

Interactive comment on “A Monte Carlo simulation of multivariate general Pareto distribution and its application” by L. Yao et al.

Anonymous Referee #2

Received and published: 11 February 2015

General comments:

The manuscript presents a novel method, called MGPD, to the world of Ocean Engineering. It presents some of the theory of MGPD, and gives an example how to model the extremes of wind and waves, taking their correlation into account. I find the theoretical background hard to understand, as will be the case for most of the readers of Ocean Science Discussions as well (I guess). On the other side, the theoretical background is not comprehensive enough to be self-explaining. The same holds for the wind/wave example, which does not convince me that the MGPD works as expected.

My conclusion is that the manuscript is hard to understand, and will not invite readers to use the MGPD approach in Ocean Science.

C1335

Specific comments:

p 2734, line 13: "rather small" should be avoided: either quantify the error, or remove the word "rather".

p 2736, line 13: Change into "Rootzen and Tajvidi (2006) suggest, based on the research by Tajvidi (1996), that..."

p 2736, line 16: either use the abbreviation MGPD, or write all words fully. Now, it is a mixed of full words and abbreviations.

p 2736, line 19: "more stable calculation results". One of the disadvantages of (M)POT over (M)GEV/Gumbel is that one has to choose a threshold, which may influence the fit considerably. It is thus not trivial that (M)POT results in more stable results than (M)GEV/Gumbel

p 2739, line 14: "...were not proposed..." Sentence seems rather cripple to me.

p 2741, line 8: The selection procedure is not clear to me. If a 5-day interval is chosen, the interval between two extremes cannot be (less than) 2 day, so I don't understand line 8. Additionally, is the extreme wind chosen, and the corresponding wave height at the same moment added, or is the extreme wind in a 5-day interval connected to the extreme wave height in the same 5-day interval? Please clarify.

p 2741, line 19. Add "probability plot of THE marginal distributionS." p 2741, Eq. 13. Why is the Pearson Type III chosen to model the return levels, and not the GEV distribution (which is mentioned several times to be the natural distribution for extremes)?

p 2742, line 1. The motivation for the choice of the bivariate logistic generalized Pareto distribution is missing.

p 2742, line 16: "The simulation results are in agreement with the actual situation". I disagree with the conclusion, as Figure 4 shows that the observations show much more 40m/s events than the stochastic simulation indicates. Additionally, these 40m/s

C1336

events correspond with much lower wave heights than the simulation indicates.

p 2743, line 11. Please add the return levels to the table as well (p 2747, table 2)

p 2744, line 10. I disagree with the conclusion that, by Monte Carlo simulation, a better estimate can be made than based on the traditional extreme-value methods, as no extra information is added.

p 2747, Table 2&3. I think that addition of $V=30$ and $V=40$ (instead of $V=1$ and $V=2$) makes the table more interesting. I suspect that especially at these high levels, the difference between the analytic solution and the simulation increases. If not, it would be worthwhile to intercompare the direct observation with the simulation as well.

p 2750, Figure 1. Add a note why the empirical distributions are discrete. I assume that the observations are stored with a limited amount of digits.

p 2750, Figure 2. Please add at the upper X-axis the return level values.

p 2750, Figure 2. It seems to me that the wind distribution shows a 'kink' around 20m/s, which might be attributed to a different phenomenon that causes the wind: normal depressions for lower winds, typhoons for higher winds. If so, this violates the assumption of EVT that all extremes originates from a single parent distribution.

Interactive comment on Ocean Sci. Discuss., 11, 2733, 2014.