Study on the joint probability distribution of irrigation water volume and irrigation water efficiency

Yong Zhao, Jinping Zhang and Weihua Xiao

ABSTRACT

Using the copula method, the joint probability distribution of irrigation water volume and efficiency is constructed, and their joint return period is also described to reveal the encounter probability of irrigation water volume and efficiency. Furthermore, the conditional probability of irrigation water efficiency with different water volumes is built to show the quantitative effects of flow on irrigation water efficiency. The results show that the copula-based function can present the encounter risk and conditional probability of irrigation water volume and efficiency at their different magnitudes.

Key words | copula, irrigation water efficiency, irrigation water volume, joint probability distribution

INTRODUCTION

Irrigation water efficiency is an important evaluation index for irrigation water use. There are many factors which have an influence on irrigation water efficiency, including meteorological factors, underlying surface conditions, irrigation water volume, irrigation mode, farming practices and so on. To reduce irrigation water loss and enhance irrigation water use efficiency, many researchers have studied the relationship between irrigation water efficiency and the factors that affect it (Diaz et al. 2004; Stirzaker & Hutchinson 2005; Wang & Zhang 2009; Yilmaz et al. 2009; Ali & Klein 2014; Zhang et al. 2014a, b; Dong et al. 2015). However, these study results mainly focus on either the measurement of irrigation water efficiency or the correlation analysis of irrigation water efficiency and those factors which have an influence on it. There are few studies on the effect of irrigation water volume on irrigation water efficiency.

Irrigation water volume has a profound impact on irrigation water efficiency, but this impact is not direct (Beeson 2006; Matthias & Oliver 2014). Irrigation water volume can change the water use activities and water management behaviors, and then affect irrigation water efficiency. Usually, reduced irrigation water volume means less water for agricultural irrigation. For this reason, water-saving irrigation techniques are used widely in agricultural water use management, which results in a higher irrigation water efficiency (Olga et al. 2009).

For an irrigation district relying heavily on water from a local river, the volume of irrigation water is dependent on the river flow. The river flow almost determines the volume of irrigation water available. The greater the flow, the more water can be used for irrigation, so the irrigation water efficiency may be lower. When irrigation planning is conducted, the irrigation water volume needs to be predicted. Currently, the prediction of river flow is fairly accurate, but the characteristics and probability distribution of irrigation water efficiency are still unknown, let alone the relationships of river flow and irrigation water efficiency.

Moreover, river flow and irrigation water efficiency are random variables with complex correlations, but the commonly used bivariate joint distribution models (such as exponential distribution, Gumbel logistic distributions,
etc.) assume that the two univariates have the same marginal distributions, and their correlations must be linear. To solve this problem, the copula method based on nonlinear relationships among variables is introduced to construct the bivariate joint distribution without considering the marginal distribution of univariates, and it is also used widely in the hydrologic field (Balistrocchi et al. 2011; Gyasi-Agyei 2012; Gyasi-Agyei & Melching 2012; Lian et al. 2013; Ma et al. 2013; Zhang et al. 2014a, b; Yu et al. 2014).

The aims of this paper are to: (1) build the joint probability distribution of volume and irrigation water efficiency; (2) construct the return period of volume and irrigation water efficiency; and (3) present the conditional probability of irrigation water efficiency.

**MATERIAL AND METHODS**

**Data sources**

The study was conducted in the Jinghuiqu irrigation district, which is located in the middle of the Guanzhong Plain of Shaanxi Province in China. It is also a large-scale gravity irrigation district with effective irrigated farmland of 8.8 × 10⁴ hm². Water used for irrigation in Jinghui is mainly from the Jinghe River (which is the second tributary of the Yellow River), and the water diversion project is situated at the downstream end of the Jinghe River. So, the Jinghe River flow can be considered as the available irrigation water volume for the Jinghuiqu irrigation district.

The data sources involve the Jinhe River flow and irrigation water efficiency in the Jinghuiqu irrigation district. The data series of the Jinghe River flow comes from the Zhangjiashan hydrological observation station. The Zhangjiashan hydrological observation station, also situated at the water diversion project of the Jinghuiqu irrigation district, is a control station of the lower reaches of the Jinghe River. So, the observation data series of the Jinghe River flow shows the irrigation water volume for the Jinghuiqu irrigation district. The Jinhe River flow data series is a daily one, and the annual flow data series can be achieved by summing up these daily flow series. The data series of the irrigation water efficiency is obtained from the Jinghuiqu Irrigation District Administration Bureau. Every year, the Jinghuiqu Irrigation District Administration Bureau evaluates the irrigation water efficiency and puts the data on file. Therefore, the annual flow series and the annual irrigation water efficiency series are usually calculated at the end of every year.

The data series of the Jinhe River flow, and irrigation water efficiency in the Jinghuiqu irrigation district from 1980 to 2007 are shown in Figure 1.

**Method of analysis**

**Sklar theorem**

The copula method is introduced in this study, but the Sklar theorem (Sklar 1959) is the basis of the copula theory and it proves the uniqueness of copula existence. Suppose there are two random variables \( X \) and \( Y \), \( F(x, y) \) is their joint distribution function, \( F_X(x) \) and \( F_Y(y) \) are their marginal distribution functions. If \( F_X(x) \) and \( F_Y(y) \) are continuous, a copula function \( C(u, v) \) can be determined uniquely as

\[
F(x, y) = C_\theta(F_X(x), F_Y(y)), \quad \forall \ x, y
\]

where, \( C_\theta(u, v) \) is called a copula function, and \( \theta \) is a parameter to be determined.

**Properties of a copula function**

A bivariate copula function is defined as a mapping \( C: [0, 1]^2 \rightarrow [0, 1] \) with the following properties:

(a) \( \forall u, v \in [0, 1], C(u, 0) = 0; \ C(0, v) = 0; \ C(u, 1) = u; \ C(1, v) = v; \)
(b) \( \forall u_1, u_2, v_1, v_2 \in [0, 1], u_1 \leq u_2, v_1 \leq v_2, C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \)
(c) \( \forall u, v \in [0, 1], \max (u + v - 1, 0) \leq C(u, v) \leq \min (u, v) \)
Copula functions

The one-parameter Archimedean copula families of three types are widely used in the hydrology field (Cherubini et al. 2004; Grimaldi et al. 2005): Clayton function, Frank function and Gumbel–Hougaard function, their mathematical expressions are as follows:

Clayton function

\[
C_{\theta}(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0
\]  

(2)

Frank function

\[-\frac{1}{\theta} \ln \left[ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right], \quad \theta \in \mathbb{R}
\]  

(3)

Gumbel–Hougaard function

\[
\exp \left[ -((\ln u)^\theta + (\ln v)^\theta)^{1/\theta} \right], \quad \theta \geq 1
\]  

(4)

Correlation of copula function

The Kendall’s correlation coefficient \( \tau \) is also calculated to determine the parameter \( \theta \). The calculation formula of Kendall’s correlation coefficient \( \tau \) is

\[
\tau = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{sign}[(x_i - x_j)(y_i - y_j)]
\]  

(5)

where,

\[
\text{sign}[(x_i - x_j)(y_i - y_j)] = \begin{cases} 
1 & (x_i - x_j)(y_i - y_j) > 0 \\
0 & (x_i - x_j)(y_i - y_j) = 0 \\
-1 & (x_i - x_j)(y_i - y_j) < 0
\end{cases}
\]

Actually, according to the relations of \( \tau \) and \( \theta \), the parameter \( \theta \) can be obtained using the following equations:

Clayton function

\[
\tau = \frac{\theta}{\theta + 2}
\]  

(6)

Frank function

\[
\tau = 1 - \frac{4}{\theta} \left[ -\frac{1}{\theta} \int_0^\theta \exp \left( \frac{t}{\theta} \right) - 1 \, dt - 1 \right]
\]  

(7)

Gumbel–Hougaard function

\[
\tau = 1 - \frac{1}{\theta}
\]  

(8)

Identification and goodness-of-fit evaluation of copula function

Sometimes, more than one available copula function is achieved. The Kolmogorov–Smirnov (K-S) test is used to identify the appropriate copula function, and the OLS method is applied to evaluate the goodness-of-fit of the copula function. Eventually, the copula function is selected by the minimum of OLS.

The identified statistic of the K-S test is given by

\[
D = \max_{1 \leq k \leq n} \left\{ C_k - \frac{m_k}{n}, \frac{m_k - 1}{n} - C_k \right\}
\]  

(9)

where, \( C_k \) is the value of observed \( (x_k, y_k) \) of the copula function, and \( m_k \) is the number of observed \( (x_k, y_k) \) satisfying \( x \leq x_k \) and \( y \leq y_k \).

The OLS is calculated in the following expression:

\[
\text{OLS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - P_{ei})^2}
\]  

(10)
where, $P_i$ is the calculated frequency of the joint probability distribution, $P_{ei}$ is the empirical frequency of the joint probability distribution and OLS is the minimum deviation square.

**Return period**

In a hydrological event, the exceedance probability of any design variable is often concerned. The return period is just used to determine the design value of a hydrological variable. According to the selected copula function, the marginal distribution return period of random variables $X$ and $Y$ are obtained as follows:

$$
\begin{align*}
T_x &= 1/(1 - F_X(x)) \\
T_y &= 1/(1 - F_Y(y))
\end{align*}
$$

(11)

For the return period of two variables, two kinds of joint return periods are considered as follows:

$$
\begin{align*}
T_{xy} &= 1/(1 - F(x, y)) \\
T_{xy} &= 1/(1 - F_X(x) - F_Y(y) + F(x, y))
\end{align*}
$$

(12)

where, $T_{xy}$ is the joint return period, indicating the exceedance risk of any design variable; $T_{xy}$ is the simultaneous return period, indicating the exceedance risk of all design variables.

**Conditional probability distribution**

Given the different situation of $X$, the condition probability distribution of $Y$ is expressed as follows:

$$
F_{Y|X}(X, Y) = P(Y \leq y | X \leq x_1) = \frac{F(x_1, y)}{F_x(x_1)}
$$

(13)

$$
F_{Y|X}(X, Y) = P(Y \leq y | x_1 \leq X \leq x_2) = \frac{F(x_2, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)}
$$

(14)

$$
F_{Y|X}(X, Y) = P(Y \leq y | X \geq x_2) = \frac{F(y) - F(x_2, y)}{1 - F_x(x_2)}
$$

(15)

**Analysis procedures**

In practice, the analysis procedures include the following steps: (1) determine the univariate marginal distribution functions; (2) calculate the Kendall’s correlation coefficient $r$; (3) determine $\theta$ according to the relations of $r$ and $\theta$ in the three types of copula functions; (4) select the appropriate copula function with the evaluation index to construct the bivariate joint distribution function; (5) conduct the return period analysis and conditional probability distribution with the constructed bivariate joint distribution function.

**RESULTS AND DISCUSSION**

**Joint probability distribution of volume and irrigation water efficiency**

The P-III-type distribution is used to determine the marginal distribution of volume and water use coefficient. Then the Kendall’s correlation coefficient is calculated as $-0.079$. In the three types of the one-parameter Archimedean copula families, only the Frank function can present the negative correlation of two variables, so the Frank function is selected to describe the joint probability distribution of volume and irrigation water efficiency. According to the previous mentioned formulae, the parameter $\theta$ of the Frank copula function is calculated as $-0.7146$, so the selected Frank copula function is expressed as

$$
F(x, y) = -\frac{1}{0.7146} \ln \left[ 1 + \frac{(e^{-0.7146x} - 1)(e^{-0.7146y} - 1)}{e^{-0.7146} - 1} \right]
$$

(16)

Using the K-S test, the identified statistic $D = 0.1204$. With the significant level of $\alpha = 0.05$, the corresponding fractile is 0.2499, which is larger than 0.1204, so the Frank copula function can pass the K-S test, and the OLS is calculated as 0.03001.

Figure 2 shows the scatter point distribution of the Frank copula-based calculation values and the empirical probability calculation values. It can be seen that all the scatter points fall near the 45 degree diagonal line and their
correlation coefficient is 0.9863. As the Frank copula-based calculation probability distribution has the better fit with the empirical probability distribution, the selected Frank copula function is reasonable, and it can be used to describe the joint distribution of volume and irrigation water efficiency well.

Joint return period of volume and irrigation water efficiency

Given the combination schemes of volume and irrigation water efficiency, the marginal distribution of volume $F(x)$ and irrigation water efficiency $F(y)$ can be obtained by the P-III-type distribution, and their joint probability distribution $F(x, y)$ is achieved by Equation (16). Thus, the joint return $T_{xy}$ and the simultaneous return period $T_{xy}$ are all determined with Equation (12). The contour maps for volume and irrigation water efficiency of the joint return periods and simultaneous return period are shown in Figures 3 and 4.

From Figures 3 and 4, the return period of marginal distribution of volume and irrigation water efficiency can be obtained at the point where the contour and the axes intersect. Meanwhile, the joint return period with various encounter situations of volume and irrigation water efficiency can be found easily. For example, when the volume is $12.92 \times 10^8$ m$^3$ and the irrigation water use coefficient is 0.562, their joint return period is 1.75 years and their simultaneous return period is 6.9 years. Therefore, the probability of various encounter situations of volume and irrigation water efficiency is obtained quantitatively. Usually, the return period of the marginal distribution of univariates ranges between the joint return period and the simultaneous return period, while the latter two return periods can be considered as two extreme situations of marginal distribution. Therefore, the interval estimation of the actual return periods is achieved by the joint return periods and the simultaneous return periods. As shown in Table 1, when the return period of the marginal distribution of univariates is
The frequency of irrigation water efficiency when the Jinghe River volume is low, normal or high is as follows:

As the Jinghe River volume has a great impact on the irrigation water efficiency, the conditional probability effects of volume and irrigation water efficiency is less than the exceedance probability of any design variables. The differences between design periods and joint return periods are also less than that between design periods and simultaneous return periods. Moreover, the difference tends to increase with the design return period.

**Conditional probability distribution of irrigation water efficiency**

As the Jinghe River volume has a great impact on the irrigation water efficiency, the conditional probability effects of volume on irrigation water efficiency should be considered. The frequency of $pf = 37.5\%$ and $pk = 62.5\%$ is often used as the classification of the high or low situation of volume. The conditional probability distribution of irrigation water efficiency when the Jinghe River volume is low, normal or high is as follows:

$$F_{Y|X}(X, Y_i) = P(Y_i \leq y | X \leq x_{pl}) = \frac{F(x_{pl}, y) - F(x_{pl}, y)}{1 - F_x(x_{pl})}$$  \hspace{1cm} (17)

$$F_{Y|X}(X, Y_i) = P(Y_i \leq y | x_{pl} \leq X \leq x_{pl})$$

$$= \frac{F(x_{pl}, y) - F(x_{pl}, y)}{F_x(x_{pl}) - F_x(x_{pl})}$$  \hspace{1cm} (18)

According to the marginal distribution $F(x)$ and $F(y)$ of volume, irrigation water efficiency, and their joint probability distribution $F(x, y)$, the various values of $F(x_{pl}, y)$ and $F(x_{pl}, y)$ can be calculated if the different irrigation water efficiency $y$ is given. So, the conditional probability distribution of irrigation water efficiency is also determined.

**Figure 5** describes the conditional probability distribution of irrigation water efficiency. The occurrence probabilities of different irrigation water efficiency values under different water resources conditions is seen clearly in Figure 5. It can also be seen that with the same value of irrigation water efficiency, the conditional probability of irrigation efficiency in a low volume situation is always larger than in a high volume situation. It illustrates that water volume has an impact on the irrigation water efficiency, and lower flow results in more effective irrigation water use, which is consistent with the general irrigation water use laws in irrigation districts. In practice, if the volume of water available in the irrigation district is low, the irrigation water efficiency is expected to be enhanced by some measures according to water-saving irrigation planning, including improving water use management and adjusting the crop planting programme.

Although the water resources situations of irrigation districts change to some extent, it can be seen that in the given

---

**Table 1** Design return period, joint return period and simultaneous return period of flow and irrigation water efficiency

<table>
<thead>
<tr>
<th>Design return period (year)</th>
<th>Volume (108 m³)</th>
<th>Irrigation water efficiency</th>
<th>Joint return period (year)</th>
<th>Simultaneous return period (year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>18.59</td>
<td>0.57</td>
<td>1.04</td>
<td>1.43</td>
</tr>
<tr>
<td>1.5</td>
<td>15.23</td>
<td>0.55</td>
<td>1.04</td>
<td>2.16</td>
</tr>
<tr>
<td>2</td>
<td>12.92</td>
<td>0.54</td>
<td>1.37</td>
<td>3.67</td>
</tr>
<tr>
<td>5</td>
<td>9.35</td>
<td>0.52</td>
<td>2.85</td>
<td>20.21</td>
</tr>
<tr>
<td>10</td>
<td>7.93</td>
<td>0.51</td>
<td>5.35</td>
<td>76.36</td>
</tr>
<tr>
<td>20</td>
<td>6.98</td>
<td>0.50</td>
<td>10.35</td>
<td>295.82</td>
</tr>
</tbody>
</table>

---

**Figure 5** | The conditional probability distribution of irrigation water efficiency.
conditions, the conditional probabilities in these three water resources conditions all tend to be equal when the irrigation water efficiency is up to a certain level. This means the influence that the level of water resources exerts on the irrigation water efficiency has some limits. If it is required to continuously improve the water irrigation efficiency, some other measures should be taken.

Normally, the water resources available in the future can be predicted fairly accurately. Thus, with the conditional probability distribution of irrigation water efficiency, the possibility of the expected irrigation water efficiency can be obtained. This provides the quantitative basis to improve the water use management. Moreover, compared with the irrigation water efficiency in different flow situations, the irrigation water use level or water loss can also be evaluated, which will enable irrigation water to be used more effectively.

CONCLUSIONS

With the Frank copula function, the joint probability distribution of volume and irrigation water efficiency is put forward. It reveals the encounter probability and return period of volume and irrigation water efficiency. The bivariate copula function can describe the relations of volume and irrigation water efficiency in the Jinhuiqu irrigation district well, and has no restriction to the univariate marginal distribution functions. Therefore, the copula function is a potential optimal algorithm for the probability analysis of random variables.

The return period with various encounter situations of volume and irrigation water efficiency is also demonstrated. With the same design return period, the joint return period and simultaneous return period of volume and irrigation water efficiency are different, and the simultaneous return periods are greater than the joint return periods. This shows that the exceedance risk of all design variables is lower than the exceedance risk of any one design variable. If an appropriate distribution process of volume and irrigation water efficiency of a typical year is obtained, several groups of volume and irrigation water efficiency with the same return period can also be achieved with the same frequency amplification method. Furthermore, the conditional probability of irrigation water efficiency with different volume situations is built to show the quantitative effects of volume on irrigation water efficiency. The irrigation water efficiency in a low volume situation is always greater than in a high volume situation. Therefore, the water resources condition of the Jinhuiqu irrigation district has an impact on the irrigation water efficiency. Usually, this impact is exerted on the irrigation water efficiency by water use activities and water management levels. When the volume is low, the irrigation water efficiency will certainly be improved. With the conditional probability of irrigation water efficiency with different volume situations, the possibility of improving irrigation water efficiency is apparent. Thus, a feasible plan should be made for irrigation water use in accordance with the volume of irrigation water available.

ACKNOWLEDGEMENTS

This research is supported by the National Natural Sciences Foundation of China (Project No. 51309202, 51379216, 51279183) and the Open Research Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Grant no. IWHR-SKL-201402, and also by the Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No. 13IRTSTHN030).

REFERENCES


First received 21 August 2014; accepted in revised form 5 March 2015. Available online 23 March 2015