Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-594-RC1, 2021

Hydrology and Earth System Sciences Discussions

Open Access

HESSD

Interactive comment

Interactive comment on "Design flood estimation for global river networks based on machine learning models" *by* Gang Zhao et al.

Eric Gaume (Referee)

eric.gaume@ifsttar.fr

Received and published: 19 January 2021

The manuscript presents and extension of a work published in 2015 by Smith et al. and aiming attesting methods suited for design flood estimation at global scale. The article is based on the analysis of the very rich international streamflow database GSIM. Numerous other international datasets are processed to derive climatological, physiographical and hydrological descriptors for each of the considered 12000 watersheds worldwide. Three different regression methods are tested to relate the locally estimated annual maximum discharge quantiles and the watershed descriptors, namely a power function, a support vector and a random forest model. A split sample test is used to assess the performances (bias, root mean square error) of the various implemented and tested approach. All methods are described in the manuscript. The manuscript

Printer-friendly version

Discussion paper

is overall of high quality, comprehensive and based on the best available datasets and methods. It deserves without doubt a publication in HESS. Its content could be slightly improved in several ways.

1) Some figures appear complex and difficult to understand without the explanations provided in the text. The legends and captions could be improved and enriched (see comments in the attached pdf)

2) The authors made an important effort to provide to the readers all the necessary mathematical background and equations. But notations in the equations and the indices are not consistent throughout the manuscript introducing sometimes some confusion. This could be corrected (see attached document).

3) The authors use a Bayesian MCMC inference framework to derive local discharge quantile estimates but they only make use, in fact, of the most-probable estimated value (i.e. the maximum likelihood estimate if a non-informative prior is used which is what I suspect). They do not valuate the credibility intervals for the estimated quantiles in the analysis or discussion of the manuscript. In fact, a maximum likelihood local estimation is used, even if it is through a Bayesian MCMC framework. This should be clearly stated in the manuscript and the first paragraph of section 3.2.2 has to be reformulated accordingly (see comments in the pdf).

4) The proposed method is based on a clustering approach and regressions conducted on each cluster separately. But the clusters are optimized based on the SVM regression method. It is then not really surprising that the SVM method provides the best performances if compared to the two other tested method and especially to the RF. The comparison between the approaches is not totally fair. This should be mentioned in the discussion.

5) Some stations are discarded from the analysis after clustering based on a discordancy test (first paragraph of section 4.1). This may be problematic from a methodological and a statistical point of view. First, eliminating data based on a threshold

from a statistical test is questionable from a statistical point of view and may be seen as an over-interpretation of the results of the statistical tests. At a significance level of 5%, one would expect the p-value to be exceeded for 5% of the available sample on average: i.e. if the sample is homogeneous, the p-value will be exceeded 5% of the time. Does it then make sense to eliminate these 5% of stations from the sample ? Second, the method is developed to be applied on ungauged watershed. Each ungauged site will be affected to one of the defined clusters, but it will not be possible to verify that the specific site is not discordant with the rest of the cluster. It is to be foreseen that the proportion of “discordant” sites will be equivalent or even higher for the ungauged sites than for the gauged sites (the clusters were adjusted on the gauged sites). The “discordant” sites may be discarded in the calibration process of the regression method but should be included in the validation dataset to provide a proper estimate of the performances of the proposed approach: i.e. an estimate of their performance if implemented in real-life applications.

6) Some sentences appeared a little simplistic to me (see attached commented pdf). Some nuance should probably be introduced at several places.

With the hope that this review will help the authors to improve their manuscript and looking forward to seeing this interesting paper being published in the near future in HESS.

Please also note the supplement to this comment:
https://hess.copernicus.org/preprints/hess-2020-594/hess-2020-594-RC1-supplement.pdf

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-594, 2020.