



Drought prediction in the Yunnan–Guizhou Plateau of China by coupling the estimation of distribution algorithm and the extreme learning machine

Qiongfang Li^{1,2} · Yao Du¹ · Zhennan Liu³ · Zhengmo Zhou¹ · Guobin Lu⁴ · Qihui Chen¹

Received: 19 September 2021 / Accepted: 9 April 2022 / Published online: 10 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Drought prediction is a critical non-engineering approach to mitigate their significant threats to water availability, food safety, and ecosystem health. Therefore, to improve the efficiency and accuracy of drought prediction, a novel drought prediction model was proposed by optimizing the extreme learning machine (ELM) using the estimation of distribution algorithm (EDA) (EDA-ELM) and evaluated by the comparison with the genetic algorithm-optimized ELM (GA-ELM) model, standard ELM model, and adaptive network-based fuzzy inference system (ANFIS) in drought prediction for Yunnan–Guizhou Plateau (YGP). The standardized precipitation evapotranspiration index (SPEI) in 3/6/12-month time scales was treated as the dependent variable and the primary drought driving factors as predictor variables. The results revealed that the EDA-ELM model performed best in multiscalar SPEI prediction, followed by GA-ELM, ANFIS, and standard ELM models, while the model execution time was descended by EDA-ELM, GA-ELM, ANFIS, and standard ELM models, varying from 100 to 700 s. The outputs could provide a novel approach to drought prediction and benefit drought prevention and mitigation.

Keywords EDA-ELM model · GA-ELM model · ANFIS model · Drought prediction · The Yunnan–Guizhou Plateau

Abbreviations

ANFIS	Adaptive network-based fuzzy inference system
ANN	Artificial neural network
CORR	Correlation coefficient

✉ Qiongfang Li
qfli@hhu.edu.cn

✉ Yao Du
duyao@hhu.edu.cn

¹ College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

² Yangtze Institute for Conservation and Development, Hohai University, Nanjing 210098, China

³ School of Civil Engineering, Guizhou Institute of Technology, Guizhou University, Guiyang 550000, China

⁴ Hohai University, Nanjing 210098, China

D_o/D_p	Observed/predicted drought duration
EDA	Estimation of distribution algorithm
EDA-ELM	Estimation of distribution algorithm-optimized extreme learning machine
ELM	Extreme learning machine
ENSO	El Niño and the Southern Oscillation
E_{NS}	Nash–Sutcliffe coefficient
GA	Genetic algorithm
GA-ELM	Genetic algorithm-optimized ELM
MAE	Mean absolute error
RMSE	Root mean square error
RE_s/RE_D	Relative errors of drought severity and duration
SLFN	Single-hidden-layer feedforward neural network
SPEI	Standardized Precipitation Evapotranspiration Index
S_o/S_p	Observed/predicted drought severity
WI	Willmott's index

1 Introduction

Drought, an extreme climate phenomenon with high frequency and long duration, has caused significant disruptions to agricultural production, water supply, and economic stability worldwide (Forootan et al. 2019; Spinoni et al. 2018; Zhang and Zhou 2015; Orimoloye et al. 2021; Khan et al. 2021; Yildirim and Rahman 2021). The disruptions induced by drought are even more significant in China due to its unique geographical and climatic conditions (Ma et al. 2020; Li et al. 2020; Zhao et al. 2020; Yildirim et al. 2022). In the past half-century, China has experienced at least one severe drought event a year, with drought-induced economic losses amounting to \$7 billion per year on average (Su et al. 2018). Drought has become a critical constraint on the sustainable socio-economic development of China by posing severe threats to the national food and water supply security (Yao et al. 2020; Zhang et al. 2015). Drought prediction and mitigation have always been a hot topic in China, but previous drought-related studies in China have mainly been concentrated on semi-arid and semi-humid regions in Northeastern, Northwestern, and Northern China and paid little attention to humid areas. However, with climate change, high-intensity and long-duration droughts frequently occur in humid regions (Xu et al. 2015; Zhang et al. 2017; Cheng et al. 2020). As one of the typical humid regions in China, Southwestern China, which covers Yunnan, Sichuan, Guizhou, and Guangxi Provinces, and Chongqing City, is facing a growing risk of drought (Wang et al. 2014; Zeng et al. 2019). This region has experienced several severe region-wide droughts in the past decade, including the once-in-a-century one lasting from autumn 2009 to spring 2010, which not only caused water shortage in the drinking water supply for 25 million people and replenishment to rivers and irrigation but also resulted in substantial economic losses of over 40 billion yuan (Yu et al. 2014; Sun et al. 2019). Unfortunately, the studies on droughts in Southwestern China have not attracted wide attention so far and even more so for its drought prediction. Due to the unique geographical, geomorphological, and climatological conditions of Southwestern China, the drought-driving mechanisms are quite different from other regions, which are not only closely associated with precipitation and evaporation but also the interaction between monsoons, topographic features, and large-scale atmospheric circulation (Cheng et al. 2020; Wang et al. 2014, 2015, 2020a). Therefore, the drought

prediction with traditional physically based models confront considerable challenges, and developing an artificial intelligence model for drought prediction in Southwestern China could be a better choice.

The artificial intelligence model can effectively model non-linear and high-dimensional data with intricate connections and missing values (Knudby et al. 2010; Rahmati et al. 2020), with no requirements for a thorough understanding of the interaction mechanisms between predictors and predictand (Wu and Chau 2011). As a result, artificial intelligence models have been widely used in drought prediction. By using sea surface temperature and sea level pressure data, Farokhnia et al. (2011) used the adaptive neuro-fuzzy inference system (ANFIS) to forecast possible droughts in the Tehran Plain 3, 6, and 9 months in advance. However, the ANFIS model, created by combining artificial neural network (ANN) and gradient-based learning algorithms, learns relatively slowly and easily converges to a local minimum (Dehnavi et al. 2015; Jaafari et al. 2019). Accordingly, the extreme learning machine (ELM) model, a novel single-hidden-layer feedforward neural network (SLFN) proposed by Huang et al. (2006), has gained popularity in modeling hydroclimatic variables due to its ability to avoid many complex problems, such as slow learning rate, complex circular structure, learning epochs, and local extremum, which all exist in traditional gradient-based algorithms (Huang et al. 2012). The ELM model's random-assigned input weights and hidden layer biases simplify nonlinear forecasting problems by a set of linear equations, which use Moore–Penrose generalized inverse to determine the output weights (Huang et al. 2015; Deo et al. 2017). Deo and Şahin (2015) compared the ELM model with the classical ANN model for drought prediction in eastern Australia and found that the ELM model achieved significantly less computational time and higher accuracy than the ANN model. The effectiveness and efficiency of the ELM model in drought prediction have also been confirmed in literature (Liu et al. 2018; Ali et al. 2018; Mouatadid et al. 2018). However, a standalone ELM model is prone to underfitting or overfitting some degrees due to the uncertainty induced by the random generation of the ELM model's input weights and hidden layer biases (Alencar et al. 2016; Zhu et al. 2020). Subsequently, the ELM model was improved by coupling with evolutionary algorithms to overcome the defects and has been successfully applied in various fields but seldom in drought prediction (Bui et al. 2019; Feng et al. 2019; Wu et al. 2020). Besides, the current optimization of the ELM model is mainly carried out by using individual evolutionary mechanism-based algorithms that implicitly describe the probability distribution of potential solutions. Different from the individual evolutionary mechanism-based algorithms, the estimation of distribution (EDA) algorithm (Mühlenbein et al. 1996; Larrañaga and Lozano 2002) explores the optimal solution by building and sampling explicit probabilistic models of potential solutions. Its strong global search capabilities and fast convergence speed allow it to feasibly solve optimization problems, which are notoriously difficult to be handled by most conventional evolutionary algorithms (Hauschild and Pelikan 2011; Luo et al. 2015; Wang et al. 2020b). Therefore, it is necessary to improve the ELM model by coupling the EDA algorithm.

Although several studies mentioned above confirmed the effectiveness and efficiency of the ELM in drought prediction, the applicability of improved ELM by coupling EDA algorithm (namely EDA-ELM model) for drought prediction is still to be investigated, particularly in the Yunnan–Guizhou Plateau, where few drought prediction models have been developed. Therefore, the objectives of this study are as follows:

1. To improve the ELM by coupling EDA;

2. To examine the applicability of EDA-ELM in drought prediction of the Yunnan–Guizhou Plateau;
3. To explore the advantages and disadvantages of EDA-ELM in drought prediction of the Yunnan–Guizhou Plateau over GA-ELM, standard ELM model, and AFNIS model in terms of prediction accuracy and execution time.

2 Materials and methods

2.1 Study site

The Yunnan–Guizhou Plateau, one of the four major plateaus in China, is located between 22° N and 30° N and between 100° E and 111° E, and it covers the eastern part of Yunnan Province and the whole area of Guizhou Province in Southern China with a total area of approximately 140,000 km² and an altitude varying from 1000 to 2000 m. Under the joint influence of the South Asian monsoon, the East Asian monsoon, Plateau monsoon, and westerlies, the annual average precipitation of 1500 mm to 1750 mm is approximately 1219.7 mm, with more than 85% to 95% falling within May to October. The annual precipitation variation is significant, with the maximum one to the minimum one being about 1.6. The fiercely uneven inter-annual and intra-annual distribution of precipitation, the significant decreasing trend in precipitation, and the increasing trend in evaporation in a global warming environment (Cheng et al. 2020) have caused the frequent occurrence of droughts. Seasonal droughts with different scales occur almost every year in this region, and severe droughts covering a large area occur every 5–10 years. Furthermore, the drought frequency and intensity in Yunnan–Guizhou plateau are expected to further increase with climate change (Wang et al. 2014 2020a; Wang and Chen 2014; Zeng et al. 2019). Thus, droughts are increasingly becoming a critical constraint to its socio-economic development in the region, and drought prediction is in urgently needed.

2.2 Data collection and processing

The monthly precipitation and temperature data from 1971 to 2016 at 25 meteorological stations, which scatter throughout the Yunnan–Guizhou Plateau (Fig. 1), were collected from the National Meteorological Information Center, and the data consistency and reliability had been verified before release by China Meteorological Administration (<https://data.cma.cn/en>), and 99% of the data are assured correct. The SPEI in the time scales of 3, 6, and 12 months during 1972–2016 for each of the 25 weather stations were calculated by referring to Vicente-Serrano et al. (2010). The areal SPEI in the time scale of 3/6/12 months as the dependent variable of drought prediction models were then obtained using the Tyson Polygon method. The primary drought-driving factors identified by Li et al. (2022) corresponding to SPEI in the time scale of 3/6/12 months, respectively (see Table 1), were collected as predictor variables of drought prediction models.

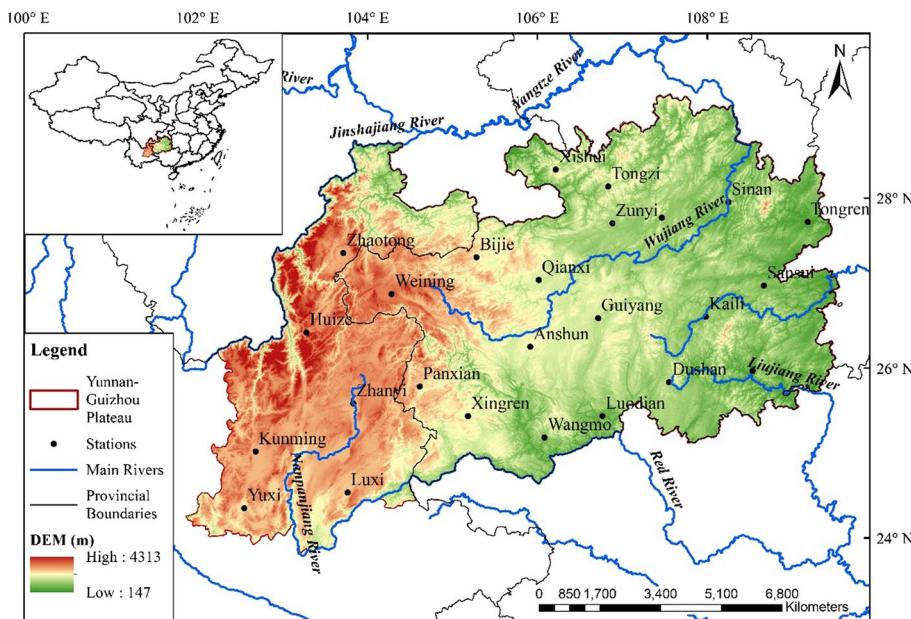


Fig. 1 Location of the meteorological stations in Yunnan–Guizhou Plateau

Table 1 Drought-driving circulation index selected by RIC-CMIDFS algorithm

Drought Indices	Circulation Indexes
SPEI-3	C1 ₅ , C2 ₆ , C3 ₈ , C2 ₅ , C1 ₉ , C1 ₁₄
SPEI-6	C1 ₇ , C1 ₉ , C4 ₂₂ , C5 ₂₁ , C6 ₂₂ , C7 ₂₁
SPEI-12	C1 ₉ , C5 ₃ , C1 ₁₂ , C8 ₂₄ , C2 ₃

C1: Atlantic Multi-decadal Oscillation Index; C2:NINO 3.4 SST Index; C3: NINO W SST Index; C4: NINO Z SST Index; C5: NINO 3 SST Index; C6: NINO C SST Index; C7: ENSO Modoki Index; C8: Cold-tongue ENSO Index. Subscript indicates lag time. SPEI-3, -6,-12 represent 3-, 6-, and 12-month accumulation values of SPEI, respectively

2.3 Development and application of EDA-ELM model

2.3.1 EDA algorithm

The EDA algorithm, first proposed in 1996, is a unique evolutionary algorithm that guides the search for the optimum by building and sampling explicit probabilistic models of promising candidate solutions, which can be updated step by step by a series of evolution, starting with the model encoding an uninformative prior over admissible solutions and ending with the model generating the global optima (Li and Aickelin 2007; Li and Zhao 2012). It is different from the most conventional evolutionary algorithms, which generate new candidate solutions using an implicit distribution defined by one or more variation operators. Thus, the EDA algorithm has unique advantages over other algorithms in maintaining the important structure of promising solutions and

Fig. 2 General process of the EDA algorithm

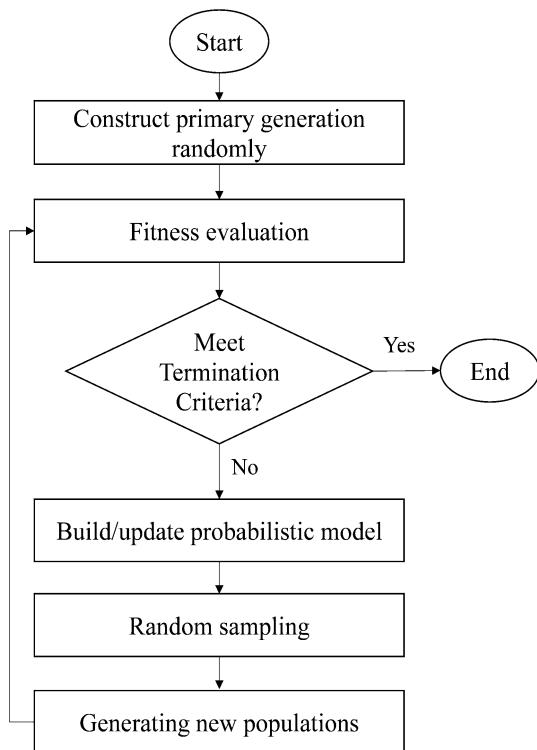


Table 2 Pseudocode of EDA algorithm

- (1) Convert problem to a searchable solution using encoding scheme
- (2) Determine parameter of the problem
- (3) Determine parameters of the algorithm including the number of iteration and population
- (4) Generate a primary population randomly
- (5) Evaluating primary population
- (6) The specific number of repetition of generation evolution process includes:
 - a. Determine elite individuals
 - b. Construct a probabilistic model with respect to elites
 - c. Generate next generation using the probabilistic model
 - d. Evaluate the population
- (7) Repeat Step 6 until meet the termination criteria

solving a wide range of problems in the discrete and continuous domains (Hauschild and Pelikan 2011; Chen et al. 2017; Gao and de Silva 2018; Liang et al. 2019). Figure 2 shows the general process of the EDA algorithm, and Table 2 presents its pseudocode.

In this paper, one of the widely used variants of the EDA algorithm, namely population-based incremental learning (Baluja 1994), was adopted, and its main characteristic is that the formula for updating the probabilistic model is special.

The algorithm assumes that $p(x) = (p(x_1), p(x_2), \dots, p(x_n))$ is a probability vector of the probabilistic models, where $p(x_i)$ ($i = 1, 2, \dots, n$) represents the probability corresponding to the i^{th} gene position with the value being one. The M individuals of each generation are randomly generated by probability vectors $p(x)$ through the iterative process of the EDA algorithm evolution. The corresponding value of each individual is calculated. Among the M individuals, the elite N ($N < M$) individuals are selected to update the probability vector $p(x)$ on the basis of the Hebb rule (Baluja 1994). Given that $p_l(x)$ represents the l^{th} probability vector, and $x_l^1, x_l^2, \dots, x_l^N$ represent the selected N ($N < M$) best individuals, the update process can be expressed as follows:

$$p_{l+1}(x) = (1 - \alpha)p_l(x) + \alpha \frac{1}{N} \sum_{k=1}^N x_l^k, \quad (1)$$

α is the learning rate.

2.3.2 EDA-ELM model

The ELM model is a widely used algorithm proposed by Huang et al. (2006) to solve single layer feedforward neural networks, and Fig. 3 illustrates its basic structure. The advantage of the ELM model lies in its random assignment of the input weights and hidden layer biases, without the need for iterative tuning of the hidden neuron parameters, which is required for the case of traditional SLFN training algorithms (e.g., back propagation algorithm). This model simplifies nonlinear problems by linear equations and performs better and faster than traditional SLFN algorithms. However, the uncertainty induced by the ELM model's random generation easily results in underfitting or overfitting (Alencar et al. 2016; Zhu et al. 2020). Hence, evolutionary algorithms have been applied to find appropriate weights and biases to advance the ELM model (Zhu et al. 2005; Guo et al. 2016; Wu et al. 2017; Feng et al. 2021).

EDA, as one of alternative optimization algorithms, was proposed to couple with ELM to optimize its input weights and hidden biases. The EDA-ELM model aims to

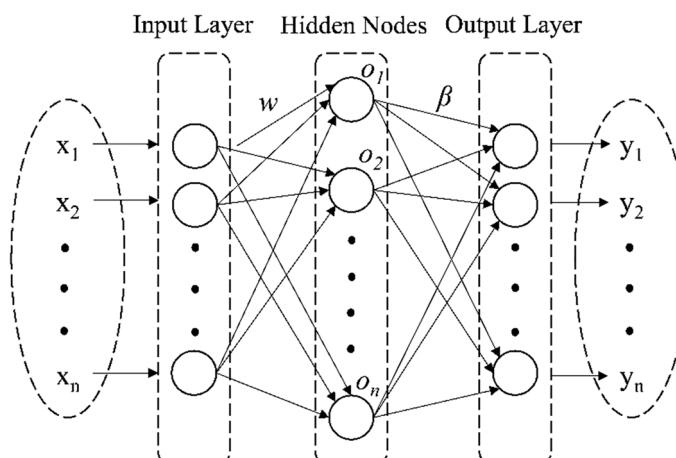


Fig. 3 Topological structure of the extreme learning machine network used in this study

solve a search problem in the space of all possible weights and hidden biases, with a search target to maximize the fitness between the predicted and the observed values (represented by fitness functions). During the search process, the search and evolution characteristics of EDA are used as function optimizers to maximize fitness functions, and the challenge is to identify and remove the weights and hidden biases with large errors.

The EDA-ELM model can be applied by following steps: (1) randomly generate the population of EDA module, and each individual in the population is composed of a set of input weights and hidden biases; (2) deliver each individual back to the ELM module as its input weights and hidden biases to calculate the fitness value; (3) conduct evolutionary process and random sampling in EDA module to create a new population after the fitness of all individuals in the population is calculated.

The iterative loops of step 2 and step 3 are repeated until the fitness of one individual satisfies the fitness function, and the corresponding input weights and hidden biases are the optimal settings for the ELM model. In this paper, the fitness function was set as the norm of the error matrix between the predicted and the observed values less than 0.00001, and random sampling uses the Monte Carlo method. The flowchart of the EDA-ELM model is shown in Fig. 4.

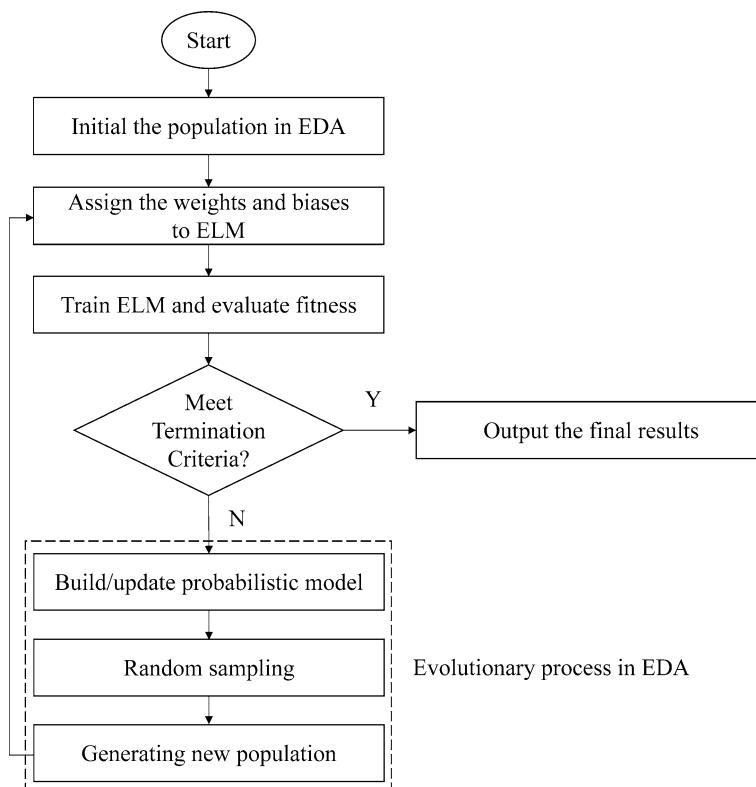


Fig. 4 Flowchart of EDA-ELM model

2.3.3 Drought prediction

In the application of the EDA-ELM model for drought prediction in the Yunnan–Guizhou Plateau, it was run in the training (1972–2006) and testing (2007–2016) periods, respectively, in a MATLAB environment provided by the i5-7200u, 2.70 GHz CPU system to test its applicability in drought prediction. The primary drought-driving factors and multiscalar SPEI were standardized by the formula to improve the efficiency of the model operation:

$$x_{\text{nor}} = \frac{(x_{\max} - x_{\min}) \cdot (x - x_{\min})}{(x_{\max} - x_{\min})} + x_{\min}, \quad (2)$$

x is the original data; x_{\max} and x_{\min} are the maximum and minimum values in the original data set, respectively; and x_{nor} is the standardized value of x .

Given that the EDA-ELM model was developed by integrating the EDA algorithm into the ELM model, their parameters can be separately determined. The number of input layer neurons in ELM was the same as that of the climate system indices in the input schemes. The output layer had only one neuron representing the predicted monthly SPEI. The number of hidden layer neurons was determined via the trial-and-error analysis, and 40 neurons were finally selected based on testing a maximum of 100 neurons in this study. According to the number of input and hidden layer neurons, the number of input weights was determined with the number of hidden biases being consistent with the number of hidden layer neurons. The excitation function was selected as “sig.” In EDA, the population size was set as 40, the learning rate as 0.01 and the maximum evolutionary algebra as 10,000.

The GA-ELM (Yang et al. 2013), standard ELM (Huang et al. 2006), and ANFIS (Jang 1993) models were also applied for the drought prediction in the Yunnan–Guizhou Plateau to reveal the effectiveness and efficiency of the EDA-ELM model by comparing their performance in drought prediction. In ELM, GA-ELM, and EDA-ELM, the same parameter values for the ELM module were adopted. In the GA algorithm module, the population size was set as 40, the cross-probability as 0.7, the probability of variation as 0.01, the generation gap as 0.85, and the maximum evolutionary algebra as 10,000. In the ANFIS model, the initial fuzzy inference system (FIS) was generated by genfis3, the Gaussian function was treated as the membership function, and the number of membership functions varies with the number of input primary drought-driving factors; the partition matrix exponent, the maximum number of iterations, and minimum improvement were set as 2, 100, and 0.00001, respectively; the times of the model training were set as 300.

2.4 Model performance evaluation

The evaluation of model performance in drought prediction was carried out by use of the following five indicators:

- (1) Correlation coefficient (CORR) expressed as:

$$\text{CORR} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

(2) Root mean square error (RMSE) expressed as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}} \quad (4)$$

(3) Mean absolute error (MAE) expressed as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |(x_i - y_i)| \quad (5)$$

(4) Willmott's Index (WI) of agreement expressed as:

$$\text{WI} = 1 - \left[\frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (|x_i - \bar{y}| + |y_i - \bar{y}|)^2} \right] 0 \leq \text{WI} \leq 1 \quad (6)$$

(5) Nash–Sutcliffe coefficient (E_{NS}) expressed as:

$$E_{\text{NS}} = 1 - \left[\frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] -\infty \leq E_{\text{NS}} \leq 1 \quad (7)$$

x_i and y_i are the predicted and the observed SPEIs in month i of the forecast period, \bar{x}, \bar{y} is their average value, and n ($= 60$) is the length of the forecast period.

In addition, the model performance was also evaluated and compared by use of the distribution characteristics of prediction errors and the relative errors of drought severity and duration:

$$\text{RE}_S = \frac{S_p - S_O}{S_O} \times 100\% \quad (8)$$

$$\text{RE}_D = \frac{D_p - D_O}{D_O} \times 100\% \quad (9)$$

S_O/S_p are observed/predicted drought severity measured by the accumulated value of the negative observed/predicted SPEI. D_O/D_p are observed/predicted drought duration determined by summing up the months when the observed/predicted SPEI is negative.

Table 3 Statistics of performance indicators for different models in the testing period

Drought index	Prediction model	CORR	RMSE	MAE	WI	E_{NS}
SPEI-3	EDA-ELM	0.944	0.341	0.254	0.974	0.906
	GA-ELM	0.901	0.470	0.345	0.946	0.821
	ANFIS	0.773	0.711	0.555	0.861	0.590
	ELM	0.636	0.882	0.717	0.644	0.369
SPEI-6	EDA-ELM	0.922	0.479	0.385	0.956	0.839
	GA-ELM	0.854	0.577	0.448	0.928	0.767
	ANFIS	0.757	0.730	0.558	0.868	0.626
	ELM	0.649	0.974	0.794	0.607	0.335
SPEI-12	EDA-ELM	0.913	0.500	0.386	0.951	0.833
	GA-ELM	0.849	0.648	0.499	0.909	0.719
	ANFIS	0.796	0.764	0.632	0.843	0.609
	ELM	0.619	0.959	0.817	0.718	0.384

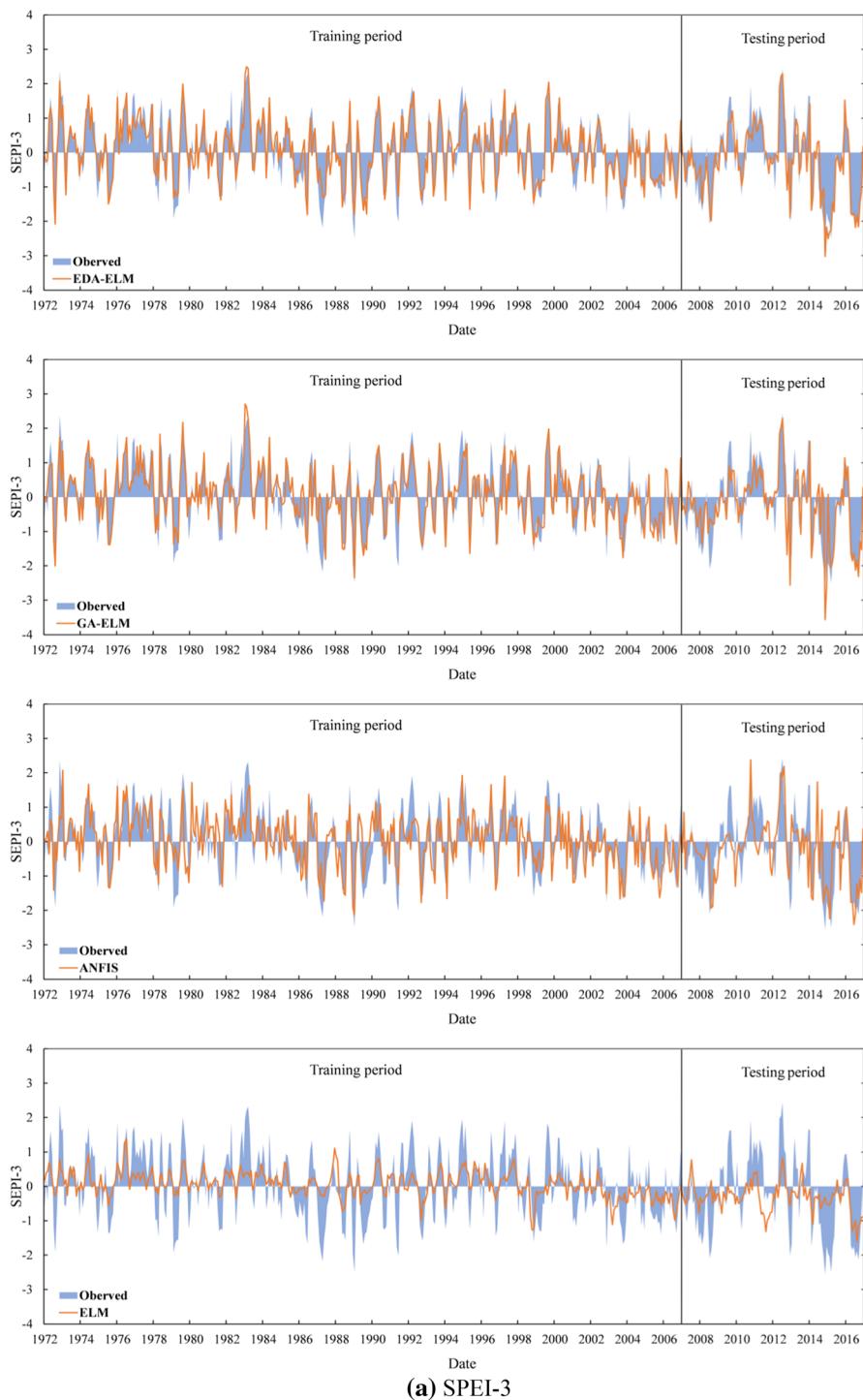
3 Results

3.1 Model performance in terms of CORR, RMSE, MAE, WI, and E_{NS}

The performance indicators of the EDA-ELM, GA-ELM, ANFIS, and standard ELM models are summarized in Table 3 to examine their performance in drought prediction. Table 3 shows that the EDA-ELM model performed best in predicting all multiscale SPEI during the testing period in CORR, RMSE, MAE, WI, and E_{NS} , followed by GA-ELM, ANFIS, and standard ELM models. The prediction performance of EDA-ELM, GA-ELM, and ELM varied with timescales of drought indices (i.e., the model performance is good when the time scale is short). However, the EDA-ELM model performed more stable for different time scalar drought indices than GA-ELM and ELM models. The ELM model exhibited much poorer performance than the other three models, and it can be inferred that the ELM should be improved by using a certain type of optimization algorithm when applied in drought prediction. The ANFIS model, which was more commonly used in drought prediction than the GA-ELM model, achieved poorer accuracy than the GA-ELM model but performed better than the ELM model. This finding implied that the ELM model could gain higher accuracy in drought prediction than ANFIS by optimization algorithms.

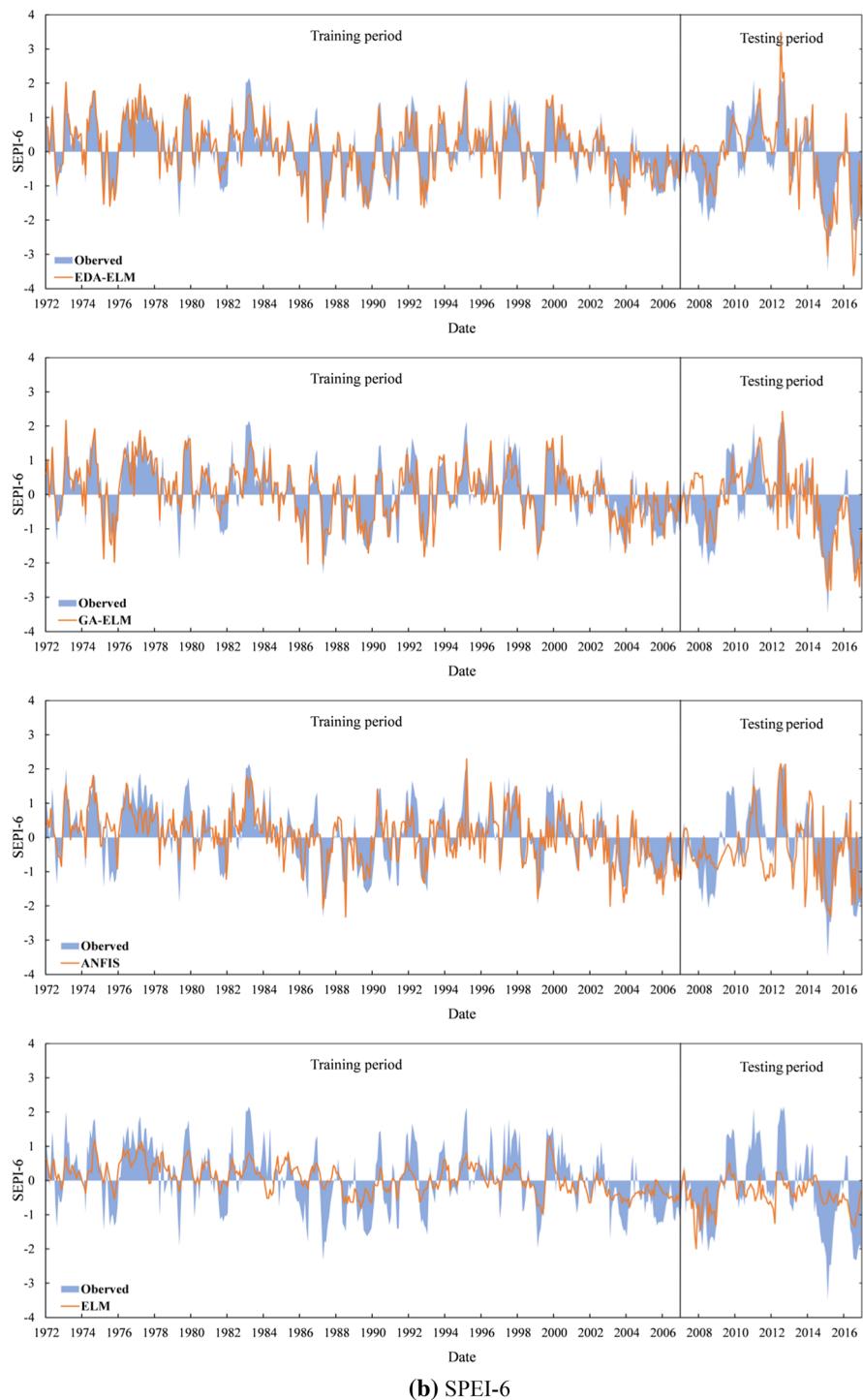
The multiscale SPEI values obtained by the EDA-ELM, GA-ELM, ELM, and ANFIS models and the observed ones are illustrated in Fig. 5. Figure 5 demonstrates that the model performance varied with the model types and timescales of SPEI in the training and testing periods. The multiscale SPEI values obtained by the EDA-ELM model best matched the observed ones, whereas the ELM model performed worst among all models. For EDA-ELM, GA-ELM, and ELM models, the best match between the predicted and the observed SPEI values was at the 3-month time scale, and the worst was at the 12-month time scale, which is consistent with Table 3.

The regression analysis between the predicted multiscale SPEI and observed ones for the testing period of 2007 to 2016 is illustrated in Fig. 6. The closer to 1 the gradient of the regression line or the regression coefficient, the better the model



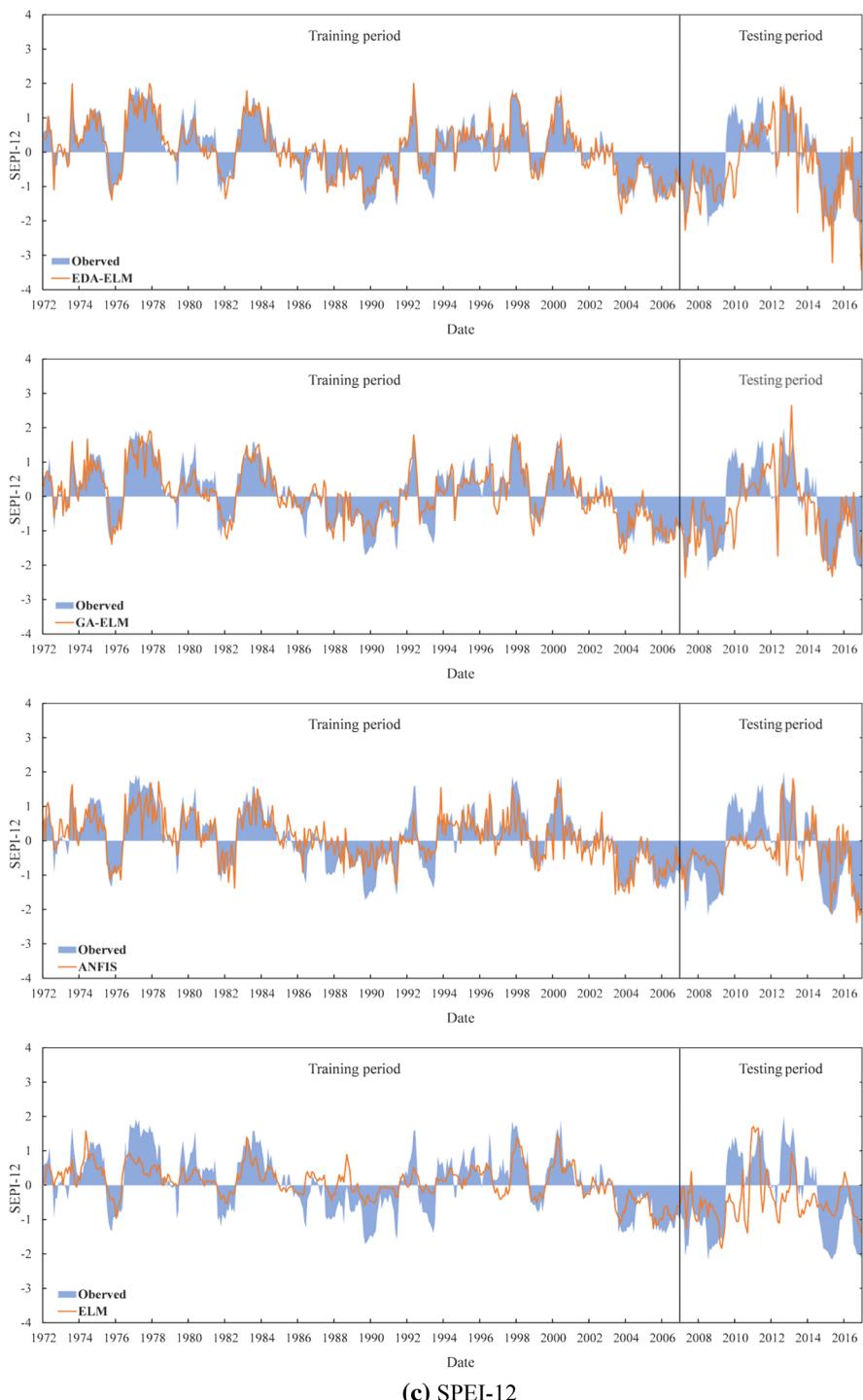
(a) SPEI-3

Fig. 5 Variations in the observed and computed monthly multiscalar SPEI during 1972–2016



(b) SPEI-6

Fig. 5 (continued)



(c) SPEI-12

Fig. 5 (continued)

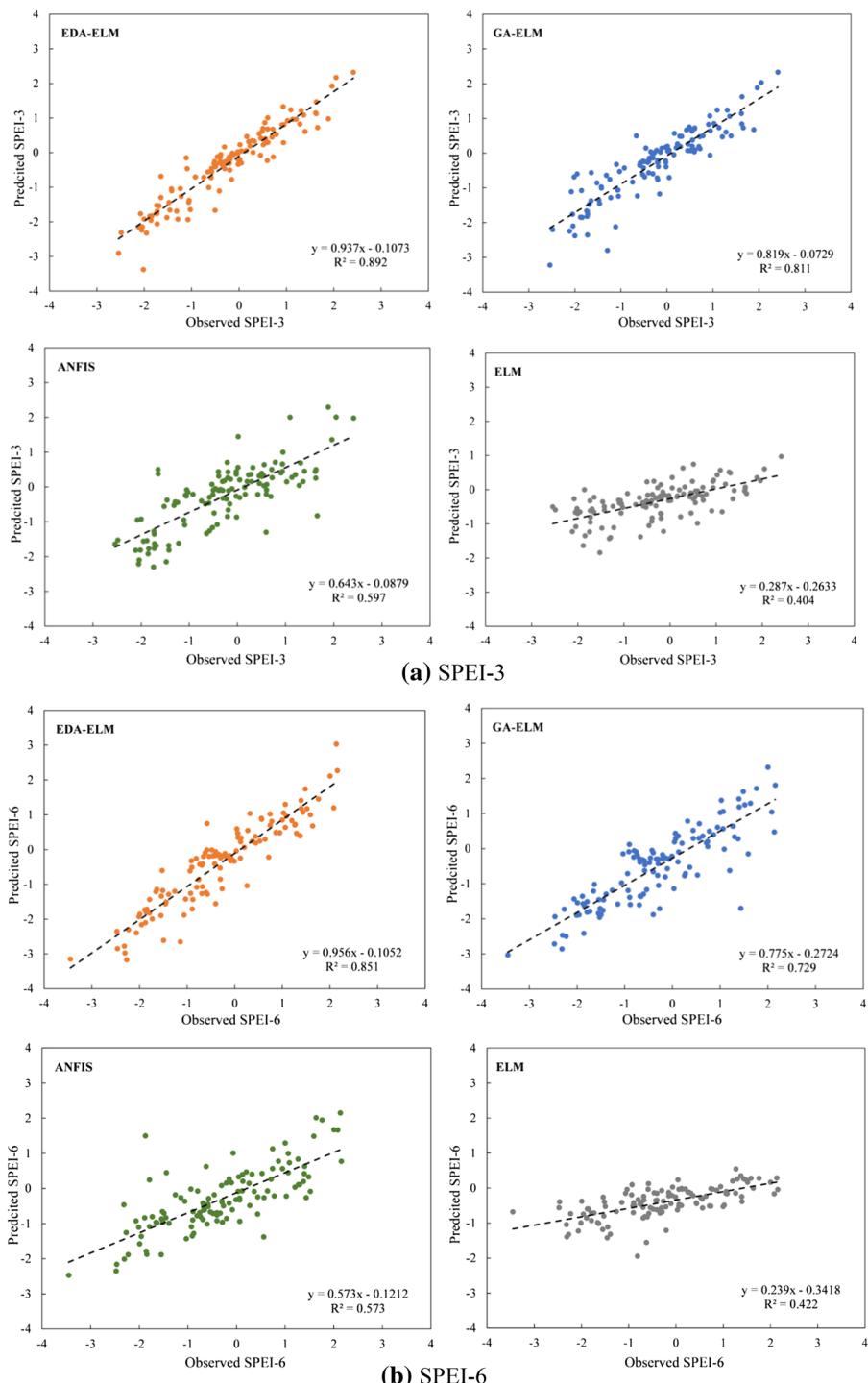


Fig. 6 Correlation between the predicted and the observed multiscalar SPEIs for the testing period of 2007 to 2016

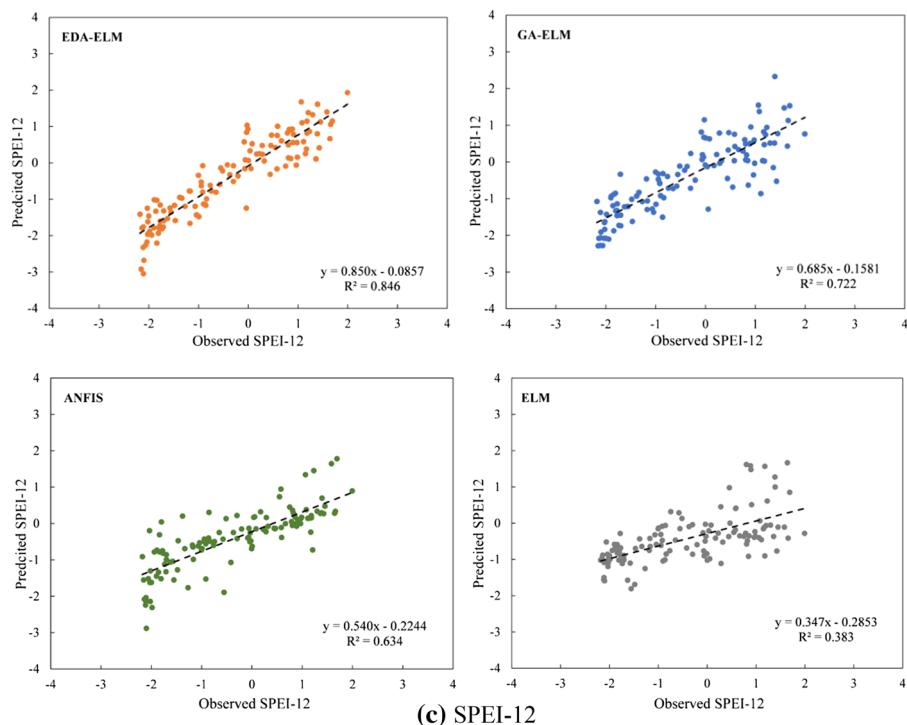


Fig. 6 (continued)

performance. For the EDA-ELM model, the gradient of the regression lines for different multiscalar SPEIs is not lower than 0.85, and the regression coefficient between the predicted and the observed multiscalar SPEI (R^2) is above 0.84. For the ELM model, the gradient varied from 0.239 to 0.347 and R^2 from 0.383 to 0.422, much lower than those of the EDA-ELM model. For GA-ELM and ANFIS models, the gradient and R^2 were bigger than those of the ELM model but smaller than those of the EDA-ELM model. Therefore, the regression analysis between the predicted and the observed multiscalar SPEI also revealed that the EDA-ELM model performed best, followed by the GA-ELM, ANFIS, and ELM models. The ELM model failed to capture the low values of the multiscalar SPEI, which represents the severer droughts.

3.2 Distribution characteristics of the multiscalar SPEI prediction errors

The frequencies of the absolute errors between the predicted and the observed monthly multiscalar SPEI during the test period for the EDA-ELM, GA-ELM, ELM, and ANFIS models are presented in Table 4. For the EDA-ELM model, the percentage of the multiscale SPEI errors falling within ± 0.2 varied between 39.2% and 51.7%, and this percentage changed to a range of 58.4%–81.7% when the errors range was set as -0.4 to $+0.4$. These two percentages became 25.8%–42.5% and 52.5%–68.3% for the GA-ELM model, 16.7%–24.2% and 33.4%–51.7% for the ANFIS model, and 12.5%–17.5% and 21.7%–36.0% for the ELM model. The smaller the percentage with small absolute errors,

Table 4 Frequencies of absolute errors between predicted and observed monthly multiscale SPEIs falling within ranges for the EDA-ELM versus GA-ELM, ANFIS, and ELM models

Range of predicted errors $ PE $	SPEI-3			SPEI-6			SPEI-12			
	EDA-ELM	GA-ELM	ANFIS	EDA-ELM	GA-ELM	ANFIS	EDA-ELM	GA-ELM	ANFIS	ELM
$0.0 \leq PE < 0.2$	51.7	42.5	23.3	17.5	46.7	25.8	24.2	15.8	29.2	16.7
$0.2 \leq PE < 0.4$	30.0	25.8	23.3	18.3	20.0	28.3	27.5	18.3	19.2	23.3
$0.4 \leq PE < 0.6$	5.0	12.5	15.0	12.5	16.7	18.3	10.0	8.3	17.5	15.8
$0.6 \leq PE < 0.8$	5.8	9.2	13.3	15.0	3.3	11.7	12.5	9.2	7.5	11.7
$0.8 \leq PE < 1.0$	5.8	3.3	14.2	10.8	9.2	5.0	7.5	13.3	12.5	7.5
$1.0 \leq PE < 1.2$	0.8	3.3	4.2	6.7	1.7	5.0	10.0	10.8	3.3	4.2
$1.2 \leq PE < 1.4$	0.8	2.5	2.5	2.5	1.7	0.8	2.5	10.8	0.8	5.0
$1.4 \leq PE < 1.6$	0.0	0.8	0.8	8.3	0.8	1.7	0.0	4.2	0.0	1.7
$1.6 \leq PE < 1.8$	0.0	0.0	0.0	5.0	0.0	1.7	1.7	3.3	0.0	0.0
$1.8 \leq PE < 2.0$	0.0	0.0	0.8	2.5	0.0	0.8	2.5	2.5	0.0	2.5
$ PE \geq 2.0$	0.0	0.0	2.5	0.8	0.0	0.8	1.7	3.3	0.0	0.0

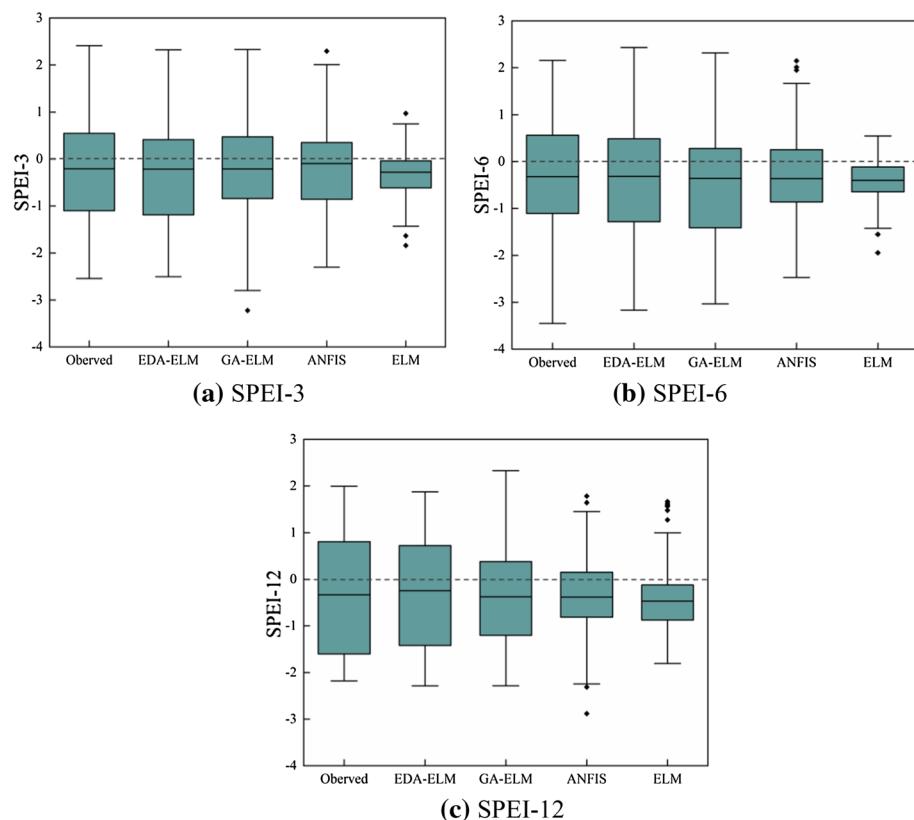


Fig. 7 Boxplot of the predicted SPEI values by EDA-ELM, GA-ELM, ANFIS, and ELM and observed ones during the testing period of 2007–2016

the higher the failure frequency of drought prediction. The EDA-ELM model recorded a significantly higher percentage with minor absolute errors than the ELM model. The frequency distribution of the absolute errors revealed that the EDA-ELM model performed best in drought prediction, followed by GA-ELM, ANFIS, and ELM models. This result also indicated that the improved ELM model using the optimization algorithm has potential advantages over the ANFIS model in drought prediction.

The spread of the observed and predicted multiscalar SPEI within the testing period is illustrated in Fig. 7 by using boxplots. In each boxplot, the outliers representing the extreme magnitude of the predicted multiscalar SPEI were denoted by “◆” and the smallest nonoutlier, lower quartile Q1, median Q2, upper quartile Q3, and the largest nonoutlier values were also demonstrated. The EDA-ELM model overestimated the smallest nonoutliers of SPEI-3 and SPEI-6, and the largest nonoutlier of SPEI-6, while it underestimated the largest nonoutlier of SPEI-3 and SEPI-12, and the smallest nonoutlier of SPEI-12. Nevertheless, the EDA-ELM model made their spreads more identical with the observed ones with a relatively closer match in lower quartile Q1, median Q2, and upper quartile Q3 than the other three models. The medians of the predicted SPEI by all models were close to the observed ones, but the difference between the observed lower/upper quartile statistics and predicted ones by GA-ELM, ANFIS, and ELM models enlarged, especially for SPEI-12.

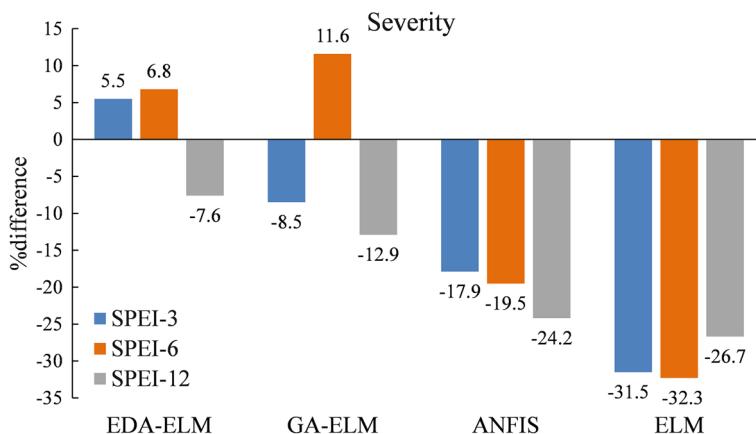


Fig. 8 Relative errors of drought severity monitored by multiscalar SPEI for the EDA-ELM, GA-ELM, ANFIS, and ELM models during the testing period of 2007–2016

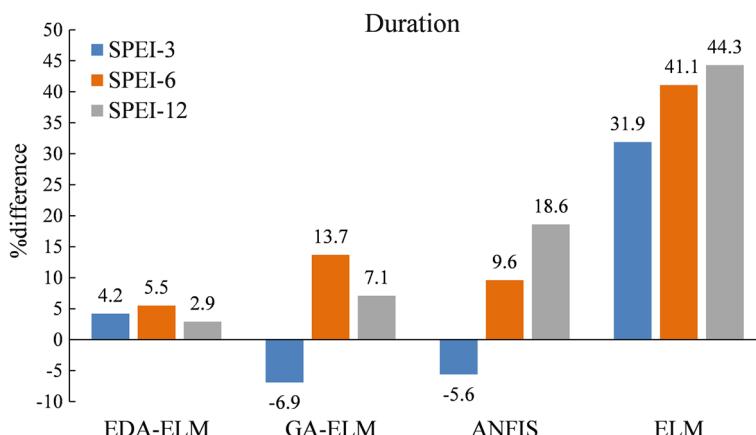


Fig. 9 Average relative errors in drought duration monitored by the multiscalar SPEI for the EDA-ELM, GA-ELM, ANFIS, and ELM models during the testing period of 2007–2016

This finding implied that the EDA-ELM model performed better than GA-ELM, ANFIS, and ELM models in terms of the statistical values of the boxplot.

3.3 Relative errors of drought severity and duration predicted

The relative errors of drought severity and duration for the EDA-ELM, GA-ELM, ANFIS, and ELM models during the testing period of 2007–2016 are presented in Figs. 8 and 9, respectively. The EDA-ELM model demonstrated the best prediction for drought properties in drought severity and duration. For SPEI-3, the EDA-ELM model overestimated drought severity and duration by 5.5% and 4.2%, and those values for SPEI-6 changed to 6.8% and 5.5%, respectively. However, for SPEI-12, it underestimated drought severity by 7.6% and

overestimated duration by 2.9%. In contrast with EDA-ELM, the ELM model poorest performed in drought severity and duration prediction with SPEI-3, SPEI-6, and SPEI-12 having the relative errors in drought severity being –31.5%, –32.3%, and –26.7% and the ones in drought duration being 31.9%, 41.1%, and 44.3%, respectively. The GA-ELM and ANFIS models performed better than the ELM model but poorer than the EDA-ELM model. Overall, the EDA-ELM model showed robust ability in capturing drought properties.

3.4 Execution time of different models

Computational time is a key factor for the selection of prediction models in addition to the prediction accuracy. The operating time of the four models for drought prediction is presented in Table 5. For all multiscalar SPEIs, the EDA-ELM model consumed the longest computational time (698.50–744.13 s), followed by GA-ELM (503.27–528.26 s), ANFIS (198.78–215.96 s), and ELM (100.18–117.68 s). For the same timescale SPEI, EDA-ELM and GA-ELM showed longer execution time than ANFIS and ELM models, and the EDA-ELM model consumed approximately seven times as much time as the ELM model did. Since the longest execution time of the EDA-ELM model is around 700 s, it can be concluded that the EDA-ELM model with much higher prediction accuracy than the other three models is absolutely acceptable.

3.5 Evolution of the prediction errors induced by different models with iterations

The evolution of the prediction errors in SPEI-12 for the EDA-ELM and GA-ELM models is shown in Fig. 10. After the ELM model was coupled with different evolutionary algorithms, the prediction errors decreased with evolution iterations. Figure 10 shows that the convergence rate of the EDA-ELM model was significantly slower than that of the GA-ELM model during the early stage of evolution. However, once the EDA-ELM model converged, its prediction errors were smaller than those of GA-ELM. The result further verified the conclusion drawn in Sect. 3.1.

Table 5 Total execution time of different models during the model training and testing periods

Drought indices	Predictive model	Time (s)
SPEI-3	EDA-ELM	722.16
	GA-ELM	503.27
	ANFIS	209.33
	ELM	109.31
SPEI-6	EDA-ELM	698.50
	GA-ELM	512.20
	ANFIS	198.78
	ELM	100.18
SPEI-12	EDA-ELM	744.13
	GA-ELM	528.26
	ANFIS	215.96
	ELM	117.68

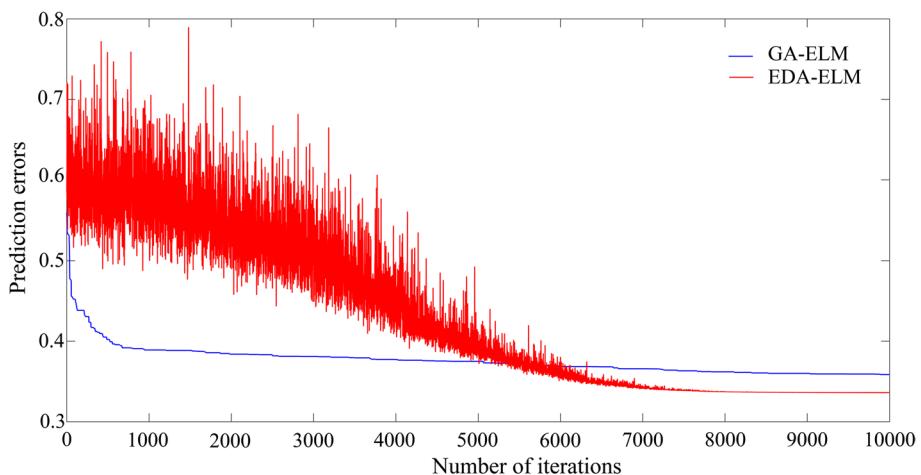


Fig. 10 Evolution of the prediction errors by the EDA-ELM and GA-ELM models for SPEI-12

4 Discussion

4.1 Limits of ELM in drought prediction

As one of the state-of-the-art models, the ELM model, which randomly assigns the weights and biases of hidden layers, has been widely applied in hydrological prediction. (Taormina and Chau 2015; Kourgialas et al. 2015; Lima et al. 2017; Prasad et al. 2018; Yaseen et al. 2019). However, when facing a complex relationship between predictors and predictand, the random assignment of the ELM model is prone to bring non-optimal parameters, increase redundant hidden nodes, and result in a degree of over-fitting or under-fitting (Miche et al. 2010; Cao et al. 2012; Huang et al. 2015). A standalone ELM model incurs challenges in handling seasonality and non-stationarity of climate-based inputs, and complex hydrological processes further exacerbate the challenges (Nahvi et al. 2016; Feng et al. 2021 2022; Han et al. 2021). For example, Barzegar et al. (2019) executed the ELM model 1000 times and averaged the results to reduce the effect of randomization when predicting river ice thickness with water flow, snow depth, and mean air temperature. This case proved that the ELM model, which does not optimize its neurons, weights, and thresholds, is sometimes inefficient in dealing with environmental problems. As non-stationarities and complex nonlinear relationships exist between primary drought driving factors and drought occurrence (Guo et al. 2016; Dariane and Azimi 2018), the standalone ELM algorithm performed poorest in drought prediction of Yunnan–Guizhou Plateau.

4.2 Impacts of the ELM optimization algorithms on the drought prediction performance

According to Lima et al. (2015, 2016), the ELM model with appropriate neurons, weights, and biases is critical to improving efficiency and accuracy. Meanwhile, the range of the input weights and hidden biases in ELM can be optimized as a parameter to achieve the

best prediction results (Parviainen and Riihimäki 2013; Lima et al. 2015). The optimization process can be considered an evolution of input weights and hidden layer biases toward an optimal set defined by a fitness function. Since evolutionary algorithms with global search procedures can effectively explore the most appropriate parameter set that approaches global optimum (Ding et al. 2013), the optimized ELM model by EDA algorithm (EDA-ELM) provides a novel approach for drought prediction.

EDA-ELM and GA-ELM models performed much better in drought prediction and the depiction of drought characteristics than the standalone ELM model. The success of the EDA-ELM and GA-ELM models can be attributed to the automatic and proper optimization of the ELM input weight and hidden biases, thereby avoiding the underfitting or overfitting problem caused by the random parameter assignment process of the ELM model. This finding is consistent with the results achieved by previous studies (Wu et al. 2017; Feng et al. 2021; Han et al. 2021).

Although both EDA-ELM and GA-ELM models originated from the ELM model, the GA-ELM model performed poorly in drought prediction than the EDA-ELM model. This result is mainly due to different evolutionary mechanisms being adopted. GA algorithm relies on the crossover and mutation of individuals to implicitly describe the probability distribution of potential solutions, but it cannot effectively maintain the valuable individuals representing optimal solutions. Different from the GA algorithm, the EDA algorithm constructs a probabilistic space to explicitly describe the distribution of the candidate solutions and preserve valuable individuals (optimal solutions) during the evolutionary process (Faraji Amiri and Behnamian 2020). As illustrated in Fig. 10, the EDA-ELM model searches the optimal global solution on a broader area (reflected by the fluctuations in the figure) than the GA-ELM model, and therefore, the GA-ELM model exhibited a faster convergence speed than the EDA-ELM model, but unfortunately, the optimal solution of the GA-ELM model eventually converged to a local one (Shim et al. 2011; Pang et al. 2019).

In addition, some studies indicated that the learning speed of models could be improved after optimization (Yaseen et al. 2019), while others obtained opposite conclusions (Deo et al. 2017; Prasad et al. 2018). In this study, the computational time of the EDA-ELM model was significantly longer than those of the ELM and ANFIS models for all multiscalar SPEIs. This is because the EDA evolutionary algorithm requires extensive iterative processes. However, the much higher prediction accuracy achieved by the EDA-ELM model than by the other three models made its running time (approximately 700 s) acceptable. Fortunately, the running time of the EDA-ELM model could be shortened by filtering invalid information through the pre-processes of the model inputs (Li and Cheng 2014; Nourani et al. 2017; Barzegar et al. 2019).

4.3 Influence of the SPEI time scales on the drought prediction performance

Several studies have involved the influence of the drought indices' time scales on the performance of drought prediction by machine learning methods, but different researchers presented different conclusions due to differences in drought characteristics of the study areas (Arikan et al. 2017; Mokhtarzad et al. 2017; Mouatadid and Adamowski 2017; Agana and Homaifar 2018). Taking the Huaihe River basin in eastern China as an example, where short-period droughts dominate, Zhu et al. (2016) disclosed that SPEI and SPI on shorter time scales generally performed better than those on longer time scales in drought detection. A similar conclusion was drawn that the smaller time scale index is better in

identifying the dynamic evolution of seasonal drought while the longer time scale index is better in judging the evolution of multi-season consecutive drought (Yu et al. 2016; Yildirim and Rahman 2021). Similar to the Huaihe River basin, Yunnan–Guizhou Plateau is also a region where most droughts are short-period (1–3 months) ones (Li et al. 2019), and its seasonal characteristics of droughts are prominent with summer, autumn, and winter showing apparent drying trends during the past decades (Cheng et al. 2020 2018; Yan et al. 2017). Moreover, the influence on droughts induced by large-scale atmospheric circulation also varies with seasons in the Yunnan–Guizhou Plateau (Cheng et al. 2020): severe and extreme spring and autumn droughts mainly occur in the El Niño years; severe and extreme summer droughts in both El Niño and La Niña years; severe and extreme winter droughts in La Niña years. The dominance of short-period droughts and the evident seasonal characteristics of droughts in the Yunnan–Guizhou Plateau imply that the SPEI-3, which represents 3-month accumulation values of SPEI, should be more appropriate for describing the dynamic evolution of seasonal droughts in the Yunnan–Guizhou Plateau than SPEI-6 and SPEI-12. This coincided with our findings that SPEI-3 performed better in drought prediction of Yunnan–Guizhou Plateau than SPEI-6 and SPEI-12. A similar conclusion was also drawn by Lu et al. (2016) that SPEI on shorter time scales (1 and 3 months) well identified once-in-a-century severe seasonal drought events from 2009 to 2010 with those on longer time scales (≥ 6 months) failing to do so.

5 Conclusion

In this study, a novel EDA-ELM model based on the capabilities of the EDA algorithm and ELM modeling technique was proposed. The advantages of the model in forecasting the multiscalar monthly SPEI in Yunnan–Guizhou Plateau, China, over GA-ELM, standard ELM, and ANFIS models were evaluated by taking the primary drought driving factors as input variables. The results revealed that the EDA-ELM model achieved the best prediction accuracy in drought prediction performance indicators, followed by GA-ELM, ANFIS, and ELM models. The execution times of the models descended in the order of EDA-ELM, GA-ELM, ANFIS, and standard ELM models, but they are all acceptable. The proposed EDA-ELM model provides a promising method for hydrological systems to predict droughts and a non-engineering measure for rational water resource allocation and drought mitigation to reduce agricultural and economic losses.

Author contributions Yao Du and Zhennan Liu involved in conceptualization, methodology, software. Qiongfang Li took part in data curation, writing—original draft preparation. Zhengmo Zhou and Guobin Lu involved in writing—reviewing and editing. Qihui Chen took part in software, validation.

Funding Financial support is gratefully acknowledged from the National Natural Science Foundation Commission of China under Grant numbers 51879069 and 41961134003, Guizhou Science Foundation-ZK [2021] General 295 and the Jiangsu Provincial Collaborative Innovation Center of World Water Valley and Water ecological civilization, China.

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Agana NA, Homaifar A (2018) EMID-based predictive deep belief network for time series prediction: an application to drought forecasting. *Hydrology* 5:18. <https://doi.org/10.3390/hydrology5010018>
- Alencar ASC, Rocha Neto AR, Gomes JPP (2016) A new pruning method for extreme learning machines via genetic algorithms. *Appl Soft Comput J.* <https://doi.org/10.1016/j.asoc.2016.03.019>
- Ali M, Deo RC, Downs NJ, Maraseni T (2018) An ensemble-ANFIS based uncertainty assessment model for forecasting multi-scalar standardized precipitation index. *Atmos Res.* <https://doi.org/10.1016/j.atmosres.2018.02.024>
- Arikan BB, Universitesi IM, Kahya E (2017) Drought forecasting using genetic programming in association with southern oscillation indices DROUGHT FORECASTING USING GENETIC PROGRAMMING IN 4–9.
- Baluja S (1994) Population-based incremental learning: A method for integrating genetic search based function optimization and competitive learning. *Tech Rep C Sch Comput Sci, Carnegie Mellon Univ, Pittsburgh, PA*
- Barzegar R, Ghasri M, Qi Z et al (2019) Using bootstrap ELM and LSSVM models to estimate river ice thickness in the Mackenzie River Basin in the Northwest Territories, Canada. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2019.06.075>
- Bui DT, Ngo PTT, Pham TD et al (2019) A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *CATENA.* <https://doi.org/10.1016/j.catena.2019.04.009>
- Cao J, Lin Z, Bin HG, Liu N (2012) Voting based extreme learning machine. *Inf Sci (ny).* <https://doi.org/10.1016/j.ins.2011.09.015>
- Chen Y, Sun X, Gong D et al (2017) Personalized search inspired fast interactive estimation of distribution algorithm and its application. *IEEE Trans Evol Comput.* <https://doi.org/10.1109/TEVC.2017.2657787>
- Cheng Q, Gao L, Chen Y et al (2018) Temporal-spatial characteristics of Drought in Guizhou Province China based on multiple drought indices and historical disaster records. *Adv Meteorol.* <https://doi.org/10.1155/2018/4721269>
- Cheng Q, Gao L, Zhong F et al (2020) Spatiotemporal variations of drought in the Yunnan–Guizhou Plateau, southwest China, during 1960–2013 and their association with large-scale circulations and historical records. *Ecol Indic.* <https://doi.org/10.1016/j.ecolind.2019.106041>
- Dariane AB, Azimi S (2018) Streamflow forecasting by combining neural networks and fuzzy models using advanced methods of input variable selection. *J Hydroinformatics.* <https://doi.org/10.2166/hydro.2017.076>
- Dehnavi A, Aghdam IN, Pradhan B, Morshed Varzandeh MH (2015) A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. *CATENA.* <https://doi.org/10.1016/j.catena.2015.07.020>
- Deo RC, Şahin M (2015) Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmos Res.* <https://doi.org/10.1016/j.atmosres.2014.10.016>
- Deo RC, Tiwari MK, Adamowski JF, Quilty JM (2017) Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch Environ Res Risk Assess.* <https://doi.org/10.1007/s00477-016-1265-z>
- Ding S, Li H, Su C, Yu J, Jin F (2013) Evolutionary artificial neural networks: a review. *Artif Intell Rev* 39:251–260. <https://doi.org/10.1007/s10462-011-9270-6>
- Faraji Amiri M, Behnamian J (2020) Multi-objective green flowshop scheduling problem under uncertainty: estimation of distribution algorithm. *J Clean Prod* 251:119734. <https://doi.org/10.1016/j.jclepro.2019.119734>
- Farokhnia A, Morid S, Byun HR (2011) Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. *Theor Appl Climatol.* <https://doi.org/10.1007/s00704-010-0317-4>
- Feng ZK, Niu WJ, Zhang R et al (2019) Operation rule derivation of hydropower reservoir by k-means clustering method and extreme learning machine based on particle swarm optimization. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2019.06.045>
- Feng Z, Niu W, Tang Z et al (2021) Evolutionary artificial intelligence model via cooperation search algorithm and extreme learning machine for multiple scales nonstationary hydrological time series prediction. *J Hydrol* 595:126062. <https://doi.org/10.1016/j.jhydrol.2021.126062>

- Feng Z, Shi P, Yang T et al (2022) Parallel cooperation search algorithm and artificial intelligence method for streamflow time series forecasting. *J Hydrol* 606:127434. <https://doi.org/10.1016/j.jhydrol.2022.127434>
- Forootan E, Khaki M, Schumacher M et al (2019) Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections. *Sci Total Environ.* <https://doi.org/10.1016/j.scitotenv.2018.09.231>
- Gao S, de Silva CW (2018) Estimation distribution algorithms on constrained optimization problems. *Appl Math Comput.* <https://doi.org/10.1016/j.amc.2018.07.037>
- Guo W, Xu T, Lu Z (2016) An integrated chaotic time series prediction model based on efficient extreme learning machine and differential evolution. *Neural Comput Appl.* <https://doi.org/10.1007/s00521-015-1903-2>
- Han S, Zhu K, Wang R (2021) Improvement of evolution process of dandelion algorithm with extreme learning machine for global optimization problems. *Expert Syst Appl* 163:113803. <https://doi.org/10.1016/j.eswa.2020.113803>
- Hauschild M, Pelikan M (2011) An introduction and survey of estimation of distribution algorithms. *Swarm Evol Comput.* <https://doi.org/10.1016/j.swevo.2011.08.003>
- Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. *Neurocomputing.* <https://doi.org/10.1016/j.neucom.2005.12.126>
- Huang GB, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst Man Cybern Part B Cybern.* <https://doi.org/10.1109/TSMCB.2011.2168604>
- Huang G, Bin HG, Song S, You K (2015) Trends in extreme learning machines: a review. *Neural Netw.* <https://doi.org/10.1016/j.neunet.2014.10.001>
- Jaafarai A, Zenner EK, Panahi M, Shahabi H (2019) Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agric Meteorol.* <https://doi.org/10.1016/j.agrmet.2018.12.015>
- Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23:665–685. <https://doi.org/10.1109/21.256541>
- Khan MZK, Rahman A, Rahman MA, Renzaho AMN (2021) Impact of droughts on child mortality: a case study in Southern African countries. *Nat Hazards* 108:2211–2224. <https://doi.org/10.1007/s11069-021-04776-9>
- Knudby A, Brenning A, LeDrew E (2010) New approaches to modelling fish-habitat relationships. *Ecol Model.* <https://doi.org/10.1016/j.ecolmodel.2009.11.008>
- Kourgialas NN, Dokou Z, Karatzas GP (2015) Statistical analysis and ANN modeling for predicting hydrological extremes under climate change scenarios: the example of a small Mediterranean agro-watershed. *J Environ Manage.* <https://doi.org/10.1016/j.jenvman.2015.02.034>
- Larrañaga P, Lozano JA (2002) Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation
- Li BH, Zhao Q (2012) Communications in computer and information science: preface. *Commun Comput Inf Sci.* <https://doi.org/10.1007/978-3-642-34387-2>
- Li BJ, Cheng CT (2014) Monthly discharge forecasting using wavelet neural networks with extreme learning machine. *Sci China Technol Sci.* <https://doi.org/10.1007/s11431-014-5712-0>
- Li J, Aickelin U (2007) BOA for nurse scheduling. *Stud Comput Intell.* https://doi.org/10.1007/978-3-540-34954-9_14
- Li Q, Han X, Liu Z et al (2022) A novel information changing rate and conditional mutual information-based input feature selection method for artificial intelligence drought prediction models. *Clim Dyn.* <https://doi.org/10.1007/s00382-021-06104-0>
- Li Q, He P, He Y et al (2020) Investigation to the relation between meteorological drought and hydrological drought in the upper Shaying River Basin using wavelet analysis. *Atmos Res.* <https://doi.org/10.1016/j.atmosres.2019.104743>
- Li Y, Wang Z, Zhang Y et al (2019) Drought variability at various timescales over Yunnan Province, China: 1961–2015. *Theor Appl Climatol.* <https://doi.org/10.1007/s00704-019-02859-z>
- Liang Y, Ren Z, He M et al (2019) An efficient estimation of distribution algorithm with rank-one modification and population reduction. *BioSystems.* <https://doi.org/10.1016/j.biosystems.2019.04.001>
- Lima AR, Cannon AJ, Hsieh WW (2016) Forecasting daily streamflow using online sequential extreme learning machines. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2016.03.017>
- Lima AR, Cannon AJ, Hsieh WW (2015) Nonlinear regression in environmental sciences using extreme learning machines: a comparative evaluation. *Environ Model Softw.* <https://doi.org/10.1016/j.envsoft.2015.08.002>

- Lima AR, Hsieh WW, Cannon AJ (2017) Variable complexity online sequential extreme learning machine, with applications to streamflow prediction. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2017.10.037>
- Liu ZN, Li QF, Nguyen LB, Xu GH (2018) Comparing machine-learning models for drought forecasting in vietnam's cai river basin. *Polish J Environ Stud.* <https://doi.org/10.15244/pjoes/80866>
- Lu, JL., Xu XH., W GP., Y MX., L XL., 2016. Development and application of the short-term/long-term composite drought index in the Southwestern China. *J Disaster Prev Mitig Eng* 1672–2132(2016)36:4<681:XNDQDQ>2.0.TX;2–8 (in Chinese)
- Luo J, Qi Y, Xie J, Zhang X (2015) A hybrid multi-objective PSO-EDA algorithm for reservoir flood control operation. *Appl Soft Comput J.* <https://doi.org/10.1016/j.asoc.2015.05.036>
- Ma B, Zhang B, Jia L, Huang H (2020) Conditional distribution selection for SPEI-daily and its revealed meteorological drought characteristics in China from 1961 to 2017. *Atmos Res.* <https://doi.org/10.1016/j.atmosres.2020.105108>
- Miche Y, Sorjamaa A, Bas P et al (2010) OP-ELM: Optimally pruned extreme learning machine. *IEEE Trans Neural Networks.* <https://doi.org/10.1109/TNN.2009.2036259>
- Mokhtarzad M, Eskandari F, Jamshidi Vanjani N, Arabasadi A (2017) Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. *Environ Earth Sci.* <https://doi.org/10.1007/s12665-017-7064-0>
- Mouatadid S, Adamowski J (2017) Using extreme learning machines for short-term urban water demand forecasting. *Urban Water J.* <https://doi.org/10.1080/1573062X.2016.1236133>
- Mouatadid S, Raj N, Deo RC, Adamowski JF (2018) Input selection and data-driven model performance optimization to predict the Standardized Precipitation and Evaporation Index in a drought-prone region. *Atmos Res.* <https://doi.org/10.1016/j.atmosres.2018.05.012>
- Mühlenbein H, Bendasch J, Voigt HM (1996) From recombination of genes to the estimation of distributions II. Continuous parameters. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp 188–197
- Nahvi B, Habibi J, Mohammadi K, Shamshirband S (2016) Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Comput Electron Agric* 124:150–160. <https://doi.org/10.1016/j.compag.2016.03.025>
- Nourani V, Andalib G, Sadikoglu F (2017) Multi-station streamflow forecasting using wavelet denoising and artificial intelligence models. In: Procedia Computer Science. <https://doi.org/10.1016/j.procs.2017.11.287>
- Orimoloye IR, Oloade OO, Belle JA (2021) Satellite-based application in drought disaster assessment using terra MOD13Q1 data across free state province, South Africa. *J Environ Manage.* <https://doi.org/10.1016/j.jenvman.2021.112112>
- Pang S, Li W, He H et al (2019) An EDA-GA hybrid algorithm for multi-objective task scheduling in cloud computing. *IEEE Access.* <https://doi.org/10.1109/ACCESS.2019.2946216>
- Parviainen E, Riihimäki J (2013) A Connection between Extreme Learning Machine and Neural Network Kernel. In: Communications in Computer and Information Science. pp 122–135
- Prasad R, Deo RC, Li Y, Maraseni T (2018) Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. *Geoderma.* <https://doi.org/10.1016/j.geoderma.2018.05.035>
- Rahmati O, Panahi M, Kalantari Z et al (2020) Capability and robustness of novel hybridized models used for drought hazard modeling in southeast Queensland, Australia. *Sci Total Environ.* <https://doi.org/10.1016/j.scitotenv.2019.134656>
- Shim VA, Tan KC, Chia JY, Chong JK (2011) Evolutionary algorithms for solving multi-objective travelling salesman problem. *Flex Serv Manuf J.* <https://doi.org/10.1007/s10696-011-9099-y>
- Spinoni J, Vogt JV, Naumann G et al (2018) Will drought events become more frequent and severe in Europe? *Int J Climatol.* <https://doi.org/10.1002/joc.5291>
- Su B, Huang J, Fischer T et al (2018) Drought losses in China might double between the 1.5 and 2.0 °C warming. *Proc Natl Acad Sci U S A.* <https://doi.org/10.1073/pnas.1802129115>
- Sun S, Li Q, Li J et al (2019) Revisiting the evolution of the 2009–2011 meteorological drought over Southwest China. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2018.10.071>
- Taormina R, Chau KW (2015) Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2015.08.022>
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim.* <https://doi.org/10.1175/2009JCLI2909.1>
- Wang L, Chen W (2014) A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. *Int J Climatol.* <https://doi.org/10.1002/joc.3822>

- Wang L, Chen W, Zhou W (2014) Assessment of future drought in Southwest China based on CMIP5 multimodel projections. *Adv Atmos Sci.* <https://doi.org/10.1007/s00376-014-3223-3>
- Wang L, Chen W, Zhou W, Huang G (2015) Teleconnected influence of tropical Northwest Pacific sea surface temperature on interannual variability of autumn precipitation in Southwest China. *Clim Dyn.* <https://doi.org/10.1007/s00382-015-2490-8>
- Wang P, Wu X, Hao Y et al (2020a) Is Southwest China drying or wetting? Spatiotemporal patterns and potential causes. *Theor Appl Climatol.* <https://doi.org/10.1007/s00704-019-02935-4>
- Wang X, Han T, Zhao H (2020b) An estimation of distribution algorithm with multi-leader search. *IEEE Access.* <https://doi.org/10.1109/ACCESS.2020.2975468>
- Wu CL, Chau KW (2011) Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2011.01.017>
- Wu L, Huang G, Fan J et al (2020) Hybrid extreme learning machine with meta-heuristic algorithms for monthly pan evaporation prediction. *Comput Electron Agric.* <https://doi.org/10.1016/j.compag.2019.105115>
- Wu T, Yao M, Yang J (2017) Dolphin swarm extreme learning machine. *Cognit Comput* 9:275–284. <https://doi.org/10.1007/s12559-017-9451-y>
- Xu K, Yang D, Yang H et al (2015) Spatio-temporal variation of drought in China during 1961–2012: a climatic perspective. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2014.09.047>
- Yan Z, Zhang Y, Zhou Z, Han N (2017) The spatio-temporal variability of droughts using the standardized precipitation index in Yunnan, China. *Nat Hazards.* <https://doi.org/10.1007/s11069-017-2904-5>
- Yang H, Yi J, Zhao J, Dong ZY (2013) Extreme learning machine based genetic algorithm and its application in power system economic dispatch. *Neurocomputing* 102:154–162. <https://doi.org/10.1016/j.neucom.2011.12.054>
- Yao N, Li L, Feng P et al (2020) Projections of drought characteristics in China based on a standardized precipitation and evapotranspiration index and multiple GCMs. *Sci Total Environ.* <https://doi.org/10.1016/j.scitotenv.2019.135245>
- Yaseen ZM, Sulaiman SO, Deo RC, Chau K-W (2019) An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *J Hydrol.* 569:387–408. <https://doi.org/10.1016/j.jhydrol.2018.11.069>
- Yildirim G, Rahman A (2021) Spatiotemporal meteorological drought assessment: a case study in south-east Australia. *Nat Hazards.* <https://doi.org/10.1007/s11069-021-05055-3>
- Yildirim G, Rahman A, Singh VP (2022) A Bibliometric analysis of drought indices, risk, and forecast as components of drought early warning systems. *Water (switzerland).* <https://doi.org/10.3390/w14020253>
- Yu M, Li Q, Hayes MJ et al (2014) Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010? *Int J Climatol.* <https://doi.org/10.1002/joc.3701>
- Yu M, Liu X, Wei L et al (2016) Drought assessment by a short-/long-term composited drought index in the upper huaihe river basin, China. *Adv Meteorol.* <https://doi.org/10.1155/2016/7986568>
- Zeng Z, Wu W, Li Z et al (2019) Agricultural drought risk assessment in Southwest China. *Water (switzerland).* <https://doi.org/10.3390/w11051064>
- Zhang L, Zhou T (2015) Drought over East Asia: a review. *J Clim.* <https://doi.org/10.1175/JCLI-D-14-00259.1>
- Zhang Q, Gu X, Singh VP et al (2015) Spatiotemporal behavior of floods and droughts and their impacts on agriculture in China. *Glob Planet Change.* <https://doi.org/10.1016/j.gloplacha.2015.05.007>
- Zhang Y, You Q, Chen C, Li X (2017) Flash droughts in a typical humid and subtropical basin: a case study in the Gan River Basin, China. *J Hydrol.* <https://doi.org/10.1016/j.jhydrol.2017.05.044>
- Zhao J, Zhang Q, Zhu X et al (2020) Drought risk assessment in China: evaluation framework and influencing factors. *Geogr Sustain.* <https://doi.org/10.1016/j.geosus.2020.06.005>
- Zhu B, Feng Y, Gong D et al (2020) Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. *Comput Electron Agric.* <https://doi.org/10.1016/j.compag.2020.105430>
- Zhu Q-Y, Qin AK, Suganthan PN, Huang G-B (2005) Evolutionary extreme learning machine. *Pattern Recognit* 38:1759–1763. <https://doi.org/10.1016/j.patcog.2005.03.028>
- Zhu Y, Wang W, Singh VP, Liu Y (2016) Combined use of meteorological drought indices at multi-time scales for improving hydrological drought detection. *Sci Total Environ* 571:1058–1068. <https://doi.org/10.1016/j.scitotenv.2016.07.096>