtering is commonly used when dealing with this kind of data in wind energy (particularly turbulence), especially when mentioning different intervals (e.g. 600s).*

AC Thank you very much, for this valuable suggestion. We renamed the section 2.1 to "Data selection and filtering" and changed the term 'filtering' to 'data filtering' in the text. For the latte, please see the track of changes.

General comment

RC10 *The labelling of peak correlations (between power fluctuations for turbine pairs) as "correlation states" is contentious, since 'states' implies different physical scenarios or flow/operational-regimes—particularly if you have not described anything like the latter. If the 'states' are basically different groups of peak correlations (magnitudes), then why not call them that? This is safer, because for different flow regimes having larger/smaller turbulence length scales (and/or other farms having different spacing, surfaces, or even hub height), then the magnitudes or groups could be quite different. Also, in the conclusion it should be mentioned how/why seemingly insignificant peak correlations (e.g. 0.2 or less) are meaningful (compared to commonly understood statistical significance being ρ_{AB} > 0.5); i.e. the relative values are significant given the 'noisy' turbulent flow, in addition to values consistent with others' LES results.*

AC We use the term states indeed to describe a group of certain correlations or correlation curves. However, these states are identified especially for this wind farm and the here considered time period. To emphasize this, we adjusted the definition of correlation states in the text as follows:

p. 3, ll. 64f: In this paper, we investigate the influencing factors on the correlation of power output fluctuations of wind turbine pairs and introduce parameters to distinguish different correlation curves, herein called correlation states. A state defines a group of similar correlation curves. Note that the states found here refer to this specific wind farm and the considered time period.

Hence, our understanding of the term correlation states is not a 'global' but a 'local' one. We do not expect to find exactly the same correlation states for other time periods or other wind farms. However, the process of the identification of correlation states stays the same and applies to any other condition.
The evaluation of the peak correlations in the conclusion was revised as follows:

> p. 17, ll. 368f: In the context of the considered highly varying flow conditions, peak correlations around 0.21 or 0.16 are still considered significant. The cause for these relatively low peak correlations lies in the varying flow conditions or noisiness of the flow within the wind farm.

Specific comments

RC30 *l.5, 7: the challenge is spatially variable flow, not just "highly" variable flow.*

AC You are right, it is spatial variable but also temporally variable. To shorten the expression we stick to highly variable. We revised the text as follows:

> p.1, ll. 4f: Here, we present an approach to investigate space-time correlations of power output fluctuations of streamwise-aligned wind turbine pairs based on high-resolution SCADA data. The proposed approach overcomes the challenge of spatially variable and temporally variable flow conditions within the wind farm.

RC31 *l.10-11: "decrease towards spanwise pairs" doesn't quite make sense. If the correlations decrease with angle between mean wind direction and pair separation vector, why not write that?*

AC Thank you for noting this. We revised the text as follows:

> p. 1, ll. 8f: First, we asses the effect of the wind direction on the correlations of power output fluctuations of wind turbine pairs. We show that the correlations are highest for the streamwise-aligned wind turbine pairs and decrease when the mean wind direction changes its angle to be more perpendicular to the pair.

RC32 *l.13-17: "the correlation of streamwise aligned wind turbine pairs" should be 'power correlations between streamwise-aligned wind turbine pairs'.*

AC We agree and revised the text as follows:

> p. 1, ll. 11f: Further, we show that the correlations for streamwise-aligned wind turbine pairs depend on the location of the wind turbines within the wind farm and on their inflow conditions (free stream or wake).

RC33 *l.18: sorting is accomplished by a "k-means clustering algorithm", not "clustering algorithm k-means" (e.g. Likas, Vlassis, & Verbeek 2003).*

AC Thank you for pointing this out. We revised the text accordingly:

> p. 1, ll. 15f: For this, we employ the data-driven k-means clustering algorithm to cluster the standard deviations of the power output fluctuations of the wind turbines and the normalised power difference of the wind turbines in a pair.

RC34 *.19-20: the sentence "These groups are here referred to as correlation states." is not needed in an abstract.*

AC We agree and erased this sentence in the abstract:

> p. 1, ll. 12f: Our primary result is that the standard deviations of the power output fluctuations and the normalised power difference of the wind turbines in a pair can characterise the correlations of power output fluctuations of streamwise-aligned wind turbine pairs.

RC35 *l.20 repeats l.7-8.*

AC The last sentences of the abstract were adjusted as follows:

> p. 1, ll. 15f: For this, we employ the data-driven k-means clustering algorithm to cluster the standard deviations of the power output fluctuations of the wind turbines and the normalised power difference of the wind turbines in a pair. Thereby, wind turbine pairs with similar power output fluctuation correlations are clustered independently from their location. With this, we account for the highly variable flow conditions inside a wind farm, which unpredictably influence the correlations.

RC36 *l.18-22: "these parameters" is repeated three times; the final point is also somewhat of a repetition of l.16-17...The abstract can be cleaned up (there are more English errors in it as well).*

AC Thank you for pointing this out. We totally agree and revised the whole abstract. Please see the track of changes for details.

RC37 *l.71: Use of "correlation curves" here to mean "states" is ambiguous and confusing; however, in l.194 and after it is used reasonably, to refer to the actual $R(\tau)$ curves. Here in this context of "states" you are really referring to the peak correlation, as seen later in e.g. Figs.7-8.*

AC Thank you for noting this. We agree and revised the text as follows:

> p. 3, ll. 67f: The parameters introduced to characterise correlation curves are then evaluated with a data-driven clustering algorithm to group the data according to the underlying correlation curves.

RC38 *l.151: sentence is not finished.*

AC Thank you for noting this. The sentence was finished as follows:

> p. 6, ll. 142f: As mentioned before, these effects have a limited effect on the power output fluctuations of the wind turbines.

RC39 *l.196: "noted as" –> 'denoted by'*

AC Revised.

RC40 *l.212 and elsewhere: "dependency" should be 'dependence'*

AC Revised. Please find all adjustments in the track of changes.

RC41 *l.242: how does τ_{norm} > 1 have any meaning, if τ_{norm} is arbitrary due to its definition via the artificial Umax?*

AC Our answer to **RC7** most likely already helps to clarify this. As τ_{norm} is only used to bin the shrunk or stretched correlation curves, $\tau_{norm} = 1$ still represents the case when the advection speed of the fluctuations equals the wind speed measured at wind turbine B. The binning only slightly affects the peaks of the correlations due to the averaging of the values inside the respective bins, which also applies to the bin around $\tau_{norm} = 1$.

RC42 *l.381: exactly which standard deviation?*

AC We clarified this sentences as follows:

> p. 17, ll. 366f: Further, the standard deviation of the power output fluctuations of the wind turbines in a pair was larger for 270° than for 90°.

References

Bossuyt, J., Howland, M. F., Meneveau, C., and Meyers, J.: Measurement of unsteady loading and power output variability in a micro wind farm model in a wind tunnel., Experiments in Fluids, 58, 1–17, https://doi.org/10.1007/s00348-016-2278-6, 2017.

Dai, J., Cao, J., Liu, D., Wen, L., and Long, X.: Power fluctuation evaluation of large-scale wind turbines based on SCADA data, IET Renewable Power Generation, 11, 395–402, https://doi.org/10.1049/iet-rpg.2016.0124, 2017.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96, p. 226–231, AAAI Press, 1996.

Kaufman, L. and Rousseeuw, P.: Partitioning Around Medoids (Program PAM), chap. 2, pp. 68–125, John Wiley & Sons, Ltd, https://doi.org/10.1002/9780470316801.ch2, 2008.

Lukassen, L. J., Stevens, R. J. A. M., Meneveau, C., and Wilczek, M.: Modeling space-time correlations of velocity fluctuations in wind farms, Wind Energy, 21, 474–487, https://doi.org/10.1002/we.2172, 2018.

Ramírez, L., Fraile, D., and Brindley, G.: Offshore Wind in Europe - Key trends and statistics 2019, https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Offshore-Statistics-2019.pdf, last access 21.01.2021, 2020.

Taylor, G. I.: The Spectrum of Turbulence, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 164, 476–490, https://doi.org/10.1098/rspa.1938.0032, https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1938.0032https://royalsocietypublishing.org/doi/10.1098/rspa.1938.0032, 1938.