Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Technical Note: Design flood under hydrological uncertainty

Anna Botto[1*], Daniele Ganora[1], Pierluigi Claps[1], and Francesco Laio[1]

[1]Department of Environment, Land and Infrastructure Engineering - Politecnico di Torino - Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy

[*]Now at the Department of Civil, Environmental and Architectural engineering - Università di Padova - Via Marzolo, 9 - Padova, Italy

Correspondence to: Daniele Ganora (daniele.ganora@polito.it)

Abstract. Planning and verification of hydraulic infrastructures demands for a design estimate of hydrologic variables, usually provided by frequency analysis, neglecting hydrologic uncertainty. However, when hydrologic uncertainty is accounted for, the design flood value is no longer a deterministic value, but should be treated as a random variable itself. As a consequence, the design flood is no longer univocally defined, making the design process undetermined.

Botto et al. (2014), with the development of the Uncertainty Compliant Design Flood Estimation (UNCODE) procedure, have shown that it is possible to fix the ambiguity in the selection of the design flood under uncertainty by considering an additional constraint based on a cost-benefit criterion. This paper contributes with an easy-to-use framework to implement the UNCODE procedure without resorting to numerical computation, but using a correction coefficient that modifies the standard (i.e., uncertainty-free) design value on the basis of sample length and return period only. The procedure is robust and parsimonious, as it does not require additional parameters with respect to the traditional uncertainty-free analysis.

Simple equations to compute the correction term to the standard estimate are provided for a number of probability distributions commonly used to represent the flood frequency curve. This new design tool provides a robust way to manage the hydrologic uncertainty and to go beyond the use of traditional safety factors. With all the other parameters being equal, an increase of the sample length reduces the correction factor, and thus the construction costs, still keeping the same safety level. This improvement is shown to be more effective when short samples are extended.

1 Introduction

The flood frequency curve is commonly used to derive the design flood as the quantile Q_T corresponding to a fixed return period T. The design value Q_T is defined as a single value when the frequency distribution and its parameters are known without uncertainty. When uncertainty in the parameters or in the probabilistic model is accounted for, this propagates to the quantile; this means that for the same return period T the quantile is no longer a single value, but should be treated as a random variable itself. As a consequence, the design flood is no longer univocally defined. However, the design of an hydraulic infrastructure demands for a single design value to be selected. A gap therefore exists between theory and practice. Quantitative methods to measure the uncertainty associated to the quantiles of the flood frequency curve (e.g., through their variance or probability distribution) have been proposed (e.g., Cameron et al., 2000; De Michele and Rosso, 2001; Brath et al.,

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

Hydrology and Earth System Sciences Open Access EGU Discussions

2006; Blazkova and Beven, 2009; Laio et al., 2011; Liang et al., 2012; Viglione et al., 2013), but very few suggestions are provided about how to obtain a single design value from the probability distribution of possible design values.

Botto et al. (2014), with the development of the Uncertainty Compliant Design Flood Estimation (UNCODE) procedure, have shown that it is possible to select meaningful flood quantiles from their distribution by considering an additional constraint based on a cost-benefit criterion. Hence, the output is a unique design flood value Q_T^*. For theoretical and practical aspects of the procedure the reader is referred to the original paper, whereas here we recall only the core concepts of the UNCODE approach.

Botto et al. (2014) primarily demonstrated that the design flood Q_T obtained with the standard flood frequency analysis (without uncertainty) is equivalent to the design flood obtained with a cost-benefit analysis with specific damage and cost curves. Examples of cost-benefit analysis in the hydrologic/hydraulic context can be found in the literature (Bao et al., 1987; Ganoulis, 2003; Jonkman et al., 2004; Tung, 2005), but only a few of them include hydrologic uncertainty (Al-Futaisi and Stedinger, 1999; Su and Tung, 2013). Cost and damage curves obtained by Botto et al. (2014) are piecewise linear functions with slope c and d respectively that, combined with the probability density function p of the flood values Q, give the total cost function (i.e., actual costs plus damages):

$$C_{TOT} = c \cdot Q^* + \int_{Q^*}^{\infty} d \cdot (Q - Q^*) \cdot p(Q|\boldsymbol{\Theta})\,\mathrm{d}Q \tag{1}$$

where Q^* is the generic design flood value and $\boldsymbol{\Theta}$ is the vector of parameters of the probability distribution, which depends on the hydrologic characteristic of the site. Parameters c and d are also generally site-specific, being influenced by topography and land use among others. Taking the derivative of C_{TOT} with respect to Q^* and setting it to 0 gives the optimal design flood value of the (uncertainty-free) cost-benefit framework, and leads also to the equivalence

$$\frac{d}{c} = \frac{1}{1 - P(Q^*|\boldsymbol{\Theta})} = T, \tag{2}$$

where $P(\cdot)$ is the cumulative distribution function of the flood values. Equation (2) links the standard flood frequency analysis to the cost-benefit approach.

In uncertain condition $\boldsymbol{\Theta}$ becomes a random vector; hence, hydrologic uncertainty should be included in the cost benefit analysis by compounding C_{TOT} over all the possible values of $\boldsymbol{\Theta}$. In mathematical terms, the cost-benefit framework with uncertainty is summarized by the equation

$$Q_T^* = \arg\min_{Q^*} \left[\int_{\boldsymbol{\Theta}} C_{TOT}(Q^*|c, d, p(\boldsymbol{\Theta})) \cdot h(\boldsymbol{\Theta})\,\mathrm{d}\boldsymbol{\Theta} \right], \tag{3}$$

where $h(\boldsymbol{\Theta})$ is the joint pdf of the parameters of the flood frequency curve.

Equation (3) is valid in general upon specification of c and d, which are usually unavailable. However, the inherent equivalence between the cost-benefit and the quantile-based approaches defined by Eq. (2) reduces the degrees of freedom of the cost-benefit framework, as c and d are not independent, but are related through the known value of the return period T. The

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

Hydrology and Earth System Sciences
Open Access
Discussions

remaining free parameter can be shown to affect only the magnitude of the integral in Eq. (3), but not the position of its minimum. As a consequence, the UNCODE framework does not add any further parameter with respect to the standard design flood procedure, but it allows one to frame the uncertainty analysis into a cost-benefit framework.

We have shown in Botto et al. (2014) that the UNCODE design flood, Q_T^*, is always larger than its corresponding standard value Q_T. However, computation of Q_T^* requires application of a numerical methods. To simplify the UNCODE application, we provide here an approximated, though reliable, method to estimate Q_T^* starting from Q_T. Other than a useful practical tool for design purposes, the analysis reported in this note also provides a method to quantify the "value" of newly available hydrological information or the effect of data scarcity on Q_T^* due to uncertainty.

2 Practical estimation of the UNCODE design flood

To compute the UNCODE design flood Q_T^*, we consider the equation:

$$Q_T^* = (1+y) \cdot Q_T \tag{4}$$

where y is a-non negative coefficient which depends on the probability distribution used to fit the flood frequency curve, and Q_T is the standard design flood.

As noted by Botto et al. (2014), the relative distance between Q_T^* and Q_T increases with the return period (as the quantile uncertainty increases) as well as, for fixed T, with the standard deviation of the probability distribution of Q_T (i.e., with the uncertainty of Q_T). We propose here to model y according to the equation

$$y = 10^{-2} \cdot \exp\left[a_0 + a_1\sqrt{n} + a_2 \ln T\right] \tag{5}$$

where T is the return period and n is the sample length which can be considered as a proxy of the standard deviation of Q_T (n can be computed from at-site records or as an equivalent sample length from the regional estimate of Q_T). Coefficients a_0, a_1 and a_2 are reported in Table 1 for different fitting distributions commonly used in the hydrological practice to compute the design flood (Hosking and Wallis, 1997). Table 1 clearly shows the effect of increasing the sample length n, which reduces the difference between Q_T^* and Q_T due to the negative value of the coefficient a_1.

The coefficients a_0, a_1 and a_2 have been evaluated through an extensive simulation study (details are reported in the Supplementary Material) in which the full UNCODE procedure has been systematically applied to many simulated records, considering a number of different combinations of parent and fitting distribution, sample length (with n from 30 to 100) and return periods (with T from 50 to 1000). Different forms of Eq. (5) have been also tested. For each run, the empirical y value has been recorded. The coefficients a_j have been finally estimated through linear regression on the log-transformed terms of Eq. (5). Table 1 reports some diagnostics of the regressions used to estimate the coefficients.

The reliability of the approximated correction factor y estimated with the regression model has been evaluated by comparing the Q_T^* value obtained through Eq. (4) and (5) with its exact counterpart calculated with the full UNCODE procedure. As a reference, time series listed in Botto et al. (2014, Table 1) with at least 30 years of record length have been analyzed, assuming the LN3 and the GEV as possible fitting distributions and different return periods. Results show a very good agreement between the exact and the approximated Q_T^* values, as reported in Fig. 1, where each panel shows the estimates for all series and all the return periods. Panel a) refers to the LN3, while panel b) to the GEV distribution.

A synthesis of quantitative results is shown in Fig. 2, where the values of y from Eq. (5) have been reported for the studied distributions, based on a set of typical sample length and return period values. As mentioned, a direct comparison of the results between different distributions is not possible, but it is relevant to observe that for all the distributions y evolves in the same way for varying n and T values. In general, the correction factor does not exceed 10% of the standard value Q_T for intermediate return periods (e.g., $T = 200$ years) even for small samples, although a significative variability is associated to the distribution type. It is also 10% around $T = 500$ years with sample length values ($n = 50$) commonly available at many gauged stations. On the other hand, the sample length plays an important role: for example, considering $T = 500$ years, the GEV distribution and varying the sample size, the reduction of the y value is about 0.075 between $n = 30$ and $n = 50$, while it drops to 0.040 between $n = 50$ and $n = 70$.

3 Discussion of the application conditions

The UNCODE approach to flood frequency analysis provides an answer to the ambiguity due to the uncertainty in the quantile estimation. Application of the full UNCODE procedure may be cumbersome and computationally demanding. For a quick estimation of the design value an approximated but reliable framework has been proposed here to easily compute the UNCODE flood starting from the standard design value.

The extensive simulation analysis at the base of this study shows that the coefficients a_j relating the UNCODE value Q_T^* to the standard value Q_T are distribution-dependent. For the most used distributions they have been computed and provided. The choice of the distribution is a problem of model selection and depends on the preliminary flood frequency analysis. The obtained results demonstrate that an increase in the length of relatively short samples has a noticeable impact in terms of reduction of y and of the UNCODE estimate Q_T^*. This implies that, while the infrastructure keeps the same safety level, additional data reduce construction costs as the actual design value is reduced. The mentioned results agree with findings recently obtained by Ganora and Laio (2016) in a study on the relative role of regional and at-site flood frequency modeling approaches, where the value of at-site data has been highlighted and regarded as a reliable way to improve regional predictions, even with short records. Under this perspective, the correction factor can be used as a metric for uncertainty comparison and quantification, thus providing a further tool to combine different modeling approaches, similarly to the applications of Kjeldsen and Jones (2007) and Ganora et al. (2013) who, with different methodologies, have exploited measures of hydrologic uncertainty to merge regional and at-site information. The coefficient y can be considered a measure of the value of data. In fact, with all other parameters being equal, increasing n leads to a reduced y value and, consequently, to a reduced UNCODE design flood Q_T^*. As a consequence, while the design value is still based on the same return period, costs will reduce.

Finally, the correction factor is a new and easy-to-implement design tool which provides a quantitative way to determine the design flood value accounting for hydrologic uncertainty, while keeping the same design hazard level considered in standard uncertainty-free analyses. This is a novel approach when compared to the common engineering practice, which accounts for hydrologic uncertainty by considering, for instance, the hydraulic freeboard. The use of the freeboard is equivalent to increase the design flood value, but without accounting for the size of the system (e.g., the basin area), nor for the hydrologic information available at the section (i.e., observed of equivalent record length used to compute the standard design flood); this approach is thus not tailored to the specific case study. The correction factor represents an advance with respect to the use of "all-encompassing" safety factors and towards a clearer way to manage the different sources of uncertainty in hydrological and hydraulic design.

Acknowledgements. Funding from the ERC Consolidator Grant 2014 "CWASI - Coping with water scarcity in a globalized world" is acknowledged. D. Ganora also acknowledges the RTD Starting Grant from Politecnico di Torino. The work is based on simulated data.

References

Al-Futaisi, A. and Stedinger, J.: Hydrologic and Economic Uncertainties and Flood-Risk Project Design, Journal of Water Resources Planning and Management, 125, 314–324, doi:10.1061/(ASCE)0733-9496(1999)125:6(314), http://dx.doi.org/10.1061/(ASCE)0733-9496(1999)125:6(314), 1999.

Bao, Y., Tung, Y., and Hasfurther, V.: Evaluation of uncertainty in flood magnitude estimator on annual expected damage costs of hydraulic structures, Water Resources Research, 23, 2023–2029, doi:10.1029/WR023i011p02023, http://dx.doi.org/10.1029/WR023i011p02023, 1987.

Blazkova, S. and Beven, K.: A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, Water Resources Research, 45, n/a–n/a, doi:10.1029/2007WR006726, http://dx.doi.org/10.1029/2007WR006726, 2009.

Botto, A., Ganora, D., Laio, F., and Claps, P.: Uncertainty compliant design flood estimation, Water Resources Research, 50, doi:10.1002/2013WR014981, 2014.

Brath, A., Montanari, A., and Moretti, G.: Assessing the effect on flood frequency of land use change via hydrological simulation (with uncertainty), Journal of Hydrology, 324, 141 – 153, doi:http://dx.doi.org/10.1016/j.jhydrol.2005.10.001, http://www.sciencedirect.com/science/article/pii/S0022169405004816, 2006.

Cameron, D., Beven, K., Tawn, J., and Naden, P.: Flood frequency estimation by continuous simulation (with likelihood based uncertainty estimation), HYDROLOGY AND EARTH SYSTEM SCIENCES, 4, 23–34, 2000.

De Michele, C. and Rosso, R.: Uncertainty Assessment of Regionalized Flood Frequency Estimates, Journal of Hydrologic Engineering, 6(6), 2001.

Ganora, D. and Laio, F.: A comparison of regional flood frequency analysis approaches in a simulation framework, Water Resources Research, pp. n/a–n/a, doi:10.1002/2016WR018604, http://dx.doi.org/10.1002/2016WR018604, 2016.

Ganora, D., Laio, F., and Claps, P.: An approach to propagate streamflow statistics along the river network, Hydrological Sciences Journal, 58, 41 – 53, 2013.

Ganoulis, J.: Risk-based floodplain management: A case study from Greece, International Journal of River Basin Management, 1, 41–47, 2003.

Hosking, J. and Wallis, J.: Regional Frequency Analysis: An Approach Based on L-Moments, Cambridge University Press, 1997.

Jonkman, S., Brinkhuis-Jak, M., and Kok, M.: Cost benefit analysis and flood damage mitigation in the Netherlands, Heron, 49 (1), 2004.

Kjeldsen, T. and Jones, D.: Estimation of an index flood using data transfer in the UK, Hydrological Sciences Journal-Journal des Sciences Hydrologiques, 52, 86–98, 2007.

Laio, F., Ganora, D., Claps, P., and Galeati, G.: Spatially smooth regional estimation of the flood frequency curve (with uncertainty), Journal of Hydrology, 408, 67–77, doi:10.1016/j.jhydrol.2011.07.022, 2011.

Liang, Z., Chang, W., and Li, B.: Bayesian flood frequency analysis in the light of model and parameter uncertainties, Stochastic Environmental Research and Risk Assessment, 26, 721–730, doi:10.1007/s00477-011-0552-y, http://dx.doi.org/10.1007/s00477-011-0552-y, 2012.

Su, H. and Tung, Y.: Incorporating uncertainty of distribution parameters due to sampling errors in flood-damage-reduction project evaluation, Water Resources Research, 49, 1680–1692, doi:10.1002/wrcr.20116, http://dx.doi.org/10.1002/wrcr.20116, 2013.

Tung, Y.: Flood Defense Systems Design by Risk-Based Approaches, Water International, 30, 50–57, doi:10.1080/02508060508691836, 2005.

Viglione, A., Merz, R., Salinas, J. L., and Blöschl, G.: Flood frequency hydrology: 3. A Bayesian analysis, Water Resources Research, 49, 675 – 692, 2013.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

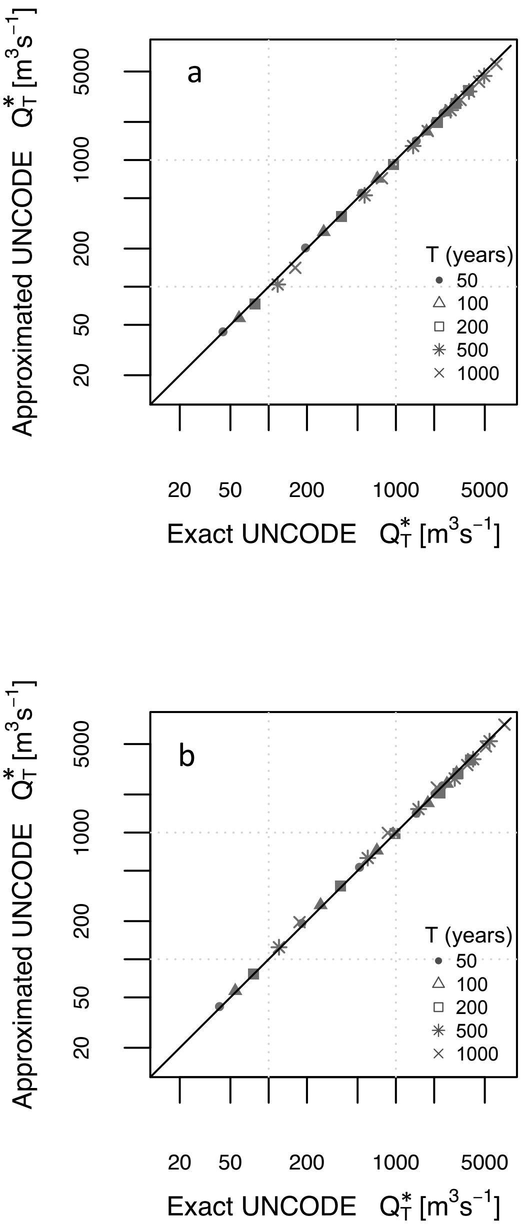

Figure 1. Comparison between the exact and the approximated UNCODE estimators of the design flood, Q_T^*, for a pool of 6 flood series considered in Botto et al. (2014, Table 1) with at least 30 years of data. Different return periods are listed in the legend. The reference distribution used for this flood frequency analysis is the 3-parameter log-Normal (LN3) in panel a) and the generalized extreme value (GEV) in panel b).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

Hydrology and Earth System Sciences Discussions
Open Access

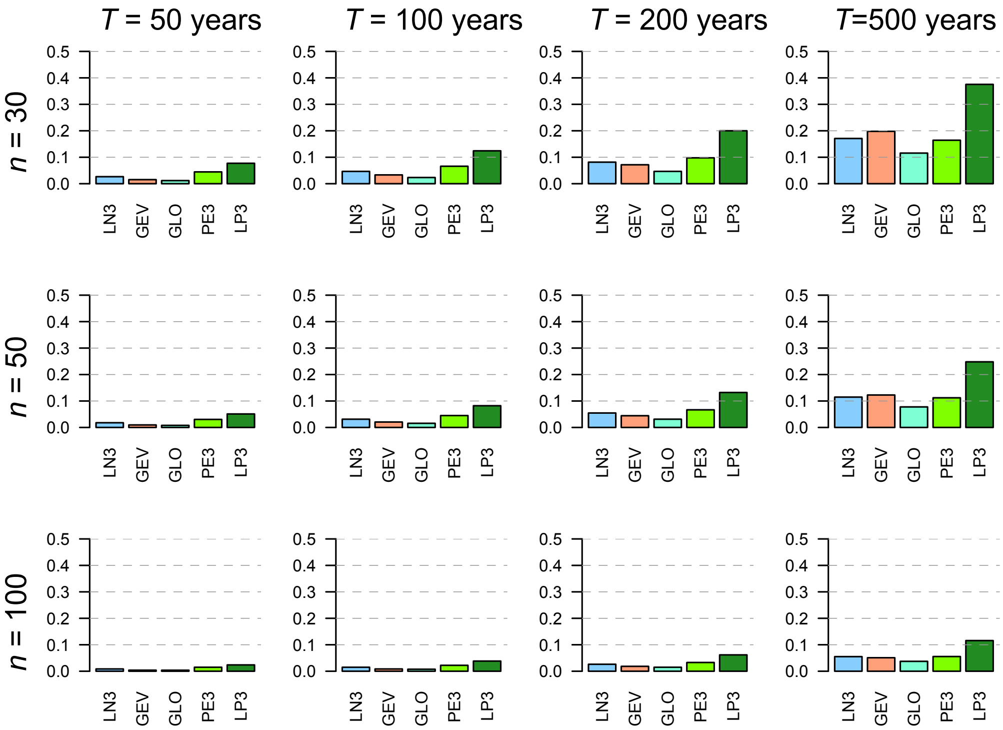

Figure 2. Values of the correction factor y from Eq. (5) for some values of the sample length n and return period T and for different three-parameter fitting distributions (LN3 = log-Normal; GEV= generalized extreme value; GLO = generalized logistic; PE3 = Pearson type III; LP3 = log Pearson type III). Details on the distributions can be found in Hosking and Wallis (1997); the LP3 corresponds to the PE3 with log-transformed variate.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-637, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 2 December 2016

Hydrology and Earth System Sciences
Discussions
Open Access

Table 1. Coefficients to be used to estimate y based on the sample length n and the return period T (eq. 5) and corresponding regression diagnostics, for different 3-parameter fitting distributions (LN3 = log-Normal; GEV= generalized extreme value; GLO = generalized logistic; PE3 = Pearson type III; LP3 = log Pearson type III). Details on the distributions can be found in Hosking and Wallis (1997); the LP3 corresponds to the PE3 with log-transformed variate.

	a_0	a_1	a_2	R^2_{adj}	MAE	$RMSE$
LN3	-0.82	-0.25	0.809	0.94	0.0107	0.0160
GEV	-2.27	-0.3	1.110	0.85	0.0190	0.0321
GLO	-2.36	-0.25	0.994	0.85	0.0096	0.0145
PE3	0.59	-0.24	0.567	0.96	0.0080	0.0115
LP3	0.78	-0.26	0.687	0.89	0.0235	0.0363