Estimating propagation probability from meteorological to ecological droughts using a hybrid machine learning-Copula method

Tianliang Jiang[1,2], Xiaoling Su[1,2], Gengxi Zhang[3], Te Zhang[1,2], Haijiang Wu[1,2]

[1] College of Water Resources and Architectural Engineering, Northwest A & F University, Yangling, Shaanxi, 712100, China

[2] Key Laboratory for Agricultural Soil and Water Engineering in Arid Area of Ministry of Education, Northwest A & F University, Yangling, Shaanxi, 712100, China

[3] College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou 225012, China

Correspondence to: Xiaoling Su (xiaolingsu@nwafu.edu.cn)

Abstract. The impact of droughts on vegetation is essentially manifested as the transition of water shortage from the meteorological to ecological stages. Therefore, understanding the mechanism of drought propagation from meteorological to ecological drought is crucial for ecological conservation. This study proposes a method for calculating the probability of meteorological drought to trigger ecological drought at different magnitudes in Northwestern China. In this approach, meteorological and ecological drought events during 1982–2020 are identified using the three-dimensional identification method; the propagated drought events are extracted according to a certain spatio-temporal overlap rule; and propagation probability is calculated by coupling the machine learning model and C-vine copula. The results indicate that: (1) 46 drought events are successfully paired by 130 meteorological and 184 ecological drought events during 1982–2020; ecological drought exhibits a longer duration, but smaller affected area and severity than meteorological drought; (2) Quadratic Discriminant Analysis (QDA) classifier performs the best among the 11 commonly used machine learning models which ~~is~~are combined with four-dimensional C-vine copula to construct drought propagation probability model; (3) the hybrid method considers more drought characteristics and more detailed propagation process which addresses the limited applicability of the traditional method to regions with large spatial extent.

1 Introduction

Drought is a multivariable and complex natural hazard with the characteristics of slow evolution, wide impact, and spatial extent (Feng et al., 2021; Wu et al., 2021; Zhang et al., 2021a; Zhang et al., 2021b). Conventionally, drought can be classified into meteorological drought, hydrological drought, agricultural drought, and socioeconomic drought. It is commonly accepted that all types of drought originate from meteorological drought (Mishra and Singh, 2010). Crausbay et al (2017) argued that existing ~~drought types~~drought types are described through a "human-centric" lens to characterize a range of effects generated by meteorological drought. This implies that the response of ecosystem to drought ~~are~~is generally ignored in policy development, which in turn ~~elicit~~elicits water use conflicts between humans and ecosystems (Zhang et al., 2021c). Ecological Drought Working Group of Science for Nature and People Partnership (SNAPP) proposed a framework of ecological drought from an "ecology-centric" lens, which incorporates ecological, meteorological, and hydrological information (Crausbay et al., 2017). Ecological drought was thus defined as an episodic deficit in water availability that drives ecosystems beyond thresholds of resilience into a vulnerable state, impacts ecosystem services, and triggers ~~feedbacks~~feedback in natural and/or human systems (Bradford et al., 2020; Crausbay et al., 2017; McEvoy et al., 2018; Munson et al., 2021; Raheem et al., 2019).

Vegetation is among the most important components in terrestrial ecosystems, and the distribution and growth of vegetation are largely influenced by meteorological factors (Wang et al., 2021; Zeng et al., 2020; Zhang et al., 2021d). Developments in remote sensing technology have facilitated the application of vegetation indices to reflect the response of vegetation to climate change (Lawal et al., 2021). For example, a simple linear relationship was found between the standardized precipitation evapotranspiration index (SPEI) and normalized difference vegetation index (NDVI) at the global scale (Vicente-Serrano et al., 2012). The correlation between SPEI and NDVI showed a positive relationship in most regions of northwestern China (NWC), with the exception of a few regions such as the western parts of the Tarim Basin, Qaidam Basin, and southeastern part of the area (Jiang et al., 2018). Actually, the impact of drought on vegetation is manifested as the transition of water shortage from the meteorological stage to the ecological stage. Therefore, this impact should be analyzed by quantifying the effect of decreasing precipitation on the variation of available ecological water, i.e., from the perspective of drought propagation.

Drought propagation refers to the transition of one drought type to another, and it is vital for drought monitoring and prediction (Fang et al., 2020; Warter et al., 2021). Accordingly, drought propagation has become a hot topic in meteorology and hydrology fields (Apurv et al., 2017; Guo et al., 2020). Approaches to drought propagation analysis are broadly divided into model simulations and statistical methods (Han et al., 2019). In the former approach, hydrological responses to meteorological drought are analyzed by using physical based models that are considered to be effective in representing relevant hydrological processes (Yu et al., 1999). Nevertheless, this approach involves labor-intensive calibration processes and is not suitable at large spatial scales (Huang et al., 2017). In contrast, statistical methods with fewer assumptions are easier to use at different spatial scales (Huang et al., 2017). However, in such methods, the propagation process was analysed using the time series of an average value of drought index in a region or subregion (explained in the Discussion section). In other words, ~~In other words,~~ the temporal connection between two drought types is ~~was~~ only considered in the traditional statistical methods, but their spatial overlap is ignored, which may result in the miscalculation of drought propagation in regions with large spatial extent.

~~Probabilistic model has been proved to be a better way to quantify the relationship between different types of droughts (Ayantobo et al.,~~ The~~2018; Chang et al., 2016; Das et al., 2020). This is due to that the~~ probability information of one type of successive drought events is contained in another type of ~~drought~~ associated ~~with it~~drought (Wu et al., 2021). Therefore, a number of studies have attempted to assess the propagation relationships between the two drought types based on the probabilistic method. A Bayesian network is a probabilistic model that acquires probabilistic inferences over interacting variables of interest based on a graphical structure. Therefore, this method has been proven to be suitable for quantifying the probability relationship between different drought types (Ayantobo et al., 2018; Chang et al., 2016; Das et al., 2020). For example, Guo et al. (2020) calculated the occurrence probability of hydrological drought based on different intervals of duration and ~~severity~~severities of meteorological drought. Sattar et al (2019) identified the occurrence probability of different classes and lag ~~time~~times of hydrological drought according to the intensity of meteorological drought. ~~Nevertheless, the number of drought characteristics considered in these studies are relatively few.~~ Xu et al. (2021) found that the probability of agricultural drought severity increased synchronously with meteorological drought in different regions of China. Jehanzaib et al. (2020) concluded that in the Korean Peninsula, the probability of meteorological drought propagating into hydrological drought increased significantly under climate change. In general, these studies primarily focused on the relationship between duration and severity between the two drought types but ignored the relationships among affected areas. Xu et al. (2015a) found that the probability of drought occurrence ~~probability~~ would be underestimated if ~~fewer~~ drought ~~characteristics were~~affected areas are not considered. Therefore, the traditional ~~drought~~ probabilistic model of drought propagation can be improved by introducing the three-dimensional ~~drought identification~~clustering method, which ~~provides~~would provide more drought information (Liu et al., 2019).

Taking a typically ecological fragile region, Northwestern China (NWC), as an example, the motivation of this study is to identify meteorological drought and ecological drought during 1982–2020 in the NWC from a three-dimensional perspective, and propose a novel method to investigate the response probability of ecological drought to meteorological drought. The remainder of the current paper is organized as follows: Section 2 briefly overviews the geographic information of NWC~~.~~ ~~Section 3~~ and describes the datasets used in this paper, and the procedure for estimating propagation probability from meteorological to ecological drought. The results and the comprehensive analysis of the proposed approach are presented in section ~~4~~3 and ~~5~~4, respectively. Finally, the conclusions are given in section ~~6~~5.

~~2~~2 Materials and methods

2.1 Study area

Northwestern China (NWC; 31°35′N–49°15′N, 73°25′E–111°15′E) includes the provinces of Shaanxi, Gansu, Qinghai, and the Autonomous Regions of Xinjiang Uyghur and Ningxia Hui, covering a total surface area of 3.1 million km^2 ~~(Fig. 1)~~(Figure 1~~Figure 1~~) (Zheng et al., 2021). The terrain of NWC constitutes mountains, basins, and the Gobi. The altitude ranges from –156 m to 6647 m, showing the characteristics of “west high and east low”. Four climatic divisions, including humid, semi-humid, semi-arid, and arid ~~area~~areas were demarcated, based on the dryness index (Zhang et al., 2021d). As NWC is located at ~~the~~ upstream of the Yangtze, Yellow, and other large rivers, it is significant to study the impact of drought on its ecosystem (Liu et al., 2021).

~~3 Materials and methods~~

~~3.1~~2.2 Datasets

Monthly meteorological data, including surface reflectance, temperature, relative humidity, atmospheric pressure, downward shortwave radiation, wind speed, and longwave radiation, was obtained from the ERA5-land reanalysis dataset (https://cds.climate.copernicus.eu) issued by the European Centre for Medium-Range Weather Forecasts (ECWMF), which has a spatial resolution of 0.1° × 0.1° and covers the period of 1981–2021. Root soil moisture data were obtained from the hydrological dataset, simulated by the Noah model of the Global Land Data Assimilation System (GLDAS, 0.25°×0.25°; https://ldas.gsfc.nasa.gov/gldas), covering the period of 1948–2021. NDVI data covering the period 1981–2021 were obtained from the National Centers for Environmental Information (NCEI) (https://www.ncei.noaa.gov/), with a spatial resolution of 0.05° × 0.05°. Land use type data (LUTD) with a spatial resolution 1 km was downloaded from China's multi-period land use/cover change monitoring dataset (http://www.resdc.cn); it includes the years of 1980, 1990, 1995, 2000, 2005, 2010, 2015, 2018, and 2020. In order to uniform the spatial resolution of ~~Root~~root soil moisture, all spatial datasets were resampled to 0.25°×0.25° using the bilinear interpolation method. The temporal range of all datasets ~~were~~was extracted from January 1982 to December 2020.

2.3~~.2~~ Meteorological and ecological drought index

Previous studies found that the standardized precipitation evaporation index (SPEI) overestimated the meteorological drought in NWC where actual atmospheric water demand is determined by precipitation variation (Ayantobo and Wei, 2019; Zhang et al., 2019a; Zhang et al., 2021b). Additionally, precipitation is the main water resources for vegetation growth in most regions of NWC due to the ~~deep phreatic buried~~great depth to groundwater (Cao et al., 2021). Standardized precipitation index (SPI) was thus used in the current study to represent meteorological drought. SPI at different time scales was calculated by aggregating *n*-month moving sums~~ (McKee et al., 1993). For example, SPI–3 in March was calculated by accumulating the series of precipitation in January, February, and March.~~, allowing the identification of various drought types (McKee et al., 1993). At short time scales, drought events are characterized by high frequency and short duration, while at long time scales, they have longer duration and lower frequency. SPI–3 has been reported to be highly representative of the impacts of meteorological conditions on vegetation as ~~the~~ vegetation variation is sensitive to ~~three months accumulated~~ precipitation accumulated over three months (McKee et al., 1993; Vicente-Serrano et al., 2012; Vicente-Serrano et al.,~~ 2010; Vicenteserrano et al.,~~ 2010). Therefore, SPI-3 was used to characterize meteorological drought in this study. Further details on SPI calculation are available in ~~(~~McKee et al~~.,~~ (1993).

Commonly used drought indices indirectly reflect the influence of drought on ecosystems, and they do not comprehensively reflect the homeostasis between ecological water consumption and requirement in drought evolution (Jiang et al., 2021). Additionally, decreases in vegetation coverage are not only caused by a persistent deficit in available water for ecosystems but also other aspects, such as wildfire, hail, flood, and human activities (Bento et al., 2020). This limited the ability of vegetation indices to reflect drought conditions. Therefore, a new drought index, the standardized ecological water deficit index (SEWDI), was constructed to monitor terrestrial ecological drought in our previous study (Jiang et al., 2021). SEWDI follows a similar procedure as SPI~~. Ecological~~, which includes the calculation of ecological water deficit (EWD~~)~~), the selection of an optimal distribution for fitting monthly EWD series, and the inverse normal transformation of the cumulative density distribution of EWD. EWD is the difference between ~~FAO-based~~ ecological water requirement (EWR) and ~~SEBS-based~~ ecological water consumption (EWC) (Chi et al., 2018; Jiang et al., 2021). Among them, EWR was calculated using the single crop coefficient method recommended by the Food and Agriculture Organization (FAO). EWC equals the actual evapotranspiration, which is derived from latent heat fluxes calculated by the surface energy balance system (SEBS) algorithm. Therefore, SEWDI ~~reflects~~can reflect the dynamics of energy and water balance under human activities and climate change. Additionally, the standardization method facilitates the same threshold and evaluation criteria in monitoring two drought types (Peng et al., 2019; Zang et al., 2020), which reduces the influence of other algorithms on final results and guarantees spatio-temporal comparability (Liu et al., 2017). The procedure in calculating SEWDI calculation is detailed in Jiang et al (2021).

~~3.3~~2.4 Drought propagation probability method

Since a reliable understanding of the drought propagation process is beneficial for drought forecasting, research interest in the probability of drought propagation from meteorological droughts to other types of droughts has been increasing (Zhou et al., 2021). The current study thus proposed a novel method coupling spatial and temporal connection method of two type droughts, with machine learning model, and C-vine copula to investigate the relationship between meteorological and ecological drought. A flow diagram of the method is depicted in ~~Fig. 2.~~Figure 2~~Figure 2~~.

Investigating the relationship between the characteristics of the two drought types is key to constructing a probability model. The approach is summarized in ~~two~~three steps as follows:

Step 1: Meteorological and ecological drought events were identified from a three-dimensional perspective, respectively (Section ~~3.3~~2.4.1).

Step2: The two drought types with a genetic relationship were paired on the basis of a certain spatio-temporal matching rule ~~to extract propagated drought events~~. Their drought characteristics, including drought affected area, drought severity, and drought duration, were calculated according to the method described in Section ~~3.3~~2.4.2.

Step 3: Taking the characteristics of meteorological drought extracted in step 2 as inputs, and propagation results as outputs, the optimal model was selected from 11 machine learning classification models to calculate the propagated probability (P1) of meteorological drought (Section ~~3.3~~2.4.3). Then, a conditional probability model of the paired meteorological and ecological drought events was constructed based on the C-vine copula (Section ~~3.3~~2.4.4). According to the severities of all identified ecological drought events, cumulative probabilities of 0.5, 0.75, and 0.9 were selected to demarcate moderate, ~~serious~~severe, and extreme drought, respectively (Guo et al., 2020). The probabilities of ecological drought at different magnitudes triggered by meteorological drought were obtained by multiplying P1 with their conditional probability

~~3.3~~2.4.1 ~~Three-dimensional drought~~Drought identification based on the three-dimensional clustering method

According to Andreadis et al (2005), the evolution of a drought event should be viewed as a spatio-temporal continuum (longitude, latitude, and time). Different from the traditional one- or two-dimensional drought identification method, the three-dimensional array of SPI-3 and SEWDI-3 were extracted to characterize the degree of meteorological and ecological droughts. The extraction procedure involves two steps ~~(Fig. 2)~~(Figure 2~~Figure 2~~) (Andreadis et al., 2005; Xu et al., 2015a; Xu et al., 2015b): firstly, the clustering method was used to identify drought patches in each month; secondly, the drought continuum was constructed by selecting the overlapping areas of drought patches between two adjacent months, which was greater than 1.6% of the total area (~~explained in~~See Section 3.1 for the ~~Discussion section).~~ reason).

For each drought event, three drought characteristics were extracted as follows: (1) affected area was calculated by cumulating the area affected by drought in each month during the entire drought period. (2) Duration denotes the time of a drought event persisted. (3) Severity is a cumulative value of SEWDI-3 or SPI-3 for the entire drought duration and areal extent, and equals ~~to~~ the volume of the three-dimensional continuum.

~~3.3~~2.4.2 Spatio-temporal connection of two drought types

Liu et al (2019) developed a new method for identifying the propagation between two related drought types based on the drought identification method from a three-dimensional perspective. The current study employed this method to identify the propagation from meteorological to ecological drought. The key to this method is the determination of the temporal and spatial connection between two drought types. The specific steps are as follows.

Firstly, the identified meteorological and ecological drought events are sorted in ~~the~~ chronological order. Secondly, whether the two drought types overlap in time is judged according to Eq. (1)–Eq. (2).

$$Overlap_t = \begin{cases} 1 \text{ if } \begin{cases} MBT_i \le EBT_j \text{ and } \min\left(MET_i, EET_j\right) - \max\left(MBT_i, EBT_j\right) \ge 2 \\ MBT_i > EBT_j \text{ and } \min\left(MET_i, EET_j\right) - \max\left(MBT_i, EBT_j\right) \ge \alpha \end{cases} \\ 0 \text{ if } MBT_i \ge EBT_j \text{ and } \min\left(MET_i, EET_j\right) - \max\left(MBT_i, EBT_j\right) < \alpha \end{cases} \tag{1}$$

$$\alpha = \min\left(\frac{MDD_i}{3}, \frac{EDD_j}{3}\right) \tag{2}$$

where 1 and 0 denote the existence and absence of time overlap between two drought types, respectively; MBT_i and EBT_j represent the beginning time of the i-th meteorological and j-th ecological drought events, respectively. Similarly, MET_i and EET_j represent the end time of the i-th meteorological and j-th ecological drought events, respectively; MDD_i and EDD_j indicate the duration of the i-th meteorological and j-th ecological drought events, respectively.

~~Secondly~~Thirdly, whether the meteorological and ecological drought patches connecting at a spatial scale is judged according to Eq. (3) and Eq. (4)

$$Overlap_s = \begin{cases} 1 \text{ if } MDA_i \cap EDA_j \geq \beta \\ 0 \text{ if } MDA_i \cap EDA_j < \beta \end{cases} \quad (3)$$

$$\beta = \max\left(1.6\% \cdot A_{NWC}, \min\left(MDA_i, EDA_j\right) \cdot b\right) \quad (4)$$

where 1 and 0 denote the existence and absence of spatial overlap between two drought types; A_{NWC} represents the total area of the NWC; MDA_i and EDA_j represent the projected area of the i-th meteorological and j-th ecological drought events, respectively. b is set as 15% in the current study (See ~~Discussion~~Section 3.1 for the reason~~)~~).

~~Thirdly~~Fourthly, successfully matched drought events are encoded following ~~the~~ chronological order. Cells in ~~Fig. 2~~Figure 2~~Figure 2~~ represent the relationship between preliminarily identified events of the two drought types. The propagation type from meteorological to ecological drought can be classified into four categories: one ecological drought event induced by one meteorological drought event (one–to–one), multiple ecological drought events induced by one meteorological drought event (one–to–many), one ecological drought event induced by multiple meteorological drought events (many–to–one), and multiple ecological drought events induced by multiple meteorological drought events (many–to–many).The codes of cells are identical if the propagation type belong to one–to–many, many–to–one, and many–to–many.

Finally, the characteristics of meteorological and ecological drought events that belong to the same paired drought event are integrated, respectively. Among them, total duration is the difference between the latest-ending and earliest-starting drought events; total affected area is the projected area of all individual drought events; total severity is the sum of severities of individual drought events.

~~3.3~~2.4.3 Drought propagation identification based on the machine learning model

The purpose of this part is to identify whether a meteorological drought event has the potential to trigger ecological drought. Eleven commonly used machine learning classification models, including the k-neighbors classifier (KN) (Parzen, 1962), support vector machine (SVM) classifier (Ben-Hur et al., 2000), Gaussian Process (GP) classifier (Chen et al., 2020), Decision Tree (DT) classifier (Quinlan, 1986), Multi-layer Perceptron (MP) classifier (Cybenko, 1989), AdaBoost (AB) classifier (Freund and Schapire, 1997), Gaussian Naïve Bayes (GNB) (Chan T.F., 1982), Quadratic Discriminant Analysis (QDA) (Cover, 1965), Gradient Boosting (GB) classifier (Friedman, 2001), XGBoost (XGB) classifier (Chen and Guestrin, 2016), and Random Forest (RF) classifier (Pal, 2005), were employed for propagation ~~judgement~~judgment. Drought duration, severity, and affected area of meteorological drought were set as the model inputs ~~(Fig. 2).~~(Figure 2~~Figure 2~~). 1 and 0 were set as model target which ~~represent~~represents propagation occurrence and non-occurrence, respectively. ~~The classifiers of~~ In this study, each ~~model were trained and validated~~binary classifier was constructed using a Python package called PyCaret, which wraps several machine-learning libraries, including scikit-learn, XGBoost, LightGBM, CatBoost, spaCy, Optuna, and Hyperopt(Ali, 2020). The tune_model() function in the PyCaret package offers simple selection of optimal hyperparameters of each model. A 5-fold cross-validation~~. The~~ was used to train and validate the classifiers in each model by setting "fold=5" in the create_model() function. In using the compare_models() function, the classifier with the highest summation of accuracy, precision, recall, F1 score, and Matthews correlation coefficient was selected as the optimal model. To avoid overfitting and maintain high calculation efficiency, the L2 regularization method was selected for each model by setting the parameter "penalty='l2'".

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$precision = \frac{TP}{TP + FP} \tag{6}$$

$$recall = \frac{TP}{TP + FN} \tag{7}$$

$$F_1\ score = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{8}$$

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \tag{9}$$

where *TP* and *FN* represent actual positives that are correctly and wrongly predicted, respectively; *TN* and *FP* represent actual negatives are correctly and wrongly predicted, respectively.

~~3.3~~2.4.4 Drought propagation probability model based on C-vine copula

Five univariate distributions, including ~~Johnsonsb~~Johnson S_B (Soukissian, 2013), Gamma (THOM, 1958), Exponential (Marshall and Olkin, 1967), Pearson III (Wallis and Wood, 1985), and Weibull distribution (Thoman et al., 1969), were used to fit affected area, duration, and severity of meteorological drought and severity of ecological drought. The optimal distribution was selected according to the goodness of fit (GOF), which was estimated with the Kolmogorov–Smirnov (KS) test (Marsaglia et al., 2003) and Root Mean Square Error (RMSE).

Commonly used Copulas, including elliptical Copula (~~Guassian~~Gaussian) and four Archimedean Copulas (Clayton, Gumbel, Frank, and Joe), were used to join two marginal distributions (Chang et al., 2016). The GOF of these Copulas was estimated with RMSE and Cramer-von Mises (~~CM~~CvM) test (Genest et al., 2009).

The Vine copula function is an effective tool for integrating different bivariate distributions and calculating the conditional probability of multiple variables (Ni et al., 2020). In a vine copula, an *n*-dimensional multivariate density is decomposed into *n*(*n*−1)/2 bivariate copula densities and arranged into *n*-1 trees. Among numerous vine Copula structures, the C-vine copula has a relatively simple structure and good robustness for constructing multivariate distributions (Wu et al., 2021). Therefore, it was of primary significance to this study. The GOF of C-vine Copulas was estimated with RMSE and CvM test. ~~the~~The joint density function of an n-dimensional C-vine Copula is expressed as Equation (10).

$$f(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_i(x_i) \times \prod_{i=1}^{n-1}\prod_{j=1}^{n-i} c_{i,i+j|1:(i-1)}\left\{F(x_i \mid x_1,\ldots,x_{i-1}), F(x_{i+j} \mid x_1,\ldots,x_{i-1})\right\} \tag{10}$$

where ~~C~~$f(x_1,\ldots,x_n)$ represents the joint density function. c represents bivariate Copula densities, which includes Gumbel, Gaussian, Frank, and Clayton Copula functions; F represents cumulative distribution function ~~(CDF) of joint distribution; F represents CDF~~ of marginal distribution. i and j represent root nodes. More detailed information about the n-dimensional C-vine copula can be referred to Wu et al (2021). By this means, the conditional probabilities of ecological drought at different magnitudes under impacts of meteorological drought are calculated using Equation (11).

$$
\begin{aligned}
F\left(X>x \mid D>d, A>a, S>s\right) &= \frac{F\left(D>d, S>s, A>a, X>x\right)}{F\left(S>s, A>a, D>d\right)} \\
&= \frac{\begin{array}{l}1-F(d)-F(s)-F(a)-F(x)+C\left(F_D(d),F_S(s)\right)+C\left(F_D(d),F_A(a)\right) \\ +C\left(F_D(d),F_X(x)\right)+C\left(F_A(a),F_S(s)\right)+C\left(F_A(a),F_X(x)\right) \\ +C\left(F_X(x),F_S(s)\right)-C\left(F_D(d),F_S(s),F_A(a)\right)-C\left(F_D(d),F_S(s),F_X(x)\right) \\ -C\left(F_D(d),F_A(s),F_X(x)\right)-C\left(F_S(s),F_A(a),F_X(x)\right) \\ +C\left(F_D(d),F_S(s),F_A(a),F_X(x)\right)\end{array}}{\begin{array}{l}1-F_D(d)-F_A(a)-F_S(s)+C\left(F_D(d),F_A(a)\right)+C\left(F_D(d),F_S(s)\right)+ \\ C\left(F_A(a),F_S(s)\right)-C\left(F_D(d),F_A(a),F_S(s)\right)\end{array}}
\end{aligned}
\tag{11}
$$

where D, A, and S represent duration, area, and severity of propagated meteorological drought, respectively; X represents ecological drought at moderate, ~~serious~~severe, and extreme magnitudes, which equals the cumulative probability of 0.5, 0.7, and 0.9, respectively. C represents ~~CDF~~the cumulative distribution function of the joint distribution.

~~4~~3 Results

3.1 Threshold selection

Determining overlapping areas of drought patches between two adjacent months is critical in the identification of drought events from a three-dimensional perspective. Sheffield et al (2009) used 500,000 km^2 as the area threshold in global scales. For mainland China, 150,000 km^2 was used as the area threshold in some studies (Wang et al., 2011; Xu et al., 2015b). Liu et al (2019) took 1.5% of the total area as the threshold in the Loess Plateau. To determine an optimal area threshold, the number of meteorological and ecological drought events as well as the ratio of minor drought events were calculated under different area thresholds, respectively. Here, a minor drought is defined as a drought event with 2 months duration and average SPI/SEWDI larger than -1. As shown in Figure 3~~Figure 3~~, 1328 and 2305 meteorological and ecological drought events were identified with an area threshold of 0.48% of the total area of the NWC, and the proportions of minor drought events were 44% and 32%, respectively. The number of drought events and the proportion of minor drought events decreased with increasing area threshold. However, this trend gradually stabilized when the area threshold was set to be larger than 1.6% of the total area of the NWC, indicating that most minor drought events with the relatively small area were excluded. Therefore, 1.6% of the total area of the NWC was used as the area threshold in this study.

Similarly, the sensitivity of b in Eq.(4) for matching two drought types was tested. The binding mode of absolute and relative thresholds was employed to extract spatial intersection. b is set as 10%, 15%, 30%, 50%, 70%, and 90% to match two drought types. Although some of the successful matching drought events may be merged into one drought event under larger *b*, the number of successful matching drought events showed little difference under different *b* (Table 1~~Table 1~~). In the current study, *b* =15% was set because the most paired drought events could be identified for fitting machine learning models and C-vine copula.

~~4.1~~3.2 Top ten meteorological and ecological drought events according to ~~DS~~drought severity

A total of 130 meteorological drought events were identified based on SPI-3 from a three-dimensional perspective. The first ten meteorological drought events in terms of severity in NWC during 1982–2020 are shown in ~~Table 1.~~Table 2~~Table 2~~. Meteorological drought events with longer duration exhibited relatively larger affected area, were mainly concentrated between 1982 and 2000. Zou et al (2005) estimated meteorological droughts with the Palmer drought severity index (PDSI) from 1951 to 2003 in China and found that most parts of NWC experienced severe droughts during 1997–2003, which is similar to the results of this study. As shown in ~~Table 1,~~Table 2~~Table 2~~, two of ten meteorological drought events occurred during this period. Moreover, according to the historical record, Xinjiang and Gansu ~~province~~ experienced severe meteorological drought during 1985–1986 (Zhang et al., 2019b). The three-dimensional identification method could sensitively capture these events. Event No.9 started from southern Gansu in August 1985 and ended in May 1987, and ranked 1st.

A total of 184 ecological drought events during 1982–2020 were identified using the three-dimensional identification method. ~~Table 2~~Table 3~~Table 3~~ lists the top ten ecological drought events in terms of severity. The most severe ecological drought event started in April 1982 and originated from central Gansu, which was induced by the persistent meteorological drought No.0 and No.3. Compared with the characteristics of meteorological droughts, ecological droughts showed longer duration and smaller affected area. This reveals that a longer recovery time is required for the mitigation of ecological droughts.

~~4.2~~3.3 Identifying propagation from meteorological to ecological drought

A total of 46 paired drought events were successfully matched based on the spatio-temporal connection criterion. As shown in ~~Fig. 3,~~Figure 4~~Figure 4~~, points representing paired drought events were mainly distributed along a diagonal, illustrating a relatively high consistency between the two types of droughts on the temporal scale. The number of one–to–one, many–to–one, one–to–many, and many–to–many were 8, 8, 4, and 26, accounting for 17.4 %, 17.4 %, 8.7 ~~%~~%, and, 56.5 % of the total number of paired drought events, respectively. Meteorological drought of OMT type showed a longer duration, a larger affected area, and a greater severity than ecological drought. However, this is contrary to type MTO. Simultaneously, ecological drought of type MTO showed a longer duration, a larger affected area, and a greater severity than those of type OTM (Figure 5~~Figure 5~~).

Paired drought event No.36, comprising meteorological drought event No.87 and ecological drought event No.127, was taken as an example to show their spatio-temporal continuums ~~(Fig. 4).~~(Figure 6~~Figure 6~~). The affected area of meteorological and ecological drought in each month ~~were~~was extracted to show their temporal variation. Meteorological drought No.87 ~~(Fig.~~

~~4a~~(Figure 6~~Figure 6~~a) started two months ahead of ecological drought ~~(Fig. 4b~~(Figure 6~~Figure 6~~b), and its effects lasted for two months. It is noteworthy that the most severe meteorological and ecological droughts mainly occurred in central Xinjiang. The affected area and severity of meteorological drought event No.87 and ecological drought event No.127 maintained a similar trend of increase–decrease ~~(Fig. 5).~~(Figure 7~~Figure 7~~). Among them, the peaks of the meteorological drought event appeared two months ahead (December 2007) that of the ecological drought (February ~~2007~~2008). In terms of drought trajectory ~~(Fig. 6),~~(Figure 8~~Figure 8~~), they all originated from the Yili Basins and showed a ~~counter clockwise~~counterclockwise shift.

3.4~~.3~~ Propagation probability from meteorological to ecological drought

To estimate the propagated potential of meteorological drought, 11 commonly used machine learning models were trained based on characteristics of 81 integrated meteorological drought events. ~~Table 3~~As can be seen in Figure 9~~Figure 9~~ propagated meteorological droughts have greater severity, larger affected area, and longer duration than non-propagated droughts. Table 4~~Table 4~~ lists the evaluation results of five-fold cross-validations, including accuracy, precision, recall, F1 Score, and MCC metrics~~ of five-fold cross-validations.~~. The closer these values are to 1, the higher precision of the model. Therefore, the five metrics were summed to compare the performances of the 11 models. Most models showed good performance except for ~~GP~~Gaussian Process and ~~MP.~~Multi-layer Perceptron. The QDA classifier with maximum total value was chosen as the best model to identify the propagation potential of meteorological drought.

The reliability of copula function is highly dependent on the dependence between two variables, which was measured by Kendall's τ and Spearman's ρ (Chang et al., 2016; Feng et al., 2021). The τ and ρ between affected area (M_Area), duration (M_Duration), severity (M_Severity) of meteorological drought, and severity of ecological drought (E_Severity) both reached significance at 0.01 level (Table S.~~1~~ 1-S.2). The optimal marginal distributions of M_Area, M_Duration, M_Severity, and E_Severity are listed in ~~Table 4.~~Table 5~~Table 5~~. All the distributions passed the KS test and their RMSE ~~were~~was small. Similarly, the parameters of bivariate distribution were estimated using the itau method~~, and~~. The Copula estimation can be eased by the itau method, which inverts Kendall's tau method (Demarta and McNeil, 2005). CvM test and RMSE were used to evaluate their goodness of fit ~~(Table 5).~~(Table 6~~Table 6~~). The selected bivariate copulas also demonstrated a well applicability. In the end, the C-vine copula was constructed centered on E_Severity. The CvM test, RMSE ~~(Table 6),~~(Table 7~~Table 7~~), and P–P plots (Fig. S.1) indicated that the distribution can be used in probability analysis. The copula structure of M_Area-M_Duration-M_Severity-E_Severity was shown in ~~Table 6.~~Table 7~~Table 7~~.

Conditional probability is helpful in providing valuable information for the effective allocation of water resources under a certain drought level (Guo et al., 2020). In the current study, the occurrence probabilities of ecological drought at different levels were determined according to the characteristics of meteorological drought ~~(Fig. 7).~~(Figure 10~~Figure 10~~). For example, the occurrence probabilities of moderate, ~~serious~~severe, and extreme ecological drought events were 80%, 63%, 14.7%, respectively, when M_DA > 17.6×10^5 km^2 ∩ M_DD > 11.8 month ∩ M_DS > 7.5×10^6 month·km^2. Furthermore, the occurrence probability was found to increase more rapidly with increasing M_DS and M_DD compared with M_DA, indicating that the duration and severity of meteorological drought had stronger effects on ecological drought than affected area. Additionally, meteorological drought events with a duration of two months but great severity has a high potential to trigger ecological drought. This may be attributable to water shortage induced by meteorological droughts with extremely high intensity (intensity is the drought severity divided by the product of drought duration and affected area).

For comparison, ternary linear and ternary quadratic models were constructed based on 46 pairs of meteorological-ecological drought events (Table 8~~Table 8~~). The comparisons were made in terms of three independent variables, M_DS, M_DD, and M_DA, and one dependent variable, E_DS. As shown in Table 8~~Table 8~~, the R^2 of the ternary quadratic model was evidently higher than that of the ternary linear model, whereas the RMSE, AIC, and BIC were lower. This illustrates that M_DS, M_DD, M_DA, and E_DS follow a nonlinear relationship, and that the ternary quadratic model is more suitable for simulating their relationship. According to the ternary quadratic model, E_DS equals 1.4×10^6 month·km^2 when M_DA > 17.6×10^5 km^2 ∩ M_DD > 11.8 month ∩ M_DS > 7.5×10^6 month·km^2. These values correspond to the thresholds of moderate (1.7×10^6 month·km^2), severe (2.4×10^6 month·km^2), and extreme (4.6×10^6 month·km^2) ecological drought.

4 Discussion

4.1 Advantages of the proposed approach

Many studies have linked meteorological drought to hydrological drought at different time scales (Ding et al., 2021; Fang et al., 2020; Feng and Su, 2020; Han et al., 2019; Huang et al., 2017; Ma et al., 2019). In these studies, propagated drought events were identified on the basis of the time series between two drought types, and they focused on their lagging, attenuation, lengthening, and pooling ~~(Fig. 9a). Spatial~~(Figure 11~~Figure 11~~a). The spatial and temporal drought propagation identification method used in the current study not only preserved the characteristics identified by the low dimensional method, but also considered the spatial overlap of two drought types ~~(Fig. 9b).~~(Figure 11~~Figure 11~~b). Using this method, two types of drought events without spatial connection would be excluded~~,~~ (only 103 out of 184 ecological drought events were induced by 81 out of 108 meteorological drought events), and more drought characteristics, such as affected area, and migration path could be extracted. This addresses the limited applicability of the traditional method to regions with large spatial extent, and provides more ~~realistic drought propagation~~reliable information~~.~~ for quantifying the relationship between characteristics of meteorological drought and ecological drought. Additionally, we improved the method for calculating the affected area and duration of paired drought events developed by Liu et al (2019), represented by a simple sum of characteristics of multiple drought events. However, this method overestimates the duration and affected area of some paired drought events, which is inconsistent with the real situation. In this study, the enhanced method could reflect the characteristics of paired drought during the propagation process more accurately.

The conditional probability model was constructed based on paired meteorological and ecological drought events; it is not suitable for calculating the probability of ecological drought at different levels according to meteorological drought events without propagation potential. For example, the probability of moderate ecological drought was 63.3% if the characteristics of meteorological drought event No.122 (M_Area=5.1×10^5 km^2, M_Duration=6 month, M_Severity=1.89×10^6 month·km^2) was directly input to the conditional probability model. In reality, this meteorological drought event did not trigger ecological drought. The QDA model added before the C-vine copula was used to address this issue, which could estimate the propagation potential of the corresponding meteorological drought. After this modification, the probability of the propagation of meteorological drought event No.122 to moderate ecological drought changed to 24.8%.

~~5.3~~4.2 Uncertainty of the model and its improvement measures

QDA model could well simulate the propagation potential of most meteorological drought events ~~(Table 3).~~(Table 4~~Table 4~~). However, some errors occurred in humid southern Shaanxi. For example, meteorological drought ~~events~~event No. 22 showed the potential to trigger ecological droughts, which were incorrectly classified as propagation occurrence. This could be attributed to the compensation of rich water resources for short-term ecological water deficit. Additionally, this paper provides a method for estimating the occurrence probability of ecological drought under the condition of a certain precipitation deficit. The effects of human activities and climate change on ecological drought were not distinguished in the current study. The proposed method may not be accurate for regions with complex water supply systems and strong anthropogenic impacts on vegetation growth.

To improve the accuracy of the method, future studies should consider the non-consistence of ecological drought to quantify the impacts of human activities on drought propagation. Moreover, SPI can be replaced by PDSI or scPDSI to represent meteorological drought through which multiple water balance processes are considered to analyze their relationship with ecological drought (Altunkaynak and Jalilzadnezamabad, 2021). However, such modification may lead new problem associated with spatio-temporal incomparability. Nevertheless, this approach is worth applying in the ecological drought warning. For example, when a meteorological drought event occurs, its characteristics can be applied as input to the tuned model to estimate propagation probability from meteorological to ecological drought in different degrees.

~~6~~5. Conclusions

This study proposed a method in identifying the propagation probability of meteorological drought events to trigger ecological drought in different magnitudes. Taking NWC as an example, 130 meteorological drought and 185 ecological drought events during 1982–2020 were extracted using the three-dimensional identification method. Compared with meteorological drought, ecological drought events exhibited longer duration, but smaller affected area and severity, suggesting that a longer recovery time is required for mitigating ecological drought.

A total of 46 drought events were successfully matched according to a certain spatio-temporal connection principle. The paired drought events were divided into four categories, including one-to-one, many-to-one, one-to-many, and many-to-many. The four categories accounted for 17.4 %, 17.4 %, 8.7 %, and 56.5 % of the total number of paired drought events, respectively. Then, a drought propagation probability model was constructed by coupling QDA and C-vine copula. Compared with the traditional propagation probability model, the proposed model intuitively provides more objective probabilities of ecological drought at different magnitudes.

The current study certainly provides a more robust method for estimating propagation probability from meteorological to ecological drought in similar ecologically fragile regions.

Data availability

Monthly meteorological data, including surface reflectance, temperature, relative humidity, atmospheric pressure, downward shortwave radiation, wind speed, and longwave radiation, was obtained from the ERA5-land reanalysis dataset (https://cds.climate.copernicus.eu) issued by the European Centre for Medium-Range Weather Forecasts (ECWMF), which has a spatial resolution of 0.1° × 0.1° and covers the period of 1981–2021. Root soil moisture data were obtained from the hydrological dataset, simulated by the Noah model of the Global Land Data Assimilation System (GLDAS, 0.25°×0.25°; https://ldas.gsfc.nasa.gov/gldas), covering the period of 1948–2021. NDVI data covering the period 1981–2021 were obtained from the National Centers for Environmental Information (NCEI) (https://www.ncei.noaa.gov/), with a spatial resolution of 0.05° × 0.05°. Land use type data (LUTD) with a spatial resolution 1 km was downloaded from China's multi-period land use/cover change monitoring dataset (http://www.resdc.cn); it includes the years of 1980, 1990, 1995, 2000, 2005, 2010, 2015, 2018, and 2020.

Author contribtion

Tianliang Jiang: Conceptualization, Methodology, Software, Visualization, Writing - original draft. Xiaoling Su: Data curation, Validation, Investigation, Funding acquisition, Supervision, Formal analysis. Gengxi Zhang: Writing - review & editing, Supervision. Te Zhang: Formal analysis, Investigation. Haijiang Wu: Data curation, Investigation.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgment

The study was supported by the National Natural Science Foundation in China (Grants 52079111 and 51879222).

References

Ali, M.: PyCaret, PyCaret: An open source, low-code machine learning library in Python, 2020.

Altunkaynak, A. and Jalilzadnezamabad, A.: Extended lead time accurate forecasting of palmer drought severity index using hybrid wavelet-fuzzy and machine learning techniques, Journal of Hydrology, 601, 126619, 10.1016/j.jhydrol.2021.126619, 2021.

Andreadis, K. M., Clark, E. A., Wood, A. W., Hamlet, A. F., and Lettenmaier, D. P.: Twentieth-Century Drought in the Conterminous United States, Journal of Hydrometeorology, 6, 985-1001, 10.1175/jhm450.1, 2005.

Apurv, T., Sivapalan, M., and Cai, X.: Understanding the role of climate characteristics in drought propagation, Water Resources Research, 53, 9304-9329, 10.1002/2017WR021445, 2017.

Ayantobo, O. O. and Wei, J.: Appraising regional multi-category and multi-scalar drought monitoring using standardized moisture anomaly index (SZI): A water-energy balance approach, Journal of Hydrology, 579, 124139, 10.1016/j.jhydrol.2019.124139, 2019.

Ayantobo, O. O., Li, Y., Song, S., Javed, T., and Yao, N.: Probabilistic modelling of drought events in China via 2-dimensional joint copula, Journal of Hydrology, 559, 373-391, 10.1016/j.jhydrol.2018.02.022, 2018.

Ben-Hur, A., Horn, D., Siegelmann, H. T., and Vapnik, V.: A support vector clustering method, Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, 3-7 Sept. 2000, 724-727, 10.1109/ICPR.2000.906177,

Bento, V. A., Gouveia, C. M., DaCamara, C. C., Libonati, R., and Trigo, I. F.: The roles of NDVI and Land Surface Temperature when using the Vegetation Health Index over dry regions, Global and Planetary Change, 190, 103198, 10.1016/j.gloplacha.2020.103198, 2020.

Bradford, J. B., Schlaepfer, D. R., Lauenroth, W. K., and Palmquist, K. A.: Robust ecological drought projections for drylands in the 21st century, Glob Chang Biol, 26, 3906-3919, 10.1111/gcb.15075, 2020.

Cao, S., He, Y., Zhang, L., Chen, Y., Yang, W., Yao, S., and Sun, Q.: Spatiotemporal characteristics of drought and its impact on vegetation in the vegetation region of Northwest China, Ecological Indicators, 133, 108420, 10.1016/j.ecolind.2021.108420, 2021.

Chan T.F., G. G. H., LeVeque R.J.: Updating Formulae and a Pairwise Algorithm for Computing Sample Variances. In: Caussinus H., Ettinger P., Tomassone R. (eds) COMPSTAT 1982 5th Symposium held at Toulouse 1982, https://doi.org/10.1007/978-3-642-51461-6_3, 1982

Chang, J., Li, Y., Wang, Y., and Yuan, M.: Copula-based drought risk assessment combined with an integrated index in the Wei River Basin, China, Journal of Hydrology, 540, 824-834, https://doi.org/10.1016/j.jhydrol.2016.06.064, 2016.

Chen, T. and Guestrin, C.: Xgboost: A scalable tree boosting system, Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794, 2016.

Chen, Z., Wang, B., and Gorban, A. N.: Multivariate Gaussian and Student-t process regression for multi-output prediction, Neural Computing and Applications, 32, 3005-3028, 10.1007/s00521-019-04687-8, 2020.

Chi, D., Wang, H., Li, X., Liu, H., and Li, X.: Estimation of the ecological water requirement for natural vegetation in the Ergune River basin in Northeastern China from 2001 to 2014, Ecological Indicators, 92, 141-150, 10.1016/j.ecolind.2017.04.014, 2018.

Cover, T. M.: Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition, IEEE Transactions on Electronic Computers, EC-14, 326-334, 10.1109/PGEC.1965.264137, 1965.

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., Betancourt, J. L., Colt, S., Cravens, A. E., Dalton, M. S., Dunham, J. B., Hay, L. E., Hayes, M. J., McEvoy, J., McNutt, C. A., Moritz, M. A., Nislow, K. H., Raheem, N., and Sanford, T.: Defining Ecological Drought for the Twenty-First Century, Bulletin of the American Meteorological Society, 98, 2543-2550, 10.1175/bams-d-16-0292.1, 2017.

Cybenko, G.: Approximation by superpositions of a sigmoidal function, Mathematics of Control, Signals and Systems, 2, 303-314, 10.1007/BF02551274, 1989.

Das, J., Jha, S., and Goyal, M. K.: Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India, Journal of Hydrology, 580, 124356, 10.1016/j.jhydrol.2019.124356, 2020.

Demarta, S. and McNeil, A. J.: The t Copula and Related Copulas, International Statistical Review / Revue Internationale de Statistique, 73, 111–129, 2005.

Ding, Y., Xu, J., Wang, X., Cai, H., Zhou, Z., Sun, Y., and Shi, H.: Propagation of meteorological to hydrological drought for different climate regions in China, Journal of Environmental Management, 283, 111980, 10.1016/j.jenvman.2021.111980, 2021.

Fang, W., Huang, S., Huang, Q., Huang, G., Wang, H., Leng, G., and Wang, L.: Identifying drought propagation by simultaneously considering linear and nonlinear dependence in the Wei River basin of the Loess Plateau, China, Journal of Hydrology, 591, 125287, 10.1016/j.jhydrol.2020.125287, 2020.

Feng, K. and Su, X.: Spatiotemporal response characteristics of agricultural drought to meteorological drought from a three-dimensional perspective, Transactions of the Chinese Society of Agricultural Engineering 36, 103-113, 10.11975/j.issn.1002-6819.2020.08.013, 2020.

Feng, K., Su, X., Singh, V. P., Ayantobo, O. O., Zhang, G., Wu, H., and Zhang, Z.: Dynamic evolution and frequency analysis of hydrological drought from a three-dimensional perspective, Journal of Hydrology, 600, 126675, 10.1016/j.jhydrol.2021.126675, 2021.

Freund, Y. and Schapire, R. E.: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, Journal of Computer and System Sciences, 55, 119-139, https://doi.org/10.1006/jcss.1997.1504, 1997.

Friedman, J. H.: Greedy function approximation: A gradient boosting machine, The Annals of Statistics, 29, 1189-1232, 1144, 2001.

Gao, T. and Jojic, V.: Degrees of Freedom in Deep Neural Networks, https://doi.org/10.48550/arXiv.1603.09260, 2016.

Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, Insurance: Mathematics and Economics, 44, 199-213, https://doi.org/10.1016/j.insmatheco.2007.10.005, 2009.

Guo, Y., Huang, S., Huang, Q., Leng, G., Fang, W., Wang, L., and Wang, H.: Propagation thresholds of meteorological drought for triggering hydrological drought at various levels, The Science of the total environment, 712, 136502, 10.1016/j.scitotenv.2020.136502, 2020.

Han, Z., Huang, S., Huang, Q., Leng, G., Wang, H., Bai, Q., Zhao, J., Ma, L., Wang, L., and Du, M.: Propagation dynamics from meteorological to groundwater drought and their possible influence factors, Journal of Hydrology, 578, 124102, 10.1016/j.jhydrol.2019.124102, 2019.

Huang, S., Li, P., Huang, Q., Leng, G., Hou, B., and Ma, L.: The propagation from meteorological to hydrological drought and its potential influence factors, Journal of Hydrology, 547, 184-195, https://doi.org/10.1016/j.jhydrol.2017.01.041, 2017.

Janson, L., Fithian, W., and Hastie, T.: Effective Degrees of Freedom: A Flawed Metaphor, https://doi.org/10.1093/biomet/asv019, 2013.

Jehanzaib, M., Sattar, M. N., Lee, J.-H., and Kim, T.-W.: Investigating effect of climate change on drought propagation from meteorological to hydrological drought using multi-model ensemble projections, STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT, 34, 7–21, https://doi.org/10.1007/s00477-019-01760-5, 2020.

Jiang, T., Su, X., Singh, V. P., and Zhang, G.: A novel index for ecological drought monitoring based on ecological water deficit, Ecological Indicators, 129, 107804, https://doi.org/10.1016/j.ecolind.2021.107804, 2021.

Jiang, Y., Wang, R., Peng, Q., Wu, X., Ning, H., and Li, C.: The relationship between drought activity and vegetation cover in Northwest China from 1982 to 2013, Natural Hazards, 92, 145-163, 10.1007/s11069-018-3282-3, 2018.

Jehanzaib, M., Sattar, M. N., Lee, J.-H., and Kim, T.-W.: Investigating effect of climate change on drought propagation from meteorological to hydrological drought using multi-model ensemble projections, STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT, 34, 7–21, https://doi.org/10.1007/s00477-019-01760-5, 2020.

Lawal, S., Hewitson, B., Egbebiyi, T. S., and Adesuyi, A.: On the suitability of using vegetation indices to monitor the response of Africa's terrestrial ecoregions to drought, The Science of the total environment, 792, 148282, 10.1016/j.scitotenv.2021.148282, 2021.

Liu, H., Jia, J., Lin, Z., Wang, Z., and Gong, H.: Relationship between net primary production and climate change in different vegetation zones based on EEMD detrending – A case study of Northwest China, Ecological Indicators, 122, 107276, 10.1016/j.ecolind.2020.107276, 2021.

Liu, M., Xu, X., Xu, C., Sun, A. Y., Wang, K., Scanlon, B. R., and Zhang, L.: A new drought index that considers the joint effects of climate and land surface change, Water Resources Research, 53, 3262-3278, 10.1002/2016WR020178, 2017.

Liu, Y., Zhu, Y., Ren, L., Singh, V. P., Yong, B., Jiang, S., Yuan, F., and Yang, X.: Understanding the Spatiotemporal Links Between Meteorological and Hydrological Droughts From a Three-Dimensional Perspective, Journal of Geophysical Research: Atmospheres, 124, 3090-3109, 10.1029/2018jd028947, 2019.

Ma, F., Luo, L., Ye, A., and Duan, Q.: Drought Characteristics and Propagation in the Semiarid Heihe River Basin in Northwestern China, Journal of Hydrometeorology, 20, 59-77, 10.1175/jhm-d-18-0129.1, 2019.

Marsaglia, G., Tsang, W. W., and Wang, J.: Evaluating Kolmogorov's distribution, Journal of statistical software, 8, 1-4, https://doi.org/10.18637/jss.v008.i18, 2003.

Marshall, A. W. and Olkin, I.: A Multivariate Exponential Distribution, Journal of the American Statistical Association, 62, 30-44, 10.1080/01621459.1967.10482885, 1967.

McEvoy, J., Bathke, D. J., Burkardt, N., Cravens, A. E., Haigh, T., Hall, K. R., Hayes, M. J., Jedd, T., Poděbradská, M., and Wickham, E.: Ecological Drought: Accounting for the Non-Human Impacts of Water Shortage in the Upper Missouri Headwaters Basin, Montana, USA, Resources, 7, 14, 10.3390/resources7010014, 2018.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, Proceedings of the 8th Conference on Applied Climatology, 179-183,

Mishra, A. K. and Singh, V. P.: A review of drought concepts, Journal of hydrology, 391, 202-216, https://doi.org/10.1016/j.jhydrol.2010.07.012, 2010.

Munson, S. M., Bradford, J. B., and Hultine, K. R.: An Integrative Ecological Drought Framework to Span Plant Stress to Ecosystem Transformation, Ecosystems, 24, 739-754, 10.1007/s10021-020-00555-y, 2021.

Ni, L., Wang, D., Wu, J., Wang, Y., Tao, Y., Zhang, J., Liu, J., and Xie, F.: Vine copula selection using mutual information for hydrological dependence modeling, Environ Res, 186, 109604, 10.1016/j.envres.2020.109604, 2020.

Pal, M.: Random forest classifier for remote sensing classification, International Journal of Remote Sensing, 26, 217-222, 10.1080/01431160412331269698, 2005.

Parzen, E.: On Estimation of a Probability Density Function and Mode, The Annals of Mathematical Statistics, 33, 1065-1076, 10.1214/aoms/1177704472, 1962.

Peng, J., Wu, C., Zhang, X., Wang, X., and Gonsamo, A.: Satellite detection of cumulative and lagged effects of drought on autumn leaf senescence over the Northern Hemisphere, Global Change Biology, 25, 2174-2188, https://doi.org/10.1111/gcb.14627, 2019.

Quinlan, J. R.: Induction of decision trees, Machine Learning, 1, 81-106, 10.1007/BF00116251, 1986.

Raheem, N., Cravens, A. E., Cross, M. S., Crausbay, S., Ramirez, A., McEvoy, J., Zoanni, D., Bathke, D. J., Hayes, M., Carter, S., Rubenstein, M., Schwend, A., Hall, K., and Suberu, P.: Planning for ecological drought: Integrating ecosystem services and vulnerability assessment, Wiley Interdisciplinary Reviews: Water, e1352, 10.1002/wat2.1352, 2019.

Sattar, M. N., Lee, J.-Y., Shin, J.-Y., and Kim, T.-W.: Probabilistic Characteristics of Drought Propagation from Meteorological to Hydrological Drought in South Korea, Water Resources Management, 33, 2439-2452, 10.1007/s11269-019-02278-9, 2019.

Sheffield, J., Andreadis, K., Wood, E. F., and Lettenmaier, D.: Global and continental drought in the second half of the twentieth century: Severity–area–duration analysis and temporal variability of large-scale events, Journal of Climate, 22, 1962-1981, https://doi.org/10.1175/2008JCLI2722.1, 2009.

Soukissian, T.: Use of multi-parameter distributions for offshore wind speed modeling: The Johnson SB distribution, Applied Energy, 111, 982-1000, https://doi.org/10.1016/j.apenergy.2013.06.050, 2013.

THOM, H. C. S.: A NOTE ON THE GAMMA DISTRIBUTION, Monthly Weather Review, 86, 117-122, 10.1175/1520-0493(1958)086<0117:Anotgd>2.0.Co;2, 1958.

Thoman, D. R., Bain, L. J., and Antle, C. E.: Inferences on the Parameters of the Weibull Distribution, Technometrics, 11, 445-460, 10.1080/00401706.1969.10490706, 1969.

Vicente-Serrano, S. M., Beguería, S., López-Moreno, J. I., Angulo, M., and El Kenawy, A.: A New Global 0.5° Gridded Dataset (1901–2006) of a Multiscalar Drought Index: Comparison with Current Drought Index Datasets Based on the Palmer Drought Severity Index, Journal of Hydrometeorology, 11, 1033-1043, 10.1175/2010jhm1224.1, 2010.

Vicente-Serrano, S. M., Beguer?a, S., Lorenzo-Lacruz, J., Camarero, J. s. J., L?pez-Moreno, J. I., Azorin-Molina, C., Revuelto, J. s., Mor?n-Tejeda, E., and Sanchez-Lorenzo, A.: Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications, Earth Interactions, 16, 1-27, 10.1175/2012ei000434.1, 2012.

Vicenteserrano, S. M., Begueria, S., and Lopezmoreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, Journal of Climate, 23, 1696-1718, https://doi.org/10.1175/2009JCLI2909.1, 2010.

Wallis, J. R. and Wood, E. F.: Relative Accuracy of Log Pearson III Procedures, Journal of Hydraulic Engineering, 111, 1043-1056, doi:10.1061/(ASCE)0733-9429(1985)111:7(1043), 1985.

Wang, A., Lettenmaier, D. P., and Sheffield, J.: Soil Moisture Drought in China, 1950–2006, Journal of Climate, 24, 3257-3271, 10.1175/2011jcli3733.1, 2011.

Wang, Y., Fu, B., Liu, Y., Li, Y., Feng, X., and Wang, S.: Response of vegetation to drought in the Tibetan Plateau: Elevation differentiation and the dominant factors, Agricultural and Forest Meteorology, 306, 108468, 10.1016/j.agrformet.2021.108468, 2021.

Warter, M. M., Singer, M. B., Cuthbert, M. O., Roberts, D., Caylor, K. K., Sabathier, R., and Stella, J.: Drought onset and propagation into soil moisture and grassland vegetation responses during the 2012–2019 major drought in Southern California, Hydrology and Earth System Sciences, 25, 3713-3729, 10.5194/hess-25-3713-2021, 2021.

Wu, H., Su, X., Singh, V. P., Feng, K., and Niu, J.: Agricultural Drought Prediction Based on Conditional Distributions of Vine Copulas, Water Resources Research, 57, e2021WR029562, 10.1029/2021wr029562, 2021.

Xu, K., Yang, D., Xu, X., and Lei, H.: Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China, Journal of Hydrology, 527, 630-640, 10.1016/j.jhydrol.2015.05.030, 2015a.

Xu, K., Yang, D., Yang, H., Li, Z., Qin, Y., and Shen, Y.: Spatio-temporal variation of drought in China during 1961–2012: A climatic perspective, Journal of Hydrology, 526, 253-264, 10.1016/j.jhydrol.2014.09.047, 2015b.

Xu, Y., Zhang, X., Hao, Z., Singh, V. P., and Hao, F.: Characterization of agricultural drought propagation over China based on bivariate probabilistic quantification, JOURNAL OF HYDROLOGY, 598, https://doi.org/10.1016/j.jhydrol.2021.126194, 2021.

Zang, C. S., Buras, A., Esquivel-Muelbert, A., Jump, A. S., Rigling, A., and Rammig, A.: Standardized drought indices in ecological research: Why one size does not fit all, Global Change Biology, 26, 322-324, https://doi.org/10.1111/gcb.14809, 2020.

Zeng, J., Zhang, R., Lin, Y., Wu, X., Tang, J., Guo, P., Li, J., and Wang, Q.: Drought frequency characteristics of China, 1981-2019, based on the vegetation health index, Climate Research, 81, 131-147, 10.3354/cr01616, 2020.

Zhang, B., AghaKouchak, A., Yang, Y., Wei, J., and Wang, G.: A water-energy balance approach for multi-category drought assessment across globally diverse hydrological basins, Agricultural and Forest Meteorology, 264, 247-265, 10.1016/j.agrformet.2018.10.010, 2019a.

Zhang, G., Su, X., Ayantobo, O. O., and Feng, K.: Drought monitoring and evaluation using ESA CCI and GLDAS-Noah soil moisture datasets across China, Theoretical and Applied Climatology, 144, 1407-1418, 10.1007/s00704-021-03609-w, 2021a.

Zhang, G., Su, X., Singh, V. P., and Ayantobo, O. O.: Appraising standardized moisture anomaly index (SZI) in drought projection across China under CMIP6 forcing scenarios, Journal of Hydrology: Regional Studies, 37, 100898, 10.1016/j.ejrh.2021.100898, 2021b.

Zhang, J., Chen, H., and Zhang, Q.: Extreme drought in the recent two decades in northern China resulting from Eurasian warming, Climate Dynamics, 52, 2885-2902, 10.1007/s00382-018-4312-2, 2019b.

Zhang, J., Liu, B., Yang, L., He, L., Cao, X., and Shao, G.: Ecological drought and its state assessment: a case study in the Yellow River estuary, Journal of Water and Climate Change, jwc2021175, 10.2166/wcc.2021.175, 2021c.

Zhang, Z., Ju, W., and Zhou, Y.: The effect of water stress on net primary productivity in northwest China, Environ Sci Pollut Res Int, 28, 65885-65898, 10.1007/s11356-021-15314-2, 2021d.

Zheng, S., Zhang, B., Peng, D., Yu, L., Lin, B., Pan, Y., and Xie, Q.: The trend towards a warmer and wetter climate observed in arid and semi-arid areas of northwest China from 1959 to 2019, Environmental Research Communications, 3, 115011, 10.1088/2515-7620/ac39f7, 2021.

Zhou, Z., Shi, H., Fu, Q., Ding, Y., Li, T., and Liu, S.: Investigating the propagation from meteorological to hydrological drought by introducing the nonlinear dependence with directed information transfer index, Water Resources Research, 57, e2021WR030028, 10.1029/2021wr030028, 2021.

Zou, X., Zhai, P., and Zhang, Q.: Variations in droughts over China: 1951–2003, Geophysical Research Letters, 32, L04707, https://doi.org/10.1029/2004GL021853, 2005.

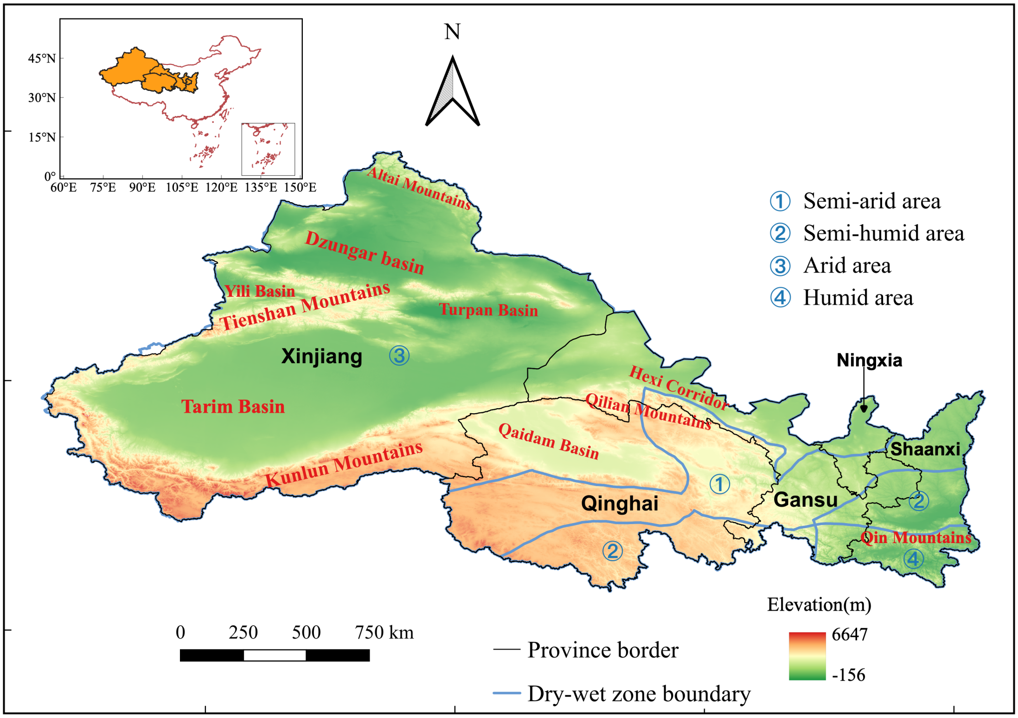


Figure ~~1:~~1: Elevation and four divisions of northwestern China.

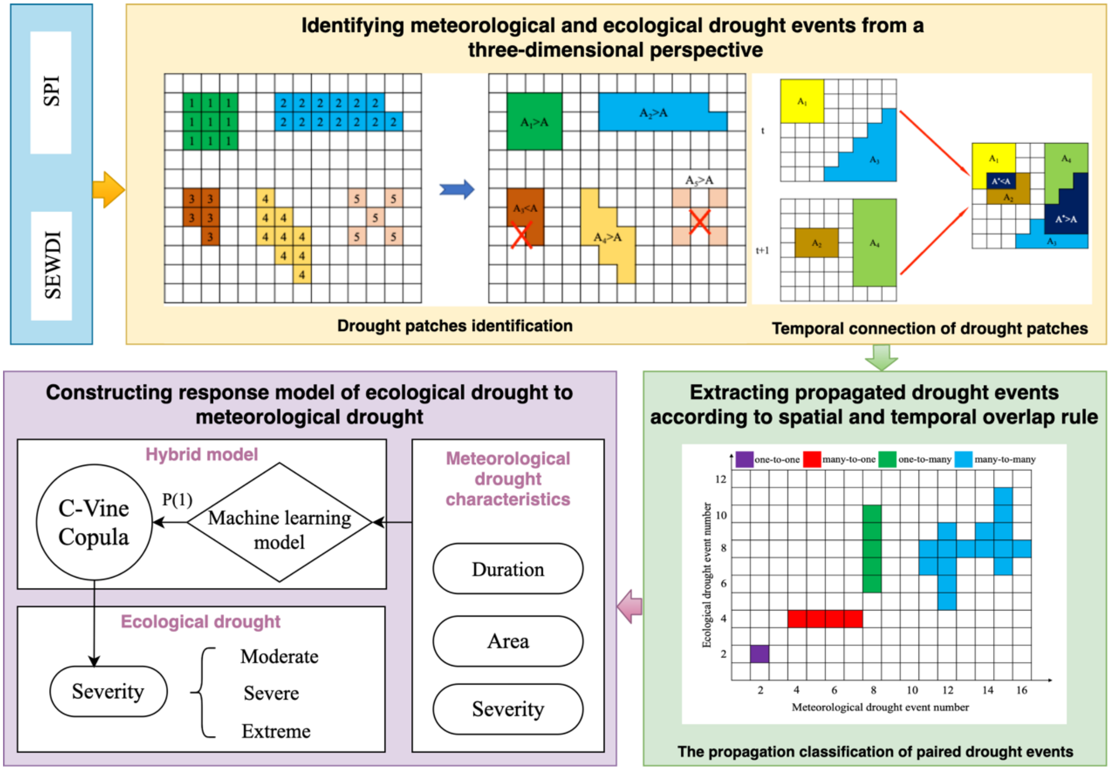


Figure ~~2:~~2: A schematic diagram illustrating the procedure of the drought propagation identification method.

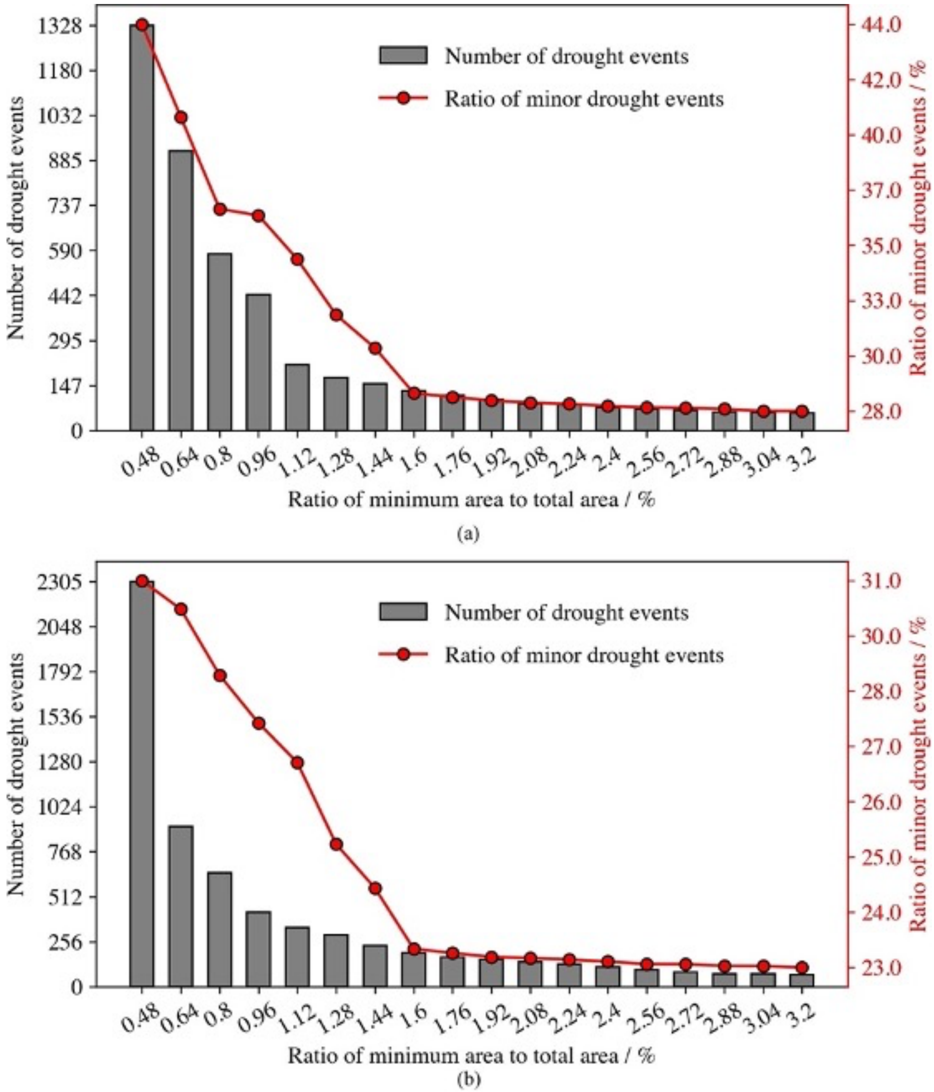


Figure 3: Sensitivity test of overlapping areas of drought patches between two adjacent months.

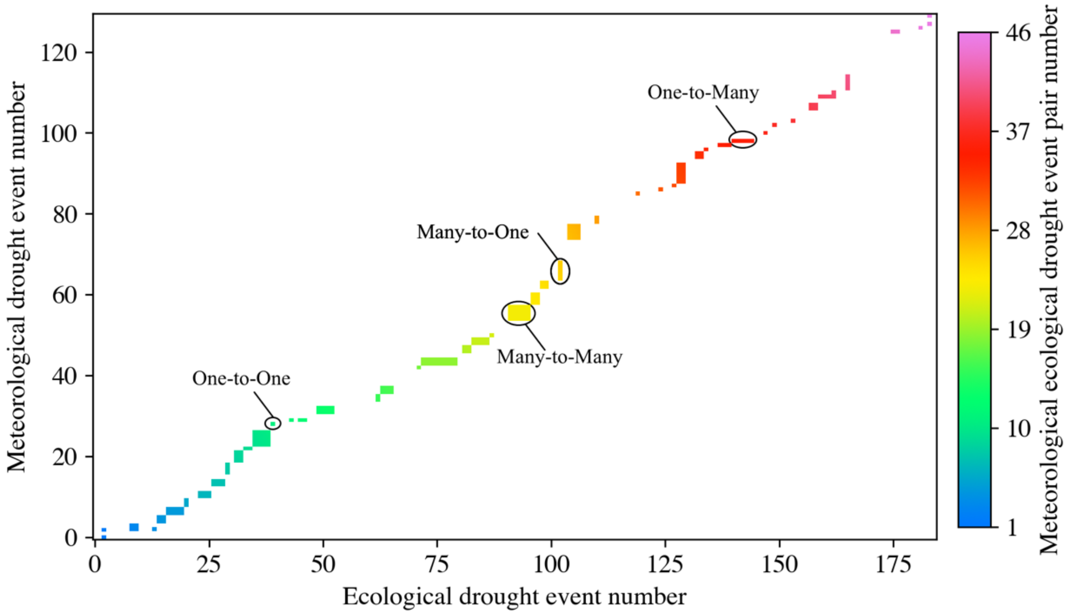
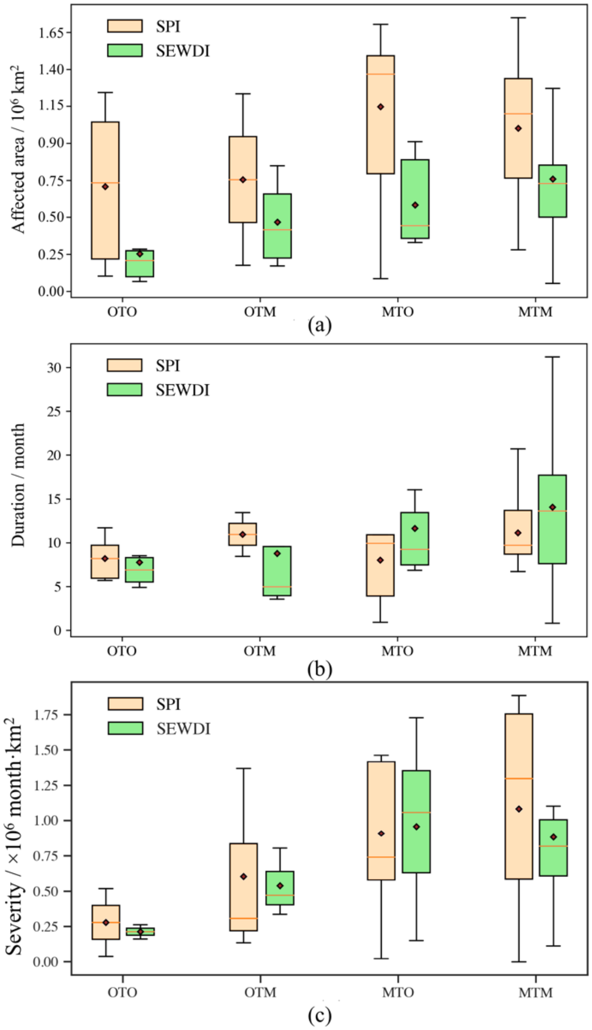


Figure 4: Identification results of paired meteorological and ecological drought events.



625 **Figure 5. A box plot showing the intensity, duration, and affected area of paired meteorological-ecological drought among different types**

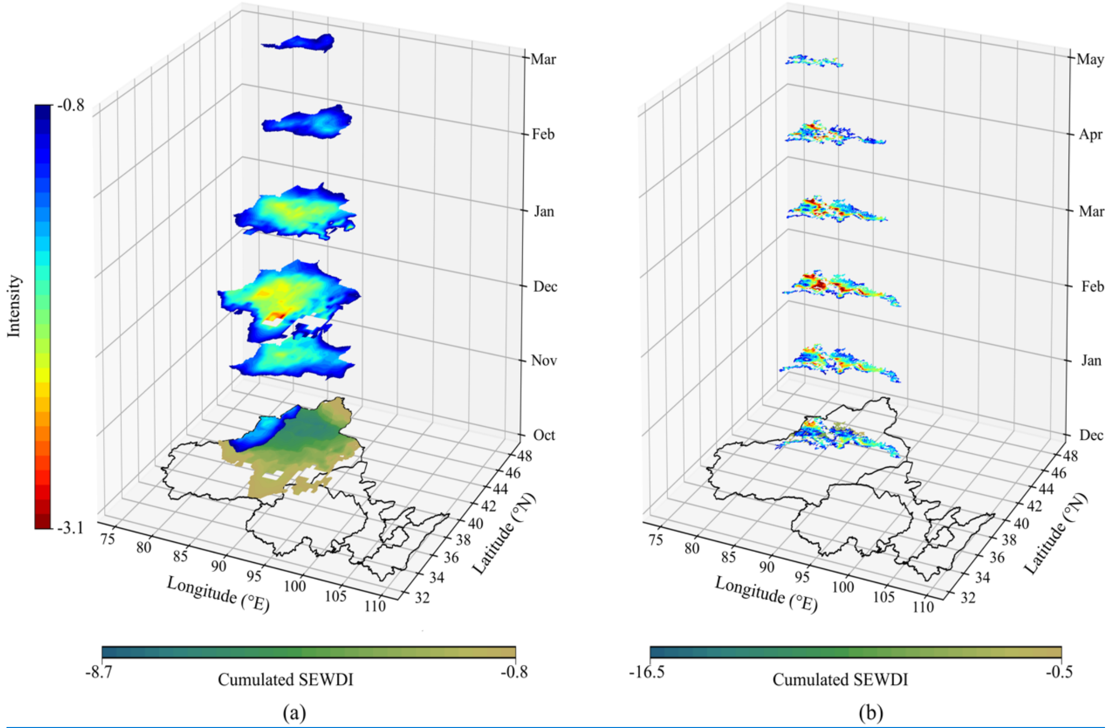
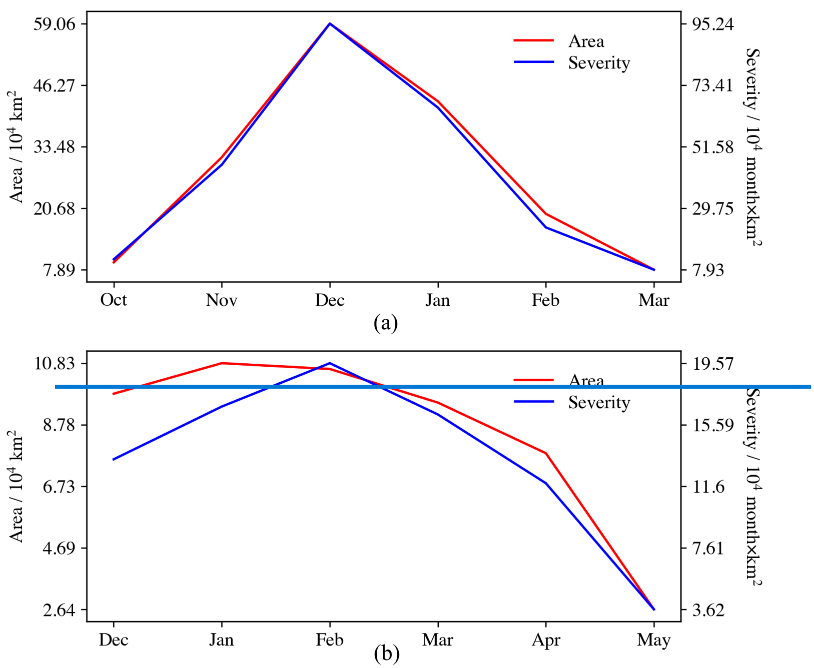


Figure 6: Spatio-temporal continuums of (a) meteorological drought event No. 87 and (b) ecological drought event No. 127.



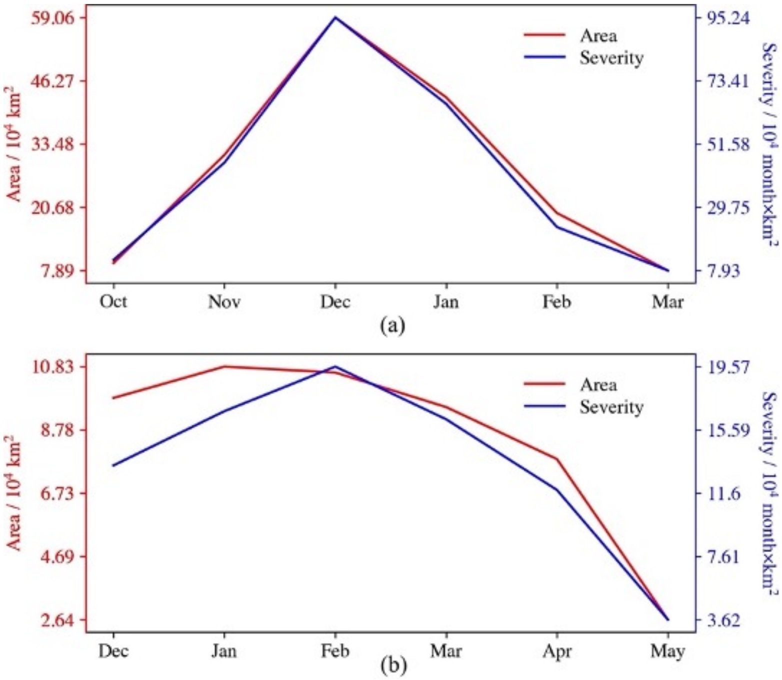


Figure ~~5:~~7: Temporal evolution of DS and DA of (a) meteorological drought event No. 87 and (b) ecological drought event No. 127.

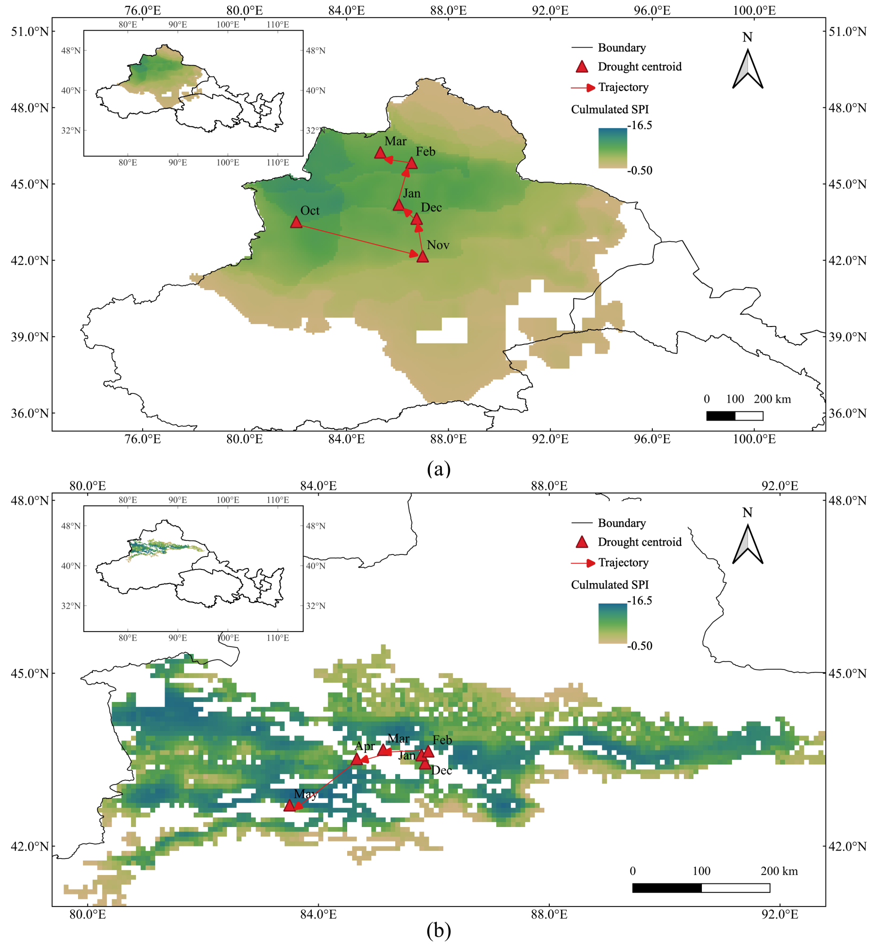


Figure ~~6:~~8: Cumulative SPI/SEWDI and migration trajectory of (a) meteorological drought event No.87 and (b) ecological drought event No.127.

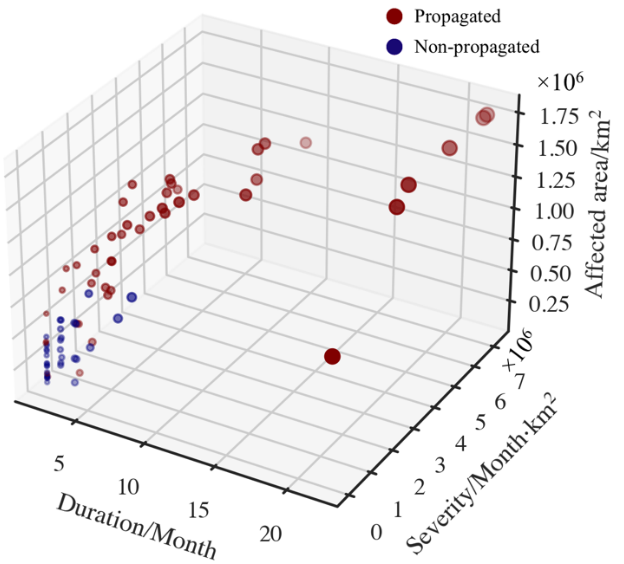


Figure ~~7:~~9: Three-dimensional diagram showing characteristics of meteorological drought events. Larger circles indicate greater severity.

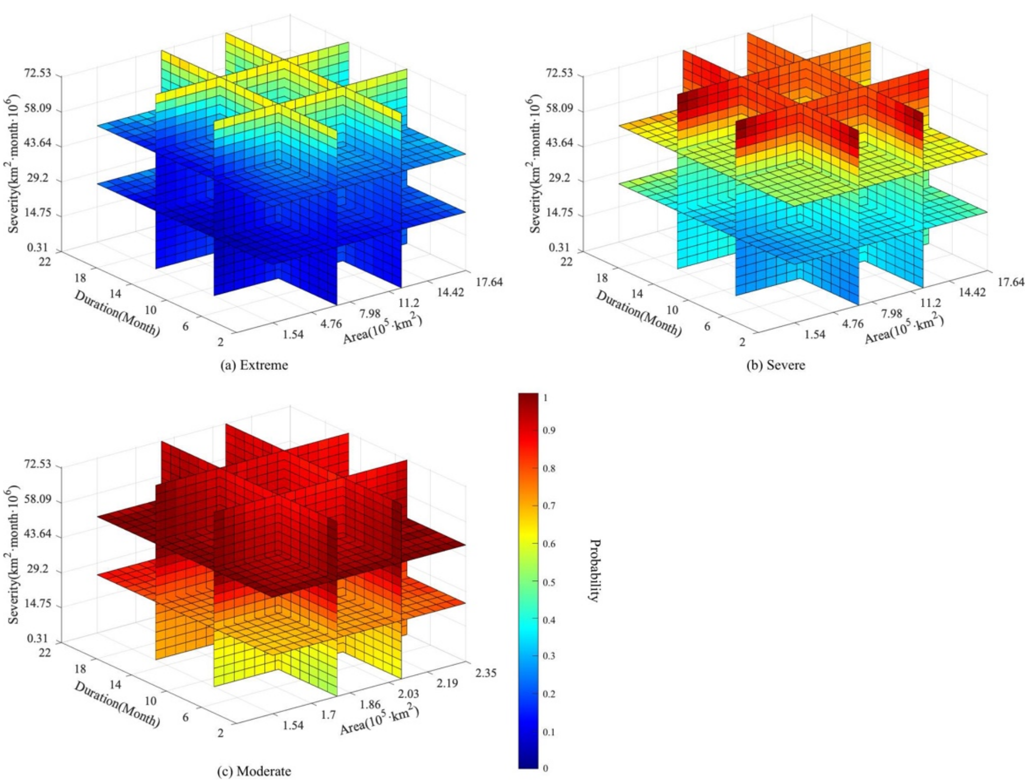


Figure 10: Conditional probability of ecological drought at ~~different~~(a) extreme, (b) severe, and (c) moderate levels, given that characteristics of meteorological drought ~~exceeding~~exceed a certain value.

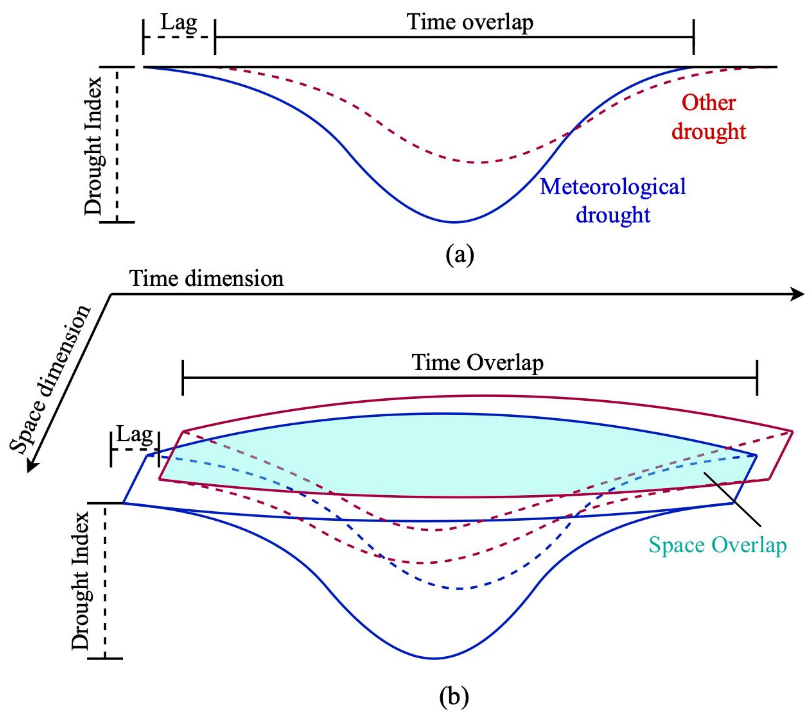


Figure 11: Conceptual graph depicting (a) traditional and (b) spatial and temporal connectivity rule of two drought types.

Table 1 Sensitivity test of parameter *b*

Threshold	Number of paired drought events
min(AM_i,AH_j)*90%	23
min(AM_i,AH_j)*70%	32
min(AM_i,AH_j)*50%	36
min(AM_i,AH_j)*30%	39
min(AM_i,AH_j)*15%	46
min(AM_i,AH_j)*10%	46

Table ~~1~~2 Top ten meteorological drought events according to severity

No.	**Affected area (km^2)**	**Duration (month)**	**Severity (month · km^2)**	**Start time (year–month)**	**End time (year–month)**
9	1764139.2	22	70730086	1985–08	1987–05

15	1675168.9	12	35741813	1997–01	1997–12
74	1511945.1	8	29594926	2001–02	2001–09
88	1610084.0	6	28099641	2008–04	2008–09
0	1613507.4	8	21190554	1982–02	1982–09
46	1407943.0	6	19184637	1995–03	1995–08
3	1553813.5	6	17818477	1983–07	1983–12
64	1194351.2	4	17346922	2000–02	2000–05
120	954552.3	8	2642278	2017–10	2018–05
115	471019.5	5	1915975	2015–12	2016–04

Table ~~2~~3 Top ten ecological drought events according to the severity

No.	Affected area (km^2)	Duration (month)	Severity (month・km^2)	Start (year–month)	End (year–month)
2	390824.2	25	4446071	1982–04	1984–04
35	347893.4	24	4197729	1986–07	1988–06
50	407626.7	30	4182267	1990–06	1992–11
37	348522.9	21	4047585	1986–10	1988–06
3	371975.6	18	3732552	1982–04	1983–09
59	407626.7	21	3566368	1991–03	1992–11
49	399717.3	27	3555634	1990–06	1992–08
56	391178.4	23	3360346	1991–01	1992–11
58	399638.6	18	3124954	1991–03	1992–08
55	120839.9	20	3085452	1991–01	1992–08

Table ~~3~~4 Estimations of 11 machine learning models in identifying the potential of meteorological drought to trigger ecological drought

Classifier	Accuracy	Precision	Recall	F1 Score	MCC	Total
KN	0.89	0.89	0.91	0.89	0.80	4.38
SVM	0.80	0.84	0.83	0.80	0.67	3.94
GP	0.43	0.22	0.50	0.30	0.00	1.45
DT	0.83	0.84	0.84	0.82	0.68	4.02
MP	0.62	0.40	0.59	0.46	0.17	2.24
AB	0.82	0.83	0.82	0.81	0.65	3.92
GNB	0.85	0.88	0.88	0.85	0.75	4.21
QDA	**0.93**	**0.93**	**0.94**	**0.93**	**0.87**	**4.58**
GB	0.83	0.83	0.84	0.82	0.67	4.00

XGB	0.85	0.86	0.87	0.85	0.72	4.15
RF	0.87	0.87	0.89	0.86	0.76	4.25

Table ~~4~~5 Goodness of fit of the marginal distribution

Distribution	Marginal distribution	RMSE	KS-test	
			Statistics	P-value
M_Area	~~Johnsonsb~~Johnson S_B	0.044	0.129	0.963
M_Duration	~~Johnsonsb~~Johnson S_B	0.068	0.161	0.823
M_Severity	Pearson III	0.057	0.226	0.413
E_Severity	~~Johnsonsb~~Johnson S_B	0.079	0.194	0.615

Table ~~5~~6 Goodness of fit of the bivariate distribution

—Joint variables	Copula	RMSE	CvM_test	
			Statistic	P-value
M_Area–M_Duration	Frank	0.005	0.086	0.373
M_Area–M_Severity	Gaussian	0.052	0.098	0.605
M_Area–E_Severtiy	Gumbel	0.032	0.042	0.933
M_Duration–M_Severity	Gaussian	0.057	0.102	0.585
M_Duration–E_Severity	Gaussian	0.053	0.087	0.663
M_Severity–E_Severity	Frank	0.054	0.105	0.570

Table ~~6~~7 Goodness of fit of the multivariate distribution

—Joint variables	RMSE	CvM_test	
		Statistic	P-value
M_Area–M_Duration–M_Severity–E_Severity	0.079	0.073	0.398

~~Table 7~~Table 8 E_DS with polynomial functions based on meteorological drought characteristics

Model types	Expression	Assessment metrics			
		RMSE	AIC	BIC	R^2
Ternary linear model	$E\_DS=4.85\times10^5+0.15M\_DS+4099.35M\_DD-1.20M\_DA$	9.24×10^5	1350.67	1357.89	0.58
Ternary quadratic model	$E\_DS=1.54-0.05M\_DS-16.91M\_DD-0.08M\_DA-1319.23M\_DD^2+0.03M\_DD\times M\_DA$	7.29×10^5	1085.75	1100.20	0.85