Spatial-Temporal Change Analysis for Multivariate Drought Risk Based on Bayesian Copula: Application to the Balkhash Lake Basin

Xin Yang
Beijing Normal University
Yongping Li (yangxinkaka@vip.qq.com)
Beijing Normal University

Research Article

Keywords: Balkhash Lake Basin, Bayesian copula, multivariate drought risk, self-calibrating PDSI, spatial-temporal analysis

Posted Date: October 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-939186/v1

Spatial-temporal change analysis for multivariate drought risk based on Bayesian copula: Application to the Balkhash Lake Basin

X. Yang[1], Y. P. Li[1, 2, *]

[1] State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Beijing Normal University, China

[2] Institute for Energy, Environment and Sustainable Communities, University of Regina, Regina, Sask. S4S 0A2, Canada

[*]Corresponding author: Y. P. Li (E-mail: yongping.li@iseis.org)

E-mail addresses: j20_yangxin@163.com (X. Yang)

Abstract

In this study, a spatial-temporal Bayesian copula (SBC) method is developed through integrating spatial-temporal analysis and Bayesian copula into a general framework. SBC method can help model dependence structures of variable pairs and handle the uncertainty caused by parameter in copulas, and SBC can reveal the spatial and temporal changes of drought events. SBC is applied to the Balkhash Lake Basin (in Central Asia) to analyze spatial-temporal characteristic and drought risk in 1901-2020. Several findings can be summarized: (1) Balkhash Lake Basin suffered 53 drought events in 1901-2020, and five typical severe drought events occurred in 1916-1920, 1943-1945, 1973-1977, 1995-1998 and 2007-2009; (2) the most severe drought event lasted for 40 months (1973.10-1977.1), affecting 335,800 km^2 of the study basin; (3) drought usually develops from east to west, and Ili River delta and alluvial plain has the highest frequency of drought (47.2%), following by plateau desert (28.3%) and arid grassland in north Balkhash Lake (24.5%); (4) drought shows significant seasonality in the study basin, which usually begins in spring and summer (64.2%) and ends in summer and autumn (66.0%); and drought risk of middle and lower reaches of Ili River is highest in spring and summer; (5) in Balkhash Lake Basin, multivariate characteristics (duration, severity and affected area) significantly affect drought risk; (6) the range of drought risk is [1.9%, 18.1%], [3.7%, 33.1%], [8.7%, 46.0%], [16.0%, 55.1%] and [27.6%, 59.8%] when guarantee rate is 0.99, 0.98, 0.95, 0.90 and 0.80, respectively.

Keywords: Balkhash Lake Basin, Bayesian copula, multivariate drought risk, self-calibrating PDSI, spatial-temporal analysis

1. Introduction

Drought is the most widely affected natural disaster in the world and has adverse impact on agriculture, industrial production, urban water supply and ecological environment (Neisi et al., 2020). Over the past twenty years, climate change led to a 46% deterioration in drought conditions worldwide, which caused economic loss of 124 billion dollars, affecting more than 1.5 billion people (Ben et al., 2019; Ault, 2020). Drought generally includes four types: meteorological, agricultural, hydrological and socio-economic drought (Anne et al., 2016). Despite the widespread impact, drought identification and risk analysis are still challenging because of its different definitions, multivariate characteristics and spatial-temporal variability (Guo et al., 2018). Therefore, it is necessary to conduct monitoring and assessment in drought-prone areas to determine drought characteristic, spatial-temporal variation and multivariate interaction.

Over the past decades, many efforts have been devoted to drought monitoring and assessment, and more than fifty drought indices have been developed which are applicable in different regions (Wang et al., 2019). The most frequently-used indices include standard precipitation index (SPI), standardized precipitation evapotranspiration index (SPEI), hydrological drought index (HDI) and Palmer drought severity index (PDSI) (Zhao et al., 2017; Bohn et al., 2020; Fatemeh et al., 2021). SPI and SPEI are developed based on the discrepancy between precipitation and water balance, which are widely applied to meteorological drought (Hamal et al., 2020). HDI is developed by meteorological indicators and runoff, which represents a drought that river runoff is below the normal level; and HDI is usually applied to hydrological drought (Yang et al., 2020). On a large regional scale, areas with scarce precipitation and intense evapotranspiration are usually characterized by meteorological drought. However, watersheds that are significantly affected by seasonal changes in runoff are usually characterized by hydrological drought (Hu et al., 2019; Dehghan et al., 2020). Thus, it is complicated to analyze drought in a watershed with large seasonal variation in runoff which is located in arid area. Both meteorological factors (e.g., precipitation and evaporation) and underlying surface factors (e.g., soil moisture and runoff) needed to be considered. PDSI provides a water balance model that includes precipitation, evapotranspiration, runoff and soil moisture to describe drought of the watershed in arid area comprehensively (Palmer, 1965). The two limitations of PDSI are strong dependency on data calibration and shortcomings in spatial comparability (Wells et al., 2004). To overcome the limitations, self-calibrating PDSI (scPDSI) was developed and gradually being widely used (Liu et al., 2018; Akinwale et al., 2019; Zger et al., 2020).

Generally, drought is a three-dimensional spatial-temporal phenomenon, and the variation of a drought evolves both static and dynamic factors (Herrera et al., 2017; Diaz et al., 2020). Specifically, duration, severity and peak are static factors of a drought, and centroid, displacement direction and affected area are static factors of a drought. The analysis methods based on drought indices mainly analyze the changes and characteristics of drought events in two ways: one is to analyze the temporal changes of drought within a fixed area; the other is to analyze the spatial distribution of drought within a fixed period (Benjamin, 2012; Vernieuwe et al., 2019). For example, Xu et al. (2015) developed a 3-dimensional clustering method to identify drought events in China from 1961 to 2012 based on three indices, and five static factors were characterized. Guo et al. (2018) integrated principle components analysis, varimax rotation, Sen's slope and modified Mann-Kendall methods into a framework to identify the dynamic factors of drought in Central Asia from 1966 to 2015. Either way requires to reduce the three-dimensional spatial-temporal structure into a subspace (one-dimensional or two-dimensional space), which destroys the original structure and dilutes many inherent characteristics (Mellak et al., 2020; Yue et al., 2020). These analysis ways have a significant drawback, that is, although the multivariate characteristic are simplified in dimensionality reduction, the spatial-temporal correlation of drought is diluted. Therefore, more robust method is desired for accurately describing a drought event from both static and dynamic perspective, as well as quantitatively analyzing the interaction between multivariate factors.

Copula can provide a statistical way to model the dependence structure of multivariate factors (Zeroual et al., 2018; Foo et al., 2019; Soumia et al., 2020). For instance, Foo et al. (2019) described the correlation and dependency between drought variables through a trivariate copula model, and results disclosed drought properties of the peninsular Malaysia. Soumia et al. (2020) used Archimedean copula to fit severity–duration–frequency and severity–area–frequency curves, and results revealed the multidimensional drought characteristics in northern Algeria.

Copula has the main advantage of reveal drought risk by quantifying the correlation among factors which affect drought event, and it is convenient when modeling marginal distributions and multivariate dependence structures (Liu et al., 2020). However, copula also suffers several drawbacks such as verification of the optimal marginal distribution, enormous uncertainty of parameter estimation. Recently, to overcome the drawbacks, a number of researchers improved copula approaches with different statistical tools (Arbel et al., 2019; Zhao et al., 2020; Liu et al., 2021). For instance, Sadegh et al. (2017) developed a new multivariate copula analysis toolbox, which employed a Bayesian framework for inferring copula parameters and estimating the underlying uncertainties. Jin et al. (2019) proposed a Bayesian parameter identification approach for applying to advanced soil models, and its robustness and effectiveness were verified based on multiple independent calculations. Yang et al. (2020) combined maximum entropy principle, Bayesian copula into a general framework, which provided an efficient and accurate method for fitting optimal marginal distribution. Overall, using Bayesian inference to improve copula can minimize the uncertainty in parameter estimation.

This study aims to develop a spatial-temporal Bayesian copula (SBC) method for analyzing drought risk, through integrating spatial-temporal analysis and Bayesian copula into a general framework. The main novelty and contribution of this study can be listed as: (1) this is the first attempt to develop an integrated SBC method for analyzing multivariate (duration, severity and affected area) drought risk; (2) SBC is capable of modeling dependence structures of variable pairs and dealing with the uncertainty caused by parameter in copulas; (3) SBC can reveal the spatial and temporal changes of drought events; (4) SBC is applied to the Balkhash Lake Basin (in Central Asia) for drought risk analysis from 1901 to 2020; (5) the findings will be helpful to disclose drought risk of Balkhash Lake Basin in the past century.

2. Methodology

The SBC method integrates spatial-temporal analysis and Bayesian copula into a general framework (Figure 1). In detail, drought variables (e.g., duration, severity, affected area) are identified by using scPDSI and runs-theory. The correlation between drought variables are tested based on Pearson, Kendall and Spearman coefficients. Marginal distribution of drought variable is fitted by gamma, generalized extreme value, inverse Gaussian, log logistic, lognormal and Weibull. Four Archimedean copulas (i.e., Clayton, Frank, Gumbel, Joe) are employed to model dependence structures of variable pairs. The optimal marginal distribution and copula can be selected based on goodness-of-fit tests. Bayesian inference is used for dealing with uncertain parameters in copulas. Drought centroid and displacement direction are used for revealing the spatial-temporal changes of drought. Multivariate drought risk of the Balkhash Lake Basin is analyzed based on joint return periods and joint probabilities at different guarantee levels.

Place Figure 1 here

2.1 Spatial-temporal analysis

Drought is a natural disaster phenomenon linked with spatiality and temporality, while the drought duration, severity and affected area are static factors. This study introduces dynamic factors to analyze drought characteristics. Dynamic factors can describe the development and variation of drought spatially and temporally, including monthly drought centroid (*DC*) and drought displacement direction (*DD*). Monthly drought centroid refers to two-dimensional weighted centroid of monthly drought pattern shape, and its weight is determined by the absolute value of grid drought index. The rasterized drought index is imported into ArcGIS software, and monthly drought centroid can be obtained and visualized by using spatial analysis tools. Drought displacement direction is a basic description of drought path, which is determined by the fitting direction of monthly drought centroid. The longitude and latitude of the start point (P_s) are calculated average value of the longitude and latitude of monthly drought centroid of the first half of a drought event. The longitude and latitude of the end point (P_e) are calculated average value of the longitude and latitude of monthly drought centroid of the second half of a drought.

Drought displacement direction can be determined based on the start point and the end point, and the angle of displacement direction (θ) can be expressed as:

$$\theta = \arctan\left[\frac{|lonP_s - lonP_e|}{|latP_s - latP_e|}\right] \tag{1}$$

The longitude and latitude of P_s and P_e can be expressed as:

$$\begin{cases} P_s^{lat} = \dfrac{1}{2D}\sum_{t=1}^{\frac{1}{2}D} latP_t \\ P_s^{lon} = \dfrac{1}{2D}\sum_{t=1}^{\frac{1}{2}D} lonP_t \end{cases}, t \in \left(T_S, T_S + \frac{1}{2}D\right) \tag{2}$$

$$\begin{cases} P_e^{lat} = \dfrac{1}{2D}\sum_{t=1}^{\frac{1}{2}D} latP_t \\ P_e^{lon} = \dfrac{1}{2D}\sum_{t=1}^{\frac{1}{2}D} lonP_t \end{cases}, t \in \left(T_E - \frac{1}{2}D, T_E\right) \tag{3}$$

where T_s and T_e represents the start time and end time of a drought event, respectively; P_t is the monthly drought centroid (Herrera et al., 2017; Guo et al., 2018).

2.2 Bayesian copula and multivariate risk

Copula is applied to model dependence structures among correlated variable pairs. Based on Sklar theory, for a n-dimensional distribution function F, with univariate marginal F_1, …, F_n, a multivariate copula function C exists:

$$F(x_1, x_2, ..., x_n) = C(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \tag{4}$$

where $x_1, x_2, \ldots, x_n$ are measured values of $X_1, X_2, \ldots, X_n$; $F_1(x_1), F_2(x_2), \ldots, F_n(x_n)$ refer to the cumulative density functions of vectors $(X_1, X_2, \ldots, X_n)$. A unique copula exists when all marginal distributions are continuous and differentiable (Nelsen, 2006):

$$C\left(u_1,\ u_2,\ldots,u_n\right)=F\left(F_1^{-1}\left(u_1\right),\ F_2^{-1}\left(u_2\right),\ldots,F_n^{-1}\left(u_n\right)\right) \tag{5}$$

The probability density of a copula can be expressed as:

$$c\left(u_1,\ u_2\right)=\frac{\partial^2 C\left(u_1,\ u_2\right)}{\partial u_1 \partial u_2} \tag{6}$$

and the joint probability density of variable pairs can be express as:

$$f\left(x_1,\ x_2\right)=\frac{\partial^2 C\left(u_1,\ u_2\right)}{\partial x_1 \partial x_2}=\frac{\partial^2 C\left(u_1,\ u_2\right)}{\partial u_1 \partial u_2}\frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_2}=f_{X_1}\left(x_1\right)f_{X_2}\left(x_2\right)c\left(u_1,\ u_2\right) \tag{7}$$

Four Archimedean copulas are used to screen out the optima one for multivariate dependence structures modeling. The cumulative probability U_1 (when $U_2 = u_2$) can be expressed as:

$$C_{U_1|U_2=u_2}\left(u_1\right)=P\left(U_1 \le u_1 \middle| U_2=u_2\right)=\frac{\partial}{\partial u_2}C\left(u_1,\ u_2\right) \tag{8}$$

Similarly, the cumulative probability U_2 (when $U_1 = u_1$) can be expressed as:

$$C_{U_1|U_2\le u_2}\left(u_1\right)=P\left(U_1 \le u_1 \middle| U_2 \le u_2\right)=\frac{C\left(u_1,\ u_2\right)}{u_2} \tag{9}$$

Based on Bayesian inference, MCMC simulation is applied to take samples from high-dimensional distributions. Bayesian inference indicates that model uncertainties come from the parameters, and the posterior distribution of parameters can be expressed as (Haario et al., 2006):

$$p\left(\theta,Y\right)=\frac{p\left(\theta\right)p\left(Y,\theta\right)}{p\left(Y\right)}\propto p\left(\theta\right)p\left(Y,\theta\right) \tag{10}$$

where $p(\theta)$ and $p(\theta, Y)$ signify prior and posterior distribution of parameters, respectively. $p(Y, \theta)$ denotes likelihood function, and $p(Y)$ is coned evidence (Yang et al., 2020). Then, according to the parameter distribution, the estimated parameter in the 95% confidence interval was selected as the calculation input of the copula function.

Drought return period (T) is a common reference for designing drought defense infrastructure. In multivariate risk analysis, T can be extends to the joint return periods (T_{and} and T_{or}) (Montaseri et al., 2018):

$$T_{u_1,u_2}^{AND} = \frac{E(L)}{1 - u_1 - u_2 + C_{U_1,U_2}(u_1, u_2)} \tag{11}$$

$$T_{u_1,u_2}^{OR} = \frac{E(L)}{1 - C_{U_1,U_2}(u_1, u_2)} \tag{12}$$

where $E(L)$ represents the mean interval time of two consecutive drought events. Therefore, the bivariate risk indictor R is defined as:

$$R_{u_1,u_2} = 1 - \left(1 - \frac{1}{T_{u_1,u_2}^{AND}}\right)^n \tag{13}$$

where $R_{u1,u2}$ is the joint risk of u_1 and u_2, and n is the design life of drought defense infrastructures.

Since the correlation of random variables would significantly affect the result of copula function, it is essential to examine the dependence structure of random variables before apply copula to joint probability distribution. Three correlation tests, including Pearson (γ), Kendall (τ) and Spearman (ρ) are used. Kendall and Spearman correlation coefficients are suitable for ordinal variables that do not meet the normal distribution hypothesis, while Pearson correlation coefficient is applicable to continuous variables (Sheng et al., 2002; Sedgwick, 2012). Variable pairs would be used for dependence structure modeling if their correlation is significant. Several measures, such as Akaike information criterion (AIC), and Bayesian information criterion (BIC), root mean square error (RMSE) are used to test goodness-of-fit of marginal distributions and copula functions.

3. Case study

3.1 Study area

Balkhash Lake is a closed terminal lake located at 73°20'E-79°12'E, 45°00'N-46°44'N in Central Asia. The lake stretches from east to west over 600 km, and width of the lake is 9-19 km in the eastern part and 74 km in the western part. The surface area of Balkhash Lake is fluctuant which ranges 17,000-22,000 km^2, and the average depth of the lake is 6 m (Isbekov et al., 2019). The supplement of lake water consists of surface runoff, precipitation and groundwater, among which the main volume of water flowing is supplied by the river runoff (over 70%). In recent decades with the gradual drying up of the Aral Sea, Balkhash Lake has become the largest lake in Central Asia (Aizhan, 2020). Balkhash Lake Basin covers an area of 413,000 km^2, and the principle part is located in Kazakhstan (86%), and the rest is in China (14%). The basin lies in an arid and semi-arid zone with an annual mean precipitation of 110 mm and an annual mean temperature of 17.5 °C in the past century (Duan et al., 2020). Since 1970, a substantial runoff decrease in Ili river (main supply, 78%) has led to a drawdown of water reaching the Balkhash Lake, resulting in numerous environmental problems (e.g., drought, desertification, salinization). Because the basin is situated in a desert area, with little precipitation and intense evaporation, the species survival and social development are facing serious challenges. For example, meteorological, hydrological and agricultural droughts occur frequently which significantly affect industrial and agricultural production and human life. This study concerns the principle part of the Balkhash Lake Basin in Kazakhstan (Figure 2), because the catchment of the rest part is small, and the socio-economic conditions and water resources management between Kazakhstan and China are quite different. The study basin includes the territories of Almaty, south-eastern Karaganda, south-western East Kazakhstan and eastern Zhambyl Oblast, which totally cover an area of 355,000 km^2. The terrain is high in the southeast and low in the northwest, and the study basin can be divided into three regions: (1) arid grassland in north Balkhash Lake; (2) Ili River delta and alluvial plain; (3) plateau desert. Water resources in Balkhash Lake Basin are mainly from Ili River (11 km^3/year) and other mountain rivers (3 km^3/year). Ili River is the most important supplement of Balkhash Lake and also the main source of social production and living water. Runoff primarily originates from rainfall and the melting of snow and ice, which are vulnerable to climate change, leading to increasing drought risk.

Place Figure 2 here

3.2 Data collection

In this study, the topographic characteristic of Balkhash Lake Basin is depicted based on digital elevation model (DEM), and DEM data can be download in National Tibetan Plateau Third Pole Environment Data Center (TPDC, https://data.tpdc.ac.cn/en/). The gridded monthly self-calibrating PDSI (0.5°×0.5°) is used for identifying the drought, which can be downloaded at Royal Netherlands Meteorological Institute (KNMI) Climate Explorer website (http://climexp.knmi.nl). Self-calibrating PDSI (scPDSI) is a kind of raster data initially, which needs to be extracted by operating an MATLAB program and visualized by using ArcGIS software (Appendix).

4. Result and discussion

4.1 Spatial-temporal change analysis of drought

Figure 3 illustrates the temporal evolution of monthly self-calibrating PDSI and frequency histograms of dry and wet months in 1901-2020. Monthly histograms show that frequent variation between dry and wet periods is fluctuant without a regular pattern. In historical period, dry and wet periods account for 51.6% and 23.2%, respectively, and normal periods account for 25.1%. The amount of dry periods is significantly more than wet periods, which indicates that the Balkhash Lake Basin was dominated by drought in the historical period. Based on runs-theory, 53 drought events occurred in the Balkhash Lake Basin in 1901-2020, and Table 1 shows the characteristics of each drought event, which includes initial/terminal time, duration, severity and drought area ratio. The characteristic information identified in Table 1 would be helpful to understand the historical period of the drought in the Balkhash Lake Basin, and is also the basis for the next step of multivariate analysis.

Place Table 1 here

Five typical severe drought events (SDE) are highlighted in Table 1, which indicates the duration, severity and affected area are significantly greater than other drought events. The spatial accumulation of scPDSI of five severe drought events are shown in Figure 4. The left, middle and right panel in each typical severe drought periods represents the distribution of scPDSI at the beginning, middle and end of each drought period respectively. For example, in Figure 4(a), the left panel represents the distribution of scPDSI in August in 1916, the middle panel represents the distribution of scPDSI in July in 1918, and the end panel represents the distribution of scPDSI in July in 1920. SDE 1 occurred in August in 1916 to July in 1920, among which June to October in 1918 was an extreme drought (S_{max} = 3.72, A_{max} = 217,000 km^2).

Plateau desert in the basin and the mountainous region in the upper reaches of Ili River were the areas where the drought was concentrated (in deep red grids). SDE 2 occurred in September in 1943 to October in 1945, among which May to September in 1944 was an extreme drought (S_{max} = 3.84, A_{max} = 244,000 km^2). Drought concentrated in the east and west ends of arid grassland in the north of Balkhash Lake. SDE 3 occurred in October in 1973 to January in 1977, among which April in 1975 to June in 1976 was an extreme drought (S_{max} = 3.66, A_{max} = 336,000 km^2). Drought gradually developed from arid grassland in the north of Balkhash Lake to Ili River delta and alluvial plain. SDE 4 occurred in February in 1995 to February to 1998, among which July to October in 1997 was an extreme drought (S_{max} = 3.79, A_{max} = 251,000 km^2). Almost the entire Balkhash Lake Basin was affected by the drought. SDE 5 occurred in August in 2007 to July in 2009, among which May to September in 2008 was an extreme drought (S_{max} = 3.71, A_{max} = 294,000 km^2). The western part of Ili River delta and alluvial plain and middle of arid grassland in the north of Balkhash Lake were affected. Generally, extreme droughts occurred within the periods of the severe droughts. By comparing the mean self-calibrating PDSI of each decade, the periods of 1911-1920, 1921-1930, 1931-1940, 1961-1970, 1971-1980 and 1991-2000 were in drought state, because the average annual self-calibrating PDSI of each period was less than -1.0. During the period from 1931 to 1940, the drought was the most serious, with annual average self-calibrating PDSI of -1.26, which indicates that the Balkhash Lake Basin was in the state of slight drought almost every year. This is mainly because of the rapid development of human activities in this region since the 20th century, which makes the situation of drought caused by water shortage further aggravated. In only two periods, 1951-1960 and 2011-2020, Balkhash Lake Basin was non-drought with an average annual PDSI of +0.36 and +0.64, respectively. In 2011-2020, increased climate change caused accelerated snow melt from the upstream glaciers, and protection measures in Balkhash Lake Basin since the end of the 20th century have reduced drought caused by water scarcity.

Place Figures 3 and 4 here

Runoff of the rivers is an important factor affecting the drought in Balkhash Lake Basin that is significantly affected by the seasonal variation. The analysis of the distribution of dry period in each season would be helpful for reflecting the seasonal characteristics of drought. From the seasonal histograms of the beginning and ending time of drought, most drought events occur in the spring and summer (34 times), accounting for about 64% of the total amount (Figure 5). In spring and summer, the number of drought events lasting 1-6 months, 6-12 months, 12-24 months and longer than 24 months are 10, 3, 2 and 2, respectively. Most of the drought events end in summer and autumn (35 times), accounting for about 66% of the total amount. In summer, the number of drought events lasting 1-6 months, 6-12 months, 12-24 months and longer than 24 months are 11, 2, 3 and 3, and in autumn the numbers are 9, 4, 2 and 1, respectively. In Balkhash Lake Basin, the beginning and ending of drought events were the least in winter and the most in summer. Therefore, summer is the crucial period of drought prevention in the study basin. From the regional spatial distribution of drought times, there are 25 drought events in the Ili River delta and alluvial plain, accounting for 47.2% of the total amounts, among which the droughts with four different duration scale are the most by comparing with other areas. There are 15 drought events in plateau desert area, accounting for 28.3%. Drought events occurred 13 times in arid grassland in the north of Balkhash Lake, accounting for 24.5%. This indicates that the drought is most severe in the Ili River delta and alluvial plain, where the human activity is very intense.

Place Figure 5 here

The spatial characteristics of drought can describe the development and change of a drought event. The most representative characteristics are the drought centroid and the displacement direction of drought. In this study, the centroids of 53 drought events was made spatial statistics according to 1-6 months, 6-12 months, 12-24 months and longer than 24 months, and the development direction of each drought event was obtained (Figure 6). The development direction of drought events in Balkhash Basin is mainly "southeast to northwest", accounting for about 50% of the total drought times. This development direction is significantly correlated with the flow direction of Ili River, especially in the alluvial plain and delta area. Since the water resources of Ili River account for about 80% of the total water inflow into Balkhash Lake, it that the water quantity variation of Ili River will greatly affect the agricultural drought in the Balkhash Lake Basin.

Place Figure 6 here

4.2 Multivariate risk analysis of drought

4.2.1 Dependence structure modeling based on Bayesian copula

Duration, severity and area can be regarded as three dimensional attributes characterizing a drought event. Based on the drought characteristics analysis, Bayesian copula is applied to model the dependence structure of drought variables in order to reveal the influence of multivariate characteristics interaction on drought risk. Since only variables with significant correlation can be analyzed as dependence structures, the correlation test of variables should be verified first. In this study, two non-parametric measures (Kendall's τ and Spearman's ρ) and one linear correlation measure (Pearson's γ) were used to test the correlation among drought duration (D), severity (S) and area (A). Table 2 presents the correlation test results, which indicates that the correlation coefficient of variable pairs of duration-severity is the highest, followed by severity-area and duration-area. All these three pairs pass the significant test at 5% level. Consequently, it is necessary to consider the influence of the interaction among variables when analyzing drought risk, otherwise the results are likely to be biased.

Place Table 2 here

Bayesian copula has the main advantage that marginal distribution and dependence structure modeling are separate parts which cannot interfere each other. Marginal distribution should be quantified first, and Figure 7 illustrates the fitted marginal distributions through Gamma, general extreme value (GEV), inverse Gaussian (INGAU), log logistic (LOL), lognormal (LOGN) and Weibull (WBL). To all these distributions, both probability and cumulative distribution functions show good agreement between the theoretical and empirical distributions. Thus, AICc test was applied to select the optimal distribution for Bayesian copula. Table 3 shows that GEV is the optimal distribution of duration, severity and area, because the AICc values of GEV are always minimum by comparing with other distributions. Through cumulative distribution functions, the threshold of duration, severity and area in different return periods (corresponding to different guarantee rates) can be defined, which can be used as a reference for drought design under univariate scenario.

Place Table 3 and Figure 7 here

Four common-used copulas, including Clayton, Frank, Gumbel and Joe, were applied to model dependence structures of variables, and the unknown parameters in copulas were estimated by using MCMC simulation. The first step of selecting the optimal copula is to determine whether posterior parameters of the alternative copulas were well constrained. If the estimated parameter of a copula merge to the bounds, there is a chance that this copula not a good fit. In drought variable pairs of duration-severity, the parameters of Frank and Joe copulas (blue cross on the bottom of each plot) are converging to the parameter boundaries. In drought variable pairs of duration-area, the parameters of Frank Gumbel, and Joe copulas (blue cross on the bottom of each plot) are all converging to the parameter boundaries. In drought variable pairs of severity-area, the parameters of Clayton, Frank, Gumbel and Joe are in the center of the distribution histograms (Figure 8). Copula functions with inappropriate parameter distribution are not considered as an option for further analysis. The second step is to determine the optimal copula according to AIC and BIC (Table 4). By comparing the AIC and BIC values among the copulas, Clayton is the optimal copula to model dependence structure of duration-severity due to the minimum AIC (-239.60) and BIC (-237.63). Similarly, Clayton is also the optimal copula to model dependence structures of duration-area and severity-area.

Place Figure 8 and Table 4 here

4.2.2 Multivariate risk analysis based on joint return period

Figure 9 illustrates the joint distributions of duration-severity, duration-area and severity-area which obtained through the Clayton copula, and the corresponding contour plots are listed. The blue point represents the empirical value of drought pair (observed data), which was identified by using runs-theory. The color contour lines represent the theoretical copula through $C(u, v)=(u^{-\theta}+v^{-\theta}-1)^{-1/\theta}$. According to Table 5, among variable pairs of duration-severity, duration-area and severity-area, the values of $C(u, v)$ would be less than the corresponding P-level. For example, at p-level of 0.95, the joint probabilities of drought variable pairs are 0.93, 0.91 and 0.91, respectively. $T_{\text{and}}(u, v)$ is much longer than the corresponding return period, while $T_{\text{or}}(u, v)$ is shorter than the corresponding return period. For example, when return periods of duration and severity are both 100-year (T=100), $T_{\text{and}}(d, s)$ would be 500 years, while $T_{\text{or}}(d, s)$ is 56 years.

$T_{and}(d, a)$ is 4485 years and $T_{and}(s, a)$ is 5308 years, which both are much longer than $T_{and}(d, s)$ due to the lower correlation of duration-area and severity-area. Besides, $T_{or}(d, a)$ and $T_{or}(s, a)$ are shorter than $T_{or}(d, s)$. The same results can be concluded by comparing with the other p-levels, which indicate that univariate return period is significantly different from joint return period, and the univariate return period cannot reflect the real situation of drought. In multivariate situations, different correlations among variables would also lead to different joint return periods. Therefore, the variables should be selected according to the main characteristics of the specific drought conditions in the study area. If the drought includes more than two typical characteristic variables, the multiple variables can be coupled into different pairs, and the maximum and minimum values of the joint return periods based on different variables pairs can be used as the upper and lower bounds of the actual return periods respectively.

With the decrease of p-level, the deviation between joint return period (both T_{and} and T_{or}) and univariate return period also shrinks. For example, when the p-level decreases from 0.99 to 0.80, T_{and} of duration-severity drops from 125 years to 6 years, and T_{or} drops from 6 years to 5 years. Univariate return period becoming shorter indicates the drought risk would increase. The return periods ranged from 100-year to 5-year infer that univariate drought risk increased from 1% to 20%. However, univariate drought risk based on p-level is inadequate to reveal the actual risk. In this study, multivariate drought risk based on Bayesian copula analyzed. Taking the interaction of drought variables into account, drought risk of duration-severity pair would be modified as 18.1%, 33.1%, 46.0%, 55.1% and 59.8% when p-level is 0.99, 0.98, 0.95, 0.90 and 0.80, respectively. To duration-area pair, drought risk would be modified as 2.2%, 4.3%, 9.5%, 18.3% and 31.0%, and to severity-area pair, drought risk would be modified as 1.9%, 3.7%, 8.7%, 16.0% and 27.6%. Obviously, multivariate risk is significantly higher than univariate risk at each p-level, which discloses that the univariate risk underestimates the actual drought risk. If univariate risk results were applied to drought management, it would lead to an inability to accurately estimate drought risk, resulting in the losses of social economy. Multivariate drought risk analysis results in Table 5 further indicate that in the Balkhash Lake Basin, the concurrence of a drought with long duration (> 43 months), high severity (> 87) and area (> 84%) would not occur frequently, while the concurrence of a drought with short duration (> 15 months), low severity (> 26) and peak (> 69%) would occur quite often.

Place Figure 9 and Table 5 here

Therefore, joint probability of duration, severity and area can be used to analyze the drought risk. The actual drought risk would be underestimated if only $T_{\mathrm{and}}(u, v)$ is considered, whereas the risk would be overestimated if only $T_{\mathrm{or}}(u, v)$ is considered. In practical application, drought risk of $T_{\mathrm{and}}(u, v)$ can be regarded as the upper bound of actual situation, and drought risk of T $T_{\mathrm{or}}(u, v)$ can be regarded as the lower bound. In Balkhash Lake Basin, the range of drought risk would be [1.9%, 18.1%], [3.7%, 33.1%], [8.7%, 46.0%], [16.0%, 55.1%] and [27.6%, 59.8%] when guarantee rate is 0.99, 0.98, 0.95, 0.90 and 0.80, respectively. In general, guarantee rate of 0.95 can meet the demand of drought resistance. These findings suggest that considering the interaction of variables can reduce calculation errors when analyzing drought risk. The expected value of typical drought characteristics under the frequent occurrence and not frequent occurrence would be helpful for reflecting the drought situation of Balkhash Lake Basin from a general perspective.

5. Conclusions

In this study, a spatial-temporal Bayesian copula (SBC) method has been developed through integrating spatial-temporal analysis and Bayesian copula into a general framework. SBC method can help model dependence structures of variable pairs and handle the uncertainty caused by parameter in copulas, and SBC can reveal the spatial and temporal changes of drought events. A case study of the Balkhash Lake Basin has been used for demonstrating the applicability of SBC. Drought risk in the historical period (1901-2020) is analyzed based on self-calibrating Palmer drought severity index.

Some major findings can be summarized as: (1) Balkhash Lake Basin suffered 53 drought events in 1901-2020, and five typical severe drought events occurred in 1916-1920, 1943-1945, 1973-1977, 1995-1998 and 2007-2009, respectively; (2) the most severe drought event occurred in October in 1973 to January in 1977, lasting for 40 months and developing to an extreme drought during April in 1975 to June in 1976, affecting 95% of the study basin (335,800 km^2); (3) most of the drought event in Balkhash Lake Basin developed in the direction of east to west; drought frequency is different in three sub regions; Ili River delta and alluvial plain were the most (47.2%), following by plateau desert area (28.3%) and the arid grassland in north Balkhash Lake (24.5%); (4) drought has significant seasonality in the study basin, which begins in spring and summer (64.2%) and ends in summer and autumn (66.0%) frequently; and drought risk of the middle and lower reaches of Ili River is highest in spring and summer; (5) in Balkhash Lake Basin, multivariate characteristics (i.e., duration, severity and affected area) significantly affect drought risk; (6) the range of drought risk is [1.9%, 18.1%], [3.7%, 33.1%], [8.7%, 46.0%], [16.0%, 55.1%] and [27.6%, 59.8%] when guarantee rate is 0.99, 0.98, 0.95, 0.90 and 0.80, respectively.

Acknowledgements

This research was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA20060302).

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Statement

This research was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA20060302).

Author's Contribution

X. Yang: Conceptualization, Methodology, Data curation, Writing original draft, Writing review & editing. Y.P. Li: Supervision, Writing review & editing, Project administration, Funding acquisition.

Availability of data and material

The data sets supporting the results of this article are included within the article. The datasets generated during and/or analyzed during the current study are available in the National Tibetan Plateau Third Pole Environment Data Center [https://data.tpdc.ac.cn/en/], and Royal Netherlands Meteorological Institute [http://climexp.knmi.nl].

Code availability

MATLAB program for self-calibrating PDSI extraction and visualization.

Step 1: Model sample data

```
1. ncdisp('H:\Global\PDSI\scPDSI.cru.3.25.bams2018.GLOBAL.1901.2017.nc');
2. data1=ncread('H:\Global\PDSI\scPDSI.cru.3.26.bams2018.GLOBAL.1901.2017.nc','scpdsi');
```

```
3. data3=data1(:,:,1);
4. data4=rot90(data3);
5. data5=flipud(data4);
6. data5(isnan(data5))=-999;
7. dlmwrite('sample_1.txt',data5,'\t',1,1)
```

Step 2: Add latitude and longitude information to sample_1.txt

```
1. ncols 720
2. nrows 360
3. xllcorner -180
4. yllcorner -90
5. cellsize 0.5
6. NODATA_value -999
```

Step 3: Rasterize the sample_1.txt by ASCII code in ArcGIS, and output it as sample_1.tif

Step 4: Load a raster file with projection information, and define the projection on the example_1.tif

Step 5: Batch processing

```
1. [aaaaa,R]=geotiffread('H:\Global\PDSI\example_1.tif');
2. info=geotiffinfo('H:\Global\PDSI\example_1.tif');
3. data=ncread('H:\Global\PDSI\scPDSI.cru.3.26.bams2018.GLOBAL.1901.2017.nc','scpdsi');
4. for year=1901:2017
5.          data1=data(:,:,1+12*(year-1901):12*(year-1900));
6.          data3=sum(data1,3)/12;
7.          data4=rot90(data3);
8.          data5=flipud(data4);
9.          filename=strcat('H:\Global\PDSI\yearly_pdsi\global',int2str(year),'yearly_PDSI.tif');
10.         geotiffwrite(filename, data5, R, 'GeoKeyDirectoryTag', info.GeoTIFFTags.GeoKeyDirectoryTag);
```

```
    for mon=1:12
        data2=data1(:,:,mon);
        data4=rot90(data2);
        data5=flipud(data4);
        filename=strcat('H:\Global\PDSI\monthly_pdsi\global', int2str(year), '_', int2str(mon), 'monthly_PDSI.tif');
geotiffwrite(filename, data5, R, 'GeoKeyDirectoryTag', info.GeoTIFFTags.GeoKeyDirectoryTag);
    end
end
```

Ethics approval

We the undersigned declare that this manuscript entitled "*Spatial-temporal change analysis for multivariate drought risk based on Bayesian copula: Application to the Balkhash Lake Basin*" is original, has not been published before and is not currently being considered for publication elsewhere.

Consent to participate

Not applicable

Consent for publication

Not applicable

References

Aizhan, U., 2020. Save Kazakhstan's shrinking Lake Balkhash. *Science*, 370(6514), 303.

Akinwale, T. O., Phillip, G. O., David, A. O., 2019. Drought spatiotemporal characterization using self-calibrating Palmer Drought Severity Index in the northern region of Nigeria. *Results in Engineering*, 3, 100088.

Ali, M., Deo, R. C., Downs, N. J., 2018. Multi-stage hybridized online sequential extreme learning machine integrated with Markov Chain Monte Carlo copula-Bat algorithm for rainfall forecasting. *Atmospheric Research*, 213, 450-464.

Anne, F., Kerstin, S., Giuliano, D. S., 2016. Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches. *Hydrology and Earth System Sciences*, 20(9), 3631-3650.

Arbel, J., Crispino, M., Girard, S., 2019. Dependence properties and Bayesian inference for asymmetric multivariate copulas. *Journal of Multivariate Analysis*, 174, 104530.

Ault, T. R., 2020. On the essentials of drought in a changing climate. *Science*, 368(6488), 256-260.

Ben, E., Matthew, G., Boyd, H., 2019. The social and economic impacts of drought. *Australian Journal of Social Issues*, 54(1), 22-31.

Benjamin, L. H., 2012. A spatio-temporal structure-based approach to drought characterization. *International Journal of Climatology*, 32(3), 406-418.

Bhatti, M. I., Do, H. Q., 2019. Recent development in copula and its applications to the energy, forestry and environmental sciences. *International Journal of Hydrogen Energy*, 44(36), 19453-19473.

Bohn, V. Y., Raúl, R., Varni, M., 2020. Using SPEI in predicting water table dynamics in Argentinian plains. *Environmental Earth Sciences*, 79(19), 1-16.

Dehghan, S., Salehnia, N., Sayari, N., 2020. Prediction of meteorological drought in arid and semi-arid regions using PDSI and SDSM: A case study in Fars province, Iran. *Journal of Arid Land*, 12(2), 318-330.

Diaz, V., Perez, G., Van L. H., 2020. An approach to characterize spatio-temporal drought dynamics. *Advances in Water Resources*, 137, 103512.

Duan, W., Zou, S., Chen, Y., 2020. Sustainable water management for cross-border resources: the Balkhash Lake basin of Central Asia, 1931–2015. *Journal of Cleaner Production*, 263, 121614.

Fatemeh, H., Farzad, H., Mostafa, Y., 2021. Evaluation of drought characterization using SPI and SC-PDSI drought indices in baseline and upcoming periods in Birjand region. *Arabian*

Journal of Geosciences, 14(11), 939.
Foo, H., Fadhilah, Y., Zulkifli, Y.. 2019. Trivariate copula in drought analysis: a case study in peninsular Malaysia. *Theoretical and Applied Climatology*, 138(1-2), 657-671.
Guo, H., Bao, A., Liu, T., 2018. Spatial and temporal characteristics of droughts in Central Asia during 1966–2015. *Science of the Total Environment*, 624, 1523–1538.
Haario, H., Marko, L., Antonietta, M., 2006. Dram: efficient adaptive MCMC. *Statistics & Computing*, 16(4), 339-354.
Hamal, K., Sharma, S., Khadka, N., 2020. Assessment of drought impacts on crop yields across Nepal during 1987-2017. *Meteorological Applications*, 27(5), 1-18.
Herrera, J. E., Satoh, Y., Sheffield, J., 2017. Spatiotemporal dynamics of global drought. *Geophysical Research Letters*, 44, 2254–2263.
Hu, Z. Y., Chen, X., Chen, D. L., 2019. "Dry gets drier, wet gets wetter": A case study over the arid regions of central Asia. *International Journal of Climatology*, 39(2), 1072-1091.
Isbekov, K. B., Tsoy, V. N., Crétaux, J., 2019. Impacts of water level changes in the fauna, flora and physical properties over the Balkhash Lake watershed. *Lakes & Reservoirs*, 24(2), 195-208.
Jia, H., Xu, T., Liang, S., 2018. Bayesian framework of parameter sensitivity, uncertainty, and identifiability analysis in complex water quality models. *Environmental Modelling & Software*, 104, 13-26.
Jin, Y. F., Yin, Z. Y., Zhou, W. H,. 2019. Identifying parameters of advanced soil models using an enhanced transitional Markov chain Monte Carlo method. *Acta Geotechnica*, 14(6), 1925-1947.
Liu, Y. R., Li, Y. P., Ma, Y., 2020. Development of a Bayesian-copula-based frequency analysis method for hydrological risk assessment – The Naryn River in Central Asia. *Journal of Hydrology*, 580, 124349.
Liu, Y. R., Li, Y. P., Yang, X., 2021. Development of an integrated multivariate trend-frequency analysis method: Spatial-temporal characteristics of climate extremes under global warming for Central Asia. *Environmental Research*, 195, 110859.
Liu, Z., Zhang, X., Fang, R., 2018. Multi-scale linkages of winter drought variability to enso and the arctic oscillation: A case study in Shaanxi, North China. *Atmospheric Research*, 200, 117-125.

Mellak, S., Souag, G. D., 2020. Spatio-temporal analysis of maximum drought severity using copulas in northern Algeria. *Journal of Water and Climate Change*, 11, 68-84.

Montaseri, M., Amirataee, B., Rezaie, H., 2018. New approach in bivariate drought duration and severity analysis. *Journal of Hydrology*, 559, 166-181.

Neisi, M., Bijani, M., Abbasi, E., 2020. Analyzing farmers' drought risk management behavior: Evidence from Iran. *Journal of Hydrology*, 590, 125243.

Nelsen, R. B., 2006. An Introduction to Copulas. Springer, New York.

Palmer, W.C.. 1965. Meteorological Drought. Research Paper No. 45. US Weather Bureau, Washington, DC.

Rana, S. M., Boccelli, D. L., Scott, D. T., 2019. Parameter uncertainty with flow variation of the one-dimensional solute transport model for small streams using Markov chain Monte Carlo. *Journal of Hydrology*, 575, 1145-1154.

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.

Sedgwick, P., 2012. Pearson's correlation coefficient. *BMJ*, 345, 4483-4483.

Sheng, Y., Paul, P., George, C., 2002. Power of the Mann-Kendall and spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, 259, 254-271.

Soumia, M., Doudja, S., 2020. Spatio-temporal analysis of maximum drought severity using Copulas in Northern Algeria. *Journal of Water and Climate Change*, 11(1), 68-84.

Vernieuwe, H., Baets, B. D., Verhoest, N. E., 2019. A mathematical morphology approach for a qualitative exploration of drought events in space and time. *International Journal of Climatology*, 40, 1-14.

Wang, L., Yu, H., Yang, M. L., 2019. A drought index: The standardized precipitation evapotranspiration runoff index. *Journal of Hydrology*, 571, 651-668.

Xu, K., Yang, D. W., Yang, H. B., 2015. Spatio-temporal variation of drought in China during 1961–2012: A climatic perspective. *Journal of Hydrology*, 526, 253-264.

Yang, X., Li, Y. P., Liu, Y. R., 2020. A MCMC-based maximum entropy copula method for bivariate drought risk analysis of the Amu Darya River Basin. *Journal of Hydrology*, 590, 125502.

Yue, P., Lee, H., 2020. Drought-induced spatio-temporal synchrony of plague outbreak in

Europe. *Science of The Total Environment*, 698, 134138.

Zeroual, A., Bouabdelli, S., Meddi, M., 2018. Trivariate Copulas for Characterization of Past and Future Droughts Over North-West Algeria. Euro-Mediterranean, Conference for Environmental Integration. Springer, Cham.

Zger, M., Baakn, E. E., Ekmekciolu, M., 2020. Comparison of wavelet and empirical mode decomposition hybrid models in drought prediction. *Computers and Electronics in Agriculture*, 179, 105851.

Zhang, J. L., Li, Y. P., Huang, G. H., 2016. Assessment of parameter uncertainty in hydrological model using a Markov-Chain-Monte-Carlo-based multilevel-factorial-analysis method. *Journal of Hydrology*, 538, 471-486.

Zhang, Y., Huang, S., Huang, Q.. 2019. Assessment of drought evolution characteristics based on a nonparametric and trivariate integrated drought index. *Journal of Hydrology*, 579, 124230.

Zhao, H., Gao, G., An, W., 2017. Timescale differences between sc-PDSI and SPEI for drought monitoring in China. *Physics & Chemistry of the Earth*, 102, 48-58.

Zhao, Z. Y., Wang, H. R., Chen, Y., 2020. Changes in spatiotemporal drought characteristics over northeast China from 1960 to 2018 based on the modified nested Copula model. *The Science of the total environment*, 739, 140328.

List of Table Captions

Table 1 Characteristics of the total 53 drought events occurred in 1901-2020

No.	Initial time	Terminal time	Duration (month)	Severity	Affected area ratio
1	1904.5	1904.11	7	9.06	43%
2	1906.3	1906.6	4	3.52	35%
3	1909.5	1911.7	27	53.66	79%
4	1912.8	1912.11	4	4.29	40%
5	1916.5	1916.5	1	1.37	43%
6	1916.8	1920.7	48	109.00	61%
7	1922.4	1922.5	2	2.79	51%
8	1922.9	1923.2	6	8.01	30%
9	1923.5	1927.7	51	94.50	57%
10	1929.11	1929.11	1	1.02	30%
11	1931.9	1931.11	3	3.50	32%
12	1932.1	1934.3	27	43.63	55%
13	1935.5	1935.5	1	1.05	55%
14	1935.8	1937.4	21	33.99	61%
15	1937.7	1940.1	31	58.21	61%
16	1940.3	1940.9	7	11.00	54%
17	1943.5	1943.5	1	1.26	43%
18	1943.7	1943.7	1	1.65	50%
19	1943.9	1945.10	26	62.43	69%
20	1948.8	1948.8	1	1.42	52%
21	1948.10	1948.11	2	2.79	52%
22	1950.9	1951.9	13	21.39	54%
23	1955.6	1956.1	8	11.58	67%
24	1956.7	1957.6	12	24.72	78%
25	1961.5	1961.5	1	1.10	68%
26	1962.2	1963.7	18	28.13	64%
27	1965.1	1965.7	7	11.53	49%
28	1967.11	1968.10	12	22.09	68%
29	1970.3	1970.7	5	7.47	49%
30	1971.9	1971.11	3	3.59	32%
31	1973.10	1977.1	40	111.67	95%
32	1977.3	1978.4	14	23.41	65%
33	1978.7	1978.10	4	7.00	69%
34	1980.8	1981.3	8	9.87	52%
35	1981.12	1983.11	24	39.64	55%
36	1984.1	1984.2	2	2.21	45%
37	1984.5	1985.1	9	13.92	57%
38	1985.6	1985.7	2	2.56	64%

39	1986.7	1986.7	1	1.50	44%
40	1990.6	1990.6	1	1.37	36%
41	1991.4	1992.7	16	30.12	78%
42	1992.11	1992.11	1	1.38	58%
43	1994.7	1994.7	1	1.18	95%
44	1994.9	1994.10	2	2.49	85%
45	1995.2	1998.2	37	76.65	71%
46	1999.5	1999.5	1	1.08	52%
47	2000.4	2000.8	5	6.81	51%
48	2001.6	2001.6	1	1.06	23%
49	2005.10	2006.9	12	21.20	69%
50	2007.8	2009.7	24	56.45	83%
51	2012.5	2012.5	1	1.01	48%
52	2013.11	2013.11	1	1.09	52%
53	2014.6	2014.8	3	4.10	39%

Table 2 Correlation test of drought variable pairs

pairs	correlation coefficient						significant
	Kendall τ	P-value	Spearman ρ	P-value	Person γ	P-value	at 5%
D-S	0.926	0.000	0.985	0.000	0.979	0.000	Yes
D-A	0.359	0.002	0.504	0.004	0.438	0.001	Yes
S-A	0.369	0.002	0.523	0.000	0.482	0.000	Yes

Table 3 Statistical tests of marginal distribution fitting

distributions	AICc		
	duration	severity	area
Gamma	357.8	416.5	357.8
Generalized extreme value	251.5	356.4	251.6
Inverse Gaussian	345.4	400.0	345.4
Log logistic	356.4	394.7	356.5
Lognormal	351.1	412.1	351.2
Weibull	356.6	413.8	356.6

Table 4 Statistical test and parameter estimation of drought variable pairs

pairs	distributions	AIC	BIC	parameter	95% range	RMSE	NSE
D-S	**Clayton**	**-239.60**	**-237.63**	28.83	[8.36, 34.51]	0.74	0.86
	Frank	-239.10	-237.13	34.57	[12.55, 34.99]	0.75	0.87
	Gumbel	-239.55	-237.58	6.36	[3.75, 34.67]	0.74	0.87
	Joe	-238.72	-236.75	34.62	[7.02, 34.64]	0.75	0.87
D-A	**Clayton**	**-231.82**	**-229.84**	1.25	[0.50, 6.17]	0.80	0.81
	Frank	-229.59	-227.62	3.25	[1.08, 11.90]	0.82	0.80
	Gumbel	-228.99	-227.02	1.42	[1.14, 7.47]	0.83	0.80
	Joe	-227.67	-225.69	1.55	[1.19, 10.37]	0.83	0.79
S-A	**Clayton**	**-333.83**	**-331.86**	0.90	[0.59, 1.39]	0.31	0.98
	Frank	-331.34	-329.37	2.87	[1.92, 3.90]	0.31	0.97
	Gumbel	-327.94	-325.97	1.39	[1.23, 1.62]	0.32	0.97
	Joe	-322.05	-320.08	1.57	[1.36, 1.796]	0.34	0.97

Table 5 Comparison of univariate and bivariate return periods of drought variables

Return period	T=100	T=50	T=20	T=10	T=5
P-level	99%	98%	95%	90%	80%
Duration	105	73	43	27	15
Severity	246	163	87	51	26
Area / %	96%	91%	84%	77%	69%
$C(d, s)$	0.98	0.97	0.93	0.88	0.78
$C(d, a)$	0.98	0.96	0.91	0.82	0.67
$C(s, a)$	0.98	0.96	0.90	0.82	0.66
P(D>d, S>s)	0.002	0.008	0.03	0.08	0.18
P(D>d, A>a)	<0.001	<0.001	0.01	0.02	0.07
P(S>s, A>a)	<0.001	<0.001	<0.001	0.02	0.06
$T_{\text{and}}(D\text{-}S)$	500	125	33	13	6
$T_{\text{and}}(D\text{-}A)$	4485	1138	200	50	14
$T_{\text{and}}(S\text{-}A)$	5308	1339	220	58	16
$T_{\text{or}}(D\text{-}S)$	56	32	15	9	5
$T_{\text{or}}(D\text{-}A)$	51	26	11	6	4
$T_{\text{or}}(S\text{-}A)$	50	25	10	5	3
Risk(D-S)	18.1%	33.1%	46.0%	55.1%	59.8%
Risk(D-A)	2.2%	4.3%	9.5%	18.3%	31.0%
Risk(S-A)	1.9%	3.7%	8.7%	16.0%	27.6%

Note: duration (month), T (year)

List of Figure Captions

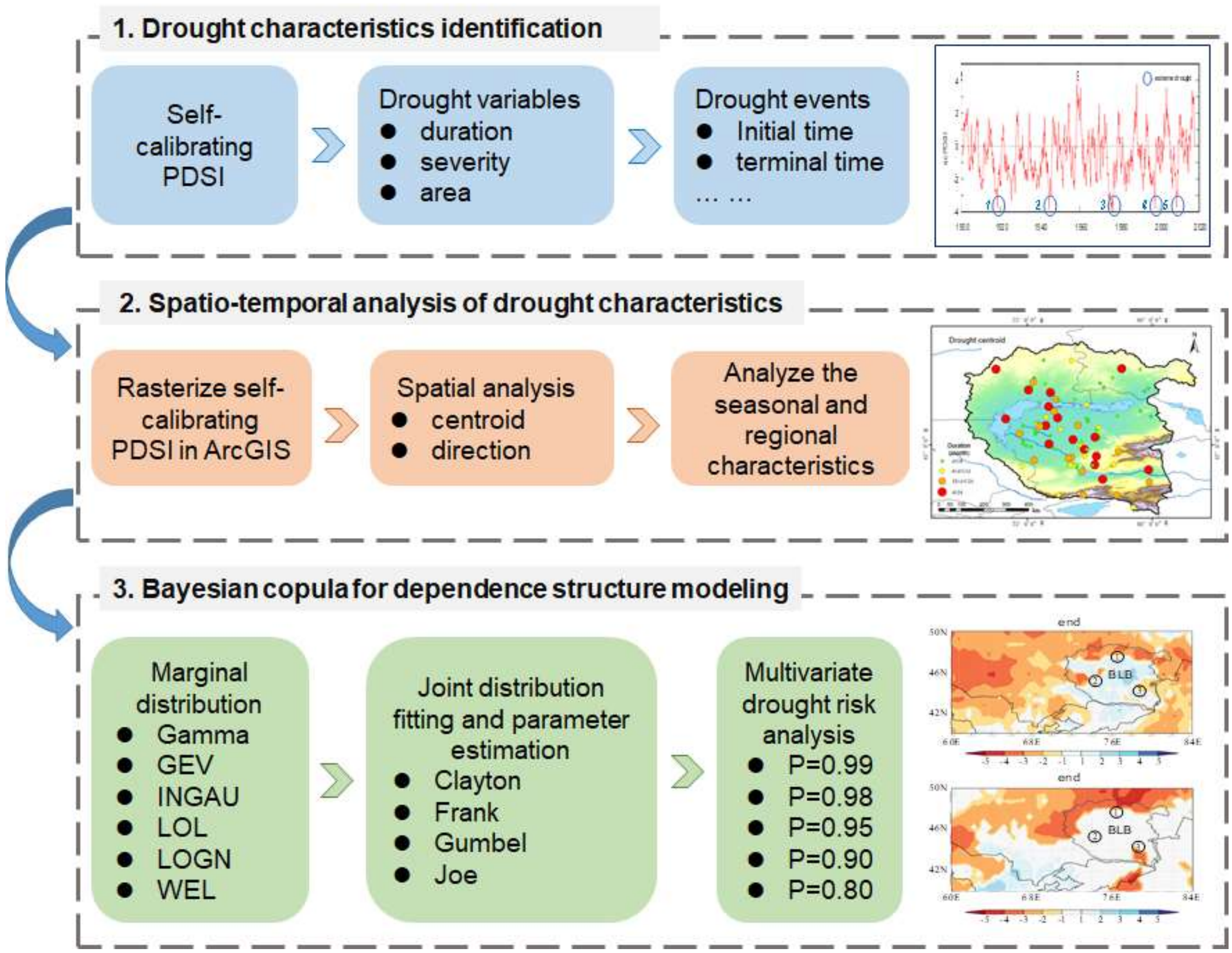
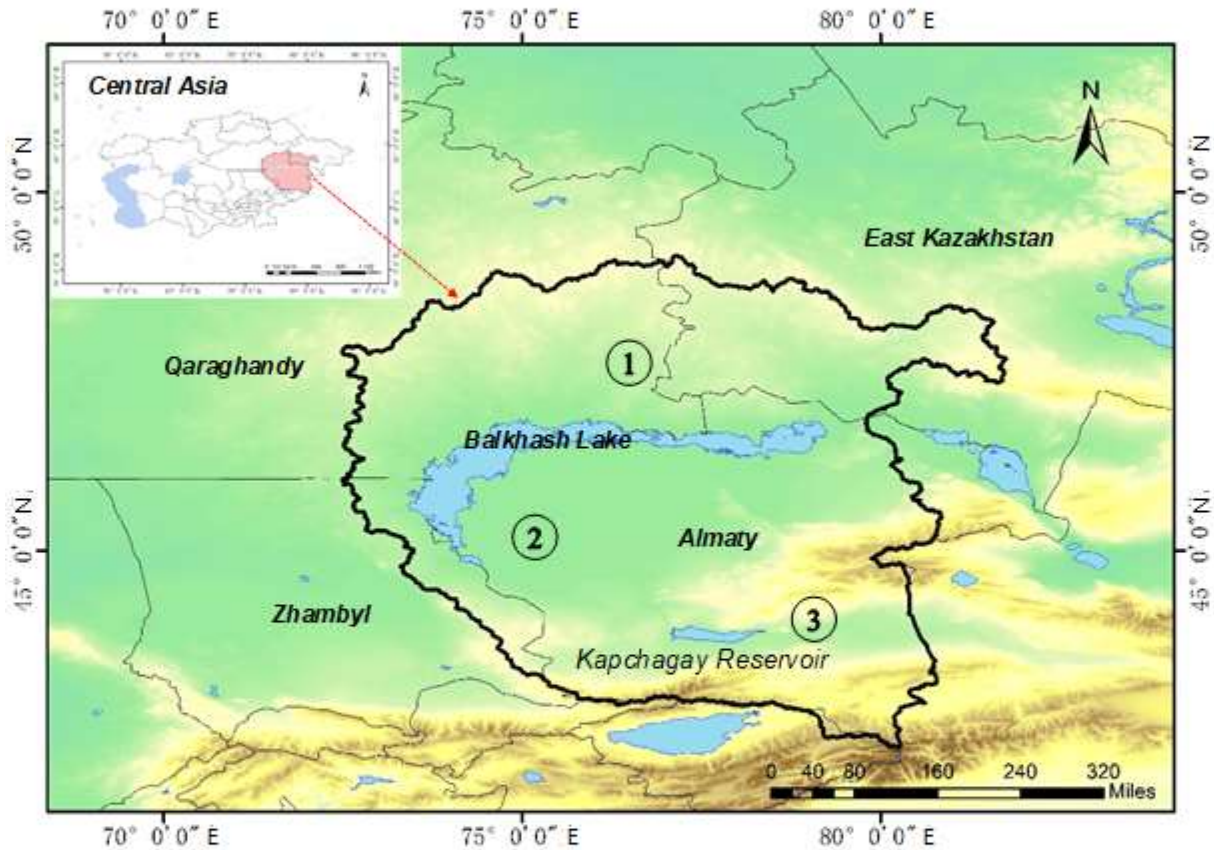


Figure 1 Framework of SBC method



① arid grassland in north Balkhash Lake; ② Ili River delta and alluvial plain; ③ plateau desert

Figure 2 Topographic characteristics of Balkhash Lake Basin

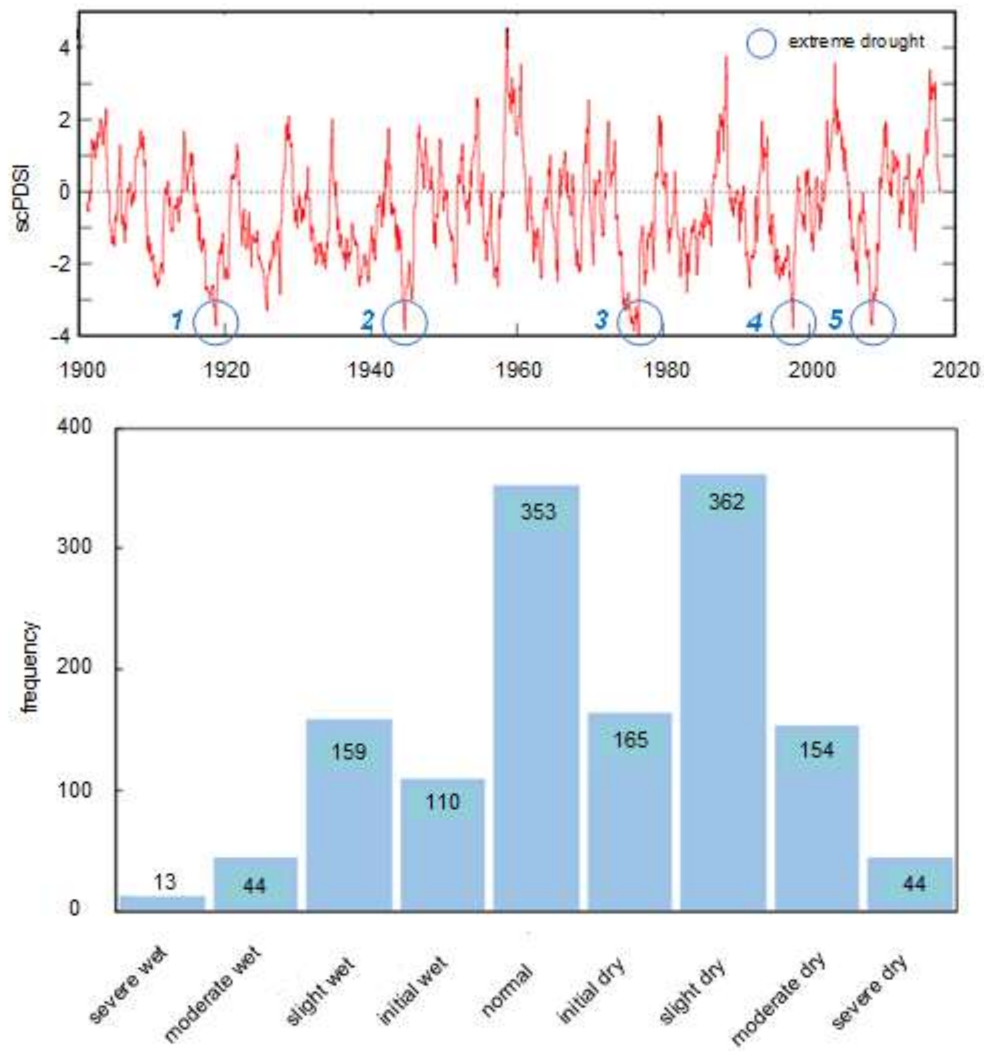


Figure 3 Temporal evolution of scPDSI, and frequency of dry and wet months in 1901-2020

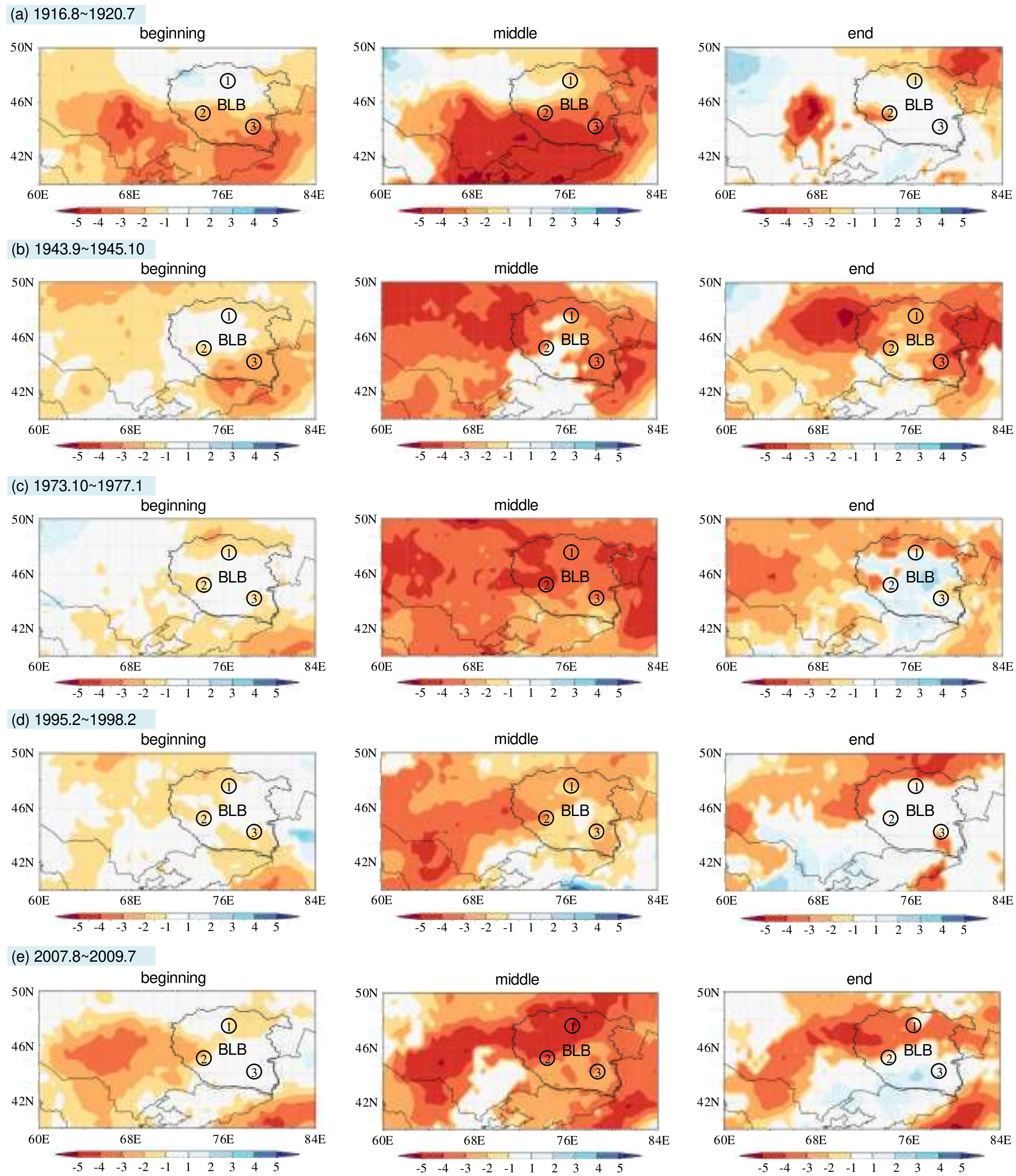


Figure 4 Spatial distribution of scPDSI of five typical severe drought events

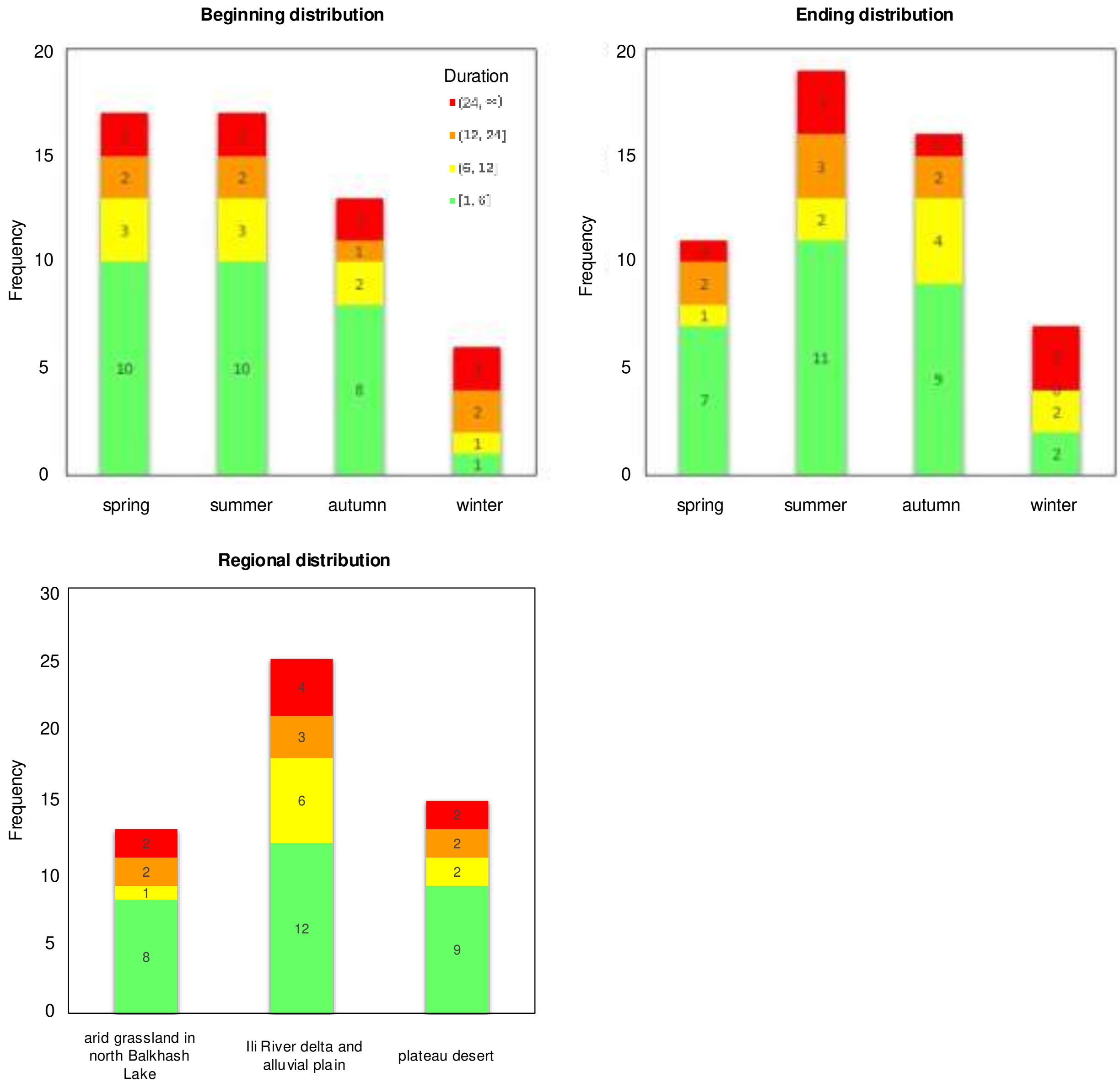


Figure 5 Seasonal and regional distributions of drought events

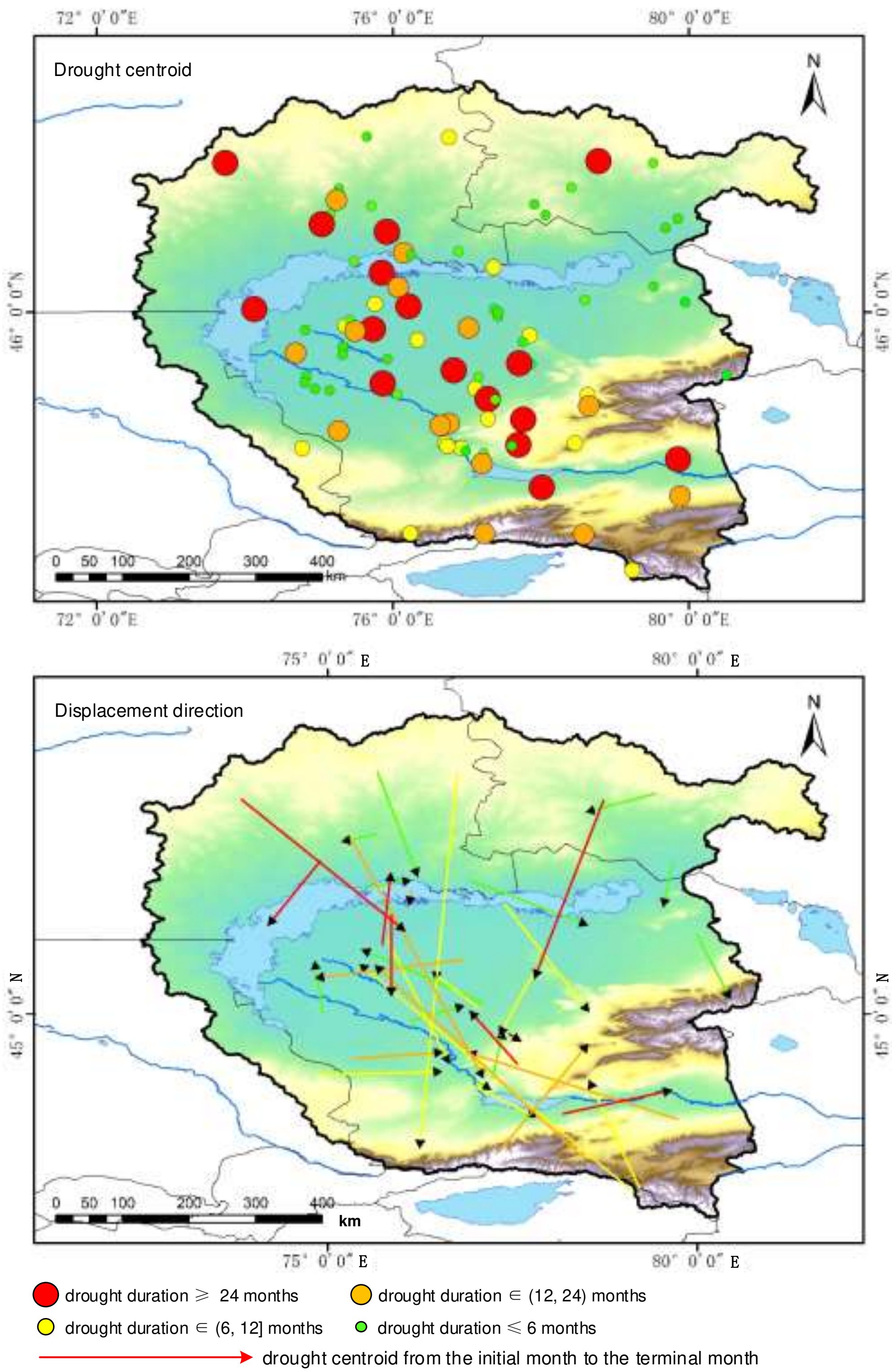


Figure 6 Centroids and displacement directions of 53 drought events

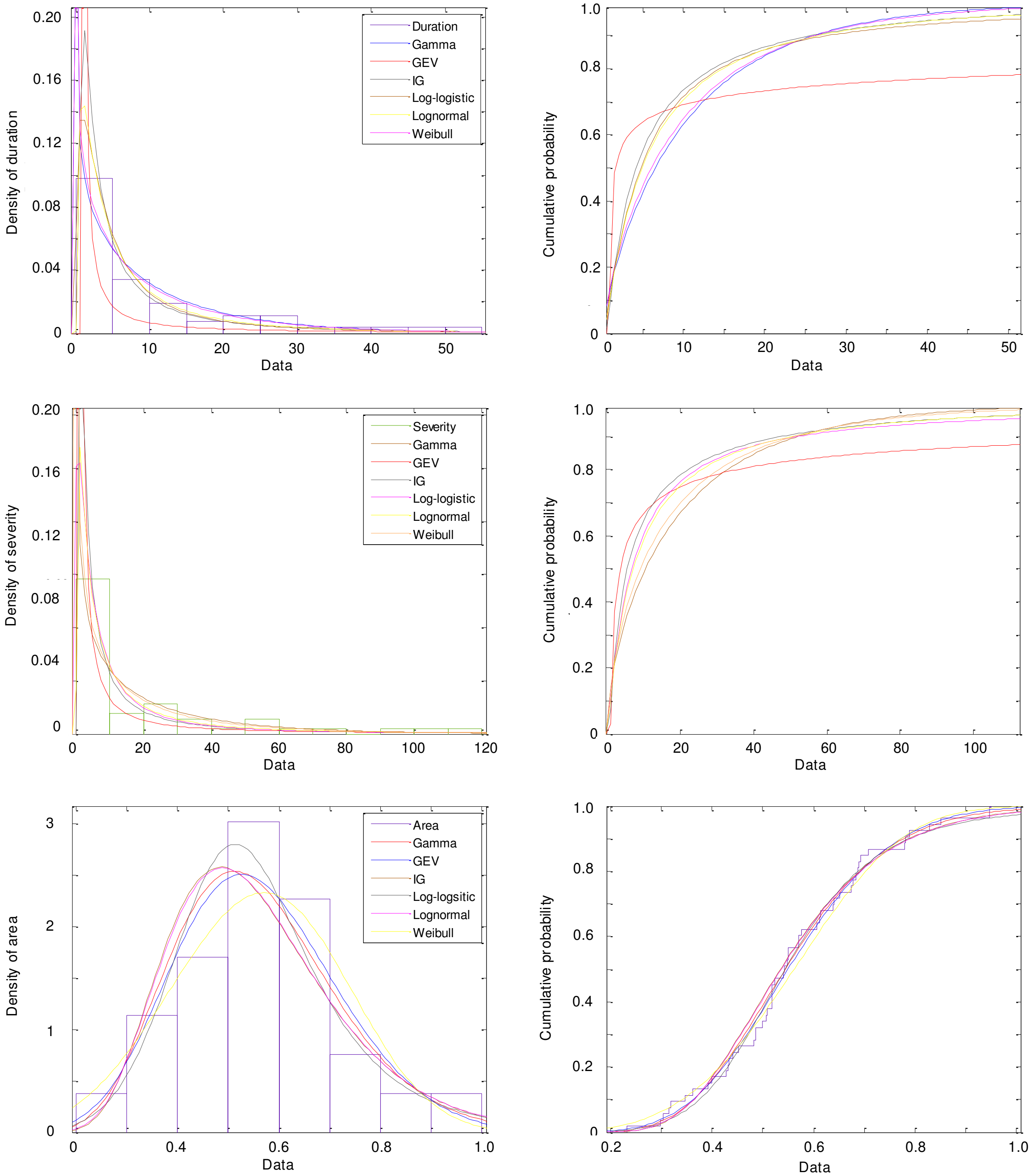


Figure 7 Marginal distribution fitting of drought duration, severity and area

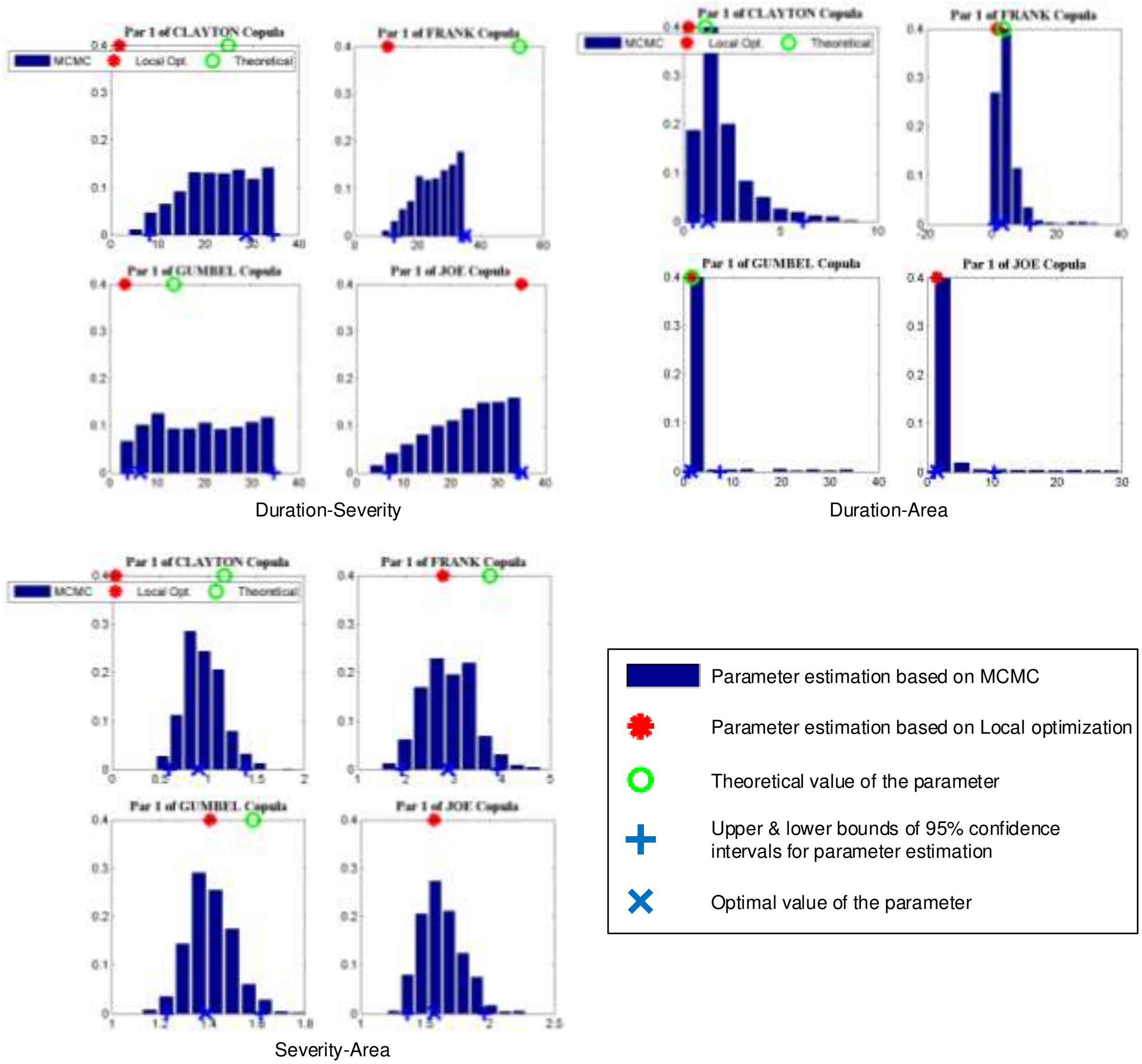


Figure 8 Parameter estimation of four copula functions

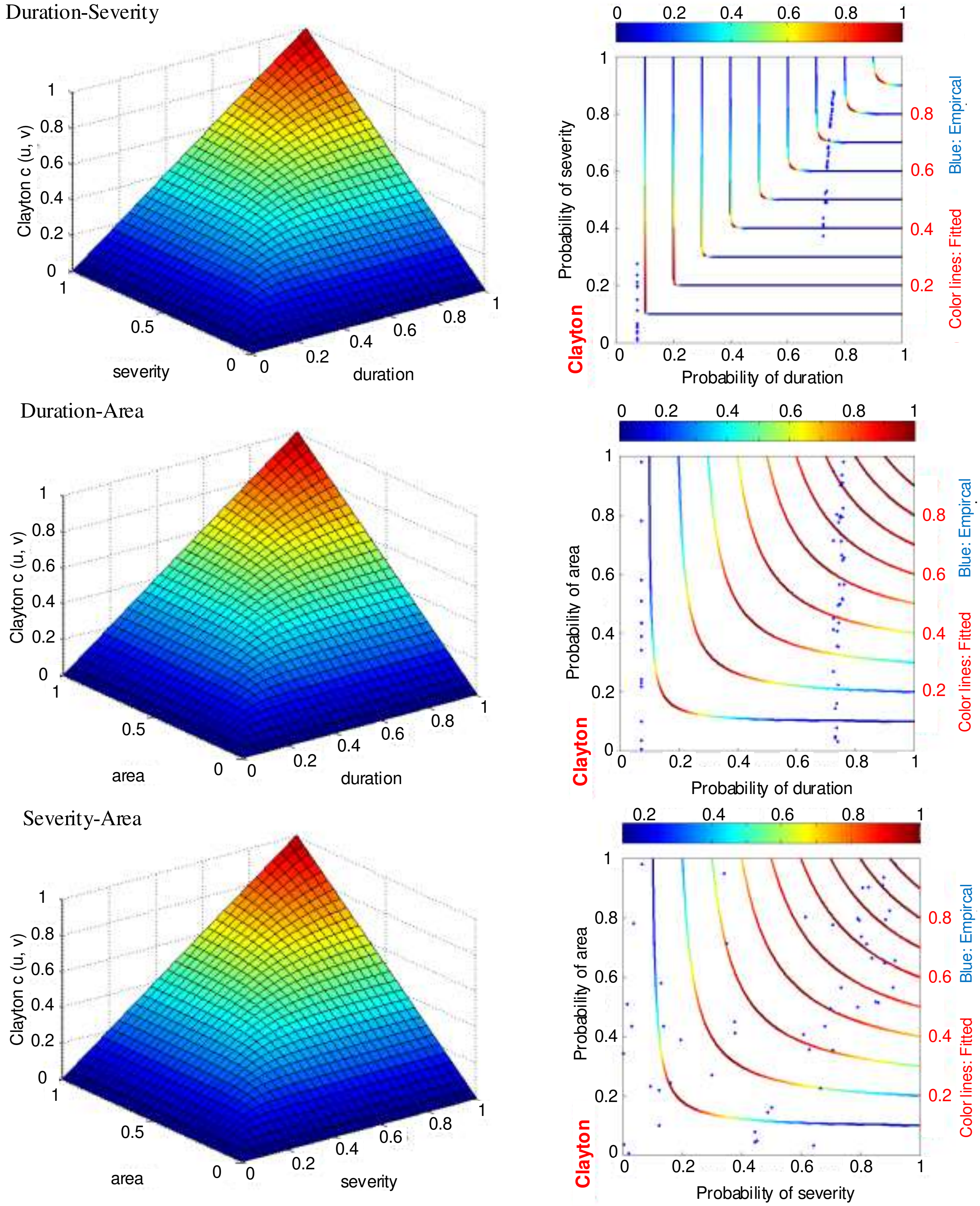


Figure 9 Joint probability and contour plot of drought variable pairs