Ecological-environmental quality estimation using remote sensing and combined artificial intelligence techniques

Vahid Nourani, Ehsan Foroumandi, Elnaz Sharghi and Dominika Dąbrowska

ABSTRACT

Ecological-environmental quality was evaluated for Tabriz and Rasht cities (in Iran) with different climate conditions using artificial intelligence (AI) and remote sensing (RS) techniques. Sampling sites were surveyed and ecological experts assigned eco-environment background values (EBVs) of sites. Then, eco-environmental attributes were extracted as RS derived, and meteorological attributes were observed. Three AI-based models, artificial neural network (ANN), support vector regression (SVR), and adaptive neuro-fuzzy inference system (ANFIS) were then applied to learn the relationship between a target set of known EBVs and eco-environmental attributes as inputs. According to the results of the single models, none of the models could evaluate EBV appropriately for all regions and classes. Thereafter, three combining techniques were applied to the outputs of single models to enhance spatial evaluation of EBV. It was observed that the modeling for Tabriz led to more accurate results. It seems that the better network performance for Tabriz may be due to a more heterogeneous dataset in this kind of climate. Furthermore, results indicated that SVR led to better performance than both ANN and ANFIS models, but the models' combining techniques were shown to be superior. Combining techniques enhanced performance of single AI modeling up to 26% in the verification step.

Key words | artificial intelligence, combining technique, ecological-environmental quality, Rasht, remote sensing, Tabriz

Vahid Nourani (corresponding author)
Ehsan Foroumandi
Elnaz Sharghi
Center of Excellence in Hydroinformatics, Faculty of Civil Engineering,
University of Tabriz,
Tabriz,
Iran
E-mail: *nourani@tabrizu.ac.ir*

Vahid Nourani
Faculty of Civil and Environmental Engineering,
Near East University,
Near East Boulevard, 99138, via Mersin 10,
Nicosia,
Cyprus

Dominika Dąbrowska
Faculty of Earth Sciences,
University of Silesia,
Bedzinska 60, 41-200 Sosnowiec,
Poland

HIGHLIGHTS

- EEQ was estimated for regions with different kinds of climate by AI methods.
- Ensemble techniques were used to improve the modeling performance.
- Modeling for semi-arid region led to accurate results regrade to humid region.
- Non-linear ensemble improved performance of single models up to 26%.
- Modeling for spring season led to more accurate results.

INTRODUCTION

Human beings are required to fulfill their economic and productive activities by increasing the population, performed mainly in the realm of the environment, and are generally detrimental to environmental mechanisms. These types of activities increase the pollution of air, water, and soil, reduce the amount of available water, and destroy environmental resources. The assessment of Ecological-Environment Quality (EEQ) index is a way of managing, planning, and decision-making in environmental studies. Investigation of EEQ helps identify, predict, and evaluate the environmental impacts of human and non-human activities (Niu & Harris 1996). With recent advances in

doi: 10.2166/hydro.2020.048

geographic sciences such as geographic information system (GIS) and remote sensing (RS), the methods for estimating environmental indicators (EIs) have progressed significantly (Nichol & Wong 2018). Given the importance of evaluating environmental conditions, several studies have used EEQ indicators to assess environmental conditions. Shi & Li (2007) utilized back propagation (BP) artificial neural network (ANN) and three types of environmental data to estimate the EEQ index. They used RS-based attributes and meteorological data for this purpose. Ying *et al.* (2007) estimated EEQ based on the combination of nature, eco-environment, disasters, environmental pollution, and socio-economic factors in Henan province of China. Other researchers have also assessed EEQ using different EIs and ways for diverse climate and land conditions (e.g., see Sarkar *et al.* 2007; Miao *et al.* 2016; Chai & Lha 2018). In this way, physical-based, black-box, and conceptual models are different ways of estimating the nonlinear structural behavior of the environmental processes (Nourani 2017). Although conceptual and physical-based models can be used for investigation of the actual physics of the phenomenon, some limitations are seen in their application. As physical perception has less importance than accurate predictions, black-box models' usage can be more beneficial. Large amounts of noisy data from nonlinear and dynamic systems can be handled with such black-box methods (Quej *et al.* 2017). Nowadays, the more relative black-box approaches known as artificial intelligence (AI) methods are widely used for assessment of environmental processes. In various aspects of environmental engineering, AI models such as adaptive neural fuzzy inference system (ANFIS), feed forward neural network (FFNN), and support vector regression (SVR) have been used. Recent literature reported the numerous applications of ANNs for monitoring and estimation of environmental aspects (e.g., see Dzeroski 2001; Moisen & Frescino 2002; Shi & Li 2007; Singh *et al.* 2009). On the other hand, SVR minimizes the structural risk that aims at minimizing a bound on the generalization error (Kecman 2001) and has been recently used to evaluate environmental attributes (e.g., see Yang *et al.* 2018; Zarei *et al.* 2018). The ANFIS model can potentially learn the basic relations of numerical data. Fuzzy systems are able to integrate information processing with the properties of mathematics (Setnes *et al.* 1998). Some studies applied ANFIS for estimating EIs (e.g., see Tabari *et al.* 2012; Awan & Bae 2016; Srivastava *et al.* 2016). There are a few papers that used various AI approaches for estimating hydrological and environmental attributes and to compare the efficiencies of the models (e.g., see Olyaie *et al.* 2015; Alizadeh *et al.* 2018). However, to the best of our knowledge, none of the presented studies have used SVR and ANFIS models for EEQ assessment.

There are various models for solving a specific issue, for which they provide different results with high reliability. Therefore, providing more accurate patterns for these issues can be achieved by integrating the results of different models using the combining techniques, as a post-processing method. The idea of combining techniques from single models was explored in statistics and econometrics more than 30 years ago (see Bates & Granger 1969; Dickinson 1973). Recently, this method has been used in hydraulic studies and time-series predictions (Zhang 2003; Sharghi *et al.* 2018; Shamshirband *et al.* 2019), but unfortunately, it feels as though this method is not used enough in environmental studies to estimate EEQs.

The present paper applies the concept of AI-based combining technique to estimate the EEQ index for two study areas in Iran. In this regard, first, feed-forward ANN, SVR, and ANFIS models are used for modeling the EEQ index of EBV. Then, combining techniques were designed using outputs of the mentioned single AI-based models to improve the overall performance of modeling. For each study area, three combining techniques, including simple linear averaging, weighted linear averaging, and nonlinear neural combining are used to combine the outputs of single AI-based models. In this way, the required data for training the single AI-based models were obtained by the RS tools and ground meteorological stations.

MATERIALS

Case studies

The case studies of this research included parts of East Azerbaijan Province, around Tabriz city (Figure 1(a)), located in northwest Iran, as well as areas in Gilan Province, around Rasht city (Figure 1(b)), located in the north of the

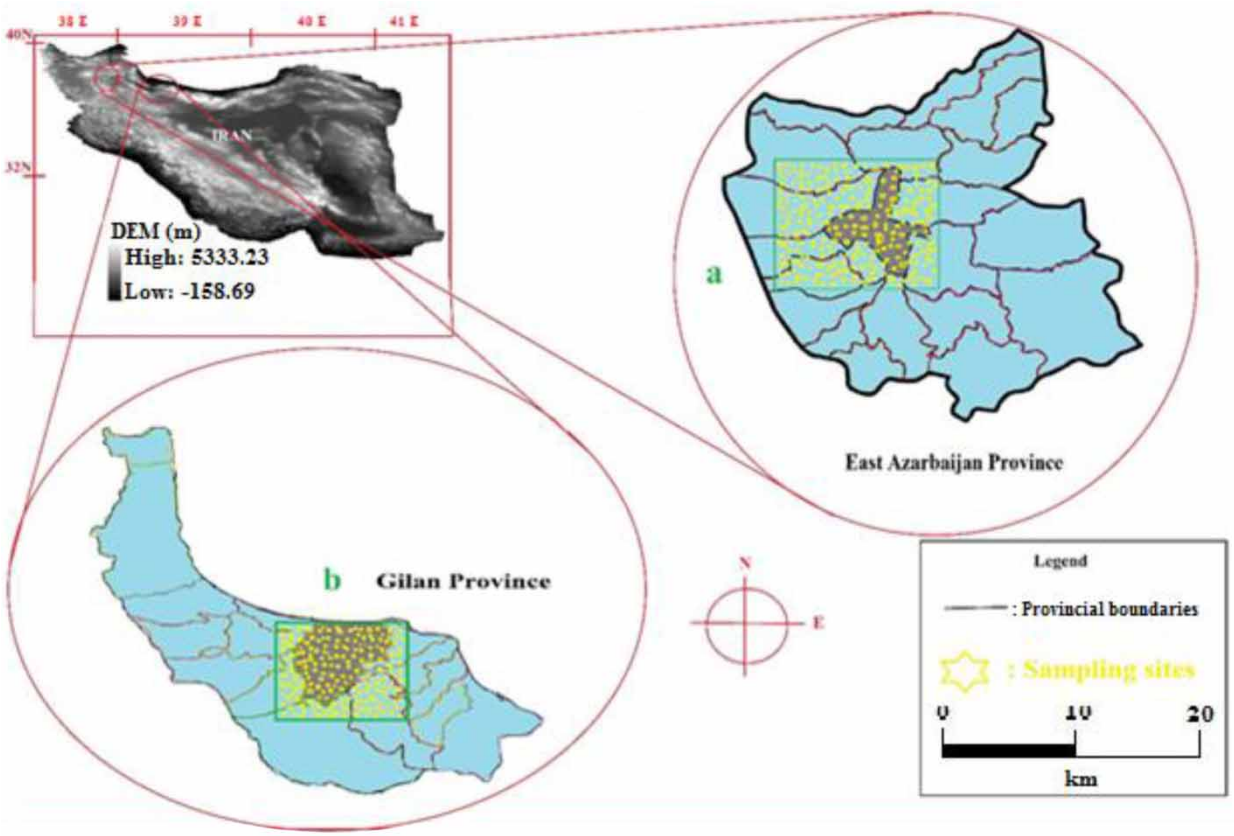

Figure 1 | Iran's DEM; the study areas, (a) case study in East Azerbaijan Province (Tabriz), (b) case study in Gilan Province (Rasht).

country. The study areas have different climates, ecological environment, and topography conditions.

East Azerbaijan Province is considered to be the largest and most populous province in northwestern Iran, with cold weather. The central city of East Azerbaijan Province is Tabriz metropolis. Tabriz is located at 46° 25′ east longitude and 38° 02′ north latitude and it is composed of mountainous and plain areas. The first part of the study was conducted on 300 sampling sites around Tabriz city. For each site, the used evaluation unit was a 300 by 300 m square. In most parts of this province, the vegetation cover is in the form of natural and small wild plants. Forest trees, natural pasture plants, and various forests and thickets also cover this region. East Azerbaijan has a semi-arid climate. The average annual rainfall is 284.8 mm in Tabriz. Tabriz has an average of 85 days of frost in a year (IRIMO (Iran Meteorological Organization) 2006). The second part of the study was conducted on 300 sampling sites, each with a 300 by 300 m square evaluation unit, around Rasht city in Gilan Province. Rasht is located at 49° 36′ east longitude and 37° 16′ north latitude. The relative humidity and the average rainfall in this city are 80% and 1,359 mm per year, respectively. Gilan Province is covered with forest and has a moderate and humid climate. Gilan consists of the western end of the Alborz mountain range and the western part of the plains around the Caspian Sea. The whole of Gilan Province is humid and green; the northern–southern dominant stream flows are wetted over the sea and are forced to go up by the strong barrier of the Alborz mountains, resulting in abundant rainfall on the plains and the northwest foothills of the mountain range. Gilan Province has high-density vegetation due to its distinct climate (IRIMO 2006).

Ground base data

Ground reference data about various eco-environmental types and zones were collected for evaluating EEQ by the

field investigation for all sites. The EEQ values of the studied areas were measured using the Eco-Environment Background Value (EBV) index, which is one of the most important EEQs (e.g., see Shi & Li 2007).

Meteorological requisite data (monthly precipitation and temperature time series) were obtained from 19 meteorological stations in East Azerbaijan Province and 13 stations in Gilan Province. All RS-based geospatial data were georeferenced and the Gauss–Krueger Projection was used with relevant geodetic datum and Krassovsky ellipsoid for each case study. A method of third-order polynomial was exerted to imagery geometrical correction (Zareie *et al.* 2016). In total, 30 recognizable ground control marks such as intersections of roads and/or railway were selected from a topographical map. A satisfactory mean squared error of 9.3 m was obtained after the corrections.

Remotely sensed data

The basic data used in this paper were RS-derived attributes and observed meteorological attributes. The RS dataset used in this study included: (i) 144 freely available scenes of Landsat8 Operational Land Imager (OLI) data which were received for every month for 2013 to 2018 separately (for deriving vegetation index, moisture index, and land surface temperature), (ii) two digital elevation model (DEM) images from ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) satellite (for elevation information) with an accuracy of 30 m (1 arc second).

METHODS

Proposed methodology

In this paper, EBVs of study areas were estimated using the three single AI-based models of SVR, FFNN, and ANFIS; then, combining techniques were used as post-processing methods. In this way, first, field survey and observations were done for 300 sampling sites for each study area by ecological experts and every sampling site was given an EBV according to Table 1. Results of the determined EBVs for the sites were then used as targets for the single AI-based models where the RS-based data (for deriving Normalized

Table 1 | The scoring system of the eco-environmental quality value (Feizizadeh & Blaschke 2013)

Rank	EBV range	Eco-environment type	Topographical condition	Vegetable cover/Landcover	Humidity and heat
I	450–500	High coverage ratio, broadleaf, conifer and broadleaf mixed eco-environment type	Moderate mountains (1,000–1,200 m)	Evergreen broadleaf forest; tree cover >90%	70–80% shady and cool
II	400–450	Moderate coverage ratio, conifer and broadleaf mixed eco-environment type	Low mountains (500–1,000 m)	Subtropical conifer and broadleaf, mixed forest, deciduous, broadleaf forest; tree cover is 80–90%	60–70% shady and cool
III	350–400	Low coverage ratio, conifer, eco-environment type	Low mountains (500–1,000 m)	Conifer forest: fir, horsetail forest	40% shady
IV	300–350	Shrub eco-environment type	Various hilly and mountainous lands	Evergreen broadleaf forest, shrub	50–60% warm
V	250–300	Meadow and brushwood, eco-environment type	Top of moderate mountains (> 1,200 m)	Meadow, bunch grasses (Gramineae)	30% cool
VI	200–250	Economic forest, eco-environment type	Costal low hills and uplands, valley	Orchard garden, tea garden, bamboo	30% warm
VII	150–200	Wetland eco-environment type	Plain farmland	Irrigation farming	50% dry and heat
VIII	100–150	Arid and barren land, eco-environment type	Peninsula, sand beach	Dry farming	20% dry and heat
IX	50–100	Artificial construction, eco-environment type	Plain	Residential area, buildings	20% dry and heat

Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), land surface temperature (LST), and elevation level (EL)) and meteorological observations (seasonal precipitation (SP) and seasonal temperature (ST)) were imposed on the model as inputs. Second, the combining techniques were created to combine the outputs of the single AI-based models as a post-processing method. Thereafter, the spatial distribution of EBV for all areas (of the case studies) was derived using the outputs of the combining techniques. In this regard, the proposed method in the current study includes four major steps (see Figure 2):

(1) Data collection: field investigating and EBV assigning for 300 sampling sites for each case study and acquisition of environmental attributes using the RS data and meteorological observations.
(2) Single AI-based models: training and testing three single models (ANN, SVR, ANFIS) for evaluation of EBV using available input data.
(3) Combining techniques as post-processing: forming three combining techniques via outputs of the single models.
(4) Determination of the spatial distribution of EBV over the whole region (for each case study).

Data collection

As the first stage, a total of 300 sampling sites for each study area were randomly selected and positioned on the base map of topography. For each site, the evaluation unit of

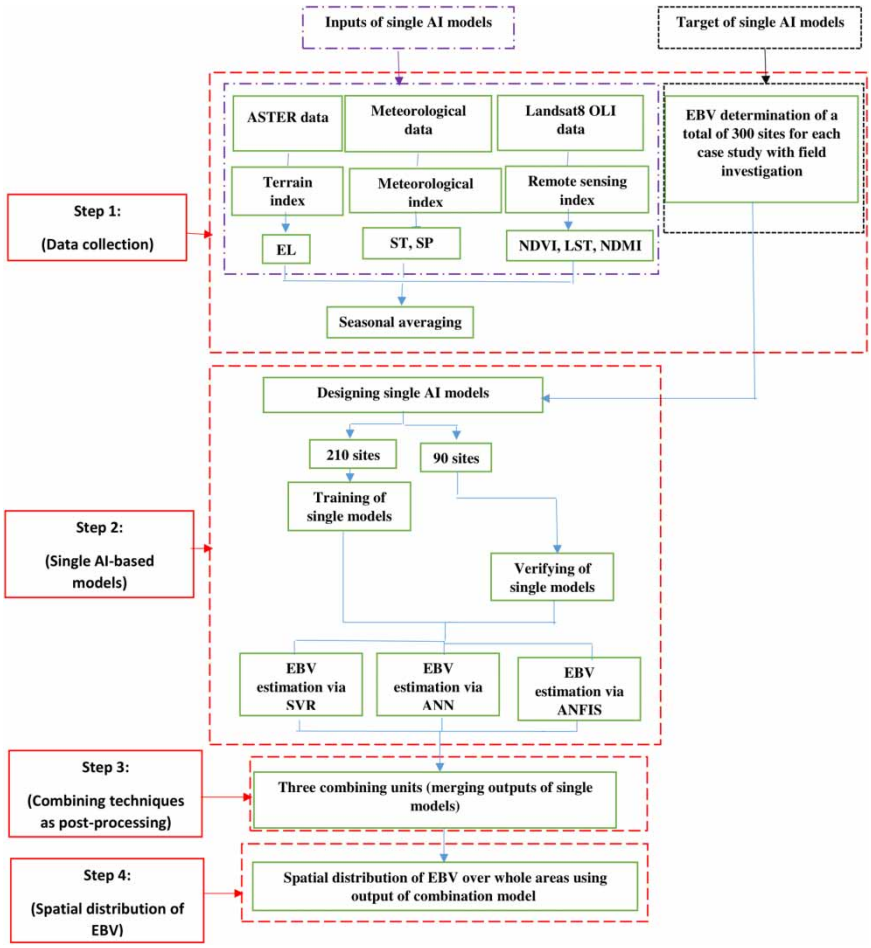

Figure 2 | Flowchart of the proposed methodology.

300 by 300 m square was used and the Global Positioning System (GPS) device was used to record the positions. These sampling sites were investigated *in situ* and EIs including soil types, topographical conditions, and vegetation details were documented. According to Table 1, EEQ rank and corresponding EBV were manually determined for each site by ecological experts. As shown in Table 1, there were nine classes of EEQ and their basic features about ecological environment conditions. A range of scores was given to each presented class, which is called EBV. Given the existing classifications, the higher the value of this index, the better the EEQ of the area.

It has been reported that the most important attributes for natural balance and stability of ecological environment are temperature, precipitation, vegetation detail, elevation, and their spatial and temporal patterns (Choi 2004). The RS tools were used to gather the spatial data required to evaluate the EBV in this study. The temporal climate dataset was provided using the meteorological station observations. First, environmental indices were derived from RS, including vegetation index (NDVI), moisture index (NDMI), and LST. RS-based indexes were extracted using images of Landsat8 (OLI) satellite mapper, as well as elevation data extracted from the DEM of ASTER mapper. For this purpose, 12 images (monthly) of Landsat8 (OLI) were retrieved for each year for the period of 2013–2018 and one image of ASTER for DEM. All 72 images of Landsat8 (OLI) for each study area were detrended to remove the dependency of models on the annual data. Also, monthly precipitation and monthly temperature time series were gathered for 2013–2018 by meteorological observations to calculate ST and SP. To examine the effect of seasonality on the outputs of the models, the data for the four seasons of each year (as the seasonal mean values of the data of six years in each study area) were used separately in the modeling because each season has different environmental conditions. Thus, three kinds of single AI-based models were used to estimate EBV of the four seasons on each case study (in total 24 single AI models) which were fed by seasonal mean values of the data of six years, gathered using the RS tools and the meteorological observations.

NDVI is one of the most widely used indexes to evaluate vegetation. The structural basis of this index is the presence of chlorophyll in different plants, which absorbs red light; also, the mesospheric layer of the leaf reflects near-infrared (NIR) light. This index exhibits well the response to photosynthetic effects, with higher values indicating denser and fresher vegetation, which greatly influences environmental parameters (Pettorelli *et al.* 2005). Plants and their roots affect the physical properties of soil, such as moisture content, infiltration rate, and shear strength, which play a significant role in environmental conditions (Gyssels *et al.* 2005). The general formula of the NDVI is (Pettorelli *et al.* 2005):

$$NDVI = (NIR - RED)/(NIR + RED) \quad (1)$$

In Equation (1), *NIR* and *RED* denote the near-infrared band and the red band, respectively.

In Figure 3, which illustrates the average NDVI data for the past six years for both study areas in spring, it is clear that the vegetation rate in pixels expressing urban areas in Figure 3(a) is about 0. It was also observed that the vegetation in the western part of Tabriz is richer than in the other areas. These areas include the flatlands of Tabriz plain. This index has lower values in the east and northeast parts of the image, which is related to the mountains of Onibnali and Bababaghi. But dense vegetation of the western areas is related to irrigation farming that can lead to lower EBV than eastern parts that have meadow and bunch grasses. Moreover, the vegetation level in the southern part of the city, which is located in the Sahand mountain range, is low, which causes lower EBV than in northern areas. For Gilan Province, as shown in Figure 3(b), vegetation is poor in pixels related to the urban areas. Other pixels show dense vegetation that is related to evergreen broadleaf forests. Orchard and tea gardens of this zone are ranked in VI and IV classes of the EBV. Rasht and its surrounding areas are part of the Caspian Sea plains and Gilan plain. Comparing the numerical indices of the vegetation images of Tabriz and Rasht in spring, the maximum NDVI rate in Rasht is mostly equal to 1; however, the maximum value of this index for Tabriz is about 0.55, indicating the denser vegetation quality compared to Tabriz and consequently better EEQ of this area.

Warmth and wetness are the major factors that control global vegetation distribution. Variability in both weather

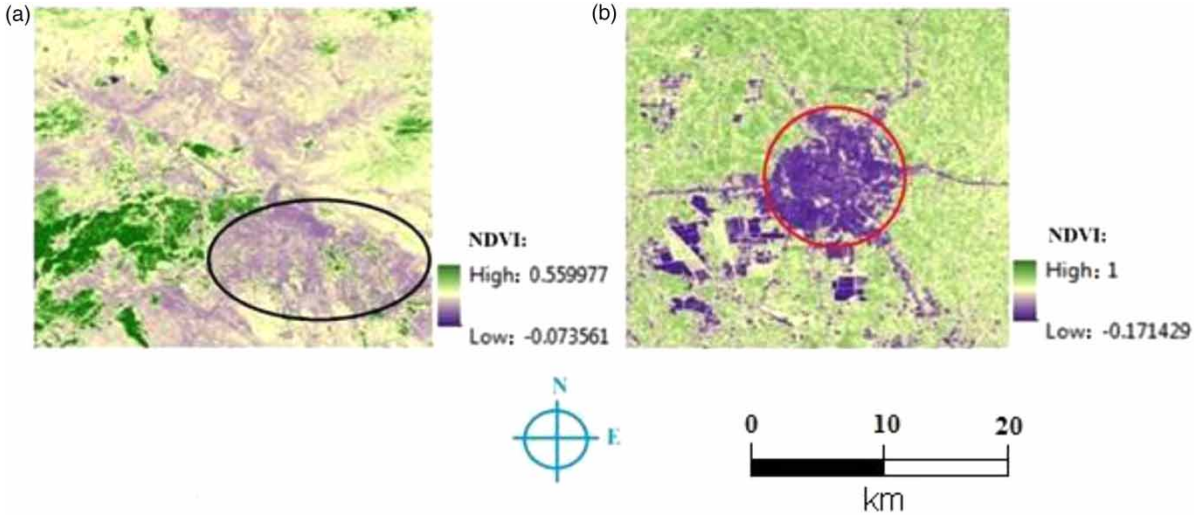

Figure 3 | NDVI map of the study areas in spring for (a) Tabriz, (b) Rasht.

and climate are therefore influenced by the moisture state. This issue also plays a vital role in the natural cycle of water, that has a significant role in environmental conditions, especially in the distribution of rainwater between surface runoff and penetration rate. Moisture is a principal factor that controls vegetation productivity, transpiration, evaporation, and rainfall-runoff. Therefore, in this study, for investigating the *EBV*, the *NDMI* parameter was used as Equation (2), which illustrates moisture as one of the thermal indicators and can be used with the *NDVI* index to determine the soil moisture content (Nguyen *et al.* 2016):

$$NDMI = (NIR - SWIR)/(NIR + SWIR) \quad (2)$$

In Equation (2), the *NIR* and short-wave infrared (*SWIR*) are associated with the near-infrared spectrum and the infrared spectrum with a short wavelength, respectively.

Investigating the images associated with the NDMI seasonal average of six years for Tabriz and Rasht and the surrounding areas shown in Figure 4(a) indicates that the amount of this index is higher in Tabriz plain in comparison to other areas. This figure illustrates that the rate of this index in the pixels of the urban areas has the lowest value. Figure 4(b) indicates a low value of this index in Rasht city concerning the other parts of the image, most likely due to the presence of urban terrains. In urban areas, because of the presence of buildings, there is less crop cover and higher water stress level.

LST is another key parameter in environmental studies, especially drought monitoring and variables affecting the environment. This index is an important parameter in land surface physics and energy flux between the earth and the atmosphere. LST is charged mainly due to irradiation of natural covering such as soil, water, snow, and vegetation. This indicator is a variable applied in a wide range of earth and environmental science and studies, particularly in projects in which a wide view of the location is needed. Surface temperature is very important in environmental studies because it measures the temperature in the air close to the earth surface. Air gets warm owing to radiation exchange from land to the atmosphere (Sobrino *et al.* 1991). LST obtained from RS tools can be used to monitor climate and understand the environmental conditions.

Examining the LST maps for Tabriz and its surrounding areas (Figure 5(a)) revealed that the temperature of the northern part of the region is lower than in the other parts. In addition, the highest temperature is associated with the northeastern and western parts of the image. It causes this zone to be classified as cool in some pixels and as warm heat in a few pixels, according to Table 1. Moreover, by examining Figure 5(b), which indicates the six-year average of the LST in the spring for Rasht and its surrounding areas, it can be seen that the temperature of Rasht city is higher than in the

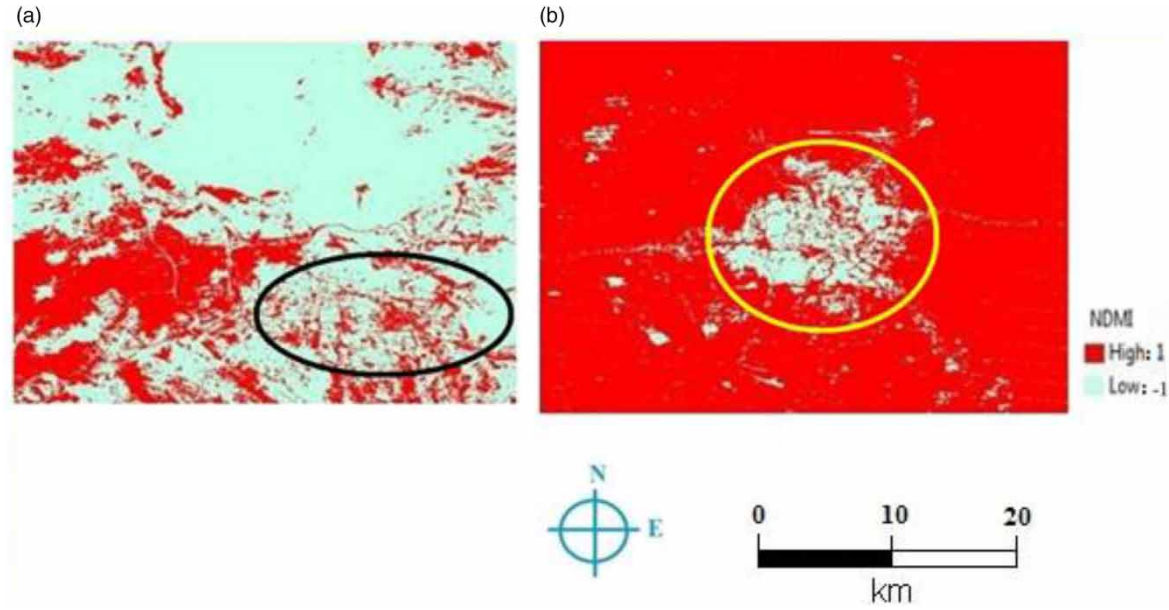

Figure 4 | NDMI map of the study areas in spring for (a) Tabriz, (b) Rasht.

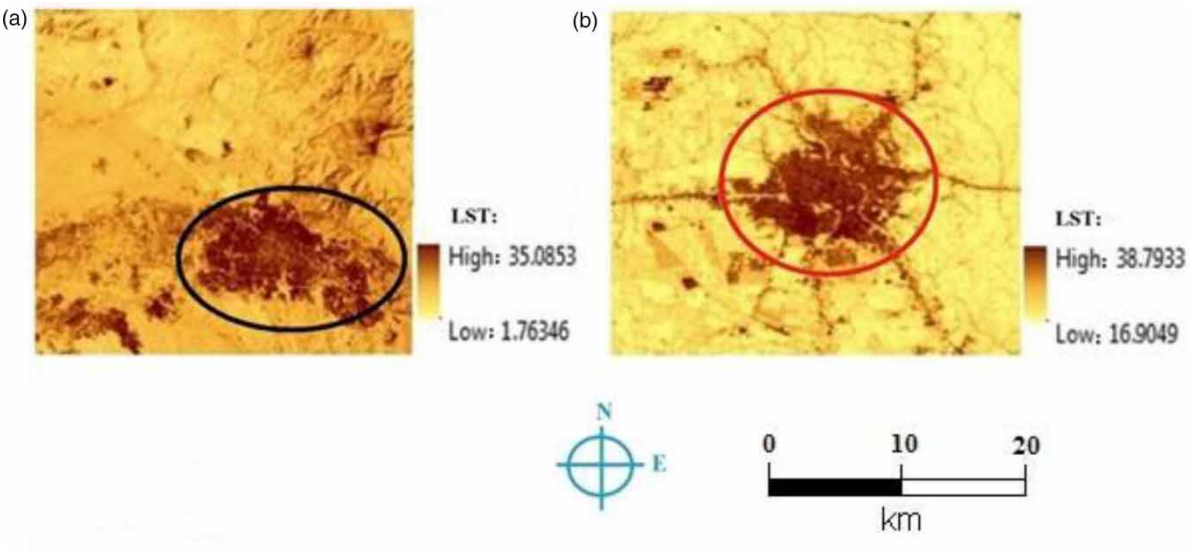

Figure 5 | LST map of the study areas in spring for (a) Tabriz, (b) Rasht.

other areas of the LST map for this case study. Investigating LST of this case study shows warm weather for approximately all parts. It causes this area to be classified as a warm region, according to Table 1. Furthermore, by examining both images in Figure 5, it can be concluded that the temperature of the pixels related to the cities of Tabriz and Rasht is higher than that of their surrounding areas, which is due to the human activities concentrated in the cities that cause heat production and temperature rise, which in turn can harm EEQ. Human activities are sources of heat production. People in urban areas perform a lot of activities such as running businesses, burning fuels, and factories. These activities not only produce heat, but also produce greenhouse gases that trap the heat near the ground and make it warmer.

DEM is a set of elevation measurements for locations distributed on the ground. DEMs are important inputs for topography for the accurate modeling of flood hydrodynamics, and floods have a key role in the natural environment (Zhang *et al.* 2013).

Investigation of the DEMs for both regions revealed that the areas in East Azerbaijan have higher elevation changes in comparison to the areas in Gilan. The western parts of Tabriz, which are located in the Tabriz plain, have lower elevations compared to the other parts. The northern, eastern, and southern parts that have higher elevations highlight the mountains around Tabriz. Most parts of the area in Tabriz, that have elevations higher than 1,200 m, can be classified as class V according to Table 1. Furthermore, in Gilan, the northern parts of the region, which are close to the Caspian Sea, are lower than other regions and have lower elevations. In the southern and western parts, approaching the Alborz mountain, the height of the areas increases. According to various hills and coastal low hills and uplands of this study area, they can be classified as classes VI and IV according to Table 1.

For ground-based data, SP and ST values were calculated and used as meteorological data using monthly temperature and precipitation time series observed in 19 meteorological stations in East Azerbaijan Province and 13 meteorological stations in Gilan Province. These two attributes were estimated and interpolated using the Kriging method (Hudson & Wackernagel 1994; Yin *et al.* 2018). All mentioned EIs were resampled with a cell size of 300 m at a grid level in both study areas. According to Table 1, we chose some attributes and analyzed their impacts on the performance of models. The results led us to choose these six attributes as inputs. These EIs had the most considerable effects on the performance of models among some environmental attributes that could be used as inputs. These attributes were used as inputs for AI-based models to learn the relationship between these inputs and output (EBV).

Used AI methods

In this paper, EBV of the case studies was estimated via the three single AI-based models of SVR, FFNN, and ANFIS. In the next sub-sections, the details of the models are presented.

Artificial neural network (ANN)

In recent decades, the unique features of the human brain have led researchers to simulate the abilities of the human brain with computers. ANNs are dynamical systems that transfer the knowledge or rule behind data to a network structure through processing the empirical data. Hence, they are AI models. The input, hidden, and output layers are the layers of an ANN. So far, various types of ANNs have been introduced, given the type of transfer functions, the network forming layers, and the impact of the weights on the inputs (Schalkoff 1997). It has been already shown that the most commonly used neural network is a feed forward neural network (FFNN) (Nourani 2017).

Adaptive neural fuzzy inference system (ANFIS)

ANFIS as a neuro-fuzzy model tries to use the benefits of fuzzy logic concepts and the neural network by combining them within a unique framework (Farhoudi *et al.* 2010). Fuzzy systems contain three basic parts, fuzzification, fuzzy database, and defuzzification, whereas a fuzzy rule base and an inference system are two sections of the fuzzy database part. Different fuzzy inference systems (FISs) can be used for fuzzy operation, and the current research utilized the Sugeno system (Aqil *et al.* 2007).

Support vector regression (SVR)

SVR based on the support vector machine (SVM) concept is used for estimating nonlinear regression objects. SVM-based methods such as SVR presume to minimize operational risk as the objective function instead of error minimization between the estimated and evaluated values. As the first stage in SVR, linear regression is applied to the dataset and then a nonlinear kernel is applied to the outputs to capture the nonlinear pattern of the data. The general SVR function due to give a set of training data $\{(x_i,\ d_i)\}_i^N$ (x_i is the input vector, d_i is the actual value and N is the total number of data variables), is (Wang *et al.* 2013):

$$y = f(x) = w\varphi(x_i) + b \tag{3}$$

where $\varphi(x_i)$ illustrates attribute spaces, nonlinearly mapped from input vector x. Regression variables of b and w may be determined by minimizing the objective function and assigning positive values for the slack parameters of ξ and ξ^* (Wang *et al.* 2013):

$$\text{Minimize: } \frac{1}{2}\|w^2\| + c\left(\sum_i^N (\xi_i + \xi_i^*)\right) \tag{4}$$

$$\text{Subject to: } \begin{cases} w_i\varphi(x_i) + b_i - d_i \le \varepsilon + \xi_i^* \\ d_i - w_i\varphi(x_i) - b_i \le \varepsilon + \xi_i^* \\ \xi_i,\ \xi_i^* \ge 0 \end{cases}$$

where $(1/2)\|w^2\|$ is the weights vector norm and c is called the regularized constant which defines the tradeoff between the regularized term and the empirical error. ϵ is referred to as the tube size and is equivalent to the estimation accuracy positioned within the training datasets. Lagrange multipliers α_i and α_i^* are defined to change the mentioned optimization problems to the dual quadratic optimization problem. Vector w in Equation (3) can be calculated after solving the quadratic optimization issue as (Wang *et al.* 2013):

$$w^* = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)\varphi(x_i) \tag{5}$$

Thus, the final formation of SVR can be illustrated as (Wang *et al.* 2013):

$$f(x,\ \alpha_i,\ \alpha_i^*) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)K(x_i, x_j) + b \tag{6}$$

where α_i and α_i^* are Lagrange multipliers, $K(x_i, x_j)$ is the kernel function applying the nonlinear mapping into feature space, and b is the bias term. Gaussian RBF (radial basis function) kernel is a commonly used kernel function which works as (Haghiabi *et al.* 2017):

$$K(x_1,\ x_2) = \exp(-\gamma x_1 - x_2^2) \tag{7}$$

In Equation (7), γ is the kernel parameter.

Efficiency criteria

To assess the performance of the models, the well-known efficiency criteria: determination coefficient (DC), root mean square error (RMSE), and Akaike information criterion (AIC) were used as Equations (8), (9), and (10), respectively (Nourani 2017; Ahmadi *et al.* 2019).

$$DC = 1 - \frac{\sum_{i=1}^{n} (EBV_{eval_i} - EBV_{est_i})^2}{\sum_{i=1}^{n} (EBV_{eval_i} - \overline{EBV}_{eval})^2} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (EBV_{eval_i} - EBV_{est_i})^2}{n}} \tag{9}$$

$$AIC = N \times \ln\left(\frac{R_m}{N}\right) + 2 \times P \tag{10}$$

where n, EBV_{eval_i}, $\overline{EBV}_{eval}$, EBV_{est_i}, N, P, and R_m are, respectively, pixel number, evaluated EBV (via site investigation), averaged value of the evaluated EBV (via site investigation), estimated EBV, the numbers of instances and model parameters, and the loss function value.

Combining unit

A post-processing approach can improve overall estimation by combining different estimations. Combining different estimations, via different models, has the advantage that it is not under the effect of a specific choice of models among the several used models. In this way, applying the combination estimate is preferred to using a sole method. It has been suggested that the overall efficiency of estimation can be enhanced by combining the outputs of various models (Sharghi *et al.* 2018).

In this study, three combining techniques were utilized to combine the outputs of three single AI-based models (ANN, ANFIS, SVR) to improve EEQ estimation as a) the simple linear averaging method, b) the linear weighted averaging method, and c) the nonlinear neural combining technique. In the nonlinear neural combining technique, another FFNN is designed to combine the outputs of the single AI-based models in contrast to linear combination methods. Simple averaging is utilized as

(Sharghi *et al.* 2018):

$$\bar{f}(t) = \frac{1}{N}\sum_{i=1}^{N} f_i(t) \tag{11}$$

where $\bar{f}(t)$ is the output of the simple combining technique, $f_i(t)$ is the output of the ith single model (here outputs of FFNN, ANFIS, and SVR) and N is the number of single AI-based models (here, $N = 3$).

The weighted averaging model is done as (Sharghi *et al.* 2018):

$$\bar{f}(t) = \sum_{i=1}^{N} w_i f_i(t) \tag{12}$$

where w_i is the exerted weight on the ith model that can be specified based on the model performance as:

$$w_i = \frac{DC_i}{\sum_{i=1}^{N} DC_i} \tag{13}$$

The performance efficiency of DC_i (e.g., determination coefficient) is the efficiency of the ith single AI-based model (Sharghi *et al.* 2018).

In the neural combining technique, an FFNN can be used to apply nonlinear averaging. Figure 6 shows the designed FFNN combining model for the current paper.

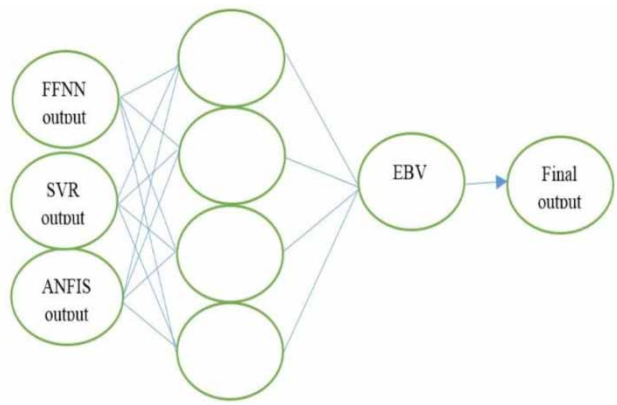

Figure 6 | Schematic of proposed neural combining model.

RESULTS AND DISCUSSION

For training models to estimate EBV, six potential input attributes (NDVI, NDMI, LST, DEM, ST, and SP) as the logic units were used in the input layer. Investigated EBV value by ecological experts of each sampling pixel was used in the output layer as the single logic unit of AI-based models for four seasons. Other EIs can be used as inputs for estimating EEQ indexes, such as radiance, soil brightness, etc., but in this study, such data were not available. EBVs in Tabriz and Rasht cities were modeled using FFNN, SVR, and ANFIS, separately. The data were first normalized, as this operation has a significant effect on increasing the accuracy and improving the results of prediction (Elshorbagy *et al.* 2000). Thereafter, three combining techniques were utilized to combine the results of single AI-based models. Finally, spatial distribution of EBV over the whole study area was derived by outputs of the combining technique.

According to recent studies, in this paper, the dataset was divided into two parts for AI models as training and verification sets; the first division as 70% of total data was used as the training set and the remaining 30% of data was used for the verification purpose. Although some studies divide data into three subsets (calibration, verification, and test) when developing AI models, such data division is optional and it is sufficient to divide data into the two subsets of calibration and verification. When dividing data into three subsets, the training (calibration) set is used to calibrate the data until the epoch (iteration) so that the RMSE reaches to minimum value for the test dataset; at this point, the calibrated AI-based models and parameters of the model are verified by the third subset of data. However, in dividing data into two subsets, when the DCs for the calibration and verification sets reach the maximum values, and the differences of these DCs reach the optimum values, it would be expected that overtraining is avoided. Such data division into two subsets has already been employed in AI-based modeling by many studies for hydrological modeling (e.g., see Asadi *et al.* 2013; Komasi & Sharghi 2016).

Results of single AI-based models

The structure of the single AI-based models with the best results in terms of the RMSE, AIC, and DC could be used

as a desirable and optimal structure for modeling. In the training phase, the AI models were trained to learn the relationship between input and output datasets. In both regions, among 300 sampling pixels, 210 and 90 pixels were randomly selected and utilized in calibration and verification steps, respectively. The FFNNs were trained using a scaled conjugate gradient scheme of the back propagation (BP) algorithm considering tangent sigmoid as activation functions (Haykin 1994). In the training process, the learning algorithm tries to minimize the error between the targets and the outputs by redistributing the error back through the model. This job is achieved through several iterations and the cycles are known as epochs (Singh *et al.* 2009). A trial and error procedure was used to determine the best number of neurons in the hidden layer and epoch number. For ANFIS models, a Sugeno FIS was utilized. An ANFIS model consists of several membership functions (MFs) and rules. In the current study, Gaussian showed the best performance in the modeling procedure. Furthermore, constant MF was applied for the output layer of the ANFIS models. The alteration of the training epoch was also investigated to gather the optimum ANFIS models in addition to examining the number of MFs. Epoch number and number of MFs were determined through trial and error. Finally, SVR models were designed by the RBF kernel. The tuning parameters of the RBF kernel are fewer than the parameters in the two sigmoid and polynomial kernels. Also, considering smoothness assumptions causes this kernel to lead to better performance (Noori *et al.* 2011). To achieve the best performance efficiency, the parameters of the RBF kernel were tuned. The statistical performance measures and obtained results for the best structure models are presented for the Tabriz and Rasht case studies in Table 2 for both calibration and verification datasets.

According to the obtained results (see Table 2), it is clear that among the models for the four seasons in both regions,

Table 2 | Results of single AI-based models for both case studies for four seasons

				DC		RMSE[b]		
Case study	**Season**	**Model**	**Model structure[a]**	**Calibration**	**Verification**	**Calibration**	**Verification**	**AIC**
Tabriz	Spring	SVR	0.268-0.125-20	0.901	0.797	0.069	0.091	−0.804
		FFNN	6-8-1	0.876	0.755	0.081	0.098	−0.677
		ANFIS	Gaussian-3	0.853	0.729	0.113	0.123	−0.549
	Summer	SVR	0.250-0.112-15	0.875	0.783	0.073	0.101	−0.663
		FFNN	6-9-1	0.845	0.737	0.089	0.103	−0.643
		ANFIS	Gaussian-3	0.802	0.721	0.114	0.126	−0.527
	Autumn	SVR	0.200-0.125-20	0.868	0.781	0.075	0.104	−0.672
		FFNN	6-8-1	0.844	0.723	0.093	0.107	−0.630
		ANFIS	Gaussian-3	0.789	0.718	0.115	0.129	−0.464
	Winter	SVR	0.5-0.380-25	0.814	0.701	0.092	0.108	−0.605
		FFNN	6-12-1	0.787	0.673	0.101	0.114	−0.589
		ANFIS	Gaussian-2	0.766	0.662	0.117	0.132	−0.507
Rasht	Spring	SVR	0.310-0.145-15	0.838	0.769	0.075	0.096	−0.312
		FFNN	6-10-1	0.814	0.740	0.098	0.115	−0.257
		ANFIS	Gaussian-3	0.800	0.661	0.103	0.130	−0.225
	Summer	SVR	0.400-0.130-10	0.816	0.757	0.086	0.103	−0.209
		FFNN	6-12-1	0.788	0.727	0.107	0.123	−0.186
		ANFIS	Gaussian-3	0.731	0.682	0.126	0.136	−0.064
	Autumn	SVR	0.450-0.136-10	0.803	0.741	0.089	0.107	−0200
		FFNN	6-10-1	0.781	0.719	0.118	0.129	−0.186
		ANFIS	Gaussian-3	0.724	0.663	0.128	0.138	−0.092
	Winter	SVR	0.40-0.300-20	0.745	0.696	0.101	0.122	−0.146
		FFNN	6-14-1	0.714	0.639	0.122	0.133	−0.139
		ANFIS	Gaussian-2	0.693	0.604	0.131	0.144	−0.075

[a]Numbering of a-b-c in SVR structure represents γ, ϵ, c. The numbering of a-b-c in the structure of the neural network illustrates the number of the input layer, hidden layer, and output layer neurons. In ANFIS structure, MF-a denotes the used membership function and number of membership functions, respectively.
[b]Since all data are normalized, the RMSE has no dimension.

the models showed better performance in spring, which can be attributed to the proper atmospheric conditions such as humidity and temperature, as well as to the suitable quality of the vegetation and soil conditions in this season. In spring, the differences are more obvious than in the other seasons. For example, in spring, regions have different values of vegetation coverage and they are in their best condition; on the other hand, in winter, hardly any of the regions have a specific kind of vegetation. According to NDVI, sample pixels of both case studies showed the highest NDVI value in spring, that shows denser and better vegetation conditions in comparison to the other seasons. This fact could lead to better EBVs in spring. The minimum value of low NDVI was seen in winter because of more rain and snow cover in this season and the minimum value of high NDVI was seen in autumn. NDVI is a standardized way to measure healthy vegetation conditions, and higher NDVI causes better environmental conditions with a positive effect on EBV evaluation performance in spring. Investigation of LST map and ST for both case studies showed a moderate temperature in spring. Spring has mild high and low LST and ST values, whereas winter has cold and summer has hot weather. The environment has better quality in mild weather and this factor can lead to better performance of modeling for spring. Furthermore, the comparison of the models for the study regions indicated that the models have better performance for Tabriz. Taking into account the density and richness of vegetation in Gilan Province (highest NDVI value around 1 could be seen in spring) relative to East Azerbaijan Province (highest NDVI value around 0.56 was seen in spring) as well as the high humidity in this region, it can be concluded that the reason for the better results of the models obtained for the Tabriz case study may be due to the topographic, weather, and other environmental conditions that are more compatible with AI models. Tabriz, due to its environmental conditions, has heterogeneous data that helps AI models have a better training and learning phase. The models can experience a greater variety of data, as a result, they can perform better in the validation phase. The NDMI map showed high values (about 1) in almost all of the study pixels of Rasht and surrounding areas in spring, autumn, and winter, that can cause an unfavorable effect on the training process of the models because many of the same values of NDMI for different pixels that are used as inputs for AI models are imposed. Also, high atmospheric humidity in Rasht impairs heat exchange efficiency by reducing the rate of moisture evaporation from skin surfaces that plays an important favorable role for EEQ but causes an unfavorable effect on the performance of the models, because again, the temperature map includes many almost equal values for different pixels. Besides, comparing the DEMs of both case studies shows higher fluctuations of elevation in Tabriz with regard to Rasht. According to RS and meteorological features, Rasht has less heterogeneity of dataset which means that models should be trained by almost the same data which reduces the training performance of AI models. On the other hand, the widespread domain of data fluctuation for Tabriz having a semi-arid climate with four distinct seasons (rather than humid areas of Rasht) led to better performance of AI-based EBV modeling. Figure 7 illustrates the evaluated values of EBV and their corresponding values of EBV estimated in three designed single models for both study areas in four seasons. Among the various models that were developed, the SVR performance in the verification step showed somewhat better results than the other two single AI-based models in each season for both case studies (see Table 2). Investigation of the obtained results shown by scatter plots (see Figure 7) showed that FFNN could not estimate EBV of classes VIII and VI in some pixels and wrongly classified them as IX and VII, respectively (indicated by the circle in Figure 7(a) and 7(d)). On the other hand, SVR could estimate IX and VIII classes (indicated by the square in Figure 7(c) and 7(f)), but it could not estimate the EBV of other pixels or set them in the right classes. Also, ANFIS could estimate VII and V classes (indicated by the square in Figure 7(b) and 7(e)), but it does not show acceptable performance for the other classes. Although SVR led to better DC, AIC, and RMSE compared to the other single AI-based models, ANN and ANFIS models showed better performance for medium and high EEQs than SVR. By investigation and comparing the results of single AI models (see Figure 7 and Table 2), it is clear that none of the single AI-based models could evaluate EBV well for all classes, but some are much better for some conditions. Thus, it is expected that the overall performance of the models may be enhanced over the single AI-based models by combining the outputs via combining techniques.

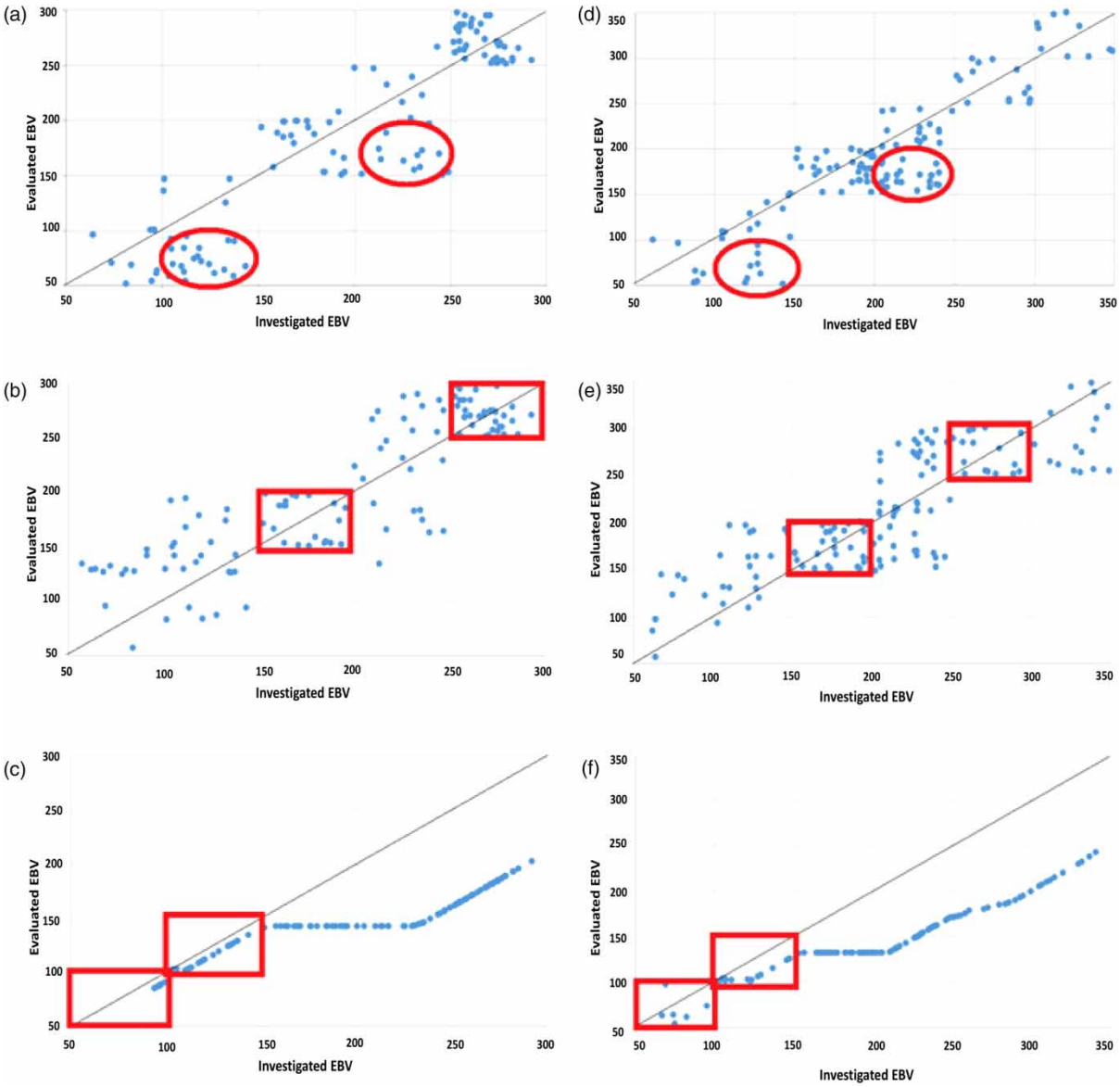

Figure 7 | The results of validation period for single models in spring for (a) ANN-Tabriz, (b) ANFIS-Tabriz, (c) SVR-Tabriz, (d) ANN-Rasht, (e) ANFIS-Rasht, (f) SVR-Rasht.

Results of combining techniques

In the next step, the outputs of single models were merged via three combining techniques to improve the performance of modeling for each case study.

To compute the variables of combining techniques, only the calibration dataset was used. In a neural combining technique, like single FFNN, the network was trained using a scaled conjugate gradient scheme of the BP algorithm, applying tangent sigmoid as activation functions of hidden and output layers, and the best epoch number and structure of the combining network were determined through the trial and error procedure. The results of combining models (via three combining techniques) in Tabriz and Rasht for four seasons are tabulated in Table 3.

The results of the combining techniques showed that almost all of the combining techniques could lead to better performance compared to the single models. Enhancement of modeling in Tabriz was up to 9, 10, and 15% in the

Table 3 | Results of combining models for both case studies for four seasons

				DC		RMSE[b]	
Case study	**Season**	**Combining model**	**Model structure[a]**	**Calibration**	**Verification**	**Calibration**	**Verification**
Tabriz	Spring	Simple averaging	–	0.91	0.81	0.068	0.054
		Weighted averaging	0.342-0.333-0.325	0.92	0.82	0.066	0.050
		Neural averaging	3-5-1	0.95	0.89	0.014	0.018
	Summer	Simple averaging	–	0.87	0.79	0.071	0.066
		Weighted averaging	0.346-0.335-0.319	0.87	0.81	0.070	0.064
		Neural averaging	3-6-1	0.93	0.89	0.038	0.045
	Autumn	Simple averaging	–	0.87	0.78	0.073	0.060
		Weighted averaging	0.346-0.338-0.316	0.88	0.80	0.072	0.060
		Neural averaging	3-4-1	0.92	0.90	0.041	0.047
	Winter	Simple averaging	–	0.83	0.71	0.091	0.0103
		Weighted averaging	0.343-0.333-0.324	0.85	0.72	0.091	0.101
		Neural averaging	3-8-1	0.91	0.86	0.057	0.062
Rasht	Spring	Simple averaging	–	0.84	0.76	0.074	0.055
		Weighted averaging	0.341-0.332-0.327	0.84	0.77	0.073	0.053
		Neural averaging		0.94	0.90	0.051	0.0.34
	Summer	Simple averaging	–	0.84	0.75	0.086	0.062
		Weighted averaging	0.348-0.337-0.315	0.85	0.76	0.084	0.061
		Neural averaging		0.91	0.90	0.046	0.047
	Autumn	Simple averaging	–	0.80	0.75	0.088	0.067
		Weighted averaging	0.348-0.338-0.314	0.81	0.75	0.086	0.067
		Neural averaging		0.91	0.87	0.052	0.049
	Winter	Simple averaging	–	0.75	0.70	0.099	0.110
		Weighted averaging	0.346-0.332-0.322	0.77	0.73	0.097	0.108
		Neural averaging		0.90	0.86	0.061	0.069

[a]Numbering of a-b-c in weighted averaging structure represents the weights of SVR, FFNN, and ANFIS models. The numbering of a-b-c in the structure of neural averaging denotes the number of the input layer, hidden layer, and output layer neurons.
[b]Since all data are normalized, the RMSE has no dimension.

calibration and up to 9, 10, and 20% in the verification step, respectively, for simple averaging, weighted averaging, and neural averaging. In Rasht, the enhancements were up to 11, 12, and 21% for the training and up to 10, 13 and 26% for the verification steps, respectively (see Figure 8). As discussed previously, models provide different performance according to their own advantages and disadvantages for different conditions. However, using the unique capability of individual models in the combining techniques contributes to a better estimation of the EBV compared to the single AI-based models. As shown in Table 3, the outputs of single models are in direct relation with the results of the weighted and simple averaging methods, thus the results of these kinds of combining techniques are almost equal.

Due to the use of nonlinear kernel in the FFNN-based combining technique, it can estimate EBV in both case studies for every season better than the other combining techniques. On the other hand, since the results of

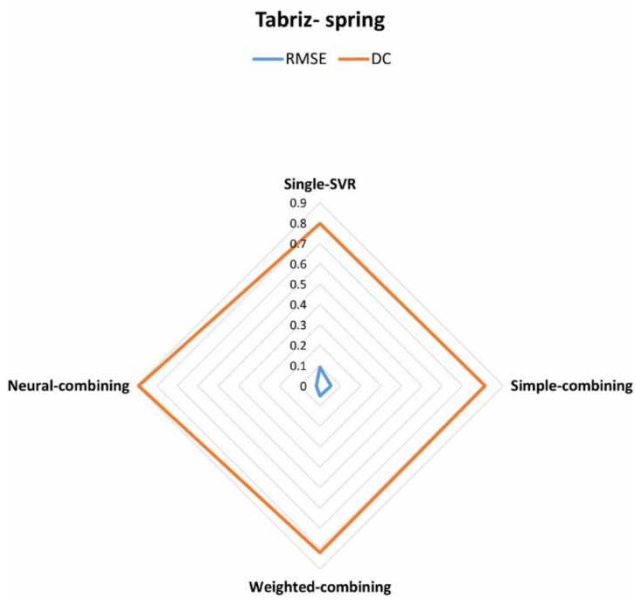

Figure 8 | Radar chart of RMSE and DC for single-SVR, simple-combining, weighted-combining, and neural-combining for verification phase, Tabriz-spring.

individual models had a direct relation with the results of both simple and weighted combining techniques, the obtained results of combining will be poor if the performance of single models is poor. In such conditions, the neural technique would be more helpful. Results of combining techniques show better performance for Tabriz city with regard to Rasht city and in both cases, spring showed better results among the seasons.

EBV distributions for whole study areas

According to Table 2 and Figure 7, none of the single AI-based models could estimate EBV for all classes, and then they could not be used to estimate EBV values for all pixels of case studies. However, combining outputs of the single AI-based models by combining techniques showed better performance for estimating EBV. Among the combining techniques, the FFNN-based technique led to much better performance (see Table 3). Hence, after training and verifying the FFNN-based combining technique using 300 sampling sites, this model was used to evaluate the spatial distribution of EBV ranks over all study area. For this purpose, NDVI, NDMI, DEM, LST, SP, and ST of all study areas were used as inputs for the accomplished model to map EBV over all study areas.

Figure 9 illustrates the spatial distribution of EBV ranking for both case studies obtained from the FFNN-based combining technique. According to the observed EBV *in situ*, although SVR as the best single model has some misestimations, the FFNN-based combined method has shown an acceptable performance. According to the observed EBV in the Tabriz case study, Tabriz city and other villages surrounding it are placed in the IX rank. The SVR model had some problems in its assessments, especially for villages around the city, and it could not have a correct estimation for them. On the other hand, the FFNN-based model could perform well in this class and it assessed residential areas except in a few small villages around the city. In the observed EBV, the agricultural lands belonging to this area, mostly located around cities and villages, were ranked in classes VII and

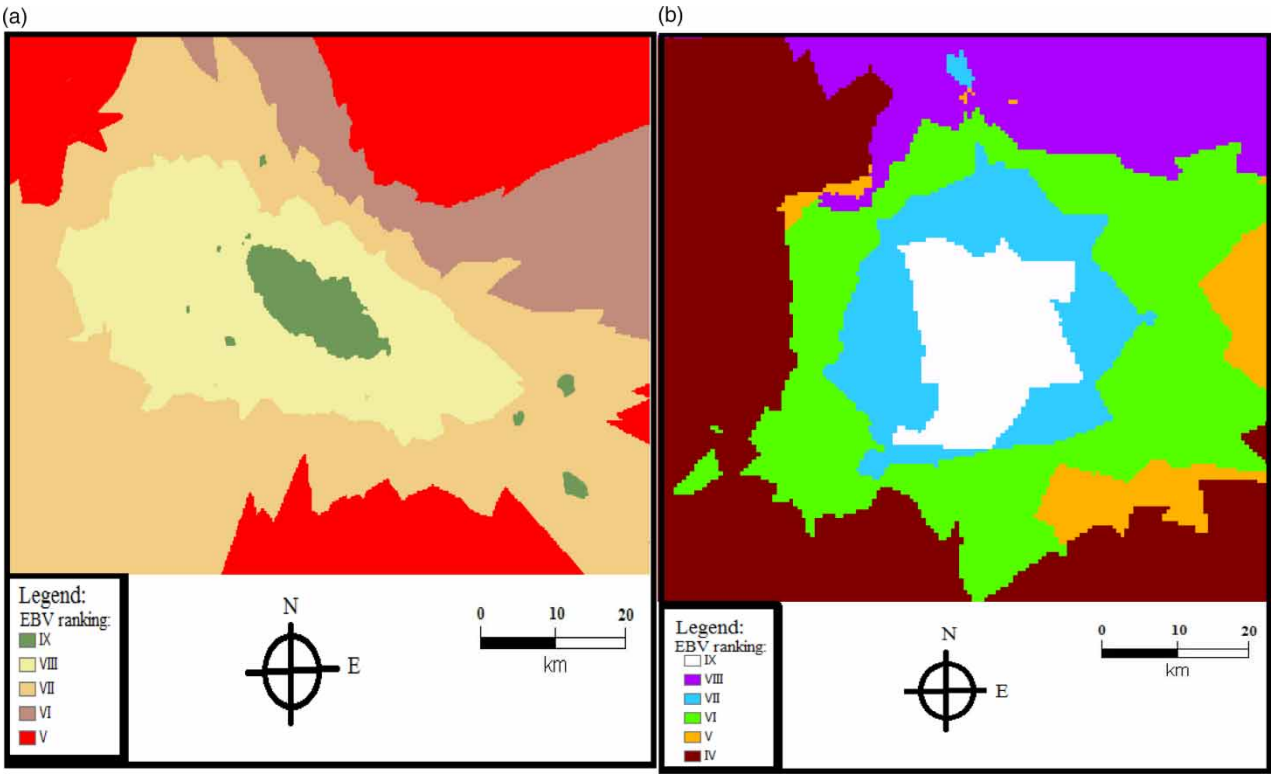

Figure 9 | EBV map of the studied areas in (a) Tabriz, (b) Rasht.

VIII. Other parts of this area in most pixels were placed in class V due to the elevation of the area (higher than 1,200 m), vegetation, the environmental species, humidity, and weather condition. In this case, the SVR model has a great number of underestimations, especially in the east and north parts of the region, but the FFNN-based model showed convincing outputs and its misestimations regarded finding the exact border between the different classes. Observed EBV for the Rasht case study indicated IX rank for urban areas. Agricultural lands in the north part of the map are in the VIII class; however, other farming pixels around Rasht city are in VII class. The other areas that are not related to agricultural lands and residential areas are mostly ranked as IV and VI classes. In this case, the SVR model had a great number of misestimations. It could not show villages, and it had some underestimations in IV, V, and VI classes. Although the FFNN-based model had better performance than SVR, it had some overestimations in estimating EBV for areas in IX class. Also, it had some problems in the north part of the map and its estimate for VII and V classes.

Investigation of these results showed that the EBV rate in urban areas is at its lowest level. Human activities are considered as an important negative factor on the EEQ.

CONCLUSIONS

Population growth, urban development, economic progress, and other human activities have endangered the environment. Monitoring and estimating of eco-environmental conditions are required to set developmental regulations and rules to prevent serious damage to the environment. In this way, EBV of Tabriz and Rasht cities of Iran was estimated via the three different single AI-based models of SVR, FFNN, and ANFIS. Thereafter, to combine the outputs of the single AI-based models, combining techniques were employed as post-processing methods to mapping and spatial distribution of EBV. The field observations were done and after absorbing the knowledge gleaned from conventional expert assessment, evaluating EBV (as targets of AI models) with documented data was done manually by ecological experts. Thereafter, ANN, ANFIS, and SVR models were used as single AI-based models to estimate EBV. For this purpose, to find the required input data, the RS tools and meteorological observations were used for all four seasons, separately, using data from 2013 to 2018. The comparison of the obtained results of single AI-based models showed that the SVR results were slightly better than the other single models, but even this model was not capable in the verification step of estimating EBV for all classes well. Therefore, to increase the EBV modeling efficiency, three combining techniques, including simple, weighted, and neural averaging techniques, were utilized to combine the outputs of the single AI-based models. Combining techniques produced better approximation than the single models for estimating EEQ. Generally, the reliability and efficiency of neural combining techniques were better than other combination methods, which could improve the performance of modeling up to 26%. Employing an FFNN model in the combining unit causes the success of the neural combining so that the modeling of the nonlinear behavior of the environment could be caught more accurately than the others using the nonlinear kernel.

The current study focused on two regions, classified as semi-arid and humid regions in the climatic divisions. Results indicate that using this methodology, especially in studies with limited pixels of the case study, is good for estimating EEQ in semi-arid regions having four distinct seasons, because this kind of region will have heterogeneous spatial (administered by RS features) and temporal (administered by meteorological features) datasets. On the other hand, humid regions with approximately equal ecological and environmental conditions in a restricted area will not have enough heterogeneity in the dataset to eventuate good results of EEQ evaluation using AI models. Hence, estimating EBV in arid regions seems to require a continuation of this study. Also, it is suggested to verify other nonlinear combining kernels such as ANFIS and SVR instead of FFNN in the combining unit.

ACKNOWLEDGEMENTS

The study has been financially supported by a grant from Research Affaires of University of Tabriz.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Ahmadi, A., Nasseri, M. & Solomatine, D. P. 2019 Parametric uncertainty assessment of hydrological models: coupling UNEEC-P and a fuzzy general regression neural network. *Hydrological Sciences Journal* **64** (9), 1080–1094.

Alizadeh, M. J., Kavianpour, M. R., Danesh, M., Adolf, J., Shamshirband, S. & Chau, K. W. 2018 Effect of river flow on the quality of estuarine and coastal waters using machine learning models. *Engineering Applications of Computational Fluid Mechanics* **12** (1), 810–823.

Aqil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 Analysis and prediction of flow from a local source in a river basin using a neuro-fuzzy modeling tool. *Journal of Environmental Management* **85** (1), 215–223.

Asadi, S., Shahrabi, J., Abbaszadeh, P. & Tabanmehr, S. 2013 A new hybrid artificial neural networks for rainfall–runoff process modeling. *Neurocomputing* **121**, 470–480.

Awan, J. A. & Bae, D. H. 2016 Drought prediction over the East Asian monsoon region using the adaptive neuro-fuzzy inference system and the global sea surface temperature anomalies. *International Journal of Climatology* **36** (15), 4767–4777.

Bates, J. M. & Granger, C. W. 1969 The combination of forecasts. *Journal of the Operational Research Society* **20** (4), 451–468.

Chai, L. H. & Lha, D. 2018 A new approach of deriving indicators and comprehensive measure for ecological environmental quality assessment. *Ecological Indicators* **85**, 716–728.

Choi, Y. D. 2004 Theories for ecological restoration in changing environment: toward 'futuristic' restoration. *Ecological Research* **19**, 75–81.

Dickinson, J. P. 1973 Some statistical results in the combination of forecasts. *Journal of the Operational Research Society* **24** (2), 253–260.

Dzeroski, S. 2001 Application of symbolic machine learning to ecological modeling. *Ecological Modeling* **146** (1–3), 263–273.

Elshorbagy, A., Simonovic, S. P. & Panu, U. S. 2000 Performance evaluation of artificial neural networks for runoff prediction. *Journal of Hydrologic Engineering* **5** (4), 424–427.

Farhoudi, J., Hosseini, S. M. & Sedghi-Asl, M. 2010 Application of the neuro-fuzzy model to estimate the characteristics of local scour downstream of stilling basins. *Journal of Hydroinformatics* **12** (2), 201–211.

Feizizadeh, B. & Blaschke, T. 2013 Land suitability analysis for Tabriz County, Iran: a multi-criteria evaluation approach using GIS. *Journal of Environmental Planning and Management* **56** (1), 1–23.

Gyssels, G., Poesen, J., Bochet, E. & Li, Y. 2005 Impact of plant roots on the resistance of soils to erosion by water: a review. *Progress in Physical Geography* **29** (2), 189–217.

Haghiabi, A. H., Azamathulla, H. M. & Parsaie, A. 2017 Prediction of head loss on cascade weir using ANN and SVM. *ISH Journal of Hydraulic Engineering* **23** (1), 102–110.

Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Englewood Cliffs, NJ, USA.

Hudson, G. & Wackernagel, H. 1994 Mapping temperature using kriging with external drift: theory and an example from Scotland. *International Journal of Climatology* **14** (1), 77–91.

IRIMO. 2006 *Iran Meteorological Organization*. Available from: https://www.irimo.ir (accessed February 2019).

Kecman, V. 2001 *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. MIT Press, Cambridge, MA, USA.

Komasi, M. & Sharghi, S. 2016 Hybrid wavelet-support vector machine approach for modelling rainfall–runoff process. *Water Science and Technology* **73** (8), 1937–1953.

Miao, C. L., Sun, L. Y. & Yang, L. 2016 The studies of ecological environmental quality assessment in Anhui Province based on ecological footprint. *Ecological Indicators* **60**, 879–883.

Moisen, G. & Frescino, T. S. 2002 Comparing five modeling techniques for predicting forest characteristics. *Ecological Modeling* **157** (2–3), 209–225.

Nguyen, A. K., Liou, Y. A., Li, M. H. & Tran, T. A. 2016 Zoning eco-environmental vulnerability for environmental management and protection. *Ecological Indicators* **69**, 100–117.

Nichol, J. E. & Wong, M. S. 2018 Assessing urban environmental quality with multiple parameters. In: *Urban Remote Sensing* (Q. Weng & D. A. Quattrochi eds). Taylor and Francis Group, LLC, CRC Press, Boca Raton, FL, USA, pp. 253–269.

Niu, W. Y. & Harris, W. M. 1996 China: the forecast of its environmental situation in the 21st century. *Journal of Environmental Management* **47** (2), 101–114.

Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A. & Gousheh, M. G. 2011 Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly streamflow prediction. *Journal of Hydrology* **401** (3–4), 177–189.

Nourani, V. 2017 An emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277.

Olyaie, E., Banejad, H., Chau, K. W. & Melesse, A. M. 2015 A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. *Environmental Monitoring and Assessment* **187** (4), 189.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J. & Stenseth, N. C. 2005 Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution* **20** (9), 503–510.

Quej, V. H., Almorox, J., Arnaldo, J. A. & Saito, L. 2017 ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *Journal of Atmospheric and Solar-Terrestrial Physics* **155**, 62–70.

Sarkar, B. C., Mahanta, B. N., Saikia, K., Paul, P. R. & Singh, G. 2007 Geo-environmental quality assessment in Jharia

coalfield, India, using multivariate statistics and geographic information system. *Environmental Geology* **51** (7), 1177.

Schalkoff, R. J. 1997 *Artificial Neural Networks, Vol. 1*. McGraw-Hill, New York, USA.

Setnes, M., Babuska, R. & Verbruggen, H. B. 1998 Transparent fuzzy modeling. *International Journal of Human-Computer Studies* **49** (2), 159–179.

Shamshirband, S., Jafari Nodoushan, E., Adolf, J. E., Abdul Manaf, A., Mosavi, A. & Chau, K. W. 2019 Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. *Engineering Applications of Computational Fluid Mechanics* **13** (1), 91–101.

Sharghi, E., Nourani, V. & Behfar, N. 2018 Earthfill dam seepage analysis using ensemble artificial intelligence-based modeling. *Journal of Hydroinformatics* **20** (5), 1071–1084.

Shi, Z. & Li, H. 2007 Application of artificial neural network approach and remotely sensed imagery for regional eco-environmental quality evaluation. *Environmental Monitoring and Assessment* **128** (1–3), 217–229.

Singh, K. P., Basant, A., Malik, A. & Jain, G. 2009 Artificial neural network modeling of the river water quality – a case study. *Ecological Modelling* **220** (6), 888–895.

Sobrino, J., Coll, C. & Caselles, V. 1991 Atmospheric correction for land surface temperature using NOAA-11 AVHRR channels 4 and 5. *Remote Sensing of Environment* **38** (1), 19–34.

Srivastava, P. K., Islam, T., Singh, S. K., Gupta, M., Petropoulos, G. P., Gupta, D. K. & Prasad, R. 2016 Soil moisture deficit estimation through SMOS soil moisture and MODIS land surface temperature. In: *Satellite Soil Moisture Retrieval* (P. K. Srivastava, G. P. Petropoulos & Y. H. Kerr eds). Elsevier, Amsterdam, the Netherlands, pp. 333–347.

Tabari, H., Kisi, O., Ezani, A. & Talaee, P. H. 2012 SVM, ANFIS, regression, and climate-based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology* **444**, 78–89.

Wang, W. C., Xu, D. M., Chau, K. W. & Chen, S. 2013 Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. *Journal of Hydroinformatics* **15** (4), 1377–1390.

Yang, K., Yu, Z., Luo, Y., Yang, Y., Zhao, L. & Zhou, X. 2018 Spatial and temporal variations in the relationship between lake water surface temperatures and water quality-A case study of Dianchi Lake. *Science of the Total Environment* **624**, 859–871.

Yin, S. Q., Wang, Z., Zhu, Z., Zou, X. K. & Wang, W. T. 2018 Using Kriging with a heterogeneous measurement error to improve the accuracy of extreme precipitation return level estimation. *Journal of Hydrology* **562**, 518–529.

Ying, X., Zeng, G. M., Chen, G. Q., Tang, L., Wang, K. L. & Huang, D. Y. 2007 Combining AHP with GIS in synthetic evaluation of eco-environment quality – a case study of Hunan Province, China. *Ecological Modeling* **209** (2–4), 97–109.

Zarei, T., Behyad, R. & Abedini, E. 2018 Study on parameters effective on the performance of a humidification-dehumidification seawater greenhouse using support vector regression. *Desalination* **435**, 235–245.

Zareie, S., Khosravi, H., Nasiri, A. & Dastorani, M. 2016 Using Landsat Thematic Mapper (TM) sensor to detect change in land surface temperature in relation to land use change in Yazd, Iran. *Solid Earth* **7** (6), 1551.

Zhang, G. P. 2003 Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175.

Zhang, Z., van Coillie, F., De Clercq, E. M., Ou, X. & De Wulf, R. 2013 Mountain vegetation change quantification using surface landscape metrics in Lancang watershed, China. *Ecological Indicators* **31**, 49–58.

First received 20 March 2020; accepted in revised form 28 September 2020. Available online 3 November 2020