

# Probabilistic analysis of the controls on groundwater depth using Copula Functions

Yifei Bai, Yi Wang, Yaning Chen and Leyuan Zhang

## ABSTRACT

Groundwater is an essential water resource in the Yarkant River Basin Irrigation District, which is the largest oasis in Xinjiang, China. This study used a novel approach to analyze the relationship between groundwater depth and three driving factors by developing Copula Functions from monthly time series data collected from 16 monitoring wells. More specifically, marginal distribution and joint distribution functions were established, and the conditional probabilities for three data ranges were calculated using both two- and three-dimensional Copula Functions. The developed statistical models showed that groundwater exploitation, runoff, and surface water irrigation significantly affected groundwater depth. The most significant effect on water table declines was associated with groundwater exploitation and lagged 1-month behind the groundwater withdrawals. The amount of runoff and irrigation were both inversely related to water table depth due to groundwater recharge. Frank Copula Functions were found to produce the best fit to the joint distribution of the variables and were used herein, allowing for the establishment of a detailed probability distribution of the change in groundwater depth under the combined effect of multiple controlling factors. The results provided key insights into how irrigation and groundwater withdrawals can be adjusted to effectively manage groundwater resources.

**Key words** | conditional probability, Copula Functions, groundwater depth, groundwater management, Yarkant River Basin

Yifei Bai

Yi Wang (corresponding author)

Yaning Chen

Leyuan Zhang

State Key Laboratory of Desert and Oasis Ecology,  
Xinjiang Institute of Ecology and Geography,  
Chinese Academy of Sciences,

Urumqi 830011,  
China

E-mail: [hywy02@foxmail.com](mailto:hywy02@foxmail.com)

Yifei Bai

Leyuan Zhang

University of Chinese Academy of Sciences, Beijing  
100049,  
China

## INTRODUCTION

The Yarkant River is the longest tributary to the Xinjiang Tarim River, which is the largest inland river in China (Chen *et al.* 2010). The basin is located in central Asia and possesses a temperate continental arid climate (Sun *et al.* 2010; Zhang *et al.* 2010) that is located between the Taklimakan Desert and the Bulgari and Togak Deserts. Although the Yarkant River Basin is subject to low and irregular rainfall, high temperatures and evaporation, as well as notable periods of drought (Duethmann *et al.* 2016), it possesses a desert oasis consisting of fragile ecosystems. It also possesses the Yarkant River Basin Irrigation District (YRBID),

which is the largest agricultural irrigation district in Xinjiang. The irrigation district is the most important area in China for the production of high-quality grain, cotton, and fruits (Chen *et al.* 2007), and agricultural water consumption is significant (Deng *et al.* 2015). The basin is a multi-ethnic area dominated by the Uighurs, and the minority population accounts for 89.85% of the total population of the basin. The agricultural population represents an important proportion of the population.

The population and cultivated land area within the Yarkant River Basin have increased significantly in recent years and have consequently increased stress on its water resources and environment. In fact, shortages in water resources have severely restricted urban and agriculture development in the area (Zhang *et al.* 2012). To help alleviate

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

doi: 10.2166/nh.2020.147

these shortages, groundwater has become an important source of water (Zhang *et al.* 2012; Xiao *et al.* 2015).

Previous research has shown that groundwater resources can help maintain the social and economic development of the oasis, by offsetting the water requirements for agricultural irrigation and domestic use (Zhou & Li 2013). However, the increased exploitation of groundwater in the area has accelerated the rate at which the depth to groundwater has increased (Wang *et al.* 2008). Moreover, research has shown that fluctuations in groundwater levels may affect hydrological processes in the entire basin, causing changes in surface water and groundwater interactions (e.g., groundwater recharge and discharge), as well as ecosystem dynamics (Mansuer *et al.* 2011). Thus, groundwater is an essential domestic, agricultural, and ecosystem resource that must be effectively managed (Du *et al.* 2016; Feng 2016).

During the latter part of the 20th century, the protection and sustainable utilization of groundwater resources in arid areas has gradually become a major concern (Mogaji *et al.* 2015). A few previous groundwater resource analyses in the Yarkant River Basin attempted to document the spatial distribution and causes of groundwater declines. These analyses were based on (1) the use of Linear Regression Methods applied to measured time series data of the changes in groundwater (water table) depths (Chen *et al.* 2016), (2) the use of the Mann–Kendall method to analyze the temporal and spatial evolution of regional groundwater systems (Shataer 2017), (3) the use of the SWAT model to predict changes in the depth to the groundwater table and to analyze the dynamic response of regional groundwater under climate change (Liang 2017), (4) the use of Monte Carlo type algorithms to predict changes in groundwater dynamics (Gulupiya 2017b), (5) the use of Wavelet analyses to determine changes and trends in water table depths (Liu *et al.* 2009), and (6) the calculation and analysis of the groundwater levels required to maintain sound ecological conditions in the area (Gulupiya 2017a). All of these studies mainly focused on annual and/or interannual variations in the depth to groundwater and helped investigators define and predict fluctuations in water table levels through time. However, most of these investigations result in a point prediction in groundwater levels, rather than providing the probability of the range in changes in water table depths using specific statistical distributions. From the point

of view of rational groundwater resources management, it is important to clarify the probability relationships between groundwater depth and the factors that determine the probability distribution of groundwater depth.

The Copula Function is an expression that links the joint cumulative distribution function of variables with the marginal cumulative distribution function of variables to explore the correlation between variables (Genest & Favre 2007). Univariate hydrological analyses cannot fully characterize the multiple controls on hydrological events, including changes in groundwater depth. Thus, the study of the frequency at which hydrological events occur by analyzing the joint distribution of multiple controlling variables has become an important focus of hydrological research (Schumann 2017).

Since the first application of the Copula Functions to the field of hydrology in 2003 (Michele 2003), they have been broadly used in hydrology to gain insights into, as well as model the relationship between, two or more parameters. Their application includes the analysis of precipitation (Khedun *et al.* 2014), the correlation between flood peak and flood volume (Szolgay *et al.* 2016), the determination of the frequency and recurrence interval of floods (Haberlandt & Radtke 2014; Stamatatou *et al.* 2018), the analysis of the frequency and recurrence intervals of droughts (De Michele *et al.* 2013; Salvadori & De Michele 2015), the assessment of multi-hazard risks (Salvadori *et al.* 2016), and environmental hydrological modeling (Vezzoli *et al.* 2017). Theoretically, groundwater depth and its controlling factors can also be analyzed using Copulas Functions, thereby providing a probability perspective of the correlative structure between the depicted variables. However, due to the unique nature of the groundwater table, the application of Copula Functions to the analysis of groundwater depth is a relatively difficult process, and there are few related studies. Nonetheless, You *et al.* (2018) used Copula Functions, along with annual time series data to analyze the correlations between groundwater depth and its controlling factors within the Jinghui Irrigation District of central China. Their research found that Symmetric Frank Copula methods were able to satisfactorily describe the probability correlations between the water table depth and the controlling factors including annual surface water irrigation and precipitation. However, they ignored seasonal fluctuations

in groundwater level and the two considered controlling factors may be unsuitable for the arid area where rainfall is a minor component for groundwater recharge.

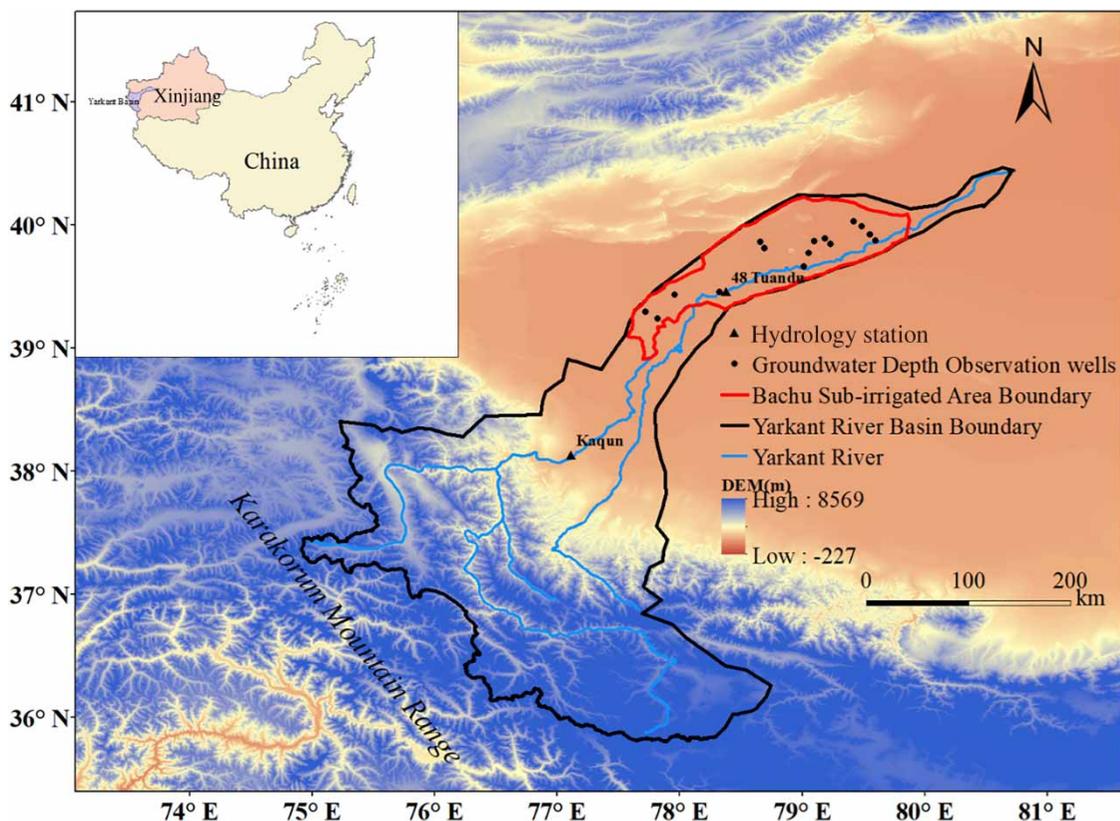
Herein, this paper expands on this earlier work by developing an explicit probabilistic Copula model of the structure of groundwater depth in an arid area, while considering more than two driving factors. More specifically, the objective of this study was to characterize the variations in groundwater depth and determine the factors controlling the depth to the water table in an arid oasis. The analyses were conducted by calculating the relevant probabilities of the controlling parameters, documenting the probability of an occurrence of groundwater depth over a range of values, and determining the appropriate probabilistic relations between the groundwater depth and the controlling variables. This paper specifically established suitable Copula Functions between groundwater depth and selected driving factors to analyze the responsiveness of the different factors as well as the combined effect of multivariate driving

factors on the probability of changes in groundwater depth by using shorter time interval (monthly) time series data and investigating the correlative structure between them. By analyzing and verifying the characteristics of probability distributions of groundwater depth, the approach provides a useful reference for groundwater resource management in irrigation districts in arid regions.

## STUDY AREA AND DATA

### Introduction to the irrigation area

The YRBID (76°40'–79°25'E to 37°20'–40°10'N) is located in the southwestern part of the Xinjiang Uygur Autonomous Region (Figure 1). It encompasses a total area of about  $3.9 \times 10^4$  km<sup>2</sup>. To the east of the irrigation district lies Takla Makan desert that is positioned on the east side of the Yarkant River. To the west of the district is the eastern



**Figure 1** | Location of the study site in China, and location of the groundwater monitoring wells within the study basin.

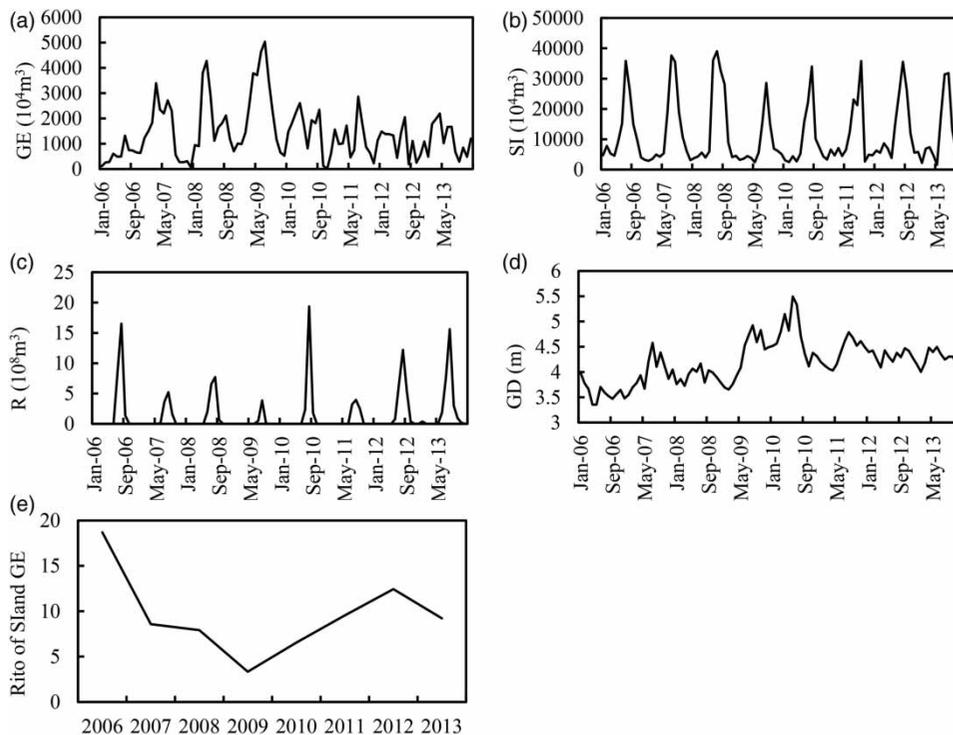
boundary of Jiashi County. The Kunlun Mountains are located to the south and north of the district. Within the study area, the average annual precipitation measures about 53.14 mm, whereas the annual potential evaporation is much larger, measuring about 2,196 mm. The annual average temperature is about 12 °C, which is typical of an inland arid climate. Bachu County is located to the north of the Yarkant River irrigation district, whose change of groundwater depth is largest in the Yarkant River irrigation area. Agriculture, industry, and household water are most dependent on groundwater resources in Bachu County.

Agriculture is the main economic activity in the region, and agricultural production depends on irrigation. Water for irrigation is derived from a combination of wells (groundwater) and canals. The irrigation area seldom produces runoff. Rather, surface water resources mainly come from snowmelt in the surrounding mountains. The uncertainty in surface runoff has led to overexploitation of groundwater resources. In addition, data show that cultivated areas within the irrigation district increased by  $7.015 \times 10^4$  ha, while population increased by 7.784 million in the past

decade. The ratio of groundwater use to total water use increased from 7.3% to 18.4% from 2006 to 2013. The ratio between surface water irrigation (SI) use and groundwater exploitation (GE) amount in the study area is shown in Figure 2(e). Groundwater exploitation, surface water irrigation, and runoff referred to GE, SI, and R, respectively.

## Data

The measurements of groundwater depth data used in the paper were collected from 16 wells in the Bachu Irrigation District between 2006 and 2013, runoff data were collected from the local hydrological station 48 TuanDu (Figure 1). A study area is located in the downstream parts of the YRBID. Monthly data series were used in this study; the monthly averaged data of groundwater depth, groundwater exploitation, surface water irrigation, and runoff were collected by and obtained from the Kashgar Hydrology Bureau and the Yarkant River Basin Authority. Figure 2 shows the temporal variations in these four variables over the monitoring period. Since the study area is in an arid region where the



**Figure 2** | Time series showing monthly variations in (a) groundwater depth, (b) groundwater exploitation, (c) surface water irrigation, (d) runoff, and (e) ratio of SI and GE between 2006 and 2013.

annual precipitation (~50 mm) is much less than the evaporation (~2,300 mm/year), groundwater recharge by precipitation is negligible. Therefore, only groundwater exploitation, surface water irrigation, and runoff (primarily from snowmelt) were considered as controlling factors for the groundwater depth analysis.

## METHODS

Groundwater depth and its controlling factors have complex interrelationships, especially for short time intervals. In order to establish proper Copula Functions between them, data preprocessing is needed to meet the required random and steady state. The procedure used to develop two- and three-dimensional (2D and 3D) Copula Functions is displayed in Figure 3 and described in detail below.

### The Kendall rank correlation coefficient method

The Kendall rank correlation coefficient, commonly referred to as the Tau coefficient, is a non-parametric test used to measure the relationship between variables. It is used in this paper to test for hysteresis relationships between the groundwater depth and the potential driving factors.

Specifically, for the time series of groundwater depth, where  $X_0 = (x_0(t))$ ,  $t = 1, 2, \dots, n$ , different time lags were assigned to the time series of the controlling factors  $X_i$ , such that  $\varepsilon = 0, 1, 2, \dots, k$ . The time series of the controlling factors with time lags can be obtained, where  $X_{i\varepsilon} = (x_i(t - \varepsilon))$ ,  $\varepsilon = 0, 1, 2, \dots, k$ . The Kendall rank correlation coefficient method with time lags is expressed as follows:

$$T_i = 1 - \frac{6 \sum (X_0 - X_{i\varepsilon})^2}{n(n^2 - 1)}, \quad (1)$$

where  $T$  ranges between  $-1$  and  $1$ . When  $T$  is  $1$ , the groundwater depth is completely consistent/varies with the driving factor, which is considered as a random variable. When  $T$  is  $-1$ , the groundwater depth varies indirectly with the driving factor. When  $T$  is  $0$ , the coefficient indicates that groundwater depth and the driving factor are independent of each other.

### The ARIMA model

The steady state of time series data is the premise of establishing Copula Function. The ARIMA model can transform a non-stationary time series into a stationary

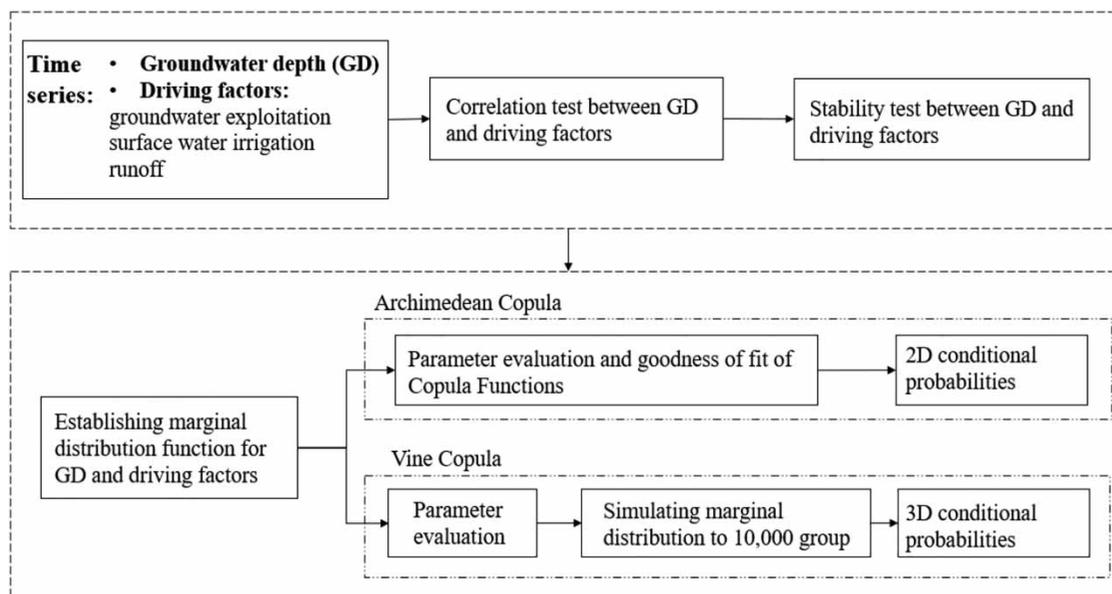


Figure 3 | Schematic diagram of the steps used for the establishment of 2D and 3D Copula Functions.

series (Momani & Nail 2009). After preprocessing the time series input data, the main steps involved in the development of an ARIMA model include pattern recognition, model parameter estimation, and performance testing (Zheng et al. 2015). The objective of pattern recognition is to analyze and test the time series data to select the most appropriate ARIMA process on the basis of the characteristics of the time series data (trend or seasonal, etc.). The intent of model parameter estimation is to determine three parameters in the model, including the autoregressive parameter ( $p$ ), the number of differencing passes ( $d$ ), and the moving average parameter ( $q$ ) (Gharde et al. 2016).

The equation used in the ARIMA model of order ( $p, d, q$ ) for a standard normal variable ( $Z_t$ ) is as follows (Choubin & Malekian 2017):

$$\varphi(B)(1 - B)^d Z_t = \theta B \varepsilon_t. \quad (2)$$

In Equation (2),  $\varphi(B)$  and  $\theta(B)$  are polynomials of degree  $p$  and  $q$ , respectively:

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p, \quad (3)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (4)$$

Through parameter evaluation, the optimal results of parameter  $d$ ,  $p$ , and  $q$  used in the ARIMA model were 2, 0, and 14, respectively. In this study, the parameters of the ARIMA model were determined based on the information criterion function method (Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC)) to obtain the optimal values of  $p$  and  $q$  (Chen 2013).

## Theory of the Copula Functions

### 2D Copula Function

Copula Functions were firstly used by Sklar in the 1950s (Sklar 1959). The Copula Function models the nonlinearity, symmetry, or asymmetry of the dependence structure of the variables. It can be used to construct a multidimensional joint distribution function by using a marginal distribution and correlation structure (Salvadori & De Michele 2007). Copula methods represent one approach from a cohort of

multivariate methods that are widely used to model the dependence structure of two (or more) random variables. Sklar's theorem states that when  $x_1, x_2, x_3, \dots, x_n$  are random variables, and  $F_1(x_1), F_2(x_2), F_3(x_3), \dots, F_n(x_n)$  are a marginal distribution function of them,  $F$  is the joint probability distribution function of the  $n$ -dimension. Therefore, there must exist a Copula Function  $C$ , which makes the following formula true:

$$F(x_1, x_2, \dots, x_n) = C[F_1(x_1), F_1(x_2), \dots, F_n(x_n)]. \quad (5)$$

where  $C$  represents the Copula Function. The Copula  $C$  captures the dependence among the variables, irrespective of their marginal distributions. If the marginal distribution functions  $F_1(x_1), F_2(x_2), F_3(x_3), \dots, F_n(x_n)$  are continuous, then the Copula Function will be unique (Sklar 1959).

Copula Functions exist in various forms. The Archimedean Copula Functions are one of the most commonly used 2D Copula Functions in hydrology (Renard & Lang 2007). These functions include the Gumbel-Hougaard (GH), Clayton, and Frank methods (Meintanis et al. 2015). These three functions were chosen for the analysis in this study; the mathematical expression of their cumulative distribution functions is shown in Table 1.

### 3D Copula Function

Vine Copula (with Archimedean pairs) was selected to construct the 3D Copula Functions (Panagiotelis et al. 2012) as it can analyze the correlation between multiple variables, and there was no need to consider the relationship between them as either positive or negative. The schematic diagram of the 3D-Vine Copula structure is shown in Figure 4 (Aas et al. 2009).

**Table 1** | 2D Archimedes Copula Functions

Copulas	Cumulative Distribution Function	Parameters
Gumbel-Hougaard	$\exp\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}\}$	$\theta$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\theta \geq 1$
Frank	$-\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right)$	$\theta \neq 1$

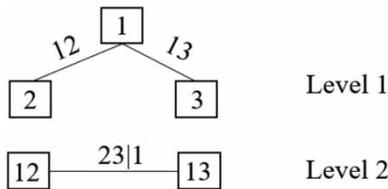


Figure 4 | Schematic diagram of the 3D-Vine Copula structure.

The Vine Copula structure is established by combining groundwater depth with any two driving factors. The joint distribution function of the 3D Copula Function under the C Vine Copula decomposition structure is given as follows:

$$f(x_1, x_2, x_3) = f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot C_{12} \cdot C_{23} \cdot C_{23|1}. \quad (6)$$

In the above expression, each Vine Copula Function can be decomposed into the product of a 2D Copula Function and an edge distribution function. The unknown parameters in the structure are generally estimated by the maximum likelihood method (Brechmann & Schepsmeier 2013).

### Establishing a marginal distribution function

Before constructing a Copula Function, the marginal distribution of multivariate variables must be calculated. The constructed marginal distribution is important because it directly affects the accuracy of the joint distribution function (Liu et al. 2018). The parametric estimation method is a commonly used approach to population probability density function analysis, and it was used in this study for constructing the marginal distribution (Xiong et al. 2015). The parameters of the marginal distributions are estimated through a maximum likelihood method. Twenty probabilistic functions (Table 2) were used to construct the marginal distribution functions for groundwater depth and its controlling factors. The Chi-square test was employed to statistically evaluate the goodness of fit of the marginal distribution function model, and the most suitable function for each variable was selected.

### Parameter evaluation and goodness of fit of Copula Functions

Markov chain Monte Carlo (MCMC) simulation within a Bayesian framework was used to calculate the parameters

Table 2 | List of empirical distribution function types

Empirical distribution function types			
Beta	Generalized extreme value	Log-normal	$t$ location-scale
Birnbaum-Saunders	Generalized Pareto	Nakagami	Weibull
Exponential	Inverse Gaussian	Normal	Binomial
Extreme value	Logistic	Rayleigh	Negative binomial
Gamma	Log-logistic	Rician	Poisson

contained in the 2D Copula Function (Madadgar et al. 2017; Sadegh et al. 2017), whereas the AIC method was used to test the goodness of fit of the distribution function (Daneshkhah et al. 2016).

AIC was calculated as follows:

$$AIC = n \log \left( \frac{RSS}{n} \right) + 2m, \quad (7)$$

where  $n$  is the number of samples, RSS is the sum of residual squares, and  $m$  is the number of parameters.

### Conditional probability distribution

For 2D variables, the marginal distribution functions of  $X$  and  $Y$  are expressed as follows (You et al. 2018):

$$F_x(x) = P(X \leq x) = u, \quad F_y(y) = P(Y \leq y) = v, \quad (8)$$

where  $X$  is the times series of controlling factors,  $Y$  is the times series of groundwater depth, and  $F$  is a marginal distribution of the time series. The conditional probability of event occurrence is given in the following equation:

$$P(y_1 < Y < y_2 | x_1 < X < x_2) = \frac{P(x_1 < X < x_2, y_1 < Y < y_2)}{P(x_1 < X < x_2)} = \frac{C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1)}{(u_2 - u_1)}. \quad (9)$$

The 3D conditional probability is calculated as follows:

$$P(z_1 < Z < z_2, x_1 < X < x_2, y_1 < Y < y_2) = \frac{P(x_1 < X < x_2, y_1 < Y < y_2, z_1 < Z < z_2)}{P(z_1 < Z < z_2)}, \quad (10)$$

where  $X$  and  $Y$  are the times series of controlling factors, and  $Z$  is the times series of groundwater depth.

## RESULTS AND DISCUSSION

### Correlation between groundwater depth and the driving factors

Groundwater exploitation, surface water irrigation, and runoff (referred to GE, SI, and R, respectively) were chosen as the driving factors to analyze the change in groundwater depth (GD) within the Bachu Irrigation District of Yarkant River Basin. Since some driving factors controlling groundwater depth may not be manifested in the same month, it is adoptable to analyze the lag effects of driving factors on groundwater depth to make the subsequent research effective.

The Kendall rank correlation coefficient method was employed to analyze the lag effects of GE, SI, and R on groundwater depth. The significance of cross-correlation between the driving factors and groundwater depth was tested by varying the lag time between 0 and 6 months (Table 3). When the groundwater depth was set to a lag of 1 month, there was a significant correlation with groundwater exploitation ( $p < 0.05$ ). When the lag time was 0, the statistical significance level of the correlation coefficient between runoff, irrigation and groundwater depth ranged between 0.05 and 0.1. This lack of significant correlation suggests that there is no obvious lag effect between them in the area. Since the groundwater monitoring wells were not used for groundwater exploitation, it is reasonable to assume that there was a 1-month lag between groundwater exploitation and a change in groundwater depth at the monitoring wells. Therefore, this study used  $GE_{+1Mo}$  (indicates

that groundwater exploitation occurs 1 month (Mo) ahead (+) of the changes in groundwater depth), SI, and R time series data as driving factors to analyze the probability distribution of groundwater depth.

### Stability test of driving factors and groundwater depth

As noted above, Copula models require that the time series for the used input variables are in a steady state (Xiong et al. 2015). It is, therefore, necessary to test the stability and autocorrelation of the time series data before using Copula Functions to establish their joint distribution. The four utilized time series datasets possessed strong autocorrelation and periodicity (Figure 5) and therefore required preprocessing to generate a stable time series for Copula modeling.

The ARIMA model was used herein to smooth non-stationary time series data, thereby generating stationary time series (Momani & Nail 2009). The autocorrelation and periodicity of the groundwater depth and driving factors data series were significantly reduced (Figure 6), creating a stable time series that could be used in the subsequent modeling efforts.

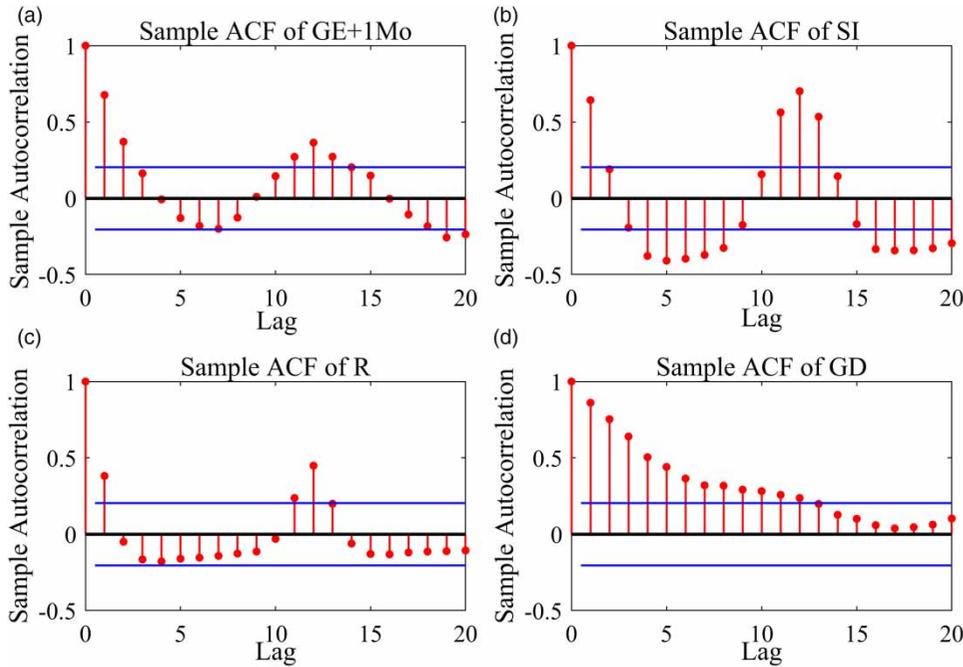
### Establishing marginal distribution functions

Selecting an appropriate marginal distribution function for each variable is required to effectively construct the 2D Copula joint distribution function. Marginal distributions were preliminarily constructed by 20 empirical distribution functions (Table 2). Subsequently, Chi-square tests were applied to estimate parameters of the marginal distribution function and assess the goodness-of-fit. Twenty probability distribution functions were selected as candidates to develop the marginal functions for groundwater depth

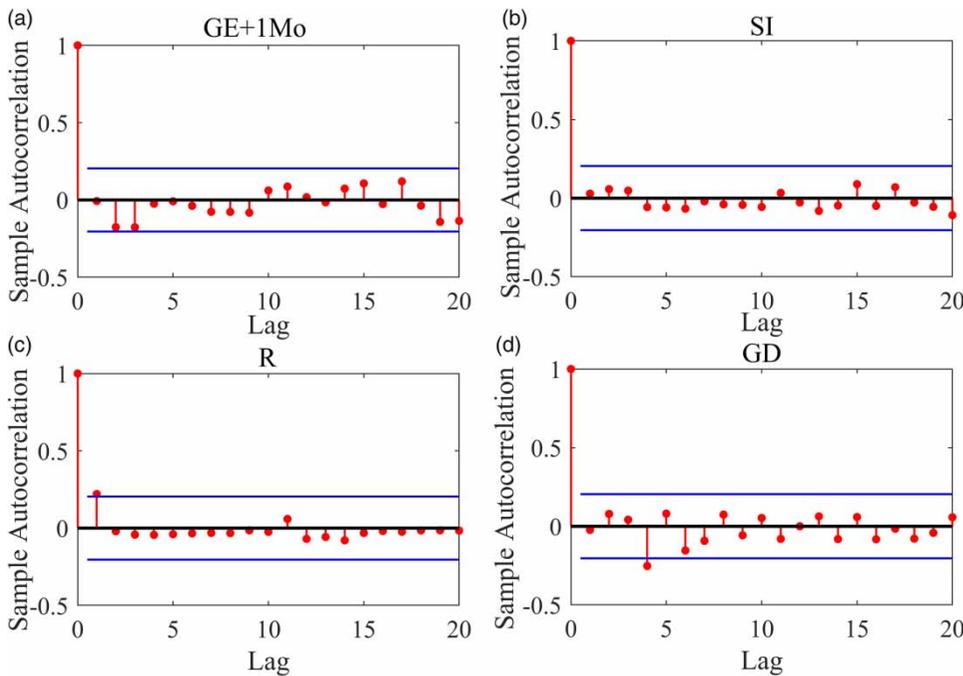
**Table 3** | Correlation test between groundwater depth and driving factors

	lag0	lag1	lag2	lag3	lag4	lag5	lag6
GE-GD	-0.06	0.21**	0.12	0.05	0.13	-	-
SI-GD	-0.13*	-0.07	-0.05	-0.09	0.00	0.00	0.07
R-GD	-0.15**	-0.02	0.10	0.11	0.04	-0.03	0.00

Note: GE is groundwater exploitation, SI is surface water irrigation, R is runoff, GD is groundwater depth, lag0-lag6 is driving factor lagging 0-6 months, respectively. \*\* $p < 0.05$ , \* $p < 0.1$ .



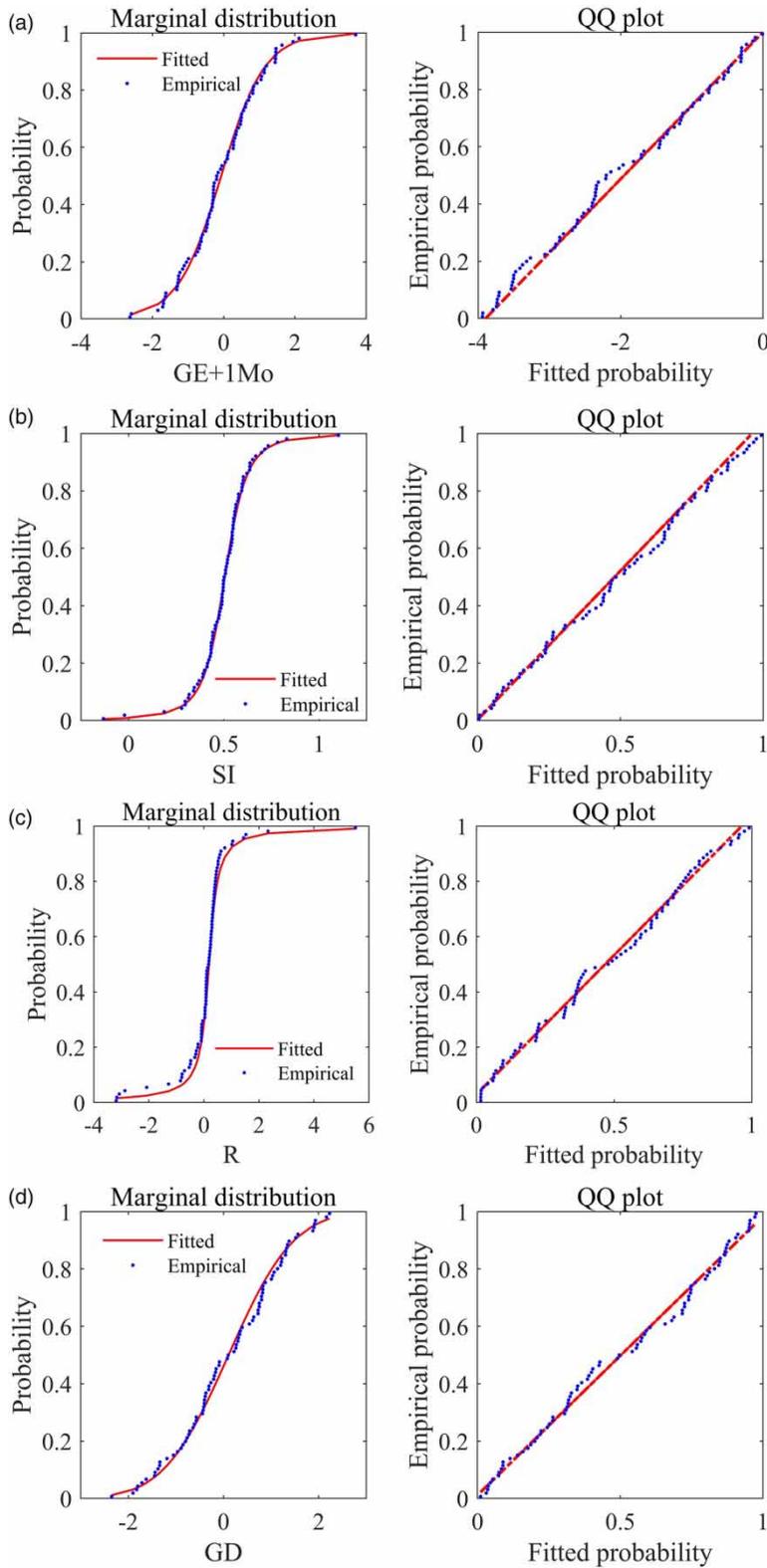
**Figure 5** | Autocorrelation test of (a) groundwater exploitation 1 month ahead, (b) surface irrigation, (c) runoff, and (d) groundwater depth.



**Figure 6** | Autocorrelation test of (a) groundwater exploitation 1 month ahead, (b) surface irrigation, (c) runoff, and (d) groundwater depth after ARIMA model processing.

(GD) and the driving factors. The optimal distribution types of marginal functions for GD,  $GE_{+1Mo}$ , SI, and R were the normal distribution, logistic distribution,  $t$  location-scale

distribution, respectively. The Chi-square test showed that the marginal distribution functions of GD,  $GE_{+1Mo}$ , SI, and R all passed the 0.05 significant level (Figure 7).



**Figure 7** | Marginal distribution curves (left) and quantile–quantile (QQ) plots (right) of (a) groundwater exploitation 1 month ahead, (b) surface irrigation, (c) runoff, and (d) groundwater depth.

## 2D Joint distribution functions

After determining the appropriate marginal distribution function for groundwater depth and its driving factors, 2D Copula Functions were constructed between them. Groundwater depth ( $X$ ) and three driving factors ( $Y$ ) were arranged in ascending order, and then, data pairs were selected from  $X_i \leq X_j, Y_i \leq Y_j$  ( $i < j = 1, \dots, n$ ) to calculate the empirical frequencies of the joint distribution functions. The marginal distribution of groundwater depth with the marginal distribution of each driving factor was then combined to construct the joint distribution. Gumble, Clayton, and Frank Copula Functions were used to fit the joint distribution functions between GD and  $GE_{+1Mo}$ , SI, and R. Copula parameters were estimated by using a maximum likelihood method. The goodness of fit was evaluated by AIC. The smaller the AIC value, the better the fit of the model. Thus, the Frank Copula Function produced the best result (Table 4) and is suitable to construct the joint distributions between  $GE_{+1Mo}$ -GD, SI-GD, and R-GD pairs. The joint distribution function of GD and  $GE_{+1Mo}$  is given as follows:

$$C_1(u_1, u_2) = \frac{1}{-2.15} \ln \left[ 1 + \frac{(e^{-2.15u_1} - 1)(e^{-2.15u_2} - 1)}{(e^{-2.15} - 1)} \right], \quad (11)$$

where  $u_1 = P(X \leq x)$ ,  $u_2 = P(Y \leq y)$ ,  $X$  is groundwater depth, and  $Y$  is groundwater exploitation.

The joint distribution function for GD and SI is given in the following equation:

$$C_2(u_1, u_3) = \frac{1}{1.06} \ln \left[ 1 + \frac{(e^{1.06u_1} - 1)(e^{1.06u_3} - 1)}{(e^{1.06} - 1)} \right], \quad (12)$$

**Table 4** | 2D Copula Function goodness of fit evaluation data

Data pair	Copula Function type	$\theta$	AIC
$GE_{+1Mo}$ -GD	Gumbel	0.62	-611.40
	Clayton	1.29	-609.90
	Frank	2.15	-618.51
SI-GD	Gumbel	0.00	-616.77
	Clayton	1.00	-616.79
	Frank	-1.06	-617.42
R-GD	Gumbel	1.00	-640.79
	Clayton	0.00	-640.81
	Frank	-1.36	-641.92

where  $u_1 = P(X \leq x)$ ,  $u_3 = P(Y \leq y)$ ,  $X$  is the depth of groundwater, and  $Y$  is the amount of surface water irrigation.

The joint distribution function of GD and R is written as follows:

$$C_3(u_1, u_4) = \frac{1}{1.36} \ln \left[ 1 + \frac{(e^{1.36u_1} - 1)(e^{1.36u_4} - 1)}{(e^{1.36} - 1)} \right], \quad (13)$$

where  $u_1 = P(X \leq x)$ ,  $u_4 = P(Y \leq y)$ ,  $X$  is groundwater depth, and  $Y$  is runoff.

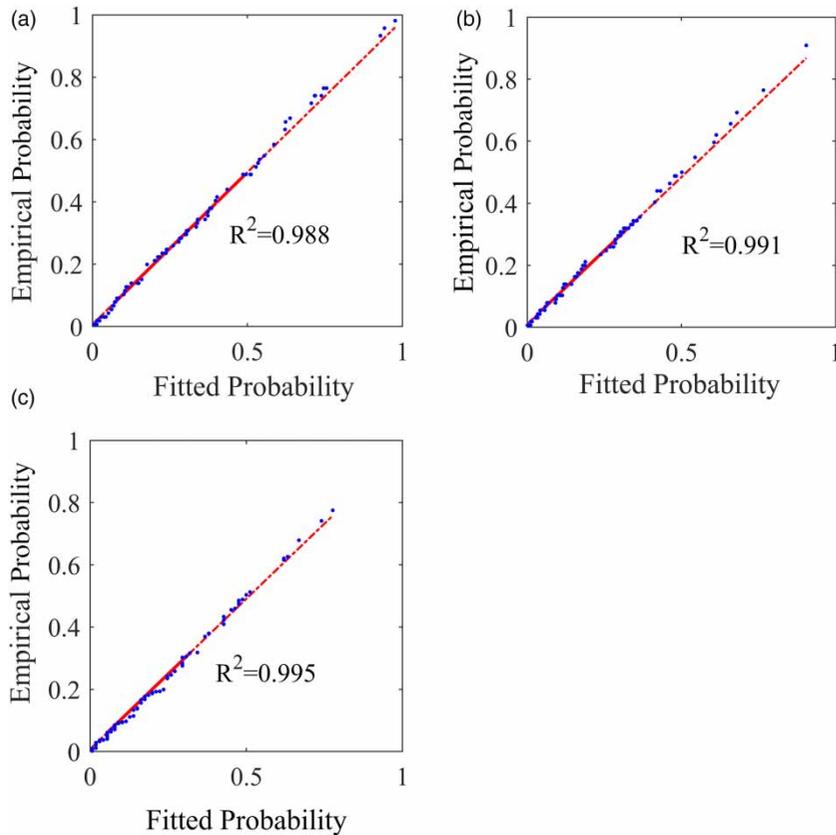
The joint probability distributions between groundwater depth and three driving factors were examined using quantile-quantile (QQ) plots (Figure 8). The correlation coefficient ( $R^2$ ) of joint distribution functions between GD and  $GE_{+1Mo}$ , SI, and R were 0.988, 0.991, and 0.995, respectively. These coefficients indicate that Frank Copula Functions developed between GD and  $GE_{+1Mo}$ , SI, and R exhibit a high level of fit, suggesting that Frank Copula Functions should be used in the study area.

The results from the above analysis and the  $\theta$  parameter (shown in Table 4) show that there was a positive correlation between groundwater exploitation and groundwater depth, while surface water irrigation and runoff were negatively correlated with groundwater depth. The absolute values of the parameters within the Frank Copula Functions are  $|\theta_{GE_{+1Mo}}| > |\theta_R| > |\theta_{SI}|$ , indicating that the correlation between groundwater exploitation and groundwater depth was the largest, followed by runoff and surface water irrigation.

## 2D Conditional probabilities

Conditional probabilities were developed by initially dividing the variable values for depth of groundwater below the surface, groundwater exploitation, surface water irrigation, and runoff into three ranges. Subdividing the data limited the effect of range length on the constructed probabilities. The probability values for different conditions were then calculated using Equations (8) and (9) (Tables 5-7).

As previously noted, there is a 1-month lag between the timing of groundwater exploitation and groundwater depth. Thus,  $GE_{+1Mo}$  was used to analyze the probability distribution of the variations in groundwater depth. The probability distribution of groundwater depth under the



**Figure 8** | 2D Copula joint distribution function fitting QQ plots between (a) groundwater exploitation 1 month ahead, (b) surface irrigation, and (c) runoff and groundwater depth.

**Table 5** | Conditional probability of water table depth falling into different ranges given a specific condition (range) of groundwater exploitation

<b>P(GD GE<sub>±1Mo</sub>)</b>	<b>0 &lt; GE<sub>±1Mo</sub> &lt; 1,677.71</b>	<b>1,677.71 &lt; GE<sub>±1Mo</sub> &lt; 3,355.41</b>	<b>3,355.41 &lt; GE<sub>±1Mo</sub> &lt; 5,033.12</b>
3.35 < GD < 4.07	0.36	0.19	0.07
4.07 < GD < 4.78	0.51	0.51	0.40
4.78 < GD < 5.49	0.13	0.30	0.53

Note: GE<sub>±1Mo</sub> is groundwater exploitation (10<sup>4</sup> m<sup>3</sup>), GD is groundwater depth (m), P is probability.

**Table 6** | Conditional probability of water table depth falling into different ranges given a specific condition (range) of the surface water irrigation

<b>P(GD SI)</b>	<b>1,631 &lt; SI &lt; 14,103.67</b>	<b>14,103.67 &lt; SI &lt; 26,576.33</b>	<b>26,576.33 &lt; SI &lt; 39,049</b>
3.35 < GD < 4.07	0.12	0.18	0.27
4.07 < GD < 4.78	0.46	0.50	0.52
4.78 < GD < 5.49	0.42	0.32	0.21

Note: SI is surface water irrigation (10<sup>4</sup> m<sup>3</sup>), GD is groundwater depth (m), P is probability.

**Table 7** | Conditional probability of water table depth falling into different ranges given a specific condition (range) of the runoff

<b>P(GD R)</b>	<b>0 &lt; R &lt; 6.45</b>	<b>6.45 &lt; R &lt; 12.91</b>	<b>12.91 &lt; R &lt; 19.36</b>
3.35 < GD < 4.07	0.10	0.20	0.30
4.07 < GD < 4.78	0.45	0.51	0.52
4.78 < GD < 5.49	0.45	0.29	0.18

Note: R is runoff (10<sup>8</sup> m<sup>3</sup>), GD is groundwater depth (m), P is probability.

influence of groundwater exploitation is shown in Table 5. Regardless of the impacts of the other factors, when the volume of groundwater exploitation was in the maximum range (3,355.41 × 10<sup>4</sup>–5,033.12 × 10<sup>4</sup> m<sup>3</sup>), the probability that groundwater depth was in the maximum range (4.78–5.49 m) reached a maximum value (0.53) after a 1-month lag time. Moreover, the larger the groundwater exploitation, the higher the probability that groundwater depth would be in the maximum depth range. Similarly, the smaller the groundwater exploitation, the smaller the probability that groundwater depth was in the maximum range of depth

values (0.13). This analysis illustrates the strong positive dependency of groundwater depth on groundwater exploitation.

Table 6 shows the probability distribution of groundwater depth under a change in surface water irrigation. When the volume of surface water irrigation was within the maximum range ( $26,576.33 \times 10^4$ – $39,049 \times 10^4 \text{ m}^3$ ), the probability that groundwater depth was in the middle depth range (4.07–4.78 m) reached a maximum probability value of 0.52. Given the maximum volume of surface water irrigation, the probability of groundwater depth falling into the maximum (deepest) range (4.78–5.49 m) became the smallest (0.21). This indicates that surface water irrigation recharges groundwater in the area.

The conditional probability distribution of groundwater depth with variations in runoff is shown in Table 7. Regardless of the impacts of the other factors, when runoff was within the maximum range ( $12.91 \times 10^8$ – $19.36 \times 10^8 \text{ m}^3$ ), the probability that groundwater depth was in the middle range (4.07–4.78 m) reached a maximum probability value (0.52). In addition, the probability that groundwater depth was in the maximum (deepest) range (4.78–5.49 m) was the smallest (0.18), while the smaller the value of runoff, the higher the probability that groundwater depth (0.45) was in the maximum depth range (4.78–5.49 m). These trends demonstrate that an increase in runoff can lead to a decrease in the depth to groundwater.

## The Vine Copula Function

The above analysis shows that the depth to groundwater was influenced by several factors in the study area. 2D Copula Functions can only assess the influence of a single parameter on the probability distribution of groundwater depth but cannot determine the combined effect of multiple parameters. Therefore, a 3D Copula Function was further constructed to analyze the multivariate correlation between the three examined controlling factors and the change in groundwater depth.

A positive correlation was found to exist between groundwater exploitation and groundwater depth, whereas a negative correlation existed between runoff and surface water irrigation with groundwater depth. 3D symmetric or asymmetric Archimedes functions can only be applied when positive correlations exist between the variables (Chen 2013). Thus, 3D-Vine Copula Functions were used to analyze the influence of  $GE_{+1Mo}/SI$ ,  $GE_{lag1}/R$ , and  $SI/R$  on groundwater depth (Tables 8–10).

The probabilities that the groundwater depth was within a specific range given a defined range of groundwater exploitation and runoff are shown in Table 8. Given the mutual effect of runoff and groundwater exploitation on groundwater depth, the first line of Table 8 indicates that when runoff is in the range of  $0$ – $6.45 \times 10^8 \text{ m}^3$ , and the groundwater exploitation is in the range of  $0$ – $1,677.71 \times 10^4 \text{ m}^3$  during the previous month, the probability that groundwater

**Table 8** | Probabilities of the groundwater depth falling into different ranges given a specific range of groundwater exploitation and runoff

	$P(GD GE_{-1Mo}, R)$	$0 < GE_{-1Mo} < 1,677.71$	$1,677.71 < GE_{-1Mo} < 3,355.41$	$3,355.41 < GE_{-1Mo} < 5,033.12$
$0 < R < 6.45$	$3.35 < GD < 4.07$	0.17	0.07	0.04
	$4.07 < GD < 4.78$	0.53	0.42	0.29
	$4.78 < GD < 5.49$	0.30	0.51	0.67
$6.45 < R < 12.91$	$3.35 < GD < 4.07$	0.29	0.15	0.08
	$4.07 < GD < 4.78$	0.53	0.51	0.44
	$4.78 < GD < 5.49$	0.18	0.34	0.48
$12.91 < R < 19.36$	$3.35 < GD < 4.07$	0.44	0.24	0.06
	$4.07 < GD < 4.78$	0.44	0.60	0.50
	$4.78 < GD < 5.49$	0.12	0.16	0.44

Note: GE is groundwater exploitation ( $10^4 \text{ m}^3$ ), R is runoff ( $10^8 \text{ m}^3$ ), GD is groundwater depth (m), P is probability.

**Table 9** | Probabilities of the groundwater depth falling into different ranges given a specific range of surface water irrigation and runoff

P(GD SI, R)		1,631 < SI < 14,103.67	14,103.67 < SI < 26,576.33	26,576.33 < SI < 39,049
0 < R < 6.45	3.35 < GD < 4.07	0.06	0.10	0.11
	4.07 < GD < 4.78	0.41	0.45	0.48
	4.78 < GD < 5.49	0.53	0.45	0.41
6.45 < R < 12.91	3.35 < GD < 4.07	0.14	0.20	0.26
	4.07 < GD < 4.78	0.48	0.50	0.51
	4.78 < GD < 5.49	0.38	0.30	0.23
12.91 < R < 19.36	3.35 < GD < 4.07	0.22	0.29	0.35
	4.07 < GD < 4.78	0.49	0.51	0.54
	4.78 < GD < 5.49	0.29	0.20	0.11

Note: SI is surface water irrigation ( $10^4 \text{ m}^3$ ), R is runoff ( $\text{m}^3$ ), GD is groundwater depth (m), P is probability.

**Table 10** | Probabilities of the groundwater depth falling into different ranges given a specific range of surface water irrigation and groundwater exploitation

P(GD GE <sub>+1Mo</sub> , SI)		1,631 < SI < 14,103.67	14,103.67 < SI < 26,576.33	26,576.33 < SI < 39,049
0 < GE <sub>+1Mo</sub> < 1,677.71	3.35 < GD < 4.07	0.19	0.28	0.36
	4.07 < GD < 4.78	0.50	0.52	0.47
	4.78 < GD < 5.49	0.31	0.20	0.16
1,677.71 < GE <sub>+1Mo</sub> < 3,355.41	3.35 < GD < 4.07	0.09	0.14	0.26
	4.07 < GD < 4.78	0.47	0.51	0.48
	4.78 < GD < 5.49	0.44	0.35	0.26
3,355.41 < GE <sub>+1Mo</sub> < 5,033.12	3.35 < GD < 4.07	0.08	0.09	0.07
	4.07 < GD < 4.78	0.28	0.41	0.52
	4.78 < GD < 5.49	0.64	0.50	0.41

Note: GE<sub>+1Mo</sub> is groundwater exploitation ( $10^4 \text{ m}^3$ ), SI is surface water irrigation ( $10^4 \text{ m}^3$ ), GD is groundwater depth (m), P is probability.

depth is in the middle range (4.07–4.78 m) is 53%. With an increase in runoff, and the same level of groundwater exploitation ( $0\text{--}1,677.71 \times 10^4 \text{ m}^3$ ), the depth to groundwater decreases, and the probability that groundwater depth is in the maximum range (4.78–5.49 m) reaches its lowest value 12%. In contrast, when groundwater exploitation increases to the maximum range during the previous month ( $3,355.41 \times 10^4\text{--}5,033.12 \times 10^4 \text{ m}^3$ ), and runoff is within the range of  $0\text{--}6.45 \times 10^8 \text{ m}^3$ , the probability that groundwater depth is in the maximum range reached its biggest value, 67%. When groundwater exploitation is within its maximum range ( $3,355.41 \times 10^4\text{--}5,033.12 \times 10^4 \text{ m}^3$ ), along with runoff ( $12.9067 \times 10^8\text{--}19.36 \times 10^8 \text{ m}^3$ ), the probability that groundwater depth is in its maximum depth range decreased (fourth line of Table 8).

Tables 8 and 9 show the probability that the groundwater depth is within different ranges under the mutual effects of surface water irrigation and runoff. The correlations between runoff, surface water irrigation, and groundwater depth are negative. When both runoff and irrigation increase, groundwater depth is reduced. In contrast, the correlation between groundwater exploitation and groundwater depth is positive. When runoff and irrigation increase, along with exploitation, the depth to groundwater will also increase. The influence of groundwater exploitation on groundwater depth is greater than that of surface water irrigation and runoff. These relationships also agree with the conclusions derived from the 2D Copula analysis.

In terms of the actual situation of the YRBID, the surface water resources of the region mainly occur in the form of ice

and snowmelt water runoff. This runoff is accessed for each subirrigation area through a channel and is insufficient to provide enough surface water irrigation to meet the needs of crop growth. As a result, surface water infiltration and groundwater recharge are small, increasing the effect of groundwater exploitation on the depth to groundwater. It can be seen that the results of time series analysis by Copula Functions in this research have also shown the consistency with the natural situation in the study area.

Theoretically, 3D symmetric or asymmetric Archimedes functions can only be applied to positively correlated variables (Chen 2013). However, in this area, groundwater depth is negatively related to parameters that increase groundwater recharge (runoff and irrigation). A question that arises is how to set up a proper 3D Copula Function for groundwater depth? You *et al.* (2018) found that the performance of a 3D Symmetric Archimedean Copula Function was unsatisfactory. In this study, we used a 3D-Vine Copula model to analyze the response of groundwater depth to groundwater exploitation, runoff, and surface water irrigation. When two factors co-vary, a reasonable fitting result was obtained and was consistent with those of the constructed 2D Copula Function. The results suggest that in addition to the commonly used Archimedean Copula Function, Vine Copula could be used to satisfactorily describe time series correlations between groundwater depth and selected controlling factors. The Archimedean Copula Functions are not the only models that can describe the correlations among the variables, suggesting that other Copula Functions require further study.

The application of multidimensional Copula analysis allows the interrelations, restrictions, and detailed changes that occurred through time in the multivariate groundwater system to be more fully understood. Moreover, the probability that groundwater depth was in a defined range could be assessed. The analysis provided critical information on the probability of how the depth of the groundwater table would change in response to selected driving factors. The results provide a power management tool in that it provides insights into how groundwater exploitation and irrigation can be adjusted to control groundwater levels in the area. For example, in the future, we will examine the ecological impacts of changing groundwater levels in the area, and establish Copula Functions to estimate the appropriate

level of groundwater exploitation given a specific runoff volume and amount of irrigation. Similarly, if a certain amount of groundwater exploitation is planned, it can be determined if groundwater levels will remain appropriate given the amount of surface irrigation.

This study used 8 years of monthly monitoring data from 16 wells in the area, and these data can be used to assess monthly and seasonal changes in groundwater depth. In future research, the data can be further expanded to obtain a more accurate functional probability relation. Furthermore, although this investigation pertains to the arid oasis, the approach could be used in other semi-arid to arid regions consisting of similarly structured groundwater flow systems.

---

## CONCLUSIONS

This paper used Copula Functions to analyze the relationship between changes in groundwater depth and three selected controlling factors from the perspective of probability. The analyses were based on a time series of monthly monitoring data collected from 16 wells in an arid oasis. The generated modeling results evaluated using the goodness of fit methods were satisfactory. Furthermore, the paper compared the results of 2D Copula Functions with 3D Copula Functions, the latter of which jointly analyzed the responsiveness of groundwater depth to changes in multiple controlling factors. The analyses showed that Copula Functions are an effective method to analyze the probability of changes in groundwater levels in an arid oasis, and the response of groundwater depth to multiple driving factors. Specifically, this study found that in YRBID, Frank Copula Functions produced the best fit for generated joint distribution functions between groundwater depth and its driving factors. A 3D-Vine Copula Function was proposed for the first time in groundwater analysis. The results showed that groundwater exploration is the main factor leading to changes in groundwater depth, whereas the influence of runoff and surface water irrigation on groundwater depth was relatively small.

Observations from the Irrigation District suggest that groundwater depth has become deeper in recent years than is typical. The model generated here can be used as a management tool to assess changes in groundwater depth as a function of groundwater exploitation, irrigation, and

runoff. For example, the data could be used to adjust the volume of surface water irrigation during a given amount of groundwater exploitation to maintain groundwater levels within a suitable (defined) range.

This study provides an effective research method for the probability analysis of groundwater depth in irrigated areas of an arid region, serving as an effective theoretical and technical approach to the sustainable utilization and management of groundwater resources in a basin.

## ACKNOWLEDGEMENTS

This research was supported by The Thousand Youth Talents' Plan (Xinjiang Project Y771071001). We would like to thank LetPub ([www.letpub.com](http://www.letpub.com)) for providing linguistic assistance during the preparation of this manuscript.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

- Aas, K., Czado, C., Frigessi, A. & Bakken, H. 2009 Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* **44** (2), 182–198.
- Brechmann, E. C. & Schepsmeier, U. 2013 Modeling dependence with C- and D-vine Copulas: the R package CDVine. *Journal of Statistical Software* **3** (52), 1–27.
- Chen, L. 2013 *Analysis and Calculation Application Research in Multivariable Hydrology Based of Copula Function Theory*. Wuhan University Press, Wuhan, China (in Chinese).
- Chen, Y. N., Li, W. H., Xu, C. C. & Hao, X. M. 2007 Effects of climate change on water resources in Tarim River Basin, northwest China. *Journal of Environmental Sciences* **19** (4), 488–493.
- Chen, Y. N., Xu, C. C., Chen, Y. P., Li, W. H. & Liu, J. S. 2010 Response of glacial-lake outburst floods to climate change in the Yarkant River Basin on northern slope of Karakoram Mountains, China. *Quaternary International* **226** (1–2), 0–81.
- Chen, Z., Yang, H. & Dong, C. 2016 Analysis of groundwater table depth changes in Yarkant plain oasis in recent 20 years and their causes. *Journal of Hydroelectric Engineering* **35** (6), 58–66 (in Chinese).
- Choubin, B. & Malekian, A. 2017 Combined gamma and M-test-based ANN and ARIMA models for groundwater fluctuation forecasting in semiarid regions. *Environmental Earth Sciences* **76** (15), 538.
- Daneshkhah, A., Remesan, R., Chatrabgoun, O. & Holman, I. P. 2016 Probabilistic modeling of flood characterizations with parametric and minimum information pair-copula model. *Journal of Hydrology* **540**, 469–487.
- De Michele, C., Salvadori, G., Vezzoli, R. & Pecora, S. 2013 Multivariate assessment of droughts: frequency analysis and dynamic return period. *Water Resources Research* **49** (10), 6985–6994.
- Deng, H. J., Chen, Y. N., Wang, H. J. & Zhang, S. H. 2015 Climate change with elevation and its potential impact on water resources in the Tianshan Mountains, Central Asia. *Global and Planetary Change* **135**. pii:S0921818115300527.
- Du, Q., Xu, H. & Zhang, G. Y. 2016 Features of eco-environmental changes in Yarkant River Basin from 1990 to 2010. *Agricultural Research in the Arid Areas* **34** (01), 252–256 (in Chinese).
- Duethmann, D., Menz, C., Jiang, T. & Vorogushyn, S. 2016 Projections for headwater catchments of the Tarim river reveal glacier retreat and decreasing surface water availability but uncertainties are large. *Environmental Research Letters* **11** (5), 054024.
- Feng, J. Y. 2016 *Analysis on the Influencing Factors of Soil and Water Environment Evolution and Its Prevention and Control Countermeasures in the Yarkant River Basin*. MD Thesis, Northwest University, Lanzhou, China (in Chinese).
- Genest, C. & Favre, A. C. 2007 Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* **12** (4), 347–368.
- Gharde, K. D., Kothari, M. & Mahale, D. M. 2016 Developed seasonal ARIMA model to forecast streamflow for Savitri Basin in Konkan Region of Maharashtra on daily basis. *Journal of the Indian Society of Coastal Agricultural Research* **34**, 110–119.
- Gulupiya, S. 2017a Determination and analysis for groundwater ecological water level of Yarkant River Watershed. *Heilongjiang Hydraulic Science and Technology* **45** (9), 17–20 (in Chinese).
- Gulupiya, S. 2017b Groundwater dynamic prediction based on Monte Carlo algorithm in the upper reaches of the Yarkant River, Xinjiang. *Groundwater* **39** (03), 210–211 (in Chinese).
- Haberlandt, U. & Radtke, I. 2014 Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows. *Hydrology and Earth System Sciences* **18** (1), 353–365.
- Khedun, C. P., Mishra, A. K., Singh, V. P. & Giardino, J. R. 2014 A copula-based precipitation forecasting model: investigating the interdecadal modulation of ENSO's impacts on monthly precipitation. *Water Resources Research* **50** (1), 580–600.
- Liang, J. 2017 Study on the influence of climate change on regional groundwater dynamics in the Yarkant River Basin of Xinjiang based on SWAT model. *Underground Water* **39** (04), 232–233 (in Chinese).
- Liu, X., Min, J. & Liu, T. 2009 Wavelet analysis of temperature and precipitation variation in the Yarkant River Basin, Xinjiang. *Journal of Desert Research* **29** (03), 566–570 (in Chinese).

- Liu, Z., Cheng, L., Hao, Z., Li, J., Thorstensen, A. & Gao, H. 2018 A framework for exploring joint effects of conditional factors on compound floods. *Water Resources Research* **54** (4), 2681–2696.
- Madadgar, S., AghaKouchak, A., Farahmand, A. & Davis, S. J. 2017 Probabilistic estimates of drought impacts on agricultural production. *Geophysical Research Letters* **44** (15), 7799–7807.
- Mansuer, S., Ablimit, A. & Riyangul, O. 2011 Cultivated land change and its ground water level response in Kashghar in last decade. *Journal of Arid Land Resources and Environment* **25** (11), 145–151.
- Meintanis, S. G., Ngatchou-Wandji, J. & Taufer, E. 2015 Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. *Journal of Multivariate Analysis* **140**, 171–192.
- Michele, C. D. 2003 A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research* **108** (D2), 4067.
- Mogaji, K. A., Lim, H. S. & Abdullah, K. 2015 Modeling of groundwater recharge using a multiple linear regression (MLR) recharge model developed from geophysical parameters: a case of groundwater resources management. *Environmental Earth Sciences* **73** (3), 1217–1230.
- Momani, P. E. N. M. & Naill, P. E. 2009 Time series analysis model for rainfall data in Jordan: case study for using time series analysis. *American Journal of Environmental Sciences* **5** (5), 599.
- Panagiotelis, A., Czado, C. & Joe, H. 2012 Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association* **107** (499), 1063–1072.
- Renard, B. & Lang, M. 2007 Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources* **30** (4), 897–912.
- Sadegh, M., Ragno, E. & AghaKouchak, A. 2017 Multivariate Copula Analysis Toolbox (MvCAT): describing dependence and underlying uncertainty using a Bayesian framework. *Water Resources Research* **53** (6), 5166–5183.
- Salvadori, G. & De Michele, C. 2007 On the use of Copulas in hydrology: theory and practice. *Journal of Hydrologic Engineering* **12** (4), 369–380.
- Salvadori, G. & De Michele, C. 2015 Multivariate real-time assessment of droughts via Copula-based multi-site hazard trajectories and fans. *Journal of Hydrology* **526**, 101–115.
- Salvadori, G., Durante, F., De Michele, C., Bernardi, M. & Petrella, L. 2016 A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. *Water Resources Research* **52** (5), 3701–3721.
- Schumann, A. 2017 Flood safety versus remaining risks – options and limitations of probabilistic concepts in flood management. *Water Resources Management* **31** (10), 3131–3145.
- Shataer, G. 2017 Spatial and temporal dynamics of groundwater and its influencing factors in the Yarkant River Basin. *Ground Water* **39** (6), 64–65.
- Sklar, M. 1959 Fonctions de repartition an dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris* **8**, 229–231.
- Stamatatou, N., Vasiliades, L. & Loukas, A. 2018 Bivariate flood frequency analysis using Copulas. *Multidisciplinary Digital Publishing Institute Proceedings* **11** (2), 635.
- Sun, C., Zhang, X., Jin, N., Du, H. & Ma, W. 2010 Spatial difference features and organization optimization of cities and towns in Tarim River Basin. *Journal of Arid Land* **2** (1), 33–42.
- Szolgay, J., Gaál, L., Bacigál, T., Kohnová, S. & Blöschl, G. 2016 Reducing uncertainty in the selection of bi-variate distributions of flood peaks and volumes using copulas and hydrological process-based model selection. In: *Egu General Assembly Conference*. EGU General Assembly Conference Abstracts.
- Vezzoli, R., Salvadori, G. & De Michele, C. 2017 A distributional multivariate approach for assessing performance of climate-hydrology models. *Scientific Reports* **7** (1), 12071.
- Wang, Y., Xiao, D., Li, Y. & Li, X. 2008 Soil salinity evolution and its relationship with dynamics of groundwater in the oasis of inland river basins: case study from the Fubei region of Xinjiang Province, China. *Environmental Monitoring and Assessment* **140** (1–3), 291–302.
- Xiao, J., Jin, Z. D., Wang, J. & Zhang, F. 2015 Hydrochemical characteristics, controlling factors and solute sources of groundwater within the Tarim River Basin in the extreme arid region, NW Tibetan Plateau. *Quaternary International* **380**, 237–246.
- Xiong, L., Jiang, C., Xu, C. Y., Yu, K. X. & Guo, S. 2015 A framework of change-point detection for multivariate hydrological series. *Water Resources Research* **51** (10), 8198–8217.
- You, Q., Liu, Y. & Liu, Z. 2018 Probability analysis of the water table and driving factors using a multidimensional copula function. *Water* **10** (4), 472.
- Zhang, Q., Xu, M. & Sun, J. 2010 *Conference on Environmental Pollution and Public Health*. Scientific Research Publishing, USA.
- Zhang, S., Gao, X., Zhang, X. & Hagemann, S. 2012 Projection of glacier runoff in Yarkant River basin and Beida River basin, Western China. *Hydrological Processes* **26** (18), 2773–2781.
- Zheng, Y. L., Zhang, L. P., Zhang, X. L., Wang, K. & Zheng, Y. J. 2015 Forecast model analysis for the morbidity of tuberculosis in Xinjiang, China. *PLoS ONE* **10** (3), e0116832.
- Zhou, H. & Li, W. 2013 The effects of oasis ecosystem hydrological processes on soil salinization in the lower reaches of the Tarim River, China. *Ecohydrology* **6** (6), 1009–1020.

First received 20 September 2019; accepted in revised form 28 December 2019. Available online 11 February 2020