RESEARCH ARTICLE

Modelling of bivariate meteorological drought analysis in Lake Urmia Basin using Archimedean copula functions

Farzad Khezri | Mohsen Irandoost | Navid Jalalkamali | Najme Yazdanpanah

Department of Water Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran

Correspondence
Mohsen Irandoost, Department of Water Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran.
Email: irandostmo@gmail.com

Abstract
The Urmia Lake Basin as the world's second largest salt lake has experienced severe drought during recent years. The purpose of this study is to analyse the bivariate characteristics of drought (i.e., duration and severity) using two indices including SPI (standard precipitation index) and SPI\textsuperscript{mod} (modified SPI) associated with copula functions. For this purpose, rainfall data of six stations were used for the period of 1971–2017. At first, the characteristics of drought were extracted using the two indices. Then, through coding in the MATLAB software environment, eight families of Archimedean copula functions were applied. The simultaneous return period and conditional and Kendall returns were also investigated. The result showed that the Joe copula function was the best predictor for the analysis of both intensity and duration of drought for the study area. The correlation coefficients of Spearman, linear correlation and Tau Kendall computed for the SPI of stations were >0.65, >0.72 and >0.48, respectively, while all of them were significant. At a given critical probability level, \( t \), the value of the Kendall return period was much greater than the standard return period, so that this difference increased with increasing \( t \) value. The results obtained from the time series of indices indicated that at least 40% of the months were dry, and the severity of droughts in the Urmia station was much higher than other stations during the studied period. Moreover, SPI\textsuperscript{mod} to a large extent eliminates the disadvantages of conventional SPI and takes into account seasonal variations of precipitation in the calculation of SPI.

KEYWORDS
Copula functions, drought, meteorological data, modelling, Urmia Lake

1 | INTRODUCTION

Drought is known as a natural phenomenon that depending on its characteristics and consequences can be studied in four categories including meteorological, agricultural, hydrological and socio-economic schemes (Dracup et al., 1980a, 1980b; Heim, 2002; Xu et al., 2015). Planning and improving the management of water resources systems under drought conditions requires estimating the period of combined and conditional drought returns. Typically, simultaneous distributions are obtained analytically by assuming that...
drought variables are independent random variables or follow the same univariate distribution and have explicit multivariate shapes (such as normal multivariate, exponential multivariate and gamma multivariate). However, these two conditions are not met in most cases because drought variables are highly correlated and may follow different univariate distributions (Wong et al., 2013).

While analysing two or more variables, copula functions incorporate the correlations between them in the calculations. These functions were first proposed by Sklar (1959) to create multivariate distributions. In hydrological studies, copula functions were first used by De Michele and Salvadori (2003) to create a two-variable distribution of storm intensity and duration. Mirabbasi et al. (2012) analysed drought bivariate using copula functions in order to study the severity and continuity of meteorological drought events at Sharaakhaneh station, Iran. They found that the Galambos copula function was the most suitable copula for the station. Amirataee et al. (2018) developed the distribution of the joint probabilities of drought intensity and the percentage of area affected by drought based on the copula functions in the catchment area of Urmia Lake in Iran and used seven different families of copula functions. Ayantobo et al. (2018) also used several copula functions to derive regional and spatial models of drought risk assessment in Mainland China for the period of 1961–2013. Pontes Filho et al. (2019) studied SPI (standard precipitation index) based on copula functions and monitored the probability of consecutive droughts in Portugal.

In the continuation of the research, various studies on droughts and indicators that have been done in the world are presented (e.g., Borji et al., 2016; Bouabdelli et al., 2020; Choubin et al., 2016; Danandeh Mehr et al., 2020; Salvadori & De Michele, 2015; Wang et al., 2020). Achour et al. (2020) performed the time-spatial analysis and drought forecasting in the plains of northwestern Algeria using SPI. They used the artificial neural network (ANN) model from different scales of the SPI. The results of their research indicated that the frequency of drought varied depending on time scale. Das et al. (2020) applied the copula function approach using the SPI to assess drought characteristics and climatic indicators in the Himalayan states in India.

Mohseni Saravi et al. (2009) studied intensity–duration–frequency (IDF) and spatial analysis of droughts using the standardized precipitation index for Karun River basin located in Iran. For this purpose, they used standard 3-, 6- and 12-month scales of precipitation index. Choubin et al. (2014) predicted drought in the Maharloo Bakhtegan semi-arid watershed in southeastern Iran using climatic signals and SPI in a fuzzy neural modelling framework. The results of their study showed that Atlantic Ocean surface temperature was inversely related to SPI. Kazemzadeh and Malekian (2016) studied the spatial and temporal characteristics of meteorological and hydrological drought in northwestern Iran and showed that most of the drought events in the base periods in the last 15 years were from 1995–1996 to 2010–2011. In addition, the driest year was 2007–2008, according to the meteorological drought index. Malekian and Kazemzadeh (2016) studied the spatio-temporal analysis of regional trends and the changes of self-correlated temperature series changes in the Urmia Lake Basin located in northwestern Iran. The results of Ljung–Box test showed a positive series correlation between mean temperature (Tmean) and maximum temperature (Tmax) for all stations at a significant level of 0.05. In the monthly series, the significant warm-up trend in the Tmean series was more noticeable than in the Tmax series. However, the Tmax trend was found to be greater than the Tmin series.

The Urmia Lake Basin as the largest interior lake in Iran has experienced severe drought during recent years. The water level in the Lake has decreased dramatically in recent years due to drought, climate change and the overuse of water resources. Up to date, no study has been done in this basin using the copula functions joint to SPI and SPI\textsuperscript{mod} for the analysis of drought. Therefore, the purpose of this study is to evaluate the bivariate analysis of drought in the basin of Urmia Lake located in northwestern Iran using two indices of SPI and SPI\textsuperscript{mod} (modified SPI) by eight categories of copula functions: Gumble Huagard, Galambos, Frank, Clyton, Placket, Farlie Gumble Morgan Stern, Ali Mikhail Haq and Joe. For this purpose, rainfall data of six synoptic stations were used for the statistical period of 1971–2017. The coding at various stages of this research was done in the MATLAB programming environment. The findings of this study will enhance our knowledge on drought vulnerability and drought impact assessment in the Urmia Lake Basin for improving the management of water resources under vulnerable conditions. Moreover, Joint and conditional return periods of drought characteristics (severity and duration) will provide useful information for drought risk assessment and analysis to managers, water resources engineers, and can design criteria for water reservoirs.
2 | MATERIALS AND METHODS

2.1 | Study area

This study was done in the Urmia Lake Basin, which is located between 37°4' to 38°17' latitudes and 45°13' to 46° longitudes in northwest Iran and covers an area of 51,800 km² (Figure 1). The lake is between two provinces of West Azarbaijan and East Azerbaijan, with an average height of 1276 m above sea level. The study of fluctuations in the water level of lakes in terms of importance, nature and location of these aquifers has become particularly important in recent years. The Urmia Lake is the largest inland lake in Iran. There are 16 wetlands around the Lake with areas ranging from 5 to 120 ha (some of which have dried up), most of them have fresh or salty water and have a high ecosystem value. Six synoptic stations were selected to acquire rainfall database for the region. Table 1 shows the general characteristics of the stations studied.

2.2 | SPI and SPI\textsuperscript{mod} indices

SPI is an index that depends on the probability of precipitation for each time and scale. This index is comparable for different time scales, and based on it, the initial warning for monitoring drought and its severity can be evaluated and calculated using the following equation (McKee et al., 1993):

\[
\text{SPI} = \frac{P - P_i}{\sigma},
\]

(1)

where SPI is the standardized precipitation index, \(P_i\) is the normalized data, \(P\) is the mean of normalized data in the desired period and \(\sigma\) is the standard deviation of normalized data in the desired period.

If \(X_w\) represents the total precipitation over a period of \(W\) months, by fitting the two-parameter gamma distribution to the time series \(X_w\), the cumulative distribution function or marginal CDF \(F_{Xw}(x_w)\) is obtained then the SPI, like any observation, is calculated with a normal inverse function or \(\phi^{-1}(u_w)\). Although the SPI calculation method seems logical, it still has its drawbacks. For example, the SPI in the annual rainfall regime cannot calculate seasonal variations. Therefore, Kao and Govindaraju (2010) suggested that the accumulated precipitation for a given time window be grouped \(X_w(t)\) based on the final month (Kao & Govindaraju, 2010). On this basis, in the modified SPI (SPI\textsuperscript{mod}), \(X_w(t)\) changes to \(X^m_w(y)\), where \(m\) represents the final month of the time window \(w\) (\(m = 1\) (Jan) ... \(m = 12\) (Dec) and \(y\) is the index of the year, where \(y = 1, 2, ..., k\) (\(k\) is the total number of years) as presented in Equation (2):

\[
X^\text{month}_w(g) = X_w(12(g - 1) + 1) = X_w(t),
\]

(2)

where, \(g\) is year index and \(m\) is month index (month number) and \(t\) is time index and is stated as \(t = 12(g - 1) + m\). By fitting the inverse distribution for each group means, \(u^m_w = F_{X^m_w}(X^m_w), w = 1, 2, ..., 12, m = 1(\text{Jan}) ... 12(\text{Jan})\) SPI\textsuperscript{mod} can be calculated through conversion of \(u^m_w\) to standard normal variable as follows:
where \( \phi^{-1}(u^m) \) is the inverse function of the standard normal distribution with mean zero and standard deviation of 1.

### 2.3 Theory of copula functions

Copula functions have been of interest to researchers in modelling (non-linear) multivariate relationships in hydrological and meteorological studies such as multivariate frequency analysis, risk assessment, drought modelling, floods and geostatistical interpolations (Schweizer & Wolff, 1981). Copula was used in hydrology because of its flexibility, since it is based on a nonlinear relationship between variables and connects both distribution and marginal functions (Joe, 1997; Schweizer & Wolff, 1981). In this study, copula function was used for the multivariate analysis of drought in the Urmia Lake Basin. The following is a brief description of the distribution functions and the copula functions used in the research.

### 2.4 Sklar theory

The first application of copula function has been attributed to Sklar (1959). In the Sklar theory, 1D distribution functions can be combined in the form of multivariate distributions (Nelsen, 2006). For continuous random variables, d dimension \( \{X_1,...,X_d\} \) with marginal CDF is \( u_j = F_{X_j}(x_j) \) that \( j = 1,...,d \) as follows in Equation 4:

\[
C_{U_1,...,U_d}(u_1,...,u_d) = H_{X_1,...,X_d}(x_1,...,x_d),
\]

where \( u_j \) is the cumulative distribution function for the variable \( j_m \) and \( H_{X_1,...,X_d}(x_1,...,x_d) \) is the same CDF with d variables for a set of variables \( \{X_1,...,X_d\} \). When copula is used to create a bivariate distribution, it is very easy to evaluate the combined probabilities. The combined distribution function defined for the two variables in the following equation indicates the probability that both \( X \) and \( Y \) are less than or equal to \( x \) and \( y \). In this case, the simultaneous probability of them is as Equation (5) (Nelsen, 2006):

\[
P(X \leq x, Y \leq y) = F_{X,Y}(x,y) = C(F_X(x), F_Y(y)).
\]

Because different combinations of \( F_X(x), F_Y(y) \) can lead to the same value of \( F_{X,Y}(x,y) \), a simple and useful way to present the simultaneous probability \( F_{X,Y}(x,y) \) is drawing equal curves of equal probabilities. In practice, the probability of both \( x \) and \( y \) exceeding their respective thresholds can also be defined in copula as follows (Shiau, 2006):

\[
P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F_{X,Y}(x,y) = 1 - F_X(x) - F_Y(y) + C(F_X(x), F_Y(y)),
\]

Now the return period is for intensity (\( I \)) and duration (\( D \)) and the return period of Kendall is in the form of Equations (7) and (8):

\[
T_{DS} = \frac{E(L)}{P(D \geq d, S \geq s)} = \frac{E(L)}{1 - F_D(d) - F_S(s) + C(F_D(d), F_S(s))},
\]

\[
T_p = \frac{E(L)}{1 - K_C(t)} = \frac{E(L)}{1 - \{u \in I^d : C(u) \leq t\}},
\]

where \( E(L) \) is the average time between the onsets of two consecutive droughts and \( t \in I \) is a critical threshold that is defined based on the following equation (Salvadori & De Michele, 2010):

\[
t = \inf\{s \in I : K_C(s) = p\} = K_C^{-1}(p),
\]

where \( \inf\{\} \) is infimum or is the largest lower bound of the set. This is clear because \( K_C \) is a distribution function, since the condition for using copula functions to analyse multivariate is the correlation between variables.

### TABLE 1 Geographical characteristics of the selected meteorological stations

| Station | Longitude | Latitude | Height | Established year |
|---------|-----------|----------|--------|-----------------|
| Urmia   | 45° 5’    | 37° 56’  | 1335   | 1951            |
| Tabriz  | 46° 14’   | 38° 07’  | 1364   | 1951            |
| Takab   | 47° 09’   | 36° 39’  | 1871   | 1885            |
| Sarab   | 47° 23’   | 37° 56’  | 1362   | 1986            |
| Maraghe | 46° 12’   | 37° 21’  | 1344   | 1983            |
| Mahabad | 45° 71’   | 37° 56’  | 1352   | 1983            |
(Nelsen, 2006). Therefore, in this study, the correlation equations of Spearman (Equation 10), Pearson (Equation 11) and Tau Kendall (Equation 12) were used as suggested by Nelsen (2006) as follows respectively:

\[
\rho = 1 - \frac{6 \sum_{i=1}^{n} (d_i)^2}{n(n^2 - 1)}, \tag{10}
\]

\[
r = \frac{E[(X - \bar{X})(Y - \bar{Y})]}{\text{Std}[X] \text{Std}[Y]}, \tag{11}
\]

where \(d_i\) is the difference between the data ranks, \(C\) is the number of coordinated pairs and \(D\) is the number of uncoordinated pairs in which \(X\) and \(Y\), respectively, are the mean of the variables of \(X\) and \(Y\), \(E[.\] is mathematical expectation and \(\text{Std}[.]\) is the standard deviation of the events.

\[
r = \frac{C - D}{C + D} = \frac{C - D}{\frac{n(n-1)}{2} \binom{n}{2} / 2^{(n-2)}} = \frac{C - D}{\frac{n^2}{2(n-2)}}, \tag{12}
\]

\[
L(x, y) = \prod_{i=1}^{n} [c(F(x_i, y_i; \theta)f(x_i)g(y_i)], \tag{16}
\]

where \(n\) is the sample size, \(C_p\) is the calculated values of the parametric copula, \(C_e\) is the observational values of the probability obtained from the experimental copula, \(C_e\) is observational values of probability obtained from the experimental copula, \(\ln ML\) the maximum value of the logarithm function and \(k\) is the number of fit parameters.

\[
\text{AIC} = -2\ln ML + 2k, \tag{13}
\]

\[
\text{NSE} = 1 - \frac{\sum_{i=1}^{n} (C_{pi} - C_{ei})^2}{\sum_{i=1}^{n} (C_{ei} - C_e)^2}, \tag{14}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} [C_{pi} - C_{ei}]^2}, \tag{15}
\]

\[
\text{where } n \text{ is the sample size, } C_p \text{ is the calculated values of the parametric copula, } C_e \text{ is the observational values of the probability obtained from the experimental copula, } \ln ML \text{ the maximum value of the logarithm function and } k \text{ is the number of fit parameters.}
\]

\[
\text{RESULTS AND DISCUSSION}
\]

\[
3.1 \quad \text{SPI and SPI}_{\text{mod}} \text{ time series}
\]

Hydrological data are usually obtained from a random process. The randomness of the data allows us to analyse it in probabilistic ways. To use possible methods for data analysis, we must make sure that the data are random. For this reason, using the Run test method, the randomness test of precipitation data for the studied stations was investigated. In order to investigate the presence of trend in statistical characteristics (such as mean, variance) of rainfall, static test (trend) of Spearman was used. The Spearman method is the most suitable method for static study of hydrological data. In the case of hydrology, the randomness of time series generally means that these data are the result of a series of natural phenomena. Run test is used to check the randomness of rainfall data. The results are presented in Table 3 for all stations indicating that the statistical data of all selected stations are random. According to Table 3, the results obtained for the Run test show that precipitation data in all stations are within the allowable range in terms of randomness testing. In Spearman test, the correlation between the rank of the events is considered as the basis.

The time series of SPI\text{mod} and SPI for the Urmia station at monthly scale are shown in Figure 2. The comparison of time series between SPI and SPI\text{mod} shows that for Urmia station and the study area, 40% of the number of months was dry. The comparison indicates that the range of changes in SPI is greater than in SPI\text{mod}, so that the most severe drought according to SPI is \(-6.07\) while for SPI\text{mod} this value is \(-2.76\). This can be attributed to the way in which the gamma distribution function is fitted.
The relationships of Archimedean copula functions (Nelsen, 2006) used in this study were adopted, whereas according to SPImod, this fitting is done on the data for the whole statistical period. The distribution coefficient parameters (alpha and beta) is turned makes the amplitude of changes in the SPI mod obtained for precipitation data. However, in the SPI, one pair of gamma distributions is separately for each month during the statistical period. Thus, for each month, one pair of gamma distributions is obtained for precipitation data in the SPI and SPImod indices. The parameters of the selected variables; therefore, the results obtained from the goodness of fit criteria showed that the superior copula function over the fitted rainfall data is the prediction of the Joe copula function, since for most stations the optimal parameter value of this copula function was 4.7. Also, the results of statistically superior distributions compared with intensity and duration data revealed that beta and exponential distribution functions were the best, respectively. The results of correlation coefficients showed that the highest correlation coefficients were related to Pearson correlation coefficient followed by Spearman and Kendall correlation coefficients. Moreover, as the condition for the application of copula functions is the existence of correlation between the selected variables; therefore, the results obtained through the correlation coefficients confirm this finding.

### Table 2: Relationships of Archimedean copula functions (Nelsen, 2006) used in this study

| Parameter domain | $C(u, v)$ | Copula function |
|------------------|-----------|-----------------|
| Gamble hogaurd   | $C(u, v) = \exp\left\{-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^2\right\}$ | $\theta \geq 1$ |
| Galambus         | $C(u, v) = u v \exp\left\{\frac{1}{\theta \ln u v} \left[\left((-\ln u)^\theta + (-\ln v)^\theta\right)^2\right]\right\}$ | $\theta \geq 0$ |
| Frank            | $C(u, v) = -\frac{1}{\theta} \ln \left[1 + \frac{e^{\theta u} - 1}{e^{\theta v} - 1}\right]$ | $\theta \neq 0$ |
| Cliton           | $C(u, v) = u^{\theta} + v^{\theta} - 1^{-1/\theta}$ | $\theta \geq \star$ |
| Plaket           | $C(u, v) = \frac{1}{\theta} \ln \left[\frac{1 + (\theta - 1)(u + v) - \theta \ln(u + v)}{(1 + \theta - 1)(u + v)^2 - 4\theta (\theta - 1)uv}^{1/\theta}\right]$ | $\theta \geq 0$ |
| Farlie Gamble Morgenstern | $C(u, v) = uv[1 + \theta(1 - u)(1 - v)]$ | $-1 \leq \theta \leq 1$ |
| Joe              | $C(u, v) = 1 - \left[(1 - u)^\theta + (1 - v)^\theta - ((1 - u)(1 - v))^\theta\right]$ | $\theta \geq 0$ |
| Ali Michaeilhagh | $C(u, v) = 1 \frac{uv}{\theta(1 - u)(1 - v)}$ | $-1 \leq \theta \leq 1$ |

### Table 3: Randomness and static tests for the rainfall data for the period of 1971–2017

| Station | Randomness test | Static test |
|---------|----------------|-------------|
|         | Run test | Domain | Trend | Domain |
| Tabriz  | 12       | (5.15)   | 1.33  | ±2.1   |
| Sarab   | 8        | (6.16)   | 0.98  | ±2.1   |
| Maraghe | 8        | (6.16)   | 2.14  | ±2.1   |
| Urmia   | 8        | (5.15)   | 0.89  | ±2.1   |
| Takab   | 9        | (5.15)   | 1.39  | ±2.1   |
| Mahabad | 8        | (6.16)   | 1.81  | ±2.1   |

With the precipitation data in the SPI and SPImod indices, Because in the SPI, only one pair of parameter gamma distribution coefficient parameters (alpha and beta) is fitted on the data for the whole statistical period is adopted, whereas according to SPImod, this fitting is done separately for each month during the statistical period. Thus, for each month, one pair of gamma distributions is obtained for precipitation data. However, in the SPI, one pair is obtained for the whole statistical period, which in turn makes the amplitude of changes in the SPImod greater than the SPI. The results of time series analysis for the studied stations on average showed that 45% of the study area had dry months (Table 4).

### 3.2 Evaluation result of the SPI and SPImod indices

In this study, the characteristics of drought are evaluated using the SPI and SPImod indices. The parameters of the fit distribution functions on the drought characteristics were extracted, and the correlation between the characteristics was examined with three criteria of Pearson, Kendall $\tau$ and Spearman $\rho$ correlation coefficient. The goodness of fit criteria were then used to determine the copula functions. The results of this section are given in Table 5. The results of Nash–Sutcliffe coefficient for multivariate analysis for all stations showed at least 0.998 values for NSE, indicating the good performance of copula functions in the analysis. In other words, events similar to historical data or more or less of them are likely to occur in future periods. Similar result was found for the values of the maximum likelihood function, RMSE, AIC for different copula-based. Accordingly, any copula function that has the lowest mean squared error and criterion has the highest maximum likelihood (ML) and is known as the superior copula function. On the other hand, any function whose Nash–Sutcliffe efficiency (NSE) coefficient is close to one is known as superior. The results obtained from the goodness of fit criteria showed that the superior copula function over the fitted rainfall data is the prediction of the Joe copula function, since for most stations the optimal parameter value of this copula function was 4.7. Also, the results of statistically superior distributions compared with intensity and duration data revealed that beta and exponential distribution functions were the best, respectively. The results of correlation coefficients showed that the highest correlation coefficients were related to Pearson correlation coefficient followed by Spearman and Kendall correlation coefficients. Moreover, as the condition for the application of copula functions is the existence of correlation between the selected variables; therefore, the results obtained through the correlation coefficients confirm this finding.
The copula theoretical and experimental values of the coincident on drought intensity and duration data for Urmia station are shown in Figure 3a. As is seen, the points are very close to the 45° angle bisector line (Line 1:1), indicating that the fitted Joe copula values have an acceptable agreement with the values obtained from the empirical copula. Also, the changes of the ML logarithm function against their dependence parameter are shown in Figure 3b. The value of the dependency parameter of the copula function for Urmia station is equal to 4.72 and its ML is 82.98.

Then, the probabilities and their return period and other characteristics related to droughts were calculated by creating bivariate distribution functions. The results of this section indicated that depending on the copula function and due to the limited nature of different events, it is possible that the appropriate copula function differs among the stations. Also, the values of the studied indices are different from each other, probably due to the coincidence of statistical distributions on rainfall data to extract drought characteristics.

### 3.3 Evaluation of the joint probabilities

In this study, by replacing the values of the density distribution of the severity margins and duration of drought in the selected copula function, the simultaneous probabilities were calculated. Since different combinations of duration and severity of drought may result in the same
probability, the simultaneous probabilities are represented by counter lines. Figure 4a shows the counter lines of the probabilities of both the severity and duration of drought at the Urmia station using a Joe superior function and observational values. From a practical point of view, it is possible that both the severity and duration of the drought are simultaneously higher than certain threshold levels, providing useful information for managers and planners to improve water resource management under drought conditions. For example, in Figure 4a, the probability of occurrence of drought with a duration of more than 6 months, that is, $D_d < 6$ months, and intensities of $D_d < 6$ at Urmia station is about 0.5. Also, a 3D representation of the probability of both intensity and duration of drought based on their superior copula function is shown in Figure 4b for Urmia station. Therefore, awareness of the occurrence of drought events can be useful information for planners in various management areas such as water resources, agriculture, environment, hydroelectric, economic and social, as well as preparedness for the destructive risks of drought.

It was also found that lower levels of probability are associated with smaller values of the characteristics, while with increasing probability, the values of the

| Index | Statistical criterion | Urmia | Tabriz | Takab | Sarab | Maraghe | Mahabad |
|-------|-----------------------|-------|--------|-------|-------|---------|---------|
| SPI   | Pearson correlation coef. | 0.935 | 0.916 | 0.94  | 0.722 | 0.946   | 0.940   |
|       | Spearman correlation coef. | 0.859 | 0.797 | 0.659 | 0.671 | 0.829   | 0.829   |
|       | Tau Kendall            | 0.704 | 0.628 | 0.829 | 0.487 | 0.667   | 0.659   |
|       | Maximum likelihood     | 82.98 | 78.75 | 62.78 | 95.01 | 75.92   | 62.78   |
|       | AIC                    | 0.16  | 0.21  | 0.4   | 0.04  | 0.24    | 0.4     |
|       | NSE                    | 0.994 | 0.999 | 0.99  | 0.992 | 0.998   | 0.997   |
|       | RMSE                   | 0.02  | 0.02  | 0.02  | 0.02  | 0.02    | 0.02    |
|       | Best copula function   | 4.72  | 5.19  | 8.2   | 1     | 6.29    | 8.2     |
| SPI$_{mod}$ | Pearson correlation coef. | 0.94  | 0.62  | 0.571 | 0.699 | 0.809   | 0.831   |
|         | Spearman correlation coef. | 0.829 | 0.542 | 0.518 | 0.55  | 0.557   | 0.573   |
|         | Tau Kendall            | 0.659 | 0.422 | 0.376 | 0.407 | 0.482   | 0.506   |
|         | Maximum likelihood     | 62.78 | 72.72 | 85.29 | 84.32 | 92.31   | 80.36   |
|         | AIC                    | 0.4   | 0.272 | 0.138 | 1480  | 0.069   | 0.19    |
|         | NSE                    | 0.997 | 0.998 | 0.999 | 0.991 | 0.996   | 0.995   |
|         | RMSE                   | 0.02  | 0.02  | 0.02  | 0.02  | 0.02    | 0.02    |
|         | Best copula function   | 8.2   | 1     | 1     | 1     | 1       | 1       |

**TABLE 5** Statistical criteria obtained for the evaluation of the SPI and SPI$_{mod}$ indices

**F I G U R E 3** Fitted theoretical and empirical copula values of intensity and dry season data and the optimal parameter values obtained based on the copula Joe function for Urmia station
characteristics also increase as is shown in Figure 4b. From mathematical point of view, this finding confirms the correct operation of copula functions in drought analysis. For instance, as mentioned above in drought time series for Urmia station, when the duration of drought was longer than 6 months and the intensity was more than 6, the probability of their simultaneous occurrence was equal to 0.5. These diagrams also can be used for the analysis of other drought characteristics such as returns period, joint and conditional probabilities. Given the above mention explanations, it is possible that the intensity values of the drought duration will be more than one threshold level at the same time, providing useful information to water resources managers, to improve water resources management, farmers, the environment and other water-related areas in a state of drought.

3.4 | Conditional probability assessment

Figure 5 shows the conditional probability of severe drought in a mode where its duration exceeds a certain threshold, \(d\), for Urmia stations. For Urmia station, the probability of occurrence of droughts for intensities less than 4 and 6, and for the duration of drought more than 3 months is 0.55 and 0.57, respectively. The result also shows that as the duration of droughts decreases, the probability of their occurrence in different intensities increases, so that the probability of occurrence for droughts with a duration of 1 month is the highest, but the probability decreases for droughts lasting 6 months or more. This, in turn, is debatable, the reason for which should be sought in the probabilistic topics of mathematics in hydrological studies. In this case, the view of the
subject will be discrete and a sample space in which any size, the smaller the port of the fraction, the higher the probability of its occurrence. Hence, according to the science of probability of their occurrence, the shorter periods of drought are much longer than the other periods, since meteorological and hydrological phenomena are periodical and there is a possibility of their occurrence in future periods. This would be applicable if we want to save a water reservoir in an area so that this reservoir can provide downstream water storage during water shortages conditions for a variety of drinking, agricultural, environmental, recreational, hydroelectric and industrial conditions. These curves, therefore, provide useful and valuable information to water resource managers and planners.

3.5 | Conditional return period evaluation

To analyse the period of drought return in the conditional state, two states can be considered: the first state is when for a certain threshold of the drought period, its intensity value exceeds a certain limit \( T_{D_d > d} \); the second case is when for a certain threshold of drought intensity, its duration exceeds a certain limit \( T_{S_d \geq S} \). As is shown in Figure 6a for Urmia station, when intensity is more than 5 \( (D_s < 5) \) and drought period is more than 4 months \( (D_d < 4) \), the return period for such conditions is about 50 years, then a water supply reservoir cannot provide enough water for drought conditions. The results for other values of intensity and duration against the return period and vice versa are presented in Figure 6b. In general, it is observed that with increasing intensity and duration of drought, the return period increases as well. In addition to our knowledge and information on water and agricultural resource systems, these diagrams are also used in rangeland or watershed management systems before and after the construction of any flood distribution system.

3.6 | Evaluation of Kendall return period

According to studies, the use of the standard return period definition in various studies in hydrology and water resources estimates the amount of return period less than its actual value. To solve this problem, the Kendall return period was calculated based on the Kendall distribution functions (Ayantobo et al., 2018; Mirabbasi et al., 2012; Salvadori & De Michele, 2010), accordingly the two return periods of standard and Kendall were compared for Urmia stations as shown in Figure 7. This diagram is plotted for the critical probability levels \( (t \in I) \) of the station. Due to the small number of observational data, the long return period was not estimated based on the Kendall return period. In fact, the fracture in the Kendall return period is due to the small number of observational data. According to Figure 7, the difference between the return period of Kendall and the standard return period for Urmia station is 10 years, while the standard return period for the same probability level is 4.5, respectively. It is revealed that with increasing time, the difference between the Kendall return period and the standard return period increases. In this context, Durante and Salvadori (2010) and Salvadori et al. (2011) reported that the Kendall return periods were significantly different from simultaneous return periods. Such differences of the return period will lead to errors in drought risk analysis. However, using the definition of Kendall return period, the results of risk analysis are more accurate for two-variable cases.
4 | CONCLUSION

In this study, multivariate analysis of drought in Urmia Lake Basin was performed using Archimedean copula functions and standard precipitation indices (SPI) and modified precipitation index (SPI_mod). Then, using these indicators, bivariate characteristics of drought, namely severity and duration, were extracted and analysed. As drought phenomenon depends on different variables, univariate analysis of drought does not comprehensively demonstrate the influence of drought characteristics. Instead, the bivariate analysis can provide a more accurate description of drought phenomenon by considering the effect of both drought characteristics, that is, severity and duration. Therefore, in order to analyse the drought-related variables, eight Archimedean copula functions (i.e., Gamble-Hoguard, Frank, Clyton, Galambos, Placket, Farli-Gumble-Morgenstern, Ali-Mikhail-Hagh and Joe) were used. The results of fitting statistical distributions on drought characteristics showed that the exponential distribution function on drought duration and the gamma distribution function on drought intensity were known as superior margin distribution functions. The results of correlation analysis showed that these characteristics have significant correlation in the studied stations. The correlation coefficients of Spearman, linear correlation and Tau Kendall computed for the SPI of stations were >0.65, >0.72 and >0.48, respectively, while all of them were significant. Also, for all stations, the Nash–Sutcliffe coefficient was 0.99, which confirms the correct operation of the copula functions for bivariate analysis.

The results of determining the best copula functions showed that the evaluated Joe copula function was recognized as the superior copula function for bivariate analysis for the study area. In general, the results of bivariate drought analysis showed that the severity of drought over the period 1971–2017 was worse at Urmia station than at other stations, and was less severe at Tabriz station than at other stations. Also, the results of the return period in different cases indicate that in future periods there is a possibility of more severe drought in the Urmia Lake Basin.

Knowledge of the return period can provide useful information to water resources engineers in order to plan to determine the volume of reservoirs in drought periods in order to meet the water needs downstream of reservoirs as well as hydroelectric needs, environmental, industrial and agricultural. The use of multivariate analysis can also provide information to agricultural engineers and to managers of watersheds and rangelands.

AUTHOR CONTRIBUTIONS

Farzad Khezri: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); writing – original draft (equal).

Mohsen Irandoost: Conceptualization (equal); formal analysis (equal); methodology (equal); supervision (equal); validation (equal).

Navid Jalalkamali: Investigation (equal); methodology (equal); software (equal); validation (equal); writing – review and editing (equal).

Najme Yazdanpanah: Conceptualization (equal); investigation (equal); methodology (equal); validation (equal);
visualization (equal); writing – review and editing (equal).

**ORCID**
Mohsen Irandoost  
https://orcid.org/0000-0003-3110-4358

**REFERENCES**
Abdi, A., Hassanzadeh, Y., Talatahari, S., Fakheri-Fard, A. & Mirabbasi, R. (2017) Regional bivariate modeling of droughts using L-cornents and copula. *Stochastic Environmental Research and Risk Assessment*, 31(5), 1199–1210.

Achour, K., Meddi, M., Zeroual, A., Bouabdelli, S., Maccioni, P. & Moramarco, T. (2020) Spatio-temporal analysis and forecasting of drought in the plains of northwestern Algeria using the standardized precipitation index. *Journal of Earth System Science*, 129(1), 1–22.

Akaike, H. (1974) A new look at statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Amirataee, B., Montaseri, M. & Rezaie, H. (2018) Regional analysis and derivation of copula-based drought severity-area-frequency curve in Urmia Lake Basin, Iran. *Journal of Environmental Management*, 206, 134–144.

Ayantobo, O.O., Li, Y., Song, S., Javed, T. & Yao, N. (2018) Probabilistic modelling of drought events in China via 2-dimensional joint copula. *Journal of Hydrology*, 559, 373–391.

Ayantobo, O.O., Li, Y. & Song, S. (2019) Copula-based trivariate drought frequency analysis approach in seven climatic sub-regions of mainland China over 1961–2013. *Theoretical and Applied Climatology*, 137(3), 2217–2237.

Borji, M., Malekian, A., Salajegheh, A. & Ghadimi, M. (2016) Multi-time-scale analysis of hydrological drought forecasting using support vector regression (SVR) and artificial neural networks (ANN). *Arabian Journal of Geosciences*, 9(19), 1–10.

Bouabdelli, S., Meddi, M., Zeroual, A. & Alkama, R. (2020) Hydrological drought risk recurrence under climate change in the karst area of northwestern Algeria. *Journal of Water and Climate Change*, 11(51), 164–188.

Choubin, B., Khalighi-Sigaroodi, S. & Malekian, A. (2014) Drought forecasting in a semi-arid watershed using climate signals: a neuro-fuzzy modeling approach. *Journal of Mountain Science*, 11, 1593–1605.

Choubin, B., Malekian, A. & Golshani, M. (2016) Application of several data-driven techniques to predict a standardized precipitation index. *Atmosfera*, 29(2), 121–128.

Danandeh Mehr, A., Sorman, A.U., Kahya, E. & Hesami Afsar, M. (2020) Climate change impacts on meteorological drought using SPI and SPEI: case study of Ankara, Turkey. *Hydrological Sciences Journal*, 65(2), 254–268.

Das, J., Jha, S. & Goyal, M.K. (2020) Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India. *Journal of Hydrology*, 580, 124356.

De Michele, C. & Salvadori, G. (2003) A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research. Atmospheres*, 108, 4067–4078.

Dracup, J.A., Lee, K.S. & Paulson, E.G., Jr. (1980a) On the statistical characteristics of drought events. *Water Resources Research*, 16(2), 289–296.

Dracup, J.A., Lee, K.S. & Paulson, E.G. (1980b) On the definition of droughts. *Water Resources Research*, 16(2), 297–302.

Durante, F. & Salvadori, G. (2010) On the construction of multivariate extreme value models via copulas. *Environmetrics: Journal of the International Environmetrics Society*, 21(2), 143–161.

Heim, R.R., Jr. (2002) A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83(8), 1149–1166.

Joe, H. (1997) Multivariate models and dependence concepts. London: Chapman & Hall, p. 399.

Kao, S.C. & Govindaraju, R.S. (2010) A copula-based joint deficit index for droughts. *Journal of Hydrology*, 380(1–2), 121–134.

Kazemzadeh, M. & Malekian, A. (2016) Spatial characteristics and temporal trends of meteorological and hydrological droughts in northwestern Iran. *Natural Hazards*, 80(1), 191–210.

Malekian, A. & Kazemzadeh, M. (2016) Spatio-temporal analysis of regional trends and shift changes of autocorrelated temperature series in Urmia Lake Basin. *Water Resources Management*, 30(2), 785–803.

McKee, T.B., Doesken, N.J. & Kleist, J. (1993) The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th conference on applied climatology*, Anaheim, California, pp. 179–184.

Mirabbasi, R., Fakheri-Fard, A. & Dinapashoh, Y. (2012) Bivariate drought frequency analysis using the copula method. *Theoretical and Applied Climatology*, 108(1), 191–206.

Mohseni Saravi, M., Safdari, A.A. & Malekian, A. (2009) Intensity-duration-frequency and spatial analysis of droughts using the standardized precipitation index. *Hydrology and Earth System Sciences Discussions*, 6, 1347–1383.

Nash, J.E. & Sutcliffe, J.V. (1970) River flow forecasting through conceptual models. A discussion of principles. *Journal of Hydrology*, 10, 282–290.

Nelson, R.B. (2006) *An introduction to copulas*. New York: Springer, p. 269.

Pontes Filho, J.D., Portela, M.M., de Carvalho, M., Studart, T. & Souza Filho, F.D.A. (2019) A continuous drought probability monitoring system, CDPMs, based on copulas. *Water*, 11(9), 1925.

Salvadori, G. & De Michele, C. (2010) Multivariate multiparameter extreme value models and return periods: a copula approach. *Water Resources Research*, 46, W10501. https://doi.org/10.1029/2009WR009040.

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., Michele, C.D. & Durante, F. (2011) On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences*, 15(11), 3293–3305.

Schweizer, B. & Wolff, E.F. (1981) On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9(4), 879–885.

Shiau, J.T. (2006) Fitting drought duration and severity with two-dimensional copulas. *Water Resources Management*, 20(5), 795–815.

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.

Wang, F., Wang, Z., Yang, H., Di, D., Zhao, Y., Liang, Q. et al. (2020) Comprehensive evaluation of hydrological drought and
its relationships with meteorological drought in the Yellow River basin, China. *Journal of Hydrology*, 584, 124751.

Wong, G., Van Lanen, H.A.J. & Torfs, P.J.J.F. (2013) Probabilistic analysis of hydrological drought characteristics using meteorological drought. *Hydrological Sciences Journal*, 58(2), 253–270.

Xu, K., Yang, D., Yang, H., Li, Z., Qin, Y. & Shen, Y. (2015) Spatio-temporal variation of drought in China during 1961–2012: a climatic perspective. *Journal of Hydrology*, 526, 253–264.

How to cite this article: Khezri, F., Irandoost, M., Jalalkamali, N., & Yazdanpanah, N. (2021). Modelling of bivariate meteorological drought analysis in Lake Urmia Basin using Archimedean copula functions. *Meteorological Applications*, 28(6), e2040. [https://doi.org/10.1002/met.2040](https://doi.org/10.1002/met.2040)