Hydrology and Earth System Sciences

Discussions

HESSD

Interactive
comment

Interactive comment on "Technical Note: Approximate Bayesian Computation to improve long-return flood estimates using historical data" *by* Adam Griffin et al.

Anonymous Referee #2

Received and published: 4 October 2018

The note proposes to use Approximate Bayesian Computations for flood frequency analysis and illustrates its application on the Severn river at Shrewsbury (UK). The idea may at first sight appear attractive but the content of the manuscript is questionable from several points of view:

1) It is first much too focused on flood frequency analysis practices in the UK. It should aim at being more general to be really useful for a large readership. The ABC method is for instance motivated by difficulties encountered when applying the classical maximum likelihood method for the calibration of generalized Logistic distributions (GLD). To my knowledge other types of distributions (Pearson III, generalized extreme value) are

preferred to the GLD distribution in many other countries. Are they affected by the same estimation problems ?

2) The main motivation for the use of the ABC method expressed on page 2 ("the maximum likelihood estimator may fail to converge for type I GLD) does not exactly correspond to the conclusions of the cited reference paper of Shao (2002). This paper explains the problem that can be encountered, but proposes also some simple solutions that can be easily implemented : include the so-called embedded distributions (Gumbel and 2-parameter reciprocal exponential distribution) in the tested solutions.

3) The implementation of the ABC method requires a measure of distance between the tested distribution defined by a set of parameters and the sample. This is mentioned, in a relatively confused way, on page 3 and 4, but the formula of this distance measure is never given. The reader understands from the abstract that this distance is based on probability weighted moments but that's all. The distance measure should be provided. Likewise, a justification for the 5% threshold should be provided somewhere, and maybe some kind of sensitivity analysis to this ad-hoc or arbitrary choice which determines the width of the "credibiliy intervals" which are computed (fig. 3 and 4).

4) The description of the ABC algorithm on top of page 4, recalls me the founding works of Hornberger and Spear (1980) on inference uncertainties. This method inspired K. Beven who developed later on the Glue method. Many researcher, and I am one of them, consider know that Bayesian inference and Monte Carlo Markov Chain algorithms introduced in hydrology 10 to 20 years ago, provide a rigorous and consistent method for inference problems, and especially for frequency analyses (see the abundant recent literature presenting implementations of Bayesian MCMC algorithms in flood frequency analysis). I have the feeling that the authors propose, under a knew name, the use of a 40-years outmoded method.

5) The cited reference list for flood frequency analysis is not up to date. The recent literature on Bayesian-MCMC flood frequency analysis is almost ignored by the note.

6) The reconstructed historical record series seems inconsistent with the recent measured series (figure 2). Most of the historic records lie beyond the adjusted rating curve for the present time (figure 1). The largest observed value over the recent 30 years period appears to have been exceeded 15 times over the past 170 years. There is a 6 percent chance for such an event to occur $(1-15/170)^{\wedge}30$. It is essential that historical records are thoroughly criticized before any implementation of statistical inference. Ideally, estimation uncertainties of historical flood discharges – that may be large, should be taken into account in the statistical inference and this is totally possible (see for instance Payrastre et al., WRR, 2011). The presented case study and implementation example is questionable and does not rely on the state of the art best practices. This is highly problematic since the note seeks to be exemplary.

7) The note is based on one single example which is probably not sufficient to demonstrate the usefulness, pertinence and robustness of the proposed method. The computed credibility intervals with two different methods appear close to one another on figure 4. But this could be the result of chance or of a calibration of the ABC threshold by the authors. It does therefore not demonstrate that the ABC credibility intervals do make sense: i.e. that they do accurately reflect the inference uncertainties. In fact, I highly suspect that they do not (see point 8). Moreover, is the proposed case study a problematic type I GLD case that would illustrate the benefit of the ABC approach? Apparently not since a MLE estimate is provided. The authors should illustrate that the non-convergence of MLE algorithms, which is the justification for the introduction of the ABC approach, is at least sometimes observed in real world.

8) The maximum likelihood confidence intervals in figure 4 are base on the so-called asymptotic Gaussian approximation which is clearly not suited (see the unrealistic decreasing lower bound on figure 4 left side). The authors should use the Bayesian MCMC algorithm, to computed credibility intervals (see for instance the BayesianMCMC command in the nsRFA package of the R software). The satisfactory outcome is the MLE estimates and the ABC posterior mean are close to one another. By the way,

the authors should explain why they selected the mean value rather than the parameter set and distribution that provided the best fit to the data, which would have been a more evident choice for the ABC algorithm. Is it because the results would have appeared less satisfactory?

9) The intervals computed with the ABC method are totally dependent on the selected threshold value. This dependence must be illustrated through a sensitivity analysis for instance as well as through the implementation of the method on various case studies. It must be acknowledged by the authors and it must be clear that no uncertainty level or probability can be attributed to these intervals. The figures and terms used (95% credibility interval is inadequate) are ambiguous and misleading.

―――――――――――――――