Research Article

Modeling Evapotranspiration Response to Climatic Forcings Using Data-Driven Techniques in Grassland Ecosystems

Xianming Dou^{1,2} and Yongguo Yang^{1,2}

¹Key Laboratory of Coalbed Methane Resources and Reservoir Formation Process of Ministry of Education, China University of Mining and Technology, Xuzhou 221116, China

²School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, China

Correspondence should be addressed to Yongguo Yang; yongguoyang@hotmail.com

Received 6 November 2017; Revised 9 February 2018; Accepted 15 March 2018; Published 22 April 2018

Academic Editor: Harry D. Kambezidis

Copyright © 2018 Xianming Dou and Yongguo Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Remarkable progress has been made over the last decade toward characterizing the mechanisms that dominate the exchange of water vapor between the biosphere and the atmosphere. This is attributed partly to the considerable development of machine learning techniques that allow the scientific community to use these advanced tools for approximating the nonlinear processes affecting the variation of water vapor in terrestrial ecosystems. Three novel machine learning approaches, namely, group method of data handling, extreme learning machine (ELM), and adaptive neurofuzzy inference system (ANFIS), were developed to simulate and forecast the daily evapotranspiration (ET) at four different grassland sites based on the flux tower data using the eddy covariance method. These models were compared with the extensively utilized data-driven models, including artificial neural network, generalized regression neural network, and support vector machine (SVM). Moreover, the influences of internal functions on their corresponding models (SVM, ELM, and ANFIS) were investigated together. It was demonstrated that most developed models did good job of simulating and forecasting daily ET at the four sites. In addition to strengths of robustness and simplicity, the newly proposed methods achieved the estimates comparable to those of the conventional approaches and accordingly can be used as promising alternatives to traditional methods. It was further discovered that the generalization performance of the ELM, ANFIS, and SVM models strongly depended on their respective internal functions, especially for SVM.

1. Introduction

Grasslands cover approximately 13% of global land surface [1]. It is universally acknowledged that grassland ecosystems play a crucial role in global water and carbon budgets as well as energy balances [2, 3]. In grassland ecosystems, the interrelations between atmosphere and evapotranspiration (ET), consisting of transpiration and evaporation from the soil and plant systems, are closely related to ecosystem productivity, particularly in arid and semiarid regions [4]. Their dynamical interactions between biophysical and biogeochemical processes are substantially affected by a variety of environmental driving factors, such as climate change, nutrient deposition, and increased atmospheric CO₂ [4–6]. Therefore, measuring and estimating ET in grasslands are

relatively essential for understanding the responses of ET to different environmental forces and offering precise ET budgets for solving the issue of water resource allocation and management. Eddy covariance technique has been considered as a reliable approach and is being used extensively to measure grassland ET in different climates [7]. A large amount of data from numerous flux tower sites, including ET and other micrometeorological variables, have been collected and archived in global flux network in recent years [8, 9], providing a solid foundation for obtaining helpful knowledge of the underlying mechanisms influencing the variation of ET across various time scales from hourly to interannual. However, accurately estimating ET of grassland ecosystems remains a great challenge due to the limited understanding of how ET responds to human practical activities (e.g., land use

change, grazing, and fertilization) and climate extreme events (e.g., extreme precipitation, wind storms, droughts, and heat waves).

A broad range of approaches are presently utilized to model and predict terrestrial ET at various spatial levels from ecosystem to regional or global level. These methods can be commonly grouped into three categories: remote sensing-based [10, 11], physically based [12, 13], and data-driven [14, 15]. Among these approaches, physically based techniques are universally used by both the ecological and hydrological communities, due to their distinct advantage of offering a physical insight into the complex hydrological process controlling the temporal and spatial variation of ET. However, temporal and spatial heterogeneity of land surface characteristics strongly affects the underlying processes of ongoing water vapor exchange, which considerably increases the complexity of elucidating the hydrological relationship between land and the atmosphere and subsequently impedes the pursuit of obtaining accurate physically based models.

Data-driven methods have been receiving increasing attention in the last decade for ET simulation and prediction at different terrestrial ecosystems. Numerous previous studies have reported the ability of machine learning techniques in the prediction of ET at terrestrial ecosystems [16, 17]. It has been demonstrated that these approaches are capable of effectively seizing most useful information stored in the available data (e.g., the measurements based on eddy covariance method), without depending on any complex underlying knowledge about the evolving process of ET. Nevertheless, to the best of our knowledge, most studies focused primarily on using three ubiquitous approaches, namely, artificial neural network (ANN), generalized regression neural network (GRNN), and support vector machine (SVM).

In recent years, many novel machine learning techniques have been proposed for handling the complex nonlinear problems, such as adaptive neurofuzzy inference system (ANFIS), extreme learning machine (ELM), and group method of data handling (GMDH). The generalization ability of these advanced approaches has been extensively acknowledged in forecasting the environmental and climatological variables. More particularly, these methods have been widely applied in the nonlinear time series prediction in several various branches of hydrology, such as reference ET forecasting [18, 19], evaporation prediction [20, 21], and soil moisture estimation [22, 23]. However, comparatively few attempts to apply these state-of-the-art approaches have been made to date for estimating the ET at terrestrial ecosystems based on the eddy covariance data.

In addition, all the techniques, including the conventional ANN, GRNN, and SVM methods, and the relative novel methods (ELM, ANFIS, and GMDH) proposed by this study remarkably differ in terms of their structures, principles, and parameters. And, therefore, their generalization ability should be different in theory. For this reason, regarding their practical application in the present study, systematically evaluating the ability of all the methods for ET prediction is particularly important with the objective of ensuring the predictive accuracy. Besides, it is also beneficial to identify

the strengths and shortcomings of each method and provide helpful suggestions for the relevant researchers.

It is a widely accepted fact that the ability of all the above-mentioned methods is strongly affected by their corresponding inner parameters [24]. However, very few studies have reported the roles of different intrinsic parameters for each approach in characterizing the actual behavior of the addressed problems, especially for the relatively new methods, ANFIS and ELM, even though these approaches have been widely applied in other fields. For example, although several recent studies have examined the effects of different membership functions on ANFIS for various hydrological time series prediction [25–27], it remains unclear how the ANFIS model is influenced by different identification algorithms for fuzzy inference system (FIS) generation for the present application. In fact, when establishing these models in practical applications, these intrinsic algorithms in most studies are commonly determined by random selection or trial-and-error procedure, which may limit, to some extent, the ability of related models. Accordingly, it is certainly worth examining the impacts of inner parameters on SVM, ANFIS, and ELM for predicting ET, which is also another major task for the current research.

Considering the presently existing above-mentioned issues, the novelty of this paper lies in the use of three state-of-the-art data-driven methods for estimating daily ET based on the available eddy covariance measurements at four different grassland sites. More specifically, the novelty of this paper is threefold as follows. First, we probe into the feasibility and effectiveness of three modern data-driven models, including GMDH, ELM, and ANFIS, for modeling and predicting the ET at the daily time scale. Second, we investigate whether these newly proposed models have the potential to sufficiently complement or replace the conventionally recognized methods, ANN, GRNN, and SVM. Finally, we particularly focus on evaluating the impacts of intrinsic algorithms involved in each model (SVM, ELM, and ANFIS) on their respective generalization ability.

2. Materials and Methods

2.1. Site Description and Data Used. In this study, four grassland ecosystems were used to explore the modeling capability of the proposed models. These sites are situated in different countries. The details of each site are shown in Table 1. They were selected for the present study mainly due to their continuous multiyear measurements of water vapor flux (ET) with the eddy covariance technique. In addition, the measurements of other variables were also undertaken at these sites, primarily including CO₂ flux, energy fluxes (latent heat and sensible heat), and meteorological variables. The utilized instruments and processing procedures of related data at each site have been, respectively, reported in previous studies as shown in Table 1.

Statistical parameters of daily measurements utilized in this study during at least 6-year period for all the sites are given in Table 2. As shown in the table, the annual mean air temperature (T_a) changes from 6.48°C at CA-Let to 10.20°C at HU-Bug; the annual mean soil temperature (T_s) is between

TABLE 1: Site characteristics used in this study.

Site	Latitude	Longitude	Elevation	MAT	TAP	Climate	Period	Reference
AT-Neu	47.12	11.32	970	6.68	669	Temperate	2004–2009	Wohlfahrt et al. [28]
CA-Let	49.71	−112.94	951	6.48	1095	boreal	1999–2007	Flanagan et al. [29]
DE-Gri	50.95	13.51	385	8.35	934	Temperate	2004–07, 2009–10	Hussain et al. [30]
HU-Bug	46.69	19.60	140	10.20	499	Temperate	2003–2008	Carvalhais et al. [31]

Note. MAT and TAP denote the mean annual temperature and total annual precipitation, respectively, and they are calculated based on the time period (period column). The units of elevation, MAT, and TAP are m, °C, and mm year^{−1}, respectively.

7.35°C at CA-Let and 10.84°C at HU-Bug; the annual mean net radiation (R_n) ranges from 4.24 mol m^{−2} day^{−1} at AT-Neu to 6.17 mol m^{−2} day^{−1} at CA-Let; the annual mean relative humidity (R_h) varies from 60.40% at CA-Let to 79.80% at AT-Neu. The other statistical characteristics of these variables are also summarized. Furthermore, Table 2 also presents the correlation coefficients between ET and environmental variables. It can be clearly seen that almost all the variables at the four sites have strong correlation with ET. The variables, T_a , T_s , and R_n , show strong positive correlation with ET, while R_h presents negative correlation with ET. Daily available measurements for at least six years were divided into three parts. At each site, the first 4-year dataset was utilized for training, the fifth year dataset was used for validation with the intention of preventing overtraining or overfitting of the training dataset, and the remaining dataset was employed for predicting.

2.2. Data-Driven Techniques

2.2.1. Artificial Neural Network. ANN model performance depends on parallel information processing system involved in a multilayer network, and the model is capable of transforming helpful information from the input data into knowledge through artificial neurons which are directly linked to active synapses with a set of weights related to different variables [32]. Its ultimate goal is to establish the nonlinear relationship between input (independent) variables and output (dependent) variables. In this study, a feed-forward ANN with a single hidden layer was used and its parameters were updated using the backpropagation learning algorithm with the intention of minimizing the error between the target values and the calculated values derived from the network. When designing ANN models in this study, hyperbolic tangent sigmoid and linear algorithms were selected as transfer functions for hidden layer and output layer, respectively, with the purpose of generating output results from each neuron. In addition, the determination of other key parameters and functions, such as the number of nodes of hidden layer and training functions, was also taken into account based on the trial-and-error approach.

2.2.2. Support Vector Machine. The SVM approach is an important soft computing method and has been utilized in many research fields. The detail of theory about SVM can be found in Vapnik [33] and its advances in the practical applications can be found in some recent reviews [33–35]. Based on the accumulated experience, it is noteworthy to

point out that the generalization ability of SVM method considerably depends on the choice of kernel function. There are many alternative kernel algorithms, such as radial basis function (RBF), polynomial, sigmoid, exponential, Laplacian, and rational quadratic. The first three kernel functions are commonly used for SVM models and therefore were compared and assessed in this study. As a matter of convenience, the SVM models with the kernel functions, RBF, polynomial, and sigmoid, were hereinafter referred to as SVM-RBF, SVM-Poly, and SVM-Sig, respectively. For a given case, the critical internal parameters involved in the training process, including the cost factor, the slack factor, and the related parameter of respective kernel algorithm, were carefully obtained through the grid search approach. All the SVM models applied in the present research were performed according to LIBSVM software written by MATLAB programming language [36].

2.2.3. Adaptive Neurofuzzy Inference System. ANFIS has received much attention for modeling and predicting the hydrological time series (e.g., evaporation, groundwater level, and stream-flow) in recent years. Hence, this method has been widely recognized as a promising alternative tool to traditional data-driven techniques for mapping the nonlinear relationship between independent and dependent variables in different research fields. Selecting a reasonable FIS is an important precondition for successfully elucidating the nonlinear processes underlying a given problem, when establishing an ANFIS model. Presently, Takagi and Sugeno [37] method is extensively adopted in diverse research areas, and, in particular, its powerful ability in estimating various hydrological variables has been proved by accumulating lines of evidence to date. Accordingly, Takagi and Sugeno-based ANFIS models were developed by the present work for all the applied cases. To achieve an optimum ANFIS model, two types of parameters, namely, the antecedent parameters which influence the effectiveness of the applied membership function and the consequent parameters which determine the quality of the system output, need to be optimized appropriately. The hybrid learning procedure is the most commonly utilized to adjust the antecedent and consequent parameters using the gradient descent and least squares methods, respectively. In addition, of particular interest is the impacts of different identification algorithms for generating FISs on the ability of developed ANFIS models. The current research also attempts to address this issue. As a convenience, the ANFIS models with three extensively used identification algorithms, namely, grid partitioning, subtractive clustering,

TABLE 2: Daily statistical parameters of flux tower measured environmental variables including air temperature (T_a), net radiation (R_n), relative humidity (R_h), soil temperature (T_s), and evapotranspiration (ET) in the whole period.

Period	Variable	$\bar{X}_{\text{mean}}$	$\bar{X}_{\text{max}}$	$\bar{X}_{\text{min}}$	$\bar{X}_{\text{sd}}$	$\bar{X}_{\text{ku}}$	$\bar{X}_{\text{sk}}$	$\bar{R}$
AT-Neu	T_a	6.68	22.96	-17.63	8.19	2.24	-0.29	0.80
	R_n	4.24	16.51	-6.44	5.33	1.97	0.16	0.88
	R_h	79.80	99.98	44.36	10.98	2.52	-0.46	-0.59
	T_s	8.71	21.28	-2.56	6.86	1.50	0.01	0.78
	ET	1.29	6.78	-0.32	1.33	2.75	0.87	1.00
CA-Let	T_a	6.48	28.57	-31.89	10.37	3.10	-0.58	0.59
	R_n	6.17	19.18	-15.83	5.91	1.93	0.38	0.76
	R_h	60.40	99.04	16.94	16.48	2.18	0.04	-0.19
	T_s	7.35	24.52	-10.06	7.99	1.81	0.18	0.61
	ET	0.81	5.96	-0.50	1.01	6.68	1.92	1.00
DE-Gri	T_a	8.35	27.83	-18.02	7.80	2.37	-0.25	0.76
	R_n	5.08	19.76	-4.25	5.34	2.17	0.55	0.90
	R_h	77.98	100.00	39.22	10.60	2.63	-0.31	-0.63
	T_s	9.68	24.32	-0.84	6.10	1.83	0.03	0.75
	ET	0.96	4.50	-0.61	0.98	3.33	1.04	1.00
HU-Bug	T_a	10.20	30.15	-18.72	8.98	2.28	-0.22	0.72
	R_n	5.81	18.26	-3.90	5.38	1.88	0.39	0.83
	R_h	76.28	99.66	38.88	12.01	2.61	-0.38	-0.46
	T_s	10.84	26.03	-3.11	7.21	1.78	0.05	0.70
	ET	1.28	5.96	-0.06	1.10	3.68	1.08	1.00

Note. The units of T_a , R_n , R_h , T_s , and ET are $^{\circ}\text{C}$, mol m^{-2} , %, $^{\circ}\text{C}$, and mm day^{-1} , respectively. $\bar{X}_{\text{mean}}$, $\bar{X}_{\text{max}}$, $\bar{X}_{\text{min}}$, $\bar{X}_{\text{sd}}$, $\bar{X}_{\text{ku}}$, and $\bar{X}_{\text{sk}}$ refer to the mean, maximum, minimum, standard deviation, kurtosis, and skewness of each variable, respectively. $\bar{R}$ refers to the correlation coefficient between each variable and ET.

and fuzzy c-means clustering, were abbreviated as ANFIS-Grid, ANFIS-SC, and ANFIS-FCM, respectively.

2.2.4. General Regression Neural Network. GRNN proposed by Specht [38] based on nonlinear regression theory is widely utilized for function approximation. It is closely associated with the radial basis function neural network, but there exists a slight distinction between these two networks in terms of topological structure. The GRNN model as a one-pass learning method is designed by using a highly parallel framework. In fact, the number of neurons in all the layers is strongly dependent upon the number of the used variables as well as the sample size of a given training dataset, and thus these neurons do not require adjustment during the learning. In addition to this, another noteworthy advantage of the GRNN approach over other data-driven techniques is that the GRNN model does not need much time to select its internal algorithms or parameters, apart from the smoothing factor that highly affects the model generalization performance. At present, there is no standard, universally accepted method for determining the optimal smoothing factor with the aim of guaranteeing the forecasting precision of the GRNN model. In this study, the proper smoothing factor that was set ranging from 0.01 to 1 was obtained by using an iterative algorithm. Moreover, to prevent the overfitting, fourfold cross validation was utilized in the training. Further details regarding the principle of the GRNN approach can be found in Raghavendra and Deka [34] and Yaseen et al. [39].

2.2.5. Extreme Learning Machine. ELM firstly proposed by Huang et al. [40] is a relatively new, modern data-driven technique and has been successfully applied to deal with the nonlinear problems in diverse areas during the last few years. The developed ELM method has a topological architecture of single-hidden-layer feed-forward neural network. Its number of hidden nodes can be randomly gained, and their relevant parameters (weights and biases) do not need to be tuned as these properties have been demonstrated to be independent of the training dataset. By contrast, for the traditional back-propagation based ANN approach, almost all the parameters are highly problem-dependent and should be carefully set or optimized by a common, time-consuming trial-and-error process. Moreover, the powerful approximation ability of the ELM approach has been strongly supported by several lines of evidence [39, 41]. The fundamental principle of ELM approach can be found in more detail in Huang et al. [40].

Furthermore, it should be emphasized that appropriate activation function is of importance for guaranteeing the modeling and generalization performance of the ELM method [42]. The sigmoid, sine, and hard limit algorithms are the three common types of activation functions applied for the hidden layer. The influences of different activation functions on the modeling ability of developed ELM models were evaluated by this study. For the sake of convenience, the ELM models with the sigmoid, sine, and hard limit algorithms were abbreviated as ELM-Sig, ELM-Sin, and ELM-Hard, respectively. Moreover, the linear activation function was adopted for the output layer for all the developed ELM models.

2.2.6. Group Method of Data Handling. GMDH approach as a typical inductive high-order regression-type algorithm is commonly subsumed under the category of feed-forward neural network [43]. The architecture of this heuristic non-linear method is established automatically based on self-organization learning algorithm, which is significantly different from other machine learning algorithms. Moreover, the GMDH method can conquer the limitation of requiring prior knowledge regarding a given problem, and redundant input variables can be eliminated effectively during the learning. Besides, the GMDH is highly robust to noise data and thus is not affected by the existing outliers in the training samples. In a well-developed GMDH network, the neuron in each hidden layer is individually connected by two inputs and a single output and acts as a transfer function for representing the results generated by the two neurons from the previous layer. Mathematically, each neuron is represented by using a quadratic polynomial equation with five weights and one bias. Ultimately, the well-trained network can be expressed by using an explicit high-order polynomial function in relation to all the neurons reserved in the network. All the weighting coefficients involved in this iterative function are calculated by a common least square method. More details about GMDH method can be found in Ivakhnenko [43] and Barzegar et al. [44].

2.3. Model Structure and Evaluation. It is clear from Table 2 that each environmental variable has high linear correlation with ET. Therefore, these variables (T_a , T_s , R_n , R_h) were used as inputs for all the proposed data-driven models. Before the training of each applied model, the input variables and ET were normalized to a range between 0 and 1. MATLAB software (version 8.2, R2013b) was utilized for the development of the models and the related statistical analysis of estimated results.

In this study, the coefficient of determination (R^2), Nash-Sutcliffe efficiency (NSE), root mean square error (RMSE), and mean absolute error (MAE) were considered as statistical indices for measuring the performance of the developed data-driven models. These statistical indices are, respectively, expressed as below:

$$\begin{aligned}
 R^2 &= \left[\frac{\sum_{i=1}^N (ET_{o,i} - \overline{ET_o})(ET_{m,i} - \overline{ET_m})}{\sqrt{\sum_{i=1}^N (ET_{o,i} - \overline{ET_o})^2 \sum_{i=1}^N (ET_{m,i} - \overline{ET_m})^2}} \right]^2, \\
 NSE &= 1 - \frac{\sum_{i=1}^N (ET_{o,i} - ET_{m,i})^2}{\sum_{i=1}^N (ET_{o,i} - \overline{ET_o})^2}, \\
 RMSE &= \sqrt{\frac{1}{N} \sum_{i=1}^N (ET_{o,i} - ET_{m,i})^2}, \\
 MAE &= \frac{1}{N} \sum_{i=1}^N |ET_{o,i} - ET_{m,i}|,
 \end{aligned} \tag{1}$$

where ET_o and ET_m denote the observed and modeled values of daily ET, respectively; $\overline{ET_o}$ and $\overline{ET_m}$ are the means of observed and modeled values, respectively; N is the number of observed values. Regarding the physical significance of these indicators, the R^2 was used to measure the proportion of total variance in the observed ET that is explained by the modeled ET. For a perfect model, the R^2 is expected to be close to unity. However, this index is limited because it is calculated based on the linear relationships between observed and simulated values and is sensitive to outliers. Alternatively, the NSE represents the level of agreement between observed and simulated ET and can offer useful knowledge about the relative estimation of the generalization performance for a given model. It is sensitive to differences in measured and calculated means and variances and accordingly was employed in the present investigation. Moreover, other valuable information involving the predictive ability of the developed models can also be provided by the RMSE and MAE indicators. Therefore, these performance criteria were employed together to obtain helpful insight into the efficiency of the developed models.

3. Results

The performance indices, including R^2 , NSE, RMSE, and MAE, are employed to evaluate the accuracy of the utilized models in predicting the daily ET in this study. R^2 and IA with larger values and RMSE and MAE with smaller values imply higher model efficiency. Table 3 shows the estimated results of all machine learning models (ANN, GRNN, GMDH, ELM, ANFIS, and SVM) for ET over the training, validation, and prediction periods at AT-Neu site. As can be seen from Table 3, the ANFIS-SC model gives the best performance in the prediction of ET in the prediction period, with the highest values of R^2 (0.9379) and NSE (0.9355) and the lowest value of RMSE (0.3308 mm day⁻¹) and MAE (0.2260 mm day⁻¹). And the ELM-Sin model provides the inferior results in terms of R^2 , NSE, RMSE, and MAE. The SVM-Sig model performs the worst among the twelve models. According to the R^2 , NSE, RMSE, and MAE metrics, the overall performance ranks of these developed models in the prediction period can be summarized as follows: ANFIS-SC, ELM-Sin, ANFIS-FCM, ANN, ELM-Sig, GMDH, SVM-RBF, ANFIS-Grid, GRNN, ELM-Hard, SVM-Poly, and SVM-Sig.

The comparisons of daily ET observed and predicted by using the data-driven models over the prediction period in the form of scatter plot at AT-Neu site are shown in Figure 1. As a convenience, the scattered estimates of the optimal ELM, ANFIS, and SVM models are exclusively compared with those of the ANN, GRNN, and GMDH models. As illustrated in Figure 1, the fit line of the SVM-RBF model is closer to the ideal line (1:1 line) than those of the other models, considering the corresponding equations of fit lines. However, it can be seen from Figure 1 and Table 3 that the ANFIS-SC model seems to have slightly higher R^2 value than the other models and obtains similar scattered estimates to those of the ANN and ELM-Sin models.

The performance of the applied data-driven models for estimating ET over the training, validation, and prediction periods at CA-Let site is shown in Table 4. As can be seen from the table, the ANN model performs the best on the whole, followed by the ANFIS-SC and ANFIS-Grid models. The SVM-Sig model provides the worst accuracy among the twelve models. Based on the R^2 , NSE, RMSE, and MAE metrics, the overall performance of these applied models in the prediction period can be ranked as follows: ANN, ANFIS-SC, ANFIS-Grid, ANFIS-FCM, SVM-RBF, ELM-Sin, ELM-Sig, SVM-Poly, GRNN, GMDH, ELM-Hard, and SVM-Sig. Figure 2 illustrates the measured and predicted ET by the optimal models for each approach in the prediction stage at CA-Let site. As shown in the figure, the ANN, ANFIS-SC, and SVM-RBF models have higher R^2 and lower RMSE than other models. In addition, the fit lines of these three models are closer to their respective ideal lines (1:1 line), in comparison with the other models.

Comparisons of the applied models are made in Table 5 for modeling daily ET at DE-Gri site. Unlike the previous sites, the ELM-Sin model has better accuracy than the other models in terms of the used indices. The ranks of the estimated models in the prediction stage are: ELM-Sin, ELM-Sig, ANN, ANFIS-Grid, GRNN, GMDH, ANFIS-FCM, ANFIS-SC, SVM-RBF, ELM-Hard, SVM-Poly, and SVM-Sig. Figure 3 illustrates the estimates of the optimal models for ET simulation in the prediction phase at DE-Gri site. It is apparent from the figure that the fit line of the ANFIS-Grid model is closer to the ideal line (1:1 line), while the ELM-Sin model has the highest R^2 and provides less scattered estimates than the other models, which has also been confirmed in Table 5.

Table 6 summarizes the accuracy of the applied models for forecasting daily ET in the prediction phase at HU-Bug site. It can be clearly seen from the table that the GMDH model achieves the best precision among the developed twelve models. Based on the utilized performance indicators, the overall model efficiency of these models in the prediction period can be ranked as follows: GMDH, ELM-Sig, ELM-Sin, ANN, ANFIS-SC, SVM-RBF, ANFIS-FCM, ANFIS-Grid, GRNN, SVM-Poly, ELM-Hard, and SVM-Sig. The scatterplot comparisons of the estimated methods with their respective optimal parameters are made in Figure 4 over the prediction stage at HU-Bug site. As shown from the figure, the slope and bias values of the fit line equation of both the ANFIS-SC and ELM-Sig models are closer to those of their corresponding exact lines (1:1 line, slope: 1, and bias: 0), compared with those of the other models. However, the GMDH model seems to be more successful than the other models from the R^2 and RMSE viewpoints.

In order to provide insights into the over- and underestimation of the used data-driven models in predicting daily ET, the measured and modeled values by the best models for each site in the whole period are demonstrated in Figure 5. It is evident from the figure that the modeled values of these best models for each site can closely follow the corresponding observed ones, which is previously confirmed in Tables 3–6. However, the peak values during the growing season in the

TABLE 3: Comparisons of data-driven model performances for evapotranspiration among the training, validation and prediction periods at AT-Neu site.

Model	Training				Validation				Prediction			
	R ²	NSE	RMSE	MAE	R ²	NSE	RMSE	MAE	R ²	NSE	RMSE	MAE
ANN	0.9298	0.9294	0.3551	0.2321	0.9407	0.9375	0.3318	0.2203	0.9355	0.9338	0.3350	0.2278
GRNN	0.9596	0.9595	0.2689	0.1635	0.9134	0.9118	0.3943	0.2640	0.9207	0.9186	0.3714	0.2473
GMDH	0.9271	0.9270	0.3610	0.2442	0.9246	0.9220	0.3707	0.2464	0.9356	0.9320	0.3394	0.2342
ELM-Sig	0.9256	0.9256	0.3644	0.2430	0.9265	0.9249	0.3639	0.2461	0.9355	0.9328	0.3376	0.2322
ELM-Sin	0.9278	0.9278	0.3591	0.2332	0.9302	0.9283	0.3554	0.2415	0.9365	0.9344	0.3335	0.2242
ELM-Hard	0.8762	0.8762	0.4702	0.3365	0.8814	0.8811	0.4577	0.3394	0.9043	0.9031	0.4053	0.2836
ANFIS-Grid	0.9407	0.9407	0.3255	0.2121	0.9237	0.9230	0.3684	0.2585	0.9247	0.9219	0.3639	0.2436
ANFIS-SC	0.9373	0.9373	0.3347	0.2138	0.9318	0.9304	0.3503	0.2402	0.9379	0.9355	0.3308	0.2260
ANFIS-FCM	0.9376	0.9376	0.3338	0.2151	0.9380	0.9358	0.3363	0.2314	0.9369	0.9343	0.3338	0.2282
SVM-RBF	0.9387	0.9373	0.3347	0.1928	0.9333	0.9328	0.3440	0.2357	0.9338	0.9321	0.3392	0.2286
SVM-Poly	0.8357	0.8321	0.5476	0.4007	0.8351	0.8341	0.5406	0.4287	0.8746	0.8696	0.4701	0.3579
SVM-Sig	0.5948	0.5061	0.9391	0.7203	0.5113	0.4352	0.9976	0.7997	0.5910	0.5063	0.9149	0.7275

Note. The units of RMSE and MAE are mm day⁻¹.

FIGURE 1: Comparisons of daily ET between eddy covariance measured and simulated by data-driven models in the prediction period at AT-Neu site. (a) ANN model; (b) GRNN model; (c) GMDH model; (d) ELM-Sin model; (e) ANFIS-SC model; and (f) SVM-RBF model.

prediction period appear to be appreciably underestimated by the optimal models for all the sites, especially for CA-Let site, which is also consistent with the scatter plots in Figures 1–4.

4. Discussion

In the following subsections, we concentrated primarily on discussing the generalization ability of all data-driven models for forecasting daily ET and then on exploring the effects of different internal functions on the ELM, ANFIS, and SVM models. Finally, we also pointed out some potential improvements for the follow-up work.

Our estimates demonstrated that all the examined models, including ANN, GRNN, GMDH, ELM, ANFIS, and SVM, can seize the nonlinear relationship between ET and environmental variables according to the combination of four different performance criteria. Moreover, our modeling results also confirmed that the new computational intelligence techniques (GMDH, ELM, and ANFIS) were capable of effectively acquiring the seasonal and interannual variation in ET driven by the environmental variables. More importantly, these three methods had significant superiority over conventional methods in terms of robustness and simplicity. Specifically, of particular interest regarding the ELM method

is the fact that its learning speed is shown to be extremely faster than that of other data-driven approaches, partly because the number of hidden nodes of an ELM model can be randomly determined. In addition, its related parameters (weights and biases) need not be tuned as these properties are problem-independent and thus are not clearly associated with the applied training dataset.

Furthermore, with the advances in machine learning in recent years, a number of new machine learning methods, such as relevance vector machine [45], M5 model tree [46], and genetic programming [47], have been proposed and successfully applied in other diverse fields, such as forecasting of meteorological time series (e.g., air temperature and precipitation) and prediction of water resource variables (e.g., rainfall–runoff, groundwater level, and drought). Consequently, it is also important to investigate the feasibility and effectiveness of these new approaches in dealing with the present ET estimation, which will be undertaken in our future work.

Our modeling results also showed that an obvious difference existed within each method (ELM, ANFIS, and SVM) with various internal functions in terms of predictive performance. Therefore, selecting appropriate internal function for

TABLE 4: Comparisons of data-driven model performances for evapotranspiration among the training, validation and prediction periods at CA-Let site.

Model	Training				Validation				Prediction			
	R ²	NSE	RMSE	MAE	R ²	NSE	RMSE	MAE	R ²	NSE	RMSE	MAE
ANN	0.7593	0.7590	0.3791	0.2541	0.8553	0.7671	0.5690	0.3546	0.7859	0.7334	0.5911	0.3607
GRNN	0.8417	0.8176	0.3298	0.1997	0.8112	0.6472	0.7002	0.4096	0.7415	0.6307	0.6957	0.3919
GMDH	0.6673	0.6673	0.4454	0.3036	0.8303	0.6887	0.6577	0.4018	0.6931	0.6343	0.6922	0.4134
ELM-Sig	0.7159	0.7159	0.4115	0.2821	0.8038	0.7045	0.6409	0.4057	0.7469	0.6907	0.6367	0.3913
ELM-Sin	0.7163	0.7163	0.4113	0.2805	0.8423	0.7570	0.5811	0.3582	0.7530	0.6998	0.6272	0.3787
ELM-Hard	0.5615	0.5615	0.5113	0.3602	0.6843	0.5572	0.7845	0.4758	0.6013	0.5446	0.7725	0.4733
ANFIS-Grid	0.7820	0.7820	0.3605	0.2418	0.8480	0.7869	0.5443	0.3473	0.7805	0.7380	0.5859	0.3527
ANFIS-SC	0.7691	0.7691	0.3711	0.2495	0.8467	0.7712	0.5639	0.3507	0.7855	0.7311	0.5936	0.3492
ANFIS-FCM	0.7649	0.7649	0.3744	0.2532	0.8628	0.7879	0.5429	0.3405	0.7805	0.7287	0.5962	0.3554
SVM-RBF	0.7936	0.7912	0.3528	0.2111	0.8453	0.7668	0.5693	0.3418	0.7762	0.7225	0.6030	0.3577
SVM-Poly	0.6888	0.6772	0.4387	0.2888	0.8489	0.7824	0.5500	0.3326	0.7528	0.7053	0.6214	0.3676
SVM-Sig	0.0244	-0.0790	290.02	241.69	0.0616	-0.1655	263.96	225.25	0.0226	-0.0104	298.60	248.78

Notes. The unit of RMSE and MAE is mm day⁻¹.

FIGURE 2: Comparisons of daily ET between eddy covariance measured and simulated by data-driven models in the prediction period at CA-Let site. (a) ANN model; (b) GRNN model; (c) GMDH model; (d) ELM-Sin model; (e) ANFIS-SC model; and (f) SVM-RBF model.

each method is essential with the objective of achieving the best modeling ability. It is highly recommended that different internal functions for each method should be evaluated in advance, when these methods are used for regression and classification problems. More specifically in the present study, for ANFIS method, the evaluated three algorithms for generating FISs had noticeable effects on the performance of ANFIS on the whole. We found that, among the three various ANFIS models, the ANFIS model with the subtractive clustering algorithm performed the best at AT-Neu, CA-Let, and HU-Bug sites, while grid partitioning based ANFIS model generated the optimal estimates at DE-Gri site. Therefore, there was no omnipotent algorithm that was appropriate for all the cases. In addition, Cobaner [48] investigated the ability of two different ANFIS methods, respectively, based on the subtractive clustering and grid partitioning algorithms, in estimating reference ET using daily climatic data, and found that subtractive clustering based ANFIS model achieved more plausible precision with fewer amounts of computation in comparison with grid partitioning based ANFIS model. Moreover, to the best of our knowledge, most studies comparing the performance of different types

of ANFIS techniques in other different fields have concentrated mainly on evaluating the roles of the membership functions under grid partitioning algorithm. It should be mentioned that selecting appropriate membership function and determining its optimal number may effectively improve the performance of grid partitioning based ANFIS model. In conclusion, our study has broadened the scope of ANFIS research and provided deep insights into the application of ANFIS method. In the follow-up work, we will investigate together the capability of the aforementioned algorithms (subtractive clustering, grid partitioning, and fuzzy *c*-means clustering), as well as different types of membership functions under grid partitioning algorithm.

According to the predictive capability and efficiency of SVM, it was found from our results that remarkable difference existed among three different kernel functions (sigmoid, polynomial, and RBF), which has been also confirmed by the previous studies in other fields [49–51]. Moreover, our estimates indicated that the RBF kernel function for SVM method performed better than the other two kernel functions (sigmoid and polynomial) in the prediction of daily ET. It concurred with the findings from numerous

TABLE 5: Comparisons of data-driven model performances for evapotranspiration among the training, validation, and prediction periods at DE-Gri site.

Model	Training					Validation					Prediction				
	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE
ANN	0.8633	0.8630	0.3470	0.2426	0.9341	0.8927	0.3411	0.2336	0.9623	0.9341	0.2769	0.1942			
GRNN	0.8612	0.8603	0.3504	0.2487	0.9294	0.8814	0.3587	0.2478	0.9585	0.9292	0.2871	0.2058			
GMDH	0.8396	0.8396	0.3755	0.2707	0.9326	0.8724	0.3719	0.2573	0.9619	0.9253	0.2949	0.2121			
ELM-Sig	0.8535	0.8535	0.3589	0.2580	0.9358	0.8885	0.3478	0.2347	0.9669	0.9392	0.2659	0.1920			
ELM-Sin	0.8546	0.8546	0.3575	0.2522	0.9477	0.9066	0.3182	0.2203	0.9711	0.9398	0.2646	0.1896			
ELM-Hard	0.8012	0.8012	0.4180	0.2911	0.8926	0.8724	0.3720	0.2662	0.9053	0.8765	0.3791	0.2534			
ANFIS-Grid	0.8772	0.8772	0.3285	0.2304	0.9173	0.8713	0.3735	0.2447	0.9574	0.9334	0.2785	0.1961			
ANFIS-SC	0.8637	0.8637	0.3461	0.2427	0.9230	0.8716	0.3732	0.2397	0.9501	0.9088	0.3258	0.2167			
ANFIS-FCM	0.8636	0.8636	0.3462	0.2439	0.9244	0.8744	0.3691	0.2378	0.9500	0.9099	0.3238	0.2159			
SVM-RBF	0.8638	0.8633	0.3466	0.2306	0.9196	0.8796	0.3614	0.2332	0.9407	0.9002	0.3408	0.2092			
SVM-Poly	0.7885	0.7852	0.4345	0.3221	0.7629	0.7043	0.5662	0.3879	0.8461	0.8199	0.4578	0.3407			
SVM-Sig	0.0170	-0.0682	43.738	37.467	0.0001	-0.3639	37.708	31.554	0.1190	-0.2045	45.388	39.515			

Note. The unit of RMSE and MAE is mm day⁻¹.

FIGURE 3: Comparisons of daily ET between eddy covariance measured and simulated by data-driven models in the prediction period at DE-Gri site. (a) ANN model; (b) GRNN model; (c) GMDH model; (d) ELM-Sin model; (e) ANFIS-Grid model; and (f) SVM-RBF model.

studies, showing the superiority of RBF over the other kernel functions in solving the regression problems [52–54]. For example, Zounemat-Kermani et al. [55] evaluated the capability of SVM model with four different kernel functions (linear, polynomial, sigmoid, and RBF) for forecasting daily suspended sediment concentrations and further pointed out that the RBF for SVM model was the best choice for modeling hydrological phenomena. Mohammadi et al. [56] investigated the ability of two different types of SVM models based on polynomial and RBF kernel functions in forecasting the horizontal global solar radiation and found that RBF for SVM was highly competent for predicting daily horizontal global solar radiation in comparison with polynomial function.

As a relatively new method, ELM exhibited strong modeling accuracy in predicting daily ET, which has been verified by previous studies for other applications, such as reference ET prediction [57] and stream-flow forecasting [39]. In particular, a very noteworthy aspect was that the ELM method presented a remarkable advantage against other data-driven approaches with respect to computational time and efficiency due to its simple network structure and nontuned mechanism [58], which is a beneficial contribution to the solution of some real-time forecasting problems,

such as rainfall and runoff time series forecasting. In brief, the ELM technique can be a promising alternative tool to traditional methods for dealing with the regression issue in the current research as well as the gap-filling problem of ET flux measured by the eddy covariance technique. On the other hand, relatively less attention has been drawn toward exploring the effects of various activation functions (e.g., sine, sigmoid, and hard limit activation function) in the hidden layer on the generalization ability of ELM. In the present study, our results demonstrated that the sine and sigmoid activation functions for ELM models played the similar roles in estimating the daily ET and dramatically outperformed the hard limit activation function. Therefore, it is important to emphasize that the sine and sigmoid activation functions are recommended as the optimal options for establishing the ELM models for ET forecasting.

Furthermore, many recent studies have found that selecting an appropriate training function is vitally essential for assuring the predictive ability and reliability of ANN in other fields [59–61]. Despite the fact that ANN has been recognized as the most popular method to simulate the ET as well as the carbon fluxes at ecosystem level based on the eddy covariance-measured data [62–64], the influences of various

TABLE 6: Comparisons of data-driven model performances for evapotranspiration among the training, validation, and prediction periods at HU-Bug site.

Model	Training					Validation					Prediction				
	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE	R^2	NSE	RMSE	MAE			
ANN	0.8765	0.8764	0.3955	0.2636	0.8038	0.8023	0.4474	0.3022	0.7969	0.7907	0.5018	0.3635			
GRNN	0.9043	0.9041	0.3483	0.2284	0.7412	0.7333	0.5196	0.3581	0.7519	0.7395	0.5598	0.3895			
GMDH	0.8135	0.8135	0.4858	0.3303	0.7773	0.7770	0.4752	0.3315	0.8165	0.8143	0.4727	0.3272			
ELM-Sig	0.8467	0.8467	0.4405	0.3030	0.7625	0.7563	0.4967	0.3566	0.8155	0.8087	0.4797	0.3402			
ELM-Sin	0.8521	0.8521	0.4327	0.2910	0.7823	0.7807	0.4712	0.3352	0.8134	0.8075	0.4813	0.3449			
ELM-Hard	0.7113	0.7113	0.6045	0.4539	0.6425	0.6291	0.6128	0.4577	0.7045	0.7017	0.5991	0.4478			
ANFIS-Grid	0.8921	0.8921	0.3695	0.2556	0.7178	0.7008	0.5504	0.3605	0.7595	0.7414	0.5578	0.3774			
ANFIS-SC	0.8721	0.8721	0.4024	0.2705	0.7761	0.7713	0.4811	0.3407	0.7893	0.7784	0.5164	0.3752			
ANFIS-FCM	0.8679	0.8679	0.4089	0.2760	0.7709	0.7655	0.4873	0.3376	0.7726	0.7636	0.5333	0.3801			
SVM-RBF	0.8917	0.8897	0.3737	0.2323	0.7529	0.7511	0.5020	0.3404	0.7699	0.7671	0.5294	0.3739			
SVM-Poly	0.7451	0.7392	0.5745	0.3901	0.7344	0.7286	0.5241	0.3737	0.7266	0.7237	0.5765	0.4114			
SVM-Sig	0.0049	-0.0610	218.15	178.95	0.0193	-0.0991	201.38	168.64	0.0090	-0.0053	207.32	171.67			

Note. The unit of RMSE and MAE is mm day⁻¹.

FIGURE 4: Comparisons of daily ET between eddy covariance measured and simulated by data-driven models in the prediction period at HU-Bug site. (a) ANN model; (b) GRNN model; (c) GMDH model; (d) ELM-Sig model; (e) ANFIS-SC model; and (f) SVM-RBF model.

training algorithms on the ANN performance have never been examined to date. Consequently, special attention will be given to the evaluation of the modeling abilities of ANN models with a variety of training algorithms for predicting ET and carbon fluxes in our follow-up work.

5. Conclusions

In recent years, many attempts mainly involving the use of soft computing modeling approaches have been made to simulate and forecast the ET in terrestrial ecosystems. However, the lack of comprehensive comparative researches related to these state-of-the-art modeling techniques largely hinders their applicability and popularity, primarily owing to the confusion in what approach should be appropriately chosen from a variety of data-driven methods in the practical applications. To overcome this hindrance, the current research first attempted to investigate the suitability and effectiveness of three newly proposed methods, GMDH, ELM, and ANFIS, for estimating daily ET at four different grassland sites based on the eddy covariance-measured data. In addition to these techniques, three traditional soft computing techniques, including ANN, GRNN, and SVM, were also employed as

benchmarks in order to compare their capability. Moreover, this study also focused on investigating the influences of internal functions on their corresponding models (ELM, ANFIS, and SVM) in terms of the generalization performance on the basis of a set of statistical indices (R^2 , NSE, RMSE, and MAE).

To summarize, the primary findings in the present study can be enumerated as follows:

- (1) It has been discovered that all the models developed in this study were capable of mapping the nonlinear processes of governing the variation of the ET between the biosphere and the atmosphere at the ecosystem level, and these novel models (GMDH, ELM, and ANFIS) produced estimates comparable to those of the conventional models.
- (2) Considering the robustness and simplicity, these novel approaches can be used as promising alternatives to traditional methods for modeling and forecasting daily ET.
- (3) The predictive accuracy of the SVM, ELM, and ANFIS models was strongly dependent on their respective

FIGURE 5: Eddy covariance measured and simulated daily ET by their respective best models for the four sites in the whole period. (a) For AT-Neu site using ANFIS-SC model; (b) for CA-Let site using ANN model; (c) for DE-Gri site using ELM-Sin model; and (d) for HU-Bug site using GMDH model.

internal functions, especially for SVM. Three different kernel functions for the SVM method were together tested and the results suggested that the RBF kernel function substantially outperformed both the polynomial and sigmoid kernel functions.

- (4) The ELM models with the sigmoidal and sine activation functions generated the similar modeling accuracy and were appreciably superior to the ELM model with the hard limit function.
- (5) For the ANFIS method, the algorithms for generating FISs had noticeable effects on the performance of ANFIS method. The optimal algorithm can be determined according to a trial-and-error procedure.

It should be noted that all the data-driven modeling techniques strongly driven by a large amount of data are often argued because the intrinsic mechanisms of these well-trained models are still not able to be represented explicitly. Accordingly, this argument is very likely to decrease the credibility of these techniques and further impede their applications. Notwithstanding this limitation, our present investigation does suggest that these data-driven methods based on soft computing can effectively complement physically based models, broaden the horizon of ecological, climatological, and hydrologic researchers, and therefore contribute to the estimates of regional and global water resources under climate change.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was financially supported by the Natural Science Fund of China (no. 41672324) and the Priority Academic Program Development of Jiangsu Higher Education Institutions. The present research used the eddy covariance data acquired and shared by the CarboEurope and Fluxnet-Canada Research Networks. The authors would like to thank the project investigators of the used site for providing the flux data.

References

- [1] P. Gong, J. Wang, L. Yu et al., "Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data," *International Journal of Remote Sensing*, vol. 34, no. 7, pp. 2607–2654, 2013.
- [2] A. Trepekli, G. Loupa, and S. Rapsomanikis, "Seasonal evapotranspiration, energy fluxes and turbulence variance characteristics of a Mediterranean coastal grassland," *Agricultural and Forest Meteorology*, vol. 226–227, pp. 13–27, 2016.
- [3] D. R. Bowling, S. Bethers-Marchetti, C. K. Lunch, E. E. Grote, and J. Belnap, "Carbon, water, and energy fluxes in a semiarid cold desert grassland during and following multiyear drought," *Journal of Geophysical Research: Biogeosciences*, vol. 115, no. 4, Article ID G04026, 2010.
- [4] P. Wagle, X. Xiao, R. L. Scott et al., "Biophysical controls on carbon and water vapor fluxes across a grassland climatic gradient in the United States," *Agricultural and Forest Meteorology*, vol. 214–215, pp. 293–305, 2015.
- [5] J. Storkey, A. J. Macdonald, P. R. Poulton et al., "Grassland biodiversity bounces back from long-term nitrogen addition," *Nature*, vol. 528, no. 7582, pp. 401–404, 2015.
- [6] L. M. Aires, C. A. Pio, and J. S. Pereira, "The effect of drought on energy and water vapour exchange above a mediterranean C3/C4 grassland in Southern Portugal," *Agricultural and Forest Meteorology*, vol. 148, no. 4, pp. 565–579, 2008.
- [7] D. Baldocchi, E. Falge, L. Gu et al., "FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities," *Bulletin of the American Meteorological Society*, vol. 82, no. 11, pp. 2415–2434, 2001.
- [8] A. J. Oliphant, "Terrestrial Ecosystem-Atmosphere Exchange of CO₂, Water and Energy from FLUXNET; Review and Meta-Analysis of a Global in-situ Observatory," *Geography Compass*, vol. 6, no. 12, pp. 689–705, 2012.
- [9] D. Baldocchi, "Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - the state and future of the eddy covariance method," *GCB Bioenergy*, vol. 20, no. 12, pp. 3600–3609, 2014.
- [10] K. Zhang, J. S. Kimball, and S. W. Running, "A review of remote sensing based actual evapotranspiration estimation," *Wiley Interdisciplinary Reviews: Water*, vol. 3, no. 6, pp. 834–853, 2016.
- [11] E. P. Glenn, P. L. Nagler, and A. R. Huete, "Vegetation Index Methods for Estimating Evapotranspiration by Remote Sensing," *Surveys in Geophysics*, vol. 31, no. 6, pp. 531–555, 2010.
- [12] A. Polhamus, J. B. Fisher, and K. P. Tu, "What controls the error structure in evapotranspiration models?" *Agricultural and Forest Meteorology*, vol. 169, pp. 12–24, 2013.
- [13] R. K. Vinukollu, E. F. Wood, C. R. Ferguson, and J. B. Fisher, "Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches," *Remote Sensing of Environment*, vol. 115, no. 3, pp. 801–823, 2011.
- [14] Y. Chen, J. Xia, S. Liang et al., "Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China," *Remote Sensing of Environment*, vol. 140, pp. 279–293, 2014.
- [15] F. Evrendilek, "Assessing neural networks with wavelet denoising and regression models in predicting diel dynamics of eddy covariance-measured latent and sensible heat fluxes and evapotranspiration," *Neural Computing and Applications*, vol. 24, no. 2, pp. 327–337, 2014.
- [16] Z. Chen, R. Shi, and S. Zhang, "An artificial neural network approach to estimate evapotranspiration from remote sensing and AmeriFlux data," *Frontiers of Earth Science*, vol. 7, no. 1, pp. 103–111, 2013.
- [17] F. Yang, M. A. White, A. R. Michaelis et al., "Prediction of continental-scale evapotranspiration by combining MODIS and AmeriFlux data through support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3452–3461, 2006.
- [18] M. Gocic, D. Petković, S. Shamshirband, and A. Kamsin, "Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine," *Computers and Electronics in Agriculture*, vol. 127, pp. 56–63, 2016.
- [19] S. Traore, Y.-M. Wang, and T. Kerh, "Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone," *Agricultural Water Management*, vol. 97, no. 5, pp. 707–714, 2010.
- [20] V. Z. Antonopoulos, S. K. Gianniou, and A. V. Antonopoulos, "Artificial neural networks and empirical equations to estimate daily evaporation: application to Lake Vegoritis, Greece," *Hydrological Sciences Journal*, vol. 61, no. 14, pp. 2590–2599, 2016.
- [21] A. Guven and Ö. Kişi, "Daily pan evaporation modeling using linear genetic programming technique," *Irrigation Science*, vol. 29, no. 2, pp. 135–145, 2011.
- [22] V. Moosavi, A. Talebi, M. H. Mokhtari, and M. R. Hadian, "Estimation of spatially enhanced soil moisture combining remote sensing and artificial intelligence approaches," *International Journal of Remote Sensing*, vol. 37, no. 23, pp. 5605–5631, 2016.
- [23] A. Elshorbagy and I. El-Baroudy, "Investigating the capabilities of evolutionary data-driven techniques using the challenging estimation of soil moisture content," *Journal of Hydroinformatics*, vol. 11, no. 3–4, pp. 237–251, 2009.
- [24] X. Dou and Y. Yang, "Estimating forest carbon fluxes using four different data-driven techniques based on long-term eddy covariance measurements: Model comparison and evaluation," *Science of the Total Environment*, vol. 627, pp. 78–94, 2018.
- [25] J. A. Awan and D.-H. Bae, "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*, vol. 36, no. 15, pp. 4767–4777, 2016.
- [26] V. Moosavi, M. Vafakhah, B. Shirmohammadi, and N. Behnia, "A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods," *Water Resources Management*, vol. 27, no. 5, pp. 1301–1321, 2013.

- [27] O. Kisi, J. Shiri, and B. Nikoofar, "Forecasting daily lake levels using artificial intelligence approaches," *Computers & Geosciences*, vol. 41, pp. 169–180, 2012.
- [28] G. Wohlfahrt, M. Anderson-Dunn, M. Bahn et al., "Biotic, abiotic, and management controls on the net ecosystem CO₂ exchange of European mountain grassland ecosystems," *Ecosystems*, vol. 11, no. 8, pp. 1338–1351, 2008.
- [29] L. B. Flanagan, L. A. Wever, and P. J. Carlson, "Seasonal and interannual variation in carbon dioxide exchange and carbon balance in a northern temperate grassland," *GCB Bioenergy*, vol. 8, no. 7, pp. 599–615, 2002.
- [30] M. Z. Hussain, T. Grünwald, J. D. Tenhunen, Y. L. Li, H. Mirzae, and T. Grünwald, "Summer drought influence on CO₂ and water fluxes of extensively managed grassland in," *Germany, Agriculture, Ecosystems Environment*, vol. 141, no. 1-2, pp. 67–76, 2011.
- [31] N. Carvalhais, M. Reichstein, G. J. Collatz et al., "Deciphering the components of regional net ecosystem fluxes following a bottom-up approach for the Iberian Peninsula," *Biogeosciences*, vol. 7, no. 11, pp. 3707–3729, 2010.
- [32] H. R. Maier and G. C. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling & Software*, vol. 15, no. 1, pp. 101–124, 2000.
- [33] A. Çevik, A. E. Kurtoglu, M. Bilgehan, M. E. Gülşan, and H. M. Albegmpri, "Support vector machines in structural engineering: A review," *Journal of Civil Engineering and Management*, vol. 21, no. 3, pp. 261–281, 2015.
- [34] S. Raghavendra and P. C. Deka, "Support vector machine applications in the field of hydrology: a review," *Applied Soft Computing*, vol. 19, pp. 372–386, 2014.
- [35] U. Okkan and Z. A. Serbes, "Rainfall-runoff modeling using least squares support vector machines," *Environmetrics*, vol. 23, no. 6, pp. 549–564, 2012.
- [36] C. Chang and C. Lin, "LIBSVM: a Library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [37] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.
- [38] D. F. Specht, "A general regression neural network," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 2, no. 6, pp. 568–576, 1991.
- [39] Z. M. Yaseen, O. Jaafar, R. C. Deo et al., "Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq," *Journal of Hydrology*, vol. 542, pp. 603–614, 2016.
- [40] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Networks*, vol. 61, pp. 32–48, 2015.
- [41] Y. Feng, N. Cui, L. Zhao, X. Hu, and D. Gong, "Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China," *Journal of Hydrology*, vol. 536, pp. 376–383, 2016.
- [42] R. C. Deo, M. K. Tiwari, J. F. Adamowski, and J. M. Quilty, "Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model," *Stochastic Environmental Research and Risk Assessment*, vol. 31, no. 5, pp. 1211–1240, 2017.
- [43] A. G. Ivakhnenko, "Polynomial Theory of Complex Systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 1, no. 4, pp. 364–378, 1971.
- [44] R. Barzegar, E. Fijani, A. Asghari Moghaddam, and E. Tziritis, "Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models," *Science of the Total Environment*, vol. 599–600, pp. 20–31, 2017.
- [45] A. F. Torres, W. R. Walker, and M. McKee, "Forecasting daily potential evapotranspiration using machine learning and limited climatic data," *Agricultural Water Management*, vol. 98, no. 4, pp. 553–562, 2011.
- [46] M. T. Sattari, M. Pal, H. Apaydin, and F. Ozturk, "M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey," *Water Resources*, vol. 40, no. 3, pp. 233–242, 2013.
- [47] J. Shiri and Ö. Kişi, "Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations," *Computers & Geosciences*, vol. 37, no. 10, pp. 1692–1701, 2011.
- [48] M. Cobaner, "Evapotranspiration estimation by two different neuro-fuzzy inference systems," *Journal of Hydrology*, vol. 398, no. 3-4, pp. 292–302, 2011.
- [49] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, no. 1, pp. 113–126, 2004.
- [50] V. Moosavi and Y. Niazi, "Development of hybrid wavelet packet-statistical models (WP-SM) for landslide susceptibility mapping," *Landslides*, vol. 13, no. 1, pp. 97–114, 2016.
- [51] L. Olatomiwa, S. Mekhilef, S. Shamshirband, and D. Petkovic, "Potential of support vector regression for solar radiation prediction in Nigeria," *Natural Hazards*, vol. 77, no. 2, pp. 1055–1068, 2015.
- [52] E. Omrani, B. Khoshnevisan, S. Shamshirband, H. Saboohi, N. B. Anuar, and M. H. N. M. Nasir, "Potential of radial basis function-based support vector regression for apple disease detection," *Measurement*, vol. 55, pp. 512–519, 2014.
- [53] S. Shamshirband, D. Petković, H. Saboohi et al., "Wind turbine power coefficient estimation by soft computing methodologies: Comparative study," *Energy Conversion and Management*, vol. 81, pp. 520–526, 2014.
- [54] M. Yousefi, B. Khoshnevisan, S. Shamshirband et al., "Support vector regression methodology for prediction of output energy in rice production," *Stochastic Environmental Research and Risk Assessment*, vol. 29, no. 8, pp. 2115–2126, 2015.
- [55] M. Zounemat-Kermani, Ö. Kişi, J. Adamowski, and A. Ramezani-Charmahineh, "Evaluation of data driven models for river suspended sediment concentration modeling," *Journal of Hydrology*, vol. 535, pp. 457–472, 2016.
- [56] K. Mohammadi, S. Shamshirband, M. H. Anisi, K. A. Alam, and D. Petković, "Support vector regression based prediction of global solar radiation on a horizontal surface," *Energy Conversion and Management*, vol. 91, pp. 433–441, 2015.
- [57] S. S. Abdullah, M. A. Malek, N. S. Abdullah, O. Kisi, and K. S. Yap, "Extreme Learning Machines: A new approach for prediction of reference evapotranspiration," *Journal of Hydrology*, vol. 527, pp. 184–195, 2015.
- [58] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [59] O. Kisi and H. Sanikhani, "Prediction of long-term monthly precipitation using several soft computing methods without climatic data," *International Journal of Climatology*, vol. 35, no. 14, pp. 4139–4150, 2015.
- [60] G. Tezel and M. Buyukyildiz, "Monthly evaporation forecasting using artificial neural networks and support vector machines," *Theoretical and Applied Climatology*, 2015.

- [61] S. Shirin Manesh, H. Ahani, and M. Rezaeian-Zadeh, "ANN-based mapping of monthly reference crop evapotranspiration by using altitude, latitude and longitude data in Fars province, Iran," *Environment, Development and Sustainability*, vol. 16, no. 1, pp. 103–122, 2014.
- [62] T. F. Keenan, E. Davidson, A. M. Moffat, W. Munger, and A. D. Richardson, "Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling," *GCB Bioenergy*, vol. 18, no. 8, pp. 2555–2569, 2012.
- [63] D. Papale and R. Valentini, "A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization," *GCB Bioenergy*, vol. 9, no. 4, pp. 525–535, 2003.
- [64] X. Dou, B. Chen, T. A. Black, R. S. Jassal, and M. Che, "Impact of nitrogen fertilization on forest carbon sequestration and water loss in a chronosequence of three douglas-fir stands in the Pacific Northwest," *Forests*, vol. 6, no. 6, pp. 1897–1921, 2015.