Nat. Hazards Earth Syst. Sci. Discuss., 2, C871–C881, 2014
www.nat-hazards-earth-syst-sci-discuss.net/2/C871/2014/

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Bayesian trend analysis of extreme wind using observed and hindcast series off Catalan coast, NW Mediterranean Sea" *by* M. I. Ortego et al.

M. I. Ortego et al.

ma.isabel.ortego@upc.edu

Received and published: 26 May 2014

The reviewer acknowledges the interest of the topic discussed in the manuscript but suggest a long list of topics that require further attention or detailed explanation. In general, we agree that the suggested topics are of interest although most of them may be out of the possibilities of the present paper, both due to the limited extension and the fact that these suggested studies require further experience on the use of Bayesian methods to estimate extremal parameters. Precisely, this paper is the result of a second step using this kind of model (two series, secular trend, using extremal data and Generalised Pareto Distribution). The first step was an analysis of ocean waves

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

where the two series, the hindcast series and the recorded at a buoy overlapped some years. As a consequence, the paper is mainly methodological and many practical but interesting points are still a matter of research or further development.

We answer the comments by reviewer number two, at the same time that we try to incorporate some of these suggestions into the manuscript. We thank Reviewer for his/her detailed reading and report.

The criteria for selecting an event (threshold of wind velocity of 15m/sec), specifying which wind velocity we are talking about (e.g. height with respect to the land or sea level) and the sensitivity of the performed analysis to different thresholds.

The deep buoy records the wind at height of 10m from the sea level. Hindcast wind is not recorded but predicted by the model at different heights. Here the level 10m over surface has been selected to assure a minimum of consistency between the two data sets. The criteria to define a wind event have been chosen to meet three hypothesis which are used in the proposed methodology. The first one is that the events are mutually independent conditional to the parameters of the model. This is, once the parameters of the model are fixed, the observations can be considered independent. Although this hypothesis cannot be checked over a short data series, autocorrelation of events should be avoided. The experience when dealing with meteorological events is that 3 days separation between events is enough to guarantee the absence of strong autocorrelations. The methodology consists of filtering the series from annual seasonality and estimate the autocovariance function. Separation between events should be of the order of the first zero crossing of the autocovariance function. Also auto-regression techniques may be used. As filtering extremal events (irregularly sampled) in order to remove seasonality involves some difficulties, the study has not been performed and it is not presented in the paper.

NHESSD

2, C871–C881, 2014

A second assumption is that events are punctual in time. This is not true as wind events can be characterised by a duration over a threshold. Therefore a standard procedure, seldom described in detail, consists of defining duration of the event between two crossings (upwards and downwards) of the threshold under the condition that the downwards crossing occur a number of days before the next upward crossing. Once the duration of the event is established, the maximum 10min wind average within the event is defined as the magnitude; if necessary or convenient, magnitude is re-defined as the logarithm of the previous one. This was the strategy followed in the pre-processing of our data. We will review our presentation in the manuscript in order to clarify the definition of event.

Finally, the threshold of 15m/s has been selected for attaining a certain degree of goodness of fit to the Generalised Pareto Distribution (GPD) in its Weibull domain. This was performed using the diagram of mean excess over threshold as a function of the threshold itself. It is known that the diagram should appear as a straight-line from the selected threshold up to larger ones. This diagram could be shown in the paper, but to the price of enlarging the manuscript detailing a preliminary pre-processing of the data set. The QQ-plots of the data, fitting the GPD are presented in the answer to Reviewer number 1. In both cases (hindcast data and observed data) the hypothesis of a GPD (Weibull domain of attraction) were not rejected at 0.05 significance.

The paper should also discuss why the magnitude of the event (maximum 10min averaged velocity during the event) is selected and again how the results would change with a different definition.

This is a long and interesting discussion. Wind and other meteorological variables, can be studied in different scales. For intrinsically positive magnitudes the relative scale is likely to be the most appropriated: the zero wind (not under detection limit) is unattainable and a logarithm place this (no movement) wind at infinity. Ratio

NHESSD
2, C871–C881, 2014

comparisons are meaningful and normally acceptable. If the relative scale is removed using logarithms, to transform it into absolute scale, the averages should be performed on the logarithm of the velocity and not on the velocity itself. Other scales, imply similar deficiency in carrying out spatial or temporal averages. For instance, the square velocity is an energy scale which would imply that spatial and temporal averages should be carried out on the square-wind. As the temporal average is performed on the velocity and the data are 10 minute averages, we guess that there may be an important influence in the results. However, the authors are not able to afford such an study as we have not access to data bases where the wind speed has been averaged for, let us say, 10s, 20s, 30s, 1min, 2min, etc. The strategy adopted in the manuscript was the only possible with the available data.

The paper should also make some reference to the randomness assumption for the time between events. The self-sustainability of convective low pressure centres in areas as the Mediterranean makes the occurrence of twin storms a relatively common pattern and that clearly goes against the hypothesis of randomness for time between events.

This statement will be addressed with the definition of wind event. Twin storms are considered different wind events whenever the wind is maintained under 15m/s for more than 3 days, what means that the *twin* character is at least doubtful. For storms more close in time than 3 days, the twin events are considered as a single event.

The paper should explain the sensitivity of the inhomogeneous Poisson process to model occurrences of events over a threshold to the functional dependence of λ with time.

NHESSD
2, C871–C881, 2014

The authors are conscious that a linear trend is not sustainable as the variation is unbounded in time, except for null slope (no linear trend). Accordingly, possible trends in parameters should be bounded. For instance, logistic trends would be an interesting option. However, this implies the introduction of a larger number of parameters to be estimated with a very limited records. Parsimony principle would recommend using the simplest model (linear trend) to the price of an inconsistent long-time prediction. Sensitivity analysis implies the introduction of alternative models including more parameters. Accordingly, this study is out of the scope of the present manuscript.

The paper should also discuss the importance of the available wind data and their quality for carrying out the proposed analysis.

The methodological character of the paper prevents a discussion on the strategies of data collecting for which authors are not qualified.

The variation of λ with time in a linear manner should be related to the projections which consider a concentration of storm events in a shorter interval during the year, and probably a reduction of the expected numbers of events with a likely increase in intensity. The inhomogeneous Poisson process should be discussed also from that perspective.

Certainly, an increase of λ represents increasing of concentration of events in a shorter time. A statement in this sense will be included. However, the suggested question on the concentration of events on time intervals within a year cannot be evaluated with the present model in which stationality is not considered. On the other hand, the increase of the size of the extremal events is directly related with the behaviour of the β-GPD parameter closely related with μ in the new parameterisation. A sentence clarifying

NHESSD
2, C871–C881, 2014

the relationship is included.

The paper should also discuss why the logarithmic scale enhances the compatibility in statistical terms of observed data and simulated data. In particular why the introduction of the logarithmic scale brings both data sets closer to the same domain of attraction in a GPD.

From our point of view, the short discussion already included in the manuscript should be enough for the reader. From the probabilistic point of view the Frechet domain of attraction is transformed into the Gumbel's domain of attraction after a logarithmic transformation of the sample. This probabilistic statement should be enough to understand why log-transformed wind-data adapts better to the Weibull domain of attraction. From an practical point of view, the use of logarithms make the large events to appear closer than in an absolute scale. An hypothetical upper limit seems to be closer to observed data than if it is considered in the absolute scale.

The paper should discuss further a physical meaning of a linear trend in time for the assumed variables. In particular for the ν variable in equation 8 so as to relate this linear variation in time with the expected projections from climatic models or even the expected qualitative change in patterns (e.g. northward shift of storm tracks) under future scenarios.

The expression of the parameter, $\nu = \log(-\xi)$, explains the meaning of the increase of the parameter. As ν increases (ξ decreases) the probability of excesses moves towards values of the excess closer to the absolute maximum. For instance, $\nu = \log(1/2) < 0$ corresponds to a triangular density of the GPD, whereas $\nu = \log(1) = 0$ corresponds to a uniform probability density which obviously is not decreasing near the absolute

NHESSD
2, C871–C881, 2014

maximum. Some sentences will be included in the manuscript to clarify these facts.

The paper should also discuss if larger records to estimate trends are needed and how long should those records should be. The practical implications for maintaining observations should also be highlighted since for instance in figure 4 it appears that larger records for δ_ν and α_ν would be desirable.

In general, extremal analysis require data records as long as possible, as extremal events are scarce by their own nature. Therefore, longer records are always desirable. Large uncertainties in the estimates are always present and the only way to reduce is to enlarge the data availability; but this is quite difficult going back to the past. The use of hindcast data is therefore important as it allows to increment the available information to the price of new difficulties. Precisely, the goal of the paper is this one: discussing the possibilities of jointly use of hindcast data and actual observations. Some sentences will be introduced in the introduction to remark these aspects.

The paper should also discuss more in detail the strength of the evidence that the upper limit for the studied variable (wind velocity) is the same for simulations and observations and what would happen if that is not the case.

A little summary of posterior distribution of the difference between the upper limits of hindcast and observed data will be introduced in the results section.

The paper should also discuss the robustness of the obtained conclusions particularly considering if the definition of storms is modified and how that would change their frequency, intensity, etc., independently of climatic change.

The statistical methods used in the paper are not robust in front of contamination of the

NHESSD
2, C871–C881, 2014

data. Presently, robust statistical methods filters out part of the information contained in the data sets to better fit central models. However, the present study deals with extreme data and standard robust methods would not be used in this context. We understand that the reviewer is referred to sensitivity to the mentioned assumptions. We are not going to introduce such a discussion that would represent doubling the extension of the paper.

The paper should rephrase carefully the sentences about the consistency of the locally obtained results with IPCC projections, since these IPCC projections are global and are not necessarily coincident with local ones such as the ones considered in the paper.

We will review these statements.

The paper should also discuss more in detail the implications of using the same averaging intervals for measurements and observations, as a function of the data availability and the variability apparent in the series.

Excuse us. We are not understanding the suggestion. Up to our understanding, any averaging of extremal data destroys most of the extremal information.

The paper should also discuss the implications of the observations suggesting a slightly larger number of events per year (about 2). How that would affect the analysis if explicitly considered in the proposed equations.

Excuse us again as we do not understand the suggestion. The increase of the number of events per year is described by the slope parameter α_λ (the increase of the number of events per year along the data record). This means that the increase of events per

NHESSD
2, C871–C881, 2014

year is directly considered in the proposed model. On the other hand the increase of two events increase refers to the difference between the buoy and hindcast data.

The paper should also consider if the joint analysis of observations and simulations could be repeated for a sub-sample of simulations corrected against the observations and how that would affect the obtained conclusions.

Assuming that *sub-sample of simulations* refers to the hindcast data set, this is again a sensitivity analysis that requires a long exposition and the subsequent report of the results. Although we consider this study is interesting or even necessary for a complete project on data assimilation (this time, assimilation of hindcast data), we would not want to add such an study in the present manuscript.

Finally the paper should be read by a native English speaker to improve some of the sentences.

The language used in the paper will be accordingly reviewed.

References

Ortego, M. I., R. Tolosana-Delgado, J. Gibergans-Báguena, J. J. Egozcue, and A. Sánchez-Arcilla (2012). Assessing wavestorm hazard evolution in the NW Mediterranean with hindcast and buoy data. *Climatic Change 113*, 713–731.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 799, 2014.

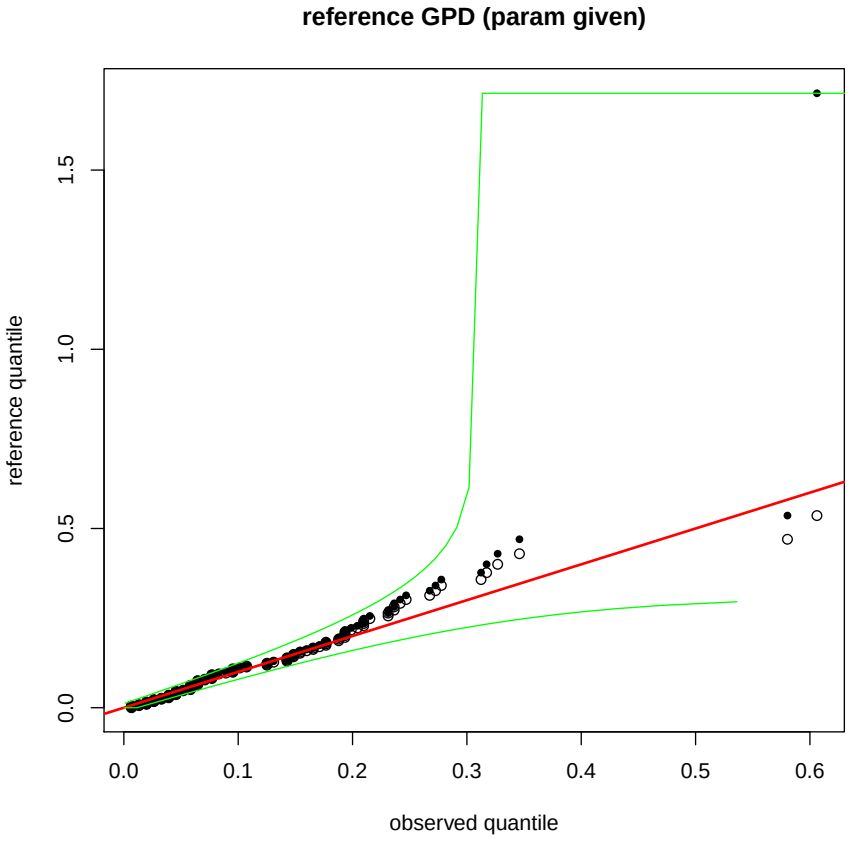

Fig. 1.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

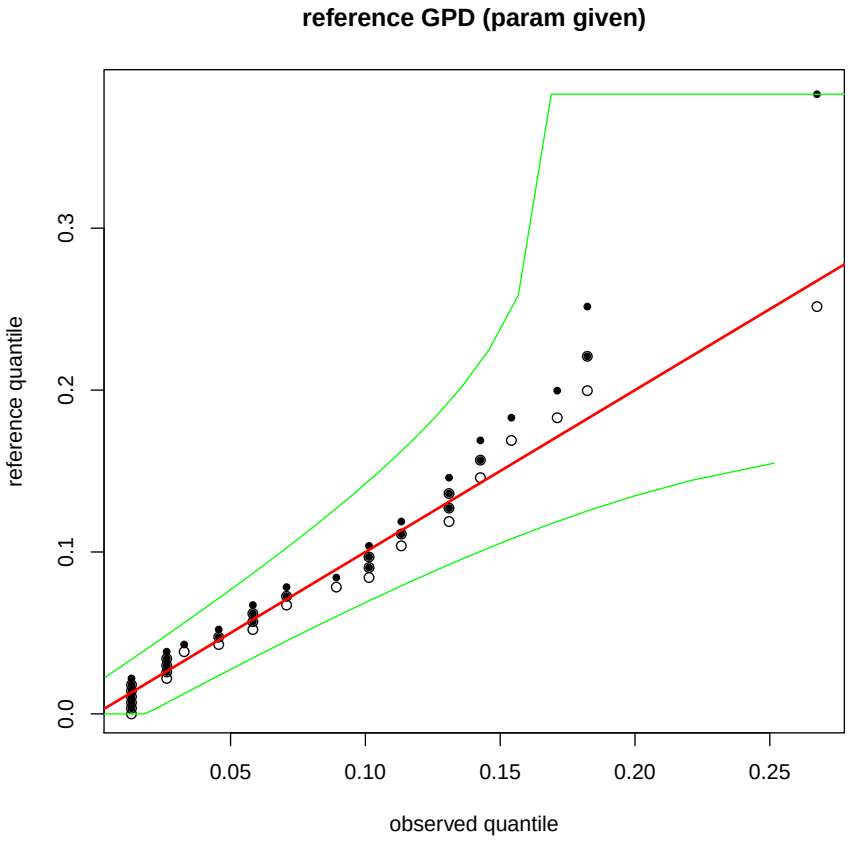

Fig. 2.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper