DE GRUYTER OPEN

Acta Geophysica
vol. 64, no. 1, Feb. 2016, pp. 206-236
DOI: 10.1515/acgeo-2015-0070

Comparison of Two Nonstationary Flood Frequency Analysis Methods within the Context of the Variable Regime in the Representative Polish Rivers

Witold G. STRUPCZEWSKI[1], Krzysztof KOCHANEK[1], Ewa BOGDANOWICZ[2], Iwona MARKIEWICZ[1], and Wojciech FELUCH[3]

[1]Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland
e-mails: wgs@igf.edu.pl, kochanek@igf.edu.pl (corresponding author), iwonamar@igf.edu.pl

[2]Project CHIHE, Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland; e-mail: ewabgd@o2.pl

[3]Warsaw University of Technology, Department of Civil Engineering in Płock, Płock, Poland; e-mail: wfeluch@gmail.com

Why do something in a simple way,
while better make it complicated to impress others?!
(overheard)

Abstract

Changes in river flow regime resulted in a surge in the number of methods of non-stationary flood frequency analysis. Common assumption is the time-invariant distribution function with time-dependent location and scale parameters while the shape parameters are time-invariant. Here, instead of location and scale parameters of the distribution, the mean and standard deviation are used. We analyse the accuracy of the two methods in respect to estimation of time-dependent first two moments, time-invariant skewness and time-dependent upper quantiles. The

Ownership: Institute of Geophysics, Polish Academy of Sciences;
© 2016 Strupczewski *et al*. This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivs license, http://creativecommons.org/licenses/by-nc-nd/3.0/.

method of maximum likelihood (ML) with time covariate is confronted with the Two Stage (TS) one (combining Weighted Least Squares and L-moments techniques). Comparison is made by Monte Carlo simulations. Assuming parent distribution which ensures the asymptotic superiority of ML method, the Generalized Extreme Value distribution with various values of linearly changing in time first two moments, constant skewness, and various time-series lengths are considered. Analysis of results indicates the superiority of TS methods in all analyzed aspects. Moreover, the estimates from TS method are more resistant to probability distribution choice, as demonstrated by Polish rivers' case studies.

Key words: non-stationary flood frequency analysis, maximum likelihood, covariance, two-stage methodology, L-moments.

1. INTRODUCTION

Until recently, most tools and techniques used in flood frequency analysis (FFA) assumed stationarity of flood processes (*e.g.*, Milly *et al.* 2008). Nowadays, it is understood and accepted that, due to the climate and land cover change and rapid development of calculation techniques, the incorporation of the "non-stationarity" factor in hydrological parametric and non-parametric modelling should be seriously considered (when necessary and when the length of the time series allows it) and has become technically possible. Thus, hydrologists face the challenge of developing new or extending existing methods of FFA by incorporating the non-stationarity of extreme hydrological events.

The research on non-stationarity in hydrology based on the flow or water level data can be generally (subjectively) divided into testing the time-series for its (non-)stationarity and searching for the non-stationarity in models' parameters. The validity of the stationarity and independence assumptions of extreme series is assessed by investigating the presence of monotonic trends using nonparametric Mann–Kendall (Kendall 1975), Pettitt (1979) and autocorrelation coefficient tests. The common use of the non-parametric tests is mainly due to their robustness to non-normality which usually appears in hydrological extreme series. The parametric student's *t*-test provides also useful information on variation of extreme series in the sense that it evaluates the significance of trend in the mean value. Diagnostic results from trend tests substantiate the implementation of the non-stationary flood frequency analysis (NFFA).

Flood frequency analysis in the presence of covariates has been the subject of countless publications and each of them contains more or less objective brief review of the literature. The relatively seldom-cited statistical publication of Davison and Smith (1990) is of great value for the development of non-stationary modelling of extreme event time-series, because, perhaps for the first time in environmental science, the idea of the maximum likelihood (ML) estimation of distribution parameters with covariates was presented there. Nowadays, the parameters of great variety of flood frequency distribution functions with the presence of time covariate can be estimated, *e.g.*, by the Generalized Additive Models for Location, Scale and Shape (GAMLSS) software (Rigby and Stasinopoulos 2005). Unfortunately, the effectiveness of the algorithms in the sense of finding the global maximum is not known and cannot be easily assessed from GAMLSS R-package.

However, the main aim of the FFA, regardless of the stationarity matter, is the proper estimation of the upper (flood) quantiles. These quantiles are particularly sensitive to the value of the shape parameter whose accurate estimation, especially by means of the ML for a short sample size, is quite a challenge even in the stationary case. The issue of the shape parameter becomes even more difficult in the NFFA where the number of parameters grows but the dataset size remains the same. Besides, the larger the number of the estimated parameters, the more likely are the elementary numerical problems with finding a global maximum of likelihood function for every time-series.

As a remedy to uncertainty of trend estimates resulted from hypothetical distributional assumption and to optimization problems arising due to large number of estimated parameters of the short time-series, the ideas of decomposition were born where the optimization tasks are carried out through the hierarchical solution schemes.

The Hybrid Two-Stage (TS) method (Kochanek *et al.* 2013) consists of a separate estimation of time-dependent mean and standard deviation, and then, on that basis, it estimates the shape parameter and time-dependent quantiles. Estimation of time-dependent mean and standard deviation is performed by the Weighted Least Squares (WLS) method (Strupczewski and Kaczmarek 1998, 2001) where the assumption of distribution model is not required. The only limitation is the existence of location and scale parameters, so the method can be used for both three-parameter distributions with lower bound parameter and two-parameter distributions with unlimited range on both sides, such as Gumbel or Normal. In the second stage of the TS, after choosing the flood model (distribution function), the shape parameter is estimated by means of the L-moment method and then the time-dependent quantiles are determined.

In this paper, the accuracy of TS method is compared by means of the Monte Carlo experiment with the accuracy of ML method with covariates. Both NFFA methods were also evaluated within the context of the case study of selected gauging stations on Polish rivers where their pros and cons are even more visible. In order not to lose the substance of presentation in the description of the NFFA implementations, here we limit ourselves to point to our own research which led to the formulation of and justification for a two-stage NFFA method.

2. TOWARDS SUBSTANTIATION OF THE HYBRID METHOD IN THE NFFA

The problem of changes of natural processes became the subject of many publications, especially in the 1980s. The primary sources of information about stochastic properties of hydrological processes are the data time series. Investigation of trend in hydrological time series was mainly focused on detection of a trend in the mean value by means of standard statistical techniques, based essentially on the hypothesis testing theory. A trend in the variance or in the auto-correlation function has been rarely analysed. In this section we present only our research which evolved towards the formulation of and justification for the two-stage (TS) method.

Strupczewski and Mitosek (1991, 1995) proposed the extension of at-site flood frequency modelling to non-stationary case. The ML method with the presence of time as the covariate was used for estimation. The climatologists who investigated non-stationarity issues in datasets concentrated on certain norms, and perhaps the deviation from these norms. Let us note that the identification of non-stationarity effect caused by changes both in land use and cover and by water-engineering activity should be based on the physical models of rainfall-runoff type which leads to estimates of changes in the mean or also in the dispersion of extreme events. Note that the main interest of the FFA is the estimation of upper quantiles of annual maxima distribution. As demonstrated in Appendix B for GEV distribution (Fig. B2) these quantiles are much more sensitive to a change in standard deviation than to a change in the mean value. This is why incorporation of the possibility of a trend for the second moment is so important in the FFA.

Given the above, Strupczewski and Mitosek (1991, 1995) have found it convenient to unify various probability distribution functions (PDFs) in terms of parameters replacing the original set of parameters of each PDF, namely: location, scale, and shape parameter by the mean value, standard deviation and possibly also skewness coefficient, respectively. This re-parameterisation is done by using the relationships between moments and parameters available in statistical literature. Consequently, after the re-parametrisation, the trend would be explicitly introduced to the moments, which is conformable to methods of trend analysis. In such a case, a time-trend is investigated in two first moments only and is assumed to be a continuous function of time. The skewness coefficient is considered to be time-invariant. Primarily, Strupczewski and Feluch (1997) and Strupczewski *et al.* (2001a) analysed four classes of time-trend in moments: (A) in the mean value, (B) in the standard deviation, (C) both in the mean value but the standard deviation was related to a constant value of the variation coefficient (CV), and (D) unrelated trend in the mean value and the standard deviation. The basic option was the time-invariable parameters, *i.e.*, the stationary option (S). If the time-dependent parameters are to be estimated from peak flow records, they shall be expressed as functions with the smallest possible number of parameters (Strupczewski and Mitosek 1995). Each of the three first classes (A-C) increases the number of parameters of the stationary PDF by the same number defined by the form of a time-trend function. Regarding the problems with solution of maximum likelihood equations, originally the two options of the form of trend were included, *i.e.*, the linear and parabolic (square trinomial) (Strupczewski and Feluch 1998, Strupczewski and Kaczmarek 1998, Strupczewski and Mitosek 1998). The ML estimation procedures of parameters of various (Distribution and Trend) models are elements of Identification of Distribution and Trend System (IDT) and the number of models had gradually increased. The procedures also covered evaluation of the standard error of parameters and time-dependent quantiles. The Akaike Information Criterion (AIC) (Akaike 1974, Hurvich and Tsai 1989) was used for the identification of an optimum model in a class of competing models. If for a given time-series the lowest AIC value corresponds to the stationary model (S), the time-series is considered as stationary. The IDT System had been extended to cover the procedure of estimation of probability distribution together with the standard error not only in respect to one year, but to a period of any length, too.

This idea has been repeatedly and scrupulously examined on simulation data and tested and applied to Polish rivers data. It was found (Strupczewski 1999, Strupczewski and Kaczmarek 2001, Strupczewski *et al.* 2001a, b) that for a hydrological size of time-series (up to max. 100 elements) the differences in Akaike criterion values of some best fitting models are rather small, but at the same time the differences of time-dependent moments and quantiles can be considerable. Sometimes the trend got from various distributional assumptions may even differ in sign and, moreover, the best fitting model may change from year to year. All these cause substantial differences in moments and quantiles when extrapolating such models out of range of record (*e.g.*, in the future). This feature is also highly confusing and raises doubts as to the reliability of the best fitting model and the utility of this methodology for short time-series.

In the NFFA, usually just one model is adopted as the only truth. This is the convenient way to hide the uncertainty of estimates, for instance, by decreeing one type of such distributions, *e.g.*, the GEV, as applicable throughout the country. The applicability of one-model concept in the NFFA assumes the knowledge of True probability distribution function. However, the simulation experiments performed for numerous pairs of True and False PDFs revealed weak power of various discrimination procedures for hydrological datasets which even drops when the number of competing distributions rises (*e.g.*, Strupczewski *et al.* 2006, Mitosek *et al.* 2006). It means that except of the random error, we have to face the model error, too. An analytical method for evaluating the robustness of the estimates of moments and quantiles by various estimation methods, including the maximum likelihood method and method of L-moments, with respect to the distribution choice has been presented by Strupczewski *et al.* (2002a, b). It was shown that the relative asymptotic bias of the ML-estimate of moments and upper quantiles can be considerable and grows rapidly with increasing value of the coefficient of variation. The relative asymptotic ML bias of the variance is often hundreds times larger than the relative asymptotic ML bias of the mean! As the number of function parameters approximating increases the accuracy of approximation and therefore decreases the model error of upper quantiles reproduction. The research briefly described above can be summarised by stating that the estimation technique in the NFFA should be relatively robust to false distributional assumptions.

3. REMARKS ON APPLICATION OF TWO-STAGE (TS) METHOD

Years of development, perfecting and testing on real data of the IDT computer package based on ML method in the presence of time as the covariance revealed its deficiencies stemming from the over-parameterization in relation to the length of hydrologic time series. It resulted in preliminary ideas of algorithms using of the Two-Stage method based on the Weighted Least Squares (WLS) (Strupczewski and Kaczmarek 1998, 2001) combined with the concept of L-moments (Kochanek *et al.* 2013). The L-moments (LM) estimation method (*e.g.*, Hosking *et al.* 1985, Hosking 1990, Hosking and Wallis 1997) requires data to be arranged in ascending order. So the procedure of the estimation of non-stationary flood quantiles was divided into two stages:

- First, the time-dependent mean (μ_t) and standard deviation (σ_t) of the annual flows time-series are estimated using the WLS (Strupczewski and Kaczmarek 1998, 2001) of the annual maximum flows time-series (see Appendix A). The so-obtained time-dependent moment values are then used to standardise (deprive of trends) the time series:

$$y_t = \left(x_t - \mu_t\right)/\sigma_t \ . \tag{1}$$

- The stationary sample is then used to estimate by the L-moments method the parameters and quantiles (stationary!) $Y(F)$ of a chosen distribution function. Afterwards, the so-calculated quantiles are re-trended:

$$X(F,t) = Y(F) \cdot \sigma_t + \mu_t \quad , \tag{2}$$

where F is the cumulative distribution.

Of course, in the second stage any parametric or non-parametric method of estimation could be used. The use of a non-parametric method, *e.g.*, the kernel probability density estimation (*e.g.*, Rosenblatt 1956, Feluch 1994) would definitely detach the trends estimation from any distribution function. It would result in "purely" data-driven technique. In this paper, due to its advantages, the L-moments method is used as the second stage of the TS approach.

4. TRENDS IN MOMENTS

Analysis of trends in datasets means that the number of parameters to be estimated from fixed-length time-series will increase. Due to the principle of parsimony, the assumed form of trends in moments or parameters ought to be as simple as possible. Therefore, one or two parameter functions, such as linear, logarithmic, exponential, trigonometric, parabolic, *etc.*, shall be first considered. We adopted here the linear form of trends in the mean (μ_t) and standard deviation (σ_t):

$$\mu_t = a \cdot t + b \quad \text{and} \quad \sigma_t = c \cdot t + d \quad , \tag{3a, b}$$

where t – time (in the years following the beginning of the flood records), a – parameter of the trend in the mean, *i.e.*, parameter of "slope", b – the parameter of mean, *i.e.*, parameter of "intercept", c – parameter trend in the standard deviation, *i.e.*, parameter of "slope", and d – the standard deviation parameter, *i.e.*, parameter of "intercept".

As one can see, instead of two parameters in stationary case (μ , σ), now there are four parameters to be simultaneously estimated in non-stationary case: a, b, c, and d. The fifth parameter, the shape, is time-independent and is estimated in the second stage of the TS method where the L-moment method is used. Note that in ML method all five parameters must be estimated together.

In the TS method, the parameters a, b, c, and d are used for the standardisation of the time series x_t of annual (or seasonal) maximum flows. As a result a sample, y_t, free of trends in the mean and standard deviation is obtained:

$$y_t = \left(x_t - a \cdot t - b\right) / \left(c \cdot t + d\right) t = 1, \ldots, N \quad . \tag{4}$$

The elements y_t of a random sample have been sorted to form increasing series suitable to calculate the L-moments for estimation of the parameters and quantiles (stationary!).

5. THE COMPARISON OF THE ML AND TS TECHNIQUES BY NUMERICAL EXPERIMENT

The ML estimates of distribution parameters are asymptotically efficient. However, this holds if and only if a random sample comes from the same distribution, which is not the case in hydrology. When the parent distribution is unknown and hypothetical (H) distribution is used instead, ML is losing its optimal properties. But even though the true distribution was known, the large number of its parameters would be impossible to estimate from short time-series. Therefore, doing our best we still deal with false distribution function (H = False, the hypothesis is false).

To compare the accuracy of the ML and TS methods, we assume the knowledge of a probability distribution (with time-dependent parameters) from which the time-series are generated, *i.e.*, H = True (the hypothesis is true); moreover, we consider the length of the time-series compatible with hydrological realities. Thus, in this way, we deliberately "give a head start" to the ML method. Theoretically, the farther the assumed distribution function from the underlying distribution (H = False), the more the advantage of the ML method over the TS should fade. Consequently, if the TS method proved better for the H = True case, it is obvious that it will be better for the H = False case, too. Taking it for granted, the presentation of the results of our simulation experiments is confined to the case when H = True.

Most commonly, the three-parameter distribution functions with the lower bound as the parameter are used as models in FFA. In the ML method, one parameter more, *i.e.*, the shape parameter, is to be estimated than in the WLS method. It just might be a factor to the detriment of the ML method even for H = True for insufficiently long time-series. Note that for two-parameter distributions unlimited on both sides (such as Gumbel and Normal), the two methods do not differ in the number of estimated parameters, so the above argument loses its strength. Furthermore, when H = Normal, the two methods are equivalent.

The Generalised Extreme Values (GEV) distribution (see Appendix B) with time-dependent moments serves here as parent distribution in simulation experiments, which are performed for various parameters of time-dependent mean (μ_t) and standard deviation (σ_t), time-invariant coefficient of asymmetry CS, and various lengths of time series, *i.e.*, $N = 50$, 100, and 200. Records smaller than 50 elements should not be qualified for NFFA while time series larger than 200 elements practically do not exist in hydrology. According to our analysis of the 38 annual maxima datasets collected over the years in gauging stations located in all bigger Polish rivers, the values of the coefficient of variation ($CV_{t=0} = d/b$) and coefficient of asymmetry (CS) selected for simulation experiments series fits perfectly within the range of values occurring in Polish rivers (compare Kochanek *et al.* 2012, Table I; Markiewicz *et al.* 2014, diagram 1). For brevity, the results will be presented only for: (i) one parent time-dependent mean, $\mu_t = 1 \cdot t + 100$ (*i.e.*, $a = 1$, $b = 100$); (ii) three time-dependent standard deviations, *i.e.*, $c = 1$ and $d =$ 25, 50, and 100; and (iii) three values of the coefficient of asymmetry, CS = 1.5, 2.0, and 3.0, which correspond to the GEV shape parameter (Eq. B4) $k = -0.053$, -0.108, and -0.176, respectively. Having three time-series lengths ($N = 50$, 100, and 200), the experiment gives altogether 27 different combinations.

Using both methods, ML and TS, for a given time-series, the estimates of time-dependent mean and standard deviation, as well as a shape parameter, can be estimated. Having all the needed parameters, the selected time-dependent quantiles $X(F, t)$ are computed (Eq. B7).

For each variant (a, $b = 1$, c, $d = 1$, CS, N) 1000 series have been generated. The absolute relative root mean square errors (RRMSE):

$$\left|\mathrm{RRMSE}\,(\hat{\theta})\right| = \left|\left[E(\hat{\theta}-\theta)^2\right]^{1/2} \Big/ \theta\right| \tag{5}$$

of the estimates of parameters of time-dependent moments, shape parameter estimators, and the upper quantiles estimators got by the both methods serv here as a goodness-of-fit criterion.

For the sake of clarity, we have decided not to go deeply into details concerning the comparison of estimators biases, although this criterion deserves a word of comment, too. Note that for our analysis the most beneficial was to demonstrate the advantages of ML methods when a true distribution function is known. Then, as in the WLS method, the ML parameter estimates will be charged only by sampling bias of ML estimation method. The estimation relative biases of both the methods represent small components of the RRMSE. Except for the estimate of the shape parameters (see Table 2), the values of bias of other estimators vary within the range of a few percent and decrease with the length of time-series. Due to the variability of sign, bias values are more difficult to interpret than the RRMSE or Standard Error. However, in reality the true distribution function is unknown and one deals with a hypothetical distribution. In such a case, a model bias can be a significant part of the overall ML bias serving as yet another argument for comparing the competitiveness of both methods. In practice, though, the size of the model bias is difficult to assess. Nevertheless, all the effects visible in the RRMSE results can be identified also in bias, though because of relatively small samples (up to 200) they are not as vivid as in terms of RRMSE. It is so because, even though the ML provides lower bias of a particular estimator, it can be marred by the standard error, so the overall RRMSE of this estimator eventually proves better for the TS (WLS) (of course, the reverse case is possible).

The conclusions from the experiment results may be generalised to the other cases. As we will show by means of numerical simulations, a simple Two-Stage methodology proved better than the ML for hydrological size of time-series.

5.1 The trends in mean and standard deviation

The RRMSE of the estimated parameters of time-dependent mean and standard deviation (*a*, *b*, *c*, and *d*) of the ML and WLS methods are shown in Table 1. The relative accuracy of the estimation of all four parameters showed similar values for both methods. The |RRMSE| increases with the trend in the relative value of σ_t, *i.e.*, with $c/d = 1/d$ and with decreasing length of time-series, while showing low sensitivity to the asymmetry of the distribution. The |RRMSE| of the initial mean $b = \hat{m}_{t=0}$ for both methods is considerably lower than |RRMSE| for three other parameters. However, the most important in Table 1 is that the position of the figures in bold clearly shows the advantage of the accuracy of estimates all four parameters obtained from the WLS method for $N = 50$. This advantage is decreasing gradually with increasing length of time-series. The bias of the trend estimators (not shown) allows to draw fairly similar conclusion to the |RRMSE|: in general the majority of the WLS estimators reveal lower |RB|, too, though the values do not follow such a stable pattern as in the |RRMSE|. Because the case considers situation when $\mathrm{H} = \mathrm{True}$ the asymptotic advantage of the accuracy of estimates by the ML method is expected. Nonetheless, it should not be forgotten that in practice we are dealing with the case $\mathrm{H} = \mathrm{False}$ that allows to anticipate that even for a very long time-series, the WLS method will retain its supremacy.

5.2 The shape parameter

The proper estimation of the coefficient of asymmetry (CS) has been the subject of countless hydrological discussions (*e.g.*, Matalas and Benson 1968, Wallis *et al.* 1974, Yevjevich and Obeysekera 1984). As the GEV belongs to the family of heavy tailed distributions, the accuracy of estimate of the shape parameter ($\hat{k}$) will be the subject of comparison (Table 2), not the coefficient of asymmetry $\hat{\mathrm{CS}}$ (see Eq. B2). Note that for some specific Monte Carlo-generated series the $\hat{\mathrm{CS}}$ may not exist.

In the non-stationary ML approach, the shape parameter *k* is estimated together with *a*, *b*, *c*, and *d* parameters by solving the maximum likelihood equations, *i.e.*, five parameters are estimated simultaneously. On the contra-

Table 1

The |RRMSE| of estimators of GEV time-dependent parameters by the ML and WLS methods

Parameters of the MC generator			RRMSE [%], H = T = GEV							
	$\mu_t = 1 \cdot t + 100$		ML				WLS			
N	σ_t	CS	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$
50	$1 \cdot t + 25$	1.5	50.48	10.41	68.49	48.32	**48.31**	10.46	**51.43**	**45.51**
		2.0	50.31	10.33	81.61	61.37	**46.72**	**10.32**	**57.90**	**52.95**
		3.0	49.13	9.99	96.67	67.12	**47.51**	10.56	**69.09**	**63.69**
	$1 \cdot t + 50$	1.5	76.17	18.34	96.18	43.48	**72.82**	**18.04**	**82.11**	**40.87**
		2.0	71.41	17.39	81.97	50.03	71.44	17.88	98.14	**45.46**
		3.0	73.20	17.23	160.11	61.66	74.01	18.26	**107.28**	**56.08**
	$1 \cdot t + 100$	1.5	131.67	37.24	512.81	65.74	**119.58**	**34.36**	**140.38**	**37.65**
		2.0	129.57	35.51	462.24	70.53	**116.22**	**33.43**	**158.71**	**46.11**
		3.0	127.63	34.89	204.11	76.02	**122.45**	**34.82**	**190.48**	**52.87**
100	$1 \cdot t + 25$	1.5	22.97	8.55	23.42	31.87	**22.80**	**8.50**	26.69	38.49
		2.0	23.80	8.17	27.27	35.96	24.39	8.64	32.85	48.30
		3.0	22.67	7.75	33.64	41.39	23.75	8.39	39.49	61.95
	$1 \cdot t + 50$	1.5	33.09	14.09	30.49	30.02	**32.48**	**13.69**	37.75	32.56
		2.0	32.39	13.57	34.71	31.45	**32.20**	**13.51**	44.09	38.76
		3.0	31.32	13.45	38.23	36.09	32.74	13.80	54.37	45.07
	$1 \cdot t + 100$	1.5	54.84	31.04	72.20	50.87	**50.25**	**28.72**	**60.36**	**31.64**
		2.0	55.66	29.10	85.22	57.34	**52.96**	**27.01**	**72.74**	**35.36**
		3.0	53.08	26.45	94.84	57.29	**51.87**	26.89	**87.21**	**46.45**
200	$1 \cdot t + 25$	1.5	12.80	7.17	13.41	26.82	12.85	7.20	15.36	36.81
		2.0	12.54	7.29	14.68	28.15	12.57	7.35	18.61	41.40
		3.0	12.18	6.69	16.17	29.21	12.67	6.81	**15.43**	**28.47**
	$1 \cdot t + 50$	1.5	17.68	11.59	17.44	26.81	**16.25**	**10.76**	18.36	**24.74**
		2.0	16.42	11.02	17.97	26.26	**15.28**	**9.79**	23.58	28.29
		3.0	14.99	10.89	18.86	25.16	15.98	**10.59**	32.30	37.37
	$1 \cdot t + 100$	1.5	28.89	26.88	30.27	26.72	**26.66**	**25.34**	**28.69**	**24.26**
		2.0	27.19	24.66	33.32	29.58	**25.78**	**23.02**	34.47	29.86
		3.0	23.97	21.14	33.52	32.58	24.22	21.45	44.28	38.89

Table 2

The |RRMSE| and RB of estimators of GEV shape parameter (k) by the ML and TS methods

Parameters of the MC generator			H = T = GEV [%]			
	$\mu_t = 1 \cdot t + 100$		|RRMSE ($\hat{k}$)|		RBias ($\hat{k}$)	
N	μ_t	k (CS)	ML	TS	ML	TS
50	$1 \cdot t + 25$	–0.053 (1.5)	240.9	**211.5**	2.5	–41.6
		–0.108 (2.0)	122.8	**105.6**	5.3	–24.4
		–0.176 (3.0)	73.7	**64.9**	2.3	–19.5
	$1 \cdot t + 50$	–0.053 (1.5)	249.1	**217.5**	14.1	–41.2
		–0.108 (2.0)	122.1	**106.6**	7.2	–26.5
		–0.176 (3.0)	82.4	**69.2**	48.2	**–23.8**
	$1 \cdot t + 100$	–0.053 (1.5)	373.5	**215.3**	271.4	**101.5**
		–0.108 (2.0)	161.4	**97.5**	111.3	**24.2**
		–0.176 (3.0)	72.7	**61.4**	1.2	–1.3
100	$1 \cdot t + 25$	–0.053 (1.5)	164.6	**162.8**	2.3	–21.3
		–0.108 (2.0)	77.6	**74.2**	0.2	–12.5
		–0.176 (3.0)	48.1	**46.0**	0.2	–11.1
	$1 \cdot t + 50$	–0.053 (1.5)	167.3	**144.9**	59.0	**6.5**
		–0.108 (2.0)	77.1	**72.2**	14.6	**–9.7**
		–0.176 (3.0)	49.6	49.8	1.7	–12.3
	$1 \cdot t + 100$	–0.053 (1.5)	354.8	**183.5**	306.6	**123.7**
		–0.108 (2.0)	157.3	**80.3**	132.6	**43.6**
		–0.176 (3.0)	76.6	**45.4**	58.2	**10.1**
200	$1 \cdot t + 25$	–0.053 (1.5)	108.4	**93.3**	46.9	**19.7**
		–0.108 (2.0)	51.9	52.2	12.0	**–1.2**
		–0.176 (3.0)	29.8	32.9	–2.7	–5.3
	$1 \cdot t + 50$	–0.053 (1.5)	158.4	**108.9**	115.0	**50.4**
		–0.108 (2.0)	63.5	**49.4**	37.6	**9.9**
		–0.176 (3.0)	31.0	33.5	3.4	**–3.4**
	$1 \cdot t + 100$	–0.053 (1.5)	323.8	**170.6**	301.7	**142.6**
		–0.108 (2.0)	132.9	**69.3**	122.7	**52.3**
		–0.176 (3.0)	56.8	**35.3**	46.5	**16.9**

ry, the TS approach allows to focus on the estimation of just the shape parameter k by means of the L-skewness estimator (see Eq. B6), *i.e.*, in the second stage of TS algorithm. In the TS method, as a result of the transformation (see Eq. 4), one deals with the i.i.d. sample with $m_y = 0$ and $\sigma_y = 1$. According to the doctrine of parsimony in parameters, the separation of the shape parameter from the a, b, c, and d estimation should positively affect its accuracy and thus the competitiveness of the TS method for small and moderate-length time-series.

As one can see from Table 2, the |RRMSE($\hat{k}$)| of both methods does not show regular variations with the trend in the standard deviation but strongly decreases with increasing skewness (CS). As far as the estimator of $\hat{k}$ is concerned, the advantage of the TS method over the ML method is vivid (see figures in bold for the RRMSE) and it extends even to longer series ($N = 200$). The values of bias of the $\hat{k}$ estimator advocate towards strong advantage of the TS method over the ML one, when the series size is big.

There is another reason why the ML approach to the NFFA should be used with special caution. Note that when a false model is assumed, the errors of ($\hat{k}$) by two approaches will definitely rise. However, the rise will be more spectacular for the ML method.

5.3 The upper quantiles

No doubt, the errors of the parameters influence the values of the |RRMSE| of upper quantiles. The vivid differences in the results by two competing methods can be observed for flood quantiles in Table 3. As an example, we discuss the selected quantile probability of non-exceedance, $F = 0.99$, which corresponds to the maximum flow of a 100-year return period. For brevity, as examples we present two moments in time (t) related to the series size (N), namely $t = N/2$ and $t = N$. These moments in time are particularly interesting when analysing the parameters of hydrologic structures within the context of time. The values of the quantiles' relative root mean square error (|RRMSE($\hat{x}_F$)|) and relative bias for the two methods show that the TS method seems to be slightly superior, for hydrological size of time series, for 100-years-flood quantiles (Table 3) regardless the size of the series (N).

To recapitulate the numerical experiment, first it should be born in mind that its design takes into account the condition of asymptotic superiority of the ML estimates, as the generated series come from known distribution with known form of time-dependent moments, *i.e.*, H = True. Despite this, it appears that the WLS exceeds the accuracy of estimates of time-dependent moments obtained by the ML method for hydrological lengths of time-series. Moreover, the TS method proved to be better in terms of stability and accu-

Table 3

The |RRMSE($\hat{x}_F$)| and RB($\hat{x}_F$) for $F = 0.99$, $t = N/2$, and $t = N$ by the ML and TS methods

Parameters of the MC generator				$x_{F=0.99}$, H = T = GEV			
$\mu_t = 1 \cdot t + 100$				RRMSE ($\hat{x}_F$) [%]		RB($\hat{x}_F$) [%]	
N	σ_t	CS	t	ML	TS	ML	TS
50	$1 \cdot t + 25$	1.5	$N/2$	20.87	**16.69**	0.94	**–0.65**
			N	28.04	**23.19**	1.95	**–0.99**
		2.0	$N/2$	23.17	**17.96**	1.86	**–0.24**
			N	30.96	**24.86**	2.97	**–1.01**
		3.0	$N/2$	24.40	**20.52**	2.01	**–0.56**
			N	32.92	**29.10**	3.26	**–2.09**
	$1 \cdot t + 50$	1.5	$N/2$	24.34	**19.20**	1.81	**–0.67**
			N	31.79	**27.06**	1.35	**–1.03**
		2.0	$N/2$	25.29	**21.72**	2.13	**–0.59**
			N	31.24	31.97	2.65	**–0.56**
		3.0	$N/2$	27.01	**23.73**	1.69	**–1.19**
			N	35.34	**33.71**	2.33	–3.24
	$1 \cdot t + 100$	1.5	$N/2$	41.14	**23.47**	21.43	**6.18**
			N	51.27	**35.20**	22.68	**6.90**
		2.0	$N/2$	38.86	**24.49**	19.04	**4.34**
			N	49.21	**34.87**	20.31	**4.10**
		3.0	$N/2$	33.94	**28.52**	12.98	**2.05**
			N	43.56	**42.06**	13.40	**1.23**
100	$1 \cdot t + 25$	1.5	$N/2$	13.64	**12.85**	0.61	**0.32**
			N	16.64	**16.52**	1.05	**0.59**
		2.0	$N/2$	15.45	**13.95**	1.18	**–0.16**
			N	18.75	18.86	1.74	**–0.40**
		3.0	$N/2$	17.46	**15.70**	1.72	**0.01**
			N	20.92	21.45	2.52	**–0.57**
	$1 \cdot t + 50$	1.5	$N/2$	15.90	**13.86**	3.24	**1.27**
			N	18.17	18.37	2.07	**1.60**
		2.0	$N/2$	17.02	**15.06**	3.93	**0.52**
			N	20.28	20.47	2.77	**1.11**

to be continued

Table 3 (continuation)

Parameters of the MC generator				$x_{F=0.99}$, H = T = GEV			
$\mu_t = 1 \cdot t + 100$				RRMSE ($\hat{x}_F$) [%]		RB($\hat{x}_F$) [%]	
N	σ_t	CS	t	ML	TS	ML	TS
100	$1 \cdot t + 50$	3.0	$N/2$	18.42	**17.88**	–0.04	0.26
			N	21.46	24.78	–0.45	**–0.34**
	$1 \cdot t + 100$	1.5	$N/2$	35.20	**17.45**	23.42	**7.03**
			N	38.39	**23.25**	23.73	**7.31**
		2.0	$N/2$	34.91	**18.18**	22.48	**5.89**
			N	39.33	**26.62**	22.90	**6.44**
		3.0	$N/2$	30.21	**19.34**	16.95	**3.23**
			N	35.53	**27.98**	16.98	**1.50**
200	$1 \cdot t + 25$	1.5	$N/2$	10.80	**9.97**	2.11	**1.45**
			N	12.21	**12.03**	2.01	**1.56**
		2.0	$N/2$	11.28	**11.22**	1.03	**0.35**
			N	12.71	13.54	0.80	**–0.01**
		3.0	$N/2$	11.37	12.57	–1.11	**0.09**
			N	12.87	15.21	–1.08	**–0.35**
	$1 \cdot t + 50$	1.5	$N/2$	13.62	**11.10**	5.94	**2.97**
			N	14.47	**13.83**	4.52	**3.24**
		2.0	$N/2$	12.81	**11.82**	2.46	**1.75**
			N	14.06	15.24	2.15	**1.95**
		3.0	$N/2$	12.49	13.81	1.25	**0.64**
			N	13.90	17.76	1.14	**0.78**
	$1 \cdot t + 100$	1.5	$N/2$	26.78	**13.88**	21.01	**7.46**
			N	28.24	**17.06**	21.01	**7.40**
		2.0	$N/2$	23.90	**14.19**	17.88	**6.48**
			N	25.68	**18.55**	17.91	**6.64**
		3.0	$N/2$	19.20	**14.84**	10.26	**4.07**
			N	20.83	**20.20**	10.16	**3.73**

racy of the shape parameter and upper quantiles calculation stemming from the ultimate simplicity of the algorithms. In addition, the TS is easier to implement in a practical calculation soft-package even for seasonal approach (see Kochanek *et al.* 2012). Therefore, bearing in mind all the drawbacks of the ML approach, regardless of its sound theoretical background, in practical situations it is safer to use the TS approach for the presented case of a three parameter pdf with trend in the two first moments, because it gives similar results but with the lower cost of implementation. This thesis will be discussed in detail also in the next section, concerning practical application of both approaches.

6. ESTIMATION OF TRENDS AND FLOOD QUANTILES FOR REPRESENTATIVE POLISH RIVERS

Two pieces of software developed by the authors, based on the ML and TS methods, were used to estimate trends in moments and quantiles for three gauging stations located on three Polish rivers differing with regard to flood regime: Warsaw–Nadwilanówka (1921-2010) at the middle course of the River Vistula, Poznań at the River Warta (1822-2010), and Nowy Targ at the River Dunajec (1921-2010). The location of the gauging stations is presented on the map of Poland in Fig. 1, whereas the values of the annual maxima are shown in diagrams in Fig. 2. The annual maxima (AM) flows of the first station originate either from summer or winter seasons, of the second station – in vast majority from winter seasons, and in the third station (representing mountain regime) from the summer seasons. The River Warta daily stage records in Poznań (from the year 1822 till now) is the longest in Poland and one of the longest in Europe. The three AM time-series do not reveal statistically significant “shifts” due to the construction of upstream reservoirs, change of the land cover, water transfers, water intakes, *etc.*

Fig. 1. Map of Poland with three selected gauging stations.

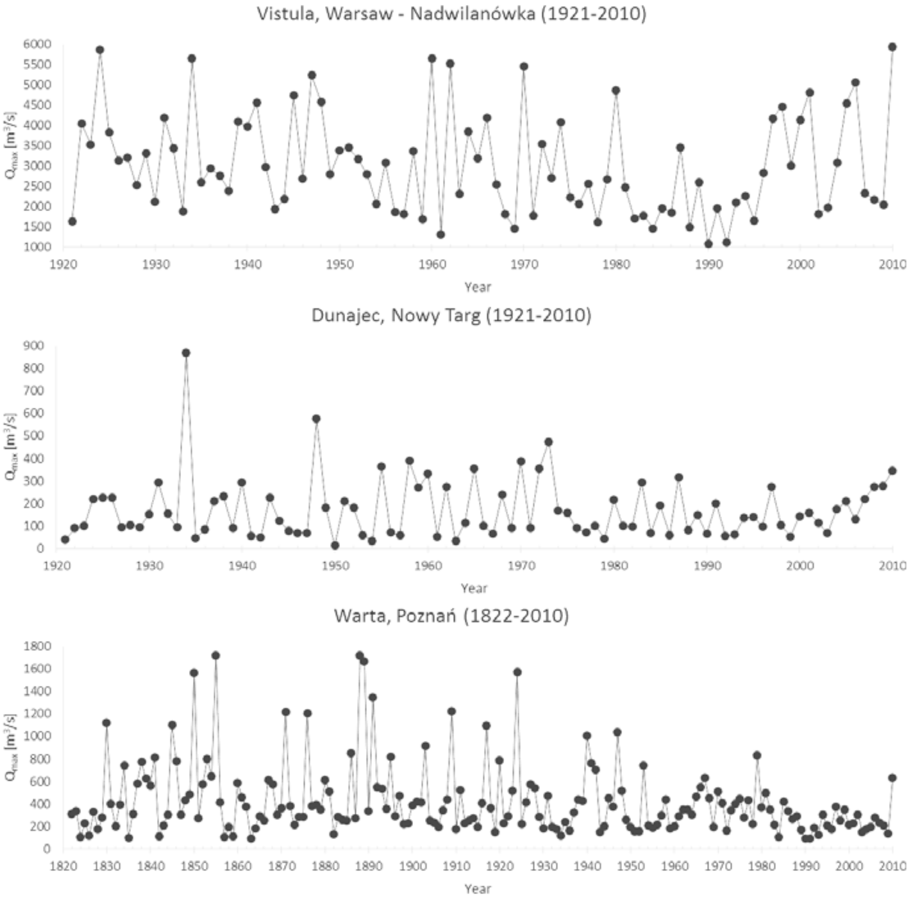

Fig. 2. Maximum annual peak flows for three selected gauging stations.

In Table 4, for the three AM time-series, the ML estimates of parameters of the non-stationary mean and standard deviation were displayed together with respective WLS estimates for the set of alternative distributions: 3p-Log-Normal, Pearson type 3, Gumbel, Generalised Extreme Value (GEV), Generalised Logistic (GLD), and Weibull. Table 4 confirms that the ML estimates of time-dependent moments are strongly distribution-dependent, which leads to the paradox: for the same time series but with different PDFs the trends differ in their values and even in directions. It is, undoubtedly, the weakness of this approach, because the model (distribution function) is usually fitted to the sample by means of more or less subjective methods, whereas the true form of the model remains unknown. Moreover, the ML estimates of slopes ($\hat{a},\hat{c}$) may even differ in sign. For example, for

Table 4

Estimators of trends in mean and standard deviation
for three exemplary Polish gauging stations

<table>
<tr><th rowspan="2">River</th><th rowspan="2">Gauging station</th><th rowspan="2">N</th><th rowspan="2">PDF</th><th colspan="4">ML</th><th colspan="4">WLS</th></tr>
<tr><th>a</th><th>b</th><th>c</th><th>d</th><th>a</th><th>b</th><th>c</th><th>d</th></tr>
<tr><td rowspan="7">Vistula</td><td rowspan="7">Warsaw Nadwilanówka</td><td rowspan="7">90</td><td>Log-Normal</td><td>–8.3</td><td>3407.8</td><td>0.6</td><td>1288.7</td><td rowspan="6">–8.1</td><td rowspan="6">3391.7</td><td rowspan="6">3.5</td><td rowspan="6">1049.3</td></tr>
<tr><td>Pearson type 3</td><td>–1.7</td><td>3022.9</td><td>3.7</td><td>1014.3</td></tr>
<tr><td>Gumbel*</td><td>–8.4</td><td>3386.6</td><td>1.5</td><td>1134.0</td></tr>
<tr><td>GEV</td><td>–8.7</td><td>3415.3</td><td>0.9</td><td>1250.7</td></tr>
<tr><td>GLD</td><td>–9.8</td><td>3585.1</td><td>–0.4</td><td>1904.7</td></tr>
<tr><td>Weibull</td><td>–1.0</td><td>3156.9</td><td>0.3</td><td>1343.9</td></tr>
<tr><td>SD</td><td>3.87</td><td>202.90</td><td>1.44</td><td>308.63</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td></tr>
<tr><td rowspan="7">Warta</td><td rowspan="7">Poznań</td><td rowspan="7">189</td><td>Log-Normal</td><td>–1.5</td><td>567.0</td><td>–1.5</td><td>455.0</td><td rowspan="6">–1.8</td><td rowspan="6">596.3</td><td rowspan="6">–2.0</td><td rowspan="6">486.3</td></tr>
<tr><td>Pearson type 3</td><td>0.0</td><td>436.4</td><td>0.0</td><td>313.5</td></tr>
<tr><td>Gumbel</td><td>–1.6</td><td>564.1</td><td>–1.4</td><td>372.8</td></tr>
<tr><td>GEV</td><td>–1.5</td><td>569.9</td><td>–1.8</td><td>550.0</td></tr>
<tr><td>GLD</td><td>–1.5</td><td>593.0</td><td>–3.0</td><td>950.0</td></tr>
<tr><td>Weibull*</td><td>–0.8</td><td>563.1</td><td>–0.8</td><td>471.1</td></tr>
<tr><td>SD</td><td>0.63</td><td>56.23</td><td>1.01</td><td>226.57</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td></tr>
<tr><td rowspan="7">Dunajec</td><td rowspan="7">Nowy Targ</td><td rowspan="7">90</td><td>Log-Normal*</td><td>0.0</td><td>172.6</td><td>–0.6</td><td>176.8</td><td rowspan="6">0.0</td><td rowspan="6">167.7</td><td rowspan="6">–1.2</td><td rowspan="6">182.4</td></tr>
<tr><td>Pearson type 3</td><td>0.0</td><td>170.0</td><td>0.0</td><td>122.1</td></tr>
<tr><td>Gumbel</td><td>0.0</td><td>163.2</td><td>-0.4</td><td>126.6</td></tr>
<tr><td>GEV</td><td>0.1</td><td>173.8</td><td>–0.5</td><td>240.9</td></tr>
<tr><td>GLD</td><td>0.1</td><td>187.3</td><td>–2.3</td><td>901.1</td></tr>
<tr><td>Weibull</td><td>–0.8</td><td>230.1</td><td>–1.0</td><td>194.2</td></tr>
<tr><td>SD</td><td>0.33</td><td>24.45</td><td>0.78</td><td>300.88</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.0</td></tr>
</table>

*)Best fitted PDF in terms of the AIC.

the Warsaw–Nadwilanówka gauge the Generalised Logistic Distribution (GLD) gives the negative trend in standard deviation (c = –0.4) while for the other models this trend is positive. Similar situation is observed for Nowy Targ and Weibull distribution function which is the only one giving a negative trend in mean (a = –0.8), and the other models show zero or minimal positive trend. Such a qualitative variability of the ML estimators undermines the credibility of the results of non-stationary flood frequency design and reveals, hidden by distributional assumption uncertainty range of the estimators. All these lead to the very dangerous conclusion that by manipulating with a variety of distribution models, one can obtain the "desired" results. In addition, there is no guarantee that the calculated trend will continue in the future, especially if its sign is negative.

Obviously, three gauging stations do not constitute a sample representative enough to draw any general conclusions, but still it is worth noting that the trend in mean for the three gauging stations is either negative or close to zero. On the other hand, the trends in standard deviation reveal bigger variability from relatively big negative value for Poznań and Nowy Targ to a big positive value for the Warsaw–Nadwilanówka.

Table 5

The coefficient of skewness estimators for three exemplary Polish gauging stations

Coefficient of skewness (CS)					
River	Gauging station	N	PDF	ML	TS
Vistula	Warsaw Nadwilanówka	90	Log-Normal	1.66	1.20
			Pearson type 3	1.66	1.08
			Gumbel	1.14	1.14
			GEV	1.65	1.22
			GLD	70.71	2.03
			Weibull	1.01	0.99
			CV (GLD excluded)	0.23	**0.08**
Warta	Poznań	189	Log-Normal	2.92	2.77
			Pearson type 3	1.49	1.95
			Gumbel	1.14	1.14
			GEV	13.89	4.50
			GLD	N.E.	43.78
			Weibull	1.97	1.94
			CV (GLD excluded)	1.26	**0.52**
Dunajec	Nowy Targ	90	Log-Normal	3.53	0.00
			Pearson type 3	1.61	0.00
			Gumbel	1.14	1.14
			GEV	N.E.	–0.02
			GLD	N.E.	–114.52
			Weibull	1.58	0.02
			CV (GEV and GLD excluded)	**0.54**	1.96

Explanation: N.E. – non existing.

The issue of significance of the detected trends within the context of climate change deserves here a few words of comment. A typical hydrological sample size implies large uncertainty and that is why the smoking gun proof of the climate-change-induced trends based on statistical analysis of floods

Table 6

The upper quantiles estimators for selected moments in time for three exemplary Polish gauging stations

Quantiles $x(F, t)$ [m^3/s]				ML				TS			
				$t = N/2$		$t = N$		$t = N/2$		$t = N$	
River	Gauging station	N	PDF	F=0.9	F=0.99	F=0.9	F=0.99	F=0.9	F=0.99	F=0.9	F=0.99
Vistula	Warsaw Nadwilanówka	90	Log-Normal	4710.2	7485.7	4371.0	7201.1	5390.9	7715.8	5336.3	7966.8
			Pearson type 3	4533.3	6666.3	4677.5	7105.4	5397.1	7490.9	5343.2	7712.4
			Gumbel	4573.7	6771.4	4281.5	6599.2	5380.3	7668.2	5324.2	7913.0
			GEV	4647.1	7372.0	4306.7	7114.0	5396.5	7787.3	5342.5	8047.7
			GLD	4847.5	9757.0	4388.5	9249.5	5334.5	8529.2	5272.4	8887.2
			Weibull	4971.3	7178.9	4942.2	7169.2	5391.5	7258.4	5336.9	7449.2
			CV	0.04	0.15	0.06	0.13	**0.00**	**0.06**	**0.01**	**0.06**
Warta	Poznań	189	Log-Normal	786.2	1604.6	480.7	929.1	1020.3	1912.6	488.4	832.3
			Pearson type 3	849.0	1536.7	849.0	1536.7	1034.5	1725.0	493.9	760.0
			Gumbel	726.6	1167.4	402.7	601.2	956.4	1469.7	463.7	661.6
			GEV	774.4	1817.6	478.5	1050.2	1003.9	2034.0	482.0	879.1
			GLD	790.5	2172.3	499.7	1288.8	981.0	2154.1	473.2	925.4
			Weibull	1028.0	1961.3	850.7	1606.1	1034.8	1720.6	494.0	758.3
			SD	0.13	0.21	0.34	0.33	**0.03**	**0.13**	**0.03**	**0.12**
Dunajec	Nowy Targ	90	Log-Normal	337.7	754.4	309.1	654.4	427.7	770.2	324.8	528.3
			Pearson type 3	332.7	579.7	332.7	579.7	431.4	709.1	326.9	492.0
			Gumbel	303.5	498.9	278.5	437.5	406.6	637.2	312.2	449.3
			GEV	335.3	914.0	324.5	848.0	423.8	809.7	322.4	551.8
			GLD	348.4	1096.5	330.9	984.4	414.3	872.0	316.8	588.8
			Weibull	396.4	701.5	305.5	525.3	431.6	697.1	327.1	484.9
			SD	0.09	0.29	0.07	0.31	**0.02**	**0.11**	**0.02**	**0.10**

data does not exist yet (Hattermann *et al.* 2013, Kundzewicz *et al.* 2014). Since there is a strong pressure to assess expected changes for engineering practice with regard to ensuring an adequate safety level of hydraulic structures, the climate change is assumed to cause trends in flood regime (even weak). To account for trends in the most reasonable way, one should find the appropriate time-dependent upper quantiles of peak flows which are the base of design procedures. The trends cannot be assessed apart from the data; it is important to avoid other sources of uncertainty, *i.e.*, the distribution choice and its parameters estimation as long as possible.

Table 5 provides a comparison of the asymmetry coefficient estimators ($\hat{\mathrm{CS}}$) of both methods for various alternative models of the three analyzed time-series. In the case of the ML method, the shape parameter of a distribution model (but Gumbel) is estimated together with time-dependent moments' parameters (a, b, c, d) by solving the ML equations and then $\hat{\mathrm{CS}}$ is computed from it.

In the TS method, after WLS estimation of time-dependent parameters of moments and then the estimation of L-moments skewness (τ_3) from the sample given by Eq. 4, the coefficient of asymmetry is computed from the functional relationship $\mathrm{CS} = \varphi(\tau_3)$ for a given distribution. As can be seen from a comparison of Table 5, the TS estimates of CS are mostly less sensitive to the distribution choice than ML estimates (compare the CS values in bold).

Table 6 shows the values of the upper quantiles estimators for the selected moments in time $t = N/2$ and $t = N$. As one can see, the TS method produces in general higher values of quantiles estimators than the ML one. Interestingly, the variability of the quantile estimates with regard to model choice is by far lower for the TS than for the ML approach (CV values in bold). This can be a fundamental fact supporting the use in the NFFA of the TS method rather than the ML one when the population model (PDF) of annual flow maxima is unknown.

7. CONCLUSIONS

In the FFA as models of probability distributions the three-parameter functions containing location, scale, and shape parameter are assumed, when the length of the sample permits. In particular, non-stationary approach to the FFA (NFFA) makes sense only when appropriately long time-series is available. Since the values of upper quantiles are more sensitive to the change in standard deviation than to the change in the mean, in the NFFA a trend in both the mean and in standard variation is to be assumed. Due to the limited length of hydrological time-series, the skewness is usually considered to be time-independent. Similarly, the adoption of time-invariant distribution function is indispensable for the NFFA. Keeping in mind the doctrine of parsimony in the number of parameters estimated to extend the FFA for non-stationarity case, only the two parameters have been added by introducing one of the possible options of a linear trend in both the mean and standard deviation.

The purpose of this study was to compare the efficiency of the two NFFA methods, namely the two-stage method (TS), and the maximum likelihood (ML) with the presence of time as a covariate. The comparison was made by means of Monte Carlo (MC) simulation experiments assuming linear trends in the mean and standard deviation as well as using the data of three representative rivers in Poland. The same distribution served in MC experiments both as the parent (True) and hypothetical (False) distribution, which constitutes the most favourable case for application of ML method. The GEV distribution was taken for the purpose and the real hydrological conditions were maintained, *e.g.*, parameters' ranges, length of typical time series natural trends in mean, and standard deviation and moderate constant skewness. The accuracy of the two methods was analysed in respect to estimates of time-dependent first two moments, time-invariant skewness, and time-dependent upper quantiles. The results of the experiment showed that the TS method generally proved to perform better than the ML one in terms of all three aspects listed above. Note, also, that although the WLS methodology (first stage of the TS) aims at samples of moderate value of skewness, the numerical experiment revealed its very good performance even for larger skewness coefficients (CS checked up to 3). Moreover, the TS is more numerically stable, *i.e.*, gives reliable results for every non-stationary time-series, while the ML sometimes fails or gives results impossible to verify in practice. What is more important, as the TS estimates of time-dependent moments do not rely on the distribution function, the estimates of skewness and, additionally, of time-dependent upper quantiles are less sensitive to distributional choice than in the ML method, as demonstrated for the Polish rivers' case studies. All these, together with the simplicity of algorithm, elevate the status of the TS approach above other methods used in the non-stationary FFA (NFFA). One can expect to get a similar result of the comparison of methods for any other three-parameter distributions with linear trend in first two moments. While the increase in the number of estimated parameters enlarges further the dominance of the TS method in all these respects. And, conversely, the reduction of the estimated parameters reduces the competitiveness of the TS method as such. For instance, the Gumbel distribution with linear trend in the mean reveals the advantage of the ML method (*e.g.*, Clarke 2002a, b). Dominance of the TS is expected to be greater in the natural conditions when we do not know the true distribution, so the hypothetic distribution is false, and asymptotic properties of the ML method do not hold. Then the ML parameter estimates are additionally biased by the model error, which may badly affects their accuracy.

For Polish rivers' case studies assuming various distributions, similarly to the MC experiments, the competition between the two methods was carried out in three criteria: dispersion of results of trends in mean and standard deviation, dispersion of the skewness coefficients, and time-dependent quantiles. In all three criteria the Polish case study also confirms the predominance of the TS over the ML approach, since the dispersion (CV, SD) of the TS results by different assumed PDFs is generally smaller than for the ML approach. The annual peak flow series of Polish rivers are mixtures of summer and winter flows, and for some cases there is no clearly dominant season. Note that both NFFA methods, *i.e.*, the ML and TS, can be extended in the seasonal approach to modelling of annual peak flows. In such cases it is advisable to use the seasonal approach to flood frequently modelling (Strupczewski *et al.* 2012, Kochanek *et al.* 2012), but its detailed description exceeds the scope of this paper.

The Polish case study revealed relatively large values of trends in mean and standard deviation. These point out that the use of stationary FFA would lead to the evident simplification, erroneous results, and decisions. So, when we know that the process is non-stationary, non-stationary methods should be also used for the analysis. On the other hand, the non-stationary modelling of complex hydrological phenomena is still difficult. Despite the gradual progress (*e.g.*, Montanari *et al.* 2013, Hall *et al.* 2014, Vogel *et al.* 2013, Machado *et al.* 2015), this area still involves huge efforts to meet the requirements of flood risk assessment in a non-stationary water regime.

Acknowledgements. This work was carried out within statutory activities No. 3841/E-41/S/2015 of the Ministry of Science and Higher Education of Poland and partially supported Norway Grants Programme and National Center for Research and Development of Poland within project Climate Change Impact on Hydrological Extremes (CHIHE) – No. DZP/POL-NOR/1873/2013.

Appendix A

The Weighted Least Squares (WLS) method

The principle of the WLS method is based on the minimisation of sums of weighted squared deviations of observed and estimated moments, where the weights are reciprocals of their expected values.

For the first moment we get

$$\sum WS_t^{(1)} = \sum_{t=1}^{T} \gamma_t^{(1)} \left(x_t - \hat{m}_t\right)^2 \ , \tag{A1}$$

where $\hat{m}_t = f_1(\mathbf{g},t)$ while the weight $\gamma_t^{(1)}$ is given by

$$\gamma_t^{(1)} = \frac{1}{\mathrm{E}\left(x_t - \hat{m}_t\right)^2} = \frac{1}{\sigma^2\left(x_t\right)} \ . \tag{A2}$$

Applying the WLS method in respect to the second central moment, we have:

$$\sum WS_t^{(2)} = \sum_{t=1}^{T} \gamma_t^{(2)} \left[\varepsilon_t^2 - \hat{\mu}_2\left(x_t\right)\right]^2 \ , \tag{A3}$$

where $\varepsilon_t = x_t - m_t$, $\hat{\mu}_2\left(x.t\right) \equiv \hat{\sigma}^2\left(x,t\right) = f_2\left(\mathbf{h},t\right)$, and $\gamma_t^{(2)}$ which for the distribution with the time invariant skewness takes the form

$$\gamma_t^{(2)} = \frac{1}{\sigma^4\left(x_t\right)\left[\varphi\left(c_s\right) - 1\right]} \ . \tag{A4}$$

The conditions of minimum of the sum of weighted squares $\left(\Sigma WS_t^{(1)}\right)$ with respect to the **g** vector of parameters are:

$$\sum_{t=1}^{T} \gamma_t^{(1)} \left(x_t - \hat{m}_t\right) \frac{d\hat{m}_t}{d\mathbf{g}} = 0 \tag{A5}$$

and of weighted squares $\left(\Sigma WS_t^{(2)}\right)$ with respect to **h** are

$$\sum_{t=1}^{T} \frac{1}{\sigma_t^4} \left(\varepsilon_t^2 - \sigma_t^2\right) \frac{d\sigma_t^2}{d\mathbf{h}} = 0 \ . \tag{A6}$$

Note that each of two sets of equations contains both time dependent mean (m_t) and variance $\left(\mu_2(t)=\sigma_t^2\right)$, *i.e.*, they would be solved jointly unless the standard deviation is assumed constant.

The WLS, being conceptually quite distinct from the ML-method, coincides with the ML method in the case of normally distributed data. In this case, a simple presentation of the WLS as a problem of the ML-estimation is possible.

Appendix B

Generalized Extreme Value (GEV) distribution and its summary statistics

The probability density distribution function (*e.g.*, Rao and Hamed 2000):

$$f(x)=\frac{1}{\alpha}\left[1-k\left(\frac{x-u}{\alpha}\right)\right]^{\frac{1}{k}-1}\cdot\exp\left\{-\left[1-k\left(\frac{x-u}{\alpha}\right)\right]^{\frac{1}{k}}\right\} \tag{B1}$$

and its cumulative distribution function (CDF):

$$F(x)=\exp\left[-\left(1-k\frac{x-u}{\alpha}\right)^{\frac{1}{k}}\right] \tag{B2}$$

where *u*, *a*, *k* are the parameters of location, scale and shape, respectively.

The first two conventional moments

$$\mu=u+\frac{\alpha}{k}\left[1-\Gamma(1+k)\right],\quad \sigma=\frac{\alpha}{|k|}\sqrt{\Gamma(1+2k)-\Gamma^2(1+k)}\ . \tag{B3a, b}$$

The coefficient of asymmetry (skewness), see also Fig. B1:

$$\mathrm{CS}=\frac{k}{|k|}\frac{\left[-\Gamma(1+3k)+3\Gamma(1+k)\Gamma(1+2k)-2\Gamma^3(1+k)\right]}{\left[\Gamma(1+2k)-\Gamma^2(1+k)\right]^{\frac{3}{2}}}\ . \tag{B4}$$

Since the trend is introduced to the two first moments (μ and σ), the parameters of the GEV distribution function should be presented as a function of those moments and a time-independent shape parameter, $k = k(\mathrm{CS})$:

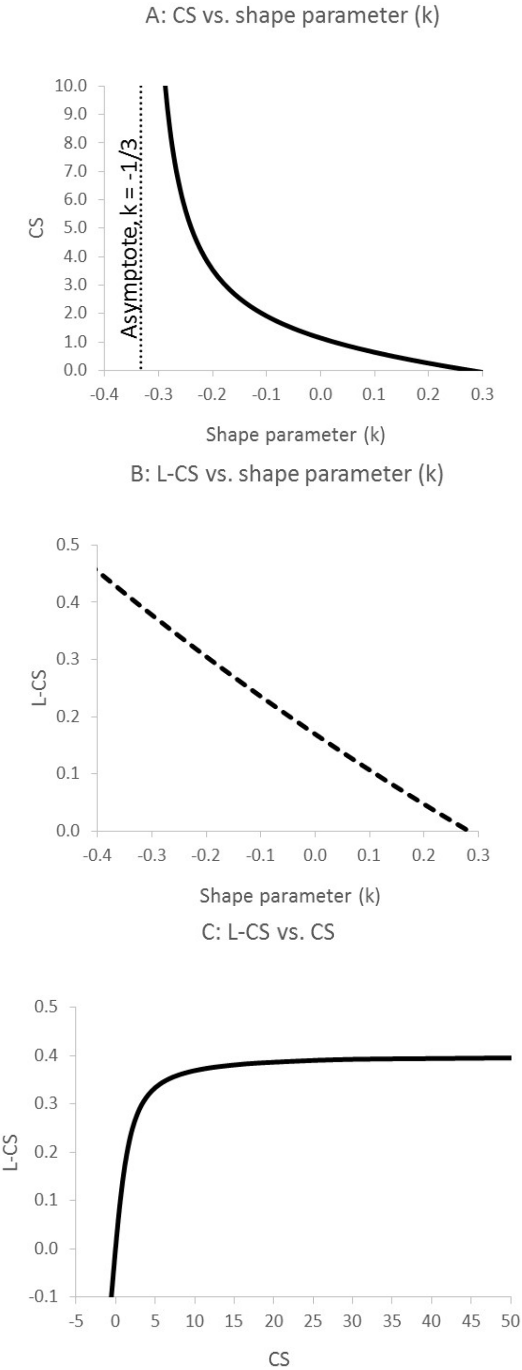

Fig. B1. Relationships of GEV skewness measures.

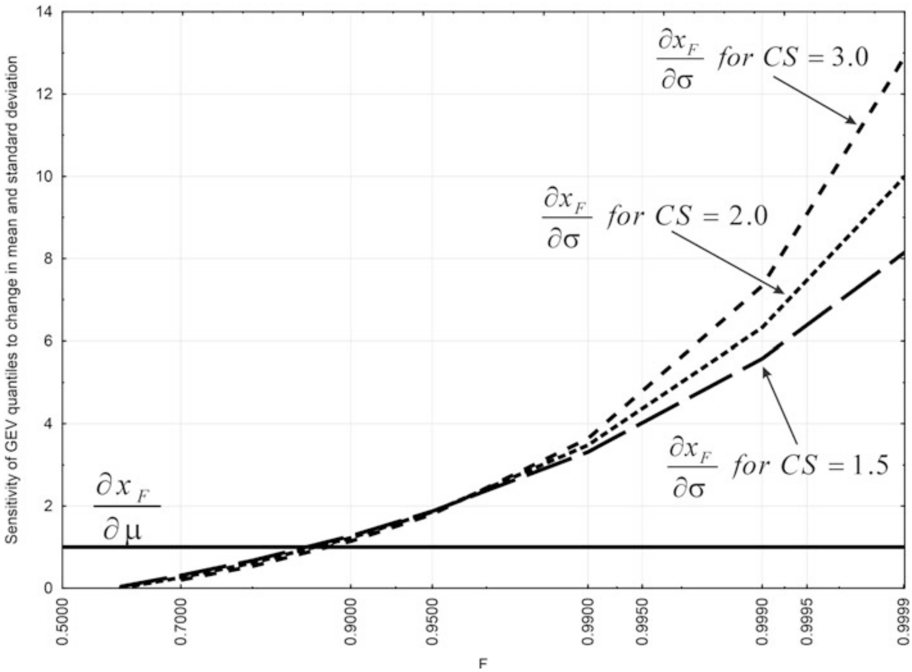

Fig. B2. The GEV quantile sensitivity to parameters regarded to the probability of non-exceedance.

$$\alpha = \sigma \cdot \varphi(k) \ , \tag{B5}$$

$$u = \mu - \sigma \cdot \varphi(k) \cdot \varphi(k) \ , \tag{B6}$$

where

$$\varphi(k) = |k| \left[\Gamma(1+2k) - \Gamma^2(1+k) \right]^{-\frac{1}{2}} \quad , \quad \phi(k) = \frac{1-\Gamma(1+k)}{k} \ .$$

The first two *L*-moments (*e.g.*, Hosking and Wallis 1997):

$$\lambda_1 = u + \frac{\alpha}{k}\left[1-\Gamma(1+k)\right], \qquad \lambda_2 = \alpha\left(1-2^{-k}\right)\Gamma(1+k)/k \ .' \tag{B7a, b}$$

The L-skewness coefficient

$$\text{L-CS:} \quad \tau_3 = 2\frac{1-3^{-k}}{1-2^{-k}} - 3 \ . \tag{B8}$$

The quantile function, *i.e.*, the inverted Eq. B2

$$x(F) = u + \frac{\alpha}{k}\left[1-(-\log F)^k\right] \ . \tag{B9}$$

The quantile function in (μ, σ, k) re-parametrisation:

$$x(F) = \mu - \sigma \cdot \varphi \cdot \phi + \sigma \cdot \varphi \cdot \mathrm{K}(k, F) \tag{B10}$$

where

$$\mathrm{K}(k, F) = \frac{1 - \left(-\ln(F)\right)^k}{k} \; .$$

The sensitivity of the quantile function to the two first moments:

$$\frac{\partial x}{\partial \mu} = 1 \; , \quad \frac{\partial x}{\partial \sigma} = \mathrm{K} \cdot \varphi - \varphi \cdot \phi \; . \tag{B11a, b}$$

For the parameters used in the Monte Carlo experiment, $k = -0.053$, -0.108, and -0.176 (CS = 1.5, 2.0, 3.0), the relationship of the sensitivity to the probability of non-exceedance is presented in Fig. B2.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Automatic Control*. **19**, 6, 716-723, DOI: 10.1109/TAC.1974.1100705.

Clarke, R.T. (2002a), Fitting and testing the significance of linear trends in Gumbel-distributed data, *Hydrol. Earth Syst. Sci.* **6**, 1, 17-24, DOI: 10.5194/hess-6-17-2002.

Clarke, R.T. (2002b), Estimating time trends in Gumbel-distributed data by means of generalized linear models, *Water Resour. Res*. **38**, 7, 16-1–16-11, DOI: 10.1029/2001WR000917.

Davison, A.C., and R.L. Smith (1990), Models for exceedances over high thresholds, *J. Roy. Statist. Soc. B* **52**, 3, 393-442.

Feluch, W. (1994), *Selected Methods of Kernel Estimation of Probability Density Function and Regression in Hydrology*, Prace Naukowe – Inżynieria Środowiska, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 85 pp. (in Polish).

Hall, J., B. Arheimer, M. Borga, R. Brázdil, P. Claps, A. Kiss, T.R. Kjeldsen, J. Kriaučiūnienė, Z.W. Kundzewicz, M. Lang, M.C. Llasat, N. Macdonald, N. McIntyre, L. Mediero, B. Merz, R. Merz, P. Molnar, A. Montanari, C. Neuhold, J. Parajka, R.A.P. Perdigão, L. Plavcová, M. Rogger, J.L. Salinas, E. Sauquet, C. Schär, J. Szolgay, A. Viglione, and G. Blöschl (2014), Understanding flood regime changes in Europe: A state-of-the-art assessment, *Hydrol. Earth Syst. Sci.* **18**, 7, 2735-2772, DOI: 10.5194/hess-18-2735-2014.

Hattermann, F.F., Z.W. Kundzewicz, S. Huang, T. Vetter, F.-W. Gerstengarbe, and P. Werner (2013), Climatological drivers of changes in flood hazard in Germany, *Acta Geophys.* **61**, 2, 463-477, DOI: 10.2478/s11600-012-0070-4.

Hosking, J.R.M. (1990), L-moments: analysis and estimation of distributions using linear combinations of order statistics, *J. Roy. Statist. Soc. B* **52**, 1, 105-124.

Hosking, J.R.M., and J.R. Wallis (1997), *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, 224 pp.

Hosking, J.R.M., J.R. Wallis, and E.F. Wood (1985), Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics* **27**, 3, 251-261, DOI: 10.1080/00401706.1985.10488049.

Hurvich, C.M., and C.-L. Tsai (1989), Regression and time series model selection in small samples, *Biometrika* **76**, 2, 297-307, DOI: 10.1093/biomet/76.2.297.

Kendall, M.G. (1975), *Rank Correlation Methods*, 4th ed., Charles Griffin, London.

Kochanek, K., W.G. Strupczewski, and E. Bogdanowicz (2012), On seasonal approach to flood frequency modelling. Part II: Flood frequency analysis of Polish rivers, *Hydrol. Process.* **26**, 5, 717-730, DOI: 10.1002/hyp.8178.

Kochanek, K., W.G. Strupczewski, E. Bogdanowicz, W. Feluch, and I. Markiewicz (2013), Application of a hybrid approach in nonstationary flood frequency analysis – a Polish perspective, *Nat. Hazards Earth Syst. Sci. Discuss.* **1**, 5, 6001-6024, DOI: 10.5194/nhessd-1-6001-2013.

Kundzewicz, Z.W., S. Kanae, S.I. Seneviratne, J. Handmer, N. Nicholls, P. Peduzzi, R. Mechler, L.M. Bouwer, N. Arnell, K. Mach, R. Muir-Wood, G.R. Brakenridge, W. Kron, G. Benito, Y. Honda, K. Takahashi, and B. Sherstyukov (2014), Flood risk and climate change: global and regional perspectives, *Hydrolog. Sci. J.* **59**, 1, 1-28, DOI: 10.1080/02626667.2013. 857411.

Machado, M.J., B.A. Botero, J. López, F. Francés, A. Díez-Herrero, and G. Benito (2015), Flood frequency analysis of historical flood data under stationary and non-stationary modelling, *Hydrol. Earth Syst. Sci. Discuss.* **12**, 525-568, DOI: 10.5194/hessd-12-525-2015.

Markiewicz, I., E. Bogdanowicz, and W.G. Strupczewski (2014), On the accuracy of stability of estimators of probable maximum flows. **In:** K. Banasik, L. Hejduk, and E. Kaznowska (eds.), *Hydrologia w Inżynierii i Gospodarce Wodnej*, Monografie KGW-PAN, Komitet Gospodarki Wodnej PAN, Warszawa, 81-93 (in Polish).

Matalas, N.C., and M.A. Benson (1968), Note on the standard error of the coefficient of skewness, *Water Resour. Res.* **4**, 1, 204-205, DOI: 10.1029/ WR004i001p00204.

Milly, P.C.D., J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer (2008), Stationarity is dead: Whither water management? *Science* **319**, 5863, 573-574, DOI: 10.1126/science. 1151915.

Mitosek, H.T., W.G. Strupczewski, and V.P. Singh (2006), Three procedures for selection of annual flood peak distribution, *J. Hydrology* **323**, 1-4, 57-73, DOI: 10.1016/j.jhydrol.2005.08.016.

Montanari, A., G. Young, H.H.G. Savenije, D. Hughes, T. Wagener, L.L. Ren, D. Koutsoyiannis, C. Cudennec, E. Toth, S. Grimaldi, G. Blöschl, M. Sivapalan, K. Beven, H. Gupta, M. Hipsey, B. Schaefli, B. Arheimer, E. Boegh, S.J. Schymanski, G. Di Baldassarre, B. Yu, P. Hubert, Y. Huang, A. Schumann, D.A. Post, V. Srinivasan, C. Harman, S. Thompson, M. Rogger, A. Viglione, H. McMillan, G. Characklis, Z. Pang, and V. Belyaev (2013), "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013-2022, *Hydrolog. Sci. J.* **58**, 6, 1256-1275, DOI: 10.1080/02626667.2013.809088.

Pettitt, A.N. (1979), A non-parametric approach to the change-point problem, *J. Roy. Statist. Soc. C* **28**, 2, 126-135, DOI: 10.2307/2346729.

Rao, A.R., and K.H. Hamed (2000), *Flood Frequency Analysis*, CRC Press, Boca Raton, 350 pp.

Rigby, R.A., and D.M. Stasinopoulos (2005), Generalized additive models for location, scale and shape, *J. Roy. Statist. Soc. C* **54**, 3, 507-554, DOI: 10.1111/j.1467-9876.2005.00510.x.

Rosenblatt, M. (1956), Remarks on some nonparametric estimates of a density function, *Ann. Math. Statist.* **27**, 3, 832-837, DOI: 10.1214/aoms/1177728190.

Strupczewski, W.G. (1999), A farewell to the ML method in flood frequency analysis. **In:** V.P. Singh, I.W. Seo, and J.H. Sonu (eds.), *Hydrologic Modeling. Proceedings of WEESHE*, Water Resources Publications, LLC, Highlands Ranch, 291-306.

Strupczewski, W.G., and W. Feluch (1997), System of identification of an optimum flood frequency model with time dependent parameters (IDT). **In:** N.B. Harmancioglu, M.N. Alpaslan, S.D. Ozkul, and V.P. Singh (eds.), *Integrated Approach to Environmental Data Management Systems*, NATO ASI Series, Vol. 31, Kluwer Acad. Publ., 291-300, DOI: 10.1007/978-94-011-5616-5_25.

Strupczewski, W.G., and W. Feluch (1998), Investigation of trend in annual peak flow series. Part I. Maximum likelihood estimation. **In:** *Proc. 2nd Int. Conf. on Climate and Water – A 1998 Perspective, 17-20 August 1998, Espoo, Finland*, Vol. 1, 241-250.

Strupczewski, W.G., and Z. Kaczmarek (1998), Investigation of trend in annual peak flow series. Part II. Weighted Least Squares estimation. **In:** *Proc. 2nd Int. Conf. on Climate and Water – A 1998 Perspective, 17-20 August 1998, Espoo, Finland*, Vol. 1, 251-263.

Strupczewski, W.G., and Z. Kaczmarek (2001), Non-stationary approach to at-site flood frequency modelling. Part II. Weighted least squares estimation, *J. Hydrol.* **248**, 1-4, 143-151, DOI: 10.1016/S0022-1694(01)00398-5.

Strupczewski, W.G., and H.T. Mitosek (1991), How to deal with nonstationary time series in the hydrologic projects. **In:** *IAHS Symposium "Mitteilungsblatt des Hydrographischen Dienstes in Osterreich"*, No. 65/66, Vienna, Austria, 36-40.

Strupczewski, W.G., and H.T. Mitosek (1995), Some aspects of hydrological design under non-stationarity. **In:** Z. Kundzewicz (ed.), *New Uncertainty Concepts in Hydrology and Water Resources*, Cambridge Univ. Press, Cambridge, 39-44.

Strupczewski, W.G., and H.T. Mitosek (1998), Investigation of trend in annual peak flow series. Part III Flood analysis of Polish rivers. **In:** *Proc. 2nd Int. Conf. on Climate and Water – A 1998 Perspective, 17-20 August 1998, Espoo, Finland*, Vol. 1, 264-272.

Strupczewski, W.G., V.P. Singh, and W. Feluch (2001a), Non-stationary approach to at-site flood frequency modelling. Part I. Maximum likelihood estimation, *J. Hydrol.* **248**, 1-4, 123-142, DOI: 10.1016/S0022-1694(01) 00397-3.

Strupczewski, W.G., V.P. Singh, and H.T. Mitosek (2001b), Non-stationary approach to at-site flood frequency modelling. Part III. Flood analysis of Polish rivers, *J. Hydrol.* **248**, 1-4, 152-167, DOI: 10.1016/S0022-1694(01) 00399-7.

Strupczewski, W.G., V.P. Singh, and S. Weglarczyk (2002a), Asymptotic bias of estimation methods caused by the assumption of false probability distribution, *J. Hydrol.* **258**, 1-4, 122-148, DOI: 10.1016/S0022-1694(01)00563-7.

Strupczewski, W.G., S. Węglarczyk, and V.P. Singh (2002b), Model error in flood frequency estimation, *Acta Geophys. Pol.* **50**, 2, 279-319.

Strupczewski, W.G., H.T. Mitosek, K. Kochanek, V.P. Singh, S. Weglarczyk (2006), Probability of correct selection from lognormal and convective diffusion models based on the likelihood ratio, *Stoch. Environ. Res. Risk Assess.* **20**, 3, 152-163, DOI: 10.1007/s00477-005-0030-5.

Strupczewski, W.G., K. Kochanek, E. Bogdanowicz, and I. Markiewicz (2012), On seasonal approach to flood frequency modelling. Part I: Two-component distribution revisited, *Hydrol. Process.* **26**, 5, 705-716, DOI: 10.1002/hyp. 8179.

Vogel, R.M., A. Rosner, and P.H. Kirshen (2013), Brief Communication: Likelihood of societal preparedness for global change: trend detection, *Nat. Hazards Earth Syst. Sci.* **13**, 7, 1773-1778, DOI: 10.5194/nhess-13-1773-2013.

Wallis, J.R., N.C. Matalas, and J.R. Slack (1974), Just a moment! *Water Resour. Res*. **10**, 2, 211-219, DOI: 10.1029/WR010i002p00211.

Yevjevich, V., and J.T.B. Obeysekera (1984), Estimation of skewness of hydrologic variables, *Water Resour. Res.* **20**, 7, 935-943, DOI: 10.1029/WR020i007 p00935.

Received 14 July 2015
Received in revised form 10 November 2015
Accepted 1 December 2015