Pak.j.stat.oper.res. Vol.20 No.3 2024 pp 553-586 DOI: http://dx.doi.org/10.18187/pjsor.v20i3.4609

Pakistan Journal of Statistics and Operation Research

Dependence of Drought Characteristics: Parametric and Non-parametric Copula Approach

Ozan Evkaya[1,*], Honygi Lu[2]

*Corresponding author

1. School of Mathematics, University of Edinburgh, United Kingdom, oevkaya@ed.ac.uk
2. NatWest Bank Groups, Edinburgh, United Kingdom, hongyi.luu@gmail.com

Abstract

Drought, which has harmful impacts both environmentally and economically, is one of the most devastating natural phenomena. In order to better understand and monitor the effects of drought, various methods have been developed in recent decades to quantify drought characteristics, with a primary focus on univariate drought indices. Understanding drought characteristics is essential for conducting an in-depth examination of its impacts on a specific area. That specifically requires examining the specific characteristics of drought such as its duration, or severity and including the association between these characteristics. In that respect, it is crucial to model the joint behavior of these drought characteristics. This study endeavors to investigate univariate and bivariate drought indices using both parametric and non-parametric copula techniques. For that purpose, drought characteristics, such as duration, severity, mean intensity and peak intensity are analysed relying on different drought indices. The dependence among the main characteristics is evaluated and corresponding bivariate return period calculations are investigated. The data set used in this study is retrieved from monthly meteorological observations collected at five different Stations in Konya, located in the Central Anatolia Region of Türkiye. As we explored, parametric or non-parametric copula usage may differ slightly based on the extreme drought cases. In that respect, the findings of the study examines the suitability of both parametric and nonparametric dependence setting for a specific region by testing across different weather stations. Besides, that comparative study indicates the importance of using multiple drought indices for different geographical reasons for extreme dry periods.

Key Words: Drought Analysis, Parametric Copulas, Non-parametric Copulas, Drought Characteristics, Exceedance Probability

Mathematical Subject Classification: 62-07, 62H20, 62P12.

1. Background

Drought is an inevitable and long-lasting natural phenomenon that often leads to environmental, societal and economic damages. It mainly arises from an insufficient amount of precipitation compared to the normal level, resulting in significant economic and environmental impacts over extended time and spatial scales. In recent years, many countries have encountered dry periods, leading to tremendous worldwide economic losses. To illustrate, droughts have had large impacts on the population, often leading to significant damages (50% of the mortality due to natural hazards). Meanwhile, nearly 7% of economic losses have been associated with the occurrence of dry periods worldwide (Núñez et al., 2011). In that respect, monitoring and understanding drought require increased attention and it is vital to design efficient drought management plans to reduce its harmful impacts.

Drought has been increasing in Türkiye in recent years, and it is predicted that there will be a significant lack of precipitation, especially in the Southeastern and Central Anatolia regions and the Aegean region. Global warming, combined with the current lack of precipitation, constitutes an important obstacle to meeting the water requirements for drinking and agriculture. Drought is occurring as a result of rising temperatures and insufficient precipitation, observed in many basins of Türkiye. The Konya Closed Basin among others, located on the southwestern edge of the Central Anatolian Plateau, plays a crucial role in Türkiye's agricultural output, encompassing 10% of the entire farmland (mainly relying on the rain-fed agriculture (Evkaya et al., 2019)). Geographically, Konya Closed Basin, which encompasses an area of 53,850 square kilometers, is situated in the Central Anatolia region of Türkiye and accounts for 7% of the country's total area. However, because of its semi-arid climate, it is one of the regions most affected by drought, and a particularly vulnerable area to drought due to its meteorological characteristics. This region is mainly classified as having a cold, semi-arid climate with low precipitation levels, mainly in winter and spring. Therefore, it is crucial to provide a detailed analysis of the performance of drought indexes in this particular region.

As a widely applied monitoring tool, univariate drought indices aim at capturing drought characteristics but may not be sufficient for quantifying risk. Generally, the benefits are limited since they are based on just one climatic variable. In the recent decades, multivariate setting is applied to construct more reliable drought indices. More clearly, the joint behavior of weather parameters are considered to derive a multivariate standardized index. To develop flexible multivariate distribution functions, copula functions played an important role recently. Mainly, this framework allows to model the dependence structure of the massive multivariate data where the joint dependence is of great interest and the usual normality assumption is violated. Copulas successfully cover both negative and positive dependence, including the case of independence and incorporating asymmetric dependence structure as well. Copulas and their multivariate extension (vine copulas) have widespread usage in various areas that require multivariate modelling, including quantitative finance, medicine, geostatistics, and hydrology. In particular, copula-oriented hydrological studies, including the use of copulas for drought analysis, have been increasingly employed in recent years. Specifically, it is vital to consider the interactions between weather variables in the design of any standardised drought index.

For many decades, researchers have proposed several indices and methods to detect and monitor droughts by different variables such as precipitation, soil moisture, stream flow, and evapotranspiration (Heim et al., 2000). One of the oldest indicators is called the Palmer Drought Severity Index (PDSI), having various recent adaptations. More widely applied ones are the Standardised Precipitation Index (SPI) (McKee et al., 1993) and Standardised Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010) by following a similar methodology. SPI is only based on the long-term precipitation amount for the certain period of time, whereas SPEI utilises precipitation and evapotranspiration data to determine dry periods. Mainly, SPEI uses the idea of the climatic water balance (precipitation minus potential evapotranspiration) to characterise dry or wet circumstances. As a common similarity, both indices are in the univariate form regarding its construction but it is observed that SPEI is more useful for incorporating the effect of climate change by finding trends in time series data (Erhardt and Czado, 2018). Recently, different versions of such univariate standardised indices have been introduce; (including but not limited to) Reconnaissance Drought Index (RDI) (Tsakiris and Vangelis, 2005), Streamflow Drought Index (SDI) (Nalbantis, 2008), Standardised Runoff Index (SRI) (Shukla and Wood, 2008), and Regional Drought Area Index (RDSI) (Fleig et al., 2011) and more. There are many other alternatives in the growing literature for drought monitoring, and a recent comprehensive review of various indices can be found in (Alahacoon and Edirisinghe, 2022). In a nutshell, the concept of standardised drought indices is generally summarised with the Standardised Drought Analysis Toolbox (SDAT), which includes the possibility of extending the idea of SPI to bivariate data using non-parametric estimates (Farahmand and AghaKouchak, 2015). Afterwards, this process is generalized and expanded for higher dimensional data so that it is possible to construct any standardized index by flexibly embedding the relationship between weather variables (Erhardt and Czado, 2018). These new tools have recently been applied to specific regions and they have mainly exploited the benefits of pair copulas for constructing multivariate drought index.

Another drought-related concept is called the return period with the corresponding probability of exceedance. Once a standardized index is created for a specific weather gauge, drought characteristics can be described using certain thresholds in terms of the run theory. Mainly, the properties of any drought event can be summarized using the characteristics such as duration, severity, peak intensity and mean intensity over a certain period of time. The relationship between these characteristics is worth exploring to gain a deeper understanding of drought based on the available data and fitted models. Specifically, different return periods and related joint exceedance probabilities are crucial quantities that decision-makers can utilize. In analogy to the use of copulas for multivariate drought indices, one can benefit from copulas to capture the dependence structure between drought characteristics.

In addition to constructing new joint indices, one can benefit from copulas to analyze drought properties more efficiently. The main motivation for this approach relies on the fact that separate univariate frequency analysis are not satisfactory (Serinaldi et al., 2009). Since one weather variable is not sufficient enough to conduct a comprehensive evaluation of drought because of the complexity of its nature (Shiau et al., 2007), considering the joint distribution of drought characteristics is more reliable in general (Mishra and Singh, 2010). As drought can be defined by its characteristics (Stambaugh et al., 2011), there have been many studies that show how copulas and the dependencies between drought features work well together. Copulas have been introduced in rainfall studies as they offer a more flexible approach that allows different types of marginal distributions to coexist and combine to generate a multivariate distribution. (Kao and Govindaraju, 2007, 2008, 2010, Serinaldi et al., 2009). Another advantage of using copula is that the method helps to relax the independence assumption, inappropriate for designing hydrological variables (De Michele and Salvadori, 2003) (Genest and Favre, 2007) (Zhang and Singh, 2007). By borrowing the same approach, it is possible to consider the use of parametric and non-parametric copulas in order to capture the dependence between the drought characteristics.

As one of the pioneering works, (Shiau, 2006) and (Shiau and Modarres, 2009) estimated the Standard Precipitation Index by fitting rainfall intensity based on the Gamma distribution. Afterwards, Abdul Rauf and Zeephongsekul investigated rainfall severity and duration patterns in the Australian state of Victoria using parametric distributions, (Abdul Rauf and Zeephongsekul, 2014a,b). However, as expected, this framework does not work well for every weather data and has a poor fit near the tails of the distribution (Haghighat jou et al., 2013). Recent related works mainly follow the implementation of non-parametric marginals, and even non-parametric copulas for the analysis of drought characteristics. To begin with, the SDAT mentioned above relies on a non-parametric framework that can be applied to different climatic variables including precipitation, soil moisture and relative humidity, without assuming typical parametric distributions. (Farahmand and AghaKouchak, 2015). In a similar vein, the multivariate tool introduced by Erhardt and Czado allows non-parametric distribution estimates for the climatic variables for the index construction (Erhardt and Czado, 2018). Generally, previous works have focused on the use of the parametric type of copulas to explore the dependence structure between pre-selected drought characteristics. For instance, conditional probabilities and multivariate return period calculations have been investigated based on the joint probabilities with copulas. Furthermore, copula functions have been said to be useful and reliable methods for probabilistic drought analysis (Mirabbasi et al., 2013) (Hesami Afshar et al., 2016). On the other side, non-parametric copula approaches have also been examined for the analysis of rainfall and severity and duration using six stations in the state of Victoria, Australia (Abdul Rauf and Zeephongsekul, 2014a). Their main findings indicate that non-parametric approaches may give better results than a pure parametric setting. In another study, the drought duration, severity and peak intensity were analysed based on only parametric copula with parametric marginals for the Türkiye case (Evkaya et al., 2019). Another work combined non-parametric margins with parametric copulas over the same study area, including similar weather stations (Vazifehkhah and Kahya, 2019).

In light of the previously mentioned works, this study aims to contribute to the analysis of drought characteristics by considering the use of both parametric and non-parametric copulas. Besides, it is aimed to use vine copula models to create a drought index and assess drought risk. For that purpose, multivariate indices are generated and the corresponding dependence emerged in drought characteristics are examined by copulas. The main objective is to examine the benefits of parametric and non-parametric copula approaches in modelling the joint distribution of drought characteristics. For that purpose, widely used drought characteristics such as duration, severity, mean intensity and peak intensity are derived using various drought indices. Mainly, we explored four different drought indices, 2 of them are univariate (classical Standardised Precipitation Index (SPI) and Standardised Precipitation and Evapotranspiration Index (SPEI) relying on univariate weather indicator), 2 joint indices relying on bivariate and trivariate settings as follows; (i) SBI(PREC-PET): bivariate case using both precipitation and evapotranspiration and (ii) STI(PREC-PET-RH): trivariate case using precipitation, evapotranspiration and relative humidity, via pair copulas. Using the PREC-PET or PREC-PET-RH variables together is reasonable for joint drought index construction, but requires careful consideration of their meaning. Since the aim is to explore the drought characteristics relationship, combination of these weather variables in the drought index adds complexity but also depth to the analysis. Higher PET or lower RH typically corresponds to more severe drought conditions, whenever PET is lower or RH is higher, the conditions are more normal compared to dryness.

For the construction of drought indices, available toolkit are exploited by considering certain parameter selections. Afterwards, the four main drought characteristics (duration, severity, mean intensity and peak intensity) are extracted with the inter-arrival time. As a next step, parametric approach is considered for the marginal distributions of drought characteristics. In order to model the dependence structure between drought characteristics, both parametric and non-parametric copula approaches are investigated over the parametrically modelled marginals. Finally, based on the fitted dependence structure of the characteristics, corresponding bivariate exceedance probabilities are derived for the specific drought events. As a case study, monthly data spanning 720 months between 1950 and 2010 from weather stations in Türkiye located within Central Anatolia Region is explored for the application of the drought analysis. After summarising the main results for the main weather station as Aksehir, other stations; Cihanbeyli, Karapinar, Cumra and Yunak are added from the same region. The rest of the paper is designed in 5 consecutive sections, Firstly, related research studies are discussed briefly in the literature review part, then copula theory simply introduced. Afterwards, modeling steps and drought index calculation part is described. In the findings, Türkiye case study is explored in detail. Finally, the main findings and limitations of the work is summarized including potential future directions.

2. Copula Theory

Previously, Sklar introduced copulas that combine individual distribution functions flexibly to generate a multivariate distribution function (Sklar, 1959). The construction of multivariate distribution, therefore, is simplified to examining the relationship between the random variables when the marginal distributions are defined. This study employs the copula method for drought analysis in two different ways; both index design and dependence of characteristics using two forms parametric (P) and non-parametric (NP).

Definition 2.1. *For $p \geq 2$, a p-dimensional copula is a multivariate distribution having marginals are all uniform over $\mathbb{I} = [0, 1]$. For a p-dimensional vector $U = (U_1, U_2, \ldots, U_p)$ on the unit hyper-cube, a copula function C is defined as:*

$$C(u_1, u_2, \ldots, u_p) = \Pr(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_p \leq u_p). \tag{1}$$

Theorem 2.1 (Sklar's Theorem)**.** *Let F be a continuous p-dimensional distribution function with univariate margins F_1, F_2, ..., F_p. Let A_j denote the range of F_j and $A_j = [-\infty, \infty]$ where $j = 1, 2, \ldots, p$. Then, there exists a copula function C such that for all $(x_1, x_2, \ldots, x_p) \in [-\infty, \infty]$ it satisfies the following:*

$$F(x_1, x_2, \ldots, x_p) = C(F_1(x_1), F_2(x_2), \ldots, F_p(x_p)). \tag{2}$$

The function C, which connects the joint distribution F with its marginals, is the copula and this is determined uniquely whenever all F_1, F_2, ..., and F_p are continuous marginal distributions.

Knowing the distribution function for many random variables is one of the most efficient ways of analyzing the association between multivariate outcomes from a probabilistic viewpoint. Copulas or copula functions are extremely promising for this goal since they focus on the joint behavior of random variables without imposing strict constraints on their marginal distributions. In that respect, there exists a certain relationship between the dependence measures and copula parameters for any multivariate random vector. Among the alternatives, the rank-based dependence measures, such as Spearman's ρ and Kendall's τ, are widely considered for the copula functions (Czado, 2019).

2.1. Parametric Copulas

Sklar's Theorem has an essential reverse implication that is often used when building multivariate models by analysing the individual behaviour of the components of a random vector and their dependence properties as described by copulas.

Theorem 2.2. *If $F_1, F_2, \ldots, F_p$ are univariate distribution functions, and if C is any p-copula, then the function $F : \mathbb{R}^p \to \mathbb{I}$ defined by Equation 2 is a p-dimensional distribution function with margins $F_1, F_2, \ldots, F_p$. In other words, a copula C can be derived from any p-variate distribution function F using Equation 2. When each F_i is continuous for every $i \in 1, 2, \ldots, p$, C can be computed using the formula:*

$$C(u_1, u_2, \ldots, u_p) = F(F^{-1}(u_1), F^{-1}(u_2), \ldots, F^{-1}(u_p)), \tag{3}$$

where F_i^{-1} is the pseudo-inverse of F_i given by $F_i^{-1}(s) = \inf \{t | F_i(t) \geq s\}$. Thus, copulas transform the random variables $(X_1, X_2, \ldots, X_p)$ into to the set of random variables $(U_1, U_2, \ldots, U_p) = (F_1(X_1), F_2(X_2), \ldots, F_p(X_p))$, which have uniform margins on $\mathbb{I}$ and preserve the dependence among the components.

Alternatively, the above theorem implies that copulas pair various univariate density distributions to generate a p-variate density distribution. This enables copulas to be used for modelling scenarios where each marginal requires a different distribution, which offers a valid alternative to classical multivariate density distributions such as Gaussian, Pareto, Gamma, among others. The Equation 3 provides a direct way to derive the copula density, $c(u_1, \ldots, u_p) = \partial^p C(u_1, \ldots, u_p) / (\partial u_1 \cdots \partial u_p)$. Then, we can express copula density function $c(u_1, \ldots, u_p)$ as follows:

$$c(u_1, \ldots, u_p) = \frac{f\left(F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p)\right)}{\prod_{j=1}^{p} f_i\left(F_j^{-1}(u_j)\right)}. \tag{4}$$

where $f, f_1, \ldots, f_p$ represents the densities corresponding to $F, F_1, \ldots F_p$, respectively. The above equation permits the use of arbitrary parametric distribution functions F to construct parametric copula families. Mainly, the parametric families can be classified into elliptical and Archimedean types based on their construction mechanism.

In particular, *elliptical copulas* are widely used families in statistical modelling, finance, and insurance to study multivariate dependence. The Gaussian and Student-t copulas are the most notable examples of elliptical copulas, which can be constructed with arbitrary parametric distribution functions F, allowing for flexibility in their use. By nature, the bivariate Student t-copula function exhibits both upper and lower tail dependence unlike the Gaussian one.

Gaussian (Normal) copula: Let Φ_Γ, be the cumulative density function of a p-dimensional vector following a multivariate normal distribution with zero means, unit variances and correlation matrix $\Gamma \in [-1, 1]^{p \times p}$. Also, let Φ be the univariate standard normal cumulative distribution function. The following provides the expression for the Gaussian copula with correlation matrix Γ: $C_\Gamma^{\text{Gauss}}(u_1, \ldots, u_p) = \Phi_\Gamma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_p))$ where Φ^{-1} denotes the inverse of the univariate standard normal cumulative distribution function, Φ.

Student t-copula: Let $t_{\nu,\Gamma}$, be the cumulative density function of a p-dimensional vector following a multivariate t-distribution with zero mean and scale parameter matrix $\Gamma \in [-1, 1]^{p \times p}$ and degrees-of-freedom parameter ν. Also, let t_ν be the univariate cumulative density function of a t-distribution with degrees-of-freedom parameter ν. The following provides the expression for the Student-t or t-copula with parameters ν, Γ: $C_{\nu,\Gamma}^{t}(u_1, \ldots, u_p) = t_{\nu,\Gamma}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_p))$

Another essential collection of parametric copula families is known as *Archimedean copulas*. To begin, let us introduce a continuous, strictly monotonic decreasing, convex $\phi : [0, 1] \to [0, \infty]$ constructive function with $\phi(1) = 0$, called as *Archimedean generator* (Nelsen, 2007). The widely known families, relying on the Archimedean generator, are simply exemplified as follows:

Frank copula: $C_\theta^{\text{Frank}}(u_1, \ldots, u_p) = -\frac{1}{\theta}\left[\log\left(1 + \frac{\prod_{i=1}^{p}(\exp(-\theta u_i) - 1)}{(\exp(-\theta) - 1)^{p-1}}\right)\right]$

Gumbel copula: $C_\theta^{\text{Gumbel}}(u_1, \ldots, u_p) = \exp\left[-\left(\sum_{i=1}^{p}(-\log(u_i))^\theta\right)^{1/\theta}\right]$

Clayton copula: $C_\theta^{\text{Clayton}}(u_1, \ldots, u_p) = \left(\sum_{i=1}^{p}(u_i)^{-\theta} - p + 1\right)^{-1/\theta}$

There are also two parameter Archimedean copulas, such as BB1, BB7 or BB8 families. To exemplify, elliptical copulas and the Frank copula are preferable to examine the symmetric dependence structures. On the other hand, Clayton and Gumbel copulas are useful to identify different tail dependencies at lower and upper quantiles, respectively. It should be noted that certain parametric families described here have restrictions in their parameter space. For example, some may only be suitable for positive dependence, and some may only accommodate either upper or lower tail-dependence. *Upper tail-dependence* occurs when the likelihood of other variables taking extreme values increases as one variable assumes an extreme value. To overcome these limitations and provide more flexibility, rotated copula families can be employed. The density functions of rotated copulas can be derived from the copula density, expressed as; (i) 90 degrees rotation: $c_{90}(u, v) := c(1 - u, v)$, (ii) 180 degrees rotation: $c_{180}(u, v) := c(1 - u, 1 - v)$ and (iii) 270 degrees rotation: $c_{270}(u, v) := c(u, 1 - v)$ (Nagler, 2014). Readers are referred to (Czado, 2019) for a comprehensive discussion on different copula families.

To keep things simple and relevant to the number of variables, all the possible parametric families are tested automatically by benefiting from the related R packages. The main families and their rotated versions are all considered using `VineCopula` CRAN and `SIndices` open-source R packages by adding new helper functions.

2.2. Non-parametric Copulas

Previously, univariate Kernel density estimation (KDE) was introduced by Rosenblatt (Rosenblatt, 1956) and Parzen (Parzen, 1962). Wand and Jones (Wand and Jones, 1993) elaborately discussed a natural extension to the multivariate case. For a more extensive introduction, interested readers can refer to (Simonoff, 2012). In the non-parametric estimation literature, as a common practice, independently and identically distributed ($i.i.d.$) copies are considered rather than $i.i.d.$ observations (Nagler, 2014). Therefore, we let $(X, Y) \in \mathbb{R}^2$ be a random vector with density f and assume we are given $i.i.d.$ copies $(X_i, Y_i)_{i=1,...,n}$. Recall that the density can be defined as

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \lim_{\epsilon_x \to 0} \lim_{\epsilon_y \to 0} \frac{F(x+\epsilon_x, y+\epsilon_y) - F(x-\epsilon_x, y-\epsilon_y)}{4\epsilon_x \epsilon_y}, \tag{5}$$

where F is the CDF corresponds to f. A natural estimator of the density could be obtained by fixing values for ϵ_x, ϵ_y in (5) and benefiting from the empirical $CDF(\widehat{F}_n)$, as an estimator for F. For simplicity, take $\epsilon_x = \epsilon_y = b$, for some small $b > 0$. The resulting estimator, using the kernel function, can be written as

$$\widehat{f}_n(x,y) = \frac{1}{nb^2} \sum_{i=1}^{n} K\left(\frac{x - X_i}{b}\right) K\left(\frac{x - Y_i}{b}\right) \tag{6}$$

where the kernel K is defined as $K(z) := \begin{cases} 1/2 & \text{if } -1 \le z \le 1 \\ 0 & \text{else} \end{cases}$.

The estimator calculates the fraction of all (X_i, Y_i) that lie in a (rectangular) neighborhood around the point (x, y) and divides it by the neighborhood's area. The parameter b controls the size of the neighborhood and is usually called the bandwidth of the estimator. This given estimator of the form (6) is known as a bivariate kernel density estimator. In the above case, the kernel K corresponds to the uniform probability density function with support on the interval $[-1, 1]$. In general, one could use any probability density function as the kernel K and the resulting estimator will be a proper probability density function. However, it is usually assumed that the kernel is bounded, i.e. $K(z) < \infty$ for $z \in \mathbb{R}$, and symmetric. The resulting estimator benefits from these features and is easier to theoretically analyze.

Definition: The kernel density estimator with bandwidth parameter $b_n > 0$ is given by

$$\widehat{f}_n(x,y) = \frac{1}{n} \sum_{i=1}^{n} K_{b_n}(x - X_i) K_{b_n}(y - Y_i).$$

for all $(x, y) \in \mathbb{R}^2$. In general, the kernel density estimator can be interpreted from both a local and a global perspective. Herein, we see a point estimate as a weighted average of frequencies in a neighborhood of that point. The weighting is determined in accordance with the kernel function K, and the bandwidth b_n is responsible for determining the size of the neighborhood. Note that the bandwidth b_n in the above definition is annotated to depend on the sample size n. Most commonly used kernel functions are Gaussian, Epanechnikov or Uniform with different characteristics. Various methods to address this issue are considered, but only the bivariate definitions and expressions presented below using the notations introduced before (Nagler, 2014).

Naive estimator: A naive estimator of a copula density $c(u, v)$ with bandwidth parameter $b_n > 0$ is given by $\widehat{c}_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} K_{b_n}(u - U_i) K_{b_n}(v - V_i)$ for all $(u, v) \in [0, 1]^2$ and we used the notation $K_b(\cdot) = 1/bK(\cdot/b)$, K is a symmetric, bounded probability density function on $\mathbb{R}^2$ and $b_n > 0$.

Mirror-Reflection estimator: The mirror-reflection estimator of a copula density $c(u, v)$ with bandwidth parameter $b_n > 0$ is given by $\widehat{c}_n^{(MR)}(u, v) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{9} K_{b_n}\left(u - \tilde{U}_{ik}\right) K_{b_n}\left(v - \tilde{V}_{ik}\right)$ for all $(u, v) \in [0, 1]^2$.

Improved Mirror-Reflection estimator: The improved mirror-reflection estimator of a copula density $c(u, v)$ with bandwidth parameter $b_n > 0$ and shrinkage function $r : [0, 1] \to \mathbb{R}$ is given by $\widehat{c}_n^{(MRS)}(u, v) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{9} K_{r(u)b_n}(u - \tilde{U}_{ik}) K_{r(v)b_n}(v - \tilde{V}_{ik})$ for all $(u, v) \in [0, 1]^2$.

Beta Kernel estimator: The beta kernel estimator of copula density $c(u, v)$ with bandwidth parameter b_n is given by

$$c^{(\beta)}(u,v) = \frac{1}{n} \sum_{i=1}^{n} K\left(U_i, \frac{u}{b_n} + 1, \frac{1-u}{b_n} + 1\right) K\left(V_i, \frac{v}{b_n} + 1, \frac{1-v}{b_n} + 1\right) \tag{7}$$

where $K(x, p, q)$ is the density of a $\mathrm{Beta}(p, q)$-distributed random variable evaluated at x, for all $(u, v) \in [0, 1]^2$.

Transformation estimator: The transformation estimator of a copula density $c(u, v)$ with bandwidth parameter b_n is given by

$$\widehat{c}_n^{(T)}(u,v) = \frac{\sum_{i=1}^{n} K_{b_n}\left(P_u - \Phi^{-1}(U_i)\right) K_{b_n}\left(P_v - \Phi^{-1}(V_i)\right)}{n\phi(P_u)\phi(P_v)} \tag{8}$$

where $P_u = \Phi^{-1}(u)$, $P_v = \Phi^{-1}(v)$ for all $(u, v) \in [0, 1]^2$, where Φ be the standard normal cdf and ϕ its density. In summary, this copula density estimator inherits all of the desirable qualities of the standard kernel density estimator. By following this transformation idea, further extensions such as the fully parametrised transformation estimator, or local Likelihood Transformation estimator based on different choices of smoothing parametrisation can be considered. As this study aims to implement such ideas for having the best non-parametric copula fit, the rest of the theoretical foundations and related properties are skipped. Interested readers can see (Nagler, 2014) for more information on kernel methods for copulas. As a computational tool, the functionalities given in `kdecopula` R package are considered.

3. Modeling Steps

The model we follow was initially proposed by Erhardt et al. (Erhardt and Czado, 2018) to compute univariate and multivariate standardised drought indices.

3.1. Identification of dry and wet conditions

The main drought-related raw variables considered in the study are *Precipitation (PREC)*, *Mean Temperature (MTemp)*, and *Relative Humidity (RH)*, which are essential for drought index construction. The *Water Balance (WB)*, a key ingredient of SPEI, can be calculated by using differences between *PREC* and *Potential Evapotranspiration (PET)*. For the calculation of PET, we are exploiting the (MTemp) values. Each arbitrary drought input variable is represented by a time series, denoted as x_{tk}, where $k = 1, \dots, N; t = 1, \dots, T$. To ensure that small values correspond to dry conditions and large values correspond to wet conditions, the sign of the time series should be reversed if it is the other way around. For instance, high temperatures signify high evapotranspiration, indicating dry conditions, and low temperatures indicate wet conditions. Thus, the (PET) series has to be multiplied by -1 while processing the data.

3.2. Elimination of seasonality

The monthly time series, x_{tk}, typically exhibit seasonal variations in their mean, variance and skewness. To eliminate this seasonal heterogeneity, we need to standardize the time series separately for each month of the year. Each time point, x_{tk} with $k = 1, \dots, N; t = 1, \dots, T$, is represented as a 2-tuple (m_k, y_k), where $m_k \in 1, \dots, 12$ ($1 \equiv$ January, . . ., $12 \equiv$ December) denotes the month, and the integer $y_k \in \mathbb{Z}$ denotes the corresponding year. The month-wise time series $x_m := (x_{t_k})_{k \in \mathcal{K}(m)} = \{x_{(m,y_k)}, k \in \mathcal{K}_m\}$, $m = 1, \dots, 12$, is then constructed, where the index set for month m is defined as $\mathcal{K}_m := \{k : m_k = m\}$.
The first step in standardizing the month-wise time series x_m, where $m = 1, \dots, 12$, is to remove any seasonal skewness. This is accomplished through continuous, monotonic increasing transformations, such as the Yeo and Johnson transformation $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (Yeo and Johnson, 2000), which is similar to the famous Box-Cox transformations. As stated in (Yeo and Johnson, 2000), the skewness in each month-wise time series x_m, $m = 1, \dots, 12$, can be reduced by estimating a parameter λ_m for each month separately using maximum likelihood estimation. The resulting transformed series is denoted as $\hat{x}_m$, $m = 1, \dots, 12$. These transformed series serve as the key inputs for the subsequent calculation of the drought index.

3.3. Elimination of temporal dependence

In addition to seasonality, time series often contain temporal dependence, which can be captured by using *auto-regressive moving average* (ARMA) models (Box et al., 2015). An $\mathrm{ARMA}(p, q)$ model, with auto-regressive order $p \in \mathbb{N}_0$ and moving average order $q \in \mathbb{N}_0$, is used to analyse a deseasonalised and homogeneous time series r_{t_k} with zero mean, where $k = 1, \dots, N$. The model is defined by the equation of $r_{t_k} = \sum_{j=1}^{p} \phi_j r_{t_{k-j}} + \sum_{j=1}^{q} \theta_j \varepsilon_{t_{k-j}} + \varepsilon_{t_k}$ where the error terms ε_{t_k} are independently and identically distributed as $N\left(0, \sigma^2\right)$. Here, ϕ_j and θ_j are the auto-regressive and moving average parameters, respectively. If p or q equals 0, the corresponding component of the model is not considered, resulting in an $\mathrm{MA}(q)$ or $\mathrm{AR}(p)$ model, respectively. If appropriate orders p and q are selected, and the corresponding parameter estimates $\hat{\phi}_j, j = 1, \dots, p$, and $\hat{\theta}_j$, $j = 1, \dots, q$, are used, the model residuals $\epsilon_{t_k} := r_{t_k} - \Sigma_{j=1}^{p} \hat{\phi}_j r_{t_{k-j}} - \Sigma_{j=1}^{q} \hat{\theta}_j \epsilon t_{k-j}, k = 1, \dots, N$ can be approximately temporally independent. Therefore, the $\mathrm{ARMA}(p, q)$ model can help to eliminate the serial dependence in the time series, and the residual terms ϵ_{t_k} can provide information about the impact on the current dryness or wetness conditions every month. As noted by Erhardt and Czado (Erhardt and Czado, 2018), an AR(1) model is usually sufficient for monthly drought statistics. In this study, we follow a similar approach to construct drought indices from weather time series using ARMA models.

3.4. Transformation to standard normal

To transform the data to a standard normal distribution, we apply the estimated distribution $\hat{F}_N$ to transform our residuals, $\epsilon_{t_k}, k = 1, \ldots, N$, into a u-scale, approximately uniformly distributed over the interval $[0, 1]$. The estimate $\hat{F}_N(x)$ is used to compare and test the effectiveness of the non-parametric empirical distribution function of the data: $\hat{F}_N(x) := (1/N)\Sigma_{k=1}^{N}\mathbb{I}\{\epsilon_{t_k} \leqslant x\}$, with residuals $\epsilon_{t_k}, k = 1, \ldots, N$ where $\mathbb{I}$ serves as an indicator function. When fitting a distribution to a sample $\epsilon_{t_k}, k = 1, \ldots, N$, whether parametric or not, it is crucial to assume that the sample is identically and independently distributed. Regarding the data used in this study, for instance, the independence assumption was already established in the previous modelling phase, where temporal dependencies were modelled and eliminated. Additionally, there is no observed variation in the distributions across different months. This is a monotonic increasing transformation, so the process is also referred to as probability integral transformation. The calculation of u_{t_k} is done by formulating $u_{t_k} := N\hat{F}_N(\epsilon_{t_k})/(N+1) = \mathrm{rank}(\epsilon_{t_k})/(N+1), k = 1, \ldots, N$. To prevent any u_{t_k} value from being equal to 1, we multiply the expression by $N/(N+1)$. Then, we transform the data to the standard normal distribution by applying the inverse probability integral transform based on the cumulative distribution function Φ of a standard normal distribution, which results in the calculation of $z_{t_k} := \Phi^{-1}(u_{t_k})$ for $k = 1, \ldots, N$. The transformed data z_{t_k}, where $k = 1, \ldots, N$, is approximately independently and identically distributed with a standard normal distribution (Erhardt and Czado, 2018).

3.5. Standardised index with different timescales

This study applies the temporal aggregation at the end of the described modelling process, to maintain the independence assumption required for fitting a probability distribution to the residuals (Erhardt and Czado, 2018). Once we have completed the corresponding modelling steps outlined earlier in this section, we can compute the index on any desired timescale without repetition. The time series, denoted as $z_{t_k}, k = 1, \ldots, N$, is assumed to be almost temporally independent and follows a standard normal distribution. Therefore, it is already a standardised index with a timescale $l = 1$. A key advantage of using a normal distribution is that the sum of independent normally distributed random variables is also normally distributed. Therefore, we can compute standardised indices for timescales $l \geqslant 1$ by exploiting this property. Specifically, by taking the sum $\Sigma_{j=0}^{l-1} z_{t_{k-j}}$ of standard normal variables, we obtain a new normal distribution with properties $N(0, l)$. Thus, a standardised index with timescale l could be represented as:

$$\mathrm{SI}_l\,(t_k) := (1/\sqrt{l})\Sigma_{j=0}^{l-1} z_{t_{k-j}},$$

for $k = l, \ldots, N$. Notably, this construction is applicable not only to univariate cases but also to multivariate standardised drought indices. The categorisation of the indices is based on dryness and wetness categories given by the quantiles of the standard normal distribution (Svoboda et al., 2002), such as the $2.05 < SI < \infty$ quantile represents the Exceptionally wet condition whereas, $-\infty < SI \leq -2.05$ shows the Exceptionally dry period. In the work of (Svoboda et al., 2002), totally 10 different categories are expressed based on the range of quantile levels.

4. Drought Index

Drought indices play a crucial role in monitoring and assessing drought events, which is essential for effective drought management. These indices enable the quantification of drought characteristics, providing a more comprehensive understanding of the impact of drought. For this study, we elaborate two types of drought indices, namely univariate and multivariate cases, and briefly describe the methods used for their calculation.

4.1. Univariate Drought Index

To compute any standardised drought index, specific steps must be taken based on the selected weather-related variable. For example, the SPI value is calculated based on monthly precipitation (PREC) (McKee et al., 1993). SPEI is an improvement of SPI and considers both PREC and evapotranspiration (PET) through the water balance function: $WB_i = PREC_i - PET_i$ for the given i'th time location. This new univariate variable serves as the input for the calculation of SPEI, similar to SPI. While the calculation of SPEI is generally robust and straightforward (Keyantash and Dracup, 2002), it requires a larger number of observations and is sensitive to the method of PET calculation. The interpretation of SPEI is comparable to that of SPI and considers both precipitation deficit and temperature variations by incorporating PET. The Thornthwaite method, which takes into account the mean temperature and latitude of the weather station, is commonly used to calculate PET. In this study, the functionalities in the SPEI CRAN package were used to derive PET values (Beguería and Vicente-Serrano, 2017). Afterwards, the functions belonging to the `SIndices` R package are adapted to construct univariate indices using raw weather variables.

4.2. Multivariate Drought Index

In addition to the widely used univariate indices, SPI and SPEI, this study evaluates the performance of multivariate indices for the comparison purposes. The standardised multivariate indices are calculated using the method outlined by (Erhardt and Czado, 2018), which can be adapted to model various types of droughts based on different drought-related variables. The method enables the combination of drought information captured in the different variables into a single standardised multivariate drought index using vine copulas, namely Pair Copula Constructions (PCCs).

By following Sklar's theorem, the margins and dependence structure can be modelled separately. As a flexible extension of limited multivariate copulas, vine copulas (vines) are d-dimensional copula constructions composed of both unconditional and conditional bivariate copulas (Aas et al., 2009). Let assume $\mathbf{u} := (\mathbf{u}_1, \ldots, \mathbf{u}_d)$ be the copula data that are obtained from the marginal models corresponding to d different drought input variables, where $\mathbf{u}_j = (u_{j,t_k})_{k=1,\ldots,N}$, $j = 1, \ldots, d$, and u_{j,t_k} is the copula data corresponding to variable j at time t_k. In the second step, a suitable vine copula is estimated for these data (Erhardt and Czado, 2018). Generally, the structure of (d-dimensional) vine copulas is organised by using a nested set of trees $\mathcal{T}_1, \ldots, \mathcal{T}_{d-1}$, which are connected graphs without cycles. The trees $\mathcal{T}_k, k = 1, \ldots, d-1$, are nested in the sense that edges of a tree (which are labelled following the scheme $i, j; \mathcal{D}_k$, where $\{i, j\}$ conditioned set, $\mathcal{D}_k$ is $(k-1)$-dimensional conditioning set) become nodes in the subsequent tree. For a more general explanation of vine copulas, see (Aas et al., 2009) and the recent book (Czado, 2019).

The computation of the multivariate drought index starts with modelling the margins of the input variables and their dependence using copulas, as they allow for separate modelling of the margins and dependence structure. The first step involves changing the sign of the input data, if required, then eliminating any seasonality and temporal dependence. Next, a non-parametric distribution of the residuals is estimated, enabling the transformation to the u-scale, which is required for copula-based dependence modelling. Mainly, there are three alternatives to create a multivariate drought index using different sources of information, as described in (Erhardt and Czado, 2018). The primary definition of the aggregation method is illustrated below;

$$\mathrm{SI}_l^{\mathcal{A}}(1, \ldots, d)\,(t_k) := \frac{1}{\sqrt{(ld)}} \sum_{i=0}^{l-1} \sum_{j=1}^{d} \Phi^{-1}\left(v_{j,t_{k-i}}\right).$$

For the simplicity, we mainly investigated the bivariate and trivariate index construction using (PREC-PET) and (PREC-PET-RH) weather variables, respectively over different time scales (3, 6 and 9 months). We primarily utilised the aggregation method (Method $\mathcal{A}$) with a timescale of $l = 3$, since there exists a certain level of dependence between considered climatic variables. It should be noted that the order in which variables are specified can influence the calculation, as noted by (Erhardt and Czado, 2018). Steps 1 to 3 involve modelling the margins and obtaining standardised time series inputs. Step 4 uses the marginal models from the previous steps to transform several univariate margins into the copula space (uniform margins). Step 5 is the calculation of classical univariate drought indices. Finally, Step 6 deals with vine-copula-based dependence modelling and provides a methodology to combine several variables into a multivariate standardised index (Erhardt and Czado, 2018).

4.3. Drought Characteristics

Once the standardised drought indices have been generated over different time scales, it is possible to classify the index values for interpretation and analysis. In this study, we set a 0-threshold for the standardised index, where positive and negative values indicate wet and dry conditions, respectively. To better understand and describe the nature of drought events, a set of drought characteristics can be derived from the behavior of the standardised index, exemplified by SPI_3 below;

- **Drought duration (D)** is defined as the number of consecutive time intervals (months) where SPI_3 remains below the threshold value selected as 0 for drought condition.
- **Drought severity (S)** is defined as a total magnitude of SPI_3 value during a drought period, which can be formulated as $S_d = |\sum_{i=1}^{D} SPI_{3_i}|$ where SPI_{3_i} denotes the SPI_3 value in the i-th month, as an example.
- **Peak Intensity (PI)** is defined as the minimum SPI_3 value within a drought period to identify the peak of drought events.
- **Mean Intensity (MI)** is defined as the average SPI_3 value within a drought period, which identifies the mean of drought events. It could be formulated as $MI = \frac{S_d}{D}$, where D is the total drought period.
- **Drought Inter-arrival Time (L_d)** is defined as the elapsed time interval from the beginning of a drought to the initiation of the following drought event (The total duration for two consecutive dry periods from the initiation of a drought event to the next one).

4.4. Return Periods

The estimated return period for drought events provides valuable advice for proper water usage (Serinaldi et al., 2009). The average elapsed time or mean interval time between the occurrences of important events, in particular, determines the drought return period (Shiau and Shen, 2001). In this regard, univariate return periods for drought events with duration, severity, and peak or mean intensity can be calculated as follows (Shiau and Shen, 2001) (Shiau, 2006).

$$Rp_D = \frac{E(L_d)}{1 - F_D(d)}, \quad Rp_S = \frac{E(L_d)}{1 - F_S(s)}, \quad Rp_{PI} = \frac{E(L_d)}{1 - F_{PI}(pi)}, \quad Rp_{MI} = \frac{E(L_d)}{1 - F_{MI}(mi)}$$

where Rp_D, Rp_S, Rp_{PI} and Rp_{MI} denotes univariate return periods for duration, severity, peak intensity and mean intensity, respectively; $F_D(\cdot), F_S(\cdot)$, $F_{PI}(\cdot)$ and $F_{MI}(\cdot)$ are percentiles of corresponding CDFs; and $E(L_d)$ denotes the expected value of inter-arrival time between consecutive drought events.
On the other hand, the return period that considers more than one variable is preferred because a multivariate approach delivers more accurate findings in assessing drought. For the sake of simplicity and without loss of generality, we now focus on the bivariate case. Especially, joint behaviour of the drought duration and severity in terms of bivariate and conditional return periods are important tools to analyse dry periods. The return period of drought duration and severity can be formulated based on two cases (Shiau, 2003).

- Case 1: When both the drought duration and severity exceed certain thresholds ($D_d \geq d \wedge D_s \geq s$), the joint return period $Rp_{D_d \wedge D_s}$, which is also referred to as the 'and' case in this study, can be calculated as follows:

$$\begin{aligned} Rp_{D_d \wedge D_s} &= \frac{E(L_d)}{P(D_d \geq d \wedge D_s \geq s)} = \frac{E(L_d)}{1 - F_D(d) - F_S(s) + F_{DS}(d,s)} \\ &= \frac{E(L_d)}{1 - F_D(d) - F_S(s) + C[F_D(d), F_S(s)]} \end{aligned}$$

- Case 2: When either the drought duration or severity exceed the respective threshold ($D_d \geq d \vee D_s \geq s$), the joint return period $Rp_{D_d \vee D_s}$, which is also referred to as the 'or' case in this study, can be calculated as follows:

$$Rp_{D_d \vee D_s} = \frac{E(L_d)}{P(D_d \geq d \vee D_s \geq s)} = \frac{E(L_d)}{1 - F_{DS}(d,s)} = \frac{E(L_d)}{1 - C[F_D(d), F_S(s)]}$$

where $d, s \in R^+$ represent the thresholds for duration and severity, respectively; $Rp_{D_d \wedge D_s}$ and $Rp_{D_d \vee D_s}$ represent the joint return period for Case 1 (the 'and' case) and Case 2 (the 'or' case), respectively. The bivariate copula function for drought duration and severity, $C\left[F_D(d), F_S(s)\right]$, is computed from the corresponding marginal CDFs, $F_D(d)$ and $F_S(s)$.
Once the suitable copula family has been selected for each characteristic pair,both univariate and bivariate return periods can be derived by exploiting the properties of drought events and the drought indices. However, the interpretation of the return period can become more prone to misinterpretation in non-stationary and multivariate settings,

compared to classical univariate frequency analysis of independent-identically distributed (iid) data (Serinaldi, 2015). This observation was pointed out by Serinaldi (Serinaldi, 2015), who argued that the return period Rp, which is often interpreted as "This event is expected to occur on average once each Rp years," can be misleading in such settings. Therefore, it is crucial to examine the joint probabilities that correspond to commonly studied bivariate return periods, which can be expressed as denominators in return period definitions: (i) $p_{\text{and}} := 1 - F_D(d) - F_S(s) + C[F_D(d), F_S(s)]$ and (ii) $p_{\text{or}} := 1 - C[F_D(d), F_S(s)]$.

The calculations presented above are based on the univariate setting of the exceedance probability concept, which represents the 'probability p_M of observing a specific event at least once in the design life period of M years'. This concept has a unique and general definition that applies to both univariate and multivariate cases. For this study, the main goal is to examine the case where $M = 1$ to understand the likelihood of certain events for the upcoming year. The calculations based on the exceedance probability concept, which represent the probability of observing a critical event at least once in the design life period. This concept has a unique and general definition that applies to both univariate and multivariate cases. The formulations for bivariate return periods and occurrence probabilities presented earlier are extensions of the exceedance probability concept in the univariate setting. Further details on the use of probabilities and the exceedance probability in place of return periods can be found in (Serinaldi, 2015).

In this study, we mainly focus on the corresponding joint exceedance probabilities instead of calculating the return periods similar to other studies. Specifically, we examine various pairs of drought characteristics, such as Duration-Severity, Duration-Peak Intensity, and Severity-Peak Intensity, for Case 1 and 2 as presented earlier. To illustrate, a set of suitable grid values are considered for the calculation of p_{and} and p_{or} cases. In the upcoming sections, mainly the findings of two settings are summarised: Parametric marginals-Parametric Copulas (PP) and Parametric marginals-Non-parametric Copulas (PN), with selected drought characteristics to see the potential impact of considered copulas.

5. Türkiye Case Study

This section first summarizes the numerical findings that belong to the monthly climatic data of the Aksehir weather station in Konya province from 1950 to 2010. Mainly, the data includes three variables: monthly precipitation (PREC), mean temperature (MTemp) and relative humidity (RH). During data preparation, a potential evapotranspiration (PET) variable is generated by using the Thornthwaite method, by using the MTemp as the input. Drought index calculation and related dependence analysis are exemplified for the Aksehir station and then extended for different weather stations in the same province, visualized in Figure 1 with their location specific information (Source: https://www.mgm.gov.tr/).

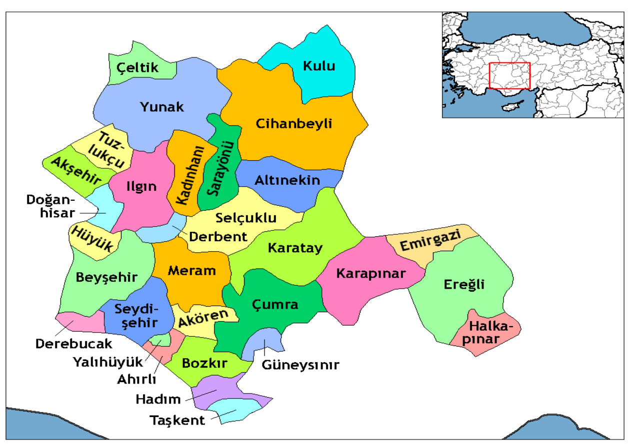


Figure 1: Map of Konya province including sub-regions (Considered stations: Aksehir (latitude 38.36, longitude 31.42), Cihanbeyli (latitude 38.65, longitude 32.92), Karapinar (latitude 37.71, longitude 33.52), Cumra (latitude 37.56, longitude 32.79) and Yunak (latitude 38.82, longitude 31.72))

5.1. Aksehir Weather Station

Regarding the constructed drought indices, including all stations that we explored, 3-month visual display is briefly presented. For simplicity, the last ten years (2000-2010) are highlighted by considering the reported drought events in the Konya closed basin. The multivariate indices (SBI and STI) demonstrated comparable performance in capturing the onset and recovery of the dry periods over time (See Figures in Appendix A). Regarding the reported drought events such as 2000-2001, 2004-2005, 2006-2007 and 2008 in (Orman Bakanligi, 2023), the multivariate indices performances aligned with the other used drought indices in that area. They behave similarly to each other but differ from the univariate indices, especially from SPI behaviour. To illustrate, for the dry period in 2000-2001, multivariate indices performance indicates more severe drought, whereas SPI has larger values and this is common in all the considered weather stations. For a different period during 2007, the reported drought severity based on specific indices, multivariate indices are aligning with the Decimal indices, commonly used in this specific region (Orman Bakanligi, 2023). Such nuances and comprehensive approach can play a crucial role regarding drought management for the Konya closed basin, and the findings of the study directly contributing to enriching the Konya Basin Drought Management Plans (Orman Bakanligi, 2023).

Mainly, the above-mentioned steps are followed to construct the related drought characteristics. The similar calculations are executed for different drought indices (SPI, SPEI, SBI and STI) over distinct time scales. To illustrate, the extracted drought events for different indices can be summarized with their descriptive statistics as in Table 1.

Table 1: Drought Characteristics Statistics Based on Different Drought Indices.

	SPI			SPEI		
Statistics	3-month	6-month	9-month	3-month	6-month	9-month
Drought Counts	102	65	62	96	69	55
$E(D)$	3.569	5.462	5.710	3.740	5.101	6.455
$E(S)$	2.781	4.297	4.414	2.989	4.067	5.165
$E(MI)$	0.621	0.568	0.452	0.644	0.540	0.454
$E(PI)$	1.005	0.954	0.747	1.025	0.896	0.822
$E(L_d)$	7.158	11.297	11.902	7.579	10.632	13.259
	SBI			STI		
	3-month	6-month	9-month	3-month	6-month	9-month
Drought Counts	95	78	61	97	70	54
$E(D)$	3.853	4.487	5.639	3.825	5.043	6.537
$E(S)$	2.91	3.553	4.598	2.864	3.890	5.145
$E(MI)$	0.574	0.46	0.437	0.612	0.497	0.462
$E(PI)$	0.933	0.774	0.728	0.999	0.840	0.762
$E(L_d)$	7.553	9.130	11.700	7.479	10.217	13.245

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity and L_d for the Interarrival time. ($E(.)$ stands for the expectation of the given characteristics.)

As one can realize from Table 1, there are certain impacts of selected time scale and the considered drought index for the summarized drought events. Firstly, when the duration is increased from 3 to 9- months, the number of detected drought events are decreased for all cases with different frequencies. Specifically, when the PET variable is taken into account under univariate index (SPEI) or joint indices (SBI or STI), the drought count numbers are comparatively smaller than SPI over 3-month scale. This pattern is preserved except the behavior of SBI over 6-month period in the lower panel of the given Table 1. When the time scale increased, for each index, the corresponding inter-arrival time between two consecutive dry periods are getting larger with different increasing rates naturally. In terms of main characteristics, such a pattern occurs for duration and severity whereas the opposite change appears for the mean and peak intensity. Briefly, considering different standardized index can result in different characteristic values so it is important to exploit more than one drought index for a comprehensive understanding.

Additionally, the dependence structure between the drought characteristics may be different for each index and this difference can play an important role for the drought analysis. As an illustration, the dependence between four main characteristics are visualized in Figures 2a, 2b, 2c and 2d to detect the change from SPI, SPEI, SBI and STI over 3-month scale, respectively.

For the parametric marginal fitting, five widely considered distributions are tested by following the considered examples in the literature. The best-fitting distribution for each drought characteristic was obtained based on criteria such as the goodness-of-fit to the observed data. Specifically, Normal, Log-normal, Weibull, Exponential and Gamma distributions are compared to detect the best option for each drought characteristic. Before doing so, it is necessary to

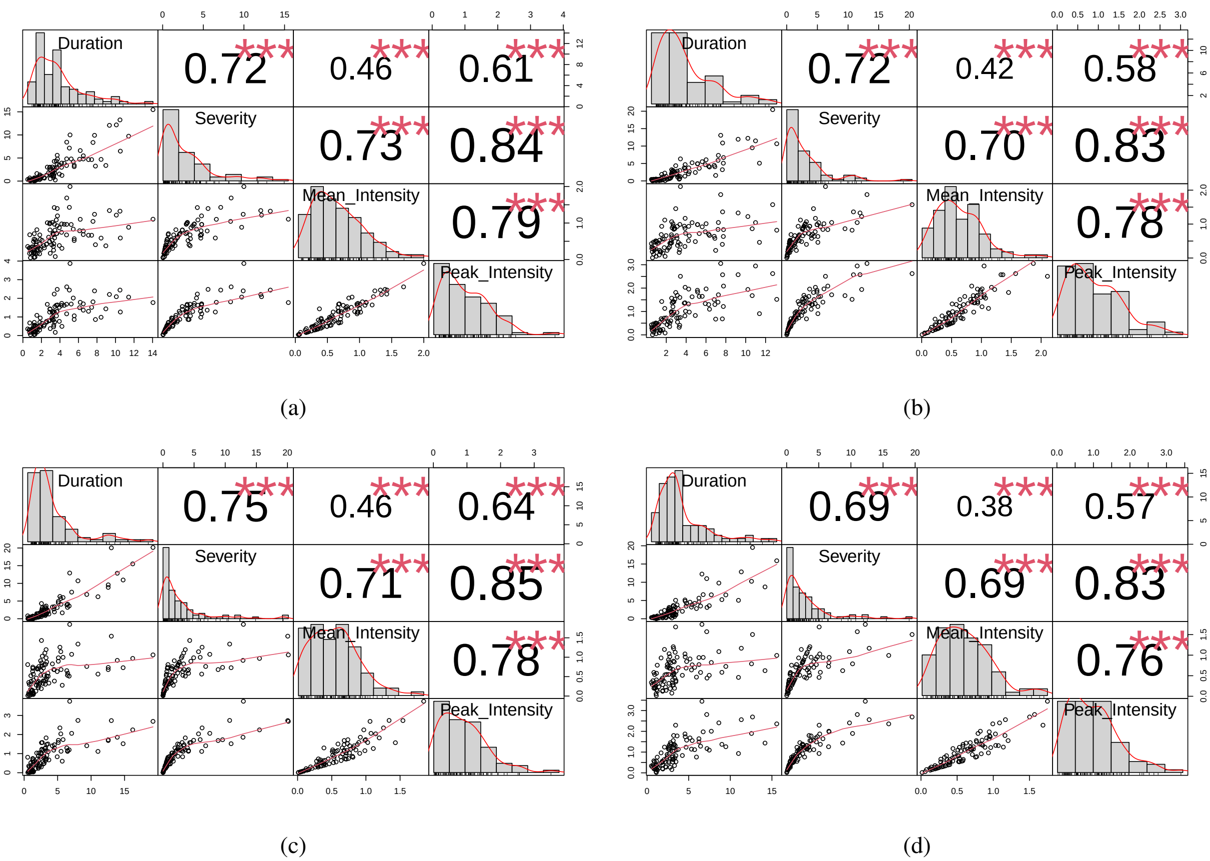


Figure 2: Kendall correlation between drought characteristics extracted from (a) SPI, (b) SPEI, (c) SBI and (d) STI over a 3-month scale. (Note that the automatically generated red stars are to simply highlight the correlation significant. Shortly, the number of red stars in the figure indicates the level of statistical significance: *: p ¡ 0.05, **: p ¡ 0.01, ***: p ¡ 0.001, as determined by Kendall's tau correlation.)

apply continuity correction to the duration data, typically a discretised version of the actual continuous measurements for drought index data with some repeated (tied) observations. When dealing with tied data (repeated values), we followed the pre-processing concept that adds random values between -0.5 and 0.5 to the values of duration, helps eliminate the issue of systematic ties in data. This approach minimally affects the inherent marginal distribution of duration, as explained by Pham et al. (Pham et al., 2016), and has been utilised in previous studies (Salvadori et al., 2007) (Pappada et al., 2017). By adding noise, duration is treated as a continuous variable, and ties are resolved as described above. As an illustration, one can see the similarities and differences of the parametric marginal distributions on 3-month scale in Table 2. As shown in Table 2, for the all considered drought indices, the characteristics of S, PI and MI followed the Weibull distribution with varying shape and scale parameters, except the case of SBI3 for MI (Normal distribution case). For Duration, again for all indices, the best selected distribution is the log-normal with very close mean and standard deviation parameters.

Table 2: Summary of selected marginal distributions for drought characteristics at a 3-month scale.

	Univariate Drought Characteristics			
	D	S	PI	MI
SPI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=1.0087	shape=0.8585	shape=1.3809	shape=1.6476
par2	sdlog=0.7455	scale=2.5664	scale=1.0997	scale=0.6953
SPEI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=0.9954	shape=0.8536	shape=1.4181	shape=1.6566
par2	sdlog=0.8255	scale=2.7486	scale=1.1233	scale=0.7185
SBI3				
Dist	Lognormal	Weibull	Weibull	Normal
par1	meanlog=1.000	shape=0.7327	shape=1.1815	mean=0.5739
par2	sdlog=0.821	scale=2.3830	scale=0.9835	sd=0.3755
STI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=1.0669	shape=0.8796	shape=1.4533	shape=1.7418
par2	sdlog=0.7497	scale=2.6754	scale=1.1018	scale=0.6865

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity.

For the selection of suitable copulas, we considered well-known copula families including all options in the VineCopula R package and the best one is selected based on the specific dependencies observed between drought characteristics and BIC value

Table 3: Summary of fitted parametric copulas for the pairs of drought characteristics (3-month, PP case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Family (Rotation)	Frank	Frank	Frank	Clayton (0)
	θ_1 (θ_2)	12.57 (-)	5.06 (-)	8.3 (-)	9.55 (-)
SPEI3	Family (Rotation)	Frank	Frank	Frank	BB7
	θ_1 (θ_2)	11.97 (-)	4.4 (-)	7.37 (-)	2.31 (12.31)
SBI3	Family (Rotation)	Gumbel	Gumbel (180)	Frank	BB7
	θ_1 (θ_2)	3.8 (-)	1.78 (-)	8.76 (-)	2.79 (16.16)
STI3	Family (Rotation)	BB8	Gaussian	Frank	Gumbel (180)
	θ_1 (θ_2)	7.64 (0.81)	0.56 (-)	7 (-)	5.75 (-)

BB7 is a bivariate case of Joe and Clayton copula; BB8 is a bivariate case of Joe and Frank copula.

For the fitted copulas, Table 3 and 4 summarizes the corresponding PP and PN scenarios for each drought index. Different standardized indices (SPI, SPEI, SBI and STI) over 3 months resulted in different parametric copula fits on the pair of characteristics. After deciding the suitable marginal distribution and attaching to two different approaches

Table 4: Summary of fitted non-parametric copulas for the pairs of drought characteristics (3-month, PN case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6852	0.4215	0.5677	0.7983
SPEI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6807	0.3825	0.5432	0.7919
SBI3	Kernel Type	T	T	T	T
	Kendall's τ	0.7197	0.421	0.6025	0.8209
STI3	Kernel Type	T	T	T	T
	Kendall's τ	0.656	0.3445	0.533	0.8033

(Parametric margin-Parametric copulas and Parametric margin-Non-parametric copula), the next step is to explore the exceedance probability calculations based on certain thresholds. To illustrate, the change across time scales, joint exceedance probabilities and applied PP-PN settings can be visualized for SPI on duration-severity pair. In Figure 3, one can see the opposite pattern change from 3-month to 9-month scales for p_{and} and p_{or} exceedance probability scenarios. To be more specific, on the left panel, when the time scale increased the PP setting dominates the PN setting for the grid values of duration and severity. On the other hand, right panel indicates the reversed relationship, where PN is getting larger when the time-scale is raised but having different behaviors for the 9-month scale.

To have a general comparison on different indices over the same time scale, only 3-month scale is investigated for the simplicity. Specifically, for the both p_{and} and p_{or} probability calculations, over PP and PN settings, a general pattern comparison can be summarized with the help of contour plots. In Figure 4a, one can see the p_{and} exceedance probability contour lines for the all indices under PP and PN setting. In a similar way, the main differences between four standardized index can be followed for p_{or} in Figure 4b.

In Figure 4a, regarding individual patterns of PP and PN setting for each index, the upper 2x2 panel is the main summary. From higher to the lower exceedance probability values (off-diagonal cursor movement), the differences between PP and PN getting more visible. Specifically, for the rare events, which copula approach is used has a moderate impact on the calculated values. In the bottom panel, the overall comparison of four indices shows that for PP setting, with low probability values, the bivariate and trivariate index behaves slightly different whereas univariate indices almost overlapped each time. This pattern quite common for the higher probability values (bottom-left corners) and certain differences appear towards to the lower probabilities (top-right corners). In PN setting, the contour lines behave similarly for rare events. This finding is reversed when the calculation of exceedance probability values are switched to the p_{or}, as in Figure 4b. Mainly, the calculated values much larger compared to p_{and} but the changing behaviour from the higher to lower (bottom-left to top-right) is preserved. However, the contour line behavior for indices seemed to be reversed at a certain level for both PP and PN settings (bottom panel of the same figure). To explore these patterns for other pair of characteristics, mainly the duration-peak intensity and severity-peak intensity combinations are summarized in terms of comparing indices all at once.

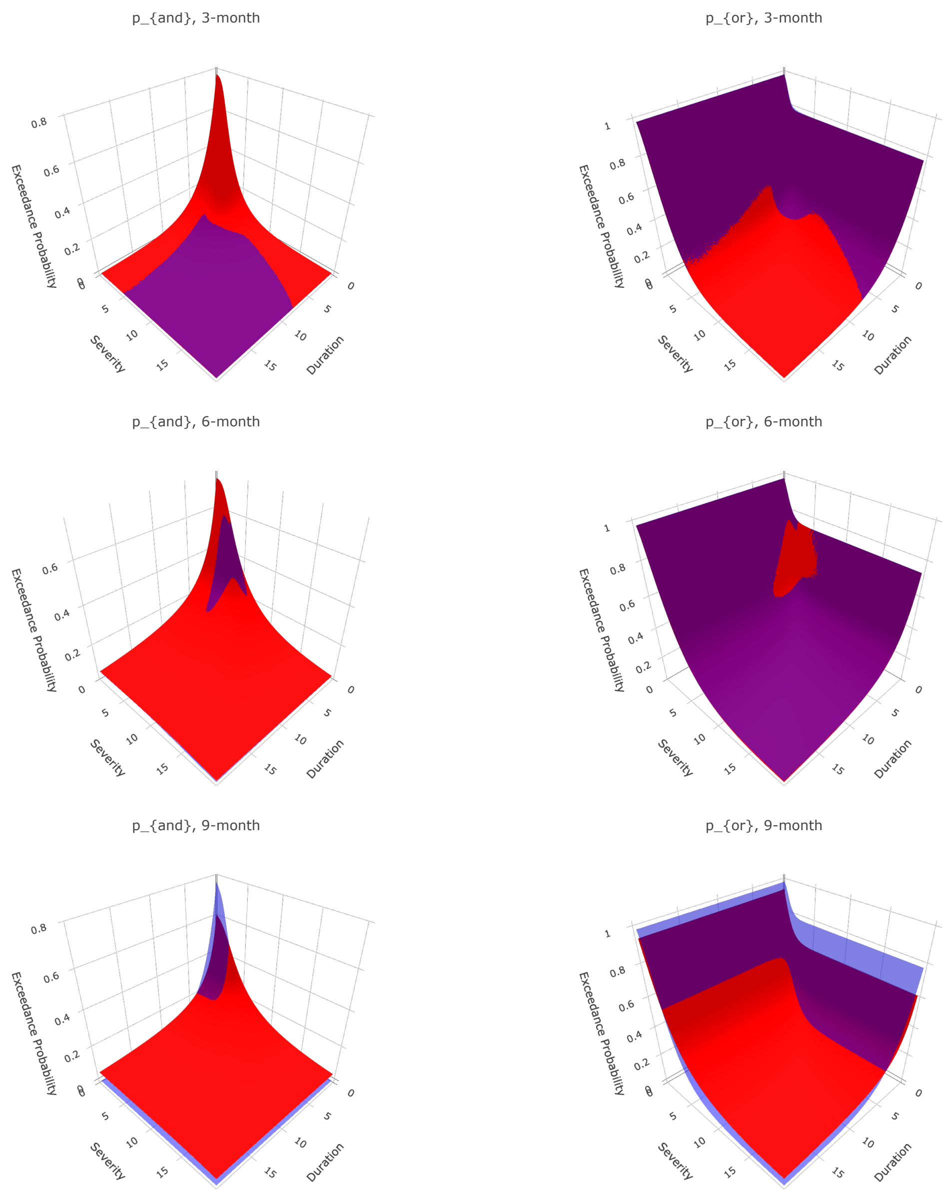
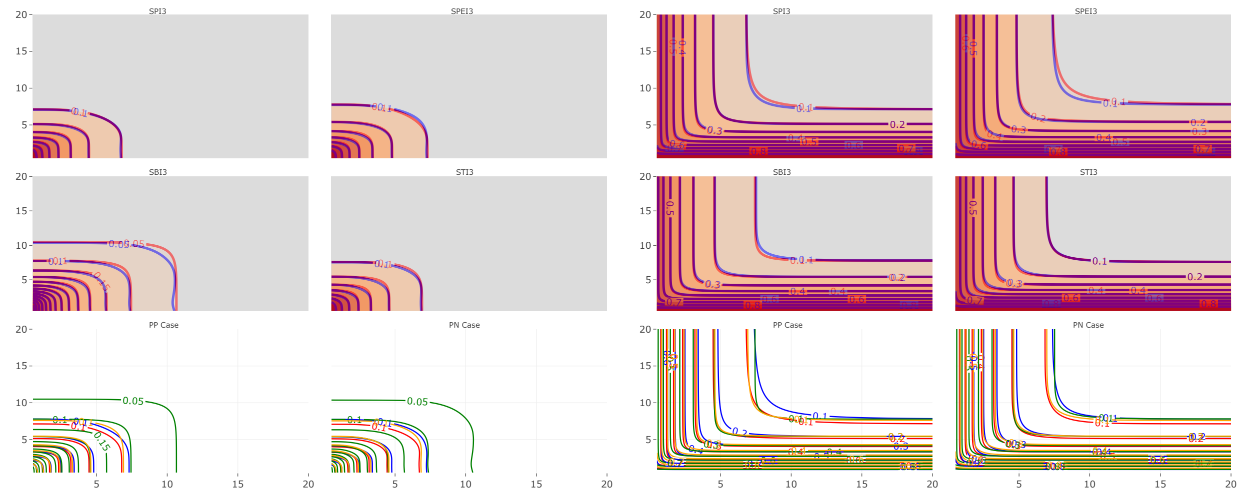
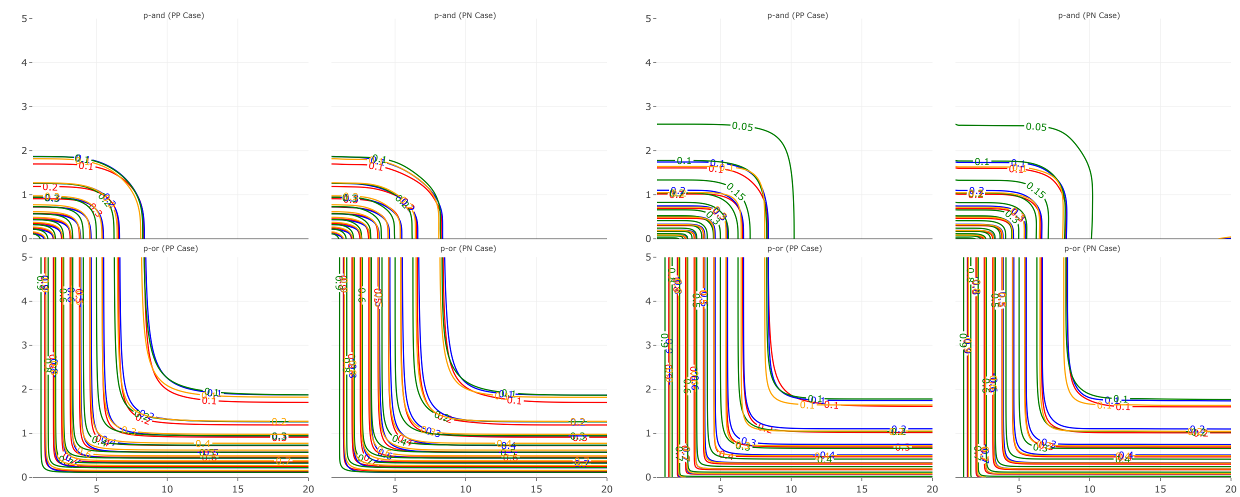


Figure 3: Left panel for p_{and} and right panel for p_{or} cases for Duration-Severity pair wit PP: Red, PN: Blue



(a) (b)

Figure 4: Exceedance Probability Contour Plots for (a) p_{and} and (b) p_{or} (Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3)



(a) (b)

Figure 5: Exceedance Probability Contour Plots for (a) Duration-Peak Intensity pair and (b) Severity-Peak Intensity pair Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3

In Figure 5a, both PP and PN scenarios are summarized graphically for the pair of Duration-Peak Intensity. For the highly likely events defined by the larger p-and case, there are slight differences in the contours presented in the upper panel. Mostly, both SPI3 and SPEI3 overlap, whereas there are some discrepancies for STI3 and SBI3. Especially, the differences between the univariate or multivariate types of drought indices are more visible when the exceedance probability values are smaller (off-diagonal movement on the contour lines). This means that the occurrence of extreme drought is associated with higher drought characteristics and this pattern is varying across different drought indices. Similar behavior can be observed in the lower panel for p-or values. To illustrate, for the value of exceedance probability of 0.1 under the PP model, both SBI3 and STI3 seem to lie above SPI3 in general. Overall, Figure 5a shows some varied behavior across PP and PN cases for p-and and p-or values, suggesting that the choice of drought index can moderately influence the relationship between drought characteristics. Such differences are more visible in Figure 5b for the pair of Severity-Peak intensity for the likelihood of joint extremes in the upper panel (towards the right upper corner). Specifically, differences or overlaps for different drought indices are presented over contour lines to demonstrate that how the index can be crucial for certain pair values of drought characteristics.

5.2. Other Stations

In addition to the Aksehir station as a main location, other weather stations with certain characteristics are examined to take a general snapshot of the region. For that purpose, four additional weather stations are investigated, including Cihanbeyli, Karapinar, Cumra and Yunak stations located in the same province. Because of the computational demand for the PN case, rather than adding all other weather stations, these four specific locations are considered. Another motivation for this is coming from the fact that these are most vulnerable places based on crop yield loss in general since various farmers got support previously (these are the cities that higher money support is given to farmers during 2002-2017). In the meantime, there are certain differences on their latitude values, compared to Aksehir station. In that sense, these new stations may serve as complementary for the overall drought analysis by underlining different locations. For the new stations, we will be presenting mainly three contour lines of p_{and}- p_{or} under PP and PN settings over only 3-month scale (short term period), by summarizing the rest of the calculations in the Appendix part. To illustrate the main differences, drought statistics based on different locations are summarized in Table 5 similar to the Table 1 only.

Table 5: Drought Characteristics Statistics Based on Different Drought Indices across four weather stations (3-month scale only).

	Cihanbeyli				Karapinar			
Statistics	SPI	SPEI	SBI	STI	SPI	SPEI	SBI	STI
Drought Counts	102	99	102	101	102	90	100	105
$E(D)$	3.539	3.586	3.49	3.386	3.539	4	3.57	3.467
$E(S)$	2.854	2.902	2.769	2.751	2.768	3.157	2.793	2.642
$E(MI)$	0.615	0.624	0.576	0.6	0.625	0.622	0.62	0.598
$E(PI)$	0.983	1.018	0.93	0.954	1.004	0.998	0.96	0.938
$E(L_d)$	7.149	7.306	7.109	7.11	7.129	8.124	7.232	6.904
	Cumra				Yunak			
	SPI	SPEI	SBI	STI	SPI	SPEI	SBI	STI
Drought Counts	92	98	93	91	99	97	101	99
$E(D)$	3.978	3.735	3.957	4	3.525	3.608	3.485	3.566
$E(S)$	3.177	3.008	3.117	3.157	2.945	2.986	2.723	2.86
$E(MI)$	0.646	0.618	0.632	0.649	0.673	0.683	0.642	0.657
$E(PI)$	1.068	0.96	1.02	1.048	1.058	1.101	1.02	1.068
$E(L_d)$	7.901	7.381	7.761	7.978	7.418	7.552	7.16	7.327

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity and L_d for the Interarrival time. ($E(.)$ stands for the expectation of the given characteristics.)

For Cihanbeyli station, for the drought characteristic pairs (Duration - Severity, Duration - Peak Intensity and Severity-Peak Intensity), the exceedance probability behaviours are summarized in Figures 6a-6c. For each characteristic pairs, one can see the main differences between PP and PN settings, specifically the discrepancy between indices, which appeared more visible for events that were less likely (more extreme events) for the upcoming year.

In Figures 6a - 6c, three different pairs of drought characteristics behavior patterns over varying joint exceedance probabilities are summarized in a similar way. To illustrate, the differences between four drought indices are more visible for the extreme events that represents why decision-makers should not rely on a single index for certain events. For specific cases, in Figure 6a, SBI3 contour line lies above the others that implies at the same level of exceedance probability, that index results in higher Duration for the fixed value of Severity (in upper-left panel). Such contour line differences are less visible for the pair of Duration-Peak Intensity in Figure 6b wheres there are certain changes in Figure 6c for the pair of Severity-Peak Intensity. As an example, for a fixed severity, the peak-intensity values are very different for the extreme events for the univariate or multivariate indices (upper panel of p-and calculations). Such a difference is narrowed down in the case of p-or by nature, given in the lower panel of Figure 6c. Especially for extreme events, both the drought index selection and the modeling approach can impact the values of drought characteristics, which implies the importance of looking at more than one drought index. In a similar manner, for the other remaining stations (Karapinar, Cumra and Yunak), similar visualizations are presented to highlight the potential differences in the calculated joint exceedance probabilities. In general, similar patterns are observed across different locations for the less extreme events (small p-and, p-or values). Generally, across different geographical locations, different drought

 DOI: http://dx.doi.org/10.18187/pjsor.v20i3.4609

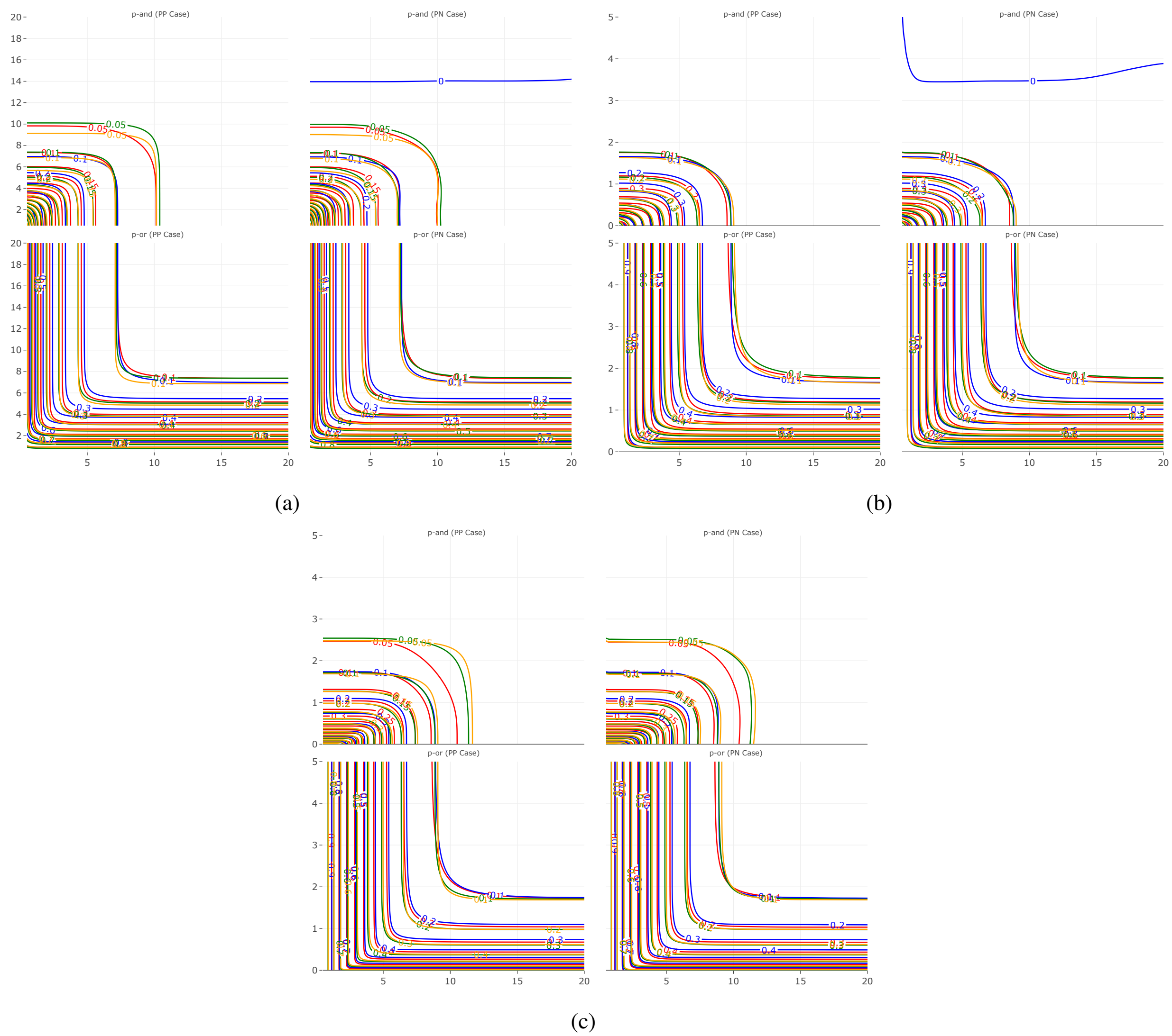
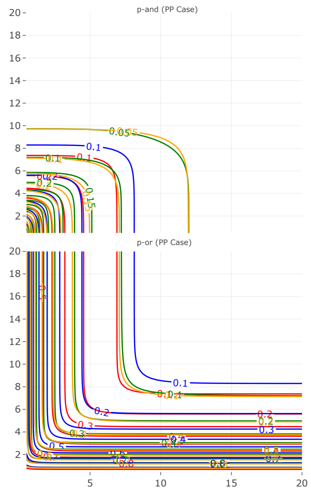
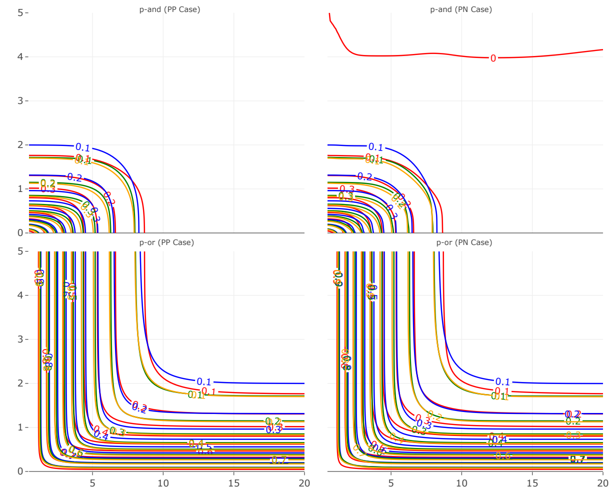


Figure 6: Exceedance Probability Contour Plots for (a) Duration-Severity, (b) Duration-Peak Intensity and (c) Severity-Peak Intensity pairs for Cihanbeyli station (Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3)

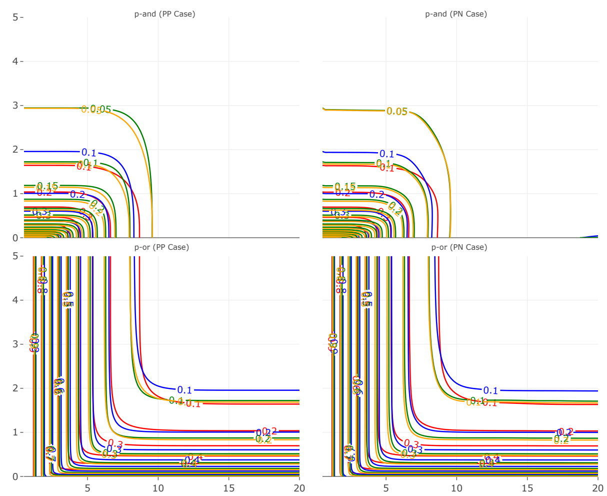
indices have a tendency to behave differently during extreme drought events.



(a)



(b)



(c)

Figure 7: Exceedance Probability Contour Plots for (a) Duration-Severity, (b) Duration-Peak Intensity and (c) Severity-Peak Intensity pairs for Karapinar station (Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3)

The importance of using different indices is elevated when the geographical characteristics are changed. To illustrate, for Karapinar, as visualized in Figures 7a-7c, the visibility of the differences of high and less extreme events is moderate. Besides, the contour line flow across different values of drought characteristics under the given exceedance probability has a tendency to change its pattern. For example, in the upper panel of Figure 7b, for the extreme event (p-and = 0.1), firstly SPEI3 contour line lies above the other indices (means higher drought characteristics) for the values of Duration between 0-5 whereas SPI3 contour lies at the top of others for higher Duration (nearby 10). Additionally, differences between univariate and multivariate drought indices are comparatively larger for the extreme events in Karapinar compared to Aksehir and Cihanbeyli in different drought characteristic pairs. This empirical evidence underlines the location-specific use of certain drought indices based on spatial properties.

For Cumra weather station, behaviors are changed again for the extreme events in each pair, with certain patterns dominated by the univeriate index. As an example, in Figure 8a, SPEI3 lies above the other contour lines in the upper panel for the case of p-and. Similar to Karapinar case, there are some turning points in the behavior of some indices regarding the change of drought characteristic values. In Figure 8b, SBI3 and STI3 are lying below the univariate cases for smaller Duration values, whereas they started to lie at the top when Duration is larger across different peak-intensity values. Especially for the contour line of SPI3 in Figure 8c is moderately different compared to the patterns in Figure 7c. These slight nuances of different indices can have certain impacts on decisions about water or agricultural

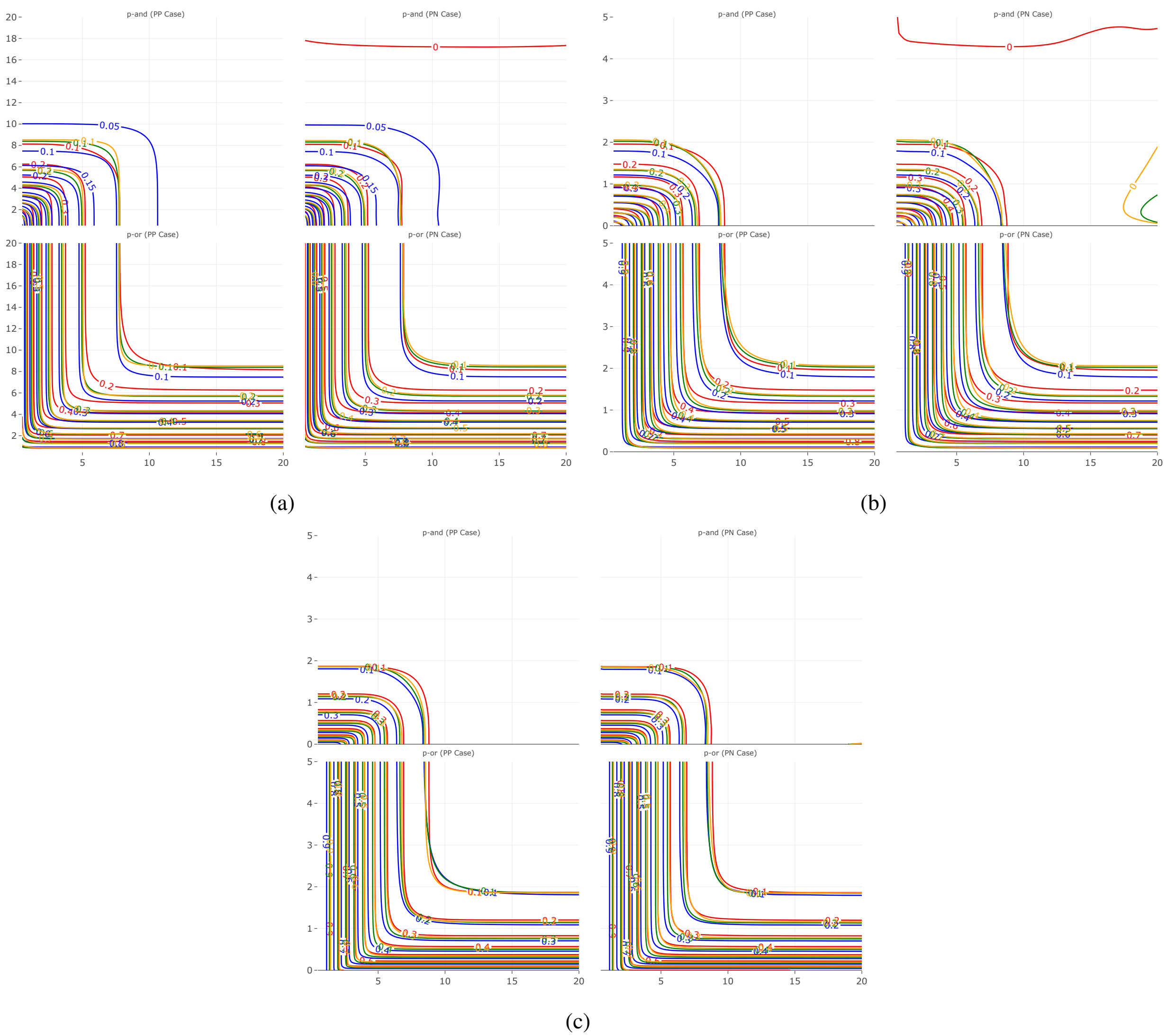
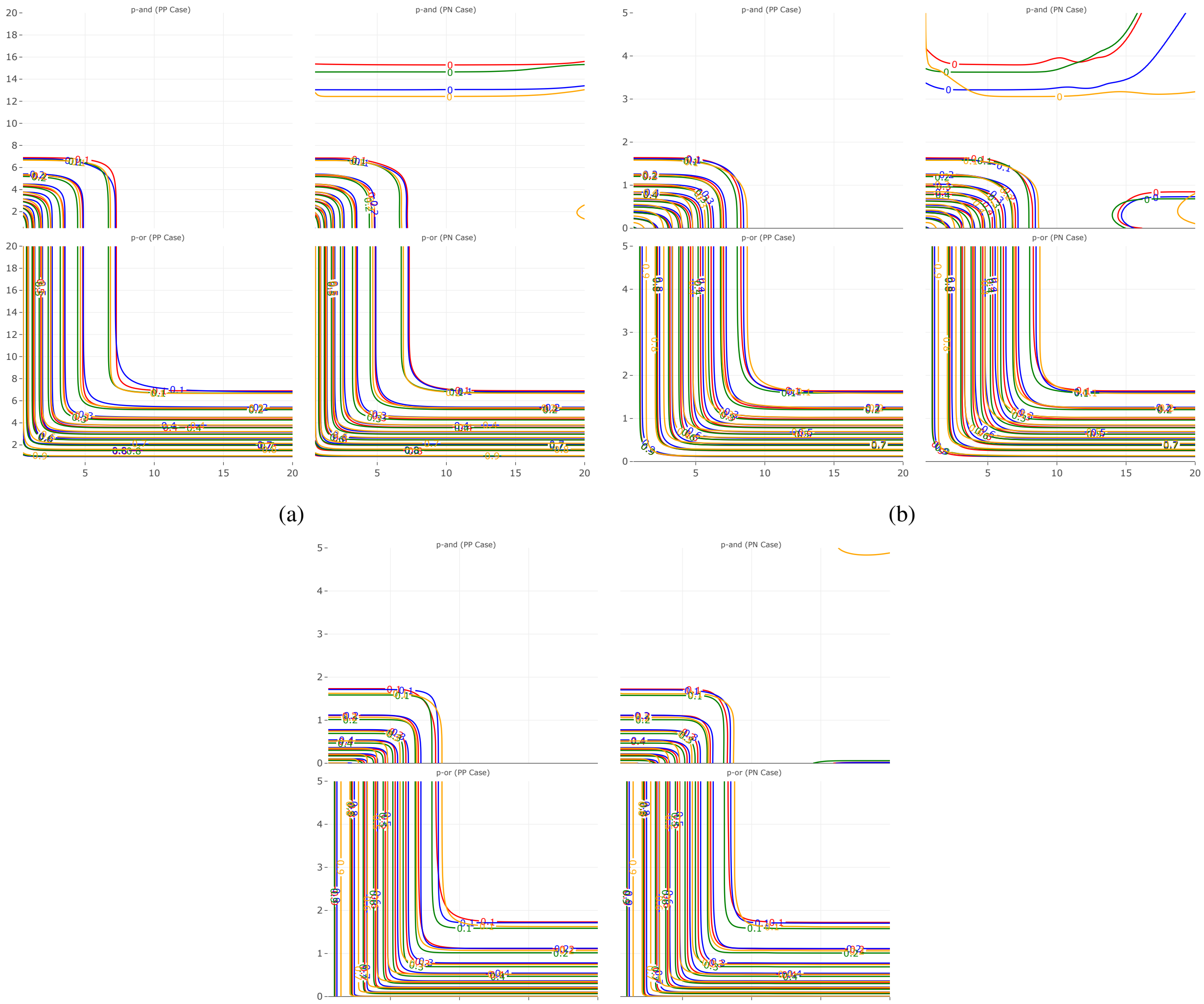


Figure 8: Exceedance Probability Contour Plots for (a) Duration-Severity, (b) Duration-Peak Intensity and (c) Severity-Peak Intensity pairs for Cumra station (Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3)

risk management in certain areas. For that reason, leveraging multiple indices at a certain location can be crucial for a comprehensive drought characteristic analysis.



(a) (b) (c)

Figure 9: Exceedance Probability Contour Plots for (a) Duration-Severity, (b) Duration-Peak Intensity and (c) Severity-Peak Intensity pairs for Yunak station (Orange: STI3, Green: SBI3, Blue: SPEI3, red: SPI3)

For a consistent but changing behavior of contour lines of four different indices, Yunak station can be different from others. In Figures 9a-9c, for the less extreme events (higher exceedance probability values), the contour lines follow a certain pattern but change across different values of p-and and p-or calculations. Regarding the nature of drought event under two calculations, whether both characteristics or at least one characteristic is the focus can have an impact on drought event conditions. To be complete, it is possible to consider both p-and and p-or values for the overall change at a specific location. For the Yunak case, as an example, STI3 lies above the other contour lines most of the time in Figure 9b whereas this is reversed in general for the case of the Duration-Severity pair given in Figure 9a. Such differences can be crucial for understanding the risks of different drought events with changing extreme levels (i.e., joint exceedance probabilities).

Overall, these presented contour plots for specific weather stations are instrumental for comparing the performance of different drought indices in capturing the joint distribution of drought characteristics. For the pairs of Duration-Severity, Duration-Peak Intensity and Severity-Peak Intensity, different modeling results (PP and PN) provides insights into how model selection affects the interpretation of drought risks and related characteristics. Additionally, empirical evidences in the same region with certain geographical differences highlight the necessity of using multiple indexes to monitor drought characteristics more effectively.

6. Concluding Remarks

In this work, both parametric and non-parametric copula approaches are investigated in detail to understand their impact on the dependence between drought characteristics. For that purpose, at the first step, the vine copula approach is exploited for the multivariate drought index construction (mainly bivariate and trivariate cases). The constructed joint indices (SBI and STI) were demonstrated to be effective tools for drought management, offering additional insight over the traditional univariate SPI and SPEI by integrating additional climatic variables (PET, RH). Regarding the geographical nature of the selected region, the complex impact from the precipitation can be captured more effectively by adding moisture availability and atmospheric demand. The findings come from different weather stations in the same region so it can be informative about their possible usage and their interpretation. Over different time scales, various exceedance probability calculations are summarized visually to depict the main similarities and differences. The main findings are summarized for 3-month scale and certain drought characteristic combinations (Duration-Severity, Duration-Peak Intensity and Severity-Peak Intensity). For the overall behavior of these characteristics, exceedance probability values are investigated in detail instead of classical return period calculations. Additionally, the use of copulas in the construction of the multivariate drought index offers promising avenues for advancing drought risk management strategies. In that sense, this study is the first comprehensive case that explores joint exceedance probabilities focusing on Türkiye's specific region.

Over a 3-month scale, SPI, SPEI, SBI and STI display similar magnitudes, but slight differences in the case of Aksehir. With regards to the copula fit of the same station, under PP setting different copula family selection is done for different indices, while PN indicates that the transformation estimator (T) is suitable and consistent in different pairs. A similar selection summary can be mentioned for the other considered weather stations; details are summarized in the Appendix. Two distinct scenarios regarding the p_{and} and p_{or} resulted slightly different results in terms of the use of four different indices. Differently from other previous works, main focus is the calculated exceedance probability values under certain models and selected drought indices rather than relying on return periods. For a certain location, the main findings indicate the importance of index selection and the resulting behavior of drought characteristics that describe the drought event properties under exceedance probability values. Specifically, from higher to lower joint exceedance probabilities (from less to most extreme cases), the differences between indices are more visible in various locations. Additionally, over different locations, patterns of drought indices are changing, which suggests the use of multiple indices for a comprehensive drought analysis. In that respect, it is crucial to leverage the benefits of both univariate (SPI, SPEI) and multivariate (SBI, STI) types of indices under different copula model settings (PP and PN).

Regarding the distribution fit or copula model selections on various weather stations, there are certain differences that can be followed in the presented detailed Tables in the Appendix B-E. To illustrate, in Cihanbeyli station, the best distribution fit results varying across different indices and drought characteristics including Gamma or Exponential one. The summary Table on the distributions indicate that why we need a distribution fit selection rather than only using one fixed family borrowing from the literature. In terms of the fitted parametric copula, certain spatial differences attached to the various tail dependence properties such as D-S pair based on SBI3 follows Clayton for Cihanbeyli station in Table 7 whereas the same D-S pair based on SBI3 follows Gumbel for Cumra station, tabulated in Table 13. Especially, for the same pair of drought characteristics, the use of different drought index can result in distinct copula families, good indicator to consider various monitoring tools together. Under the non-parametric copula approach, on the other hand, all the best fitted models follow the family of 'T' estimator.

From a practical point of view, this empirical study highlights the issue of drought index and copula modeling issues over a drought-prone region. In terms of water resources and agricultural risk management, the findings of the study could be informative for decision-makers. Such analyses are vital for enhancing drought characteristic models, aiding in the development of more effective drought risk management and mitigation strategies. Regarding the importance of the chosen region for agricultural production, the extreme event drought characteristics and the corresponding occurrence probabilities can be major forces in upgrading drought risk management policies. Especially for the behavior of multivariate drought indices, which take into account multiple weather variables, they can be more reliable for drought risk management compared to the classical univariate drought indices. Considering multiple drought indices over the same location may allow for the design of step-by-step risk action plans and territorial-based policies to cope with the harmful impacts of dry periods.

From a modeling perspective, the scenarios considered in this paper is limited up to PP and PN cases. On the other hand, as studied in the literature, it is possible to consider additional NP (non-parametric marginals and parametric copulas) and NN (non-parametric marginals and non-parametric copulas) to complete the gap. Besides, it can be aimed to decide the best model based on a benchmark cases on a certain location and choosing the most plausible model for detecting rare events. In a different direction, return period values can be considered together but for simplicity, this paper just focused on the main exceedance probability calculations. Regarding the use of weather variables in the index construction step, only three variables were taken into account but computationally, it is possible to add more variables in general, whereas the degree of being explainable can be decreased in this case. Additionally, looking at other pairs of drought characteristics (i.e. Duration-Mean Intensity), adding more weather stations from the same region and considering other type of drought indices (satellite based etc.) can polish the presented empirical findings. In order to get a more comprehensive understanding of and mitigate the impacts of drought, the first future direction on the author's agenda is to reduce the above-mentioned limitations. In a different setting, considering different threshold levels while deriving the drought characteristics can quantify the changes across different drought levels in depth. The generation of multiple drought indices relying on multiple sources can add more novelty to the drought characteristics analysis by following the work of Farahmand and AghaKouchak (2015) and the corresponding references therein. It is possible to explore the application of these indices in different climatic regions and consider integrating other relevant variables, such as soil moisture or vegetation indices, to further refine drought assessments. In a different direction, instead of a simplified parametric copula approach, copulas with varying dependence parameters based on specific weather variables can play a significant role in the drought index construction, that directly impacts the drought characteristic behaviors. Besides, the construction of the drought index can be enhanced with the help of the wavelet filtering concept by considering the short- or medium-term fluctuations in the weather variables. Since this idea is widely considered in financial market shocks, a similar analogy can be derived in the field of drought studies to capture the behaviors of weather variables at different frequency levels as long as there are enough weather data records. These ideas lie at the top of the list of potential future studies.

Competing interests

No competing interest is declared.

Author contributions statement

Ozan Evkaya led the conceptualization, formulated the primary research questions, and performed the main statistical analyses on drought patterns. Hongyi Lu worked on the statistical methodology, contributed to software tools for analysis, and created the main findings, including the summary tables and visualizations. Both authors collaborated closely in data acquisition and pre-processing, ensuring robustness in the study's results.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions to improve our research.

Data availability

The pre-processed data that motivated this work is available upon request.

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2):182–198.

Abdul Rauf, U. F. and Zeephongsekul, P. (2014a). Analysis of rainfall severity and duration in victoria, australia using non-parametric copulas and marginal distributions. *Water resources management*, 28:4835–4856.

Abdul Rauf, U. F. and Zeephongsekul, P. (2014b). Copula based analysis of rainfall severity and duration: a case study. *Theoretical and applied climatology*, 115:153–166.

Alahacoon, N. and Edirisinghe, M. (2022). A comprehensive assessment of remote sensing and traditional based drought monitoring indices at global and regional scale. *Geomatics, Natural Hazards and Risk*, 13(1):762–799.

Beguería, S. and Vicente-Serrano, S. M. (2017). *SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index*. R package version 1.7.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Czado, C. (2019). Analyzing dependent data with vine copulas. *Lecture Notes in Statistics, Springer*, 222.

De Michele, C. and Salvadori, G. (2003). A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research: Atmospheres*, 108(D2).

Erhardt, T. M. and Czado, C. (2018). Standardized drought indices: a novel univariate and multivariate approach. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 67(3):643–664.

Evkaya, O., Yozgatligil, C., and Selcuk-Kestel, A. S. (2019). Drought analysis using copula approach: a case study for turkey. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 5(3):243–260.

Farahmand, A. and AghaKouchak, A. (2015). A generalized framework for deriving nonparametric standardized drought indicators. *Advances in Water Resources*, 76:140–145.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Hannah, D. M. (2011). Regional hydrological drought in north-western europe: linking a new regional drought area index with weather types. *Hydrological Processes*, 25(7):1163–1179.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering*, 12(4):347–368.

Haghighat jou, P., Akhoond-Ali, A. M., and Nazemosadat, M. J. (2013). Nonparametric kernel estimation of annual precipitation over iran. *Theoretical and applied climatology*, 112:193–200.

Heim, R. et al. (2000). Drought indices: a review. *Drought: a global assessment*, pages 159–167.

Hesami Afshar, M., Sorman, A. U., and Yilmaz, M. T. (2016). Conditional copula-based spatial–temporal drought characteristics analysis—a case study over turkey. *Water*, 8(10):426.

Kao, S.-C. and Govindaraju, R. S. (2007). A bivariate frequency analysis of extreme rainfall with implications for design. *Journal of Geophysical Research: Atmospheres*, 112(D13).

Kao, S.-C. and Govindaraju, R. S. (2008). Trivariate statistical analysis of extreme rainfall events via the plackett family of copulas. *Water Resources Research*, 44(2).

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

Keyantash, J. and Dracup, J. A. (2002). The quantification of drought: an evaluation of drought indices. *Bulletin of the American Meteorological Society*, 83(8):1167–1180.

McKee, T. B., Doesken, N. J., Kleist, J., et al. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, pages 179–183. Boston.

Mirabbasi, R., Anagnostou, E. N., Fakheri-Fard, A., Dinpashoh, Y., and Eslamian, S. (2013). Analysis of meteorological drought in northwest iran using the joint deficit index. *Journal of Hydrology*, 492:35–48.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts. *Journal of hydrology*, 391(1-2):202–216.

Nagler, T. (2014). Kernel methods for vine copula estimation. Master's thesis, Technische Universität München.

Nalbantis, I. (2008). Evaluation of a hydrological drought index. *European water*, 23(24):67–77.

Nelsen, R. B. (2007). *An introduction to copulas*. Springer science & business media.

Núñez, J., Verbist, K., Wallis, J., Schaefer, M., Morales, L., and Cornelis, W. (2011). Regional frequency analysis for mapping drought events in north-central chile. *Journal of Hydrology*, 405(3):352–366.

Orman Bakanligi, T. (2023). *Konya Basin Drought Management Plan*. Basins and Drought Analysis.

Pappada, R., Durante, F., and Salvadori, G. (2017). Quantification of the environmental structural risk with spoiling ties: is randomization worthwhile? *Stochastic environmental research and risk assessment*, 31:2483–2497.

Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076.

Pham, M. T., Vernieuwe, H., Baets, B. D., Willems, P., and Verhoest, N. E. (2016). Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas. *Stochastic Environmental Research and Risk Assessment*, 30:2197–2214.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The annals of mathematical statistics*, pages 832–837.

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R. (2007). *Extremes in nature: an approach using copulas*, volume 56. Springer Science & Business Media.

Serinaldi, F. (2015). Dismissing return periods! *Stochastic environmental research and risk assessment*, 29:1179–1189.

Serinaldi, F., Bonaccorso, B., Cancelliere, A., and Grimaldi, S. (2009). Probabilistic characterization of drought properties through copulas. *Physics and Chemistry of the Earth, Parts a/B/C*, 34(10-12):596–605.

Shiau, J. (2003). Return period of bivariate distributed extreme hydrological events. *Stochastic environmental research and risk assessment*, 17:42–57.

Shiau, J. (2006). Fitting drought duration and severity with two-dimensional copulas. *Water resources management*, 20:795–815.

Shiau, J.-T., Feng, S., and Nadarajah, S. (2007). Assessment of hydrological droughts for the yellow river, china, using copulas. *Hydrological Processes: An International Journal*, 21(16):2157–2163.

Shiau, J.-T. and Modarres, R. (2009). Copula-based drought severity-duration-frequency analysis in iran. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, 16(4):481–489.

Shiau, J.-T. and Shen, H. W. (2001). Recurrence analysis of hydrologic droughts of differing severity. *Journal of water resources planning and management*, 127(1):30–40.

Shukla, S. and Wood, A. W. (2008). Use of a standardized runoff index for characterizing hydrologic drought. *Geophysical research letters*, 35(2).

Simonoff, J. S. (2012). *Smoothing methods in statistics*. Springer Science & Business Media.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231.

Stambaugh, M. C., Guyette, R. P., McMurry, E. R., Cook, E. R., Meko, D. M., and Lupo, A. R. (2011). Drought duration and frequency in the us corn belt during the last millennium (ad 992–2004). *Agricultural and Forest Meteorology*, 151(2):154–162.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., et al. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, 83(8):1181–1190.

Tsakiris, G. and Vangelis, H. (2005). Establishing a drought index incorporating evapotranspiration. *European water*, 9(10):3–11.

Vazifehkhah, S. and Kahya, E. (2019). Hydrological and agricultural droughts assessment in a semi-arid basin: Inspecting the teleconnections of climate indices on a catchment scale. *Agricultural Water Management*, 217:413–425.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7):1696–1718.

Wand, M. P. and Jones, M. C. (1993). Comparison of smoothing parameterizations in bivariate kernel density estimation. *Journal of the american Statistical Association*, 88(422):520–528.

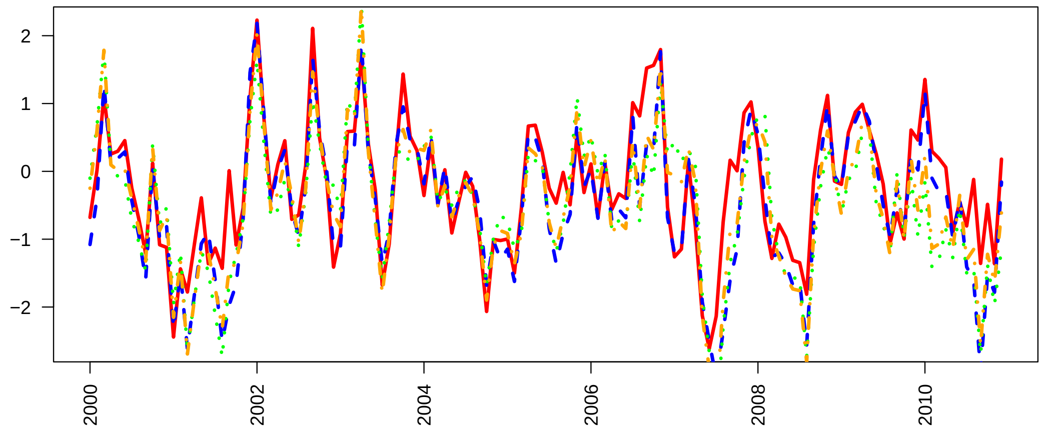
Yeo, I.-K. and Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, 87(4):954–959.

Zhang, L. and Singh, V. P. (2007). Bivariate rainfall frequency distributions using archimedean copulas. *Journal of Hydrology*, 332(1-2):93–109.

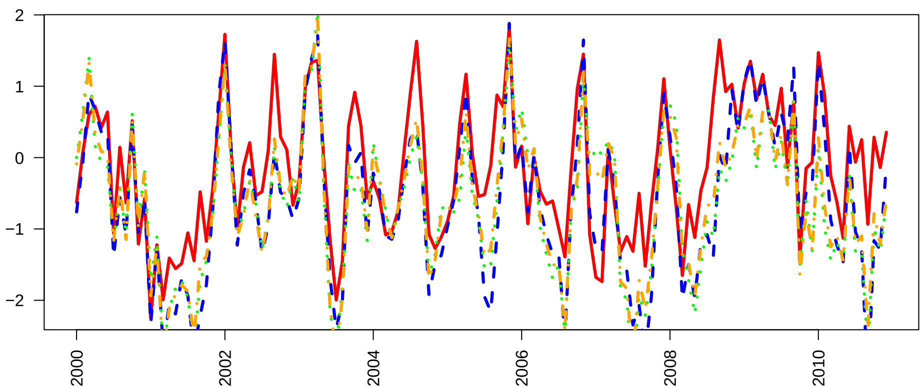
Pak.j.stat.oper.res. Vol.20 No.3 2024 pp 553-586 DOI: http://dx.doi.org/10.18187/pjsor.v20i3.4609

A. Drought Index Comparison, 3-month

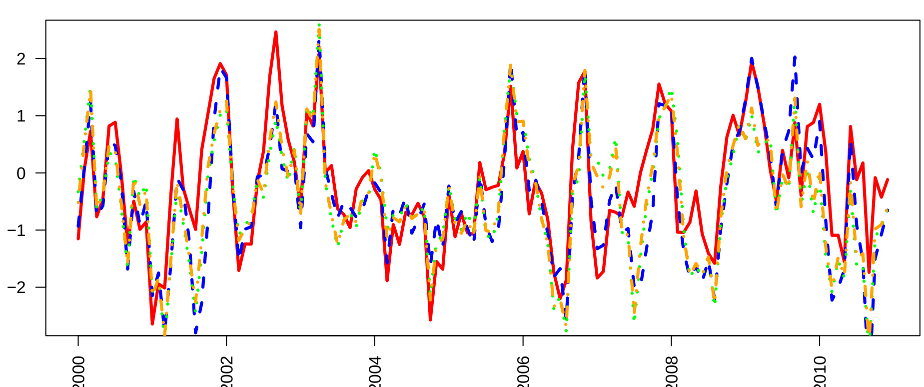
Drought Indices: SPI3 (Solid), SPEI3 (Dashed), SBI3 (Dotted), STI3 (Dash-Dotted) for different stations



(a) Akşehir

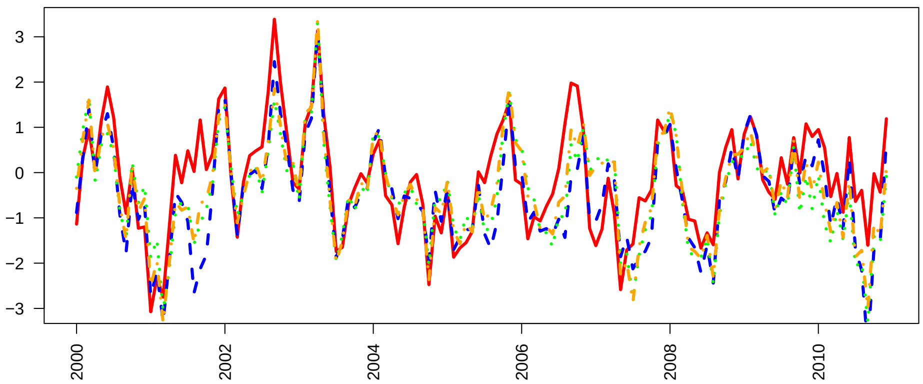


(b) Cihanbeyli

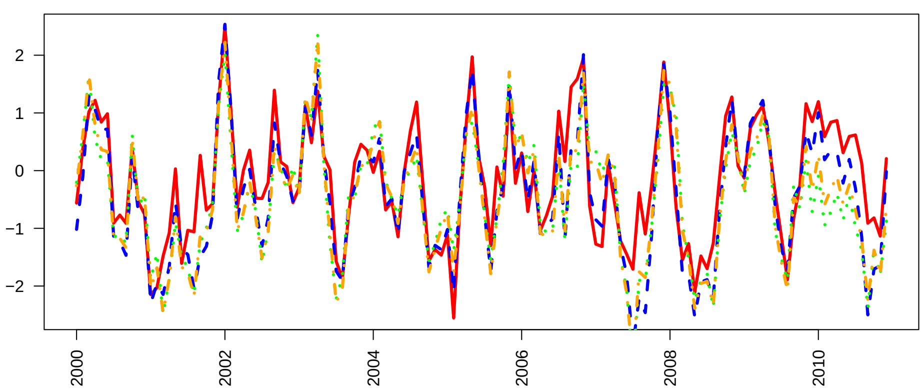


(c) Karapinar

Drought Indices: SPI3 (Solid), SPEI3 (Dashed), SBI3 (Dotted), STI3 (Dash-Dotted) for different stations



(d) Cumra



(e) Yunak

B. Cihanbeyli Station Modeling Summary

C. Karapinar Station Modeling Summary

D. Cumra Station Modeling Summary

E. Yunak Station Modeling Summary

Table 6: Summary of selected marginal distributions for drought characteristics for Cihanbeyli

	Univariate Drought Characteristics			
	D	S	PI	MI
SPI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=0.9618	shape=0.7760	shape=1.2225	shape=1.4619
par2	sdlog=0.8045	scale=2.4639	scale=1.0459	scale=0.6766
SPEI3				
Dist	Weibull	Gamma	Weibull	Weibull
par1	shape=1.4696	shape=0.7078	shape=1.2164	shape=1.3921
par2	scale=3.9487	rate=0.2439	scale=1.0805	scale=0.6814
SBI3				
Dist	Lognormal	Weibull	Exponential	Weibull
par1	meanlog=0.8905	shape=0.7053	rate=1.0751	shape=1.3201
par2	sdlog=0.8649	scale=2.1967	-	scale=0.6244
STI3				
Dist	Lognormal	Weibull	Exponential	Weibull
par1	meanlog=0.9210	shape=0.7280	rate=1.0477	shape=1.3229
par2	sdlog=0.7846	scale=2.2511	-	scale=0.6492

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity.

Table 7: Summary of fitted parametric copulas for the pairs of drought characteristics for Cihanbeyli (3-month, PP case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Family (Rotation)	BB8	Frank	Frank	Joe (180)
	θ_1 (θ_2)	8 (0.84)	5.44 (-)	8.28 (-)	10.55 (-)
SPEI3	Family (Rotation)	BB8	Gaussian	Frank	Clayton
	θ_1 (θ_2)	8 (0.78)	0.65 (-)	7.37 (-)	10.38 (-)
SBI3	Family (Rotation)	Clayton (180)	Frank	BB8	BB7
	θ_1 (θ_2)	3.45 (-)	4.95 (-)	7.78 (0.65)	2.68 (25)
STI3	Family (Rotation)	Clayton (180)	Frank	BB8	BB7
	θ_1 (θ_2)	3.59 (-)	5.01 (-)	6.97 (0.69)	3.37 (25)

BB7 is a bivariate case of Joe and Clayton copula; BB8 is a bivariate case of Joe and Frank copula.

Table 8: Summary of fitted non-parametric copulas for the pairs of drought characteristics for Cihanbeyli (3-month, PN case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6958	0.4432	0.573	0.8026
SPEI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6593	0.4255	0.5415	0.8192
SBI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6301	0.42	0.5353	0.8396
STI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6402	0.4224	0.5304	0.844

Table 9: Summary of selected marginal distributions for drought characteristics for Karapinar

	Univariate Drought Characteristics			
	D	S	PI	MI
SPI3				
Dist	Weibull	Gamma	Weibull	Weibull
par1	shape=1.2977	shape=0.7196	shape=1.2390	shape=1.4724
par2	scale=3.8744	rate=0.2600	scale=1.0693	scale=0.6865
SPEI3				
Dist	Lognormal	Lognormal	Weibull	Weibull
par1	meanlog=0.9850	meanlog=0.3250	shape=1.3878	shape=1.7391
par2	sdlog=0.8819	sdlog=1.3814	scale=1.0937	scale=0.6973
SBI3				
Dist	Lognormal	Lognormal	Weibull	Weibull
par1	meanlog=0.9038	meanlog=0.1870	shape=1.3803	shape=1.7210
par2	sdlog=0.8335	sdlog=1.3962	scale=1.0505	scale=0.6928
STI3				
Dist	Lognormal	Lognormal	Weibull	Normal
par1	meanlog=0.8744	meanlog=0.1050	shape=1.3227	mean=0.5981
par2	sdlog=0.8503	sdlog=1.4445	scale=1.0164	sd=0.3673

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity.

Table 10: Summary of fitted parametric copulas for the pairs of drought characteristics for Karapinar (3-month, PP case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Family (Rotation)	Gumbel	BB8 (180)	Frank	BB7
	θ_1 (θ_2)	3.29 (-)	3.86 (0.83)	7.37 (-)	2.13 (14.39)
SPEI3	Family (Rotation)	Clayton (180)	Frank	BB8	Gumbel (180)
	θ_1 (θ_2)	4.23 (-)	4.09 (-)	8 (0.64)	5.56 (-)
SBI3	Family (Rotation)	BB8	Frank	BB8	Gumbel (180)
	θ_1 (θ_2)	6.67 (0.92)	3.51 (-)	5.94 (0.75)	6.17 (-)
STI3	Family (Rotation)	Clayton (180)	Gaussian	BB8	BB1
	θ_1 (θ_2)	3.12 (-)	0.51 (-)	5.43 (0.75)	5.27 (1.87)

BB7 is a bivariate case of Joe and Clayton copula; BB8 is a bivariate case of Joe and Frank copula.

Table 11: Summary of fitted non-parametric copulas for the pairs of drought characteristics for Karapinar (3-month, PN case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6847	0.4192	0.5409	0.795
SPEI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6731	0.3599	0.5424	0.8021
SBI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6507	0.32	0.518	0.8113
STI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6074	0.291	0.4923	0.8234

Table 12: Summary of selected marginal distributions for drought characteristics for Cumra

	Univariate Drought Characteristics			
	D	S	PI	MI
SPI3				
Dist	Weibull	Gamma	Weibull	Weibull
par1	shape=1.3658	shape=0.7749	shape=1.3818	shape=1.6033
par2	scale=4.4160	rate=0.2440	scale=1.1628	scale=0.7179
SPEI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=0.9774	shape=0.7710	shape=1.2408	shape=1.4710
par2	sdlog=0.8075	scale=2.5574	scale=1.0274	scale=0.6818
SBI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=0.9928	shape=0.8012	shape=1.3970	shape=1.6236
par2	sdlog=0.8852	scale=2.7298	scale=1.1149	scale=0.7043
STI3				
Dist	Lognormal	Weibull	Weibull	Weibull
par1	meanlog=1.0046	shape=0.8310	shape=1.4436	shape=1.7366
par2	sdlog=0.8882	scale=2.8419	scale=1.1534	scale=0.7283

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity.

Table 13: Summary of fitted parametric copulas for the pairs of drought characteristics for Cumra (3-month, PP case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Family (Rotation)	Frank	Gumbel (180)	Frank	BB7
	θ_1 (θ_2)	12.35 (-)	1.74 (-)	7.49 (-)	2.4 (13.71)
SPEI3	Family (Rotation)	Gumbel	Frank	Frank	Clayton
	θ_1 (θ_2)	3.32 (-)	4.85 (-)	7.38 (-)	8.41 (-)
SBI3	Family (Rotation)	Gumbel	Gaussian	Frank	Clayton
	θ_1 (θ_2)	3.52 (-)	0.6 (-)	7.84 (-)	9.04 (-)
STI3	Family (Rotation)	Joe	Frank	Frank	BB7
	θ_1 (θ_2)	5.2 (-)	3.89 (-)	7.15 (-)	2.14 (10.79)

BB7 is a bivariate case of Joe and Clayton copula; BB8 is a bivariate case of Joe and Frank copula.

Table 14: Summary of fitted non-parametric copulas for the pairs of drought characteristics for Cumra (3-month, PN case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6918	0.3996	0.5436	0.7931
SPEI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6896	0.4096	0.5366	0.7909
SBI3	Kernel Type	T	T	T	T
	Kendall's τ	0.699	0.3772	0.5619	0.7973
STI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6876	0.3376	0.5187	0.765

Table 15: Summary of selected marginal distributions for drought characteristics for Yunak

	Univariate Drought Characteristics			
	D	S	PI	MI
SPI3				
Dist	Gamma	Gamma	Normal	Normal
par1	shape=2.0274	shape=0.7591	mean=1.0578	mean=0.6730
par2	rate=0.5704	rate=0.2578	sd=0.7279	sd=0.4208
SPEI3				
Dist	Weibull	Gamma	Normal	Normal
par1	shape=1.5705	shape=0.8383	mean=1.1014	mean=0.6830
par2	scale=3.9981	rate=0.2807	sd=0.7336	sd=0.4042
SBI3				
Dist	Gamma	Gamma	Normal	Normal
par1	shape=2.1362	shape=0.7722	mean=1.0197	mean=0.6416
par2	rate=0.6110	rate=0.2835	sd=0.7051	sd=0.4040
STI3				
Dist	Weibull	Gamma	Weibull	Weibull
par1	shape=1.6404	shape=0.8613	shape=1.4104	shape=1.5907
par2	scale=4.0084	rate=0.3012	scale=1.1677	scale=0.7308

Abbreviations: D for Duration, S for Severity, MI for Mean Intensity, PI for Peak Intensity.

Table 16: Summary of fitted parametric copulas for the pairs of drought characteristics for Yunak (3-month, PP case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Family (Rotation)	Gumbel	Gaussian	Frank	Clayton
	θ_1 (θ_2)	3.11 (-)	0.64 (-)	7.48 (-)	10.18 (-)
SPEI3	Family (Rotation)	Frank	Gaussian	Frank	BB7
	θ_1 (θ_2)	11.49 (-)	0.6 (-)	6.84 (-)	3.51 (15.34)
SBI3	Family (Rotation)	Gumbel	Gaussian	Frank	Gumbel (180)
	θ_1 (θ_2)	2.94 (-)	0.56 (-)	7.21 (-)	6.36 (-)
STI3	Family (Rotation)	Gumbel	Frank	Frank	BB6 (180)
	θ_1 (θ_2)	2.82 (-)	4.18 (-)	7.15 (-)	3.12 (3.19)

BB7 is a bivariate case of Joe and Clayton copula; BB8 is a bivariate case of Joe and Frank copula.

Table 17: Summary of fitted non-parametric copulas for the pairs of drought characteristics for Yunak (3-month, PN case)

		Bivariate Drought Characteristics			
		D-S	**D-MI**	**D-PI**	**S-PI**
SPI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6744	0.4013	0.5507	0.8068
SPEI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6673	0.3684	0.5203	0.7843
SBI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6525	0.342	0.5397	0.8101
STI3	Kernel Type	T	T	T	T
	Kendall's τ	0.6378	0.3595	0.5378	0.8242