



**Inundation risk for
embanked rivers**

W. G. Strupczewski et al.

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

Inundation risk for embanked rivers

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

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

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

Received: 31 January 2013 – Accepted: 19 February 2013 – Published: 8 March 2013

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

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

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

The Flood Frequency Analysis (FFA) concentrates on probability distribution of peak flows of flood hydrographs. However, examination of floods that haunted and devastated the large parts of Poland lead us to revision of the views on the assessment of flood risk of Polish rivers. It turned out that flooding is caused not only by overflow of the levees' crest but mostly due to the prolonged exposure to high water on levees structure causing dangerous leaks and breaches that threaten their total destruction. This is because, the levees are weakened by long-lasting water pressure and as a matter of fact their damage usually occurs after the culmination has passed the affected location. The probability of inundation is the total of probabilities of exceeding embankment crest by flood peak and the probability of washout of levees. Therefore, in addition to the maximum flow one should consider also the duration of high waters in a river channel.

In the paper the new two-component model of flood dynamics: "Duration of high waters–Discharge Threshold–Probability of non-exceedance" (DqF), with the methodology of its parameters estimation was proposed as a completion to the classical FFA methods. Such model can estimate the duration of stages (flows) of an assumed magnitude with a given probability of exceedance. The model combined with the technical evaluation of probability of levees breach due to the d -days duration of flow above alarm stage gives the annual probability of inundation caused by the embankment breaking.

The results of theoretical investigation were illustrated by a practical example of the model implementation to the series of daily flow of the Vistula River at Szczucin. Regardless promising results, the method of risk assessment due to prolonged exposure of levees to high water is still in its infancy despite its great cognitive potential and practical importance. Therefore, we would like to point out the need for and usefulness of the DqF model as complementary to the analysis of the flood peak flows, as in classical

HESSD

10, 2987–3025, 2013

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



FFA. The presented two-component model combined with the routine flood frequency model constitutes a new direction in FFA for embanked rivers.

1 Introduction

The most popular way of flood protection in Poland is the embankment of the rivers. In consequence of this passive way of protection, floods in Poland occur mostly due to the levee breach or to flow over the crest of dikes. Sense of security in floodplains of embanked rivers results from the belief that levees protect against the flood magnitude for which they were designed. So it creates the illusion that if the actual forecasted flood peak does not exceed the safety levels related to levee's designed value one can assume that the risk of water overtopping the dike crest is negligible and so is the risk of flooding in the protected area. The records of floods in Poland show that this is not true; more often the floods are the result of the prolonged exposure to high water on levees. The levees are weakened by water and their disruption occurs when it seems that the danger is over, so after passing culmination. This is particularly dangerous because when the staff responsible for flood protection and local residents breathe sigh of relief the worst is yet to come.

Therefore, apart from the magnitude of the peak flows another important factor should be taken into consideration, the duration of high water levels, in fact, a parameter of the wave's shape. Long-lasting high stages may weaken the levees' structure (soaking) and cause dangerous leaks, blurs and breaks that threaten their destruction. That is why the classical flood frequency analysis (FFA) concerning only the frequency of the annual maximum (AM) flows (Bogdanowicz et al., 2011) is not suitable in this case and ought to be supplemented by the analysis of the duration of flows over the given threshold. The joint risk of inundation making allowance for the two main sources of vulnerability to flood hazard for areas protected by embankments, over-crest flow and levees failure, has been proposed and defined.

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inundation risk for embanked rivers

W. G. Strupczewski et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[I ◀](#)[▶ I](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

In Poland, as in many other countries for each hydrological station two benchmark water levels, called the warning stage and the alarm stage, have been specified. Although warning and alarm stages are assigned to the places where water levels are observed, to the hydrological stations, their determination procedures as well as other inundation risk characteristics take into account, inter alia, the elevation of the embankment system for the whole river reach. So, the results of below analysis refer to the river reaches represented by data observed at hydrological stations. The frequency of annual maximum uninterrupted duration, D (in days), of flows over the flood alarm stage (Fig. 1) can be used to assess the risk of flooding due to waning of the levees' strength. The aim of this study is to introduce formal aspects of the Duration–flow–Frequency (DqF) modelling in stationary and non-stationary conditions, to use it to assess the inundation risk due to the levees breach and to combine it with the AM flow model to get the cumulative probability of inundation. In the presented statistical model, the duration is considered as a random variable while the alarm flow discharge is the fixed value.

The paper is built as follows: in the second section the concept of the inundation risk for embanked river is defined. Then a short review of literature on statistical modelling of flood shape hydrographs with emphasis on one-dimensional models is presented (Sect. 3). In the next section the Duration-Flow discharge-Frequency (DqF) model is introduced and estimations of its parameter for stationary and non-stationary case are described and discussed. Taking into account the embankment resistance, the annual probability of inundation caused by levees breaching is introduced. To illustrate the proposed way of inundation risk assessment the case study for the Szczucin gauging station at the Vistula River (Southern Poland) is presented (Sect. 5). The probability of inundation due to levees breaching is compared with the conventional probability of peak flow exceeding the levee crest and the cumulative probability of inundation are computed. The Sect. 6 concludes the paper.

2 Flood risk

Floods occur as a result of water spilling over the crest of embankment ($Q > Q_B$) or more often as a result of prolong existence of high water in the embanked river channel, so when the peak flow discharge exceeds the alarm flow (Q_A) but is lower than the overtopping flow (Q_B is the discharge that overtops levee crests) ($Q_A < Q_{\max} < Q_B$). One can also distinguish many other causes of floods, such as back water and ice-jams, etc., but they do not stem from the embankment failures and will not be considered in this study.

The annual probability of inundation for embanked river reach is expressed as the total of probability of the two exclusive events (FI stands for “flood”) (see Fig. 2):

$$P(\text{FI}) = P_1(\text{FI}) + P_2(\text{FI}) \quad (1)$$

where the first term comes from the conventional FFA

$$P_1(\text{FI}) = \rho(Q_{\max} > Q_B). \quad (2)$$

The second term of Eq. (1) defines the probability of inundation caused by levees breaching which depends on both the flood persistency and levees resistance to high water stages which in turns depends on their design and technical condition. Therefore the $P_2(\text{FI})$ is expressed as the integral of the product of the value of the hazard index $\rho(\text{FI}|d)$ which is defined as the probability of levee breaching caused by the d -days duration of flow over the flow level Q_A and of the pdf of the the d – duration, so $f(d)$ for annual peak flows in the interval $Q_A < Q_{\max}(t) < Q_B$.

$$P_2(\text{FI}) = \rho(\text{FI}|(Q_A < Q_{\max} \leq Q_B)) = \int_{0^+}^{\infty} h(\text{FI}|d) \cdot f(d) \cdot dd \quad (3)$$

where

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- $f(d)$ – pdf of the duration d of flows above the alarm stage;
- $h(F|d)$ – the hazard index being the probability of levee breaching caused by a high water of the duration d .

The value of the hazard index $h(F|d)$ tends to 0 for d going to 0 and to 1 for d going to infinity (e.g. Fig. 5). The hazard index $h(F|d)$ is determined administratively for the river reach by the Regional Water Management Board based on the technical assessment of flood embankments.

Note that collating the annual maximum high flow duration data for analysis one puts $d_t = 0$ both for $Q_{\max} \leq Q_A$ and $Q_{\max} > Q_B$, so 1 inundation yearly is considered and that caused by spilling over crest has the priority over one caused by prolonged high stages. Furthermore note that the weaker is the relationship between annual maximal values of peak flow and duration of flows above the alarm flow (Q_A) the more justified is the separate analysis of the both random variables.

The ratio of probabilities P_2 to P_1 and their total is helpful to determine the actions to reduce the risk of flooding, namely the strengthening or heighten the levees (or building parallel levees).

3 The statistical modelling of flood hydrographs shape

Due to complexity of stochastic nature of river flow process one has to accept a rational ignorance while dealing with flood risk management. In response to practical needs several simple conceptual structures are being developed for statistical modelling of flood hydrographs. The methods of constructing design flood hydrographs are most popular for modelling flood hydrographs. Their reviews is available in, e.g. Serinaldi and Grimaldi (2010), Strupczewski (1964, 1966) and Strupczewski et al. (2013). The design hydrograph $Q(t)$ with the defined return period of its peak serves both in flood-risk mapping procedures and for designing a reservoir storage capacity and other hydraulic structures sensitive for flood hydrograph magnitude and shape.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The common feature of most of the approaches to flood hydrographs analysis is an avoidance of using a joint probability distribution of parameters describing the shape of the hydrographs while limiting multi-dimensional analysis to conditional expectations further reduced to a regression. The most commonly used variables are flood peak and flood volume.

Extension of the standard FFA for statistical analysis of peak part of flood hydrographs is the one-dimensional model Flow-duration Frequency (QdF) initiated by NERC (1975) and Askhar (1980). In the nineties, Sherwood (1994), Balocki and Burgess (1994). Galea and Prudhomme (1997) laid out the foundations of the present form of the QdF method. Based on the assumption of the convergence of different flood distributions for small return periods. Javelle et al. (1999), Javelle (2001) introduced a converging approach to the QdF modeling. Here the annual mean maximum peak flood volume (or equivalently the mean excess discharge – \bar{Q}_d) corresponding to the given duration (d) is taken (Fig. 3a) as the random variable. Therefore consequently the maximum d -days annual outflow volume $V_d = d \cdot \bar{Q}_d$ is the random variable as well. In fact, the above idea of flood peaks analysis is modelled on the analyses of the Intensity-duration-Frequency (IdF) commonly used for stochastic modelling of high intensity rainfalls and of the QdF analysis of low flows.

To cater for the conventional FFA, the flow discharge (Q_A) corresponding to the alarm stage (H_A) is used here, so the upper limb of the rating curve is regarded as time invariant. The frequency of annual maximum uninterrupted duration of flows, D (in hours, days, etc.), over the flood alarm stage (H_A) (or equivalently over the alarm flow (Q_A)) but excluding floods pouring over the embankment crest (which corresponds to flows exceeding the overtopping flow Q_B) serves to assess the inundation risk of flood spilling out of river channel caused by scouring the levees (Fig. 2). Therefore, the $d_t = 0$ in the $[d]$ time-series denotes that the Q_A has not been exceeded during the t -th year ($Q_{\max}(t) < Q_A$) or that the peak flow has exceeded the overtopping flow ($Q_{\max}(t) > Q_B$) where Q_{\max} denotes the annual maximum discharge. Note that if more than one flood

appears in a year, the D and the annual peak flow (Q_{\max}) can correspond to different floods (Fig. 1).

Using multi-duration approach, by fitting the appropriate statistical distribution to the extracted samples for various durations, from the relations QdF for various d one can roughly construct the scaled Flood-duration-Frequency curve (QdF). To avoid inconsistency of the estimates of quantile $Q(d, F)$ for various d , the same distribution function is applied for all duration (Javelle et al., 1999) and the quantiles are reduced by the appropriate function $\phi(d, \nu)$ which is decreasing function of d :

$$Q(d, F) = \phi(d, \nu) \cdot Q(0, F) \text{ for } d = 0, 1, 2, \dots; \phi(0) = 1 \quad (4)$$

where the ν denotes the vector of parameters which are estimated from the data.

It means that differences in the distributions of various d values result from the differences in the mean value only. Note that $Q(0, F)$ corresponds to the distribution of annual instantaneous peak discharges. The parameters of the function $\phi(d, \nu)$ and $Q(0, F)$ (Eq. 5) are estimated separately.

Finding that flood persistence is a factor of flood hazard for embanked rivers, Bogdanowicz et al. (2008) modified the above model redefining Q as the annual maximum flow discharge (Q_d) which is continuously exceeded during the period d , wherein the d variable is still treated as a deterministic value (Fig. 3b). The applied way of determining the scaled distribution function does not differ much from the method described by Javelle et al. (1999). In parallel, the use of ML method in the presence of the d as the covariate (Strupczewski et al., 2001; Katz et al., 2002; Stasinopoulos and Rigby, 2007; Stasinopoulos et al., 2008, 2012) is demonstrated for Weibull distribution with the lower bound parameter and the constant shape parameter. Here all parameters are estimated jointly.

However to address the 1-D statistical analysis of the peak part of flood hydrographs directly to the problem of softening and breaching of river embankment, the duration (d) of high stages should be taken as a random variable rather than the mean excess

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



discharge \bar{Q}_d (Javelle, 2001) (Fig. 3a) or the the annual maximum flow discharge (Q_d) (Fig. 3b) (Bogdanowicz et al., 2008).

4 Formal aspects of the duration–flow–frequency modelling

To address the flood risks arising from softening and washing out the river embankments, Bogdanowicz et al. (2011) proposed to take as the subject of analysis the frequency of annual maximum uninterrupted duration, D (in days), of flows over the flood alarm stage (Q_A), the duration (D) is considered as a random variable while the alarm flow discharge (Q_A) is the fixed value (Fig. 1).

The time-series of annual maximum uninterrupted duration, D (in days), of flows over the flood alarm flow Q_A , $d = (d_1, d_2, \dots, d_t, \dots, d_T)$, is the subject of statistical modelling in stationary and non-stationary conditions. The $d_t = 0$, denotes that the Q_A has not been exceeded during the t -th year ($Q_{\max}(t) < Q_A$) or that the peak flow has exceeded the overtopping flow ($Q_{\max}(t) \geq Q_B$), which means that the priority of overtopping over breaching is given and we rule out the possibility of two inundation floods of the two different origins within one year. Note that the condition $Q_{\max}(t) \geq Q_B$ is equivalent to the unconditional inundation, from Eq. (2) $P_1(F|Q_{\max}(t) \geq Q_B) = 1$, while $Q_B > Q(t) \geq Q_A$ points only possible inundation (see Eq. 3).

Frequency analyses of hydrological sample with zero events have received relatively little attention. Still there are several approaches for analysis of censored data, including probability plot regression, weighted-moment estimators, maximum likelihood estimators, and conditional probability analyses (Gilliom and Helsel, 1986; Hass and Scheff, 1990; Harlow, 1989; Helsel, 1990). A consistent approach to the frequency analysis of such data requires using discontinuous probability distribution functions. Jennings and Benson (1969), Interagency Advisory Committee on Water Data (1982), Woo and Wu (1989), Wang and Singh (1995) among others developed empirical three-parameter models for frequency analysis of hydrologic data containing zero values.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



When the available water stage records have been sampled in daily intervals, the d values are integer numbers and in reality correspond to the duration ($d - 0.5$, $d + 0.5$), while in particular for $d = 0$ to the interval $(0, d + 0.5)$. If a flood starts at the end of a year and is continuing to the next year, the d value is derived for whole flood and ascribed to the year t when 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 .

4.1 Stationary conditions

As far as the probability theory is concerned, the occurrence of zero events can be expressed by placing a non-zero probability mass on a zero value: $P(D = 0) \neq 0$, where D is the random variable, and P is the probability mass (e.g. Strupczewski et al., 2002, 2003; Weglarczyk et al., 2005). Therefore, the parent distribution functions of such hydrologic series would be discontinuous (with discontinuity at 0) and, using the theorem of total probability, their forms can be written as:

$$f(d) = \beta \delta(d) + (1 - \beta) f^0(d; \mathbf{g}) \cdot 1(d) \quad (5)$$

where β denotes the probability of the zero event, $\beta = P(D = 0)$, $f^0(d; \mathbf{g})$ is the conditional probability density function (CPDF), $f^0(d; \mathbf{g}) \equiv f(d|D > 0)$, which is continuous in the range $(0, +\infty)$ with a lower bound of 0, and \mathbf{g} is the vector of parameters (containing β or not), $\delta(d)$ is the Dirac's delta function and $1(d)$ is the unit step function. Assuming the infinite upper bound for D seems acceptable and facilitates modelling. Due to discretised duration d intervals, the probability of exceeding the Q_A flow during one day only equals to

$$P(d) = \int_{d-1/2}^{d+1/2} f(d) \cdot dd.$$

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Hydrological samples with zero values are most frequently of exponential-like shape. Weglarczyk et al. (2005) model the continuous part of Eq. (5) by two-parameter distributions, namely by Generalized Pareto, Weibull and Gamma, estimating parameters by the maximum likelihood (ML) and the moments (MOM) methods.

5 4.1.1 Estimation of the weight parameter β

From the pdf of the duration d (Eq. 5) and the records $d = (d_1, d_2, \dots, d_t, \dots, d_T)$ for given alarm flow Q_A

From Eq. (5) one can write the likelihood function as:

$$L = \beta^{n_1} \cdot (1 - \beta)^{n_2} \prod_{j=1}^{n_2} f^0(d_j; \mathbf{g}) \quad (6)$$

10 where n_1 and n_2 denote the number of zeros and non-zeros values, respectively.

If $\beta \notin \mathbf{g}$, from ML-equations:

$$\frac{\partial \ln L}{\partial \beta} = \frac{n_1}{\beta} - \frac{n_2}{(1 - \beta)} = 0 \quad (7)$$

one can easily find that the ML-estimate of β is

$$\hat{\beta} = \frac{n_1}{n_1 + n_2} \quad (8)$$

15 so β and \mathbf{g} are estimated by MLM independently.

From CDF of annual maximum floods obtained from FFA

The better estimate of the β parameter in the sense of definition (Eq. 9), not its standard error, can be obtained from the CDF of annual peaks providing the selected for

Annual Maxima (AM) model fits well upper tail data. Note that the $D = 0$, denotes that the Q_A has not been exceeded during the t -th year ($Q_{\max}(t) < Q_A$) or that the peak flow has exceeded the overtopping flow ($Q_{\max}(t) > Q_B$) where Q_{\max} denotes the annual maximum discharge, therefore, probability of zero value of D

$$\hat{P}(D = 0) = \hat{P}(Q_{\max} < Q_A) + \hat{P}(Q_{\max} > Q_B) = \hat{\beta} \quad (9)$$

should be estimated from CDF of annual peak flows got from FFA rather than from the (0, 1) time series of the d record. Having derived from FFA the CDF of the annual peaks $\hat{G}(Q_{\max}) \equiv \phi(Q_{\max}, \hat{h})$ where \hat{h} is the vector of parameter estimates, one can get the estimate of β as

$$\hat{\beta} = \hat{G}(Q_{\max} = Q_A) + (1 - \hat{G}(Q_{\max} = Q_B)). \quad (9a)$$

Note that if more than one flood appears in a year it may happen that the d_t and the annual peak flow $Q_{\max}(t)$ correspond to different floods.

Floods in excess of Q_B are unique in Polish rivers, but if they were they should be in the FFA treated as of unknown magnitude over the threshold Q_B , thus one deals with first order right censored sample.

4.1.2 Estimation of parameters of the continuous part of Eq. (5)

ML estimate of the parameters (\mathbf{g}) of the continuous part of PDF (Eq. 6): the conditional probability density function (CPDF). $f^0(d; \mathbf{g}) \equiv f(d|D > 0)$ of $f^0(d; \mathbf{g})$, can be obtained by solving the ML system of equations:

$$\frac{\partial \ln L}{\partial \mathbf{g}} = \frac{\partial}{\partial \mathbf{g}} \sum_{j=1}^{n_2} \ln f^0(d_j; \mathbf{g}) = 0 \text{ for } \beta \notin \mathbf{g} \quad (10)$$

Since the samples with zero values are most frequently of exponential-like shape, the distribution functions in Table 1 are recommended as candidates for $f^0(d_j; \mathbf{g})$ model.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The detailed information on the models mentioned above with the methods of ML estimation, one can easily find in hydrological and statistical literature, e.g. in Rao and Hamed (2000) and for GE in Gupta and Kundu (2000).

4.2 Non-stationary case

5 The basic assumption in the classical Flood Frequency Analysis and the Duration-Flood-Frequency modelling is that neither the adopted distribution function nor its parameters change in time. However, the longer the hydrological series, the harder to maintain the assumption of stationarity in the face of a changing environment and climate (Milly et al., 2008). The non-stationarity of hydrological data ought to be taken
 10 into account in FFA for theoretical and empirical reasons, but practical aspects of its introduction into design and planning procedures are not so obvious and simple and pose significant ongoing challenges to the hydrological research and water management policy. One could easily accept the increasing trend in design upper quantiles, but decreasing detected trends may distort decision-making in the engineering design, evaluation of flood risk and in other flood-related issues. Especially when statistical inference is based on peak flow series of average length currently covering barely 60,
 15 70 elements or on climate change scenarios and their hydrological response that we presume, we are able to predict in a realistic manner. Herein the formal aspects of at site non-stationary Duration-Flow-Frequency modelling are presented while regional Flow-Duration-Frequency modeling being introduced by Cunderlik and Ouarda (2006).
 20

Assuming that only the values of parameters of the continuous part of the PDF may vary with time, but its form remains unchanged, the PDF f can be written as:

$$f(d|t) = \beta(t)\delta(d) + [1 - \beta(t)]f^0[d; \mathbf{g}(t)] \cdot 1(d) \quad (16)$$

Assuming the forms of trends and denoting the vectors of their parameters, respectively, as $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ we have got:

$$f(d|t) = \beta(t; \boldsymbol{\theta})\delta(d) + [1 - \beta(t; \boldsymbol{\theta})]f^0[d; t, \boldsymbol{\xi}] \cdot 1(d); \boldsymbol{\theta} \notin \boldsymbol{\xi}. \quad (17)$$

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



For compact notation let us define the dichotomous variable Y_t given by:

$$Y_t = \begin{cases} 1 & \text{for } D = 0 \\ 0 & \text{for } D > 0 \end{cases} \quad (18)$$

For the time series $d = (d_1, d_2, \dots, d_t, \dots, d_T)$ of the maximal annual duration of river flows exceeding the given threshold, the likelihood function can be expressed as:

$$L = \prod_{t=1}^T \beta(t; \boldsymbol{\theta})^{y_t} \cdot \prod_{t=1}^T (1 - \beta(t; \boldsymbol{\theta}))^{1-y_t} \cdot \prod_{t=1}^T f^0(d_t; t, \xi)^{1-y_t} \quad (19)$$

and the Log-likelihood function

$$\ln L = \sum_{t=1}^T y_t \cdot \ln(\beta(t; \boldsymbol{\theta})) + \sum_{t=1}^T (1 - y_t) \cdot \ln(1 - \beta(t; \boldsymbol{\theta})) + \sum_{t=1}^T (1 - y_t) \cdot \ln(f^0(d_t; t, \xi)) \quad (20)$$

As one can see from Eq. (20), the parameters $\boldsymbol{\theta}$ and ξ , as they are independent, can be estimated separately.

4.2.1 Estimation of parameters of the continuous part of Eq. (16) ($f^0(d; t, \xi)$)

The ML estimate of the parameters ξ of CPDF ($f^0(d; t, \xi)$) are obtained by solving the system of equations:

$$\frac{\partial \ln L}{\partial \xi} = \frac{\partial}{\partial \xi} \sum_{t=1}^T (1 - y_t) \cdot \ln(f^0(d_t; t, \xi)) = 0 \quad (21)$$

while the candidate functions f^0 are given by Eqs. (11)–(15) however with time dependent parameters in this case (Strupczewski et al., 2001, Strupczewski and Kaczmarek, 2001). The estimates can also be found by direct search for the maximum of

Log-likelihood function (the last component of Eq. 20) with respect to trend parameter vector ξ .

The consequence of making allowance for time dependent parameters of. $f^0(d; \mathbf{g})$ is an increase of the number of parameters to be estimated. Given the small number of non-zero elements in the time series $d = (d_1, d_2, \dots, d_t, \dots, d_T)$, the number of parameters which can be effectively estimated is small. Therefore, we decided to adopt the values of these parameters as independent of time. Then the only non-stationarity lies in the weighting parameter $\beta(t; \theta)$ which plays the role of the time-dependent function “switching” on and off the event of dikes’ prolonged exposure to high waters. Note here that the duration d is a parameter that describes the shape of the flood hydrograph, so we assume that the persistence of flood of magnitude $Q_A < Q_{\max} < Q_B$ is not subject to time variability.

4.2.2 Two ways of estimation the time dependent weight parameter $\beta(t; \theta)$

The estimation of parameters θ of the discrete part – weighting parameter $\beta(t; \theta)$, in the joint distribution Eq. (17) can be performed in two ways: by regression analysis and on the base of non-stationary distribution of annual maxima with time dependent parameters.

Regression analysis

As Y_t represents binary outcomes and has a binomial distribution with parameter

$$\beta(t; \theta) = P(Y_t = 1) = P(D = 0) \quad (22)$$

the trend in β can be found by means of logistic regression, which does not require the assumption that the error term is homoscedastic, nor it is normally distributed as in normal regression. In this case, it is assumed, that β is the Logistic function of time t with parameter vector $\theta = [a, b]$.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

A useful advantage of the Logistic function is that its domain is the set of all real numbers and the output range (probability) remains constrained between 0 and 1.

The Logistic regression coefficients a and b are usually determined using maximum likelihood estimation by iterative process until the improvement of the solution is minute and the procedure is said to have converged. Sometimes, when the considered flow threshold is high and thus number of “ones” greatly exceeds number of zero values of y_t , the convergence cannot be reached. The failure to converge may indicate that the trend coefficients are not significant or other methods of inference about the trend in β should be applied.

Several measures enable to evaluate the goodness of fitted trend model. Deviance, pseudo- R^2 and odds ratios confidence intervals are the most frequently used. There are two measures of deviance corresponding to the likelihood ratio. One, called model deviance, to compare fitted model to saturated model (a theoretical model with perfect fit) and second, null deviance, which represents the difference between null model (a model with only intercept, so representing the stationary case, β given by Eq. 8) and saturated model. Model deviance is given by equation:

$$D_{\text{model}} = -2 \ln \frac{\text{likelihood of the fitted model}}{\text{likelihood of the saturated model}} \quad (24)$$

and similarly, null deviance:

$$D_{\text{null}} = -2 \ln \frac{\text{likelihood of the null model}}{\text{likelihood of the saturated model}} \quad (25)$$

Note that in Logistic regression the likelihood of the saturated model ($y_t = \beta(t; \theta)$) is equal 1.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The deviance has an approximate chi-square distribution with 1 degree of freedom for each predictor, so 1 in our case. Smaller values of deviance indicates better fit what corresponds to non-significant chi-square values.

Pseudo - R^2 is calculated on the base of deviances:

$$5 \quad \text{Pseudo} - R^2 = \frac{D_{\text{null}} - D_{\text{model}}}{D_{\text{null}}} \quad (26)$$

and interpreted almost like a coefficient of determination in linear regression.

The method via annual maxima distribution with time-varying parameters

10 An alternative way of analyzing a trend in β is to use the non-stationary CDF of annual peaks with time dependent parameters. From NFFA (Strupczewski et al., 2001) one gets $G = \varphi(Q, \mathbf{h}, t)$ where \mathbf{h} – the vector of PDF parameters of the annual flood peaks distribution. Then per analogy to Eq. (9a) one can write:

$$\begin{aligned} \hat{\beta}(t) &= \hat{P}[D(t) = 0] = \hat{P}[Q_{\max}(t) \leq Q_A] + \left\{ 1 - \hat{P}[Q_{\max}(t) > Q_B] \right\} \\ &= \hat{G}(Q_A|t) + \left[1 - \hat{G}(Q_B|t) \right] \end{aligned} \quad (27)$$

15 providing the selected distribution and trend model of its parameters fits well upper tail of data. It would be advisable to compare the results of both methods. Compatibility of the results could serve as the overall test of correctness of the assumptions made.

4.2.3 Probability of inundation during the period (t_1, t_2)

Dealing with hydrologic design, due to non-stationarity, the notion of return period is no longer valid and the probability of inundation should refer to the whole period of life of a hydraulic structure, not to a single year as has been agreed in the stationary case.

20 When the parameters of DqF distribution are time dependent, consequently the annual probability of levee breach (Eq. 3) becomes time dependent: $P_2(FI, t)$. The

probability that at least once in the period (t_1, t_2) the inundation caused by levees breach occurs is expressed as:

$$P_2(\text{FI}, (t_1, t_2)) = 1 - \prod_{t=t_1}^{t_2} [1 - P_2(\text{FI}, t)] \quad (28)$$

Similarly, if the distribution of annual maximum peaks is time dependent, $G = \varphi(Q, h, t)$, the exceedance probability of overflow of the levees' crest, so the probability that (see Eq. 2), $P(Q_{\max} \geq Q_B, t) = 1 - G(Q_B, t) = P_1(\text{FI}, t)$ is time dependent. Then the probability that the inundation caused by overtopping the embankment crest occurs at least once in the period (t_1, t_2) and can be expressed as

$$P_1(\text{FI}, (t_1, t_2)) = \rho(Q > Q_B, (t_1, t_2)) = 1 - \prod_{t=t_1}^{t_2} [1 - P_1(\text{FI}, t)] \quad (29)$$

The total probability of inundation in the period (t_1, t_2) equals to:

$$P(\text{FI}, (t_1, t_2)) = P_1(\text{FI}, (t_1, t_2)) + P_2(\text{FI}, (t_1, t_2)) \quad (30)$$

5 Case study – Szczucin at Vistula River (Southern Poland)

To illustrate how the proposed approach works in practice the Szczucin gauge (southern Poland) at the Vistula River has been selected as the case study. The daily flows record covering the period 1951–2006 ($n = 56$ yr) was used in this study. At first the daily records have been controlled and tested with regard to the sharp discontinuities and jumps in data – no particular irregularities have been detected (Fig. 4).

The overtopping flow Q_B was assessed from the rating curve as $10\,500 \text{ m}^3 \text{ s}^{-1}$ which roughly corresponds to two-hundred-years return period of annual peak flow ($Q_{0.5\%}$), the base design value for the 1st class embankments. In fact, there are no annual

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

peak flows exceeding this value in the record. Therefore the Q_B value does not affect the composition of the vector of observation values (d_t). The alarm threshold for the Szczucin station $Q_A = 1690 \text{ m}^3 \text{ s}^{-1}$ (which means flow of ca. 2-yr return period, stage 660 cm), however, for completion a few other thresholds will be analysed, too, namely $Q_{Tr} = 700, 1000, 1300$ and $2000 \text{ m}^3 \text{ s}^{-1}$. The hazard index $h(\text{FI}|d)$ for $Q_A = 1690 \text{ m}^3 \text{ s}^{-1}$ (Eq. 3) was assessed as:

$$h(\text{FI}|d) = \begin{cases} 0.05 \cdot d & \text{for } d \leq 20 \text{ days} \\ 1 & \text{for } d > 20 \text{ days} \end{cases} \quad (31)$$

so the embankments cannot withstand the pressure of high waters of more than 20 days.

5.1 Stationary case

The weak correlation between the durations when $d(t) > 0$ and the respective annual maxima $Q_{\max}(t)$ indicates the variety of shapes of flood hydrographs and, as a consequence, d cannot be represented (or replaced rather) in FFA by Q_{\max} . It implies the analysis of both d and Q_{\max} by (perhaps) two different types of models. As a model for the parameters of the f^0 function Generalised Exponential (GE) distribution has been chosen (e.g. Gupta and Kundu, 2000). Among the distributions presented in Eqs. (11)–(15) the GE distribution Eq. (14) performs relatively well in terms of the AIC value and shows stability of numerical ML solutions in estimation of $f^0(d; \mathbf{g})$ parameters, regardless the Q_{Tr} threshold applied for the calculations. The list of the GE estimated parameters of the two-component DqF model and β values for different Q_{Tr} including Q_A is presented in Table 2.

The annual maxima are believed to be adequately described by the heavy-tailed distributions (e.g. Strupczewski et al., 2011), so to cater for the Flood Frequency Analysis (FFA) for extreme values (annual maxima) the β values (Eq. 8) and $P_1(\text{FI})$ (Eq. 2) by means of Q_{\max} series were calculated with the three-parameter Generalised Extreme

Value distribution:

$$G(q; \alpha, \gamma, \varepsilon) = \exp \left\{ - \left[1 - \frac{\gamma}{\alpha} (q - \varepsilon) \right]^{1/\gamma} \right\} = G_{\text{GEV}}^{\gamma}(x) \quad (32)$$

From the AM sample covering the period 1951–2006 ($n = 56$ yr) we got the ML estimates of GEV parameters equal: location $\equiv \hat{\varepsilon} = 1260.02 \text{ m}^3 \text{ s}^{-1}$, scale $\equiv \hat{\alpha} = 671.39 \text{ m}^3 \text{ s}^{-1}$ and shape $\equiv \hat{\gamma} = -0.33$. For completion note that the value of Log-likelihood function $\ln L = -463.231$ and thus $\text{AIC} = 932.461$.

Substituting for q into Eq. (32) the chosen Q_{Tr} and Q_{B} values and then putting the corresponding probabilities to Eq. (9a), one gets the estimates of the weighting parameters display in Table 2.

One can notice from the Table 2 that β got by means of Eqs. (8) and (9) are quite similar particularly for higher values of Q_{Tr} and for all cases the confidence interval for proportion β includes the value estimated from AM distribution (Eq. 9).

5.1.1 Assessment of probability of levee breach along Szczucin reach

Since the event of levee breach is conditioned by the peak flow being in the range of $[Q_{\text{A}}, Q_{\text{B}}]$, Eq. (3) can be written as (see also Fig. 5)

$$P_2(\text{FI}) = (1 - \beta) \int_{0^+}^{\infty} h(\text{FI}|d) \cdot f^0(d) \cdot dd \quad (33)$$

The pdf of GE (Eq. 15) for $Q_{\text{Tr}} = Q_{\text{A}} = 1690 \text{ m}^3 \text{ s}^{-1}$ (Table 2) takes the form

$$(1 - \hat{\beta}) \cdot f^0(d; \alpha = 3.4238, \gamma = 0.8357) = 0.423 \frac{0.2441 \cdot \exp(-d/3.4238)}{[1 - \exp(-d/3.4138)]^{0.1643}} \quad (34)$$

while the ML estimate of β equals (Table 2) 0.577. Substituting them and the hazard index function defined by Eq. (31) into Eq. (33) and integrating one gets the annual

3006

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5

10

15

20

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I ◀

▶ I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



probability of levee breaching $P_2(\text{FI}) = 0.064$. Note that at the same time, and when the same GEV distribution is used (see the Eq. 32 and its parameters below the equation), the probability of flood caused by exceeding embankment crest by annual peak flow: $P_1(\text{FI}) = P(Q_{\max} > Q_B = 10\,500 \text{ m}^3 \text{ s}^{-1}) = 1 - G(Q_B)$ is equal to 0.005, so it is almost insignificant (more than ten times smaller than P_1), hence, the overall probability of flood along Szczucin reach $P = P_1 + P_2 = 0.069$.

Variety of shapes of flood hydrographs one can evaluate by a measure of correlation strength between $Q_{\max}(t)$ and $d(t)$. Due to shape similarity of flood peak parts, a strong dependence between the peak flows (Q_{\max}) and the duration above the alarm flow (d) can take place. If it is a case, the probability $P_2(\text{FI})$ can be assessed on the base of Q_{\max} distribution $g(Q_{\max})$. Assuming that $d = \psi(Q_{\max})$ one can expressed in Eq. (33) the d variable by the Q_{\max} getting

$$P_2(\text{FI}) = p(\text{FI} | (Q_A < Q_{\max} \leq Q_B)) = \int_{Q_A}^{Q_B} h(\text{FI} | \psi(Q_{\max})) \cdot g(Q_{\max}) \cdot dQ_{\max} \quad (35)$$

where per analogy to Eq. (31) $h(\text{FI} | \psi(Q_{\max}))$ equals 0 and 1 for Q_A and Q_B , respectively. The Pearson's correlation coefficient $r(Q_{\max}, d)$ for Szczucin equals to 0.83.

Of course, when estimating the risk of a levee breach except the time of high water residence, more technical parameters of levees should be analysed, such as the construction of the levee, the material used for its building, its age, susceptibility to softening, the regime of the river, wind-induced waving and so on. All in all, those who decided to build their houses in the river's proximity behind the levees, sooner or later do experience a catastrophe.

5.2 Non-stationary case

Analysis of long series of hydrological observations on Polish rivers lead us to the conclusion that two random variables whose probability distributions have been considered

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



as components of DqF analysis show different behaviour versus time. The continuous variable – duration of water level above certain stage – in general, shows no trend. It describes the shape of the flood waves which has been stated to be rather stable and, if any trend there exists, it does not pose any effect on the final results of the DqF calculations. On the other hand, a visual assessment of records for Szczucin and other hydrological stations show that the frequency of occurrence of extreme flows (P_1) and flows above (so well below) a given threshold (Q_{Tr}) may reveal some trend. Therefore in this study we focused only on the search of trends in the probability P_1 and in the weighting factor β that plays the role of the time-dependent function “switching” on and off the event of dikes’ prolonged exposure to high waters. These trends have been estimated from the annual peak flow series and by direct analysis of $[d_t]$ vector represented by the sequence of 0 and 1 as given by Eq. (18). In both cases the maximum likelihood method (MLM) has been used for calculation, while the Logistic function (23) serves to model the (0,1) duration series.

The estimation $\beta(t)$ for the threshold corresponding to the alarm stage ($Q_A = 1690 \text{ m}^3 \text{ s}^{-1}$) in the form of the Logistic function Eq. (26) revealed the decreasing trend ($b < 0$), whereas $a = 0.405$, so the $\beta(t)$ takes the form:

$$\beta^{LO}(t) = [1 + \exp(0.002 \cdot t - 0.405)]^{-1} \quad (36a)$$

and the parameters of stationary f^0 function for selected Q_{Tr} values can be found in Table 2.

The above equation (Eq. 36a) says that the odds (the ratio of probabilities of events against nonevents: $\beta(t)/(1 - \beta(t))$) decreases in average by 0.2 % from year to year, that gives the change of β from ca. 0.60 in 1951 to about 0.58 in 2006. However this trend is not statistically significant. The model deviance D_{model} being equal to 75.8286 and the null deviance $D_{\text{null}} = 75.8372$ give the difference with p -value of 0.9264 from chi-square distribution. The value of pseudo- $R^2 = 0.046$ is close to 0. It is likely that this result points on almost stable risk of inundation caused by dike breaches for summer floods that prevail in the reach of the Vistula river represented by Szczucin hydrological

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



station, where changes in the river bed and on the floodplains have not influenced considerably the transportation of high waters. Winter floods can reveal stronger trends due to greater variability of melting condition and observed temperature rise, so, as consequence, the volume of runoff. Small catchments seem to be more susceptible for trends in β . These statements ought to be verified on the larger hydrological data set.

If instead of the Logistic we take the non-stationary Generalised Exponential distribution function (see stationary case above), assume linear trends in mean value and standard deviation (but not in the parameters of location, scale and shape) and calculate the $\beta(t)$ by means of Non-stationary Flood Frequency Analysis (e.g. Strupczewski et al., 2001, 2009) we obtain:

$$\beta^{\text{GEV}}(t) = \exp \left\{ - \frac{[t \cdot (33.067 \cdot t - 22453.3) + 3.812 \times 10^6]^{1.54}}{[1.43 \cdot t + 1.339 \cdot \sqrt{t \cdot (33.067 \cdot t - 22453.3) + 3.812 \times 10^6} - 275.269]^{3.08}} \right\} \quad (36b)$$

The comparison of the values of the non-stationary Log-likelihood function and AIC, $\ln L = -463.078$ and $\text{AIC} = 936.157$, respectively with the stationary results reveals that the supplement by two extra parameters to the model (those responsible for the linear trend in mean and standard deviation) worsen the estimation results. It means that for a given series size ($n = 56$) the detected trends are in fact weak, and perhaps addition of a few new measurements in series can dramatically change their value or even sign. The weakness of the trends in moments are confirmed by the weakness of β time-variability.

The time variability of β functions got by the two approaches are shown in the Table 1.

The above equations (Eqs. 36a and 36b) and the diagram (Table 1) point at the difference in trend sign of β between the results received by the two approaches (LO and GEV). However, there are similarities, too. The results for both cases say that the

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



value of β is practically time independent (statistically insignificant) within time period 1951 to 2006 and thus maintain the relatively constant balance between the first and the second terms of the DqF probability density function (Eq. 17). In consequence, the certain durations of water stay above Q_A described by the f^0 function are actually as frequent nowadays as they were in past. On the other hand, the probability P_1 and P_2 (and thus P) are now the functions of t . If we take the GEV-based $\beta(t)$ as an example (as more reliable than LO-based $\beta(t)$) and $t = 1$ (year 1951) one obtains $P_2 = 0.066$. Further, with the non-stationary GEV (by the same parameters as for $\beta(t)$): $P_1 = 0.007$, so in consequence $P = 0.073$. For $t = 56$ (year 2006): $P_1 = 0.004$, $P_2 = 0.064$, so $P = 0.068$, thus the probability of flood in Szczucin dropped by 7% over the half of the century – a judgement whether it is much or not we leave for the reader and decision makers. Please also note that regardless the point in time the ratio P_1/P_2 is similar to the stationary conditions.

However, the probability for the certain point in time may not carry information sufficient for flood protection authority. Therefore, it is interesting to know what is the probability of inundation over the certain period, e.g. 20 yr of the exploitation of the dikes in Szczucin. For the GEV non-stationary model (with the parameters mentioned above) and last 20 yr of the time series (1986–2006) the probability of overtopping over the levee crest is equal to $P_1 = 0.048$, whereas the dike's breach probability is more than 10 times larger: $P_2 = 0.516$. Overall risk of inundation $P = 0.563$, it is almost 10 times larger than for a single year. The reader also notes easily that again the ratio P_1/P_2 is alike the ratios for the point-in-time non-stationary case as well as for the stationary case.

One has to bear in mind, however, that the linear trend in parameters (in case of the LO) and first two moments (as it was in GEV) is just the simplest of the countless trend patterns that may be employed for the time-dependent models and application of other ways (e.g. parabolic, polynomial, exponential, etc.) usually leads to the overparametrisation and noteworthy complication of numerical calculations. It is so, because maximum likelihood estimates for time-dependant models require

multi-parameter optimisation of relatively “flat” Log-likelihood functions with use of relatively short data-series.

6 Conclusions

In the paper the new two-component model of flood waves, “duration of flooding-discharge-probability of non-exceedance” (DqF), with the methodology of its parameters estimation was proposed as a completion to the classical FFA methods. Such model can estimate the duration (d) of stages (and flows) exceeding the assumed magnitude with a certain probability which is of key importance when the river’s dikes are prone to the prolonged impact of high waters. The embankments may be weakened by the water, soak and eventually break – this is the most frequent cause of floods in Poland. However, in this study the two main causes of inundation of embanked rivers, namely over-crest flow and wash out of the levees, were combined to assess the total risk of inundation. The proposed DqF modelling approach was generalised to the non-stationary conditions. Therefore, in addition to the maximum flow one should consider also the duration of high waters above the alarm flow Q_A in a river channel. The model combined with the technical evaluation of probability of levees breach expressed by the hazard index gives the annual probability of inundation caused by the embankment failure. The probability of inundation is the total of probabilities of exceeding embankment crest by flood peak and the probability of washout of levees.

The DqF modelling is the consequence of QdF approach developed by Javelle et al. (1999, 2000, 2002) and Bogdanowicz et al. (2008) but in the first model the gravity is put on the probability of the certain duration above alarming stage/discharge (Q_A) rather than on magnitude of flood itself (Q_{max}) like in the latter case (Fig. 3).

The DqF model in the form of Eq. (5) consists of two terms: $\beta \cdot \delta(d)$ deals with the zero event, $D = 0$, whereas the latter term $(1 - \beta) \cdot f^0(d; \mathbf{g}) \cdot 1(d)$ stands for the events when the duration $D > 0$. In general both β and f^0 in non-stationary case may depend on time. The maximum likelihood method (MLM) was proposed for estimation of β and \mathbf{g}

HESSD

10, 2987–3025, 2013

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



parameters. In the non-stationary case it is convenient to describe the $\beta(t; \theta)$ by means of the logistic function Eq. (23). However, β and $\beta(t; \theta)$ can be also estimated by means of annual peak flows series, Q_{\max} , using the routine flood frequency techniques (FF) with distribution functions commonly used in FFA (e.g. GEV) for stationary and non-stationary case, respectively. Note that estimating the weighting factor β and $\beta(t; \theta)$ from the duration d time series the information (0,1) for excess the threshold level Q_{Tr} is used exclusively, while basing on the annual peak flow time-series Q_{\max} the information from whole range of recorded flood magnitude is used to assess the trend in the alarm flow Q_A . For $f^0(d_j; \mathbf{g})$ model (both stationary and non-stationary) the exponential-like shaped distribution functions are recommended, such as: Exponential, Weibull, Pareto, Generalised Exponential, Gamma and similar.

The calculations for the Szczucin at the Vistula River case study made for several threshold values (Q_{Tr}) including the alarm flow (Q_A) have showed the similar results for the weighting factor β estimated by ML method from the duration time-series and from annual peaks time-series (Table 2). The peak flows that could overtop the embankments have not been detected in the Szczucin's record (1951–2005). According to the hazard function Eq. (31) the possibility of levees breaching increases almost tenfold the probability of inundation.

Variability in the Szczucin time series of the d -duration (understood as a time-dependence of the \mathbf{g} parameters of $f^0(d_t; \mathbf{g})$) has not been subject of modelling because of the insufficient data and the conviction based on the visual judgment of the ($d(t)$ vs. t) diagram that the trend would be negligibly small. The only trend considered is the trend in the weighting factor β . The significant difference in trend estimates of $\beta(t)$ got by ML method from the direct analysis of (d_t) vector represented by the sequence of 0 and 1 (Eq. 18) assuming the logistic function (LO) of time and from GEV distributed annual peak flow series is striking. The results for both cases differ in sign (Fig. 6) and moreover they point that the value of β is practically time independent within time period 1951 to 2006. Nevertheless, as long as the change of river regime

in time is visible (regardless its origin), one should consider non-stationary modelling accepting (sadly) the fact that the tools available are in their infancy.

The DqF model proved to be the important completion to the traditional FFA concentrating on maximal seasonal or annual discharges. The DqF approach is especially useful in polish specific conditions where the flood protection infrastructure is dated and often does not survive confrontation with prolonged pressure of high waters.

Reliable data and information about floods are indispensable for better understanding the interactions between rivers and flood protection system: embankments, reservoirs and polders. Improvement of statistical models is essential for engineering design in general and in particular for implementation of flood risk mitigation procedures. Not only has the DqF modelling shown that actual flood risk is greater than the risk assessed by means of classical FFA but also provides quantitative measures which can be used in flood protection systems planning, exploitation and conservation. This measures in form of dependence of inundation risk on river flow (or water level) should be established for other hydrological stations on Polish rivers and their dimensionless versions compared. The geographic information systems technique (GIS) could be used to indicate locations prone to inundation, Also the GIS can be a helpful tool to visualisation and testing trends in the structure of river network and to the regional analysis. These results can constitute the theoretical background to a number of practical decisions in water management issues.

Acknowledgements. The authors would like to thank Janusz Żelaziński for inspiring discussions and exchange of ideas, without which this work would not have its present form and meaning. Sincere thanks for a sense of realism and keeping both feet on the ground what in face of many sources of uncertainty, so typical for hydrology, helped us to work out practical solution to the problem.

This research project was partly financed by the grant of the Polish National Science Centre titled “*Modern statistical models for analysis of flood frequency and features of flood waves*”, decision no. DEC-2012/05/B/ST10/00482, the Grant Iuventus Plus IP 2010 024570 “*Analysis*”

HESSD

10, 2987–3025, 2013

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of the efficiency of estimation methods in flood frequency modelling” and made as the Polish contribution to COST Action ES0901 “European Procedures for Flood Frequency Estimation (FloodFreq)”.

References

- 5 Ashkar, F.: Partial duration series models for flood analysis, PhD Thesis, Ecole Polytechnique de Montréal, Montréal, Canada, 1980.
- Balocki, J. B. and Burges, S. J.: Relationships between n -day flood volumes for infrequent large floods, *J. Water Res. Pl.-ASCE*, 120, 794–818, 1994.
- 10 Bogdanowicz, E., Strupczewski, W. G., and Kochanek, K.: Application of Discharge-duration-Frequency model for description of peak part of flood hydrograph, *Przegląd Geofizyczny*, LIII, 3–4, 263–288, 2008 (in Polish).
- Bogdanowicz, E., Strupczewski, W. G., and Kochanek, K.: Persistence as a factor of flood hazard for embanked rivers, EGU Leonardo Conference “Floods in 3-D”, Bratislava, 23–25 November 2011, Abstract in Proceedings, 2011.
- 15 Cunderlik, J. M. and Ouarda, T. B. M. J.: Regional flood-duration-frequency modelling in changing environment, *J. Hydrol.*, 318, 276–291, 2006.
- Galéa, G. and Prudhomme, C.: Notations de bases et concepts utiles pour la comprehension de la modélisation synthétique des régimes de crue des bassins versants au sens des modeles QdF, *Rev. Sci. Eu.*, 1, 83–101, 1997.
- 20 Gilliom, R. J. and Helsel, D. R.: Estimation of distributed parameters for censored trace level water quality data-1. Estimation techniques, *Water Resour. Res.*, 22, 1201–1206, 1986.
- Gupta, R. D. and Kundu, D.: Generalized exponential distribution: different method of estimations, *J. Stat. Comput. Sim.*, 69, 315–337, doi:10.1080/00949650108812098, 2001.
- Haas, C. N. and Scheff, P. A.: Estimation of averages in truncated samples, *Environ. Sci. Technol.*, 24, 912–919, 1990.
- 25 Harlow, D. G.: Effect of proof-testing on the Weibull distribution, *J. Material Sci.*, 24, 1467–1473, 1989.
- Helsel, D. R.: Less than obvious: statistical treatment of data below detection limit, *Environ. Sci. Technol.*, 24, 1767–1774, 1990.

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Interagency Advisory Committee on Water Data: Guidelines for determining flood flow frequency, Bulletin 17B, US Department of the Interior, Geological Survey, Office of Water Data, Reston, Va., 1982.

Javelle, P.: Caractérisation du régime des crues: le modèle débit-durée-fréquence convergent, Approche locale et régionale, PhD thesis, Camagref-Lyon, Institut National Polytechnique de Grenoble, 268 pp., 2001.

Javelle, P., Grésillon, J. M., and Galéa, G.: Discharge-duration-frequency curves modelling for floods and scale invariance. *Comptes Rendus de l'Academie des Sciences, Sciences de la terre et des planets*, 329, 39–44, 1999.

Javelle, P., Galéa, G., and Grésillon, J. M.: L'approche debit-durée-fréquence: historique et avancées, *Revue des Sciences de la terre et des planètes*, 329, 39–44, 2000.

Javelle, P., Ouarda, T. B. M. J., Lang, M., Bobée, B., Galéa, G., and Grésillon, J. M.: Development of regional flood-duration-frequency curves based on the index-flood method, *J. Hydrol.*, 258, 249–259, 2002.

Jennings, M. E. and Benson, M. A.: Frequency curve for annual flood series with zero events or incomplete data, *Water Resour. Res.*, 51, 276–280, 1969.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, *Adv. Water Resour.*, 25, 1287–1304, 2002.

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.

NERC: Flood Studies Report, Estimation of flood volumes over different durations, *Nat. Environ. Res. Council, London, Vols. 1–5*, 1100 pp., 1975.

Rao, A. R. and Hamed, K. H.: *Flood frequency analysis*, CRC Press, 2000.

Serinaldi, F. and Grinaldi, S.: Synthetic design hydrographs based on distribution functions with finite support, *J. Hydrol. Eng. ASCE*, 16, 435–446, 2010.

Sherwood, J. M.: Estimation of volume-duration-frequency relations of ungauged small urban streams in Ohio, *Water Resour. Bull.*, 30, 261–269, 1994.

Stasinopoulos, D. M. and Rigby, R. A.: Generalized additive models for location scale and shape (GAMLSS) in R, *J. Stat. Softw.*, 23, 1–46, 2007.

Stasinopoulos, M., Rigby, B., and Akantziliotou, C.: *Introductions on How to Use the GAMLSS Package in R, 2nd Edn.*, 2008.

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Stasinopoulos, M., Rigby, B., and Akantziliotou, C.: Selecting Explanatory variables using ad-term, dropterm and step AIC, in: Instructions on How to Use the GAMLSS Package in R, 2nd Edn., 2012.

Strupczewski, W.: Flood hydrograph equationk, (in Polish: Równanie fali powodziowej), Bulletin du Service Hydrologique et Meteorologique, Wyd. Komunikacji i Lacznosci, Fascicule 57, 2/1964, 35–58, 1964.

Strupczewski, W.: Statistical analysis of shapes of flood hydrographs (in Polish: Statystyczna analiza kształtów fal wezbraniowych), Ph.D., Warsaw Technical University, 181 pp., plus Table 25 and Fig. 39, 1966.

Strupczewski, W. G. and Kaczmarek, Z.: Non-stationary approach to at-site flood-frequency modelling – Part II: 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 flood-frequency modelling – Part I: Maximum likelihood estimation, J. Hydrol., 248, 123–142, 2001.

Strupczewski, W. G., Singh, V. P., and Weglarczyk, S.: Physically based model of discontinuous distribution for hydrological samples with zero values, in: Proceedings of the Int. Conf. On WaRMAR, Kuwait, Surface Water Hydrology, edited by: Singh, V. P., Al-Rashed, M., and Sherif, M. M., A. A. Balkema Publishers, Swets & Zeitlinger, Lisse, 523–537, 2002.

Strupczewski, W. G., Weglarczyk, S., and Singh, V. P.: Impulse response of the kinematic diffusion model as a probability distribution of hydrologic samples with zero values, J. Hydrol. 27, 328–351, 2003.

Strupczewski, W. G., Kochanek, K., Feluch, W., Bogdanowicz, E., and Singh, V. P.: On seasonal approach to non-stationary flood frequency analysis, Phys. Chem. Earth, 34, 612–618, 2009.

Strupczewski, W. G., Kochanek, K., Markiewicz, I., Bogdanowicz, E., Weglarczyk, S., and Singh, V. P.: On the tails of distributions of annual peak flow, Hydrol. Res., 42, 171–192, doi:10.2166/nh.2011.062, 2011.

Strupczewski, W., Bogdanowicz, E., and Kochanek, K.: Discussion of “Synthetic Design Hydrographs Based on Distribution Functions with Finite Support” by Francesco Serinaldi and Salvatore Grimaldi, J. Hydrol. Eng., 18, 121–126, doi:10.1061/(ASCE)HE.1943-5584.0000538, 2013.

- Wang, S. X. and Singh, V. P.: Frequency estimation for hydrological samples with zero values, J. Water Res. Pl. ASCE, 121, 98–108, 1995.
- Woo, M. K. and Wu, K.: Fitting annual floods with zero flows, Can. Water Resour. J., 14, 10–16, 1989.
- 5 Weglarczyk, S., Strupczewski, W. G., and Singh, V. P.: Three-parameter discontinuous distributions for hydrological samples with zero values, Hydrol. Process., 19, 2899–2914, 2005.

HESSD

10, 2987–3025, 2013

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 1. Distribution functions recommended as $f^0(d_j; \mathbf{g})$ model.

Distribution name	Probability density function	Parameters	Equation no.:
Exponential (Ex)	$f^0(d; \alpha) = \frac{1}{\alpha} \exp(-d/\alpha)$	α – scale	(11)
Weibull (We) distribution	$f^0(d; \alpha, b) = \frac{b}{\alpha} \left(\frac{d}{\alpha}\right)^{b-1} \exp(-d/\alpha)$	α – scale $b > 0$ – shape	(12)
Generalized Pareto (Pa)	$f^0(d; \alpha, k) = \frac{1}{\alpha} \left(1 - \frac{k}{\alpha} d\right)^{1/k-1}$	$\alpha > 0$ – scale $k < 0$ – shape	(13)
Generalized Exponential (GE)	$f^0(d; \alpha, \gamma) = \frac{\gamma}{\alpha} \exp(-d/\alpha) [1 - \exp(-d/\alpha)]^{\gamma-1}$	$\alpha > 0$ – scale $\gamma > 0$ – shape	(14)
Gamma (Ga)	$f^0(d; \lambda, \alpha) = \frac{1}{\alpha^\lambda \Gamma(\lambda)} d^{\lambda-1} e^{-d/\alpha}$	$\alpha > 0$ – scale $\lambda > 0$ – shape	(15)

Note that Exponential distribution is a special case of all other mentioned above distributions, Eqs. (12)–(15).

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[I ◀](#)

[▶ I](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Inundation risk for
embanked rivers

W. G. Strupczewski et al.

Table 2. The parameters of the two-component DqF model for Szczucin data.

Q_{Tr}	First component (by two methods)			Second component, f^0 is the two-parameter Generalised Exponential		
	n_2	$\beta = n_1/n$ by Eq. (8)	$\beta = \beta(Q_{Tr})$ by Eq. (9)	scale	shape	$\ln ML/n_2$
700	51	0.089	0.076	2.8799	0.2938	-2.63
1000	40	0.286	0.226	4.0392	0.5228	-2.10
1300	32	0.429	0.395	4.8616	0.7464	-1.77
1690^a	23	0.589	0.577	3.4238	0.8357	-1.62
2000	17	0.696	0.683	3.7411	0.9126	-1.54

^a $Q_{Tr} = Q_A$

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I ◀

▶ I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inundation risk for embanked rivers

W. G. Strupczewski et al.

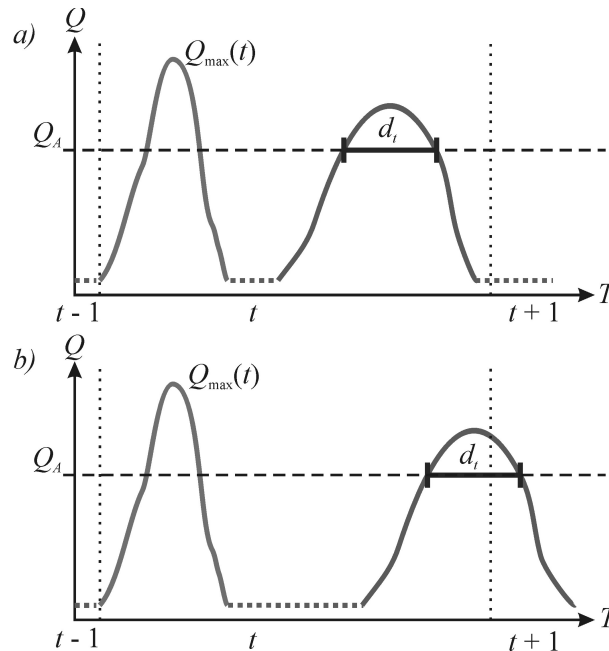


Fig. 1. Definition of the threshold flow discharge and duration in DqF model: **(a)** the flood wave of d_t duration entirely in the year t , **(b)** the flood wave starts in the year t and continues in $t+1$.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



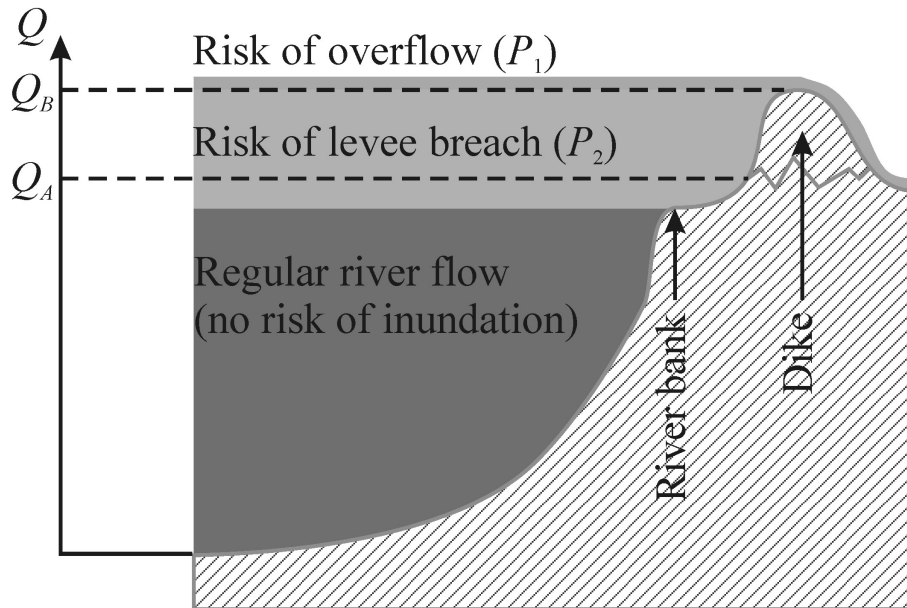


Fig. 2. Two reasons of inundation – an illustration.

Inundation risk for embanked rivers

W. G. Strupczewski et al.

[Title Page](#)

[Abstract](#) [Introduction](#)

[Conclusions](#) [References](#)

[Tables](#) [Figures](#)

[◀](#) [▶](#)

[◀](#) [▶](#)

[Back](#) [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Inundation risk for embanked rivers

W. G. Strupczewski et al.

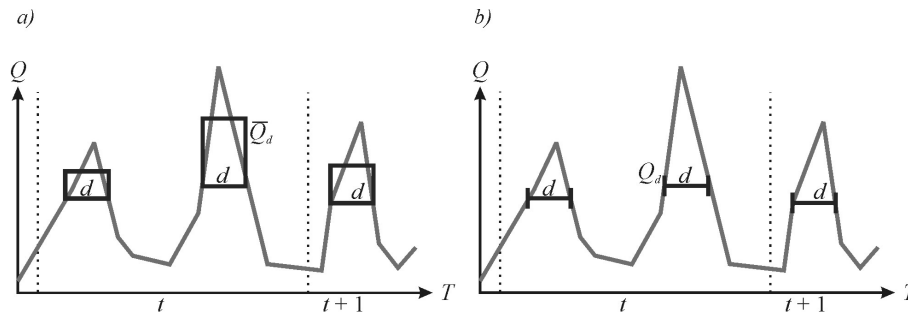


Fig. 3. Definition of the random variables in the QdF models: **(a)** the mean maximum d -days flow, **(b)** the annual maximum flow discharge (Q_d) continuously exceeded during the period d .

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



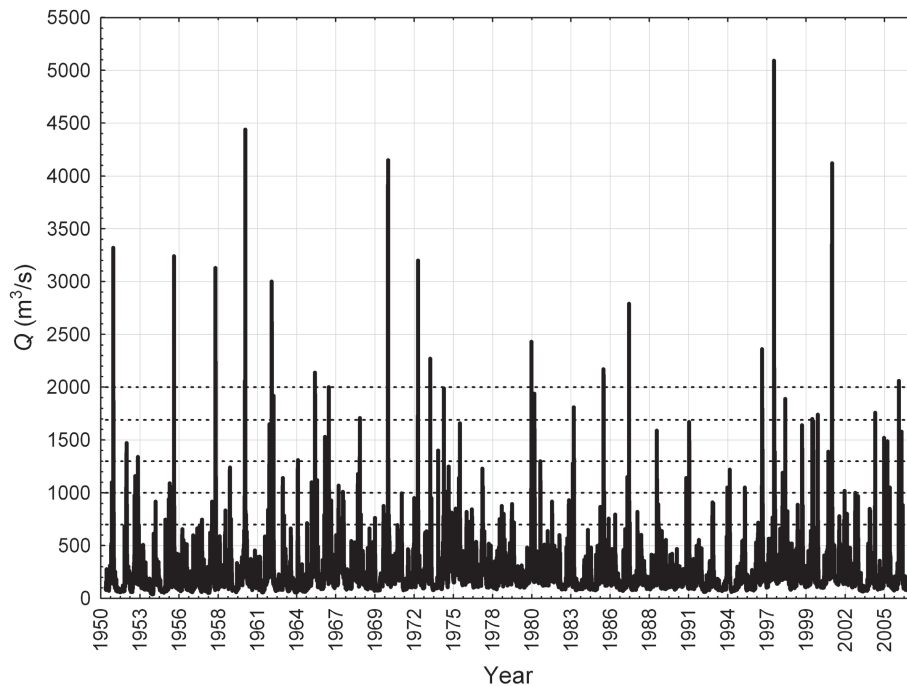


Fig. 4. Hydrograph of the daily flows at the Szczucin gauging station. Horizontal dashed lines reflect the Q_{T_r} values used in this study.

Inundation risk for embanked rivers

W. G. Strupczewski et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



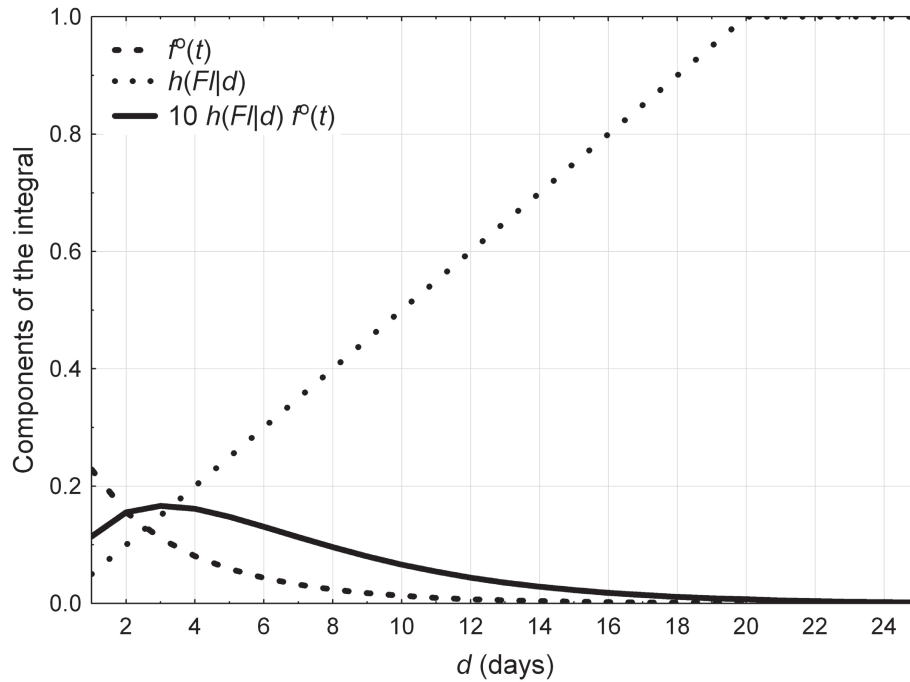


Fig. 5. Components of the integral Eq. (34).

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



**Inundation risk for
embanked rivers**

W. G. Strupczewski et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

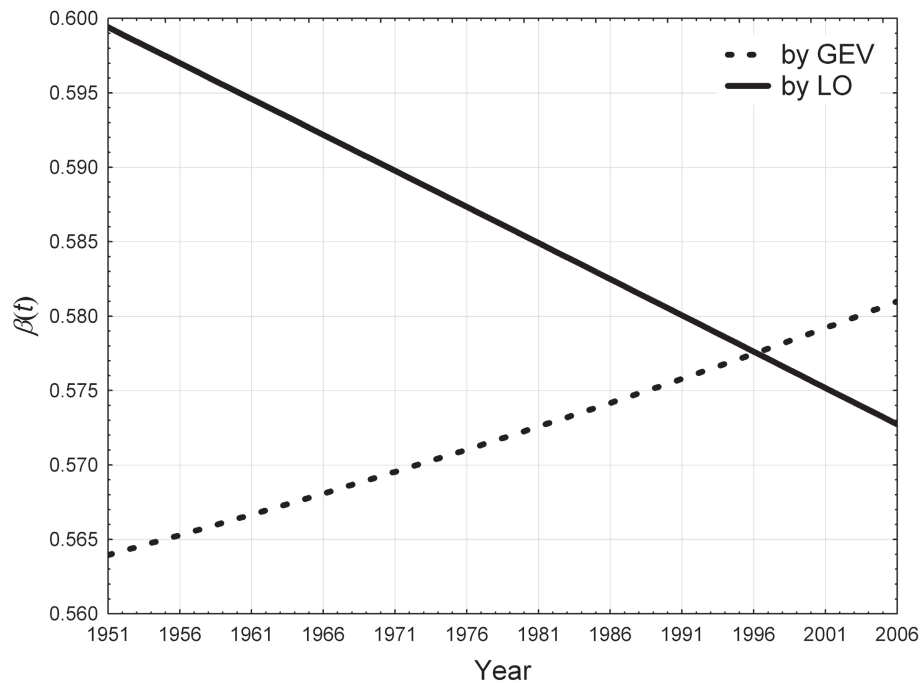
Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Fig. 6.** Non-stationary $\beta(t)$ by two approaches.