

This discussion paper is/has been under review for the journal Natural Hazards and Earth System Sciences (NHESS). Please refer to the corresponding final paper in NHESS if available.

Application of a hybrid approach in nonstationary flood frequency analysis – a Polish perspective

K. Kochanek¹, W. G. Strupczewski¹, E. Bogdanowicz², W. Feluch³, and I. Markiewicz¹

Received: 30 August 2013 – Accepted: 30 September 2013 – Published: 31 October 2013

Correspondence to: K. Kochanek (kochanek@igf.edu.pl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Discussion Pa

Discussion Paper

Discussion Pa

NHESSD

1,6001-6024,2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I₫

►I

4

•

Back

Close

Full Screen / Esc

Printer-friendly Version



¹Institute of Geophysics, Polish Academy of Sciences, Księcia Janusza 64, 01-452 Warsaw, Poland

²Institute of Meteorology and Water Management, Podleśna 61, 01-673 Warsaw, Poland

³Warsaw University of Technology, Łukasiewicza 17, 09-400 Płock, Poland

The comparative results of Monte Carlo simulations confirmed the superiority of two-stage NFFA methodology over the classical maximum likelihood one. Further analysis of trends in GEV-parent-distributed generic time series by means of both NFFA methods revealed big differences between classical and two-stage estimators of trends got for the same data by the same model (GEV or Gumbel). Additionally, it turned out that the quantiles estimated by the methods of traditional stationary flood frequency analysis equal only to those non-stationary calculated for a strict middle of the time series. It proves that use of traditional stationary methods in conditions of variable regime is too much a simplification and leads to erroneous results. Therefore, when the phenomenon is non-stationary, so should be the methods used for its interpretation.

1, 6001–6024, 2013

Paper

Discussion Paper

Discussion Paper

Discussion Pape

NHESSD

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page **Abstract**

Introduction

Conclusions Reference

Tables Figures

14

▶I

Close

Back

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Introductio

Back

Close

Printer-friendly Version

Scientists are still debating whether the observed climate changes are temporary and caused by human activity, but do not question the changes themselves. The variability of natural phenomena is permanent feature of nature which was noted already by ancient philosopher Heraclitus of Ephesus (c. 535-c. 475 BCE), the author of the famous citation: Panta Rhei, which modern man long seemed to ignore. Quite recently most of the tools and techniques used in flood frequency analysis assumed stationarity of hydrological processes in rivers (e.g. Milly et al., 2008). Nowadays, it is understood and accepted, that due to the observed change in climatic parameters and rapid development of calculation techniques the incorporation of the "non-stationarity" factor in hydrological parametric and non-parametric modelling is necessary and yet has become technically possible. Thus, hydrologists face the challenge of developing new or improving existing methods of flood frequency analysis (FFA) by taking into account the non-stationarity of extreme hydrological events. One has to bear in mind, however, that the variety of hydrological parameters that change in time and thus can (and should) be analysed within the context of time-variability is vast. These are for instance maximum annual or seasonal maxima (Q_{max}^T) , number of floods per year (N_O^T) , volume of maximal food (V_{Omax}^T) , or quite interesting centre of mass of seasonal discharge (which for a period $T_1 - T_2$ can be defined as: $\int_{T_1}^{T_2} Q \cdot T \cdot dT / \int_{T_2}^{T_2} Q \cdot dT$); each of these parameters require tailored approach. The recent research on non-stationarity in hydrology can be generally (subjectively) divided into two threads: (i) non-parametric trend(s) analysis in hydrologic data, moments and linear moments (aka L-moments) by means of statistical tests, e.g. Zhang et al. (2001), Vinnikov and Robock (2002), Burn and Hag Elnur (2002) and (ii) assumption of the parent distribution (both in local or regional scale) and detecting trends in its parameters (e.g. Khalig et al., 2006; Renard et al., 2006; El Adlouni et al., 2007; Villarini et al., 2009). The latter approach consists in treatment of time parameter as the covariate and estimation of time-dependent flood quantiles by any method of parameters estimation. Frequently

Discussion Paper

Discussion Pape

14



NHESSD

1, 6001–6024, 2013

Application of

a hybrid approach in

NFFA - a Polish

perspective

K. Kochanek et al.

Title Page

Full Screen / Esc



Discussion Paper

Discussion Paper

Figures

Introductio

14

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

cited Davison and Smith's publication (1990) is of highest recognition and, in our opinion, particularly Sect. 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. The great variety 5 of flood frequency distribution functions with the presence of time covariate can be estimated e.g. by the free-of-charge Generalized Additive Models for Location, Scale and Shape (GAMLSS) software (Rigby and Stasinopoulos, 2005).

In this article we will attempt to join these two complementary threads (i and ii) and confront the classical theoretically sound method of non-stationary quantiles estimation based on maximum likelihood (ML) with covariates with a simple hybrid, two-stage (TS) method based on Weighted Least Square (WLS) and L-moments.

The tangible results of the research are two variants of software for calculations of non-stationary flood quantiles by means of seasonal and annual maxima datasets one based on ML classical approach and the second on TS method.

The Two-Stage (TS) method

Our earlier works on issues of flood quantile estimation in non-stationary conditions and deficiencies of classical approach to non-stationary flood frequency analysis (NFFA) based on maximum likelihood method (ML) (Strupczewski et al., 2001; Strupczewski and Kaczmarek, 2001) resulted in preliminary ideas of algorithms using up-and-coming method based on the concept of L-moments (LM). This method, paradoxically, requires stationary series of independently and identically distributed components (i.i.d.). Being a modification of the method of weighted probability moments (PWM) (Greenwood et al., 1979; Hosking et al., 1985; Rao and Hamed, 2000) the L-moments method (Hosking, 1990; Hosking and Wallis, 1997) is widely used in the FFA because of its simplicity and satisfactory results for short sequences of hydrological measurements. Hosking et al. (1985) showed that for small samples the PWM estimators give better estimators of quantiles of the high probability of non-exceedance (called "flood

1, 6001–6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

NHESSD

K. Kochanek et al.

Title Page

Abstract

Conclusions Reference

Tables

Discussion Paper

Close

Printer-friendly Version

Interactive Discussion

quantiles") then other popular methods of estimation (such as moments, maximum likelihood, etc.). Moreover, the flood quantile estimates obtained by means of the Lmoments are less sensitive to the (erratic) selection of the model (the probability distribution function) than the estimated by the maximum likelihood method (e.g. Strupczewski et al., 2002a, b). Besides, as far as our experience is concerned (Strupczewski et al., 2005; Kochanek et al., 2005) the numerical methods applied to determine the maximum of the likelihood function of multi-parameter distributions often encounter the local maxima resulting in termination of the optimisation algorithm, thus the estimates of the parameters (and quantile) are far from being optimal. The difficulty to find the global maximum of the likelihood function increases with the number of model parameters, while the method of the L-moments is "indifferent" to the type and number of parameters of the probability distribution functions. In addition to the advantages already mentioned, the L-moments (Hosking, 1990; Hosking and Wallis, 1997):

- can be used for more kinds of models than the conventional moments, including models whose finite conventional moments may not exist (these models are called limited-existence-moments distributions); if the mean exists, the higher Lmoments also exist.
- are less biased than the conventional moments.
- are resistant to the outliers.

15

20

- are easy to calculate for the distributions having an explicit analytical form of quantile as a function of the cumulative distribution function, i.e.: x = x (F).

However, the LM method requires data series to be sorted from the smallest to the largest, which results in the devastation of the chronology of the subsequent flood episodes occurrence. For stationary cases the order of the elements in the sample is not important, but when we consider the non-stationary series, the sequence of the

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page

Abstract

Introductio

Conclusions

Reference **Figures**

▶

Tables

14

Back

Full Screen / Esc

Interactive Discussion

© 0 BY

measurements is crucial. Therefore, the process of the estimation of non-stationary flood quantiles was divided into two steps:

- in the first stage of the trends in the mean and standard deviation of the annual or seasonal maximum flows are estimated using the weighted least squares method (WLS) which is the equivalent to the estimation of trends in the moments of the normal distribution derived for distributions of constant and moderated skewness (Strupczewski and Kaczmarek, 2001). The calculated trend values are then used to standardise (deprive of trends) the time series;
- the resulting stationary sample is then used to estimate parameters and quantiles (stationary!) of a selected distribution function whose skewness is independent of the first two moments. Afterwards, the so calculated quantiles are re-trended.

The idea to introduce trends in the moments, rather than indirectly in parameters, nears the two-stage method to the classic analysis of time series and the techniques of trends detection. What is more important, it eliminates the estimation errors in moments, particularly large when the method of maximum likelihood is used, occurring when the selected distribution function (model) is incorrect (i.e. does not fully represent the population it describes) or when the model is different from the Normal distribution function for which the estimation errors of moments are 0. Similarly, the estimation error for the mean value is 0 when the Gamma and Inverse Gaussian distributions are used. Note that the TS method assumes that the statistical model (distribution type) does not vary in time, in other words once assumed distribution function (e.g. GEV or Gumbel) does not change over time and only its parameters (moments) are time-dependent. This is so, because the introduction of the time-variability in the model type would drastically complicate an already complex algorithms for estimating time-variant flood quantiles. We decided also to adopt the simplest, i.e. linear, trends in the mean and standard deviation (bearing in mind that other forms of time functions, like

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions References

Tables Figures

I4

•

M

Close

Back

Full Screen / Esc

Laterna et la Bianca et la

$$\mu_t = a \cdot t + b$$
 and (1a)

$$\sigma_t = c \cdot t + d \tag{1b}$$

5 where:

t – time (in the years following the beginning of the time series)

 μ_t – the average in year t

a – parameter of the trend in the mean

b – the mean parameter of

 σ_t – standard deviation in year t

c – parameter trend in the standard deviation

d – the standard deviation parameter.

As one can see, instead of two parameters in stationary case (μ, σ) now there are four parameters to be estimated in non-stationary case (a, b, c, d).

The parameters a, b, c, d are used for the standardisation of the time series x_t of annual or seasonal maximum flows (we prepared two variants of the calculation software packages: for annual and seasonal maxima). As a result a sample, y_t, free of trends in the mean and standard deviation is obtained:

$$y_t = (x_t - a \cdot t - b)/(c \cdot t + d) \tag{2}$$

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

In the second step, firstly the model parameters are calculated with the use of the method of L-moments and then quantiles (Q_{ν}^{M}) for the given probability of nonexceedance (F) for the selected probability distribution function (M). Of course, in the second stage, the method of L-moments can be replaced by any parametric or nonparametric method of estimation, but due to the advantages of L-moments this method

Discussion Pape

Back

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

1, 6001–6024, 2013

Abstract Introductio Conclusions

Reference

NHESSD

Application of

a hybrid approach in

NFFA - a Polish

perspective

K. Kochanek et al.

Title Page

Tables Figures

14 **▶**

Close

Discussion Pape

Printer-friendly Version

Interactive Discussion

is recommended. In both versions of the software package (for the seasonal and annual maxima) 7 functions were implemented: two-parameter distributions: Normal (N2) and Gumbel (Gu2), and three-parameter ones: Log-Normal (LN3), Pearson type III (Pe3), Generalised Extreme Value (GEV), Generalized Log-Logistic (GLL) and Weibull (WE3). The algebraic forms of these functions and formulas of parameters estimation by the method of L-moments can be found in Rao and Hamed (2000). Since the elements of y, can be, and usually are, both positive and negative, the distributions have to be characterised by unlimited or location-parameter limited left-hand side tails. Obviously, the list of models available in the packages does not exhaust the variety of possibilities but can be easily completed by new distribution functions. The selection of the best model for the particular series can be made by comparing AIC values (Akaike, 1974; Hurvich and Chih-Ling, 1989). In case of seasonal approach to estimation of the annual peak flows, the seasonal distributions may differ from each other (as long as they are implemented in the soft-package), and the quantile is calculated numerically for the model being the alternative of seasonal distribution functions. More on the seasonal approach to modelling of annual peak flows one can find in Strupczewski et al. (2012) and Kochanek et al. (2012). Having the stationary quantiles $Q_{\nu}^{M}(F)$, the values of parameters and trends in the mean and standard deviation estimated in the first stage of the procedure (a, b, c, d) the time-dependent quantiles $Q_{r}^{M}(F, t)$ can be calculated for selected moments of time (t):

$$Q_x^M(F,t) = (a \cdot t + b) + (c \cdot t + d) \cdot Q_v^M(F). \tag{3}$$

As a result of the two-stage algorithm, one obtains the quantile values for the given probability of non-exceedance and a given year. Similar results (only for the series of annual maxima) are obtained when using the classical approach based on the maximum likelihood functions with covariates (time). Here, to allow the comparison with the TS, the location and scale parameters are expressed in terms of timedependent mean and standard deviation. The shape parameter is assumed to be time independent. Please note, that the number of estimated parameters in the ML

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions Reference

Tables

Figures

14

▶

Close

Full Screen / Esc

approach increases by two in relation to the stationary case. This method is well known and described in the hydrological literature (e.g. Strupczewski et al., 2001; Katz et al., 2002), and its detailed description is omitted in this article.

3 The comparison of the ML and TS approach – a numerical experiment

In order to compare the results of both methods of estimating time-dependent quantiles the numerical simulations based on Monte Carlo (MC) technique were carried out. A series of non-stationary pseudo-random samples of the assumed trends in the mean and standard deviation were generated and then used to estimate trends in mean and standard deviation as well as calculate quantiles of the given probability of non-exceedance (F) and year (t). The values of mean and the standard deviation and their trends and the skewness coefficient in the generated time series reflect the maximum flow regime of Polish rivers. Knowing the true values of trends and quantiles, one can calculate the errors generated by both methods of estimation.

The selection of models for the observed series (in case of the TS approach – a fixed sample) were based on the AIC values which means that in the subsequent MC simulations in general the quantiles, but in the ML approach also trends in mean and standard deviation, can be determined for different models. For both variants of the experiment (TS and ML) three cases of model's selection were considered: (i) if the true probability distribution function of the series is unknown, and therefore every distribution (from a set of available in the software) has an equal chance to be indicated as appropriate for the sample, (ii) if we use the wrong distribution (the case reflecting the most "natural" situation), i.e. the true (generator) model is eliminated from the competition with the alternative models, and (iii) if we use only true model consistent with the generator. Calculations were performed for different variants of the generator, the mean and standard deviation, trends, skewness and sample size, probability of non-exceedance (*F*) and time horizons (*t*) of the quantile. However, for brevity only the

NHESSD

1, 6001-6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions References

Tables Figures

4 ▶

Close

Full Screen / Esc

Back

Printer-friendly Version

Interactive Discussion



most informative results will be presented in this article; the conclusions, if possible, will be generalised to the other cases.

The results of the estimation of trends for the generator of the GEV distribution with parameters mean, $\mu_t = 1000 + 2 \cdot t$, standard deviation $\sigma_t = 500 + 2 \cdot t$ and CS = 1.5 are shown in the Fig. 1. It is worth noting that the generated series rarely were "recognized" by the software as the GEV population, especially if they were short. This is because the relatively short time series do not reflect all the characteristics of the population. Additionally, in case of the TS method the structure of series was further disturbed by standardisation, so distribution function was adjusted to the samples without trends. Therefore, the results of the ML estimation for all available distributions, option (i), barely differ from those got for the set of distributions without the GEV, option (ii), the small percentage of cases when the GEV was recognised pose no significant effect on the mean scores and, therefore, the presentation of graphs for the variant (ii) is pointless. Of course, as far as the TS is concerned, the trend estimated by means of the WLS is the same for all three variants, since it does not depend on the model which is selected in the second stage of the algorithm.

The graphs show that both methods lead to a relatively good estimation of the mean and its trend with a slight predominance of the ML method. This method in variant (i) gives good approximation of the trend in small samples (N < 80), whereas, in variant (iii) the results are comparable to the WLS. As far as the trend of the standard deviation coefficient (c) is concerned, better results in terms of stability are achieved by the WLS. If, however, the general estimation of time-dependent mean, μ_t , and standard deviation, σ_t , are taken into account, the WLS is unbeatable. The shape of the ML curves reveal the unstable character of the solution which (even for very large samples N > 1000 unreal in hydrology) do not reach the true value of the population, while the parameters by the WLS tend periodically to the correct solution. In addition, the ML method is relatively erratic – a few percent of the estimation attempts fails for unknown reasons and/or the results are unreliable because we cannot be sure whether the algorithm reached global maximum of the likelihood function. The WLS method always gives

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I∢

14

►I ►

Back

Close

Full Screen / Esc

Printer-friendly Version



Considerably stronger differences in the results obtained by two competing methods, TS and ML, can be observed for flood quantiles. As an example, the discussion of the selected quantile probability of non-exceedance F = 0.9 which corresponds to the maximum flow of a 10 yr return period. The criterion of estimation errors were the *relative bias* (RB):

$$RB = \left(Q_{F \text{ est.}}^t - Q_{F \text{ teor.}}^t\right) / Q_{F \text{ teor.}}^t \cdot 100\%, \tag{4}$$

where:

 $Q_{F \text{ est.}}^{t}$ – the value of estimated quantile in *t*th moment in time, $Q_{F \text{ teor.}}^{t}$ – the value of theoretical quantile in *t*th moment in time, and *relative root mean square error*, (RRMSE):

$$RRMSE = \left\{ E \left[\left(Q_{F \text{ est.}}^t - Q_{F \text{ teor.}}^t \right)^2 \right] \right\}^{0.5} / Q_{F \text{ teor.}}^t \cdot 100 \%.$$
 (5)

The results for the variant (i) and (iii) for selected moments of time is shown in Fig. 2. Option (ii) is omitted because of the substantial similarity to (i).

It is easy to notice, that for the option (i) the bias of the quantile obtained by the TS method is lower than the one for the ML method regardless the time t. This is primarily because the ML method is more sensitive to a model error, and models identified as the best for the generated series are usually different from the population of the pseudorandom generator. A clear advantage of the method TS over the ML disappears only when quantile estimation is done using the same model as the generator model — option (iii). It should be noted, however, that such a situation does not happen in practice, because even though the real "true" model of annual maximum flows was known, it would be too complex so its parameters could be estimated from the short series of measurements usually available in hydrology.

sion Paper

scussion Pap

Discussion Paper

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

[**4** ▶]

→

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Paper

The values of the quantile's relative root mean square error (RRMSE) for the two methods are comparable but the ML method seems to be slightly superior. However, a little better results of the RRMSE for the ML do not compensate much higher values of bias.

To sum up the numerical experiment, we can conclude that although the trend estimation results using both methods are similar, the TS method proved better when calculating the time-variant flood quantiles. When we do not know the model and the parameters of the populations of the time series, it is safer to use the TS approach. In addition to the better estimation of the quantile, it is also more stable in terms of numerical methods and easier to implement in a practical calculation soft-package.

4 Estimation of trends and flood quantiles for Polish rivers

Software based on the ML and TS methods (for the latter variants for annual and seasonal maxima) was used to estimate trends and flood quantiles for 55-element time series of maximum annual (seasonal) flow measurement (years 1951–2005) for 31 gauging stations located on the largest Polish river: Vistula and its tributaries that represent the whole variety of the catchment sizes and regimes available in Poland.

The sequences of the historical annual and seasonal maxima do not reveal statistically significant "shifts" in annual/seasonal peak flows due to a sudden change in the hydrological regime of the rivers, for example, because of the construction of reservoirs upstream, water transfers and intake, etc. Such datasets can be then treated as the subject of further trend analysis. Thus, using the three methods implemented in soft-package: (1) the two stage method for a series of annual maxima, (2) the two stage method for seasonal peak flows and (3) method of maximum likelihood for annual maxima the trends in the mean and standard deviation were calculated, as well as, the upper quantiles for the selected time moments, i.e. the first year of the time series (t = 1), the middle of the series (t = 25), the end of the observation period (t = 55) and

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions References

Tables Figures

I₫

►I

Back

Close

Full Screen / Esc

Printer-friendly Version



Paper

10 yr after the last element of the series (t = 65). For brevity two distributions, namely Gumbel and GEV, were employed in the calculations.

The vast amount of measuring material and the results of the estimation made it impossible to present all the findings in the limited context of this article. Without loss 5 of the generality the detailed analysis of the results is limited to the 10 yr quantile calculated by means of annual peak flows. In order to compare the TS (WLS in terms of trends in mean value and standard deviation) and ML methods the results are presented as the ratio of a value obtained by the TS (WLS part) and ML approach (Tables 1 and 2); the absolute numbers are discussed only for one station – Warsaw-Nadwilanówka on the Vistula River.

Since the estimation results of the trends moments by means of WLS do not depend on the model (because it is always normal distribution function), the numbers in the table above support the belief that the values of trends and moments for the ML are strongly distribution dependent. 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. Interesting is, that the differences between the results are not limited only to the absolute values of the estimators but also to their signs! This means that, for example, for the Zywiec gauge (number 12) the GEV gives a negative trend in mean (a = -0.04) and the Gmbel's trend is positive (a = 0.17). Such a qualitative variability of the estimators undermines the credibility of the results of non-stationary flood frequency design and increases the margin of error resulting from the uncertainty of the estimators. In addition, there is no quarantee that the calculated trend will continue in the future, especially if its sign is negative. According to the expectations, the Gumbel model gives ML-based trend estimators nearer to the WLS method than the GEV distribution, since the skewness of the Gumbel distribution is constant (CS = 1.14) and just slightly higher than the Normal distribution (CS = 0).

Table 2 shows that, although the results of the trend estimator may vary, quantile values obtained by both methods are generally very similar (differences rarely exceed NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions Reference

Tables Figures

14

▶[

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract Conclusions

Reference

Introductio

Tables

Figures

14

▶

Back

Close

Full Screen / Esc

NHESSD

1, 6001–6024, 2013

Application of

a hybrid approach in

NFFA - a Polish

perspective

K. Kochanek et al.

Title Page

Printer-friendly Version

Interactive Discussion



a few percent), the procedure ML/GEV gives usually larger quantiles than TS/GEV one, while for ML/Gumbel they are generally smaller. Obviously, the differences grow with the probability of exceedance of a quantile and "distance" from the centre of the time series ($t \approx 25$) revealing the greatest value for the quantile extrapolated beyond the time of the time series. Analogically, it was found that time-dependent quantiles equal to stationary quantiles only near the centre of the time series, and the difference increases when moving away from the centre, and is the greatest when we make a prediction (extrapolate results) for a few years after the end of the series. This result proves that the use of traditional stationary flood quantile estimation methods for the cases where the variation of hydrological regime of rivers is evident is a farreaching simplification and leads to erroneous results and decisions. So, when the process is known to be non-stationary, also non-stationary methods should be used for its analysis. On the other hand, the capability of modelling the non-stationary complex hydrological phenomena is still very poor. One has to admit that it is difficult to identify a combination of method/model that gives results close to reality. However, basing on the experience drawn from the numerical experiment you can incline toward the results obtained by the TS, while the choice of the model depends on the characteristics of the time series and the preferences of the analyst.

Results for gauge Nadwilanówka Warsaw on the Vistula by different methods are presented in Table 3. According to the maximum likelihood criterion the best-fitting distribution to the sequences of annual maxima and summer maxima on the Vistula River in Warsaw is three three-parameter Pearson Type III distribution, and for the winter maxima the Weibull distribution (Kochanek et al., 2012). For these models, the calculations were carried out.

Strikingly large differences in the estimated values of the moments (and trends) between the ML and TS are transformed into differences in flood quantiles. As mentioned above, differences in the trends consider not only the absolute value but also the sign of the mean trend (in case of summer) and standard deviation (in case of year). It means that by careful adjustment of the estimation methodology, one can get

14

Back

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the "desired" results. The quantiles by the ML are much smaller than those obtained by the TS, the difference increases with the return period of a quantile. It is interesting, that in both methods nearly always the quantile values decrease with time, even if the trend of the mean is positive (like in TS/summer) but too small (a = 0.56) to compensate large negative trend of the standard deviation (c = -10.8). The exception is TS/winter $Q_{F=0.9}^t$ and $Q_{F=0.99}^t$, where the high value of the negative trend in the mean and equal to this (in its absolute value) trend of standard deviation $(a/c \approx -1)$ result in a slight increase in value of flood quantile over the years. Similarly, in the case of the TS yr⁻¹, the ratio $a/c \approx -1$ results in little variation of the quantile values, in particular the $Q_{F=0.9}^t$ slightly decrease, whereas $Q_{F=0.99}^{t}$ slightly increase over the years. It is also interesting that the annual quantiles received by the method of TS for alternative events (the last row of Table 3) is very similar to the quantiles for the summer season. This reflects the dominant role of the summer maxima over winter ones; this issue is discussed in detail in Strupczewski et al. (2012). As in the case of GEV and Gumbel distributions it is impossible to indicate clearly the proper results and thus the estimation method that best predicts the quantiles. However, accepting the uncertainty of the estimated quantiles, it is possible to draw the conclusions about the direction of changes in the river regime in next years and prepare a water management policy assuming the possible reduction in the annual and seasonal maximum flows.

Summary and conclusions

The purpose of this study was to present a two-stage method (TS) of flood quantile estimation from non-stationary time measurement series and to confront its accuracy in the results of trends in the mean, standard deviation and time-dependent quantiles with the classical method based on the maximum likelihood (ML) function with covariates (time). In order to compare these two methods a numerical Monte Carlo experiment was carried out. The results of the experiment showed that the TS method is characterized by a greater numerical stability, which gives reliable results for almost

Abstract Conclusions

Reference

NHESSD

1, 6001–6024, 2013

Application of

a hybrid approach in

NFFA - a Polish

perspective

K. Kochanek et al.

Title Page

Tables

Figures

Introduction

▶|

Close

every non-stationary sample, while the ML approach sometimes fails or gives unreliable results difficult to verify in practice. The TS method, and more specifically its first stage, the weighted least square approach (WLS), also provides more accurate estimates of the time-dependent mean and standard deviation, even though the precision of the estimates of the trends themselves is similar in both methods. What is more important, the estimated values of moments in the TS method do not depend on the model (distribution function) choice, which is used to the estimation of time-dependent quantiles.

If the probability distribution of the population from which the measuring sequence generated is unknown (i.e. always), the TS method gives more accurate time-dependent flood quantiles than the ML, regardless of the size of the random sample (N), moment (t) and the probability of non-exceedance (F).

Both approaches (TS and ML) was used to estimate trends in the first two moments and to calculated time-dependent flood quantiles for 31 measuring series of the maximum annual and seasonal flows. The analysis of the results indicates significant differences in the assessment of trends in the mean and standard deviation. The differences between the results for the quantiles were smaller, but grow with a probability of exceedance of the quantile and the distance from the middle of the time series. Additionally, it was found that time-dependent quantiles equal to stationary quantiles when they are calculated for the time near to the centre of the sample. The difference increases when moving away from the centre and it is the largest for the quantiles predicted (extrapolated) for the period after the last observation in the series. This result is the evidence that the use of stationary FFA when we are aware of the variation of hydrological regime of rivers is far too much a simplification and leads to the erroneous results and decisions. So, when we know that the process is nonstationary, non-stationary methods should be also used for the analysis. On the other hand, the possibility of non-stationary modelling of complex hydrological phenomena is still limited.

NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introductio

Conclusions Reference

Tables Figures

I₫

►I

4

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



NHESSD

1, 6001–6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introduction

Conclusions Reference

Tables

Figures

14

▶I

Close

Back

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



According to the criterion of maximum likelihood the distribution best-fitting to the series of annual and summer maxima on the Vistula River in Warsaw is threeparameter Pearson Type III distribution and for the winter maxima Weibull distribution. Large differences in the values of the trends in moments by the ML and TS are the 5 cause of significant differences in flood quantiles.

To conclude, it is difficult to indicate the combination of a method/model that gives the results closest to reality. However, referring to the numerical experiment the TS method can be recommended, leaving the choice of the probability distribution of the analyst's experience and preferences.

Despite the fact, that the statistical techniques used in both approaches are relatively complex, still the non-stationary models are only a simplified description of the actual volatility of the quantile of maximum flow in rivers. Although we observe a significant progress in non-stationary flood frequency analysis, this area is still in its infancy and requires huge efforts of the researchers and practitioners in order to meet the requirements of flood risk assessment in a non-stationary water regime.

Acknowledgements. This research project was partly financed by the Grant of the Polish Ministry of Science and Higher Education luventus Plus IP 2010 024570 "Analysis of the efficiency of estimation methods in flood frequency modelling", Grant of the National Science Centre (contract no. 2011/01/B/ST10/06866): "Stochastic flood forecasting system (The River Vistula reach from Zawichost to Warsaw)" and made as the Polish contribution to COST Action ES0901 "European Procedures for Flood Frequency Estimation (FloodFreq)". The flow data were provided by the Institute of Meteorology and Water Management (IMGW), Poland.

References

Akaike, H.: A new look at the statistical model identification, IEEE T. Automat. Contr., 26, 358-375, 1974.

Burn, D. H. and Hag Elnur, M. A.: Detection of hydrologic trends and variability, J. Hydrol., 225, 107-122, 2002,

- NHESSD
 - 1, 6001–6024, 2013
- **Application of** a hybrid approach in NFFA - a Polish perspective
 - K. Kochanek et al.
- Title Page Abstract Introduction Conclusions References **Tables Figures** 14 **▶** Close Back
- Full Screen / Esc
 - Printer-friendly Version
 - Interactive Discussion

- Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds, J. Roy. Stat. Soc. B, 52, 393–442, 1990.
- El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R., and Bobée B.: Generalized maximum likelihood estimators of the non-stationary GEV model parameters, Water Resour. Res., 43, W03410, doi:10.1029/2005WR004545, 2007.
- Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, Water Resour. Res., 15, 1049–1054, 1979.
- Hosking, J. R. M.: L-Moments: analysis and estimation of distributions using linear combinations of order statistics, J. Roy. Stat. B Met., 52, 105-124, 1990.
- Hosking, J. R. M. and Wallis, J. R.: Regional Frequency Analysis, an Approach Based on L-Moments, Cambridge University Press, Cambridge, UK, 224 pp., 1997.
- Hosking, J. R. M., Wallis, J. R., and Wood, E. F.: Estimation of the generalized extreme-value distribution by the method of probability weighted moments, Technometrics, 27, 251-261, 1985.
- Hurvich, C. M. and Chih-Ling, T.: Regression and time series model selection in small samples, Biometrika, 76, 297-307, 1989.
- Katz, R., Parlange, M, B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25, 1287–1304, doi:10.1016/S0309-1708(02)00056-8, 2002.
- 20 Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P., and Bobée, B.: Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review, J. Hydrol., 329, 534-552, 2006.
 - Kochanek, K., Strupczewski, W. G., Weglarczyk, S., and Singh, V. P.: Are the parsimonious FF models more reliable than the true ones? Il Comparative assessment of performance of simple models versus the parent distribution, Acta Geophysica Polonica, 53, 437–457, 2005.
 - Kochanek, K., Strupczewski, W. G., and Bogdanowicz, E.: On seasonal approach to flood frequency modelling. Part II: Flood frequency analysis of Polish rivers, Hydrol. Process., 26, 717–730, 2012.
 - Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: whither water management?, Science, 319, 573-574, doi:10.1126/science,1151915, 2008.
 - Rao, A. R. and Hamed, K. H.: Flood Frequency Analysis, CRS Press LLC, Boca Raton, Florida, USA, 2000.

Paper

- NHESSD 1, 6001–6024, 2013
 - **Application of** a hybrid approach in NFFA - a Polish perspective
 - K. Kochanek et al.
 - Title Page Abstract Introduction Conclusions References **Tables Figures** 14 **▶**
 - Close
 - Full Screen / Esc

Back

- Printer-friendly Version
- Interactive Discussion

- Renard, B., Lang, M., and Bois, P.: Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data, Stoch. Env. Res. Risk. A., 21, 97-112, 2006.
- Rigby, R. A. and Stasinopoulos, D. M.: Generalized additive models for location, scale and shape, Appl. Stat., 54, 507-554, 2005.
- Strupczewski, W. G. and Kaczmarek, Z.: Non-stationary approach to at-site flood-frequency modelling. Part I I. Weighted least squares estimation, J. Hydrol., 248, 143-151, 2001.
- Strupczewski, W. G., Singh, V. P., and Feluch, W.: Non-stationary approach to at-site floodfrequency modelling. Part I. Maximum likelihood estimation, J. Hydrol., 248, 123-142, 2001.
- Strupczewski, W. G., Singh, V. P., and Weglarczyk, S.: Asymptotic bias of estimation methods caused by the assumption of false probability distribution, J. Hydrol., 258, 122-148, 2002a.
- Strupczewski, W. G., Wêglarczyk, S., and Singh, V. P.: Model error in flood frequency estimation, Acta Geophysica Polonica, 50, 279-319, 2002b.
- Strupczewski, W. G., Kochanek, K., Singh, V. P., and Weglarczyk, S.: 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, 53, 419-436, 2005.
- Strupczewski, W. G., Kochanek, K., Bogdanowicz, E., and Markiewicz, I.: On seasonal approach to flood frequency modelling. Part I: Two-component distribution revisited, Hydrol. Process., 26, 705-716, 2012.
- Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D., and Krajewski, W. F.: Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, Adv. Water Resour., 32, 1255-1266, 2009.
- Vinnikov, K. Y. and Robock, A.: Trends in moments of climatic indices, Geophys. Res. Lett., 29, 14-1-14-4, doi:10.1029/2001GL014025, 2002.
- Zhang, X., Harvey, K. D., Hogg, W. D., and Yuzuk, T. R.: Trends in Canadian streamflow, Water Resour. Res., 37, 987-998, 2001.

NHESSD

1, 6001-6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page									
Abstract	Introduction								
Conclusions	References								
Tables	Figures								
I₫	►I								
[4	ÞI Þ								
I ∢ ■ Back	►I ► Close								

Printer-friendly Version

Table 1. The ratios of the trend values in mean and standard deviation got by the WLS to the one by the ML approach.

			GE	ΞV		Gumbel						
		$a_{\rm WLS}/a_{\rm ML}$	$b_{\mathrm{WLS}}/b_{\mathrm{ML}}$	$c_{ m WLS}/c_{ m ML}$	$d_{\rm WLS}/d_{\rm ML}$	$a_{\rm WLS}/a_{\rm ML}$	$b_{\rm WLS}/b_{\rm ML}$	$c_{ m WLS}/c_{ m ML}$	$d_{\rm WLS}/d_{\rm ML}$			
1	Jawiszowice	1.17	0.97	0.27	0.29	1.01	1.03	1.75	1.14			
2	Tyniec	0.95	0.97	0.69	0.70	1.06	1.04	1.35	1.33			
3	Jagodniki	0.84	0.98	-0.42	0.70	1.02	1.03	-0.93	1.05			
4	Szczucin	1.05	0.97	0.14	0.58	1.20	1.03	0.23	1.06			
5	Sandomierz	0.98	0.98	0.74	0.74	1.02	1.01	1.05	1.04			
6	Zawichost	1.15	0.98	1.07	0.74	1.14	1.01	1.27	0.98			
7	Puławy	0.92	0.99	0.64	0.84	0.97	1.01	1.00	1.05			
8	Warsaw	0.93	0.99	-0.01	0.94	0.82	1.00	-0.02	1.11			
9	Kępa	0.99	1.00	1.34	0.99	0.99	1.00	1.12	0.97			
10	Toruń	1.01	1.00	4.31	0.99	1.01	1.00	3.92	1.00			
11	Tczew	1.03	1.00	1.53	1.02	1.01	1.01	1.43	1.07			
12	Żywiec	-2.62	0.94	0.79	0.60	0.60	1.04	1.78	1.26			
13	Sucha	0.87	1.01	0.79	0.69	0.96	1.07	2.30	1.20			
14	Wadowice	0.73	0.99	-0.07	0.67	0.89	1.05	-0.05	1.17			
15	Rudze	0.68	1.11	-0.23	1.03	0.90	1.05	-0.97	1.16			
16	Stróża	1.01	0.98	0.69	0.69	1.34	1.04	1.06	1.11			
17	Proszówki	0.17	0.79	0.26	0.59	0.16	0.80	0.29	0.72			
18	Nowy Sącz	2.32	0.92	0.27	0.17	0.73	1.04	1.83	1.24			
19	Żabno	1.36	0.95	1.03	0.53	1.07	1.03	1.49	1.19			
20	Nowy Targ	1.06	0.93	0.50	0.39	1.07	1.03	1.20	1.11			
21	Zakopane	0.53	1.05	-1.26	0.92	0.80	1.07	4.23	1.42			
22	Muszyna	0.73	0.96	2.24	1.02	0.60	0.97	2.09	1.38			
23	Stary Sącz	0.80	0.97	1.13	0.92	0.86	1.00	1.50	1.19			
24	Koszyce W.	3.63	0.80	-0.79	0.11	1.10	1.02	1.43	1.03			
25	Jarosław	1.08	1.00	3.11	1.05	0.99	1.01	2.07	1.23			
26	Radomyśl	0.42	0.98	1.69	0.84	0.51	1.00	1.63	0.96			
27	Tryncza	2.37	0.85	0.81	0.96	2.30	0.88	0.93	1.30			
28	Żółków	0.95	0.96	0.32	0.74	1.04	1.02	0.64	1.41			
29	Mielec	1.52	0.99	-0.15	0.83	0.85	1.01	-0.32	1.13			
30	Klęczany	-1.40	0.82	-6.42	0.10	1.46	1.01	1.95	0.93			
31	Wyszków	1.05	1.02	1.31	1.12	1.04	1.03	1.51	1.33			

Table 2. The ratio of estimated quantile $Q_{F=0.9}$ got by TS to ML for selected moments in time: t=1 (beginning of the time series -1951), t=25 (\sim middle of the series -1975), t=55 (end of the series -2005) and t=65 (prediction for the 10th yr after the time series -2015).

		$Q_{\rm est}^{\rm TS}/Q_{\rm est}^{\rm ML}$ for GEV					$Q_{\rm est}^{\rm TS}/Q_{\rm est}^{\rm ML}$ for Gumbel				
		<i>t</i> = 1	<i>t</i> = 25	<i>t</i> = 55	<i>t</i> = 65		<i>t</i> = 1	t = 25	<i>t</i> = 55	<i>t</i> = 65	
1	Jawiszowice	0.79	0.74	0.67	0.65		1.06	1.10	1.15	1.18	
2	Tyniec	0.87	0.80	0.56	0.33		0.93	0.84	0.56	0.32	
3	Jagodniki	0.91	0.99	1.11	1.15		1.02	1.08	1.16	1.20	
4	Szczucin	0.94	0.96	1.01	1.02		1.04	1.07	1.12	1.14	
5	Sandomierz	0.98	0.98	0.98	0.98		1.04	1.04	1.04	1.04	
6	Zawichost	0.96	0.98	1.01	1.01		1.02	1.04	1.06	1.07	
7	Puławy	0.98	1.00	1.01	1.02		1.04	1.04	1.05	1.05	
8	Warsaw	1.02	1.00	0.97	0.96		1.05	1.04	1.02	1.01	
9	Kępa		No r	esults			1.01	1.01	1.00	0.99	
10	Toruń		for	ML			1.02	1.00	0.97	0.96	
11	Tczew	1.03	1.01	0.98	0.96		1.04	1.02	1.00	0.99	
12	Żywiec	0.88	0.87	0.86	0.85		1.13	1.10	1.05	1.03	
13	Sucha	0.99	0.98	0.98	0.97		1.09	1.11	1.12	1.13	
14	Wadowice	0.98	0.98	0.98	0.98		1.09	1.11	1.13	1.14	
15	Rudze	1.28	1.05	0.89	0.85		1.13	1.05	0.97	0.95	
16	Stróża	0.92	0.92	0.94	0.94		1.08	1.07	1.07	1.07	
17	Proszówki	0.84	0.97	1.25	1.41		0.88	1.03	1.38	1.59	
18	Nowy Sącz	0.84	0.82	0.79	0.78		1.14	1.10	1.04	1.02	
19	Żabno	1.03	0.97	0.89	0.86		1.11	1.09	1.04	1.03	
20	Nowy Targ	0.92	0.87	0.77	0.71		1.09	1.07	1.01	0.97	
21	Zakopane	1.12	1.00	0.87	0.83		1.18	1.12	1.05	1.03	
22	Muszyna	1.13	1.03	0.87	0.81		1.15	1.09	0.99	0.94	
23	Stary Sacz	0.99	0.97	0.94	0.92		1.09	1.06	1.01	0.99	
24	Koszyce W.	0.73	0.86	1.01	1.07		1.04	1.07	1.10	1.11	
25	Jarosław	1.08	1.00	0.88	0.83		1.10	1.06	0.98	0.95	
26	Radomyśl	0.95	1.00	1.07	1.08		0.98	1.03	1.08	1.10	
27	Tryncza	0.94	1.00	1.05	1.07		1.00	1.06	1.11	1.12	
28	Żółków	1.05	0.98	0.93	0.91		1.16	1.09	1.03	1.02	
29	Mielec	1.01	0.99	0.96	0.95		1.07	1.06	1.04	1.04	
30	Klęczany	0.77	0.93	1.14	1.21		0.99	1.07	1.15	1.18	
31	Wyszków	1.08	1.03	0.87	0.70		1.11	1.07	0.93	0.77	

NHESSD

1, 6001-6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

4

Back

Close

▶

•

Full Screen / Esc

Printer-friendly Version



Table 3. The values of the trend coefficients and quantiles of annual and seasonal maximal flows for the 10 and 100 yr floods for the Warszawa–Nadwilanówka gauging station got by both methods of non-stationary flood frequency analysis.

Method	Best model		Values of the trend coefficients				$Q_{F=0.9}^t$				$Q_{F=0.99}^{t}$			
			а	b	С	d	t = 1	t = 25	t = 55	t = 65	<i>t</i> = 1	t = 25	t = 55	<i>t</i> = 65
ML	Year	Pe3	-3.89	2879.66	3.57	1190.24	3692	3657	3614	3599	5211	5286	5378	5409
	Summer	Pe3	-1.70	2235.67	-0.43	1477.75	3183	3136	3077	3057	5162	5100	5024	4998
	Winter	We3	-2.55	2180.96	1.11	924.52	2898	2858	2807	2790	3942	3932	3918	3914
TS	Year	Pe3	-2.79	2849.15	-0.06	1185.38	4517	4448	4362	4334	6833	6762	6672	6643
	Summer	Pe3	0.56	2165.83	-10.80	1633.51	4470	4115	3672	3524	7698	6828	5740	5378
	Winter	We3	-7.54	2320.41	6.93	744.34	3365	3417	3482	3504	4607	4934	5342	5478
	Year by se	easonal a	approach (proach (summer: Pe3, winter: We3)			4494	4323	4114	4040	7570	6749	6007	5872

NHESSD

1, 6001-6024, 2013

Application of a hybrid approach in NFFA – a Polish perspective

K. Kochanek et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

I∢

►I ►

Back

Close

Full Screen / Esc

Printer-friendly Version





Discussion Paper

Discussion Paper

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



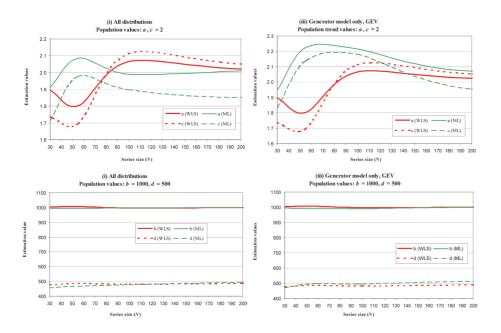


Fig. 1. Average values of the estimated trends in mean and standard deviation got by the two methods (WLS - the thicker lines and ML) in 1000 Monte Carlo simulations.

NHESSD

1, 6001-6024, 2013

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

▶I

14

Back

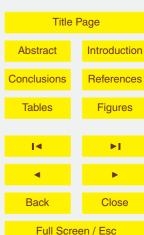


1, 6001-6024, 2013

NHESSD

Application of a hybrid approach in NFFA - a Polish perspective

K. Kochanek et al.



Printer-friendly Version



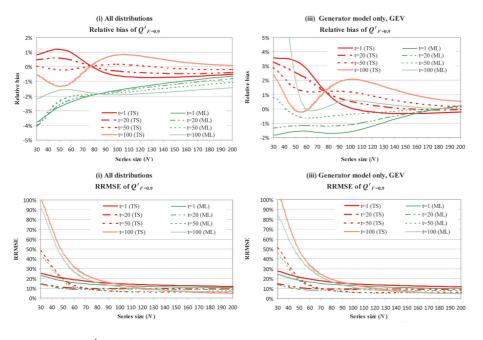


Fig. 2. The quantile $Q_{F=0.9}^t$ estimation errors got by the two methods (TS – the thicker lines and ML) and selected moments in time in 1000 Monte Carlo simulations.