

Original Research

Comparing Machine-Learning Models for Drought Forecasting in Vietnam's Cai River Basin

Zhen Nan Liu^{1,2*}, Qiong Fang Li¹, Luong Bang Nguyen³, Gui Hong Xu²

¹College of Hydrology and Water Resources, Hohai University, Nanjing, China

²School of Civil Engineering, Guizhou Institute of Technology, Guiyang, China

³Thuyloi University, Hanoi, Vietnam

Received: 11 August 2017

Accepted: 28 November 2017

Abstract

Drought occurs throughout the world, affecting people more than any other major natural hazards – especially in the agriculture industry. An effective and timely monitoring system is required to mitigate the impacts of drought. Meanwhile, extreme learning machine (ELM), online sequential extreme learning machine (OS-ELM), and self-adaptive evolutionary extreme learning machine (SADE-ELM) are rarely applied as the alternative drought-forecasting tools in the meantime. The present study aims to evaluate the ability of these models to predict drought and the quantitative value of drought indices, the standardized precipitation index (SPI), and the standardized precipitation evapotranspiration index (SPEI). For this purpose, the sea surface temperature anomalies (SSTA) events at NinoW and Nino4 zones were selected for input variables to forecast drought. The SPI/SPEI values may contain a one/three/six-month dry and a one/three/six-month wet period in short-term periods, and this causes instability. For this reason, 4 models for SPI/SPEI (12 months) were trained and tested by these methods, respectively. According to two statistical indices (RMSE and CORR) and stability of these methods, the SADE-ELM models perform the best, and the performance of the OS-ELM models are better than the ELM models.

Keywords: drought indices, drought forecast, extreme learning machine, online sequential extreme learning machine, self-adaptive evolutionary extreme learning machine, sea surface temperature anomalies

Introduction

Drought often occurs and is the costliest one of all natural disasters over the world, leading to significant societal, economic, and ecologic impacts [1-2]. Drought usually affects human lives more than any other form of natural hazards, and is widely considered to be the most

complex and least understood of all the natural hazards [3-4]. Drought not only affects agricultural systems but also has a serious impact on the environment. Therefore, drought monitoring and assessment and so on, are hot topics among hydrologists and meteorologists, and attract world-wide attention [5-9].

In order to prevent and mitigate the effects of future occurrences of drought, a number of drought indices exist that have been used to represent different types of drought, including meteorological or climatological,

*e-mail: lewzhenan@foxmail.com

agricultural and hydrological drought, such as the Palmer index [10], the crop moisture index [11], the standardized precipitation index (SPI) [12], and the standardized precipitation evapotranspiration index (SPEI) [13] and so on. The Palmer index and SPI are traditionally the most welcome indices for forecasting drought due to their standardization. For the purposes of comparing drought conditions from different areas, which often have different hydrological balances, the most important characteristic of a drought index is its standardization [14]. Standardization of a drought index ensures independence from a geographical position, as the index in question is calculated with respect to the average precipitation in the same location [15]. The most important difference between Palmer and SPI is that Palmer has a complex structure with a very long memory, while SPI is an easily interpreted and simple-moving average process [16]. This characteristic has made SPI the most universal drought index. The SPEI index is based on precipitation and evapotranspiration, and the SPEI is used in risk analysis and decision-making analysis such as SPI, because it is simple, spatially invariant in its interpretation, and probabilistic. For the above reasons, we chose SPI and SPEI for drought forecasting in this paper.

Due to the impact of drought, an effective and timely monitoring system is required, the effective monitoring of drought can aid in the development of an early warning system [17]. At present a lot of models have been accepted as efficient tools for the modeling of complex hydrologic systems widely employed for forecasting, such as artificial neural networks (ANNs) and autoregressive moving average approach (ARMA) [18-20]. The ARMA technique assumes the time series to be stationary and to follow the normal distribution [21]. However, drought time series is usually characterized by features of both nonlinearity and unstableness; thus, linear-related time series forecasting techniques are not sufficient to capture the characteristics of hydrological time series [22]. Nevertheless, ANNs possess the strong ability of nonlinear function approximation and self-organizing and self-adaptive function. Some specific applications of ANNs to hydrology include the hybrid models for water quality prediction [23], the modeling of the rainfall-runoff process [24], hydrologic time series modeling [25], sediment concentration estimation [26], and runoff and sediment yield modeling [27]. Although the utility of the ANNs approach for long-term forecasting of drought events has been certified [28], slow learning convergence speed is one of defects of ANNs for two key reasons: 1) the slow gradient-based learning algorithms are extensively used to train neural networks, and 2) all the parameters of networks are tuned iteratively by using such learning algorithms. In order to overcome the drawback, an extreme learning machine (ELM) was proposed, and then online sequential extreme learning machine (OS-ELM) and self-adaptive evolutionary extreme learning machine (SADE-ELM) were developed in the following years. ELM has been used to predict water electrical conductivity (EC) [29], groundwater level [30], fluoride

contamination [31], monthly mean streamflow water level [32], wind speed [33-36], and drought [37]. OS-ELM has been used to recognize the different positions of arm movement [38], predict river water discharge [39], and diagnose the sensor fault of an aero engine [40]. SADE-ELM has been used to predict the scour depth around bridge piers [41] and electricity price [42], and estimate soil temperature [43]. However, ELM, OS-ELM, and SADE-ELM are rarely used in drought prediction applications in the meantime. Furthermore, to date no research has been published that uses an OS-ELM or SADE-ELM model to predict drought. The main objective of the present study is to investigate the ability of the ELM, OS-ELM, and SADE-ELM to predict drought in the Cai River Basin of east Vietnam.

Material and Methods

Case Study Site and Data Collection

Case Study Site

The basin that we studied lies between latitudes $12^{\circ}02'49''$ and $12^{\circ}36'13''$ N and between longitudes $108^{\circ}40'03''$ - $109^{\circ}11'38''$ E in Khanhhoa Province, Vietnam, with a total area of 1,889 (Fig. 1). This basin is located in a tropical monsoon zone that exhibits rather unique deformational features, and an oceanic climate. The average annual temperature and evaporation in the period 1982-2012 was 26.9°C and 1,310 mm, respectively. The average rainfall in the period 1982-2012 was 1,616 mm, and there are two distinct seasons: rainy and dry. The rainy season usually lasts from May to December, and rainfall is largely concentrated in September, October, and November, comprising 55% of the average annual

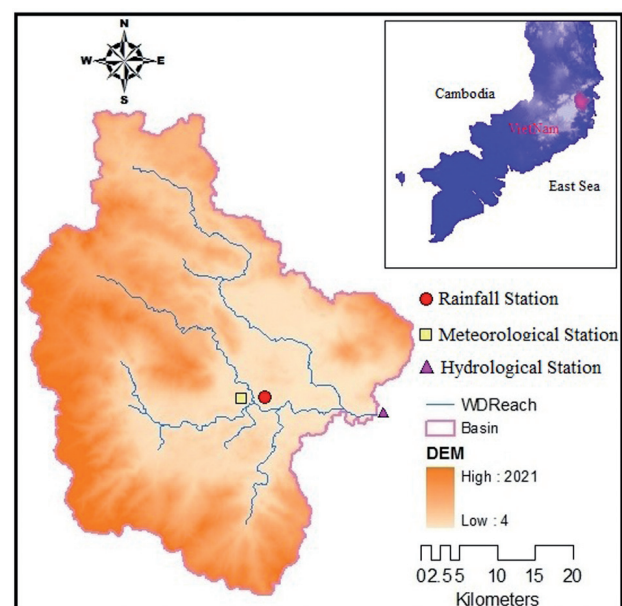


Fig. 1. Cai River basin, Vietnam.

rainfall. During many dry-season months no rain falls. With a high temperature foundation, the average annual temperature during many years is 26.7°C, and the difference in temperature between the months is relatively small. The potential evaporation in this region is high, averaging approximately 1,200-1,600 mm/year.

Data Collection

The daily 1983-2012 time series of rainfall and temperature datasets were collected from Khanhvinh Meteorological Station in Khanhhoa Province, Vietnam. The homogeneity and reliability of the datasets have been checked and firmly controlled by the Vietnam Institute of Meteorology and Hydrography. The Four SSTA datasets in the Nino12, Nino3, Nino4, and NinoW zones were obtained from the Comprehensive Ocean Atmosphere Data Set (COADS) online [44-45]. The data was standardized by dividing standard deviations and collected from 1982 to 2012 – a duration of 31 years. Due to the purposes of this research, the field of SSTA used in the Nino12, Nino3, Nino4, and NinoW zones was within the corresponding coordinates of (0°N-10°S; 90°E-80°W), (5°N-5°S; 150°W-80°W), (5°N-5°S; 160°E-150°W), and (15°N-0°S; 130°E-150°E). A 2° latitude × 2° longitude resolution was applied, by using averaged computations from the original 1 × 1 grid point datasets.

Methods

Extreme Learning Machine

An extreme learning machine (ELM) proposed by [46] has been applied in order to study some problems, such as regression analysis and classification recognition. The ELM model is based on a single-hidden layer feed-forward neural network (SLFN), which overcomes the shortcomings of training speed slowly, obtaining a local optimal solution easily, and being affected by the learning rate, when the traditional SLFN is training [47-49]. The primary advantage of the ELM is that the algorithm

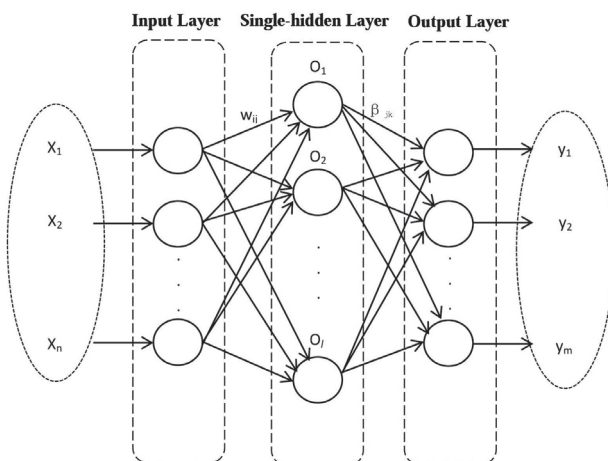


Fig. 2. The structure of the ELM model.

randomly produces the connection weight between the input layer and the hidden layer and the threshold of the hidden layer neuron, there is no need to adjust as much as the traditional neural network during training – we only need to set the number of the hidden layer neurons, and the only optimal solution can be obtained [48, 50]. The ELM model used in this study is shown in Fig. 2. The circle nodes are used to reflect different neurons. A three-layer structure of ELM divides into input layer, one hidden layer, and output layer. The individual neurons are connected between two layers; there is only a connection weight between adjacent two layers.

Assume data set X (model inputs variables) and Y comprised of N training data samples, i.e., $(x_i, y_i) \in R^n \times R^m$, $i = 1, 2, \dots, N$, and the hidden layer has l neurons. w is connection weight between the input layer neurons and the hidden layer neurons. β is connection weight between the hidden layer neurons and the output layer neurons. An ELM model can be expressed as follow [51]:

$$\sum_{i=1}^l \beta_i g_i(x_i) = H\beta \quad (1)$$

$$H = \begin{bmatrix} g(w_1 \cdot x_1 \cdot b_1) & L & g(w_l \cdot x_1 \cdot b_l) \\ M & O & M \\ g(w_1 \cdot x_N \cdot b_1) & L & g(w_l \cdot x_N \cdot b_l) \end{bmatrix}_{N \times l} \quad (2)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ M \\ \beta_l^T \end{bmatrix}_{l \times m} \quad T = \begin{bmatrix} t_1^T \\ M \\ t_l^T \end{bmatrix}_{N \times m} \quad (3)$$

...where H represents the output matrix of the hidden layer and T is the label matrix.

It has been proven that if $l = N$, for any w and b randomly chosen, the training samples can be approximated with zero error by ELM and

$\sum_{j=1}^N \|t_j - y_j\| = 0$. While N is larger, in order to reduce the

amount of calculation we usually select l as smaller than N , and the training error by ELM can approximate any

$\varepsilon (> 0)$ and $\sum_{j=1}^N \|t_j - y_j\| < \varepsilon$, where $t = [t_1, t_2, \dots, t_m]^T$ is the value of test samples.

Hence, when the activation function $g(x)$ is infinitely differentiable, the parameters of ELM do not need to be fully adjusted, while w and b can be randomly chosen and remain unchanged during training. Meanwhile, β can be obtained by the method of minimum norm least-squares and the solution of $\min_{\beta} \|H\beta - T\|$ is unique, which is $\hat{\beta} = H^+ T$, where H^+ is the Moore-Penrose generalized inverse of Matrix H .

As analyzed in the above works, the modeling steps of ELM generally include the following:

- Step 1. Determine the number of hidden layer neurons and randomly set the values of w and b
- Step 2. Choose an activation function $g(x)$, which is infinitely differentiable, and calculate Matrix H
- Step 3. Calculate $\hat{\beta} : \hat{\beta} = H^+ T'$

Online Sequential Extreme Learning Machine

An online sequential extreme learning machine (OS-ELM), as a new algorithm, was proposed by [52] for training sequential data. Compared to the ELM, the regularization parameters are used in the optimization ELM to increase accuracy and generation performance [51]. OS-ELM is composed of an initialization phase and sequential learning phase [53].

In the initialization phase, the output weight of the feed-forward neural networks is obtained through a few training samples. The sequential learning phase included using a single sample or a block of samples to update the output weight of the feed-forward neural networks acquired during the initialization phase. As described above, the aim of ELM is

to calculate β_0 to make $\min_{\beta} \|H_0 \beta - T_0\|$ and $H^+ = (H^T H)^{-1} H^T$. $\beta_0 = P_0 H_0^T T_0$, where $P_0 = (H_0^T H_0)^{-1}$ and $K_0 = H_0^T H_0 = P_0^{-1}$. For the sake of simplicity, set $k = 0$, where k is a parameter that shows the number of data sets that are presented to the network. When $k = 1$, in order to calculate β_1 as:

$$\min_{\beta} \left\| \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \beta - \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \right\| \quad (4)$$

... and

$$\beta_1 = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \quad (5)$$

$$K_1 = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = \begin{bmatrix} H_0^T & H_1^T \end{bmatrix} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = K_0 + H_1^T H_1 \quad (6)$$

...and

$$\begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = K_0 K_0^{-1} H_0^T + H_1^T T_1 = K_0 \beta_0 + H_1^T T_1 = K_1 \beta_0 - H_1^T H_1 \beta_0 + H_1^T T_1 \quad (7)$$

Therefore, Eq.(5) is expressed:

$$\beta_1 = \beta_0 + K_1^{-1} H_1^T (T_1 - H_1 \beta_0) \quad (8)$$

When $k = k + 1$, the recursive formula is obtained as follow:

$$\beta_{k+1} = \beta_k + K_{k+1}^{-1} H_{k+1}^T (T_{k+1} - H_{k+1} \beta_k) \quad (9)$$

From Eq. (9), in order to calculate β_{k+1} , K_{k+1}^{-1} is obtained using the Woodbury formula.

$$K_{k+1}^{-1} = (K_k + H_{k+1}^T H_{k+1})^{-1} = K_k^{-1} - K_k^{-1} H_{k+1}^T (I + H_{k+1} K_k^{-1} H_{k+1}^T)^{-1} K_k^{-1} H_{k+1} \quad (10)$$

And $P_{k+1} = K_{k+1}^{-1}$, Eq.(9) is modified by using (10)

$$\beta_{k+1} = \beta_k + P_{k+1} H_{k+1}^T (T_{k+1} - H_{k+1} \beta_k) \quad (11)$$

As analyzed above, the modeling steps of OS-ELM generally include the following:

- Step 1) Calculate β_0 in the way same as ELM.
- Step 2) Use a single sample or a block of samples to update β_0 to β_1 .
- Step 3) Return the final β .

Self-Adaptive Evolutionary Extreme Learning Machine

The self-adaptive evolutionary extreme learning machine (SADE-ELM) was developed by [54], which chooses trial vector generation strategies and some relative control parameters adaptively. A self-adaptive differential evolution algorithm is used to optimize the network input weights and hidden node biases, and the extreme learning machine is used to derive the network output weights [55-57]. We summarize SADE-ELM in the following steps.

Step 1. Initialization

Given a set of training data and l hidden nodes with an activation function $g(x)$, a set of NP vectors where each one includes all the network hidden node parameters are initialized as the populations of the first generation

$$\theta_{k,G} = [a_{l,(k,G)}^T, \mathbf{L}, a_{l,(k,G)}^T, b_{l,(k,G)}^T, \mathbf{L}, b_{l,(k,G)}^T] \quad (12)$$

...where a_j and b_j ($j = 1, 2, \dots, l$) are randomly generated, G represents the generation and $k = 1, 2, \dots, NP$.

Step 2. Calculate output weights and root mean square error

Calculate the network output weight matrix and root mean square error (RMSE) with respect to each population vector with the following equations, respectively:

$$\beta_{k,G} = H_{k,G}^+ T \quad (13)$$

$$RMSE_{k,G} = \sqrt{\frac{\sum_{i=1}^N \left\| \sum_{j=1}^I \beta_j g(a_{j,(k,G)}, b_{j,(k,G)}, x_i - t_i) \right\|^2}{m \times N}} \quad (14)$$

...where $H_{k,G}$ is shown in Eq. (15) and $H_{k,G}^+$ is the Moore-Penrose generalized inverse of $H_{k,G}$:

$$H_{k,G} = \begin{bmatrix} g(a_{1,(k,G)}, b_{1,(k,G)}, x_1) & L & g(a_{I,(k,G)}, b_{I,(k,G)}, x_1) \\ M & O & M \\ g(a_{1,(k,G)}, b_{1,(k,G)}, x_N) & L & g(a_{I,(k,G)}, b_{I,(k,G)}, x_N) \end{bmatrix} \quad (15)$$

$$\theta_{k,G} = \begin{cases} u_{k,G+1} & \text{if } RMSE_{\theta_{k,G}} - RMSE_{u_{k,G+1}} > \varepsilon \cdot RMSE_{\theta_{k,G}} \\ u_{k,G+1} & \text{if } RMSE_{\theta_{k,G}} - RMSE_{u_{k,G+1}} < \varepsilon \cdot RMSE_{\theta_{k,G}} \\ & \text{and } \|\beta_{u_{k,G+1}}\| < \|\beta_{\theta_{k,G}}\| \\ \theta_{k,G} & \text{else} \end{cases} \quad (16)$$

In the first generation, the population vector with the best RMSE is stored as $\theta_{best,1}$ and $RMSE_{\theta_{best,1}}$.

Step 3. Mutation and crossover

The trial vector produces a strategy from a candidate pool for each target vector according to probability $p_{t,G}$, where $p_{t,G}$ represents the probability that the strategy t ($t = 1, 2, 3, 4$) should be chosen at the G th generation. It defines a fixed number of iterations (generations) as the learning period (LP), and the probability $p_{t,G}$ is updated in the following ways.

1. When $G \leq LP$, all strategies have equal probability of

$$\text{selection, i.e., } p_{t,G} = \frac{1}{4}$$

2. When $G > LP$, $p_{t,G} = \frac{S_{t,G}}{\sum_{t=1}^4 S_{t,G}}$ with

$$S_{t,G} = \frac{\sum_{g=G-LP}^{G-1} ns_{t,g}}{\sum_{g=G-LP}^{G-1} ns_{t,g} + \sum_{g=G-LP}^{G-1} nf_{t,g}} + \varepsilon \quad (17)$$

...where $ns_{t,G}$ denotes the number of trial vectors produced by the t th strategy at g th generations (iterations) that can successfully enter the next generation while $nf_{t,g}$ is the number of trial vectors generated by t th strategy at g th generations (iterations) that are discarded in the next generation. LP generations' success and failure numbers of trial vectors are stored. Once the iterations go beyond the initial LP generations, the earliest records are removed

and the new numbers in the current generation are stored. ε is a small positive constant value to avoid the possible null success rate. Meanwhile, a set of control parameters F and crossover rate CR are randomly generated for each target vector according to the normal distributions $N(0.5, 0.3)$ and $N(0.5, 0.1)$. For a given problem, proper values of CR usually fall into a small range, the mean value of CR is gradually adjusted using the values of previous CR that have successfully generated trail vectors and are able to enter the next generation.

Step 4. Evaluation

All the trail vectors $u_{k,G+1}$ generated at the $(G+1)$ th generation are evaluated using the equation above, where ε represents the preset small positive rate of tolerance. A smaller weight is used to make the neural network perform better in this respect and one more criteria is added as a norm of the output weight β . Steps 3 and 4 continue until the finally set goal is reached or the maximum learning iterations are achieved.

Applying the Models in Forecasting Drought

Calculating SPI and SPEI

SPI was established by [12]. The calculation formula is:

$$SPI = S \left(t - \frac{c_0 + c_1 + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) \quad (18)$$

...where $t = \sqrt{\ln 1/F^2}$, $S = 1$ when $F > 0.5$, $S = -1$ when $F \leq 0.5$, F is an accumulative probability function, $c_0 = 2.515517$, $c_1 = 0.8022853$, $c_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, and $d_3 = 0.001308$.

In order to calculate SPI, a long-term precipitation record is required to fit the gamma probability density function to the observed data. Based on the precipitation data from 1983 to 2012, SPI of different time scales (i.e., 1, 3, 6, and 12 months) was calculated by using the SPI program online [58].

In order to calculate SPEI, which was introduced by [13], index (D) is cited to distinguish rainfall (P) and potential evaporation (PET). Index (D) indicates the redundancy or shortage of humidity, from which we can determine wet or dry conditions. Each period of increase or decrease in water discharge can be defined as:

$$D_i = P_i - PET_i \quad (19)$$

Here, the Thornthwaite method is used to calculate PET. In the incremental or decremental series of water discharge, a negative value may occur. Therefore, SPEI uses 3 parameters of log-logistic probability distribution function to describe the probability of an event. The form of accumulative probability function is expressed as:

$$F(x) = \left[1 + \left(\frac{\alpha}{x - \gamma} \right)^\beta \right]^{-1} \quad (20)$$

...where parameters α , β , and γ may use the linear moments method to determine conformance. SPEI was calculated in the same manner as SPI.

According to the World Meteorological Organization (WMO), SPI and SPEI were classified by climatic conditions (drought or wet) and are shown in Table 1 [59].

Selection of Forecast Factors

To obtain reliable forecast results, it is important that the model chooses the reasonable predictors. The correlation matrices between SSTAs at four ENSO activities observation zones, which are Nino12, Nino3, Nino4, and NinoW with SPI/SPEI used to analyze the selection of potential factors serving the drought forecast.

Correlation equations are a way of studying the association between two or more random variables. We chose the correlation equation between SSTA and SPI/SPEI for determining correlation coefficients. If the correlation coefficient is positive (reflecting a covariant relationship), then the absolute value of correlation coefficient is greater, and the level of linear dependence between SSTA and SPI/SPEI is greater. The reverse holds if the correlation coefficient is negative (reflecting a counter covariant relationship).

In order to confirm a correlation matrix between SSTA and SPI/SPEI, we will employ the correlation between an SSTA data chain $x\{x_1, x_2, \dots, x_n\}$ and an SPI/SPEI data chain $y\{y_1, y_2, \dots, y_n\}$, with n observatory value pairs $\{x_i, y_i\}$ ($i = 1, 2, \dots, n$). The data chains used for our calculation are described as follows:

- 1) Our SSTA data chain includes factors from SSTA in the Nino12, Nino3, and Nino4, and NinoW zones with time scales of 1, 3, 6, and 12 months. These are designated for the respective zones as $N12_j$, $N3_j$, $N4_j$, and NW_j . The index signs $j = 1, 2, L, 12$ indicate SSTA data at the j^{th} time ($j = 1$ at SSTA indicates the same period as SPI/SPEI, $j = 2$ indicates SSTA at one month ahead of SPI/SPEI, and $j = 12$ indicates SSTA at 11 months prior to SPI/SPEI).

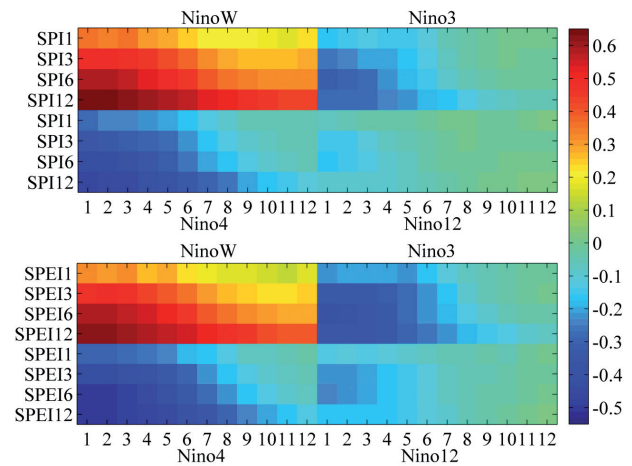


Fig. 3. Correlation between SSTA and SPI/SPEI.

- 2) The comparative element series is given by the SPI/SPEI value series at the case study site, whose objects are denoted by SPI_k and $SPEI_k$, where index $k = 1, 2, 3, 4$. Here, $k = 1$ when SPI/SPEI is for one-month time scales, $k = 2$ when SPI/SPEI is for 3-month time scales, $k = 3$ when SPI/SPEI is for 6-month time scales, and $k = 4$ when SPI/SPEI is for 12-month time scales.

The correlation matrix results between SSTA and SPI/SPEI are shown in Fig. 3. The correlation between SSTA and SPI is greater than the correlation between SSTA and SPEI. However, there is one notable shared characteristic. That is, the correlation between SSTA and SPI/SPEI decreases gradually from NinoW to Nino4, then to Nino3, and is smallest in Nino12 zones. The correlation between SSTA and SPI/SPEI in NinoW is positive, while it is negative in most other cases. The SSTA and SPI/SPEI over 12-month time scales is most often the greatest, while it is smallest over 1-month time scales.

The results for the correlation coefficient value between SSTA and SPI/SPEI indices show that the correlations at Nino3 and Nino12 zones are weaker than the correlation at NinoW and Nino4 zones, and when SPI/SPEI is larger than SSTA from a range of 1 to 3 months, the highest correlation coefficient always occurs. We can find that the linear dependence between SSTA and SPI/SPEI in the NinoW and Nino4 zones is high and close. Hence, we will choose SSTA data from these zones with a high coefficient of correlation with SPI/SPEI as forecast factors.

Table 1. Standard Precipitation Index (SPI) and Standard Precipitation Evapotranspiration Index (SPEI) classifications.

SPI/SPEI	Classification	SPI/SPEI	Classification
2.00 or more	Extremely	-0.50 to -0.99	Mild drought
1.50 to 1.99	Very wet	-1.00 to -1.49	Moderate drought
1.00 to 1.49	Moderately wet	-1.50 to -1.99	Severe drought
0.50 to 0.99	Mildly wet	-2.0 or less	Extreme drought
-0.49 to 0.49	Normal		

Table 2. Structures of forecasting models.

Model	Input structure	Output
M1	SSTAw(t-1),SSTAw(t-2),SSTAw(t-3),SSTAw(t-4),SSTAw(t-5),SSTAw(t-6)	SPI(t),SPEI(t)
M2	SSTAw(t-1),SSTAw(t-2),SSTAw(t-3),SSTAw(t-4),SSTAw(t-5),SSTA4(t-1)	SPI(t),SPEI(t)
M3	SSTAw(t-1),SSTAw(t-2),SSTAw(t-3),SSTA4(t-1),SSTA4(t-2),SSTA4(t-3)	SPI(t),SPEI(t)
M4	SSTAw(t-1),SSTAw(t-2),SSTAw(t-3),SSTAw(t-4),SSTA4(t-1),SSTA4(t-2)	SPI(t),SPEI(t)

Input Variables

As mentioned above, SSTA data at the two zones NinoW and Nino4 with high correlation coefficient with SPI/SPEI are used as input variables. The SPI/SPEI values may contain a 1/3/6-month dry and a 1/3/6-month wet period in short-term periods, and this causes instability. Passages between positive and negative values occur more frequently, and this also results in instability. For this sake, the SPI/SPEI outputs for 12 months, as a long-term period, are considered to generate these models with ELM, OS-ELM, and SADE-ELM methods in this study. In the construction of the estimation models, different models are generated for the SPI/SPEI (12 months) outputs. The datasets are divided into 2 sub sets: training and testing datasets. The training dataset includes data records measured between the years 1983 and 2008. In order to achieve a more reliable evaluation and comparison, the models are tested by evaluation if a dataset was not used during the training process. The testing dataset consists of data records observed between 2009 and 2012.

Data Normalization

In order to accelerate the learning speed of the models, we need to carry out normalization processing with the value of SSTA and SPI/SPEI. Retaining original data makes the values in [-1, 1]. The normalization formula is as follows:

$$x_{nor} = \frac{(x_{max} - x_{min}) \cdot (x - x_{min})}{x_{max} - x_{min}} + x_{min} \quad (21)$$

...where x , x_{max} , and x_{min} are the original data, the maximum data, and the minimum data of the original data, and x_{nor} is the normalized data. The *mapminmax* function is used to achieve this process in MATLAB.

Structures of the Models

The model structures have a great influence on the prediction results, and one of the most important steps is especially the selection of the input variables. Therefore, different estimation models are constructed for each phase. According to the research of [60], the lagged observations of the index (SPI) itself and precipitation were used as input variables to build the SPI forecasting

model, and good predicting results were obtained. Hence, in this research, the lagged observations of SSTA data are used similarly. The models for 12 months are named SPI/SPEI-12. Four models with different input numbers and structures are constructed for each phase using these variables. In this study, forecasting models based on various combinations of SSTA events in NinoW and Nino4 are constructed (Table 2). The SSTA events in NinoW and Nino4 are named $SSTA_w$ and $SSTA_4$.

Table 3. Optimal hidden neurons of all the models.

Neurons Networks	Models	Hidden neurons
ELM	SPI12	1
		2
		3
		4
	SPEI12	1
		2
		3
		4
OS-ELM	SPI12	1
		2
		3
		4
	SPEI12	1
		2
		3
		4
SADE-ELM	SPI12	1
		2
		3
		4
	SPEI12	1
		2
		3
		4

Table 4. The average results of all the models during training and testing.

Models	Training set		Testing set		Times
	RMSE	CORR	RMSE	CORR	
M1(for SPI-12)					
ELM	0.5255	0.6686	0.5843	0.3540	0.0012
OS-ELM	0.5068	0.6909	0.5457	0.5372	0.1370
SADE-ELM	0.5317	0.7112	0.5406	0.5469	3.2906
M1(for SPEI-12)					
ELM	0.6305	0.6541	0.6998	0.3971	0.0431
OS-ELM	0.5087	0.6801	0.5204	0.6134	0.1267
SADE-ELM	0.5874	0.5791	0.6155	0.5267	
M2(for SPI-12)					
ELM	0.5301	0.6638	0.5264	0.4096	0.0210
OS-ELM	0.5084	0.6873	0.5982	0.5635	0.1241
SADE-ELM	0.4613	0.4809	0.4747	0.5500	3.0625
M2(for SPEI-12)					
ELM	0.5431	0.6394	0.5793	0.4824	0.0312
OS-ELM	0.5861	0.7036	0.6046	0.6205	0.1052
SADE-ELM	0.4621	0.7273	0.5274	0.6251	3.4750
M3(for SPI-12)					
ELM	0.4711	0.4543	0.5202	0.2382	0.0341
OS-ELM	0.7415	0.6820	0.7671	0.6786	0.0918
SADE-ELM	0.4511	0.6353	0.4603	0.5708	3.5469
M3(for SPEI-12)					
ELM	0.5028	0.4887	0.5227	0.4758	0.0217
OS-ELM	0.7146	0.6735	0.7002	0.7610	0.1093
SADE-ELM	0.4506	0.7997	0.4933	0.7079	3.2110
M4(for SPI-12)					
ELM	0.5251	0.6694	0.5850	0.3450	0.0311
OS-ELM	0.5639	0.7334	0.6061	0.5670	0.1121
SADE-ELM	0.4339	0.7093	0.4372	0.6521	3.4610
M4(for SPEI-12)					
ELM	0.5362	0.6472	0.5673	0.5308	0.0112
OS-ELM	0.6038	0.7550	0.6300	0.6047	0.1301
SADE-ELM	0.4803	0.7501	0.5125	0.6525	3.2656

The performance of the models for training and testing datasets is evaluated according to statistical criteria, such as the correlation coefficient (CORR), coefficient of determination (R^2), and root mean square error (RMSE). CORR is a commonly used statistic, providing information on the strength of the linear relationship between the observed and computed values. R^2 is one of

the most widely employed statistics for evaluating model performance. The value of CORR and R^2 close to 1.0 indicates good model performance. The RMSE statistic indicates a model's ability to predict a value separated from the mean, and a value of RMSE close to 0 indicates a good model performance.

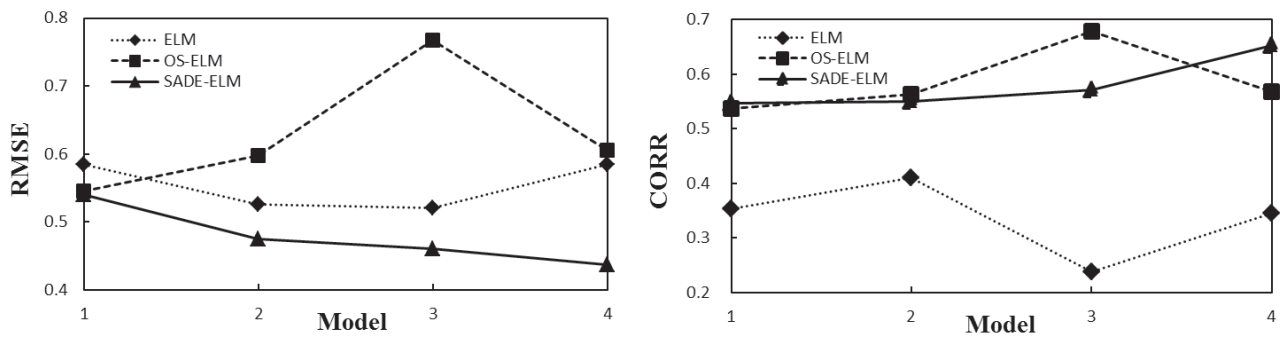


Fig. 4. Comparison of the average performances of the ELM, OS-ELM, and SADE-ELM models for SPI-12.

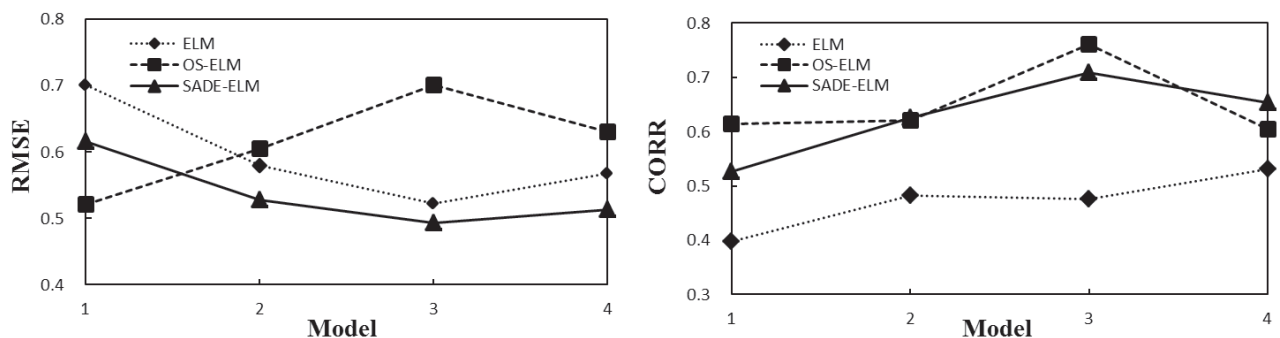


Fig. 5. Comparison of the average performances of the ELM, OS-ELM, and SADE-ELM models for SPEI-12.

Development of the Models

The ELM, OS-ELM, and SADE-ELM models are developed in a MATLAB environment. The computations related to all of the models including ELM, OS-ELM, and SADE-ELM were implemented in a computer with Inter core i5, 2.3-GHz CPU, and 8GB of RAM. The number of hidden nodes of each model is changeable for better accuracy, and the number of hidden neurons is selected via a trial and error method. The number of neurons between 1 and 200 is tested in hidden layers of these models. In each trail, the number of nodes in the hidden layer is increased gradually until the optimal nodes are achieved. There are a lot of popular activation functions, such as sigmoid, sine, hard-limit, and radial basis. The “sigmoid” activation function is chosen to develop the ELM models. The focus of the article is to compare the ability of the ELM and the ELM optimized by online sequence and self-adaptive evolutionary algorithm in drought prediction, so the OS-ELM and SADE-ELM models select the same activation function as ELM models for objective performance analysis and comparison. The optimal hidden neurons are listed in Table 3. Each model is run 10 times for the instability of the ELM, OS-ELM, and SADE-ELM methods, and all these results are collected by averaging multiple trails in the following section. According to the real world application, the number of populations NP is set to be 8 in the SADE-ELM method.

Results and Discussion

The average performance of a model (M1) with input variables including SSTaw at zone NinoW (the highest correlation SSTa and SPI/SPEI) and models (M3-M4) with input variables including SSTaw and SSTa₄ trained by the ELM, OS-ELM, and SADE-ELM methods for predicting drought in both training and testing stages is shown in Table 4, respectively. From Table 4, it is seen that the SPI index delivers a similar performance to SPEI in all models. Higher prediction errors obtained by models in the testing dataset, compared with the training dataset, indicate that these models exhibit relatively better generalization as compared with the predictions. Meanwhile, the computer time required of OS-ELM and SADE-ELM is much more than the single ELM without the updating and optimization processes.

It is clear from Figs 4 and 5 that almost all SADE-ELM forecasting models for SPI/SPEI obtain the smallest RMSE, and the biggest CORR compared with the other models, and the ELM forecasting models for SPI/SPEI are superior to the OS-ELM forecasting models in terms of RMSE. However, it turns out just the opposite in terms of CORR. From the average results of all models, we can see that M2 by ELM, M1 by OS-ELM, and M4 by SADE-ELM for SPI and M3 by ELM, M1 by OS-ELM, and M3 by SADE-ELM for SPEI are the best performance because the models have comparatively small RMSE and big CORR.

When the results for all models are compared for every trial, the models with highest performance are M2 (for SPI) and M3 (for SPEI) as trained by the ELM method, M1 (for SPI) and M1 (for SPEI) as trained by the OS-ELM method, and M4 (for SPI) and M3 (for SPEI)

as trained by the SADE-ELM method. Figs 6-8 show the best performance in testing stages. R^2 for the ELM, OS-ELM, and SADE-ELM models for SPI are 0.6243, 0.6966, and 0.7122, respectively. Meanwhile, The R^2 for the ELM, OS-ELM, and SADE-ELM models for SPEI

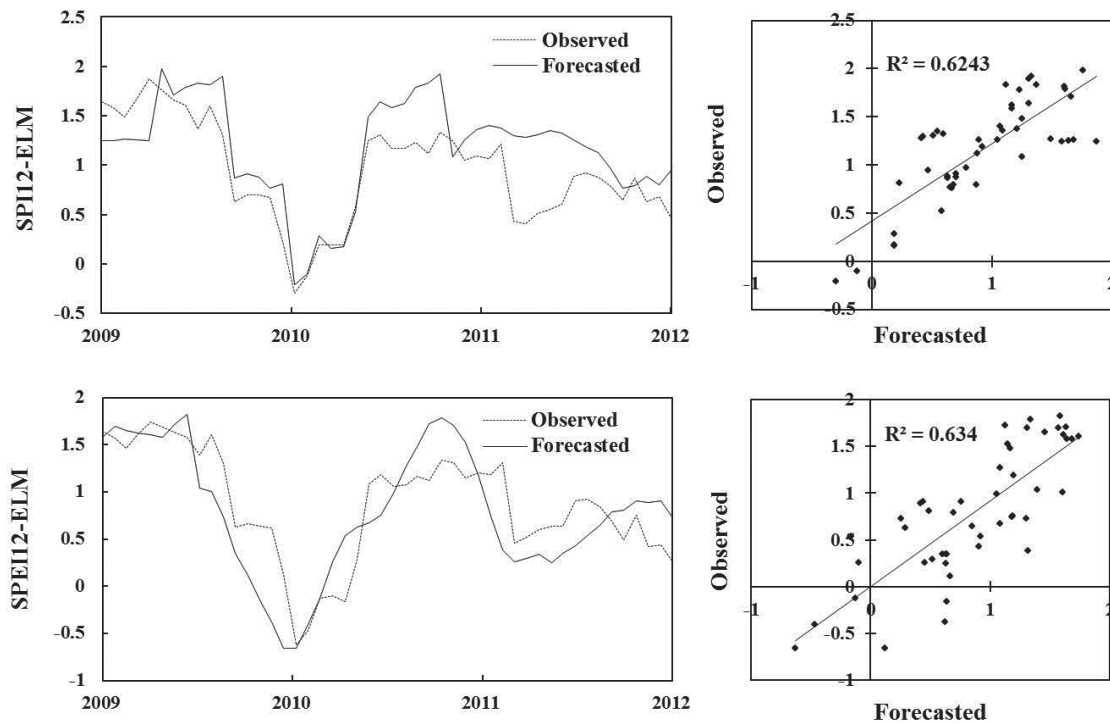


Fig. 6. The best results of ELM model for SPI/SPEI.

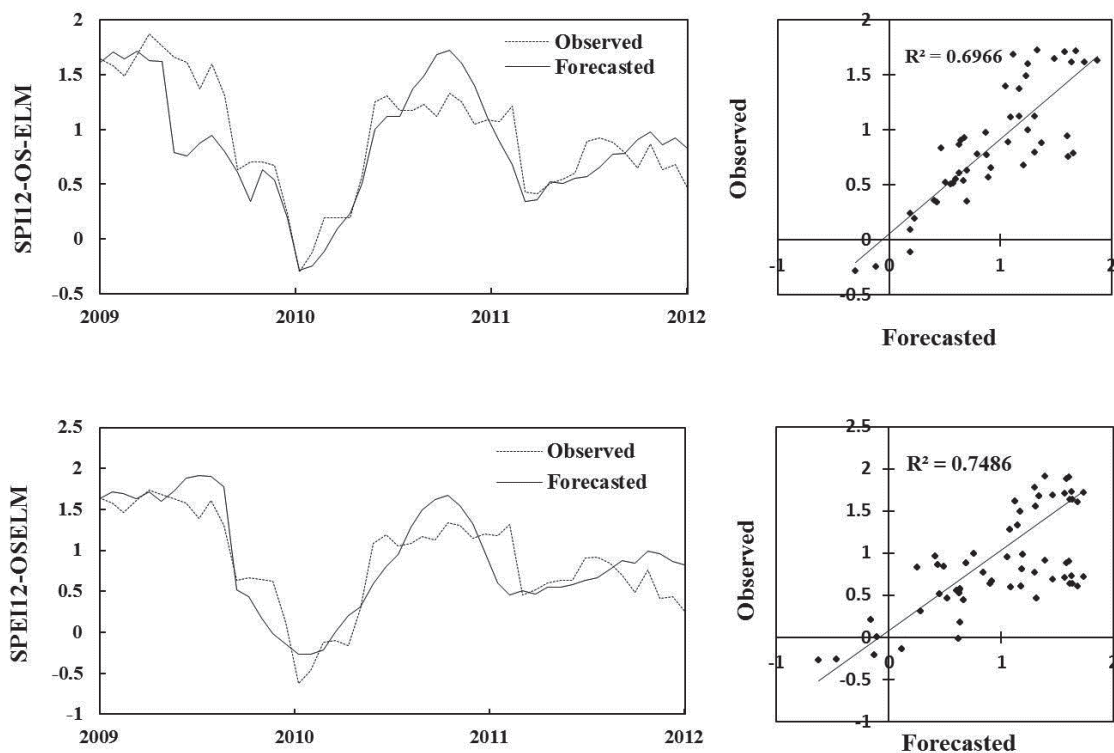


Fig. 7. The best results of OS-ELM model for SPI/SPEI.

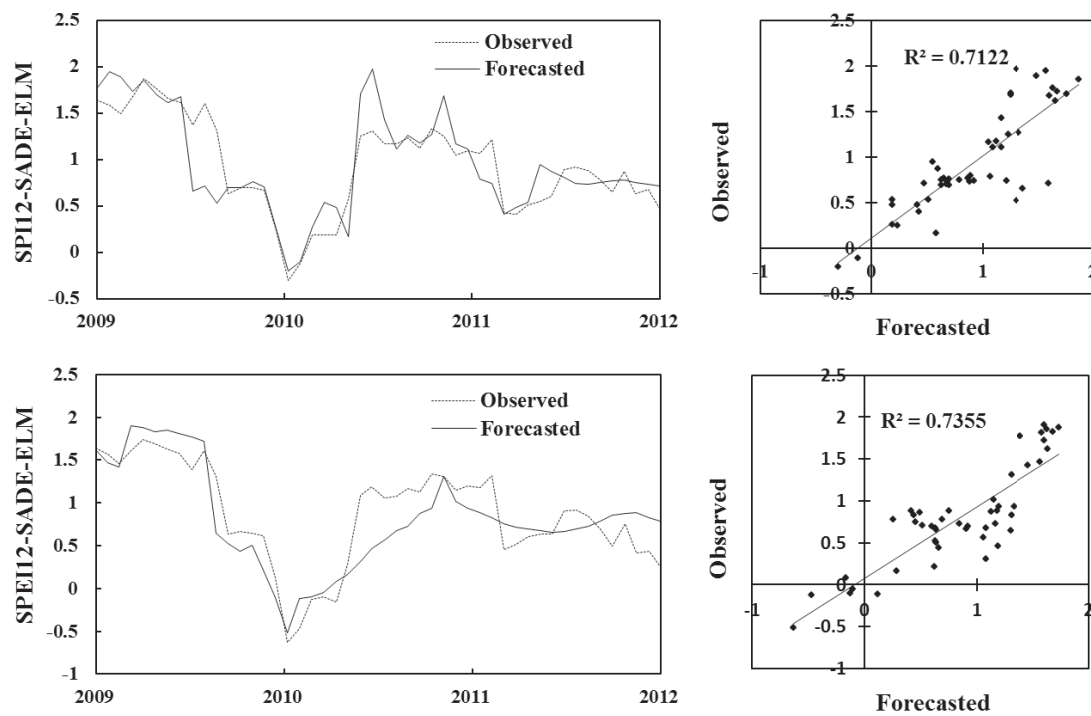


Fig. 8. The best results of SADE-ELM model for SPI/SPEI.

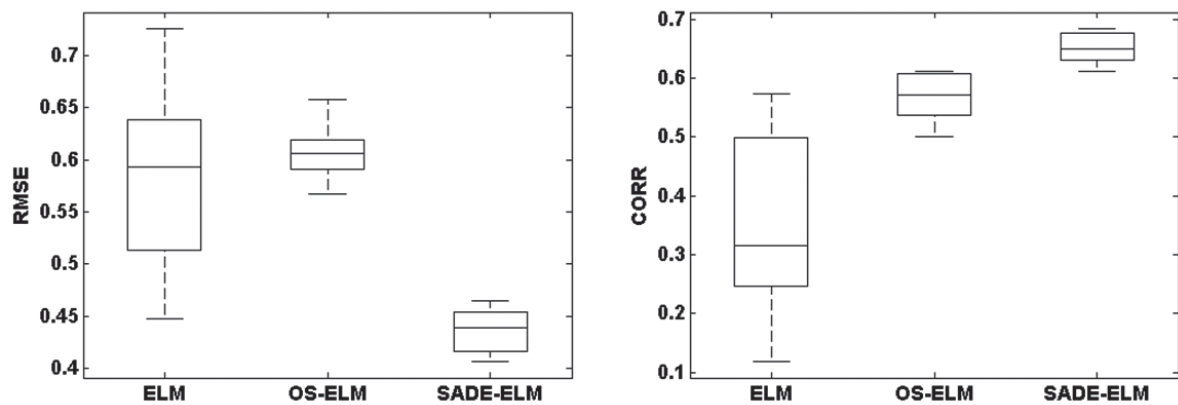


Fig. 9. The boxplot of 10 times test of M4 for SPI.

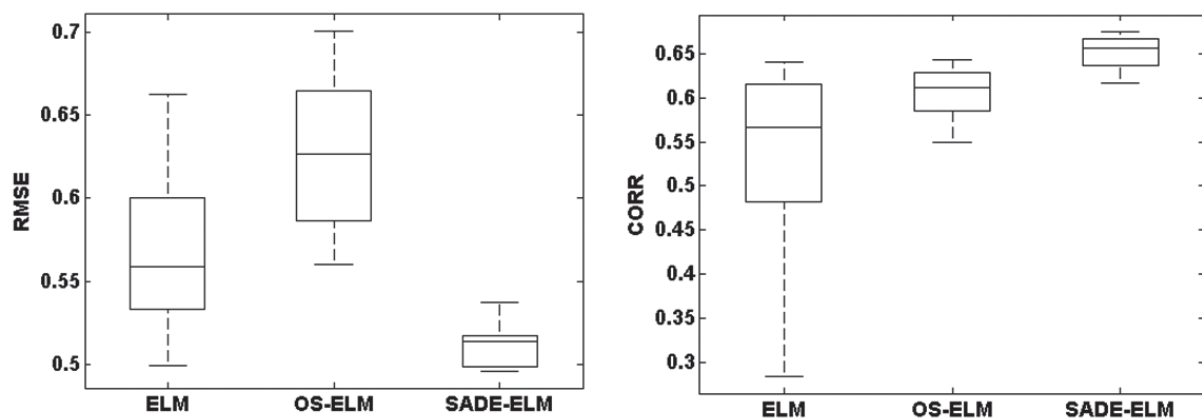


Fig. 10. The boxplot of 10 times test of M4 for SPEI.

are 0.6340, 0.7486, and 0.7355, respectively. We can find that the above-mentioned models trained by OS-ELM and SADE-ELM perform better than the models by ELM via the value of R^2 .

In order to compare the stability of ELM, OS-ELM, and SADE-ELM methods, and as it may require too much space to show the results for each phases, only a boxplot is used to present the results of 10 times test of the M4 model by these methods. These results of M4 model are presented in Figs 9 and 10. Fig. 9 shows that the OS-ELM, and SADE-ELM for SPI are more stable than ELM. The reason is that the initial weights and thresholds of ELM are randomly chosen to make the results quite different. From Fig. 10, it is seen that SADE-ELM is still most stable, and OS-ELM is more stable than ELM in terms of CORR, but the stability of OS-ELM is almost the same as ELM in terms of RMSE. For the result of RMSE, the reason may be as follow: When each batch dataset was input to the model trained for updating the output weight by OS-ELM method, the data itself may lead to poor stability of the model.

As a result, it can be concluded from the forecasting results of the experiments of this study that the models trained by ELM for drought forecasting cannot give a satisfactory performance. Additionally, the OS-ELM method can achieve better performance than ELM. Meanwhile, the SADE-ELM method can be successfully applied and provide the highest accuracy and stability for drought forecasting.

Conclusions

This study investigated the ability of the ELM, OS-ELM, and SADE-ELM models to predict drought in Khanhhoa Province of Vietnam. The results demonstrated that the ELM, OS-ELM, and SADE-ELM models can be successfully applied for drought forecasting, but the ELM models cannot obtain satisfactory performance compared with OS-ELM and SADE-ELM models for drought prediction. Furthermore, the SADE-ELM models can provide higher accuracy and more stability for drought forecasting than OS-ELM models, and the OS-ELM models can achieve better performance than ELM. Additionally, this study found that using the lagged observations of SSTA data as input variables can predict drought in the study area effectively. Meanwhile, it was obtained that the computation time of ELM models is less than OS-ELM and SADE-ELM models, without a batch input and optimized process.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 51508121) and the Guizhou Province Science and Technology Fund (Nos. LH [2016]7096, J [2015]2063, and LH [2014]7374).

References

1. AGHAKOUCHAK A., CHENG L., MAZDIYASNI O., FARAHMAND A. Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters*, **41** (24), 8847, **2014**.
2. HAO Z., AGHAKOUCHAK A., NAKHJIRI N., FARAHMAND A. Global integrated drought monitoring and prediction system. *Scientific Data*, **1**, 140001, **2014**.
3. WILHITE DA. Drought as a natural hazard: Concepts and definitions. *Drought: A Global Assessment*, **69** (1), 3, **2000**.
4. DAI A., TRENBERTH K.E., QIAN T. A Global Dataset of Palmer Drought Severity Index for 1870-2002: Relationship with Soil Moisture and Effects of Surface Warming. *Journal of Hydrometeorology*, **5** (6), 1117, **2004**.
5. YU M., LI Q., HAYES M.J., SVOBODA M.D., HEIM R. R. Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010?. *International Journal of Climatology*, **34** (3), 545, **2014**.
6. JAIN S.K., KESHRI R., GOSWAMI A., SARKAR A. Application of meteorological and vegetation indices for evaluation of drought impact: a case study for Rajasthan, India. *Natural Hazards*, **54** (3), 643, **2010**.
7. TODISCO F., MANNOCCHI F., VERGNI L. Severity-duration-frequency curves in the mitigation of drought impact: an agricultural case study. *Natural Hazards*, **65** (3), 1863, **2013**.
8. SOHRABI M. M., RYU J. H., ABATZOGLOU J., TRACY J. Climate extreme and its linkage to regional drought over Idaho, USA. *Natural Hazards*, **65** (1), 653, **2013**.
9. OSORIO J. G., GALIANO S. G. Non-stationary analysis of dry spells in monsoon season of Senegal River Basin using data from Regional Climate Models (RCMs). *Journal of Hydrology*, **450** (11), 82, **2012**.
10. WAYNE C. PALMER. Meteorological Drought. Research Paper, **45**, 58, **1965**.
11. WAYNE C. PALMER. Keeping track of crop moisture conditions nationwide: The new Crop Moisture Index. *Weatherwise*, **21**, 156, **1968**.
12. MCKEE T.B., DOESKEN N.J., KLEIST J. The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, **17** (22), 179, **1993**.
13. VICENTE-SERRANO S.M., BEGUERÍA S., LÓPEZ-MORENO J.I. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate*, **23** (7), 1696, **2010**.
14. BONACCORSO B., BORDI I., CANCELLIERE A., ROSSI G., SUTERA A. Spatial variability of drought: an analysis of the SPI in Sicily. *Water Resources Management*, **17** (4), 273, **2003**.
15. CACCIAMANI C., MORGILLO A., MARCHESI S., PAVAN V. Monitoring and forecasting drought on a regional scale: Emilia-Romagna region. *Methods and Tools for Drought Analysis and Management*, **62** (2), 29, **2007**.
16. TSAKIRIS G., VANGELIS H. Towards a drought watch system based on spatial SPI. *Water Resources Management*, **18** (1), 1, **2004**.
17. NGUYEN L. B., LI Q. F., NGOC T. A., HIRAMATSU K. Adaptive Neuro-Fuzzy Inference System for Drought Forecasting in the Cai River Basin in Vietnam. *Journal of*

- the Faculty of Agriculture Kyushu University, **60** (2), 405, **2015**.
18. KHASHEI M., BIJARI M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*, **11** (2), 2664, **2011**.
 19. ZHAO X. H., CHEN X. Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting. *Water Resources Management*, **29** (8), 2913, **2015**.
 20. ZOUNEMAT-KERMANI M. Investigating Chaos and Nonlinear Forecasting in Short Term and Mid-term River Discharge. *Water Resources Management*, **30** (5), 1851, **2016**.
 21. BOX G. BOX, JENKINS. Time series analysis forecasting and control. *A Very British Affair: Six Britons and the Development of Time Series Analysis during the 20th Century*, **11** (27), 161, **2012**.
 22. WEI S., ZUO D., Song, J. Improving prediction accuracy of river discharge time series using a Wavelet-NAR artificial neural network. *Journal of Hydroinformatics*, **14** (4), 974, **2012**.
 23. BARZEGAR R., ADAMOWSKI J., ASGHARI MOGHADDAM A. Application of wavelet-artificial intelligence hybrid models for water quality prediction: a case study in Aji-Chay River, Iran. *Stochastic Environmental Research and Risk Assessment*, **30** (7), 1797, **2016**.
 24. DAWSON C.W., WILBY R. An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, **43** (1), 47, **1998**.
 25. JAIN A., KUMAR A.M. Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, **7** (2), 585, **2007**.
 26. NAGY H.M., WATANABE K.A.N.D., HIRANO M. Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering*, **128** (6), 588, **2002**.
 27. AGARWAL A., MISHRA S.K., RAM S., SINGH J.K. Simulation of runoff and sediment yield using artificial neural networks. *Biosystems Engineering*, **94** (4), 597, **2006**.
 28. MORID S., SMAKHTIN V., BAGHERZADEH K. Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology*, **27** (15), 2103, **2007**.
 29. BARZEGAR R., ASGHARI MOGHADDAM A., ADAMOWSKI J., OZGA-ZIELINSKI B. Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model. *Stochastic Environmental Research and Risk Assessment*, <http://dx.doi.org/10.1007/s00477017-1394-z>, **2017**.
 30. BARZEGAR R., FIJANI E., ASGHARI MOGHADDAM A., TZRITIS, E. Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. *Science of the Total Environment*, **599** (600), 20, **2017**.
 31. BARZEGAR R., MOGHADDAM A.A., ADAMOWSKI J., FIJANI E. Comparison of machine learning models for predicting fluoride contamination in groundwater. *Stochastic Environmental Research and Risk Assessment*, **1**, **2016**.
 32. DEO RAVINESH C., MEHMET SAHIN An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environmental Monitoring and Assessment*, **188** (2), 90, **2016**.
 33. WU S.I., YOUYI WANG, SHIJIE CHENG. Extreme learning machine based wind speed estimation and sensorless control for wind turbine power generation system. *Neurocomputing*, **102**, 163, **2013**.
 34. WANG, JIANZHOU, JIANMING HU. A robust combination approach for short-term wind speed forecasting and analysis—Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. *Energy*, **93**, 41, **2015**.
 35. NIKOLIĆ V., MOTAMEDI S., SHAMSHIRBAND S., PETKOVIĆ D., CH S., ARIF M. Extreme learning machine approach for sensorless wind speed estimation. *Mechatronics*, **34**, 78, **2016**.
 36. ZHANG C., ZHOU J., LI C., FU W., PENG T. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Conversion and Management*, **143**, 360, **2017**.
 37. DEO R.C., TIWARI M.K., ADAMOWSKI J.F., QUILTY J.M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stochastic Environmental Research and Risk Assessment*, **31** (5), 1211, **2017**.
 38. ANAM, KHAIRUL, ADEL AL-JUMAILY. Adaptive myoelectric pattern recognition for arm movement in different positions using advanced online sequential extreme learning machine. *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the IEEE*, 900, **2016**.
 39. YADAV B., CH S., MATHUR S., ADAMOWSKI J. Discharge forecasting using an online sequential extreme learning machine (OS-ELM) model: a case study in Neckar River, Germany. *Measurement*, **92**, 433, **2016**.
 40. LU J., HUANG J., LU F. Sensor fault diagnosis for aero engine based on online sequential extreme learning machine with memory principle. *Energies*, **10** (1), 39, **2017**.
 41. EBTEHAJ I., SATTAR A.M., BONAKDARI H., ZAJI A.H. Prediction of scour depth around bridge piers using self-adaptive extreme learning machine. *Journal of Hydroinformatics*, **19** (2), 207, **2017**.
 42. XIAO C., DONG Z., XU Y., MENG K., ZHOU X., ZHANG X. Rational and self-adaptive evolutionary extreme learning machine for electricity price forecast. *Memetic Computing*, **8** (3), 223, **2016**.
 43. NAHVI B., HABIBI J., MOHAMMADI K., SHAMSHIRBAND S., AL RAZGAN O.S. Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers and Electronics in Agriculture*, **124**, 150, **2016**.
 44. REYNOLDS R.W., RAYNER N.A., SMITH T.M., STOKES D.C., WANG W. An Improved In Situ and Satellite SST Analysis for Climate. *Journal of Climate*, **15** (13), 1609, **2002**.
 45. APDRC (2013) APDRC LAS7 for public (SSTA). Available online: <http://apdrc.soest.hawaii.edu/las/v6/constrain?var=295> (accessed on 10 Oct 2013).
 46. HUANG G.B., ZHU Q.Y., SIEW C.K. Extreme learning machine: theory and applications. *Neurocomputing*, **70** (1), 489, **2006**.
 47. HUANG G.B. An Insight into Extreme learning machines: random neurons, random features and kernels. *Cognitive Computation*, **6** (3), 376, **2014**.

48. LI S., WANG P., GOEL L. A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. *IEEE Transactions on Power Systems*, **31** (3), 1788, **2016**.
49. SUN Z., CHOI T., AU K., YU Y. Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, **46** (1), 411, **2008**.
50. ZHANG R., DONG Z., XU Y., MENG K., WONG K. Short-term load forecasting of Australian National Electricity Market by an ensemble model of extreme learning machine. *IET Generation, Transmission and Distribution*, **7** (4), 391, **2013**.
51. HUANG G.B., WANG D.H., LAN Y. Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, **2** (2), 107, **2011**.
52. LIANG N.Y., HUANG G.B., SARATCHANDRAN P., SUNDARAJAN N. A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Transactions on neural networks*, **17** (6), 1411, **2006**.
53. SALIM HEDDAM, OZGUR KISI. Extreme learning machines: a new approach for modeling dissolved oxygen (DO) concentration with and without water quality variables as predictors. *Environmental Science and Pollution Research*, **24** (20), 16702, **2017**.
54. CAO J., LIN Z., HUANG G.B. Self-adaptive evolutionary extreme learning machine. *Neural Processing Letters*, **36** (3), 285, **2012**.
55. ZHANG Y., CAI Z., GONG W., WANG X. Self-adaptive differentialevolution extreme learning machine and its application in water quality evaluation. *Journal Compute Inference System*, **11** (4), 1443, **2015**.
56. DAS S., RAM G., MANDAI D., KAR R., GHOSHAL S.P. Self adaptive differential evolution based thinning of concentric circular antenna arrays. *International Conference on Digital Object Identifier*, 1, **2013**.
57. ZHAO Z., HU M. Multi-level forecasting model of coal mine water inrush based on self-adaptive evolutionary extreme learning machine. *Applied Mathematics and Information Sciences Letters*, **2** (3), 103, **2014**.
58. BEGUERÍA S., VICENTE-SERRANO SM (2009) SPEI and SPI calculator. Available online: <http://digital.csic.es/handle/10261/10002> (accessed on 15 Feb 2014).
59. WMO Standardized precipitation index user guide. In: WMO-No. 1090. World Meteorological Organization, Geneva 2, Switzerland. **2012**.
60. BACANLI U., FIRAT M., DIKBAS F. Adaptive Neuro-Fuzzy InferenceSystem for drought forecasting. *Stochastic Environmental Research and Risk Assessment*, **23** (8), 1143, **2009**.