

Research Article

Integration of Multiple Models with Hybrid Artificial Neural Network-Genetic Algorithm for Soil Cation-Exchange Capacity Prediction

Mahmood Shahabi,¹ Mohammad Ali Ghorbani,¹ Sujay Raghavendra Naganna ,² Sungwon Kim,³ Sinan Jasim Hadi ,⁴ Samed Inyurt,⁵ Aitazaz Ahsan Farooque,^{6,7} and Zaher Mundher Yaseen ^{8,9,10}

¹Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

²Department of Civil Engineering, Siddaganga Institute of Technology, Tumakuru 572103, Karnataka, India

³Department of Railroad Construction and Safety Engineering, Dongyang University, Yeongju 36040, Republic of Korea

⁴Department of Real Estate Development and Management, Faculty of Applied Sciences, Ankara University, Ankara, Turkey

⁵Faculty of Engineering and Architecture, Department of Geomatics Engineering, Tokat Gaziosmanpaşa University, Tokat, Turkey

⁶Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE C1A4P3, Canada

⁷School of Climate Change and Adaptation, University of Prince Edward Island, Charlottetown, PE C1A4P3, Canada

⁸New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq

⁹Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia

¹⁰Adjunct Research Fellow, USQ's Advanced Data Analytics Research Group, School of Mathematics Physics and Computing, University of Southern Queensland, Toowoomba, QLD 4350, Australia

Correspondence should be addressed to Zaher Mundher Yaseen; yaseen@alayen.edu.iq

Received 20 February 2022; Revised 23 April 2022; Accepted 5 May 2022; Published 13 June 2022

Academic Editor: Gonzalo Farias

Copyright © 2022 Mahmood Shahabi et al. 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.

The potential of the soil to hold plant nutrients is governed by the cation-exchange capacity (CEC) of any soil. Estimating soil CEC aids in conventional soil management practices to replenish the soil solution that supports plant growth. In this study, a multiple model integration scheme supervised with a hybrid genetic algorithm-neural network (MM-GANN) was developed and employed to predict the accuracy of soil CEC in Tabriz plain, an arid region of Iran. The standalone models (i.e., artificial neural network (ANN) and extreme learning machine (ELM)) were implemented for incorporation into the MM-GANN. In addition, it was tested to enhance the prediction accuracy of the standalone models. The soil parameters such as clay, silt, pH, carbonate calcium equivalent (CCE), and soil organic matter (OM) were used as model inputs to predict soil CEC. With the use of several evaluation criteria, the results showed that the MM-GANN model involving the predictions of ELM and ANN models calibrated by considering all the soil parameters (e.g., Clay, OM, pH, silt, and CCE) as inputs provided superior soil CEC estimates with a Nash Sutcliffe Efficiency (NSE) = 0.87, Root Mean Square Error (RMSE) = 2.885, Mean Absolute Error (MAE) = 2.249, Mean Absolute Percentage Error (MAPE) = 12.072, and coefficient of determination (R^2) = 0.884. The proposed MM-GANN model is a reliable intelligence-based approach for the assessment of soil quality parameters intended for sustainability and management prospects.

1. Introduction

Cation-exchange capacity (CEC) refers to the extent of soil's capacity to preserve exchangeable cations, the like of which have a direct bearing on the soil fertility triangle [1]. Soil CEC is a sensitive indicator of natural and human-induced perturbations over soil profile and groundwater [2]. Monitoring changes in soil CEC can assist in predicting whether soil quality has degraded, improved, or sustained under diverse agricultural or forestry schemes. In the course of conventional soil management practices to replenish the soil solution that supports plant growth, the negatively charged clay particles and organic substances adsorb and hold on positively charged soil nutrients (e.g., NH_4^+ , K^+ , Mg^{2+} , Ca^{2+} , etc.) via electrostatic forces [3, 4]. The preferential adsorption of cations is as per the sequence: $\text{Al}^{3+} > \text{Ca}^{2+} > \text{Mg}^{2+} > \text{K}^+ = \text{NH}_4^+ > \text{Na}^+$ [5]. Depending on the soil structure, CEC clearly demonstrates the shrink-swell potential of any soil; a high CEC value (>40 meq/100 g) denotes that a soil structure will recuperate gradually and sometimes can show expansive behavior. In contrast, a soil with low CEC value (<10 meq/100 g) will have a reduced capacity to hold water and end up being acidic rapidly [6]. Soil CEC can fluctuate according to clay percentage, soil pH, ionic strength, soil-to-solution ratio, clay type, and changing organic matter composition. It is sometimes affected by the redistribution of cations (exchange kinetics) in the soil attributed to soil solution buffering and solute transport. CEC also enables the categorization of certain soils including oxisols, vertisols, alfisols, mollisols, and ultisols [7]. In general, the organic matter enriches soils and, usually, clays (except kaolinite) have a high CEC, while sands have no CEC. For agriculture, the preferred value of CEC is >10 meq/100 g for exchange between plant root hairs and soils [8]. The leaching of contaminants into the underlying aquifer system is usually affected by CEC and percent base saturation which are eloquent indices of soil fertility and nutrient retention capacity. In areas of intensive irrigation, the continuous use of inorganic fertilizers (in excess) inundates the soil profile with more nutrients and thereby flushes a plume of contaminants into the groundwater [9]. Hence, in the early stages of agriculture, it is necessary to estimate CEC for determining the supplemental nutrient needs or to remove excess salts which influence soil structure and agricultural productivity. Soil CEC is a sensitive indicator of natural and human-induced perturbations over soil profile and groundwater. Monitoring changes in soil CEC can assist in predicting whether soil quality has degraded, improved, or sustained under diverse agricultural or forestry schemes [10].

Various methods for direct measurement of soil CEC have been reported and extensively discussed in the literature [11–13]. Multiple comparison of CEC estimation techniques is presented by Conradie and Kotze [14]. In addition, there exist several ancillary approaches such as pedotransfer functions (PTF) for estimating CEC based on easily measured soil's physical and chemical properties [15–18]. Several other researchers conducted studies on the functional relationships between CEC, water retention, and

particle-size distribution. Lambooy investigated the influence of CEC on the water retention characteristics of soils [19]. Implementing multiple regressions, Parfitt et al. estimated CEC by taking into account soil organic carbon and clay content [20]. Krogh et al. modeled CEC rates of Danish soils by using clay and organic matter content as input variables through multiple linear regression analysis [21]. The actual CEC of agricultural soils was found to be directly related to the estimated charge of clay and organic carbon in the soil mass at the actual pH [22]. Using soil organic matter and noncarbonate clay contents as predictors, Seybold et al. explained the variation in CEC for several soil horizons based on soil pH, mineralogy class, taxonomic family, and CEC-activity class [23]. Fooladmand derived PTFs using multiple linear regression between CEC and soil textural data including sand content, clay content, geometric mean particle size diameter, the soil particle-size distribution, and soil organic matter content [24]. Several PTFs relating soil CEC with soil's sand, silt or clay fractions, and soil organic carbon content were evaluated by Khodaverdiloo et al. taking into account calibration dataset size on the prediction accuracy of soil CEC [25]. These classical pedotransfer function-based approaches often suffer from a high degree of inaccuracy due to spatial scale dependence, nonlinear relationships among variables, and incompetence to handle mixed data [26]. Hence, the motivation of the current state of the art is directed toward a new research era where more intelligent models should be explored in this field.

Recent research studies have focused on improving the estimation accuracy of soil CEC by means of artificial intelligence (AI) techniques. Artificial neural network (ANN) based PTFs have become popular to predict/estimate soil CEC of different soil types under diverse climatic zones [27–31]. Kalkhajeh et al. conducted the accurate prediction of soil CEC using different data-driven models [32]. They compared the performance of multiple linear regressions (MLR), adaptive neurofuzzy inference system (ANFIS), multilayer perceptron (MLP), and radial basis function (RBF) based ANN models for predicting the soil CEC using the bulk density, calcium carbonate, organic carbon, clay, and silt content (%) of the soil as input variables. The MLP model gave the most reliable prediction of soil CEC. A set of AI models along with empirical PTFs were developed and evaluated by Ghorbani et al. [33]; the authors found the most influential soil properties that influence soil CEC through sensitivity analysis. The ANFIS model provided the superior performance to RBF, MLP, MLR, and empirical PTFs while estimating soil CEC. Arthur [2] presented an ANN based methodology for estimating CEC from soil water content at different relative humidity ranges. Relatively few studies utilize a support vector machine (SVM), random forests (RF), genetic expression programming (GEP), multivariate adaptive regression splines (MARS), and a subtractive clustering algorithm based ANFIS for estimating soil CEC using readily measured soil properties as inputs [34–38]. A hybrid model integrating ant colony optimization (ACO) algorithm with ANFIS improved the prediction accuracy of soil CEC accompanied by an optimal choice of input subset which comprised soil properties (e.g., soil organic matter,

clay, silt, pH, and bulk density) [39]. Although there has been noticeable progress in AI implementation in the field of geoscience, the enthusiasm for developing and exploring more reliable intelligent predictive models is still an ongoing research era. In addition, the applications of hybrid AI models have been observed remarkably reported in the literature and for diverse engineering and sciences domains [40–43]. As a result, the inspiration for developing multiple learning intelligent models is investigated here for modeling soil CEC.

Soil CEC is a sensitive indicator of natural and human-induced perturbations over soil profile and groundwater. Monitoring changes in soil CEC can assist in predicting whether soil quality has degraded, improved, or sustained under diverse agricultural or forestry schemes. Hybrid soft computing approaches involving evolutionary algorithms coupled with AI techniques facilitate the development of more sophisticated models with higher prediction accuracy. Hence, in the present study, a hybrid approach involving the multilayer perceptron neural network optimized with a genetic algorithm (GANN) was developed and employed to enhance the prediction efficiency of soil CEC in Tabriz plain, an arid region of Iran. In addition, a multiple model integration scheme supervised with hybrid GANN (MM-GANN) was also simulated and verified to improve the soil CEC prediction efficiency. This multiple model integration scheme supervised with the GANN approach is a unique form of a hybrid model for soil CEC prediction. Standalone MLP artificial neural network (ANN) and extreme learning machine (ELM) models were also implemented for incorporation into the multiple model integration scheme and for reasonable evaluation with MM-GANN model predictions.

2. Theoretical Overview

2.1. Artificial Neural Network (ANN). The multilayer perceptron (MLP), a class of feedforward ANN, is one of the most versatile algorithms that has proven able to simulate highly complex and nonlinear relationships between a set of input variables (predictors) and the output data (predictand) [44]. A multilayer perceptron (MLP) neural network with 1 hidden layer is shown in Figure 1. The network is trained to learn a function, $f(\cdot): P^d \rightarrow P^o$ on a set of training data, where “ d ” denotes the number of input dimensions and “ o ” denotes the number of output dimensions of the model [45]. The Levenberg-Marquardt backpropagation (BP) algorithm fine-tunes the weights and parameters of the MLP network. The network architecture involving the input layer consists of a set of processing units (neurons) $\{p_i | p_1, p_2, \dots, p_n\}$ signifying the model input features and every hidden layer neuron performs a nonlinear transformation of the inputs from the previous layer via weighted linear summation of inputs $(w_1 p_1 + w_2 p_2 + \dots + w_n p_n)$. A nonlinear activation function (σ) is then applied to each hidden unit to make a specific topology of weighted links more flexible following the affine transformation [46]. The neurons of the final layer receive connections from hidden layers of the network and are referred to as the output layer that produces a refined output. Some of the commonly used activation functions

include hyperbolic tangent (\tanh) and sigmoid (logsig) functions. There are no general rules for choosing training algorithms and adjusting associated parameters of the MLP architecture to maximize the efficiency of the network. A good introduction and mathematical concepts of ANN and its applications are provided in the following literature [47–51].

2.2. Extreme Learning Machine (ELM). The extreme learning machine (ELM) model proposed by Huang et al. [52] for a single layer feedforward network (SLFN) has been widely used for the prediction, forecasting, and estimation in many engineering fields [53–55]. Previous research studies have proved the outstanding advantages of the ELM model over the traditional AI techniques [56–58]. In addition, the ELM model can be implemented easily and has improved features such as fast learning speed [59], superior generalization performance [60], and utilization of activation functions (of nondifferentiable form) for training SLFN [52, 61]. Figure 2 portrays the general network structure of an ELM model. For N arbitrary distinct input samples $(X_i, Y_i) \in R^n \times R^n$, the standard SLFN with “ L ” hidden layer nodes can be described as follows:

$$\sum_{i=1}^N \beta_i g(X_i) = \sum_{i=1}^N \beta_i g(W_i \cdot X_i + c_i) = Y_k \quad k = 1, 2, 3, \dots, N, \quad (1)$$

where $c_i \in R$ is the assigned bias of the i^{th} hidden node, $w_i \in R$ is the assigned input weight connecting the i^{th} hidden and input layer nodes, β_i is the weight connecting the i^{th} hidden and output layer nodes, and $g(X_i)$ is the output of the i^{th} hidden layer node with respect to the input X_i . Each input is assigned to the hidden nodes in the ELM model. The output weights can be derived by finding the least square solutions to the linear system. The main difference between the ELM model and traditional AI techniques is that the parameters of the feedforward network including its input weights and the hidden layer biases are randomly selected without any adjustment in the ELM model. For good introduction and mathematical concepts of ELM and its architectures, refer to Huang et al. [62], Martínez-Martínez et al. [63], Wang et al. [64], and Ding et al. [53].

2.3. Hybrid Genetic Algorithm-Neural Network (GANN). Genetic algorithm (GA) belongs to a class of search iterative approaches based on the “Darwinian” theory of natural selection and genetics that provide optimum solutions for combinatorial optimization, heuristic search, or process planning problems [65, 66]. GA implements genetic operators like reproduction, crossover, and mutation for upgradation and search for the best population by imitating the natural evolution process artificially. The genetic algorithm is initiated with individuals, an initial population of possible solutions, with a specified objective (fitness function) wherein every single individual is symbolized using a chromosome, a distinct form of encoding [67]. The chromosomes of a population are nominated for

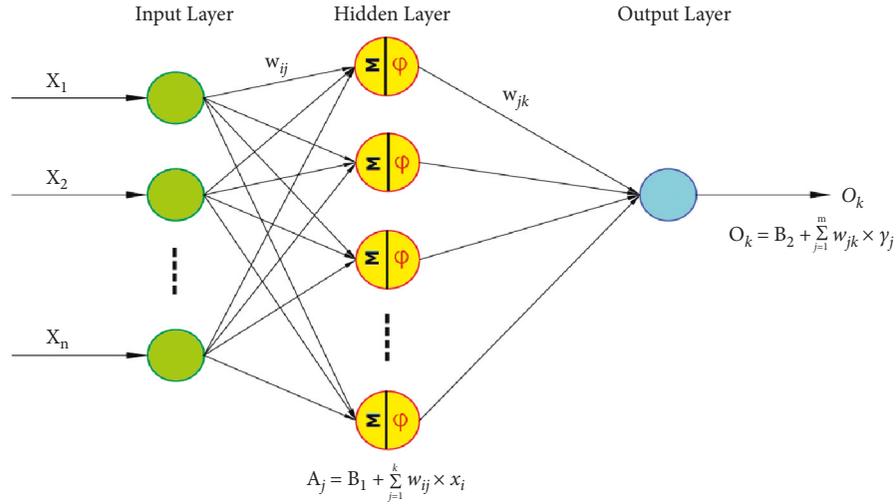


FIGURE 1: Architecture of MLP network.

reproduction based on the fitness value and the fittest individuals so selected are manipulated using crossover and mutation. The rudimentary idea here is the hope that superior parents can probabilistically produce superior offspring. The offspring of the next generation are generated by applying the GA operators crossover and mutation, upon the selected parents. The iteration process continues until the search converges to the termination criterion [65, 68]. The schematic illustration of the GA cycle is represented in Figure 3. The advantages of GA include (1) rapid convergence to the global optima, (2) superior multidirectional global search even in complex search surfaces, (3) use of probabilistic transition rules, and (4) the not deterministic ones in the search spaces where the gradient information is missing. The training of an MLP network, which is a type of neural network (NN), is somewhat a cyclic process. However, in the present case of the hybrid genetic algorithm-neural network (GANN), the intelligent search technique (GA) allows the user to configure the weight initialization range and the number of hidden layer neurons and update the weights and bias terms of an MLP network. Eventually, GA is used to learn the best hyperparameters for an MLP network. Even though the weights of the MLP network are initialized randomly, GA does not adhere to a simple random walk. Based on the parameter settings, it effectively exploits the information to gamble on fresh search points for anticipated improved performance [69]. GA selects the primary superlative solution with the best fitness values iteratively and recombines it with mutation and crossover operators to introduce offspring into the population. This process continues until the optimal solution with the highest fitness value is found based on any stopping criterion. Thus, the population's most fit MLP network is determined.

2.4. Multiple Model Integration Scheme Supervised with Hybrid GANN Model. The proposed multiple-model integration scheme involves the development of ANN and ELM

models individually using input combinations as defined in their model structures. The discrete outputs (predicted series) of individual ANN and ELM models are then unified as inputs for the GANN model to obtain superior soil CEC predictions. The implementation of this multiple-model scheme involves two phases. In the first phase, the best-performing ANN and ELM models are identified by simulating all possible combinations of inputs. Later, in the second phase, the discrete outputs (predicted series) of the best ANN and ELM models are unified as inputs to simulate the GANN model. The GA optimizes the number of hidden layer neurons and updates the weights and bias terms of an ANN. The final output derived from this proposed scheme is referred to as integrated multiple models supervised with a hybrid GANN (MM-GANN) strategy (Figure 4).

3. Case Study and Data Description

The study area (Tabriz plain) considered encompasses an area of 150000 hectares (between $45^{\circ}25'$ – $46^{\circ}12'$ E, $37^{\circ}50'$ – $38^{\circ}20'$ N) and is located in the East Azerbaijan province of Iran. The surface topography of the area comprises rugged, mountainous rims, and the study area is sited toward the north-eastern part of Urmia Lake (Figure 5). Tabriz plain is a high-altitude location (1360 m above mean sea level) characterized by cooler, wetter winters and hot summers with a tropical and subtropical steppe climate. The study area never receives greater than 40 mm of rainfall in any of the months, and the annual mean precipitation is around 360.7 mm. The geology of the area includes recent alluvium, fine elastic sediments, and red conglomerate with an alternation of sandstone and red marl. The method of ammonium saturation as mentioned in Chapman [70] was used for the cation-exchange capacity determination. The descriptive statistics of soil CEC and other soil parameters of the study area under consideration are tabulated in Table 1. The spatial distribution of observed soil CEC is presented in Figure 6. The clay and soil organic matter were positively

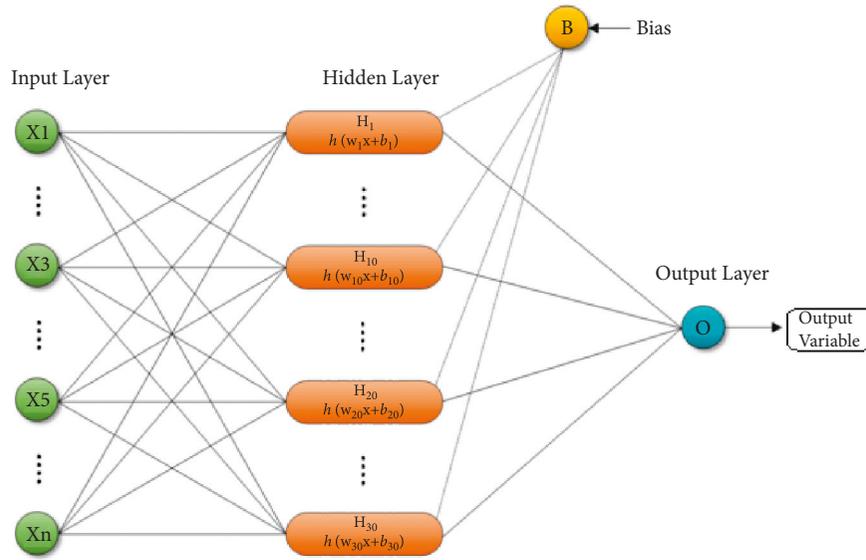


FIGURE 2: Architecture of ELM model.

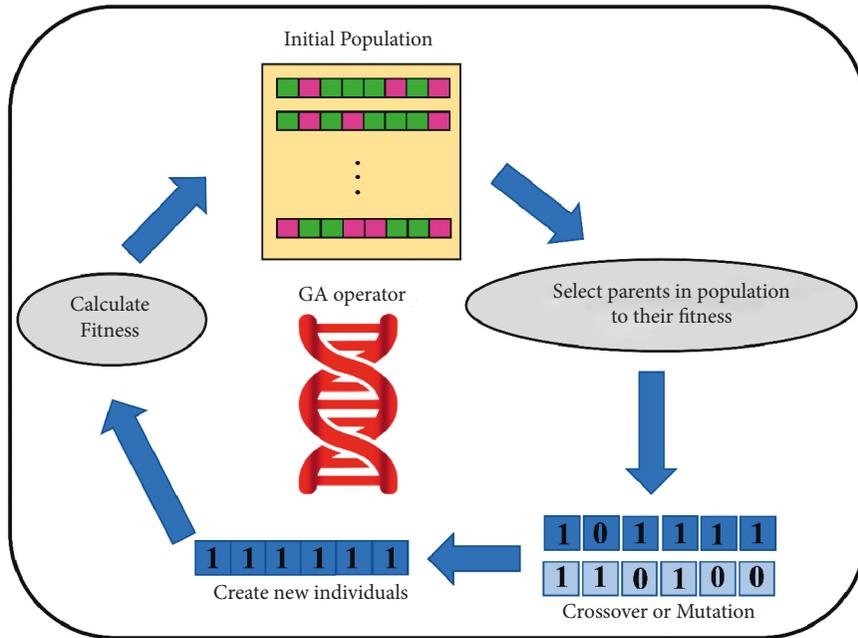


FIGURE 3: Schematic diagram of GA cycle.

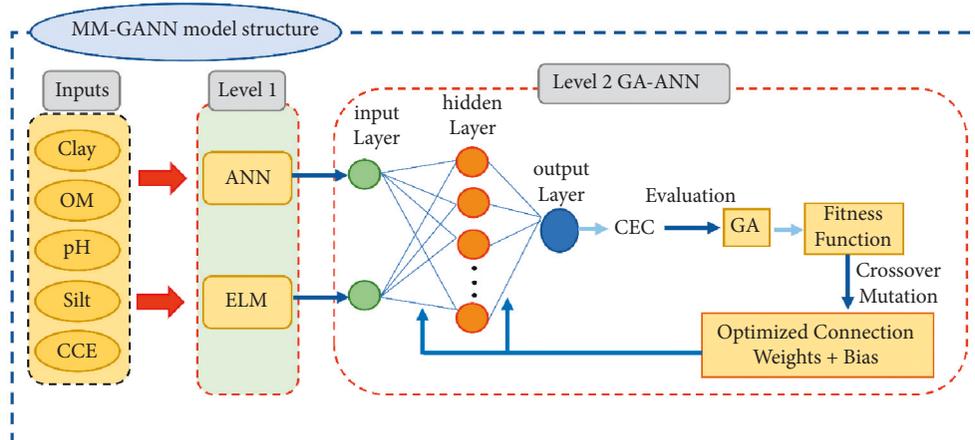


FIGURE 4: Structure of MM-GANN model.

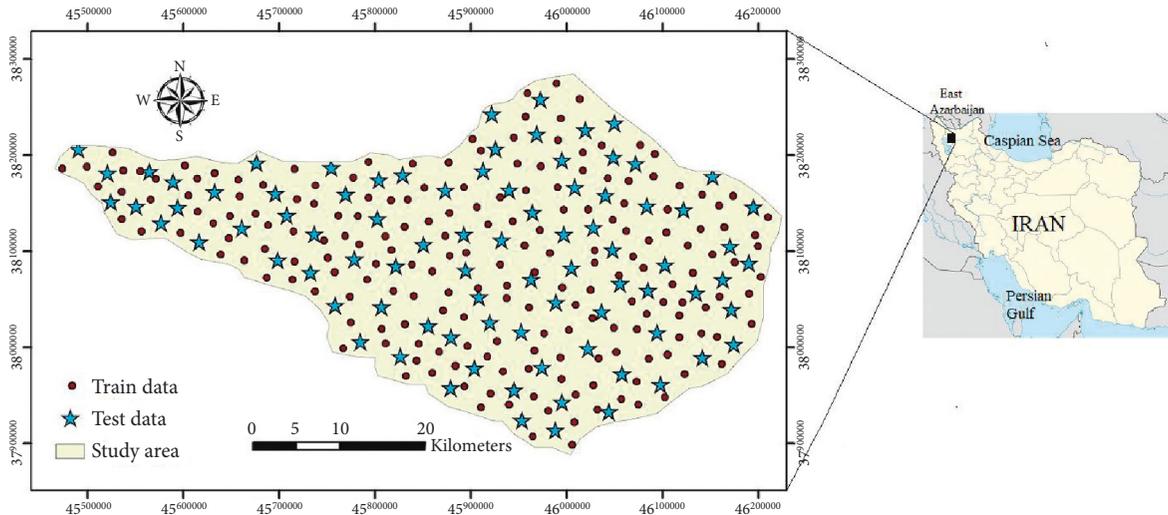


FIGURE 5: Location of the study area along with sampling points.

TABLE 1: Descriptive statistics of CEC and the soil parameters at the sampling points (training ($n = 195$) and testing ($n = 85$)).

Data set	Units	min		max		Mean		Standard deviation		Correlation with CEC
		Train	Test	Train	Test	Train	Test	Train	Test	
Clay	%	3.80	4.80	70.5	69.3	37.77	36.42	16.38	16.31	0.630
Silt	%	6.00	2.00	68.0	66.6	39.06	40.98	12.07	13.21	0.117
Sand	%	1.30	3.00	81.8	88.2	23.08	22.58	19.80	18.04	-0.605
OM	%	0.03	0.06	1.64	1.40	0.47	0.48	0.30	0.32	0.534
pH	—	7.00	6.90	8.70	8.20	7.65	7.60	0.35	0.30	0.284
CCE	%	0.55	1.60	26.55	45.0	13.02	13.82	5.00	6.38	0.220
CEC	cmol(+) kg ⁻¹	6.20	5.90	45.00	42.0	21.89	21.38	9.23	8.05	—

CCE: carbonate calcium equivalent; OM: organic matter, CEC: cation-exchange capacity.

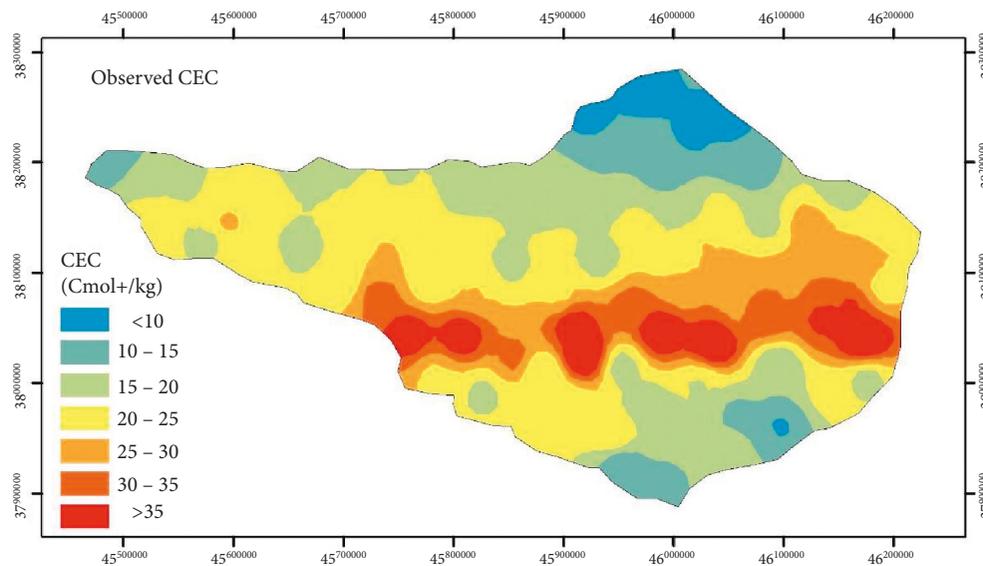


FIGURE 6: The spatial distribution map of observed soil CEC.

correlated, while the sand was found negatively correlated with soil CEC. The silt, pH, and carbonate calcium equivalent (CCE) parameters were not so significantly correlated with the soil CEC.

4. Modeling Development

Based on different combinations of soil parameters, the framework of model input-output scenarios was set for the development of ANN and ELM models with soil CEC as the output parameter. The input-output scenarios put on trial are listed in Table 2. The performance of the developed models was assessed based on the multiple statistical indices, namely, Root Mean Square Error (RMSE), Mean Absolute

TABLE 2: Input-Output structures for model development.

Models	Input combination	Output
Model 1	Clay	CEC
Model 2	Clay + pH	CEC
Model 3	Clay + OM	CEC
Model 4	Clay + OM + pH	CEC
Model 5	Clay + OM + pH + silt	CEC
Model 6	Clay + OM + pH + silt + CCE	CEC

Error (MAE), Nash Sutcliffe Efficiency (NSE) [71], Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2).

$$\text{Root Mean Square Error (RMSE)} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}},$$

$$\text{Mean Absolute Error (MAE)} = \frac{\sum_{i=1}^n |y_i - x_i|}{n},$$

$$\text{Nash Sutcliffe Efficiency (NSE)} = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (2)$$

$$\text{Mean Absolute Percentage Error (MAPE)} = \frac{100}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right|,$$

$$\text{coefficient of determination } R^2 = \left(\frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \right)^2,$$

where x_i is the actual value, y_i is the model estimated value, \bar{x} is the mean of true values, \bar{y} is the mean of the model estimated values, and n is the number of data points.

5. Results and Discussion

5.1. Performance of ANN and ELM Models. The ANN and ELM models were simulated for predicting the soil CEC based on the input-output combinations as mentioned in Table 2. The model structure (input nodes-hidden layer nodes-output nodes) and performance metrics of the ANN model for each input combination are presented in Table 3. In this study, the proposed ANN, ELM, and MM-GANN models were developed using MATLAB interface coding. The input-output scenario involving all the soil parameters (i.e., Clay + OM + pH + silt + CCE) provided the virtuous estimates of soil CEC with an NSE = 0.842. The input-output scenario involving four soil parameters (i.e., clay, OM, Ph, and silt) also offered relatively good soil CEC estimates with an NSE = 0.826. Despite having a significant correlation between clay and soil CEC, the single input-output ANN model (i.e., clay-CEC) failed to provide good soil CEC predictions. The spatial distribution map of ANN predicted

soil CEC is presented in Figure 7. The ability of the MLP network to formulate a priori explicit hypotheses about a possible nonlinear relationship among several input variables makes it illustrious from other AI methods.

The performance metrics of ELM models for each input-output scenario are tabulated in Table 4. The scenario involving all the soil parameters (i.e., clay + OM + pH + silt + CCE) provided the virtuous predictions of soil CEC with an NSE = 0.835. The ELM model efficiency was slightly lesser than that of ANN. The ELM model simulated with four inputs (i.e., clay, OM, pH, and silt) had reasonably substandard performance when compared to that of the ANN model with a similar input structure. The spatial distribution map of ELM predicted soil CEC is shown in Figure 8. The scatter plots presented in Figure 9 of the three efficient models display the accounted linear relationship between the observed and estimated soil CEC by ANN and ELM models. According to Figure 9, the ELM outperformed ANN although they have very close performance in terms of the statistical indices (Tables 3 and 4). The ELM is known for its superior learning speed and virtuous generalization performance than the ANN architecture.

TABLE 3: The performance criteria of the ANN model for six input combinations.

Input combination	Output	Structure	Train					Test				
			MAE	RMSE	MAPE	NSE	R^2	MAE	RMSE	MAPE	NSE	R^2
Model 1	CEC	1-4-1	5.593	6.904	27.784	0.438	0.438	4.599	6.162	23.598	0.407	0.409
Model 2	CEC	2-6-1	3.913	5.186	20.008	0.683	0.683	4.425	6.016	23.007	0.435	0.442
Model 3	CEC	2-5-1	4.571	5.619	24.081	0.627	0.628	4.101	5.326	22.161	0.557	0.577
Model 4	CEC*	3-7-1	2.664	3.361	13.911	0.866	0.866	3.105	3.818	16.504	0.787	0.787
Model 5	CEC**	4-5-1	2.807	3.496	15.701	0.855	0.855	2.736	3.338	15.043	0.826	0.827
Model 6	CEC***	5-6-1	2.352	2.927	12.492	0.899	0.899	2.585	3.177	13.695	0.842	0.843

Note: the unit of MAE and RMSE is cmol (+) kg^{-1} .

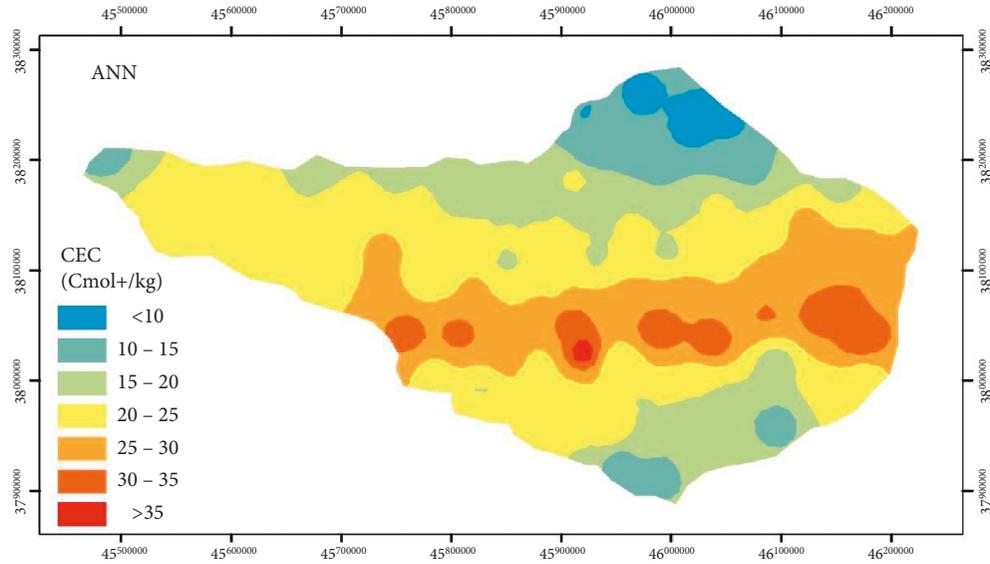


FIGURE 7: The spatial distribution map of ANN predicted soil CEC (ANN model with 5 inputs).

TABLE 4: The performance criteria of the ELM model for six input combinations.

Input combination	Output	Train					Test				
		MAE	RMSE	MAPE	NSE	R^2	MAE	RMSE	MAPE	NSE	R^2
Model 1	CEC	5.626	6.966	28.784	0.428	0.428	4.609	6.202	23.698	0.399	0.408
Model 2	CEC	3.965	5.310	20.108	0.667	0.668	4.744	6.384	23.307	0.364	0.432
Model 3	CEC	4.659	5.740	24.181	0.611	0.612	4.261	5.425	22.261	0.540	0.560
Model 4	CEC*	2.541	3.210	13.711	0.878	0.878	2.977	3.840	16.554	0.769	0.804
Model 5	CEC**	2.942	3.656	15.801	0.842	0.842	3.064	3.674	16.043	0.789	0.808
Model 6	CEC***	2.019	2.597	10.844	0.920	0.920	2.463	3.248	13.640	0.835	0.858

Note. The unit of MAE and RMSE is cmol (+) kg^{-1} .

5.2. Performance of MM-GANN Models. The soil CEC estimates of ANN and ELM models were employed as new inputs to the GANN model to predict soil CEC. To select the optimal input combinations in further modeling steps, examples from previous literature were referred to for enhancing the accuracy of models based on the different fields [72–75]. Within this category, it is worth mentioning that only the three highest performed combinations were considered in this hybrid model. The parameters of the genetic algorithm for adjusting the weights and bias terms of the ANN are presented in Table 5. Also, the performance statistics of MM-GANN models are shown in Table 6. The MM-GANN models involving the

predictions of ELM and ANN models calibrated by considering all the soil parameters (i.e., clay, OM, pH, silt, and CCE) as inputs provided superior performance with an $\text{NSE} = 0.87$ in the test phase. The ANN parameters of the best MM-GANN model are charted in Table 7. The multiple model integration scheme imparts potential to the hybrid GANN model through the usage of standalone model outputs as inputs to the model. Thus, the hybrid GANN model exploits the previously learned information to improvise the predictive power of the model. The multiple model integration scheme is apparently enhancing the learning process of the hybrid model, where the output of the standalone models is used as relative informative

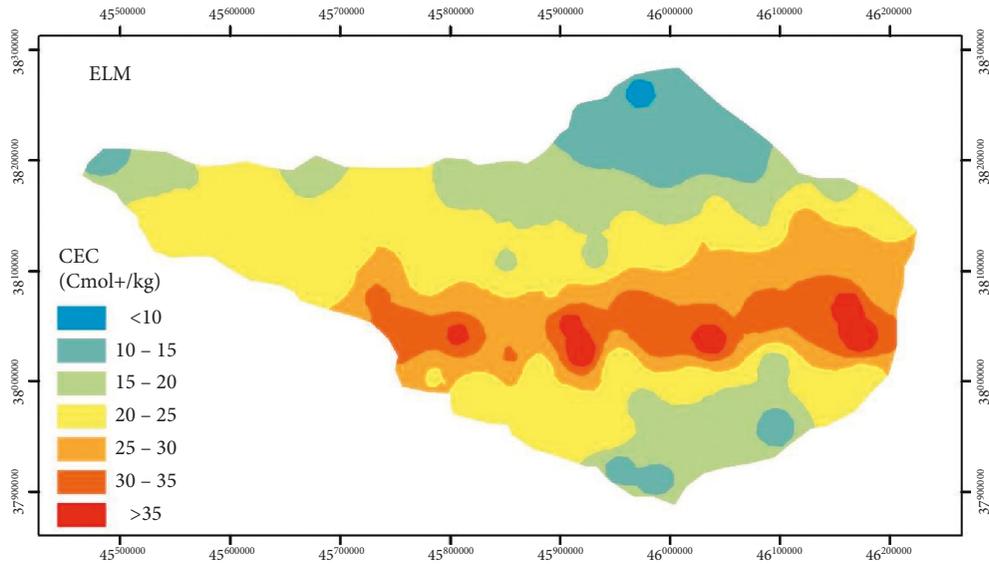


FIGURE 8: The spatial distribution map of ELM predicted soil CEC (ELM model with 5 inputs).

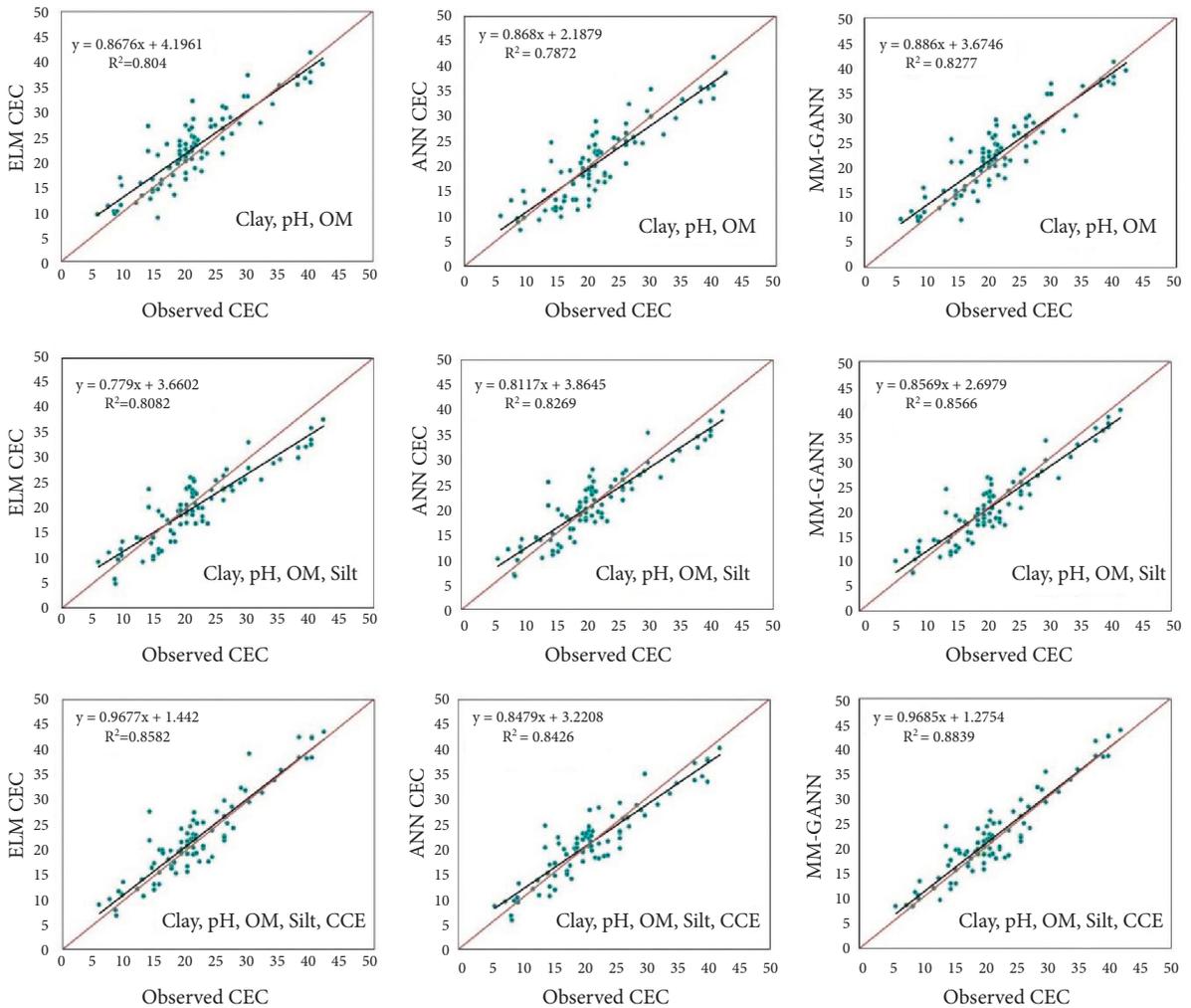


FIGURE 9: The observed and estimated soil CEC in the test stage: ELM, ANN, and MM-GANN models.

TABLE 5: Parameters of GA-Toolbox model.

Inputs	Outputs	Population	Generation	Crossover fraction	Mutation operator
ANN weights + bias	Best weights + bias	1000	500	0.7	Gaussian

TABLE 6: The performance of MM-GANN models for the three best input combinations.

Input combination	Output	ANN-structure	Train				Test					
			MAE	RMSE	MAPE	NSE	R^2	MAE	RMSE	MAPE	NSE	R^2
CEC*(ELM), CEC*(ANN)	CEC	2-5-1	2.465	3.107	13.620	0.886	0.878	2.816	3.583	14.802	0.800	0.828
CEC**(ELM), CEC**(ANN)	CEC	2-6-1	2.640	3.320	13.893	0.870	0.866	2.491	3.054	13.020	0.854	0.857
CEC*** (ELM), CEC*** (ANN)	CEC	2-5-1	1.942	2.486	10.236	0.927	0.929	2.249	2.885	12.072	0.870	0.884

Note: The unit of MAE and RMSE is cmol (+) kg^{-1} .

TABLE 7: ANN Parameters of the best MM-GANN model.

Inputs	Output	Structure	Hidden layer function	Output layer function	Hidden layer neurons	Training algorithm
CEC*** (ELM), CEC*** (ANN)	CEC	2-5-1	Tansig	Linear	5	Trainlm

predictors. The spatial distribution map of MM-GANN predicted soil CEC is presented in Figure 10 which is very much similar to that of the observed soil CEC map. The MM-GANN models developed with the predictions of ELM and ANN models calibrated by considering three and four soil parameters as inputs also offered convincingly good soil CEC predictions with $\text{NSE} = 0.80$ and 0.854 , respectively. The scatter plots of MM-GANN models shown in Figure 9 depict the goodness of fit of the model predictions against the actual soil CEC values. In Figure 9, it is evident that the third combination of the MM-GANN model indicated a very close linearly fitted line to the 1 : 1 line, especially for the combination that had all the parameters.

Table 8 compares the performances of the best model of ANN, ELM, and MM-GANN models based on the statistical measures during the training and testing phases. This table shows that the performance accuracy of the hybrid model is higher than the ELM and ANN models, respectively, based on all the criteria values. The Taylor diagrams plotted for the best ANN, ELM, and MM-GANN models are shown in Figure 11. According to the Taylor diagram, it is very much evident that the multiple-model scheme (MM-GANN) offered relatively accurate estimates of soil CEC compared to the ELM and ANN models based on three statistical metrics (RMSD, standard deviation, and correlation coefficient). The MM-GANN model was the closest to the observed/actual data. The point density plots presented in Figure 12 also supported the above statement by exposing the tradeoff between observed soil CEC against the modeled.

5.3. Validation with Published Research Studies. Validating the results of current research with reliable published literature within the context of a similar kind of study area (i.e., semiarid region) is worthwhile. The correlation coefficient (R^2) indices were selected as an indicator of

the prediction capability. The best R^2 obtained for MM-GANN, ELM, and ANN models is $R^2 \approx 0.88, 0.85,$ and 0.84 . In one of the earliest research performed on the soil CEC simulation along the Zayandehroud River in Isfahan, Iran, Amini et al. [27] established two classical ANN algorithms (i.e., feed-forward neural network and generalized regression neural network). The applied models were performed with poor prediction results with $R^2 \approx 0.69$ and 0.66 . Another study was conducted by Emamgolizadeh et al. to predict soil CEC on collected soil information from Semnan, Mashahad, and Taybad cities of Iran [35]. The authors developed two new data intelligence models, namely, genetic expression programming (GEP) and multivariate adaptive regression spline (MARS). GEP and MARS models attained an $R^2 \approx 0.80$ and 0.86 . Overall, the current study showed a convincing correlation performance over the state-of-the-art research studies.

Although the current research was the solitary approach to develop and assess the multiple model integration scheme supervised with hybrid GANN (MM-GANN), the certified limitations should be addressed in future research. It is evident from tables and figures that the MM-GANN model can improve the prediction accuracy of soil CEC when the inputs involving the predictions of ELM and ANN models calibrated by considering all the soil parameters (e.g., clay, OM, pH, silt, and CCE) are provided. However, one of the disadvantages of the MM-GANN model lies in the selection of the best standalone model for enhancing the prediction accuracy of soil CEC. Therefore, it is recommended to incorporate the prediction results of other data-driven models as the inputs of the MM-GANN model which can enhance the model's performance. In addition, this concept can be expanded and applied to other engineering fields such as structural, hydrologic, water resources, climatic, and different time series prediction/forecasting.

TABLE 8: Comparing performances of the best models of ANN, ELM, and MM-GANN.

Input combination	Output	Model structure	Train					Test				
			MAE	RMSE	MAPE	NSE	R^2	MAE	RMSE	MAPE	NSE	R^2
ANN model 6	CEC***	5-6-1	2.352	2.927	12.492	0.899	0.899	2.585	3.177	13.695	0.842	0.843
ELM model 6	CEC***		2.019	2.597	10.844	0.920	0.920	2.463	3.248	13.640	0.835	0.858
CEC*** (ELM), CEC*** (ANN)	CEC	2-5-1	1.942	2.486	10.236	0.927	0.929	2.249	2.885	12.072	0.870	0.884

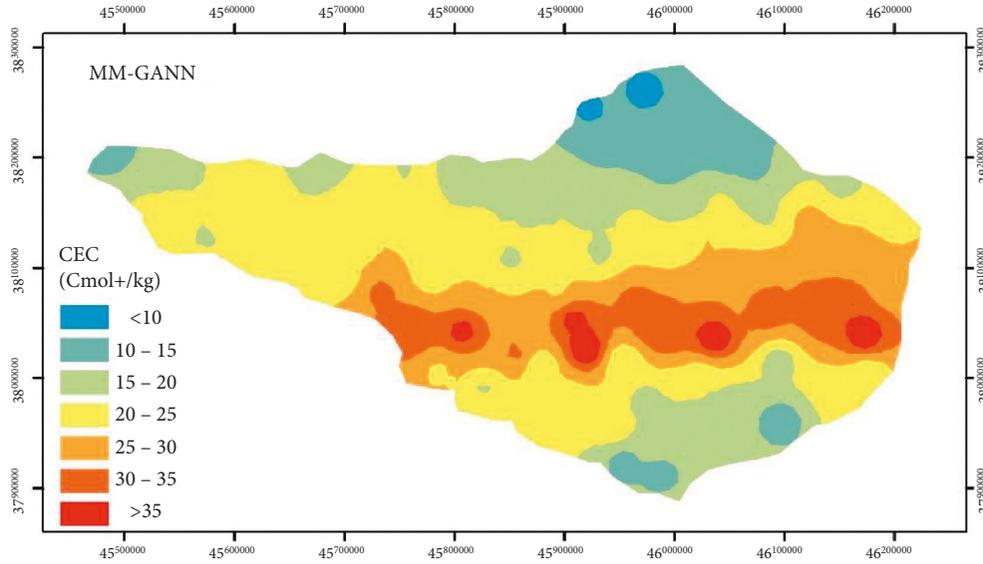


FIGURE 10: The spatial distribution map of MM-GANN predicted soil CEC (MM-GANN model calibrated with CEC*** (ELM) and CEC*** (ANN) as inputs).

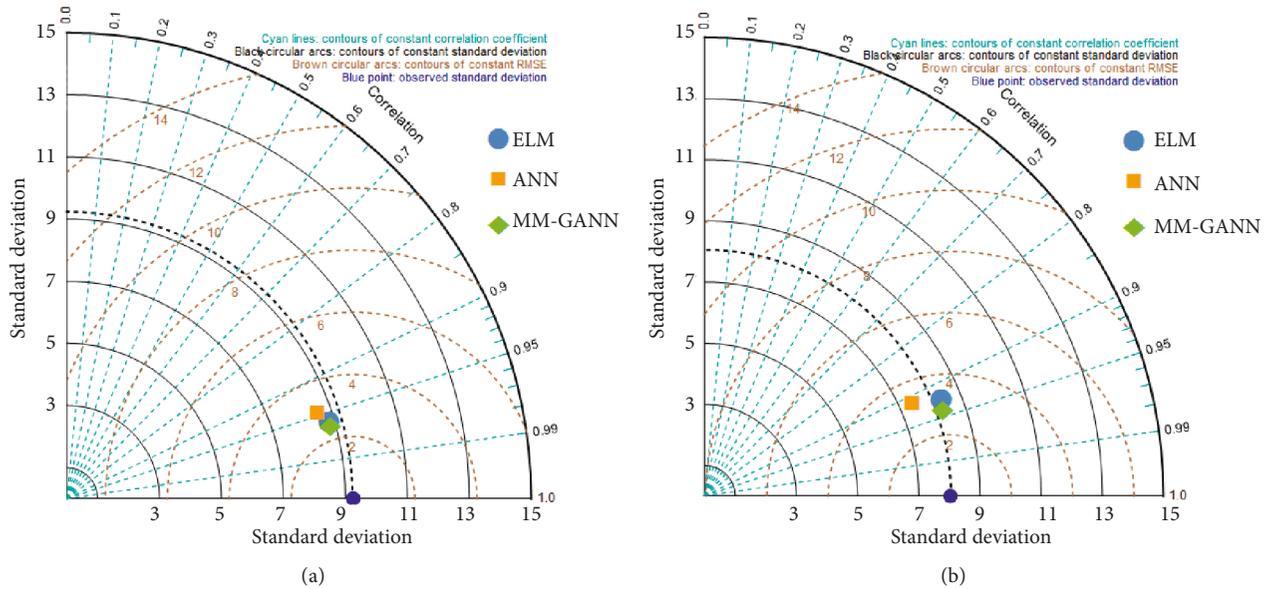


FIGURE 11: Taylor diagrams of the models at the training (a) and testing (b) phases.

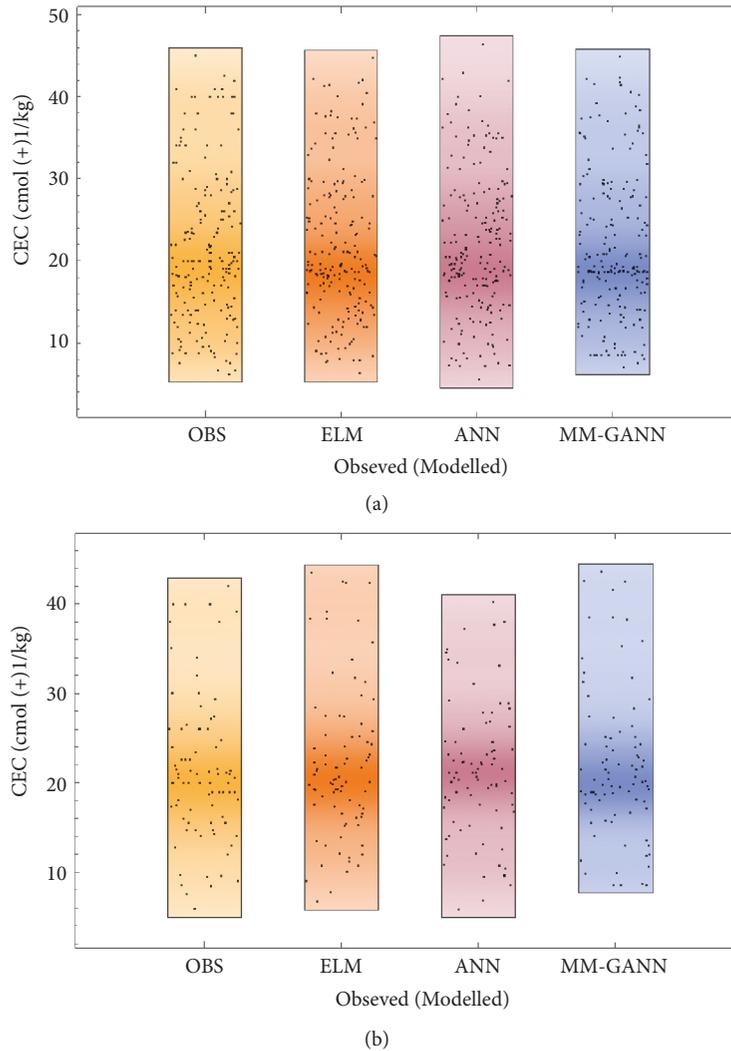


FIGURE 12: Point density plots of the models at the training (a) and testing (b) phases.

6. Conclusion

Over the past two decades, there is a noticeable demand for soil data assessment with regard to pollution and land degradation. The new era of soil process modeling using data intelligence models has been rapidly boosted. The current study was to develop a hybrid machine intelligence model based on the multimodel genetic algorithm-neural network for soil cation-exchange capacity. Two classical artificial intelligence models, namely, the ANN and ELM, were developed to evaluate their performance in estimating soil CEC along with the proposed hybrid MM-GANN model. Several correlated soil parameters including clay, silt, pH, carbonate calcium equivalent (CCE), and soil organic matter (OM) were used in the form of input attributes to the proposed and the comparable machine intelligence models. In particular, the hybrid MM-GANN model which received the predicted values of ANN and ELM as input attributes performed well in the estimation of soil CEC. Overall, the proposed multiple model integration scheme supervised with hybrid GANN model functions as an efficient pedotransfer function to predict or estimate soil CEC using readily available

soil parameters (i.e., clay, OM, pH, silt, and CCE) as input variables. In particular, the conclusions of the current investigation are as follows:

- (i) Based on the applied evaluation metrics, the ELM model provided superior CEC estimates than ANN.
- (ii) The proposed hybrid MM-GANN model outperforms both standalone ANN and ELM models in terms of all the statistical metrics.
- (iii) The proposed integrated hybrid machine intelligence scheme (MM-GANN) proved to be a reliable modeling strategy for modeling the soil cation-exchange capacity of the study area.

Before this end, it is worth stating the possibility for future research. As a fact, soil CEC is influenced by several morphological parameters [76, 77]; thus, integrating a feature selection as a prior modeling phase for the prediction process is highly recommended to be established. In addition, owing to the associated variability with each soil CEC type, it is an ideal proposition to estimate each type individually.

Data Availability

The datasets are available. Data can be shared upon request from the corresponding author.

Disclosure

This article has been previously published as preprint as follows: <https://doi.org/10.21203/rs.3.rs-471256/v1> [78].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. Wolf, "Cation exchange capacity and anion exchange," in *The fertile triangle :the Interrelationship of Air, Water, and Nutrients in Maximizing Soil Productivity*, pp. 177–188, Food Products Press, New York, NY, USA, 1999.
- [2] E. Arthur, "Rapid estimation of cation exchange capacity from soil water content," *European Journal of Soil Science*, vol. 68, no. 3, pp. 365–373, 2017.
- [3] Q. Ketterings, S. Reid, and R. Rao, *Cation Exchange Capacity*, Cornell University Cooperative Extension, Agronomy Fact Sheet #22, Ithaca, NY, USA, 2007.
- [4] B. M. Tucker, "Displacement OF ammonium IONS for cation exchange capacity measurements," *Journal of Soil Science*, vol. 25, no. 3, 1974.
- [5] D. Carroll, "Ion exchange in clays and other minerals," *GSA Bull*, vol. 70, no. 6, pp. 749–779, 1959.
- [6] P. J. Thomas, J. C. Baker, and L. W. Zelazny, "An expansive soil index for predicting shrink–swell potential," *Soil Science Society of America Journal*, vol. 64, no. 1, pp. 268–274, 2000.
- [7] M. Pansu and J. Gautheryrou, "Cation exchange capacity," in *Handbook of Soil Analysis*, pp. 709–754, Springer Berlin Heidelberg, Berlin, Germany, 2006.
- [8] D. B. Mengel, *Fundamentals of Soil Cation Exchange Capacity (CEC)*, Purdue University Cooperative Extension Service, Agronomy Guide AY-238, West Lafayette, Indiana, 1993.
- [9] J.-K. Böhlke, "Groundwater recharge and agricultural contamination," *Hydrogeology Journal*, vol. 10, no. 1, pp. 153–179, 2002.
- [10] E. Arthur, M. Tuller, P. Moldrup, and L. W. Jonge, "Clay content and mineralogy, organic carbon and cation exchange capacity affect water vapour sorption hysteresis of soil," *European Journal of Soil Science*, vol. 71, no. 2, 2020.
- [11] R. Dohrmann, "Cation exchange capacity methodology I: an efficient model for the detection of incorrect cation exchange capacity and exchangeable cation results," *Applied Clay Science*, vol. 34, no. 1–4, pp. 31–37, 2006.
- [12] R. Dohrmann, "Cation exchange capacity methodology II: a modified silver–thiourea method," *Applied Clay Science*, vol. 34, no. 1–4, pp. 38–46, 2006.
- [13] L. Delavernhe, M. Pilavtepe, and K. Emmerich, "Cation exchange capacity of natural and synthetic hectorite," *Applied Clay Science*, vol. 151, pp. 175–180, 2018.
- [14] M. Conradie and W. A. G. Kotze, "Comparison of methods for estimating cation exchange capacity of orchard soils," *South African Journal of Plant and Soil*, vol. 6, no. 2, pp. 136–137, 1989.
- [15] M. Khorshidi and N. Lu, "Determination of cation exchange capacity from soil water retention curve," *Journal of Engineering Mechanics*, vol. 143, no. 6, Article ID 04017023, 2017.
- [16] K. Liao, S. Xu, and Q. Zhu, "Development of ensemble pedotransfer functions for cation exchange capacity of soils of Qingdao in China," *Soil Use & Management*, vol. 31, no. 4, pp. 483–490, 2015.
- [17] L. A. Manrique, C. A. Jones, and P. T. Dyke, "Predicting cation-exchange capacity from soil physical and chemical properties," *Soil Sci. Soc. Am.* vol. 55, no. 3, pp. 787–794, 1991.
- [18] S. E. Obalum, Y. Watanabe, C. A. Igwe, M. E. Obi, and T. Wakatsuki, "Improving on the prediction of cation exchange capacity for highly weathered and structurally contrasting tropical soils from their fine-earth fractions," *Communications in Soil Science and Plant Analysis*, vol. 44, no. 12, pp. 1831–1848, 2013.
- [19] A. M. Lambooy, "Relationship between cation exchange capacity, clay content and water retention of Highveld soils," *South African Journal of Plant and Soil*, vol. 1, no. 2, pp. 33–38, 1984.
- [20] R. L. Parfitt, D. J. Giltrap, and J. S. Whitton, "Contribution of organic matter and clay minerals to the cation exchange capacity of soils," *Communications in Soil Science and Plant Analysis*, vol. 26, no. 9–10, pp. 1343–1355, 1995.
- [21] L. Krogh, H. Breuning-madsen, and M. H. Greve, "Cation-exchange capacity pedotransfer functions for Danish soils," *Acta Agriculturae Scandinavica Section B Soil and Plant Science*, vol. 50, no. 1, pp. 1–12, 2000.
- [22] P. J. van Erp, V. J. G. Houba, and M. L. van Beusichem, "Actual Cation Exchange Capacity of Agricultural Soils and its relationship with pH and content of organic carbon and clay," *Communications in Soil Science and Plant Analysis*, vol. 32, no. 1–2, pp. 19–31, 2001.
- [23] C. A. Seybold, R. B. Grossman, and T. G. Reinsch, "Predicting cation exchange capacity for soil survey using linear models," *Soil Science Society of America Journal*, vol. 69, no. 3, pp. 856–863, 2004.
- [24] H. R. Fooladmand, "Estimating cation exchange capacity using soil textural data and soil organic matter content: a case study for the south of Iran," *Archives of Agronomy and Soil Science*, vol. 54, no. 4, pp. 381–386, 2008.
- [25] H. Khodaverdiloo, H. Momtaz, and K. Liao, "Performance of soil cation exchange capacity pedotransfer function as affected by the inputs and database size," *Clean - Soil, Air, Water*, vol. 46, no. 3, Article ID 1700670, 2018.
- [26] K. V. Looy, J. Bouma, and M. Herbst, J. Koestel, B. Minasny, U. Mishra, C. Montzka, A. Nemes, A. Y. Pachepsky, J. Padarian, M. G. Schaap, B. Tóth, A. Verhoef, J. Vanderborght, M. J. V. D. Ploeg, L. Weihermüller, S. Zacharias, Y. Zhang, H. Vereecken, "Pedotransfer functions in earth system science: challenges and perspectives," *Reviews of Geophysics*, vol. 55, no. 4, pp. 1199–1256, 2017.
- [27] M. Amini, K. C. Abbaspour, H. Khademi, N. Fathianpour, M. Afyuni, and R. Schulin, "Neural network models to predict cation exchange capacity in arid regions of Iran," *European Journal of Soil Science*, vol. 56, no. 4, pp. 551–559, 2005.
- [28] H. Bayat, N. Davatgar, and M. Jalali, "Prediction of CEC using fractal parameters by artificial neural networks," *International Agrophysics*, vol. 28, no. 2, pp. 143–152, 2014.
- [29] J. Seyedmohammadi, L. Esmaelnejad, and H. Ramezani, "Determination of a suitable model for prediction of soil cation exchange capacity," *Model. Earth Syst. Environ.* vol. 2, no. 3, p. 156, 2016.

- [30] L. Tang, G. Zeng, F. Nourbakhsh, and G. L. Shen, "Artificial neural network approach for predicting cation exchange capacity in soil based on physico-chemical properties," *Environmental Engineering Science*, vol. 26, no. 1, pp. 137–146, 2009.
- [31] A. A. Zolfaghari, R. T. Mehrjardi, A. R. Moshki et al., "Using the nonparametric k-nearest neighbor approach for predicting cation exchange capacity," *Geoderma*, vol. 265, pp. 111–119, 2016.
- [32] Y. K. Kalkhajeh, R. R. Arshad, H. Amerikhah, and M. Sami, "Comparison of multiple linear regressions and artificial intelligence-based modeling techniques for prediction the soil cation exchange capacity of Aridisols and Entisols in a semi-arid region," *Aust. J. Agric. Eng.*, vol. 3, no. 2, p. 39, 2012.
- [33] H. Ghorbani, H. Kashi, N. Hafezi Moghadas, and S. Emamgholizadeh, "Estimation of soil cation exchange capacity using multiple regression, artificial neural networks, and adaptive neuro-fuzzy inference system models in golestan province, Iran," *Communications in Soil Science and Plant Analysis*, vol. 46, no. 6, pp. 763–780, 2015.
- [34] S. I. C. Akpa, S. U. Ugboje, T. F. A. Bishop, and I. O. A. Odeh, "Enhancing pedotransfer functions with environmental data for estimating bulk density and effective cation exchange capacity in a data-sparse situation," *Soil Use & Management*, vol. 32, no. 4, pp. 644–658, 2016.
- [35] S. Emamgolizadeh, S. M. Bateni, D. Shahsavani, T. Ashrafi, and H. Ghorbani, "Estimation of soil cation exchange capacity using genetic expression programming (GEP) and multivariate adaptive regression splines (MARS)," *Journal of Hydrology*, vol. 529, pp. 1590–1600, 2015.
- [36] A. A. Jafarzadeh, M. Pal, M. Servati, M. H. FazeliFard, and M. A. Ghorbani, "Comparative analysis of support vector machine and artificial neural network models for soil cation exchange capacity prediction," *International journal of Environmental Science and Technology*, vol. 13, no. 1, pp. 87–96, 2016.
- [37] A. Keshavarzi, F. Sarmadian, J. Shiri, M. Iqbal, R. Tirado-Corbalá, and E.-S. E. Omran, "Application of ANFIS-based subtractive clustering algorithm in soil Cation Exchange Capacity estimation using soil and remotely sensed data," *Measurement*, vol. 95, pp. 173–180, 2017.
- [38] K. Liao, S. Xu, J. Wu, Q. Zhu, and L. An, "Using support vector machines to predict cation exchange capacity of different soil horizons in Qingdao City, China," *Journal of Plant Nutrition and Soil Science*, vol. 177, no. 5, pp. 775–782, 2014.
- [39] H. Shekofteh, F. Ramazani, and H. Shirani, "Optimal feature selection for predicting soil CEC: comparing the hybrid of ant colony organization algorithm and adaptive network-based fuzzy system with multiple linear regression," *Geoderma*, vol. 298, pp. 27–34, 2017.
- [40] A. A. Ewees, A. A. Mohammed, A. Laith, and O. Diego, "Boosting arithmetic optimization algorithm with genetic algorithm operators for feature selection: case study on cox proportional hazards model," *Mathematics*, vol. 9, no. 18, p. 2321, 2021.
- [41] O. Maciel, A. Valdivia, D. Oliva, E. Cuevas, D. Zaldívar, and M. Pérez-Cisneros, "A novel hybrid metaheuristic optimization method: hypercube natural aggregation algorithm," *Soft Computing*, vol. 24, pp. 8823–8856, 2020.
- [42] H. Tao, A. A. Ewees, A. O. A. Sultani et al., "Global solar radiation prediction over North Dakota using air temperature: development of novel hybrid intelligence model," *Energy Reports*, vol. 7, pp. 136–157, 2021.
- [43] A. A. Bidgoli, H. E. Komleh, and S. J. Mousavirad, "Seminal Quality Prediction Using Optimized Artificial Neural Network with Genetic Algorithm," in *Proceedings of, Bursa, Turkey*, 2016.
- [44] K. Mehrotra, C. Mohan, and S. Ranka, *Elements of Artificial Neural Networks*, The MIT Press, Cambridge, MA, USA, 1996.
- [45] S. K. Bhagat, K. Pyrgaki, S. Q. Salih et al., "Prediction of copper ions adsorption by attapulgite adsorbent using tuned-artificial intelligence model," *Chemosphere*, vol. 276, Article ID 130162, 2021.
- [46] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 842 pages, Prentice-Hall, Hoboken, NY, USA, 1999.
- [47] T. Kohonen, "An introduction to neural computing," *Neural Networks*, vol. 1, no. 1, pp. 3–16, 1988.
- [48] G. Dreyfus, *Neural Networks*, Springer-Verlag, Berlin/Heidelberg, Germany, 2005.
- [49] V. Kotu and B. Deshpande, *Data Science: Concepts and Practice*, Elsevier, Amsterdam, Netherlands, 2019.
- [50] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, CRC Press, Boca Raton, FL, 2018.
- [51] F. Cui, Z. A. Al-Sudani, G. S. Hassan, H. A. Afan, S. J. Ahammed, and Z. M. Yaseen, "Boosted artificial intelligence model using improved alpha-guided grey wolf optimizer for groundwater level prediction: comparative study and insight for federated learning technology," *Journal of Hydrology*, vol. 606, Article ID 127384, 2022.
- [52] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine : a new learning scheme of feedforward neural networks," *IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 985–990, 2004.
- [53] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, "Extreme learning machine: algorithm, theory and applications," *Artificial Intelligence Review*, vol. 44, no. 1, pp. 103–115, 2015.
- [54] S. Naganna, P. Deka, M. Ghorbani, S. Biazar, N. Al-Ansari, and Z. Yaseen, "Dew Point Temperature Estimation: Application of Artificial Intelligence Model Integrated with Nature-Inspired Optimization Algorithms," *Water*, vol. 11, no. 4, 2019.
- [55] M. Leuenberger and M. Kanevski, "Extreme Learning Machines for spatial environmental data," *Computers & Geosciences*, vol. 85, pp. 64–73, 2015.
- [56] B. Deng, X. Zhang, W. Gong, and D. Shang, "An overview of extreme learning machine," in *Proceedings of thin 2019 4th international conference on control, robotics and cybernetics (crc)*, pp. 189–195, Tokyo, Japan, September 2019.
- [57] J. X. Chen, "Applications of extreme learning machines," *Computing in Science & Engineering*, vol. 21, no. 5, pp. 4–5, 2019.
- [58] J. Alavi, A. A. Ewees, S. Ansari, S. Shahid, and Z. M. Yaseen, "A new insight for real-time wastewater quality prediction using hybridized kernel-based extreme learning machines with advanced optimization algorithms," *Environmental Science & Pollution Research*, vol. 29, pp. 20496–20516, 2022.
- [59] H. T. Huynh, Y. Won, and J.-J. Kim, "An improvement of extreme learning machine for compact single-hidden-layer feedforward neural networks," *International Journal of Neural Systems*, vol. 18, no. 05, pp. 433–441, 2008.
- [60] A. Akusok, *Extreme Learning Machines: Novel Extensions and Application to Big Data*, University of Iowa, Iowa City, IA, USA, 2016.
- [61] Z. M. Yaseen, M. Ali, A. Sharafati, N. Al-Ansari, and S. Shahid, "Forecasting standardized precipitation index using

- data intelligence models: regional investigation of Bangladesh,” *Scientific Reports*, vol. 11, no. 1, 2021.
- [62] G.-B. Huang, D. H. Wang, and Y. Lan, “Extreme learning machines: a survey,” *Int. J. Mach. Learn. Cybern.* vol. 2, no. 2, pp. 107–122, 2011.
- [63] J. M. Martínez-Martínez, P. Escandell-Montero, E. Soria-Olivas, J. D. Martín-Guerrero, R. Magdalena-Benedito, and J. Gómez-Sanchis, “Regularized extreme learning machine for regression problems,” *Neurocomputing*, vol. 74, no. 17, pp. 3716–3721, 2011.
- [64] Y. Wang, F. Cao, and Y. Yuan, “A study on effectiveness of extreme learning machine,” *Neurocomputing*, vol. 74, no. 16, pp. 2483–2490, 2011.
- [65] D. E. G. A. J. H. Holland, *GUEST EDITORIAL Genetic Algorithms and Machine Learning*, pp. 95–99, Springer, New York, NY, USA, 1988.
- [66] J. H. Holland, “Genetic algorithms,” *Scientific American*, vol. 267, no. 1, pp. 66–72, 1992.
- [67] D. E. Goldberg and K. Deb, “A comparative analysis of selection schemes used in genetic algorithms,” *Foundations of genetic algorithms*, vol. 1, pp. 69–93, 1991.
- [68] A. H. Wright, “Genetic algorithms for real parameter optimization,” in *Foundations of genetic algorithms* vol. 1, , pp. 205–218, Elsevier, 1991.
- [69] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-We, Boston, MA, 1989.
- [70] H. D. Chapman, “Cation-exchange capacity,” *Methods soil Anal. Part 2 Chem. Microbiol. Prop.* vol. 9, pp. 891–901, 1965.
- [71] Z. M. Yaseen, “An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: review, challenges and solutions,” *Chemosphere*, vol. 277, Article ID 130126, 2021.
- [72] S. Kim and H. S. Kim, “Neural Networks and Genetic Algorithm Approach for Nonlinear Evaporation and Evapotranspiration Modeling,” *Journal of Hydrology*, vol. 351, pp. 299–317, 2008.
- [73] S. Kim, S. Seo, M. Rezaie-Balf, O. Kisi, M. A. Ghorbani, and V. P. Singh, “Evaluation of Daily Solar Radiation Flux using Soft Computing Approaches based on Different Meteorological Information: Peninsula vs Continent,” *Theoretical and Applied Climatology*, vol. 137, pp. 693–712, 2019.
- [74] S. Kim and V. P. Singh, “Modeling Daily Soil Temperature using Data-Driven models and Spatial Distribution,” *Theoretical and Applied Climatology*, vol. 118, pp. 465–479, 2014.
- [75] S. Kim, V. P. Singh, C. J. Lee, and Y. Seo, “Modeling the Physical Dynamics of Daily Dew Point Temperature using Soft Computing Techniques,” *KSCE Journal of Civil Engineering*, vol. 19, pp. 1930–1940, 2015.
- [76] A. Sharma, D. C. Weindorf, D. D. Wang, and S. Chakraborty, “Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC),” *Geoderma*, vol. 239, pp. 130–134, 2015.
- [77] K. H. Tan and P. S. Dowling, “Effect of organic matter on CEC due to permanent and variable charges in selected temperate region soils,” *Geoderma*, vol. 32, no. 2, pp. 89–101, 1984.
- [78] M. Shahabi, M. I. Ghorbani, S. R. Naganna et al., “Modeling Soil Cation Exchange Capacity in Arid Region of Iran,” *Application of Novel Hybrid Intelligence Paradigm*, 2021.