# Imprecise probabilistic estimation of design floods with epistemic uncertainties

**Wei Qi[a,b], Chi Zhang[a*], Guangtao Fu[b] and Huicheng Zhou[a]**

[a] School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

[b] Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK

[*] Corresponding author: email: czhang@dlut.edu.cn; Tel: +86 411 84708517; Fax: +86 411 84708517

**key points:**

An imprecise probabilistic approach is developed for design flood estimation

A new robustness criterion for design flood selection is proposed

Interactions among various sources of uncertainty affect the total cost considerably

**Abstract:** An imprecise probabilistic framework for design flood estimation is proposed on the basis of the Dempster-Shafer theory to handle different epistemic uncertainties from data, probability distribution functions and probability distribution parameters. These uncertainties are incorporated in cost-benefit analysis to generate the lower and upper bounds of the total cost for flood control, thus presenting improved information for decision making on design floods. Within the total cost bounds, a new robustness criterion is proposed to select a design flood that can tolerate higher levels of uncertainty. A variance decomposition approach is used to quantify individual and interactive impacts of the uncertainty sources on total cost. Results from three case studies, with 127-, 104- and 54-year flood data sets respectively, show that the imprecise probabilistic approach effectively combines aleatory and epistemic uncertainties from the various sources and provides upper and lower bounds of the total cost. Between the total cost and the robustness of design floods, a clear trade-off which is beyond the information that can be provided by the conventional minimum cost criterion is identified. The interactions among data, distributions and parameters have a much higher contribution than parameters to the estimate of the total cost. It is found that the contributions of the various uncertainty sources and their interactions vary with different flood magnitude, but remain roughly the same with different return periods. This study demonstrates that the proposed methodology can effectively incorporate epistemic uncertainties in cost-benefit analysis of design floods.

**Keywords:** Decision making; Dempster-Shafer theory; Frequency analysis; Hydraulic design; Imprecise probability; Uncertainty

## 1. Introduction

Estimation of design flood discharge related to a specific return period plays a crucial role in flood management: for example, the design of hydraulic structures. Conventionally, Flood Frequency Analysis (FFA) is used to estimate design floods, i.e., fitting Probability Distribution Functions (PDFs) to observed flood data and deriving a design flood discharge through the extrapolation of the upper distribution tail to specified low exceedance probabilities [Merz and Blöschl, 2008]. Recently, cost-benefit analysis has been incorporated into FFA to compare different design floods and obtain a cost effective design value [Tung and Mays, 1981; Bao et al., 1987; Ganoulis, 2003; Jonkman et al., 2004; Abrishamchi et al., 2005; Rossi et al., 2005; Su and Tung, 2013a; Su and Tung, 2013b; Botto et al., 2014]. It has been proven that the design flood value calculated using cost-benefit analysis with the assumption of liner damage and cost functions is equivalent to the flood value from the conventional FFA method [Botto et al., 2014].

A key issue in design flood estimation is to quantify and reduce the various uncertainties from different sources [Wood and Rodríguez-Iturbe, 1975a; Wood and Rodríguez-Iturbe, 1975b; Bodo and Unny, 1976; Stedinger et al., 1993; Tanaka and Takara, 2002; Pandey et al., 2004; Beguería, 2005; Merz and Thieken, 2005; Su and Tung, 2013b]. The inherent variability of flood events which is of aleatory uncertainty in nature is represented using a PDF. Significant uncertainties exist regarding the PDF derivation, such as the use of insufficient historical data, selection of PDFs and estimation of PDF parameters; most of these uncertainties are epistemic in nature and related to imprecise and incomplete knowledge about flood systems [Merz and Thieken, 2005; Fu et al., 2011; Su and Tung, 2013b].

Previous research has analysed the respective effects of epistemic uncertainties in data, PDF selection and distribution parameters on design flood estimation. It has been illustrated that a longer length of data could reduce the uncertainty in design flood estimation and project benefits [Su and Tung, 2013a; Su and Tung, 2013b; Botto et al., 2014]. However, the uncertainties related to the use of the Peak Over Threshold (POT) series selection have not been analysed in terms of cost-benefit. PDF selection has been widely recognised as a major uncertainty source for flood frequency analysis [Kidson and Richards, 2005; Calenda et al., 2009; Rahman et al., 2013]. Many PDFs, such as Generalized Extreme Value (GEV) and 3-parameter Lognormal (LN3), have been used for comparison [Botto et al., 2014], but the overall uncertainty from these PDFs which cannot be rejected using statistical tests has not been quantified and compared to other uncertainty sources. Parameter uncertainty of PDFs arises from parameter estimation approaches or data used in estimation, and has been represented using normal distribution [Su and Tung, 2013b] or using other distributions derived from the Monte Carlo approach [Tung and Mays, 1981; Bao et al., 1987; Botto et al., 2014]. Consequently, there is lack of understanding of combined and interactive contributions of different uncertainty sources to design flood estimates and lack of understanding of overall benefits of design options.

The various uncertainties should be represented and handled in a more holistic and coherent framework which will allow for a more realistic design flood estimation considering multiple uncertainty sources. Most recently, efforts have been made to systematically represent and quantify multiple uncertainty sources in a chain of models, such as investigating climatic impacts on hydrological systems and water resources management [Steinschneider et al., 2012], investigating impacts of precipitation and hydrological model uncertainties on discharge simulation uncertainty [Qi et al., 2016a], and investigating influence of parameter uncertainties on algorithm performance [Qi et al., 2016b]. Uncertainties in emission scenarios, global circulation models, downscaling methods and hydrological models have been quantified and their respective contributions to the overall output uncertainty have been compared [Vrugt et al., 2005; Wilby and Harris, 2006; Kay et al., 2009; Prudhomme and Davies, 2009; Bosshard et al., 2013]. In those prior studies, Monte Carlo based probabilistic approaches or sensitivity analysis approaches have been used. This holistic framework allows for identification of predominant sources of uncertainty and provides a more complete understanding of uncertainties in the modelling chain. To the best of our knowledge, a holistic framework has not been developed for design flood estimation, which requires simultaneous handling of multiple aleatory and epistemic uncertainty sources.

The overall aim of this paper is to develop an Imprecise Probabilistic Design Flood (IPDF) approach that can effectively handle various aleatory and epistemic uncertainties through cost-benefit analysis. In this approach, an imprecise probabilistic approach, based on Dempster-Shafer theory [Dempster, 1967; Shafer, 1976], is used to combine the epistemic uncertainties of data, probability distributions and their parameters. As a result, the lower and upper bounds of cumulative probabilities of flood can be generated and incorporated in cost-benefit analysis. The lower and upper total cost, including construction cost and expected flood damage cost, is then estimated explicitly. To select a robust design flood within the range of lower and upper total cost, a new criterion is proposed and contrasted with the conventional minimum total cost criterion. The individual and interactive contributions of different uncertainty sources to the overall uncertainty in estimating the total cost are quantified using a variance-based sensitivity analysis approach. Three case studies, with 54-year, 104-year and 127-year flood data respectively, were used to test the newly proposed IPDF approach. Similar to Botto et al. [2014], flood series were assumed to be stationary, i.e., the probability of occurrence of an extreme event in the current or any future year is the same [Olsen et al., 1998]. In each case study, three probability distributions were selected on the basis of the Anderson-Darling (A-D) test and different data sets were generated using POT and Annual Maximum (AM) methods to represent epistemic uncertainties. In this paper, there are three advancements from the present state of knowledge: (1) a new and holistic imprecise probabilistic estimation approach for design flood estimation is proposed and demonstrated to integrate aleatory and epistemic uncertainties; (2) a new robustness criterion is proposed and demonstrated to select a design flood that can tolerate higher levels of uncertainty, and a clear trade-off between the total cost and the robustness of design floods is identified, which is beyond the information that can be provided by previous research; and (3) the variance decomposition approach is used to quantify individual and interactive impacts of uncertainty sources on total cost, and it is found that the interactions among data, distributions and parameters affect the total cost considerably. These findings represent state-of-the-art knowledge in design flood estimation. The research of this paper could be used to evaluate design options and guide efforts to reduce the uncertainty from multiple epistemic uncertainties in design flood estimation.

This paper is divided into seven sections. Section 2 provides an overview of three case studies and the related epistemic uncertainties in data selection, probability distribution fitting and probability distribution parameters. Section 3 introduces the IPDF approach. The numerical procedures to implement the IPDF approach are described in section 4. Applications of the IPDF approach to real-world cases are presented in section 5. Discussion and conclusions are presented in section 6 and section 7.

## 2. Case studies

Three case studies of different flood record lengths are used in this study. These are selected from three rivers of different climates and different catchment areas: Yangtze River in south China, Songhuajiang and Biliu rivers in northeast China. Yangtze River is the largest rive in China with a catchment area of 2 million km$^2$, and is dominated by a sub-tropical humid monsoon climate with abundant rainfall. Songhuajiang is the third largest river in China with a catchment area of 0.56 million km$^2$, and is characterized by a temperate monsoon climate with long winter, aridness and low temperature. Biliu is a medium scale basin with a catchment area of 2814 km$^2$, and is characterized by a temperate monsoon marine climate. The daily flow records are from 1882 to 2008 (127 years) at the Three Gorges gauge station of Yangtze River, from 1898 to 2001 (104 years) at the Harbin site of Songhuajiang, and from 1958 to 2011 (54 years) at the Biliu gauge of Biliu.

## 2.1 Data selection uncertainty

Hydrological data are generally associated with different sources of uncertainties including data quality, representative data period selection, AM or POT series selection and length of time series, as summarised by Merz and Thieken [2005]. In this study, data uncertainty arises from the selection of historical data, represented by different data sets generated using AM and POT methods, and other data uncertainties are not considered.

Selection of a threshold value is normally based on expert judgement [Beguería, 2005], and its impact on design flood estimation is not fully understood [Tanaka and Takara, 2002; Pandey et al., 2004; Beguería, 2005]. To analyse the data uncertainty, a series of threshold values were adopted to generate flood series of different sizes, as shown in Table 1. The dependence between flood flows at different time steps was not considered, similar to other studies [Coles et al., 2003; Kidson and Richards, 2005; Calenda et al., 2009; Xu et al., 2009], as the impacts caused by ignoring the dependence seem negligible [Rosbjerg, 1985; Xu et al., 2009].

## 2.2 Probability distribution uncertainty

In FFA, the probability $\mathrm{P}\left[Q \geq q_T\right]$ of a *T*-year flood $q_T$ (the flood is exceeded once in *T* years on average) can be defined as

$$\mathrm{P}\left[Q \geq q_T \mid \boldsymbol{\theta}\right] = \int_{q_T}^{\infty} f\left(q \mid \boldsymbol{\theta}\right) dq \tag{1}$$

where *Q* denotes the random flow variable and $f\left(q \mid \boldsymbol{\theta}\right)$ denotes a PDF corresponding to the Cumulative Distribution Function (CDF) $F\left(q \mid \boldsymbol{\theta}\right)$. With the AM series, the sampling interval of observed floods is one year, so the number of events is automatically one per year. With the POT series, the number of occurrences of events in a given year is a random variable. Assuming a Poisson process [Cunnane, 1979; Onoz and Bayazit, 2001], the return period of *Q*, in years, can be calculated as [Rosbjerg, 1985; Rosbjerg et al., 1992; Madsen et al., 1997; Beguería, 2005; Bhunya et al., 2012; Bhunya et al., 2013]:

$$T = \frac{1}{\lambda \cdot \mathrm{P}\left[Q \geq q_T \mid \boldsymbol{\theta}\right]} \tag{2}$$

where $\lambda$ is the mean number of occurrences per year, and $T$ is the return period.

Many probability distributions have been proposed to simulate the true, unknown probability distribution of flood in the literature [Stedinger et al., 1993; Kidson and Richards, 2005]. There are three main approaches for distribution selection: official recommendation [Kidson and Richards, 2005], experience knowledge based selection [Merz and Thieken, 2005; Viglione et al., 2013] and statistical test based selection using methods such as L-Moments and A-D test [Chowdhury et al., 1991; Di Baldassarre et al., 2009; Kjeldsen and Prosdocimi, 2015]. A single distribution is often recommended for use in an entire country due to simplicity and practicality, and this approach is used by many countries in the world, though there is no theoretical basis [Calenda et al., 2009]. The selection based on goodness-of-fit tests is not conclusive and this does not support the view that only one candidate distribution should be selected as there may be several distributions that pass statistical tests [Kidson and Richards, 2005; Calenda et al., 2009; Laio et al., 2009; Rahman et al., 2013]. In the sense that many candidate distributions cannot be rejected, each can be considered as a possible distribution. The uncertainty resulting from probability distribution selection is referred to as distribution uncertainty hereafter.

A-D test is normally used to assess the goodness-of-fit of different distributions and it is suggested that it has good performance for extreme events as it gives more weight to the tails than the Kolmogorov-Smirnov test [Palynchuk and Guo, 2008; Calenda et al., 2009; Haddad and Rahman, 2010]. The null hypothesis is that the data follow a specified distribution. This hypothesis is rejected at the chosen significance level if the test statistic, $A^2$, is greater than the relevant critical value.

The A-D test results are shown in Table 2. Three distributions - GEV, Generalized Logistic (GLO) and LN3 - are shown. The maximum likelihood method was used to estimate distribution parameter values, and their probability density functions and cumulative probability density functions are shown in Appendix A. It should be noted that critical values of the A-D test vary with probability distribution types, distribution parameters and significance levels. D'Agostino and Stephens [1986; Table 4.32] introduced two approaches to calculate the critical values: empirical distribution function based approach [Stephens, 1974; Stephens, 1976; Stephens, 1977; Stephens, 1979; Ahmad et al., 1988] and normalized spacing based approach [Lockhart et al., 1986b]. It is argued that the latter is better than the former as it does not depend on a specific parameter estimation method [Lockhart et al., 1986a]. According to D'Agostino and Stephens [1986; Table 4.32], the critical values based on normalized spacings for GEV, Logistic and Normal distributions are 3.00, 3.41 and 2.73 at a significance level of 0.01, respectively. These critical values are used as a reference, as other statistics available for three-parameter distributions are not reliable [Laio, 2004]. All the distributions in Table 2 pass the test.

## 2.3 Parameter uncertainty

After the selection of data sets and probability distributions, the parameter uncertainty will arise in distribution parameter value estimation because of the limited length of the data sets. For considering the parameter uncertainty, parameter uncertainty bounds were first defined. Many methods are available to define parameter bounds of probability distributions: for example, subjective definition of an interval or perturbation around optimal estimates to generate lower and upper parameter bounds [Blazkova and Beven, 2002; Liang et al., 2011; Le Coz et al., 2014]; subjective definition of parameter distribution with known parameters [Reis and Stedinger, 2005; Ribatet et al., 2006; Lee and Kim, 2008; Su and Tung, 2013b]; and using regional information to define mean and variance of parameters [Perreault et al., 2000]. In this paper, because no prior information on distribution parameters was known, the perturbation method was used. The percentage perturbation of parameters was derived through trial and error to ensure all the observed extreme flow data were bracketed by the resulting lower and upper flow bounds. In this study, the posterior probability is calculated according to the Generalized Likelihood Uncertainty Estimation (GLUE) approach [Beven and Binley, 1992; Beven and Freer, 2001], which is equivalent to the importance sampling approach [Nott et al., 2012]. The incorporation of the parameter uncertainty into total cost calculation will be presented in Eqs. (8), (11), (14) and (19) in Section 3.

## 3. Imprecise probabilistic framework for design flood estimation

The IPDF approach is illustrated in Fig. 1. This new approach includes five components. (a) The first is uncertainty characterisation of different sources, i.e., different probability distributions, their distribution parameters and different data thresholds (three thresholds $T_1$, $T_i$ and $T_n$ are shown for illustration). (b) The second is uncertainty combination using evidence theory, which results in lower and upper bounds of probabilities. (c) The third is cost-benefit analysis to show the variations of total cost. The uncertainty of the total cost is propagated from the imprecise probabilities of flow. (d) The fourth is sensitivity analysis to quantify individual and interactive contributions of different uncertainty sources using a variance decomposition method: the ANalysis Of VAriance (ANOVA). ANOVA can identify important uncertainty sources, and guide efforts to reduce uncertainty. (e) The fifth component is a new robustness criterion to select design flood. As shown in Fig. 1c, the *T*-year design flood falls in an interval

($[D_1, D_n]$). In this interval, different design floods can be selected and compared based on their robustness which is evaluated by measuring the variations of total cost with different uncertainty levels (as illustrated by $\mu_0$, $\mu_i$ and $\mu_n$). Steps (a) - (d) can be repeated to reduce the uncertainty of total cost when new data or distribution models are included. Compared to previous methods [e.g., *Su and Tung*, 2013a, b; *Botto et al.*, 2014], this new IPDF approach provides the upper and lower bounds of minimum total cost for a specific *T*-year design flood, as a result of considering different types of epistemic uncertainties. Details of each component are presented in the following subsections.

### 3.1 Dempster-Shafer theory of evidence

Dempster-Shafer theory of evidence is a kind of set-valued and evidence-based theory and can describe overall uncertainties of stochastic and epistemic nature. It can handle uncertainties from different aleatory and epistemic sources [Hall, 2003; Hall and Lawry, 2004; Hall et al., 2004; Fu et al., 2011]. This theory has been used in many fields, such as water distribution system design [Fu and Kapelan, 2011], evaluation of sewer flooding [Fu et al., 2011], groundwater flow and transport simulation [Ross et al., 2009], reliability analysis [Tonon et al., 2000], climate change [Hall et al., 2007], and rainfall-runoff modelling [Maskey et al., 2004]. One main difference from the Bayesian theory is that the Dempster-Shafer theory admits imprecision in probability (e.g., a probabilistic interval), whilst the Bayesian theory assumes that uncertainty should always be measured by a single probability [Walley, 1991; Hall, 2003; Fu and Kapelan, 2011]. From this point of view, the Dempster-Shafer theory can be regarded as a generalization of probability theory to cope with a problem for which information is not enough for an assignment of a single probability. Many uncertainties in design flood estimation are epistemic and do not allow the assignment of a single probability value due to insufficient information or conflicting evidence, thus it is promising to apply the evidence theory to handle various uncertainties.

In Dempster-Shafer theory the minimum and maximum amounts of evidence can be taken into consideration to construct probability. For example, suppose that based on evidence $\Psi_1$, the probability of a set of states $\mathbf{\Phi} = \{\Phi_1, \cdots, \Phi_n\}$ which relate to interests $\Omega$ (e.g., the probability of flood events) can be assigned as $\mathrm{P}[\Phi_i \mid \Psi_1] = p_i$, while using another evidence $\Psi_2$, the probability of a set of states $\mathbf{\Theta} = \{\Theta_1, \cdots, \Theta_k\}$ which also relate to interests $\Omega$ can be assigned as $\mathrm{P}\left[\Theta_j \mid \Psi_2\right] = q_j$. The Dempster-Shafer theory of evidence can be used to combine $p_i$ with $q_j$ to give $m^{\Psi_1,\Psi_2}$ which represents beliefs assigned to interests $\Omega$ based on evidence $\Psi_1$ and $\Psi_2$.

Let *X* be a universal nonempty set containing all the possible values of a variable *x*, and *P*(*x*) is the power set of *X*, i.e., the set of all the subsets of *X*. Dempster-Shafer theory of evidence can be defined as a pair ($\xi$, *m*), where $\xi$ is the family of nonempty element of *P*(*x*) and *m* is mapping

$$m : \xi \to [0,1] \tag{3}$$

Such that $m(\varnothing) = 0$ and

$$\sum_{A \in \xi} m(A) = 1 \tag{4}$$

where $A \in P(X)$ and *m* is called the basic probability assignment. The related imprecision of probability can be bounded at the lower end by a belief function

$$Bel(E) = \sum_{A \subset E} m(A) \tag{5}$$

and at the upper end by a plausibility function

$$Pl(E) = \sum_{A \cap E \neq \varnothing} m(A) = 1 - Bel(\overline{E}) \tag{6}$$

where $\overline{E}$ is the complement of *E*. The $Bel(E)$ measures the minimum amount of evidence that fully supports $x \in E$ . $Pl(E)$ measures the maximum amount of evidence that could be linked with the event *E*.

## 3.2 Annual expected damage cost estimation

For hydrologic structures, the Annual Expected Damage Cost (AEDC) can be defined below using the probability $\mathrm{P}\left[Q \geq q_T\right]$ in Eq. (1)

$$E\left(D \mid q_T^*, \boldsymbol{\theta}\right) = \int_{q_T^*}^{\infty} D(q) \cdot f\left(q \mid \boldsymbol{\theta}\right) dq \tag{7}$$

where $q_T^*$ is the *T*-year flow capacity of hydraulic structures; $D(q)$ represents the flood-damage function corresponding to a flood magnitude of $q$ and $\boldsymbol{\theta}$ represents the parameters. When parameter uncertainty is considered, the expected damage Eq. (7) can be written as [Bao et al., 1987]

$$E\left(D \mid q_T^*, \boldsymbol{\theta}, S\right) = \int_{q_T^*}^{\infty} D(q) \cdot \int_{\Theta} f\left(q \mid \boldsymbol{\theta}\right) \cdot \pi_{\boldsymbol{\theta}|S}\left(\boldsymbol{\theta} \mid S\right) d\boldsymbol{\theta} dq \tag{8}$$

where $\Theta$ is the parameter space of a probability distribution; *S* is the sample of flow data. The sampling distribution $\pi_{\boldsymbol{\theta}|S}\left(\boldsymbol{\theta} \mid S\right)$ can be calculated based on Bayesian theory as described below

$$\pi_{\boldsymbol{\theta}|S}\left(\boldsymbol{\theta} \mid S\right) = \frac{\pi\left(\boldsymbol{\theta}\right) \cdot l_{S|\boldsymbol{\theta}}}{\int_{\Theta} \pi\left(\boldsymbol{\theta}\right) \cdot l_{S|\boldsymbol{\theta}}} \tag{9}$$

where $\pi\left(\boldsymbol{\theta}\right)$ is the prior probability of parameter $\boldsymbol{\theta}$; $\pi_{\boldsymbol{\theta}|S}\left(\boldsymbol{\theta} \mid S\right)$ is the posterior probability of parameter $\boldsymbol{\theta}$; $l_{S|\boldsymbol{\theta}}$ is likelihood function

$$l_{S|\boldsymbol{\theta}} = \prod_{i=1}^{N} f\left(x_i \mid \boldsymbol{\theta}\right) \tag{10}$$

where *N* represents the total number of sampled flow data.

Based on the total probability theorem the predictive distribution is obtained

$$f^{PD}(q) = \int_{\Theta} f(q \mid \boldsymbol{\theta}) \cdot \pi_{\theta|S}(\boldsymbol{\theta} \mid S) d\boldsymbol{\theta} \tag{11}$$

An analytical solution of Eq. (11) can be derived only for a few probability distributions [Stedinger, 1983; Kuczera, 1999; Fawcett and Walshaw, 2015], and in practice, Monte Carlo method can be used to calculate the integral in Eq. (11). When data uncertainty and distribution uncertainty are considered, $f^{PD}(q)$ is not unique and the lower and upper bounds of imprecise probabilities, $\underline{f^{PD}(q)}$ and $\overline{f^{PD}(q)}$, can be defined as

$$\underline{f^{PD}(q)} = \inf_{e} f_{e}^{PD}(q) \tag{12}$$

$$\overline{f^{PD}(q)} = \sup_{e} f_{e}^{PD}(q) \tag{13}$$

where *e* represents the *e*th probability function. $\underline{f^{PD}(q)}$ and $\overline{f^{PD}(q)}$ correspond to $Bel(E)$ and $Pl(E)$ in the sense of the Dempster-Shafer theory of evidence. The lower probability $\underline{f^{PD}(q)}$ measures the minimum probability that evidence can fully support, i.e., the minimum probability calculated from selected data and probability distributions, and the upper probability $\overline{f^{PD}(q)}$ measures the maximum probability that evidence can potentially support, i.e., the maximum probability calculated from selected data and probability distributions. The interval formed with Eqs. (12) and (13) provides a bracketing of a series of probabilities and its spread represents the extent of incomplete knowledge and imprecise information about the unknown but true distribution. Combining Eq. (8) with Eqs. (12) and (13), AEDC considering epistemic uncertainties can be defined as follows

$$\begin{cases} \underline{E\left(D \mid q_T^*, S\right)} = \int_{q_T^*}^{\infty} D(q) \cdot \underline{f^{PD}(q)} dq \\ \overline{E\left(D \mid q_T^*, S\right)} = \int_{q_T^*}^{\infty} D(q) \cdot \overline{f^{PD}(q)} dq \end{cases}, \tag{14}$$

where $\underline{E\left(D \mid q_T^*,S\right)}$ and $\overline{E\left(D \mid q_T^*,S\right)}$ represent the minimum and maximum values of AEDC estimation.

### 3.3 Imprecise probabilistic estimation of design floods

Assuming that the construction cost and damage functions are linear and represented as $c \cdot q_T^*$ and $d \cdot \left(q - q_T^*\right)$, respectively, the total cost function is described below according to Eq. (8)

$$C_{total}\left(q_T^* \mid \boldsymbol{\theta},S\right) = c \cdot q_T^* + \int_{q_T^*}^{\infty} d \cdot \left(q - q_T^*\right) \cdot \int_{\Theta} f\left(q \mid \boldsymbol{\theta}\right) \cdot \pi_{\theta|S}\left(\boldsymbol{\theta} \mid S\right) d\boldsymbol{\theta} dq \tag{15}$$

and can also be written as

$$C_{total}\left(q_T^* \mid S\right) = c \cdot q_T^* + \int_{q_T^*}^{\infty} d \cdot \left(q - q_T^*\right) \cdot f^{PD}\left(q\right) dq \tag{16}$$

where $c$ and $d$ are parameters.

There exists a deterministic relationship between parameters $c$ and $d$ and return period $T$ under the linear cost and damage assumption according to Botto et al. [2014]

$$\frac{d}{c} = \frac{1}{\mathrm{P}\left[Q \geq q_T^* \mid S\right]} = T \tag{17}$$

which is derived by minimizing Eq. (16) using AM data, i.e., taking the derivative of the total cost function with respect to $q_T^*$ and setting it to 0. When POT data sets are considered, the above relationship becomes

$$\frac{d}{c} = \frac{1}{\lambda \cdot \mathrm{P}\left[Q \geq q_T^* \mid S\right]} = T \tag{18}$$

Eq. (18) is a generalization of Eq. (17), i.e., $\lambda$ equals to 1 for AM data set. In this paper, the flood-damage data are available for Biliu, and a linear function is fitted with $d$ =1.891. For Three Gorges and Harbin case studies, the same value of $d$ is assumed since damage data are not available.

When data uncertainty and distribution uncertainty are considered, the total cost function Eq. (16) becomes

$$\begin{cases} \underline{C_{total}\left(q_T^* \mid S\right)} = c \cdot q_T^* + \int_{q_T^*}^{\infty} \lambda(q) \cdot d \cdot \left(q - q_T^*\right) \cdot \underline{f^{PD}(q)} dq \\ \overline{C_{total}\left(q_T^* \mid S\right)} = c \cdot q_T^* + \int_{q_T^*}^{\infty} \lambda(q) \cdot d \cdot \left(q - q_T^*\right) \cdot \overline{f^{PD}(q)} dq \end{cases} \tag{19}$$

which shows the lower and upper bounds of the total cost, incorporating epistemic uncertainties from data, probability distribution, and parameter uncertainties into the aleatory uncertainty of flood.

## 3.4 Robustness criterion

The minimum total cost criterion can be used to select the design flood. This criterion is employed in the case of one single total cost curve generated as in the study of Botto et al. [2014]. In the case of imprecise probabilities, a range of total costs can be obtained, bounded by the lower and upper curves. The total cost intervals provide an indication of the magnitude of total cost uncertainty which is faced by the decision maker when selecting a design flood, and with these intervals, the selection of a design flood depends on the preference of the decision maker or the use of decision criteria. However, the minimum total cost criterion [Botto et al., 2014] or the expected opportunity loss criterion [Su and Tung, 2013a; Su and Tung, 2013b] can be used for cases where parameter uncertainty is considered only.

A robustness criterion is proposed here to analyse the differences of design floods. The robustness, defined in the sense of the Info-gap theory [Ben-Haim, 2006; Hine and Hall, 2010], seeks a design value that can make a system maintain its prescribed functions over a range of uncertainty levels. In design flood estimation, robustness involves connecting $C_{total}$ with decision variation $q_T$ under an uncertainty level of $\partial$:

$$\hat{\partial}(q_T, r_c) = \max\left(\partial : \min C_{total}(q_T|\partial) \geq r_c\right) \quad (20)$$

where $r_c$ is a critical level of $C_{total}$. This critical level can be assumed to be the minimum $C_{total}$ under an uncertainty level of $\partial$, thus robustness can be interpreted as the variation of the minimum $C_{total}$ under many discrete uncertainty levels of $\partial$ [Matrosov et al., 2013]. The smaller the variations at different uncertainty levels, the more robust the design flood.

## 3.5 Variance decomposition

ANOVA is used to analyse the respective contributions of data, distributions, distribution parameters and their interactions to the overall uncertainty in total cost, $C_{total}$. Fig. 2 depicts the combinations employed in the uncertainty decomposition. To relate $C_{total}$ to the uncertainty sources, the superscripts *j*, *k* and *l* in $C_{total}^{j,k,l}$ are used to represent a combination of data set *j*, distribution *k* and parameters *l*. Two cases, without parameter uncertainty (using the estimated optimal parameters) and with parameter uncertainty (using predictive probability distributions), are considered.

### 3.5.1 Subsampling approach

It has been argued that the ANOVA approach is based on a biased variance estimator that underestimates the variance when the sample size is small [Bosshard et al., 2013]. To reduce the effect of the biased estimator on quantification of variance contribution, Bosshard et al. [2013] proposed a subsampling method, which was used in this paper. In each subsampling iteration, *i*, we select two data sets out of all data sets analysed, and the superscript *j* (data set) in calculating $C_{total}^{j,k,l}$ is replaced with $\mathbf{g}(h,i)$. In the case of Three Gorges, the time series is divided into nine non-overlapping subsets resulting in 9!/(2!(9-2)!)=36 possible combinations of two elements, and correspondingly the superscript **g** is a 2×36 matrix as follows

$$\mathbf{g} = \begin{pmatrix} 1 & 1 & \cdots & 1 & 2 & 2 & \cdots & 6 & 6 & 6 & 7 & 7 & 8 \\ 2 & 3 & \cdots & 9 & 3 & 4 & \cdots & 7 & 8 & 9 & 8 & 9 & 9 \end{pmatrix} \tag{21}$$

Similarly the superscript **g** is a 2×28 matrix in the case study of Harbin and a 2×15 matrix in the case study of Biliu.

### 3.5.2 ANOVA approach

Based on the ANOVA method, the total sum of squares (SST) of $C_{total}$ can be divided into sums of squares of the individual effects (with SSA, SSB and SSC corresponding to the contribution of data, probability distributions and parameters respectively) and of their interactions (SSI) as follows:

$$\mathrm{SST} = \mathrm{SSA} + \mathrm{SSB} + \mathrm{SSC} + \mathrm{SSI} \tag{22}$$

The terms can be estimated using the subsampling procedure as follows [Bosshard et al., 2013]:

$$\mathrm{SST}_i = \sum_{h=1}^{H}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(C^{\mathbf{g}(h,i),k,l} - C^{\mathbf{g}(o,i),o,o}\right)^2 \tag{23}$$

$$\mathrm{SSA}_i = K \cdot L \cdot \sum_{h=1}^{H}\left(C^{\mathbf{g}(h,i),o,o} - C^{\mathbf{g}(o,i),o,o}\right)^2 \tag{24}$$

$$\mathrm{SSB}_i = H \cdot L \cdot \sum_{k=1}^{K}\left(C^{\mathbf{g}(o,i),k,o} - C^{\mathbf{g}(o,i),o,o}\right)^2 \tag{25}$$

$$\mathrm{SSC}_i = H \cdot K \cdot \sum_{l=1}^{L}\left(C^{\mathbf{g}(o,i),o,l} - C^{\mathbf{g}(o,i),o,o}\right)^2 \tag{26}$$

$$\mathrm{SSI}_i = \sum_{h=1}^{H}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(C^{\mathbf{g}(h,i),k,l} - C^{\mathbf{g}(h,i),o,o} - C^{\mathbf{g}(o,i),k,o} - C^{\mathbf{g}(o,i),o,l} + 2 \cdot C^{\mathbf{g}(o,i),o,o}\right)^2 \tag{27}$$

where symbol $^{o}$ indicates averaging over the particular index. Then the contribution of each uncertainty source $\eta^2$ is calculated as follows:

$$\eta^2_{\mathrm{data}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\mathrm{SSA}_i}{\mathrm{SST}_i} \tag{28}$$

$$\eta^2_{\text{distribution}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\text{SSB}_i}{\text{SST}_i} \tag{29}$$

$$\eta^2_{\text{parameter}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\text{SSC}_i}{\text{SST}_i} \tag{30}$$

$$\eta^2_{\text{interaction}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\text{SSI}_i}{\text{SST}_i} \tag{31}$$

$\eta^2$ has a value between 0 and 1, which represent 0% and 100% of contribution to the overall uncertainty of the total cost respectively.

## 4. Implementation of the proposed new approach

This section is devoted to describe how the new proposed IPDF approach can be applied in practice. The numerical procedures are implemented according to the following main steps:

1. Uncertainties are clearly identified: e.g., data selection, probability distribution and distribution parameter uncertainty, which forms the basis of the implementation (see Fig. 1a).
2. The imprecise probabilities are quantified on the basis of Eqs. (12) and (13). The upper and lower probability bounds can be obtained (see Fig. 1b).
3. Once the uncertainties are quantified, Eq. (19) is applied to calculate total costs. Two total cost curves can then be obtained: lower and upper total cost bounds (see Fig. 1c).
4. The ANOVA approach is applied to quantify the contributions of uncertainties to the total cost uncertainty on the basis of Eqs. (21)-(31) (see Fig. 1d).
5. The robustness criterion is applied to evaluate flood value estimates based on Eq. (20) (see Fig. 1e). This criterion could be provided to decision maker for informed decision making.

It should be noted that, in the first step, although the identification of the uncertainties is subjective, this procedure enables a rigorous evaluation of the respective and combined impacts of the uncertainties and thus provides an enhanced understanding of their impacts on the selection of design floods. Its application to three real-world cases is described below in Section 5.

## 5. Application to real-world cases

In this section, the newly proposed methodology is demonstrated step by step in the subsections from 5.1 to 5.4. Section 5.1 first shows imprecise probability characteristics of flood through integration of the uncertainties in data selection, probability distributions and parameters, derived from Eqs. (12) and (13). Section 5.2 shows the total cost derived from Eq. (19). Section 5.3 discusses the use of a new robustness criterion for design flood selection. Section 5.4 discusses the contributions of different uncertainty sources to the overall uncertainty in total costs using the variance decomposition method.

### 5.1 Imprecise probability

In this study, a Monte Carlo based method was used to compute the posterior distributions using GLUE. 2000 parameter sets for each distribution were sampled within the parameter uncertainty bounds using Latin Hypercube Sampling. Sampling (larger parameter sets were also used obtaining similar results) 2000 parameter sets were used in the research.

Fig. 3 shows the sampling distributions of a specific design flood obtained on the basis of the posterior parameter distributions of GEV, GLO, and LN3 in Biliu, Three Gorges and Harbin. Note that the posterior distributions are reduced to a single curve when integrated via Eq. (11). Cumulative probability curves in each panel represent different data sets under the same probability distribution. In the case of Biliu, the curves span a large range, while most curves from the other two case studies are closer to each other, except for one curve (i.e., AM in the case of Three Gorges and T3 in the case of Harbin). The big departures of AM in Three Gorges and T3 in Harbin imply high uncertainties in flood estimation when the corresponding data sets are considered only. Recall that the specific distributions cannot be rejected under each of the data sets using the A-D test. However, the spread of the distribution curves clearly shows the epistemic uncertainties in the selection of data sets. Similarly, comparing the differences in each panel reveals the significant epistemic uncertainties in the selection of distributions.

Fig. 4 shows the imprecise cumulative probability distributions of Biliu, Three Gorges and Harbin, respectively, when data and distribution uncertainties are incorporated with parameter uncertainty. For each individual probability distribution (GEV, GLO and LN3), as shown in the panels (*a*-*i*) of the first three rows, the probability of each flood value is calculated based on predictive distributions and the intervals are derived from the selected data sets, i.e., 6 data sets, 9 data sets and 8 data sets for Biliu, Three Gorges and Harbin, respectively. The overall CDFs in the panels (*j*-*l*) of the fourth row result from the selected data sets listed in Table 1, probability distributions (GEV, GLO and LN3) and calculated predictive distributions related to parameter uncertainty using Eq.(11). In the case of Biliu, the overall probability bounds are roughly the same as those of each individual distribution, implying the distribution uncertainty has less impact than the data uncertainty. However, in the case of Three Gorges, the overall bounds are primarily determined by the bounds of LN3, implying that the distribution uncertainty is the dominating uncertainty source. The case of Harbin shows a mixed impact from both data and distribution uncertainties. This is compared with the study of Botto et al. [2014] where only one predictive distribution was generated when considering the uncertainty of distribution parameters only. Theoretically this predictive distribution should lie within the grey areas, i.e., bracketed by the lower and upper probabilistic bounds, because in this research data selection, probability distribution and parameter uncertainties all are considered and the bounds represent the minimum and maximum probabilities.

## 5.2 Imprecise probabilistic estimation of total cost

Fig. 5 illustrates the lower and upper total cost bounds for the three case studies when data, distribution and parameter uncertainties are considered. For any design flood value shown on *x*-axis, the lower and upper bounds of the total cost are represented by the two curves in each panel. For each individual distribution in the panels (*a*-*i*), the intervals illustrate the uncertainties in data and distribution parameters; for the cases of overall uncertainty in the panels (*j*-*l*), the intervals illustrate the uncertainties in data, distributions and distribution parameters. This is compared with the study of Botto et al. [2014] where only one curve was generated when considering the uncertainty of distribution parameters only.

In the case of Three Gorges, the differences of individual distributions in upper and lower total cost bounds are remarkable, and in the cases of Biliu and Harbin, the differences are also obvious, as shown in the first three rows of panels in Fig. 5. The overall upper and lower total cost bounds are notably larger than those of each distribution in all the three cases, in particular, in the cases of Biliu and Harbin. In the case of Three Gorges, the total cost bounds are mainly affected by the uncertainty in the selection of distributions, while in the cases of Biliu and Harbin, the influence of data sets and distributions on total cost bounds are all important. In total cost calculation, the lower and upper probability bounds are multiplied by flood damage and flood values, resulting in rather different total cost bounds due to their highly nonlinear relationships (e.g., as shown in Eq. (19)).

## 5.3 Design flood selection using a robustness criterion

In this study, 300 uncertainty levels were used. This means the uncertainty intervals from the median CDF towards lower and upper bounds ( $\overline{f^{PD}(q)}$ and $\underline{f^{PD}(q)}$ ) in Fig. 4 were discretized into 300 subintervals. The variations of minimum total cost are thus calculated for each uncertainty level as shown in Fig. 6. The minimum total cost within each uncertainty level is shown on the *x*-axis, and robustness is shown on the *y*-axis under a set of uncertainty levels ($\alpha$ %). Under each uncertainty level (except when $\alpha$ equals 1), the parameter $\lambda$ (recall Eq. (19)) was unknown, and the minimum value of $\lambda$ out of all selected data sets in each case study was used for calculating the minimum total cost. Two design flood selection criteria are compared in Fig. 6: the minimum total cost approach [Botto et al., 2014] and the robustness based approach. An $\alpha$ value of 0 means that the probability of a flood $q$ is determined by the median CDF, while $\alpha$=1 represents the maximum deviation degree: upper and lower probability bounds. The minimum total cost curve corresponding to $\alpha$=0 is shown by the dashed lines in Fig. 5.

The existence of robust decisions depends on both the degree of uncertainty and the richness of the available decision options [Lempert and Collins, 2007]. In this research, we did not try to find the robust decisions but to assess the robustness of options, thus the richness of options doesn't matter. The upper and lower total cost bounds correspond to different design floods with minimum total cost, and the optional design floods fall in an interval. To make an informed decision, the decision maker is presented with the intervals represented by the upper and lower total cost curves, though the exact design flood is unknown. Thus, for comparison with D1 which represents the results of the traditional minimum total cost criterion, D2 and D3 were selected within the interval, representing two possible design floods that might be selected by decision makers. It should be noted that D2 and D3 correspond to the minimum total cost of two total curves respectively. In Biliu, D1, D2 and D3 are 5700 $m^3$/s, 8000 $m^3$/s and 10,000 $m^3$/s respectively. In Three Gorges, they are 73,900 $m^3$/s, 80,000 $m^3$/s and 95,000 $m^3$/s respectively. In Harbin, they are 23,800 $m^3$/s, 35,000 $m^3$/s and 40,000 $m^3$/s respectively. The selected flood values are marked in Fig. 5.

In Fig. 6, each curve represents a design flood, and its slope describes the variation of minimum total cost with uncertainty ($\alpha$ %). The steeper the slope is, the more robust the design is. If a curve is on the right hand side of another, it has a larger minimum total cost. In the case of Biliu, the curve of D1 is gentler than the other two designs, thus fewer changes in uncertainty can result in larger perturbation in total cost. The robustness curves become steeper with an increase in design floods from D1 to D3, thus the robustness increases, but the smallest minimum total cost increases as well, when $\alpha$=1. Similarly, in the cases of Three Gorges and Harbin, D1 options are less robust compared with D2 and D3, but the smallest minimum total cost of D2 and D3 is larger than D1. Between total cost and robustness there is a clear trade-off which decision makers need to balance in the decision making process. Under some uncertainty levels, D1 has a larger minimum total cost than D2 or D3: for example, for $\alpha$=0 in Biliu, D1 is larger than D2 but is smaller than D3, and in Three Gorges D1 is larger than D2 and D3, which results from the differences in the total cost curve corresponding to $\alpha$=0 (shown in Fig. 5 as dashed lines), and implies that smaller design floods do not mean smaller total cost and larger total cost does not mean robust designs. In Biliu and Harbin, the curves corresponding to $\alpha$=0 are close to the lower total cost bound, while in Three Gorges the result is different: the curves corresponding to $\alpha$=0 is close to the upper total cost bound. The differences may be because of the variations in upper and lower probability bounds and parameter $\lambda$ (recall Eq. (19)).

Although only three design flood values are selected for comparison, the results reveal the patterns: with an increasing design flood magnitude, more uncertainties can be tolerated while still guaranteeing the calculated total cost varies only slightly; thus the robustness increases, but the minimum total cost increases as well. Likewise, although 300 uncertainty levels were used, the results show the different robustness of design floods. Larger uncertainty level numbers were also analysed resulting similar robustness analysis results.

### 5.4 Contributions of uncertainty sources

Fig. 7 shows the breakdown results when applying ANOVA, i.e., the total cost curves of different data and probability distribution combinations under two cases: with parameter uncertainty (using predictive probability distributions – the first row panels) and without parameter uncertainty (using the estimated optimal parameters – the second row panels). It can be seen that total cost curves change with the variations of data set and probability distribution combinations. For example, in Biliu, there are $6\times3$ different total cost curves and these curves span large areas. Comparing the panels (*a*-*c*) and (*d*-*f*), it can be seen that the total cost curves are different in the two cases with and without parameter uncertainty. For example, in Three Gorges, when considering parameter uncertainty, the total cost curve of the AM-LN3 combination (the most upper total cost curve in Fig. 7b) moves up compared with the optimal parameter case (the most upper total cost curve in Fig. 7e).

As shown in Eqs. (16)-(19), the total cost is a function of return period $T$. Thus, the total cost is different for different return period floods. On the basis of Eqs. (21)-(31), Fig. 8 shows the contributions of individual uncertainty sources, i.e., data selection, distribution and parameter uncertainties, and their interactions to the overall uncertainty in total cost in Biliu, Three Gorges and Harbin for three return periods, i.e., 500-, 1000- and 2000-year, respectively. The contributions of uncertainty sources are represented by the strips varying with flood values on *x*-axis.

In Biliu, regarding the 500 years return period, the contributions of data and distribution uncertainty sources varies slightly with flood magnitude. Interactions which cannot be considered in conventional FFA have a much higher contribution than parameter uncertainty, and approximately have the same contribution as distributions. Other return periods in Biliu show the same tendency. Similarly, in Three Gorges and Harbin, the contributions of uncertainty sources vary significantly with flood magnitude but almost have no changes in different return periods. The contribution of interactions is larger than parameter uncertainty in Three Gorges and Harbin also. In Three Gorges and Harbin which have much longer flow records than Biliu, with flow increases, contribution from interactions decreases and contribution from distributions increases. Comparing the differences in the data contribution among the three cases, the longer the data record, the less impact it has. For example, Three Gorges with the longest data record (23-year longer than Harbin and 73-year longer than Biliu) has the least impact from data uncertainty: the uncertainty contributions from data at most are 10.7% in Three Gorges, 38.9% in Harbin and 45.1% in Biliu. The similar contributions of different uncertainty sources in different return periods imply that the return periods have little influence on the relative influences of different uncertainty sources.

## 6. Discussion

Botto et al. [2014] incorporated parameter uncertainty in the design flood estimation through cost-benefit analysis, however, epistemic uncertainties from other sources, e.g., data and probability distribution uncertainties, were not incorporated. Several studies have compared the separate influence of data and distribution epistemic uncertainties in flood estimation [Beguería, 2005; Merz and Thieken, 2005]. Bao et al. [1987] studied the influence of the number of data and four different probability distributions on annual expected damage cost separately. Su and Tung [2013b] studied the influence of different parameter estimation methods on flood damage. However, combining data, distribution and parameter uncertainties in design flood estimation has not been investigated in the previous literature. The approach proposed in this paper systematically combines the above mentioned aleatory and epistemic uncertainties (data, probability distribution and distribution parameter uncertainties) in a holistic framework.

Every single curve in Fig. 7 represents the results of total costs should the previous approach proposed by Botto et al. [2014] be used. In the approach proposed by Botto et al. [2014], which effectively addressed distribution parameter uncertainty in a cost-benefit analysis approach, one single total cost curve is generated to find the optimal design flood estimate. Building on this work, our approach can take other uncertainty sources (such as probability distribution and data selection uncertainties) into consideration, and thus generate uncertainty intervals (the grey areas in Fig. 7). It can also be seen that the grey areas are different from the uncertainty ranges spanned by all the total cost curves: for example, the lower bounds of the grey areas can be larger (e.g., Figs. 7a and 7b) or smaller (e.g., Fig. 7c) than the ranges of all the total cost curves. These differences are because the total cost calculation (Eq. (19)) considers data selection uncertainty, probability distribution uncertainty and parameter uncertainty and is very different from previous approaches: for example, only considering parameter uncertainty [Botto et al., 2014]. In addition, before a decision maker makes a decision, the design flood value is a range using the newly developed approach, but the traditional approaches, such as FFA and the approach proposed by Botto et al. [2014], provide decision maker a precise deign flood value. The design flood values obtained from the newly proposed approach in this research are shown to be no smaller than results using FFA and the UNcertainty COmpliant DEsign (UNCODE) approach proposed by [Botto et al., 2014]. For example, Table 3 shows the design flood estimates from the newly proposed IPDF approach, the UNCODE approach and FFA. In IPDF, the design flood intervals correspond to the minimum total costs in the lower and upper total cost bounds (to clearly show the minimum total costs in the lower total cost bounds, the minimum total costs in the upper total cost bounds are not shown in Fig. 7); in UNCODE, the design floods correspond to the minimum total costs among all the total cost curves shown in Fig. 7; in FFA, the design floods correspond to the minimum values among all the data set and distribution combinations. In the case of Three Gorges, the minimum 1000-year flood from the newly proposed IPDF approach is 73,900 $m^3/s$, but it is 73,700 $m^3/s$ and 73,400 $m^3/s$ from UNCODE and FFA respectively (i.e., 0.3% and 0.7% smaller than the IPDF result respectively); in the cases of Harbin and Biliu, the minimum design floods of IPDF are no less than those from UNCODE and FFA. In addition, as shown in Table 3, IPDF provides design flood intervals which are not available in UNCODE and FFA. These intervals result from the considered uncertainties in data selection, probability distributions and distribution parameters.

In the previous research, design flood selection was based on either return periods according to flood frequency analysis or minimum total cost criterion according to cost-benefit analysis. Compared with previous studies, in our research, a robustness criterion is introduced. This criterion can allow decision makers to analyse the sensitivity of calculated total cost to the variations of uncertainties. This information is particularly useful because it can be incorporated in the decision making process to select the most robust design floods under deep uncertainties where data is scarce and distribution is unknown.

Sensitivity analysis can be conducted to explicitly evaluate the impact of uncertainty sources on decision making [Van-Waveren et al., 2000; Xu and Tung, 2009], however, the newly developed framework in our research can quantify the individual and interactive impacts of uncertainty sources in design flood estimation. As shown in this study, interactive influence among different uncertainties can be significant (e.g., interactive contribution in Biliu is up to 45.1%), and the importance of an uncertainty source can be underestimated without considering its interactions with other sources. It should be noted that the uncertainty contribution fractions obtained in this study are case-specific and might vary depending on the specific uncertainty sources included. Further research on more case studies is required to understand how the contribution fractions are affected by different uncertainty sources.

In addition to the epistemic uncertainties considered in this paper, other epistemic uncertainties can be explored using the new IPDF approach. For example, the number of distribution parameter sets used in the calculation of the sampling distributions and, as pointed out by Laio et al. [2009], the selection of significance level in PDF fitting. The higher the significance level, the more difficult the probability distributions can pass the test. In this study, with a significance level of 0.05, GLO and LN3 for the T1 data and LN3 for the AM data in Three Gorges could be rejected. Recall that these distributions cannot be rejected with a significance level of 0.01. The inclusion or exclusion of a specific distribution might have an impact on the lower and upper bounds of flood probabilities and thus on the ranges of total cost. However, this IPDF approach provides a quantitative means to measure the impacts and thus can better inform decision making.

It should be noted that the use of the GLUE approach to calculate the sampling distributions, and the use of a trial and error approach to define the uncertainty bounds of the probability distribution parameters are not necessary. Other approaches, such as importance sampling, Metropolis–Hastings algorithm, Gibbs sampling, and the use of regional information to define parameter bounds, can be applied as well.

## 7. Conclusions

Accurate estimation of design flood plays a crucial role in flood management: for example, the design of hydraulic structures. However, the estimation is influenced by various uncertainties: for example, aleatory and epistemic uncertainties. The state-of-the-art methodologies in design flood estimation did not account for the aleatory and epistemic uncertainties simultaneously, evaluating the overall benefits of design flood options, and providing quantitative information about aleatory and epistemic uncertainty contributions and their interactive contributions to design flood estimation uncertainty. These have posed a long term challenge to hydrologists and engineers. This paper presents a state-of-the-art progress to meet the challenge. A holistic and coherent framework to allow for realistic design flood estimation under multiple uncertainties is developed. To illustrate the proposed methodology, three case studies with 127-year, 104-year and 54-year flood data sets were employed. Three distributions were selected using the A-D test, and different data sets generated using AM and POT methods from historical flood data were considered. The major findings from this study are presented as follows.

First, an imprecise probabilistic approach for design flood estimation is proposed. This approach effectively combines aleatory and epistemic uncertainties from data, probability distribution functions, and parameters on the basis of the Dempster-Shafer theory. It also presents upper and lower bounds of total cost faced by decision makers when selecting a design flood.

Second, a robustness criterion for decision support in design flood selection is proposed. The design flood corresponding to the smallest minimum total cost can tolerate lower uncertainties, thus is not robust. With an increasing design flood magnitude, more uncertainties can be tolerated while still guaranteeing the calculated total cost varies only slightly, thus the robustness increases, but the minimum total cost increases as well. Between total cost and robustness, there is a clear trade-off which decision makers need to balance in the decision making process. This trade-off quantitatively provides the overall benefits of design flood options, which provides an objective tool for decision makers to balance conflicting concerns.

Third, the interactions among data, distributions and parameters are significant and have a much higher contribution than parameters to the uncertainty in total cost. The contributions of data, distributions and parameters to the overall uncertaity in total cost vary with flood magnitude. However, the contributions are almost the same for different return periods. This information implies that the overall uncertainty in estimated design floods could be underestimated if the interactions are disregarded, and therefore interactions should be considered in design flood estimation.

The approach proposed in this study could provide a blueprint for pragmatic flood frequency analysis under multiple epistemic uncertainties. Future research is encouraged to examine the applicability of the approach in other regions. In addition, climate change could influence flood frequency analysis, and future research should focus on incorporating climate change impacts into design flood estimation.

## Acknowledgements:

This study was supported by the National Natural Science Foundation of China (Grant No. 51320105010 and 51279021). The first author gratefully acknowledges the financial support provided by the China Scholarship Council. The authors are deeply indebted to editors, Dr

Francesco Serinaldi and another anonymous reviewer for their valuable time and constructive suggestions that greatly improved the quality of this paper. The data of Three Gorges were obtained from the China Three Gorges Corporation. The data of Biliu were obtained from the Biliu reservoir administration. The data of Harbin were obtained from the Harbin hydrology bureau. These data are available as in Supporting Information Data Set which includes Data Set S1, Data Set S2 and Data Set S3. Data Set S1 corresponds to Three Gorges; Data Set S2 corresponds to Biliu; Data Set S3 corresponds to Harbin.

## Appendix A: Probability density functions (*f*) and cumulative probability density functions (*F*)

The probability density functions (*f*) and cumulative probability density functions (*F*) used in this paper are given in Eqs. (A1)-(A3):

Generalized Extreme Value distribution

$$F(x)=\begin{cases} exp\left(-\left(1+kz\right)^{-1/k}\right), k\neq 0 \\ exp\left(-exp\left(-z\right)\right), k=0 \end{cases}, z\equiv\frac{x-\mu}{\sigma} \tag{A1}$$

where $k$, $\sigma$ >0 and $\mu$ are shape, scale and location parameter, respectively.

Generalized Logistic distribution

$$F(x)=\begin{cases} \dfrac{1}{1+\left(1+kz\right)^{-1/k}}, k\neq 0 \\ \dfrac{1}{1+exp\left(-z\right)}, k=0 \end{cases}, z\equiv\frac{x-\mu}{\sigma} \tag{A2}$$

where $k$, $\sigma$ >0 and $\mu$ are shape, scale and location parameter, respectively.

3-parameter Log-Normal distribution

$$f(x)=\frac{exp\left(-\frac{1}{2}\left(\frac{ln\left(x-\gamma\right)-\mu}{\sigma}\right)^2\right)}{x\sigma\sqrt{2\pi}} \tag{A3}$$

where $\mu$, $\sigma$ and $\gamma$ are, shape, scale and location parameter, respectively.

## References

Abrishamchi, A., Ebrahimian, A., Tajrishi, M., Mariño, M., (2005). Case Study: Application of Multicriteria Decision Making to Urban Water Supply. *Journal of Water Resources Planning and Management*, 131(4): 326-335. DOI:10.1061/(ASCE)0733-9496(2005)131:4(326)

Ahmad, M.I., Sinclair, C.D., Spurr, B.D., (1988). Assessment of flood frequency models using empirical distribution function statistics. *Water Resources Research*, 24(8): 1323-1328. DOI:10.1029/WR024i008p01323

Bao, Y., Tung, Y.-K., Hasfurther, V.R., (1987). Evaluation of uncertainty in flood magnitude estimator on annual expected damage costs of hydraulic structures. *Water Resources Research*, 23(11): 2023-2029. DOI:10.1029/WR023i011p02023

Beguería, S., (2005). Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology*, 303(1-4): 215-230. DOI:10.1016/j.jhydrol.2004.07.015

Ben-Haim, Y., 2006. Info-Gap Decision Theory: Decisions Under Severe Uncertainty, 2nd ed., Academic, London.

Beven, K., Binley, A., (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3): 279-298. DOI:10.1002/hyp.3360060305

Beven, K.J., Freer, J.E., (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4): 11–29. DOI:10.1016/S0022-1694(01)00421-8

Bhunya, P.K., Berndtsson, R., Jain, S.K., Kumar, R., (2013). Flood analysis using negative binomial and Generalized Pareto models in partial duration series (PDS). *Journal of Hydrology*, 497: 121-132. DOI:10.1016/j.jhydrol.2013.05.047

Bhunya, P.K., Singh, R.D., Berndtsson, R., Panda, S.N., (2012). Flood analysis using generalized logistic models in partial duration series. *Journal of Hydrology*, 420-421: 59-71. DOI:10.1016/j.jhydrol.2011.11.037

Blazkova, S., Beven, K., (2002). Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). *Water Resources Research*, 38(8): 14-1-14-14. DOI:10.1029/2001wr000500

Bodo, B., Unny, T.E., (1976). Model uncertainty in flood frequency analysis and frequency-based design. *Water Resources Research*, 12(6): 1109-1117. DOI:10.1029/WR012i006p01109

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., Schär, C., (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, 49(3): 1523-1536. DOI:10.1029/2011wr011533

Botto, A., Ganora, D., Laio, F., Claps, P., (2014). Uncertainty compliant design flood estimation. *Water Resources Research*, 50(5): 4242-4253. DOI:10.1002/2013wr014981

Calenda, G., Mancini, C.P., Volpi, E., (2009). Selection of the probabilistic model of extreme floods: The case of the River Tiber in Rome. *Journal of Hydrology*, 371(1-4): 1-11. DOI:10.1016/j.jhydrol.2009.03.010

Chowdhury, J.U., Stedinger, J.R., Lu, L.H., (1991). Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research*, 27(7): 1765-1776. DOI:10.1029/91wr00077

Coles, S., Pericchi, L.R., Sisson, S., (2003). A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273(1-4): 35-50. DOI:10.1016/s0022-1694(02)00353-0

Cunnane, C., (1979). A note on the Poisson assumption in partial duration series models. *Water Resources Research*, 15(2): 489-494. DOI:10.1029/WR015i002p00489

D'Agostino, R.B., Stephens, M.A., 1986; Table 4.32. Goodness-of-Fit-Techniques. Taylor & Francis.

Dempster, A.P., (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics*, 38(2): 325-339. DOI:10.2307/2239146

Di Baldassarre, G., Laio, F., Montanari, A., (2009). Design flood estimation using model selection criteria. *Physics and Chemistry of the Earth*, 34(10-12): 606-611. DOI:10.1016/j.pce.2008.10.066

Fawcett, L., Walshaw, D., (2015). Sea-surge and wind speed extremes: optimal estimation strategies for planners and engineers. *Stochastic Environmental Research and Risk Assessment*. DOI:10.1007/s00477-015-1132-3

Fu, G., Butler, D., Khu, S.-T., Sun, S.A., (2011). Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory. *Water Resources Research*, 47(2): W02534. DOI:10.1029/2009wr008944

Fu, G., Kapelan, Z., (2011). Fuzzy probabilistic design of water distribution networks. *Water Resources Research*, 47(5). DOI:10.1029/2010wr009739

Ganoulis, J., (2003). Risk‐based floodplain management: A case study from Greece. *International Journal of River Basin Management*, 1(1): 41-47. DOI:10.1080/15715124.2003.9635191

Haddad, K., Rahman, A., (2010). Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment*, 25(3): 415-428. DOI:10.1007/s00477-010-0412-1

Hall, J., Fu, G., Lawry, J., (2007). Imprecise probabilities of climate change: aggregation of fuzzy scenarios and model uncertainties. *Climatic Change*, 81(3-4): 265-281. DOI:10.1007/s10584-006-9175-6

Hall, J.W., (2003). Handling uncertainty in the hydroinformatic process. *Journal of Hydroinformatics*, 5(4): 215-232.

Hall, J.W., Lawry, J., (2004). Generation, combination and extension of random set approximations to coherent lower and upper probabilities. *Reliability Engineering & System Safety*, 85(1-3): 89-101. DOI:10.1016/j.ress.2004.03.005

Hall, J.W., Rubio, E., Anderson, M.G., (2004). Random sets of probability measures in slope hydrology and stability analysis. *Zeitschrift Angewandte Mathematik und Mechanik*, 84(10-11): 710-720. DOI:10.1002/zamm.200410146

Hine, D., Hall, J.W., (2010). Information gap analysis of flood model uncertainties and regional frequency analysis. *Water Resources Research*, 46(1). DOI:10.1029/2008wr007620

Jonkman, S.N., Brinkhuis-Jak, M., Kok, M., (2004). Cost-benefit analysis and flood damage mitigation in the Netherlands. *Heron*, 49: 95-111.

Kay, A.L., Davies, H.N., Bell, V.A., Jones, R.G., (2009). Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change*, 92(1-2): 41-63. DOI:10.1007/s10584-008-9471-4

Kidson, R., Richards, K.S., (2005). Flood frequency analysis: assumptions and alternatives. *Progress in Physical Geography*, 29(3): 392-410. DOI:10.1191/0309133305pp454ra

Kjeldsen, T.R., Prosdocimi, I., (2015). A bivariate extension of the Hosking and Wallis goodness-of-fit measure for regional distributions. *Water Resources Research*, 51(2): 896-907. DOI:10.1002/2014WR015912

Kuczera, G., (1999). Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resources Research,* 35(5): 1551-1557. DOI:10.1029/1999wr900012

Laio, F., (2004). Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*, 40(9): W09308. DOI:10.1029/2004wr003204

Laio, F., Di Baldassarre, G., Montanari, A., (2009). Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research*, 45(7). DOI:10.1029/2007wr006666

Le Coz, J., Renard, B., Bonnifait, L., Branger, F., Le Boursicaud, R., (2014). Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach. *Journal of Hydrology*, 509: 573-587. DOI:10.1016/j.jhydrol.2013.11.016

Lee, K.S., Kim, S.U., (2008). Identification of uncertainty in low flow frequency analysis using Bayesian MCMC method. *Hydrological Processes*, 22(12): 1949-1964. DOI:10.1002/hyp.6778

Lempert, R.J., Collins, M.T., (2007). Managing the risk of uncertain threshold responses: comparison of robust, optimum, and precautionary approaches. *Risk Analysis*, 27(4): 1009-26. DOI:10.1111/j.1539-6924.2007.00940.x

Liang, Z., Chang, W., Li, B., (2011). Bayesian flood frequency analysis in the light of model and parameter uncertainties. *Stochastic Environmental Research and Risk Assessment*, 26(5): 721-730. DOI:10.1007/s00477-011-0552-y

Lockhart, R.A., O'Reilly, F.J., Stephens, M.A., (1986a). Test for the extreme value and weibull distributions based on normalized spacings. *Naval Research Logistics Quarterly*, 33(3): 413-421. DOI:10.1002/nav.3800330307

Lockhart, R.A., O'Reilly, F.J., Stephens, M.A., (1986b). Tests of Fit Based on Normalized Spacings. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3): 344-352. DOI:10.2307/2345431

Madsen, H., Rasmussen, P.F., Rosbjerg, D., (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research*, 33(4): 747-757. DOI:10.1029/96WR03848

Maskey, S., Guinot, V., Price, R.K., (2004). Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. *Advances in Water Resources*, 27(9): 889-898. DOI:10.1016/j.advwatres.2004.07.001

Matrosov, E.S., Woods, A.M., Harou, J.J., (2013). Robust Decision Making and Info-Gap Decision Theory for water resource system planning. *Journal of Hydrology*, 494: 43-58. DOI:10.1016/j.jhydrol.2013.03.006

Merz, B., Thieken, A.H., (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, 309(1-4): 114-132. DOI:10.1016/j.jhydrol.2004.11.015

Merz, R., Blöschl, G., (2008). Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research*, 44(8). DOI:10.1029/2007WR006744

Nott, D.J., Marshall, L., Brown, J., (2012). Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection? *Water Resources Research*, 48(12): W12602. DOI:10.1029/2011wr011128

Olsen, J.R., Lambert, J.H., Haimes, Y.Y., (1998). Risk of Extreme Events Under Nonstationary Conditions. *Risk Analysis*, 18(4): 497-510. DOI:10.1111/j.1539-6924.1998.tb00364.x

Onoz, B., Bayazit, M., (2001). Effect of the occurrence process of the peaks over threshold on the flood estimates. *Journal of Hydrology*, 244(1-2): 86-96. DOI:10.1016/s0022-1694(01)00330-4

Palynchuk, B., Guo, Y., (2008). Threshold analysis of rainstorm depth and duration statistics at Toronto, Canada. *Journal of Hydrology*, 348(3-4): 535-545. DOI:10.1016/j.jhydrol.2007.10.023

Pandey, M.D., Van Gelder, P.H.A.J.M., Vrijling, J.K., (2004). Dutch case studies of the estimation of extreme quantiles and associated uncertainty by bootstrap simulations. *Environmetrics*, 15(7): 687-699. DOI:10.1002/env.656

Perreault, L., Bernier, J., Bobée, B., Parent, E., (2000). Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology*, 235(3–4): 221-241. DOI:10.1016/S0022-1694(00)00270-5

Prudhomme, C., Davies, H., (2009). Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate. *Climatic Change*, 93(1-2): 197-222. DOI:10.1007/s10584-008-9461-6

Qi, W., Zhang, C., Fu, G., Sweetapple, C., Zhou, H., (2016a). Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations. *Hydrology and Earth System Sciences*, 20(2): 903-920. DOI:10.5194/hess-20-903-2016

Qi, W., Zhang, C., Fu, G., Zhou, H., (2016b). Quantifying dynamic sensitivity of optimization algorithm parameters to improve hydrological model calibration. *Journal of Hydrology*, 533: 213-223. DOI:10.1016/j.jhydrol.2015.11.052

Rahman, A.S., Rahman, A., Zaman, M.A., Haddad, K., Ahsan, A., Imteaz, M., (2013). A study on selection of probability distributions for at-site flood frequency analysis in Australia. *Natural Hazards*, 69(3): 1803-1813. DOI:10.1007/s11069-013-0775-y

Reis, D.S., Stedinger, J.R., (2005). Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology*, 313(1-2): 97-116. DOI:10.1016/j.jhydrol.2005.02.028

Ribatet, M., Sauquet, E., Grésillon, J.-M., Ouarda, T.B.M.J., (2006). A regional Bayesian POT model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment*, 21(4): 327-339. DOI:10.1007/s00477-006-0068-z

Rosbjerg, D., (1985). Estimation in partial duration series with independent and dependent peak values. *Journal of Hydrology*, 76(1–2): 183-195. DOI:10.1016/0022-1694(85)90098-8

Rosbjerg, D., Madsen, H., Rasmussen, P.F., (1992). Prediction in partial duration series with generalized pareto-distributed exceedances. *Water Resources Research*, 28(11): 3001-3010. DOI:10.1029/92WR01750

Ross, J.L., Ozbek, M.M., Pinder, G.F., (2009). Aleatoric and epistemic uncertainty in groundwater flow and transport simulation. *Water Resources Research*, 45(12): W00B15. DOI:10.1029/2007wr006799

Rossi, G., Cancelliere, A., Giuliano, G., (2005). Case Study: Multicriteria Assessment of Drought Mitigation Measures. *Journal of Water Resources Planning and Management*, 131(6): 449-457. DOI:10.1061/(ASCE)0733-9496(2005)131:6(449)

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton, N. J.

Stedinger, J.R., (1983). Design events with specified flood risk. *Water Resources Research*, 19(2): 511-522. DOI:10.1029/WR019i002p00511

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. Handbook of Hydrology. McGraw-Hill, New York.

Steinschneider, S., Polebitski, A., Brown, C., Letcher, B.H., (2012). Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change. *Water Resources Research*, 48(11): W11525. DOI:10.1029/2011wr011318

Stephens, M.A., (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, 69(347): 730-737. DOI:10.1080/01621459.1974.10480196

Stephens, M.A., (1976). Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters. 357-369. DOI:10.1214/aos/1176343411

Stephens, M.A., (1977). Goodness of Fit for the Extreme Value Distribution. *Biometrika*, 64(3): 583-588. DOI:10.2307/2345336

Stephens, M.A., (1979). Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function. *Biometrika*, 66(3): 591-595. DOI:10.2307/2335180

Su, H.-T., Tung, Y.-K., (2013a). Flood-Damage-Reduction Project Evaluation with Explicit Consideration of Damage Cost Uncertainty. *Journal of Water Resources Planning and Management*, 139(6): 704-711. DOI:10.1061/(asce)wr.1943-5452.0000291

Su, H.-T., Tung, Y.-K., (2013b). Incorporating uncertainty of distribution parameters due to sampling errors in flood-damage-reduction project evaluation. *Water Resources Research*, 49(3): 1680-1692. DOI:10.1002/wrcr.20116

Tanaka, S., Takara, K., 2002. A study on threshold selection in POT analysis of extreme floods, Extremes of the Extremes: Extraordinary Floods. IAHS Publication. Int Assoc Hydrological Sciences, Wallingford, pp. 299-304.

Tonon, F., Bernardini, A., Mammino, A., (2000). Reliability analysis of rock mass response by means of Random Set Theory. *Reliability Engineering & System Safety*, 70(3): 263-282. DOI:10.1016/S0951-8320(00)00059-4

Tung, Y.-K., Mays, L.W., (1981). Optimal risk-based design of flood levee systems. *Water Resources Research*, 17(4): 843-852. DOI:10.1029/WR017i004p00843

Van-Waveren, R.H., Groot, S., Scholten, H., Geer, F.C.v., Wosten, J.H.M., Koeze, R.D., Noort, J.J., 2000. Good Modelling Practice Handbook. STOWA, Utecht, The Netherland.

Viglione, A., Merz, R., Salinas, J.L., Blöschl, G., (2013). Flood frequency hydrology: 3. A Bayesian analysis. *Water Resources Research*, 49(2): 675-692. DOI:10.1029/2011wr010782

Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., Verstraten, J.M., (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, 41(1): W01017. DOI:10.1029/2004wr003059

Walley, P., 1991. Statistical Reasoning with Imprecise Probabilities. Taylor & Francis.

Wilby, R.L., Harris, I., (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, 42(2): W02419. DOI:10.1029/2005wr004065

Wood, E.F., Rodríguez-Iturbe, I., (1975a). A Bayesian approach to analyzing uncertainty among flood frequency models. *Water Resources Research*, 11(6): 839-843. DOI:10.1029/WR011i006p00839

Wood, E.F., Rodríguez-Iturbe, I., (1975b). Bayesian inference and decision making for extreme hydrologic events. *Water Resources Research*, 11(4): 533-542. DOI:10.1029/WR011i004p00533

Xu, Y.-P., Booij, M.J., Tong, Y.-B., (2009). Uncertainty analysis in statistical modeling of extreme hydrological events. *Stochastic Environmental Research and Risk Assessment*, 24(5): 567-578. DOI:10.1007/s00477-009-0337-8

Xu, Y., Tung, Y., (2009). Decision Rules for Water Resources Management under Uncertainty. *Journal of Water Resources Planning and Management*, 135(3): 149-159. DOI:doi:10.1061/(ASCE)0733-9496(2009)135:3(149)

Table 1 Flood discharge series generated using different thresholds and Annual Maximum (AM) approach from flood records in three cases

| Three Gorges | | | Harbin | | | Biliu | | |
|---|---|---|---|---|---|---|---|---|
| Symbol | Threshold level ($m^3/s$) | Number of data | Symbol | Threshold level ($m^3/s$) | Number of data | Symbol | Threshold level ($m^3/s$) | Number of data |
| T1 | 52,000 | 270 | T1 | 6500 | 264 | T1 | 500 | 105 |
| T2 | 53,000 | 229 | T2 | 7000 | 227 | T2 | 700 | 62 |
| T3 | 54,000 | 190 | T3 | 7800 | 178 | T3 | 737 | 54 |
| T4 | 55,000 | 169 | T4 | 8500 | 124 | T4 | 1100 | 19 |
| T5 | 56,000 | 135 | T5 | 9000 | 104 | T5 | 1300 | 13 |
| T6 | 56,300 | 127 | T6 | 9500 | 88 | - | - | - |
| T7 | 57,000 | 109 | T7 | 10,000 | 77 | - | - | - |
| T8 | 58,000 | 85 | - | - | - | - | - | - |
| AM | - | 127 | AM | - | 104 | AM | - | 54 |

Table 2 Anderson-Darling test results of three probability distributions in three cases

| | Three Gorges | | | Harbin | | | Biliu | | |
|---|---|---|---|---|---|---|---|---|---|
| | GEV | GLO | LN3 | GEV | GLO | LN3 | GEV | GLO | LN3 |
| T1 | 1.57 | 2.46 | 2.02 | 1.29 | 1.97 | 1.39 | 0.34 | 0.41 | 0.47 |
| T2 | 0.98 | 1.56 | 1.46 | 1.49 | 2.18 | 1.33 | 0.73 | 0.87 | 0.23 |
| T3 | 0.62 | 1.14 | 0.56 | 0.90 | 1.07 | 0.59 | 0.60 | 0.73 | 0.33 |
| T4 | 0.91 | 1.36 | 1.21 | 1.05 | 0.95 | 1.29 | 0.24 | 0.27 | 0.29 |
| T5 | 0.54 | 0.91 | 0.45 | 0.89 | 0.77 | 1.06 | 0.18 | 0.20 | 0.18 |
| T6 | 0.63 | 1.00 | 0.49 | 0.66 | 0.70 | 0.93 | - | - | - |
| T7 | 0.67 | 0.99 | 0.51 | 0.90 | 1.07 | 0.59 | - | - | - |
| T8 | 0.47 | 0.68 | 0.46 | - | - | - | - | - | - |
| AM | 0.30 | 0.77 | 2.23 | 0.32 | 0.44 | 0.35 | 0.25 | 0.25 | 0.38 |

Note: GEV represents Generalized Extreme Value distribution; GLO represents Generalized Logistic distribution; LN3 represents 3-parameter Log-Normal distribution. AM represents annual maximum approach; the symbols from T1 to T8 represent different thresholds.

Table 3 Design flood estimates from the newly proposed Imprecise Probabilistic Design Flood (IPDF) approach, the UNcertainty COmpliant DEsign (UNCODE) approach proposed by [Botto et al., 2014] and Flood Frequency Analysis (FFA).

| | IPDF($m^3/s$) | UNCODE($m^3/s$) | FFA($m^3/s$) |
|---|---|---|---|
| Three Gorges | [73,900; 113,200] | 73,700 | 73,400 |
| Harbin | [23,800; 63,800] | 23,800 | 22,800 |
| Biliu | [5700; 17,300] | 5700 | 5700 |

Note: in IPDF, the design flood intervals correspond to the minimum total costs in the lower and upper total cost bounds; in UNCODE, the design floods correspond to the minimum total costs among all the total cost curves shows in Fig. 7; in FFA, the design floods correspond to the minimum values among all the data set and distribution combinations.

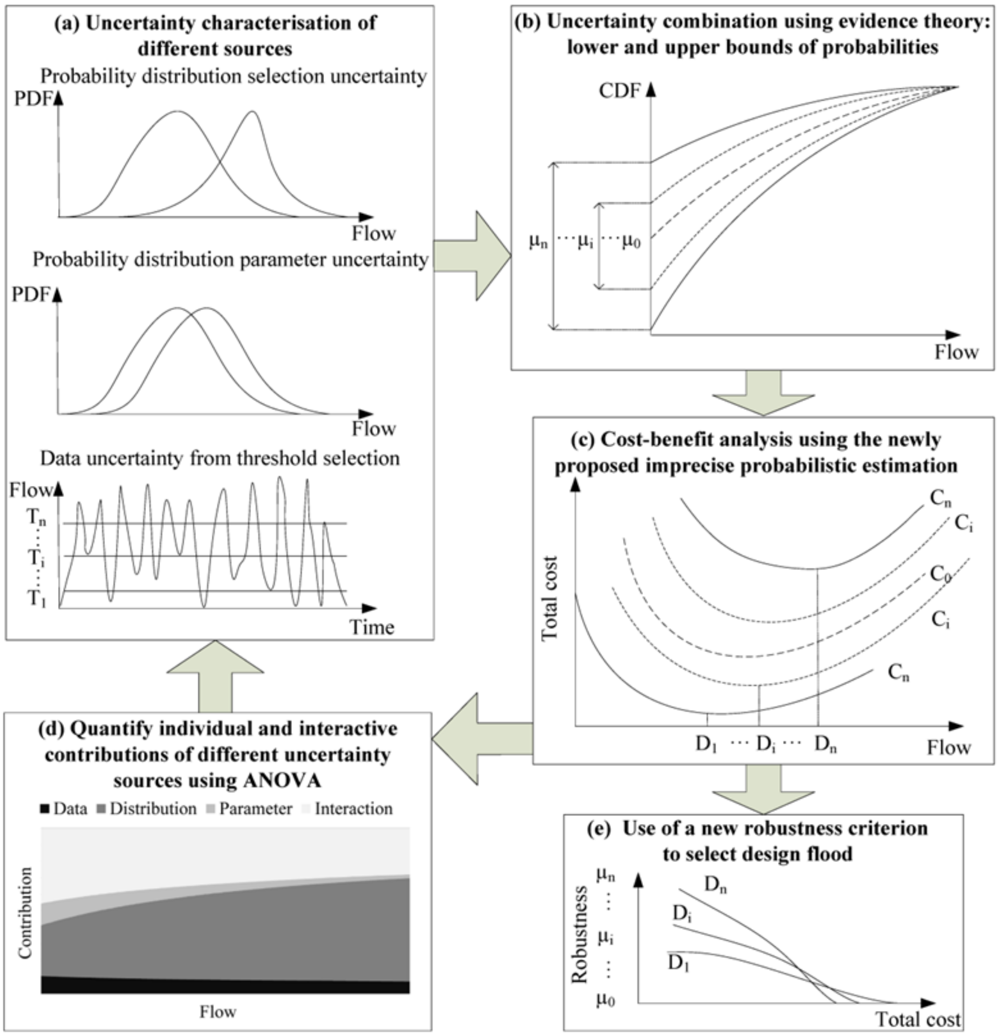


Fig. 1 Diagrammatic representation of the proposed imprecise probabilistic framework for design flood estimation with epistemic uncertainties. Three thresholds $T_1$, $T_i$ and $T_n$ are shown for illustration; ANOVA represents the analysis of variance approach; $\mu_0$, $\mu_i$ and $\mu_n$ represent three different uncertainty levels; PDF represents Probability Distribution Functions; CDF represents Cumulative Distribution Function; $D_1$, $D_i$ and $D_n$ represent three different flood values; $C_1$, $C_i$ and $C_n$ represent three different total cost values.

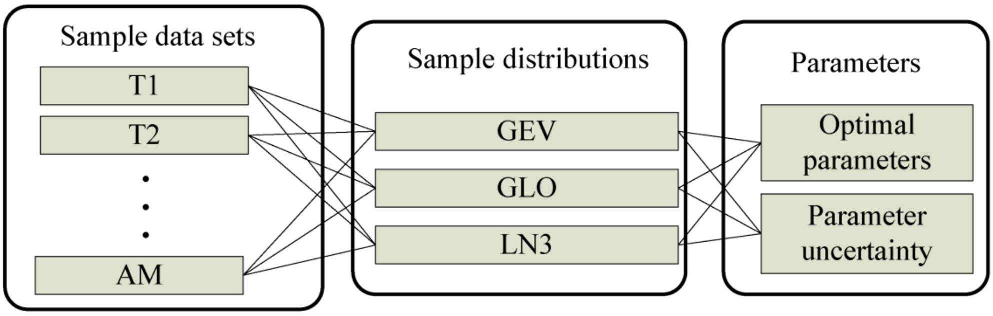


Fig. 2 The combinations of data sets, distributions and parameters. T1, T2 and AM represent three selected data sets.

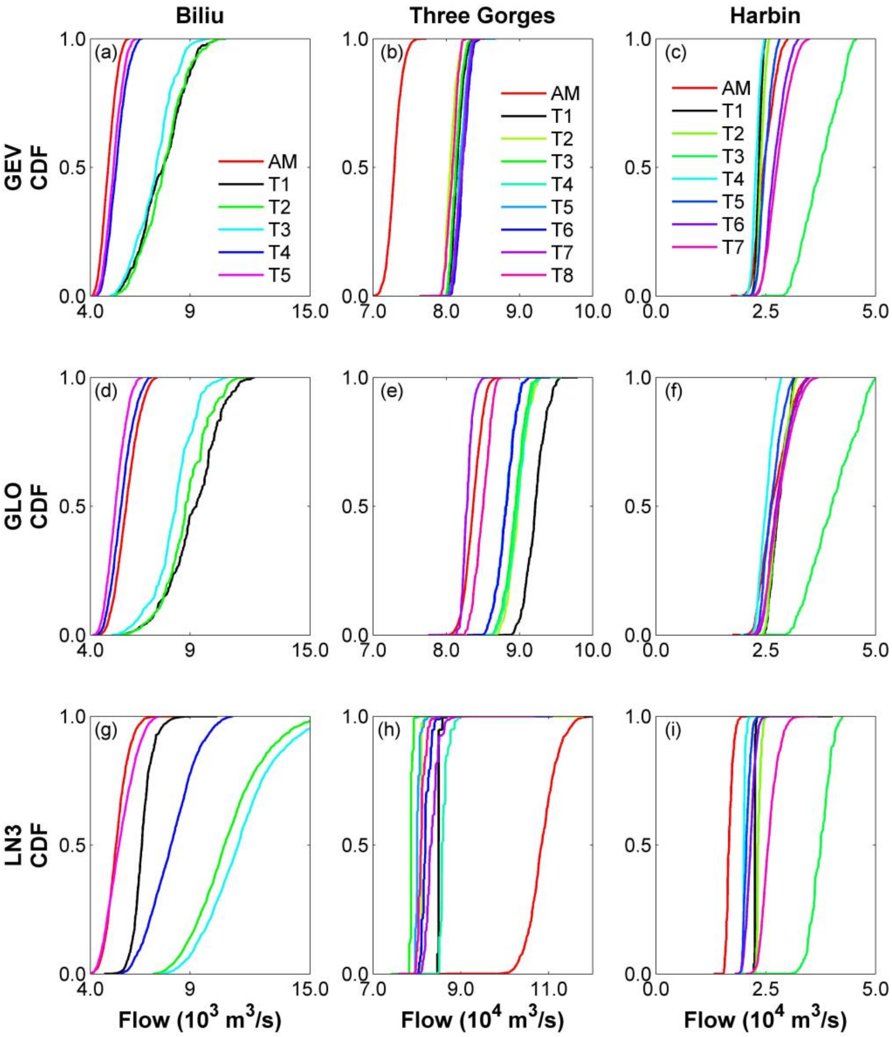


Fig. 3 Sampling distributions of a specific design flood obtained using the posterior parameter distributions of GEV, GLO and LN3 in Biliu (a, d and g), Three Gorges (b, e and h) and Harbin (c, f and i).

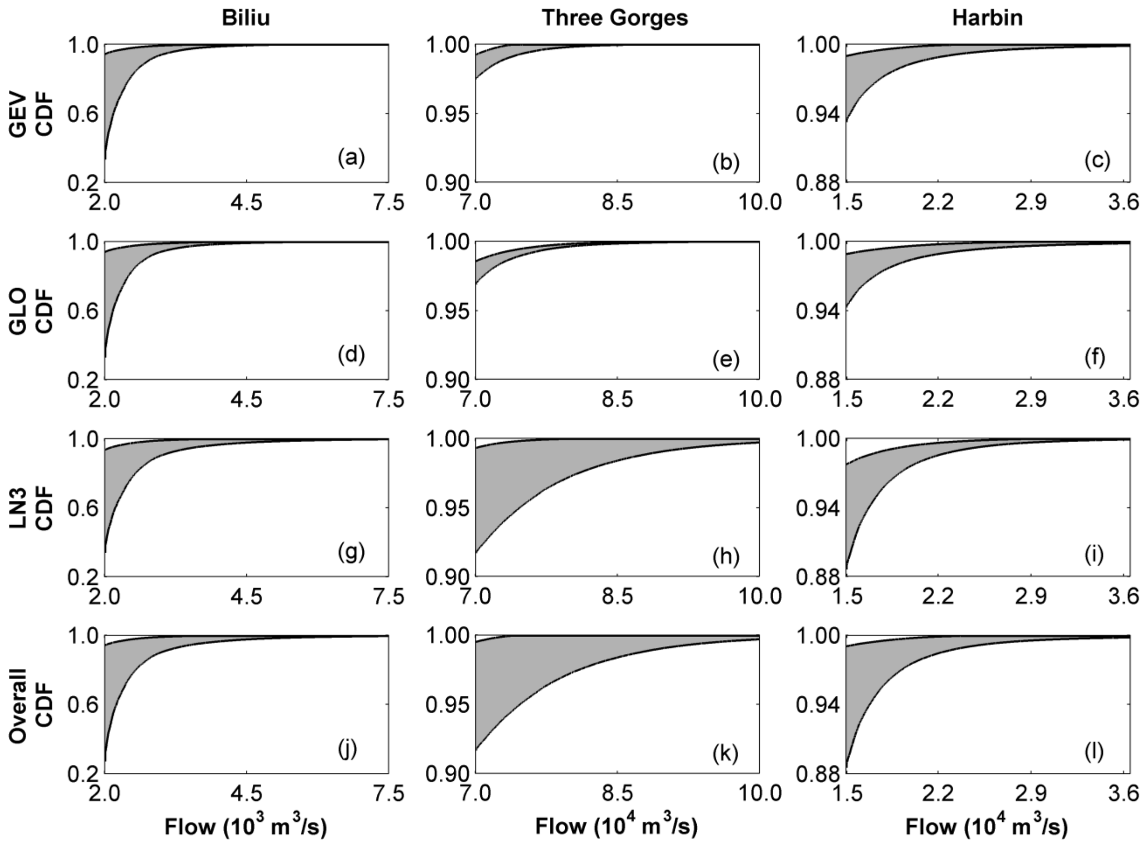


Fig. 4 Lower and upper bounds of cumulative probabilities of flood for GEV, GLO, LN3 and combined distributions in Biliu (a, d, g and j), Three Gorges (b, e, h and k) and Harbin (c, f, i and l) respectively. For each individual probability distribution (GEV, GLO and LN3), the probability of each flood value is calculated based on predictive distributions and the intervals are derived from the selected data sets. The combined CDFs (j, k and l) result from selected data sets, probability distributions and calculated predictive distributions related to parameter uncertainty.

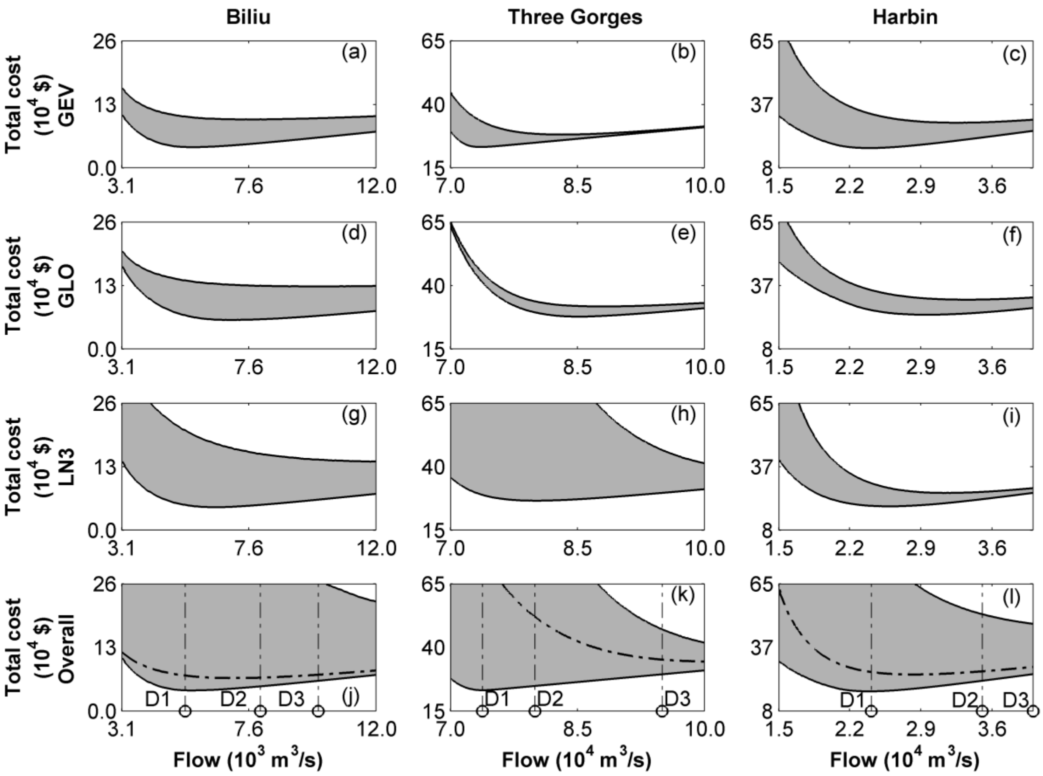


Fig. 5 Lower and upper total cost bounds for 500-year, 1000-year and 500-year design flood in Biliu (a, d, g and j), Three Gorges (b, e, h and k) and Harbin (c, f, i and l) respectively. D1, D2 and D3 represent three design floods. For each individual probability distribution (GEV, GLO and LN3), the uncertainty in total cost results from data selection uncertainty and probability distribution parameter uncertainty; the overall uncertainty in total cost results from selected data sets, probability distributions (GEV, GLO and LN3) and calculated parameter uncertainty.

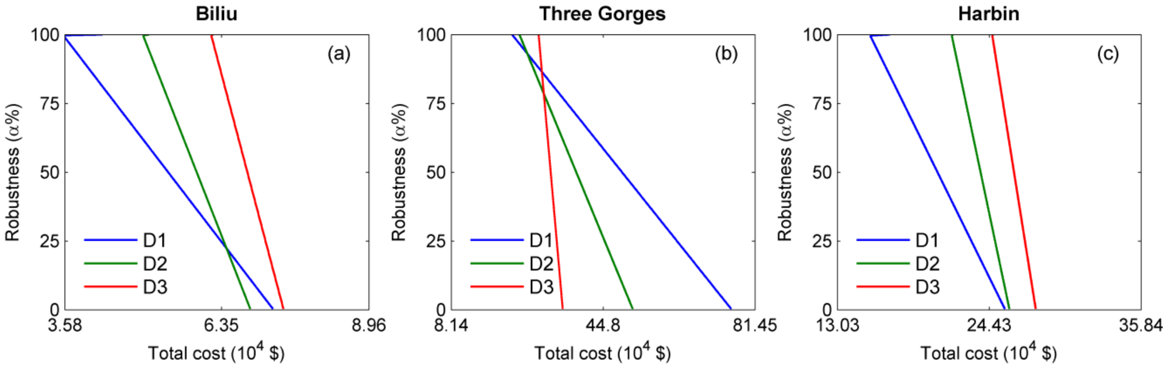


Fig. 6 Robustness of design floods under different uncertainty levels in the three case studies: Biliu (a), Three Gorges (b) and Harbin (c). Each curve represents a design flood, and its slope describes the variation of minimum total cost with uncertainty. D1, D2 and D3 represent three design flood values.

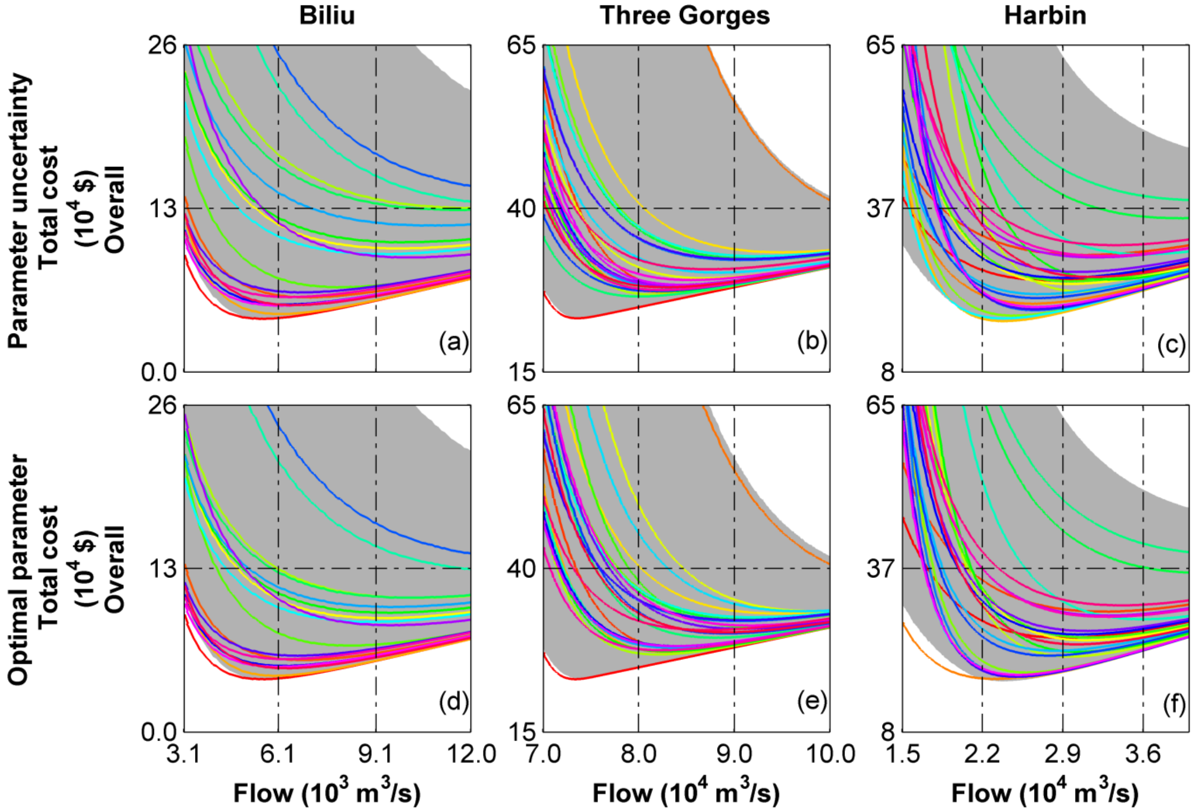


Fig. 7 Total cost curves of different data and probability distribution combinations (the solid lines) under two cases: with parameter uncertainty using predictive probability distributions (a, b and c) and without parameter uncertainty using the estimated optimal parameters (d, e and f); total cost uncertainty bounds (the grey areas) resulting from data selection uncertainty, probability distribution uncertainty and distribution parameter uncertainty.

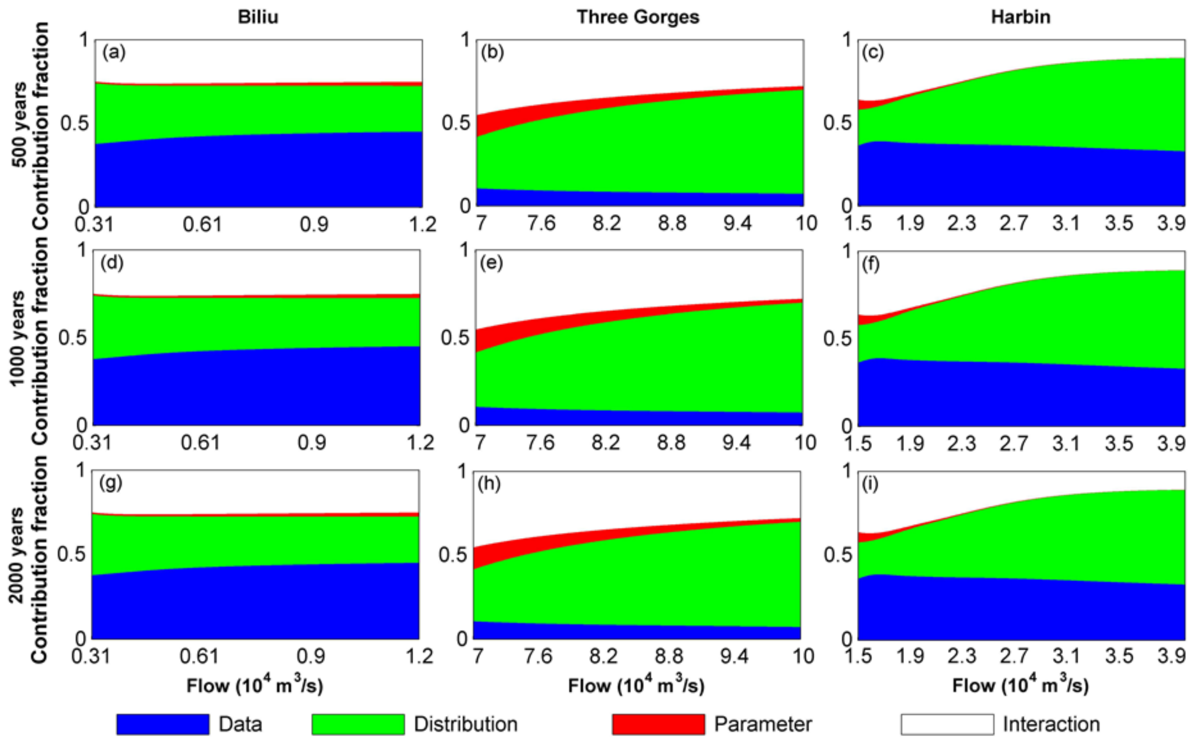


Fig. 8 Contributions of uncertainty sources to the total costs of three return periods: 500, 1000 and 2000 years in Biliu (a, d and g), Three Gorges (b, e and h) and Harbin (c, f and i). The contributions of uncertainty sources are represented by the widths of the relevant strips varying with flood values on *x*-axis.