Response to 2nd review of referee #1

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

Thank you for another comprehensive and useful list of comments. We feel that they were very helpful for removing the last flaws in the paper. The most important changes to the paper include:

1. The profiles in the 2nd column of Fig. 2 and 11 are now better explained.
2. Section 4.3 includes an explanation on why only 1 cluster groups the unstable profile shapes.
3. A lot of unclear relationships in the text are clarified.

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:

1. [AR] = author's response
2. [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).
The manuscript improved considerably from the previous submission. Its content is more focused and its language is much clearer than before. Following are some general comments and language corrections.

Language:

The language improved a lot since the previous submission and previous ambiguities and mistakes have been removed. The article seems rather wordy. I could not get the exact word count, but copy-pasting the text into a text editor resulted in over 12000 words. Certain sentence structures with the word ´´respectively" as well as gerund forms repeat fairly often. I therefore recommend rewriting and shortening some paragraphs by combining sentences and simplifying long and complicated sentences. Some examples can be found in section 2.

[AR] Improved the language as suggested by the reviewer in section 2.

Figures:

The added figures 9 and 13 do not add significant information and could be summarized as text boxes in figure 8 or 12 which they refer to, e.g. don't mention L value, but associated stability bin in text box in each profile plot 1-8

[AR] These plots were added to depict how the L found for each shape compares to the others, this information would be lost by using the approach suggested by the reviewer. This information is used throughout e.g. Sec. 4.2.

Wind data: The usage of Ri, L and Ψ to assess atmospheric stability throughout the paper is rather confusing. Since Ri and L are interchangeable, it might make sense to just use one of them.

[AR] We only present results in terms of L. Ri is only introduced for explaining how we get to L from the ERA5 data.

2 Specific comments including technical corrections

2.1 Abstract

page 1

line 4: ´´vertical variation" sounds like vertical (w) component of wind velocity. Maybe clarify by writing: variation in height or variation of horizontal wind speed.

[AC] Changed to: "... variation ... with height."

line 6: Why introduce AWE and AEP, but not DOWA abbreviation in abstract?

[AR] I don't use DOWA multiple times in the abstract, in contrast to AWE and AEP.

line 7-8 + 12: The abstract should include a summary of information found in the paper. It should not provide results or conclusions.

[AR] To my understanding the abstract should include information of the whole paper, including results and conclusions.

line 10: Add: " ... for each wind speed profile shape".

[AC] Changed

line 12: 4 cluster error relative to what?

[AC] Changed to: "... the difference in AEP with respect to the converged value is within three percent for four or more clusters."

2.2 Introduction

page 1

line 14: Grammar: Comparative between altitude and turbines is not parallel; Comparison of stronger and more persistent winds to what? Rewrite: "AWE systems use tethered flying devices to harness energy at higher altitudes, typically heights above 150 m (Malz et al., 2019; Salma et al., 2019), where wind is generally stronger and more persistent (steady) than at heights of (reachable to) tower-based wind turbines."

[AC] Rephrased

line 20: Add reference to validity of log and power law beyond surface layer

[AR] Such references are given below seperately for the two relationships.

equation 1+2: Only time u used for wind speed. Other times v or is $\hat{v}$ only used for normalised wind speeds?

[AC] Changed to v

page 2

line 13: Wording: "assumptions... frequently violated in practice...". Not the assumptions are violated, it's the range of validity that is ignored.

[AC] Rephrased

page 3

line 1: Change sentence order. Before you write about log fit. Rephrase to 1 sentence: "The wind direction can also vary substantially with height in the lower atmosphere."

[AC] Rephrased

line 2: What are "...scalar quantities..."? Remove before comma

[AC] Removed

2.2.1 Wind dataset

page 4

line 9: "..., any dataset containing time series..." (singular and add any)

[AC] Corrected

line 10: Rephrase: "..we focus on sensitivity of the AWE system power production ..."; add "..to the wind profile shape ..."

[AC] Corrected

line 13: What does typically refer to ? power assessment of wind turbines?

[AC] Rephrased

line 20: The spatial resolution needed for this study (country wide) does not exist for lidar. This is a justification to use model data instead.

[AR] As is implied.

line 21: Are these “typical” sites? Did you compare them to other sites? Maybe just write “representative” or “exemplary”.

[AC] Left out typical

line 24+25: Repetitive use of “grass-land”. Maybe merge sentences; spelling: grassland

[AC] Corrected

line 26: Could mention that “down-scaled” means higher resolution

[AC] Sentence removed

line 27: Remove sentence and put in parenthesis in previous sentence to reduce wordiness of text: “The sites (shown in figure 1) were chosen...”. Add reference to studies using these locations wind data.

[AC] Corrected

2.2.2 ERA5

page 5

line 2: ERA5 reference linked to https://cds.climate.copernicus.eu/cdsapp#!/home. Could instead use:

i.
```
\begin{small}
@misc{website:era5,
Author = {Hans Hersbach and Dee Dick},
Month = {November},
Title = {{ERA5} reanalysis is in production},
url = {http://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production},
Year = {2016},
Publisher = {{ECMWF} - European Center for Medium Range Weather Forecast},
note = {last accessed: 22.10.2019}}
```

[AC] Cited as requested by ECMWF: https://confluence.ecmwf.int/display/CKB/ERA5

line 2: Add: reference to ECMWF

[AR] Is considered not essential after having cited ERA5.

line 7: “... is performed on the DOWA data...”

[AC] Modified

line 7: Remove: “As is explained later on”

[AC] Removed

2.2.3 DOWA

page 5

line 10: Remove: “The”; spelling of “downscaled” different from above.

[AC] Modified

line 13: Add “grid spacing” and “grid points” for clarity.

[AC] Rephrased

page 6

line 1: Add reference to improved DOWA performance.

[AR] The last sentence covers such a reference: Kalverla 2019

line 3: Clarify “routine weather stations”

[AC] the KNMI’s network of automated weather stations.

2.3 Clustering procedure

2.3.1 Prepossessing of the wind data

page 6

line 11: Remove: “the” before “wind speed” similar to no “the” before “direction”

[AC] Removed

line 11: Remove: “the” after “Therefore”, because general wind profiles and not specific ones.

[AC] Removed

line 13: Separate into 2 sentences. Explain what you mean by processed using its own properties.

[AC] Removed

line 14: Rewrite for clarity; for example: “The wind speed components are expressed as parallel and perpendicular components relative to their reference wind velocity at 100 m, similar to Kalverla et al. (2017) and Malz et al. (2020a), thereby making them independent of wind directions.”

[AC] Rephrased, also taking into account comments of referee 1

line 18 + 19: Active voice and clarity; for example: “Wind velocity components are normalised using each profile’s 90th percentile wind speed, because it reduces the amount of outliers in comparison to using the maximum speed of each profile. These normalised and decomposed samples are referred to as wind profile shape. ”

[AC] Rephrased

line 21: Use irregular,atypical or unconventional instead of eccentric, add: "...low wind speeds."

[AC] used irregular

line 24: Clarify: wind resource representation. Is this the wind speed probability distribution similar to a Weibull distribution? ;

[AC] removed

2.3.2 Principal component analysis of the wind profile shape dataset

page 6

line 27: Replace: "while" with "which"

[AR] Would result in incorrect language. I don't see why I could not use "while" here.

page 7

line 7: Replace: "layer" with "height"

[AC] Rephrased

line 14: Explain Fig 2 PC unit vectors. Are these the unit vectors in PC1 and PC2 domain which have a length not equal to one in $\hat{v}, height$ domain? Is it that the length ($\sqrt{PC1_{\parallel}^2 + PC1_{\perp}^2} = 1$) of the parallel (orange) and perpendicular (blue) component at each height is one? Are the orange and blue lines the direction of the PC at each height? Add: $\hat{v}$ is normalised wind speed magnitude in text.

[AC] Explanation expanded

line 16 + 17: Word order: mostly characterises.

[AC] modified

line 18: What shows the large contribution of both PCs? Is it that the $\hat{v}$ magnitude is high? Does it make sense that most variance is top and bottom? I would expect that profiles within one cluster have similar wind speeds at high altitudes as they are probably driven by similar large scale weather phenomenon or is this lost due to normalisation and PCA?

[AC] A large PC coefficient. We explain it better in the text now.

[AR] We expect most variance at both ends. For a large part this is explained by the pre-processing approach that is used, i.e., the choice of the reference height.

"I would expect that profiles within one cluster have similar wind speeds at high altitudes as they are probably driven by similar large scale weather phenomenon or is this lost due to normalisation and PCA?" - I'm not sure why you would expect this. But indeed I would expect the correlation between the clusters and large scale phenomena is weakened by the two steps in the preprocessing, so both dissecting the reference wind direction and normalisation.

line 20: Explain wind profiles shapes along PC1 and PC2 more. Why did you chose minus and plus one standard deviation as multipliers? What do you want to show with these profiles?

[AC] Explanation expanded

[AR] We want to illustrate how the PCs can be physically interpreted.

line 22: What are the eigenvalues of PCs? What are the retained PCs?

[AR] They follow from the PC analysis. It's believed to be beyond the scope of the paper to explain the details on this.

[AC] Left out sentence that mentions the eigenvalues.

line 32: add: "Figure 4g which shows onshore data will be ..."

[AC] modified

page 8

figure 2: What are the orange and blue dashed lines in column 2 PC1 and PC2? I think you did not explain $\hat{v}$ anywhere in the text or the caption. What does it mean that both PC1 and PC2 have negative parallel components below 300m?

[AC] Dashed lines were explained in last sentence. $\tilde{v}$ is now introduced in text, it is not deemed necessary to also state it in the caption. That the parallel wind speed component at this height decreases for an increase in the PC. The PC coefficients are now also better explained in the text.

2.3.3 Choosing the number of clusters

page 8

line 2: Replace: applied instead of employed

[AC] Changed

line 3: Replace: "...represented by its centroid ..."

[AC] Changed

line 9: Do wind profiles have such a structure?

[AC] No, expanded in text.

page 9

figure 4: Mention for offshore a) and onshore b) in caption. Put figs 2 and 11 in parenthesis? replace: "deviation away from their mean" ?

[AC] Corrected

line 1+2: Remove “Moreover”. Add reference to elbow and silhouette method find appropriate k.

[AR] Methods are explained in next paragraphs so reference is deemed unnecessary.

[AC] Removed moreover

line 3: Replace: applied instead of employed. Can add on how evaluated: by comparing the estimated AEP

[AC] Replaced

line 5: Reference: appropriate choices of k?

[AR] Not deemed necessary

page 10

figure 5 a): Remove: “the” before cost function and cluster cohesiveness. The error between which data? mean wind speed error between each profile and its respective cluster centroid?

[AR] See text for error definition

[AC] "The"’s removed

figure 5: Add: “k-means” before clustering ; remove: point before (b)

[AC] Corrected

line 5: Maybe use scaling instead of denormalisation; remove “of the cluster to which it is assigned”; replace “of” with “used in pre-processing”

[AC] Corrected, kept de-normalisation as it is more informative than scaling.

line 7: Add : ε_{ij} here to make equation easiert to understand. Define between which data the error is calculated.

[AC] Reformulated

line 8: Is the representation accuracy shown in Fig 5?

[AC] Yes, reference to figure added in text.

page 11

line : Is the error only based on different 5 Ψ values?

[AR] Actually only the values for L are restricted to 5 values. Ψ is a function of L and the height.

line 5: Sentence needed?

[AC] Reformulated

line 6: Add reference to Ψ function; replace every with each

[AR] Introduced in introduction

line 14: Grammar; add (bottom) to Fig 5a

[AR] Grammar?

[AC] Added "The lower panel of Fig. 5a ..."

line 15: Is “ Note that...” sentence necessary here? Maybe move to where you introduce WCSS?

[AC] Sentence removed

line 19: Why not just wind speed error? $\overline{\varepsilon_i}$ never used in any equation

[AC] Reformulated. Expression removed.

line 20: Replace top and bottom with above and below.

[AC] Added "... of the vertical grid."

line 21: Replace vertical grid with dataset or rewrite

[AC] Rewritten "... of the vertical grid points of DOWA."

page 12

line 5: Limited to what? Rather a fixed number of clusters.

[AC] Rephrased

line 6: What do you mean by: “ ... and our aim to present a meaningful analysis and interpretation of the resulting clusters.” and how did this affect your choice of k?

[AR] Using many clusters would make it hard to find relations between clusters and atmospheric phenomena manually.

[AC] Rephrased a little

line 8: Isn’t it just the silhouette score and not the mean silhouette score in both sentences?

[AR] No, the mean is calculated over all samples in a cluster.

2.4 Cluster wind resource representation

page 13

figure 7: What does the asterisk mean? Maybe rename axis label to $PC1_{onshore}$ / $PC1_{offshore}$?

[AR] It refers to the average of the two locations.

[AC] Rephrased, hopefully that clarifies it better.

line 5: Even though clusters are not the same between on and offshore.

[AC] Rephrased a little

line 6: Replace: full with entire

[AC] Replaced

2.4.1 Cluster representation for the offshore location

page 13

line 9: Replace: of with at; add: their centroids

[AC] Corrected

line 15: Add: backwards transformed from PC space to normalised wind velocity

[AR] This is implied by the definition of the cluster-mean wind profile shape introduced in Sec. 3.3.

line 15: Explain: stability function is varied continuously?

[AC] Reformulated

page 14

table 2: Replace: at with in?

[AC] Rephrased

line 3: Is this sentence about a) ? remove: is, and & represented " In contrast to the other panels, the absolute frequency on the y-axis serves to show which part of the total dataset is presented "

[AR] Rephrased

line 6: Replace: for with of

[AC] Replaced

line 7: Where is "here"? Figure 9, 10? Isn't it only one distribution?

[AC] Rephrased

line 9: Which data did you use? L is always calculated using surface data (see: http://glossary.ametsoc.org/wiki/Obukhov_length)

[AR] We infer L from $\mathrm{Ri_B}$. The latter we calculate over an elevated layer. So the approximation of L is not based on surface data. However, it is only an approximation.

page 16

figure 9: You could replace figure 9 and 13 with legend or text box in figure 8 and 12 to save space. Using a figure to show this seems unnecessary

[AR] These plots were added to depict how the L found for each shape compares to the others, this information would be lost by using the approach suggested by the reviewer. This information is used throughout e.g. Sec. 4.2.

page 17

figure 10: Add a),b),c),d),e),f) to caption; Describe what is shown in a,b,c. What is the benefit of using this v_{100} binning or what is the reason for having equal bin size for wind speed?

[AR] Difference between a,b,c is easily inferred from the legend.

[AC] Letters added.

2.4.2 Interpretation of the offshore cluster representation

page 18

line : Maybe makes more sense to put interpretation in the same paragraph where figure is described? This way you have to flip back and forth a lot as

[AC] The current order is used to have a nice flow of information.

line 5: Add: magnitude of profiles? replace: well-described

[AC] Modified

line 6: Replace: shear and veer

[AC] Replaced

line 8: Shorten: North-West direction

[AR] Current text is more precise.

line 10: Equilibrium profile meaning a unstable shape?

TBD

line 25: Remove: “The frequent”; Rewrite: Winds at this location with a southerly component...

[AC] Rewritten

line 29: Remove: “... and gradually”

[AC] Removed

page 19

line 1: Replace: “abrupt kink” with “sharp bend”

[AC] Replaced

line 5: Replace: “rather than” with “and less often with”

[AC] Rephrased "... more often for winds with a westerly than southerly component."

2.4.3 Comparing the on- and offshore cluster representations

page 19

line 8: Rewrite: The onshore data at the met mast Cabauw is clustered using the same methodology.

[AC] Rewritten

line 9: Compared how and where?

[AC] Rephrased

line 12: Rewrite: “the mean profile shape below 200 m is in accordance with a stable logarithmic profile.”

[AC] Rewritten

line 15: Increased relative to what?

[AR] The offshore mean shape as implied by context.

line 16: Add reference to where they are plotted

[AR] Referred to at the start of the paragraph

line 16: If they share the same coordinate system does that mean that both locations have the same PCs?

[AR] No, the location average PCs were used for the coordinate system.

[AC] Rephrased

line 24: “...clustering algorithm...”; “.... for each cluster.”

[AC] Corrected

line 25: Replace: more or less with onshore closely resembles offshore

[AC] Rephrased - align covers statement better

line 28: Remove: again

[AC] Removed

line 29: Obukhov length determined how?

[AC] Rephrased

line 30: Meaning? “show an increase in wind shear”. Increasing wind gradient from 1 to 3?

[AR] Yes, the reader is ought to understand at this point.

line 34: Replace: “turning” with “rotation” and “kink” with “sharp bend”

[AC] Rephrased

page 20

line 4: Rephrase: The frequency distribution over the first five onshore clusters is more balanced than the offshore clusters which show one distinct dominant cluster.

[AC] Rephrased

line 9: Rephrase: Convection occurs in the presence of daytime solar irradiation which leads to the development of well-mixed wind speed profiles.

[AC] Rephrased

line 12: Explain: “Patterns in times of occurrences” of what?

[AC] Added reference to figs.

line 13: Explain: Meaning of “ almost identical bin distributions” if wind speed distribution is different?

[AC] Rephrased a little.

line 14: Explain: Reason for choosing wind speed bin limits this way?

[AC] Added "... thereby the distributions of the individual clusters are easily related to the uniform general distribution and compared among one another."

line 15: Rephrase: Total frequency of each bin is roughly the same over the entire dataset. Explain: What is the benefit of this approach?

[AR] See above

line 16: How can they show similar stability distributions if figure 9 and 13 are totally different?

[AC] Added more precise description: "In the case of the stability distributions, the onshore location shows a tendency to more stable conditions for all clusters."

line 22: Replace: “... start to be observed” with “are observed.”

[AC] Replaced

page 21 + 22

figure 11+12: See same figures for offshore

[AR] No Figure 12 offshore equivalent (Figure 8) are given.

page 23

figure 13: Very strange that 6 out of 8 profile shapes are stable. Add to caption: fitted up to 200m

[AR] Not so strange: there is little diversity in the shape of the unstable profiles, therefore all the associated samples are grouped together by the clustering. Added more text on this in Sect. 4.3

page 24

figure 14: Same as for previous similar figure. Rephrase: “Frequency distributions by time of occurrence...”

[AC] Corrected

2.4.4 Spatial frequency distribution of wind profile shape clusters

page 25

line 2: Replace: “are” with “were”

[AC] Corrected

line 3: Replace: “for” with “from”

[AC] Corrected

line 10: Remove: “So”

[AC] Removed

line 12: Replace: “inherent” with “synonymous” Inherent turns the relationship between reducing error and increasing n clusters around.

[AC] Rephrased: "increasing the number of clusters reduces the error"

line 21:Can you quantify how much they “look alike”?

[AR] You could, but I don’t think such a metric would be very intuitive.

line 22: 21.7% ?Reference Table 3 here already

[AC] Pushed introducing the table more forward.

line 27: Which terrain features? Elevation? Forests, cities?

[AC] "Terrain features" replaced by "orographic features"

page 26

table 3: How do you quantify similarity?

[AR] I don’t.

page 28

figure 16: Mention in caption: Each sample of every grid point is assigned to the cluster with the closest centroid.

[AR] Should be clear from the text already.

2.5 Efficient AWE production estimation using the cluster representation

page 29

line 6: Add: power curve and AEP

[AR] I don’t think this would help the reader understand. If so more explanation would be required as we actually promote using multiple power curves.

line 7: Replace: “winds higher up” with “winds at higher altitudes”

[AC] Replaced

line 11: Add: “wind speed distribution within the corresponding cluster”; Replace: “of the clusters” with “of each cluster”

[AC] Phrased as "Each power curve together with the corresponding cluster-specific wind speed distribution yields the AEP contribution of the respective cluster."

2.5.1 Constructing the power curves

page 29

line 13: Never heard “construction” in the context of power curves. Replace throughout the document with derive or determine ?

[AC] Using derive now.

line 15: Fix grammar (reel-out twice); mention change in angle of attack;

[AC] Rephrased

page 30

line 2: What is V3 kite?

[AC] V3 removed

line 4: Add: “average energy production”

[AC] Rephrased: "The proposed AEP estimation requires characterising the maximum mean cycle power for a large variety of wind conditions."

page 31

line 1: Replace: “are exceeded” with “would be exceeded”? Simplify sentence: “Reel-out” repetition in same sentence. Why is reel-out duration important?

[AC] Rephrased: "the reeling speed is kept zero as long as the tether force does not exceed its limit" and "During reel-out, the tether force should yield a high power, while increasing the fraction of time spent producing energy."

line 6+8: What are cycle settings? Are these the fixed constraint in table 5?

[AC] Rephrased: "Table 5 lists the cycle setting parameters, which are used as optimisation variables, together with their respective limits."

line 15-18: a bit wordy, could be shortened to 1 or 2 sentences.

[AC] Rephrased

line 17: Add “reference wind speed”; Remove: “to yield the absolute wind profile”

[AC] Removed

table 5: Is tether diameter and associated drag considered in the model?

[AR] Yes, fixed as it is a design parameter.

line 20: Replace: “smallest” with “lowest”; “whole” with “entire”

[AC] Replaced

page 32

line 9: Does this intersection depend on tether length?

[AR] Yes, this is implied by relating the intersection height to the kite height earlier.

figure 18: Replace: “from ” with “in”; also reference which section describes these profiles in fig 12; Add: “pre-determined” before cut-in

[AC] Modified

line 12: Simplify with active voice: “The depicted trajectories highlight operational changes occurring at different reference wind speeds”; Specify how the approach / attitude changes. Is this the reason for the strange v_{ref} in legend of Fig 19?

[AC] Rephrased

[AR] The remaining of the paragraph describes these approach changes.

line 13: The fact that the tether length constraint is always active indicates that the global optimum is beyond this constraint. Can you comment on whether you tried out different settings or what the reasons could be for maxing out this constraint?

[AR] The constraint was chosen as it is a representative maximum value for the Kitepower system. The use of a different maximum was not extensively assessed.

line 14: Same is true for lower wind seeds though?!

[AC] replace "for" with "below"

line 16: criterion is also a constraint?

[AC] Yes, rephrased

line 19: They are similar. Are they cubic as we would expect from $P \sim \rho v^3$

[AR] Doesn’t appear to be cubic.

page 33

figure 19: Add: “same normalised wind speed profile shape”; Approach = attitude? Specify what you mean by change in approach.

[AC] Added "cluster-mean wind profile shape"

[AR] Explained in text

figure 20: Frequency better in %; Add: “scaled cluster-mean...”; Why did you aggregate 4 bins into 1? Not sure if you have to mention this here.

[AC] Added "scaled"

[AR] It’s only for illustrative purposes, therefore I felt that the figure was the right place to mention this.

line 1: LLJ benefit because of high wind speed at low elevation angle hence low cosine losses?

[AR] In the case of a LLJ and a fixed minimum and maximum tether length (which is more or less the case in my study as the pumping tether length coincides with its upper bound for all wind speeds), the kite does not necessarily see higher wind speeds for an increase in elevation angle, in contrast to monotonic profiles. The cosine loss does not play a big role here, as we do not observe a tendency to increase the reel-out power by increasing the elevation angle (at least not at low wind speeds). At a certain point the elevation angle is increased, but this is driven by keeping the duty cycle high.

2.5.2 Estimating the annual energy production

page 34

line 2: Replace “Constructed” with “derived” ?; Grammar: “... are used to calculate the average ...”

[AC] Replaced

line 4: Is P_i the power curve or power?

[AC] Replaced by "maximal mean cycle power"

line 8: Is the numerical error so significant that you need to use 100 bins instead of the way more common 1m/s or 0.5m/s bins?

[AR] It is significant for the AEP convergence analysis.

line 10: Isn’t it that you just use the reference wind speed v_{100m} to calculate the frequency and they make up these bins?

[AR] No the frequency is determined on the basis of the normalisation wind speeds.

[AC] Rewritten paragraph

line 19: What do you mean by “...inaccuracy of the cluster wind resource representation”? Does it mean that fewer clusters lead to higher inaccuracies because of averaging or misrepresentation of the entire wind resource?

[AC] Yes, however, I wanted to emphasize here that it’s not just the wind resource that introduces errors, therefore rephrased the sentence.

line 22: Rewrite: It's hard to understand how MMC and ML relate to each other and what 16 and 32 clusters mean. Does it mean that 16 MMC clusters have the same difference to converged value as 32 ML clusters? Doesn't this indicate that you need twice as many clusters to achieve similar quality using the ML approach? Which makes sense as it combines on and offshore locations and therefore different flow regimes?

[AC] Rephrased: "The AEP error at 16 clusters for the MMC representation is similar to the AEP error at 32 clusters for the ML representation, which suggests that the ML representation needs twice the number clusters to yield the same accuracy as the MMC representation."

line 28: Reference figure 20 here again. 50 optimisations per power curve because of the step size of Δv_{scale} you chose? Maybe mention step size for clarification.

[AC] Rephrased

line 29: Remove of: "...half the number"; Replace; ".. can be used" with "yield similar results with half the computational cost"

[AC] Corrected

line 30: Grammar: "used for generating" with "used to generate"; Rewrite "in comparison to brute force, where 8760 optimisations ...."

[AC] Rewritten

page 35

line 1:"orders of magnitude lower"

[AC] Corrected

figure 21:"Comparison of AEP conversion between MMC onshore and ML offshore"

[AR] I believe "conversion" should be "convergence"

[AC] Rephrased "Comparison of the AEP convergence with increasing number of clusters for the onshore reference location using the MMC and ML cluster wind resource representations."

2.5.3 Conclusion

page 35

line 1: Remove: "have"; add: "a set of normalised wind profile shapes"

[AC] Removed "have"

[AR] Shape already implies normalised

line 6: Did you quantify the occurrence of LLJs? There are also other ways of doing that. You approach allows the inclusion of wind profile shapes into the wind resource description (equivalent to wind speed distribution/ Weibull for conventional turbines).

[AC] Rephrased

line 7: Grammar: “We demonstrated this methodology for two reference locations on and offshore based on the DOWA dataset.”

[AC] Changed

line 9: What do you mean by “expressed in terms of wind velocity at 100 m?” and where did you do it?

[AC] Rewritten

line 10: “... profile shape variance.”

[AC] added "... in the dataset"

line 15: Rewrite: “The DOWA dataset is partitioned using k-means clustering. The resulting cluster-mean wind profile shapes are used to represent the wind resource, thereby reducing the wide range of wind conditions to a reasonable number of wind profile shapes.”

[AC] Rewritten

line 18: Rewrite: “Although some variability is lost by only using the mean cluster profiles, ...”

[AC] Sentence removed

line 21: Order: “... 8 offshore clusters show 3 monotonic ...”

[AC] Not sure what’s meant exactly, but rewritten a bit.

page 36

line 1: Replace: “entire DOWA domain”

[AC] Replaced

line 3: Replace: “... between the profile shape and terrain.”

[AC] Replaced

line 3: Replace: “... in contrast to 8760 for an ...”

[AR] "against" works here

line 16: Grammar: “...ML enables an assessment...”

[AC] Corrected

line 18: Replace: “...in estimating the AEP...”

[AC] Replaced

Response to 2nd review of referee #2

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

Thank you for another useful list of comments. We feel that they were very helpful for removing the last flaws in the paper. The most important changes made to the paper based on the comments of referee #1 are:

1. The profiles in the 2nd column of Fig. 2 and 11 are now better explained.
2. Section 4.3 includes an explanation on why only 1 cluster groups the unstable profile shapes.
3. A lot of unclear relationships in the text are clarified.

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:

1. [AR] = author's response
2. [AC] = author's changes in manuscript

Referee comments

1. I am confused by the use of parentheses as a way to soften the meaning of words or to make some of the words feel "optional". According to my understanding such use of parentheses is not considered good style. Please consider removing the parentheses in sentences such as P18L16 "we also see frequent (very) stable conditions". I think that here "we also see frequent stable or very stable conditions" is much more precise.

   [AC] Agreed, parentheses used in this way are removed.

2. The overbar in formula (5) is hard to see. Please check that this gets corrected in the final proof.

   [AC] Increased spacing between overbar and the fraction deliminator.

3. I would suggest adding a subscript for the symbol for the reference height z in (6), to avoid confusing that with the z in formula (2), where z has the meaning of a coordinate.

   [AC] We use $\bar{z}$ now instead.

4. I think that some additional sentences are still needed in describing the sample normalization procedure (P6L16-L19), because I still struggled to understand the details. I would suggest explicitly defining reference wind velocity for each

sample as the wind speed and direction of that sample at 100 m height, explicitly noting that each sample has its own reference wind velocity, because my first association is that “reference” is something that is common for all samples. I would suggest rephrasing P6L17 as “the value of the perpendicular component of the wind speed profile is 0 for each sample and for all averaged profiles”.

[AC] Rephrased, now only referring to a reference height in stead of a reference velocity: "... the wind speed components are expressed as parallel and perpendicular components relative to the wind velocity at a reference height, which we have chosen to be 100 m. As a result, the value for the perpendicular wind speed at 100 m is zero and the reformatted wind profile is independent of the wind direction at 100 m."

5. Caption of Figure 8. I would suggest: “In the hodograph, the lowest level is indicated by the dotted line connecting the lowest level to the origin of coordinates”.

   [AC] Rephrased to: "In each hodograph, the lower end of the profile is indicated by the dotted line connecting the lowest height point to the origin."

6. I would rephrase the statement that MMIJ-5 has “relatively large deviations from logarithmic profile” P9L21, because later authors state that its shape is well described by unstable logarithmic profile, similarly in the caption for Figure 8 I would like it to be pointed out that these are stability adjusted logarithmic fits. I am stressing this point because a slightly distracted reader might confuse what exactly is meant by “logarithmic profile” because in some cases this term describes only the neutrally stable profile shape.

   [AC] Replaced by: "Because wind speed increases monotonically with height in the logarithmic wind profile relationship, it can not describe these type of profile shapes.". Specified in the captions of figs. 2, 8, 11, and 12 that non-adiabatic log profiles are fitted.

Clustering wind profile shapes to estimate airborne wind energy production

Mark Schelbergen[1], Peter C. Kalverla[2], Roland Schmehl[1], and Simon J. Watson[1]

[1]Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands
[2]Meteorology and Air Quality Section, Wageningen University, PO box 47, 6700 AA Wageningen, The Netherlands

Correspondence: Mark Schelbergen (m.schelbergen@tudelft.nl)

Abstract. Airborne wind energy (AWE) systems harness energy at heights beyond the reach of tower-based wind turbines. To estimate the annual energy production (AEP), measured or modelled wind speed statistics close to the ground are commonly extrapolated to higher altitudes, introducing substantial uncertainties. This study proposes a clustering procedure for obtaining wind statistics for an extended height range from modelled datasets that include the ~~vertical~~ variation of the wind speed and direction with height. K-means clustering is used to identify a set of wind profile shapes that characterise the wind resource. The methodology is demonstrated using the Dutch Offshore Wind Atlas for the locations of the met masts IJmuiden and Cabauw, 85 km off the Dutch coast in the North Sea and in the centre of the Netherlands, respectively. The cluster-mean wind profile shapes and the corresponding temporal cycles, wind properties, and atmospheric stability are in good agreement with literature. Finally, it is demonstrated how a set of wind profile shapes ~~and their statistics are~~ is used to estimate the AEP of a small-scale pumping AWE system located at Cabauw, which ~~require~~ requires the derivation of a separate power curve for each wind profile shape. Studying the relationship between the ~~calculated~~ estimated AEP and the number of ~~site specific~~ site-specific clusters used for the calculation shows that the ~~AEP error is within three percent of~~ difference in AEP relative to the converged value is less than three percent for four or more clusters.

1 Introduction

Airborne wind energy (AWE) systems ~~use~~ employ tethered flying devices to harness energy ~~at higher altitudes than~~ above the operational height range of tower-based wind turbines~~, typically heights~~. Typically these devices operate above 150 m (Malz et al., 2019; Salma et al., 2019), where wind is generally stronger and more persistent ~~. For estimating~~ than in the surface layer. To estimate the annual energy production (AEP), measured or modelled wind speed statistics close to the ground are commonly extrapolated to higher altitudes to obtain the wind speed statistics in the full operational height range of the AWE system ~~are commonly approximated by extrapolating the wind speed distribution close to the ground to higher altitudes~~ using either the wind profile power law or the logarithmic profile (e.g., Heilmann and Houle, 2013). This way of representing the wind resource introduces substantial uncertainties since the aforementioned wind profile relationships are not strictly valid beyond the surface layer. Moreover, within this layer, not all wind profiles can be described well with these relationships.

The power law is a simple empirical relationship which can be used to relate the wind speed ~~u~~ v at one height z_1 to that at a different height z_2 and has the form:

$$\underline{u}v(z_2) = \underline{u}v(z_1)\left(\frac{z_2}{z_1}\right)^{\alpha} \quad , \tag{1}$$

where α is an empirical shear exponent factor related to the surface properties. The power law is normally applied up to around 100–200 m (Peterson and Hennessey, 1978), and does not offer enough flexibility to describe the variety of measured wind profiles (e.g., Park et al., 2014).

The logarithmic wind profile is frequently used to estimate the variation in wind speed with height over a flat surface. This profile is based on physical arguments and a form of the profile has been well established based on Monin-Obukhov similarity theory (Monin and Obukhov, 1954). In this non-adiabatic form, the mean wind speed ~~u~~ v at height z is given by:

$$\underline{u}v(z) = \frac{u_*}{\kappa}\frac{v_*}{\kappa}\left[\ln\left(\frac{z}{z_0}\right) - \Psi\left(\frac{z}{L}\right)\right] \quad , \tag{2}$$

in which ~~u_*~~ v_* is the friction velocity, κ is the von Karman constant, z_0 is the roughness length, Ψ is a stability correction function, and L is the Obukhov length that is often used to evaluate atmospheric stability and is positive (negative) for stable (unstable) stratification and infinite for neutral stratification. Holtslag et al. (2014) proposes the following stability correction functions:

$$\Psi\left(L \leq 0\right) = 2\ln\left(\frac{1+x}{2}\right) + \ln\left(\frac{1+x^2}{2}\right) - 2\arctan\left(x\right) + \frac{\pi}{2}, \qquad x = \left(1 - 19.3\frac{z}{L}\right)^{\frac{1}{4}} \quad , \tag{3}$$

$$\Psi\left(L \geq 0\right) = -6.0\frac{z}{L} \quad . \tag{4}$$

Although the ~~assumptions underlying the~~ logarithmic wind profile ~~are frequently violated in practice (e.g., Optis et al., 2014), it is widely used, particularly for~~ is less accurate under stable stratification above the surface layer (e.g., Optis et al., 2014), the relationship is often applied to any condition in wind resource estimation.

The value of L is not easily measured or derived from model data and is generally inferred indirectly. One way to do this is to fit a functional form of the logarithmic wind profile with stability correction to the wind velocity magnitude profile. Such an approach is outlined by Basu (2018), using three levels of wind speed. Another common way of estimating L, is by inferring it from the gradient Richardson number, $\mathrm{Ri_G}$. We approximate this number using a finite difference, yielding the bulk Richardson number, $\mathrm{Ri_B}$, which expresses the ratio between the temperature stratification and the wind shear:

$$\mathrm{Ri_B} = \frac{\frac{g}{\bar{\theta}_\nu}\Delta\theta_\nu\Delta z}{\Delta u^2 + \Delta v^2}\frac{g}{\bar{\theta}_\nu}\frac{\Delta\theta_\nu\Delta z}{\Delta v^2} \quad , \tag{5}$$

in which g is the gravitational acceleration, $\bar{\theta}_\nu$ is the mean virtual potential temperature, and $\Delta\theta_\nu$ ~~, Δu, and Δv~~ and Δv are the virtual potential temperature and the ~~two horizontal wind component differences~~horizontal wind speed difference, respectively, determined over the height difference Δz. Positive (negative) $\mathrm{Ri_B}$ values indicate stable (unstable) stratification

and values close to zero indicate neutral stratification. By assuming a functional form of the stability correction, L can be derived from $\mathrm{Ri_B}$ (Holtslag et al., 2014):

$$\frac{z}{L}\frac{\tilde{z}}{L} = \begin{cases} \frac{\mathrm{Ri_B}}{1-5\mathrm{Ri_B}}, & \text{if } \mathrm{Ri_B} \geq 0 \\ \mathrm{Ri_B}, & \text{otherwise} \end{cases}, \tag{6}$$

in which ~~z~~ $\tilde{z}$ is a reference height which is commonly taken as either the arithmetic or geometric mean of the heights used to determine the temperature and wind speed differences.

The wind direction can vary substantially with height ~~, also~~ in the lower atmosphere (e.g., Brown et al., 2005; Floors et al., 2015). ~~As they are scalar quantities,~~ A limitation of both the power law and logarithmic profile ~~wind speed relationships~~ is that they provide no information about any wind direction dependence with height. In addition, the relationships assume that wind speed increases monotonically with height. In practice, low-level maxima in wind speed, with decreasing wind speed above (low-level jets), are likely to occur, which is also observed in reanalysis data (e.g., Ranjha et al., 2013; Kalverla et al., 2019). To extend the validity of wind profile relationships to higher altitudes, several modifications have been proposed (e.g., Gryning et al., 2007; Holtslag et al., 2017). However, these theoretical formulations are only validated up to heights relevant for conventional, tower-based wind turbines.

Alternatively, computationally expensive ~~,~~ brute-force energy production calculations do not assume any wind profile relationship and are performed using historical wind data for the full operational height range. Bechtle et al. (2019) use ERA5 reanalysis data to map out the wind resource available to AWE systems over a large part of Europe, but do not touch upon the respective power production of an AWE system. Ranneberg et al. (2018) combine COSMO-DE reanalysis data with power curves for multiple heights, that are independent of the wind profile shape, to estimate the AEP. This is a valid approach if the system is operating at a nearly constant height. However, the wind profile shape has to be considered if the system operates in a larger height range, as is the case for a flexible-kite AWE system (Van der Vlugt et al., 2019). AEP calculations become more computationally expensive if the wind profile shape is considered, especially when identifying the optimal cycle settings for all time points. Malz et al. (2020a) use three months of three-hourly MERRA-2 reanalysis data and speed up the computation by a factor 20 by using the solution of the previous optimisation to initialise the next. In a follow-up study, Malz et al. (2020b) use this approach to determine the AEP of an AWE system for 16 locations in Europe. The current state-of-the-art is lacking a methodology that can be confidently used to make efficient AEP calculations for a pumping AWE system that sweeps a non-negligible height range.

Previously, clustering techniques have been used for identifying wind profile patterns. Sommerfeld et al. (2019) applies k-means clustering to subdivide stable and unstable wind profile datasets from lidar observations into two clusters for a location in a mostly flat area in northern Germany. Duran et al. (2019) use self-organising maps to characterise wind profile data for two locations (Cabauw in the centre of the Netherlands and the FINO-1 platform in the North Sea, 45 km north of the German/Dutch coast) from Weather Research and Forecasting modelled data using 2300 clusters. The clusters are used for forecast verification and to investigate diurnal and seasonal cycles.

This study proposes a clustering procedure for obtaining representative wind profile shapes from measured or modelled data that include the vertical variation of the wind speed and direction. The data is partitioned into a small number of clusters and the corresponding cluster-mean wind profile shapes are determined. We have chosen an empirical approach for identifying these shapes such that they are not restricted by physical assumptions. Nevertheless, we try to physically interpret the observed features. In contrast to earlier studies that use clustering, we investigate normalised wind profiles, as these are often described by wind profile relationships and yield a more compact wind resource representation. Moreover, the variation of the wind direction with height is included as it affects the operation of an AWE system.

The following sections of this paper outline the process of making an efficient AEP estimation for an AWE system based on historical wind data. Section 2 introduces the Dutch Offshore Wind Atlas (DOWA) and ERA5 datasets. Section 3 discusses the data processing and clustering techniques, complemented by interim results. Section 4 first addresses the clustering of DOWA data ~~, for which the results are presented~~ and presents the results for an on- and offshore location. Subsequently, the DOWA data of 45 other locations is clustered altogether to generate a generalised set of wind profile shapes that is applicable for an area which includes a wide range of location types. Although the resulting wind resource representation can be used for other applications, we illustrate its use for estimating the AEP of pumping AWE systems. Section 5 demonstrates the AEP estimation for a flexible-kite AWE system and assesses how many wind profile ~~shapes (clusters)~~ shape clusters are required for an accurate estimation. Finally, Sect. 6 summarises the conclusions of this study.

2 Wind datasets

In principle, ~~datasets~~ any dataset containing time series of wind speeds and directions for multiple altitudes can be used as input for the proposed methodology. For the AEP calculation, we focus on the sensitivity of the ~~power production of an AWE system~~ AWE system power production to the wind profile, which is assumed to be ~~steady-state~~ non-varying in the calculation. 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)~~ adverse effect on the power production, these effects ~~are typically~~ can be superimposed on a ~~steady-state~~ mean time-invariant wind profile using separate models for assessing, e.g., the associated loss in power production (Fechner, 2016). This power loss is device-specific and depends on control strategy, and is therefore not considered here. The first commercial AWE initiatives envisage a maximum operational height of 500 m because operation at higher altitudes requires more complex system designs (Watson et al., 2019, p. 4) and legislative procedures (Salma et al., 2018). For the wind resource representation for AWE, it is thus desirable to have wind data at least up to this height. The vertical resolution should be adequate to assess the shape of the wind profile with sufficient detail for the performance calculations. Both long-term lidar observations and modelled data qualify as input. This study focuses on using modelled data, which provides good spatial and temporal coverage.

~~A typical~~ An on- and offshore location in the Netherlands and the North Sea, respectively, are selected for demonstrating the methodology. The offshore location, that of the met mast IJmuiden, is located 85 km off the Dutch coast in the North Sea. The onshore location, namely, the met mast Cabauw, is located in the centre of the Netherlands. The area directly surrounding

the mast is flat open ~~grass-land~~ grassland for at least 400 m in all directions and up to 2 km in the dominant wind direction, i.e. west-south-west. ~~Within~~ Furthermore, within a radius of 20 km, the terrain is predominantly ~~grass-land and is~~ grassland and virtually flat. ~~Note that no anemometer or lidar measurements are analysed. We only analyse down-scaled reanalysis data for these sites. The sites were chosen as~~ The met mast sites, shown in Fig. 1, are selected because they are well-known in the literature. ~~Figure 1 shows where in the Netherlands they are situated~~We do not use the anemometer or lidar measurements of the met masts in this study. The other 45 depicted locations are used to evaluate the full DOWA domain and are selected such that onshore, coastal, and offshore locations are equally represented. The datasets for the met masts Cabauw and IJmuiden and the 45 locations are referred to as the onshore, offshore, and multi-location datasets, respectively.

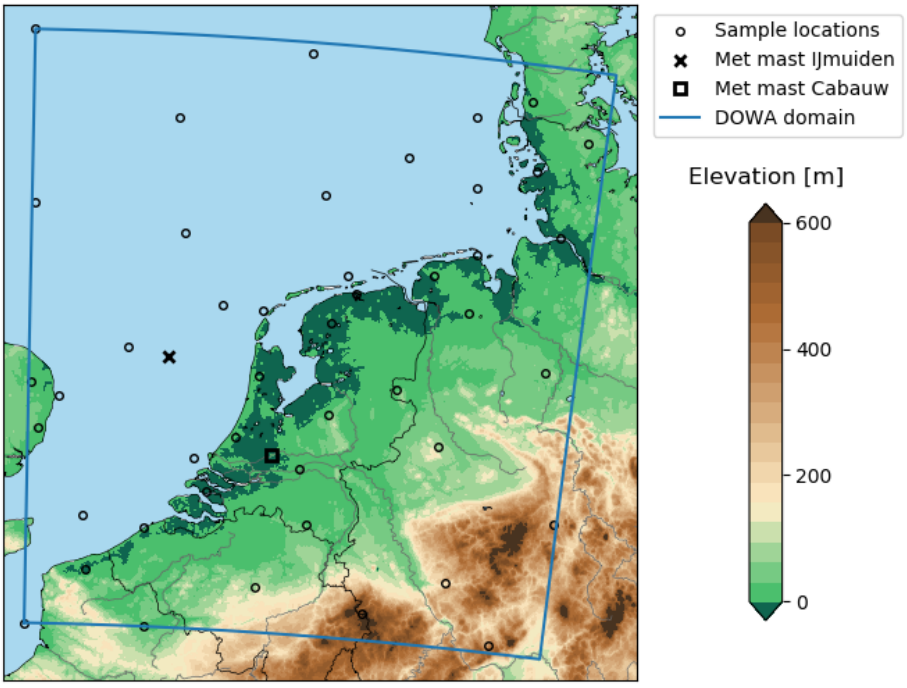


Figure 1. The DOWA domain, framed by the blue line, covers the Netherlands, a substantial part of the North Sea, and adjoining coastal areas. The ×, □, and ○ markers depict the locations analysed in Sect. 4.1, 4.3, and 4.4, respectively. The sea is depicted in light blue and the colour scale shows the elevation of the land surface (Amante and Eakins, 2009).

2.1 ERA5

ERA5 ~~(Copernicus Climate Change Service (C3S), 2017)~~ (Copernicus Climate Change Service, 2017) is a global reanalysis produced by ~~ECMWF using their integrated forecasting system, namely ECMWF's~~ the European Centre for Medium-Range Weather Forecasts (ECMWF) using their atmospheric model and data assimilation system. At the time of writing, ERA5 data is available from 1979 to the present time. The data includes hourly modelled values of a large number of atmospheric vari-

ables on a 30 km horizontal grid with 137 vertical pressure levels up to a height of roughly 80 km. The level heights are time dependent and interpolation is needed to obtain the wind data for fixed heights as required by the presented methodology. The clustering of the wind ~~profiles is performed using~~ profile shapes is performed on the DOWA data, which is described in the following section. ~~As explained later on,~~ ERA5 is only used to determine the atmospheric stability at the time and location of the analysed wind profiles.

2.2 Dutch Offshore Wind Atlas

~~The~~ DOWA (Wijnant et al., 2019) is ~~based on the~~ produced by the Royal Netherlands Meteorological Institute (KNMI) by downscaling ERA5 ~~reanalysis and downscaled~~ data to a finer-resolution surface grid using ~~the~~ their mesoscale weather model HARMONIE-AROME (Bengtsson et al., 2017). The downscaled reanalysis is performed for 10 years from 2008 until 2017. Hourly values for temperature, wind speed and direction, pressure and relative humidity are made available on ~~2.5 grid (~~a 217 x 234 ~~points)~~ grid with 2.5 km spacing and 17 heights between 10 and 600 m. The DOWA domain is illustrated in Fig. 1. Due to the higher resolution and non-hydrostatic nature of HARMONIE-AROME, DOWA benefits from an improved representation of the coastline, land surface heterogeneity, and mesoscale circulations, such as the sea breeze. Furthermore, additional observations from ~~routine~~ the KNMI's network of automated weather stations, satellite retrievals (ASCAT), and aircraft sensors (MODE-S EHS) have been assimilated by the HARMONIE-AROME model. Kalverla (2019) shows that DOWA improves on ERA5 in terms of wind speed, wind shear, and directional accuracy, as well as the representation of anomalous events such as low-level jets.

3 Clustering procedure

This section illustrates the clustering procedure for the offshore location. The data is filtered and normalised and its dimensions are reduced using a principal component (PC) analysis. Next, the clustering performance is analysed and the number of clusters is chosen for the wind resource representations analysed in Sec. 4.

3.1 Preprocessing of the wind data

The operation of an AWE system is affected by the variation of ~~the~~ wind speed and direction with height. Therefore, ~~the~~ wind profile shapes are studied with both these features included. Each wind profile sample consists of ~~wind velocities~~ easterly and northerly wind speed components for multiple heights (vertical grid points) at a given time and location and is processed ~~using its own properties~~ in two steps to obtain its shape. ~~The wind velocities are originally expressed by their easterly and northerly wind speed components.~~ Firstly, similar to Kalverla et al. (2017) and Malz et al. (2020a), the wind speed components are expressed as ~~components~~ parallel and perpendicular ~~to the reference~~ components relative to the wind velocity at a reference height, which we have chosen to be 100 m~~for every sample, which makes the wind profiles independent of their reference wind directions~~. As a result, the value for the perpendicular wind speed ~~profiles are zero~~ at 100 m is zero and the reformatted wind profile is independent of the wind direction at 100 m. Secondly, the wind speed components are normalised using the

90th percentile of the sample's wind velocity magnitudes~~is used to normalise the wind speed components~~. Using the percentile makes the normalisation less sensitive to outliers than using the maximum value. The normalised parallel and perpendicular wind speeds together form the *wind profile shape* of a sample~~is referred to as the wind profile *shape* throughout this paper. 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~~. However~~, however, it is prone to producing ~~eccentric~~ irregular wind profile shapes for low winds. Therefore, the wind profiles that have a mean wind speed below 5 $\mathrm{m\,s^{-1}}$ are filtered out before clustering. Note that the low wind conditions have a small contribution to the AEP of a wind energy system and their wind profile shapes are thus of small importance for the AEP calculation. Although the results presented in Sects. 3 and 4 do not account for the the low wind samples, the ~~wind resource representation for the~~ AEP calculation in Sect. 5 does.

3.2 Principal component analysis of the wind profile shape dataset

~~Prior to clustering, a PC analysis is used to reduce the dimensionality of the dataset, while preserving most of the variance. This reduces the computational effort and thus speeds up the clustering. The PC analysis specifies a transformation from the original to the PC coordinate system with its origin at the mean of the dataset. The first axis is oriented such that it accounts for most of the variance in the data. Subsequent axes are perpendicular to their predecessors and oriented such that they account for as much of the variance as possible. As a result, the last axis accounts for least of the variance. The PCs are unit vectors in the direction of the positive PC axes.~~

The mean wind profile shape for the offshore location is ~~shown~~ illustrated in the upper left panel of Fig. 2 by plotting the normalised wind speed $\tilde{v}$ against height using profiles for the parallel and perpendicular velocity components. As expected for an offshore location, the mean shape exhibits low wind shear. ~~Moreover, in~~ The hodograph, in the lower left panel of Fig. 2, shows how the normalised wind velocity changes with height by plotting the parallel and perpendicular normalised wind speed ($\tilde{v}_\parallel$ and $\tilde{v}_\perp$) for every height. In accordance with Ekman theory, the mean shape shows wind veer (wind direction turns clockwise with height)~~, which can be observed in the hodograph (top-view) in the lower left panel of Fig. 2~~. A logarithmic profile with roughness length z_0=0.0002 m, a representative value for open water (Wijnant et al., 2015), is fitted to the ~~lowest~~ lower 200 m of the mean shape. We use ~~this layer~~ 200 m as a proxy for the ~~'surface layer'~~ top of the surface layer, though in very stable situations the surface layer could be considerably smaller. Consequently, we consider applying the logarithmic profile relationship up to 200 m to be valid. Following the approach recommended by Kelly and Gryning (2010), the profile is fitted by varying the friction velocity ~~u_*~~ v_* and the stability function Ψ, which we constrain to the functional forms given in Eqs. 3 and 4. From this, a mean value of the Obukhov length L can be inferred. The best fit profile corresponds to a value L=-3391 m, implying a neutral logarithmic profile in the surface layer (assuming neutral conditions if $|L| > 500$). Above 200 m, the fit slightly deviates from the mean profile.

~~The first two PC unit vectors are shown~~ Prior to clustering, a PC analysis is used to reduce the dimensionality of the dataset, while preserving most of the variance. This reduces the computational effort and thus speeds up the clustering. The PC analysis specifies a transformation from the original to the PC coordinate system with its origin coinciding with the mean of the dataset.

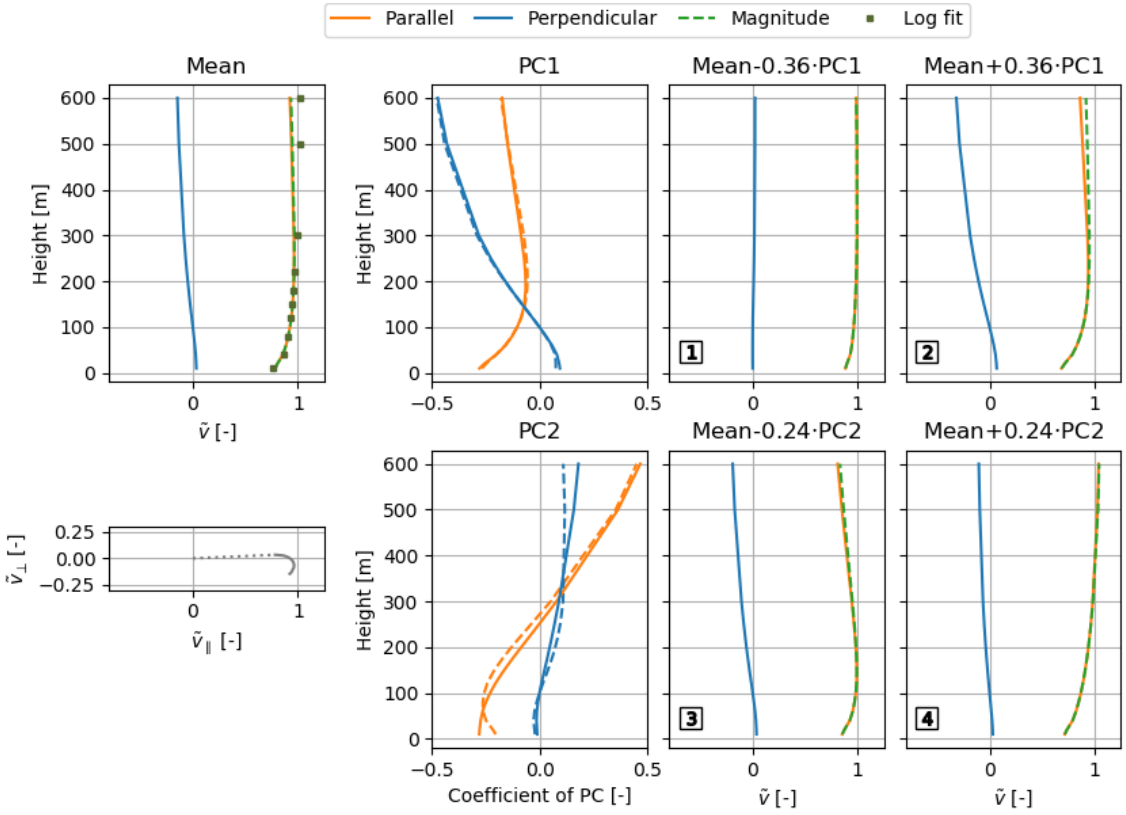


Figure 2. Mean wind profile shape and corresponding non-adiabatic logarithmic profile fit (top-left) and corresponding hodograph (bottom-left), composition of the first and second PCs (second column), and PC multiplicands superimposed on the mean wind profile shape (using minus and plus one standard deviation as multipliers for the third and fourth columns, respectively) for the filtered offshore dataset. The wind profile shape numbers 1–4 refer to the markers in Fig. 4a. The average of the PC profiles ~~of~~ from the two reference locations ~~are~~ is plotted alongside the offshore ~~PCs~~ PC profile using the dashed ~~lines~~ line (second column).

The first axis is oriented such that it accounts for most of the variance in the data. Subsequent axes are perpendicular to their predecessors and oriented such that they account for as much of the variance as possible. As a result, the last axis accounts for least of the variance. The PCs are unit vectors in the direction of the positive PC axes.

The compositions of the first two 34-dimensional PCs of the offshore dataset are illustrated in the second column of Fig. 2~~,~~ ~~which illustrates how much each~~ . The coefficients of each PC describe the relation between the PC and the parallel and perpendicular normalised wind speed ~~component contributes to the unit vectors~~components at the 17 heights. The absolute values of the PC coefficients quantify the contribution of the respective normalised wind speed components to the PC. The contributions of the perpendicular ~~wind speed~~ components account for most of PC1, indicating that PC1 ~~characterises mostly~~ mostly characterises wind veer. In contrast, the contributions of the parallel ~~wind speed~~ components account for most of PC2, indicating that PC2 ~~characterises mostly~~ mostly characterises wind shear. Both PCs show large contributions at both ends of the

height range, which indicates that most variance in the dataset is found at these heights. ~~The features of the dataset are no longer (normalised) physical properties, but contain contributions of all the normalised wind speed components.~~ In the PC-space, the data is expressed by multiplicands of the PCs superimposed on the mean wind profile shape. The third and fourth columns of Fig. 2 illustrate how to physically interpret the PCs by depicting two variations of the wind profile ~~shapes vary with respect to the mean~~ shape along PC1 and PC2 ~~. The~~ using minus and plus one standard deviation as multipliers. This means that 68 % of the PC1 (PC2) values lie between the values used for generating wind profile shape 1 and 2 (3 and 4). Indeed, the wind veer differs substantially between wind profile shape 1 and 2 and the wind shear between 3 and 4.

The percentage of variance retained after dimensionality reduction depends on how many PCs are used to express the data~~and is calculated by the sum of the eigenvalues of the retained PCs divided by the sum of the eigenvalues of all components.Figure 3shows that the~~ . The relation between the percentage of variance retained and the number of PCs follows from the PC analysis and is shown in Fig. 3. The first four PCs already account for more than 90 % of the variance in the offshore dataset. Since the wind velocities of neighbouring vertical grid points are highly correlated, most of the variance in the data is retained using a limited number of PCs. We consider retaining 90 % or more acceptable for our application. Since the variance retained still increases a few percent between four and five PCs, we opt for using five PCs. The preprocessed data is mapped onto the PC1–5-space and used as input for the clustering.

Figure 4a shows the frequency distribution of the wind profile shapes in the PC1, PC2-space. The PC1, PC2-projection of the wind profile shapes in the third and fourth columns of Fig. 2 are indicated with the markers. By visual inspection, two relatively dense groups of data points are identified: a confined group and a less confined group which resembles a tail extending from the first group, marked with the left and right ellipses, respectively. Figure 4b shows results for the onshore location and will be discussed in Sect. 4.3.

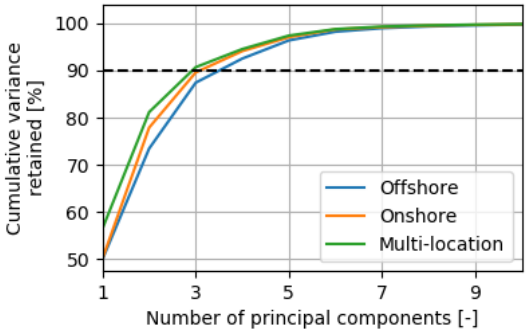
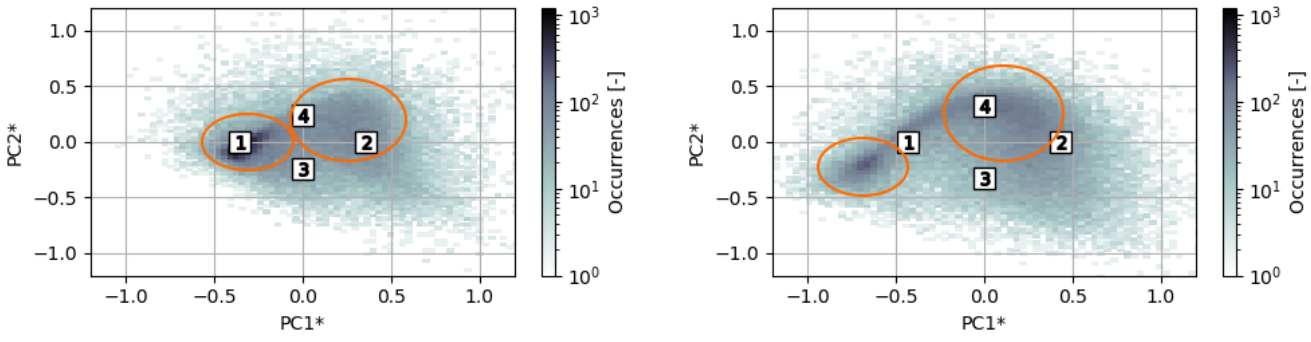


Figure 3. Relationship between the percentage of variance retained and the number of PCs for the filtered offshore, onshore, and multi-location datasets analysed in Sect. 4.1, 4.3, and 4.4, respectively.



(a) (b)

Figure 4. Sample frequency distributions in the PC1, PC2-space for the offshore (a) and onshore (b) locations. The origins coincide with the mean wind profile shapes and the markers with the wind profile shapes numbered 1–4 in Figs. 2 and 11~~are situated exactly one standard deviation away from the means~~. The orange ellipses indicate the visually identified clusters. The ~~coordinate system represents~~ x-axis (y-axis) is aligned with the average ~~PC profiles~~ of the PC1 (PC2) unit vectors from the two reference locations, which are denoted by an asterisk and their profiles shown in Figs. 2 and 11.

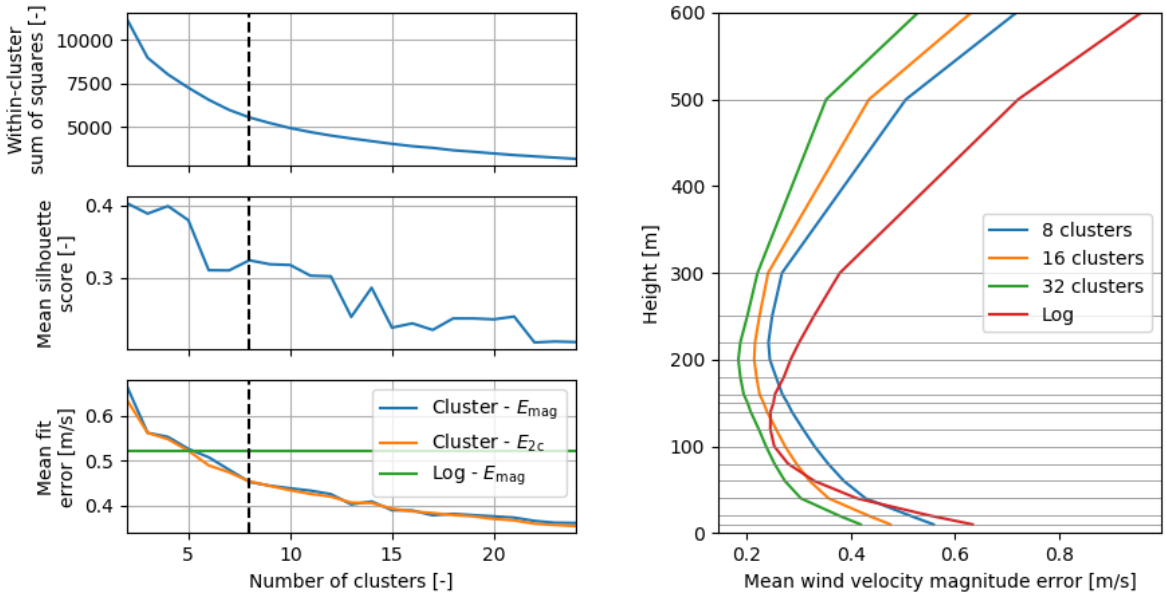
3.3 Choosing the number of clusters

K-means clustering (Pedregosa et al., 2011) is ~~employed~~ applied to identify the set of wind profile shapes that are used for representing the wind resource. Each cluster is represented by ~~a~~ its centroid and each sample is assigned to the cluster with the nearest centroid. The clustering algorithm iteratively searches for the positions of the centroids that minimise the sum of the ~~(squared)~~ squared Euclidean distances between the centroids and their associated samples. This cost function is also referred to as the within-cluster sum of squares (WCSS). 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.

K-means clustering is always able to produce a result, which makes it very powerful but also potentially deceptive. The algorithm tends to produce spherical clusters with equal radius and sample size, and works best on data with such a structure. The previous visual analysis of Fig. 4 revealed a different structure type for the wind profile shape datasets with two unevenly sized groups of data points. ~~Moreover, the result is highly sensitive to the choice of the~~ The number of clusters k generated by the algorithm needs to be specified by the user and it is often not evident how many clusters to choose. The elbow and silhouette method are used for finding an appropriate number for k. Moreover, the choice for k is evaluated in the context of ~~employing~~ applying the cluster-mean wind profile shapes to represent the wind resource.

The elbow method investigates the trend of WCSS against k. Increasing the number of clusters is equivalent to reducing the WCSS. Kinks in the trend indicate appropriate choices for k. The elbow plot in the upper left panel of Fig. 5a shows no distinct kinks for more than three clusters.

The silhouette score expresses the similarity of a sample to the other samples in its cluster ~~with respect~~ relative to its similarity to the nearest neighbouring cluster's samples. The dimensionless score ranges from -1 to 1: a negative value suggests



(a) Cost function of the clustering algorithm (top), cluster cohesiveness metric (middle), and the mean wind speed fit error (bottom) against the number of clusters. The dashed vertical lines depict the final choice for eight clusters.

(b) Mean wind velocity magnitude error with height for the cluster and logarithmic representations.

Figure 5. Sensitivity of the k-means clustering performance to the number of clusters over the full vertical grid (a) and for each height ~~point~~ (b) for the filtered offshore dataset.

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. The middle left panel of Fig. 5a shows the mean silhouette score is highest for two clusters. The division of the dataset into two clusters thus yields the most cohesive clusters, which is in agreement with the visual inspection of Fig. 4a. The decreasing trend of silhouette score with k implies that, in general, a small number of clusters should be used to maintain cluster cohesiveness.

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~~ absolute vertical wind speed profile is approximated by ~~de-normalising the~~ scaling the associated cluster-mean wind profile shape ~~of the cluster to which it is assigned~~ using the normalisation wind speed ~~of~~ used in the pre-processing. We assess the accuracy of ~~the~~ this cluster representation using the mean ~~sample fit error~~ fit error over all filtered samples. The fit error of ~~an individual sample is assessed by the magnitude and two-component~~ the j^{th} sample is calculated by the root mean square ~~fit errors over all the 17 DOWA vertical grid points. The representation accuracy is evaluated by the mean of all filtered samples for the two sample fit errorexpressions~~of the errors at each vertical grid point. Two different expressions are used to evaluate the error at the i^{th} vertical grid point: the wind velocity magnitude error $\varepsilon_{i,j}$

and that which includes both the parallel and perpendicular wind speed errors $\varepsilon_{\parallel,i,j}$ and $\varepsilon_{\perp,i,j}$. The resulting magnitude and two-component forms of the mean fit error, E_{mag} and $E_{2\mathrm{c}}$, ~~respectively~~are given by:

$$E_{\mathrm{mag}} = \frac{1}{n_{\mathrm{s}}} \sum_{j=1}^{n_{\mathrm{s}}} \left(\sqrt{\frac{1}{n_{\mathrm{h}}} \sum_{i=1}^{n_{\mathrm{h}}} \varepsilon_{i,j}^2} \right) \tag{7}$$

and

$$E_{2\mathrm{c}} = \frac{1}{n_{\mathrm{s}}} \sum_{j=1}^{n_{\mathrm{s}}} \left(\sqrt{\frac{1}{2n_{\mathrm{h}}} \sum_{i=1}^{n_{\mathrm{h}}} \left({\varepsilon_{\parallel,i,j}}^2 + {\varepsilon_{\perp,i,j}}^2 \right)} \right) \quad , \tag{8}$$

in which n_h is the number of heights, n_s is the number of samples~~, $\varepsilon_{i,j}$ is the wind velocity magnitude error, and $\varepsilon_{\parallel,i,j}$ and $\varepsilon_{\perp,i,j}$ are the parallel and perpendicular wind speed errors, respectively, at each vertical grid point i for the j^{th} sample~~. The relation between both mean fit errors and the number of clusters is shown in the lower panel of Fig. 5a.

~~Alternatively, logarithmic profiles can be used~~ We consider the use of the cluster representation valid when it yields a higher accuracy than a representation that uses logarithmic profiles to approximate the vertical ~~wind variations of the samples in a wind resource representation. This logarithmic~~ variation of the horizontal wind speed. The logarithmic wind resource representation is obtained by fitting logarithmic profiles with roughness length z_0=0.0002 m to ~~every~~ each sample. Here, the ~~value of~~ Obukhov length L passed to the Ψ stability function is restricted to ~~one of the values corresponding to~~ the representative values of the five stability classes, ~~which follow from the Obukhov lengths L~~ listed in the third column of Table 1. Moreover, the fit is performed to the full height range, i.e. 10–600 m, as we aim to minimise the fit error of the wind resource representation and, therefore, allow applying the logarithmic profile relationship beyond the surface layer. As the logarithmic representation does not include information about the wind direction variation with height, ~~it~~ its accuracy is only assessed using E_{mag}. We evaluate the fit error of the cluster representation in relation to the number of clusters and compare it to the fit error of the logarithmic representation. The lower panel of Fig. 5a shows that whether the fit error of the cluster representation is evaluated using E_{mag} or $E_{2\mathrm{c}}$ makes little difference. The cluster representation is more accurate than the logarithmic representation when using three clusters or more.

~~We consider the usage of the cluster representation valid when it yields a higher accuracy than the logarithmic representation. To this end, we evaluate the fit error of the cluster representation in relation to the number of clusters and compare it to the fit error of the logarithmic representation. Figure 5a shows that whether the error of the cluster representation is evaluated using E_{mag} or $E_{2\mathrm{c}}$ makes little difference. Note that the WCSS scales quadratically with the mean fit error of the cluster representation. The cluster representation is more accurate than the logarithmic representation when using three clusters or more.~~

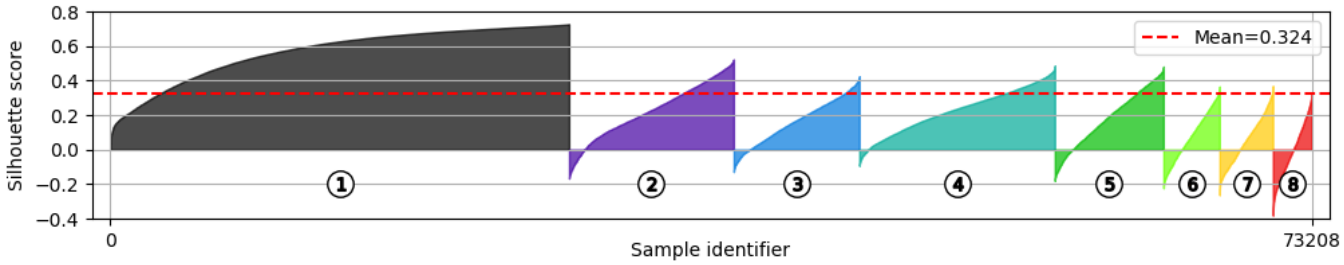
~~The fit error is also evaluated for all vertical grid points separately to evaluate their contributions to the mean fit error. The mean value of the~~ The wind resource representations do not yield the same accuracy for each vertical grid point. To investigate the height dependency, the mean wind velocity magnitude ~~errors of~~ error over all filtered samples ~~($\bar{\varepsilon}_i$) is shown~~ is calculated

Table 1. Stability classes in terms of the Obukhov length, adapted from Holtslag et al. (2014).

Class name	Class boundaries [m]	Representative L [m]
Very unstable (VU)	$-200 \leq L < 0$	-100
Unstable (U)	$-500 \leq L < -200$	-350
Neutral (N)	$\|L\| > 500$	10^{10}
Stable (S)	$200 < L \leq 500$	350
Very stable (VS)	$0 < L \leq 200$	100

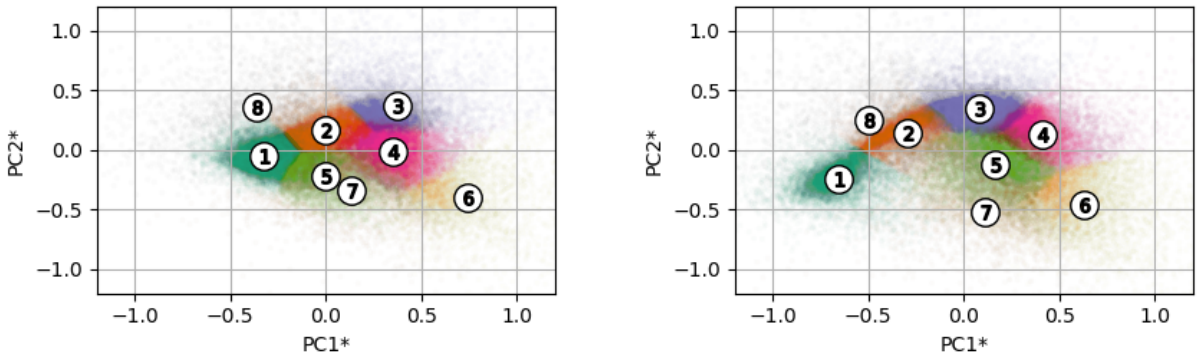
for each vertical grid point. The results are shown in Fig. 5b, in which the horizontal lines depict the 17 heights of the vertical grid points of DOWA. The fits have a relatively low ~~mean wind speed~~ error around 150 m height and a higher error at the top and bottom ~~. The horizontal lines depict the 17 heights of the DOWA~~ of the vertical grid. Around 150 m height, the grid is relatively fine, which is equivalent to allocating more weight to the 100–200 m interval for the logarithmic profile fitting procedure. As a result, the fitting favours minimising the errors in this interval over those at both ends of the height range. Note that the sensitivity of the cluster representation to the grid spacing is limited by the PC analysis prior to the fitting. As stated before, 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. Although the error of the cluster representations at 100 m is higher than that of the logarithmic representation, on average they perform substantially better.

The ~~analysis in Sect. 4 requires a limited~~ choice for the number of clusters used to represent the wind resource ~~. The choice for eight clusters~~ depends on the type of analysis. Eight clusters are chosen for investigating their characteristics in Sect. 4. This choice follows from a trade-off between the mean wind profile fit error, the silhouette score, representation validity, and our aim to present a ~~meaningful analysis and~~ concise analysis and meaningful interpretation of the resulting clusters. To get more insight in the structures of the eight offshore clusters (MMIJ-1–8), the mean silhouette score is calculated for each cluster. The higher the mean silhouette score, the more likely that a cluster is representing a natural structure in the data. Figure 6 shows that a large fraction of the samples have high silhouette scores for MMIJ-1–4, indicating that MMIJ-1–4 are relatively cohesive clusters. The silhouette score distributions of MMIJ-5–8 indicate less uniform sets of samples, especially that of MMIJ-8. Note that MMIJ-1 is roughly a factor 2.5 larger than the second biggest cluster despite the tendency of k-means clustering to produce equally sized clusters.



Cluster label	MMIJ-1	MMIJ-2	MMIJ-3	MMIJ-4	MMIJ-5	MMIJ-6	MMIJ-7	MMIJ-8	overall
Mean silhouette score	0.567	0.226	0.164	0.226	0.184	0.080	0.056	-0.009	0.324

Figure 6. The silhouette scores of the individual samples grouped by cluster and in ascending order for the filtered offshore dataset. The numbered markers and filled area colours indicate to which cluster the sample belongs. The overall mean score is indicated by the dashed line and the table below the figure states the mean score for each cluster.



(a) Offshore location

(b) Onshore location

Figure 7. Projection of the samples onto the PC1, PC2-space. The colour indicates to which cluster a sample belongs and the markers represent the cluster-centroids. The ~~coordinate system represents~~ x-axis (y-axis) is aligned with the average ~~PC profiles~~ of the PC1 (PC2) unit vectors from the two reference locations, which are denoted by an asterisk and their profiles shown in Figs. 2 and 11.

4 Cluster wind resource representation

This section discusses ~~the~~ our physical interpretation of the cluster representations. Firstly, the clusters and their cluster-mean wind profile shapes that result from the offshore dataset are presented. For each cluster, patterns in the times of occurrences are studied together with their association to wind properties at 100 m and atmospheric stability. The analysis is then repeated for the onshore location. ~~Comparing the~~ The cluster sets for both reference locations ~~sheds~~ are compared to shed some light on the ~~their general applicability~~similarities and differences between them. Finally, data from 45 locations are combined to obtain a single set of clusters that is applicable for the ~~full~~ entire DOWA domain. For each of the resulting clusters, a map is generated depicting the cluster frequency distribution over the DOWA domain.

4.1 Cluster representation for the offshore location

The clustering of the dataset for the offshore location ~~of~~ at the met mast IJmuiden yields eight clusters (MMIJ-1–8), which are represented by their centroids shown in ~~Figure~~Fig. 7a. The clusters are well spread over the PC1, PC2-space, with the exception of MMIJ-5 and 7, which are relatively close to each other. Note that only two axes of the five-dimensional PC-space are shown. Table 2 lists all five PC-coordinates of the cluster-centroids and confirms that the PC1 and PC2 coordinates of MMIJ-5,7 are similar, in contrast to their PC3 coordinates: the centroids are furthest apart along PC3. The PC4 and PC5 values have a substantially smaller range than that for PC1–3 and are superfluous for distinguishing between the eight clusters.

The cluster-mean wind profile shapes of the offshore clusters are shown in Fig. 8. Logarithmic profiles with roughness length z_0=0.0002 m are fitted to the magnitude profiles and shown for comparison. Here, the Obukhov length used in the stability function is varied ~~continuously~~ freely and the fit is restricted to the lower 200 m. The values ~~found for the corresponding Obukhov lengths are assigned to the stability classes in Table 1 and~~ for the Obukhov lengths inferred from the fits are plotted

Table 2. Principal component coordinates of the cluster-centroids for the filtered offshore dataset. The ~~centroids~~ centroid positions in the PC1, PC2-space are depicted in Fig. 7a ~~at their PC1, PC2-coordinates~~ with the numbered markers.

Cluster label	PC1	PC2	PC3	PC4	PC5
MMIJ-1	-0.33	-0.05	-0.04	-0.02	0.01
MMIJ-2	0	0.17	-0.08	0.05	0.02
MMIJ-3	0.38	0.38	-0.01	0.05	0
MMIJ-4	0.35	-0.02	0.04	-0.09	-0.06
MMIJ-5	0	-0.22	-0.16	0.07	-0.04
MMIJ-6	0.74	-0.4	0.02	0	0.12
MMIJ-7	0.14	-0.33	0.44	0.09	0.03
MMIJ-8	-0.36	0.36	0.45	0.04	0.02

as 500 m/L in Fig. 9 and categorised using the stability classes in Table 1. The comparison serves to show to what extent the cluster shapes deviate from non-adiabatic logarithmic profiles, particularly above the surface layer.

To investigate the characteristics of each cluster, Fig. 10a–c show how the clusters are distributed over the years, months, and hours of the day. ~~The upper panel~~ Figure 10a 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 ($\sim$20 years). ~~In contrast to the other panels, the absolute frequency is~~ The absolute frequency on the y-axis ~~and~~ serves to show ~~which part of the total dataset is represented by each of the clusters. Fig.~~ the cluster sizes. Figure 10d–f show the relative frequency of each cluster for different conditions in terms of wind speed, wind direction, and atmospheric stability. As for the logarithmic profile fits, the stability ~~for~~ of each sample is classified using Table 1.

~~Here~~For generating Fig. 10f, we derive the stability class distributions using the bulk Richardson number ~~,~~ $\mathrm{Ri_B}$ ~~,~~converted to the Obukhov length ~~,~~L ~~,~~using Eqs. 5 and 6. 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~~ $\bar{z}$ in order to convert $\mathrm{Ri_B}$ to L.

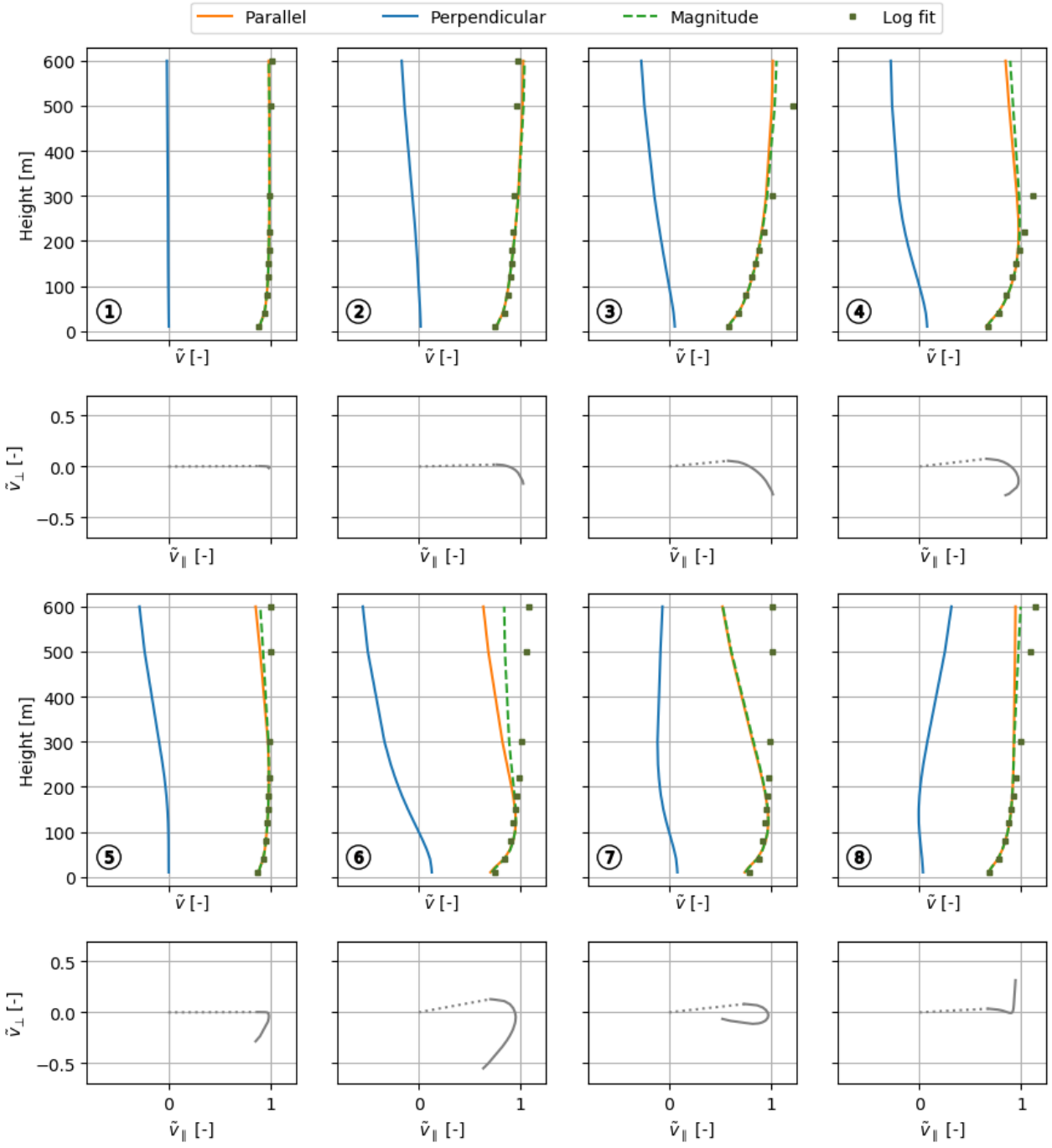


Figure 8. The eight cluster-mean wind profile shapes of the offshore clusters (MMIJ-1–8). Each shape is depicted by the normalised wind speed components with height (first and third rows) with the corresponding hodograph below (second and fourth rows). ~~Logarithmic~~ Non-adiabatic logarithmic profile fits are plotted alongside the shapes. In ~~the hodographs~~each hodograph, the ~~lowest points are connected to~~ lower end of the profile is indicated by the ~~origins with~~ dotted ~~lines and~~ line connecting the ~~highest points are~~ lowest height point to the ~~loose ends~~origin. All plots share the same x-axis.

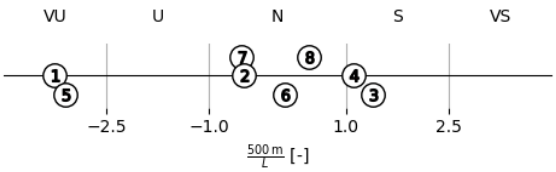


Figure 9. Obukhov lengths (plotted as $500\,\mathrm{m}/L$) found by fitting logarithmic profiles to the offshore cluster-mean wind profile shapes in Fig. 8. The stability classes are adopted from Table 1.

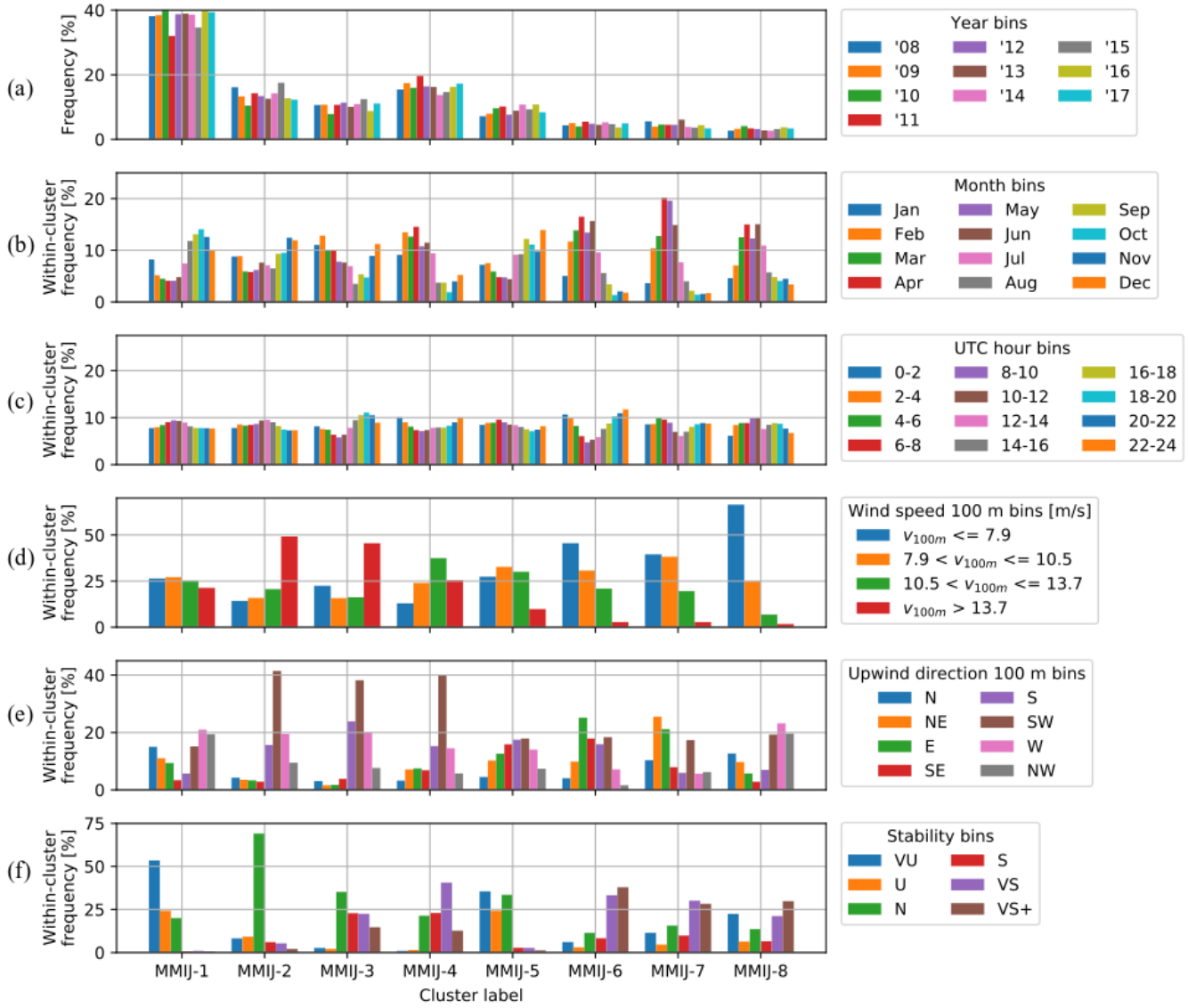


Figure 10. Frequency distributions broken down into bins by time of occurrence (a, b, and c), ~~reference~~ wind speed and direction at 100 m (d and e), and atmospheric stability (f) for the filtered offshore dataset. The wind speed bin limits are chosen such that the frequency over all clusters for each bin is roughly the same. The stability bins correspond to the classes in Table 1 together with the VS+ bin ($\mathrm{Ri_B} \geq 0.2$). The other distributions have equal bin widths.

4.2 Interpretation of the offshore cluster representation

By examining Figs. 8 and 9, we can see how the cluster-mean wind profile shapes differ from standard logarithmic profiles, particularly above the surface layer. Moreover, by referring to Fig. 10, it is possible to investigate the conditions under which each of the clusters occur, and to gain insight into their physical origins.

Figure 8 shows that the MMIJ-1 and 2 ~~profile shapes are described well~~ magnitude profiles are well-described with logarithmic profiles. The MMIJ-1 profile shape suggests a well-mixed convective ~~(unstable)~~ profile with little wind shear ~~or~~ and veer. MMIJ-1 occurs predominantly in autumn and is slightly more frequent in the morning hours. The wind is more frequently weak or moderate than strong and mostly coming from the westerly, north-westerly or northerly directions. Furthermore, this cluster occurs predominantly during ~~(very)~~ unstable conditions. These observations make sense as in autumn, the relatively warm sea water favours neutral to unstable stratification; the dominant wind directions have long fetches over sea which ~~allows~~ allow the boundary layer to ~~develop an equilibrium profile.~~reach an equilibrium state due to the relatively constant surface forcing.

The MMIJ-2 and 3 profile shapes ~~display increased levels of shear compared to that of~~ show an increase in wind shear relative to MMIJ-1. The MMIJ-2 profile shape closely resembles a neutral logarithmic profile up to 600 m, whereas that of MMIJ-3 only shows a good fit with a stable logarithmic profile in the surface layer. These clusters occur typically during strong winds, predominantly from the south-west. Strong south-westerly winds are characteristic of the wind climate at this mid-latitude location, which is dominated by the frequent passage of low-pressure systems. The relatively strong winds explain why we see the highest occurrence of near-neutral conditions, especially for MMIJ-2. For MMIJ-3, we also see frequent ~~(very)~~ stable conditions. Simultaneously, we observe that MMIJ-2 occurs more often in the late autumn and MMIJ-3 in winter and the start of the spring. The colder sea water in spring favours the formation of stable stratification, which explains the difference in stability distribution between the two clusters. Stable stratification suppresses turbulent mixing, which helps to sustain a strong wind shear, consistent with the increasing wind shear and veer seen in Figs. 7a and 8.

The MMIJ-4–7 profile shapes are all ~~(slightly)~~ jet-like~~with relatively large deviations from logarithmic profiles at greater altitudes~~. Because wind speed increases monotonically with height in the logarithmic wind profile relationship, it can not describe these type of profile shapes. The wind direction and stability distributions associated with the MMIJ-4 cluster are correlated with south-westerly winds and stable stratification. The seasonal cycle is very pronounced and peaks in spring, when stable stratification is frequent. The winds recorded for MMIJ-4 are mostly moderate to strong. The distributions associated with the MMIJ-5 cluster are very similar to those of MMIJ-1, with the exception of the wind direction distribution, which shows an opposite trend. The ~~frequent~~ winds with a southerly component are dominant for MMIJ-5 and typically have shorter fetches over sea than the north and westerly winds seen for MMIJ-1. The hodograph of the MMIJ-5 profile shape indicates a rather abrupt kink around 140 m, suggesting a discontinuity such as ~~an (internal)~~ a boundary-layer top. The MMIJ-6 profile shape shows a maximum at 120 m. Although the magnitude profiles of MMIJ-5 and 6 look somewhat similar, the MMIJ-6 profile shape veers more~~and gradually~~. The MMIJ-7 profile shape shows the most pronounced jet-like shape, also peaking around 120 m. MMIJ-6 and 7 occur almost exclusively for very stable conditions in spring and for weak wind situations. Both clusters occur predominantly for ~~(north)easterly winds~~ winds with an easterly component and show a diurnal cycle with fewer

occurrences around noon. Such a diurnal cycle is in agreement with various studies that have linked low-level jets and the diurnal variation of both the land-sea temperature difference and the intensity of turbulent mixing (e.g., Burk and Thompson, 1996; Parish, 2000; Mahrt et al., 2014; Shapiro et al., 2016).

The hodographs of MMIJ-5 and 8 both show ~~an abrupt kink~~ a sharp bend around 140 m. However, the wind direction turns anticlockwise with height above the ~~kink~~ bend for MMIJ-8, which is opposite to the veering of the other profile shapes. Despite the peculiarities of the wind direction profiles, the magnitude profiles of MMIJ-5 and 8 are described reasonably well below 200 m with very unstable and neutral logarithmic profiles, respectively. MMIJ-8 occurs mostly in spring, under stable conditions, and more often for winds with a westerly rather than a southerly component. Note that this shape belongs to an incohesive cluster and, therefore, gives a relatively poor representation of the cluster samples.

4.3 Comparing the on- and offshore cluster representations

The ~~clustering is repeated using the~~ dataset for the onshore location ~~of~~ at the met mast Cabauw is clustered using the same approach. The eight resulting clusters are referred to as MMC-1–8. The results of the PC analysis of the onshore dataset are shown in Fig. 11~~and compared~~, which we will compare to those of the offshore dataset, shown in Fig. 2. A logarithmic profile with roughness length z_0=0.1 m, a representative value for the area surrounding the mast (Verkaik, 2006), is fitted to the mean wind profile shape as before. With a stability function value corresponding to L=476 m, the mean profile shape ~~shows a good fit~~ below 200 m is in accordance with a stable logarithmic profile~~below 200~~. Above that, the fitted logarithmic profile rapidly diverges from the mean shape. A higher wind shear is observed than for the offshore location due to the higher surface roughness. The hodograph in the lower left panel shows that also the wind veer is substantially increased. Despite the apparent differences in mean shape, the PC1 and PC2 profiles are very similar for both reference locations. The average ~~PC profiles of the two locations are~~ of the PC profiles from the two reference locations is plotted alongside the onshore PC ~~profiles~~ profile using the dashed ~~lines. The coordinate system of~~ line. To enable a direct comparison between results, the same coordinate system is used for Figs. 4a and b~~represent the average PC profiles. Both figures share the same coordinate system, which enables a direct comparison between the~~. The x-axis (y-axis) is aligned with the average of the PC1 (PC2) unit vectors from the two reference locations.

The distribution in Fig. 4b shows a similar pattern to Fig. 4a: a dense, confined group of samples, marked with the left ellipse, with a tail of samples extending from this group at around 45 degrees towards the right ellipse. In general, the samples of the onshore dataset are more spread out than the offshore samples, particularly along the PC1 axis. Also the confined group is less dense for the onshore location. Figure 7 shows that, for both locations, the samples of these confined groups belong to the on- and offshore clusters with number 1. The remaining onshore clusters with monotonic wind speed and veering profiles, MMC-2–4, account for most of the tail, see Fig. 12. Note that the ~~cluster~~ clustering algorithm produces arbitrary labels for each ~~class~~cluster. We have manually renumbered them such that the ~~numbering is more or less aligned between onshore and offshore clusters~~onshore cluster numbers align with the offshore cluster numbers. This allows us to draw parallels between them and show that the resulting profiles are very similar between both locations, e.g., the first offshore clusters (MMIJ-1–3) also have monotonic profiles.

~~Again, logarithmic~~ Logarithmic profiles with roughness length z_0=0.1 m are fitted to the cluster-mean wind profile shapes, and plotted alongside them in Fig. 12. The values ~~found~~ for the Obukhov lengths ~~and the corresponding stability classes are presented~~ inferred from the fits are shown in Fig. 13 ~~.~~ and categorised by stability class. For the offshore location, Fig. 9 shows that six out of eight logarithmic profiles found by fitting are neutral or stable and those for MMIJ-1 and MMIJ-5 are more unstable. Figure 13 shows that only one unstable logarithmic profile is found for the onshore location, next to six stable and one neutral logarithmic profile. Since there is little diversity in the shape of the unstable profiles, all the associated samples are grouped together by the clustering. The fact that this type of profile is well-mixed with little shear and a relatively high boundary layer height explains why the diversity is small. By contrast, the neutral and stable profiles can have a wide range of shear and in addition, particularly in stable conditions, the boundary layer height can be quite low which will have a strong influence on wind shear. This means that a greater diversity of profile shapes is to be expected under neutral or stable conditions.

The profile shapes for MMC-1 to MMC-3 show an increase in wind shear. Between ~~MMIJ-3 and MMIJ-4~~MMC-3 and MMC-4, we see an increased wind veer, though reduced wind shear. The profile shapes for MMC-5–7 are ~~(slightly)~~ jet-like, as is the case for the offshore clusters MMIJ-4–7. MMC-5 and MMC-6 have similar wind velocity magnitude profiles with a relatively weak wind speed maximum around 200 m, but MMC-6 shows a much stronger wind veer. MMC-7 shows the strongest fall-off above 200 m. Like its offshore counterpart, MMC-8 is characterised by ~~anticlockwise turning and an abrupt kink in the wind profile~~ an anticlockwise-turning profile with a sharp bend, which is most clearly visible in the hodograph~~view~~. Recall that the offshore wind profile shape for MMIJ-5 also showed a ~~kink~~sharp bend, albeit, in combination with clockwise turning. We do not see these features for any of the MMC profile shapes.

Figure 14 shows that MMC-3 is the most frequent cluster in the ~~(filtered)~~ filtered onshore dataset, with a frequency of 20.6 %. ~~Where for the offshore location one distinct dominant cluster is identified, the frequency distribution over the~~ The first five onshore clusters ~~is more balanced~~have similar total frequencies, whereas MMIJ-1 dominates for the offshore location. As for the offshore clusters MMIJ-6–8, the onshore clusters MMC-6–8 are less frequent.

Figure 14 shows clusters that typically occur during spring/summer (MMC-1, MMC-7 and MMC-8) or autumn/winter (MMC-2–6). The diurnal cycles of the onshore location are highly pronounced in contrast to those for the offshore location. This effect is caused by the lower heat capacity of the land surface which promotes a more immediate heat transfer to or from the atmosphere. ~~In the presence of daytime solar irradiation , convection is created, and well-mixed profiles are expected, whereas at night, stable stratification is more frequent.~~ Convection created by solar irradiation leads to more turbulent mixing during the day than at night. Indeed, MMC-1 and MMC-2 show mixed profiles and predominantly occur during the day~~and~~ , whereas MMC-3–8 show profiles with less mixing and predominantly occur during the night. Note that low-level jets, and stable conditions in general, occur almost exclusively at night. ~~The patterns in the times of occurrences indicate~~ Figure 14c indicates a pronounced diurnal cycle in atmospheric stability for the onshore location, whereas for the offshore location the seasonal cycle, shown in Fig. 10b, is more pronounced. ~~Figs.~~Figures 10d and 14d ~~have almost identical bin distributions , but different bin widths, and thus different~~ display almost identical frequency distributions over the bins, however, the actual wind speed distributions differ due to the different bin limit values. Note that the ~~wind speed bin limits are chosen such that~~

~~the frequency over all clusters~~ chosen limits give the same total frequency for each bin~~is roughly the same, yielding different bin widths for the two reference locations~~, thereby the distributions of the individual clusters are easily related to the uniform general distribution and compared with one another. Also the wind direction ~~and stability~~ distributions show similar patterns for both locations. ~~Only the stability distribution of MMC-5 and MMIJ-5 differ substantially.~~ In the case of the stability distributions, the onshore location shows a tendency to more stable conditions for all clusters.

In conclusion, we see that very similar cluster-mean wind profile shapes have been identified for the on- and offshore reference locations. Moreover, similar profiles seem to be related to similar conditions in terms of wind speed, wind direction, and atmospheric stability. The strongest winds typically act to neutralise the stratification, leading to monotonic profiles with relatively little veer. These profiles are relatively well captured by logarithmic wind profiles. For weaker winds, atmospheric stability acts to enhance wind shear and veer, up to the point where low-level jets ~~start to be~~ are observed. However, whereas stability at the offshore location is governed by a clear seasonal cycle in the underlying sea surface, stability over land is regulated by the relatively rapid diurnal heating cycle of the land surface. Over sea, the wind direction also seems to play a more pronounced role, since it controls the characteristics of the prevailing fetch.

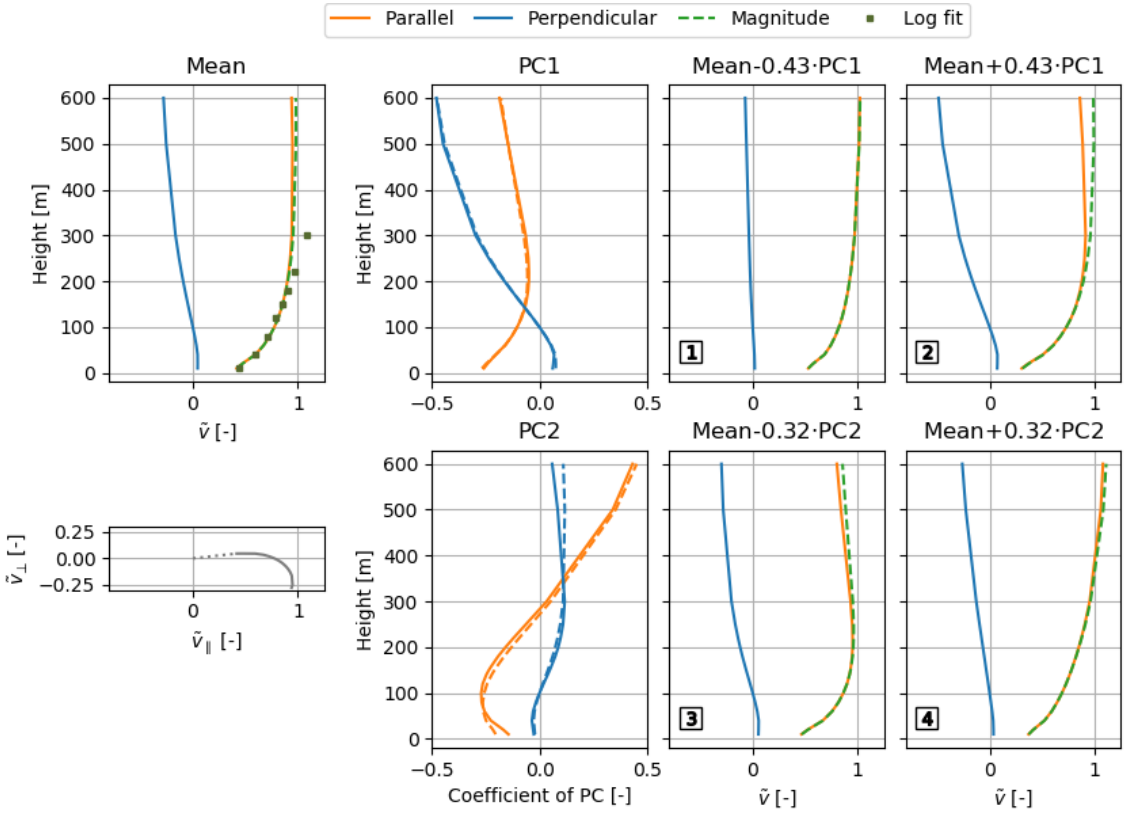


Figure 11. Mean wind profile shape and corresponding non-adiabatic logarithmic profile fit (top-left) and corresponding hodograph (bottom-left), composition of the first and second PCs (second column), and PC multiplicands superimposed on the mean wind profile shape (using minus and plus one standard deviation as multipliers for the third and fourth columns, respectively) for the filtered onshore dataset. The wind profile shape numbers 1–4 refer to the markers in Fig. 4b. The average of the PC profiles ~~of~~ from the two reference locations ~~are~~ is plotted alongside the onshore ~~PCs~~ PC profile using the dashed ~~lines~~ line (second column).

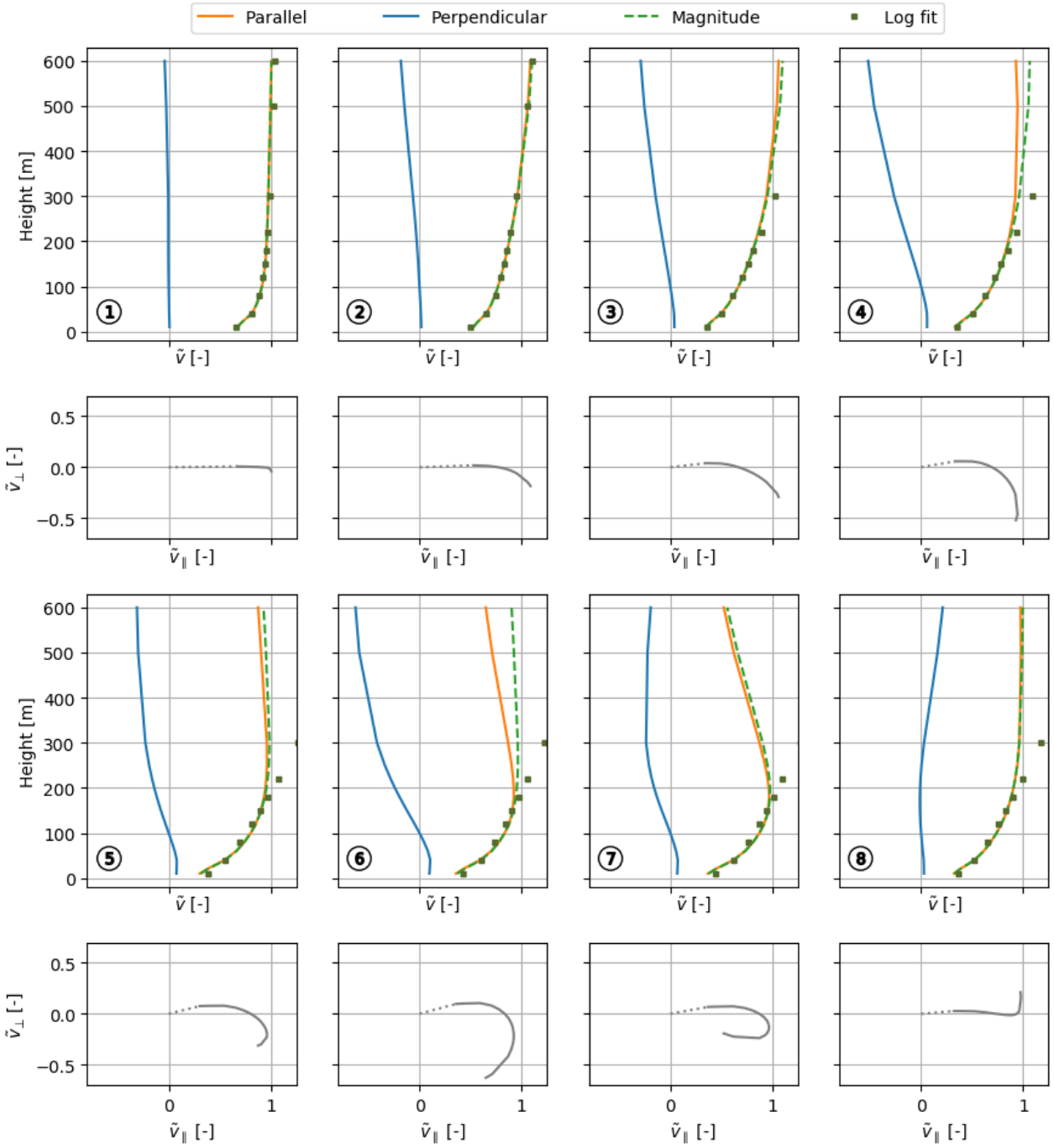


Figure 12. The eight cluster-mean wind profile shapes of the onshore clusters (MMC-1–8). Each shape is depicted by the normalised wind speed components with height (first and third rows) with the corresponding hodograph below (second and fourth rows). ~~Logarithmic~~ Non-adiabatic logarithmic profile fits are plotted alongside the shapes. In ~~the hodographs~~each hodograph, the ~~lowest points are connected to~~ lower end of the profile is indicated by the ~~origins with~~ dotted ~~lines and~~ line connecting the ~~highest points are~~ lowest height point to the ~~loose ends~~origin. All plots share the same x-axis.

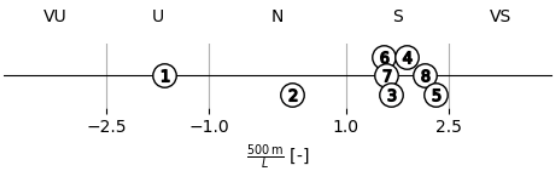


Figure 13. Obukhov lengths (plotted as $500\,\mathrm{m}/L$) found by fitting logarithmic profiles to the onshore cluster-mean wind profile shapes in Fig. 12. The stability classes are adopted from Table 1.

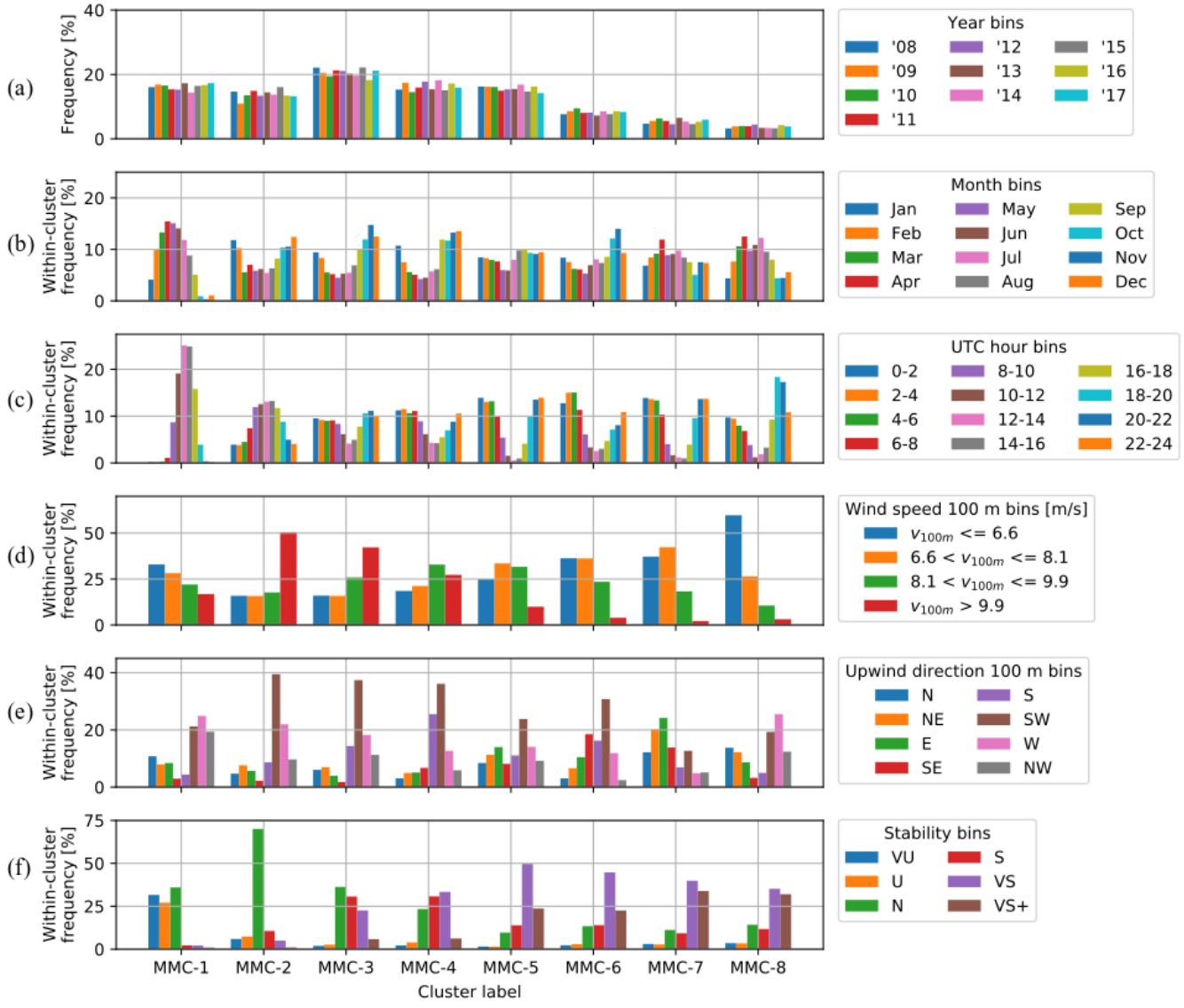


Figure 14. Frequency distributions broken down into bins ~~for~~ by time of occurrence (a, b, and c), ~~reference~~ wind speed and direction at 100 m (d and e), and atmospheric stability (f) for the filtered onshore dataset. The wind speed bin limits are chosen such that the frequency over all clusters for each bin is roughly the same. The stability bins correspond to the classes in Table 1 together with the VS+ bin ($\mathrm{Ri_B} \geq 0.2$). The other distributions have equal bin widths.

4.4 Spatial frequency distribution of wind profile shape clusters

So far, we have shown that similar clusters and, consequently, similar wind profile shapes ~~are~~ were identified for both onshore and offshore locations. Here, we apply our clustering algorithm on a dataset that includes wind data ~~for~~ from a variety of locations. 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). The sampled grid points are marked on the map in Fig. 1. Our aim is to give some insight into the spatial variability of wind profile characteristics, in particular to see how the clustering approach highlights profile characteristics of the on- and offshore environments. In principle, the multi-location approach gives a set of profile shapes that could be used for an AEP assessment, though a ~~site specific~~ site-specific set would be better suited if a more accurate assessment is required. ~~So, whilst~~ Whilst we increase the applicability of the cluster representation to a larger area by keeping the number of clusters the same, we compromise on the accuracy. As mentioned before, increasing the number of clusters ~~is inherent to reducing~~ reduces the error. The number of clusters can be increased until a suitable accuracy is attained. Here, we still use eight clusters, as it suffices to give an impression of the spatial variability of the wind profile shapes. The eight resulting multi-location clusters are referred to as ML-1–8.

Figure 15 shows the cluster-mean wind profile shapes for each of the multi-location clusters. Each sample of every grid point in the DOWA domain is assigned to the cluster with the closest centroid. For each cluster, a map is generated showing the spatial distribution of its frequency of occurrence, see Fig. 16. Note that the colour scale is different for each map ~~so~~ such that spatial patterns are easier to observe. Table 3 lists the frequency of each cluster at the on- and offshore reference locations. It is interesting to compare the multi-location clusters with the site-specific clusters identified earlier. With a frequency of 48.5 %, ML-1 is dominant at the met mast IJmuiden. Therefore, we expect it to be similar to MMIJ-1, the dominant cluster resulting from the offshore analysis. Comparing Figs. 8 and 15 indeed shows that the cluster-mean wind profile shapes of ML-1 and MMIJ-1 look alike. Similarly, ML-7 has the highest frequency at the met mast Cabauw, i.e. 21.7 %, and has a profile shape somewhere in between those of MMC-3 and 4, the most frequent clusters resulting from the onshore analysis. 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.

The maps in Fig. 16 show a distinct division between clusters that mostly occur over sea (ML-1–3) and over land (ML-4–8). The latter group is sub-divided into coastal and onshore clusters, see Table 3. The sharply defined patterns in the frequency maps of ML-5–8 coincide with ~~terrain~~ orographic features and thus suggest a strong relationship between the clusters and ~~terrain~~orography. Other site characteristics such as recurring weather systems and land cover also affect the clusters and thus the frequency maps. Over land the frequency maps of ML-5 and ML-6 suggest an inverse relationship: the frequency of ML-5 peaks at high elevations, whereas that of ML-6 is highest in the river valley in the lower right corner of the DOWA domain. A similar inverse relationship is observed between ML-7 and ML-8. Also, the frequency maps of ML-7 and ML-8 show contours

Table 3. Classification of the multiple-location clusters and frequencies of occurrence of the clusters at the on- and offshore reference locations (met masts Cabauw and IJmuiden).

Cluster label	Class	Similar single location cluster(s)	Frequency at offshore location	Frequency at onshore location
ML-1	offshore	MMIJ-1	48.5 %	5.8 %
ML-2	offshore	MMIJ-2, 3	22.0 %	4.2 %
ML-3	offshore	MMIJ-4, 6	14.1 %	4.2 %
ML-4	coastal	MMC-1	8.6 %	16.1 %
ML-5	onshore/coastal	MMC-2	3.0 %	17.4 %
ML-6	onshore	MMC-6	2.0 %	13.3 %
ML-7	onshore	MMC-3, 4	1.2 %	21.7 %
ML-8	onshore	MMC-5	0.6 %	17.3 %

coinciding with the elevation map, though the relationship between the frequency and elevation is not as direct as for ML-5 and ML-6.

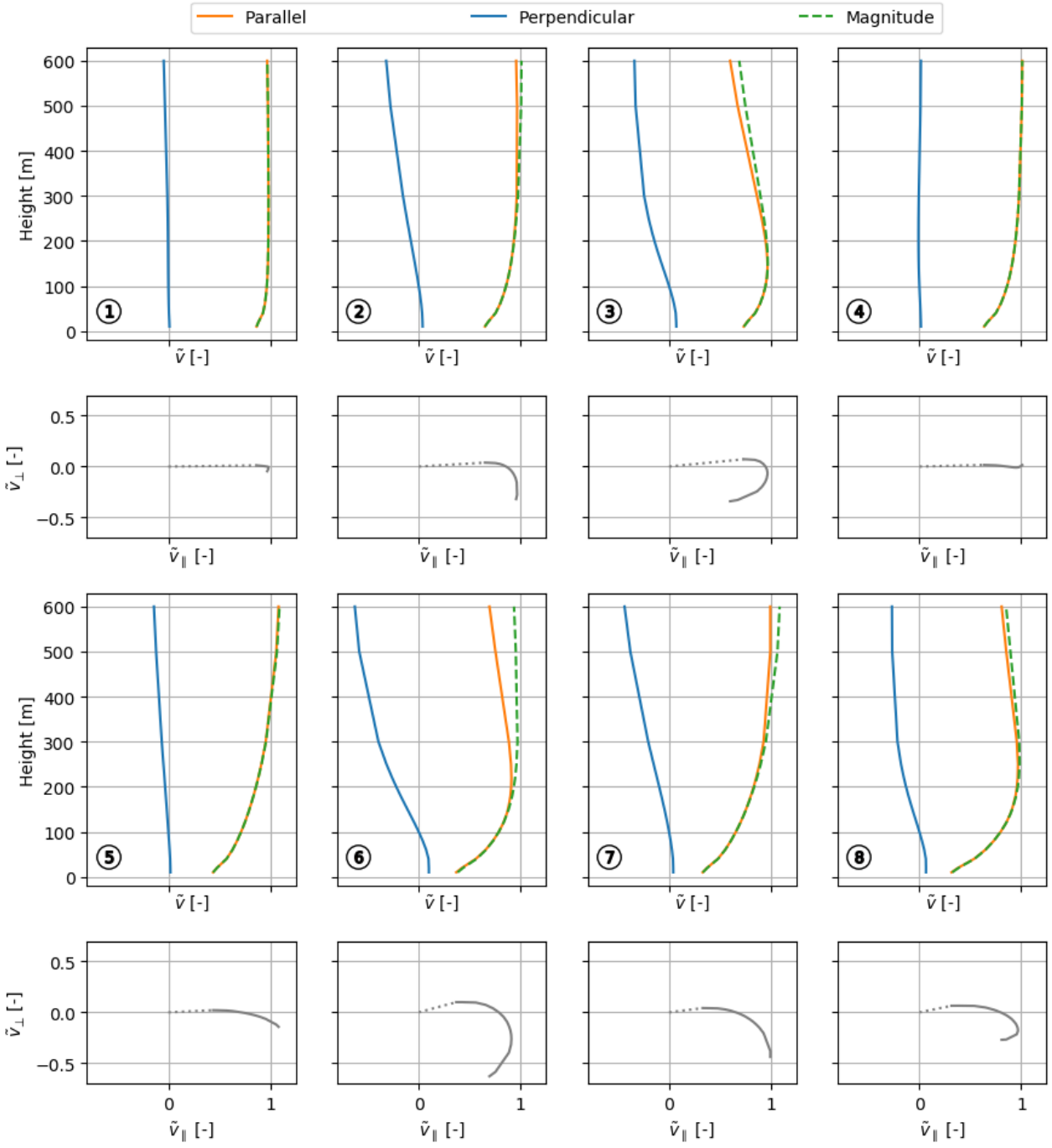


Figure 15. The eight cluster-mean wind profile shapes of the multi-location clusters (ML-1–8). Each shape is depicted by the normalised wind speed components with height (first and third rows) with the corresponding hodograph below (second and fourth rows). In ~~the hodographs~~each hodograph, the ~~lowest points are connected to~~ lower end of the profile is indicated by the ~~origins with~~ dotted ~~lines and~~ line connecting the ~~highest points are~~ lowest height point to the ~~loose ends~~origin. All plots share the same x-axis.

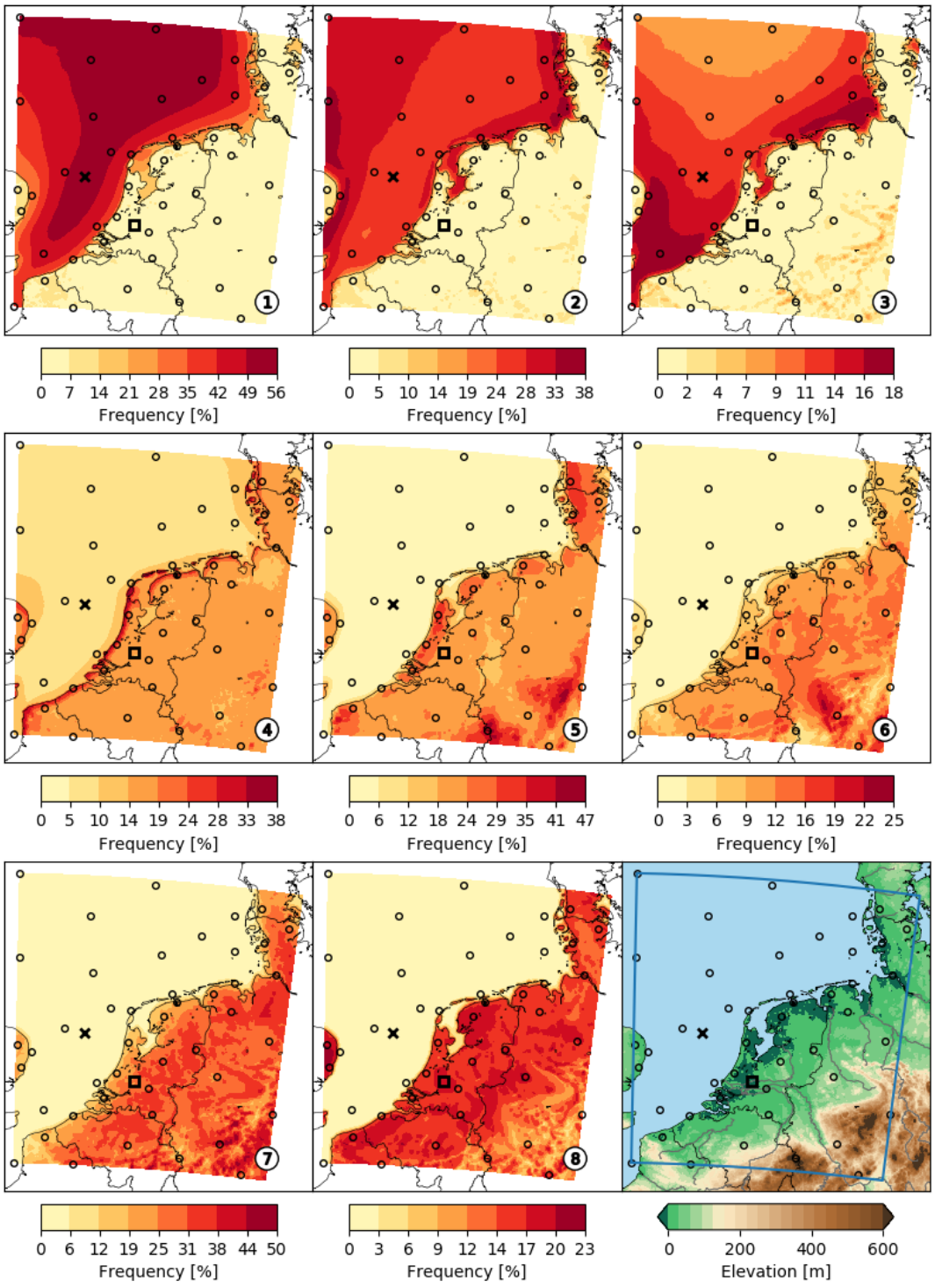


Figure 16. Frequency of occurrence of each multiple-location cluster (ML-1–8) mapped over the DOWA domain. The × and □ markers depict the reference locations of the met masts IJmuiden and Cabauw. The ○ markers show the sampled grid points. The lower right plot is a repetition of Fig. 1.

5 Efficient AWE production estimation using the cluster representation

With AWE technology maturing and approaching the deployment stage, the community is debating how to uniformly define the performance of AWE systems (Van Hussen et al., 2018). A generally applicable set of wind profile shapes is considered to be an important step to facilitate the standardisation of wind conditions for which AWE systems are rated in terms of power production. In this section, we demonstrate how the wind resource representations obtained using the clustering procedure can be used for estimating the AEP of a pumping AWE system. An advantage of AWE systems over tower-based wind turbines is that they have access to winds ~~higher up~~at higher altitudes. This advantage is limited when low-shear wind profiles are frequent at the installation site, as is the case offshore, but unusual for onshore locations. Deploying 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). A separate power curve is generated for each cluster using its cluster-mean wind profile shape. ~~The power curves~~ Each power curve together with the corresponding ~~wind speed distributions yield the AEP contributions of the clusters~~cluster-specific wind speed distribution yields the AEP contribution of the respective cluster. Finally, the sensitivity of the total AEP to the number of clusters is evaluated.

5.1 ~~Constructing the~~ Deriving power curves for a pumping AWE system

A pumping AWE system is alternating between reeling the tether in and out and thereby consuming and producing power, respectively. The relatively high lift force generated by the kite during reel-out and long reel-out phase yields a positive net energy output. The specific operational approach differs between AWE concepts and may require different performance models for calculating the generated power. We evaluate a flexible-kite system using the quasi-steady model (QSM) developed by Van der Vlugt et al. (2019) specifically for this concept. The cluster representation can in principle be used together with any performance model for estimating the AEP.

Flexible-kite systems typically sweep a large height range during the pumping cycle, which requires pronounced transitions between the reel-out and reel-in phases (Salma et al., 2019). Figure 17 shows the distinct phases of a pumping flexible-kite system. During the reel-out phase, the kite flies figure-of-eight manoeuvres in a fast cross-wind motion. After the reel-out phase, the kite stops flying cross-wind, de-powers, and flies towards zenith. Once reeled back in, the kite steers down, flies towards the starting position of the reel-out phase, and starts a next cycle. The QSM idealises and represents the pumping cycle using three phases: the reel-in, transition, and reel-out phase. The transition between the reel-out and reel-in phases is not modelled separately, but is included in the reel-in phase. The model does not resolve the cross-wind flight manoeuvres during the reel-out phase but represents them by an average cross-wind flight state with constant values for the elevation, azimuth, and course angle. The motion of the kite is approximated by moving it along the idealised flight path according to the computed steady-state kite speed.

The QSM assumes a steady wind field with a constant wind direction and only a vertical variation in the wind speed. Therefore, we consider only the magnitude profiles of the cluster-mean wind profile shapes in the calculations. The unidirectional

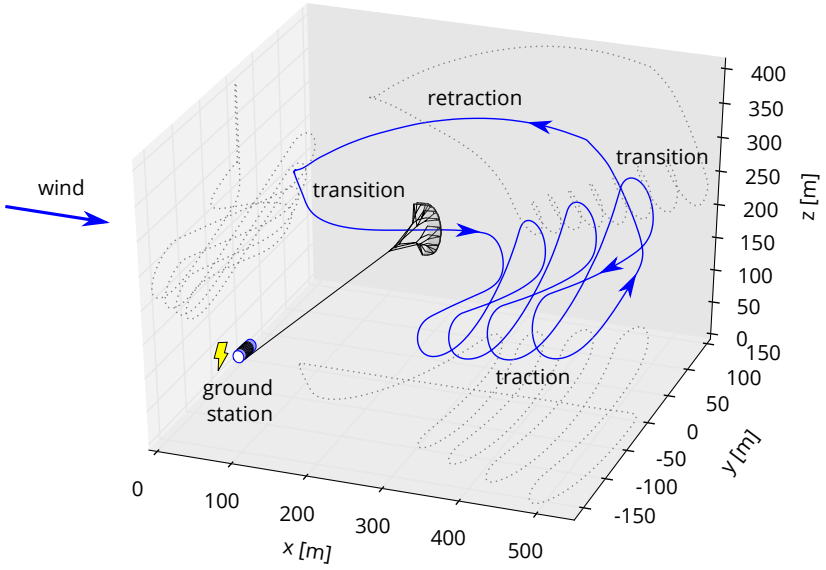


Figure 17. Flight path of the flexible-kite, pumping AWE system (kite & drum not to scale) adapted from Fechner (2016).

wind profile approximation is equivalent to hypothetically knowing the wind direction profile and steering the kite to correct for direction changes. We use the system properties of the 20 kW technology demonstrator of Delft University of Technology given in Table 4. 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).

Table 4. Constant system properties that are required as model input for the QSM. C_L and C_D stand for lift and drag coefficients, respectively. The kite and tether properties follow from the work of Oehler and Schmehl (2019). The other properties are chosen by judgment of the authors for being representative for the analysed system.

Kite properties		Tether properties		Operational limits		Representative reel-out state	
Projected area	19.75 m^2	Density	724 $\mathrm{kg\,m^{-3}}$	Min. reeling speed	2 $\mathrm{m\,s^{-1}}$	Azimuth angle	13°
Mass	22.8 kg	Diameter	4 mm	Max. reeling speed	10 $\mathrm{m\,s^{-1}}$	Course angle	100°
$C_{L,\mathrm{powered}}$	0.9	$C_{D,\mathrm{tether}}$	1.1	Min. tether force	300 N		
$C_{D,\mathrm{powered}}$	0.2			Max. tether force	5000 N		
$C_{L,\mathrm{depowered}}$	0.2						
$C_{D,\mathrm{depowered}}$	0.1						

~~Calculating the energy production of an AWE system requires characterising the (maximum)~~ The proposed AEP estimation requires the characterisation of the maximal mean cycle power for a large variety of wind conditions. The mean cycle power

depends on the operational settings that control the cycle trajectory and phase durations. These cycle settings include the forces applied to the tether during reel-in and reel-out. The values of the cycle settings are chosen such that they yield ~~maximum~~ maximal 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~~as long as the tether force does not exceed its limit. During reel-out, the tether force should yield a high ~~energy production, while letting the reel-out phase comprise most of the cycle duration~~power, while increasing the fraction of time spent producing energy in the cycle. 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 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 $C_{L,\text{powered}}$. 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~~. The minimum tether length is fixed at 200 m.

We use numerical optimisation to determine the cycle settings that maximises the mean cycle power. Table 5 lists the ~~optimisation variables and~~ cycle setting parameters, which are used as optimisation variables, together with their respective limits. Imposing a lower bound on the tether force ensures that the kite stays tensioned, as required for a flexible-kite. The upper bound corresponds to the maximum allowed tether force. The remaining limits are chosen by judgment of the authors. The optimisation uses the sequential quadratic programming algorithm (SLSQP) that is part of pyOpt (Perez et al., 2012). This class of algorithms is generally seen as a good general-purpose method for differentiable constrained non-linear problems. 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~~ the wind speed at any height can be used as a scaling parameter. We use the wind speed at 100 m ~~. By stepping through~~ as scaling parameter. A power curve is derived for each of the clusters, by determining the maximal mean cycle power for a range of wind speeds at 100 m between cut-in and cut-out~~, a power curve is constructed for each of the clusters. At~~ . In 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~~ de-normalised, followed by an optimisation using the resulting wind profile as input.

Table 5. ~~Optimisation variables used~~ Cycle setting parameters which are varied for maximising the mean cycle power and their corresponding limits defining the search space. The limits are chosen by judgment of the authors for being representative for the analysed system.

Parameter	Lower bound	Upper bound
Reel-out force	300 N	5000 N
Reel-in force	300 N	5000 N
Reel-out elevation angle	25°	60°
Pumping length tether	150 m	250 m

Prior to performing the optimisations, we determine the cut-in and cut-out wind speeds at 100 m for each wind profile shape. The cut-in limit is assumed to be the ~~smallest~~ lowest wind speed for which, along the ~~whole~~ entire reel-out path, feasible

steady flight states are found with the QSM. The cut-out limit is determined by the criterion that the pumping cycle should complete at least one figure-of-eight manoeuvre (at an elevation angle of 60 degrees). This criterion becomes more critical at high wind speeds, as the reel-out phase gets shorter. The QSM as presented by Van der Vlugt et al. (2019) does not resolve the cross-wind flight motion. 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.

Scaling each wind profile shape such that the wind speed at 100 m equals the previously determined cut-in and cut-out wind speeds yields the respective absolute wind profiles, shown in Fig. 18. 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. 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~~ well-defined at 80 m. At 100 m, MMC-3 and MMC-7 show the lowest and highest cut-out wind speed, respectively.

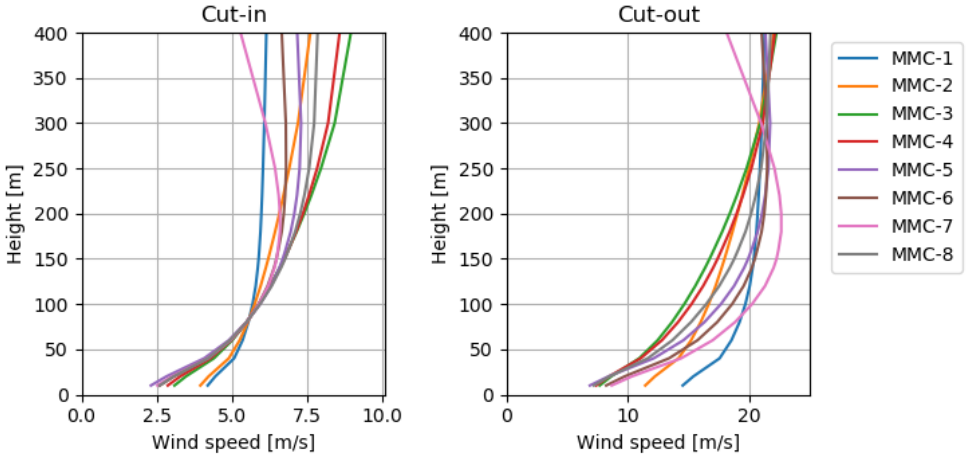


Figure 18. The cut-in (left) and cut-out (right) wind profiles that follow from scaling the onshore profile shapes ~~from~~ in Fig. 12 ~~such that the wind speed at 100 equals~~ (Sect. 4.3) using the calculated cut-in and cut-out wind speeds at 100 m, respectively.

Figure 19 shows the idealised cycle trajectories that follow from the optimisations for the cluster-mean wind profile shape of MMC-1. The ~~wind speeds for which the trajectories are depicted~~ depicted trajectories highlight changes in the operational approach at wind speeds at 100 m between cut-in and cut-out, which are discussed next. The optimal pumping tether length coincides with its upper bound for all wind speeds. The reel-out elevation angle of the flight trajectory ~~for $v_{100\mathrm{m}}$~~ below

v_{100m}=10.5 m s^{-1} coincides with its lower bound. For higher wind speeds, an increased inclination of the reel-out path yields a higher mean cycle power. At roughly ~~v_{100m}~~ v_{100m}=16 m s^{-1}, the ~~maximum~~ maximal mean cycle power is reached with the kite completing only one cross-wind pattern. Above 16 m s^{-1} wind speed, the ~~criterion of completing~~ constraint that requires completing at least one cross-wind pattern is driving the elevation angle to higher values until reaching its upper bound for ~~v_{100m}~~ v_{100m}=19.7 m s^{-1}, above which no feasible solution exists.

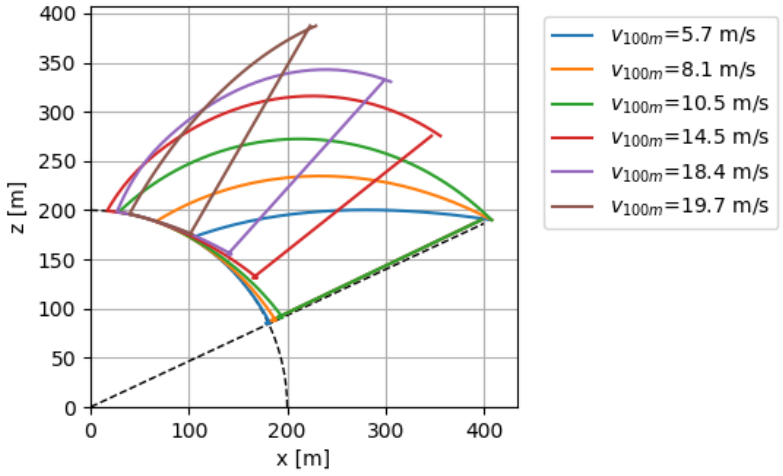


Figure 19. The optimal idealised cycle trajectories for six wind profiles with the same cluster-mean wind profile shape, i.e., that of cluster MMC-1, but different scaling. The wind speeds for which the trajectories are depicted highlight changes in the operational approach. The angle of inclination of the straight dotted line is the minimum elevation angle. The radius of the dotted quarter circle shows the fixed minimum tether length.

The calculated power curves are shown in Fig. 20. 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. Up to roughly ~~v_{100m}~~ v_{100m}=8.5 m s^{-1}, all the power curves are similar. Above this wind speed, the curves flatten off and become different from one another. The MMC-3 and MMC-7 curves show the lowest and highest ~~maximum~~ maximal mean cycle power, respectively. In conclusion, a pronounced low-level jet is favoured over a high-shear wind profile shape in terms of the power production of an AWE system.

5.2 Estimating the annual energy production

The previously ~~constructed~~ derived power curves are used ~~for calculating~~ to calculate the average generated power of the AWE system:

$$\bar{P} = \sum_{i=1}^{n_c} \int_0^{\infty} p_i(v_{\mathrm{norm}}) \cdot P_i(v_{100\mathrm{m}})\, \mathrm{d}v_{\mathrm{norm}} \approx \sum_{i=1}^{n_c} \sum_{j=1}^{n_b} \frac{f_{i,j}}{n_s} \cdot P_i(v_{100\mathrm{m}\,j,100\mathrm{m}}) \quad , \tag{9}$$

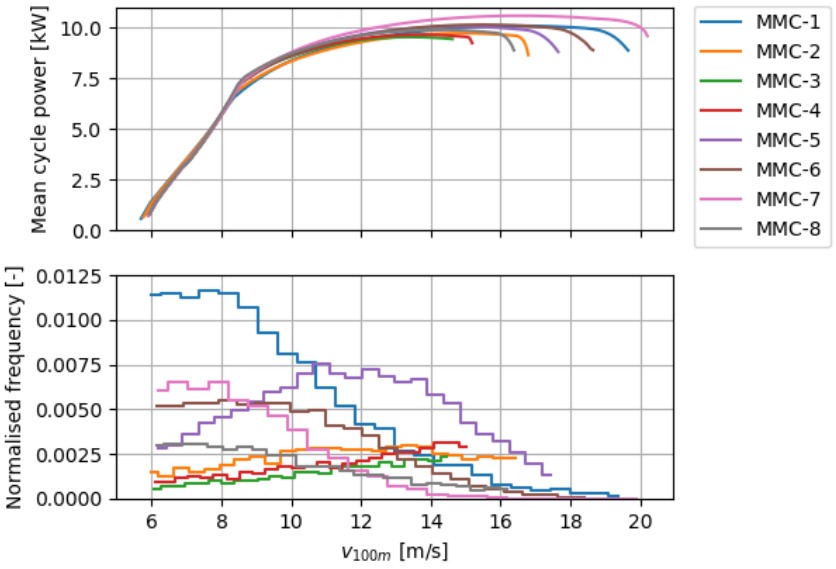


Figure 20. Power curves obtained by performance optimisations using the scaled cluster-mean wind profile shapes of MMC-1–8 (top). Wind speed distribution of the samples within each cluster using the full onshore dataset (bottom). Only the frequencies between cut-in and cut-out are depicted and ~~, for illustrative purposes,~~ every four wind speed bins are aggregated purely for illustrative purposes and shown as a single bin.

in which ~~p_i is~~ the wind speed probability ~~for cluster i, P_i is the power curve for the wind profile shape of cluster i, v_{norm} is~~ p_i is a function of the normalisation wind speed ~~, $v_{100\mathrm{m}}$~~ v_{norm} used in the pre-processing, the maximal mean cycle power P_i is a function of the wind speed at 100 m height ~~, n_{c} is the number of clusters, n_{b} is the number of wind speed bins and n_{s} is the total number of samples~~$v_{100\mathrm{m}}$, both functions apply to the i^{th} cluster, and n_{c} is the number of clusters. The integral in the expression is solved numerically using ~~100~~ $n_{\mathrm{b}} = 100$ wind speed bins between the cut-in and cut-out ~~wind speeds at 100~~ of equal width. A large number of bins is used to mitigate numerical errors. In the resulting right hand side expression, ~~p_i is approximated by normalising the sample~~ the number of samples n_{s} is used to normalise the frequency $f_{i,j}$ ~~for cluster i and wind speed bin j in the onshore dataset. A large number of bins is used to mitigate numerical errors. The probability of each cluster is characterised using~~ of the i^{th} cluster and j^{th} bin, which is determined on the basis of the normalisation wind speeds. Consequently, the normalisation wind speed ~~of the pre-processing. The equivalent~~ is used to express the bin limits. The argument of P_i is the equivalent wind speed at 100 m height ~~is calculated to determine the frequency in the wind speed bin, using: $v_{100\mathrm{m}} = v_{\mathrm{norm}} \cdot \hat{v}_{i,100\mathrm{m}}$~~at the center of the j^{th} bin, which is derived using: $v_{j,100\mathrm{m}} = v_{j,\mathrm{norm}} \cdot \tilde{v}_{i,100\mathrm{m}}$, in which ~~$\hat{v}_{i,100\mathrm{m}}$~~ $\tilde{v}_{i,100\mathrm{m}}$ is the normalised wind speed ~~of the wind profile shape of cluster i~~ at 100 m height of the i^{th} mean-cluster wind profile shape. The resulting wind speed distributions based on the normalised bin frequencies are shown in the lower panel of Fig. 20. Multiplying the average generated power by the hours in a year gives the AEP estimate.

The AEP at the onshore location is evaluated for the MMC and ML cluster representations from Sects. 4.3 and 4.4, respectively. Moreover, the number of clusters used for the representations is varied to assess how many clusters are needed for the

AEP to converge to a steady value, see Fig. 21. The trend for the MMC representation converges to around 36 MWh for a large number of clusters. In the following, we refer to the difference relative to the AEP at 32 clusters calculated using the MMC representation as the AEP error. For four or more clusters, the AEP error is within three percent ~~of the converged value~~ and for 14 or more clusters, there is virtually no more variation in the AEP and the steady solution is reached. The error can be mostly ~~explained by the inaccuracy of the cluster~~ attributed to the wind resource representation, but also ~~by~~ the numerically obtained power curves and ~~inaccuracies in the numerical integration~~ numerical integration introduce errors. The AEP trend for the MMC representation converges faster than that for the ML representation, since the former is generated specifically for the evaluated location. The ~~MMC and ML trends show a similar difference to the converged value at~~ AEP error at 16 ~~and~~ clusters for the MMC representation is similar to the AEP error at 32 clusters ~~, respectively, suggesting~~ for the ML representation, which suggests that the ML representation needs twice the number clusters to yield the same accuracy as the MMC representation. Note that assumptions in the performance model also affect the convergence, e.g., neglecting the change of wind direction with height is expected to increase the convergence rate. How many clusters to use depends on the application of the AEP calculation. In a preliminary design optimisation, where the computational cost is critical, four MMC-clusters may be a sensible choice. For more detailed design studies, 14 MMC-clusters would be more suitable.

~~More~~ Previously, at more than 50 ~~optimisations are used~~ wind speeds between cut-in and cut-out performance optimisations were performed to obtain each ~~power curve, yielding highly detailed curves. With only a small compromise on the~~ of the highly detailed power curves in Fig. **??**. Half the number of optimisations yield similar level of detail ~~, half of the number of optimisations can be used, which thus halves~~ with half the computational cost. Assuming that a four-cluster representation provides sufficient accuracy and 25 optimisations are used ~~for generating~~ to generate a single power curve, 100 performance optimisations are ~~required~~ needed for the AEP calculation. In ~~case of a~~ comparison, an hourly brute force calculation ~~, in which for every hour a separate optimisation is performed,~~ needs 8760 optimisations ~~are required~~ per year. ~~When only evaluating a single year,~~ Already for a one-year calculation the number of optimisations required by the presented methodology is ~~already~~ two orders of magnitude ~~less. Also, it~~ lower. The calculation for a longer period does not require more optimisations~~when evaluating a longer period~~, however, it does increase the computational effort for the clustering~~increases~~.

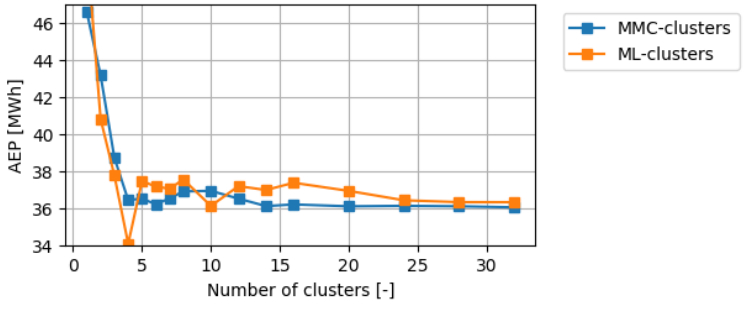


Figure 21. Comparison of the AEP convergence ~~for~~ with increasing number of clusters at the onshore reference location ~~using~~ for the MMC and ML cluster wind resource representations.

6 Conclusions

We have presented a ~~data-driven methodology~~ methodology for including multiple wind profile shapes in a wind resource description. A data-driven approach is used to identify a set of wind profile shapes ~~using clustering that can be used to characterise~~ that characterises the wind resource~~for AWE systems~~. These shapes go beyond the height range for which conventional wind profile relationships are developed, such as the logarithmic profile~~, are developed~~. Moreover, they ~~allow the quantification of phenomena such as the occurrence of~~ include non-monotonic wind profile shapes such as low-level jets. ~~The Dutch Offshore Wind Atlas is used for demonstrating the~~ We demonstrated this methodology for an on- and offshore reference location ~~.~~ using DOWA data. Subsequently, the resulting cluster wind resource representation for the onshore location has been used to estimate the AEP of a pumping AWE system.

~~Prior to a principal component analysis, the DOWA wind profiles are expressed in terms of the~~ To obtain the wind profile shapes of the DOWA samples, the wind profile of each sample is expressed relative to its wind velocity at the 100 m reference height and normalised. ~~The~~ A PC analysis shows that three PCs already account for about 90 % of the variance in the dataset. The first and second PCs are very similar for the datasets of the onshore and offshore locations. The first PC ~~characterises mostly~~ mostly characterises wind veer, whereas the second PC ~~characterises mostly~~ mostly characterises wind shear. Moreover, the analysis reveals a natural structure of the data in the principal component space with two relatively dense groups of data points. The data points for the onshore location are more spread out, indicating a larger variety of wind profile shapes.

The dataset is partitioned using k-means clustering~~and the~~ . The resulting cluster-mean wind profile shapes are ~~determined. In~~ used to approximate the vertical variation of the wind, yielding the cluster wind resource representation~~, each sample's vertical wind variation is approximated using these shapes, thereby,~~ . This representation reduces the wide variety of wind conditions in the DOWA dataset ~~is reduced~~ to a reasonable number of wind profile shapes. ~~Although some variability remains in the profile shapes of the samples assigned to each cluster, the cluster-mean shapes nonetheless allow a better representation of the vertical wind variation than relying on a monotonic logarithmic profile. Indeed, the~~ The accuracy of the representation using three or more clusters is already higher than that of ~~the~~ a representation using logarithmic wind profiles~~up to 600 . The~~ . The eight cluster-mean wind profile shapes of the offshore ~~eight-cluster representation shows~~ representation include three monotonic profiles, four jet-like profiles, and an anticlockwise-turning, ~~kinked~~ sharply-bent profile. Very similar cluster-mean wind profile shapes have been identified for the onshore location occurring under similar conditions. A single set of clusters is generated that is representative for the ~~whole~~ entire DOWA domain and used to analyse the spatial variability of the frequency of occurrence of the clusters. The cluster frequency maps indicate a clear distinction between onshore and offshore clusters. The sharply defined patterns in the frequency maps of the onshore clusters coincide with ~~terrain~~ orographic features and thus suggest a strong relationship between the ~~clusters and terrain~~wind profile shape and orography.

The AEP of a flexible-kite, pumping AWE system is estimated using the onshore cluster representation. For each cluster-mean wind profile shape, a power curve is ~~obtained by using the~~ derived by using a quasi-steady model in power production optimisations. The highest power is found for the shape with a pronounced low-level jet. Together with the respective wind speed distributions, the power curves yield the AEP contributions of the clusters. The relationship between the estimated AEP

and the number of ~~site specific~~ site-specific clusters shows that the ~~AEP error is within three percent of~~ difference in AEP relative to the converged value ~~when using~~ is less than three percent for four or more clusters. For 14 or more clusters, there is virtually no more variation in the AEP estimation. For a four-cluster representation and using 25 optimisations for ~~constructing~~ deriving the power curve of a single cluster, 100 optimisations are required for the AEP estimation against 8760 for an hourly brute-force calculation. The ~~proposed methodology~~ AEP estimation using clusters is thereby roughly two orders of magnitude faster.

The ~~methodology presented~~ presented methodology has the capability to produce a single set of wind profile shapes that is valid for a large area. Such a set can facilitate the standardisation of wind conditions for which AWE systems are rated in terms of power production. Moreover, the multi-location cluster representation ~~enable~~ enables an assessment of which installation site is best for an AWE system in terms of its AEP, which makes this methodology a very powerful tool for project developers. In future work, the role of the performance model in ~~making the AEP calculation~~ estimating the AEP is further investigated.

Author contributions. Analysis and preparation of the paper were performed by MS, under the supervision of RS and SJW. PCK provided input on meteorology and data analysis and contributed to writing the result section.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Without the publicly available Dutch Offshore Wind Atlas by the KNMI and the ERA5 dataset by the ECMWF, this work would not have been possible. Mark Schelbergen and Roland Schmehl have received financial support by the project REACH (H2020-FTIPilot-691173), funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 691173, and AWESCO (H2020-ITN-642682) funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 642682.

References

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA, https://doi.org/10.7289/V5C8276M, 2009.

Basu, S.: A simple recipe for estimating atmospheric stability solely based on surface-layer wind speed profile, Wind Energy, 21, 937–941, https://doi.org/10.1002/we.2203, 2018.

Bechtle, P., Schelbergen, M., Schmehl, R., Zillmann, U., and Watson, S.: Airborne wind energy resource analysis, Renew. Energ., 141, 1103 – 1116, https://doi.org/https://doi.org/10.1016/j.renene.2019.03.118, 2019.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, X., and Køltzow, M. Ø.: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System, Mon. Weather Rev., 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2017.

Brown, A. R., Beljaars, A. C. M., Hersbach, H., Hollingsworth, A., Miller, M., and Vasiljevic, D.: Wind turning across the marine atmospheric boundary layer, Q. J. Roy. Meteor. Soc., 131, 1233–1250, https://doi.org/10.1256/qj.04.163, 2005.

Burk, S. D. and Thompson, W. T.: The Summertime Low-Level Jet and Marine Boundary Layer Structure along the California Coast, Mon. Weather Rev., 124, 668–686, https://doi.org/10.1175/1520-0493(1996)124<0668:TSLLJA>2.0.CO;2, 1996.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), https://cds.climate.copernicus.eu/cdsapp#!/home, 2017.

Duran, P., Basu, S., Meissner, C., and Adaramola, M.: Automated classification of simulated wind field patterns from multi-physics ensemble forecasts, Wind Energy, https://doi.org/10.1002/we.2462, 2019.

Fechner, U.: A Methodology for the Design of Kite-Power Control Systems, Ph.D. thesis, Delft University of Technology, https://doi.org/10.4233/uuid:85efaf4c-9dce-4111-bc91-7171b9da4b77, 2016.

Floors, R., Peña, A., and Gryning, S.-E.: The effect of baroclinicity on the wind in the planetary boundary layer, Q. J. Roy. Meteor. Soc., 141, 619–630, https://doi.org/10.1002/qj.2386, 2015.

Gryning, S.-E., Batchvarova, E., Brümmer, B., Jørgensen, H., and Larsen, S.: On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer, Bound.-Lay. Meteorol., 124, 251–268, https://doi.org/10.1007/s10546-007-9166-9, 2007.

Heilmann, J. and Houle, C.: Economics of Pumping Kite Generators, in: Airborne Wind Energy, edited by Ahrens, U., Diehl, M., and Schmehl, R., pp. 271–284, Springer, Berlin Heidelberg, https://doi.org/10.1007/978-3-642-39965-7_15, 2013.

Holtslag, M., Bierbooms, W., and van Bussel, G.: Extending the diabatic surface layer wind shear profile for offshore wind energy, Renew. Energ., 101, 96 – 110, https://doi.org/https://doi.org/10.1016/j.renene.2016.08.031, 2017.

Holtslag, M. C., Bierbooms, W. A. A. M., and van Bussel, G. J. W.: Estimating atmospheric stability from observations and correcting wind shear models accordingly, J. Phys. Conf. Ser., 555, 012 052, https://doi.org/10.1088/1742-6596/555/1/012052, 2014.

Kalverla, P. C.: Characterisation of offshore winds for energy applications, Ph.D. thesis, Wageningen University, https://edepot.wur.nl/498797, 2019.

Kalverla, P. C., Steeneveld, G.-J., Ronda, R. J., and Holtslag, A. A.: An observational climatology of anomalous wind events at offshore meteomast IJmuiden (North Sea), J. Wind Eng. Ind. Aerod., 165, 86 – 99, https://doi.org/https://doi.org/10.1016/j.jweia.2017.03.008, 2017.

Kalverla, P. C., Duncan Jr., J. B., Steeneveld, G.-J., and Holtslag, A. A. M.: Low-level jets over the North Sea based on ERA5 and observations: together they do better, Wind Energ. Sci., 4, 193–209, https://doi.org/10.5194/wes-4-193-2019, 2019.

Kelly, M. and Gryning, S.-E.: Long-Term Mean Wind Profiles Based on Similarity Theory, Bound.-Lay. Meteorol., 136, 377–490, https://doi.org/10.1007/s10546-010-9509-9, 2010.

Mahrt, L., Vickers, D., and Andreas, E. L.: Low-Level Wind Maxima and Structure of the Stably Stratified Boundary Layer in the Coastal Zone, J. Appl. Meteorol. Clim., 53, 363–376, https://doi.org/10.1175/JAMC-D-13-0170.1, 2014.

Malz, E., Koenemann, J., Sieberling, S., and Gros, S.: A reference model for airborne wind energy systems for optimization and control, Renew. Energ., 140, 1004 – 1011, https://doi.org/https://doi.org/10.1016/j.renene.2019.03.111, 2019.

Malz, E., Hedenus, F., Göransson, L., Verendel, V., and Gros, S.: Drag-mode airborne wind energy vs. wind turbines: An analysis of power production, variability and geography, Energy, 193, 116 765, https://doi.org/https://doi.org/10.1016/j.energy.2019.116765, 2020a.

Malz, E., Verendel, V., and Gros, S.: Computing the power profiles for an airborne wind energy system based on large-scale wind data [Submitted], Renew. Energ., 2020b.

Monin, A. and Obukhov, A.: Basic laws of turbulent mixing in the ground layer of the atmosphere (in Russian), Trudy Geofiz. Inst. Akad. Nauk SSSR, 151, 163–187, 1954.

Oehler, J. and Schmehl, R.: Aerodynamic characterization of a soft kite by in situ flow measurement, Wind Energ. Sci., 4, 1–21, https://doi.org/10.5194/wes-4-1-2019, 2019.

Optis, M., Monahan, A., and Bosveld, F. C.: Moving Beyond Monin–Obukhov Similarity Theory in Modelling Wind-Speed Profiles in the Lower Atmospheric Boundary Layer under Stable Stratification, Bound.-Lay. Meteorol., 153, 497–514, https://doi.org/10.1007/s10546-014-9953-z, 2014.

Parish, T. R.: Forcing of the Summertime Low-Level Jet along the California Coast, J. Appl. Meteorol., 39, 2421–2433, https://doi.org/10.1175/1520-0450(2000)039<2421:FOTSLL>2.0.CO;2, 2000.

Park, J., Basu, S., and Manuel, L.: Large-eddy simulation of stable boundary layer turbulence and estimation of associated wind turbine loads, Wind Energy, 17, 359–384, https://doi.org/10.1002/we.1580, 2014.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Perez, R. E., Jansen, P. W., and Martins, J. R. R. A.: pyOpt: A Python-Based Object-Oriented Framework for Nonlinear Constrained Optimization, Struct. Multidiscip. O., 45, 101–118, https://doi.org/10.1007/s00158-011-0666-3, 2012.

Peterson, E. W. and Hennessey, J. P.: On the Use of Power Laws for Estimates of Wind Power Potential, J. Appl. Meteorol., 17, 390–394, https://doi.org/10.1175/1520-0450(1978)017<0390:OTUOPL>2.0.CO;2, 1978.

Ranjha, R., Svensson, G., Tjernström, M., and Semedo, A.: Global distribution and seasonal variability of coastal low-level jets derived from ERA-Interim reanalysis, Tellus A, 65, 20 412, https://doi.org/10.3402/tellusa.v65i0.20412, 2013.

Ranneberg, M., Wölfle, D., Bormann, A., Rohde, P., Breipohl, F., and Bastigkeit, I.: Fast Power Curve and Yield Estimation of Pumping Airborne Wind Energy Systems, in: Airborne Wind Energy: Advances in Technology Development and Research, edited by Schmehl, R., pp. 623–641, Springer, Singapore, https://doi.org/10.1007/978-981-10-1947-0_25, 2018.

Salma, V., Ruiterkamp, R., Kruijff, M., Paassen, M. M. R. v., and Schmehl, R.: Current and Expected Airspace Regulations for Airborne Wind Energy Systems, in: Airborne Wind Energy – Advances in Technology Development and Research, edited by Schmehl, R., Green Energy and Technology, chap. 29, pp. 703–725, Springer, Singapore, https://doi.org/10.1007/978-981-10-1947-0_29, 2018.

Salma, V., Friedl, F., and Schmehl, R.: Reliability and Safety of Airborne Wind Energy Systems, Wind Energy, 23, 340–356, https://doi.org/10.1002/we.2433, 2019.

Shapiro, A., Fedorovich, E., and Rahimi, S.: A Unified Theory for the Great Plains Nocturnal Low-Level Jet, J. Atmos. Sci., 73, 3037–3057, https://doi.org/10.1175/JAS-D-15-0307.1, 2016.

Sommerfeld, M., Crawford, C., Monahan, A., and Bastigkeit, I.: LiDAR-based characterization of mid-altitude wind conditions for airborne wind energy systems, Wind Energy, 22, 1101–1120, https://doi.org/10.1002/we.2343, 2019.

Van der Vlugt, R., Bley, A., Noom, M., and Schmehl, R.: Quasi-Steady Model of a Pumping Kite Power System, Renew. Energ., 131, 83–99, https://doi.org/10.1016/j.renene.2018.07.023, 2019.

Van Hussen, K., Dietrich, E., Smeltink, J., Berentsen, K., van der Sleen, M., Haffner, R., and Fagiano, L.: Study on challenges in the commercialisation of airborne wind energy systems, Tech. Rep. ECORYS Report PP-05081-2016, European Commission, Brussels, https://doi.org/10.2777/87591, 2018.

Verkaik, J. W.: On Wind and Roughness over Land, Ph.D. thesis, Wageningen University, 2006.

Watson, S., Moro, A., Reis, V., Baniotopoulos, C., Barth, S., Bartoli, G., Bauer, F., Boelman, E., Bosse, D., Cherubini, A., Croce, A., Fagiano, L., Fontana, M., Gambier, A., Gkoumas, K., Golightly, C., Latour, M. I., Jamieson, P., Kaldellis, J., Macdonald, A., Murphy, J., Muskulus, M., Petrini, F., Pigolotti, L., Rasmussen, F., Schild, P., Schmehl, R., Stavridou, N., Tande, J., Taylor, N., Telsnig, T., and Wiser, R.: Future emerging technologies in the wind power sector: A European perspective, Renew. Sust. Energ. Rev., 113, 109 270, https://doi.org/10.1016/j.rser.2019.109270, 2019.

Wijnant, I. L., van den Brink, H. W., and Stepek, A.: North Sea wind climatology part 1: a review of existing wind atlases, Tech. rep., KNMI, 2015.

Wijnant, I. L., van Ulft, B., van Stratum, B., Barkmeijer, J., Onvlee, J., de Valk, C., Knoop, S., Kok, S., Marseille, G. J., Klein Baltink, H., and Stepek, A.: The Dutch Offshore Wind Atlas (DOWA): description of the dataset, Tech. rep., KNMI, 2019.