

Assessment of extreme discharges of the Vltava River in Prague

M. Holický, K. Jung & M. Sýkora

Klokner Institute, Czech Technical University in Prague, Czech Republic

Abstract

Damage of construction works due to flooding in 1997 and 2002 in the Czech Republic initiated investigations of structural failures and reassessment of available data for discharge extremes. In this study, hydrological data for 166 annual maximum discharges of the Vltava River in Prague since 1827 are analysed using various statistical methods. Moment characteristics of the measurements – the mean, standard deviation and skewness – are estimated and the enhancing effect of an exceptional observation in 2002 is detected. The annual maxima are described by two- or three-parameter lognormal distributions and the extreme value distributions of the type I and II. Standard statistical Kolmogorov and chi-square tests are applied to assess goodness of fit of the theoretical models. It appears that a two-parameter lognormal distribution may be the most suitable theoretical model. Assuming this distribution, extreme discharges corresponding to characteristic and design values are estimated. It is shown that the partial safety factor estimated from the measurements significantly differs from the recommended value of 1.5. The discharge in 2002 corresponds to an exceptionally long return period. It is concluded that statistical methods provide a valuable background for evaluation and prediction of discharges. However, the presented analysis should be further improved to include non-statistical aspects that influence discharges such as the effects of water management and deforestation.

Keywords: discharge, probabilistic assessment, extremes, failure, statistical methods.



1 Introduction

A number of structures in the Czech Republic were affected by the flooding in July 1997 in Moravia and in August 2002 in Bohemia. In particular damage and destruction caused to structures in the historic city of Prague in 2002 was on an unprecedented scale.

Main observed causes of structural damage have been subdivided into geotechnical and structural aspects. The geotechnical causes include:

- Insufficient foundation (depth, width),
- Underground transport of sediments and man-made ground (propagation of caverns),
- Increased earth pressure due to elevated underground water.

The major structural causes cover:

- Insufficient structural robustness (no ring beams as indicated in Figure 1),
- Use of inadequate construction materials (unfired masonry units),
- Material property changes caused by moisture (volume, strength).



Figure 1: Failure of a structure with insufficient robustness.

Water levels recorded in Prague and its surroundings during the flooding seem to be exceptionally high. However, was the flooding really so exceptional and unpredictable? What was the actual return period corresponding to the measured amount of water?

Annual maximum discharges of the Vltava River in Prague are assessed to answer these questions. Available measurements are analysed using various statistical methods with a particular focus on influence of the measurement Q_{2002} . Characteristic and design values of discharges are then determined on the basis of extreme discharges corresponding to specified probabilities. Partial safety

factors are then derived as the ratio of the design value over the characteristic value, EN 1990 [1]. Finally a return period corresponding to the discharge Q_{2002} is estimated.

2 Statistical evaluation of annual maximum discharges

Annual maximum discharges Q_i of the Vltava River in Prague recorded by the Czech Hydrometeorological Institute since 1827 are analysed using basic statistical methods provided by Ang and Tang [2]. Statistical characteristics of the discharges are initially estimated by the classical method of moments for which prior information on the type of an underlying distribution is not needed. The resulting characteristics are given in Table 1 for the samples without and with the observation Q_{2002} .

Table 1: Sample characteristics of the annual maxima in m^3/s (sample size $n = 165$ or 166).

Sample characteristic	Formula used in the analysis	Without Q_{2002}	With Q_{2002}	Enh. factor
Mean	$m = \frac{1}{n} \sum_{i=1}^n Q_i$	1197	$\frac{122}{1}$	1.02
Standard deviation	$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Q_i - m)^2}$	787	846	1.07
CoV	$v = \frac{s}{m}$	0.66	0.69	1.05
Coefficient of skewness	$w = \frac{n}{(n-1)(n-2)s^3} \sum_{i=1}^n (Q_i - m)^3$	1.43	1.74	1.22

It appears that the sample mean, standard deviation and coefficient of variation are influenced by the discharge Q_{2002} rather insignificantly (the enhancing factor varies from 1.02 up to 1.07). However, the coefficient of skewness seems to be considerably affected by Q_{2002} (the enhancing factor is 1.22).

3 Probabilistic distributions

The characteristics provided in Table 1 indicate that the annual maxima might be well described by a two-parameter lognormal distribution having the lower bound at the origin (LN0) or more universal three-parameter lognormal distribution (LN) having the lower bound (for a positive skewness) generally different from zero. Other possible theoretical models are extreme value distributions: the type II called also the Fréchet distribution (F) or type I, a popular Gumbel distribution (G) with the constant skewness of 1.14.



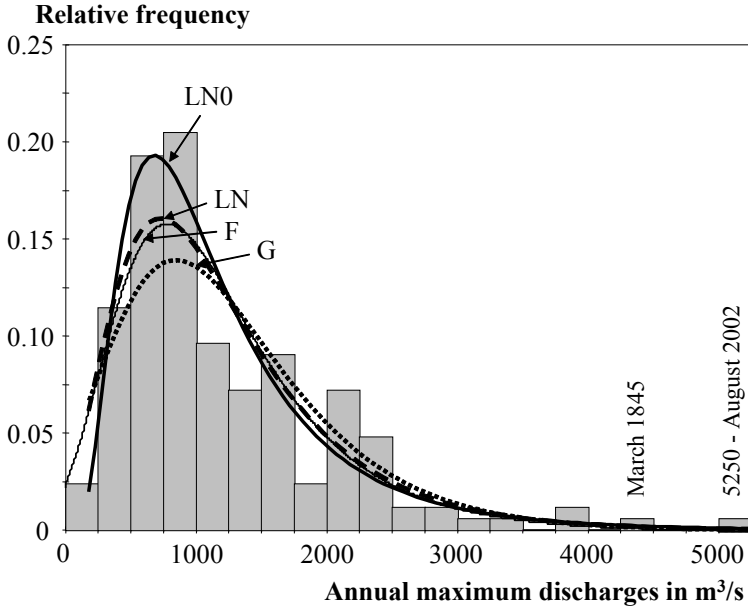


Figure 2: Histogram of the annual maxima and the selected probabilistic distributions.

Probability density functions of the considered theoretical models and a histogram of the analysed measurements are shown in Figure 2. It follows that the lognormal distribution LN0 fits the investigated sample very well. To compare goodness of fit of the considered distributions, Kolmogorov-Smirnov and χ^2 -tests are applied. A hypothesis that a theoretical distribution fits well the sample distribution should be accepted under the condition:

$$K_r = K_0 / K_p \leq 1 \quad (\chi_r^2 = \chi_0^2 / \chi_p^2 \leq 1) \tag{1}$$

Otherwise the hypothesis should be rejected. In eqn. (1) K_0 denotes a test value; K_p critical value and K_r relative test value of the Kolmogorov-Smirnov test. Analogous symbols are used for the chi-square test.

Relative test values are listed in Table 2 for the samples without and with the discharge Q_{2002} .

Table 2: Results of the Kolmogorov-Smirnov and chi-square tests.

Probabilistic distribution	Without Q_{2002}		With Q_{2002}	
	K_r	χ_r^2	K_r	χ_r^2
Lognormal distribution LN0	0.53	1.11	0.49	1.01
Lognormal distribution LN	0.73	1.26	0.65	1.16
Fréchet distribution F	0.82	1.35	0.73	1.30
Gumbel distribution G	0.85	1.43	0.86	1.63



It follows that all the applied distributions meet the condition (1) in accordance with the Kolmogorov-Smirnov test. However, the chi-square test indicates that the measured frequencies significantly differ from theoretical values for all the considered distributions. It appears that the lognormal distribution LN0 is the most suitable model. Less favourable test results are observed for the three-parameter lognormal LN and Fréchet distribution F, and the worst test results are obtained for the Gumbel distribution (the fixed skewness of 1.14 may be rather low). If the discharge Q_{2002} is involved, the tests provide more favourable results for all the distributions, except for the Gumbel distribution.

It should be noted that the test results are indicative only. Suitable models should not be solely selected on the basis of the statistical tests only, but also taking into account experience with discharges measured at other localities. Experience of the Czech Hydrometeorological Institute indicates that the lognormal distribution LN0 could be a suitable model. Therefore, this distribution is further considered in estimation of extreme discharges.

4 Parameter estimation

The method of moments applied in Section 2 to estimate the sample characteristics is often considered to be rather inefficient. Assuming that the underlying distribution of the sample is the lognormal distribution LN0, the sample characteristics can be improved by the maximum-likelihood method, which is considered as the most efficient method for parameter estimation, particularly for large samples. The maximum-likelihood estimators \hat{q} of unknown parameters θ of the distribution (here mean m and standard deviation s) are obtained maximizing the logarithm of a likelihood function:

$$\max_{\mathbf{q}} \ln[L(\mathbf{q}|\mathbf{Q})] \rightarrow \hat{\mathbf{q}} \tag{2}$$

where $\mathbf{Q} = (Q_1, \dots, Q_n)$ is the sample, \mathbf{q} realization of the vector of the parameters θ and $L(\mathbf{q}|\mathbf{Q})$ is the likelihood function:

$$L(\mathbf{q}|\mathbf{Q}) = \prod_i^n f(Q_i|\mathbf{q}) \tag{3}$$

where $f(\bullet)$ denotes the probability density function of the underlying distribution. Here the non-linear conjugate-gradient method implemented in the software package Mathcad[®] is applied. Comparison of the distribution parameters estimated by the method of moments and the maximum-likelihood method is indicated in Table 3.

Table 3: Estimated parameters in m³/s.

Method	Without Q_{2002}		With Q_{2002}	
	m	s	m	s
Moments	1200	790	1220	850
Maximum-likelihood	1210	870	1230	910



It appears that the estimates of the mean are nearly independent of the applied method (differences about 1%). However, the standard deviations estimated by the maximum-likelihood method are systematically greater than those obtained by the method of moments (differences about 10%).

5 Estimation of extreme values

Upper fractiles Q_p of the lognormal distribution LN0 are further estimated using the classical coverage method for the given confidence level γ , see e.g. ISO 12491 [3]:

$$P(Q_{p,cov} > Q_p) = \gamma \tag{4}$$

In accordance with EN 1990 [1], the characteristic value Q_k is obtained as the 0.98 fractile of annual maxima while the design value Q_d is the fractile of the life-time maxima corresponding to the probability:

$$p_d = 1 - \Phi(\alpha_E \times \beta) = 1 - \Phi(-0.7 \times 3.8) = 1 - 0.0039 \tag{5}$$

where Φ denotes the cumulative distribution function of the standardised normal distribution, α_E is the FORM sensitivity factor (considering the recommended value of -0.7 for the leading action) and β is the reliability index equal to 3.8 for the reference period of 50 years.

Assuming statistical independence of the annual maxima, the design value is estimated as follows:

$$Q_d = F^{-1}\left(\sqrt[50]{p_d} \mid \hat{\mathbf{q}}\right) \tag{6}$$

where $F^{-1}(\bullet)$ denotes the inverse cumulative distribution function of the underlying distribution of the annual maxima.

Partial safety factor γ_Q for unfavourable effects of a variable action is consequently obtained as the ratio Q_d / Q_k . Estimated extreme discharges and partial factors are summarized in Table 4.

Table 4: Estimated extreme discharges in m³/s.

Method	Without Q_{2002}		With Q_{2002}		Partial factor γ_Q
	Expect.	$\gamma = 0.75$	Expect.	$\gamma = 0.75$	
	Characteristic value (p_k)				
Moments	3430	3640	3630	3860	
Maximum-likelihood	3710	3950	3830	4090	
Method	Design value (p_d)				Partial factor γ_Q
Moments	9650	10640	10700	11840	2.82 – 3.07
Maximum-likelihood	11360	12620	11960	13310	3.07 – 3.26



It is indicated that the extreme values predicted from the available data including the discharge Q_{2002} are greater than those estimated without this discharge (by about 9% for the method of moments and 4% for the maximum-likelihood method). It also appears that the extreme discharges predicted by the maximum-likelihood method are greater than those obtained by the method of moments (by about 6–9% for the characteristic value and 12–19% for the design value). Furthermore, the upper fractiles estimated considering the commonly accepted 0.75 confidence level are greater than the expected upper fractiles (by about 6% in case of the characteristic values and 11% in case of the design values). The partial safety factor $\gamma_Q \approx 3.0$ derived from the data seems to be significantly greater than the recommended value 1.5.

6 Return period of the discharge Q_{2002}

Expected return periods T corresponding to the discharge Q_{2002} are derived using the relationship:

$$T = 1 / [1 - F(Q_{2002} | \hat{q})] \quad (7)$$

where $F(\bullet)$ denotes the cumulative distribution function of the underlying distribution. Expected return periods are listed in Table 5.

Table 5: Expected return periods in years.

Method	Without Q_{2002}	With Q_{2002}
Moments	350	240
Maximum-likelihood	210	180

It appears that the return period is considerably affected by the fact whether the discharge Q_{2002} is taken into account or not. In addition the estimates based on the maximum-likelihood method are significantly lower than those obtained using the method of moments. Considering the data without Q_{2002} and the maximum-likelihood method, the observed discharge $Q_{2002} = 5250 \text{ m}^3/\text{s}$ corresponds to the exceptionally long return period of 210 years. Obviously the discharge Q_{2002} could have been hardly expected. Note that estimates of the return period may also enormously vary with a type of the applied distribution as indicated by Holicky and Sykora [4].

It is emphasized that the presented analysis is based on statistical methods only. More detailed analysis should also consider non-statistical influences that may have evolved during the period covered by the measurements (since 1827). In particular discharges may be strongly dependent on a river management including modifications of depth, width and roughness of a river channel and removal of vegetation. Effects of deforestation and other man-made interventions in environment should be also taken into account.

7 Conclusions

Investigation of structural failures due to flooding in Moravia (1997) and Bohemia (2002) indicates that main causes of structural damage may be



subdivided into geotechnical and structural reasons. Statistical analysis of available data for annual discharge maxima shows that:

- Discharges may be well described by a two-parameter lognormal distribution LN0.
- The characteristic and design discharges predicted using data including the discharge in 2002 are greater than those estimated without considering the discharge in 2002 (by about 5%).
- Extreme discharges predicted by the maximum-likelihood method are greater than those by the method of moments (by about 10%).
- The recommended value of the partial safety factor $\gamma_Q = 1.5$ is considerably lower than the value derived from the available data ($\gamma_Q \approx 3.0$).
- The discharge observed in 2002 corresponds to an exceptionally long return period and, therefore, could have been hardly expected.

It appears that statistical methods provide a valuable background for evaluation and prediction of discharges. It is further noted that the presented results are indicative only since purely statistical methods are used in the assessment. Effects of water management and man-made interventions in environment should be also taken into account.

Acknowledgements

This study has been conducted at the Klokner Institute, Czech Technical University in Prague, with financial support of the project GACR 103/06/1521 Reliability and risk assessment of structures in extreme conditions.

References

- [1] EN 1990 Eurocode: *Basis of structural design*, CEN, 2002.
- [2] Ang, A.H.-S. & Tang, W.H., *Probabilistic concepts in engineering planning and design*, John Wiley and Sons: New York, 1975.
- [3] ISO 12491 *Statistical methods for quality control of building materials and components*, ISO, 1997.
- [4] Holický, M. & Sykora, M, Probabilistic Evaluation and Prediction of Discharges on the Vltava River in Prague, *CTU Reports (Vol. 8, No. 3), Proc 3rd Czech/Slovak Symposium Theoretical and Experimental Research in Structural Engineering*, CTU Publishing House: Prague, p. 47–52, 2004.

