Response to referee #1

Mark Schelbergen, Peter C. Kalverla, Roland Schmehl, and Simon J. Watson

Thank you for the comprehensive comments. We feel that they were very helpful for increasing the quality of the paper to the current level. Your comments, together with those of referee #2, led to a thorough revision of the paper. The most important changes to the paper include:

1. Including information on orography
2. Discarding the lidar discussion
3. Using one stability metric: the Obukhov length, and corresponding classification for identifying stability trends within the clusters
4. Fitting logarithmic profiles only to the lower part, i.e. <200 m, of the mean profile shapes

We respond to the referee comments by including our answers below the original comments. Our answers are preceded by one or both of the following labels:

AR = author's response

AC = author's changes in manuscript

1 General comments

This paper is a useful contribution to better understand the wind energy potential of airborne wind energy systems (AWESs). The investigated onshore and offshore wind regimes make it especially interesting for regions close to the shore such as the Netherlands which this papers wind data is based on. Simulated Dutch Offshore Wind Atlas (DOWA) data is normalized, transformed using principal component analysis (PCA) and clustered to generate generalized wind profiles which are then scaled and fed into a quasi-steady AWES model to estimate power curves and annual energy production (AEP). However, this paper would benefit from a more detailed explanation, validation and justification of the described process. Following are some of the general questions and comments that require further explanation.

Language: Please revise the writing of this paper with regards to the usage of active voice to form shorter more concise sentences. Avoid sentences such as: " Finally, it is demonstrated how a set of wind profile shapes and their statistics can be used to estimate the AEP of a pumping AWE system" (page 1, line 9). Section 2 highlights some but not all text passages with

passive voice. Remove repetitive sentences, combine sentences where possible. Avoid filler words and obvious wording such as “vertical wind profile”. Some line breaks seem unnecessary as both the preceding and following paragraph are related or continue the same topic.

[AR] We have tried to incorporate the active voice as much as possible based on your comments. It is however not always used, as sometimes it feels less appropriate. We feel that the language has improved substantially, e.g., repetitive sentences and obvious words are omitted and the phrasing has become more precise.

Figures: Many figures are very similar. The paper might benefit from focusing on one location and moving the figures showing the other locations to the appendix. More detailed captions would improve the understanding of your paper, especially for people who are skimming through the text or are reading the paper for the first time. Try to be consistent with the labels within figures as some figures use circles and others use squares.

[AR] The clusters of the offshore, onshore and multi-location analyses are considered to be the core of the paper, therefore, putting them in the appendix would not be appropriate. Please find more explanation later on in this document.

[AC] A more consistent labelling is introduced. Now, only the clusters are referred to using the acronyms: MMIJ, MMC, and ML. We have made the captions a lot more informative. Furthermore, we have discarded the lidar discussion as it was not a strong validation and introduced an additional 1-page plot with 8 profile shapes.

Wind data: You are using several different data sets in your analysis which occasionally confuses the reader which could be because of the naming convention you chose: MMIJ, MMC and ML. A simpler naming such as “onshore” and “offshore” could help, especially when skimming through the paper.

[AC] Now, the locations of the met masts IJmuiden and Cabauw are referred to as the off- and onshore reference locations.

I am sceptical about combining offshore and onshore wind data into one data set. It is my understanding that you want to simplify the energy estimation of AWES by creating a general purpose set of wind profile clusters. This could lead to results that are so generalized that their application is not valid in either situation. This averaging effect probably aggravates due to the small number of clusters and the temporal resolution of the data set. Comparing the mean normalized wind speed profiles in Fig. 2 and 10 (as well as Fig. 8 and 11) shows that these profiles are in deed fairly different especially at altitudes up to 300m which is well within the operating range shown in Fig. 18. This difference is further supported by the map in Fig. 15. Please show a validation of your approach (e.g. compare power outputs or wind profile shapes reconstructed against actual simulation) or explain in the text how you validated it.

[AR] The aim with using the larger combined onshore and offshore dataset was to highlight how the prevailing profiles do vary depending on the terrain. We have tried to show this by relating the cluster profiles so obtained with those of the individual onshore and offshore analyses. For a detailed performance assessment, indeed we would suggest a more tailored (offshore or onshore) clustering approach which is why we used the conditions specifically at Cabauw for illustrating the AWE AEP assessment.

[AC] We have made some changes in the text at the beginning of Section 4.4 to try and clarify the rationale for the large area clustering analysis.

The usage of data sets in your analysis is as follows:

- Section 3: Normalization, transformation and clustering process with **2D offshore** wind profiles
- Section 4: Compare normalized clustered **2D offshore, onshore, lidar** (7 lines of text) and entire domain
- Section 5: Calculate power curve for **1D scaled onshore** wind profiles
- Section 5: Compare AEP for **1D onshore** and entire domain

This leads to the repetition of similar plots, e.g. Fig.8, 11, 13, 14 which take up a lot of space and hardly additional information content. Therefore, my recommendation is to focus on either the onshore or offshore location, and compare final results such as power curve and AEP with all data sets. You can choose to move the other figures to the appendix or remove them entirely. Similarly, I recommend removing the lidar section from this paper, because you do not use it in any analysis other than "... cluster profile shapes for both datasets are very similar." A deeper analysis of the differences and commonalities between the data sets would justify keeping all these figures if you choose to keep them.

[AR] Although, the methodology is developed with the aim of using the cluster wind resource representation for AWE production estimations, we consider obtaining and analysing the cluster representation the most important contribution of this paper. The last section is merely to show an application of the cluster representation and thereby presenting a complete story line: from data to application.

[AC] Lidar discussion is discarded. Onshore and multi-location analyses are expanded.

While you go into great detail explaining the process of normalizing, applying PCA and clustering your wind data, very little detail is given on the denormalization of the wind profiles so they can be used for power estimation. Did you use the cluster centroids i.e. the mean of all the profiles which is not an actual profile that occurs in your data set? Similarly one can argue that the profiles derived from the $mean \pm PC * std$ shown in Fig. 2 and 10 are not actual profiles in the data set. Which range and step size did you use to denormalize the wind speed profiles? Is it possible to determine the wind speed range of each cluster or is this information lost due to normalization? If so it would be interesting to see this range in power curve plot. Validating your approach against actual simulated profiles by comparing energy or power curves would add to the credibility of this paper. This could include a comparison against non-normalized, clustered wind profiles or not PCA transformed profiles or standard log profiles with assumed Weibull distribution etc... Please justify why you go through all this process for 2D wind profiles if you just use the wind speed as an input into your AWES model.

[AR] Indeed they are not actual profiles, but averages. 100 bins are used between the cut-in and cut-out speed prescribed at 100 m. Basically, we only use the magnitude profiles because the performance model is not compatible with a profile that is veering. The methodology is however not developed specifically for coupling to this performance model. For different

applications, the two-component profiles may be useful. We believe that the two-component analysis does provide useful insights in terms of profile veering which relate to stability.

[AC] We have included a more precise description of the wind profile shapes reflecting the centroids: "the cluster-mean wind profile shapes". Also, we introduced a more precise description of obtaining the cluster representation: "After obtaining the cluster-mean wind profile shapes, they are used for constructing the *cluster representation* of the wind resource. Each sample's vertical wind variation is approximated by de-normalising [scaling] the cluster-mean wind profile shape of the cluster to which it is assigned using the normalisation wind speed of the pre-processing." Furthermore, the power curves are complemented by the wind speed distributions.

PCA: Which PC and std do you use to derive the profiles column 3 and for of Fig. 2 and 10? Maybe a better representation would be to plot the mean, $mean \pm PC * std$ in one plot with a shaded area in between to highlight the range of possible speeds / velocities? It is my understanding that the justification for using PCA is that it accelerates the clustering process. Is clustering the bottle neck of your analysis or is it the optimization? How does the end result vary comparing transformed and not transformed profiles?

[AR] For each PC a row is reserved, the std is an output of the PC analysis. In our case, the clustering is relatively cheap. The PC analysis however reveals some interesting features of the datasets. Clustering the data in the 'physical' space would probably yield similar results.

[AC] More precise panel titles are added.

Clustering: Please justfy in the text why you only use the cluster centroids for your analysis and how which uncertainties / inaccuracy this causes? How much does the profile shape within one cluster vary and is this variation reduces by normalizing the data? Clustering labels are not in order. How did you sort them or determine analogies between clusters of different data sets (see table 2)? Comparing clusters like this gives the impression that they have the same profile shape. Cluster centroids are the arithmetic mean of all the data points within this cluster. Therefore, adding or removing a data point changes the outcome of the entire process. How can you compare these clusters in table 2? If only 8 clusters are enough, then why not use existing stability classification (like table 1 in https://www.adv-sci-res.net/6/155/2011/asr-6-155-2011.pdf or https://www.wind-energ-sci.net/4/563/2019/wes-4-563-2019.pdf) based on Obukhov length or Richardson number which are widely accepted and a common meteorological classification?

[AR] We aimed in this work for a compact representation and therefore use normalisation, which inevitably comes at the cost of precision. Standard wind resource assessment relies on the use of a monotonic logarithmic profile which at best incorporates the effect of stability. We believe that our compact cluster mean profile methodology represents a significant advance on this approach particularly where the wind resource extends well beyond any surface layer. One could look at variability of profiles in the cluster as a measure of uncertainty, but this is beyond the scope of this work and would cloud the central aim of the work. Indeed, we now do try and relate the properties of the surface layer profile to the stability conditions and use a standard stability classification (Table 1).

[AC] We have expanded the text on the labelling of the clusters. While discussing Figure 7, we state: "Note that the cluster algorithm produces arbitrary labels for each class. We have manually renumbered them such that the numbering is more or less aligned between onshore and offshore clusters." Furthermore, the original Table 2 is omitted. We added some text in the conclusions to acknowledge variability in the profile shape in each cluster. We added a table of fairly standard stability classes.

2 Specific comments including technical corrections

Title: I recommend a more specific title including information such as: ground-gen / pumping mode, mesoscale wind data, but understand that it is personal preference.

[AR] This would not really reflect what we see as our main contribution: the data-driven methodology.

2.1 Abstract

page 1

line 1: Why not use AWES abbreviating for Airborne Wind Energy System as commonly used in the community and literature

[AR] We prefer to stick to using solely AWE as a convention.

line 6: Introduce abbreviation DOWA here.

[AR] We consider it as good practice to avoid acronyms in abstract.

line 10: Mention the derivation of power curves.

[AR] Added

line 10: Mention the location where you estimate AEP or add AEP analysis for all sites since it sounds like multiple AEP estimates are compared.

[AR] Added

line 11: Define or rewrite the sentence so that it is clear that you compare the amount of clusters necessary to estimate AEP and use concrete results rather than “within a few percent”.

[AC] Replaced by "within three percent"

2.2 Introduction

page 1

line 13: AWE already defined in Abstract in line 1.

[AR] We consider it as good practice to reintroduce acronyms in the body of the paper.

line 14: Many different concepts exist with varying (anticipated) operating altitudes. Either remove subordinate clause "typically in a range up to 500 m" or explain / reference what this assumption is based on.

[AC] Rephrased to "above 150 m" and references added.

line 14: Add reference for stronger and more persistent winds.

[AR] Considered superfluous as this is described by the relationships in the next paragraph.

line 18: This is not the definition of surface layer. I would stay away from concrete number as the height of the surface layer varies a lot depending on atmospheric condition. Remove or replace with: in the order of tens of meters depending on atmospheric stability (http://glossary.ametsoc.org/wiki/Surface_boundary_layer)

[AC] Left out the 200 m

page 2

line 1: Low-level jets are a known phenomena and have been discovered a while ago. Rewrite: "Recent studies have identified numerous low-level jets"

[AC] Rephrased

line 4: Remove: "more"

[AR] Done

line 9: What kind of "performance calculation"?

[AC] replaced by "power production"

line 9: Add reference to COSMO-DE

[AR] Only listed because used in the reference. Since the indirect reference, we do not deem it necessary to include a separate reference to the dataset.

line 10 & 11: Add: "shape" after wind profile

[AR] Done

line 14: Add reference to MERRA-2

[AR] Only listed because used in the reference. Since the indirect reference, we do not deem it necessary to include a separate reference to the dataset.

line 14 & 16: Malz et. al b) referenced before a).

[AR] Modified

line 15: Sentence on the reduction of computational cost and choice of initialization seem unnecessary for this paper.

[AR] We find it useful as it emphasises that there is a demand for cheaper AEP calculations.

line 20: Add reference and expansion of WRF

[AC] Expanded WRF

line 26: Remove “such a historical dataset”. Does not contribute to the understanding or quality of this paper. What makes this data set so special?

[AC] Rephrased

line 34: Add reference to DOWA, ERA5 and expansion of LiDAR (also, check WES house standards on capitalization of abbreviations: https://www.wind-energy-science.net/for_authors/manuscript_preparation.html. One feedback I got was: “WRF is a defined model name (the same applies to the ARW), but lidar is a general term and is therefore not capitalized according to our house standards.”

[AC] Reference included in the data section

page 3

line 1: Is the DOWA data set comprised of met mast data? I thought it is derived from reanalysis models which had various measurements assimilated.

[AC] Rephrased

line 2: Sentence is a little confusing as it suggests that met mast data is somehow involved. Maybe chose a different naming convention e.g. onshore, offshore.

[AC] Rephrased

line 2: “The procedure” is confusing as the previous sentence describes comparing DOWA and LiDAR and not generating a set of representative wind speed profiles from clusters

[AC] Rephrased

line 5: More accurate than log profiles or uniform wind? Is it actually more accurate and did you validate the improvement?

[AC] Rephrased

line 7: Active voice

[AC] Rephrased

2.3 Wind data sets

page 3

line 10: Combine first 2 sentences.

[AC] Rephrased

line 16: Repetitive sentence: “Both long-term ...”

[AC] Rephrased

line 17: In what way “sparser” lidar? spatial, temporal?

[AC] Rephrased

line 21: Repetitive sentence: "... coast in the North Sea". Also shorten the sentence e.g. "The selected offshore location, met mast IJmuiden, is located 85 km off the Dutch coast."

[AC] Rephrased

line 21: Write out the wind direction as it doesn't safe lots of space but adds clarity

[AC] Rephrased

line 21: Why not use the Cabauw or IJmuiden met masts for validation as well? I thought that was one of the reasons why you chose this location.

[AC] Rephrased — stressed that we don't use met mast data (anymore)

2.3.1 ERA5

page 3

line 26: Define and add reference to ERA5 and ECMWF.

[AC] Reference added

line 27: I don't think the year in brackets as well as the information that more data from years back will be available soon is necessary.

[AC] Rephrased

line 31: Replace outputs with a different verb: produces, uses, calculates and rewrite sentence to be more concise

[AC] Rephrased

page 4

figure 1: More explanation in caption

[AC] Caption expanded

figure 1: How and why did you choose the sample location? Are these representative locations?

[AC] Explained in text

figure 1: Is it important that you use met mast locations as you don't you met mast data at all.

[AR] No, the reason for doing so is that they are well covered in literature.

line 1: Remove sentence: The long-term wind climate is not important for this study as no long-term predictions were made.

[AC] Removed

line 2: You use ERA5 to determine atmospheric stability. Does DOWA not provide the necessary data?

[AR] It does, however DOWA did not give satisfactory results for calculating the Richardsons number.

[AC] We clarify that we used ERA5 rather than DOWA as it gave more satisfactory stability values/

line 2: Stability is only used for cluster statistics in Fig. 9 and similar. Please expand this analysis and the relationship between clusters and stability.

[AR] Parallels have been drawn between the clusters and stability in Sect. 4.2, however, because we only feed in the normalised wind profile data, the relationship will be indirect.

[AC] We have added more analysis of relationship between stability and cluster mean profiles.

line 3: DOWA defined on page 2

[AC] Rephrased

2.3.2 Dutch offshore wind atlas

page 4

title : Capitalization different from previous DOWA

[AC] Capitalized

line 6: How downscaled? Is it not more the extraction of specific values from a data set?

[AR] Mesoscale model downscaling is a widely accepted way to increase the resolution of a coarser scale meteorological dataset. It is not just the extraction of values - the mesoscale model captures the physics of finer scale processes.

line 6: Add reference to model and / or move link to footnotes

[AC] Reference added

line 10: Add reference

[AR] Same ref as previous

line 11: Remove 2nd “than” and “alone”

[AC] Removed

line 12: Define or paraphrase “ASCAT” and “mode-s ehs”. Add reference

[AC] Put between parentheses

line 13: Sentence about website unnecessary

[AC] Removed

line 14: Remove “additionally” and conjugate “show” or add et. al

[AC] Removed

line 15: LLJs are not anomalous

[AR] It is defined as such in https://www-sciencedirect-com.tudelft.idm.oclc.org/science/article/pii/S0167610516307061 .

2.3.3 lidar observations

[AC] Subsection removed

page 5

title : Check capitalization and abbreviation rules for lidar

line 1: Rewrite. I don't know what you want to say with this sentence.

line 4: ECN?

line 5: Add link as reference and / or footnotes

line 6: Where is difference between "clock hour" and hourly?

line 7: Active voice

line 8: Only time you use "data set" instead of "dataset"

2.4 Clustering procedure

page 5

line 12: Mention that you do the same for MMC

[AR] The onshore analysis is part of Sect. 4 and therefore only introduced there.

line 13: How normalized? Add PCA abbreviation

[AR] Described in following subsection.

line 14: rewrite "for choosing", active voice e.g. the number of clusters is chosen based on the clustering performance...

[AC] Rephrased

2.4.1 Preprocessing of the wind data

page 5

line 16: Shorten sentence, mention use of all time steps, entirety sounds like more than it is.

[AC] Rephrased

line 20: Is that 90th percentile of each time step or of all data points of this altitude or of all data points at 100 m?

[AR] Each time step

[AC] Rephrased

line 20: Line break not necessary

[AC] Removed

line 21: Why did you choose this normalization and why normalize in general if you expect that it will not lead to good results for low winds?

[AR] Normalisation is used as we aim for a compact representation.

[AC] Rephrased: "Fewer outlying wind profile shapes result when the 90th percentile instead of, e.g., the maximum value is taken as normalisation value. The normalisation yields a more compact wind resource representation."

line 21: Active voice. Do you only expect eccentric profiles or did you actually observe them?

[AR] We observed them mostly in the lidar dataset due to outliers. For the modelled data it is less urgent to use a percentile instead of the max value.

line 24: How do you implement them if they are omitted? [AC] Rephrased

2.4.2 Principal component analysis of the wind profile shape dataset

page 5

line 27: Defined above

[AC] Rephrased

page 6

line 4: “for every wind direction” = omnidirectional

[AC] Rephrased

line 4: The difference to ...

[AC] Rephrased

line 5:“logarithmic profile representation of the wind environment” = logarithmic wind profile. Remove one “logarithmic profile”. Introduce z_0 here. Why is $z_0 = 0.0002$?

[AC] Rephrased, added reference for roughness length.

line 6: “vertical” is understood, wind speed is always magnitude

[AC] Rephrased

line 6: Add variable name $u(z)$ and reference to equation

[AC] Moved to introduction, and introduced before stating the equation.

line 8: Add reference or explanation to Obukhov length

[AC] Explained in introduction.

line 13: Explain relationship of L and atmospheric stability. Does this sentence mean that the average of multiple years of wind data fits an unstable log profile? Are you fitting wind speed or just parallel component?

[AR] Fitting to the magnitude profiles, not the parallel component profile.

[AC] Explained in introduction.

line 13: Add reference to “Theory”. Add “mean wind profile...”, remove “in the” & “direction”.

[AR] Ekman theory is considered text-book knowledge and therefore we chose not to add a reference here.

[AC] Rephrased to "In accordance with Ekman theory, the mean wind profile shape veers .."

line 13: Add “(top-view in the bottom left panel)”

[AC] Referred to lower left panel in text.

line 14: Remove “of the wind profile shape”

[AC] Removed

line 22 & 23: Remove text in parentheses

[AC] Removed and rephrased

line 24: Add reference

[AC] Rephrased: "We consider retaining 90 % or more acceptable for our application."

line 28: Active voice

[AC] Rephrased

line 31: Isn’t this a general feature of PC? Remove “Note that”?

[AC] Removed

page 7

figure 2: What do the number 1-4 in bottom boxes mean? Add normalized to vertical wind profile. Does second column show the mean or is it the orientation of PC axis with height? What does it mean to multiply PC with std? Which std do you multiply with? It might make sense to show the std of each PC in a table. Plots in column 2 are not on the same x-axis. Please explain how PC1 and 2 rotate over altitude.

[AR] Second column is not the mean: it illustrates the unit vector defining the direction of the PCs. Not completely clear what you mean with: "Please explain how PC1 and 2 rotate over altitude." Refrained from adding another table, the panel titles should be more self explaining now.

[AC] Added: "The wind profile shape numbers 1–4 refer to the markers in Fig. 4a." Changed x-axes of column 2.

2.4.3 Number of clusters

page 7

line 6: Either active voice or general statement in which case you should use “a” instead of “the”

[AC] Reordered/rephrased, reference shouldn’t have been at the end of the sentence.

line 8: Remove parenthesis. Replace “all” with “each”

[AC] Rephrased

line 10: Add reference

[AR] Rephrased a little, such that is more generic and does not need a reference.

page 8

figure 3: Explain legend abbreviations and subset in caption. What is the MMIJ subset?

[AC] Changed legend

figure 4: why * on x- & y-axis label? Are the markers the orange circles? Do the numbers 1-4 correspond to the clusters shown in Fig. 2? If so mention this in the caption.

[AC] Expanded caption. "The coordinate system represents the average PC profiles of the two reference locations, denoted by an asterisk." Also reference to Fig. 2's profiles explained.

figure 4: Maybe replace "wind profile shape" with wind velocity profile.

[AC] Shape is used consistently throughout paper, so also preferred here.

line 1: What is "itk" ? Number of clusters is usually the k in k-means clustering

[AR] Typo, should be italic k.

[AC] Corrected

line 5: Delete if not shown. If you keep it in explain "over-fitting" in this context. Probably not happening because of (relatively long) temporal averaging.

[AC] Removed

line 7: How is the silhouette score calculated? Adding silhouette score and WCSS equation would make it easier to understand. Is a score of 1, 0 or -1 better?

[AR] The silhouette score calculation is quite complex and not considered crucial for understanding the assessment and therefore left out.

[AC] Added explanation on score values.

line 9: Remove repeated subject "mean silhouette score". How do you interpret the decreasing trend? Explain if you mention it, the fact that it is decreasing is obvious.

[AC] Rephrased

page 9

figure 5: Which data set do these results belong to: MMIJ, MMC, ML? Are these results the same for all data sets? Expand on the captions.

[AC] Added ".. for filtered offshore dataset"

figure 5 a): y-axis label: Isn't it the "sum of squared distances" ?

[AC] Changed to WCSS, is considered a more precise description.

figure 5 a): Add a vertical line at $k = 8$ to highlight your choice.

[AC] Added vertical line

figure 5 b): Explain non-linear grid lines in caption. Why no vertical grid lines? "...error over height for logarithmic profiles...(...and four different stability...)", "...three exemplary clusters". How is the error defined?

[AC] Expanded in text.

line 1: So it would be best to use only 2 clusters to represent the many different wind conditions even though the sum of square distance is way higher and 4 log-fits out perform it? What defines a good silhouette score? Wouldn't it be a fair comparison to also vary the number of log-fits over the x-axis? Why 4 log-fits?

[AR] Purely based on the silhouette score: yes, but overall: no. Score close to 1 is desirable. 5 log shapes are used now: 1 for each stability class. Is considered a reasonable comparison when considering 5 stability corrections - often the wind resource representation in AEP estimation is not corrected for stability at all.

[AC] Rephrased

line 2: Why do you show the different approaches to get best k and then use AEP without showing it here?

[AC] Rephrased

line 5: Shorten to"filtered wind profile data". Parentheses are repeating what has been said before, remove.

[AC] Rephrased

line 5: Remove: "Next, it is..."

[AC] Removed

line 8: Equation would make it easier to understand.

[AC] Equations added

line 11: More combined than individually, otherwise 2 lines.

[AC] Rephrased

line 16: How did you choose these L values? What are the assumed ranges of L associated with certain atmospheric stability?

[AC] Table added with stability classes, including reference.

page 10

figure 6: Why vertical grid line in center of plot? Showing the x-axis as percentage of total data would be better. Briefly explain what positive and negative values mean?

[AR] We feel it is justified to use "identifiers" on the x-axis, as they point to specific samples and thereby emphasize that the order of the samples is not random.

[AC] No specific reason for the middle grid line, therefore discarded.

line 5: Is this weighting intended? Why not linear interpolation in z?

[AR] We did not tailor the weighting, but it could be justified: the weight is higher around 100–200 m, in which the reel-out phase is mostly taking place.

line 6: It seems like this is not the only reason. Resolution is high up to 200 m.

[AR] Agreed, we had a closer look and found out that model deficiencies also contribute to this effect.

[AC] Added: ".. the PC1 and PC2 profiles show that most variance in the dataset is found at both ends of the height range. Due to the relatively high variance and fit model deficiencies, the fit error is also expected to be largest at these heights."

line 9: I assume that a low number of clusters is enough to capture variations within the hourly data set which is why I would recommend mentioning that this choice is specific for your temporal resolution.

[AR] For this type of analysis we are not interested in atmospheric phenomena with a time scale lower than an hourly one.

[AC] Added to start of Sect. 2 to address the latter: "An hourly temporal resolution of the datasets suffices for capturing the diurnal cycle of the wind profile. While smaller scale atmospheric phenomena might have an (adverse) effect on the power production, these effects are typically superimposed on a steady-state wind profile using separate models for assessing, e.g., the associated loss in power production (Fechner, 2016)."

line 14: Move the last sentence up to the section where you first introduce clusters.

[AC] Sentence removed

2.5 Wind resource representation based on clusters

2.5.1 Interpretation of prevailing wind profile shapes at MMIJ

page 11

line 10: Are these profiles the centroids of the clusters you calculated before?

[AC] Rephrased this sentence. Moreover, added "The resulting centroids reflect the cluster-mean wind profile shapes in the dataset, which follow from back-transforming the cluster centroids from the PC to physical space." to Sect. 3.3 to be more clear about how the centroids relate to the shapes.

line 13: Are these the Obukhov lengths? Which ranges do you associate with each stability condition?

[AC] Stability classes given in Table 1.

line 12: Replace “...moving between...”

[AC] Rephrased

line 13: What are the number in parentheses? Add variable.

[AC] Obukhov lengths found for the mean-cluster shapes are now presented in Figure 9.

line 13: Move line break as veering clearly refers to the previous paragraph.

[AC] Text is rearranged

line 21: Fig 7 is mentioned after Fig 8. Active voice

[AC] Figure ordering corrected

line 22: Also active voice: “Examining the five PC coordinates in table 1 ....”

[AC] Rephrased

line 26: How can that be deduced? Define filter e.g. “low-speed filtered...” as readers might have forgotten or skipped the previous sections.

[AC] Discussion on filtered vs full dataset removed as requested by other referee.

line 28: Define “calm wind”

[AR] See previous bullet

line 29: These 2 sentences are more general and introduce the data set. Move up before previous sentences.

[AR] See previous bullet

page 12

equation 2: Why are you using different Δz for u. v, θ within this equation? Are u. v, θ on different altitudes?

[AR] No, on the same altitudes.

[AC] Equation moved to introduction and reformulated. Now, we only use a single Δz.

line 6: Why not rewrite Eq. 2 to include absolute temperature instead? Explain how it is used. I guess instead of virtual potential temperature.

[AR] We changed the stability metric. Currently, we now determine Ri_B between 10–31 m.

[AC] Added to Sect. 4.1: "Here, we derive the stability class distributions using the bulk Richardson number, $\mathrm{Ri_B}$, converted to the Obukhov length, L, using Eqs. **??** and **??**. The data from either ERA5 or DOWA could be used to derive $\mathrm{Ri_B}$, however, we found that using the data from the two lowest ERA5 model levels, i.e., $\sim$10–31 m yields the most realistic values. We use the arithmetic mean of the model level heights for z in order to convert $\mathrm{Ri_B}$ to L."

line 6: Isn’t humidity high close to the shore / offshore and would the effect not be considerable?

[AR] The newly used expression used for $\mathrm{Ri_B}$ includes humidity.

line 8: Add “Positive Ri_B values”

[AC] Added

table 1: Do these values relate to Figure 2?

[AC] Added: "The centroids are depicted in Fig. 7a at their PC1, PC2-coordinates with the numbered markers."

page 13

table 2:The cluster centroid shape will change based on the underlying data, i.e. filtered. It seems that you assume that clusters between data sets are same, e.g. cluster 1 is same through out all data sets. Did you sort or determine similarity between cluster centroids of different data sets? How similar are PC transformed and normalized wind data? Can you quantify how similar / different they are?

[AR] The clusters are not the same, only the number assigned to them is. Labelling of the clusters is done manually. The implications of PC transformation is explained at length in Sect. 3.2, but I don't think this is what you're asking about. If I'm correct, you are asking about the sensitivity of the clusters to the data pre-processing? Such a sensitivity analysis could be done, however, we feel this is out of the scope of this paper. At an early stage, we found that the PC transformation does not have a large effect on the resulting clusters. The PCA however helps us to understand the data structure and therefore is a helpful contribution to to the paper.

[AC] Added to Sect. 4.3: " Note that the cluster algorithm produces arbitrary labels for each class. We have manually renumbered them such that the numbering is more or less aligned between onshore and offshore clusters. This allows us to draw parallels between them .."

figure 7: Nice visualization. Why * on x and y label?

[AR] The location-average of the PCs is used for the coordinate system, enabling a direct comparison between the two subfigures.

[AC] Added: The coordinate system represents the average PC profiles of the two reference locations, denoted by an asterisk and shown in Figs. 2 and 11.

page 14

figure 8: Is this the filtered or unfiltered data set? Do all x-axis use the same range, because different label? Explain more in caption, i.e. why $\hat{v}$ over variable, where is the origin of the hodograph / increasing z direction? Better spread out all the graphs so that they appear after mentioned in text. Do you need all these graphs here? Maybe move some to the appendix.

[AR] Filtered. Yes, same x-axis. Yes, we discuss all shapes in the body of the paper, so therefore it would also be best to present them here.

[AC] Caption expanded, added: "The eight cluster-mean wind profile shapes of the offshore clusters (MMIJ-1–8). Each shape is depicted by the normalised wind speed components with height (first and third rows) with the corresponding hodograph below (second and fourth rows). Logarithmic profile fits are plotted alongside the shapes. In the hodographs, the lowest points are connected to the origins with dotted lines and the highest points are the loose ends. All plots share the same x-axis."

page 15

figure 9: Nice way of comparing the data set. Might be a bit overwhelming at first though. Is this the entire data set or filtered? Where are these Ri bins coming from? Add reference. What do you mean by "Bins have the same overall frequency" ? Wouldn't that mean that all the bars have same height? Do you mean bin width? Use the same comprehensible bin width for all subfigures (especially wind speed and Ri seem arbitrarily chosen) in figures 9 and 12. Maybe add an "offshore" and "onshore" to data set labels for clarity. A better and easier understandable metric would be atmospheric stability (i.e stable, unstable) instead of using Ri number.

[AR] Filtered data set. Bin sizes are the same over all clusters, but not within one cluster.

[AC] Rephrased and expanded caption, on/offshore added. Stability distribution changed corresponding to earlier introduced classes.

page 16

line 1: Change figure order or reference which figure you mean after 3 pages of figures. "... over the 10 year timeframe".

[AC] Reference added

line 2: Why is this a prerequisite? You have temporal variations on all time scales. But I guess this way you make sure that your data is not based on outliers.

[AC] Text expanded: "The upper panel shows that the inter-annual variability is limited, which asserts that the results can safely be generalised to the lifetime of a wind energy system (20 years)."

line 3: Remove sentence and line break. Maybe change order. Why first talk about 9 a) and then all the other figures separately ?

[AC] New subsection introduced specifically for the interpretation of the results.

line 5: Rewrite: "... not so frequently strong". According to figure 9, cluster 1 has an almost even occurrence through all wind speed ranges.

[AC] Rephrased

line 5: "are more frequent".

[AC] Rephrased

line 7: Add reference to figure which shows cluster shape.

[AC] Reference added

line 9: "well mixed". How about shear?

[AC] Corrected

line 26: Remove line break as both paragraphs are about cluster 6,7.

[AC] Removed

2.5.2 Comparison with an onshore location

page 17

line 5: Vertical is understood. Better write: “The mean normalized wind speed profile...”.

[AR] Earlier in the paper we define the shape as being a normalised profile. We feel that this definition should be clear at this stage of the paper.

[AC] "Vertical" removed

line 8: Same as line 5

[AR] See above

line 11: Removing “of the two locations ... ” makes the sentence easier to understand.

[AC] Rephrased

line 19: Add the variable to values in parentheses

[AC] Results now presented in Figure.

line 20: So is there no cluster that corresponds to stable stratification without LLJ? If so wouldn’t that be unusual?

[AC] Stable stratification is recorded for MMC-3/4. MMC4 cluster-mean shape only shows a weak jet shape.

line 22: “however” and “just” seem like filler words. You could shorten the sentence

[AC] Rephrased

line 24: Active voice.

[AC] Rephrased

line 26: Sentence hard to read. Shorten i.e. “This affect is caused by the lower heat capacity of the land surface which promotes a more immediate heat transfer to or from the atmosphere.”

[AC] Rephrased as suggested

line 30: What is the point of comparing diurnal to seasonal cycles? They are caused by entirely different effects and play out over vastly different timescales.

[AC] Rephrased: "The patterns in the times of occurrences indicate a pronounced diurnal cycle in atmospheric stability for the onshore location, whereas for the offshore location the seasonal cycle is more pronounced."

line 32: What does overall bin frequency mean and why does it have to be equal? It sounds like you are varying your classification to get a certain result rather than try to characterize actual physical effects.

[AR] Not applicable anymore for stability distributions. Bin size is the same over all clusters, not within one cluster.

[AC] Rephrased: "Note that the wind speed bin limits are chosen such that the frequency over all clusters for each bin is roughly the same, yielding different bin widths for the two reference locations."

page 18

figure 10,11,12: More explanation for readers who just skim through the text or just look at figures, i.e. onshore, normalised wind speed, numbers and explain PC*std. Consider moving some figures to the appendix.

[AC] Captions expanded

2.5.3 Validation with LiDAR observations

page 21

line 1: This paragraph seems rather unnecessary. Very short, no mayor inside, only that results are similar.

[AR] Agreed

[AC] Subsection removed

line 2: replace “investigate” with “show”..

2.5.4 Spatial distribution of wind profile shapes

page 21

line 10: How did you chose these sample locations in Fig.1? Are 45 grid points the entire domain? Does your selection affect / bias the results? Line break not needed.

[AC] Text expanded: "The multi-location dataset (filtered to exclude low wind samples) includes wind data from 45 DOWA grid points that are selected such that onshore, coastal, and offshore locations are equally represented. For each location type, 15 grid points are chosen (pseudo-randomly) to yield a good coverage of the full DOWA domain (50778 grid points in total)."

line 12: Explain cluster mapping. Why only 8 clusters again? reference previous chapter.

[AC] Added: "Each sample of every grid point in the DOWA domain is assigned to the cluster with the closest centroid."

line 13: How can you apply the same mapping to a new data set?

[AR] With mapping, we ment the assignment of samples to a cluster, see upper bullet.

[AC] "Mapping" is a bit confusing, therefore, rephrased.

line 16: Remove “be” and shorten / rewrite sentence.

[AC] Rephrased

line 18: Remove sentence “Since cluster 1...”

[AC] Rephrased

page 22

table 3: Is this a necessary table? offshore / onshore is not a sufficient classification of wind. How did you match clusters from different data sets? Are they the exact same clusters and if not how similar are they?

[AR] We expanded the analysis in this subsection and feel that it is now justified to leave the table in. We agree that it is only a very crude classification scheme, however, e.g. IEC standards also differentiate between on- and offshore classes, so we feel that justifies using these classes. Matching is done manually. In practice, clusters are never the same for different datasets.

[AC] Added: "Every multi-location cluster is manually linked to the single location clusters based on resemblance of their cluster-mean wind profile shapes, see Table 3."

page 23

figure 13: Is this a necessary figure? Dashed lines missing in legend. Is this based on hourly average lidar data? Do these numbers in circles have any meaning?

[AC] Figure removed

page 24

figure 14: Rewrite: many “the” in caption. Maybe a bit more explanation, i.e. which locations are represented.

[AC] Caption expanded

page 25

figure 15: Add details to captions. Add ML abbreviation. Add reference to fig. 14 for info on clusters. How did you ´´map” to a new data set? Meaning of numbers? Use consistent frequency ranges for comparison or justify why you did not.

[AR] Mapping is explained in text.

[AC] Caption expanded. Text added: "Note that the colour scale is different for each map so that spatial patterns are easier to observe."

figure 15: This clear division between on- and off-shore profiles would justify separating the analysis. Wouldn’t this lead to better more detailed results and insights? Please validate and quantify how much information you lose by clustering everything this way rather than off- and onshore individually.

[AR] The aim with using the larger combined onshore and offshore dataset was to highlight how the prevailing profiles do vary depending on the terrain. We have tried to show this by relating the cluster profiles so obtained with those of the individual onshore and offshore analyses. For a detailed performance assessment, indeed we would suggest a more tailored (offshore or onshore) clustering approach which is why we used the conditions specifically at Cabauw for illustrating the AWE AEP assessment.

[AC] We have made some changes in the text at the beginning of Section 4.4 to try and clarify the rationale for the large area clustering analysis.

2.6 Fast AWE production estimation based on historical wind data

page 26

title : What is “fast” about this analysis?

[AR] The number of optimisations needed is reduced substantially w.r.t. ’brute-force’ calculations.

line 1: Rather “estimating” than “calculating”

[AC] Replaced

line 2: You went through all the process of explaining the clustering process using MMIJ data set than introduce MMC, lidar and ML to now only use MMC? Why not focus on MMC entirely or apply your power estimation to all data sets?

[AC] Expanded in text: "An advantage of AWE systems over tower-based wind turbines is that they have access to winds higher up. This advantage is limited when low-shear wind profiles are frequent at the installation site, as is the case offshore, but this is not usual for onshore locations. Employing an AWE system at an onshore location thus requires a more variable operational approach. For this reason, we demonstrate the AEP estimation for the met mast Cabauw location using the eight clusters from the single location analysis (Sect. 4.3)."

2.6.1 Determining power curves for AWE systems operated in pumping mode

page 26

line 10: “... differs between ...”. Add reference to flexible-kite sentence.

[AC] Reference added

line 15: Rewrite sentence: “...is ended ... is depowered... is steered ....”. Remove: “the” in front of zenith. Add: “... to the starting position of the traction phase”

[AC] Rephrased to active voice and added suggestions.

line 17: Check capitalization rules for abbreviations.

[AC] Uncapitalized

line 18: Rewrite sentence: “... moves the kite along an idealised flight path conform a series...” ?

[AC] Rephrased: "The motion of the kite is approximated by moving it along the idealised flight path according to the computed steady-state kite speed."

line 19: Doesn’t the limitation to lightweight membrane kites mean that the approach is only applicable to soft kites and not “any kind of pumping AWE system” (line 12)?

[AC] Rephrased: "The specific operational approach differs between concepts and may require different performance models for calculating the generated power. We evaluate a flexible-kite system using the quasi-steady model (QSM) .."

line 24: I agree that it is justified to use only magnitude wind speed profiles, but why go through all the process of clustering 2D profiles in the first place?

[AR] The methodology is not specifically developed to be coupled to this performance model. Other applications might require the two-component profiles.

line 26: Explain in more detail how do you derive the power curve and how you scale the normalized profiles.

[AC] Expanded: "The power curves required for the AEP estimation relate the mean cycle power to the scaling parameter used for de-normalising the cluster-mean wind profile shapes of MMC-1–8. Given the profile shape, this scaling parameter can be prescribed as a wind speed at any height. We use the wind speed at 100 m. By stepping through a range of wind speeds between cut-in and cut-out, a power curve is constructed for each of the clusters. At each step, the profile shape is scaled using the respective wind speed to yield the absolute wind profile. An optimisation is then performed using this wind profile as input."

line 28: Active voice. Mention that aero coefficients are assumed to be constant.

[AC] Expanded: "The QSM uses constant values for the lift and drag coefficients of the powered and de-powered kite. In reality, the coefficients vary and representative values of the leading edge inflatable V3 kite are selected based on the experiment of Oehler and Schmehl (2019)."

line 31: What means “sufficiently high”? How is the tether reeling speed during reel-in and -out? Is it a constraint or output of the optimization?

[AR] Speed follows from the steady state calculation, given the tether force.

[AC] Rephrased: "The values of the cycle settings are chosen such that they yield maximum mean cycle power. The reel-in tether force should allow a fast retraction of the kite, while limiting the energy consumption. During the transition phase, the reeling speed is kept zero unless tether force limits are exceeded. During reel-out, the tether force should yield a high energy production, while letting the reel-out phase comprise most of the cycle duration."

line 31: Rewrite to include tether force constraint and that it corresponds to setting a fixed max tether diameter.

[AR] I’m not considering any design variations here. Tether diameter is constant as given in Table 4.

page 27

table 4: More descriptive caption.

[AC] Caption expanded

line 1: What are your control variables? Table 5 contains constraints that you are keeping constant I assume. Maybe write out the optimization formulation?

[AC] Table 5 lists the optimisation variables and their limits. Note that we are not dealing with an optimal control problem here. We are maximising the mean cycle power by varying this confined set of variables. The stated algorithm implementation uses these limits as input. On the background the limits are converted to constraints,

however, as a user of the algorithm you don't have to deal with this. Therefore, we don't think this information is needed for either the user of the reader of this paper.

line 2 & 3: What is the message of these sentences? Optimizations basically always have active constraints. Why do you have to lower the tether force? What if an increase in elevation angle leads to an increase in wind speed and therefore force?

[AC] Expanded text: "For high wind speeds, the system runs into its maximum tether force and reeling speed limits. Increasing the elevation angle of the reel-out path generally alleviates the tether force and expands the wind speed range that allows safe operations."

line 3: Define effective pumping length.

[AC] Expanded text: ".. the effective pumping length of the trajectory is the difference between the minimum and maximum tether length during reel-out and is included as a cycle setting."

line 8: Rewrite so that you describe where cut-in and cut-out comes from in this sentence or same paragraph.

[AR] The procedure for determining the cut-in and -out conditions are quite elaborate and therefore a separate paragraph is reserved for that.

page 28

table 5: What are these constraints based on? Is actual tether length $l_{min} + l_{pumping}$? Seems like a list of constraints. What are all the constraints? Are these realistic values (add reference)?

[AR] Yes it is. See reply to [page] 27/[line] 1.

[AC] Added to caption: "The limits are chosen by judgment of the authors."

line 3: How do you define "steady flight states"? What are the states? Active voice.

[AR] Flight states without acceleration - these follow from the QSM. Recall earlier sentence: "The motion of the kite is approximated by moving it along the idealised flight path according to the computed steady-state kite speed."

line 4: What about reducing lift / flight speed to achieve one figure-of-eight? To which ground station reeling speeds do your constraints correspond? Explain why you chose this constraint.

[AR] Lift is reduced indirectly, we don't allow $C_{L,\mathrm{powered}}$ to change. In practice however this would be feasible. Flight speeds follow from, amongst others, kite position and C_L. The ground station used is not representing a real one.

[AC] - Added: "Increasing the elevation angle of the reel-out path generally indirectly de-powers the kite and alleviates the tether force. Controlling the elevation angle can thereby expand the wind speed range that allows safe operations. Although not considered here, the kite could also be de-powered directly by controlling $C_{L,\mathrm{powered}}$."
- Added to caption of table 5.: "The limits are chosen by judgment of the authors."

line 6: Explain the developed module and assumptions etc.

[AR] We try to refrain from going into to many technicalities as it will be to distracting here.

[AC] Expanded text "However, this motion can also be approximated as a transition through steady flight states, yielding an approximate duration of the figure-of-eight manoeuvre. Dividing the total duration of the reel-out phase by the average duration of a figure-of-eight manoeuvre yields the number of cross-wind manoeuvres flown."

line 9: This sentence explains cut-out limit again, same as line 4.

[AC] Removed last sentence

line 11: Rewrite sentence. “The corresponding cut-in and cut-out wind profiles are shown in figure 17” or so. How did you scale the normalized wind speeds?

[AC] Rewritten: Scaling each wind profile shape such that the wind speed at $100\,\mathrm{m}$ equals the previously determined cut-in and cut-out wind speeds yields the respective absolute wind profiles, shown in Fig. 18.

line 12: Is the critical height of 80 m related to the minimal tether length, size of device or other parameters?

[AC] Expanded: "The cut-in profiles have the same wind speed at roughly 80 m, which is the kite height at the start of the reel-out phase for the minimum elevation angle employed at low winds. This indicates that, for every wind profile, the cut-in criterion is critical at the start of the reel-out phase rather than at the end. The cut-out profiles exhibit roughly the same wind speed at 300 m, which is the kite height at the end of the reel-out phase for the maximum elevation angle and tether length employed at high winds.

line 19 & figure 24: If wind speeds at 80 and 300m are a sensible choice why do you use v_{100} ?

[AC] Rephrased to clarify the point we want to make: "The cut-out wind conditions for an AWE system are ambiguous when defined by wind speeds at a certain height without defining the profile shape. However, since the cut-out profiles all intersect at roughly 300 m, characterising the cut-out wind speed at this height yields a reasonably precise definition for all profile shapes. Similarly, the cut-in wind speed is well defined at 80 m."

line 25: add “... power curves...”

[AC] Rephrased: "Note that plotting the mean cycle power against the wind speed at 300 m would yield curves that end at roughly the same wind speed."

line 26: Check comma placement

[AC] Rephrased, see above

page 29

figure 17: You sometimes use left and right and other times a and b for sub-figures, choose one. More descriptive caption. How did you scale the profiles?

[AR] We use the letter sub-labels only when we need to explicitly refer to subfigures.

[AC] Scaling better explained

figure 18: Change to “Height” instead of “z” as you always used height before. Add black dashed lines in legend. Why disconnected lines at end of traction phase? Why did you choose this strange v_{100} values? Why not v_{80} instead? Remove: “traction” and “constant” from captions.

[AR] We use x, y, z if the plots comprises multiple spatial dimenstions. Jump in the lines is an artifact of the QSM. We use v_{100} for the power curve, so also here.

[AC] Removed suggested words. Added to caption: "The wind speeds for which the trajectories are depicted highlight changes in the operational approach."

line 1: It would be interesting to see reeling speeds, tether force and other variables during one production cycle.

[AR] We chose to leave this out, since we already have a lot of figures as it is and we value the existing figures more than the suggested figure.

line 4: Why is a profile with LLJ the last to reach cut-out speeds? I would have expected it to cut-out earlier. Is it because of the height of the LLJ?

[AR] Because of its shape and because the cut-out speed is prescribed at 100 m. If prescribed at 300 m, all shapes would roughly reach cut-out at the same time.

2.6.2 Estimating the Annual Energy Production

page 30

figure 19: Seems like the profile shape has almost no impact especially at lower wind speeds where the power ramps up. Remove: “... that are ...”. What is the actually wind speed range of each non-normalized cluster? How come the power curve does not plateau and bend down before cut-out?

[AR] What do you mean with non-normalized cluster? The curves do show such a trend.

[AC] Removed suggested. Wind speed distributions added for each cluster.

line 4: f for frequency of occurrence rather than p for probability as I assume it is based on the data you used and not a model like Weibull. It would be great if you could show the distribution of wind speed frequency. Remove apostrophe: “... is the systems power curve....”

[AC] Equation 9 expanded with numerical approximation. Wind speed distributions added to figure 20.

line 7: No line break needed.

[AC] Removed

line 8: How constructed? Equation?

[AC] Expanded in text: "The probability of each cluster is characterised using the normalisation wind speed of the pre-processing. The equivalent speed at 100 m height is calculated to determine the frequency in the wind speed bin, using:"

page 31

figure 20: How about MMMIJ? How does the AEP and power curve compare to log profiles with Weibull distribution?

[AR] We consider such comparison out of the scope of this paper. Remember that the aim of this section is illustrating the AWE AEP assessment.

line 3: Did you use 50 calculations to get the power curves of 8 clusters, i.e. more than 6 wind speeds per cluster?

[AR] No, 50 per curve.

[AC] Clarified in text.

2.7 Conclusions

page 31

line 7: What is “fast” about the calculation? Do you mean simplified? Rewrite e.g.: “... used to estimate AEP for a simplified pumping-mode AWES ...”

[AR] The number of optimisations needed is reduced substantially w.r.t. ’brute-force’ calculations: two orders of magnitude faster.

line 11: Shorten to: “...simple logarithmic profiles...”. Would be good to compare power curve and AEP against these log profiles.

[AC] Rephrased the conclusions

line 12: For hourly average profiles. What could be the impact of higher resolution data?

[AR] Such data would capture smaller than wind profile scale phenemona in which we are not interested. I would expect these to be filtered out by the PC analysis.

line 13 & 14: Shorten: Both locations show similar results.

[AC] Rephrased, drawing PCA conclusions for both locations at the same time.

line 14: Which samples do you refer to, all, MMC, MMIJ?

[AC] Rewritten: "The data points for the onshore location are more spread out, indicating a larger variety of wind profile shapes."

line 21: I am not convinced by this conclusion. Why does profile shape similarity proof that clustering is able to differentiate between atmospheric conditions? Also which conditions? If only stable and unstable two clusters might be enough.

[AC] Removed

line 21: Is your process able to determine atmospheric stability (with a certain confidence) solely based on wind profile shape? If so that would be a great addition to your analysis.

[AR] Conclusion was phrased bluntly. There's some sort of relation, but it won't be very strong or useable as suggested.

page 32

line 3: Which wind resource presentation?

[AC] Added "onshore"

line 4: How did you get a distribution from profiles? Would be interesting to see which cluster/ time of year or day contributes how much to AEP.

[AR] Question not completely clear. However the constructing the wind speed distribution is now better explained in the previous section. We don't think the specifics about that are needed here in the conclusions.

line 7: How do 25 optimizations relate to 4 clusters or wind speeds?

[AR] 25 per cluster, so 4 x 25

[AC] Expanded text: "25 optimisations for constructing the power curve of a single cluster"

line 11: How high is the error in comparison to single location clustering?

[AR] Looking at Figure 21, the ML-line at 28 clusters has a similar error as the MMC-line for 14. So you would roughly need twice the clusters to get a similar accuracy.

line 12: Add line break in front of "In the future..."

[AR] Currently this would leave a single sentence paragraph and therefore we choose not to do so.

Response to referee #2

Mark Schelbergen, Peter C. Kalverla, Roland Schmehl, and Simon J. Watson

Thank you for the comprehensive comments. We feel that they were very helpful for increasing the quality of the paper to the current level. Your comments, together with those of referee #1, led to a thorough revision of the paper. The most important changes to the paper include:

1. Including information on orography
2. Discarding the lidar discussion
3. Using one stability metric: the Obukhov length, and corresponding classification for identifying stability trends within the clusters
4. Fitting logarithmic profiles only to the lower part, i.e. <200 m, of the mean profile shapes

We respond to the referee comments by including our answers below the original comments. Our answers are preceded by one or both of the following labels:

AR = author's response

AC = author's changes in manuscript

1 General comments

Paper "Clustering wind profile shapes to estimate airborne wind energy production" describes statistical analysis of wind profiles in order to identify typical shapes of wind profiles that are later used to optimize Airborne Wind Energy systems. The paper is interesting, has scientific novelty and describes some interesting, meaningful and useful results. However, I would suggest that the paper would benefit from some careful editing. The overall quality of English is good, but the presentation of information could use some improvement. I had to apply some effort to follow the argument set out by the authors. Some important procedures (normalization of the wind profiles, fitting of the Obukhov length) are not explained sufficiently, while other aspects of the work that do not lead to any significant conclusions are explained at length.

[AR] Thank you for your comments. We have thoroughly revised the paper using your feedback and that of referee #1 (which had quite some comments). The most important changes include:

1. Including information on orography

2. Discarding the lidar discussion

3. Using one stability metric: the obukhov length, and using classes in terms of the Obukhov length

4. Doing the log fit only for the lower 200 m

2 Specific comments

- P2L6: “Computationally expensive, brute force calculations”. How expensive are the calculations? Can they be done in a few hours on a desktop computer, or is HPC required? An approximate estimation could be added here for the general reader.

  [AR] For the study of Malz, three months of three-hourly MERRA-2 reanalysis data was used which originally took ten days. We decided not to put any concrete numbers here as it is very much dependent on the machine.

- P3L10-L25: Here the word “reanalysis” is used to refer both to ERA5 and DOWA. Although formally DOWA fits the definition of reanalysis, in practice, at least in my experience, the word “reanalysis” is used exclusively for global reanalysis, such as ERA5, ERA-Interim, MERRA, etc. I would suggest using term “modelled datasets” or something similar, to refer to ERA5 and DOWA at the same time, to avoid confusing the reader.

  [AC] Changed "reanalyis" in "modelled" as suggested.

- P5L6: Hourly averages are calculated for the LiDAR data. Why? For the rest of the paper modelling results with hourly resolution are used, which are usually interpreted as the “instantaneous” wind speed, of course taking into account the fact that the mesoscale model cannot represent the turbulent fluctuations. Are the hourly averages comparable with model data?

  [AR] In "Low-level jets over the North Sea based on ERA5 and observations: together they do better", P.C. Kalverla et al., various averaging methods are compared and it is shown that hourly averages align best with the instantaneous reanalysis data. However we discarded the lidar discussion completely since it was not adding much to the paper.

  [AC] Discarded lidar discussion.

- P6L9: In the paper vertical wind speed profiles are fitted using logarithmic profile that has been corrected for stability. From the context I understand that Obukhov length L is the parameter that is being changed during the fitting procedure, but a more precise description of the methodology how the Obukhov length is acquired would be beneficial. If I understand correctly, the same procedure is applied for fitting the instantaneous and long-term mean profiles in Fig. 2. I would like to see some arguments why long-term means can be treated in the way authors do it here. I refer the authors to the paper “Kelly, Mark, and Sven-Erik Gryning. "Longterm mean wind profiles based on similarity theory." Boundary-layer meteorology 136.3 (2010): 377-390.” for a discussion about long-term stability correction of wind profiles.

[AR] We are effectively fitting a mean value of the stability function (not using a mean value of L) which is the correct way to make the mean long term profile. The Kelly paper makes the point that it is not appropriate to calculate a mean L (based on flux measurements, etc) and then produce a mean profile.

[AC] Added: ".. following the approach recommended by Kelly et al. From this, a mean value of the Obukhov length L can be inferred." to second paragraph of Sect 3.2.

- P5L18: Do I understand correctly that the parallel and perpendicular components are calculated for a different reference wind direction in each sample, namely, in each sample the wind velocity at 100 m will have only the parallel component? If so, please expand the description of this procedure in the text.

  [AR] Correct

  [AC] Added: "As a result, the perpendicular wind speed profiles are zero at 100 m."

- P5L19-20 “each sample is normalized by the 90th percentile of its wind speeds at each height”. This sentence is highly confusing, because one interpretation could be that at each height the distribution of wind speeds is constructed, and each height has its own normalization speed calculated as the 90th percentile of all wind speeds in given height. I suggest expanding the description of normalization to remove ambiguity.

  [AC] Rephrased to: ".. the 90th percentile of the sample’s wind velocity magnitudes is used to normalise the wind speed components."

- I think that Figure 2 is not successful in conveying the important information in an effective manner. The horizontal axis for PC1 and PC2 are not the same. The hodograph is small and the markings (dotted line vs uninterrupted line) are not explained. However, my biggest problem is the fact that the x-axis is chosen in such a way that the plots for parallel component and magnitude of the wind, which arguable are the most important features of the profile, are very small. In order to understand the results and conclusions I had to look at profiles that had the width of less than a centimeter (when printed out). I would like the most important features of profile to be plotted large and easy to understand. I would recommend rearranging Figure 2, dividing it into separate figures. To keep paper at a reasonable length, I suggest editing and shortening further sections of the manuscript. These issues continue in further Figures, such as Fig. 8. etc. I recommend focusing on the important parts of the graphs, making them large and omitting less important information for brevity.

  [AR] We have experimented a great deal with how to present these figures. We agree that they are not the best for depicting details in the profiles. However, they allow our results to be shown in a consistent and compact manner, while still showing the overall trends of the profiles. Our conclusions are mostly about the latter and therefore we feel that it is justified to keep the current lay-out of the figures.

  [AC] Changed the horizontal axes for PC1 and PC2 to be the same. Hodograph dotted line explained in caption.

- Section 3.3. requires some editing. Many complicated metrics are introduced, but only briefly, which does not allow the reader to follow the discussion easily. For instance, the fact that higher values of silhouette score mean better similarity within cluster, should be mentioned when the metric is first introduced (P8L10), or at the start of the discussion paragraph (e.g. P10L12)

  [AC] Added: "The dimensionless score ranges from -1 to 1: a negative value suggests that the sample is assigned to the wrong cluster, a value around zero indicates that the sample lies between two clusters, and a high value indicates that the sample is assigned to a distinct cluster". Also added expressions for the fit error metrics.

- P8L1: Why is “itk” used as a parameter name for the number of clusters? It is used twice and then never used again, for instance, in Fig. 5.a. Maybe it is better not to introduce such variable at all.

  [AR] Sorry for the typo. It should have been an italic k, where k refers to the number of clusters (k-means).

  [AC] Corrected

- Fig.5.(a). If the “cluster mag” and “cluster 2d” lines are identical, then maybe one of them can be omitted?

  [AR] They are not identical, but similar. We think it is helpful to show how the two error metrics for the same representation compare, as both metrics are used for drawing conclusions: ’mag’ is used for comparing to the log fit and ’2c’ is equivalent to WCSS.

- Figure 9 and Figure 12: (a) panel could be omitted without loss of important information – all years are the same. Please make the rest of the plots larger – again I was forced to concentrate on very small portions of plots to arrive at conclusions. Maybe Figures 9 and 12 can be combined to save space for legends. I would also suggest using the same Richardson number bins for both figures to make interpretation easier.

  [AR] In contrast to the other panels, the absolute frequency is on the y-axis and serves to show which part of the total dataset is represented by each of the clusters. This is deemed necessary since the table listing the cluster frequencies is discarded (see next comment item).

  [AC] To make interpretation of the stability easier, the samples have been binned using stability classes in terms of the Obukhov length.

- The discussion of filtered vs full dataset, e.g. Table 2, could be omitted to shorten the paper.

  [AR] Agreed

  [AC] Table discarded

- Two different measures of stability are used in the paper – L (Obukhov length) and bulk Richardson number. For clusters (2) and (3) in Fig.8. L values that correspond to neutral stability conditions are reported (P11L13). From Fig.9. and the discussion (P16L11-L17) they are associated with stable conditions. How do authors explain this discrepancy? Additionally, I would like to point out that clusters (2) and (3) are associated with higher windspeeds, and higher windspeeds typically are associated with higher frequency of neutral conditions, see for instance, Holtslag et al. 2014.

[AR] Due to these discrepancies between the fitted L to cluster-mean profiles and the recorded samples stabilities, we decided to change the fitting procedure such that the log profiles are fitted only to the wind speeds in the lower 200 m. This led to both approaches yielding more consistent results. Also the log profile is used in a more valid manner, i.e., the layer < 200 m approximates the surface layer.

[AC] Changed the fitting procedure. Converted the bulk Richardson number to L, allowing direct comparison.

- P11L13: Fitted Obukhov lengths are only reported for clusters (1)- (3). I understand why they should not be reported for cases where the fit cannot reproduce the shape of the profile, e.g. cluster (7). But what are the fitted Obukhov lengths for other clusters, such as (5) and (8)?

  [AC] Added figures 9 and 11, showing how the Obukhov lengths found compare. We think stating the exact values of L is not necessary, as the fits serve most importantly to show to what extent the cluster-mean wind profile shapes deviate from non-adiabatic logarithmic profiles.

- The authors claim that clusters 4-7 indicate potential low-level jets. I do not disagree, but in my opinion, the evidence presented is not strong enough and is somewhat circumstantial. Mostly, because the amplitude of wind speed maxima in Figures 8 and 11 is quite small (except for cluster (7) where I agree that the jet-like profile is quite distinct). Sometimes in literature it is required that wind speed in jets is at least 3 m/s higher than the surrounding flow. Due to the normalization procedure it is hard to estimate how pronounced are these wind speed maxima. The authors' position could be strengthened if more robust evidence to associate the clusters with jets could be provided – a case study or some wind speed profiles that are not normalized and show how distinct the jet actually is.

  [AR] Our conclusion was stated somewhat boldly. The "low-level jets" description only served to describe the shape of the resulting profiles. We rather refrain from discussing what defines a low-level jet and therefore phrased our conclusions more mildly.

  [AC] Rephrased "low-level jets" to "jet-like shapes".

- If the conditions for cluster (5) MMIJ DOWA are analyzed (Fig. 9), than one could conclude that they are very similar to cluster (1) for MMIJ DOWA, with the exception of wind direction, so all the interpretation for conditions for cluster (1) MMIJ DOWA could apply to cluster (5) as well, with the exception that the upwind flow is advected over the water for shorter time, as the prevailing directions for cluster (5) are from South. In fact, in PC1-PC2 axis (Fig. 7) the cluster (5) is next to cluster (1). The problem is that for MMC such conditions are impossible and therefore cluster (5) for MMC is not the same as cluster (5) for MMIJ, nor in the shape of the profile and nor in the location of cluster in the PC1-PC2 axis. Therefore, I cannot agree with the statement that cluster (5) for MMC is similar to MMIJ (P17L20).

  [AR] Our statement was not phrased carefully and therefore misinterpreted.

  [AC] Rephrased to: "The profile shapes for MMC-5–7 are (slightly) jet-shaped, as is the case for the offshore clusters MMIJ-4–7." Also, the parallels drawn by the reviewer for the offshore clusters MMIJ-1,5 are now explicitly described in Sect. 4.2.

- Table 3, column “Most frequent area”. What is this classification based on? Are the orography and land use data used in DOWA available? If so, then adding this information to Figure 15 as a plot would strengthen the authors’ argument, because currently, although I agree that features seen in Figure 15 are probably related to orography, again I would argue that more evidence is needed to support the authors’ claims. I would not expect the general reader to be familiar with the topography of the Netherlands and Germany

  [AR] DOWA does not include such data, at least it is not published. The classification was based on ’quick and dirty’ observations, so not precise and well supported by data. We agree that the topography was missing.

  [AC] We have discarded the "Most frequent area" column in the table, as they were blunt conclusions. We have included information on the surface elevation to Figure 1 and 16, such that the claims in the text are now better supported.

3 Technical comments

- Paper referenced as “Sommerfeld et al.”: “Improving mid-altitude mesoscale wind speed forecasts using LiDAR based observation nudging for Airborne Wind Energy Systems” does not seem to contain anything related to k-means clustering. Maybe a different paper with the same first author is meant instead: Sommerfeld, Markus, et al. "LiDAR based characterization of mid altitude wind conditions for airborne wind energy systems." Wind Energy 22.8 (2019): 1101-1120.

  [AR] Correct

  [AC] Changed reference

- Order of figures. Figure 8 is referenced in text before Figure 7. Figure 8 is referenced in P11L10. Figure 7 is referenced in P11L21.

  [AC] Corrected

- P1L19: “Deviating profiles are likely to occur”. I suggest “deviations from the expected profile shape are likely to occur”.

  [AC] Replaced with: "Moreover, within this layer, not all wind profiles can be described well with these relationships."

- P4L11: “Better representation of the coastal morphology”. “Coastal morphology” is the study of natural processes that change the shape of coastline, e.g., erosion or sediment transport. I suggest using “coastline” or “better resolution of coastline” instead.

  [AC] Replaced with “coastline”

- P16L9: “the wind profile is typically will mixed”. Probably, “well mixed”.

  [AC] Corrected