Variance of design floods from the single-site FEH method: the GLO distribution

Thomas Rodding Kjeldsen, MSc, PhD

Senior Lecturer

Department of Architecture and Civil Engineering, University of Bath, BA2 7AY, Bath, United Kingdom

https://orcid.org/0000-0001-9423-5203

Word Count: 1885

Abstract

An easy-to-use equation is presented for calculating the variance of a design flood estimated using a generalised Logistic (GLO) distribution with model parameters estimated using single site analysis as described in the Flood Estimation Handbook. The equation is applicable for return periods in the range $T = 2$ to $T = 1000$ years and for annual maximum flood series with L-skewness in the interval $-0.45 \leq t_3 \leq 0.45$ which is considered representative of most UK flood data and practical uses.

List of symbols

α	GLO scale parameter
β	GLO normalised scale parameter
κ	GLO shape parameter
ξ	GLO location parameter
AMS	Annual maximum series
FEH	Flood Estimation Handbook
GLO	Generalised Logistic distribution
n	Record length
QMED	Median annual maximum flood (m^3/s)
Q_T	T-year design flood (m^3/s)
T	Return period (years)
t_2	*L*-CV
t_3	*L*-skewness
y_L	Logistic reduced variate, $\ln(T-1)$

Introduction

Flood frequency analysis at a gauged site typically involves fitting a statistical distribution, F, to an annual maximum series (AMS) of peak flow events obtained from a gauging station. The design flood Q_T of the required return period, T, is derived as the $(1 - 1/T)$ quantile of the chosen distribution, i.e.

$$Q_T = F^{-1}\left(1 - \frac{1}{T}\right)$$

In the United Kingdom, guidelines for flood frequency analysis and design flood estimation were published by the Institute of Hydrology (1999) as part of the Flood Estimation Handbook (FEH), and later updated by the Environment Agency (2008). The FEH recommends the three-parameter generalised logistic (GLO) distribution for which the T-year design flood is derived as

$$Q_T = \xi\left[1 + \frac{\beta}{\kappa}\left(1 - (T-1)^{-\kappa}\right)\right]$$

where ξ, β, κ are the location, normalised scale and shape parameter, respectively, which all need to be estimated from the available AMS. The FEH recommends that the location parameter ξ is estimated as the median of the AMS, and is denoted QMED. The normalised scale (β) and the shape (κ) parameters are estimated using the method of *L*-moments as described in the FEH and summarised in the Appendix here.

Stedinger et al. (1993) summarised procedures for estimating the variance of *T*-year events for a range of distributions commonly used in flood frequency analysis, e.g. the Gumbel, the generalised extreme value (GEV), and the log-Pearson type 3 (LP3) distributions. However, no such guidance was published as part of the FEH or its updates. For a single site analysis based on fitting a GLO distribution as described in the FEH Vol. 3, the variance of the *T*-year event will depend on: the return period T, the record-length n, and the three GLO model parameters. Kjeldsen and Jones (2004) derived an approximate analytical solution for the case of a single site analysis based on the GLO distribution, resulting in a complex set of equations involving specialised mathematical function, and thus unlikely to be useful in practice. In addition, the approximate solution is valid for very large sample sizes (>1000 years) and can perform poorly for data series of a duration typically encountered in flood hydrology (i.e. between 30-50 years) and with a high degree of skewness. To facilitate calculation of uncertainty with the FEH methodology, this note contains a set of simple polynomial approximations that will allow the standard deviation of the *T*-year design event to be estimated for any return period between 2 to 1000 years without the need to invoke specialised mathematical functions.

Variance of design flood from a GLO distribution

Based on the theory of asymptotic variance of *L*-moments outlined by Hosking (1986) and results by Kjeldsen and Jones (2004) the variance of the *T*-year event, as derived from a GLO distribution, using the FEH procedure can be approximated as:

$$Var(\hat{Q}_T) = \frac{(QMED \times \beta)^2}{n}\ g(T,\kappa) \qquad (1)$$

where n is the record-length, QMED, β and κ are GLO parameters, T is the return period, and g is a complex function involving specialised mathematical functions. Following the example by Stedinger et al. (1993), a simple-to use empirical approximation to Eq.(1) can be defined as:

$$Var(\hat{Q}_T) = \frac{(QMED \times \beta)^2}{n} exp[a_0 + a_1 y_L + a_2 y_L^2 + a_3 y_L^3] \qquad (2)$$

where a_0, a_1, a_2, and a_3 are constants that will be defined based on Monte Carlo simulations, and y_L is the logistic reduced variate recommended by the FEH for use with the GLO distribution and defined as:

$$y_L = \ln(T-1) \qquad (3)$$

An extensive set of Monte Carlo experiments were conducted to produce a set of parameter values a_0, a_1, a_2, and a_3 with practical application. Three practical considerations were needed, and reflect trade-offs between operational usefulness and precision.

***L*-skewness, t_3:** Figure 1 shows a histogram of *L*-skewness created using 811 AMS with a minimum of 20 years of record from the National River Flow Archive (NRFA) peak flow dataset. The majority (96%) of data series have *L*-skewness values in the range -0.45 to 0.45.

Return periods, T: It was decided that the range of return period range from $T = 2$ to $T = 1000$ years. It is hoped that this will cover the majority of return periods required in practice.

Record lengths, n: A nominal record-length of $n = 50$ years was chosen as the basis for the Monte Carlo simulations. This is compatible with method for calculating the variance of a GEV distribution presented by Stedinger et al. (1993) based on a record-length of $n = 40$ years.

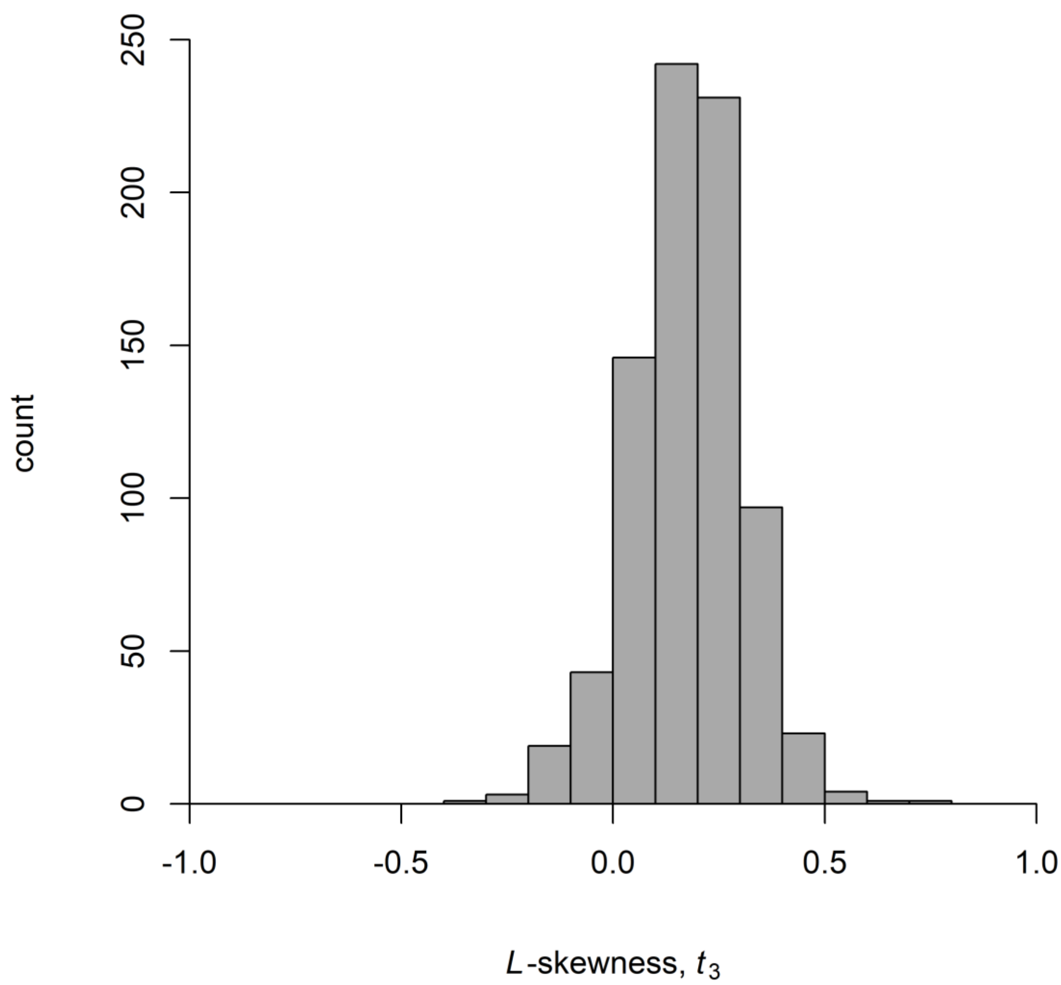

Figure 1: Histogram of *L*-skewness derived from 811 AMS of peak flow from UK catchments with a minimum record-length of 20 years.

Monte Carlo simulations

A set of Monte Carlo simulations were conducted to enable the estimation of the parameters a_0, a_1, a_2, and a_3.

1. Define a GLO distribution. As the normalised variance derived from in Eq. (1)

$$\frac{n\,Var(Q_T)}{(QMED \times \beta)^2} = g(T, \kappa) \tag{4}$$

depends only on the shape parameter and the return period, and thus values of the location and scale parameter were chosen by fixing the first *L*-moment and the second order *L*-moment ratio at $l_1 = 1.0$ and $L - CV = 0.20$.

2. Set the shape parameter to a value between $-0.45 \leq \kappa \leq 0.45$. Note that for the GLO distribution the shape parameter is given by the *L*-skewness directly as $\kappa = -t_3$ (see Appendix).

3. Choose a return period from T = 2 to 1000 years.

4. Simulate 10,000 AMS of peak flow of length n = 50 years with the specifications outlined in steps 1 to 3. For each AMS, fit a GLO distribution as outlined in the FEH and calculate the corresponding *T*-year design events.

5. Calculate the variance of the 10,000 design events and normalise according to Eq. (4).

6. Repeat step 4 to 5 for each return period (T = 2 to T = 1000 years).

7. Repeat step 3 - 6 for each specified value of the shape parameter κ.

Eq. (2) was fitted to each set of results representing a specific value of *L*-skewness. The fitting was done using ordinary least square and achieved an R^2 value in excess of 0.999 for all transects. The resulting values of the model parameters a_0, a_1, a_2, and a_3 are shown in Table 1.

Table 1: Parameter values for use in Eq. (2) to derive variance of *T*-year design event.

	L-skewness, t_3									
	-0.45	**-0.35**	**-0.25**	**-0.15**	**-0.05**	**0.05**	**0.15**	**0.25**	**0.35**	**0.45**
a_0	1.2888	1.3441	1.3668	1.3805	1.3757	1.3348	1.2970	1.2819	1.2903	1.2831
a_1	-0.4104	-0.4287	-0.3687	-0.2282	-0.0464	0.1785	0.4152	0.6006	0.7375	0.9136
a_2	0.2114	0.2172	0.2082	0.1885	0.1760	0.1684	0.1607	0.1743	0.1928	0.1901
a_3	-0.0181	-0.0177	-0.0164	-0.0146	-0.0139	-0.0140	-0.0137	-0.0150	-0.0161	-0.0155

Example of application:

A continuous record consisting of AMS of 49 peak flow events is available from the gauging station 53018 River Avon @ Bathford. The median, the sample *L*-moment ratios and the resulting GLO parameters, as derived using the FEH procedure, are shown in Table 2.

Table 2: Median, *L*-CV, *L*-skewness and GLO parameters as estimated for 53018 Avon @ Bathford.

Median (m^3/s)	*L*-CV	*L*-skewness	QMED	Scale β	Shape κ
167.2	0.132	0.084	167.2	0.133	0.084

The values of the parameters a_0, a_1, a_2, and a_3 are obtained through simple linear interpolation between the two columns in Table 1 on either side of the sample value of $t_3 = 0.084$. The resulting parameter values are shown in Table 3

Table 3: Linear interpolation of model parameters from Table 1.

	$t_3 = 0.05$	$t_3 = 0.084$ interpolated	$t_3 = 0.15$
$\boldsymbol{a_0}$	1.3348	1.3220	1.2970
$\boldsymbol{a_1}$	0.1785	0.2590	0.4152
$\boldsymbol{a_2}$	0.1684	0.1657	0.1607
$\boldsymbol{a_3}$	-0.0140	-0.0139	-0.0137

For each return period between *T*=2 and T=1000, the standard deviation of the design flood can now be estimated by combining Eq. (2) with the parameter values reported in Table 2 and 3 as:

$$\widehat{Var}\left(\hat{Q}_T\right) = \frac{(167.2 \times 0.133)^2}{49} exp[1.3220 + 0.2590 y_L + 0.1657 y_L^2 - 0.0139 y_L^3] \qquad (5)$$

Finally, assuming the *T*-year design flood is approximately normal distributed, the approximate 95% confidence interval can be plotted alongside the flood frequency curve as

$$\left[Q_T - 2\sqrt{Var(\hat{Q}_T)}\,; Q_T + 2\sqrt{Var(\hat{Q}_T)}\right] \tag{6}$$

The results are visualised in Figure 2. A second set of 95% confidence intervals were derived by generating 10000 AMS from a GLO distribution with the parameters specified in Table 3. As can be observed, a good agreement between the equation developed in this note and the Monte Carlo simulations has been achieved. Visual inspection across all AMS in the NRFA peak flow database confirmed this result. However, it should be noted that for large return periods it is possible for the confidence intervals to include negative flow value, which lacks a physical interpretation.

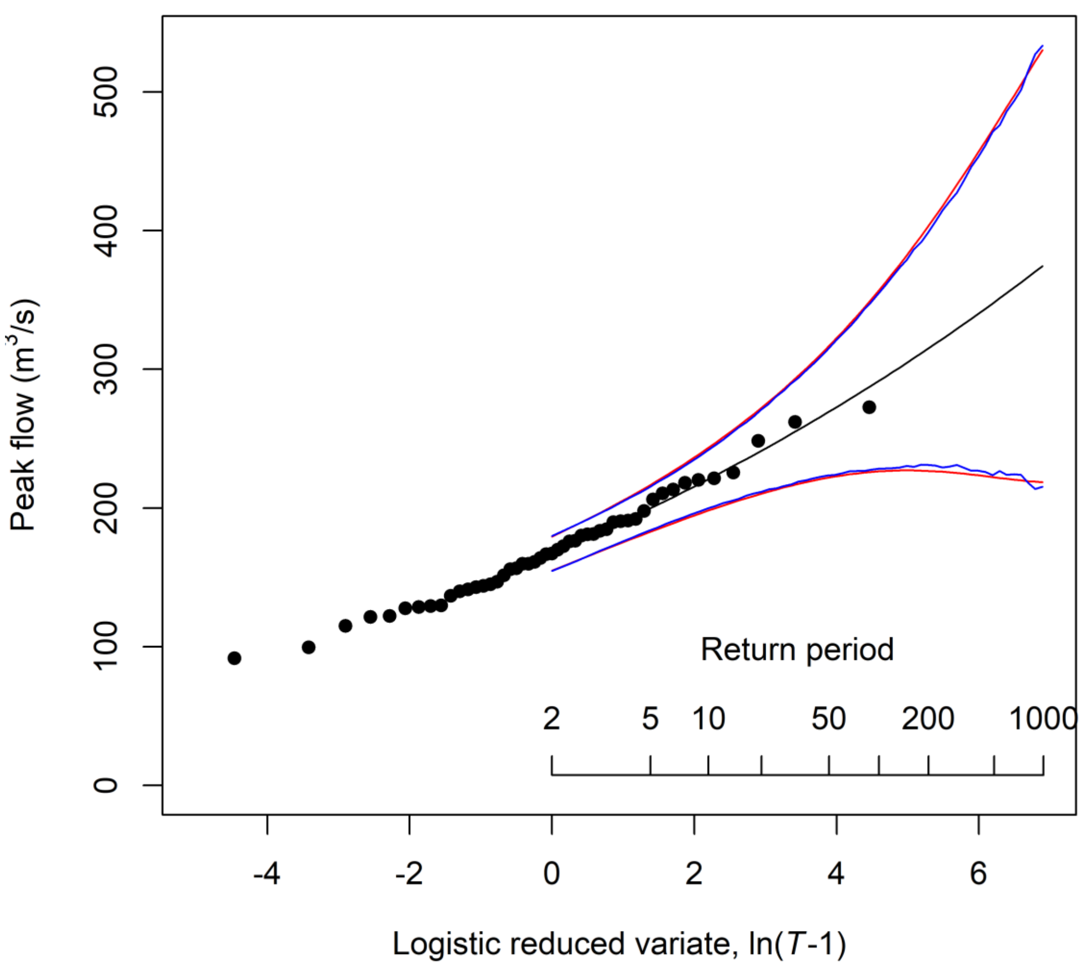

Figure 2: GLO flood frequency curve (solid line) fitted to AMS from 53018 Avon @ Bathford. 95% confidence interval calculated using Eq. (5) shown in red, and the corresponding 95% confidence interval derived from 10000 Monte Carlo generated AMS shown in blue.

Discussion and conclusion

The results presented in this paper allow estimates of the variance of design floods obtained using a Generalised Logistic (GLO) distribution for return periods in the range between T = 2 – 1000 years. The proposed method is applicable only when the three parameters of the GLO distribution are fitted to a single AMS series using the method of L-moments as presented in the Flood Estimation Handbook (FEH), i.e. FEH single site analysis. Caution should be exerted if the method is extrapolated to consider the variance of design events with a return period in excess of T=1000 years, as the lower bound, in particular, might exhibit counter intuitive behaviour such as inclusion of negative flow values. Finally, the method presented here could be extended to consider the variance of design flood events derived from other distributions commonly used in flood frequency analysis such as the Generalised Extreme Value (GEV), generalised Pareto or the Kappa distribution.

Acknowledgements: The annual maximum series of peak flow used in this study were made available by the National River Flow Archive (NRFA)

References

Hosking, J. R. M., 1986. The theory of probability weighted moments. Res. Rep. RC12210, Res. Div., IBM, Yorktown Heights, N. Y.

Institute of Hydrology (1999) Flood Estimation Handbook, 5 Volumes, Institute of Hydrology Wallingford, UK.

Kjeldsen, T.R. and Jones, D.A., 2004. Sampling variance of flood quantiles from the generalised logistic distribution estimated using the method of *L*-moments. *Hydrology and Earth System Sciences*, 8(2), pp.183-190.

Environment Agency, 2008. Improving the FEH statistical procedures for flood frequency estimation: *Defra/Environment Agency Science Report: SC050050*.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), Frequency analysis of extreme events. *In Handbook of Hydrology*, edited by D. Maidment, chap. 18, McGraw-Hill, New York.

Appendix: FEH single site analysis using the GLO distribution

The *T*-year design flood from a GLO distribution is estimated from the quantile function as

$$Q_T = \xi \left[1 + \frac{\beta}{\kappa}\left(1 - (T-1)^{\kappa}\right)\right]$$

Consider an AMS series containing n observations: $q_1, q_2, \cdots, q_n$ and the corresponding *L*-CV (t_2) and *L*-skewness (t_3).

The location, normalised scale and shape parameters are estimated according the FEH (Institute of Hydrology, 1999) as:

$$\xi = median[q_1, q_2, \cdots, q_n]$$

$$\beta = \frac{t_2 \kappa \sin(\pi\kappa)}{\pi\kappa(\kappa + t_2) - t_2 \sin(\pi\kappa)}$$

$$\kappa = -t_3$$