

I thank the authors for their work to make the manuscript better. I think that great improvements have been made. I still have some minor suggestions to improve the readability and make the paper more understandable for the general reader.

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.
2. The overbar in  $\bar{\theta}_v$  in formula (5) is hard to see. Please check that this gets corrected in the final proof.
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.
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”.
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”.
6. I would rephrase the statement that MMJ-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.

# **Clustering wind profile shapes to estimate airborne wind energy production - 2nd Referee Comment**

Markus Sommerfeld<sup>1</sup>

<sup>1</sup>Institute for Integrated Energy Systems, University of Victoria, British Columbia, Canada

**Correspondence:** Markus Sommerfeld (msommerf@uvic.ca)

## 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, trans-  
5 formed 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.

### 10 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 sen-  
15 tences. Some examples can be found 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  
20 which they refer to, e.g. don’t mention L value, but associated stability bin in text box in each profile plot 1-8

Wind data: The usage of  $Ri$ ,  $L$  and  $\Psi$  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.

## 2 Specific comments including technical corrections

### 2.1 Abstract

25 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.

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

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

line 10: Add: “ ... for each wind speed profile shape”.

line 12: 4 cluster error relative to what?

## 2.2 Introduction

page 1

35 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."

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

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

page 2

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

page 3

45 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."

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

### 2.2.1 Wind dataset

page 4

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

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

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

55 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.

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

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

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

60 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.

## 2.2.2 ERA5

page 5

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

```
65     i.         \begin{small}
                @misc{website:era5,
                Author = {Hans Hersbach and Dee Dick},
                Month = {November},
                Title = {{ERA5} reanalysis is in production},
70     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}}
```

75 line 2: Add: reference to ECMWF

line 7: "... is performed on the DOWA data..."

line 7: Remove: "As is explained later on"

## 2.2.3 DOWA

page 5

80 line 10: Remove: "The"; spelling of "downscaled" different from above.

line 13: Add "grid spacing" and "grid points" for clarity.

page 6

line 1: Add reference to improved DOWA performance.

line 3: Clarify "routine weather stations"

## 85 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"

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

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

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.”

95 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. ”

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

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

### 100 2.3.2 Principal component analysis of the wind profile shape dataset

page 6

line 27: Replace: “while” with “which”

page 7

line 7: Replace: “layer” with “height”

105 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.

line 16 + 17: Word order: mostly characterises.

110 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?

115 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?

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

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

page 8

120 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?

### 2.3.3 Choosing the number of clusters

page 8

line 2: Replace: applied instead of employed

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

125 line 9: Do wind profiles have such a structure?

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" ?

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

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

line 5: Reference: appropriate choices of k?

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?

135 figure 5: Add: "k-means" before clustering ; remove: point before (b)

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

line 7: Add :  $\varepsilon_{ij}$  here to make equation easier to understand. Define between which data the error is calculated.

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

140 page 11

line : Is the error only based on different 5  $\Psi$  values?

line 5: Sentence needed?

line 6: Add reference to  $\Psi$  function; replace every with each

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

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

line 19: Why not just wind speed error?  $\bar{\varepsilon}_i$  never used in any equation

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

line 21: Replace vertical grid with dataset or rewrite

page 12

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

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?

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

## 2.4 Cluster wind resource representation

155 page 13

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

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

line 6: Replace: full with entire

### 2.4.1 Cluster representation for the offshore location

160 page 13

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

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

line 15: Explain: stability function is varied continuously?

page 14

165 table 2: Replace: at with in?

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 ”

line 6: Replace: for with of

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

170 line 9: Which data did you use?  $L$  is always calculated using surface data (see: [http://glossary.ametsoc.org/wiki/Obukhov\\_length](http://glossary.ametsoc.org/wiki/Obukhov_length))

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

175 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?

## 2.4.2 Interpretation of the offshore cluster representation

page 18

- 180 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
- line 5: Add: magnitude of profiles? replace: well-described
- line 6: Replace: shear and veer
- line 8: Shorten: North-West direction
- 185 line 10: Equilibrium profile meaning a unstable shape?
- line 25: Remove: “The frequent”; Rewrite: Winds at this location with a southerly component...
- line 29: Remove: “... and gradually”

page 19

- line 1: Replace: “abrupt kink” with “sharp bend”
- 190 line 5: Replace: “rather than” with “and less often with”
- line 6: Replace: shear and veer
- line 8: Shorten: North-West direction
- line 10: Equilibrium profile meaning a unstable shape?
- line 25: Remove: “The frequent”; Rewrite: Winds at this location with a southerly component...
- 195 line 29: Remove: “... and gradually”

## 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.
- line 9: Compared how and where?
- 200 line 12: Rewrite: “the mean profile shape below 200 m is in accordance with a stable logarithmic profile.”
- line 15: Increased relative to what?
- line 16: Add reference to where they are plotted
- line 16: If they share the same coordinate system does that mean that both locations have the same PCs?
- line 24: “...clustering algorithm...”; “... for each cluster.”

205 line 25: Replace: more or less with onshore closely resembles offshore  
line 28: Remove: again  
line 29: Obukhov length determined how?  
line 30: Meaning? “show an increase in wind shear”. Increasing wind gradient from 1 to 3?  
line 34: Replace: “turning” with “rotation” and “kink” with “sharp bend”

210 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.

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

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

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

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

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

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

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

page 21 + 22

figure 11+12: See same figures for offshore

page 23

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

page 24

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

#### **2.4.4 Spatial frequency distribution of wind profile shape clusters**

page 25

230 line 2: Replace: “are” with “were”

line 3: Replace: “for” with “from”

line 10: Remove: “So”

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

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

line 22: 21.7% ?Reference Table 3 here already

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

page 26

table 3: How do you quantify similarity?

240 page 28

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

## **2.5 Efficient AWE production estimation using the cluster representation**

page 29

line 6: Add: power curve and AEP

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

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

### **2.5.1 Constructing the power curves**

page 29

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

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

page 30

line 2: What is V3 kite?

255 line 4: Add: “average energy production”

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?

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

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

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

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

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

line 1: Replace: “are exceeded” with “would be exceeded”?

265 page 32

line 9: Does this intersection depend on tether length?

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

270 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?

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?

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

line 16: criterion is also a constraint?

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

page 33

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

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.

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

## 2.5.2 Estimating the annual energy production

285 page 34

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

line 4: Is  $P_i$  the power curve or 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?

290 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?

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?

295 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?

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

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

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

page 35

line 1: "orders of magnitude lower"

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

### 2.5.3 Conclusion

page 35

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

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

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

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

315 line 10: "... profile shape variance."

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."

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

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

page 36

line 1: Replace: “entire DOWA domain”

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

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

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

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