Authors' response to comments by referee #2

Manuscript title: The selection of directional sectors for the analysis of extreme wind speed. We would like to thank the referee for all the valuable comments that will help us to improve our manuscript and future research. Comments from the referee are written in italics, followed by our response.

Referee #2

The paper presents an original method to be employed for the selection of directional sectors for the analysis of the extreme wind speed in order to develop design of structures exposed to wind action. I found the subject of the manuscript relevant for NHESS and I think that it could be of interest for other scientific and engineering field such as, for example, coastal and offshore engineering. The formulation of the problem, the presentation and the discussion of the results are clear and adequately extended. I have some comment and observation about some aspect of the analysis presented by the authors:

a) *The authors decided to employ a criterion for the selection of directional sectors based on different statistical requirements and indicators. What could be the difference by the use with some sort of clustering techniques (k-means for example, or similar)? Could the authors comment on this point and eventually add some discussion in the paper*

   Starting from the information of wind direction and speed, a cluster analysis methodology would allow to determine subsets of similar data in terms of some distance measure that could be defined in terms of these two variables. This can be seen as a way to define directional sectors for the extreme analysis. However, cluster analysis does not usually include metrics to ensure that data subsets are homogeneous and independent of each other. In turn, the use of cluster analysis requires defining the clustering methodology, the distance measure between data and the number of clusters.

   It is possible that the methodological approach suggested in this article can be adapted to guide the analyst in the definition of these three points in order to obtain directional groups that meet the requirements imposed for the extremal analysis (i.e. homogeneous and independent sectors such that the variance of estimated extremes is minimized).

b) *I would make dimensionless the global indicator: instead of varying between 0 and sqrt(3) I would make it varying between 0 and 1*

   We agree with the referee that varying the global indicator between 0 and 1 leads to a more direct understanding of its value and it does not change the method or the results in any way. For that reason, we have changed equation 5 according to the referee's proposal.

c) *I would add some plot about the minimum data for sectors and subsectors (first paragraph of section 3.4)*

   The power curves for the Anderson Darling and Kolmogoroff-Smirnoff tests have been obtained as indicated in Appendix A and have been added to the manuscript (section 3.4), along with a more detailed description of their use in this article.

d) *I would put the x scale of figure 4 for criterion C0 varying from 0 to 360 in order to have a visual comparison with Criterion 45 and 90*

Varying the x scale of C0 from 0 to 360 divides sector S3 into two parts (one on the right and the other on the left of the figure). In our opinion, this makes it more difficult to understand the figure and to place the boxplots in a consistent manner. For this reason, although the referee's proposal has some advantages, we believe it is better to maintain the figure 4 in its current state.

e) *May be I miss something but I do not understand why sigma_s and u are in [m] in table 3*

We appreciate the referee's comment. The units for "sigma_s" and "u" are [m/s]. We have changed the units in columns 6 and 7 in table 3 accordingly.

f) *I would add bound conditions in table 3 in order to have a clear defined picture of the quantities involved*

We agree with the referee that the inclusion in table 3 of a column with the upper bound of the wind velocity in each sector helps to illustrate the differences between the extreme-value models for each criterion, as indicated in section 4.1.

g) *There is any effect on the results on the choice of different inter arrival time (different from 5 days)?*

The inter arrival time between independent storms should have a physical/statistical sense and can be chosen either by rational methods or experience. In any case, performing a sensitivity analysis for testing the impact of this parameter on the results is recommended.

In the case study, an inter arrival time of 5 days between the peaks of consecutive storms has been adopted, which leads to 270 storms. Table 1 shows the results for inter arrival times of 2, 5, 7 and 10 days. In all these cases, the number of directional sectors obtained according to the method is three. First column shows the inter arrival time; second one the number of storms; from third to fifth column, the directional sectors; and from sixth to eighth column, the 100 year return value in each sector.

Δt	N_s	S_1	S_2	S_3	$u_{100,S1}$	$u_{100,S2}$	$u_{100,S3}$
[days]	[-]	[deg]	[deg]	[deg]	[m/s]	[m/s]	[m/s]
1	294	140-235	**235-290**	290-140	21.8	22.1	20.4
2	288	**125-235**		**290-125**	21.7	22.2	20.4
5	**270**				**21.8**	**22.1**	**20.5**
7	256				21.6	21.9	20.7
10	238				21.4	21.8	20.3

TABLE 1

Results show how, in the case study, the method is very little sensitive to changes in the inter arrival time. The directional sectors are the same for the inter arrival times of 2, 5, 7 and 10 days and differ in 15 degrees for the inter arrival time of 1 day. The 100 year return values are also consistent, with variations of less than 2.1% in all sectors.

h) It is not clear to me how the threshold has been chosen (and it would be nice to have indication about changes in the results depending on the threshold)

In this study, the threshold used to define the extreme data was chosen by applying the method in Solari et al. (2017), i.e. the threshold that maximizes the p-value of the Anderson-Darling test, under the assumption that the omnidirectional data come from a Generalized Pareto Distribution.

As shown in Table 2, directional sectors are sensitive to changes in the threshold u (first column), i.e., different definitions of what is an extreme value result in different directional sectors. However, for thresholds above the chosen one, the sectors are quite stable (in particular S2). Therefore, if there is no strong criterion for an a priori selection of the threshold, a sensitivity analysis of the results is recommended. In this situation, p0 (now defined as suggested in (b) in column 9) could serve as an indicator for the final choice of the threshold.

These considerations have been included in the last paragraph of section 4.2.

u	N_s	S_1	S_2	S_3	$u_{100,S1}$	$u_{100,S2}$	$u_{100,S3}$	p_0
[m/s]	[-]	[deg]	[deg]	[deg]	[m/s]	[m/s]	[m/s]	[-]
14.0	306	20-140	140-195	195-20	19.9	21.5	23.0	0.9489
14.3	**270**	**125-235**	**235-290**	**290-125**	**21.8**	**22.1**	**20.5**	**0.8822**
14.5	251	145-235	235-285	285-145	22.5	22.0	20.5	0.8799
15.0	203	200-235	235-290	290-200	23.0	20.7	21.5	0.8668

TABLE 2

Solari, S., Egüen, M., Polo, M. J., & Losada, M. A. (2017). Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value. Water Resources Research, 53(4), 2833-2849.

i) in last line of page 7 there is a typo "rfrg"

We have corrected this typo.