

A joint-probabilistic programming method for water resources optimal allocation under uncertainty: a case study in the Beiyun River, China

Xueping Gao, Yinzhu Liu and Bowen Sun

ABSTRACT

In recent years, the lower reaches of the Beiyun River have suffered from growing water resource shortages due to the reduction of upstream water resource and drying up of the stream channel. More reasonable and scientifically based water allocation plans should be developed and implemented; however, uncertainties exist regarding the determination of water supply availability and spillage of extra water. To assess and manage regional water shortage, the combined effects of multiple water supply sources as well as the joint probability of typical events should be considered. The joint probability of water supply, considering upstream and local water supplies, was estimated through the copula functions. A multi-objective optimization model was then developed and solved by improved genetic algorithms to plan water resources allocation within a multi-source environment containing multiple competitive users. The framework is demonstrated, and represents a range of different water supply scenarios in terms of different probabilities of occurrence and constraint violations. The results showed that water allocation was greatly influenced by uncertainties, especially in upstream-local water supply. In addition, violating water-allocation constraint posed an extra uncertainty. This study facilitates the proposition of adaption allocation plans for uncertain environments, aiming to balance the shortage, economy, and reliability.

Key words | joint probability, multi-objective optimization, uncertainty, water resources allocation, water shortage crisis

Xueping Gao
Yinzhu Liu
Bowen Sun (corresponding author)
State Key Laboratory of Hydraulic Engineering
Simulation and Safety,
Tianjin University,
Tianjin 300072,
China
E-mail: bwsun@tju.edu.cn

INTRODUCTION

With the rapid increase in water demand driven by socio-economic development in recent years, water shortages have become increasingly common (Alavian *et al.* 2009). In addition, an imbalanced water distribution could intensify the shortages (Fan *et al.* 2012). To solve the contradictions between the supply and demand of water resources and to mitigate losses, water resource diversion and allocation have been developed. The optimal allocation of available water supply is quite necessary in areas of severe water shortage in China. However, several uncertain factors may affect the performance of water allocation (Zeng *et al.* 2016), such as hydrological system complexities, parameters

and their interactions. For natural impacts, the characteristics of available water resources are assumed to follow a probability distribution, which is affected by the flows of rivers, streams, varied precipitation, and runoff levels, etc. The water demand associated with the population increase and economic development is uncertain as well. Water allocation plan development is not straightforward due to the limited available water resources and uncertainties. Previous research into optimal water resources has mostly focused on embodying various characteristics in deterministic values and intervals, but little work has been done to consider the various complexities in terms of relationship

doi: 10.2166/hydro.2017.032

of multiple water resources. Correspondingly, more efficient management associated with water supply probabilities should be considered, depending on the origins of the uncertainties.

As for uncertainties, copulas are being increasingly employed in the analysis of joint probability. The copula method has been developed by Sklar (1959) and has received much attention regarding environmental and water resources in recent years (van den Berg *et al.* 2011). For example, the application of copula functions in hydrology has been mainly employed in the extreme hydrological events modeling (Yazdi *et al.* 2014), frequency analysis of hydrological events (Salvadori & De Michele 2004), and hydrological prediction (Liu *et al.* 2016). Copulas are functions that combine univariate probability distributions to obtain multivariate probability distributions (Lian *et al.* 2013; Yazdi *et al.* 2015). Additionally, copulas can describe nonlinear relations among variables. Compared with other methods, the marginal properties and dependence structure of random variables can be investigated separately (Grimaldi & Serinaldi 2006). Different families of copulas have been proposed and are described by Nelsen (1999). The Archimedean copula family is more desirable for hydrological analysis, because it can be easily constructed. The widely used Archimedean copulas include the Clayton, Frank, and Gumbel copulas. Proof of their detailed properties has been presented by Genest & MacKay (1986) and Nelsen (1999). For that mentioned above, the one-parameter Archimedean copulas were applied in this study for determining the joint probability distribution of correlated water resource variables.

The allocation of water resources is an optimization process that groups water resources to ensure that limited water resources are reasonably and fully exploited. Today, many optimization techniques are available for such problems. One is the use of mathematic programming algorithms, such as linear programming, non-linear programming, dynamic programming and progressive optimality algorithms (Rani & Moreira 2009). However, water allocation is a multi-objective system problem, which is a challenge due to both limited water resources and water shortage risk. Although supported by rigorous mathematical theory, these classical optimization methods cannot successfully address multidimensional problems with increasing numbers of state variables (Azamathulla *et al.* 2008).

In recent decades, other optimization techniques, including swarm intelligence algorithms and evolutionary algorithms, have been popularly employed to plan the optimal allocation of water resources, such as particle swarm optimization (Reddy & Kumar 2009), ant colony algorithm (Liu *et al.* 2012), harmony search algorithm (Manshadi *et al.* 2015), and the genetic algorithm (GA) (Reddy & Kumar 2006), as well as its improved algorithm. As a global heuristic search, the GA is one of the popular approaches due to its robustness, good adaptability, and overall optimization. It shows a significant advantage in solving complicated optimization problems that have two or more contradictory objective functions that should be simultaneously considered. Similar to the process of Darwinian evolution, GA can obtain a series of solutions that converge to the optimum by use of random initialization and a stochastic algorithm. It has been successfully employed in the allocation of various water resources (Zhou *et al.* 2015; Lalehzari *et al.* 2016). In deficit allocation studies, due to the increased quantity of water resources crises, it is necessary to fully use water resources to minimize allocation cost under conditions of both fewer or more water resources. The planned measures have to change to fit the changing environment to further strengthen their efficiency. Therefore, it is indispensable to obtain solutions under different water supply conditions, revealing a trade-off between system cost and shortages.

As an extension of the above studies, the objective of this study is to develop a joint-probabilistic multi-objective optimization method for planning of regional water resources systems. Due to the complexity of water resources, the main contribution of this paper is the representation of copula functions that consider the dependence structure of water supply variables. The statistical copula functions obtained the joint probability distribution for water resources by separating their dependence from their marginal distributions. Meanwhile, the research can tackle random parameters in the left- and right-hand sides of constraints considering spillage of extra water. The results obtained from the application of the methodology to a water shortage area showed the applicability of the method for analyzing a variety of water supply scenarios. This provides managers with a range of alternatives under various system conditions, so that managers will spend the

least under the uncertain water availabilities. In conclusion, the method can help water resources managers to make decisions about water allocation considering uncertainty and tradeoffs between economy and reliability.

MATERIALS AND METHODS

Case study and data

The case study is the lower reaches of the Beiyun River (Figure 1). The ten river systems in the downstream are Beiyun, Longfeng, Qinglongwan, Chaobaixin, and Yongdongxin rivers and Qinhuang, Yundong, Qingwu, Qingpai, Jintang canals. Beiyun River and Longfeng River are the upper reaches of the river system, thus the available water resources from the two rivers. The Beiyun River, which runs 186 km from Tong county of Beijing through Wuqing county and the Hongqiao district of Tianjin, is the main channel of the Haihe River water system. The Longfeng River, which runs 77.5 km from the Ganggou River of

Beijing, inflows into Tianjin. Kuangergang is a hydro-junction where the Beiyun River and Longfeng River flow together and bifurcate again. The river system connects the northern river of Tianjin and the main stream of the Haihe River; for water allocation, flood control, environment, and landscape, it plays an influential role in the development of urban areas and the Binhai New Area. In recent years, due to both natural and human influences, the runoff in the upstream of the Beiyun River has been less and less. The dried up stream channel, silting river course, and poor drainage have led to serious degeneration of its ecological functions.

The spatial distribution and variation in water is significantly influenced by various water resources and their changing probabilities. Meanwhile, due to the extremely uneven distributions of precipitation, the available water supply showed a distinctive seasonal variation (Gibbs et al. 2013). In the wet season, the total water demands of users could be satisfied because sufficient water supply was recharged. In the drought season, the decreased availability of water supply potentially caused competing water users to

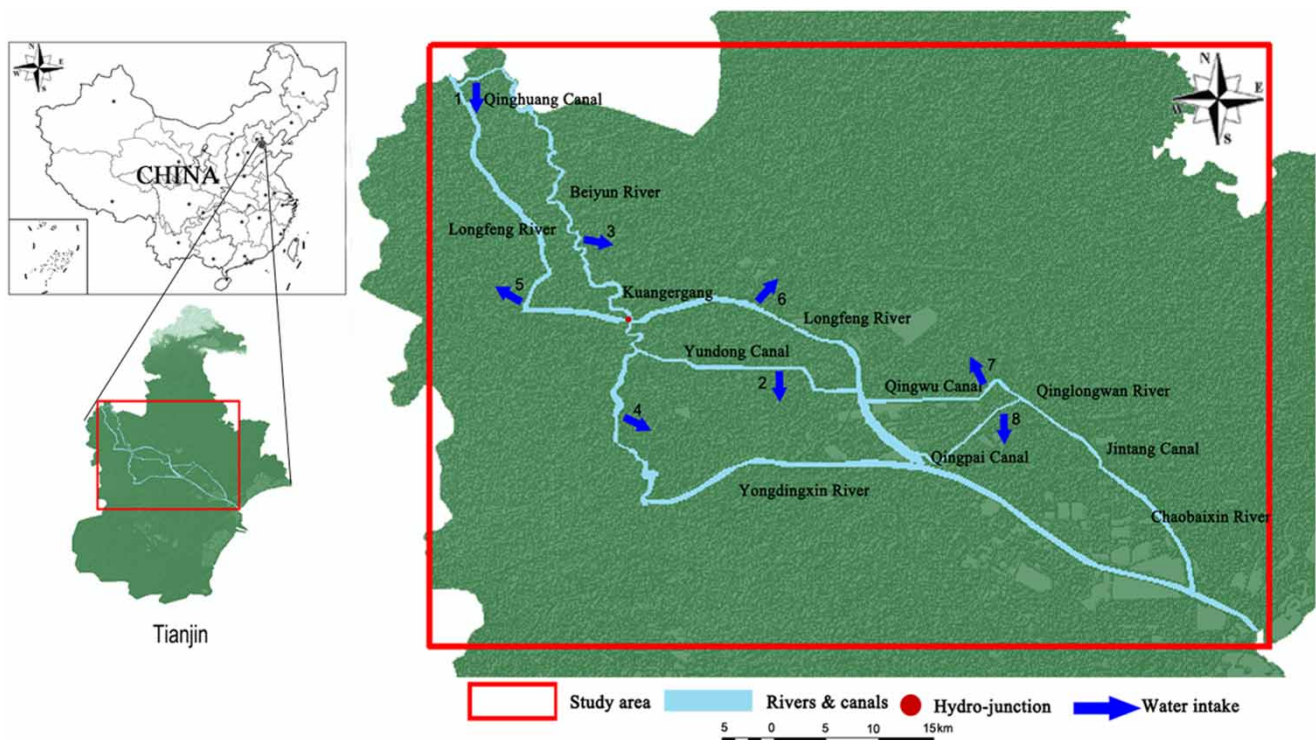


Figure 1 | Map of the region showing location of streams, hydro-junction, and water intakes.

face serious crises. Due to the reduction of the water supply, there may be an impact on water resource allocation. Therefore, the nature uncertainty factor in this study is the uncertain water supply of the Beiyun and Longfeng Rivers. As for the artificial factor, the uncertainty is whether the managers think the water conveyance losses acceptable. Consequently, the uncertainties that would influence the objective achievement are presented in terms of different acceptable levels of constraint violation and joint probabilities expressed in the water supply.

This paper considers a case in which managers have the responsibility of allocating water from two rivers (Beiyun River and Longfeng River) to four user groups: municipality, industry, agriculture, and ecology. The lower reaches of Beiyun River are the agricultural districts where agriculture accounts for a large proportion of the national economy. Thus, considering the specific situations, the order of the four groups is municipality, agriculture, industry, and ecology. Upstream and local water resource data were based on long-term observed hydrologic data collected from hydrologic stations on the river system, monthly from 1954 to 2009. The double-mass method has been employed to check the consistency of data, and the results show that the data series used in this study is consistent. The mean annual upstream water resources of the Beiyun River is $113.35 \times 10^6 \text{ m}^3$, more than those of the Longfeng River ($100.52 \times 10^6 \text{ m}^3$). In contrast, for mean annual local water resources, the value of the Beiyun River ($8.01 \times 10^6 \text{ m}^3$) is less than the Longfeng River ($12.59 \times 10^6 \text{ m}^3$). The water consumption data were recorded from 2000 to 2009 by the Tianjin Institute of Hydro-Technical and statistical yearbook (Table 1). The need for water leads to competition among the four user groups. Moreover, because of the effects of the different water demand targets in these competing water users, the guarantee ratios and priorities of assigned water are different. The quantity of municipality, industry, and ecological water demand used in this study is based on the estimation of the average proportion of water consumption with consideration of the population and economic changes.

Methodology

The joint-probabilistic programming method for water resources optimal allocation is based on integrating the

Table 1 | Water consumption (10^6 m^3)

Year	Total	Industry	Municipality	Ecology	Agriculture
2000	1.726	1.709	0.018	0	1.674
2001	1.610	1.590	0.020	0	1.570
2002	1.680	1.650	0.030	0	1.630
2003	1.800	1.760	0.040	0	1.680
2004	1.840	1.840	0.000	0	1.800
2005	1.870	1.844	0.003	0.023	1.768
2006	1.867	1.835	0.003	0.030	1.750
2007	1.916	1.882	0.000	0.034	1.766
2008	1.825	1.782	0.002	0.041	1.668
2009	1.731	1.681	0.044	0.006	1.610

uncertainty analysis and optimal water allocation, requiring the following components. First, the paired observations of yearly upstream water resource U and yearly local water resource L are chosen to establish the joint distribution by copula functions and to obtain water availabilities under the associated probabilities of occurrence. Depending on the joint distribution function, the joint probability of any combinations of U and L can be estimated. Second, the GA takes the available water combinations of upstream and local water supply as input, and outputs the optimal allocation of water resources. In the optimal model, different joint probability levels represent different likelihoods of occurrence and can reflect uncertain water supply, with the purpose of making proper and more productive judgments. Simultaneously, to investigate the acceptable level of violating the constraint and to generate a series of decision alternatives, three conditions corresponding to different acceptable levels were considered. As a result, different water resource regulation schemes should be created that consider upstream water resources, local water resources, and water conveyance losses. The study framework of the paper is shown in Figure 2.

Copula function

To ensure that the data set could capture the temporal runoff process accurately, the paired data from upstream and local water resources during 1954–2009 were collected as samples to build the joint distribution of the two

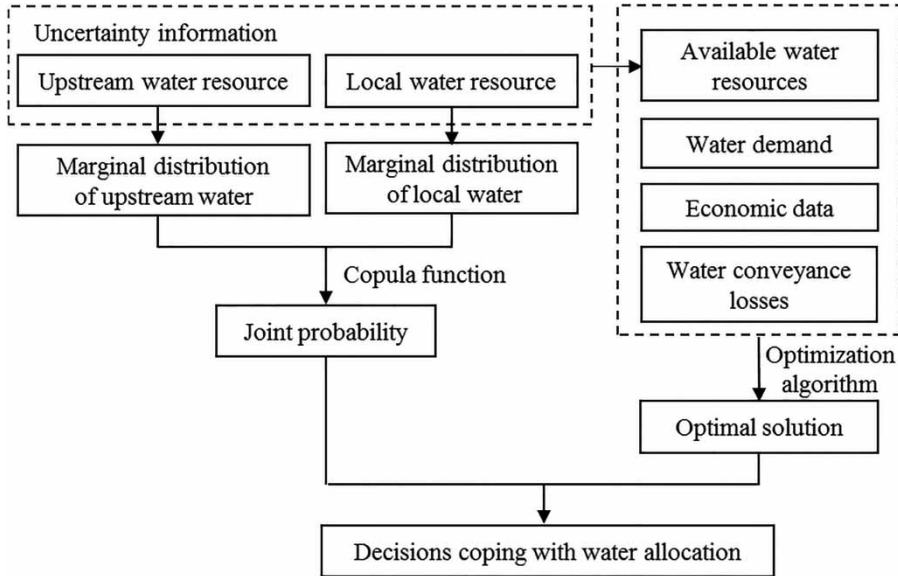


Figure 2 | Study framework of this paper.

variables. The joint impact of upstream and local water resources on the allocation was estimated using joint probability as the evaluation index.

In this study, copulas were used to describe the relevant relationships, which can make it easier to formulate multivariate models compared to other limited and complex models. Copulas are functions that link univariate marginal distribution to form a multivariate distribution. According to Sklar’s theorem, if the marginal distributions $F_U(u)$ and $F_L(l)$ are determined, the joint cumulative distribution function (CDF) F with random variables x and y can be expressed with the function C as follows:

$$F(x, y) = C(F_U(u), F_L(l)) \tag{1}$$

where C is the copula function that connects the marginal and joint distribution. If $F_U(u)$ and $F_L(l)$ are continuous

functions, the joint distribution can be uniquely determined by the marginal distribution and structure correlation. Three widely used one-parameter Archimedean copulas, Clayton, Gumbel, and Frank, are compared to choose the best fit for the study. Applying copulas to problem solving is usually a two-step process when the marginal distribution is determined: estimation of copula parameters followed by selection of a copula using goodness-of-fit tests.

The parameters of the copulas were estimated by Kendall correlation coefficient τ (Zhang 2006). Table 2 shows the expressions of the three copulas and gives the range of parameters θ and the relationship between θ and τ .

The Kendall correlation coefficient τ was defined as:

$$\tau = (C_n^2)^{(-1)} \sum_{i < j} \text{sign}[(x_i - x_j)(y_i - y_j)] \tag{2}$$

Table 2 | Definition of the widely used one-parameter copula families

Copula type	Copula function	Range of parameter	Relationship between θ and τ
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta \geq 0$	$\tau = \frac{\theta}{\theta + 2}$
Frank	$-\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right]$	$\theta \neq 0$	$\tau = 1 - \frac{4}{\theta} \left[1 - \frac{1}{\theta} \int_0^\theta \frac{t}{\exp(t) - 1} dt \right]$
Gumbel	$\exp \left\{ - \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\}$	$\theta \geq 1$	$\tau = 1 - \frac{1}{\theta}$

where (x_i, x_j) is the joint observation values; $sign(\cdot)$ is the sign function if $(x_i - x_j)(y_i - y_j) > 0$, $sign = 1$, if $(x_i - x_j)(y_i - y_j) < 0$, $sign = -1$, and if $(x_i - x_j)(y_i - y_j) = 0$, $sign = 0$.

The next step is to identify an appropriate copula. The root mean square error (RMSE) was used to assess the goodness-of-fit of the probability distributions. RMSE can be expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [p_c(i) - p_0(i)]^2} \quad (3)$$

where $p_c(i)$ denotes the i th theoretical value; n is the number of observations; $p_0(i)$ denotes the i th empirical value calculated according to the following equation:

$$p_0(i) = \frac{(m_i - 0.44)}{(n + 0.12)} \quad (4)$$

where m_i is the number of joint observations meeting the conditions $U < u_i$ and $L < l_i$ in the joint observation sample.

The Akaike information criterion (AIC) (Fan et al. 2016) was employed to select the best fitting copula in this study. AIC can be obtained either by calculating the maximum likelihood or by calculating the mean square error of the model. Thus, AIC is expressed as follows:

$$AIC = n \ln(\text{MSE}) + 2k \quad (5)$$

where k is the number of fitted parameters, and

$$MSE = \frac{1}{n} \sum_{i=1}^n [p_c(i) - p_0(i)]^2 \quad (6)$$

The best fitting copula is the one that has the minimum AIC value. Additionally, the Kolmogorov–Smirnov (K-S) goodness-of-fit test can be applied to evaluate the fitting degree of empirical distribution function and hypothetical overall distribution function. The statistic D was calculated as follows:

$$D = \max_{1 \leq i \leq n} \left\{ \left| F(x_i, y_i) - \frac{m(i) - 1}{n} \right|, \left| F(x_i, y_i) - \frac{m(i)}{n} \right| \right\} \quad (7)$$

where $F(x_i, y_i)$ is the joint distribution value.

When the best fitting copula is chosen, the joint behavior of x and y with some specific values can be calculated. Some probabilities should be focused on, as follows:

1. The probability that x and y both exceed specific values, denoted as $P \cap (x, y)$.
2. The probability that one of the two variables exceeds specific values, denoted as $P \cup (x, y)$.
3. The probability that the random variable falls within a range, denoted as $P(X > x | Y > y)$.

These probabilities were obtained as follows:

$$\begin{aligned} P \cap (x, y) &= P((X > x) \cap (Y > y)) \\ &= 1 - F_x(x) - F_y(y) + F(x, y) \end{aligned} \quad (8)$$

$$P \cup (x, y) = P((X > x) \cup (Y > y)) = 1 - F(x, y) \quad (9)$$

$$P(X > x | Y > y) = \frac{P \cap (x, y)}{P(Y > y)} = \frac{1 - F_x(x) - F_y(y) + F(x, y)}{1 - F_y(y)} \quad (10)$$

Multi-objective optimization

In the case of the optimal allocation of water resources, with the key purpose of allocating water resources scientifically and rationally, managers may wish to obtain better comprehensive effects while weighing the economic benefits, ecosystem environment, social factors, and so on. Optimization has been formulated as a multi-objective problem with many conditions, and the appropriate allocation plans depend on their utilization efficiency and economic efficiency. Therefore, the regulated objectives in this paper included two objectives: the minimization of water shortage and allocation cost, simultaneously. Typically, these objectives are mutually conflicting and restricting: less cost increases shortage, so there is no single optimal solution. The problem in multi-objective optimization analysis is, at its core, the optimization of all objectives at the same time. Therefore, a possible optimal solution that harmonizes all of the objectives, called the Pareto optimal solution, should be obtained.

Proposed algorithm

In this paper, the non-dominated sorting genetic algorithm-II (NSGA-II) is introduced to solve the problems of optimal water resources allocation. As a multi-objective optimization method, the NSGA-II has provided a new way to solve complex problems. The algorithm provides an efficient non-dominated sorting scheme for classifying the population into different fronts and a good diversity preserving mechanism by using crowding distance functions in the population.

In the model, planning water allocation is simulated as evolutionary events, regarding both the amount of water from changing sources and the amount that was assigned to various water users as decision variables. First, an initial solution (which is called population) is randomly created with encoding these optimized variables with binary coding. To ensure the feasibility and diversity of the individuals, the initial population is randomly generated to be feasible according to the constraint conditions. The fitness function, which is transformed by the objective function, makes use of a ranking method to compare the solutions. As a result, the optimum overall performance of an individual can lead to a better fitness value and then provide more opportunities to participate in the evolution. The evolution of a solution (namely population) relies on some genetic operators acting on the current solution to produce a new generation of solutions. The genetic operations include selection, crossover and mutation, and they play a decisive role in the performance of the algorithm. The optimization process iterates until a Pareto optimal allocation solution is obtained. Finally, the best solution can be selected among the alternatives by using the marginal rate of substitution between objectives. The optimal point has the maximum marginal rate of substitution, which corresponds to maximum slope for the two-objective Pareto front.

Thus, the problem can be expressed as follows.

1. Allocation cost:

$$\min f_1 = \sum_{i=1}^I \sum_{j=1}^J \beta_j x_{ijt} c_{ij} \quad (11)$$

2. Water shortage:

$$\min f_2 = \sum_{i=1}^I \sum_{j=1}^J (D_{ij} - \beta_j x_{ijt}) \quad (12)$$

where i denotes water intake; j denotes water user; t represents the various scenarios that the available water is formed by upstream and local water supplies, that is, t represents different joint probability levels; x_{ijt} is the amount of water supply of different intake water i distributed to different water users j under scenario t ; c_{ij} is the water supply cost of different intake water i distributing to different water users j (RMB/m³, ¥); D_{ij} is the fixed target needs of water for user j at water intake i ; β_j represent the different priority of water resources. In the study, the water demand is assumed to be the same in all scenarios.

These were subject to the following.

1. Restriction of the water supply amount:

$$\begin{cases} W_B - W_{1t} - W_{2t} - W_{3t} - W_{4t} + W_{9t} > 0 \\ W_L - W_{5t} - W_{6t} - W_{7t} - W_{8t} + (W_{1t} - w_{1t}) \\ \quad + (W_{2t} - w_{2t}) > 0 \\ W_{it} - w_{it} > 0 \quad i = 3, 4, 5, 6 \\ (1 - \alpha)W_{it} - w_{it} > 0 \quad i = 1, 2, 7, 8 \end{cases} \quad (13)$$

where W_B and W_L are the quantity of upstream and local water supplies, respectively, W_{it} ($i = 1, 2, 7, 8$) is the flow of downstream reach where the water intake i sit under scenario t , which in order are Qinhuang, Yundong, Qingwu, and Qingpai canals; W_{it} ($i = 3, 4, 5, 6$) is the flow of Beiyun and Longfeng Rivers where the water intake i sit under scenario t ; W_{9t} is the allocation water diverted from the Longfeng River to the Beiyun River at the Kuangergang hydro-junction under scenario t ; w_{it} is the allocation water at the water intake i under scenario t ; and α denotes the part of the river flow served as minimum ecological runoff, here it is set to 0.2 (the data comes from the research report of water regulation schemes for river improvement at the lower reaches of Beiyun River).

2. Restriction of water demand amount:

$$B_{j\max} \geq \sum_{i=1}^I x_{ijt} \geq B_{j\min} \quad (14)$$

where $B_{j\max}$ and $B_{j\min}$ are the maximum and minimum water demand amount of water user j .

3. Restriction of water-carrying capacity:

$$Pr \left\{ \sum_{j=1}^J x_{ijt}(1 + \lambda) \leq w_{it} \right\} \geq 1 - q \tag{15}$$

$$\lambda \in (\mu, \sigma^2) \tag{16}$$

In the constraint condition, the constraints were satisfied at a certain probability $(1 - q)$, where q stood for a predetermined probability set by managers, denoting the acceptable level to violate the constraint (Zhuang et al. 2015). This constraint prevents the spill of water by exceeding the conveyance capacity. The water loss rate λ of the constraints was treated as a normally distributed random parameter (μ is the expectation and σ is the standard variation). The water loss rate was $N(0.04, 0.062)$ in this study.

4. Non-negative constraint:

$$x_{ijt} \geq 0 \tag{17}$$

RESULTS AND DISCUSSION

Establishment of joint probability distribution

For upstream water supply U and local water supply L , the marginal distributions of $F_U(u)$ and $F_L(l)$ obeyed a Person type-III (P-III) distribution (Peng & Xu 2010). The probability density function and CDF are:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x - \alpha_0)^{\alpha-1} e^{-\beta(x-\alpha_0)} \tag{18}$$

$$F(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{\alpha_0}^x (x - \alpha_0)^{\alpha-1} e^{-\beta(x-\alpha_0)} dx \tag{19}$$

where $\Gamma(\alpha)$ is an α function and α, β , and α_0 are the shape, scale and position parameters, respectively, of the P-III distribution. For U and L , the parameters of the marginal distribution are estimated on the statistics reported in Table 3.

Table 3 | Estimated parameters of P-III distribution for annual water resource series

Parameters	Upstream water resource		Local water resource	
	Beiyun River	Longfeng River	Beiyun River	Longfeng River
α	1.2484	1.1317	5.1653	3.6281
β	0.0137	0.0128	0.8107	0.4323
α_0	22.1634	11.7630	1.6384	4.1967

The marginal function could be estimated based on the parameters in Table 3. The comparisons of the empirical and theoretical marginal distributions for U and L (Beiyun River) are plotted in Figure 3. The marginal distribution can be acceptable by the K-S test with the significant level of 0.05. From Figure 3, it can also be seen that the theoretical marginal distributions fit the observational data very

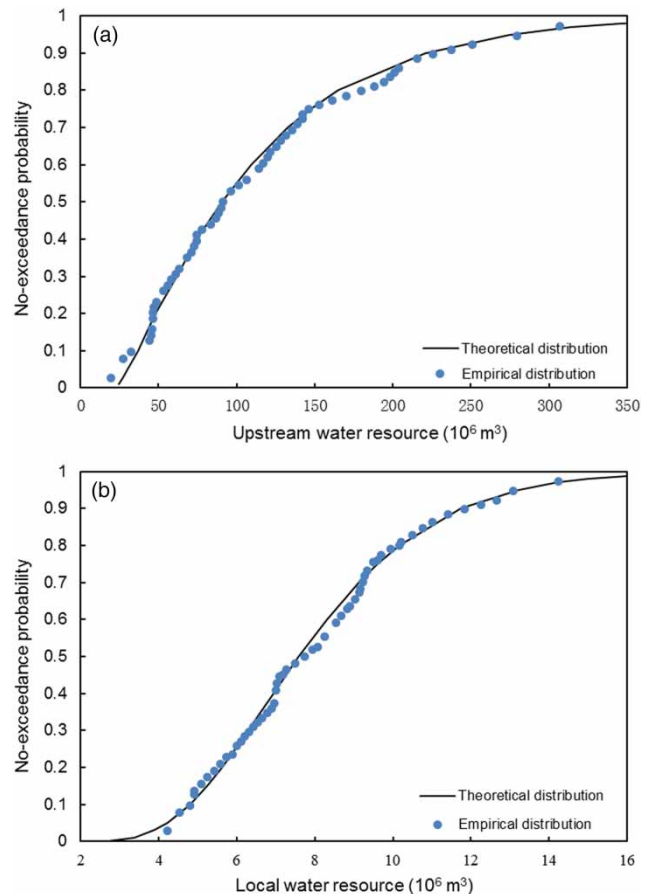


Figure 3 | Comparisons of the empirical and theoretical marginal distributions for the upstream and local water resources (Beiyun River), respectively: (a) upstream water resource; (b) local water resource.

well. According to the CDF of water resources, the exceedance feature probability P_1 or P_2 for any quantity of water can be estimated, and vice versa. For example, for a local water supply quantity of $11.54 \times 10^6 \text{ m}^3$ in the Beiyun River, its exceedance feature probability P_2 in terms of the CDF of L is 20%.

To choose the best fitting copula, the values of Kendall's τ , parameter θ , K-S D and RMSE and AIC of the three copulas (the Clayton, Frank, and Gumbel copulas) given in Table 4 were compared. The dependence of water resources variables also can be ascertained using Kendall coefficient. The values for the Beiyun River and the Longfeng River were 0.5288 and 0.4098, respectively.

The best fitting copula is the one that has the minimum value of RMSE and AIC. The value D for each copula function can pass the K-S test when the level of significance is 5%. Obviously, the RMSE and AIC value of the Clayton copula were the smallest of all of the copulas for both the Beiyun and Longfeng Rivers, which means that the Clayton copula was the most appropriate copula to describe the joint probability distribution of U and L . For example, with the Beiyun River, the AIC value of the Clayton copula of -164.093 was less than -158.812 of the Frank copula and -152.373 of the Gumbel copula.

Then, with reference to the parameters in Table 4 and the equation in Table 2, the Clayton copula constructing the joint distribution $F(x, y)$ can be expressed as:

$$F(x, y) = \left[F_x(x)^{-\theta} + F_y(y)^{-\theta} - 1 \right]^{-1/\theta} \quad (20)$$

For water allocation projects, it is necessary to predict the quantity of total water supply; therefore, the probability

of encountering rich or poor water resources was analyzed. This study noted the joint probability that typical events would occur, that is, the probability that variables would exceed some feature values. For individual probabilities, the design guaranteed efficiency of use of upstream water supply, including 10%, 25%, 50%, 75%, and 95%. Meanwhile, at each level of inflow, four local water supply level probabilities ranging from 5% to 50% were considered. Here, the probabilities reflect that the quantity of water supply exceeded the set points, e.g., medium flows are those with the 50% of exceedance probability. Tables 5 and 6 respectively show 20 types in terms of the established joint distribution model for different combinations of U and L . To intuitively illustrate the connection between the two variables, the joint probabilities ($P \cap (u, l)$) of upstream water supply and local water supply for the two rivers are plotted in Figure 4. Tables 5 and 6 and Figure 4 illustrate that the joint behavior of the Longfeng River appeared to have the same characteristics as the Beiyun River. Thus, the subsequent analysis of allocation quantity mainly focused on the Beiyun River. The correspondence analysis is described in detail below.

Joint impact on water allocation

The model was solved by the solution algorithm described in the section Methodology. The developed model considered variable q in the allocation plans. The variable q , set by managers, reflects the acceptable levels of violating the constraint. In this study, an increased q demonstrates a raised risk of violating constraint and, at the same time, a decreased satisfactory level in meeting the water demand.

Table 4 | Parameters and goodness-of-fit test of copulas

	Kendall correlation coefficient		Clayton	Frank	Gumbel
Beiyun River	0.5288	Parameter θ	2.2445	6.2858	2.1222
		K-S D	0.0832	0.1031	0.1099
		RMSE	0.0329	0.0367	0.0418
		AIC	-164.0929	-158.8177	-153.3733
Longfeng River	0.4098	Parameter θ	1.389	4.298	1.694
		K-S D	0.1036	0.1059	0.111
		RMSE	0.0372	0.0392	0.0416
		AIC	-158.0548	-155.6126	-151.6933

Table 5 | Probabilities of different encounter conditions for Beiyun River

Encounter conditions							
Upstream water resources		Local water resources		$F(u, l)$	$P \cap (u, l) (\%yr^{-1})$	$P(U > u L > l) (\%yr^{-1})$	$P \cup (u, l) (\%yr^{-1})$
$U (10^6 m^3)$	$P_1 (%)$	$L (10^6 m^3)$	$P_2 (%)$				
220.71	10	13.00	5	0.8639	1.3871	27.7427	13.6129
220.71	10	11.23	10	0.8265	2.6538	26.5383	17.3462
220.71	10	9.34	20	0.7483	4.8329	24.1644	25.1671
220.71	10	6.46	50	0.4879	8.7947	17.5893	51.2053
147.51	25	13.00	5	0.7296	2.9563	59.1261	27.0437
147.51	25	11.23	10	0.7075	5.7505	57.5053	29.2495
147.51	25	9.34	20	0.6581	10.8126	54.0628	34.1874
147.51	25	6.46	50	0.4625	21.2458	42.4916	53.7542
90.58	50	13.00	5	0.4944	4.4369	88.7371	50.5631
90.58	50	11.23	10	0.4879	8.7947	87.9465	51.2053
90.58	50	9.34	20	0.4722	17.2165	86.0827	52.7835
90.58	50	6.46	50	0.3859	38.5857	77.1714	61.4143
54.8	75	13.00	5	0.2494	4.9397	98.7943	75.0603
54.8	75	11.23	10	0.2487	9.8688	98.6880	75.1312
54.8	75	9.34	20	0.2468	19.6841	98.4206	75.3159
54.8	75	6.46	50	0.2334	48.3422	96.6844	76.6578
36.9	90	13.00	5	0.1000	4.9969	99.9381	90.0031
36.9	90	11.23	10	0.0999	9.9932	99.9324	90.0068
36.9	90	9.34	20	0.0998	19.9835	99.9177	90.0165
36.9	90	6.46	50	0.0991	49.9066	99.8131	90.0934

Here, three levels of variable q (1%, 5%, and 10%) were considered comprehensively by referring to Equations (15) and (16). Meanwhile, 12 water supply probability levels were considered emphatically, when the design guaranteed efficiency levels of upstream water supply were 25% (at high flow), 50% (at medium flow), and 75% (at low flow). In summary, 36 types of representative scenarios were obtained.

Figures 5 and 6 present the total amount of water allocation and water shortage under different joint probability levels with the same q at 1%. It is observed that various joint probabilities result in different amounts of available water supply, and the water shortages are related to water availability. Deficits occur if the available water supply did not meet the users' demands. In order to meet people's daily requirements, the allotment to the municipality should be administratively guaranteed, regardless of water supply levels, followed by allotment to agriculture, industry,

and ecology. In the case of insufficient water resources, there would be almost no water deficit for the municipal user group. In comparison, for the agricultural users, which had the largest percentage of water demand, allotments could not reach the demand. However, the percentage of agricultural shortage is small relative to industry and ecology. This is because of its low cost, and the lower reaches of Beiyun River is the agricultural district where agriculture accounts for a large share of the national economy. The optimized water demand of agricultural, industry, and ecology will be satisfied in the high flow level period. As seen in Table 1, there is no water left over to allocate to ecological sectors before 2005, for the reason that humans have ignored the ecological environment. Thus, even though the society regularly experiences shortages, the ecological value of water has been more and more valued.

Table 6 | Probabilities of different encounter conditions for Longfeng River

Encounter conditions							
Upstream water resources		Local water resources		$F(u, l)$	$P \cap (u, l) (\%yr^{-1})$	$P(U > u/L > l) (\%yr^{-1})$	$P \cup (u, l) (\%yr^{-1})$
$U (10^6 m^3)$	$P_1 (\%)$	$L (10^6 m^3)$	$P_2 (\%)$				
196.43	10	17.30	5	0.8608	1.0808	21.6169	13.9192
196.43	10	14.51	10	0.8210	2.0981	20.9814	17.9019
196.43	10	11.54	20	0.7394	3.9411	19.7055	26.0589
196.43	10	7.02	50	0.4794	7.9398	15.8795	52.0602
134.41	25	17.30	5	0.7243	2.4349	48.6974	27.5651
134.41	25	14.51	10	0.6977	4.7665	47.6652	30.2335
134.41	25	11.54	20	0.6410	9.1027	45.5134	35.8973
134.41	25	7.02	50	0.4418	19.1787	38.3574	55.8213
83.21	50	17.30	5	0.4901	4.0088	80.1760	50.9912
83.21	50	14.51	10	0.4794	7.9398	79.3977	52.0602
83.21	50	11.54	20	0.4553	15.5345	77.6724	54.4655
83.21	50	7.02	50	0.3536	35.3563	70.7125	64.6437
47.87	75	17.30	5	0.3457	4.5732	91.4649	65.4268
47.87	75	14.51	10	0.3410	9.1040	91.0402	65.8960
47.87	75	11.54	20	0.3301	18.0132	90.0658	66.9868
47.87	75	7.02	50	0.2780	42.8035	85.6070	72.1965
26.96	90	17.30	5	0.0998	4.9783	99.5666	90.0217
26.96	90	14.51	10	0.0995	9.9539	99.5389	90.0461
26.96	90	11.54	20	0.0989	19.8945	99.4723	90.1055
26.96	90	7.02	50	0.0955	49.5493	99.0987	90.4507

Analyses of water shortages for the agricultural user group under various scenarios are presented below. The results for other users could be similarly interpreted. The results indicated that any changes in the joint probability levels would lead to varied water shortages in different cases (denoted as cases 1–12 in the order as in Table 5, $q = 1\%$), as shown in Figure 7. In general, under a certain upstream or local water supply level, the shortage quantity was in proportion to the joint probability and inversely proportional to the supply quantity, which meant that the deficits would decrease with incremental increases in available water. For example, when the upstream water supply level is medium, under the worse case scenario (case 8, joint probability = 38.59%), the allocated water for agricultural users would be $148.73 \times 10^6 m^3$ with local water supply frequency of 50%, indicating a serious shortage ($33.63 \times 10^6 m^3$). In comparison, under a

better case scenario (case 5, joint probability = 4.44%), the value would be $159.05 \times 10^6 m^3$ with local water supply frequency of 5%, also indicating a significant shortage ($23.31 \times 10^6 m^3$, but less serious than that under the worse case scenario). Consequently, the users would have to obtain water from other sources to satisfy demands, at high price, such as underground water and reclaimed water, and then additional costs will arise. On the other hand, for the joint behavior of upstream and local water supplies, the increased value of the encounter probability means that there is an increased crisis of water allocation, in other words, a decrease in available water. For example, as seen in Table 5, the encounter probabilities of $p \cap (u, l)$ with $P_1 = 25\%$ and $P_2 = 5\%$ were 2.96% per year, indicating the higher level of water supply quantities. The encounter probabilities of $p \cap (u, l)$ with $P_1 = 75\%$ and $P_2 = 50\%$ were 48.34% per year, indicating the

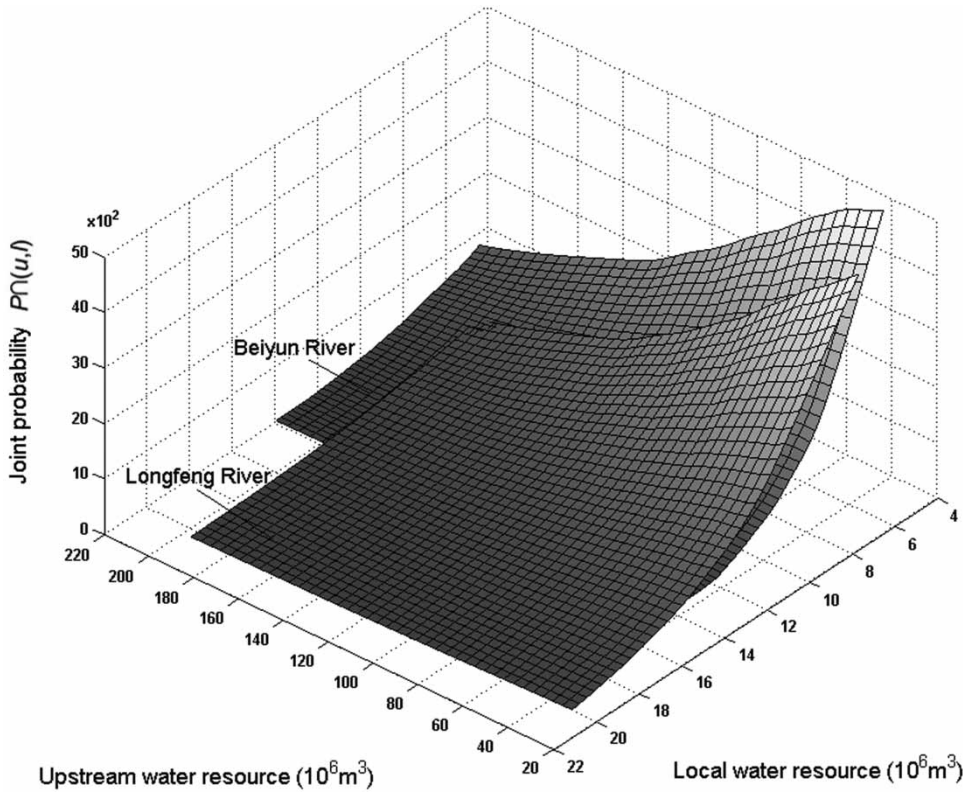


Figure 4 | Joint probability between upstream water resource and local water resource in term of absolute values.

lower level of water supply quantities. Obviously, the frequency of potential drought period occurrences was relatively high.

The complex uncertainties have significant impacts on water shortage, either a joint influence or single factors. From Figure 7, when the upstream water supply was

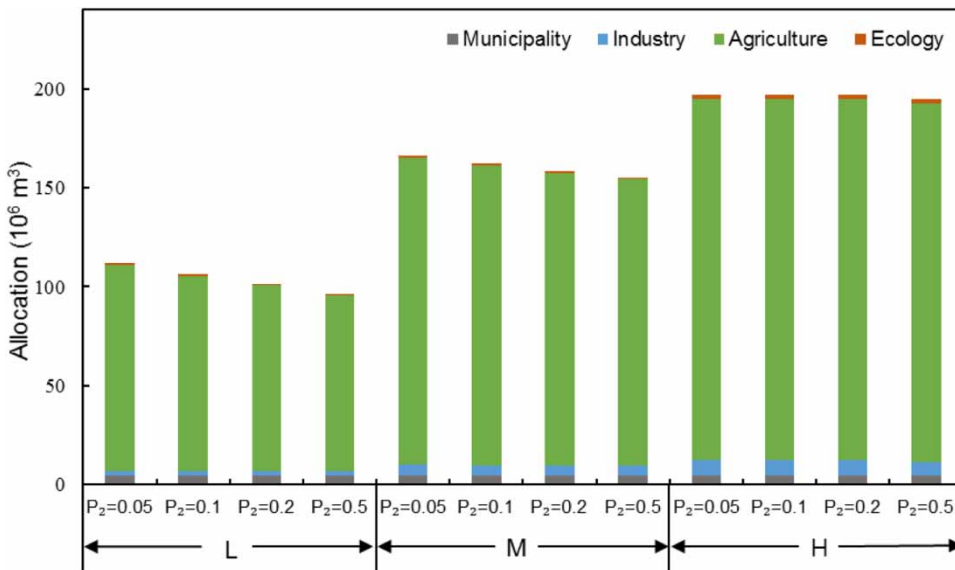


Figure 5 | Optimized water-allocation patterns under different joint probabilities.

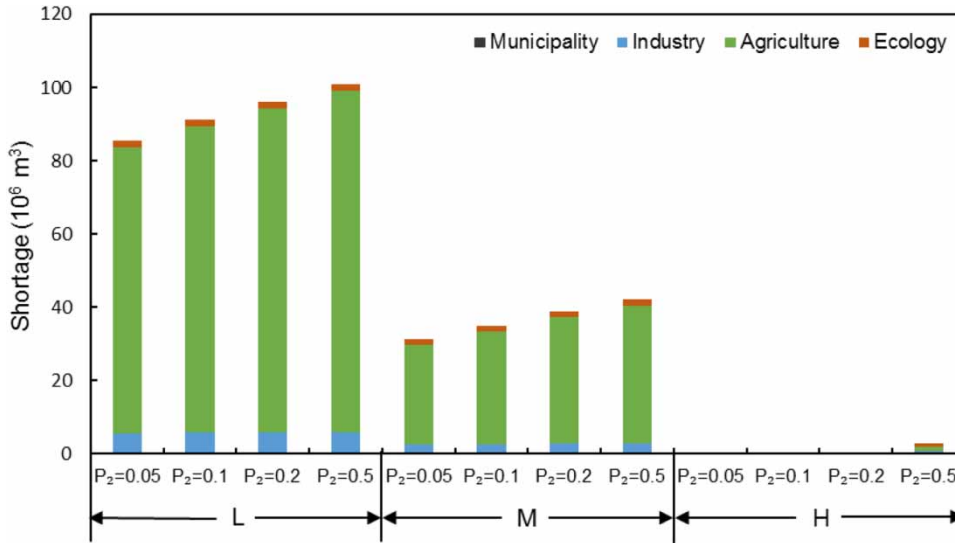


Figure 6 | Water shortage under different joint probabilities.

medium and accompanied by local water supply from 5% (case 5) to 10% (case 6), shortage for agricultural users would increase from $23.31 \times 10^6 m^3$ to $26.94 \times 10^6 m^3$, respectively; when the local water supply frequency was 5%, accompanied by upstream water supply level from medium (case 5) to low (case 9), shortage would vary from $23.31 \times 10^6 m^3$ to $74.21 \times 10^6 m^3$, respectively. This is because the upstream and local water supply quantity have

an effect on the water allocation. Combining Tables 5 and 6, the column $P(U > u | L > l)$ indicates that the exceedance probability of U , P_1 was less than the conditional probability. Thus, a positive dependence existed between U and L , that is, a quantity of upstream water supply was usually accompanied by a similar level of local water supply. It also can be found that the values of $P \cup (u, l)$ were also slightly larger than the exceedance probability of U , P_1 ,

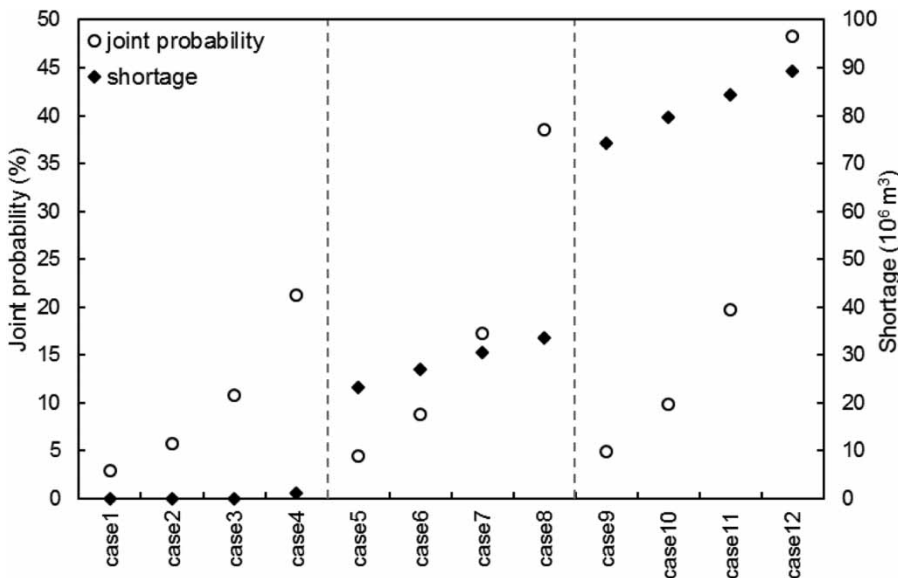


Figure 7 | Different water shortages and joint probability for agricultural users in different cases.

illustrating that local water supply created another factor influencing water allocation. Therefore, by considering as many as possible water-availability scenarios, the developed method can help identify an appropriate water allocation plan that balances various conflicting water allocation goals under extreme conditions.

Additionally, a conclusion could be made from comparing the water deficits at three q levels: a greater water shortage would exist under lower q levels, while a smaller water shortage corresponded to high q levels. Figure 8 shows the different water shortages for agricultural users under changing q levels. It can be observed that when the value of q is 1% and 10% with the joint probability of 4.44%, the shortages for agriculture would be $27.45 \times 10^6 \text{ m}^3$ and $23.31 \times 10^6 \text{ m}^3$, respectively. The constraint limits the possible spillage of water during delivery in the model, and its acceptable level could be adjusted by managers. It can be explained that the relaxations (a higher q , meaning a smaller freeboard in the canals) of constraint would lead to raised available water supply and reduced absolute water losses, and correspondingly smaller water shortage.

Figure 9 presents the trend of the total cost (objective function) under different joint probability levels. Generally, the cost of the system gradually decreases as the available water supply decreases, thus correspondingly higher joint probability. Figure 10 shows the different system costs under changing q levels. The solutions demonstrate that a

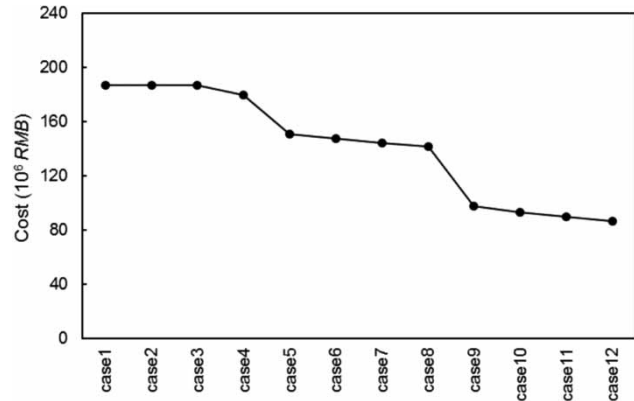


Figure 9 | System cost in different cases.

lower q value would result in a lower system cost and a lower constraint-violation risk, leading to enhancement of the system's stability. For example, the total cost would be $151.00 \times 10^6 \text{ RMB}$, $150.77 \times 10^6 \text{ RMB}$, and $149.24 \times 10^6 \text{ RMB}$ under q levels of 10%, 5%, and 1% when joint probability = 4.44% (the upstream water supply is medium and the probability of local water supply is 5%). From the results, it also reveals that although the cost would have a trend with various q , the q level has few significant impacts on the system objective. Generally, the results demonstrate that high constraint-violation risk and system cost have to be faced if managers aim is less shortage; however, more serious shortage could improve the system reliability and reduce the system cost. A trade-off among the system

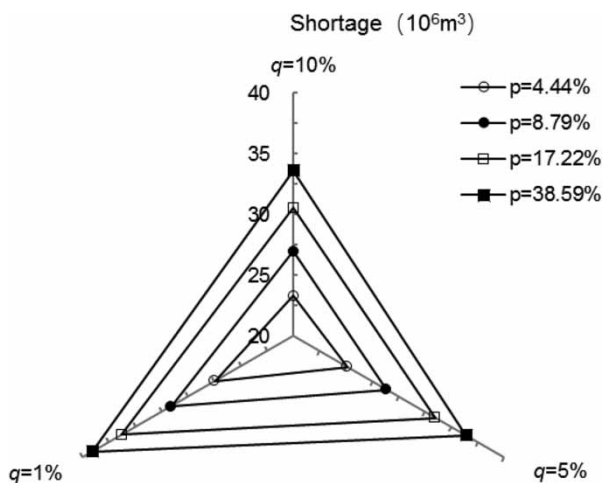


Figure 8 | Different water shortages for agricultural users under changing q levels. Higher q lever demonstrates a raised risk of violating constraint and smaller water shortages.

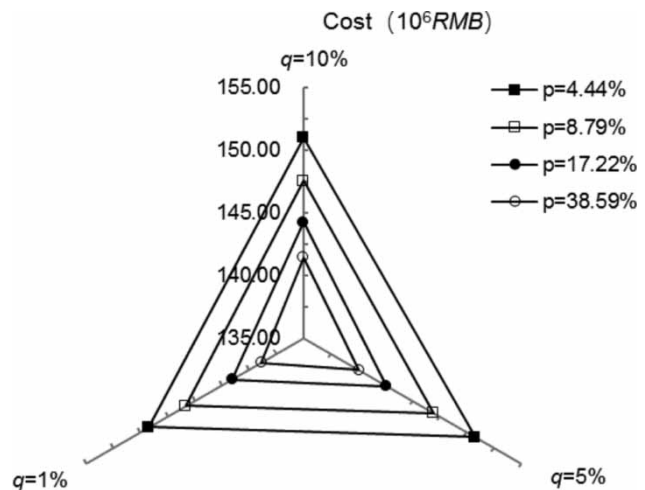


Figure 10 | Different system cost under changing q levels. Higher q level demonstrates a raised risk of violating constraint and higher system cost.

allocation demand, system cost, and reliability could be considered to help managers to obtain an in-depth insight into the water resource management.

Discussion

This paper presents a joint-probabilistic programming method for water resources optimal allocation under uncertainty. Complexities and uncertainties exist in the water resources allocation system; any change in the available water can cause a considerable variation of the model outputs. Therefore, it is indispensable to obtain solutions under different scenarios in allocating water to users, revealing a trade-off among the system allocation demand, system cost economy, and reliability. However, the traditional multi-objective optimization can hardly address uncertain information. By comparison, the joint-probabilistic programming method proposed here has the advantages of tackling joint probability as well as preference to accept risks of water shortage, improving the ability of handling uncertainties in optimization. More specifically, it can: (1) deal with the water resources uncertainties of both upstream and local area supplies, expressed as various joint probabilities that represent different water supply quantities; (2) reflect random uncertainties in the left and right sides of the constraints (the left side of the constraint expressed as the permitted water conveyance loss, and the right side of the constraint expressed as the acceptability of violating the constraint). In conclusion, the optimization model can provide various representative scenarios due to the occurrence of different joint probabilities and the stochastic feature on the coefficients of the constraint.

In other words, the complex uncertain factors, the different water flow levels, and acceptable levels of water shortage, had a significant effect on the modeling outputs. The water flow amount under different levels influences the total water allocation and water shortage, and the cost of per water supply can play a role in determining different water allocation amount for users. Compared with local water resources, the upstream water resource plays a more significant role in the total amount of water allocation. As for acceptability of water shortage, a variability measure has been combined in the constraint, and its effect on the model outputs can be adjusted according to the preference

of managers. This, in fact, reflects that the constraint allowed the spillage of water during delivery in the model. In this study, the model can provide the trade-off among allocation demand, system cost, and reliability by making various water policies. Therefore, different policy scenarios can be obtained to cope with practical situations. Generally, the water shortage would decrease and result in an increased allocation cost as q increases. This would lessen the system feasibility and enhance the system crisis. The joint probability of a rainy year is relatively small, and managers should increase the acceptability of the system crisis to reduce the water shortage; instead, they can implement a compromised scheme in a demanding environment. In many practical problems, the probabilities of the worse case scenarios are relatively high. The occurrence of such demanding environmental conditions, especially in an extreme drought period, will lead to serious consequences due to the low level of water supply amount. Therefore, by considering as many water-availability scenarios as possible, the method can help managers to develop in-depth insights into their water resource allocation systems and to put forth a compromise water allocation plan that balances shortage crisis and system reliability. Managers can modify their allocation plans according to the potential available water, aiming to reduce the shortage crisis and cost of the system.

Generally, although this model is able to effectively address uncertainties, there are several limitations. First, two factors affecting the water availability: upstream and local water resources were considered in this paper. However, the problem can become more complicated when more hydrologic regions and variables are involved. From a broader perspective, the three-dimensional copula can be used to solve high-dimensional copula-based multivariate probability distribution problems. Second, the predetermined distribution approach cannot fully handle water deficits. Particularly in drought periods, when water supply is low, the contradiction between water supply and demand still exists due to limited water availabilities and ever-increasing water demands. Therefore, a more extensive study of the interactions between the multi-water resources system and human activities in dry periods should be undertaken. Third, there are other possible sources of uncertainty in water availability, such as natural climatic variability, hydrological model parameter uncertainty, water demand

uncertainty related to population and economy. Especially for the water demand and cost, they were assumed to be constant throughout the optimization period. Indeed, it is suggested to discuss how these uncertainties have been taken into account for water resources allocation and management in a follow-up study. Furthermore, only typical events were presented in this study. It might be more valuable to generate large-scale scenarios using a Monte Carlo sampling method to provide overall management options. The method is not suitable to evaluate measures that change the system (e.g., the construction of a dam), because the copulas are no longer valid for the modified water system.

SUMMARY AND CONCLUSIONS

In this study, a joint-probabilistic programming optimization model was developed for regional water allocation management under multiple uncertainties. The information provided by this method can provide technical advice to managers. By coupling the copula function with an optimization algorithm, the proposed model was able to address water source uncertainties of both upstream and local water supplies, as described by a copula function with water loss as the constraint. A case of water resources allocation was optimized to demonstrate the applicability of the proposed model. The whole system included two parts, namely, the water source and distribution, containing two water sources and four water user groups. It was able to generate a set of representative scenarios under different probabilities of occurrence. Results show that the two water sources had a common influence on water shortage, as well as preference of managers to accept risks of water shortage. The local water supply level was closely related to the upstream water supply level. For a constant q , an increased joint probability meant an increased risk of water shortage, and more water shortage amounts would exist under lower q levels. Therefore, there is a trade-off among water shortage, economic objective and constraint-violation risk. The lack of significant research of uncertainty in water resources allocation persuaded us to do this research. This paper demonstrated that it is necessary to consider the joint probability of upstream and local water supply when planning water allocation. The various scenarios can help managers

to formulate appropriate water allocation plans according to practical uncertain situations. The formulation of the water allocation problem and its application considering different joint probability levels, is a general formulation which can be extended to other scenarios like design scenarios for strategic planning or climate resilience studies.

ACKNOWLEDGEMENT

The research was supported by the National Science and Technology Support Program of China (2015BAB07B02), Science Fund for Creative Research Groups of the National Natural Science Foundation of China (Grant no. 51621092).

REFERENCES

- Alavian, V., Qaddumi, H. M., Dickson, E., Diez, S. M., Danilenko, A. V., Hirji, R. F., Puz, G., Pizarro, C., Jacobsen, M. & Blankespoor, B. 2009 *Water and Climate Change: Understanding the Risks and Making Climate-Smart Investment Decisions*. World Bank, Washington, DC, pp. 1–174.
- Azamathulla, H. M., Wu, F.-C., Ab Ghani, A., Narulkar, S. M., Zakaria, N. A. & Chang, C. K. 2008 *Comparison between genetic algorithm and linear programming approach for real time operation*. *Journal of Hydro-Environment Research* 2 (3), 172–181.
- Fan, Y., Huang, G., Guo, P. & Yang, A. 2012 *Inexact two-stage stochastic partial programming: application to water resources management under uncertainty*. *Stochastic Environmental Research and Risk Assessment* 26 (2), 281–293.
- Fan, Y. R., Huang, W. W., Huang, G. H., Li, Y. P., Huang, K. & Li, Z. 2016 *Hydrologic risk analysis in the Yangtze River basin through coupling Gaussian mixtures into copulas*. *Advances in Water Resources* 88, 170–185.
- Genest, C. & MacKay, J. 1986 *The joy of copulas: bivariate distributions with uniform marginals*. *The American Statistician* 40 (4), 280–283.
- Gibbs, M. S., Dandy, G. C. & Maier, H. R. 2013 *Assessment of the ability to meet environmental water requirements in the Upper South East of South Australia*. *Stochastic Environmental Research and Risk Assessment* 28 (1), 39–56.
- Grimaldi, S. & Serinaldi, F. 2006 *Asymmetric copula in multivariate flood frequency analysis*. *Advances in Water Resources* 29 (8), 1155–1167.
- Lalehzari, R., Nasab, S. B., Moazed, H. & Haghghi, A. 2016 *Multiobjective management of water allocation to sustainable irrigation planning and optimal cropping pattern*.

- Journal of Irrigation and Drainage Engineering* **142** (1), 05015008.
- Lian, J. J., Xu, K. & Ma, C. 2013 Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: a case study of Fuzhou City, China. *Hydrology and Earth System Sciences* **17** (2), 679–689.
- Liu, D., Guo, S., Chen, X., Shao, Q., Ran, Q., Song, X. & Wang, Z. 2012 A macro-evolutionary multi-objective immune algorithm with application to optimal allocation of water resources in Dongjiang River basins, South China. *Stochastic Environmental Research and Risk Assessment* **26** (4), 491–507.
- Liu, Z., Törnros, T. & Menzel, L. 2016 A probabilistic prediction network for hydrological drought identification and environmental flow assessment. *Water Resources Research* **52** (8), 6243–6262.
- Manshadi, H. D., Niksokhan, M. H. & Ardestani, M. 2015 A quantity-quality model for inter-basin water transfer system using game theoretic and virtual water approaches. *Water Resources Management* **29** (13), 4573–4588.
- Nelsen, R. B. 1999 *An Introduction to Copulas, Volume 139 of Lecture Notes in Statistics*. Springer-Verlag, New York.
- Peng, D. Z. & Xu, Z. X. 2010 Simulating the impact of climate change on streamflow in the Tarim River basin by using a modified semi-distributed monthly water balance model. *Hydrological Processes* **24** (2), 209–216.
- Rani, D. & Moreira, M. M. 2009 Simulation–optimization modeling: a survey and potential application in reservoir systems operation. *Water Resources Management* **24** (6), 1107–1138.
- Reddy, M. J. & Kumar, D. N. 2006 Optimal reservoir operation using multi-objective evolutionary algorithm. *Water Resources Management* **20** (6), 861–878.
- Reddy, M. J. & Kumar, D. N. 2009 Performance evaluation of elitist-mutated multi-objective particle swarm optimization for integrated water resources management. *Journal of Hydroinformatics* **11** (1), 79–88.
- Salvadori, G. & De Michele, C. 2004 Frequency analysis via copulas: theoretical aspects and applications to hydrological events. *Water Resources Research* **40** (12), 1–17.
- Sklar, M. 1959 *Distribution Functions of n Dimensions and Their Margins*. Institute of Statistics, University of Paris, Paris, France, pp. 229–231.
- van den Berg, M. J., Vandenberghe, S., De Baets, B. & Verhoest, N. E. C. 2011 Copula-based downscaling of spatial rainfall: a proof of concept. *Hydrology and Earth System Sciences* **15** (5), 1445–1457.
- Yazdi, J., Salehi Neyshabouri, S. A. A. & Golian, S. 2014 A stochastic framework to assess the performance of flood warning systems based on rainfall-runoff modeling. *Hydrological Processes* **28** (17), 4718–4731.
- Yazdi, J., Lee, E. H. & Kim, J. H. 2015 Stochastic multiobjective optimization model for urban drainage network rehabilitation. *Journal of Water Resources Planning and Management* **141** (8), 04014091.
- Zeng, X. T., Li, Y. P., Huang, G. H. & Liu, J. 2016 Modeling water trading under uncertainty for supporting water resources management in an arid region. *Journal of Water Resources Planning and Management* **142** (2), 04015058.
- Zhang, L. S. V. P. 2006 Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering* **11** (2), 150–164.
- Zhou, Y., Guo, S., Xu, C.-Y., Liu, D., Chen, L. & Ye, Y. 2015 Integrated optimal allocation model for complex adaptive system of water resources management (I): methodologies. *Journal of Hydrology* **531** (3), 964–976.
- Zhuang, X. W., Li, Y. P., Huang, G. H. & Zeng, X. T. 2015 An inexact joint-probabilistic programming method for risk assessment in water resources allocation. *Stochastic Environmental Research and Risk Assessment* **29** (5), 1287–1301.

First received 28 February 2017; accepted in revised form 20 November 2017. Available online 21 December 2017