



Three-dimensional copula framework for early warning of agricultural droughts using meteorological drought and vegetation health conditions

Mehdi H. Afshar, Ali Unal Şorman & Fatih Tosunoglu

To cite this article: Mehdi H. Afshar, Ali Unal Şorman & Fatih Tosunoglu (2024) Three-dimensional copula framework for early warning of agricultural droughts using meteorological drought and vegetation health conditions, Hydrological Sciences Journal, 69:5, 592-606, DOI: [10.1080/02626667.2024.2326187](https://doi.org/10.1080/02626667.2024.2326187)

To link to this article: <https://doi.org/10.1080/02626667.2024.2326187>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 10 Apr 2024.



[Submit your article to this journal](#)



Article views: 483



[View related articles](#)



[View Crossmark data](#)

Three-dimensional copula framework for early warning of agricultural droughts using meteorological drought and vegetation health conditions

Mehdi H. Afshar^a, Ali Unal Şorman^b and Fatih Tosunoglu^c

^aInstitute of Geo-Hydroinformatics, Technical University of Hamburg, Hamburg, Germany; ^bDepartment of Civil Engineering, Middle East Technical University, Ankara, Turkey; ^cDepartment of Civil Engineering, Erzurum Technical University, Erzurum, Turkey

ABSTRACT

This study develops an early warning system for crop yield (CY) failure based on meteorological drought and vegetation health conditions. The framework combines three drought indices – the Standardized Precipitation Evapotranspiration Index (SPEI), standardized Normalized Difference Vegetation Index (stdNDVI), and standardized CY (stdCY) values – using copulas. Datasets for five major wheat-producing cities in Turkey between 2000 and 2022 are used for analysis. Results indicate that the time periods used to calculate SPEI and NDVI are critical in determining agricultural drought and CY conditions. The critical threshold values for SPEI and NDVI, with a 10% probability of causing agricultural drought, are found to be ~0.28 and ~0.42, respectively. Using a three-dimensional copula model resulted in more precise CY simulations than a two-dimensional model. The validation efforts showed that all of the observed CYs fell within the simulated range, indicating the robustness of the methodology in capturing drought impacts on CY conditions.

ARTICLE HISTORY

Received 5 May 2023
Accepted 8 February 2024

EDITOR

K. Soulis

ASSOCIATE EDITOR

G. Mascaro

KEYWORDS

Drought; crop yield; SPEI; NDVI; vegetation health; copula

1 Introduction

Drought events are one of the major climatic hazard types that can significantly impact agriculture and food security (Hameed *et al.* 2020). Agricultural drought assessment is an important tool for understanding the magnitude of droughts and identifying their risks to crop yield (CY) conditions (Rojas *et al.* 2011). Such assessments and early warning systems developed using different data sources can help decision makers mitigate the negative impacts of climate extremes by providing them with timely information about the magnitude of upcoming drought events (Merz *et al.* 2020).

One practical method of agricultural drought monitoring is to combine meteorological drought indices, vegetation indices, and CY amounts to assess the impact of meteorological droughts on CYs and analyse the existing interactions among them. In recent years, the use of multiple indices simultaneously has been seen as a superior alternative to the use of a single index in drought monitoring analysis. Different studies have shown that using multiple drought indices can provide a more robust and comprehensive understanding of drought events (Khan *et al.* 2020, Afshar *et al.* 2021, 2022). Moreover, combining indices from different sources, such as meteorological drought conditions, soil moisture anomalies, vegetation health conditions, and CY amounts, can capture different aspects of agricultural droughts and provide a better representation of them (Bayissa *et al.* 2018).

Three commonly used indices for agricultural drought monitoring are the Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano *et al.* 2010), the Normalized

Difference Vegetation Index (NDVI) (Tucker 1979), and CY anomalies (Kogan 1995, Dutta *et al.* 2015, Sun *et al.* 2020, Trnka *et al.* 2020, Javed *et al.* 2021). SPEI is a meteorological drought index that uses both precipitation and evapotranspiration information to measure the severity of drought events. NDVI is a vegetation index that uses remote sensing observations to measure global warming (Vicente-Serrano *et al.* 2010) and the state and health of crops (Bezdan *et al.* 2019). CY anomalies, in turn, are used to measure the deviation of CYs from their long-term averages and can provide realistic information about the impact of agricultural drought events on crop productivity. These indices together can be used to understand the impact of drought events on agricultural production.

Copula functions are powerful tools that have recently been used more frequently in modelling the dependency structure between multiple variables (de Melo E Silva Accioly and Chiyoshi 2004), including bivariate (Shiau 2006, Reddy and Ganguli 2012, Sraj *et al.* 2015) and trivariate (Zhang and Singh 2007, Wong *et al.* 2010). Copulas are mathematical functions that are used to describe the association between variables, regardless of their marginal distributions. In the context of integrating different drought indices such as SPEI, NDVI, and CY anomalies, copulas can provide a flexible framework for quantifying the degree of association between these indices and understanding how changes in one index may impact the other indices.

Overall, probabilistic modelling of the relationships between meteorological and agricultural drought indices using copula functions has several advantages over traditional

methods such as correlation analysis or linear regression analysis (Wang *et al.* 2021a, Kamali *et al.* 2022). Copula functions can account for non-linear relationships and capture the tail dependencies among multiple variables, making them particularly beneficial in modelling complex systems such as droughts. Moreover, probabilistic approaches allow the modelling framework to represent the uncertainty within the modelling (Brigode *et al.* 2013, Huang and Fan 2021), which is important in dealing with the modelling of some variables under the influence of multiple different factors (Zampieri *et al.* 2017). On the other hand, regression models assume a linear relationship between the variables involved within the modelling framework which may not be the case in real-world scenarios such as relating CY anomalies to meteorological droughts.

While probabilistic modelling of the relationships between drought indices using copulas has several advantages over traditional regression methods and correlation analysis, some challenges also need to be considered. One of the challenges of using copulas in such a modelling framework is related to their complexities. Copula functions, in comparison to regression methods, are difficult to understand, apply, and interpret, which makes them difficult to use in practice (Afshar and Yilmaz 2017). Another challenge in using copulas is related to the availability of the number of observations considered in the analysis. Copulas, particularly those functions that require calibration and fitting efforts in their functionality process, need large, global datasets (Spinoni *et al.* 2019, Ionita and Nagavciuc 2021), and gathering such data can be difficult in the context of relating CYs with environmental factors. Moreover, copulas are computationally intensive and have high computational costs for generating uncertainties associated with relationships among meteorological, and agricultural drought indices (Hasan *et al.* 2019). Overall, however, despite the computational demands and other challenges associated with copulas in drought analyses, the opportunity they provide to accurately capture the tail dependency structure among drought indices using time series clustering (de Luca and Zuccolotto 2021) makes them a more robust and comprehensive choice over correlation analysis and regression models and leads them to generate improved drought predictions.

Several studies have used copulas in hydrological studies, such as precipitation–streamflow relationships (Deb and Kiem 2020, Jaiswal *et al.* 2020, Yisehak and Zenebe 2021), modelling the dependence between different hydrological variables (Ni *et al.* 2020, Qian *et al.* 2020), modelling the dependence between groundwater and surface water (Tahroudi *et al.* 2020), modelling maximum flows under climate extremes (Goodarzi *et al.* 2020, Wang *et al.* 2021b), the analysis of bivariate frequencies for hydrological modelling (Tosunoglu and Can 2016, Tosunoglu and Onof 2017, Wu *et al.* 2021, Zhong *et al.* 2022) and measuring drought characteristics in a drought monitoring scheme (Afshar *et al.* 2016, 2020).

In the context of agricultural drought analysis, copulas have demonstrated their utility in modelling the interdependencies between drought indices and CYs in various countries, such as Poland, the USA, Canada, and North China (Heim 2002, Quiring and Papakryiakou 2003,

Łabędzki and Bąk 2015, Liu *et al.* 2018). These studies have consistently shown that copulas offer an accurate representation of these intricate relationships. Likewise, several investigations in Europe and Spain have explored the connections between CY conditions and diverse environmental factors (Iglesias and Quiroga 2007, Iglesias *et al.* 2012), as well as climate conditions affecting drought risk (Larsen *et al.* 2013, Filipa Silva Ribeiro *et al.* 2020, Yoon *et al.* 2020, Bali and Singla 2022). These studies concluded that copulas serve as valuable tools for predicting CYs under varying environmental and drought conditions. However, there remains an evident gap in the literature, with no prior attempts to employ copulas in a three-dimensional framework to simultaneously address the interdependencies between meteorological drought, vegetation health conditions, and CY anomalies.

A significant novelty of this work lies in its pioneering approach, which introduces a novel approach to CY prediction studies. This approach harnesses copula functions to compute the anticipated CY under various drought scenarios through the use of conditional joint probability, denoted as $P(CY = cy \mid SPEI = spei \ \& \ NDVI = ndvi)$. To determine the associated probabilities of specific CY levels under varying environmental and climate conditions, this study relies on the numerical calculation of the density of conditional copula functions. This represents a departure from traditional copula applications that typically provide cumulative distribution functions (CDFs) of joint probabilities. This innovation constitutes a substantial contribution to time series prediction and holds immense promise for CY forecasting. Additionally, within the regional context, this study stands as the first of its kind in the region to predict CY as a function of meteorological drought and vegetation health conditions, highlighting its unique and distinctive nature.

Hence, the aim of this study is to develop a framework that integrates three distinct drought indices from disparate sources, namely SPEI, NDVI, and CY values, and facilitates the timely provision of early warnings regarding CY failure probabilities. In line with this objective, SPEI, NDVI anomalies, and wheat CY anomalies of Turkey's five main wheat-producing regions have been sourced and interconnected using a novel copula-based framework. Moreover, to validate the developed framework, we test the performance of the approach by comparing the simulated and observed wheat CY conditions during severe drought events that occurred in the study area throughout the study period between 2000 and 2022.

2 Methodology

2.1 Study area

The study area covers five cities – Ankara, Eskisehir, Konya, Kayseri, and Cankiri – located in the so-called Central Anatolian (CA) region of Turkey in the centre of the country (Fig. 1). The region encompasses 11 cities covering an area of approximately 65 000 km². The region is characterized by a semi-arid climate with hot summers and cold winters. The

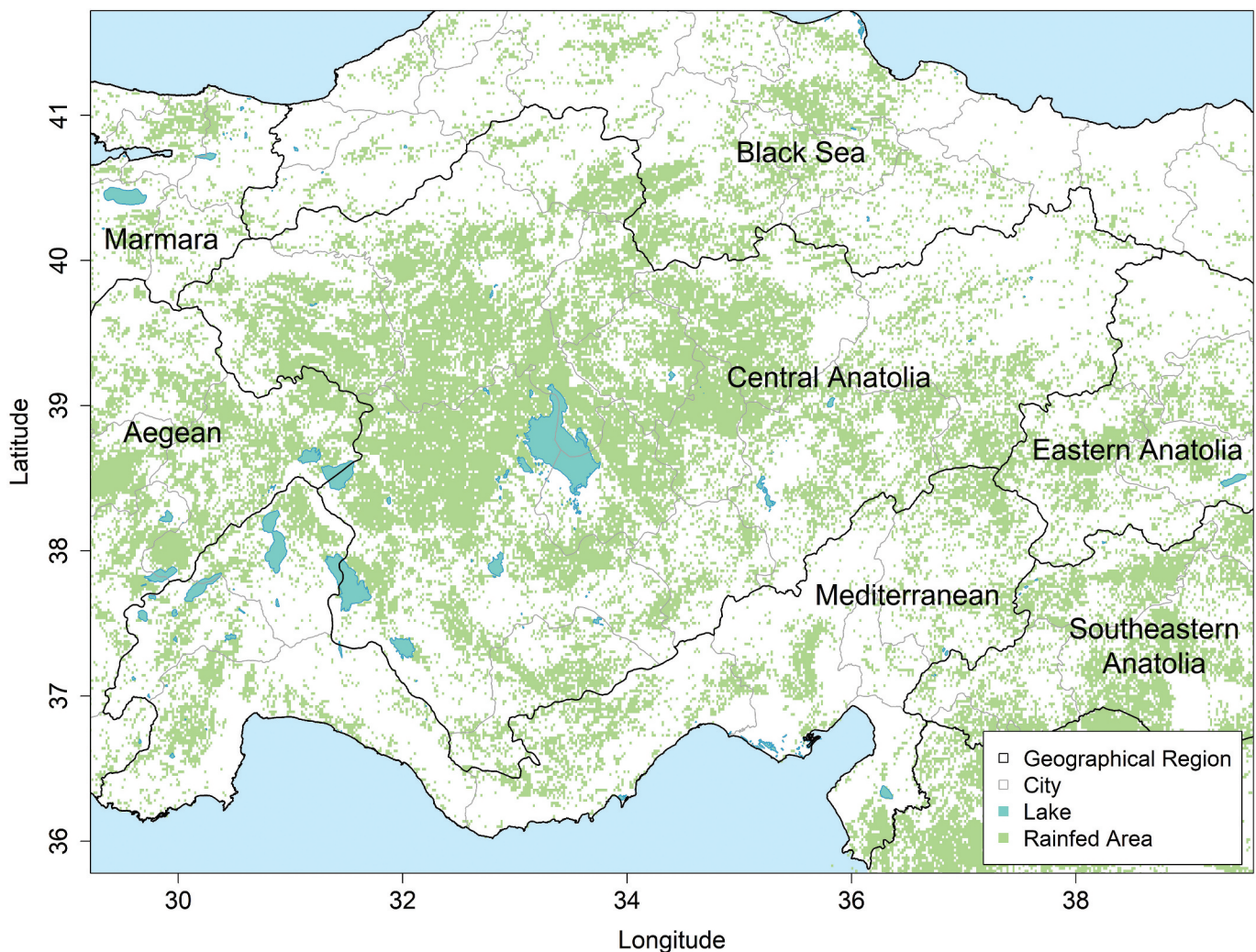


Figure 1. Location of the study area (Central Anatolia (CA) region of Turkey) and rainfed areas over it.

topography of the region is dominated by rolling plains and high plateaus, with elevations ranging from 650 to 2 600 m above sea level.

The CA region of Turkey is a significant producer of wheat and other crops such as barley, corn, and sunflower, due to its fertile soil and favourable agricultural conditions, making it a key contributor to the country's agronomic economy. However, the region is prone to drought (Hesami Afshar *et al.* 2016, Afshar *et al.* 2020, Danandeh Mehr *et al.* 2020) which can affect CYs, especially for rainfed crops. Data collected by the State Statistical Institute (TÜİK 2019) throughout 2000–2019 shows that the CA region provided around 24% of the total wheat production in Turkey, with the cities of Konya and Ankara being the largest producers, accounting for 1.9 million tons and 1.0 million tons in 2019 alone. This highlights the important role of the region in the analysis of agricultural drought, particularly in terms of wheat crop production in Turkey (Bulut 2021).

The datasets used in this study contain multiple environmental variables. Daily precipitation, temperature, and solar radiation values for the five abovementioned cities are retrieved from European Environment Agency (ERA5)

datasets between the years 1980 and 2019 to calculate the SPEI. Moreover, remotely sensed observations of Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua and Terra satellites (MOD09GA.006 and MYD09GA.006 products; (Vermote and Wolfe 2015a, 2015b) are used to calculate NDVI values over rainfed areas of the considered cities in the study using Copernicus global land cover layers (CGLS-LC100 Collection 3) datasets (Buchhorn *et al.* 2020) within the Google Earth Engine environment (Gorelick *et al.* 2017). Finally, to evaluate the CY amounts, the winter wheat CY values of these cities are collected from TÜİK, considering only rainfed products among both rainfed and irrigated statistics.

While the datasets collected from ERA5 (Hersbach *et al.* 2020) that are used in the calculation of SPEI were collected between the years 1980 and 2019 to achieve a better understanding of drought events, the NDVI and CY datasets were collected between the years 2000 and 2019, as the early years of wheat yields (between 1980 and 2000) were not steady and showed drastic changes in the products. Moreover, the available trend in NDVI and CY time series for the five considered cities is removed linearly to be more consistent with climatic conditions, and the SPEI time series.

2.2 Drought indices

Within the scope of this study, three indices coming from different sources, namely SPEI, standardized NDVI (stdNDVI), and standardized CY (stdCY) values, are integrated to evaluate agricultural droughts over the considered study area. The SPEI uses the difference between precipitation and potential evapotranspiration (PET) amounts (referred to as water balance amounts) to represent the available and under-demand water amounts and assess drought events accordingly. The SPEI values can be calculated at different time scales (accumulated balance amounts) ranging from daily to multi-year scales. The choice of time scale for SPEI depends greatly on the aim of its utilization, and while balance amounts for short periods (e.g. weekly) can be used to analyse flash droughts, the longer time scales can provide insights into seasonal dynamics and agricultural drought assessments. In this study, the SPEI has been calculated on a daily basis over a time frame that spans between specific dates. Notably, this time frame does not follow calendar months but has been tailored to align with the phenology of the wheat crop, aiming to maximize the correlation between SPEI, NDVI, and CY anomalies. For more detailed information, please refer to the final paragraph of [Section 2.2](#).

There are multiple ways to calculate PET amounts, such as Thornthwaite (Thornthwaite 1948), Food and Agriculture Organization of the United Nations (FAO)-56 methodology (Allen *et al.* 1998), or Hargreaves-Samani (Hargreaves and Samani 1985). In this study, the Penman-Monteith equation-based PET estimation using the dataset published by Singer *et al.* (2021) are utilized to calculate SPEI values over the regions. The other climatic data comprising precipitation used in the calculation of SPEI values over the regions are retrieved from ERA5 datasets. In this study, the calculation of SPEI is followed by the methodology proposed in the SPEI package (Beguería and Vicente-Serrano 2013) in R environment (R Core Team 2021), with some modifications that helped this study perform these calculations daily.

It is noteworthy that in this study, the choice of SPEI over alternative indices, such as the Standardized Soil Moisture Index (SSI), offering potentially richer insights into agricultural droughts is influenced by practical limitations in accessing daily soil moisture data. Daily datasets derived from satellite-based soil moisture, which are commonly regarded as more reliable for global soil moisture measurements (Bulut *et al.* 2019), often include gaps in their time series (Afshar *et al.* 2022). Considering our focus on daily accumulations of drought indices and dataset constraints, in this study, the SPEI is preferred to represent meteorological drought aspects of the analysis.

The next index used in this study is stdNDVI, which considers the deviation of a certain period's NDVI values from the long-term average. The NDVI is a remotely sensed vegetation index that uses red (with the electromagnetic spectrum centred at 858 nm) and near-infrared (with the electromagnetic spectrum centred at 645 nm) bands' reflectance to assess the vegetation dynamics and the healthiness of vegetation cover in a given area. The NDVI values usually vary between 0 and 1, representing the bare soil to dense vegetation cover conditions, respectively, in a specific region.

The stdNDVI values, on the other hand, can be calculated by subtracting the long-term average of NDVI values for a specific time and location from the normal NDVI values and then dividing the calculated deviation at the previous step by the long-term standard deviation of the considered NDVI values. The stdNDVI values can provide valuable information about agricultural droughts in different locations, as this normalization process removes the range of NDVI values and turns it into a normally distributed time series, making it easier to compare over different periods and locations.

In this study, the selection of NDVI over other vegetation indices was guided by its widespread utility in remote sensing applications. The NDVI's simplicity and straightforward interpretation of values make it a practical indicator for assessing the condition of vegetation around the globe. While other indices may have their unique merits, NDVI's ease of calculation, robustness, and compatibility with various sensors and platforms make it a reliable option for drought analysis.

The third index considered in this study is the stdCY. The process of stdCY calculation is very similar to the calculation of stdNDVI, with the difference that the CY values, unlike NDVI or SPEI values, are obtained annually and, hence, their standardization does not include the selection of a particular time period.

The periods for which the SPEI and stdNDVI are being calculated are found by considering the linear relationship between SPEI and stdNDVI, and stdNDVI vs. stdCY values. For this purpose, correlation analysis has been conducted to find the best window which provides the highest correlation between SPEI, stdNDVI and stdCY values over the five selected cities (Ankara, Cankırı, Eskisehir, Kayseri, and Konya) in the CA region of Turkey. The analysis has been carried out by considering different durations of the accumulated water balance (precipitation minus PET) from 1 October until May 30 (approximately equivalent to the sowing date and flowering time, respectively, of wheat crops in Turkey, mainly in CA).

2.3 Copula functions

The modelling of multivariate hydrological processes is generally achieved using joint distributions that describe the present correlation among variables. Despite the potential usefulness of using joint distribution functions, the implementation of multivariate joint distribution functions in the modelling of hydrological processes has remained limited due to the challenges posed by fitting correlated variables to the same type of distribution. To address these difficulties, copulas, a concept developed by Sklar (1959), have been widely adopted in multivariate distribution analysis. The key advantage of using copulas is the ability to dissociate dependence effects from marginal distribution effects, thereby providing greater flexibility in the selection of univariate marginal distributions (Shiau 2006).

Copulas are mathematical tools used for characterizing and analysing the complex interdependence between multiple variables. These functions provide a way to represent multivariate distributions in univariate or multivariate forms. The concept

of copulas involves linking two or more random variables, X, Y, \dots, Z through a unique function, C , that is expressed as cumulative joint distribution function $F(x, y, \dots, z)$. Considering a situation with two random variables, Sklar's Theorem states that it is a two-dimensional CDF with marginal CDFs and then there exists a copula C such that:

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)) \quad (1)$$

Under the assumption that the marginal distributions are continuous with probability density functions (PDF), the joint PDF then becomes:

$$f_{XY}(x, y) = c(F_X(x), F_Y(y))f_X(x)f_Y(y) \quad (2)$$

where c is the density function of C , defined as

$$c(x, y) = \frac{\partial^2 C(x, y)}{\partial x \partial y} \quad (3)$$

Multiple copula functions are available and each is suitable for a specific kind of study, depending on the nature of the existing relationships among the considered variables. The Archimedean (Clayton, Frank, and Gumbel) and elliptical classes of copulas (Gaussian and T) are the families commonly used in hydrological applications. In this study, since SPEI, stdNDVI and stdCY values follow a normal distribution, the Gaussian copula is used to model the joint relationships among them. The appropriate form of the copula function (true values of parameters) can be determined through methods such as error formulation or maximum likelihood estimation. In this study, to determine the existing correlation between SPEI, stdNDVI, and stdCY values, the time series of five cities are combined to create a longer time series, resulting in a more robust copula function model.

is a fundamental aspect of the analysis conducted in the study. To facilitate clarity and comprehension of the formulations and illustrations of the numerical method used in this study, some notations are employed. These notations involve the introduction of a buffer, $x + \Delta x$, enabling the computation of $P(X = x)$ as the difference between $P(x + \Delta x)$ and $P(x - \Delta x)$. In the interest of clarity, $x + \Delta x$ is referred to as x_p (representing "x positive") and $x - \Delta x$ as x_n (representing "x negative"). These notations may vary according to the nature of the equation, whether transitioning from x to y or z , depending on the presentation of bivariate or trivariate conditional probabilities.

One of the relevant probabilities in the context of agricultural droughts is related to the probability of variable X being less than x given that Y is equal to y , denoted by $P(X < x | Y = y)$. In this formulation, as an example, considering variable X to be stdCY amounts and Y to be SPEI, or stdNDVI values, would help agricultural decision makers to make informed predictions about CYs and assess the risk of crop failure when encountering meteorological drought events or critical NDVI vegetation index values. Equation (4), which computes this conditional probability, is extended to a trivariate case in Equation (5) to consider the risk of obtaining stdCY values less than threshold values when SPEI and stdNDVI values have already reached specific thresholds. It is noteworthy that in following Equations (4) – (11), the variable X can be considered to represent stdCY values, and Y , and Z , depending on the nature of problem, can be considered to represent the SPEI and stdNDVI values.

$$\begin{aligned} P(X < x | Y = y) &= \frac{P(X < x, Y = y)}{P(Y = y)} \\ &= \frac{P(X < x, Y < y_p) - P(X < x, Y < y_n)}{P(Y < y_p) - P(Y < y_n)} \end{aligned} \quad (4)$$

$$\begin{aligned} P(X < x | Y = y, Z = z) &= \frac{P(X < x, Y = y, Z = z)}{P(Y = y, Z = z)} \\ &= \frac{P(X < x, Y < y_p, Z < z_p) - P(X < x, Y < y_n, Z < z_p) - P(X < x, Y < y_p, Z < z_n) + P(X < x, Y < y_n, Z < z_n)}{P(Y < y_p, Z < z_p) - P(Y < y_n, Z < z_p) - P(Y < y_p, Z < z_n) + P(Y < y_n, Z < z_n)} \end{aligned} \quad (5)$$

2.4 Univariate and multi-variate copula models using conditional probabilities

In this study, to analyse the interactions between drought indices and develop a framework for assessing agricultural drought events, several different formulations based on conditional probability functions and their densities have been developed.

In this manner, the approach used in this study hinges on harnessing copula functions, which have the ability to calculate joint CDFs. This capacity is essential for the purpose of calculating the conditional probabilities like $P(X = x | Y = y)$, which

where as $y_p = y + \Delta y$, $y_n = y - \Delta y$, $z_p = z + \Delta z$, $z_n = z - \Delta z$, and Δy , and Δz are small values such as 0.005 that can be added to y or z values to satisfy the numerical derivation of the conditional PDF.

Equations (4) and (5) are particularly important in determining the thresholds for SPEI and NDVI that trigger CY loss. They enable interested parties to pinpoint specific environmental conditions that significantly elevate the risk of CY failure associated with agricultural drought.

In contrast, the probabilities such as $P(X = x | Y < y)$ and $P(X < x | Y < y)$ can offer another dimension of this relationship. These conditional probabilities consider the impact of

drought across a range of SPEI values, rather than focusing on a single specific value as seen in $P(X < x|Y = y)$. The above-mentioned probabilities can inform drought analysis about the likelihood of obtaining a specific yield or a yield below a certain threshold in the context of current drought conditions. To compute these probabilities, the below equations

(Equations (6) and (8) for two-dimensional and Equations (7) and (9) for three-dimensional cases) can be employed. These pivotal equations (Equations (6) – (9)) can aid managers and also stakeholders in quantifying the relationship between expected drought conditions and CY outcomes when precise conditions are uncertain.

$$P(X = x|Y < y) = \frac{P(X = x, Y < y)}{P(Y < y)} = \frac{P(X < x_p, Y < y) - P(X < x_n, Y < y)}{P(Y < y)} \quad (6)$$

$$P(X = x|Y < y, Z < z) = \frac{P(X = x, Y < y, Z < z)}{P(Y < y, Z < z)} = \frac{P(X < x_p, Y < y, Z < z) - P(X < x_n, Y < y, Z < z)}{P(Y < y, Z < z)} \quad (7)$$

$$P(X < x|Y < y) = \frac{P(X < x, Y < y)}{P(Y < y)} = \frac{P(X < x, Y < y)}{P(Y < y)} \quad (8)$$

$$P(X < x|Y < y, Z < z) = \frac{P(X < x, Y < y, Z < z)}{P(Y < y, Z < z)} = \frac{P(X < x, Y < y, Z < z)}{P(Y < y, Z < z)} \quad (9)$$

where as $x_p = x + \Delta x$, $x_n = x - \Delta x$, $y_p = y + \Delta y$, $y_n = y - \Delta y$, $z_p = z + \Delta z$, $z_n = z - \Delta z$, and Δx , Δy , and Δz are small values such as 0.005 that can be added to x , y , or z values to satisfy the numerical derivation of the conditional PDF.

Another type of conditional probability that contributes to a more precise understanding of the relationship between stdCY amounts and SPEI or NDVI, either individually or simultaneously in a trivariate scenario, is $P(X = x|Y = y)$. This probability represents the likelihood of X being exactly equal to x given that Y is exactly equal to y . This specific conditional probability offers advantages over $P(X = x|Y < y)$ or $P(X < x|Y < y)$ as it directly considers the specific value of Y and its impact on X , rather than just considering that the overall trend of Y is less than a certain value. In cases involving more than two variables, a trivariate copula can be used to describe their interdependence. The full representation of the interdependence between the variables can be achieved through the calculation of pairwise correlations between all the variables. The considered probability can be calculated using Equations (10) and (11) for bivariate and trivariate cases, respectively:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{P(X < x_p, Y < y_p) - P(X < x_n, Y < y_p) - P(X < x_p, Y < y_n) + P(X < x_n, Y < y_n)}{P(Y < y_p) - P(Y < y_n)} \quad (10)$$

$$\begin{aligned} P(X = x|Y = y, Z = z) &= \frac{P(X = x, Y = y, Z = z)}{P(Y = y, Z = z)} \\ &= \frac{P(X < x_p, Y < y_p, Z < z_p) - P(X < x_p, Y < y_p, Z < z_n) - P(X < x_p, Y < y_n, Z < z_p) + P(X < x_p, Y < y_n, Z < z_n)}{P(Y < y_p, Z < z_p) - P(Y < y_n, Z < z_p) - P(Y < y_p, Z < z_n) + P(Y < y_n, Z < z_n)} \\ &\quad - \frac{P(X < x_n, Y < y_p, Z < z_p) - P(X < x_n, Y < y_p, Z < z_n) - P(X < x_n, Y < y_n, Z < z_p) + P(X < x_n, Y < y_n, Z < z_n)}{P(Y < y_p, Z < z_p) - P(Y < y_n, Z < z_p) - P(Y < y_p, Z < z_n) + P(Y < y_n, Z < z_n)} \end{aligned} \quad (11)$$

where as $x_p = x + \Delta x$, $x_n = x - \Delta x$, $y_p = y + \Delta y$, $y_n = y - \Delta y$, $z_p = z + \Delta z$, $z_n = z - \Delta z$, and Δx , Δy , and Δz are small values such as 0.005 that can be added to x , y , or z values to satisfy the numerical derivation of the conditional PDF. Equations (10) and (11) play a pivotal role in predicting CYs and other desired time series. They can quantify the probability of achieving a specific CY under predetermined environmental conditions; and given the time frame of data collection for independent indices in this study, this predictive capacity can also serve as an effective early warning system, potentially assisting in various applications beyond agriculture.

2.5 Validation

In this study, several different conditional probability formulations have been used to find the relationships between SPEI, stdNDVI values, and stdCY amounts. To validate these formulations, the latest considered scenario (i.e. Equations (10) and (11)) is used to simulate the seven drought events ranging from moderate to severe drought types that occurred over the region between the years 2000 and 2022. Using Equations (10) and (11) and conditioning stdCY values on the occurrence of SPEI and stdNDVI, during validation efforts, we attempted to simulate ranges of CY values and find their associated conditional probabilities, and, finally, to compare the most probable CY values with the observed CY amount during the selected drought event. It is noteworthy here that for easier

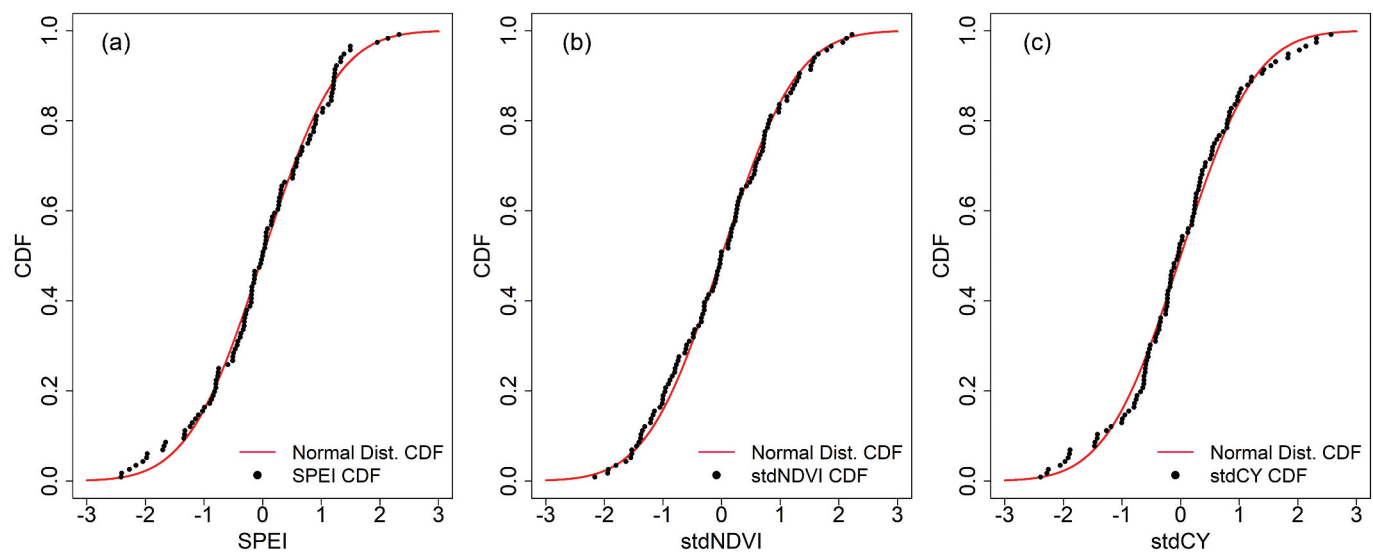


Figure 2. Comparison of the cumulative distribution function (CDF) of the normal distribution and different drought indicators of (a) Standardized Precipitation Evapotranspiration Index (SPEI); (b) standardised Normalized Difference Vegetation Index (stdNDVI); and (c) standardised Crop Yield (stdCY).

understanding of CY conditions and relating them to SPEI and NDVI values, in the Results and discussion section, the real values of NDVI and CY amounts are used instead of their standardized values to provide a better indication of their conditions.

3 Results and discussion

The fitting of two- and three-dimensional Gaussian copulas to the datasets was performed to investigate the dependency structure between the drought indices (SPEI, stdNDVI, and stdCY values) and to create a framework for providing early warning of CY failure probabilities.

The results from this correlation analysis showed that a window with a duration of 23 days (between 11 May and 3 June, approximately overlapping with the period in which NDVI time series reach their peak values over the region) for retrieving NDVI values in it can provide the highest correlation between averaged stdNDVI over the desired window and stdCY values (correlation value equal to 0.6). Moreover, the results of correlation analyses between SPEI and stdNDVI values suggested that the accumulated daily water balances for the calculation of SPEI would provide a better correlation with NDVI if they were calculated between 28 December and 23 May (correlation value equal to 0.75). Hence, for linking these three components (i.e. SPEI, stdNDVI, and stdCY values) over the CA region, the SPEI time series are calculated considering accumulated water balance values between 28 December and 23 May of the following year, the NDVI values are averaged between 11 May and 3 June of each year, and CY values have been obtained in a yearly based rotation (single crop season per year).

A more in-depth exploration of correlations between SPEI, stdNDVI, and stdCY values reveals the importance of the seasonal dynamics of agricultural processes and environmental factors. For instance, the strong correlation between SPEI and stdNDVI indicates how soil moisture, as represented by SPEI, influences the vegetation dynamics captured by NDVI. This observation aligns with the region's characteristics, as the study area (CA)

generally receives sufficient solar radiation during the wheat growing season and often experiences water deficit as the primary limiting factor. Furthermore, an analysis of the temporal lag between meteorological drought events and their impact on CYs can offer valuable insights into the lead time required for the effective implementation of mitigation strategies. In this context, this lead time is estimated to be approximately 1–2 months.

Since the drought indices (SPEI, stdNDVI, and stdCY) were standardized using the quantile matching method, they were expected to conform to a normal distribution. After verifying the normality of these indices by comparing their CDFs with the CDF of a normal distribution, in this study, the Gaussian copula is selected for conducting further probabilistic analysis of the interactions between SPEI, stdNDVI, and stdCY (Fig. 2).

The parameter estimates for the Gaussian copula were obtained using the maximum likelihood method, and the results showed that the copula parameters varied depending on the combination of the variables. For the SPEI and stdNDVI combination, the parameter value was estimated to be 0.74, while for SPEI and stdCY, it was found to be 0.48. For the stdNDVI and stdCY combination, the parameter value was estimated to be 0.6, and for the three-variable combination of SPEI, stdNDVI, and stdCY, the parameter value was estimated to be equal to the combination of three previously fitted copulas. These findings suggest that there is a moderate to strong dependence between the variables, which agrees with the results of the correlation analysis.

3.1 Determination of critical ranges of SPEI and NDVI values to produce standardized CYs values of less than -1 and -2

The threshold limits for the SPEI and the NDVI values are determined through the use of conditional probabilities represented in Equations (4) and (5). The analysis was conducted in two stages, firstly through the application of a two-dimensional copula model to determine the limits for each

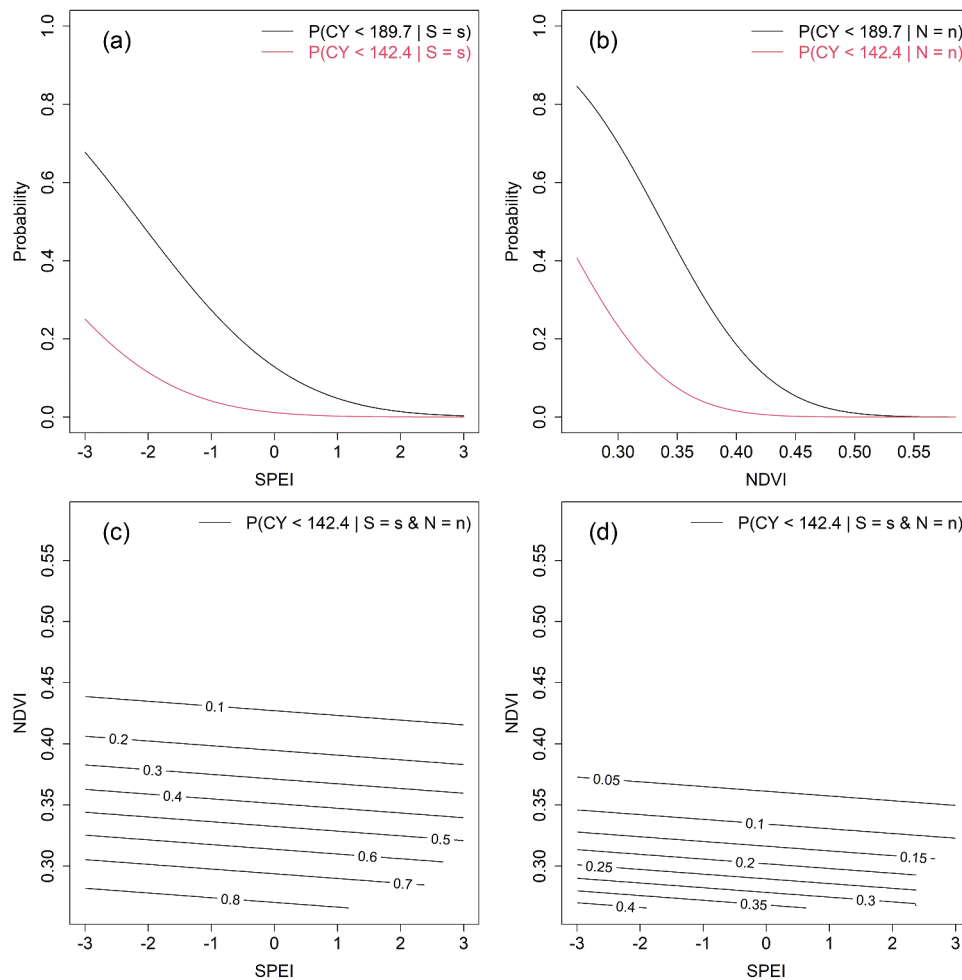


Figure 3. The probabilities of moderate and severe agricultural drought occurrences based on different ranges of SPEI and NDVI values (the probabilities are calculated using Equations (4) and 5). CY: crop yield; S: SPEI; and N: NDVI.

variable (i.e. SPEI and NDVI) individually, followed by the implementation of a three-dimensional copula model to obtain the joint range for both SPEI and NDVI. Application of this analysis resulted in the estimation of the probabilities for obtaining stdCY values less than -1 ($CY < 189.7$ kg/da equivalent to moderate agricultural drought) and less than -2 ($CY < 189.6$ kg/da equivalent to extreme agricultural drought) for ranges of SPEI and NDVI values (Fig. 3). The critical threshold for SPEI to cause agricultural drought ($CY < 189.7$ kg/da) with a 10% probability was found to be ~ 0.28 (represented with black colour in Fig. 3(a)). Similarly, the critical thresholds for NDVI that were associated with a 10% probability of CY values below 189.7 kg/da or 142.4 kg/da were found to be approximately 0.42 or 0.33, respectively, as depicted in Fig. 3(b).

Moreover, the results associated with the three-dimensional copula (trivariate) showed that the dual impact of SPEI and NDVI values can change their critical threshold values. For example, the critical threshold value for NDVI can be increased to 0.44 when the SPEI value drops below -2 (the contour with probability of 0.1 in Fig. 3(c)). Considering the probability of moderate agricultural drought occurrences for the joint happening of some values linked with NDVI and SPEI amounts together, it can be seen that with the same

SPEI values (e.g. -1), when the NDVI values are decreased from 0.43 to 0.30 (moving downward vertically), the probability of moderate agricultural drought increases from 0.1 to 0.7, while if the NDVI values are held constant and the SPEI values are decreased from $+3$ to -3 (moving leftward horizontally), the probability of moderate agricultural drought occurrence increases by 0.1 which is much less than 0.6 (for the case of keeping SPEI constant; Fig. 3(c)). A very similar trend can be also observed for severe agricultural drought conditions ($CY < 142.4$ kg/da; Fig. 3(d)); that the change in conditional probabilities is mostly vertical rather than horizontal reveals that the dominant variable in controlling agricultural drought in the region is NDVI.

The determination of critical thresholds for SPEI and NDVI presented in this study holds significant practical implications for agricultural management and decision making. Farmers and policymakers can leverage these thresholds as valuable tools to anticipate and mitigate the impacts of agricultural drought on CYs. The identified critical threshold for SPEI, approximately 0.28, implies that when this value is surpassed, there is a 10% probability of moderate agricultural drought ($CY < 189.7$ kg/da). Similarly, NDVI thresholds of 0.43 and 0.30, associated with a 10% probability of CY values below 189.7 kg/da or

142.4 kg/da, respectively, highlight the sensitivity of CY to vegetation health. Understanding these critical thresholds and obtaining them for different crop growth stages enables stakeholders to implement proactive measures, such as adjusting irrigation practices or considering alternative crops, when conditions approach these limits. Moreover, the three-dimensional copula analysis reveals the interplay between SPEI and NDVI, offering a nuanced understanding of their joint impact on agricultural drought. This insight allows for more targeted interventions, as changes in NDVI values appear to have a more pronounced effect on agricultural drought occurrences.

3.2 Discussion of copula model runs using two- and three-dimensional and their comparisons

To compare the results of two models (two- and three-dimensional copulas) in terms of expected NDVI and CY values for different levels of SPEI, the most probable values of NDVI as well as CYs are searched for using Equations (6) and (7) considering different drought scenarios (for the occurrence of SPEI to be less than 0, -0.5, -1.0, -1.5, -2.0). The results of these analyses are tabulated in Tables 1 and 2 for two- and three-dimensional cases, respectively, where the two-dimensional conditions (represented in Table 1) reflect different levels of drought scenarios, defined by SPEI values ranging from 0 to -2, and the corresponding expected NDVI and CY values with their one-standard-deviation uncertainty range. The three-dimensional conditions, on the other hand, are represented in Table 2 and reflect the impact of adding the third dimension to the analysis by integrating NDVI values with SPEI values, to predict the most probable CY.

The results show that as the SPEI values decrease and approach more severe drought conditions, the most probable

NDVI values decrease as well. The same trend is also visible over CY values with a larger decrease observed in conditions of more severe drought. Moreover, the three-dimensional conditions, incorporating both SPEI and NDVI, result in more precise predictions of CY (with narrower uncertainty range), with the CY values consistently lower than the corresponding two-dimensional conditions.

These results indicate that using a three-dimensional copula model results in lower CYs compared to a two-dimensional conditional model. Specifically, for SPEI values less than or equal to -2, the three-dimensional copula model yields approximately 170.8 kg/da, while the two-dimensional model yields 185 kg/da (the expected CY drops around 14 kg/da).

The revelation of the superiority of the three-dimensional copula model over its two-dimensional model marks a significant stride in enhancing the practical applications of agricultural drought prediction. This advanced model not only refines the precision of CY predictions but also opens a gateway to considering additional variables, such as solar radiation or other limiting factors, that play crucial roles in agricultural productivity. This is especially important in the context of meeting food demands, as the multi-dimensional approach empowers stakeholders with a more comprehensive toolkit for decision making, offering the potential to integrate a broader spectrum of variables into future analyses.

3.3 Expected most probable CYs considering a wider range of drought scenarios

For a more robust estimate of expected CYs in the CA region of Turkey under different drought scenarios, a broader range of SPEI and NDVI can be considered (Table 3). This table presents the expected CY based on values of the SPEI less than (0.0 to -2.0) and NDVI less than (0.43 to 0.32). The expected CY values are given in the table along with their ranges of uncertainty covering a one-standard-deviation boundary for the most expected CY values.

The table indicates that as the values of SPEI and NDVI decrease, the expected CY also decreases. For example, when SPEI is less than 0 and NDVI is less than 0.43, the expected CY is 213 with a range of uncertainty of 172–254. On the other hand, when SPEI is less than -2 and NDVI is less than 0.32, the expected CY is 166 with a range of uncertainty of 130–202, which is the lowest expected CY among all the combinations. It can also be observed that the range of uncertainty becomes narrower as the values of both SPEI and NDVI decrease, implying that the expected CY is more certain when both indices have lower values.

Additionally, it is worth noting here that the expected CY values remain the same for certain combinations of SPEI and NDVI values – for instance, when NDVI is less than -0.32 and SPEI is less than -1.5 or -2, the expected CY is 166 with a range of uncertainty of 130–202, implying that the expected CY is not affected by small changes in the SPEI values within a certain range.

The comparison of two- and three-dimensional copula runs are also done using a continuous range of CY values represented in Fig. 4, where the probability of CYs for the SPEI being less than -1 and -2 and NDVI being less than 0.37 and

Table 1. The expected values of Normalized Difference Vegetation Index (NDVI) and Crop Yield (CY) values considering different drought scenarios using two-dimensional copulas (Equation (6)). NDVI range and CY range show the uncertainty associated with the expected NDVI and CY values covering one standard deviation around the expected CY. S: Standardized Precipitation Evapotranspiration Index (SPEI); CY: crop yield.

Two-dimensional condition	Expected NDVI	NDVI range	Expected CY	CY range
$S < 0$	0.4	0.36–0.44	218.1	156.6–251.3
$S < -0.5$	0.38	0.34–0.42	213.4	151.9–237.1
$S < -1$	0.37	0.33–0.41	203.9	147.1–232.4
$S < -1.5$	0.35	0.31–0.39	194.5	137.7–222.9
$S < -2$	0.33	0.29–0.37	185	128.2–213.4

CY: Crop Yield; SPEI: Standardized Precipitation Evapotranspiration Index; NDVI: Normalized Difference Vegetation Index; stdNDVI: standardised Normalized Difference Vegetation Index; stdCY: standardised Crop Yield; CA: Central Anatolia.

Table 2. Three-dimensional scenarios (Equation (7)) for generating expected CYs considering different SPEI and their corresponding expected NDVI values. CY range shows the uncertainty associated with the expected CY values covering one standard deviation around the expected CY. S: SPEI, N: NDVI; CY: crop yield.

Three-dimensional condition	Expected CY (kg/da)	CY range
$S < 0$ and $N < 0.4$	203.9	164.1–243.7
$S < -0.5$ and $N < 0.38$	194.5	155.3–233.7
$S < -1$ and $N < 0.37$	189.7	151.2–228.2
$S < -1.5$ and $N < 0.35$	180.3	142.6–218.1
$S < -2$ and $N < 0.33$	170.8	133.9–207.6

Table 3. Most probable CY values (kg/da) considering different drought scenarios using (0, -0.5, -1, -1.5, -2) for both SPEI and stdNDVI indices (using Equation (7)). Values in parentheses show the uncertainty associated with the expected CY values covering one standard deviation around the expected CY. S: SPEI; N: NDVI.

S/N	N < 0.43	N < 0.4	N < 0.37	N < 0.35	N < 0.32
S < 0	213 (172–254)	204 (164–244)	194 (156–233)	180 (143–218)	171 (135–207)
S < -0.5	209 (168–250)	199 (159–239)	194 (156–233)	180 (143–218)	171 (135–207)
S < -1	204 (163–245)	199 (159–239)	190 (151–228)	180 (143–218)	171 (135–207)
S < -1.5	194 (154–235)	190 (150–230)	185 (146–224)	176 (138–213)	166 (130–202)
S < -2	185 (145–225)	185 (145–225)	180 (142–219)	176 (138–213)	166 (130–202)

0.32 are explored using two-dimensional copulas and combination of these considerations in three-dimensional copulas. The cumulative probabilities represented in the bottom panel of Fig. 4(b) show that when the expected CY is found to be below 189.7 kg/da (StdCY < -1), the cumulative joint probability computed using Equation (9) ranges from 0.5 to 0.6 with a mean of 0.55 for NDVI values less than 0.37, and from 0.71 to 0.74 with a mean of 0.72 for NDVI values less than 0.32. On the other hand, when the expected CY is found to be below 142.4 kg/da, (stdCY < -2) the cumulative probability drops to a mean range of 0.15 to 0.27 with interval values of 0.12 to 0.18 and 0.25 to 0.29 for NDVI values of less than 0.37 and 0.32, respectively. The cumulative probabilistic results to obtain an estimate of CYs are tabulated in

Table 4 as well to present a better understanding of the above discussion considering both changes in SPEI (<-1 and <-2) and NDVI (<0.37 and <0.32) together using the three-dimensional model.

Overall, the results represented in Fig. 4(a) and (b) show outcomes consistent with the results presented in Tables 1 and 3 as well as Table 4, confirming that the expected CY values decrease with the utilization of three-dimensional copulas. These results highlight the complex relationship between meteorological and agricultural droughts and the necessity of adding more dimensions (such as adding soil moisture drought index or evapotranspiration amounts measured over the regions) into the copula simulations to give more robust results.

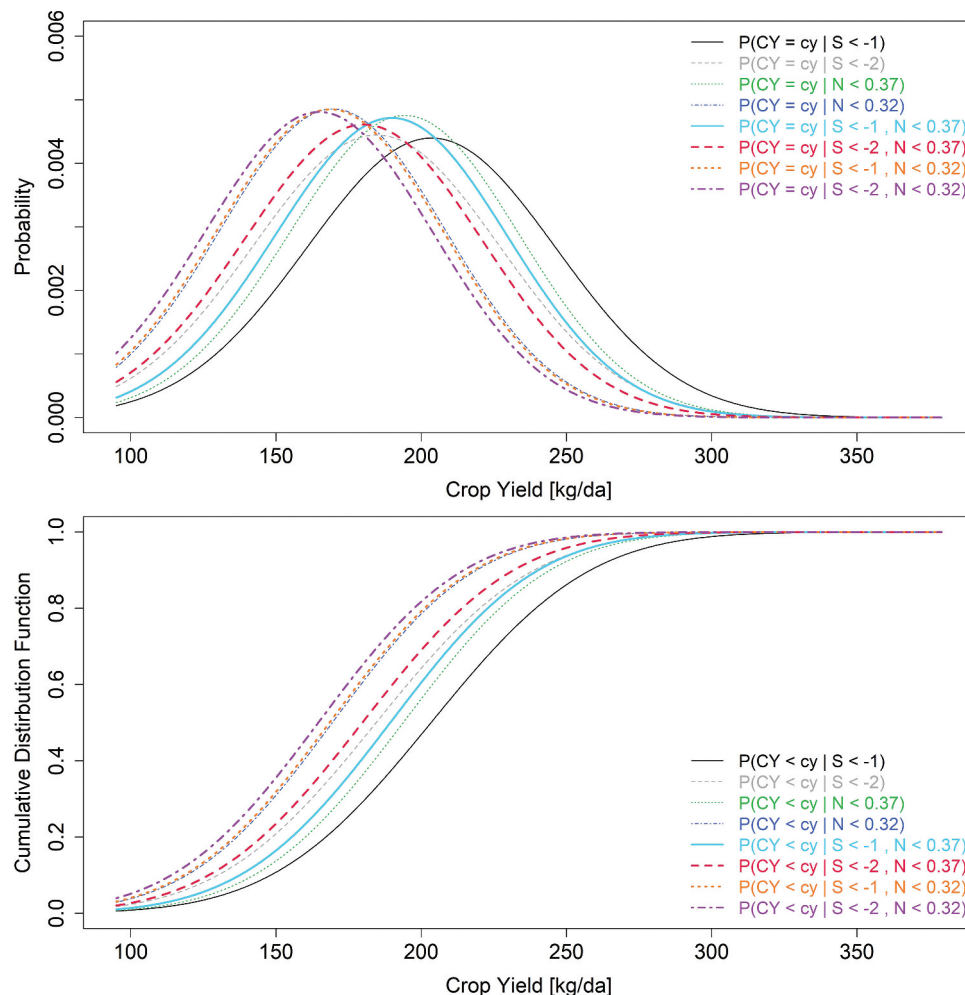


Figure 4. The PDF and CDF curves of two- and three-dimensional copula runs for a continuous range of CY values over different scenarios of SPEI less than -1 and -2 as well as NDVI values less than 0.37 and 0.32 (the probabilities are calculated using Equations (6) – 9)). PDF: probability density function, CDF: cumulative distribution function, CY: crop yield.

Table 4. The conditional three-dimensional cumulative probability values considering different drought scenarios (SPEI less than -1 and -2 , NDVI less than 0.37 and 0.32 , and CY values less than 190 and 142 kg/da). S: SPEI, N: NDVI, CY: crop yield (the probabilities are calculated using Equations (9) and (10)).

SPEI	NDVI	Conditional cumulative probability	
		$P(CY < 190 S < s, N < n)$	$P(CY < 142 S < s, N < n)$
$S < -1$	$N < 0.37$	0.50	0.12
	$N < 0.32$	0.71	0.25
$S < -2$	$N < 0.37$	0.60	0.18
	$N < 0.32$	0.74	0.29

3.4 Validation – probabilistic analysis of CYs for different ranges of drought events and comparisons with the observed records

The proposed methodology has been validated through a comparison of simulated values with observed CY values during drought years. To this aim, Equations (10) and (11) are utilized to calculate the corresponding probabilities for a range of CY values covering the observed values of SPEI and NDVI indices. The results of this comparison, presented in Table 5 and Fig. 5, provide compelling evidence of the accuracy and effectiveness of the proposed methodology.

The comparison of observed CY data with the developed framework's outputs for various drought years provides valuable insights into the model's performance and its implications (Table 5). In the year 2007, for instance, where the observed SPEI was -1.01 and NDVI was 0.38 , the most probable CY was estimated at 194.5 kg/da. However, the observed CY was lower at 174.9 kg/da, indicating a potential limitation in the model's ability to capture the true impact of drought conditions. These discrepancies highlight the need for a nuanced understanding of local factors influencing CYs during drought events. The results underscore the importance of considering additional variables, such as detailed crop type maps and higher-resolution weather data, to enhance the model's accuracy.

The validation of the proposed methodology is also presented in Fig. 5, where the CYs for the drought years of 2014 and 2021 are compared with observed CYs using both two- and three-dimensional copula model approaches. Focusing on the moderate drought year 2014, as illustrated in the top panel of Fig. 5, the observed SPEI was -1.07 , and NDVI was 0.4 . The comparison of the observed CY of 210.9 kg/da with the most probable CY through simulation results, estimated at 208.7 kg/da, indicates a close alignment, showcasing the model's capacity to accurately predict CYs during drought conditions. The consistent agreement between the observed and the most probable CYs during the 2021 drought year

further supports the robustness of the model. These results highlight the potential of the proposed methodology as a reliable tool for early warning systems and decision making processes in agricultural drought management, providing farmers and policymakers with valuable insights for more informed and timely actions.

The findings of this study are consistent with previous studies (such as Huang *et al.* 2014 and Nagy *et al.* 2021) which also found that the best time for reliable estimation of NDVI values is around the flowering time of crops. The research of Serinaldi *et al.* (2009) highlights the necessity of using higher-order dimensional copulas to correctly model the stochastic structure of variables, which is also reflected in the current study's use of a three-dimensional copula for generating more robust estimations of expected CYs. The results of this study confirm the outcomes of Li *et al.* (2021), which used vine copula and linked SPEI-NDVI and found similar results with the current study in terms of correlation analysis. These findings highlight the potential utility of copulas and demonstrate the importance of considering multiple factors in better development of copula models for agricultural drought management efforts.

Ribeiro *et al.*'s (2019) research on wheat CYs, similar to the present study, identified SPEI and vegetation condition index (derived from NDVI) as the dominant indicators within a two-dimensional copula framework. However, the current study takes this a step further by utilizing a three-dimensional copula to generate more robust estimations of expected CYs. This approach is particularly useful for developing early warning systems for agricultural drought management and forecasting the probabilities of different ranges of CY conditions based on observed SPEI and NDVI conditions a few months before harvest time. This differs from previous studies such as Fang *et al.* (2019) and Li *et al.* (2021), which used probability values, as the current study employs the density of conditional copulas and considers various conditions of SPEI and NDVI values to generate expected CYs amounts. Overall, the present study contributes to the ongoing effort to enhance agricultural drought forecasting methodologies for improved food security and sustainable agriculture.

Although the results of the current study present valuable insights into the early detection of agricultural drought using a combination of SPEI, NDVI, and CY data, several limitations should be noted. Firstly, the study did not include a detailed crop type map, which is a crucial factor in accurately assessing the impact of drought on CY. The land cover classification used for

Table 5. The expected CY values and their associated uncertainty range (covering one standard deviation around the expected CY) considering three-dimensional copula models using observed SPEI and NDVI values for the recorded droughts in the CA region. CY: crop yield.

Drought year	Observed SPEI	Observed NDVI	Expected CY (kg/da)	Expected CY range (kg/da)	Observed CY (kg/da)
2001	-0.71	0.41	222.9	185.2–260.6	203.9
2007	-1.01	0.38	194.5	157.4–231.6	174.9
2008	-1.13	0.39	199.2	161.9–236.5	243.9
2013	-0.43	0.44	256	218.3–293.7	275.7
2014	-1.07	0.4	208.7	171.2–246.2	210.9
2018	-0.52	0.43	241.8	204.1–279.5	240.7
2021	-2.29	0.38	189.7	152.7–226.7	181.7
2022	-1.88	0.38	194.5	157.3–231.7	218.4

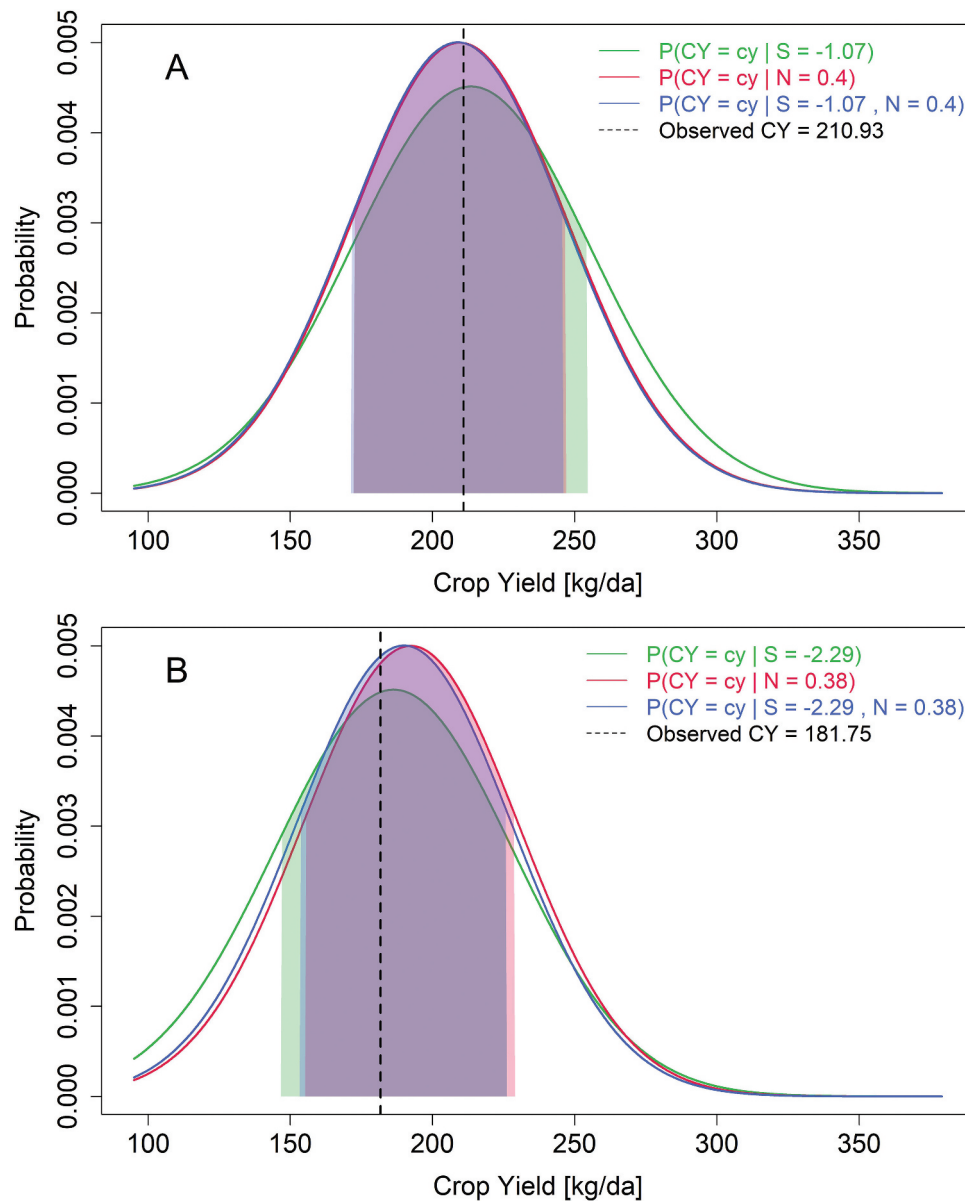


Figure 5. Comparison of observed and expected CY values under two drought conditions occurred in the years (a) 2014 and (b) 2021. Shaded areas cover one standard deviation around the expected CY. The probabilities for the range of CY values are calculated using Equations (10) and (11). CY: crop yield, S: SPEI, and N: NDVI.

determining rainfed areas may not necessarily be representative of winter wheat crops, leading to potential inaccuracies in the analysis. Secondly, the weather data used in this study had a resolution of 0.25° , which may not be sufficient for such an analysis. A denser station dataset with higher resolution could provide more accurate and reliable results. Finally, the reliability of CY values is crucial for the development of an effective early warning system. For example, in this study, there were instances where both SPEI and stdNDVI indicated severe drought, while the stdCY values did not reflect drought conditions. Such discrepancies add uncertainty to the analysis and increase the error rate of developed early warning systems. Thus, future research is needed to address these limitations to improve the accuracy and reliability of early detection of agricultural drought.

4 Concluding remarks

Agricultural drought is highly influenced by a multitude of factors, such as meteorological, hydrological, and vegetation health conditions. This research aimed to investigate the effect of various drought factors on CY, considering the temporal variability of these factors, and to develop an early warning system for CY conditions through the implementation of a conditional probabilistic approach with two- and three-dimensional analyses.

To this aim, this study utilized copula functions to combine three drought indices (SPEI, stdNDVI, and stdCY values) between the years 2000 and 2022 in Turkey and create a framework for providing early warning and information regarding CY failure probabilities based on meteorological drought and vegetation health conditions.

The main outcomes of the study suggest that the timing and threshold limits play an important role in determining agricultural drought conditions and CYs. In this study, the window between 11 May and 3 June was identified as the best period for retrieving NDVI values to obtain the highest correlation with stdCY values (0.60), while the time frame between 28 December and 23 May was found to be the best for calculating SPEI with the highest correlation with stdNDVI values (0.75). Moreover, the critical threshold values for SPEI and NDVI for causing agricultural droughts with 10% probability were determined to be ~ 0.28 and ~ 0.42 respectively. Additionally, the results of this study also highlighted that the dual impact of SPEI and NDVI values can change their critical threshold values, with NDVI being the dominant variable in controlling agricultural drought.

This study also found that utilizing a three-dimensional copula model results in more precise CY simulations than a two-dimensional conditional model, providing CYs with narrower uncertainty ranges. Furthermore, the validation efforts of this study through a comparison of simulated CY values with observed CY values during drought years showed that all of the observed CY amounts fell within the simulated expected range, demonstrating the robustness of the methodology in capturing the impact of drought conditions on CY.

Overall, the outcome of this study provides an important toolkit for decision makers to measure the impact of meteorological drought and vegetation health conditions during critical stages of crop growth and develop early warning systems and preventative measures during drought events to mitigate their effects in meeting food demands.

Acknowledgements

We thank the Turkish Statistical Institute for the wheat CY values, ECMWF for the ERA5 datasets, the ESA Climate Change Initiative Consortium for the land cover classification, and the Land Processes Distributed Active Archive Center (LP DAAC; "lpdaac.usgs.gov") for providing access to the MODIS data used in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Mehdi H. Afshar  <http://orcid.org/0000-0002-4411-3299>

References

- Afshar, M.H., Al-Yaari, A., and Yilmaz, M.T., 2021. Comparative evaluation of microwave L-Band VOD and optical NDVI for agriculture drought detection over Central Europe. *Remote Sensing*, 13 (7), 1251. doi:10.3390/rs13071251.
- Afshar, M.H., et al., 2022. Global spatiotemporal consistency between meteorological and soil moisture drought indices. *Agricultural and Forest Meteorology*, 316, 108848. doi:10.1016/j.agrformet.2022.108848
- Afshar, M.H., Sorman, A., and Yilmaz, M.T., 2016. Conditional copula-based spatial-temporal drought characteristics analysis—a case study over Turkey. *Water (Basel)*, 8 (10), 426. doi:10.3390/w8100426.
- Afshar, M.H., et al., 2020. Climate change impact assessment on mild and extreme drought events using copulas over Ankara, Turkey. *Theoretical and Applied Climatology*, 141 (3–4), 1045–1055. **Under Rev.** doi:10.1007/s00704-020-03257-6.
- Afshar, M.H. and Yilmaz, M.T., 2017. The added utility of nonlinear methods compared to linear methods in rescaling soil moisture products. *Remote Sensing of Environment*, 196, 224–237. doi:10.1016/j.rse.2017.05.017
- Allen, R.G., et al., 1998. Crop evapotranspiration - guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. *Irrigation and Drainage*. doi:10.1016/j.eja.2010.12.001.
- Bali, N. and Singla, A., 2022. Emerging trends in machine learning to predict crop yield and study its influential factors: a survey. *Archives of Computational Methods in Engineering*, 29 (1), 95–112. doi:10.1007/s11831-021-09569-8.
- Bayissa, Y.A., et al., 2018. Developing a satellite-based combined drought indicator to monitor agricultural drought: a case study for Ethiopia. *GIScience & Remote Sensing*, 56 (5), 718–748. doi:10.1080/15481603.2018.1552508.
- Beguieria, S. and Vicente-Serrano, S.M., 2013. SPEI: calculation of the standardised precipitation-evapotranspiration index. *R Package Version*, 1 (6).
- Bezdan, J., et al., 2019. SPEI-based approach to agricultural drought monitoring in Vojvodina region. *Water*, 11 (7), 1481. doi:10.3390/w11071481.
- Brigode, P., Oudin, L., and Perrin, C., 2013. Hydrological model parameter instability: a source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology (Amst)*, 476, 410–425. doi:10.1016/j.jhydrol.2012.11.012
- Buchhorn, M., et al., 2020. Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing*, 12 (6), 1044. doi:10.3390/rs12061044.
- Bulut, B. (2021, June) *Estimation of wheat yields by using remotely sensed and modeled agro-meteorological data-driven statistical and crop growth models*. Doctoral Program. Ankara: Middle East Technical University.
- Bulut, B., et al., 2019. Evaluation of remotely-sensed and model-based soil moisture products according to different soil type, vegetation cover and climate regime using station-based observations over Turkey. *Remote Sensing (Basel)*, 11 (16), 1875. doi:10.3390/rs11161875.
- Danandeh Mehr, A., et al., 2020. Climate change impacts on meteorological drought using SPI and SPEI: case study of Ankara, Turkey. *Hydrological Sciences Journal*, 65 (2), 254–268. doi:10.1080/02626667.2019.1691218.
- de Luca, G. and Zuccolotto, P., 2021. Hierarchical time series clustering on tail dependence with linkage based on a multivariate copula approach. *International Journal of Approximate Reasoning*, 139, 88–103. doi:10.1016/j.ijar.2021.09.004
- de Melo E Silva Accioly, R. and Chiyoshi, F.Y., 2004. Modeling dependence with copulas: a useful tool for field development decision process. *Journal of Petroleum Science and Engineering*, 44 (1–2). doi:10.1016/j.petrol.2004.02.007.
- Deb, P. and Kiem, A.S., 2020. Evaluation of rainfall-runoff model performance under non-stationary hydroclimatic conditions. *Hydrological Sciences Journal*, 65 (10), 1667–1684. doi:10.1080/02626667.2020.1754420.
- Dutta, D., et al., 2015. Assessment of agricultural drought in Rajasthan (India) using remote sensing derived Vegetation Condition Index (VCI) and Standardized Precipitation Index (SPI). *Egyptian Journal of Remote Sensing and Space Science*, 18 (1), 53–63. doi:10.1016/j.ejrs.2015.03.006.
- Fang, W., et al., 2019. Bivariate probabilistic quantification of drought impacts on terrestrial vegetation dynamics in mainland China. *Journal of Hydrology (Amst)*, 577, 123980. doi:10.1016/j.jhydrol.2019.123980.
- Filipa Silva Ribeiro, A., et al., 2020. Risk of crop failure due to compound dry and hot extremes estimated with nested copulas. *Biogeosciences*, 17 (19). doi:10.5194/bg-17-4815-2020.
- Goodarzi, M.R., Fatehifar, A., and Moradi, A., 2020. Predicting future flood frequency under climate change using copula function. *Water and Environment Journal*, 34 (S1), 710–727. doi:10.1111/wej.12572.

- Gorelick, N., *et al.*, 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. doi:10.1016/j.rse.2017.06.031
- Hameed, M., Ahmadalipour, A., and Moradkhani, H., 2020. Drought and food security in the Middle East: an analytical framework. *Agricultural and Forest Meteorology*, 281, 107816. doi:10.1016/J.AGRFORMET.2019.107816
- Hargreaves, G.H. and Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, 1 (2). doi:10.13031/2013.26773.
- Hasan, K.N., Preece, R., and Milanović, J.V., 2019. Existing approaches and trends in uncertainty modelling and probabilistic stability analysis of power systems with renewable generation. *Renewable and Sustainable Energy Reviews*, 101, 168–180. doi:10.1016/j.rser.2018.10.027
- Heim, R.R., Jr, 2002. A review of 20th-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83 (8), 1149–1166. doi:10.1017/CBO9781107298002.002.
- Hersbach, H., *et al.*, 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146 (730), 1999–2049. doi:10.1002/qj.3803.
- Hesami Afshar, M., Sorman, A., and Yilmaz, M., 2016. Conditional copula-based spatial-temporal drought characteristics analysis—a case study over Turkey. *Water (Basel)*, 8 (10), 426. doi:10.3390/w8100426.
- Huang, J., *et al.*, 2014. Analysis of NDVI data for crop identification and yield estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (11), 4374–4384. doi:10.1109/JSTARS.2014.2334332.
- Huang, K. and Fan, Y.R., 2021. Parameter uncertainty and sensitivity evaluation of copula-based multivariate hydroclimatic risk assessment. *Journal of Environmental Informatics*, 38 (2). doi:10.3808/jei.202100462.
- Iglesias, A., *et al.*, 2012. A regional comparison of the effects of climate change on agricultural crops in Europe. *Climate Change*, 112 (1), 29–46. doi:10.1007/s10584-011-0338-8.
- Iglesias, A. and Quiroga, S., 2007. Measuring the risk of climate variability to cereal production at five sites in Spain. *Climate Research*, 34 (1), 47–57. doi:10.3354/cr034047.
- Ionita, M. and Nagavciuc, V., 2021. Changes in drought features at the European level over the last 120 years. *Natural Hazards and Earth System Sciences*, 21 (5). doi:10.5194/nhess-21-1685-2021.
- Jaiswal, R.K., Ali, S., and Bharti, B., 2020. Comparative evaluation of conceptual and physical rainfall-runoff models. *Applied Water Science*, 10 (1). doi:10.1007/s13201-019-1122-6.
- Javed, T., *et al.*, 2021. Performance and relationship of four different agricultural drought indices for drought monitoring in China's mainland using remote sensing data. *Science of the Total Environment*, 759, 143530. doi:10.1016/j.scitotenv.2020.143530.
- Kamali, B., *et al.*, 2022. Probabilistic modeling of crop-yield loss risk under drought: a spatial showcase for sub-Saharan Africa. *Environmental Research Letters*, 17 (2), 024028. doi:10.1088/1748-9326/ac4ec1.
- Khan, M.I., *et al.*, 2020. Assessment of spatiotemporal characteristics of agrometeorological drought events based on comparing standardized soil moisture index, standardized precipitation index and multivariate standardized drought index. *Journal of Water and Climate Change*, 11 (1S), 1–17. doi:10.2166/wcc.2020.280.
- Kogan, F.N., 1995. Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin - American Meteorological Society*, 76 (5), 655–668. doi:10.1175/1520-0477(1995)076<0655:DOTLIT>2.0.CO;2.
- Łabędzki, L. and Bąk, B., 2015. Meteorological and agricultural drought indices used in drought monitoring in Poland: a review. *Meteorology Hydrology and Water Management*, 2 (2), 3–14. doi:10.26491/mhwm/34265.
- Larsen, R., *et al.*, 2013. The use of copulas in explaining crop yield dependence structures for use in geographic diversification. *Agricultural Finance Review*, 73 (3), 469–492. doi:10.1108/AFR-02-2012-0005.
- Li, H.W., *et al.*, 2021. Quantifying effects of compound dry-hot extremes on vegetation in Xinjiang (China) using a vine-copula conditional probability model. *Agricultural and Forest Meteorology*, 311, 108658. doi:10.1016/j.agrformet.2021.108658
- Liu, X., *et al.*, 2018. Performance of different drought indices for agriculture drought in the North China Plain. *Journal of Arid Land*, 10 (4), 507–516. doi:10.1007/s40333-018-0005-2.
- Merz, B., *et al.*, 2020. Impact forecasting to support emergency management of natural hazards. *Reviews of Geophysics*, 58 (4). doi:10.1029/2020RG000704.
- Nagy, A., *et al.*, 2021. Wheat yield forecasting for the tisa river catchment using landsat 8 ndvi and savi time series and reported crop statistics. *Agronomy*, 11 (4), 652. doi:10.3390/agronomy11040652.
- Ni, L., *et al.*, 2020. Vine copula selection using mutual information for hydrological dependence modeling. *Environmental Research*, 186, 109604. doi:10.1016/j.envres.2020.109604.
- Qian, L., Wang, X., and Wang, Z., 2020. Modeling the dependence pattern between two precipitation variables using a coupled copula. *Environmental Earth Sciences*, 79 (21). doi:10.1007/s12665-020-09233-7.
- Quiring, S.M. and Papakryiakou, T.N., 2003. An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and Forest Meteorology*, 118 (1–2), 49–62. doi:10.1016/S0168-1923(03)00072-8.
- R Core Team. 2021. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: <https://www.R-project.org/>
- Reddy, M.J. and Ganguli, P., 2012. Bivariate flood frequency analysis of upper Godavari River flows using archimedean copulas. *Water Resources Management*, 26 (14), 3995–4018. doi:10.1007/s11269-012-0124-z.
- Ribeiro, A.F.S., *et al.*, 2019. Copula-based agricultural drought risk of rainfed cropping systems. *Agric. Water Manag.*, 223, 105689. doi:10.1016/j.agwat.2019.105689
- Rojas, O., Vrieling, A., and Rembold, F., 2011. Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment*, 115 (2), 343–352. doi:10.1016/j.rse.2010.09.006.
- Serinaldi, F., *et al.*, 2009. Probabilistic characterization of drought properties through copulas. *Physics and Chemistry of the Earth*, 34 (10–12). doi:10.1016/j.pce.2008.09.004.
- Shiau, J.T., 2006. Fitting drought duration and severity with two-dimensional copulas. *Water Resources Management*, 20 (5), 795–815. doi:10.1007/s11269-005-9008-9.
- Singer, M.B., *et al.*, 2021. Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present. *Science Data*, 8 (1). doi:10.1038/s41597-021-01003-9.
- Sklar, A., 1959. *Fonctions de Répartition à n Dimensions et Leurs Marges*. Paris, France: Publications de l'Institut de Statistique de L'Université de Paris.
- Spinoni, J., *et al.*, 2019. A new global database of meteorological drought events from 1951 to 2016. *Journal of Hydrology, Regional Studies*, 22. doi:10.1016/j.ejrh.2019.100593.
- Sraj, M., Bezak, N., and Brilly, M., 2015. Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrological Processes*, 29 (2), 225–238. doi:10.1002/hyp.10145.
- Sun, B., *et al.*, 2020. Comparison and evaluation of remote sensing indices for agricultural drought monitoring over Kazakhstan. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43. 10.5194/isprs-archives-XLIII-B3-2020-899-2020.
- Tahroudi, M.N., *et al.*, 2020. Analyzing the conditional behavior of rainfall deficiency and groundwater level deficiency signatures by using copula functions. *Hydrology Research*, 51 (6), 1332–1348. doi:10.2166/nh.2020.036.
- Thorntwaite, C.W., 1948. An approach toward a rational classification of climate. *Geographical Review*, 38 (1), 55. doi:10.2307/210739.
- Tosunoglu, F. and Can, I., 2016. Application of copulas for regional bivariate frequency analysis of meteorological droughts in Turkey. *Natural Hazards*, 82 (3), 1457–1477. doi:10.1007/s11069-016-2253-9.

- Tosunoğlu, F. and Onof, C., 2017. Joint modelling of drought characteristics derived from historical and synthetic rainfalls: application of generalized linear models and copulas. *Journal of Hydrology, Regional Studies*, 14, 167–181. doi:10.1016/j.ejrh.2017.11.001
- Trnka, M., et al., 2020. Czech drought monitor system for monitoring and forecasting agricultural drought and drought impacts. *International Journal of Climatology*, 40 (14), 5941–5958. doi:10.1002/joc.6557.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8 (2), 127–150. doi:10.1016/0034-4257(79)90013-0.
- TÜİK. (2019) Turkish Statistical Institute - Data Portal. <https://data.tuik.gov.tr/>.
- Vanderlinden, K., Giráldez, J.V., and van Meirvenne, M., 2004. Assessing reference evapotranspiration by the hargreaves method in Southern Spain. *Journal of Irrigation and Drainage Engineering*, 130 (3), 184–191. doi:10.1061/(asce)0733-9437(2004)130:3(184).
- Vermote, E. and Wolfe, R., 2015a. MYD09GA MODIS/Aqua surface reflectance daily L2G global 1kmand 500m SIN grid V006 [Data set]. NASA EOSDIS Land Processes DAAC.
- Vermote, E. and Wolfe, R., 2015b. MOD09GA MODIS/Terra surface reflectance daily L2G global 1kmand 500m SIN grid V006 [Data set]. NASA EOSDIS Land Processes DAAC.
- 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. doi:10.1175/2009JCLI2909.1.
- Wang, L., et al., 2021a. Vegetation vulnerability and resistance to hydro-meteorological stresses in water- and energy-limited watersheds based on a Bayesian framework. *Catena (Amst)*, 196. doi:10.1016/j.catena.2020.104879.
- Wang, S., et al., 2021b. Future projections of flooding characteristics in the Lancang-Mekong River Basin under climate change. *Journal of Hydrology (Amst)*, 602, 126778. doi:10.1016/j.jhydrol.2021.126778
- Wong, G., et al., 2010. Drought analysis using trivariate copulas conditional on climatic states. *Journal of Hydrologic Engineering*, 15 (2), 129–141. doi:10.1061/(asce)he.1943-5584.0000169.
- Wu, C., et al., 2021. Copula-based risk evaluation of global meteorological drought in the 21st century based on CMIP5 multi-model ensemble projections. *Journal of Hydrology (Amst)*, 598, 126265. doi:10.1016/j.jhydrol.2021.126265.
- Yisehak, B. and Zenebe, A., 2021. Modeling multivariate standardized drought index based on the drought information from precipitation and runoff: a case study of Hare watershed of Southern Ethiopian Rift Valley Basin. *Modeling Earth Systems and Environment*, 7 (2), 1005–1017. doi:10.1007/s40808-020-00923-6.
- Yoon, D.H., et al., 2020. Agricultural drought assessment in East Asia using satellite-based indices. *Remote Sensing (Basel)*, 12 (3). doi:10.3390/rs12030444.
- Zampieri, M., et al., 2017. Wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales. *Environmental Research Letters*, 12 (6), 064008. doi:10.1088/1748-9326/aa723b.
- Zhang, L. and Singh, V.P., 2007. Gumbel–hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering*, 12 (4). doi:10.1061/(asce)1084-0699(2007)12:4(409).
- Zhong, L., et al., 2022. Identification and risk assessment of flash drought in the Pearl River basin based on the standardized evaporative stress ratio. *Theoretical and Applied Climatology*, 150 (3–4), 1513–1529. doi:10.1007/s00704-022-04228-9.