STAHY2017-13
STAHY 2017 workshop,
21–22 September 2017, Warsaw, Poland

Bias of Maximum Likelihood Estimates of Flood Quantiles Based Completely on Historical Flood Records.

Krzysztof Kochanek (1), Witold G. Strupczewski (1), and Ewa Bogdanowicz (2)

(1) Institute of Geophysics, Polish Academy of Sciences, Księcia Janusza 64, 01-452 Warsaw, Poland, (2) Polish Society of Geophysics, Podleśna 61, 01-673 Warsaw, Poland

To improve the accuracy of the Flood Frequency Analysis the systematic series of the annual flow peaks are often completed by historical (pre-instrumental) floods.
These two samples, i.e. complete systematic dataset and pre-instrumental heavy censored sample, represent the same but unknown population. Therefore, they are jointly used for selection of proper distribution model and estimation of its parameters and quantiles of interest.
To examine the impact of historical and paleo-flood data on bias of design quantile we analyse the case where only the historical data are available (this situation is characteristic for e.g. ungauged rivers). So, we deal with a lower-bounded censored sample (type 2 censoring) of size M having known its k largest elements. As two-parameter models, namely Gumbel, Log-Gumbel and Weibull, are considered, the k = 2 is the smallest number that allow estimation of parameters. Both the True and False hypothesis are considered.
The investigation is focused on the maximum likelihood (ML) estimation bias (and also, to the extend, mean square error) of 100-years flood. The calculation was made for k = 2, 3,. . . M and for the range from k/M = 2/M up to 1 for various values of M and coefficient of variation (CV). While the quantile or the least squares estimation methods give for synthetic censored samples unbiased estimates of parameters and hence, quantiles, the ML method behaves differently and produces relative bias of flood parameters and quantiles in particular one with return period 100 years.
For the True case the lowest absolute value (but non-zero) of the ML bias for 100-year flood estimator is obtained when M = 100, regardless the k and CV. Also for each k/M ratio the quantile bias is relatively small (a few percent at maximum) and tends to zero when M tends to infinity. It is interesting, that the bias quickly falls to zero with the growth of the number of known historical floods (k) for each M except M = 100. The False case involves also the model bias, which may be of different sign and compensate the value of sampling bias induced by the method of estimation. Peculiarity of the latter case stems from the fact that it was not always possible to predict (in qualitative terms) the results of calculations and thus they required diagnostic pre-investigation.