

Flood Frequency Analysis supported by the largest historical flood.

W. G. Strupczewski¹, K. Kochanek¹ & E. Bogdanowicz²

1) Institute of Geophysics Polish Academy of Sciences, Ksiecia Janusza 64; 01-452 Warsaw, Poland,
e-mails: wgs@igf.edu.pl, kochanek@igf.edu.pl

2) Institute of Meteorology and Water Management, Podlesna 61, 01-673 Warsaw, Poland,
e-mail: ewa.bogdanowicz@imgw.pl

ABSTRACT

The use of non-systematic flood data for statistical purposes depends on reliability of assessment both flood magnitudes and their return period. The earliest known extreme flood year is usually the beginning of the historical record. Even if one properly assess the magnitudes of historic floods, the problem of their return periods remains unsolved. The matter in hand is that only the largest flood (XM) is known during whole historical period and its occurrence marks the beginning of the historical period and defines its length (L). It is common practice to use the earliest known flood year as the beginning of the record. It means that the L value selected is an empirical estimate of the lower bound on the effective historical length M . The estimation of the return period of XM based on its occurrence (L), i.e. $\hat{M} = L$, gives severe upward bias. The problem arises that to estimate the time period (M) representative of the largest observed flood XM .

From the discrete uniform distribution with support $1, 2, \dots, M$ of the probability of the L position of XM one gets $\hat{L} = M/2$. Therefore $\hat{M} = 2L$ has been taken as the return period of XM and as the effective historical record length as well this time. As in the systematic period (N) all its elements are smaller than XM , one can get $\hat{M} = 2(L + N)$.

The efficiency of using the largest historical flood (XM) for large quantile estimation (i.e. one with return period $T = 100$ years) has been assessed using the ML method with various length of systematic record (N) and various estimates of historical period length \hat{M} comparing accuracy with the case when systematic records alone (N) are used only. The simulation procedure used for the purpose incorporates N systematic record and one largest historic flood (XM_i) in the period M which appeared in the L_i year backward from the end of historical period. The simulation results for selected distributions, values of their parameters, different N and M values are presented in terms of bias and RMSE of the quantile of interest are more widely discussed.

Keywords: Flood frequency analysis; Historical information; Error analysis, Maximum Likelihood, Monte Carlo simulations.

1. INTRODUCTION

Flood engineering usually deals with the determination of the flood of a given return period T years, i.e. the flood quantile X_T or the design flood. The problems with the assessment of these parameters result from short time series ($N < T$), unknown probability distribution function of annual peaks, error corrupted data, the simplifying assumptions as of identical independently distributed (i.i.d.) data and, in particular, the assumption of stationarity of relatively long data series. All these account for high uncertainty of the upper quantile estimate. The effect of sample size is widely documented for various distribution models and estimation methods, thus, it is obvious that due to a short sample the confidence interval of the design flood estimate is already very broad. In addition to Flood Frequency Analysis (FFA) other sources of error would result in increasing uncertainty in the design flood estimate. This feature is not appreciated by the designers as they want to have only one value for designing flood related structures. Conversely, efforts to improve the accuracy of estimates of the hydrologic design value by specifying the various sources of uncertainty and incorporating them in the analysis produce the opposite effect from the one intended.

To improve the accuracy of estimates of upper quantiles all possible sources of additional information and 'statistical tricks' are used, such as: independent peaks above the threshold, seasonal approach, regional analysis, record augmentation by correlation with longer nearby records and, finally, augmentation of the systematic records by historical and paleo-flood data.

Frequency analysis of flood data arising from systematic, historical, and paleo-flood records has been proposed by several investigators (a review *Stedinger and Baker, 1987, Frances et al. 1994, MacDonald, 2013*). The use of non-systematic flood data for statistical purposes depends on reliability of assessment both flood magnitudes and their return period. If the historical record is available, the information about the floods larger than prevailing majority of floods reported in the systematic record can be introduced to the datasets and, if we

53 are lucky, the unique information about the largest reported floods. Serious difficulties relate to the
54 (un)availability and (not-) exhaustiveness of historical information, the (low) quality and (in)accuracy of
55 historical sources. As if it was not enough, depending on the number of parameters and their method of
56 estimation, the estimates of high quantiles are more or less sensitive to the largest observed floods.

57 The earliest and simplest procedures for employing historical and paleo-flood data were based on plotting
58 positions and graphical concepts (Zhang, 1982, 1985, Bernieur et al., 1986, Wang and Adams, 1984, Hirsch,
59 1985, Cohn, 1986). The PWM method and L -moment method were introduced by Ding & Yang (1988), Wang
60 (1990, 1996) and Hosking (1995). To deal with historical and paleo-floods Hosking and Wallis (1986 a, b)
61 applied the maximum likelihood (ML) as the estimation method. Recently the Bayesian estimation paradigm has
62 been incorporated (Vigilione et al., 2013, Parent and Bernier, 2003, Reis and Stedinger, 2005). It takes into
63 account that the historical floods are known with uncertainty, for instance with lower and upper bounds (in fact
64 the effect of corrupted historical flood magnitudes was investigated by Hosking and Wallis via MLE mentioned
65 as early as in 1986 a, b) The subject of historical floods currently constitutes one of the main scientific threads in
66 flood frequency analysis (MacDonald, 2013, Payraastre et al., 2011, 2013). **It is important to add, that the
67 inclusion of historical information is recommended in a number of national and international policy documents
68 e.g. EU Flood Directive.** The log Gumbel, Weibull and Gamma distributions together with maximum likelihood
69 method were considered by Frances et al. (1994) to tackle systematic and historical or paleo-flood data in FFA.
70 To assess the potential statistical derived from historical information the asymptotic variances of the quantile
71 estimates from the systematic records alone and the combined time-series were compared by means of computer
72 simulation experiments. The study performed to define the length (M) of historical period indicate that value of
73 historical data for estimating flood quantiles can vary depending on only three factors: the relative magnitudes of
74 the length of the systematic record (N) and the length of the historical period (M); the return period (T) of the
75 flood quantile of interest; and the probability threshold defining the historical floods.

76 Most often it is the first historical large flood that is the most remembered (and described in historical
77 sources) and, therefore, it is usually not considered as important (or simply not known) what had happened
78 before (Girguś and Strupczewski, 1965). **In other words, only the largest (paleo-)historical flood is usually
79 known for either it was best remembered (and thus recorded) because of its destructive character and taking a
80 toll on many lives or, in case of pre-historical time, the largest inundation swept away any evidence of smaller
81 floods that occurred earlier.** The date of the first recorded historical flood is taken as the historical memory
82 length L , i.e. L becomes the duration of non-systematic period commencing on the large flood. Even if one
83 properly assess the magnitudes of historic floods, the problem of their return periods remains unsolved. In most
84 literature examples (specially Benson, 1950, Dalrymple, 1960, IACWD, 1982, Zhang, 1982 and NERC, 1975,
85 p.177) one reads that effective length of historical record M used for frequency analysis is always taken to be the
86 period from the first extraordinary flood to the beginning of the systematic record, i.e. L .

87 The matter in hand is that only the largest flood (XM) is known during the whole historical period and its
88 occurrence marks the beginning of the historical period and defines its length (L) (Fig.1). That is because the
89 beginning of the historical period was somehow forced by the appearance of the largest flood (XM) but in fact its
90 unusual magnitude corresponds rather to a longer return period than L (or, if in systematic record all
91 observations are smaller than XM , to $(L+N)$ -period), i.e. the probability that the actual return period of XM is
92 longer than the L is greater than fifty percent.

93 Attempts to eliminate or lessen this error lead us to the estimation the time period (M) representative of the
94 largest observed flood XM as accurately as possible. In order to do so, we will carry out the evaluation of the
95 efficiency of using the largest historical flood (XM) for large quantile estimation and its comparison with the
96 case when systematic records alone (N) are used. To keep and preserve the unspoiled genuine information
97 contained in the observation (XM, L), the return period (\hat{M}) of the largest observed historical flood (XM) should
98 be assessed without data from the systematic record **providing that it does not contain elements larger than XM
99 values.**

100 It is obvious that the return period of the historical flood assessed on the base of the year of occurrence (L)
101 represents just the lower limit of its real empirical return period (M). Of course, there is an upper **empirical**
102 limit as well, which however, can not be estimated unambiguously. This is so because, if the occurrence of a large
103 flood was reported in a given year, for sure a similar or more serious flood a year before would have been also
104 noted and commented in historical sources (Hirsch and Stedinger, 1987). The same can be stated for horizon of
105 two, three, four, etc. years. If we could identify this time span, we would have determined the upper limit of the
106 empirical return period.

107 The estimation of M based on the date of the first extraordinary flood occurrence exacerbates an already
108 severe imprecision. By defining as historical floods all floods during the M period above a given threshold and
109 taking four different plotting position formulas, Hirsch and Stedinger (1987) calculated (with the use of Monte
110 Carlo experiment) the magnitude of the upward bias of the plotting position of the largest sample elements
111 occurring when L is taken as the beginning of the historical record. Doing so they noticed that L is a random

112 variable dependent on the flood-producing process itself; this would be a violation of the assumption of the
113 plotting position formulas.

114 Similarly, *Hosking and Wallis* (1986 a, b) use Monte Carlo (MC) computer simulation to assess whether a
115 single paleo-flood estimate, when included in a single-site Maximum Likelihood (ML) flood frequency analysis
116 procedures, gives a worthwhile increase in the accuracy of estimates of extreme floods. They found that the main
117 factors affecting the utility of this kind of paleological information are the specification of the fitted flood
118 frequency (whether it has two or three unknown parameters) and the size of the measurement error of paleo-
119 discharge estimates. Errors in estimating the date of the paleo-flood are considered to be of minor importance.
120 For distributions with higher CV or skewness the difference between the effects of the errors of the magnitude of
121 paleo-flood and its return period is smaller.

122 Note that the randomness of the systematic records time series of i.i.d. variable can also be sometimes
123 questioned and undermined, e.g. when the largest value XM of a time series intentionally terminates the N -
124 elements' systematic record. Then the XM is the last element of the N -element time-series. Such a case may arise
125 when a water gauge was swept away by a heavy flood (XM) and not restored, or the intentional movement of the
126 hydrological station. As before, the use of such a series in FFA with $\hat{M} = N$ will lead to an overestimation of
127 large quantiles.

128 2. PROBLEM FORMULATION

129 The object of the paper is to assess by use of the maximum likelihood (ML) method whether there is any impact
130 of the largest flood terminating the time series assuming its magnitude (XM) is known. Therefore, the case of
131 systematic data completed by largest flood is compared with the case where records contain systematic data
132 only. These two variants are examined by comparing the bias (B) and the root mean square error ($RMSE$) of
133 flood quantiles. The two two-parameter distributions, namely Gumbel and Weibull were used when applying the
134 simulation experiments. The emphasis is put on the effect of misspecification of the return period (M) of the
135 largest historical (paleo-) flood (XM) and on the proper assessment of the M estimate on the basis of XM
136 occurrence (L). So far, the results of such research has not been presented in the hydrological literature.

137 The theoretical framework of our research is based on Maximum Likelihood estimation which has been
138 generally found to have desirable properties for combine systematic and historical information (*Frances et al.*,
139 1994, *Stedinger and Cohn*, 1986, *Naulet et al.*, 2005). It is assumed that the annual maximum floods are
140 independent and identically distributed.

141 Assessment of the return period M of the XM flood

142 *Hirsch and Stedinger* (1987) considered that the time of occurrence of the earliest documented historical flood L
143 is the random variable defining a lower bound of the sample size used for computation of plotting positions. The
144 position L of the largest in M period element (XM) (Fig. 1) is the random variable being discretely uniformly
145 distributed in the M period, i.e. $p_t = 1/M$ for $t = 1, 2, \dots, M$. Obviously the magnitude of the largest element (XM) is
146 also a random variable. It can correspond in the population to a smaller or larger value of the exceedance
147 probability than $1/M$ defining the effective return period (M_R) of XM , Therefore the difference ($M_R - L$) is not
148 restricted in sign.

149 Assume that the return interval (M) of XM is known. As L is uniformly distributed variable in the M length
150 time series with support $L \in [0, 1, \dots, M]$, one gets $E(L) = M/2$ and $V(L) = M^2/12$. In reality M is not known and its
151 assessment is our goal. Taking the observed L value as the estimate of the expecting value, i.e. $L = E(L)$ we get
152 the M estimate equal $\hat{M} = 2L$. Because regardless of the estimation method the quantile estimators are not in
153 general linear function of \hat{M} , the minimum bias of quantile $B(\hat{x}_p) = E[\hat{x}_p(\hat{M}) - x_p]$ does not necessarily
154 correspond to the zero-bias of \hat{M} , i.e. to $\hat{M} = 2L$. If in the systematic period (N) all its elements are smaller
155 than XM , one can get $\hat{M} = 2(L + N)$. Note that usually $N \ll L$.

156 3. SIMULATION PROCEDURE

157 The simulation procedure incorporates N systematic record and one largest historic flood (XM) in the period M
158 which appeared in the L year backward from the end of historical period (Fig. 1). Obviously, the systematic
159 record and both magnitude (XM) and year of occurrence (L) randomly vary from simulation to simulation. As an
160 estimate of the length of the historical period shall be successively $\hat{M} = L, 2L$ and the actual value $\hat{M} = M$, i.e.
161 the length of the period M in simulation experiment.

162 First, generate a gauged record x_1, x_2, \dots, x_N of independent random variates from the assumed (two-
 163 parameter) flood-like distribution $[F(x)]$ with parameters chosen to give specified values of CV . Then generate
 164 historical series of the same distribution of the length M , i.e. y_1, y_2, \dots, y_M , and find the maximum event (XM) of
 165 the historical series denoting the time (L) of its occurrence. Since the random variables (XM) and L are mutually
 166 independent the XM can be generated from the distribution of the largest element in a M -element series, i.e.
 167 $F(M) = F_{1,M}(y) = F^M(y)$, while the corresponding time of its occurrence (L) from the discrete uniform distribution
 168 with support $\{1, 2, \dots, M\}$.

169 A flood frequency distribution fitted by the method of maximum likelihood has a distribution function $F(x, \theta)$
 170 and a density function $f(x, \theta)$, where θ is a vector of unknown parameters, then the likelihood function (L) is
 171 taken to be

$$172 \quad \mathbf{L}(\theta; x, y) = F_x^{\hat{M}-1}(y = XM; \theta) \cdot f_x(y = XM; \theta) \cdot \left\{ \prod_{i=1}^N f_x(x_i; \theta) \right\}, \quad (1)$$

173 i.e., the use of incomplete data likelihood, where $\hat{M} = L, 2L$ and M , and for systematic record only

$$174 \quad \mathbf{L}(\theta; x) = \prod_{i=1}^N f_x(x_i; \theta). \quad (2)$$

175 Calculate quantile estimates $\hat{X}_T = F^{-1}(1-1/T; \hat{\theta})$ for $\hat{M} = L, 2L$ and M and the systematic record (N) only

176 (i.e. when $\hat{M} = 0$), where F^{-1} is the inverse distribution function of the fitted flood frequency distribution, $\hat{\theta}$ is
 177 the maximum likelihood estimate of θ , and T is the return period of interest.

178 Repeat the above steps a large number of times (i) and calculate the mean and variance of \hat{X}_T , and hence the
 179 relative bias RB and relative $RMSE$ of \hat{X}_T taking $\hat{M}_i = L_i, 2L_i$ and M and the systematic record (N) only
 180 ($\hat{M} = 0$) considered as an estimator of the true quantile $X_T = F^{-1}(1-1/T; \theta)$. If in a generated series one gets
 181 $\max(x_1, x_2, \dots, x_N) \geq XM$ such simulation is ignored which allows us to assume $\hat{M} = 2L$.

182 4. SIMULATION RESULTS

183 The concise frame of this paper made us to limit the number of models we took into consideration in our
 184 calculations. In order to lessen the number of the figures for all the combinations of CS and CV values we
 185 resigned from three-parameter distributions such as generalised extreme value (GEV) and turned into its two-
 186 parameter special forms, namely Gumbel (Gu) and Weibull (We). Another cause was also that, however
 187 theoretically sound, the GEV working perfectly for large samples often fails in far-from-asymptotic samples
 188 which we examine in this study. We scrutinised a number of two- and three-parameter distribution functions in
 189 terms of their best fit to hydrological annual and seasonal peak flows in Poland and it turned out that despite the
 190 regime of the river other models were preferred rather than GEV (*Strupczewski et al, 2012, Kochanek et al,*
 191 *2012*). However, the crucial argument after the choice of the parent distribution was the pioneering works of
 192 *Frances et al. (1994)* that we wanted to continue and develop. Results of simulation experiments are shown for
 193 Gu and We distributions with four values of the coefficient of variation $CV = 0.25, 0.5, 0.75, 1.0$, with two
 194 different lengths of systematic records $N = 15, 50$ and the length of effective historical period $M = N \exp(a)$
 195 where $a \in [0, 3]$. Due to the limited capacity of this paper without the loss of generality, only the selected results
 196 were presented in Figs. 2-5, namely for $CV=0.25$ and 1.0 ; the results for $CV=0.5$ and 0.75 locate themselves
 197 between those presented in the figures. Results for the correct value of the return period ($\hat{M} = M$) are compared
 198 with those got for $\hat{M} = L_i, 2L_i$. For completion the results for the systematic record only (i.e. $\hat{M} = 0$) were
 199 presented in all figures (solid line). Of course, for this case the results do not depend on M and in consequence
 200 on $\log(M/N)$.

201 5. DISCUSSION OF THE RESULTS

- 202 ▪ The shorter the gauged record (N) is, the more useful the historical information.
- 203 ▪ Using as the estimate of the true return period of largest historical flood (XM) the historical memory length (L)
 204 results in considerable upward bias RB of 1% quantile far exceeding the bias for the systematic record only. Its
 205 value increases with C_V (and C_S) and with the M/N ratio.
- 206 ▪ Using in ML estimation the $\hat{M} = 2L$ instead of $\hat{M} = L$ considerably reduces the bias and further reduction is
 207 obtained for the $\hat{M} = M$, i.e. for the return period (M) of the largest historical flood XM .

- 208 ▪ Although the use of $\hat{M} = 2L$ instead of $\hat{M} = L$ reduces the bias more than twice, it is still *circa* 40% larger
- 209 than the bias of a known return period M of XM , and comparable or lower than the bias from systematic record
- 210 (N).
- 211 ▪ As far as the relative root mean square error (RRMSE) of 1% quantile is concerned, for both Gumbel and
- 212 Weibull models one can notice some reduction in its values when one uses L , $2L$ or M return periods in
- 213 comparison to the systematic sample. The worst reduction of RRMSE one gets for L , better for $2L$ and the best
- 214 for M which means that it is worth, at least, considering using a historical measurement XM in upper quantile
- 215 estimation and then set the return period of XM to $2L$ rather than L if we do not know M .
- 216 ▪ The reduction in RMSE for both models (Gumbel and Weibull) rises generally with M/N ratio. In other words:
- 217 the bigger M (compared to N), the higher distance between RRMSE values got for the sample with additional
- 218 historical information and the systematic series. It goes without saying, that for $N = 15$ one gets better
- 219 reduction than for $N = 50$.
- 220 ▪ For the Gumbel model, regardless of the sample return period, L , $2L$ or M , the relative reduction in RRMSE
- 221 compared to systematic samples does not depend on CV . It does not hold for Weibull where the reduction
- 222 decreases with CV , e.g. between $CV = 0.25$ and 1.0 there is usually a few-percent difference which is minimal
- 223 (almost marginal) for $\hat{M} = M$.
- 224 ▪ For Gumbel model reduction in comparison to systematic sample for $\log(M/N) = 3$, $CV = 0.25$ and $N = 15$ the
- 225 reduction gets up to 2.2, 3.6 and 5.3% for L , $2L$ and M respectively. For $N = 50$ these numbers are roughly
- 226 three times smaller.
- 227 ▪ For Weibull the gain in RRMSE is more spectacular and for $\log(M/N) = 3$, $N = 15$ and $CV = 0.25$ equals to 3.4,
- 228 4 and 4.9% for L , $2L$ and M respectively (when $CV = 1.0$ the gain is c.a. four times lower). For $N = 50$ the
- 229 general trend for Weibull remains the same as for $N = 15$ but the reduction of RRMSE is smaller.
- 230 ▪ To sum up the RRMSE issues, the inclusion of the largest historical flood in FFA with $\hat{M} = 2L$ (i.e. the
- 231 effective historical record length) gives a few-percent reduction in RRMSE of extreme flood estimates.
- 232 However, the reduction is *circa* 20 up to 60% lower than if we took M as the length of simulation period. The
- 233 true value of M is not available in reality, so one is doomed to use $2L$ instead.
- 234 ▪ Therefore, to benefit from the largest historical observation every effort should be made to establish M
- 235 accurately.
- 236 ▪ In the absence of any information about the period preceding the occurrence of XM one should put \hat{M} equal
- 237 $2L$ or $2(L+N)$.
- 238 ▪ The benefit from including the largest historical flood of a given value is measured by the reduction of
- 239 RRMSE. It depends on:
 - 240 i. the length of systematic record (N),
 - 241 ii. the ratio of the true return period of XM , i.e. M to N ,
 - 242 iii. the ratio of N to the return period of quantile of interest,
 - 243 iv. the CV and skewness of the parent distribution.

244 6. CONCLUSIONS

245 Errors in historical data reduce, of course, the utility of the data for improvement of the estimation of flood
 246 magnitude at a given return period. In the simulations (Figs. 2-5) it was assumed that the magnitude of the
 247 largest historical flood (XM) was measured without error and the same was assumed for the systematic record. It
 248 is realistic to suppose that the XM flood was measured much less accurately than the gauged record. Error in
 249 estimating the largest historical magnitude (XM) is much more important than error in estimating the date of its
 250 occurrence (e.g. *Hosking and Wallis*, 1986 a, b). It is significant that inspired by the practice of efforts to
 251 improve the accuracy of estimates of flood quantiles through more realistic assumptions and a fuller use of the
 252 information they give just the opposite effect leads to increased uncertainty of flood estimates.

253 The next step should be to refer to the general problem of historical information when the applied distribution
 254 model is false, which is always the case (*Strupczewski et al.*, 2002). On the other hand, the uncertainty of the
 255 paleo-historical floods (both in terms of their magnitude and return period) combined with considerable
 256 increases in the complexity of the problem (when compared to analysis of systematic data only) provokes a
 257 fundamental question, whether the whole operation is worth a candle. Therefore, whether to include the paleo-
 258 historical information or turn a blind eye to it, is a matter of conscience.

259 All these generate two important practical problems which we leave for further study, namely:

- 260 1. What is the theoretical upper limit of accuracy of high quantile estimation when the theoretical value (i.e.
- 261 taken from the parent distribution) of return period for XM is known?
- 262 2. Here in our simulation experiment we assumed the knowledge of the true (parent) distribution function. The
- 263 role of historical information when the assumed distribution serves as the model of the true distribution
- 264 remains, for the time being, unknown.

265 Only the solutions to these two problems completed by the consideration of the observation errors in FFA brings
266 us closer to the answer to the fundamental question stated above, i.e. whether the available paleo-historical
267 record can give worthwhile improvement in flood estimates.

268
269 **Acknowledgements.** This research project was partly financed by the grant of the Polish National Science
270 Centre titled ‘*Modern statistical models for analysis of flood frequency and features of flood waves*’, contract nr
271 UMO-2012/05/B/ST10/00482 and made as the Polish voluntary contribution to COST Action ES0901
272 ‘*European Procedures for Flood Frequency Estimation (FloodFreq)*’.

273 8. REFERENCES

- 274 Benson, M.A.: Use of historical data in flood-frequency analysis, *EOS Trans. AGU*, 31(3), 419-424, 1950.
275 Bernieur, I., Miquel, J., Lebosse, A. and Griffet, A.: Use of additional historical information for estimation and
276 goodness of fit of flood frequency model, *Int. Symp. On Flood Frequency and Risk Analysis, L.S.U., Baton*
277 *Rouge, May 14-17, 1986*.
278 Cohn, T.A.: Flood Frequency Analysis with Historical Information, *Ph.D. Thesis, Cornell University, N.Y.*, June
279 1986.
280 Dalrymple, T.: Flood frequency analysis, *U.S. Geol. Surv. Water Supply Pap.*, 1543-A, 1960.
281 Ding, J. and Yang, R.: The determination of probability weighted moments with the incorporation of
282 extraordinary values into sample data and their application to estimating parameters for the Pearson type
283 three distribution. *Journal of Hydrology* 101, pp. 63-81, 1988.
284 Frances F., Salas, J.D. and Boes, D.C.: Flood frequency analysis with systematic and historical or paleoflood
285 data based on the two-parameter general extreme value models. *Wat. Resour. Res.* 30 (6), pp. 1653-1664,
286 1994.
287 Girguś, R. and Strupczewski, W.: Excerpts from the historical sources dealing with extraordinary hydro-
288 meteorological phenomena on the Polish territories from X - XVI c. *Instr. i Podr.* 87(165). Wyd. Kom. i
289 Laczn., pp.216 (in Polish), Warsaw 1965.
290 Hirsch, R. M., and Stedinger, J. R.: Plotting positions for historical floods and their precision, *Water Resour.*
291 *Res.*, 23(4), 715-727, 1987.
292 Hirsch, R.M.: Probability plotting positions for flood records with historical information, paper presented at
293 U.S., China Bilateral Symposium on the Analysis of Extraordinary Flood Events, Oct. 21-23, Nanjing, China,
294 1985.
295 Hosking U. R. M. and Wallis, J.R.: The value of Historical Data in Flood Frequency Analysis. *Water Resources*
296 *Research*, 22, 11, 1506-1612, 1986a.
297 Hosking U. R. M. and Wallis, J. R.: Paleoflood Hydrology and Flood Frequency Analysis. *Water Resources*
298 *Research*, 22, 4, pp.543-550, 1986b.
299 Hosking, J. R. M.: The use of *L*-moments in the analysis of censored data. In *Recent Advances in Life-Testing*
300 *and Reliability*, edited by N. Balakrishnan, pp. 545-64. CRC Press, Boca Raton, Fla. 1995
301 Interagency Advisory Committee on Water Data (IACWD) and U.S. Water Research Council Hydrology
302 Committee,. Guidelines for determining flood flow frequency, Bull 17B, (revised) Hydrol Subcomm, Office
303 of Water Data Coord., U.S. Geol. Surv., Reston, Va.U.S. Gov. Print. Off. Washington D.C., 1982
304 Kochanek, K., Strupczewski, W. G. and Bogdanowicz, E.,: On seasonal approach to flood frequency modelling.
305 Part II: flood frequency analysis of Polish rivers. *Hydrol. Process.*, 26: 717–730. doi: 10.1002/hyp.8178,
306 2012
307 MacDonald, N.,: Reassessing flood frequency for the River Trent through the inclusion of historical flood
308 information since AD 1320. *Hydrology Research*, Volume: 44 Issue: 2 Pages: 215-233 DOI:
309 10.2166/nh.2012.188, 2013
310 Natural Environment Research Council (NERC), Flood Studies Report, vol.1, London 1975.
311 Naulet, R., Lang, M., Ouarda, T. B. M. J., Coeur, D., Bobee, B., Recking, A. and Moussay, D.: Flood frequency
312 analysis on the Ardèche river using French documentary sources from the last two centuries. *J. of Hydrol.*
313 313, pp. 58-78, 2005.
314 Parent, E., and Bernier, J.: Bayesian POT modeling for historical data, *J. Hydrol.*, 274(1–4), 95–108,
315 doi:10.1016/S0022-1394(02) 00396-7, 2003.
316 Payrastré, O., Gaume, E. and H. Andrieu, Usefulness of historical information for flood frequency analyses:
317 Developments based on a case study, *Water Resour. Res.*, 47, W08511, doi:10.1029/2010WR009812, 2011
318 Payrastré, O., Gaume, E. and Herve, A.: Historical information and flood frequency analyses: which optimal
319 features for historical floods inventories? *Houille blanche-revue internationale de l'eau*, Issue: 3 Pages: 5-11
320 DOI: 10.1051/1hb/2013019, published: JUN 2013

321 Reis, D. S. J., and Stedinger, J. R.: Bayesian MCMC flood frequency analysis with historical information, *J. of*
322 *Hydrol.*, 313(1–2), 97–116, 2005.

323 Stedinger, J. R. and Cohn, T. A.: Flood frequency analysis with historical and paleoflood information. *Wat.*
324 *Resour. Res.* 22(5), pp. 785-793, 1986.

325 Stedinger, J. R., and Baker, V. R.: Surface Water Hydrology: Historical and Paleoflood Information. *Review of*
326 *Geophysics*, V.25, 2, pp. 119-124. U.S. National Report to IUGG 1983-1986, 1987.

327 Strupczewski, W. G., Kochanek, K., Bogdanowicz, E. and Markiewicz, I.: On seasonal approach to flood
328 frequency modelling. Part I: Two-component distribution revisited. *Hydrol. Process.*, 26: 705–716.
329 doi: 10.1002/hyp.8179, 2012

330 Strupczewski, W. G., Singh, V. P. and Weglarczyk, S.: Asymptotic bias of estimation methods caused by the
331 assumption of false probability distribution. *J. of Hydrol.* 258/1-4, pp. 122-148, 2002.

332 Vigilione, A., Merz, R. and Salinas, J. L.: Flood Frequency Hydrology. 3. A Bayesian analysis. *Water Resour.*
333 *Res* 49, 2, pp. 675-692. DOI: 10.1029/2011WR010782, 2013

334 Wang Q. J.: Unbiased estimation of Probability Weighted Moments and Partial-Probability Weighted Moments
335 from Systematic and Historical Flood Information and Their Application to Estimating the GEV
336 Distribution. *J. of Hydrol.* 120, pp. 115-124, 1990.

337 Wang, Q. J.: Using partial probability weighted moments to fit the extreme value distributions to censored
338 samples. *Water Resour. Res.*, vol. 32 no. 6., pp. 1767-1771, 1996.

339 Wang, S. X. and Adams, B. J.: Parameter Estimation in flood frequency analysis. Publ. 84-02. Dep. Of Civ.
340 Eng., Univ. of Toronto, April 1984.

341 Zhang, Y.: Plotting positions of annual flood extremes considering extraordinary values. *Water Resour.*
342 *Res.*,18(4), pp. 859-64, 1982.

343 Zhang, Y.: On the role and treatment of outliers in probability estimation method of flood frequency analysis.
344 Paper presented at U.S. – China Bilateral Symposium on the Analysis of Extraordinary Flood Event, Oct. 21-
345 25, Nanjing, China, 1985

346

347 **List of figures:**

348

349 **Figure 1.** The case of N systematic and one largest flood in the beginning of historical period.

350 **Figure 2.** Relative bias (RB) and relative root mean square error (RRMSE) of $\hat{X}_{T=100}$ as a function of gauge record length N
351 and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$. Parent distribution Gumbel with CV equal: 0.25 and 1.0 and $N=15$. Fitted
352 distribution Gumbel.

353 **Figure 3.** RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$.
354 Parent distribution Gumbel with CV equal 0.25 and 1.0 and $N=50$. Fitted distribution Gumbel.

355 **Figure 4.** RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$.
356 Parent distribution Weibull with CV equal 0.25 and 1.0 and $N=15$. Fitted distribution Weibull.

357 **Figure 5.** RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$.
358 Parent distribution Weibull with CV equal 0.25 and 1.0 and $N=50$. Fitted distribution Weibull.
359

360 **Figures:**
361

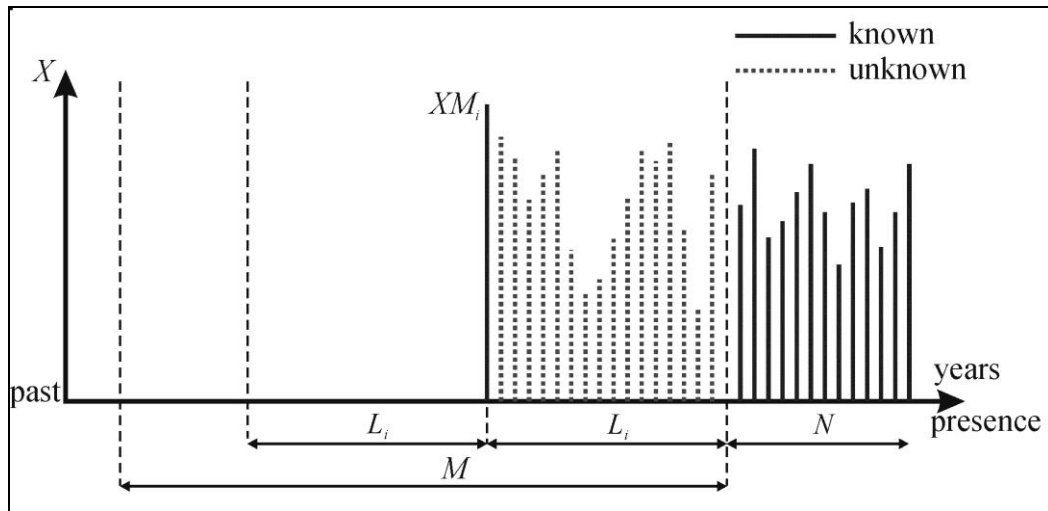


Figure 1. The case of N systematic and one largest flood in the beginning of historical period.

362
363
364

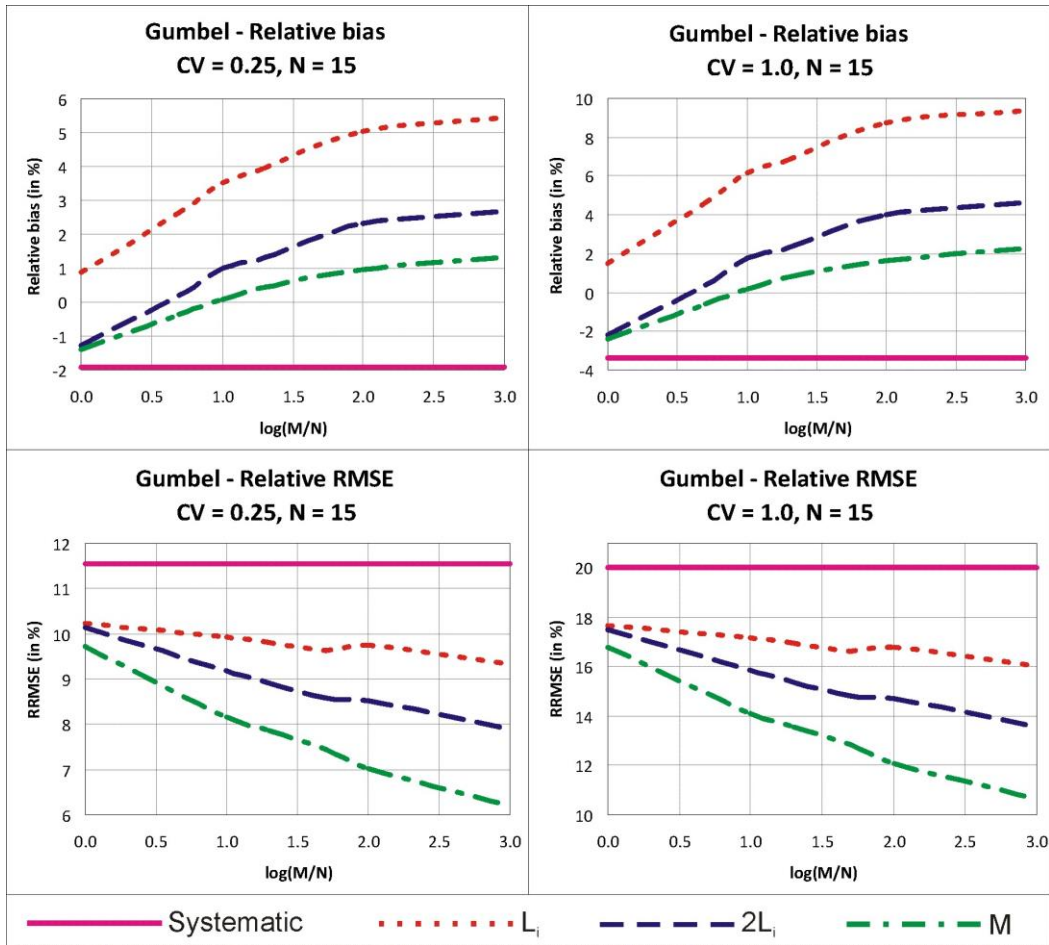
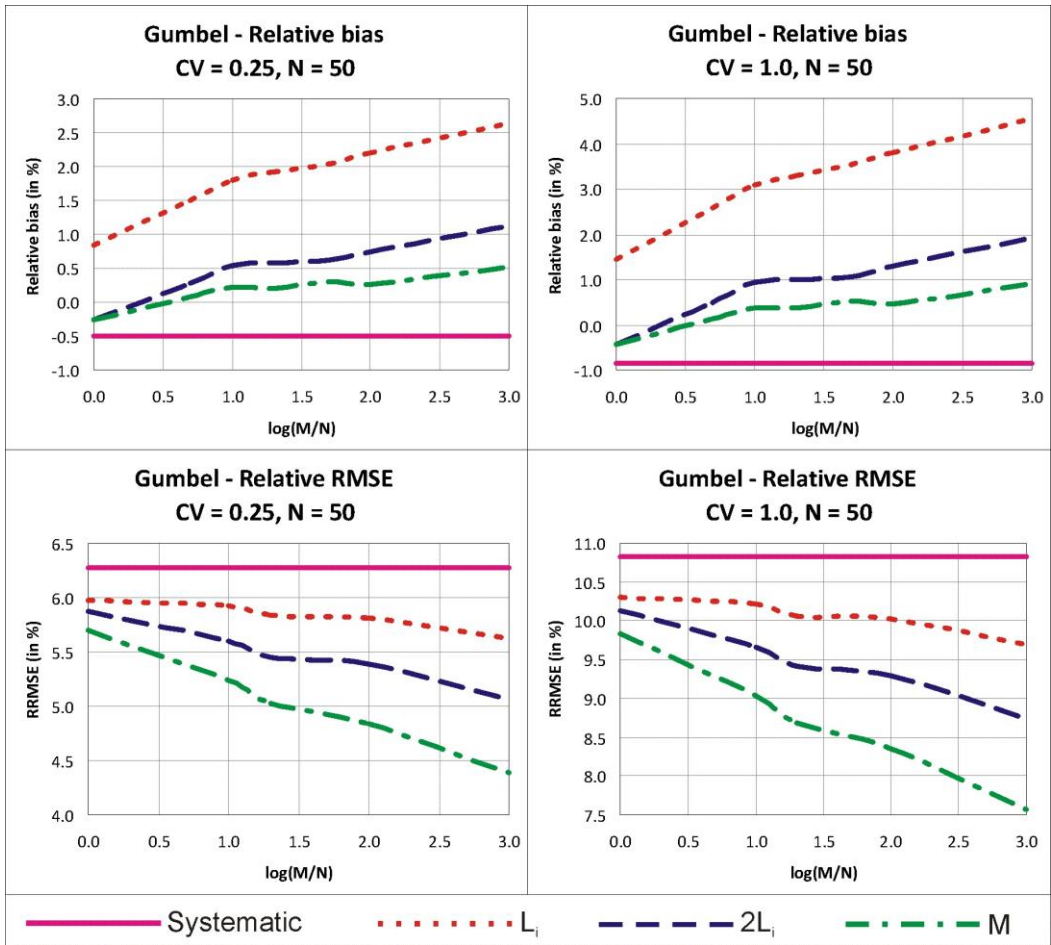


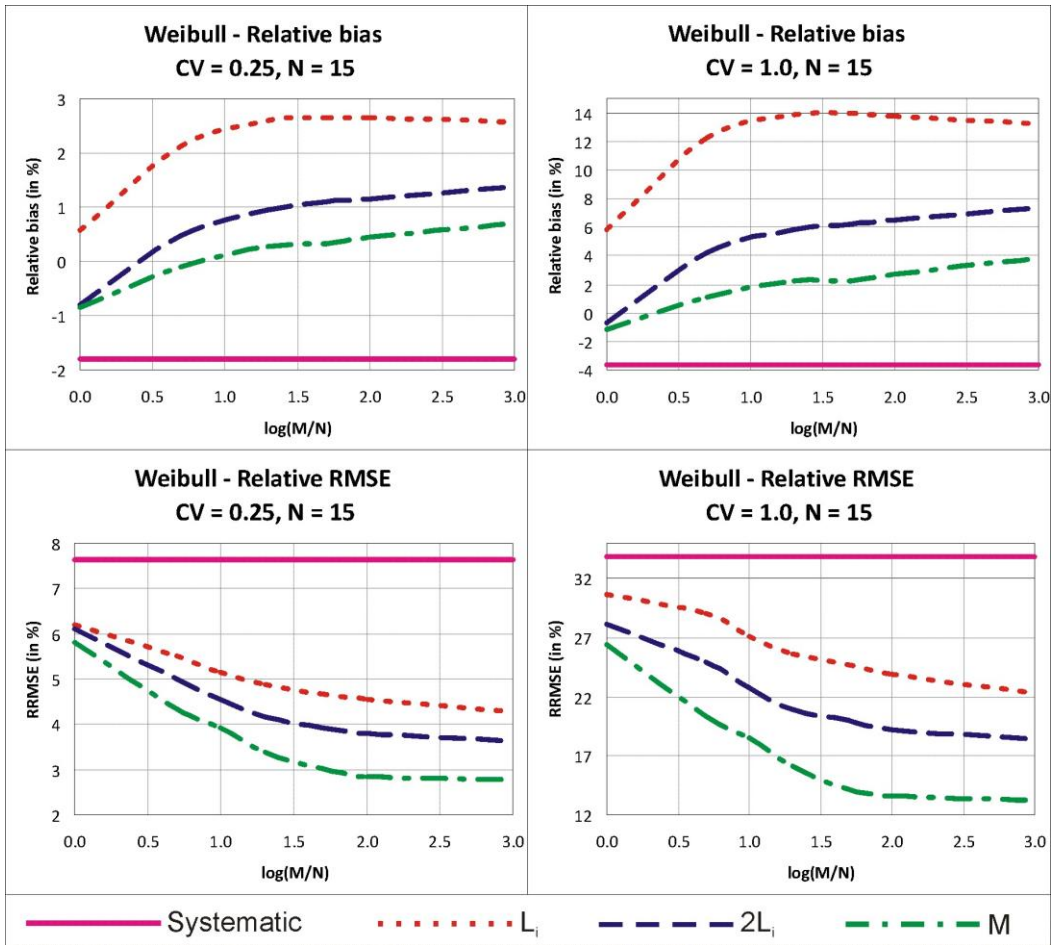
Figure 2. Relative bias (RB) and relative root mean square error (RRMSE) of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$. Parent distribution Gumbel with CV equal: 0.25 and 1.0 and $N=15$. Fitted distribution Gumbel.

365
366
367
368
369



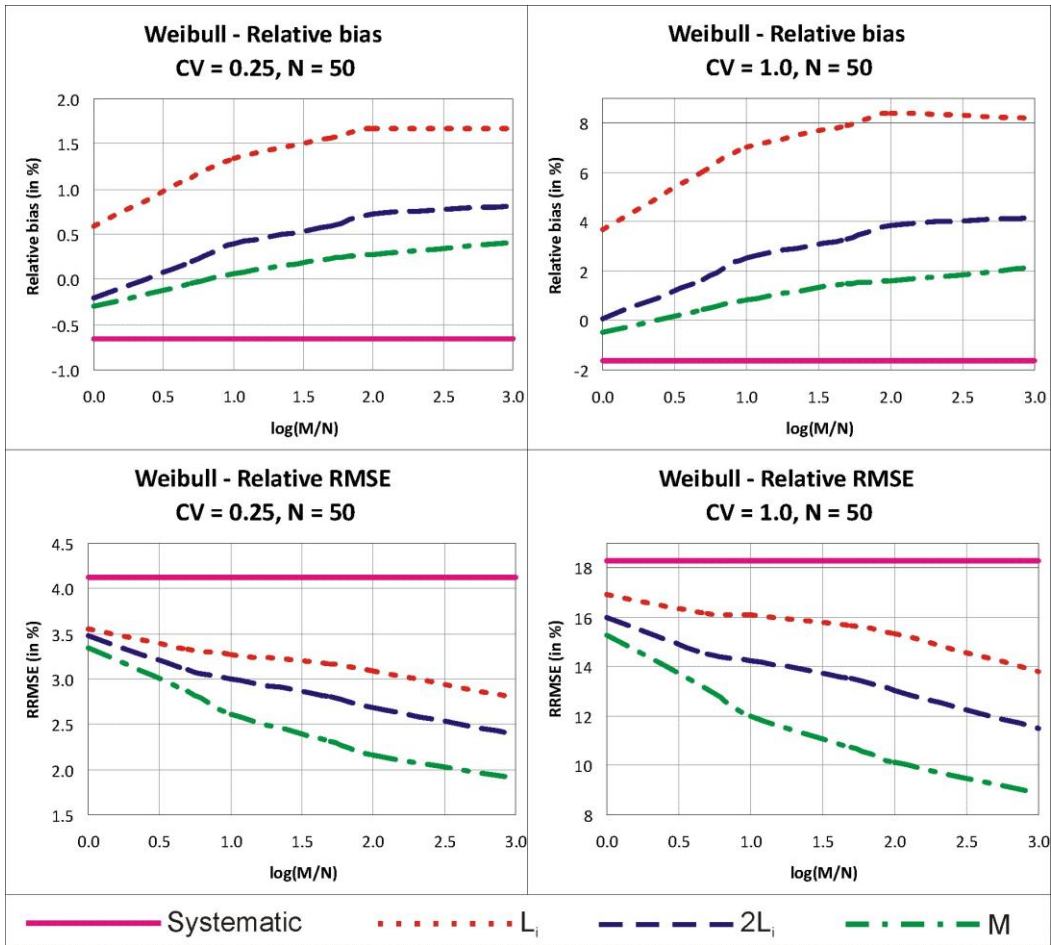
370
 371
 372
 373

Figure 3. RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$. Parent distribution Gumbel with CV equal 0.25 and 1.0 and $N=50$. Fitted distribution Gumbel.



374
 375
 376

Figure 4. RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$. Parent distribution Weibull with CV equal 0.25 and 1.0 and $N=15$. Fitted distribution Weibull.



377
 378
 379

Figure 5. RB and RRMSE of $\hat{X}_{T=100}$ as a function of gauge record length N and historic period M for $\hat{M}_i = 0, L_i, 2L_i, M$. Parent distribution Weibull with CV equal 0.25 and 1.0 and $N=50$. Fitted distribution Weibull.