Applied Mathematical Modelling 102 (2022) 371–384

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Tracing design flood hydrograph uncertainty in reservoir flood control system

Aijun Guo[a], Jianxia Chang[a,*], Yimin Wang[a], Bin Wu[a], Yunyun Li[b]

[a] State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an 710048, China
[b] Mianyang Teachers' College, Mianyang 621000, China

ARTICLE INFO

Article history:
Received 30 May 2021
Revised 8 September 2021
Accepted 10 September 2021
Available online 29 September 2021

Keywords:
Design flood hydrograph uncertainty
Optimal reservoir operation for flood control
Uncertainty quantification and propagation
Information theory

ABSTRACT

Increasing studies have highlighted that the remarkable inherent uncertainty in design flood hydrograph (DFH) can potentially undermine flood management decisions. In order to quantitatively trace the propagation of DFH uncertainty in reservoir flood control system, we propose a novel methodological framework including three corn parts. First, a copula-based DFH estimation model integrating Bayes' theorem is presented to estimate DFH under model parameter uncertainty. Second, we perform an optimal reservoir operation model for flood control ($OROM_{FC}$) with uncertain DFH as input variable to derive reservoir flood control operations (i.e. output variable). Third, an information theory-based model is designed to trace the DFH uncertainty propagation in reservoir flood control system. A reservoir flood control system in the Han River basin in China is selected as case study. Related results indicate that uncertainty in reservoir flood control operations reduces in comparison with the remarkable uncertainty in DFH due to the performance of the $OROM_{FC}$. Specifically, uncertainty in reservoir flood control operations in periods close to peak flow is much smaller compared with that in other periods. The phenomenon further highlights the importance of reservoir flood control operations during periods prior to peak flow. Additionally, we find that uncertainty in flood peak of DFH is the dominant factor affecting reservoir flood control operations compared with that in flood volume of DFH. Moreover, we explore the impact of reservoir flood control capacity on reservoir flood operations in the context of DFH uncertainty. An interesting linear expression is found and fitted for identifying design flood events inducing reservoir overtopping under specific reservoir flood control capacity.

© 2021 Published by Elsevier Inc.

1. Introduction

1.1. Confronting uncertainty in design flood hydrograph (DFH) and its impact on reservoir flood control operation

Uncertainty, such as sampling uncertainty, model uncertainty and parameter uncertainty, not desirably but inevitably, accompanies design flood estimation [1–5]. As crucial input of reservoir flood control system (RFCS), the uncertain design flood is bound to influence the system output, i.e., reservoir flood control operations (such as reservoir release or water level in different periods) [6,7]. Exploring the uncertainties in the aforementioned input and output variables in RFCS and

* Corresponding author.
E-mail address: chxiang@xaut.edu.cn (J. Chang).

https://doi.org/10.1016/j.apm.2021.09.012
0307-904X/© 2021 Published by Elsevier Inc.

disclosing the propagation of uncertainty from input to output are greatly beneficial and indispensable for flood control decision-making under uncertainty. Above considerations motivate our present work. To carry out aforementioned analysis, following three interconnected parts should be implemented sequentially: (1) estimating design flood in the context of uncertainty; (2) establishing optimal reservoir operation model for flood control ($OROM_{FC}$); and (3) designing methodological framework for uncertainty analysis in RFCS. Among them, issues involved in the second part have been investigated quite intensively to date. Therefore, in this paper, we restrict our attention to the first and third parts.

As for the design flood estimation, over the last 20 years, copula-based joint distribution functions have received widespread attention [8–11]. A significant advantage of copula-based approach is that the copula function allows for modeling the inherent dependence between (among) flood-related features, such as the flood peak (Q_p), flood volume (V) [12–14]. Therefore, this paper focuses on copula-based multivariate flood frequency analysis approach. However, noteworthy is that multiple uncertainties come along with design flood estimation. As highlighted by Montanari [15], Serinaldi and Kilsby [16], Michailidi and Bacchi [7], these uncertainties should be carefully defined and taken into account in flood control decision-making, hydrologic engineering design, etc; otherwise, it would induce incomplete or even misleading conclusions in practical flood control. Given above, Serinaldi [2] performed pioneer work putting light on multivariate design flood estimation under uncertainty. Following his work, a few studies have been performed to investigate the pair of design variable quantiles (Q_p, V) under uncertainty in recent years [7,10,13,17]. These studies all highlight that overlooking the uncertainty in design flood would easily result in deceiving results in flood management.

Apart from the design variable quantiles (Q_p, V), hydrograph shape is another indispensable element in practical flood control [11,18]. Different hydrograph shapes may result in different costs of hydraulic structures and influence flood control decision. In view of this, Serinaldi and Grimaldi [18] proposed one dimensionless hydrograph shape approach to estimate DFH in the form of a Beta density with design variables Q_p and V. Brunner et al. [11] estimated DFHs with the concepts developed by Serinaldi and Grimaldi [18] and further considered the dependence between Q_p and V. Above works both elucidate that design flood variables Q_p and V greatly determine the shape of flood hydrograph. Along this line of consideration, DFHs would be uncertain in the context of Q_p and V uncertainty. Therefore, motivated by work of Brunner et al. [11], we present a statistical approach for DFHs estimation taking into account both the uncertainty and dependence of design flood variables Q_p and V.

Furthermore, it is worthy of note that inputting the uncertain DFH into RFCS generates uncertain reservoir flood control operations consequently [19]. Given this, the chief issue that should be solved is to design one framework to quantify the uncertainty in DFH (i.e. input variable in RFCS) and reservoir flood control operations (i.e. output variable in RFCS), and investigate uncertainty propagation from input to output in RFCS. In this paper, we employ Shannon Entropy based information theory [20] to develop a methodological framework for aforementioned uncertainty analysis. To date, information theory has been extensively applied to characterize uncertainty in various fields [21–24]. Despite these progresses, research on information theory-based uncertainty analysis is still under investigation as well in methodological aspects as in concrete applications. Gong et al. [25] presented a novel information theoretic approach to estimate epistemic and aleatory uncertainties in hydrologic modeling; Goodwell and Kumar [26,27] developed an information partitioning approach to identify nonlinear interactions between source and target variables. Undoubtedly, previous studies provide tremendous potential for uncertainty quantification analysis through information theory. Specifically, uncertainty quantification analysis is concerned with how we measure the uncertainties in input and output variables, how the uncertainty in input variable is propagated to the output variable through one model, etc. However, to date, there is still a lack of a comprehensive method for performing aforementioned uncertainty quantification analysis.

1.2. Objectives

Given above, this paper aims to establish a methodological framework for aforementioned uncertainty quantification analysis in RFCS triggered by DFH uncertainty. In addition to the optimal reservoir flood control model, the methodological framework integrates two closely related steps, i.e. (1) estimating DFH under uncertainty and (2) performing uncertainty quantification analysis in RFCS through information theory-based approach. Here, uncertainty in DFH involves uncertainties in Q_p, V and time to peak (t_p). In particular, we employ Bayes' theorem to explore uncertainty in design variable quantiles (Q_p, V) stemming from parameter uncertainty of copula-based design flood estimation model. Neglecting uncertainties in probability distribution model and parameter estimation method is primarily due to their relatively less impact in comparison with parameter uncertainty [10]. t_p uncertainty is considered through setting discrete schemes ranging from whole period according to practical flood characteristics in study region.

1.3. Outline of the presentation

This paper is organized as follows. Section 2 presents the methods used in this study. Section 3 describes the study area and data. Results and discussion are presented in Section 4. Section 5 summarizes the main conclusions drawn from the study.

2. Methodology

2.1. Methodological framework

This section briefly describes the outline of the proposed methodology. First, the DFH under uncertainty is estimated by an integrated framework. Specifically, the integrated framework includes two parts, i.e., copula-based bivariate design flood estimation model and synthetic DFH estimation model. Second, we construct the $\mathrm{OROM_{FC}}$. Following, the uncertain synthetic DFHs are inputted to $\mathrm{OROM_{FC}}$. Reservoir flood control operations (i.e., reservoir release or water level) are outputs. Finally, an information theoretic approach is designed to quantify uncertainties in DFH and reservoir flood control operations and uncover uncertainty propagation from input to output. Methods mentioned above will be described in detail below.

2.2. Copula-based DFH estimation model

2.2.1. Copula-based bivariate design flood estimation model

According to Sklar's theorem, given random continuous variables X_1, X_2, ..., X_d ($d \geq 2$) with marginal distribution of $u_1=F_1(X_1|\alpha_1)$, $u_2=F_2(X_2|\alpha_2)$, ..., $u_d = F_d(X_d|\alpha_d)$, the copula-based d-dimensional multivariate joint distribution function $H_C(\bullet)$ is defined as

$$\mathrm{H}_C(X_1,\cdots,X_\mathrm{d}|\Theta) = \mathrm{C}((\mathrm{F}_1(X_1|\alpha_1),\cdots,\mathrm{F}_\mathrm{d}(X_\mathrm{d}|\alpha_\mathrm{d}))|\theta) = \mathrm{C}(\mathrm{u}_1,\cdots,\mathrm{u}_\mathrm{d}|\theta) \tag{1}$$

where $\Theta = (\alpha_1,\cdots,\alpha_\mathrm{d},\theta)$ means the parameter vector of $H_C(\bullet)$; θ is parameter of copula function $\mathrm{C}(\mathrm{u}_1,\cdots,\mathrm{u}_\mathrm{d}|\theta)$. The joint probability density $h_C(\bullet)$ takes the form

$$\mathrm{h}_C(X_1,\cdots,X_\mathrm{d}|\Theta) = \mathrm{c}(\mathrm{u}_1,\cdots,\mathrm{u}_\mathrm{d}|\theta)\prod_{\mathrm{i}=1}^{\mathrm{d}}\mathrm{f}_{X_\mathrm{d}}(X_\mathrm{d}|\alpha_\mathrm{d}) \tag{2}$$

where $\mathrm{c}(\mathrm{u}_1,\cdots,\mathrm{u}_\mathrm{d}|\theta)$ is the probability density of copula function $\mathrm{C}(\mathrm{u}_1,\cdots,\mathrm{u}_\mathrm{d}|\theta)$; $\mathrm{f}_{X_\mathrm{d}}(X_\mathrm{d}|\alpha_\mathrm{d})$ is probability density distribution of $F_d(X_d|\alpha_d)$.

Bayes' theorem-based differential evolution adaptive metropolis algorithm [28] is employed to uncover parameter uncertainty of joint distribution function $\mathrm{H}_C(X_1,\cdots,X_\mathrm{d}|\Theta)$. Mathematically,

$$\mathrm{p}(\Theta|X) = \frac{\mathrm{f}(\mathrm{x}|\Theta)\pi(\Theta)}{\int \mathrm{f}(\mathrm{x}|\Theta)\pi(\Theta)\mathrm{d}(\Theta)} \propto \mathrm{f}(\mathrm{x}|\Theta)\pi(\Theta) \tag{3}$$

where $\mathrm{p}(\Theta|X)$ and $\pi(\Theta)$ represent the posterior and prior parameter probability density distribution respectively; $\mathrm{f}(\mathrm{x}|\Theta)$ means the likelihood function quantifying the probability of observations $\boldsymbol{X}$ given Θ.

In terms of copula function selection, large numbers of copula functions have been developed, which can be generally classified into elliptical, Archimedean, extreme value and other copula families [29–31]. Among them, Archimedean copulas (such as Frank, Clayton and Gumbel copulas) have received widespread attention within the hydrologic community, because they can be easily generated and effectively capture different types of dependence structures. Specifically, the Clayton copula and Gumbel copula can characterize negative and positive tail dependence between variables of interest, respectively. The Frank copula exhibits no tail dependence. In this study, the inherent positive dependence between flood peak and flood volume indicates Gumbel copula is the appropriate one. The Gumbel copula density is defined as

$$\begin{aligned}\mathrm{c}^{\mathrm{Gumbel}}(\mathrm{u}_1,\mathrm{u}_2|\theta) = &\left\{\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{\left(\frac{1-\theta}{\theta}\right)^2} + (\theta-1)\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{\frac{1-2\theta}{\theta}}\right\} \\ &\times \frac{(\ln u_1 \ln u_2)^{\theta-1}}{u_1 u_2}\exp\left\{-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{\frac{1}{\theta}}\right\}\end{aligned} \tag{4}$$

As for the marginal distribution selection, five popularly used distributions in hydrology is selected as candidates, i.e., generalized extreme value distribution, Pearson type III distribution, lognormal distribution, normal distribution and gamma distribution. The Kolmogorov–Smirnov (K–S) test is employed to perform goodness-of-fit test. The corrected Akaike information criterion and Bayes information criterion are used together to select the most appropriate distribution.

Once the copula-based joint distribution is constructed, following step is to estimate multivariate design flood event for a given return period. This paper applies a most-likely design event estimation approach (proposed by Salvadori et al. [8]) to determine the multivariate design flood event. The most-likely design realization δ_{ML} refers to the joint design value with maximum joint probability density lying on critical layers L_s^F for a critical level s. δ_{ML} is defined as

$$\delta_{ML} = \underset{(x,y)\in L_s^F}{\arg\max}\, h_C(x,y|\Theta), s \in (0,1) \tag{5}$$

Here, L_s^F is defined as $L_s^F = \{(x,y) : H_C(x,y|\Theta) = s\}$. Uncertain parameter vector $\Theta = (\alpha_1,\cdots,\alpha_\mathrm{d},\theta)$in $h_C(x,y|\Theta)$ will produce uncertain outputs, i.e. uncertain design flood in this paper. As for the type of joint return period, the Kendall's return period T_k (Salvadori and De Michele [32]) is employed in this study.

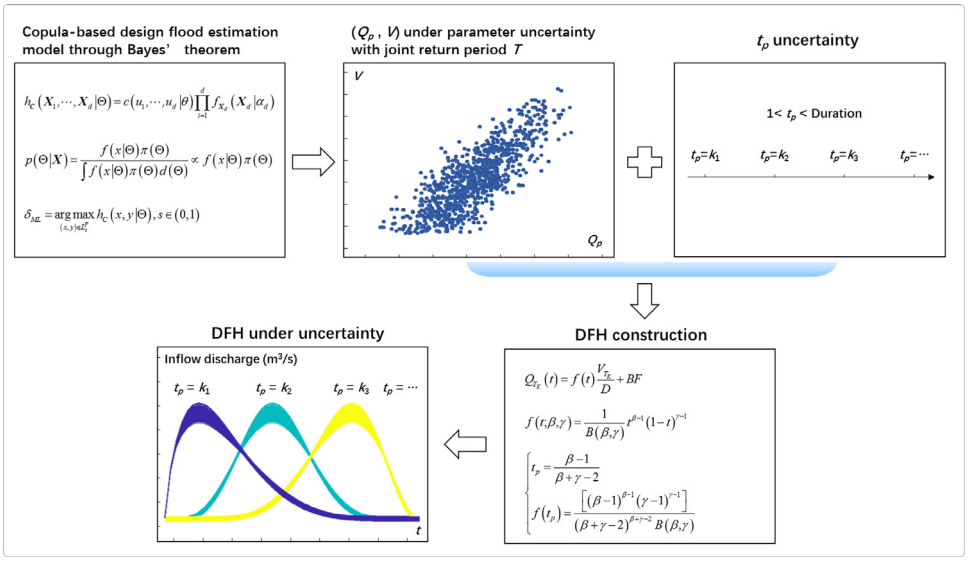

Fig. 1. Framework for copula-based DFH estimation model under uncertainty.

2.2.2. Synthetic DFH estimation model

Apart from Q_p and V, hydrograph shape is an important characteristic of a flood event. Different hydrological shapes may result in different costs of hydraulic structures and different flood control schemes. In this study, we perform the synthetic DFH estimation model (Brunner et al. [11]) to derive design hydrographs based on design variable quantiles $Q_{p_{T_K}}$ and V_{T_K}. Specifically, the synthetic design hydrograph employs the probability density function PDF ($f(t)$) to model the dimensionless hydrographs. Then, the synthetic design hydrograph is given by

$$Q_{T_K}(t) = f(t)\frac{V_{T_K}}{D} + BF \tag{6}$$

where D and BF represent the duration and base flow of a flood event. In one word, the synthetic design hydrographs ($Q_{T_K}(t)$) for a given return period can be expressed by $f(t)$ times the mean discharge (i.e. V_{T_K}/D) plus BF. Hence, D is set as a fixed value according to historical flood events in study region.

Following Serinaldi and Grimaldi [18], the beta PDF ($f(t)$) with support [0,1] is employed to model the dimensionless hydrograph, which is given by

$$f(t;\beta,\gamma) = \frac{1}{B(\beta,\gamma)}t^{\beta-1}(1-t)^{\gamma-1} \tag{7}$$

where $\beta>0$, $\gamma>0$; $B(\beta, \gamma)$ represents the complete beta function. Parameters β and γ are linked to t_p and related $f(t_p)$ by following relationships

$$\begin{cases} t_p = \frac{\beta-1}{\beta+\gamma-2} & \text{(8a)} \\ f(t_p) = \frac{\left[(\beta-1)^{\beta-1}(\gamma-1)^{\gamma-1}\right]}{(\beta+\gamma-2)^{\beta+\gamma-2}B(\beta,\gamma)} & \text{(8b)} \end{cases}$$

On the basis of Eq. (8), $Q_{p_{T_K}}$ can be expressed as

$$Q_{p_{T_K}} = f(t_p)\frac{V_{T_K}}{D} + BF \tag{9}$$

Here, the sum of $Q_{T_K}(t)$ in different periods should be equal to V_{T_K}. Based on this constraint and Eqs. (8a),(9), the unknown parameters β, γ and BF are obtained by using sequential quadratic programming optimization algorithm. Then, substituting β, γ and BF into Eq. (6), the synthetic DFH can be obtained.

For a more thorough understanding, the schematic diagram of the copula-based DFH estimation model is shown in Fig. 1.

2.3. Optimal reservoir operation model for flood control

In practice, different optimal objectives of reservoir flood control operation are generally set according to the flood event magnitudes. For example, we primarily consider ensuring the safety of the dam and its upstream area given high magnitude flood event. In contrast, we consider ensuring the safety at a downstream control point on condition of low magnitude flood event. This paper focuses on deriving optimal flood control operations under high magnitude flood event. Hence, ensuring

the safety of the dam and its upstream area is adopted as objective. The essence of this objective is to encourage reservoir operations to keep the reservoir water level as low as possible. Therefore, this objective F_1 is defined as

$$\min \quad F_1 = \max_{t\in[2,N]} Z_t \tag{10}$$

where Z_t represents reservoir water level at time t (t = 2, 3, 4, ..., N).

Primary constraints include:

(1) Mass balance constraints:

$$V_t = V_{t-1} + (Q_t + Q_{t-1})\Delta t/2 - (q_t + q_{t-1})\Delta t/2 \tag{11}$$

(2) Reservoir storage constrains:

$$V^{\min} \le V_t \le V^{\max} \tag{12}$$

(3) Reservoir release constraints:

$$q_t \le {q_t}^{\max} \tag{13}$$

(4) Reservoir release fluctuation limits:

$$|q_t - q_{t-1}| \le \Delta q_m \tag{14}$$

(5) Initial and boundary conditions:

$$Z_1 = Z_N = Z_C \tag{15}$$

(6) Non-negative constraints: All of the above variables are non-negative.

where V_t denotes reservoir storage at time t (m^3); Q_t and q_t represent reservoir inflow and outflow at time t, respectively (m^3/s); Δt means time interval (s); $V^{\min}$ and $V^{\max}$ are the minimum and maximum allowable reservoir storage, respectively; ${q_t}^{\max}$represents maximum reservoir release capacity; Δq_m denotes the allowable discharge variation between adjacent time periods (m^3/s); Z_C means the initial Z_1 and end Z_N which is generally set to flood control water level.

The particle swarm optimization (PSO) algorithm (proposed by Kennedy and Eberhan [33]) based on the hunting behavior law of birds is employed to solve the model. To date, the algorithm has attracted widespread attention due to its advantages of less parameters, high search efficiency and easy implementation.

2.4. Information theoretic approach for uncertainty quantification and propagation

2.4.1. Concepts of information theory

As a fundamental concept of information theory, entropy (Shannon [20]) has been widely used to quantify uncertainty of random variable(s). The entropy of a random variable X is defined as

$$H(X) = -\int f(x)\log f(x)dx \tag{16}$$

where X indicates a continuous random variable; $f(x)$ represents the probability density function of X. Here, the base of the logarithm is set to 2 manifesting the unit of entropy is bits. The $\int$ should be replaced by Σ if X is discrete random variable.

In terms of two random variables X_1= (x_{11}, x_{12}, ..., x_{1N}) and X_2= (x_{21}, x_{22}, ..., x_{2N}), the joint entropy $H(X_1, X_2)$ is given by

$$\begin{aligned} \mathrm{H}(X_1, X_2) &= -\int\int f(x_1, x_2)\log f(x_1, x_2)dx_1dx_2 \\ &= -\int\int c(F(x_1), F(x_2))f(x_1)f(x_2)\log c(F(x_1), F(x_2))f(x_1)f(x_2)dx_1dx_2 \end{aligned} \tag{17}$$

Here, $f(x_1, x_2)$ is the joint probability density function of X_1 and X_2. The joint entropy is sub-additive, i.e. $\mathrm{H}(X_1, X_2) \le \mathrm{H}(X_1) + \mathrm{H}(X_2)$, with equality if X_1 and X_2 are statistically independent.

Moreover, for the multi-variables, the joint entropy $\mathrm{H}(X_1, X_2, ..., X_k)$ can be computed with the aid of following chain rule,

$$\mathrm{H}(X_1, X_2, \cdots, X_k) = \mathrm{H}(X_1) + \sum_{i=2}^{k} \mathrm{H}(X_i|X_1, \cdots, X_{i-1}) \tag{18}$$

where $\mathrm{H}(X_i|X_1, ..., X_{i-1})$ denotes the conditional entropy.

The conditonal entropy $\mathrm{H}(Y|X_1, X_2, ..., X_k)$ can be used to measure the remaining uncertainty of variable Y given knowledge of (X_1, X_2, ..., X_k) [34],

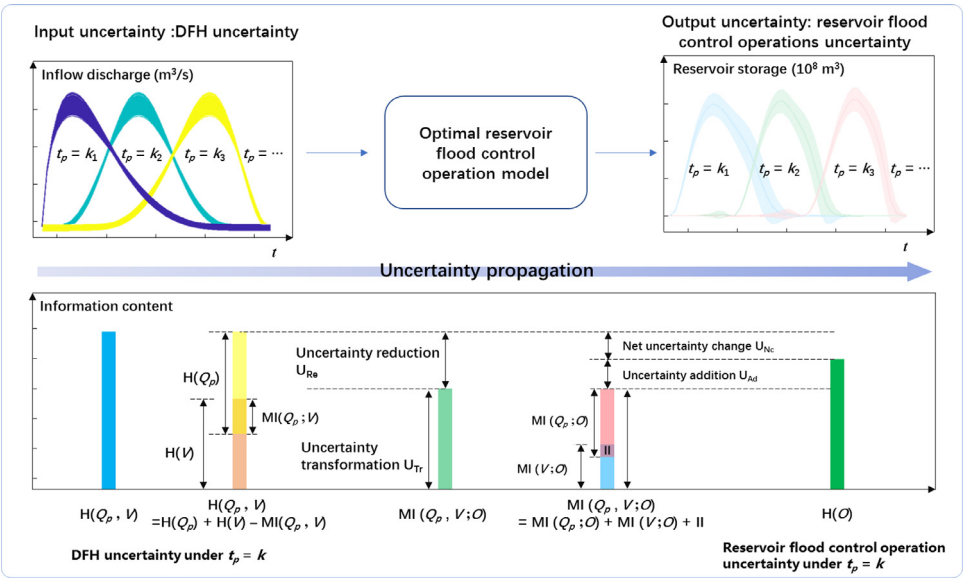

Fig. 2. Uncertainty quantification and propagation in reservoir flood control system under DFH uncertainty.

$$\begin{aligned} \mathrm{H}(Y|X_1, X_2, \cdots, X_k) &= \mathrm{H}(Y) - \mathrm{MI}(X_1, X_2, \cdots, X_k; Y) \\ &= -\iint f(x_1, x_2, \cdots, x_k, y) \log f(x_1, x_2, \cdots, x_k|y)\, dx_1 dx_2 \cdots dx_k dy \end{aligned} \tag{19}$$

where MI (X_1, X_2, ..., X_k;Y) is the mutual information which is also called transformation (Amorocho and Espildora [35]). MI indicates how much information is shared between (X_1, X_2, ..., X_k) and Y or how much uncertainty is cnveyed from (X_1, X_2, ..., X_k) to Y,

$$\mathrm{MI}(X_1, X_2, \cdots, X_k; Y) = \mathrm{H}(X_1, X_2, \cdots, X_k) + \mathrm{H}(Y) - \mathrm{H}(X_1, X_2, \cdots, X_k, Y) \tag{20}$$

Apart from MI, interaction information (denoted as II) is another index measuring the amount of shared information between variables, but is not present in any feature subset. In the case of $k = 2$, II can be expressed as

$$II(X_1; X_2; Y) = \mathrm{MI}(X_1, X_2; Y) - \mathrm{MI}(X_1; Y) - \mathrm{MI}(X_2; Y) \tag{21}$$

II can be negative, positive or zero. Positive II means synergy dominates over redundancy in the partitioning of shared information [26]. Negative II denotes the dominance of redundant relationship among variables.

2.4.2. Uncertainty quantification and propagation in reservoir flood control system

Aforementioned information theory is used to characterize uncertainty in reservoir flood control system under DFH uncertainty. Here, input is the uncertain DFH and output is reservoir water level Z or reservoir storage (Fig. 2a). The uncertain DFH is characterized by uncertain Q_p, V and t_p.

With the aid of information theory, the input uncertainty (denoted as U_{IN}) in reservoir flood control operation system is quantified in terms of entropy

$$Input \quad \text{uncertainty} : U_{IN} = \begin{cases} \mathrm{H}(Q_p), & t_p = k \\ \mathrm{H}(V), & t_p = k \\ \mathrm{H}(Q_p, V), & t_p = k \end{cases} \tag{22}$$

Similarly, the output uncertainty (denoted as $\mathrm{U_{OUT}}$) is quantified as following

$$Output \quad \text{uncertainty} : U_{OUT} = \mathrm{H}(Z), \qquad t_p = k \tag{23}$$

MI is used to measure the uncertainty transformed (denoted as $\mathrm{U_{Tr}}$) from DFH to Z

$$Uncertainty \quad transformation : U_{Tr} = \mathrm{MI}(Q_p, V; Z), \qquad t_p = k \tag{24}$$

Apart from the transformed uncertainty from DFH to Z, the untransformed uncertainty due to $\mathrm{OROM_{FC}}$ is defined as the difference between the $\mathrm{H}(Q_p, V)$ and U_{Tr} which can be considered as the lost uncertainty or reduced uncertainty (denoted as U_{Re}). Mathematically, the U_{Re} can be computed by

$$Uncertainty \quad reduction : U_{Re} = \mathrm{H}(Q_p, V) - U_{Tr}, \qquad t_p = k \tag{25}$$

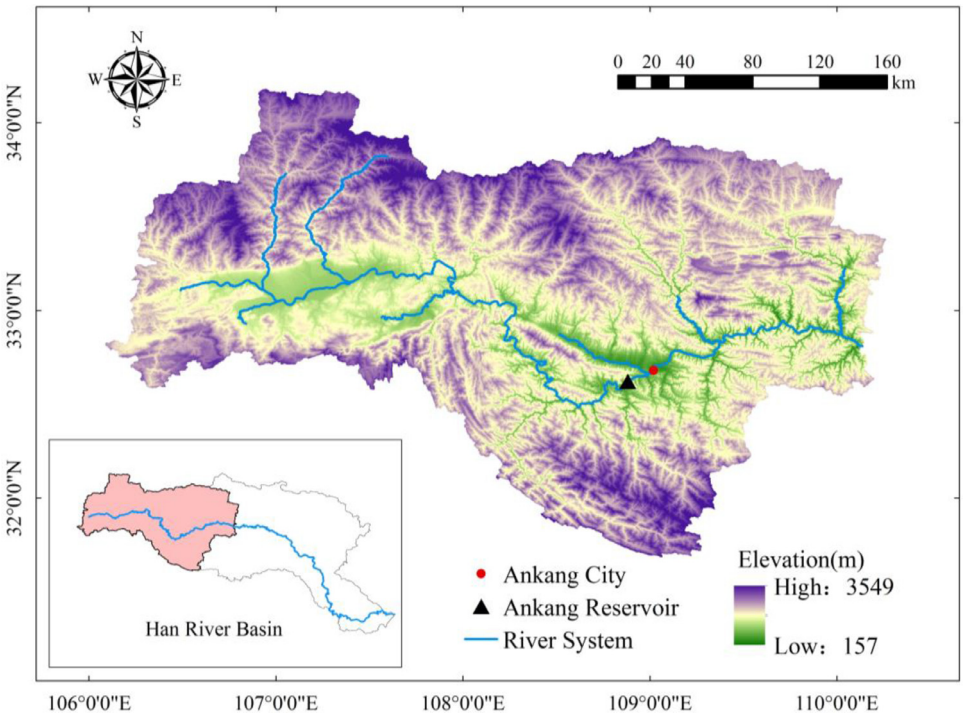

Fig. 3. Map of han river basin and locations of Ankang reservoir.

Additionally, the randomness of optimization algorithm would add additional uncertainty in reservoir operations. If optimization algorithm can always search the optimal solution, the uncertainty addition ($\underline{U}_{\mathrm{Ad}}$) from DFH to Z will be zero. The formula of U_{Ad} is expressed as

$$Uncertainty\ addition: U_{Ad} = \mathrm{H}(Z) - U_{Tr}, \qquad t_p = k \tag{26}$$

Optimization algorithm can be regarded as an inseparable part of $\mathrm{OROM_{FC}}$. Therefore, the net uncertainty change due to $\mathrm{OROM_{FC}}$ can be computed as the difference between U_{Re} and U_{Ad}.

$$Uncertainty\ net-change: U_{Nc} = \mathrm{H}(Q_p, V) - \mathrm{H}(Z) = U_{Re} - U_{Ad}, \qquad t_p = k \tag{27}$$

More specifically, aforementioned relationships for uncertainty quantification and propagation in reservoir flood control operation system are exhibited in Fig. 2.

3. Introduction to the case study

As the largest tributary of the Yangtze River, the Han River in China is selected as study region in this paper. Ankang reservoir located in the upper reaches of the Han River is a critical element of flood control system (Fig. 3). The total storage capacity 3.2 billion m^3. The normal water level of the reservoir is 330 m with storage capacity 2.2 billion m^3. The flood control water level is 325 m with storage capacity 1.8 billion m^3. Exemplarily, this paper aims to study reservoir flood control operations when facing a flood event with a return period of 200 years. The primary designed flood control object of Ankang reservoir is Ankang city with a population size of 2,675,000 and GDP of 118.21 billion CNY.

Annual maximum flood discharge Q_p and annual maxima 3-day flood volume V (1956–2013) observed at Ankang station are utilized to estimate bivariate design flood. Results of marginal distribution selection indicate that Gam distribution is the best one to fit Q_p and V series (Table S1). In this paper, the flood process is set to 46 periods with an interval of 3 h. Hence, the estimated V under specific return period is transformed to flood volume during the whole flood process according to the relationship between maxima 3-day flood volume and flood volume during the whole flood process on July, 1983. Moreover, according to shape of historical flood events observed at AnKang station, present study sets t_p changing from the 7th to 34th period.

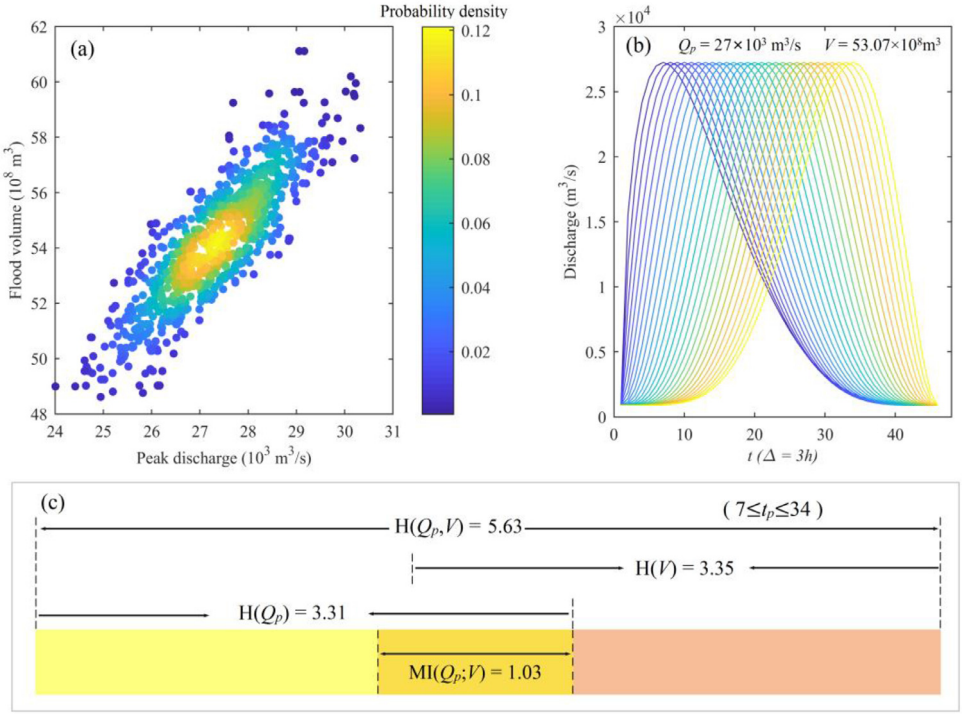

Fig. 4. Design flood for the AnKang station with 200-year return period under uncertainty.

4. Results and discussion

4.1. Input uncertainty quantification: DFH uncertainty

Fig. 4a displays design flood with 200-year return period under uncertainty for the AnKang station. It can be found that design flood displays enormous difference in the context of uncertainty. Specifically, the peak discharge with 200-year return period spans from 24.01 × 10^3 m^3/s to 30.32 × 10^3 m^3/s, while flood volume spans from 48.61 × 10^8 m^3 to 61.11 × 10^8 m^3. The difference between the lowest and highest flood volume (12.50 × 10^8 m^3) is larger than the flood control volume (9.75 × 10^8 m^3) of AnKang reservoir. Fig. 4b exemplarily shows the DFH (Q_p = 27 × 10^3 m^3/s, V = 53.07 × 10^8 m^3) under t_p uncertainty with return period of 200 years.

To quantify the uncertainty in design flood, information entropy introduced in Section 2.4.2 is applied in this section. The higher the entropy, the more uncertain the variable(s) is (are). The bar chart Fig. 4c illustrates that uncertainty in (Q_p, V) (i.e. H(Q_p, V)) reaches to 6.11 under 7 ≤t_p ≤ 34. To be specific, uncertainty in V (i.e. H(V)) is close to that in Q_p (i.e. H(Q_p)). Furthermore, the mutual information MI(Q_p;V) is applied to measure the amount of shared information between Q_p and V . It can be found that MI(Q_p;V) accounts for 25.73% and 24.72% of H(Q_p) and H(V), respectively. The large MI(Q_p;V) also implies the co-dependencies between Q_p and V should be seriously taken into account.

4.2. Output uncertainty quantification: reservoir flood control operation uncertainty

In this paper, we select 3000 DFHs with 200-year return period as inputs to the $OROM_{FC}$. Therefore, 3000 groups of optimal flood control operations of AnKang reservoir can be determined by the action of $OROM_{FC}$ on the 3000 DFHs. Fig. 5 exemplarily displays flood control operations of AnKang reservoir under DFH uncertainty with t_p = 7, 20 and 34, respectively. It can be found from the right panels of Fig. 5 that the width of 95% confidence region of reservoir operations before the time of MRS is narrower than that after the time of MRS. The narrower confidence region of reservoir operations before the time of MRS implies less candidate reservoir operation scheme in this stage. That is to say, reservoir operations before the time of MRS is more certain than that after the time of MRS under DFH uncertainty. Fig. 6 can further confirm this conclusion. It can be found that, similar to Fig. 5, uncertainties in reservoir flood control operations are smaller in periods close to t_p compared with that in other periods. Reservoir flood control operations are more certain in periods close to t_p compared with that in other periods.

Maximum reservoir storage (MRS) is a critical indicator in reservoir flood control system. The smaller MRS means the safer of the reservoir and its upstream. To explore the relationship between design flood and MRS, we do the scatter plot in this section (Fig. 7). From the lower panel of Fig. 7, it can be easily found that t_p uncertainty brings about a little impact on

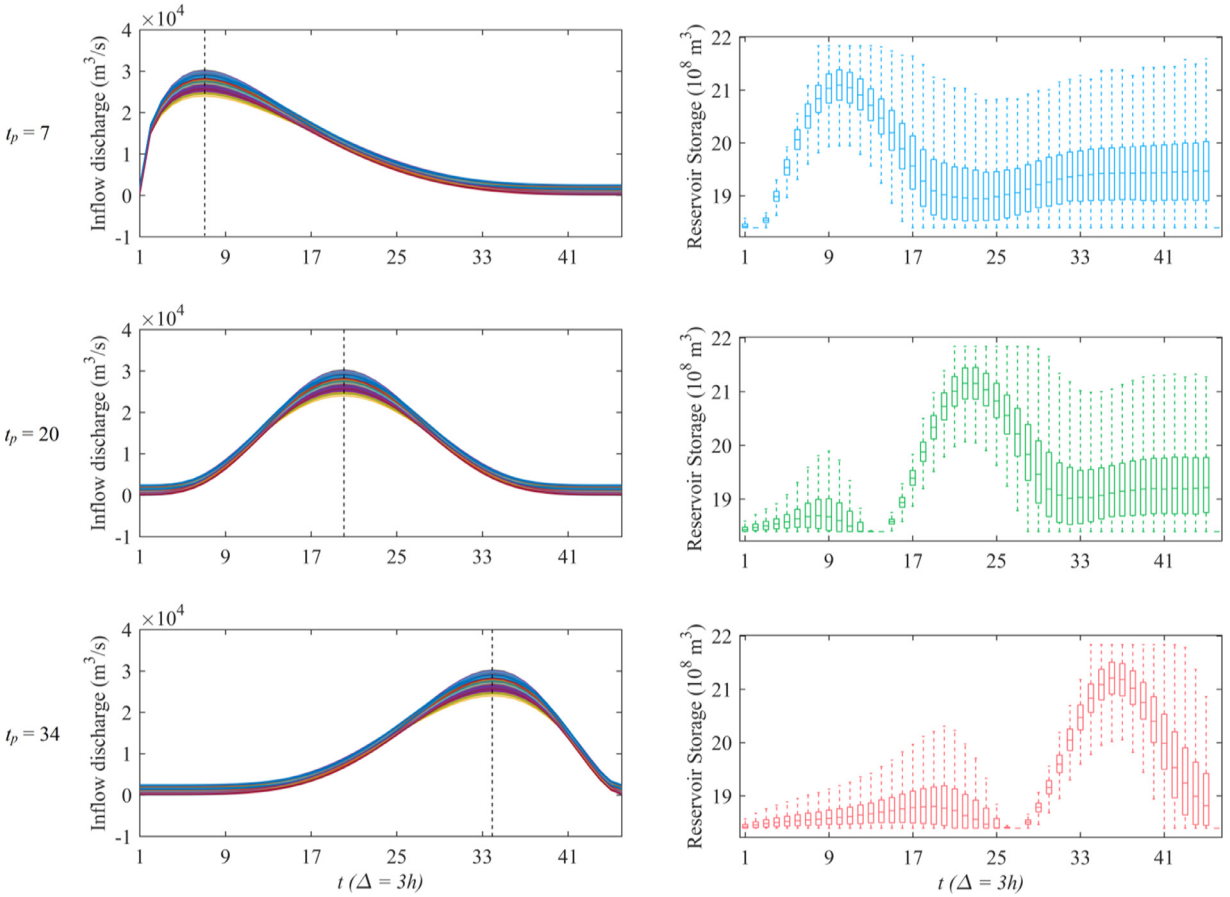

Fig. 5. Flood control operations of AnKang reservoir under DFH uncertainty.

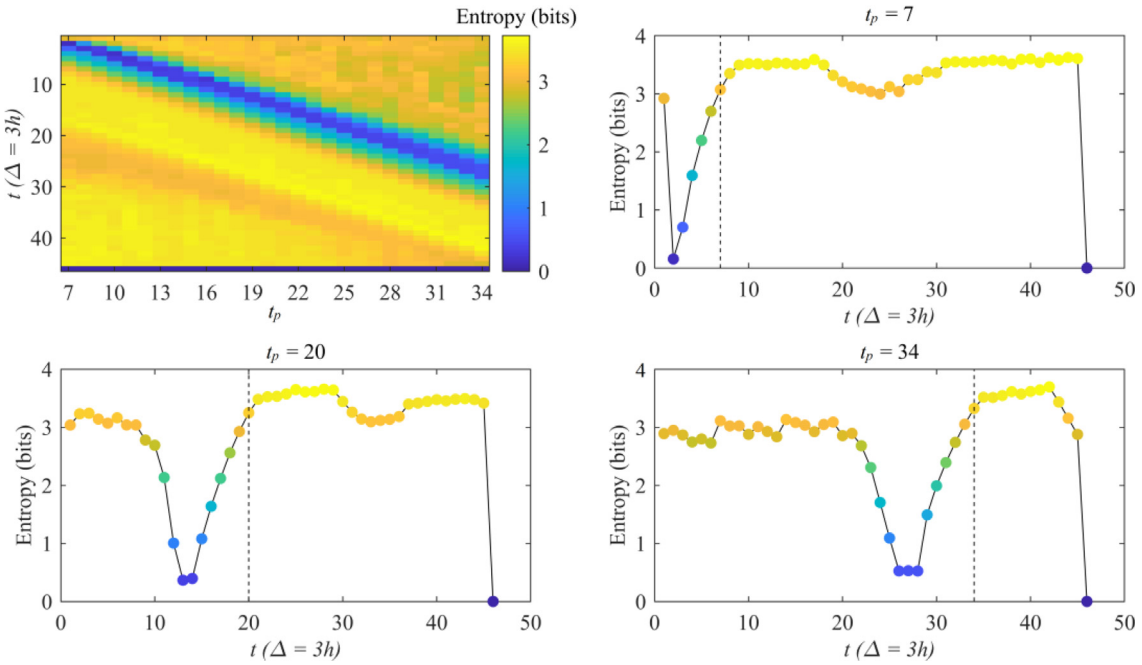

Fig. 6. Uncertainty in reservoir flood control operations.

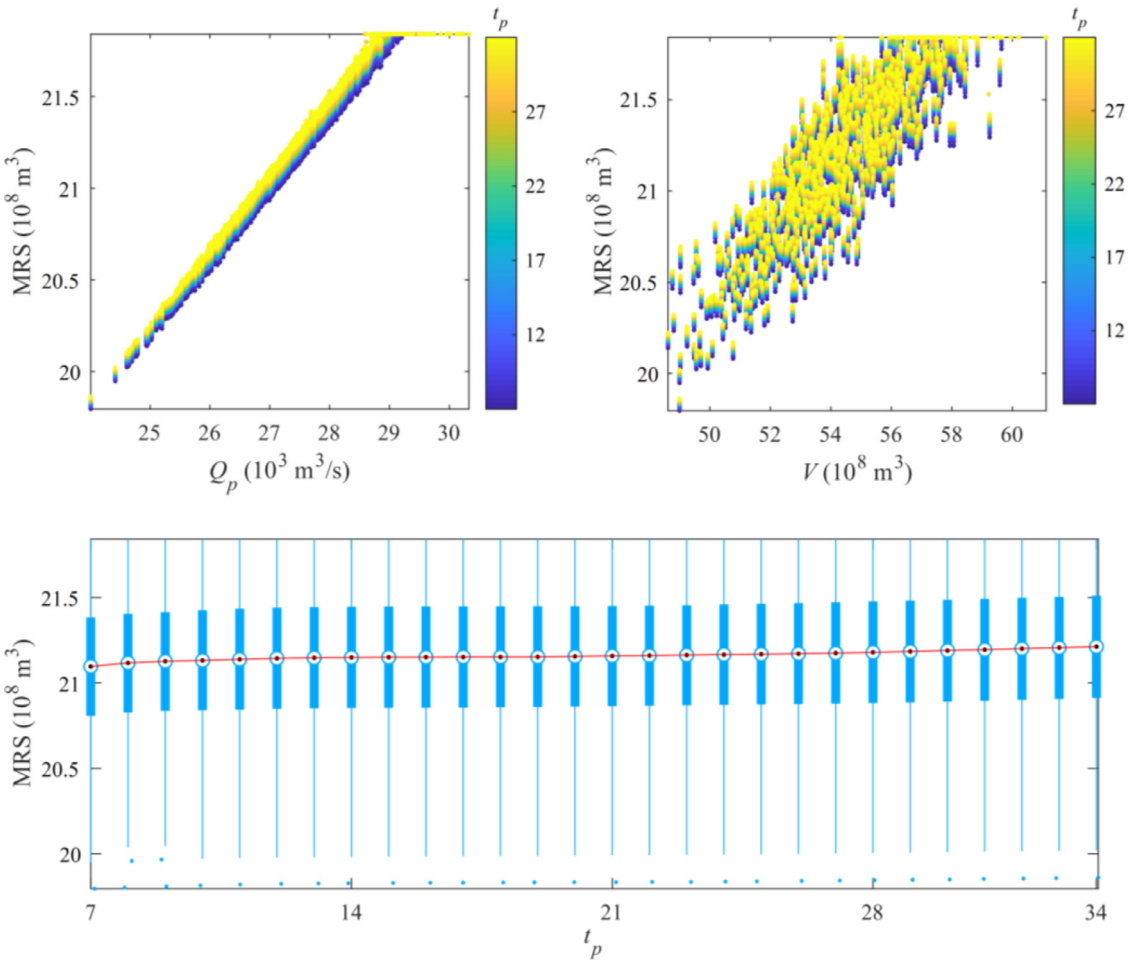

Fig. 7. Maximum AnKang reservoir storage under DFH uncertainty.

MRS. Unlike t_p, uncertainties in Q_p and V leads to larger impact on MRS. Moreover, top panels of Fig. 7 highlight the striking positive linear relationship between Q_p and MRS, and the positive nonlinear relationship between V and MRS.

4.3. Uncertainty propagation: from DFH to reservoir flood control operation

This section focuses on disclosing the propagation of uncertainty from DFH to reservoir operations. Due to the importance of MRS, we employ the MRS to perform uncertainty propogation analysis instead of reservoir storage for each period. Related results are shown in Fig. 8. From Fig. 8, it can be found that the output uncertainty (U_{OUT}) spans from 3.51 to 3.54 under $7 \leq t_p \leq 34$. The small difference of U_{OUT} under different t_p further indicates t_p uncertainty has little influence on reservoir operations. Similarly, uncertainty transformation U_{Tr} from DFH to reservoir operations presents a slight variation under $7 \leq t_p \leq 34$ which ranges from 2.76 to 2.84. Specifically, U_{Tr} accounts for 49.02–50.44% of U_{IN} which indicates nearly half amount of uncertainty in DFH is propagated to the output variable, MRS. In other words, reservoir can effectively decrease uncertainty in DFH through its operation.

Moreover, it can be found from Fig. 8 that the reduced uncertainty (U_{Re}) spans from 2.79 to 2.87 under $7 \leq t_p \leq 34$ which is slightly higher than U_{Tr}. Additionally, as aforementioned in methodology part, the randomness of optimization algorithm would bring additional uncertainty in reservoir operations. The uncertainty addition U_{Ad}, as shown in Fig. 8, is relatively small ranging from 0.69 to 0.76. It denotes that optimization algorithm can find the optimal solution in most cases. Taking the U_{Re} and U_{Ad} into account together, we can obtain the net uncertainty change (U_{Nc} in Fig. 8). It can be found that U_{Nc} is close to U_{Re}.

Fig. 9 presents results on information contribution by individual design flood variables (Q_p and V) to the fluctuation of MRS. Specifically, MI(Q_p; MRS) ranges from 2.60 to 2.73 under $7 \leq t_p \leq 34$ which is 2.73 to 2.90 times larger than MI(V; MRS). It indicates that Q_p uncertainty exhibits largest contribution to the fluctuation of MRS. Additionally, the negative II denotes

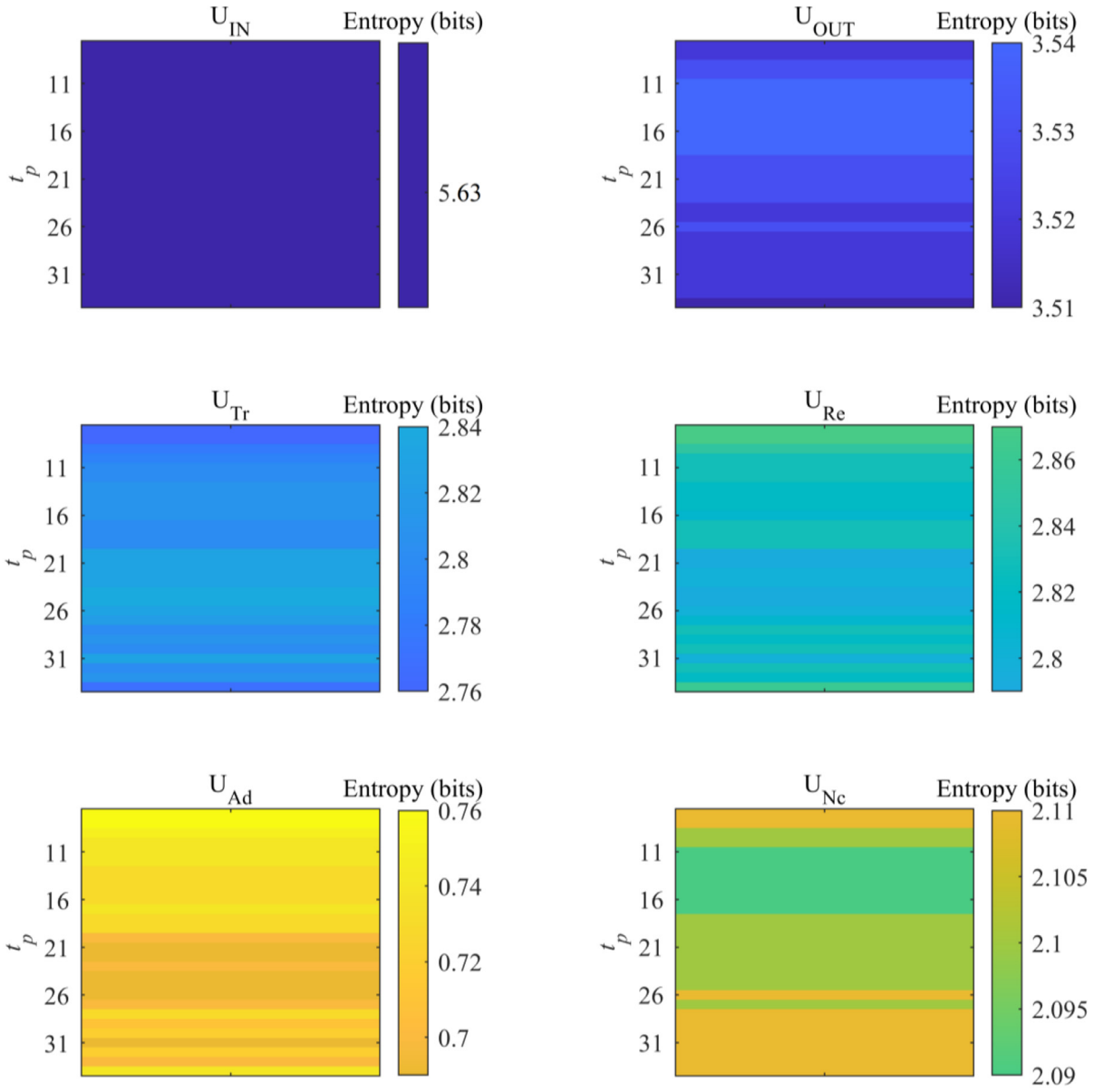

Fig. 8. Uncertainty quantification and propagation in reservoir flood control system under DFH uncertainty.

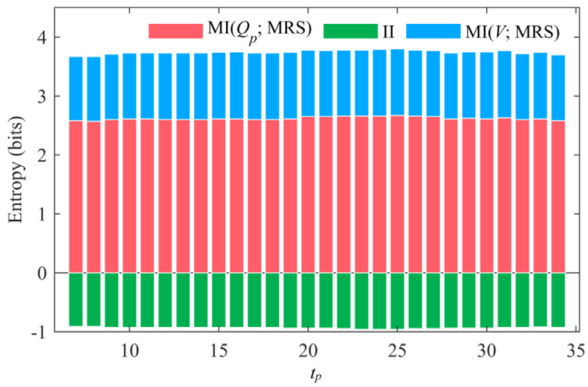

Fig. 9. Illustration of MI(Q_p, V; MRS) partitioning results.

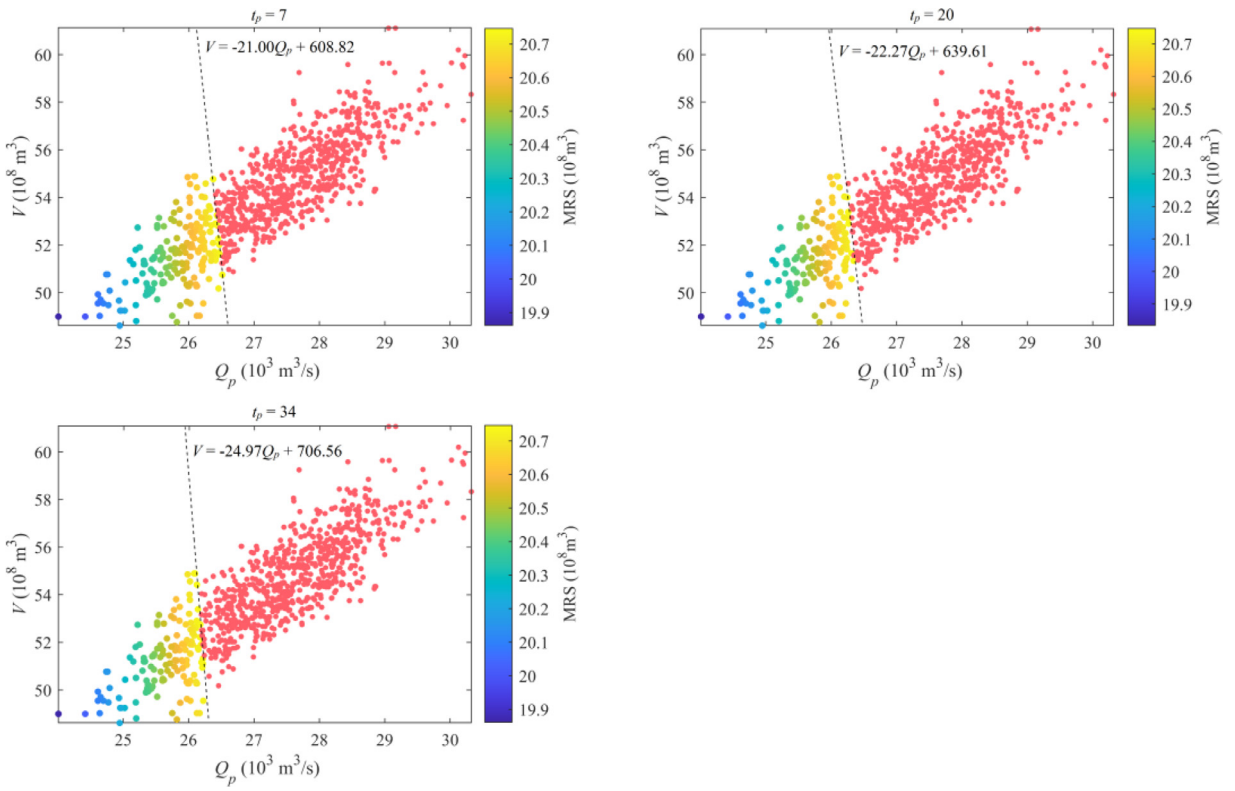

Fig. 10.. Relationship between MRS and Q_p, V under 95% of original reservoir flood control capacity.

that the information offered by the combination of Q_p and V is less than the sum of information provided by Q_p and V, respectively. There is redundant dominantly information provided by Q_p and V.

4.4. Discussion on impact of reservoir flood control capacity

As shown in Fig. 7, MRS is equal to the largest reservoir storage (i.e., 21.84 × 10^8 m^3) when Q_p and V are larger than about 28.6 × 10^3 m^3/s and 54.25 × 10^8 m^3, respectively. This result infers that there is potential risk of reservoir overtopping under DFH uncertainty. In view of this, this section further investigates impact of reservoir flood control capacity on reservoir flood control operations under DFH uncertainty. Here, we set the max reservoir storage to 20.75 × 10^8 m, i.e., 95% of original max reservoir storage (95% × 21.84 × 10^8 m^3 = 20.75 × 10^8 m^3). Subsequently, we perform the OROM_{FC} under DFH uncertainty with 100-year return period when t_p = 7, 20 and 34.

Related results are shown in Fig. 10. Red points represent hydrological pairs of design flood peak and volume inducing reservoir overtopping. Like to Fig. 6, both Q_p and V (Fig. 10) present positive relationship with MRS. Larger Q_p and V result in larger MRS, i.e., the larger risk of reservoir overtopping. In addition, Fig. 6 illustrates that red points appears more along with larger t_p. That is to say, apart from Q_p and V, larger t_p induces larger risk of reservoir overtopping. The primary reason is that reservoir reaches to the extremum water level before t_p while reservoir release capacity is lower than reservoir inflow during the following time. Reservoir hasn't enough storage ability to control following flood discharge. Moreover, as is shown in Fig. 6, hydrological pairs of Q_p and V are classified into two parts by corresponding critical line. Specifically, AnKang reservoir can effectively control DFHs characterized by hydrological pairs of Q_p and V in the lower-left part of Fig. 6, while DFHs characterized by (Q_p, V) pairs in the upper-right part of Fig. 6 lead to reservoir overtopping. The corresponding formulas (such as V = -21.00Q_p + 608.82, t_p = 7) for aforementioned critical lines are shown in Fig. 6. In practice, flood control department can roughly determine whether reservoir is able to control the future flood discharge or not with the aid of these formulas. Additionally, analysis on uncertainty propagation from DFH to reservoir flood control operations are performed and related results are shown in Tables S2 and S3.

5. Summary and conclusions

Design flood hydrograph (DFH) harbours inherent uncertainties. Inevitably, the uncertain DFH correspondingly affects flood control operation of reservoir. To enable better flood control decision making under uncertainty, it is necessary to

quantify uncertainty in DFH and related reservoir flood control operation, and disclose DFH uncertainty propagation in reservoir flood control system (RFCS). Hence, in this paper, we present a methodological framework to perform above mentioned uncertainty investigation. In this framework, the DFH is estimated by a copula-based model while the Bayes' theorem is integrated to analyze the uncertainty in DFH (Fig. 1). Subsequently, the uncertain DFHs are inputted to the optimal reservoir operation model for flood control ($OROM_{FC}$). Based on the outputted optimal reservoir flood control operations, we devise an information theory-based approach (Fig. 2) to measure the uncertainties in input and output variables in RFCS and explore the propagation of uncertainty from input to output (Eqs 22-27).

Application of this framework is exemplified in AnKang RFCS in Han River basin in China. We find that uncertainty in reservoir flood control operations triggered by DFH uncertainty decreases and then increases with time (Figs. 5 and 6). Specifically, the turning point occurs prior to the time of peak flow occurrence. In other words, reservoir flood control operations are relatively certain around the turning point in comparison with that in the other periods. The phenomenon further highlights the cruciality of flood control operations during periods prior to the time of peak flow occurrence. To further understand the process of DFH uncertainty propagation in reservoir flood control system, we employ the maximum reservoir storage (MRS) as indicator to represent the output of $OROM_{FC}$ (Fig. 7). Results indicate that uncertainty in output variable is smaller than that in DFH in RFCS (Fig. 8). Additionally, we find that uncertainty in flood peak of DFH dominates the uncertainty in MRS compared with uncertainty in flood volume of DFH (Figs. 8 and 9). Furthermore, we explore the impact of reservoir flood control storage on the relationship between input and output variables in reservoir flood control system under DFH uncertainty. When reservoir flood control storage is 95% of the original one, there exists remarkable critical line between (Q_p, V) pairs inducing reservoir overtopping or not under DFH uncertainty (Fig. 10). The specific formula describing the critical line also is derived in this paper.

As aforementioned, this paper is triggered by the remarkable uncertainty in DFH. Tracing the DFH uncertainty in reservoir flood control system constitutes the primary target of this paper. A high magnitude flood event, i.e. a flood event with 200-year return period at AnKang reservoir, determines choosing "ensuring the safety of the dam and its upstream area" as the optimal objective of $OROM_{FC}$. If readers are interested in exploring the impact of optimal objective (such as ensuring the safety at a downstream control point on condition of low magnitude flood event) on results of this paper, they can re-apply the proposed framework shown in Figs. 1 and 2. Additionally, this paper just explores the impact of DFH uncertainty on single reservoir flood control operations. The designed information theory-based framework should be further improved to apply to more uncertainty sources (such as initial water level uncertainty) and cascade reservoirs. How to calculate the joint entropy of continues and discrete variables and how to consider the compensating operation between reservoirs need to be investigated in future study.

Data availability

All the data used in this research are available upon request.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Aijun Guo: Conceptualization, Visualization, Investigation, Writing – original draft, Methodology. **Jianxia Chang:** Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing. **Yimin Wang:** Funding acquisition, Writing – review & editing. **Bin Wu:** Formal analysis, Methodology. **Yunyun Li:** Formal analysis, Methodology.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Granted Nos: 51679187, 51909207, 51879214), National Key Research and Development Program of China (2016YFC0400906), Teachers' Doctoral Research Initiation Fund in Xi'an University of Technology (104/256081916). The authors extend special thanks to editors and numerous anonymous reviewers.

References

[1] D. Hine, J.W. Hall, Information gap analysis of flood model uncertainties and regional frequency analysis, Water Resour. Res. 46 (2010) W01514, doi:10.1029/2008WR007620.
[2] F. Serinaldi, An uncertain journey around the tails of multivariate hydrological distributions, Water Resour. Res. 49 (2013) 6527–6547, doi:10.1002/wrcr.20531.
[3] A. Botto, D. Ganora, F. Laio, P. Claps, Uncertainty compliant design flood estimation, Water Resour. Res. 50 (2014) 4242–4253, doi:10.1002/2013WR014981.
[4] G.H. Steinbakk, T.L. Thorarinsdottir, T. Reitan, L. Schlichting, S. Hølleland, K. Engeland, Propagation of rating curve uncertainty in design flood estimation, Water Resour. Res. 52 (2016) 6897–6915, doi:10.1002/2015WR018516.
[5] M.I. Brunner, A.E. Sikorska, R. Furrer, A.C. Favre, Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments, Water Resour. Res. 54 (2018) 1493–1512, doi:10.1002/2017wr021129.

[6] W. Qi, C. Zhang, G. Fu, H. Zhou, Imprecise probabilistic estimation of design floods with epistemic uncertainties, Water Resour. Res. 52 (2016) 4823–4844, doi:10.1002/2015WR017663.
[7] E.M. Michailidi, B. Bacchi, Dealing with uncertainty in the probability of overtopping of a flood mitigation dam, Hydrol. Earth Syst. Sci. 21 (2017) 2497 2017, doi:10.5194/hess-21-2497-2017.
[8] G. Salvadori, C. De Michele, F. Durante, Multivariate design via copulas, Hydrol. Earth Syst. Sci. Discuss. 8 (2011) 5523–5558, doi:10.5194/hessd-8-5523-2011.
[9] E. Volpi, A. Fiori, Hydraulic structures subject to bivariate hydrological loads: return period, design, and risk assessment, Water Resour. Res. 50 (2014) 885–897, doi:10.1002/2013WR014214.
[10] N.V. Dung, B. Merz, A. Bárdossy, H. Apel, Handling uncertainty in bivariate quantile estimation-An application to flood hazard analysis in the Mekong delta, J. Hydrol. 527 (2015) 704–717, doi:10.1016/j.jhydrol.2015.05.033.
[11] M.I. Brunner, D. Viviroli, A.E. Sikorska, O. Vannier, A.-C. Favre, J. Seibert, Flood type specific construction of synthetic design hydrographs, Water Resour. Res. 53 (2017) 1390–1406, doi:10.1002/2016WR019535.
[12] M.I. Brunner, J. Seibert, A.C. Favre, Bivariate return periods and their importance for flood peak and volume estimation, Wine Econ. Policy 3 (2016) 819–833, doi:10.1002/wat2.1173.
[13] A. Guo, J. Chang, Y. Wang, Q. Huang, S. Zhou, Flood risk analysis for flood control and sediment transportation in sandy regions: a case study in the loess plateau, Chin. J. Hydrol. 560 (2018) 39–55, doi:10.1016/j.jhydrol.2018.02.076.
[14] Y. Fan, K. Huang, G. Huang, Y. Li, F. Wang, An uncertainty partition approach for inferring interactive hydrologic risks, Hydrol. Earth Syst. Sci. 24 (2020) 4601–4624 2020, doi:10.5194/hess-24-4601-2020.
[15] A. Montanari, What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology, Hydrol. Process 21 (2007) 841–845, doi:10.1002/hyp.6623.
[16] F. Serinaldi, C.G. Kilsby, Stationarity is undead: uncertainty dominates the distribution of extremes, Adv. Water Resour. 77 (2015) 17–36, doi:10.1016/j.advwatres.2014.12.013.
[17] A. Guo, J. Chang, Y. Wang, Q. Huang, Y. Li, Uncertainty quantification and propagation in bivariate design flood estimation using a Bayesian information-theoretic approach, J. Hydrol. 584 (2020) 124677, doi:10.1016/j.jhydrol.2020.124677.
[18] F. Serinaldi, S. Grimaldi, Synthetic design hydrographs based on distribution functions with finite support, J. Hydrol. Eng. 16 (2011) 434–446, doi:10.1061/(ASCE)HE.1943-5584.0000339.
[19] J. Yin, S. Guo, Z. Liu, G. Yang, Y. Zhong, D. Liu, Uncertainty analysis of bivariate design flood estimation and its impacts on reservoir routing, Water Resour. Manag. 32 (2018) 1795–1809, doi:10.1007/s11269-018-1904-x.
[20] C. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x.
[21] N. Lüdtke, S. Panzeri, M. Brown, D.S. Broomhead, J. Knowles, M.A. Montemurro, D.B. Kell, Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks, J. R Soc. Interface 5 (2008) 223–235, doi:10.1098/rsif.2007.1079.
[22] Y. Qian, J. Liang, F. Wang, A new method for measuring the uncertainty in incomplete information systems, Int. J. Uncertain. Fuzz. 17 (2009) 855–880, doi:10.1142/S0218488509006303.
[23] Y. Chen, K. Wu, X. Chen, C. Tang, Q. Zhu, An entropy-based uncertainty measurement approach in neighborhood systems, Inf. Sci. 279 (2014) 239–250, doi:10.1016/j.ins.2014.03.117.
[24] A. Sharma, R. Mehrotra, An information theoretic alternative to model a natural system using observational information alone, Water Resour. Res. 50 (2014) 650–660, doi:10.1002/2013WR013845.
[25] W. Gong, H.V. Gupta, D. Yang, K. Sricharan, A.O. Hero, Estimating epistemic and aleatory uncertainties during hydrologic modeling: an information theoretic approach, Water Resour. Res. 49 (2013) 2253–2273, doi:10.1002/wrcr.20161.
[26] A.E. Goodwell, P. Kumar, Temporal information partitioning networks (TIPNets): a process network approach to infer ecohydrologic shifts, Water Resour. Res. 53 (2017) 5899–5919, doi:10.1002/2016WR020218.
[27] A.E. Goodwell, P. Kumar, Temporal information partitioning: characterizing synergy, uniqueness, and redundancy in interacting environmental variables, Water Resour. Res. 53 (2017) 5920–5942, doi:10.1002/2016WR020216.
[28] J.A. Vrugt, C.J.F. ter Braak, C.G.H. Diks, B.A. Robinson, J.M. Hyman, D. Higdon, Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, Int. J. Nonlinear Sci. Num. 10 (2009) 273–290, doi:10.1515/IJNSNS.2009.10.3.273.
[29] A. Bardossy, Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res. 42 (2006) W11416, doi:10.1029/2005WR004754.
[30] C. Genest, A.C. Favre, J. Béliveau, C. Jacques, Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data, Water Resour. Res. 43 (2007) W09401, doi:10.1029/2006WR005275.
[31] A. Guo, J. Chang, Y. Wang, Q. Huang, Z. Guo, Y. Li, Development of a partial copula-based algorithm for disclosing variability of dependence structures between hydro-meteorological factors under consideration of covariate-effect, J. Hydrol. 583 (2020) 124570, doi:10.1016/j.jhydrol.2020.124570.
[32] G. Salvadori, C. De Michele, Frequency analysis via copulas: theoretical aspects and applications to hydrological events, Water Resour. Res. 40 (2004) W12511, doi:10.1029/2004WR003133.
[33] J. Kennedy, R.J. Eberhan, Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks, 4, 1995, pp. 1942–1948, doi:10.1109/ICNN.1995.488968.
[34] T.G. Chapman, Entropy as a measure of hydrologic data uncertainty and model performance, J. Hydrol. 85 (1986) 111–126, doi:10.1016/0022-1694(86)90079-X.
[35] J. Amorocho, B. Espildora, Entropy in the assessment of uncertainty in hydrologic systems and models, Water Resour. Res. 9 (1973) 1511–1522, doi:10.1029/WR009i006p01511.