Ocean Science

Discussions

Open Access

Interactive comment on "Evaluation of Peaks-Over-Threshold Method" *by* Soheil Saeed Far and Ahmad Khairi Abd. Wahab

Anonymous Referee #2

Received and published: 15 August 2016

General evaluation

Although the manuscript addresses an interesting topic, it has some problems and as such I do not recommend its publication.

Comments

1) The authors claim that the GPD and the POT are two different models. The GPD plays a crucial role in extreme value theory as the distribution of the sample of excesses above a sufficiently high threshold, method known as the peaks-over-threshold (POT) - see, for instance, Davison and Smith (1990), Pickands (1975), Embrechts et al (1997). The POT and the GPD are thus closely connected.

I do not understand the meaning of "to select the best fitting distribution for the dataset"

(page 1, line 23). The issue is to fit a GPD distribution to the excess (or exceedance) data. Obviously, the Generalized Extreme Value (GEV) or the Gumbel, Fréchet and Weibull are candidate distributions if the Annual Maxima Method is considered. For using this method the data has to exhibit a temporal structure (and obviously there is no threshold involved).

(Some references of using the POT in waves context - Teena et al (2012), Thevasiyani and Perera (2014) and Cañellas (2007))

2) The authors, in page 6, present the hybrid method that starts with the choice of the threshold through the mean excess (ME) plot. In applications, the choice of the threshold just by looking at the ME plot is generally very difficult if not impossible (in some cases). Due to the practical difficulties in choosing the threshold, and considering the importance of this step in the method, some complementary approaches could have been mentioned (see, for instance, Beirlant et al (2004)). In page 7-line 11, the authors say that Weibull distributions were fitted to the wave data. Later on, the same is done for the Gumbel and for the Fréchet distributions. Due to the fact that "... the exceeded data over a certain threshold are employed in the analysis" (page 9, line 10) then the GPD should have been used instead because it is the proper distribution for modelling the excesses above the threshold, as was stated in my previous comment.

3) The data lacks a convenient explanation. The authors seem to consider wave heights recorded at time intervals of at least three days apart (page 2, lines 28-29). Then, the authors state that they do linear interpolation "to fill in the large gaps" (line 29, page 2). I wonder why that is needed. The exceedances are supposed to be the focus. Additionally, it would have been nice to see the plot of the data which was analysed.

Technical correction

Expression (4) is not consistent with the parameterization of the GPD the authors indicated in (1) – see Coles (2001).

References

Beirlant J, Goegebeur Y, Segers J, Teugels J (2004). Statistics of Extremes – Theory and Applications. Wyley, Chichester.

Davison AC and Smith RL (1990). Models for exceedances over high thresholds. J Roy Stat Soc B 52:393-442.

Embrechts P, Klüppelberg C, Mikosch, T (1997). Modelling Extremal Events for Insurance and Finance. Springer-Verlag Berlin Heidelberg.

Pickands III, J (1975). Statistical inference using extreme order statistics. Ann Stat 3:119-131. Teena, NV, Kumar S, Sudheesh K, Sajeev R (2012) Statistical analysis on extreme wave height. Natural Hazards, Volume 64, Issue 1, pp 223-236.

Thevasiyani T, Perera, K (2014) Statistical analysis of extreme ocean waves in Galle, Sri Lanka. Weather and Climate Extremes, Volumes 5–6, October 2014, Pages 40–47.

Cañellas B, Orfila A, Méndez FJ, Menéndez M, and Tintoré J (2007) Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). Journal of Coastal Research, Special Issue 50.

Printer-friendly version

Discussion paper