Modelling temporal trend within bivariate generalized Pareto models of logistic type

I would like to thank the authors for writing an interesting paper that is very relevant for the readership of the Environmetrics journal. The authors make use of the bivariate generalised Pareto distribution to jointly model wind speed extremes at pairs of sites in Germany, and investigate the use of non–stationary forms of this model that allow the model parameters – including one which quantifies the extent of extremal dependence – to change through time. The authors then use these models to construct, and compare, prediction regions for extreme wind speeds, before addressing the issue of model validation.

This very applied paper would be of great interest to practitioners of extreme value theory. I do, however, have a few issues that I would like the authors to respond to before this paper is considered for publication.

1. **Length of manuscript**
   I would like to thank the authors on the clarity of this paper. Very rarely is a paper so accessible on it's first read! However, I do feel that, in places, some of the material might be a little *over*–explained, resulting in an article that is unnecessarily long. For example, are both the abstract and the introduction a little too long? And, for example, do the authors really need to explain to the reader what is meant by "efficiency gains" (page 4, lines 25–27)? I believe there are many other places in the manuscript where the authors could leave a little more to the imagination, thus reducing the overall length of the paper.

   Do the authors think it necessary to include *all* of the figures that appear in this first submission?

2. **Comparison with other bivariate threshold excess techniques**
   The authors make use of the bivariate GPD for modelling threshold excesses. I just wondered if they'd considered how inferences for parameters and prediction regions, as given in their paper, compare to inferences on equivalent quantities from the 'usual' approach for modelling bivariate threshold excesses, i.e. as provided by the function `fbvpot` in the `R` package `evd`? How do these approaches differ? Do the authors know how their inferences compare to that from a more general point process approach to modelling? Both a bivariate threshold excess approach and a point process approach, as applied to bivariate extremes, are explained in detail in Coles (2001), Chapter 8.

3. **Asymptotic dependence**
   The authors make us of the bivariate logistic model for threshold excesses. I wonder if they checked the assumption of asymptotic dependence here?

4. **Other issues**

   – Page 20, line 52: the authors refer to Table 4, and tell us that setting μ_1 to be zero "...has a negligible effects upon the parameter estimates for ξ_1,... and upon the maximum value of the log–likelihood...". Firstly, I think this should say "... has a

negligible effect..." *or* "... has negligible effects...". Secondly – unless I'm missing something – I don't think the maximum values of the log–likelihoods are given in Table 4.

- Table 5: Should this table show "Fitted" and "Bootstrapped" for BGPD-3 with fixed μ_1 also, analogous to Table 4? As well as maximum values of the corresponding log–likelihoods? I think the text at the top of page 21 suggests that this table should show more than it does.
- On page 10, line 27, the authors tell us that it "... can be difficult to transform back onto the original scale". I have certainly seen this done before, so why might this be difficult? Could a reference be given here?
- In Section 5, the authors focus on four pairings: Bremerhaven–Hamburg, Bremerhaven–Hannover, Fehmarn–Hannover and Fehmarn–Bremerhaven. Why this specific choice of pairings? Perhaps there are good practical reasons, but if so, I think these should be stated.
- In Section 5.1 the authors state that, in the plots comparing empirical margins of the exceedances with the quantiles of the BGPD–0 models, there is "substantial" over–estimation of the extremely high quantiles. What, if any, criteria have been used to assess the degree of over–estimation? Is this over–estimation *really* 'substantial'?
- In Section 5.1 (page 16, lines 50–51) the authors state that the changes in parameter estimates could "... be within the (unknown) variation associated with each estimate". My immediate thought at this point was "why have they not quantified this variation? Where are the standard errors for these estimates?". On further reading, the authors tell us that the Hessian of the likelihood function is almost singular... is this why we do not have standard errors? The authors later use bootstrapping to obtain standard errors. Could bootstrap standard errors not be shown for ξ_1, ξ_2, α in Table 1?
- On page 16, line 55, the authors tell us that they "do not present equivalent results for the μ_1, μ_2 and σ_1, σ_2 parameters" as they "have no natural interpretation within the context of the BGPD model". Why is this the case? Is a bit more needed here?
- On page 17, lines 29 and 31, the authors discuss the results for Bremerhaven, and remark that, when applied to the Bremerhaven–Hamburg data, the BGPD "...shows a decrease in the 75% quantile at Bremerhaven"; when applied to the Bremerhaven–Hannover data, we have "an increase in this quantile". Have the authors got this mixed up? I think the table of results shows an *increase* in the 75% quantile using the Bremerhaven–Hamburg data ($15.8 \rightarrow 16.1$) and a *decrease* in the 75% quantile using the Bremerhaven–Hannover data ($16.1 \rightarrow 15.7$).
- It strikes me that an extension to non–linear time trends would be quite simple to implement. Have the authors considered investigating the potential benefits of allowing the parameters to vary in a non–linear way? Perhaps there was empirical evidence to suggest that linear trends would suffice?
- Various (univariate) studies have investigated the use of functional forms for extreme value model parameters to account for seasonal variation in wind speeds, for example. Have the authors considered seasonal variability here?

– The authors mention that they make use of a cluster of machines at Edinburgh University to run their tasks in parallel. I think it would be interesting to include information on how long the code took to run – for example, the simulation studies or the bootstrap procedures – and how this compares to running these tasks on a single machine.

5. Other minor issues

– In Section 2.1, should the equivalence between the GEV and GPD shape parameters ξ be made clear? Similarly, should the link between the GPD and GEV scale parameters be stated, i.e. $\tilde{\sigma} = \sigma + \xi(u - \mu)$ where $\tilde{\sigma}$ is the GPD scale parameter?

– Page 7, lines 11 and 13: "... if and only if the distribution of the maximum converges to the EVD" – should it be made clear that, in fact, it *does*? Hence the conditional distribution of the exceedances *does* converge to the GPD?

– Page 4, line 17 – there is a reference to Stephenson and Tawn (2005), but this does not appear in the reference list.

– In my copy of the paper there seems to be quite a few references in the text to Figure 7, and I don't think all of these should be for Figure 7. Could the authors double check this?

– Abstract – page 1, line 46 – should the order of the hazardous events be changed to correspond to the order of the fields of application in lines 42 and 44 – i.e. "severe financial loss, flooding or mechanical failure", in that order?

– Page 3, line 10: *Should say* "Ribatet et al. 2007".

– Page 5, lines 33–39: "...construction prediction regions" *should say* "... constructing prediction regions"; "... non–stationary BVGPD model to the German wind speed in..." *should say* "... non–stationary BVGPD models to the German wind speeds in.."; "... before concluding, in Section 7, with some useful conclusions..." – "conclude" used too often?

– Page 12, line 29: "... trends in the parameters can be easily be..." *should say* "... trends in the parameters can easily be..."

– Page 22, lines 10–12: "... than have been already been..." – the word "been" has been (!) used to many times.

– Table 2: There is nothing in the caption to explain what x and y are. I suppose it might be obvious that x corresponds to the first location in each pair and y the second, but perhaps this should be made clear?