Drought prediction models driven by meteorological and remote sensing data in Guanzhong Area, China

Jianzhu Li, Siyao Zhang, Lingmei Huang, Ting Zhang and Ping Feng

ABSTRACT

Drought is an important factor that limits economic and social development due to its frequent occurrence and profound influence. Therefore, it is of great significance to make accurate predictions of drought for early warning and disaster alleviation. In this paper, SPEI-1 was confirmed to classify drought grades in the Guanzhong Area, and the autoregressive integrated moving average (ARIMA), random forest (RF) and support vector machine (SVM) model were established. Meteorological data and remote sensing data were used to derive the prediction models. The results showed that (1) the SVM model performed the best when the models were developed using meteorological data, remote sensing data and a combination of meteorological and remote sensing data, but the model's corresponding kernel functions are different and include linear, polynomial and Gaussian radial basis kernel functions, respectively. (2) The RF model driven by the remote sensing data and the SVM model driven by the combined meteorological and remote sensing data were found to perform better than the model driven by the corresponding other data in the Guanzhong Area. It is difficult to accurately measure drought with the single meteorological data. Only by considering the combined factors can we more accurately monitor and predict drought. This study can provide an important scientific basis for regional drought warnings and predictions.

Key words | drought prediction, random forest model, standardized precipitation evaporation index, support vector machine model, the integrated autoregressive moving average model

HIGHLIGHTS

- SPEI-1 was used to analyze the temporal distribution characteristics of drought and the main driving factors in Guanzhong Area, China.
- Drought grades were selected as the dependent variable, and the meteorological, geographical and vegetative factors were selected as the independent variables to establish an autoregressive integrated moving average (ARIMA) model, random forest (RF) model and support vector machine model.
- Meteorological data and remote sensing data were used as independent variables to derive prediction models, respectively.
- Comparing the models driven by remote sensing data only and the combination of meteorological and remote sensing data, the RF model driven by the remote sensing data and the SVM model driven by the combined meteorological and remote sensing data were found to perform better than the model driven by the corresponding other data in Guanzhong Area.
- This study can provide an important scientific basis for regional drought warning and prediction.

Jianzhu Li
Siyao Zhang
Ting Zhang (corresponding author)
Ping Feng
State Key Laboratory of Hydraulic Engineering Simulation and Safety,
Tianjin University,
Tianjin,
China
E-mail: *zhangting_hydro@tju.edu.cn*

Lingmei Huang
Faculty of Water Resources and Hydroelectric Engineering,
Xi'an University of Technology,
Xi'an,
China

doi: 10.2166/nh.2020.184

INTRODUCTION

As a severe natural disaster, drought not only affects economic development but also influences water resources, agriculture, ecology and environments (Mazdiyasni & Aghakouchak 2015; Malik *et al.* 2019; Zhang *et al.* 2019a, 2019b; Guo *et al.* 2020). Due to the uncertainty in when droughts begin and end, it is very difficult to predict drought. Therefore, drought has become one of the critical factors limiting the sustainable development of the economy and society in many areas (Yuan & Zhou 2004; Zhang *et al.* 2019a, 2019b; Dai *et al.* 2020). As it increases in severity, drought has a profound impact on biogeochemical processes in terrestrial ecosystems (Fang *et al.* 2018). Drought also affects the water absorption of vegetation by affecting soil water content and increases the sensitivity of vegetation to water energy, which plays an important role in terrestrial water, energy and carbon cycles (Fang *et al.* 2019; Yinglan *et al.* 2019a, 2019b). A warming climate is expected to perturb the hydrological cycle, resulting in changes in both the frequency and duration of drought (Han *et al.* 2018). To alleviate drought effects, it is of great significance to strengthen the study of regional drought characteristics and to make more accurate drought predictions for early warnings and disaster mitigation (Miao 2008).

The drought index is a primary factor in drought prediction, and many indices were presented in the previous studies for different drought types. The standardized precipitation index (abbreviated as SPI, see McKee *et al.* 1993; Bonaccorso *et al.* 2003) and the Palmer drought severity index (abbreviated as PDSI, refer to Palmer 1965) are two common indicators used to characterize regional drought. The SPI was proposed by McKee *et al.* (1993) and has been widely used as an indicator of drought. This index only requires long-term (generally more than 30 years) precipitation data (Che & Li 2010; Li *et al.* 2016), but it cannot reflect the seasonal distribution characteristics of precipitation (Dong & Xie 2014). The PDSI considers the balance of water resources, including precipitation and evaporation processes, additional runoff values, soil moisture content and other conditions (Zargar *et al.* 2011), and the PDSI has been extensively used to monitor long-term drought (Liu *et al.* 2004). This index also considers water shortages in crops and has shown better performances in capturing spatial and temporal characteristics (Yuan 2015). Based on SPI and PDSI concepts, Vicente-Serrano *et al.* (2010) proposed a standardized precipitation evaporation index (SPEI) (Yuan & Zhou 2004; Vicente-Serrano *et al.* 2010), which can express the dry and wet conditions of the land surface on multiple scales with the potential evaporation included. Zhuang *et al.* (2013) and Wang & Chen (2014) investigated the application of the SPEI in China and found that this index has good applicability in China and can accurately represent the occurrence of drought. Hernandez & Annette (2014) applied SPI and SPEI in conjunction with precipitation and temperature projections from two general circulation models at six major urban centers of south Texas spanning five climatic zones. Both the models predicted a progressively increasing aridity in the region throughout the 21st century. Li *et al.* (2016) investigated the spatiotemporal characteristics of drought in the Weihe River Basin by employing the SPEI index. Vega *et al.* (2019) investigated hydrological patterns in the Brazilian rainforest through a 9-month SPEI series and determined the Hurst exponents from detrended time series of days with precipitation and accumulated monthly rainfall. The researchers found that the Hurst exponent correlated positively with the monthly mean rainfall. The SPEI not only considers temperature and precipitation but also evapotranspiration (Lu 2018), and it has been confirmed to be applicable and better than other drought indices in the Guanzhong Area (Xu 2018). Therefore, the SPEI was selected as the drought index to evaluate the drought events in the Guanzhong Area in this paper.

The traditional drought index is usually based on hydrometeorological data measured at stations, and the spatial resolution of the index does not necessarily meet the requirements of drought monitoring in large-scale regions. Meteorological satellites, which are widely used in drought remote sensing monitoring (Di *et al.* 1994; Sahoo *et al.* 2015), can acquire multitemporal, multispectral, continuous and complete data (Kogan 1995; Yilmaz *et al.* 2008). At present, the most commonly used and better-performing remote sensing drought index is the normalized vegetation index

(NDVI) (Farrar *et al.* 1994; Bannari *et al.* 1995; Sun 2017), which can effectively monitor the dynamic changes in vegetation cover (Kogan 1990; Chen *et al.* 2010). Ichii *et al.* (2005) suggested that there was a relative difference in the correlation between NDVI and natural meteorological factors at different geographical latitudes and that NDVI was significantly correlated with temperature as well as regional precipitation. Yan *et al.* (2012) studied the relationship between NDVI and climate data in coastal areas of the Jiangsu Province and confirmed that both temperature and precipitation showed a significantly positive correlation with NDVI data in the region.

To reduce the losses caused by drought, drought prediction is needed on the basis of real-time drought monitoring. Commonly used prediction methods include time series analysis, artificial neural networks (Moody & Darken 1989), support vector machines (SVMs) (Ahmad *et al.* 2010; Weng 2012) and random forests (RFs) (Breiman 2001; Gislason *et al.* 2006; Cong 2017). The autoregressive integrated moving average (ARIMA) (Box & Jenkins 1976; Shumway & Stoffer 1982) model proposed by Box & Jenkins (1976) is a common model used in time series analysis (Han *et al.* 2012). Yurekli *et al.* (2005) used the ARIMA model to simulate the 5-year monthly runoff observation data of the Kelkit River (Bai *et al.* 2015). Zhang *et al.* (2017) used the ARIMA model to predict the drought of the northern Haihe River in China, indicating that the ARIMA model has a good prediction accuracy. Machine learning has been widely used to predict drought. For example, Fan *et al.* (2011) established a drought prediction model in autumn in the Zhejiang Province based on the SVM method with a radial basis kernel function and a cross-validation approach. The optimal model parameters were then determined, and they concluded that the developed model has high prediction accuracy (Fan *et al.* 2011). Wu *et al.* (2016) employed the RF model to analyze the drought grades of 21 representative stations from 1962 to 2012 in the Huaihe River Basin. The researchers found that the overall average prediction accuracy is higher than the weather system's weather prediction accuracy, suggesting acceptable prediction results. ARIMA, RF and SVM have been used in drought prediction, but the performance of these three methods needs to be further explored to improve the accuracy of drought prediction in the Guanzhong Area.

The main aims of this study are (1) to analyze the temporal distribution characteristics of drought and the main driving factors in the Guanzhong Area, China; (2) to build an ARIMA model, RF model and SVM model, and predict drought grades by using gauge-based and remote sensing monitoring data, respectively; and (3) to select the best prediction method and the corresponding predictors.

STUDY AREA AND DATA

Study area

The Guanzhong Area is located in the middle of the Shaanxi Province, between the Loess Plateau and the Qinling Mountains, China. The geographical coordinates are 33°34′N–35°52′N, 106°18′E–110°38′E, with an area of 36,000 km^2. The administrative divisions in the Guanzhong Area include the cities of Xi'an, Xianyang, Weinan, Baoji and Tongchuan (Wei *et al.* 2015), and this area is an important area connecting the eastern and western parts and the northern and southern parts of China (Qiao 2015). Additionally, this area is a key construction area of the Belt and Road Initiative.

The topography and geomorphology of the Guanzhong Area are complex with elevations between 270 and 2,439 m. The altitudes are high in the western, northern and southern regions and low in the eastern and middle areas (Figure 1). The topography descends in steps from the mountainous area to the center of the basin and in turn includes the piedmont alluvial plain, the loess plateau and the terrace of the river valley (Qiao 2015).

The Guanzhong Area is located in the transitional zone between arid and humid regions (Li *et al.* 2016), which belongs to a continental monsoon climate with cold winters, hot summers and distinct seasonal features. During the same period of heat and rain, drought is more likely to occur in the Guanzhong Area. The average annual precipitation is 500–700 mm, with precipitation concentrated mostly in summer and autumn and little precipitation occurs in winter (Qiao 2015).

The special topography results in climatic conditions of high temperature and low rainfall. The area is mostly plain formed by loess sedimentation and river alluvial sediments

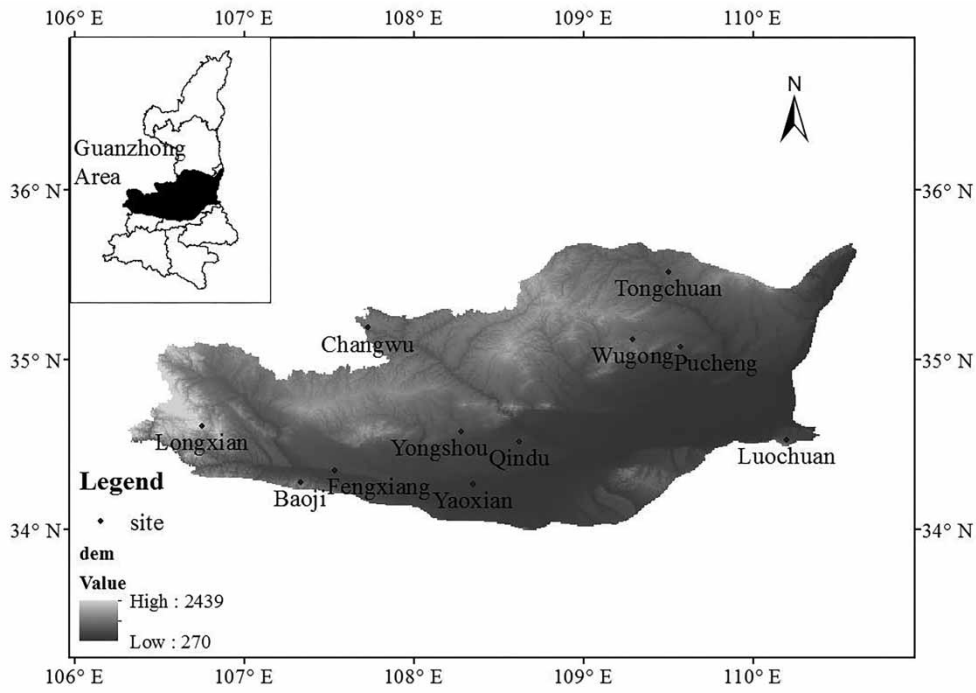

Figure 1 | Location of the Guanzhong Area and the meteorological stations.

with soft soil, poor water retention capacity, and therefore, the area is prone to drought. In terms of agriculture, approximately 70% of the land on the Guanzhong Plain is covered by cropland, which mainly includes grain crops, fruit woodlands and vegetables. More than 50% of the croplands in this area are rainfed. There are many irrigated croplands in the western and middle parts, while most of the croplands in the east are rainfed (Zhou *et al.* 2019). Due to the dense population and developed agriculture, the increasing industrial and agricultural water supply led to a more severe drought situation. Thus, the Guanzhong Area is known by the saying, '9 drought years out of 10 years' (Yu 2015).

According to the 'Record of Natural Disasters in the Shaanxi Province', from the 2nd century to 1949 AD, there were more than 600 drought events in the Shaanxi Province, and 326 in the Guanzhong Area, accounting for 54% of the total drought records (Meteorological Station of Shaanxi Meteorological Bureau 1976). There were 22 large-scale drought events in the Guanzhong Area from 1949 to present, especially the continuous severe droughts that occurred in 1959–1961, 1966–1967, 1971–1972, 1976–1980 and 1994–1997 (Gao 2015). Droughts in the Guanzhong Area occurred mostly in spring and summer, and continuous drought events frequently occurred, lasting from spring to summer and from summer to autumn (Wu 2018).

Data

The data used for drought prediction in the Guanzhong Area include gauge-based meteorological data and remote sensing data.

The gauge-based meteorological data include the daily wind speed, precipitation, air temperature, air pressure, sunshine hours and relative humidity of 11 meteorological stations in the Guanzhong Area, covering the period from January 1960 to December 2016. The data were downloaded from the Meteorological Data Sharing Service System of the China Meteorological Administration (http://data.cma.cn/).

The remote sensing data are the land surface temperature of the Guanzhong Area and the NDVI from July 2002 to June 2011. These data were downloaded from the International Scientific Data Mirror website of the Chinese

Academy of Sciences Computer Network Information Center (http://www.gscloud.cn). The digital elevation model (DEM) was also obtained from this website. Detailed information on the remote sensing data used in this study is shown in Table 1.

METHODS

Calculation of the SPEI-1 series

The SPEI index is calculated as follows:

(1) Calculate the monthly water surplus and deficiency D_i, $D_i = P_i - \mathrm{PET}_i$, with i indicating the month, P_i indicating the monthly precipitation (mm) and PET_i the monthly potential evaporation (mm), which is calculated using the Morton-type Penman formula (Mei 2012).
(2) Calculate the probability distribution of the monthly water surplus and deficit D_i by using the log–logistic probability distribution function.
(3) Standardize the log–logistic probability distribution function $F(x)$ for monthly accumulated water surplus and deficit D_i (Li *et al.* 2016).

The severity of drought can be graded according to the value of SPEI, as shown in Table 2 (Zhang *et al.* 2014).

According to the multi time scale characteristics of the SPEI, the SPEI on different time scales can reflect changes in humidity and dryness in different periods (Yang *et al.* 2018). SPEI-1 refers to the SPEI index on a monthly time scale. The trend of SPEI values on different time scales is consistent overall, but the trend of humidity and dryness reflected by SPEI values on short time scales is more specific. Therefore, SPEI-1 is selected in this paper.

Table 1 | Detailed information on the remote sensing data

Production name	Spatial resolution	Temporal resolution
China 1 km land temperature monthly synthetic products (MYDLT1M)	1 km	month
China 500 m NDVI synthetic products (MYDND1M)	500 m	month
30 m resolution DEM data (GDEMV2)	30 m	–

Table 2 | The classification of drought based on SPEI

SPEI	Drought grades	Number
(−0.5,0)	no drought	1
(−1, − 0.5]	light drought	2
(−1.5, − 1]	moderate drought	3
(−2, − 1.5]	heavy drought	4
− ∞, − 2]	severe drought	5

Predictive models

ARIMA model

The ARIMA model is one of the most commonly used models in time series analyses (Zhao *et al.* 2012). The data series are formed by the prediction index with time regarded as a random sequence. The dependence relation of this random series reflects the continuity of the original data in time, which has both the influence of external factors and its own change laws (Wu 2014). The ARIMA modeling steps are as follows:

(1) Determine if the time series is stationary. The time series can be differentiated to obtain a stationary series if it is a non-stationary series. The difference operators can be explained as each observation minus the previous one, in which d represents the difference times. The autoregressive moving average (ARMA) model can be described as ARMA (p,q), and the ARIMA model can be described as ARIMA (p,d,q). When the time series is stationary, $d = 0$, and the ARIMA model becomes the ARMA model (Ghashghaie & Nozari 2018).
(2) Determine p and q (p represents the lag order of the autoregressive processes, and q represents the lag order of the moving average processes (Ghashghaie & Nozari 2018)). The autocorrelation function (ACF) graph and the partial ACF (PACF) graph are drawn, and the ARMA(p,q) models are judged based on the tailing and truncation of the ACF and PACF graph (Wu 2014).
(3) Fit the model and perform the normality test, autocorrelation test and white noise test on the residuals.
(4) Use the ARIMA model to inspect and predict the monitoring series. For the seasonal non-stationary time series,

seasonal differences are also needed to obtain the stationary time series and form the ARIMA (p,d,q) $(P, D,Q)^S$ model, where D represents the order of seasonal differences; P and Q represent the order of seasonal autoregression and the order of seasonal moving averages, respectively; and S represents the length of the seasonal period.

RF model

The RF model uses the bootstrap resampling method to extract multiple samples from the original samples, builds the decision tree for each bootstrap sample, and then combines the predictions of multiple decision trees to obtain the final prediction results by voting (Breiman 2001; Fang *et al.* 2011). Specifically, the RF model is a combined classification model composed of many decision tree classification models ($h(X,\theta_i)$, $i = 1, 2, \ldots, k$), and the parameter set (θ_i) is an independent and identically distributed random vector. Given the independent variable X, each decision tree classification model selects the optimal classification results by one-vote voting rights (Wang 2017). The modeling steps for the RF model are as follows:

(1) Randomly extract M samples from the fitting data set T, and then fit the extracted data set.
(2) For each instance, generate a decision tree, and at each node of the tree:
    a. Randomly extract a subset of m variables from the p valid overall features;
    b. Choose the best variables and the best partition from the set of m variables; and
    c. Continue until the tree is fully generated (Wang 2017).
(3) Perform RF predictions using all trees.

SVM model

The SVM model seeks the best compromise between the complexity of the model (learning-intensive reading of specific training samples) and the ability to learn (the ability to identify any samples without error) based on limited sample information for better promotion ability (Wang 2017).

The SVM was originally proposed for the problem of binary pattern classification in the case of linear separability. Given a set of observation samples $S = \{(x_1, y_1), (x_2, y_2)\ldots\} \subset X \times \{-1, 1\}$, $X \subset R_n$ is called the input space or the input characteristic space, and $y_i \in \{-1,1\}$ is the sample class tag. The purpose of classification is to find a classification hyperplane that completely separates the two classes. Let $G = \{\omega \cdot x + b = 0 |\ \omega \in R_n,\ x \in X,\ b \in R\}$ be all hyperplane sets that can completely and correctly classify S. In all hyperplanes, the maximum interval classifier is looking for an optimal hyperplane that satisfies the two types of classification intervals (the sum of the sample-to-hyperplane distance closest to each other from the hyperplane). The linear inseparable problem of the input space can be transformed into a linear separable problem by using the appropriate kernel functions to operate in two classes. Common kernel functions are linear kernel functions (Altman 1992), polynomial kernel functions, Gaussian radial basis kernel functions and sigmoid kernel functions (Wang 2017; Burges 1998).

RESULTS AND DISCUSSION

Calculation of the SPEI-1 series

The SPEI-1 series of 11 meteorological stations in the Guanzhong Area are obtained based on monthly meteorological data. Here, we show only the results of Baoji Station in Figure 2.

The SPEI-1 series of the Guanzhong Area was calculated according to the SPEI-1 series at the 11 stations, as shown in Figure 3. The Guanzhong Area is prone to drought disasters because of the frequent alternation of dry and wet conditions. In 1962, 1963, 1967, 1969, 1976, 1979, 1994–2002 and 2007, the Guanzhong Area suffered severe drought events as shown in Figure 3 based on the drought classification in Table 2.

The occurrence frequency and percentage of drought at all grades in 12 months at Baoji Station are acquired and shown in Table 3. Moderate drought mainly occurs in autumn, while the occurrence of heavy drought is concentrated in spring and summer. In addition, Baoji Station experienced no severe drought during 1960–2016.

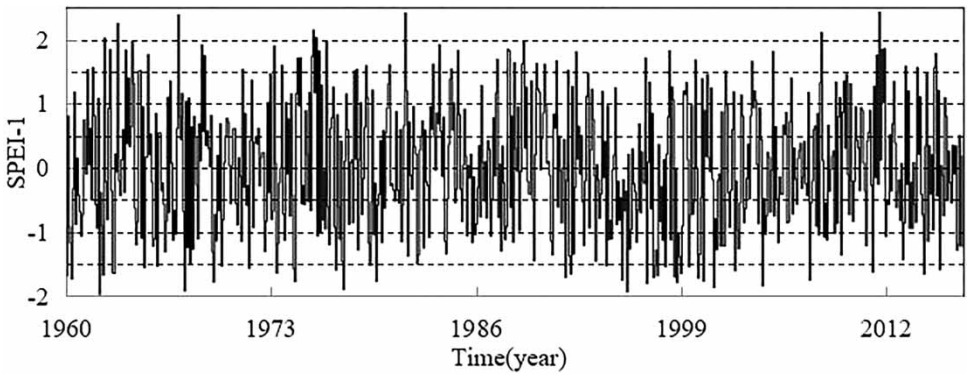

Figure 2 | SPEI-1 series of the Baoji Station.

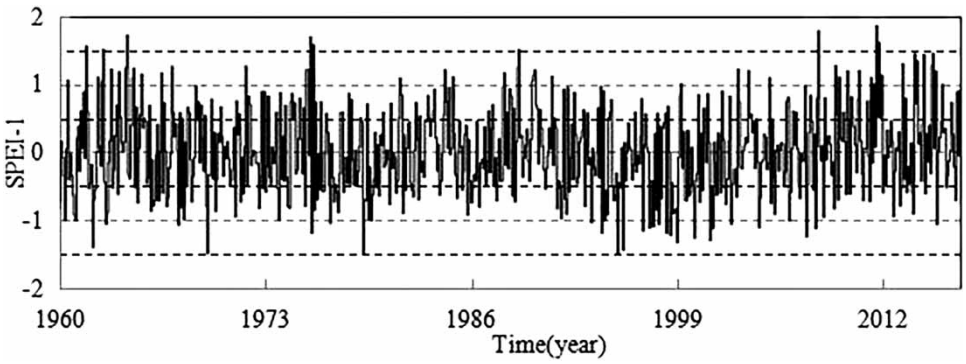

Figure 3 | SPEI-1 series of the Guanzhong Area.

Table 3 | The number and proportion of droughts at all grades in 12 months of Baoji Station

Month	No drought		Light drought		Moderate drought		Heavy drought		Severe drought	
1	34	7.49%	13	12.26%	7	9.21%	2	4.76%	0	–
2	41	9.03%	6	5.66%	5	6.58%	4	9.52%	0	–
3	37	8.15%	11	10.38%	5	6.58%	4	9.52%	0	–
4	35	7.71%	12	11.32%	6	7.89%	3	7.14%	0	–
5	40	8.81%	6	5.66%	5	6.58%	5	11.90%	0	–
6	39	8.59%	7	6.60%	6	7.89%	4	9.52%	0	–
7	37	8.15%	10	9.43%	5	6.58%	5	11.90%	0	–
8	40	8.81%	6	5.66%	7	9.21%	4	9.52%	0	–
9	41	9.03%	5	4.72%	6	7.89%	4	9.52%	0	–
10	37	8.15%	10	9.43%	8	10.53%	2	4.76%	0	–
11	36	7.93%	9	8.49%	10	13.16%	2	4.76%	0	–
12	37	8.15%	11	10.38%	6	7.89%	3	7.14%	0	–

In the 'History of Natural Disasters in the Shaanxi Province' and 'Report on China's Disaster Situation 1949–1995' (Liu 2012), the Guanzhong Area suffered flood disasters in the mid-1960s and early 1980s, and among them, 1964, 1983 and 1984 were serious flooding years. There were drought events in the early 1960s, mid-late 1970s and 1990s, among which, the continuous drought events that occurred in 1959–1961, 1966–1967, 1971–1972, 1976–1980, 1994–1997 and 1999–2002 were severe. Drought and floods occurred alternately from the mid-1980s to the early 1990s, and the climate was more humid in the 21st century (Cai *et al.* 2013; Lei *et al.* 2016). The wet and dry conditions in Figure 3 are consistent with previous studies. Therefore, the SPEI-1 can reasonably reflect the occurrence of drought in the Guanzhong Area.

The number and proportion of droughts at all grades were calculated according to the SPEI-1 series at the 11 stations, as shown in Table 4. The proportion of droughts at 11 stations is nearly 35%, mainly light and moderate droughts, and severe droughts hardly occur. In addition, the number of droughts at all grades decreases with the severity of drought.

Drought prediction by different models

The three models previously mentioned in the 'Predictive models' section were used to predict the drought grades in the Guanzhong Area, with a predicted length of 12 months.

Drought prediction by the ARIMA model

The SPEI-1 series of the 11 meteorological stations in the Guanzhong Area were used for ARIMA modeling and prediction. The ARIMA model was fitted using the SPEI-1 series from January 1960 to December 2015 and predicted drought grades from January to December 2016 with a length of 12 months. The ARIMA model parameters corresponding to each station are shown in Table 5, and the fitting and prediction results of Baoji Station are shown in Figure 4. The ARIMA model parameters at different stations have regional heterogeneity, depending on the natural conditions, such as the underlying surface, geographical location and climatic characteristics.

According to Figure 4(a), the SPEI-1 fitting series of the ARIMA model at each station is mostly consistent with the measuring series, which shows that the ARIMA model is basically in line with reality and that the fitting performs well. The prediction SPEI-1 series shown in Figure 4(b) fluctuates in the range of −1 to 1, suggesting that the Guanzhong Area was in a no drought or light drought state in 2016. There is no sudden increase or decrease in SPEI-1 values in 2016, and the change is relatively small. Among them, the prediction SPEI declined in 2016. The SPEI-1

Table 4 | The number and proportion of droughts at all grades

Station	No drought		Light drought		Moderate drought		Heavy drought		Severe drought	
Baoji	454	66.57%	108	15.84%	77	11.29%	43	6.30%	0	0.00%
Fengxiang	451	66.13%	106	15.54%	90	13.20%	35	5.13%	0	0.00%
Luochuan	440	64.42%	122	17.86%	89	13.03%	30	4.39%	2	0.29%
Tongchuan	459	67.30%	106	15.54%	77	11.29%	38	5.57%	2	0.29%
Wugong	447	65.54%	108	15.84%	93	13.64%	33	4.84%	1	0.15%
Yaoxian	458	67.06%	110	16.11%	73	10.69%	41	6.00%	1	0.15%
Changwu	456	66.86%	102	14.96%	87	12.76%	34	4.99%	3	0.44%
Pucheng	457	67.01%	105	15.40%	84	12.32%	33	4.84%	3	0.44%
Longxian	461	67.60%	108	15.84%	87	12.76%	23	3.37%	3	0.44%
Yongshou	456	66.86%	99	14.52%	92	13.49%	34	4.99%	1	0.15%
Qindu	452	66.08%	118	17.25%	81	11.84%	26	3.80%	7	1.02%
average	454	66.49%	108	15.88%	85	12.39%	34	4.93%	2	0.31%

Table 5 | The ARIMA model parameters corresponding to each station

Station	ARIMA$(p,d,q)(P,D,Q)^S$	Station	ARIMA$(p,d,q)(P,D,Q)^S$
Baoji	ARIMA$(5,4,3)(1,0,1)^{12}$	Fengxiang	ARIMA$(5,4,3)(1,0,2)^{12}$
Luochuan	ARIMA$(3,1,2)(0,1,0)^{12}$	Tongchuan	ARIMA$(2,2,3)(0,1,0)^{12}$
Wugong	ARIMA$(3,1,2)(0,1,0)^{12}$	Yaoxian	ARIMA$(1,1,2)(0,1,0)^{12}$
Changwu	ARIMA$(0,1,2)(1,1,0)^{12}$	Pucheng	ARIMA$(1,2,2)(0,1,0)^{12}$
Longxian	ARIMA$(1,1,2)(0,1,1)^{12}$	Yongshou	ARIMA$(4,3,3)(2,1,0)^{12}$
Qindu	ARIMA$(1,1,2)(0,1,0)^{12}$	Guanzhong Area	ARIMA$(0,1,2)(0,1,0)^{12}$

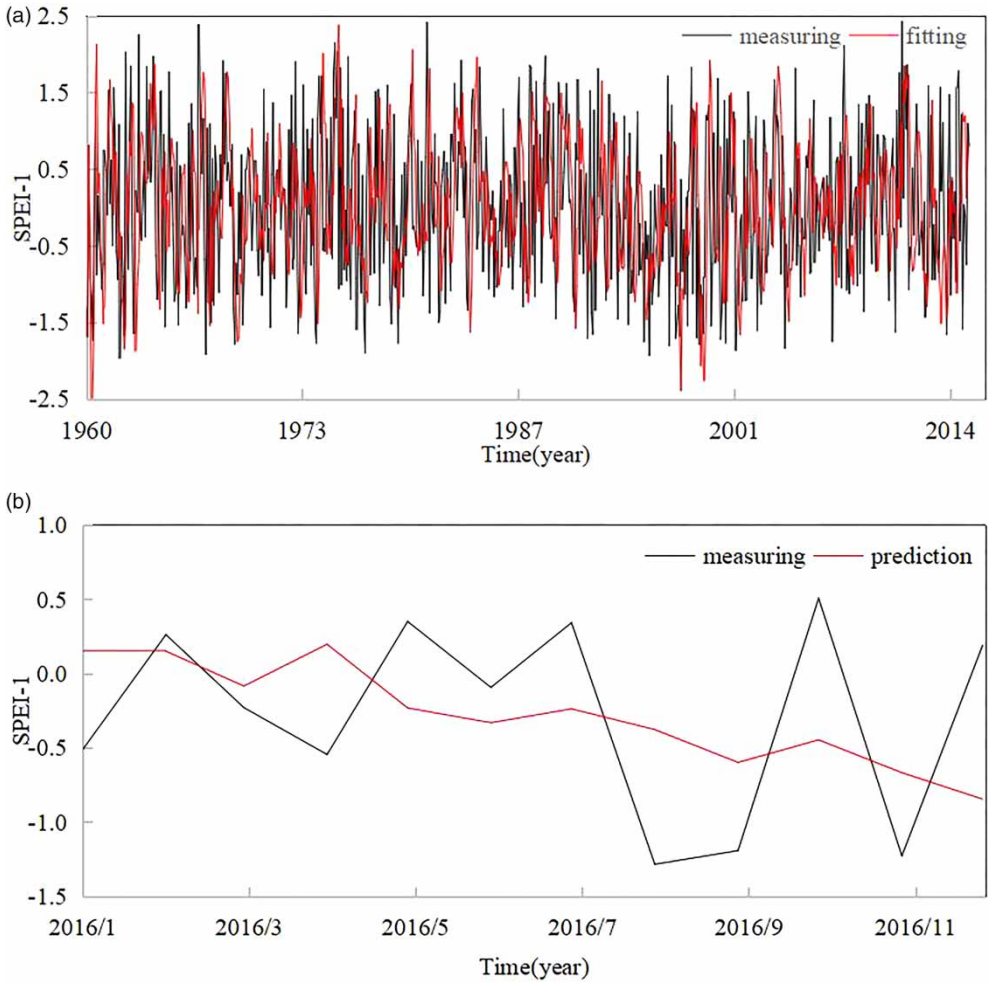

Figure 4 | Fitting and prediction results of the ARIMA model at Baoji Station. (a) Fitting results of the ARIMA model in Baoji Station. (b) Prediction results of the ARIMA model in Baoji Station.

was lower after August 2016 when the drought was serious. The SPEI-1 was at a higher level before August 2016, that is, the drought degree was lighter.

The fitting and prediction of SPEI values were graded according to Table 2, and the qualified rate of the drought grade was calculated. The qualified fitting and prediction

rates of the ARIMA model at each station are shown in Table 6. The drought prediction grades of the ARIMA model can represent the actual situation to some extent, and the average qualified fitting and prediction rates are 0.5337 and 0.6061, respectively. Except for the qualified prediction rate of the Longxian station, which is 0.7500 and is significantly higher than the average, the qualified rates of other stations fluctuate around the average. The case where the qualified fitting rate is lower than the qualified prediction rate may be due to the uneven length of the fitting and prediction data sets. The fitting data were used from January 1960 to December 2015 (56 years), and the prediction data were used from January to December 2016 (1 year), so that the internal variance in the fitting process was much larger than the prediction process, that is, the fitting process had a large error with respect to the prediction process.

Drought prediction by the RF model

RF model driven by meteorological data and accuracy assessment. SPEI-1 grade (no drought: 1, light drought: 2, moderate drought: 3, heavy drought: 4, special drought: 5) was selected as the dependent variable. The monthly mean values of daily maximum temperature, daily minimum temperature, average humidity, average wind speed, sunshine hours, rainfall and potential evaporation of the 11 meteorological stations in the Guanzhong Area were selected as the independent variables. The data from January 1960 to December 2015 were used as the fitting data, and the data from January 2016 to December 2016 were used as the prediction data with a predicted length of 12 months.

The qualified fitting and prediction rates corresponding to the 11 stations of the RF model constructed by the meteorological data were calculated and are shown in Table 7. The average qualified fitting and prediction rates are 0.7623 and 0.6515, respectively. The qualified prediction rates are significantly different from the average qualified prediction rates at some stations, such as Baoji, Tongchuan, Wugong and Yongshou with prediction accuracies of 0.5000, 0.7500, 0.8333 and 0.5000, respectively. The qualified fitting and prediction rates of the other stations fluctuate around the average. Due to the noise in the fitting data, the overfitting phenomenon is obvious in the model.

RF model driven by remote sensing data and accuracy assessment. After classifying the SPEI-1 grades of 11 stations in the Guanzhong Area, kriging spatial interpolation was performed to obtain the SPEI-1 spatial distribution. The DEM,

Table 6 | Qualified fitting and prediction rate of the ARIMA model

Station	Fitting	Prediction	Station	Fitting	Prediction
Baoji	0.5417	0.5000	Changwu	0.5372	0.5833
Fengxiang	0.5461	0.6667	Pucheng	0.4940	0.6667
Luochuan	0.4717	0.5000	Longxian	0.6741	0.7500
Tongchuan	0.5134	0.6667	Yongshou	0.5655	0.5833
Wugong	0.4985	0.5833	Qindu	0.5149	0.5833
Yaoxian	0.5134	0.5833	Average	0.5337	0.6061

Table 7 | Drought qualified fitting and prediction rates of the RF models in 11 stations (meteorological data)

Station	Fitting	Prediction	Station	Fitting	Prediction
Baoji	0.7887	0.5000	Changwu	0.7708	0.5833
Fengxiang	0.7723	0.5833	Pucheng	0.7827	0.8333
Luochuan	0.7426	0.5833	Longxian	0.7455	0.6667
Tongchuan	0.7753	0.7500	Yongshou	0.7307	0.5000
Wugong	0.7857	0.8333	Qindu	0.7321	0.6667
Yaoxian	0.7589	0.6667	Average	0.7623	0.6515

slope, slope direction, rainfall, NDVI, daytime land temperature (LTD) and nighttime land temperature (LTN) were used as independent variables, and SPEI-1 grades were used as the dependent variables. The RF model was constructed with the data from July 2002 to June 2010 as the fitting data and the data from July 2010 to June 2011 as the prediction data. The average qualified fitting and prediction rates of the RF model constructed by remote sensing data in the Guanzhong Area are 0.6746 and 0.6836, respectively. The independent variables corresponding to the 11 stations were extracted, and then, the qualified fitting and prediction rates were calculated as shown in Table 8.

As seen from Table 6, the qualified fitting and prediction rates of the 11 stations are 0.6913 and 0.5616, respectively. Except for the qualified fitting rate of Luochuan and the qualified prediction rate of Qindu, the qualified rates of other stations fluctuate around the average qualified rates. The model works well for the 11 stations, but the RF model built by using remote sensing data also presents a significant overfitting problem due to data noise.

RF model driven by combined meteorological and remote sensing data and accuracy assessment. The monthly mean values of daily maximum temperature, daily minimum temperature, average humidity, average wind speed, sunshine hours, rainfall and potential evaporation obtained from 11 meteorological stations in the Guanzhong Area were used as the independent variables, as well as DEM, slope, and slope direction, rainfall, NDVI, LTD and LTN obtained from remote sensing products. SPEI-1 grades were used as the dependent variable. The data from July 2002 to June 2010 were used for fitting the model, and the data from July 2010 to June 2011 were used for prediction. The qualified fitting and prediction rates corresponding to the 11 stations are shown in Table 9. The average qualified fitting and prediction rates are 0.6316 and 0.5388, respectively. Except for the qualified prediction rate of Longxian station (0.6667), which is significantly higher than the average qualified prediction rate, the qualified rates of the other stations fluctuate around the average.

Drought prediction by the SVM model

SVM model driven by meteorological data and accuracy assessment. By using the same fitting data and prediction data as in the 'RF model driven by meteorological data

Table 8 | Drought qualified fitting and prediction rates of the RF models at 11 stations (remote sensing data)

Station	Fitting	Prediction	Station	Fitting	Prediction
Baoji	0.6979	0.6563	Changwu	0.6458	0.4792
Fengxiang	0.6667	0.5625	Pucheng	0.7396	0.5313
Luochuan	0.5833	0.4896	Longxian	0.6979	0.6250
Tongchuan	0.7188	0.5000	Yongshou	0.6771	0.4896
Wugong	0.7500	0.6042	Qindu	0.7813	0.6771
Yaoxian	0.6458	0.5625	Average	0.6913	0.5616

Table 9 | Drought qualified fitting and prediction rates of the RF models in 11 stations (meteorological and remote sensing data)

Station	Fitting	Prediction	Station	Fitting	Prediction
Baoji	0.6667	0.6042	Changwu	0.5625	0.4583
Fengxiang	0.5938	0.4688	Pucheng	0.6667	0.5729
Luochuan	0.5625	0.5208	Longxian	0.6875	0.6667
Tongchuan	0.6042	0.4479	Yongshou	0.6354	0.4896
Wugong	0.6875	0.5938	Qindu	0.6875	0.6250
Yaoxian	0.5938	0.4792	Average	0.6316	0.5388

and accuracy assessment' section, four common kernel functions (linear kernel function, polynomial kernel function, Gaussian radial basis kernel function and sigmoid kernel function) were used to construct the SVM model as well as fit and predict the drought grades. The average qualified fitting and prediction rates at the 11 stations are shown in Table 10.

As shown in Table 10, the polynomial kernel function and the linear kernel function perform best in the fitting and prediction, respectively. The Gaussian radial basis kernel function ranks second for both the qualified fitting and prediction rates. The qualified fitting and prediction rates of the sigmoid kernel function differ by 0.0364, which is the smallest and the most stable among the four kernel functions, but the qualified rates of fitting and prediction are the lowest. The qualified rate of linear kernel function fitting is 0.7532, which is only 0.0437 smaller than the highest Gaussian radial basis kernel function, and the qualified rate is the highest in the prediction process. The difference between qualified fitting and prediction rates is 0.0498, which is only larger than the sigmoid kernel function (ranks first). Therefore, the linear kernel function is considered to be more suitable for drought monitoring in the Guanzhong Area.

SVM model driven by remote sensing data and accuracy assessment. Using the same fitting data and the prediction data shown in the 'RF model driven by remote sensing data and accuracy assessment' section, four kernel functions were used to construct the SVM model as well as fit and predict the drought grades. The qualified fitting and prediction rates of the four kernel functions are shown in Table 11.

The polynomial kernel function and the linear kernel function perform best in the fitting and prediction process, respectively. The qualified rate of the polynomial kernel function in the prediction ranks second, and it differs from the first-ranked linear kernel function by only 0.0081. In addition, the differences in the qualified fitting and prediction rates of the four kernel functions are less than 0.1, which shows that they are relatively stable. Therefore, the polynomial kernel function is more suitable for drought prediction with remote sensing data driving the SVM model.

Table 10 | The average qualified rates of four common kernel functions (meteorological data)

Kernel functions	Linear kernel function	Polynomial kernel function	Gaussian radial basis kernel function	Sigmoid kernel function
Fitting	0.7532	0.7969	0.7676	0.6273
Prediction	0.8030	0.6970	0.7121	0.5909
Difference	0.0498	0.0999	0.0555	0.0364

Table 11 | The qualified rates of four common kernel functions (remote sensing data)

Kernel functions	Linear kernel function	Polynomial kernel function	Gaussian radial basis kernel function	Sigmoid kernel function
Fitting	0.7247	0.7795	0.7523	0.6680
Prediction	0.7269	0.7188	0.6548	0.7016
Difference	0.0022	0.0607	0.0975	0.0336

SVM model driven by combined meteorological and remote sensing data and accuracy assessment. Using the same fitting data and prediction data as in the 'RF model driven by combined meteorological and remote sensing data and accuracy assessment' section, four kernel functions were used to construct the SVM model as well as fit and predict the drought grades. The qualified fitting and prediction rates of the four kernel functions are shown in Table 12.

The Gaussian radial basis kernel function and the linear kernel function perform best in the fitting and prediction process, respectively. The Gaussian radial basis kernel function ranks second with a qualified prediction rate of 0.7652 and only differs from the linear kernel function (ranks first) by 0.0075. That is, the Gaussian radial basis kernel function performs well in both fitting and prediction. Therefore, the Gaussian radial basis kernel function is more suitable for drought prediction with the SVM model driven by combined meteorological and remote sensing data.

Table 12 | The qualified rates of four common kernel functions (meteorological and remote sensing data)

Kernel functions	Linear kernel function	Polynomial kernel function	Gaussian radial basis kernel function	Sigmoid kernel function
Fitting	0.8532	0.7727	0.9441	0.6799
Prediction	0.7727	0.7348	0.7652	0.7273

Comparison of the models

Comparison of different models driven by meteorological data

Comparing the qualified fitting and prediction rates of the ARIMA model, the RF model and the SVM model driven by meteorological data, it is found that among the three models, the SVM model performs best in drought prediction, where the polynomial kernel function and linear kernel function perform best in the fitting and prediction process, respectively. The qualified fitting rate of the linear kernel function is 0.7532, which is only 0.0437 smaller than the highest Gaussian radial basis kernel function, and the qualified rate in the prediction is the highest. The difference between qualified fitting and prediction rates is 0.0498, which is relatively stable. Therefore, it is considered that the SVM model (linear kernel function) is more suitable for drought monitoring in the Guanzhong Area.

Comparison of the models driven by remote sensing data

Comparing the qualified rates of the RF model and SVM model based on remote sensing data, the SVM model performs better in the prediction of drought, in which the polynomial kernel function and linear kernel function perform best in the fitting and prediction, respectively. The polynomial kernel function ranks second in the qualified prediction rate and only differs from the linear kernel function (ranks first) by 0.0081, and the difference between the qualified fitting and prediction rates is only 0.0607, which is relatively stable. Therefore, the SVM model (polynomial kernel function) driven by remote sensing data is more suitable for drought monitoring in the Guanzhong Area.

Comparison of models driven by combined meteorological and remote sensing data

Comparing the qualified rates of the RF model and SVM model based on combined meteorological and remote sensing data, the SVM model performs better than the RF model in the prediction of drought, in which the Gaussian radial basis kernel function and linear kernel function perform best in the fitting and prediction process, respectively. The qualified prediction rates of the Gaussian radial basis kernel function rank second, the differences from the linear radial basis kernel function which ranks first are small, and the performance is excellent. Therefore, the SVM model (Gaussian radial basis kernel function) driven by combined meteorological and remote sensing data is more suitable for drought prediction in the Guanzhong Area.

According to the 'Comparison of different models driven by meteorological data, Comparison of the models driven by remote sensing data, and Comparison of models driven by combined meteorological and remote sensing data' sections, the SVM model is superior to the other two models for the following reasons: (1) The SVM model performs better than the RF model when dealing with unbalanced data. Drought grades in the Guanzhong Area are unbalanced, with light and moderate drought as the main types, and heavy and severe droughts are rare. The samples of drought levels are unbalanced. (2) The ARIMA model is better in terms of linear prediction and less effective in terms of nonlinear prediction. The SVM model can also solve nonlinear problems through the application of kernel functions.

Comparison of the model-driven data

For the RF model, except for Longxian station, the qualified fitting and prediction rates show that the performance of remote sensing data in the RF model performed better than the combined meteorological and remote sensing data. The use of meteorological data will reduce the qualified fitting and prediction rates of the RF model. However, the qualified fitting and prediction rates of remote sensing data at the 11 stations are only slightly higher than those of the combined meteorological and remote sensing data. Therefore, the RF model driven by the remote sensing data performed well for drought monitoring in the Guanzhong Area.

In the SVM model, except for the polynomial kernel function, the qualified fitting and prediction rates of the other three kernel functions based on the remote sensing data are lower than those driven by the combined meteorological and remote sensing data. The difference in the qualified rates of the four kernel function models based on different data was calculated, which is shown in Table 13. The differences between the fitting and prediction of the

Table 13 | Difference in the qualified rates of the SVM model for different data (absolute values)

Kernel functions	Linear kernel function	Polynomial kernel function	Gaussian radial basis kernel function	Sigmoid kernel function
Fitting	0.1285	0.0068	0.1918	0.0119
Prediction	0.0458	0.0160	0.1104	0.0257

polynomial and sigmoid kernel functions are mostly less than 0.05, respectively, but large for the other kernel functions. The use of meteorological data has different effects on the qualified fitting and prediction rates for different kernel functions and has little influence on the polynomial and sigmoid kernel functions. The use of meteorological data significantly improves the qualified fitting and prediction rate of the Gaussian radial basis kernel function and the qualified fitting rate of the linear kernel function. The use of remote sensing data significantly improved the qualified fitting and prediction rates (except for the polynomial kernel function for the fitting and linear kernel functions for prediction). It is difficult to accurately measure the process of drought using a source of meteorological data. Only by considering the factors of precipitation, temperature, vegetation growth and so on, can drought be monitored and predicted more accurately. Therefore, the SVM model driven by the combined meteorological and remote sensing data is better for drought monitoring in the Guanzhong Area.

CONCLUSIONS

Based on the calculation of the SPEI series of 11 meteorological stations in the Guanzhong Area, the applicability of the SPEI to the drought characterization in the Guanzhong Area was analyzed. Three different models were used to predict the drought grades, and the main conclusions are as follows:

(1) The identified drought events based on SPEI-1 were in line with the recorded droughts, suggesting that the SPEI-1 data can reasonably reflect the occurrence of drought events in the Guanzhong Area.
(2) The SVM model (linear kernel function) performs the best, while the ARIMA model performs the worst in terms of drought prediction using meteorological data. The SVM model (polynomial kernel function) is superior if the models are driven by remote sensing data, while the SVM model (Gaussian radial basis kernel function) outperforms the other models if the predictors are a combination of meteorological and remote sensing data.
(3) When using the remote sensing data and the combination of meteorological and remote sensing data to build the RF model, the use of meteorological data has a small reduction effect on the qualified fitting and prediction rate; the use of meteorological data has a different effect on qualified fitting and prediction rates of different kernel functions in the SVM model. In conclusion, the RF model driven by the remote sensing data and the SVM model driven by the combined meteorological and remote sensing data performed better than the model driven by the corresponding other data in the Guanzhong Area.

Different types and lengths of meteorological data and remote sensing data used in this paper may affect the fitting and prediction accuracy of meteorological and remote sensing data models, which should be discussed in the future.

ACKNOWLEDGEMENT

This work is supported by National Natural Science Foundation of China (No. 51479130).

REFERENCES

Ahmad, S., Kalra, A. & Stephen, H. 2010 Estimating soil moisture using remote sensing data: a machine learning approach. *Advances in Water Resources* **33** (1), 69–80.

Altman, N. S. 1992 An introduction to kernel and nearest-neighbor non-parametric regression. *American Statistician* **46** (3), 175–185.

Bai, Z., Zhang, L., Wang, J. & Huang, Y. 2015 Research on forecast of meteorological drought in Yunnan based on ARIMA Model. *Yangtze River* **46** (15), 6–9.

Bannari, A., Morin, D., Bonn, F. & Huete, A. R. 1995 A review of vegetation indices. *Remote Sensing Reviews* **13** (1–2), 95–120.

Bonaccorso, B., Bordi, I., Cancelliere, A., Rossi, G. & Sutera, A. 2003 Spatial variability of drought: an analysis of the SPI in Sicily. *Water Resources Management* **17**, 273–296.

Box, G. E. & Jenkins, G. M. 1976 *Time Series Analysis Forecasting and Control*. Holden-Day, San Francisco.

Breiman, L. 2001 Random forests. *Machine Learning* **45** (1), 5–32.

Burges, C. J. 1998 A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* **2**, 127–167.

Cai, X., Ye, D., Li, Q., Zhang, C. & Wang, N. 2013 Temporal and spatial variation characteristics of drought in Shaanxi based on CI index. *Agricultural Research in the Arid Areas* **31** (5), 1–8.

Che, S. & Li, C. 2010 Time and space characteristics of drought in Shijiazhuang based on standardized precipitation index. *Meteorological Science and Technology* **38** (1), 66–70.

Chen, T., De, J. R. A. M., Liu, Y., Van, D. W. G. R. & Dolman, A. J. 2010 Using satellite based soil moisture to quantify the water driven variability in NDVI: a case study over mainland Australia. *Ecological Informatics* **5** (5), 400–409.

Cong, L. 2017 *Analysis of Drought Evolution Characteristics and Drought Prediction in Semi-Arid Areas*. Master's Thesis, Shenyang Agricultural University, Shenyang.

Dai, M., Huang, S. Z., Huang, Q., Leng, G. Y., Guo, Y., Wang, L. & Zheng, X. D. 2020 Assessing agricultural drought risk and its dynamic evolution characteristics. *Agricultural Water Management* **231**, 106003.

Di, L. P., Rundquist, D. C. & Han, L. H. 1994 Modeling relationships between NDVI and precipitation during vegetative growth cycles. *International Journal of Remote Sensing* **15** (10), 2121–2136.

Dong, Q. & Xie, P. 2014 Progress in hydrological drought research. *Journal of China Hydrology* **34** (4), 1–7.

Fan, G., Zhang, Y., Liu, M. & Mao, Y. 2011 Research on drought prediction based on support vector machine. *Chinese Journal of Agrometeorology* **32** (3), 475–478.

Fang, K., Wu, J., Zhu, J. & Xie, B. 2011 A review of random forest methods. *Statistics & Information Forum* **26** (3), 32–38.

Fang, Q., Wang, G., Xue, B., Liu, T. & Kiem, A. 2018 How and to what extent does precipitation on multi-temporal scales and soil moisture at different depths determine carbon flux responses in a water-limited grassland ecosystem? *Science of the Total Environment* **635**, 1255–1266.

Fang, Q., Wang, G., Liu, T., Xue, B., Sun, W. & Shrestha, S. 2019 Unraveling the sensitivity and nonlinear response of water use efficiency to the water-energy balance and underlying surface condition in a semiarid basin. *Science of the Total Environment* **699**, 134405.

Farrar, T. J., Nicholson, S. E. & Lare, A. R. 1994 The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana.I. NDVI response to soil moisture. *Remote Sensing of Environment* **50** (2), 121–133.

Gao, P. 2015 *Precipitation Forecast in Guanzhong Basin and Its Application in Drought Research*. Master's Thesis, Chang'an University, Xi'an.

Ghashghaie, M. & Nozari, H. 2018 Effect of dam construction on Lake Urmia: time series analysis of water level via ARIMA. *Journal of Agricultural Science and Technology* **20** (S), 1541–1553.

Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R. 2006 Random forests for land cover classification. *Pattern Recognition Letters* **27** (4), 294–300.

Guo, Y., Huang, S. Z., Huang, Q., Leng, G. Y., Fang, W., Wang, L. & Wang, H. 2020 Propagation thresholds of meteorological drought for triggering hydrological drought at various levels. *Science of the Total Environment* **712**, 136502.

Han, P., Wang, P. X., Tian, M., Zhang, Y., Liu, J. & Zhu, D. 2012 Application of the ARIMA models in drought forecasting using the standardized precipitation index. In: *International Conference on Computer and Computing Technologies in Agriculture*. Springer, Berlin, Heidelberg, pp. 352–358.

Han, D., Wang, G., Liu, T., Xue, B. L., Kuczera, G. & Xu, X. 2018 Hydroclimatic response of evapotranspiration partitioning to prolonged droughts in semiarid grassland. *Journal of Hydrology* **563**, 766–777.

Hernandez, E. & Annette, U. V. 2014 Standardized precipitation evaporation index (SPEI)-based drought assessment in semi-arid south Texas. *Environmental Earth Sciences* **71** (6), 2491–2501.

Ichii, K., Hashimoto, H., Nemani, R. & White, M. 2005 Modeling the interannual variability and trends in gross and net primary productivity of tropical forests from 1982 to 1999. *Global and Planetary Change* **48** (4), 286.

Kogan, F. N. 1990 Remote sensing of weather impacts on vegetation in nonhomogeneous areas. *International Journal of Remote Sensing* **11** (8), 1405–1420.

Kogan, F. N. 1995 Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin of the American Meteorological Society* **76** (5), 655–668.

Lei, X., Li, Q., Wang, J., Li, H. & Li, H. 2016 The drought and flood evolution law in the Guanzhong Area and the analyses of the present century's drought and flood characteristics. *Journal of Catastrophology* **31** (3), 101–109.

Li, J., Mo, S., Shen, B., Si, H. & Wang, Y. 2016 Analysis of drought characteristics in Weihe River Basin based on SPEI. *Journal of Xi'an University of Technology* **32** (1), 70–76.

Liu, X. 2012 *Natural Disasters and Preventive Measures in the Northwest Region Since the Founding of the People's Republic (1994–2000)*. Master's Thesis, Tianjin University of Commerce, Tianjin.

Liu, W., An, S., Liu, G. & Guo, A. 2004 Further correction of Palmer's drought pattern. *Journal of Applied Meteorological Science* **15** (2), 207–215.

Lu, J. 2018 Temporal and spatial variation characteristics of drought in Yungui area from 1960 to 2014 based on SPEI and run-length theory. *Journal of Zhejiang University (Science Edition)* **45** (3), 363–372.

Malik, A., Kumar, A. & Singh, R. P. 2019 Application of heuristic approaches for prediction of hydrological drought using multi-scalar streamflow drought index. *Water Resources Management* **33** (11), 3985–4006.

Mazdiyasni, O. & Aghakouchak, A. 2015 Substantial increase in concurrent droughts and heatwaves in the United States. *Proceedings of the National Academy of Sciences* **112** (37), 11484–11489.

Mckee, T. B., Doesken, N. J. & Kleist, J. 1993 The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*, January 17–22 Anaheim, pp. 179–184.

Mei, X. 2012 *Temporal and Spatial Distribution of Actual Evapotranspiration in Guanzhong Agricultural Region of Shaanxi Province*. Master's Thesis, Xi'an University of Technology, Xi'an.

Meteorological Station of Shaanxi Meteorological Bureau 1976 *Record of Natural Disasters in the Shaanxi Province*. Shaanxi Meteorological Bureau, Xi'an.

Miao, Z. 2008 *Analysis and Prediction of Meteorological Drought Characteristics in Baojixia Irrigation District*. Master's Thesis, Northwest A&F University, Yangling.

Moody, J. & Darken, C. J. 1989 Fast learning in networks of locally tuned processing units. *Neural Computation* **1** (2), 281–294.

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

Qiao, L. 2015 *Study on Drought Early Warning and Emergency Water Source Allocation in Guanzhong Area. Doctoral Dissertation*, Chang'an University, Xi'an.

Sahoo, A. K., Sheffield, J., Pan, M. & Wood, E. F. 2015 Evaluation of the tropical rainfall measuring mission multi-satellite precipitation analysis (TMPA) for assessment of large-scale meteorological drought. *Remote Sensing of Environment* **159**, 181–193.

Shumway, R. H. & Stoffer, D. S. 1982 An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* **3** (4), 253–264.

Sun, Q. 2017 *Temporal and Spatial Variation Characteristics of Drought in China From 1981 to 2015 Based on VCI Index*. Master's Thesis, Jiangsu Normal University, Nanjing.

Vega, H. M., Lima, J. R. & Cerniak, S. N. 2019 SPEI and Hurst analysis of precipitation in the Amazonian Area of Brazil. *Revista Brasileira de Meteorologia* **34** (2), 325–334.

Vicente-Serrano, S. M., Beguer, A. S. & Lpez-Moreno, J. I. 2010 A multi-scalar drought index sensitive to global warming: the standardized precipitation evapo-transpiration index. *Journal of Climate* **23** (7), 1696–1718.

Wang, Z. 2017 *Big Data Mining and Application*. Tsinghua University Press, Beijing.

Wang, L. & Chen, W. 2014 Applicability analysis of standardized precipitation evapotranspiration index in drought monitoring in China. *Plateau Meteorology* **33** (2), 423–431.

Wei, Z., Chen, S. & Huang, Y. 2015 Temporal and spatial characteristics of potential evapotranspiration in Shaanxi from 1981 to 2000 and its response to climate factors. *Scientia Geographica Sinica* **35** (8), 1033–1041.

Weng, Q. 2012 Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. *Remoting Sensing of Environment* **117** (2), 34–49.

Wu, J. 2014 *SPSS Statistical Analysis Starts From Scratch*. Tsinghua University Press, Beijing.

Wu, H. 2018 *Study on the Characteristics of Water Resources Change and Drought Vulnerability in Guanzhong Plain. Doctoral Dissertation*, Chang'an University, Xi'an.

Wu, J., Chen, Y. & Yu, S. 2016 Drought prediction based on random forest model. *China Rural Water and Hydropower* **11**, 17–22.

Xu, H. 2018 *Analysis of the Characteristics of Drought Time and Space Evolution and Vulnerability Assessment in Shaanxi Province*. Doctoral Dissertation, Chang'an University, Xi'an.

Yan, S., Lu, Q., Zhang, J., Zhang, Z. & Bai, S. 2012 Evolution characteristics of NDVI in coastal areas of Jiangsu Province and its response to regional climate change. *Journal of Nanjing Forestry University (Natural Sciences Edition)* **36** (01), 43–47.

Yang, S. Y., Meng, D., Li, X. J. & Wu, X. L. 2018 Multi scale response of vegetation change to SPEI meteorological drought index in North China from 2001 to 2014. *Journal of Ecology* **38** (3), 1028–1039.

Yilmaz, M. T., Hunt, E. R. & Jackson, T. J. 2008 Remote sensing of vegetation water content from equivalent water thickness using satellite imagery. *Remote Sensing of Environment* **112** (5), 2514–2522.

Yinglan, A., Wang, G., Liu, T., Shrestha, S., Xue, B. L. & Tan, Z. 2019a Vertical variations of soil water and its controlling factors based on the structural equation model in a semi-arid grassland. *Science of the Total Environment* **691**, 1016–1026.

Yinglan, A., Wang, G., Liu, T., Xue, B. L. & Kuczera, G. 2019b Spatial variation of correlations between vertical soil water and evapotranspiration and their controlling factors in a semi-arid region. *Journal of Hydrology* **574**, 53–63.

Yu, Y. 2015 *Comparison of Drought Monitoring and Drought Time and Space Characteristics in Guanzhong Region Based on CWSI and TVDI*. Master's Thesis, Shaanxi Normal University, Xi'an.

Yuan, X. 2015 *Study on Establishment of Hydrological Drought Index and Calculation Method of Drought Frequency*. Master's Thesis, Taiyuan University of Technology, Taiyuan.

Yuan, W. & Zhou, G. 2004 Theoretical analysis and research prospects of drought index. *Advance in Earth Sciences* **19** (6), 982–991.

Yurekli, K., Kurune, A. & Ozturk, F. 2005 Application of linear stochastic models to monthly flow data of Kelkit steam. *Ecological Modeling* **183**, 67–75.

Zargar, A., Sadiq, R., Naser, B. & Khan, F. I. 2011 A review of drought indices. *Environmental Reviews* **19** (17), 333–349.

Zhang, Y., Hao, Z., Wang, Y., Li, M. & Chen, E. 2014 The relationship between multi-scale drought characteristics and climate index of Taiyuan based on SPEI and SPI index. *Ecology and Environmental Sciences* **23** (9), 1418–1424.

Zhang, B., Wang, S. & Wang, Y. 2019a Copula-based convection-permitting projections of future changes in multivariate

drought characteristics. *Journal of Geophysical Research–Atmospheres* **124** (14), 7460–7483.

Zhang, X., Chen, N., Sheng, H., Ip, C., Yang, L. & Chen, Y. 2019b Urban drought challenge to 2030 sustainable development goals. *Science of the Total Environment* **693**, 133536.

Zhang, Y., Li, W., Chen, Q., Pu, X. & Xiang, L. 2017 Multi-models for SPI drought forecasting in the north of Haihe River Basin, China. *Stochastic Environmental Research and Risk Assessment* **31** (10), 1–11.

Zhao, X., Zhu, N. & Huang, L. 2012 Time series modeling algorithm and empirical analysis based on ARIMA model. *Journal of Guilin University of Electronic Technology* **32** (5), 410–415.

Zhou, X. J., Peng, X. W., Tansey, K., Zhang, S. Y., Li, H. M. & Wang, L. 2019 Developing a fused vegetation temperature condition index for drought monitoring at field scales using Sentinel-2 and MODIS imagery. *Computers and Electronics in Agriculture* **168**, 105144.

Zhuang, S., Zuo, H., Ren, P., Xiong, G. & Li, B. 2013 Application of standardized precipitation evaporation index in China. *Climatic and Environmental Research* **18** (5), 617–625.

First received 24 December 2019; accepted in revised form 29 March 2020. Available online 23 June 2020