Nat. Hazards Earth Syst. Sci. Discuss.,
doi:10.5194/nhess-2016-240-AC1, 2017

Natural Hazards and Earth System Sciences
Discussions

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

Øyvind Breivik and Ole Johan Aarnes

oyvind.breivik@met.no

Received and published: 17 January 2017

Response to Reviewer 1, R Katz:

Thank you for your thorough review. Our responses are in italics below. Please note that there is considerable overlap with the comments from the other reviewer, and we refer to our separate response to her). We propose to include a new Fig 6 which looks at the CIs as a function of number of bootstrap resamples. We will also include a new paragraph to the Discussion. This is outlined in our reply to Reviewer 2.

How to Bootstrap Extremes if You Must GENERAL COMMENTS: The focus of the manuscript is on efficient use of the bootstrap, a resampling tech- nique, to quantify uncertainty (e.g., in the form of a confidence interval) in estimated extreme statistics such as return levels. Justification is provided for a simplified boot- strap procedure

in which the resamples are generated through only drawing from the highest values in the original sample, not the entire sample. This common sense result is consistent with conventional statistical modeling of extremes, with the common as- sumption that the uncertainty in estimating the rate of exceedance of a high threshold can be ignored (e.g., Chapter 4 in Coles, 2001). Perhaps the present paper serves to place this conventional approach on firmer footing.

Nevertheless, there are a number of alternative techniques for uncertainty quantification in extreme value analysis not even mentioned in the manuscript. These alternatives include different implementations of the bootstrap, as well as ones in which no resampling need be performed (e.g., profile likelihood technique; Coles, 2001). At the least, these alternatives should be mentioned.

We agree, and we will include a section where we go through the various alternatives to non-parametric bootstrapping. However, we will argue that this is somewhat beside the point of the article as our main objective has been to investigate how tail statistics can be bootstrapped, if, as the referee says, you must. We do not necessarily argue that non-parametric bootstrapping is the best alternative, and we will make clearer where we think it is appropriate to use (see also our reply to Reviewer 2).

For this reason, I recommend that the manuscript be accepted for publication subject to minor revision. SPECIFIC COMMENTS: (1) Nonparametric versus parametric bootstrap A nonparametric bootstrap is used in which the resamples are created by drawing with replacement from the original sample. When fitting extreme value distributions (e.g., the generalized Pareto in Sec. 3.3), it has been suggested that a parametric boot- strap would be preferable for constructing confidence intervals for return levels (i.e., resamples are created by Monte Carlo simulation from the fitted distribution) (Kysely, 2008).

We agree that, especially for small samples, parametric bootstraps are probably capable of better coverage than non-parametric bootstrap techniques. However, we are

looking at very large samples, indeed at samples that are so large that we can perform in some cases in-sample estimates of 100-year return values. The paper now acknowledges the limitations of non-parametric bootstraps, and we stress that it should be seen as a study of how to efficiently handle the original data set if, as the reviewers points out, you wish to perform a non-parametric bootstrap. We will include a paragraph in the Discussion where we look at the caveats to using non-parametric bootstraps (see our reply to Reviewer 2).

(2) Refined bootstrap techniques Bootstrap-based confidence intervals can be too short, especially for return levels with long return periods. Consequently, alternative more involved bootstrap techniques (e.g., the so-called "test inversion" bootstrap) have been proposed to improve the per- formance of such confidence intervals (Schendel and Thongwichian, 2015).

This is an interesting technique, and Reviewer 2 also refers to a follow-up paper by the same authors. We will include a short paragraph in the Discussion where we outline this alternative method (see our reply to Reviewer 2).

(3) Alternatives to bootstrap When estimating the parameters of an extreme value distribution by maximum likeli- hood, an alternative technique for obtaining confidence intervals for return levels is pro- file likelihood (Coles, 2001). This technique does not require any resampling, but does require repeated fits of the extreme value distribution under parameter constraints. It is competitive with resampling for obtaining confidence intervals of return levels (e.g., Schendel and Thongwichian, 2015).

The profile likelihood technique is a well-known technique, but it falls outside the scope of this paper to investigate it as we focus strictly on efficient methods for non-parametric bootstrapping.

REFERENCES: Kysely, J.: A cautionary note on the use of nonparametric bootstrap for estimating un- certainties in extreme-value models, Journal of Applied Meteorology and Climatology, 47, 3236–3251, 2008. Schendel, T. and Thongwichian, R.: Flood

frequency analysis: Confidence interval es- timation by test inversion bootstrapping, Advances in Water Resources, 83, 1–9, 2015.

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