Nat. Hazards Earth Syst. Sci. Discuss.,
doi:10.5194/nhess-2016-240-RC2, 2016

Natural Hazards and Earth System Sciences
Open Access
Discussions

Interactive comment on "Efficient Bootstrap Estimates for Tail Statistics" *by* Øyvind Breivik and Ole Johan Aarnes

S. Caires (Referee)

sofia.caires@deltares.nl

Received and published: 29 December 2016

General comments:

The authors show that in the non-parametric bootstrap procedure for obtaining confidence intervals of estimates based on the k largest values in a sample, the computations can be carried out in a more computationally efficient way by drawing bootstrap samples from the K0 (K0>k) largest values of the sample rather than from the entire sample. They propose that K0 be fixed at a value leading to a very low probability of drawing fewer than the required k largest entries of the sample and provide the expression of that probability.

The article is concise and well-written. The suggested approach appears to be useful for applications such as those considered in examples 1 and 2 (empirical percentile).

Printer-friendly version

Discussion paper

However, I have doubts about the correctness of the non-parametric bootstrap procedure for obtaining confidence intervals of GPD return value estimates as described in Example 3.

I have two major comments that I would like the authors to address or at least consider that they, despite not being covered by the article, should also be taken into consideration when bootstrapping to obtain confidence intervals estimates related to extremes.

Major comments:

1. I was not aware of the idea of the bootstrap being applied to the entire dataset rather than to a sample of cluster peaks as in the computation of confidence intervals of Example 3. In the usual form of the parametric bootstrap one does not return to the entire sample, but considers the (much smaller) sample to which the GPD was fitted. In any case, ensuring that the coverage rates - the percentage of times that a confidence interval really contains the true parameter in (hypothetical) repetitions of the same sampling and estimation process - of bootstrap confidence intervals are sufficiently correct has, in my view, priority over the computational efficiency of those intervals. Both Coles and Simiu (2003, J. Engrg. Mech., 129 (11), 1288-1294) and Schendel and Thongwichian (2017, Adv. Water Resour., 99, 53-59, http://dx.doi.org/10.1016/j.advwatres.2016.11.011) consider the shortcomings of bootstrap intervals with respect to coverage, the first paper offering an ad hoc solution and the second suggesting the use of Test Inversion Bootstrap. I wonder if the authors could add information to the article about the coverage rates of their confidence intervals.

2. The results shown in figures 3, 4 and 5 are based on M=10,000 bootstrap replications, while those shown in Figure 8 are based on M=1,000. I wonder if the authors could say something about how M should be chosen. According to Efron and Tibshirani (1993, Monographs on Statistics & Applied Probability 57), 200 bootstrap replications are usually enough for obtaining reasonable estimates of the standard error. Could optimizing the number of bootstrap replications be a possible solution to some of the

computational problems pointed out by the authors?

Specific comments:

âĂć Page 1, Line 3: ‘confidence intervals . . . can be estimated’. I would replace ‘estimated’ with ‘obtained’ everywhere, since the intervals are random variables and not parameters.

âĂć Page 1, Line 13: In the light of my Major Comment 1, I would not say that “This is a straightforward procedure”; it is not the computational or algorithmic aspects of a method that matter most, but its validity.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-240, 2016.