Inundation risk for embanked rivers.

W. G. Strupczewski, K. Kochanek, E. Bogdanowicz and I. Markiewicz

Answer to reviews

The authors would like to thank very much the anonymous reviewers for careful reading and criticism. We found the remarks and questions thought provoking and inspiring. We believe that the alterations to the manuscript induced by the Reviewers' remarks will positively influence the quality of the paper and satisfy the Referees. Below we present the detailed answers to the remarks of both Reviewers. We also attached a separate file with the corrected manuscript where the alterations induced by the first Reviewer are marked by blue background of the text whereas the Reviewer's #2 ones by yellow.

Reviewer #1

Questions: The authors of this paper adapted and paraphrase the well known approach developed by Davison and Smith (1990) for statistical modeling the extremes of a non-stationary process of exceedances over thresholds by the usage of the generalized Pareto distribution (GPD). Due to this approach the rate and magnitude of exceedances over threshold can be modeled as binary and GPD regression models as functions of the covariates. Standard software procedures for fitting the components of these models are widely and freely available, e.g. Yee and Stephenson (2007) and Gilleland et al. (2013). However, all these papers, for some reason, are completely ignored by these authors.

Answers:
Thank you very much for this remark. We also recognize the Davison and Smith's article (1990) very interesting and inspiring. Paraphrasing the opinion about a popular book in the field of probability theory, we can state that many papers in hydrology are written by mathematicians and have the build in bias that the reader is assumed to be mathematician coming to the material for its beauty. Our paper is geared towards hydrologists-practitioners whose primary focus is neither mathematics nor statistics for their own sake (i.e. hydrologic realism is present).

It is declared goal of the paper to enforce the engineering design procedures by providing a detail background of new, in this domain, approach to flood risk assessment. This approach involves taking into account not only the magnitude of peak flow, as in classical FFA, but also the duration of high waters, the factor leading to embankments failure. Therefore, the focus of the paper is **not** about Peak Over Threshold (POT) (PDS) approach aimed on better description of the upper tail of the annual maxima distribution but about combine modelling and flood risk assessment of both: magnitude and duration of high waters, so the shape of flood waves. You write that in our work we adapted and paraphrased the well known POT approach developed by Davison and Smith (1990) for statistical modelling the extremes of a non-stationary process. Although it is not the case, what we mention on page 2, lines 14-18, your words provoked us to consider the possibility to use all independent floods above the Q_A threshold in our analysis.

First of all, there is no gain in accuracy of upper quantile estimates when applying POT to low annual arrival rate [e.g. 8]. Therefore, it may be worth a candle to apply POT only in limited area of the upland part of Poland when winter and summer floods are almost equivalent. The hydrological regime of other Polish rivers shows strong domination of winter (northern Poland) and summer (southern Poland) floods which may result in just a slight (if any) increase of informative value of the POT with regard to annual maxima series. To this effect we adopted the seasonal approach [28 and 5]. For example, in our Case Study – the Szczucin gauge (southern Poland) at the Vistula River, received the Poisson rate equal to 0.482 for floods exceeding the threshold Q_A. In the course of just 4 of the 56 years of record, the second flood occurred in excess of Q_A but it lasted no longer than one day and therefore did not affect the increased risk of flooding.

If we understand the essence of Davison's *et al*. (1990) proposal concerning the POT model parameters estimation (for both stationary and non-stationary case) by ML method, the Poisson parameter is estimated from AM series as the GEV parameter while having two other parameters of GEV estimated from the series of all exceedances over threshold,. i.e. as the parameters of the Generalized Pareto distribution. Hence, it makes the estimates of the Poisson and Generalized Pareto (GP) parameters dependent on each other. This is a significant modification and departure from the original design of the POT models. The fact in itself of getting here a higher value Likelihood function for GEV does not give preference to this modification. We estimated the Poisson process arrival rate from the time-series of annual numbers of the threshold exceedances both for stationary [7-9, 29] and non-stationary assumption [19]. In the case study considered by us the discrete-continuous probability distribution function, while using ML method, the independence of the estimates of the occurrence probability, i.e. the weighting factor $(1 - \beta)$), and the parameters of the pdf of annual maximum duration (for $d > 0$) $f^{o}(d; \mathbf{g})$ [Eq.(5)] comes from the assumed structure of $f(d)$ [Eq.(5)]. However, this may not always be so, for instance if the impulse response of the linear kinematic diffusion [24] is taken as the discrete-continuous probability function. Then the same parameters are both in the continuous and discrete part of the distribution

Indeed, frequently cited Davison and Smith's publication (1990) is of highest recognition and, in our opinion, particularly Section 3.1 of Chapter 3 – Maximum Likelihood Regression with the Appendix A, where, perhaps for the first time in hydrology, maximum likelihood (ML) estimation of distribution parameters with covariates was presented. Noteworthy, it was only one year before our first publication [10], where ML estimation in the presence of time as a covariate was presented as extension of the conventional FFA. Other ideas presented by Davison and Smith (1990) are not new to us, too. The first author's PhD thesis defended in 1966 [6-9] revolved around similar issues, though not named 'Peak Over Threshold' (POT) model. The Poisson-Exponential model with time-dependent parameters was comprehensively described in 2001 [18, Subsections 7. 1-5]. It is worth comparing our approach [20] with Davison and Smith's (1990). It seems that time-variable Poisson parameter, i.e. the time-dependent arrival rate, should be estimated from the time-series of annual numbers of the threshold exceedances as it was done in [20]. We used ML method while modelling the time-dependent Poisson parameter by exponential function of time: $\lambda = \exp(a + bt)$.

However, for obvious reasons, i.e. short hydrological series, bigger error in large flow assessment, unknown form of trends, etc., one should be cautious about trends in FFA and also in the DqF if low threshold is assumed. In the reviewed paper the trend is considered only in the β parameter. It is estimated by ML method both (1) on the base of annual flood peaks distribution with time dependent mean and standard deviation and (2) by direct binary regression of weighting parameter β, advising to compare the results of both methods as an overall test of the assumptions made. We suggest also to verify the trend significance by the AIC. It is not included in the approach recommended by you.

We would like to draw your attention to the one of the two approaches we used for estimation of the binomially distributed parameter β for stationary (equation 7) and the logistic regression for non-stationary case (equation 23). The β parameter [real, $\beta \in (0, 1)$] although resembles the POT's Poisson parameter (which is real, too, but larger than 0), e.g.

Davison and Smith (1990), it is not the one! If the arrival of floods over the threshold Q_A (or equivalently floods with duration $d > 0$ over the threshold Q_A) is modelled by the Poisson process (with mean annual arrival rate λ) then probability of non-excedeence of Q_A (or equivalently of $d = 0$) over the one-year period equals exp($-\lambda$). It corresponds with no arrival probability [Eq.(5)] of the binomial distribution and for large Poisson rate λ the estimate β from exp($-\lambda$) will be more accurate.

Were we forced to look for similarities of our model of time-independent discrete-continuous distribution with zero discrete values [22, 27, 30] to other approaches used in hydrology, we would rather compare its form of the likelihood function with one of the censored sample of annual peak flows which was introduced in Polish hydrology by Kaczmarek [1] and then published in English book [2] in 1977. This becomes particularly obvious when a discrete-continuous probability function will be the impulse response of the linear kinematic diffusion [23].

Concluding the above, we agree with you entirely: in respect to non-stationarity modelling we used and developed the old concepts but they happen to be our own ideas, too. We do believe that there is nothing wrong in using our in-depth verified soft packages for the present study. We are aware that, due to 'the Iron Curtain' Polish works were not popular among Western scientists. On the other hand, we hope that a few of our numerous papers on non-stationary flood frequency analysis (NFFA) [10-21, 27] must be known to the Reviewer. We would like to add that due to the limited capacity of the paper it was not our intention to present a list of our previous papers in the references of the reviewed article, nevertheless, now we feel obliged to put most important works related to the subject. Like the Reviewer, we are astonished that our achievements were not recognised by the Western World despite the 'Global Village Effect' we observe now in science.

Having said that, we would point your attention to the fact that in our article the stress was laid on **stationary** DqF, which has been then extended to non-stationary cases. Note, that the non-stationary approach to binomially distributed time-series is original. We detached the non-stationarity of β parameter (i.e. the probability of not exceeding the annual peak flow of the alarm threshold, Q_A) from time-variability of the shape of the wave (negligible in our case). Then we proposed two ways of estimation of the β parameter (based on binomial distribution and by means of routinely Flood Frequency Analysis – FFA). Moreover, we brought in the Generalised Exponential (GE) distribution to model the flood duration $D > 0$ which so far, to our best knowledge, was not used in hydrology at least in such a role. The GE's most interesting feature is its flexible shape – it can be unimodal distribution or exponential-like depending on the data it is fitted to. Let us note that the GPD mentioned by the Reviewer and frequently applied in POT methods does not have such properties.

As far as the software is concerned, we are absolutely aware that the equations we present can be solved by myriads of methods and software available in the market, including the ones we presented. Since 1995 we have been successfully developing in Fortran our own software which was many times checked and double checked in terms of its robustness, errors and the rate of successful solutions in Monte Carlo simulation experiments. We have been extensively using this software for calculations to our almost all papers and they have never been questioned. The basic stationary version of this soft-package was used e.g. in [24 and 3], non-stationary version in [27 and 19], stationary and non-stationary seasonal approach in [27, 28 and 5] and censored distribution in [4, 25, 26]. Basing on concepts developed by Strupczewski and Mitosek (1991) [10] (later completed and published, e.g. in [11-21]), we developed the procedures of maximum likelihood estimation of the non-stationary mean and dispersion (i.e. with time as the covariate of the mean and standard deviation, believing that the two first moments of AM distributions are finite) of various 'hydrological' distribution functions commonly used in FFA and they were used in the reviewed paper. This is exactly the same approach as is used in GAMLSS soft-package in R which in fact was presented by us at International Symposium *'Deciphering River Flood Change'* (oral presentation: '*R-package GAMLSS for trend detection and estimation in parameters of extreme value distributions*' by E. Bogdanowicz in Vienna, September 2012). Nevertheless, the soft-packages and the numerical methods implemented are of further importance to us because we treat them as tools only and we rather concentrate on the hydrology. We agree, that we can use R as well and the methods described in Yee and Stephenson (2007) and Gilleland, *et al*. (2013), but in Poland none software is particularly recommended in hydrology and R would probably give similar results with much more cost, i.e. time spend on software training, data pre- and post-processing, results comparison, etc. Our software obeys also the consistency of maximum likelihood values with the respect of the number of parameters while covariates including – the more parameters the greater value of maximum likelihood function within the same type of the distribution. The rate of successful solutions got from Monte Carlo simulation experiments helped to improve the reliability of our algorithms. We are quite happy and confident with the present version of our soft-packages. To this effect we recommend it in page 10 on lines 26-27 and are eager to share it with other scientists.

Judging by the year of publication, Gilleland *et al*. (2013) probably preparing their paper simultaneously to ours. In fact we have been waiting for the critical comparison of existing techniques and algorithms, which unfortunately was out of scope of Gilleland's *et al*. (2013) paper. As additional criteria for comparison we would recommend the ML values of compared algorithms and user friendliness of the competing procedures**.**

Q: The focus of the paper is about time series data but the likelihood factorizations (equations (5) and (19)) are presented under the assumption of temporal independence. Obviously, this should be clarified.

A: Indeed, the issue mentioned by the Reviewer deserves a word of clarification. As the Reviewer probably noticed we concentrate on a single annual longest flood that occur above the alarm threshold. It does not have to be the same flood that gives the annual flood peak. The DqF approach is the extension of the conventional FFA performed on a single annual peak flow series. The annual peak flows are usually assumed to be temporary independent what has been verified by several investigators and so is assumed here for the annual maximal durations. Besides, due to the poor measurement material we are not supposed to question the independence of the particular events – in other words – the systematic hydrological records are too short to 'play' with autocorrelation of the annual maximum duration. The clarification of this important issue were added to the text, page 3 lines 15-20.

Q: I found the presentation outdated. It is done in a such way as in the 80's if little had happened in statistical modeling in hydrology and statistical software technologies. I would like to share that during the last 20 years the Generalized Linear Models (GLMs) methodology have been well developed and widely applied in hydrology due to availability of software procedures in computer packages such as R, S-plus, Matlab, SAS or Stata, e.g., Clarke (1994) and Aitkin et al. (2009). On the other hand the books of Coles (2001), Beirlant et al. (2004) and Reiss and Thomas (2007) give excellent overview for extreme value models with time-varying parameters as well as software procedures that handle the computations. Thus the authors should concentrate on the usage and interpretation of the software procedures that can handle the computations instead of presenting estimation equations such as (10), (20) or (21). The idea is to use standard software in order to unify the computations and make comparisons.

A: We answered partly to these remarks above. We realise that the list of literature is incomplete and omits important items. Thank you very much for the identification of these significant shortcomings. We have tried to improve the text and reference list in accordance with your remarks.

However, we would like to make some observations rather than comments on the topic of the completeness of the list of references, which is essential part of your review. It is a cliché to say that in a scientific paper no one can use and quote all the previous publications of the domain in question. Nowadays, neither Bernoulli's, nor Poisson's original works are cited; similarly, the articles published later then the text was submitted to the editor (e.g. Gilleland. and al., 2013!). At a time when seminal works in the field of estimation procedures, POT analysis, truncated or censored data and non-stationarity were formed, there existed two barriers blocking the exchange of scientific information and publications. The first one, so called the 'Iron Curtain', limited the free access of Polish scientists working on FFA to the research and achievements of the Western World and blocked the possibility of presentation of important works of the Polish School of Hydrology. The second one was the language barrier. Researchers deprived of the opportunity to publishing in English-language journals disseminated their works in Poland and in Polish. This does not mean, in any case, that Polish contribution to the development of hydrology science is negligible. It is obvious that papers in Polish, even the best, are of zero probability to be noted by the non-Polish-speaking scientific community. Still it is baffling that even after being published in English, they were often omitted by the English-language authors! Concluding, your remark about incompleteness of the references list in our article seems to be mutual and concerns also Western researchers. It seems to be clear that more gladly we quote papers of our own scientific and cultural community than faraway (in many meanings) research centres. Perhaps this could be called 'the long-term effects of the Iron Curtain'.

Now the Iron Curtain is a history, so we can assure the Reviewer that we are aware of new trends and 'fashions' in hydrological sciences but some of them we accept as subjects of our research and some do not. We actively take part in international conferences (sometimes as invited speakers) and our papers are still published despite being '*outdated*'. If the Reviewer is interested he/she can easily compare the list of our latest works with other authors.

As we mentioned above, the methods of solution of our hydrological problems do not lie within our scope of interest. The problem of statistical software you mentioned needs a word of general comment. We know the computer packages possibilities and we use them occasionally (comparability with our former results and the necessity of adjusting our data formats are not much encouraging). However, some experiences we have made with R-package GAMLSS for trend detection and estimation in parameters of extreme value distributions lead us to conclude that:

- There is a danger concerning the 'ready and easy-to-use' software stemming from the fact that investigators are equipped with a very powerful tool about which they have only limited understanding. How to cope with the gap between the understanding of theoretical developments and the need to apply these methods to FFA (or NFFA)? In our opinion detailed description of the methods in research and scientific papers can bridge this gap.
- It will be extremely important and useful to compare different software packages by performing calculation on the same set of synthetic data and choose the standards. So, we fully agree with your idea about application of standard software to unify computations and make comparisons but this is a scientific (but not a hydrological) subject itself.

Reviewer #2

Question: Pag. 2992, the meaning of the symbol dt should be specified;

Answer: We added the explanation (page 3, line 10): '$d_t = 0$ (index t marks the t-th year in a series in which the particular event $d = 0$ occurred, $t = 1, 2,\ldots, T$ and T is the length of the series in years)'.

Q: Pag. 2989 line 22, other "methodologically innovative" flood frequency analysis models that we believe should be mentioned are: Eagleson (1972); Sivapalan et al. (1990); Gioia et al., 2008; Iacobellis et al., 2011.

A: Indeed, we did not considered these remarkable works in our paper. The mistake has been corrected now (page 1, line 49 and, of course, in the reference list).

Q: Pag. 2993, lines 25-28, the sentence here reported should be explained one times in the paper; the same concept is also reported at page 2995 lines 11-13 and at page 2998 lines 1-3;

A: We rephrased the fragment of the text (page 3, line 51 to page 4, line 5): 'Therefore, the $d_t = 0$ in the [**d**] time-series means that the threshold discharge, Q_A, has not been exceeded during the t-th year of the series ($Q_{\max}(t) < Q_A$) or that the peak flow has exceeded the overtopping flow ($Q_{\max}(t) > Q_B$) where $Q_{\max}(t)$ denotes the annual maximum discharge occurred in the t-th year of the sample series. In other words, there is no risk of the dike's damaging due to the prolonged exposure to the high water because the flood wave was either too small to reach the weaken construction of the levee or, the contrary, the flood is such big and sudden that the water immediately overtops the levee's crest.'

Q: Pag. 2994, line 7, regarding the probabilistic model for flow duration curves, a huge literature is omitted we suggest "at least" mentioning Castellarin et al., 2004; Iacobellis 2008; Botter et al. 2008;

A: We have read and considered the papers recommended by the Reviewer (page 4, line 10).

Q: Pag. 2996 lines 1-7 the sentence should be better explained;

A: We rephrased the whole paragraph (page 5, lines 3-10): 'When the available data represent mean daily discharge, the d values are in fact the integer numbers (the exposition can last 1, 2, 3, etc. days) but to maintain the continuity of time we treat them as real numbers and consider d as if it corresponded to the duration range (d – 0.5 day, d + 0.5 day). In particular, for $d = 0$ (beginning of the time axis) the interval corresponds to the range (0, d + 0.5 day). If a flood starts before the end of a year and is continuing to the next year, the d value is derived for the entire flood wave (from its beginning in one year to its end in the next year) but attributed to the year t when the flood culmination occurred. To get an insight into flood persistence properties, the several threshold stages (Q_T) are considered but not only the alarm stage Q_A.'

Q: Pag 3001 line 23, the reason of the use of logistic function for the evaluation of β parameter by regression analysis should be explained;

A: We added the fragment to the text (page 7 line 40 to page 8 line 16) which, we hope, will better clarifies the reason why we used the logistic distribution:

The variable Y_t represents binary outcomes and has a binomial distribution with parameter:

$$\beta(t;\boldsymbol{\theta}) = P(Y_t = 1) = P(D = 0) \tag{22}$$

However, the trend in β can not be found by means of frequently assumed linear regression. The reasons of being that

- in general linear trend may take the values of probability $\beta(t, \boldsymbol{\theta})$ outside the range from 0 to 1,
- the error term is not homoscedastic, nor it is normally distributed as in normal regression.

In order to avoid values outside the range from 0 to 1 a monotonic transformation of the interval (0,1) is performed to the range (–∞, +∞). There are many transformations with this property, but the most popular are two: probit and logit transformations. Both give similar results but logit transform is more convenient for calculations. Probit transformation consists in converting the probability to corresponding quantiles of the standard normal distribution. Logit transformation is given by:

$$\text{logit} = \ln[\beta/(\beta - 1)] \tag{22a}$$

And the trend is modelled as:

$$\text{logit} = a + bt \tag{22b}$$

Inverse transformation leads to the logistic function β of time t with parameter vector $\boldsymbol{\theta} = [a, b]$.

$$\beta(t;a,b) = \frac{1}{1+e^{-(a+bt)}} \tag{23}$$

Logistic regression is used in many disciplines, medicine, social science, econometrics, in engineering, especially for predicting the probability of failure of a system or product.

Q: Pag. Equation 27, per analogy to equation (9) should be…

A: Indeed, we made a mistake, thank you very much for pointing it. The equation (27) was fixed according to the Reviewer's remark:

$$\hat{\beta}(t) = \hat{P}\left[D(t) = 0\right] = \hat{P}\left[Q_{\max}(t) \le Q_A\right] + \left\{\hat{P}\left[Q_{\max}(t) > Q_B\right]\right\} = \hat{G}\left(Q_A|t\right) + \left[1 - \hat{G}\left(Q_B|t\right)\right]$$

Q: The application to a case study at Vistula River in Southern Poland, is described like an exercise; in my opinion this section should reorganized, simplified and better explained.

A: In fact it was our intention to show on relatively easy example how the developed procedure works in practice, therefore the 'case study section' may resemble an exercise, indeed. But following the Reviewer's recommendations, we added some explanations and rephrased slightly the text of the section 5. In order not to mislead the Readers, we changed the title, too: 'Example – Szczucin at Vistula River (southern Poland)'. Apart from these, for better understanding we added the following fragment (page 9, lines 20-38).

'To illustrate how the proposed approach works in practice the Szczucin gauge (southern Poland) at the Vistula River has been selected as an example. Recent flooding in the upper Vistula bared the weakness of the system of flood protection, especially unsatisfactory condition of the embankments in the region of Szczucin. One, but not only, major reason for the current state of flood protection infrastructure is a complex history of these lands. When Western European countries formed an effective flood protection schemes Polish south-eastern lands were periphery of three empires, two of which were among the most undeveloped countries of the continent. After regaining independence, social and economic problems associated with merging the various districts of the reborn Poland influenced the poor development of an efficient protection system. For these reasons, embankments built by the World War II do not meet current requirements which were lately even put to higher level. The Polish People's Republic period did not bring any important changes. Although, the embankments have been periodically increased and strengthened, the high cost of post-war reconstruction and industrialization carried out under conditions of socialist economy, did not allow to catch up with Western standards. Lately, the material excavated on the flood land, very often at the immediate vicinity of the embankments, was used for the re-construction. As a consequence, the top layer of inactivated meadow was damaged, what facilitated the filtration of water from the horizontal residual layer under the layer of permeable sealer coat. There are present plans to modernise the dikes and first works have been carried out. The investor claims that the modernisation will reduce the flooding risk by 80%. To assess the risk before and after modernisation (provided that the statement of the investor is right) the following analysis was performed.'

We also completed the section with additional diagram of the durations (in days) above the alarm level Q_A = 1690 m^3/s recorded on Szczucin gauging station:

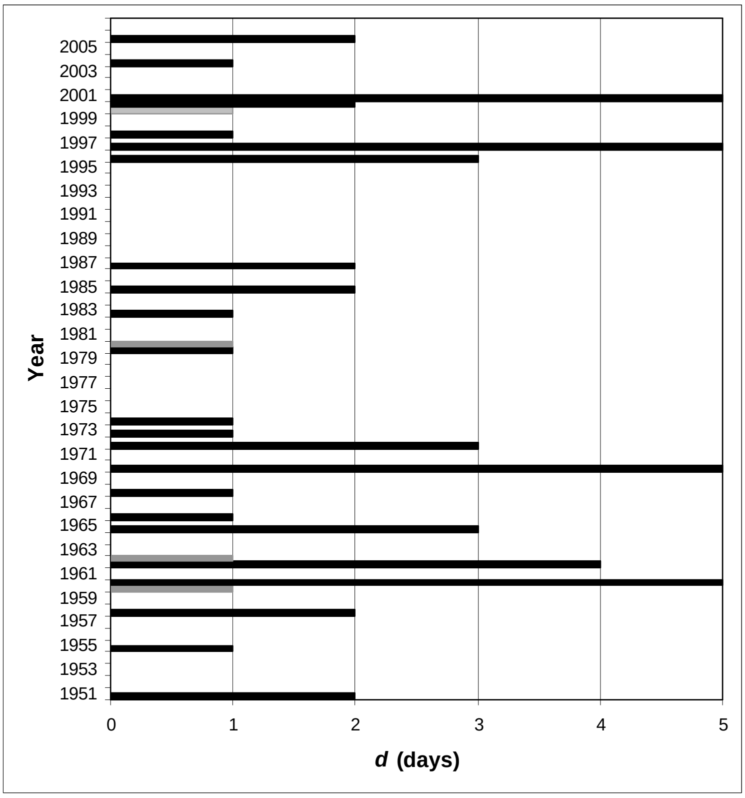

Fig. 5. The durations (in days) of the discharge above Q_A = 1690 m^3/s for Szczucin gauging station (1951–2006). The annual maximal durations are in black.

Q: Technical Comments…

A: All the mistakes spotted by the Reviewer were corrected.

The authors' references

[1] Kaczmarek Z. (1960). Confidence Interval as a Measure of Estimating Accuracy of Flood Discharges. (in Polish). *Wiadomości Służby Hydrologicznej i Meteorologicznej*, Vol. VII , 4.

[2] Kaczmarek Z. (1977). *Statistical Methods in Hydrology and Meteorology*. Sec. 4.5, 205-217.Published for the Geological Survey, U.S. Department of the Interior and the

National Science Foundation, Washington, D.C (translated of Polish book, 1970), 320 pp

[3] Kochanek K., Strupczewski W.G., Singh V.P. and Weglarczyk S. (2005) Are Parsimonious Flood Frequency Models More Reliable than the True Ones? II. Comparative assessment of the performance of simple models versus the parent distribution. *Acta Geophysica Polonica*, Vol. **53**, no 4, pp. 437-457.

[4] Kochanek K., Strupczewski W.G., Singh V.P. & Weglarczyk S. (2008) The PWM large quantile estimates of heavy tailed distributions from samples deprived of their largest element. *Hydrological Sciences–Journal–des Sciences Hydrologiques*, **53**(2), pp 367–386

[5] Kochanek, K., Strupczewski, W.G., Bogdanowicz, E. (2012) On seasonal approach to flood frequency modelling. Part II: Flood frequency analysis of Polish rivers. *Hydrological Processes*. Volume: **26**, Issue: 5, Pages: 717-730, DOI: 10.1002/hyp.8178.

[6] Strupczewski,W. (1966) Statistical Analysis Of Flood Hydrographs. D.Sc.Thesis, Warsaw Technical University, Water Eng. Dep., Warsaw, 190 Pp. (In Polish).

[7] Strupczewski, W. (1967a) Determination of the probability of repeating phenomena. *Acta Geoph. Polonica*, XV, **2**, 147-158 (in Polish).

[8] Strupczewski, W. (1967b) Determination of the annual probability distribution of some events on basis of all their occurrences. *Acta Geoph. Polonica*, XV, **3**, 247-262 (in Polish).

[9] Strupczewski, W. (1967c) Determination of the probability distribution of maximum discharges on basis of all observed floods. Publ. IAHS, 3:41-49.

[10] Strupczewski, W.G., Mitosek, H.T. (1991) How to deal with non-stationary time series in the hydrologic projects. Mitteilungsblatt des Hydrographishen Dienstes in Osterreich, Nr.65/66,36-40. Presented at IAHS Symposium, Vienna.

[11] Strupczewski, W.G., Feluch, W. (1996) Floods analysis of the Warta river by the of identification of an optimum flood frequency model with time dependent parameters (IDT). Abstract. J. Annales Geophysicae. Supplement of V.14, 51 .

[12] Strupczewski, W.G., Feluch, W. (1997a) System of identification of an optimum flood frequency model with time dependent parameters (IDT). In: Integrated Approach to Environmental Data Management Systems. Editor Harmancިouglu et al., Kluwer Acad. Publ. 291-300.

[13] Strupczewski, W.G., Feluch, W. (1997b) Floods study of the Polish rivers by the IDT soft tool. Annales Geophysicae, Part II, Hydrology, Oceans, Atmosphere & Nonlinear Geophysics, Supplement II to Volume 15, C-310.

[14] Strupczewski, W.G., Feluch W. (1997c) System of identification of an optimum flood frequency model with time dependent parameters (IDT). In: N.B. Harmancioglu et al. (ed) Integrated Approach to Environmental Data management Systems,, Kluwer Academic Publishers, 291-300. Presented at NATO-ARW "Integrated Approach to Environmental Data Management Systems" Izmir, September.

[15] Strupczewski, W.G., Feluch, W. (1998a) Flood frequency analysis under non-stationarity, Geographia Polonica, 71, 19-34.

[16] Strupczewski, W.G., Feluch, W. (1998b) Investigation of trend in annual peak flow series. Part.I. Maximum likelihood estimation. Proceedings of the Second International Conference on Climate and Water V.1, 241-250.

[17] Strupczewski, W.G., Kaczmarek, Z. (1998c) Investigation of trend in annual peak flow series. Part. II Weighted Least Squares estimation. *Proceedings of the Second International Conference on Climate and Water*, V.1, 251-263.

[18] Strupczewski, W.G., Mitosek, H.T. (1998d) Investigation of trend in annual peak flow series. Part..III Flood analysis of Polish rivers. *Proceedings of the Second International Conference on Climate and Water*, V.1, 264-272.

[19] Strupczewski, W.G., Singh, V.P., Feluch, W., (2001a). Non-stationary approach to at-site flood-frequency modelling. Part I. Maximum likelihood estimation. J. Hydrol., 248, 123-142.

[20] Strupczewski, W.G., Kaczmarek Z.,(2001b). Non-stationary approach to at-site flood-frequency modelling. Part II. Weighted least squares estimation. J. Hydrol. ,248, 143-151

[21] Strupczewski, W.G., Singh, V.P., Mitosek, H.T., (2001c). Non-stationary approach to at-site flood-frequency modelling. Part III. Flood analysis of Polish rivers. J. Hydrol., 248, 152-167

[22] Strupczewski W.G., Weglarczyk, S. , Singh, V.P., (2002) Modelling Discontinuous Distribution for Hydrological Samples with Zero Values. *5-th ICHE*, Warsaw, Poland, CD Proceedings, Track C, PDF 159, pp. 11.

[23] Strupczewski W.G., Weglarczyk, S. , Singh, V.P., (2003). Impulse response of the kinematic diffussion model as a probability distribution of hydrologic samples with zero values. *J.Hydrol.* 270, 328-351.

[24] Strupczewski W.G., Kochanek K., Singh V.P. and Weglarczyk S. (2005) Are Parsimonious Flood Frequency Models More Reliable than the True Ones? I. Accuracy of Quantiles and Moments Estimation (AQME) – Method of Assessment. *Acta Geophysica Polonica*, Vol. **53**, no 4, pp. 419-436

[25] Strupczewski W.G., Kochanek K., Weglarczyk S. and Singh V.P. (2005) On robustness of large quantile estimates of log-Gumbel and log-logistic distributions to largest element of the observation series: Monte Carlo results vs. first order approximation. *Stoch Environ Res Risk Assess*, Vol 19, No 4: 280–291, DOI 10.1007/s00477-005-0232-x.

[26] Strupczewski W.G., Kochanek K., Weglarczyk S., Singh V.P. (2007) On robustness of large quantile estimates to largest elements of the observation series. *Hydrol. Processes*. Vol. 21, issue 10, pp. 1328-1344

[27] Strupczewski W.G., Kochanek K., Feluch W., Bogdanowicz E. & Singh V.P. (2009) On seasonal approach to non-stationary flood frequency analysis. *Physics and Chemistry of the Earth* 34 pp 612–618

[28] Strupczewski, W.G., Kochanek, K., Bogdanowicz, E., Markiewicz, I. (2012) On seasonal approach to flood frequency modelling. Part I: Two-component distribution revisited. *Hydrological Processes*. Volume: **26** Issue: 5 Pages: 705-716, DOI: 10.1002/hyp.8179

[29] Strupczewski, W.G. and Napiorkowski, J. (1981) Distribution of maximum seasonal storm depths and total precipitation from sequences model. *Przegl. Geof.*,XXVI,1-2:25-36 (in Polish).

[30] Weglarczyk, S., Strupczewski, W.G. and Singh, V.P. (2005), Three-parameter discontinuous distributions for hydrological samples with zero values, *Hydrologic Processes*, **19**, 2899-2914 DOI: 10-1002/hyp.5787.

Other references

Aitkin,M., Francis,B., Hind,J. and Darnell,R. (2009) *Statistical Modelling in R*. Oxford Univ. Press, London.

Botter, G., S. Zanardo, A. Porporato, I. Rodriguez-Iturbe, and A. Rinaldo, 2008, Ecohydrological model of flow duration curves and annual minima, *Water Resour. Res.*, **44**, W08418, doi:10.1029/2008WR006814.

Castellarin, A., R. M. Vogel, and A. Brath, 2004, A stochastic index flow model of flow duration curves, *Water Resour. Res.*, 40, W03104, doi:10.1029/2003WR002524.

Clarke, R.T. (1994). *Statistical modelling in hydrology*. John Wiley & Sons. New York.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer-Verlag, London.

Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds. Journal of the *Roy. Statist. Soc.. Series B*, **52**, pp. 393-442.

Eagleson, P.S., 1972, Dynamics of flood frequency, *Water Resour. Res*., 8(4), 878-898

Gilleland, E. Ribatet, M.·and Stephenson, A.G. (2013). A software review for extreme value analysis. *Extremes* **16**, 103–119, DOI 10.1007/s10687-012-0155-0

Gioia, A., Iacobellis, V., Manfreda, S., and Fiorentino, M., 2008, Runoff thresholds in derived flood frequency distributions. *Hydrol. Earth Syst. Sci.*, **12**, 1295–1307, doi:10.5194/nhess-12-1295-2008.

Iacobellis, V., 2008, Probabilistic model for the estimation of T-year flow duration curves, *Water Resources Research*, ISSN: 0043-1397, **Vol. 44**, doi: 10.1029/2006WR005400

Iacobellis, V., Gioia, A., Manfreda, S., Fiorentino, M., 2011, Flood quantiles estimation based on theoretically derived distributions: regional analysis in Southern Italy. *Nat. Hazards Earth Syst. Sci.*, **11**, 673–695, doi:10.5194/nhess-11-673-2011.

Reiss, R.D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed. Birkhauser, New York.

Sivapalan, M., E. F. Wood, and K. J. Beven, 1990, On hydrologic similarity, 3, A dimensionless flood frequency model using a generalized geomorphologic unit hydrograph and partial area runoff generation, *Water Resour. Res.*, **26**(1), 43-58

Yee, T.W. and Stephenson, A.G. (2007). Vector generalized linear and additive extreme value models. *Extremes* **10**, 1-19.