Ain Shams Engineering Journal 15 (2024) 102510

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: www.sciencedirect.com

Bidirectional Long Short-Term Memory (BILSTM) - Support Vector Machine: A new machine learning model for predicting water quality parameters

Zahra Jamshidzadeh [a], Mohammad Ehteram [b], Hanieh Shabanian [c,*]

[a] Department of Civil Engineering, University of Kashan, Kashan, Iran
[b] Department of Water Engineering, Semnan University, Semnan, Iran
[c] Computer Science Department, Western New England University, Springfield, USA

ARTICLE INFO

Keywords:
Water quality
Deep learning
Optimization
Total dissolved solids
Bidirectional LSTM

ABSTRACT

Water pollution threatens human health, agriculture, and ecosystems. Accurate prediction of water quality parameters is crucial for effective protection. We suggest a novel hybrid deep learning model that enhances the efficiency of Support Vector Machines (SVMs) in predicting Electrical Conductivity (EC) and Total Dissolved Solids (TDS). Our model combines Bidirectional Long Short-Term Memory (BILSTM) and SVMs to extract essential features and predict output variables. We evaluated the models using input parameters (PH, Ca++, Mg++, Na+, K+, HCO3, SO4, and Cl) for one, two, and three-day predictions. Employing the Ali Baba and Forty Thieves (AFT) optimization algorithm, we identified optimal input combinations. The BILSTM-SVM model accurately estimated TDS values, with MAPE values of 2%, outperforming other models. Similarly, it successfully predicted EC values, exhibiting an R2 value of 0.94. Our proposed model processes complex relationships and captures crucial features from the data, contributing to improved water quality prediction.

1. Introduction

Rapid urbanization and industrialization often affect the water quality of urban areas. The health of aquatic ecosystems is directly affected by water quality. Therefore, the prediction of water quality parameters is crucial for assessing the health of ecosystems, identifying threats and developing appropriate management strategies [1]. While mechanical models can capture complex chemical, biological, and hydrodynamic processes, their computations can be time-consuming due to the large number of data points required [2]. Conversely, machine learning models are preferred because they can accurately identify relationships between inputs and responses. Support Vector Machine (SVM) is a well-known machine learning model that is widely used to predict water quality parameters. Several studies have applied SVM models to predict various water quality parameters (WQPs). Kisi and Parmar [3] used least square SVM (LSSVM), decision tree models, and multivariate adaptive regression splines to predict Chemical Oxygen Demand (COD). They used various combinations of water quality parameters as inputs to the models. The LSSVM was a relabel model for WQPs. Similarly, Ji et al. [4] developed SVM models for predicting dissolved Oxygen (DO) and identified ammonium-nitrogen as the most important input variable. Model inputs consisted of 11 hydro-chemical variables. The data were bimonthly measured at eight sampling sites. Their SVM model successfully predicted water quality parameters. Kamyab-Talesh et al. [5] employed SVM models to accurately predict the water quality index. The measured data were used to design the model. The data were split into two groups for training and testing. They achieved a coefficient of determination of 0.87 and a mean square error of 0.061. Najafzadeh and Niazmardi [6] developed a multiple-kernel SVM (MKSVM) to predict water quality parameters. The particle swarm optimization (PSO) algorithm was used for solving the optimization problem of the MKSVR. The MKSVM enhanced the accuracy of the SVM model. A combination of MKSVR and PSO demonstrated that the newly developed SVM technique successfully estimated water quality parameters in natural streams. Deng et al. [7] used SVM and artificial neural network (ANN) models for predicting water quality parameters. The SVM models required longer training times. Nasir et al. [8] developed SVM, ANN, and decision tree models to enhance the accuracy of predictions. They proposed the boosting algorithm for water quality classification. They stated that the artificial intelligence models

* Corresponding author.
E-mail address: Hanieh.shabanian@wne.edu (H. Shabanian).

https://doi.org/10.1016/j.asej.2023.102510
Received 10 July 2023; Received in revised form 21 September 2023; Accepted 27 September 2023
Available online 10 October 2023

achieved the high accuracy as a meta classifier. Scholars developed new SVMs model for predicting water quality parameters. They reported that the hybrid SVM model outperformed the SVM model.

Although studies have widely used SVM models for predicting water quality parameters, they have not addressed the shortcomings of the SVM model. One of these limitations is the inability of SVM models to extract significant features from data. Another limitation is the dependency of the accuracy of the model on the choice of input combinations, which requires the use of robust input selection methods. Additionally, the adjustment of SVM parameters requires robust optimization algorithms. It is essential to develop SVM models for accurate water quality prediction due to their limitations in extracting important features from time series data. they may not accurately handle complex data and may not perform optimally [9,10].

Deep learning models have emerged as powerful tools for extracting features and modeling complex phenomena [11,12]. Bidirectional long-short-term memory neural networks (BILSTM) and long-short-term memory neural networks (LSTM) are two popular deep learning models that have been successfully used for predicting water quality parameter. Unlike LSTM models, BI-LSTM models process information in both directions. The LSTM and BI-LSTM models use advanced operators to extract important features from input data. Wu and Wang [12] developed LSM models to predict water quality parameters. They reported the wavelet-LSTM model outperformed the LSTM models. Wan et al. [11] also reported that LSTM models could improve prediction precision. In another study, Valadkhan and Moghaddasi [13] successfully used the LSTM-ANN model for groundwater quality predictions. Zhang et al. [14] developed a novel BILSTM model for predicting water quality parameters. Their model consisted of a BILSTM layer and an attention layer to identify crucial features and make accurate predictions. The results revealed that the attention mechanism significantly improved the performance of the BILSTM model. Bi et al. [15] combined BILSTM and attention mechanism to predict water quality parameters. The study reported that the novel method successfully estimated the water quality parameters. These studies indicate that the LSTM and BISLTM have high capabilities for analyzing ad simulating complex problems.

The aim of this paper is to enhance the prediction precision of WQPs by leveraging the information extraction capabilities of BILSTM and LSTM models. Specifically, this study proposes two hybrid models, the LSTM-SVM and the BI-LSTM-SVM models, to predict water quality parameters. The combination of LSTM and BILSTM models with SVM enhances the feature extraction process from time series data.

This study also introduces a novel optimization algorithm to identify optimal input scenarios. Specifically, we utilize the Ali Baba and the Forty Thieves (ABFT) optimization algorithm. Braik et al. [16] introduced the ABFT as an efficient and effective solution for solving complex optimization problems. The ABFT was inspired by the famous story of Ali Baba and the forty thieves. The ABFT was tested on a variety of benchmark functions and engineering problems. The ABFT outperformed the other algorithms. By integrating ABFT with the hybrid models, this study enhances the accuracy and robustness of the proposed framework in predicting water quality parameters.

The proposed hybrid models, namely the LSTM-SVM and BI-LSTM-SVM models, are applied to predict two critical water quality parameters, namely electrical conductivity (EC) and total dissolved solids (TDS). Accurate predictions of EC and TDS are particularly vital for assessing water pollutants and aquatic life sustainability. Therefore, this study aims to enhance the performance of the proposed models in predicting these critical water quality parameters.

This research presents several notable contributions to the field of water quality prediction, including:

1. Development of the BILSTM-SVM model for extracting complex features from the data.
2. Evaluation of the performance and accuracy of the BILSTM-SVM model versus existing models such as LSTM, SVM, BILSTM, and LSTM-VM.
3. Utilization of the ABFT optimization algorithm for selecting optimal input scenarios and improving the predictive capabilities of the proposed models
4. Application of the BILSTM-SVM model for predicting one-, two-, and three-day ahead values of EC and TDS.

2. Material and methods

2.1. Structure of the SVM model

The SVM model is a popular machine-learning technique introduced by Vapnik et al. [17]. It is used to minimize an optimization problem defined based on a set of equations. The linear form of the SVM model is as follows:

$$f(in) = \delta^{Tr}.in + \varepsilon \tag{1}$$

where in is the input, ε is the bias, δ is the weight, and $f(in)$ is the output. The SVM should minimize following function:

$$Minimize \frac{1}{2}\|\delta\|^2 + \alpha \sum_{i=1}^{m} \left(\lambda_i^- + \lambda_i^+\right) \tag{2}$$

$$subject(to)(\delta_i.in_i + \varepsilon) - out_i < \gamma + \lambda_i^+ \tag{3}$$

$$out_i - (\delta_i.in_i + \varepsilon) \leq \gamma + \lambda_i^- \tag{4}$$

where α is the penalty coefficient, m is the number of training data points, λ_i^- and λ_i^+ are the penalty for the prediction errors; and out is the observed output. Equation (1) is rewritten based on the kernel functions:

$$f(in) = \delta^{Tr}.K(in, in_i) + \varepsilon \tag{5}$$

where $K(in, in_i)$: the kernel function. In modeling process, the radial basis function (RBF) is widely used as a kernel function. Previous studies indicated that the RBF gave the high accuracy [17].

$$K(in, in_i) = \exp\left(-\frac{|in - in_i|}{2\phi^2}\right) \tag{6}$$

where ϕ: the kernel parameter. This study uses a continuous version of ABFT to adjust the SVM parameters (kernel parameter and penalty coefficient).

2.2. Structure of the BI-LSTM and LSTM models

The LSTM replaces the self-connected hidden units of a traditional RNN with memory blocks [11,18]. The memory blocks contain multiple memory cells. Each memory cell stores information over a specific period of time and can be accessed and modified by different gate units. Specifically, LSTMs have three types of gates: input (IN), forget (FO), and output gates (OU), which are responsible for controlling the flow of information. The input gate regulates the flow of new information, while the forget gate determines what information should be discarded from each memory cell. Finally, the output gate determines what information should be sent to the output of the LSTM [11,14]. The outputs of different gates are calculated as follows:

$$fo_t = \sigma\left[\mu_f.[z_{t-1}, \kappa_t] + \beta_f\right] \tag{7}$$

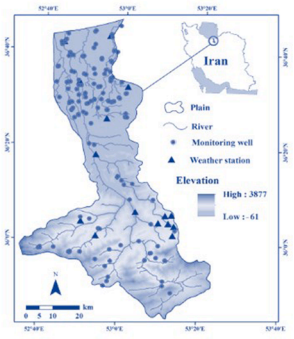
$$i_t = \sigma[\mu_i.[z_{t-1}, \kappa_t] + \beta_i] \tag{8}$$

$$\bar{\zeta}_t = \tanh(\mu_t.[z_{t-1}, \kappa_t] + \beta_c) \tag{9}$$

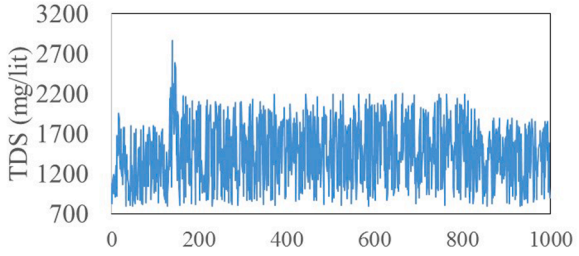
$$\zeta_t = f_t{}^*\zeta_{t-1} + i_t{}^*\bar{\zeta}_t \tag{10}$$

Table 1
Input data.

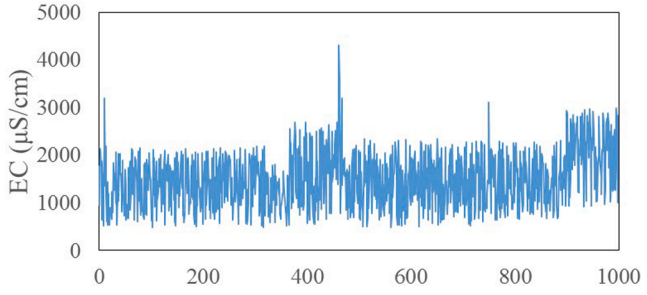
Parameter	Maximum	Average	Minimum
PH	8.3	7.0	6.5
TDS (mg/lit)	2888.1	1454.4	312.0
EC (μS/cm0	4310.0	2300.6	488.00
Mg^{++}	84.2	22.12	5.1
Ca^{++}	212.2	110.2	25.2
SO_4	435.2	54.23	3.21
HCO_3	880.2	404.2	54.12
K^+	4.2	3.32	1.23
Na^+	724.2	130.2	9.4



(a) [29]



(b)



(c)

Fig. 1. Location of the case study area and the TDS and EC data, respectively. (a) location of the case study, (b) TDS data, and (c) EC data.

$$o_t = \sigma[\mu_o.[z_{t-1}, \kappa_t] + \beta_o] \tag{11}$$

$$z_t = o_t \times \tanh(\zeta_t) \tag{12}$$

where fo_t: Output of forget gate, κ_t: Output of the input gate, o_t: Output of the output gate, σ: the activation function, μ_f, μ_o and μ_i: Weight coefficients, β_i, β_o, β_f, and β_c: Bias values of the IN gate, OU gate, FO gate, and memory cell, z_{t-1} Hidden state at time t-1, $\bar{\zeta}_t$:The new candidate value for cell state, ζ_t: Value of memory cell, z_t Output matrix. Unlike ordinary LSTMs, BILSTMs process information in both forward and backward directions [19]. The first LSTM model processes information in the forward direction, while the second LSTM model processes information in the backward direction [19]. This study uses a binary version of ABFT to adjust the LSTM and BILSTM parameters.

Table 2
The optimal parameter values for predicting TDS and EC. (a) the optimal parameter values for predicting TDS, and (b) the optimal parameter values for predicting TDS EC.

(a) PSI	OBF	MNI	OBF
45	1.12	70	1.14
90	0.89	140	0.87
135	1.34	210	0.99
180	1.45	280	1.44
(b) PSI	OBF	MNI	OBF
45	1.17	70	1.19
90	0.86	140	0.84
135	1.36	210	1.02
180	1.49	280	1.43

2.3. Structure of the ABFT for solving optimization problems

The ABFT optimization algorithm is inspired by the famous tale of Ali Baba and the forty thieves. In the story, a group of forty thieves collaborates to find the house of Ali Baba. Marjaneh is a clever woman who protects Ali Baba [16]. She employs intelligent tactics to deceive the thieves. First, the positions of the thieves are initialized, as proposed by Braik et al. [16]:

$$TH^i = lo_j + r \times (u_j - lo_j) \tag{13}$$

where TH^i:Position of the *ith* thief, r:Random number, u_j and lo_j: upper and lower bound of the decision variable. The algorithm considers three assumptions:

1- The thieves gather information in order to locate Ali Baba's house. Hence, they update their location as follows [16]:

$$TH^i_{t+1} = Gbest^t + \left[TD_t\left(best^i_t - AL^i_t\right)r_1 + TD_t\left(AL^i_t - m^t_i\right)r_2\right]\mathrm{sgn}(rand - 0.50)$$
$$r_3 > 0.50$$
$$r_4 > P_{p_t} \tag{14}$$

where TH^i_{t+1}:Location of the *ith* thief, $Gbest^t$:Global best position, $best^i_t$: Local best location, m^t_i : Marjaneh's wit level, AL^i_t:Location of Ali baba, TD_t:Tracking distance of the thieves, P_{p_t}: Perception potential of the thieves, r_1, r_2 , r_3, and r_4:Random numbers.

$$TD_t = \eta_0 e^{-\eta_1\left(\frac{t}{T}\right)} \tag{15}$$

where t:Number of iterations, T:The maximum number of iterations, η_0: Initial estimate of the tracking distance. η_1:Controller parameter.

$$P_{p_t} = \tau_0 \log\left(\tau_1\left(\frac{t}{T}\right)^{\tau_0}\right) \tag{16}$$

where τ_0 represents a final estimation of the probability that the thieves will achieve their goals, and τ is the controller parameter.

2- Marjaneh deceives thieves. The thieves realize that they have been deceived. They randomly change their location [16].

$$TH^i_{t+1} = Td_t\left[(u_j - lo_j)r + lo_j\right], r_4 \le P_{p_t}, r_3 \ge 0.50 \tag{17}$$

Table 3
The optimal parameter values of the LSTM, BI-LSTM, and SVM models.

Model	For predicting TDS	For predicting EC
LSTM	Number of LSTM layers (NLL):2 batch size:128Number of neurons: 20/3/1	Number of neurons: 20/3/1
BI-LSTM	NLL:2 Batch size:128Number of neurons: 24/3/1	NLL:2 Batch size:128Number of neurons: 24/3/1
SVM	Penalty coefficient, 37.20 Kernel parameter:1.02	Penalty coefficient, 37.20 Kernel parameter:1.02

Table 4
The optimal input scenarios for predicting TDS and EC (t-1: previous day, t-2: previous two days, t-10: previous ten days).

Output	Input
One-day ahead TDS	TDS (t-1), TDS (t-2), HCO_3 (t), HCO $_3$ (t-1) Na^{++} (t), Na^{++} (t-1), Mg^{++} (t-1) Na^{++} (t-1)
Two-day ahead TDS	TDS (t-1), HCO_3 (t), Na^{++} (t), Na^{+} (t-1), Mg^{++} (t)
Three-day ahead TDS	TDS (t-1), HCO_3 (t), Na+ (t), Mg++ (t), Na^{++}(t)
One-day ahead EC	Na^{+} (t), Na^{+} (t-1), HCO_3 (t), SO_4 (t), SO_4 (t-1), and Ca^{++} (t), EC (t-2), EC (t-1)
Two-day ahead EC	Na+ (t-1), Na+ (t), HCO_3 (t), SO_4 (t), and Ca++ (t), EC (t-1)
Three-day ahead EC	Na+ (t), Na+ (t-1), HCO_3 (t), SO_4 (t), and Ca++ (t), EC (t-1)

3- The thieves may explore other positions if a better location is not found.

$$TH_{t+1}^{i} = Gbest^{t} - \left[Td^{t}\left(best_{i}^{t} - AL_{i}^{t}\right)r_1 + Td^{t}\left(AL_{i}^{t} - m_{i}^{t}\right)r_2\right]\mathrm{sgn}(rand - 0.50) \tag{18}$$

The binary version of ABFT is employed for feature selection, utilizing Equation (19) to create the binary representation of the algorithm.

$$T\left(TH_{t+1}^{i}\right) = \tanh\left(TH_{t+1}^{i}\right) = \frac{e^{\left(2TH_{t+1}^{i}\right)} - 1}{e^{\left(2TH_{t+1}^{i}\right)} + 1} \tag{19}$$

$$\Delta\left(TH_{t+1}^{i}\right) = \begin{bmatrix} 1 \leftarrow if\left(T\left(TH_{t+1}^{i}\right)\right) \rangle \lambda_t \\ 0 \leftarrow otehrwise \end{bmatrix} \tag{20}$$

2.4. The hybrid models of LSTM-SVM, BI-LSTM-SVM

In this study, we use the BILSTM and LSTM models to predict water quality parameters. The model is created as follows:

1. The data are categorized into training and testing groups for analysis. 65% and 35% of the data were chosen for the training and testing levels. These data sizes were chosen because they yielded the lowest root mean square error (RMSE).
2. The names of the input variables and model parameters are considered decision variables.
3. A binary vector is generated by utilizing the input variable names.
4. A binary vector is represented by unselected and selected features with 0 and 1 values, respectively.
5. The binary vectors are fed into the BILTM model as inputs.
6. The model is trained using the training data to predict water quality parameters.
7. Once the stop criteria (SC) are satisfied, the process moves to step 8. If the SC criteria are unmet, the model is connected to the ABFT. The model parameters are initially defined as the population of the ABFT. The BILSTM (LSTM) model is employed to predict the output, and the accuracy of the models is evaluated using the RMSE metric. The binary vectors, which contains the names of the input variables and model parameters, are updated using equations 14–18. The optimization process continues until the Convergence Creation Criterion (CCR) is satisfied. Once the CCR is met, the models proceed to step 8.
8. The BILSTM (LSTM) model outputs are utilized as inputs to the SVM model.
9. At the training level, the SVM model is executed, and if the SC is met, the model proceeds to step 10; otherwise, it is linked to the ABFT. The model parameters are initialized as the population of the ABFT. The SVM model predicts the output, and the accuracy of the models is evaluated using the RMSE metric. Equations 14–18 are employed to update the SVM parameters. If the CCR is satisfied, the model advances to step 10; otherwise, the optimization process continues.
10. The SVM model is executed at the testing level.

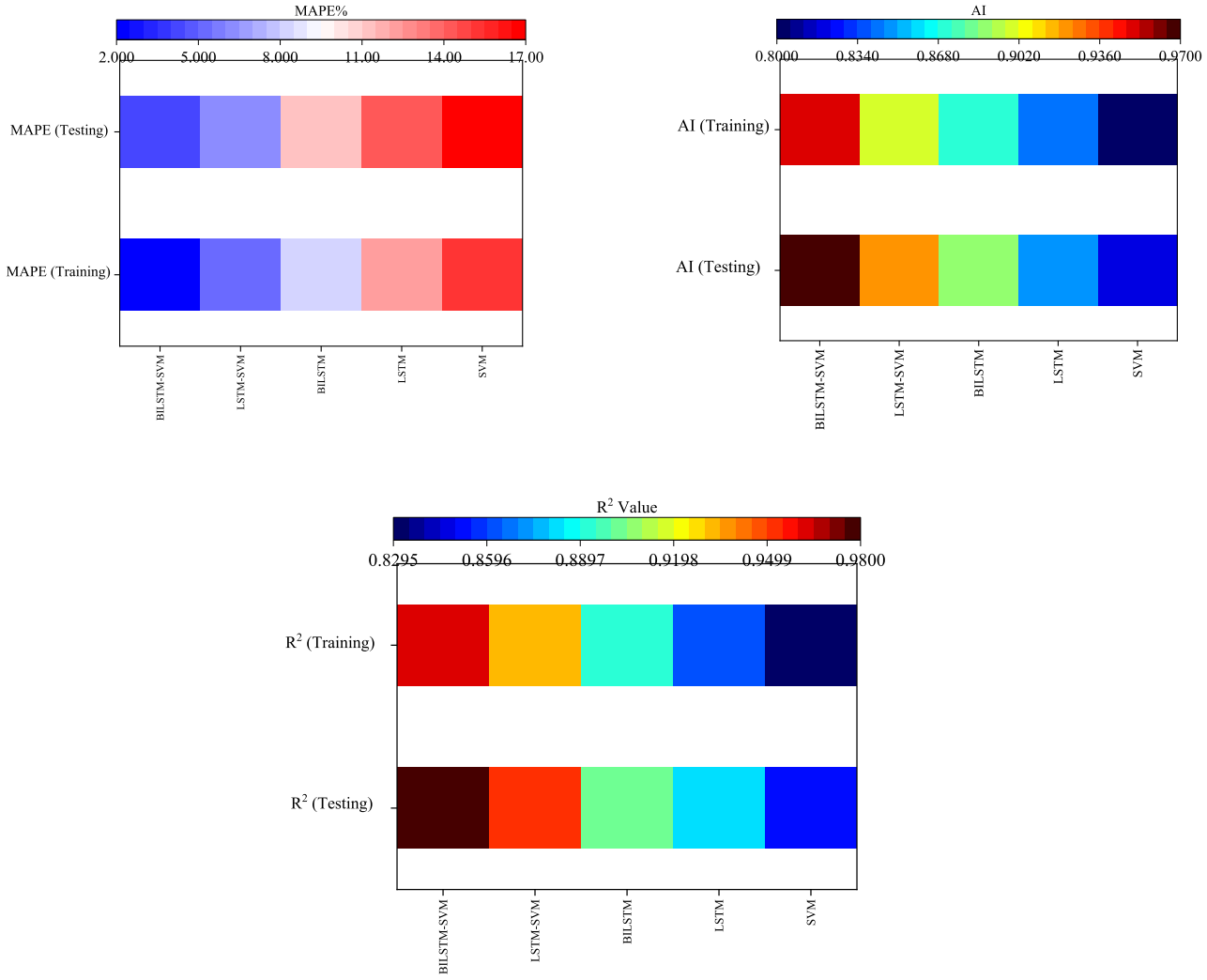
2.5. Benchmark models

2.5.1. Multilayer Perceptron models (MLP)

The Multilayer Perceptron (MLP) is a crucial type of artificial neural network (ANN) model. Within the MLP architecture, neurons serve as the fundamental processing units [20]. This architecture typically comprises three essential layers. The first layer, known as the Input Layer, plays the role of receiving and transmitting raw input data to subsequent layers. An interesting aspect of MLPs is their ability to incorporate one or more hidden layers [21]. These hidden layers are where the network's complexity and learning capabilities are derived. The final layer in the MLP structure is responsible for generating the model's predictions. The process through which the MLP produces these predictions is known as forward propagation. During forward propagation, input data traverses through the layers of the network, undergoing a series of transformative operations [22]. In this process, each neuron within a hidden layer receives the weighted sum of activations from the neurons in the preceding layer. Once this weighted sum is computed, the neuron applies a specific activation function (e.g., sigmoid, tanh, ReLU) to calculate its output. This comprehensive process defines how the MLP model operates, allowing it to make valuable predictions based on its learned patterns.

2.5.2. Radial basis function neural network models (RBFNN)

The Radial Basis Function Neural Network (RBFNN) is a specific type of Artificial Neural Network (ANN) model characterized by a three-layer architecture, as described in reference [23]. These layers include an input layer, a hidden layer equipped with radial basis functions, and an output layer. In the RBFNN's structure, the input layer is comprised of neurons that correspond to the various input features. Input data is directly supplied to these neurons. The collective activation values of all the neurons within the hidden layer together constitute the output of this hidden layer, as outlined in reference [24]. The configuration of the output layer is adaptable, depending on the specific problem being addressed [25]. It may consist of one or multiple neurons, and it plays a crucial role in generating the final predictions of the RBFNN. One of the distinctive features of the RBFNN lies in how the activation of each radial basis function neuron is determined. This activation is contingent upon the similarity or distance measurement between the input data and



(a)

Fig. 2. Error function values for one, two, and three-day ahead TDS predictions. (a) one-day ahead TDS prediction, (b) two-day ahead TDS prediction, and (c) three-day ahead TDS prediction.

the center point associated with that particular neuron, as elucidated in reference [26]. This unique mechanism underscores the RBFNN's ability to make similarity assessment predictions.

2.5.3. *Recurrent neural network models (RNN)*

Recurrent Neural Networks (RNNs) represent a specialized class of Artificial Neural Network (ANN) models meticulously designed for the analysis of sequential data. Diverging from the conventional feedforward neural networks, RNNs incorporate cyclical connections within their architecture, which equip them with the unique capacity to manage and process sequences of variable lengths, as detailed in reference [27]. Fundamental to the RNN's operation is an internal hidden state, serving as a form of memory. This hidden state undergoes continual updates at each time step and is under the influence of both the incoming input data and the hidden state from the previous time step. Conceptually, the architecture of an RNN can be envisioned as an interlinked chain of repeating modules or cells. These recurrent connections within the RNN structure facilitate the preservation of information from previous time steps. At any given time *t* the hidden state *t-1* is computed as a function of the current input *x_t* at time *t* and the preceding hidden state *h_(t-1)* at time *t-1* employing a set of learned weights and an activation function. This sequential data processing mechanism culminates in the RNN's ability to systematically process input data as it unfolds in a temporal sequence [28]. Specifically, the input *x_t* is fused with the hidden state from the previous time step *h_(t-1)* to compute the new hidden state *h_t*, a process underpinned by a set of adaptively learned weights and an activation function. Following the update of *h_t*, the RNN is poised to generate an output corresponding to the current time step *t*. This description encapsulates the essence of how RNNs operate in the context of sequential data analysis.

3. Case study

This study aims to predict the water quality parameters of a coastal aquifer in north Iran, where humid and sub-humid climates prevail. Annual rainfall varies from 718 to 880 mm in the basin. Groundwater supplies 63% of all water needs for domestic, agricultural, and industrial purposes. Groundwater covers 85% of agricultural needs and 75% of drinking water needs. Since groundwater plays a crucial role in the water supply of the basin, predicting groundwater quality is an

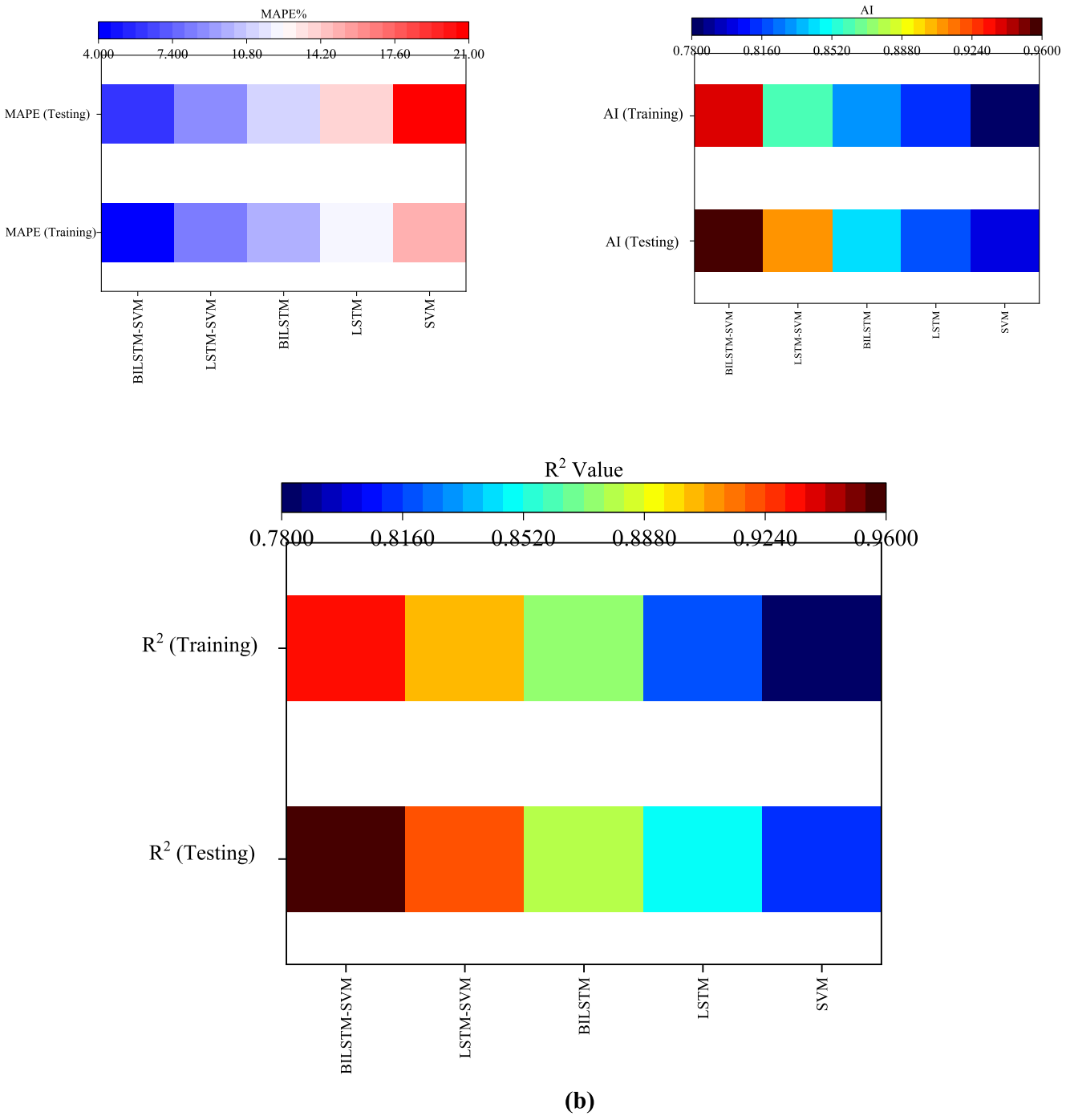


Fig. 2. (*continued*).

important topic within this basin. There are four hydrogeological units in this area that are characterized by unconsolidated sediments. The uppermost level contains a calcareous unit of sand and gravel that forms a shallow unconfined aquifer. This aquifer has an average thickness of 25 m and a mean hydraulic conductivity of 4 × 10–4 m per second. The semi-confined aquifer consists of gravel and fine to medium-sized sand and has an average thickness of 150 m and a hydraulic conductivity of 4 × 10—–5 m per second. This study focuses on predicting two of the most critical water quality parameters, namely TDS and EC. We collected daily data from 2017 to 2020. Lagged values of PH, Ca^{++}, Mg^{++}, Na^{+}, K^{+}, HCO_3, SO_4, and Cl were used to predict to predict 1-, 2- and 3-day ahead TDS and EC values. Table 1 shows detailed information about the input data. The Iranian water resource management company provided the data, and the monitoring wells were the source of daily data. Fig. 1a, 1b, and 1c illustrate the study area's location, TDS and EC data, respectively [29]. Figure 1a was extracted from [29].

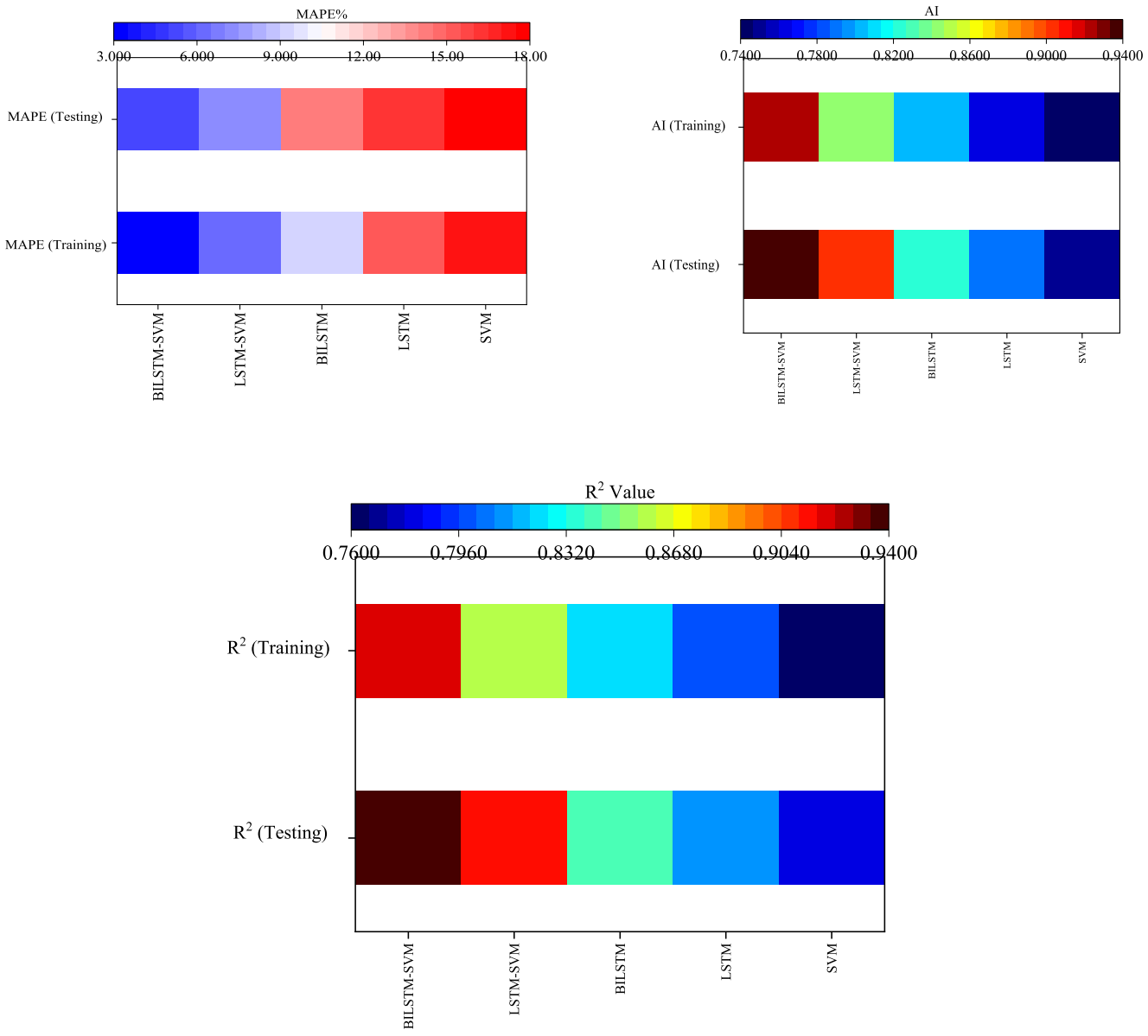
This study uses equations 21–24 to evaluate the accuracy of models.

1- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(OB_i - ES_i)^2} \tag{21}$$

2- Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|ES_i - OB_i|}{ES_i} \tag{22}$$



(c)

Fig. 2. (*continued*).

3- Index of Agreement (IA):

$$IA = 1 - \frac{\sum_{i=1}^{N}(OB_i - ES_i)^2}{\sum_{i=1}^{N}(|OB_i - \bar{OB}| + |ES_i - \bar{ES}|)^2} \quad (23)$$

4- Coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(ES - OB)^2}{\sum_{i=1}^{N}(OB - \bar{OB})^2} \quad (24)$$

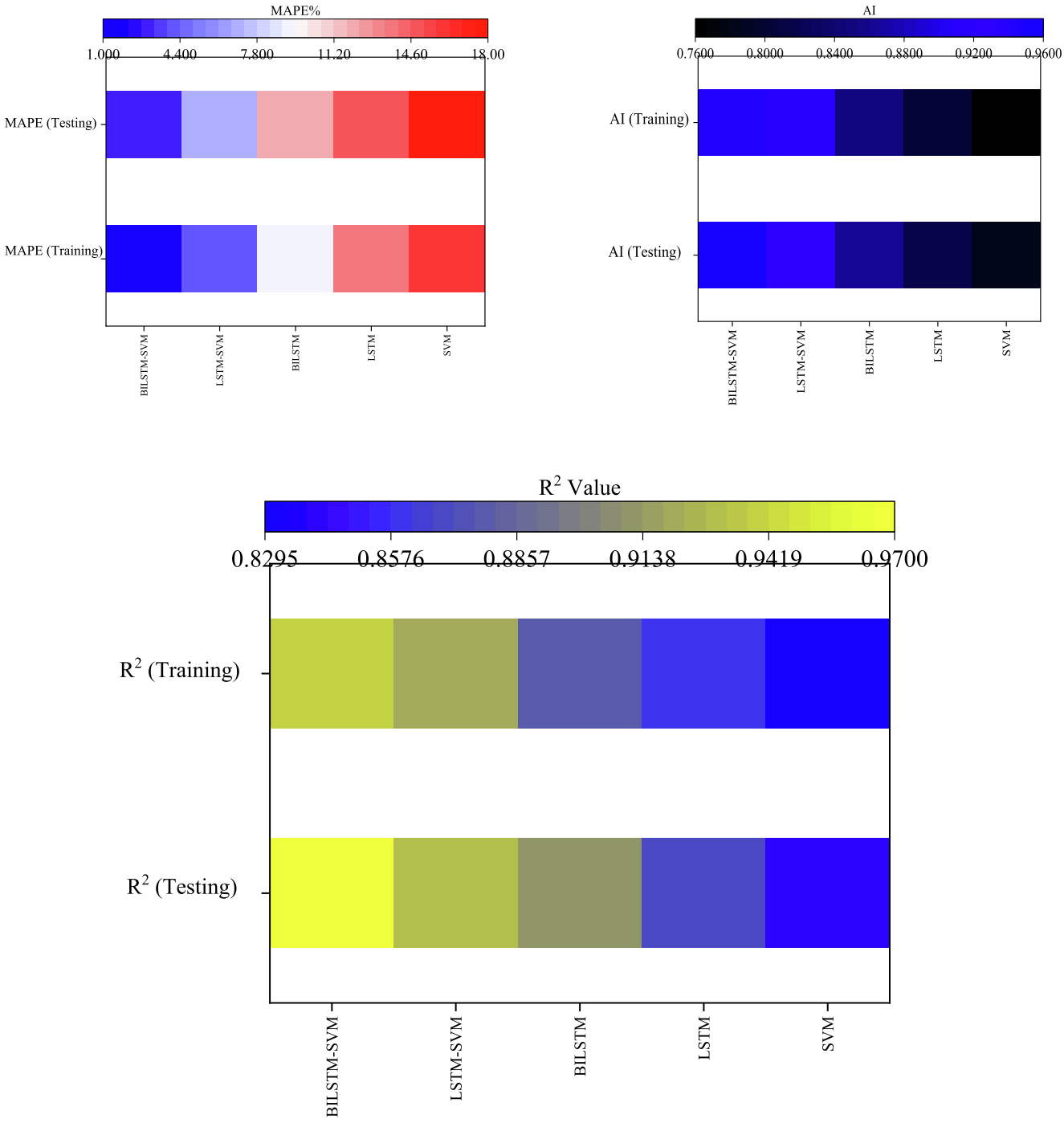
4. Results and discussion

4.1. Determination of model parameters and random parameters

The ABFT algorithm is employed to fine-tune the parameters of BILSTM, LSTM, and SVM models. The population size (PSI) and the Maximum Number of Iterations (MNI) are considered random parameters of the ABFT. Optimal values of these parameters must be determined to achieve the best results. Tables 2a and 2b present the optimal parameter values for predicting TDS and EC. The parameter values are varied within a specified range, and the optimal values are chosen to minimize the error function (RMSE). According to Table 2a, the population size varied between 45 and 180, and the and the optimal value was 90, which leaded to the lowest RMSE value. Similarly, the MNI varied between 70 and 280, and the optimal value was 140. Table 3 provides the optimal values of model parameters.

4.2. Input selection

A binary version of SSA was employed to identify the optimal input combinations for predicting TDS and EC. The names of the input variables were decision variables. The lag times varied between (t-1) and (t-10). As there were 2^{80}-1 combinations of inputs, determining the



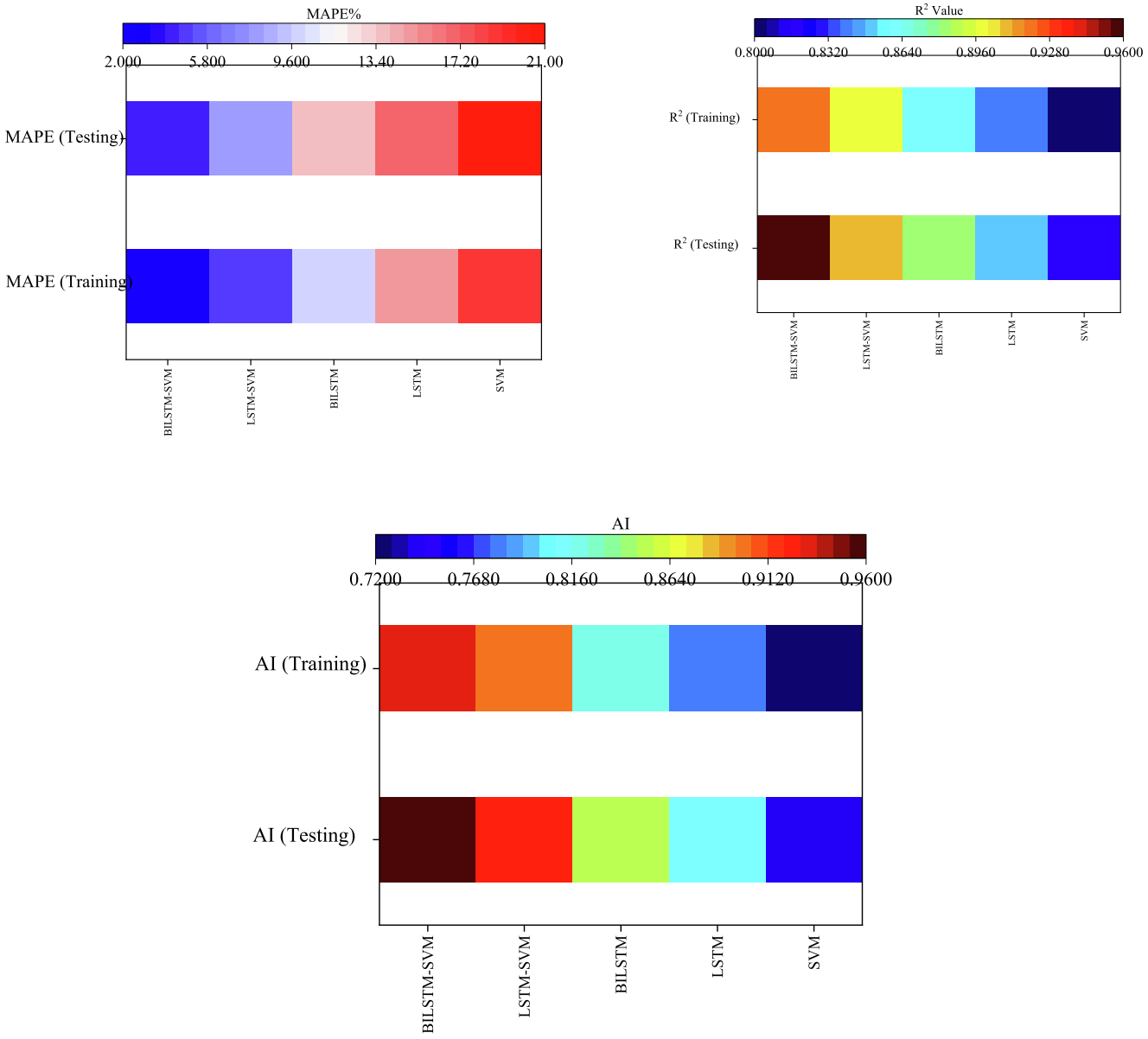
(a)

Fig. 3. Error function values for one-day, two-day, and three-day ahead EC predictions. (a) one-day ahead EC prediction, (b) two-day ahead EC prediction, and (c) three-day ahead EC prediction.

optimal input scenario was complex. The Binary SSA automatically identified the optimal input scenarios, while correlation methods and principal component analysis identified the significant inputs. The optimal input scenarios are presented in Table 4.

4.3. Accuracy assessment of models

Fig. 2a presents the evaluation of TDS prediction accuracy at the one-day prediction horizon. The MAPE values of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models were 2, 5, 8, 12, and 15, respectively, at the training level. The testing MAPE values for the same models were 4, 6, 11, 14, and 17, respectively. Additionally, the

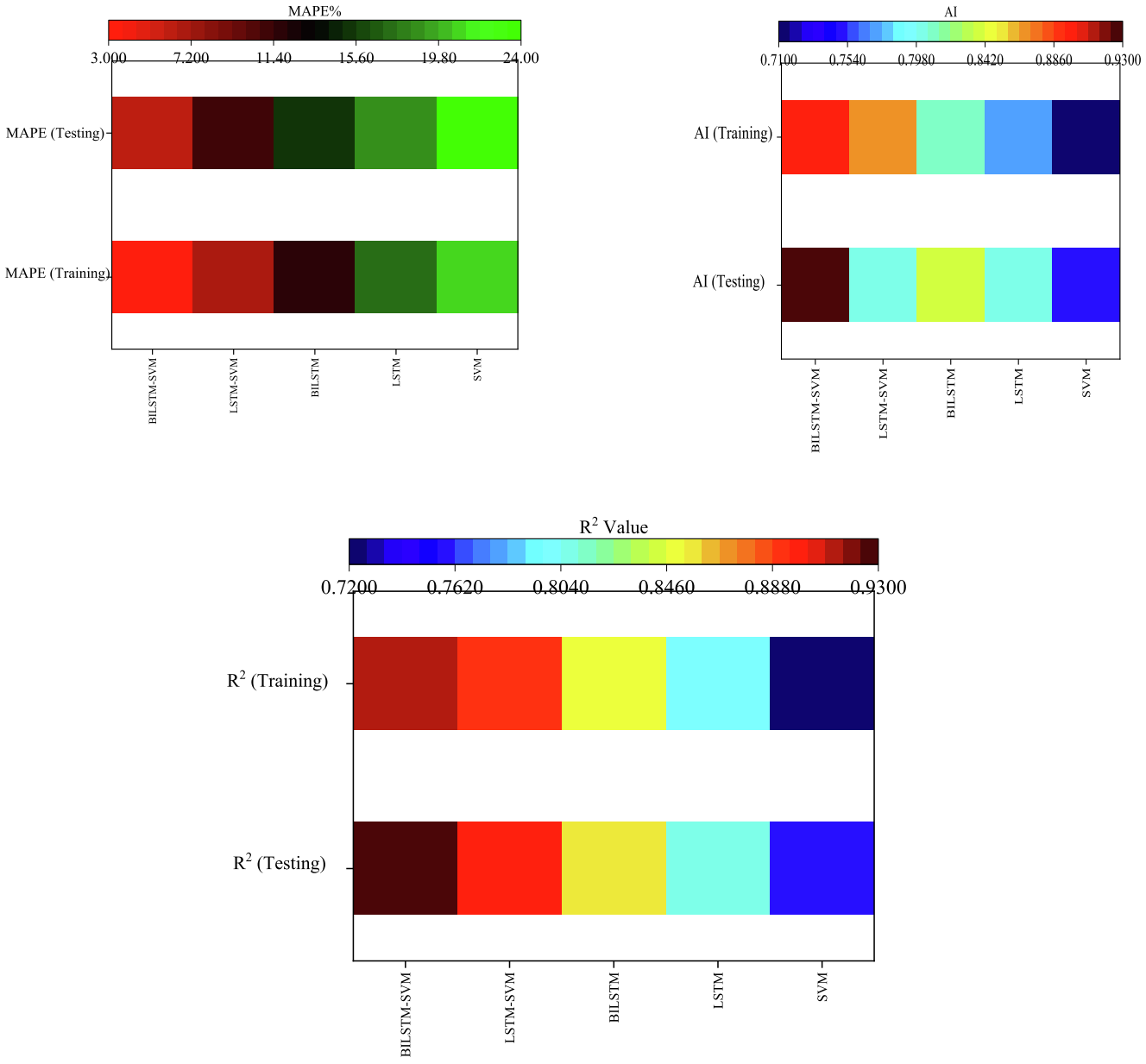


(b)

Fig. 3. (*continued*).

accuracy index (AI) values of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models were 0.97, 0.92, 0.89, 0.85, and 0.82, respectively, at the training level. The key findings are: (1) The BILSTM-SVM model performed better than the LSTM-SVM model because it considered both past and future observations. (2) The hybrid model (BILSTM-SVM and LISTM-SVM) performed better than the LSTM, BILSTM, and SVM models. (3) The SVM model had lower AI and R^2 values compared to other models, and all other models outperformed the SVM model. (4) The BILSTM model had higher AI and R^2 values compared to the LSTM model. Fig. 2b assesses the precision of TDS forecasts at the two-day prediction horizon. The key findings of this assessment are (1) The BILSTM-SVM model outperformed the LSTM-SVM model and showed the lowest MAPE values at both the training and testing levels. Specifically, the BILSTM-SVM model achieved training and testing MAPEs of 3% and 5%, respectively. (2) The MAPEs of the LSTM-SVM model were higher than those of the BILSTM-SVM model at both the training and testing levels. The LSTM-SVM model achieved training and testing MAPEs of 6% and 7%, respectively. Consequently, the BILSTM-SVM model was superior to the LSTM-SVM model. (3) The BILSTM model exhibited the greatest accuracy among standalone models. At the testing level, the AI values of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models were 0.94, 0.86, 0.83, 0.81, and 0.78, respectively.

Fig. 2c assesses the precision of TDS predictions for a three-day prediction horizon. The following results are obtained from this assessment (1) The BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models had training MAPE values of 6%, 9%, 11%, 14%, and 21%, respectively. The testing MAPE of BILSTM-SVM was 4%, 6%, 8%, and 11% lower than that of the LSTM-SVM, BILSTM, LSTM, and SVM models, respectively. The BI-LSTM model showed the best performance for predicting TDS three days ahead. (2) The BILSTM-SVM model had training and testing AI values of 0.94 and 0.92, respectively. The BI-LSTM-SVM model outperformed the LSTM-SVM models. (3) The SVM model had a testing R^2 value of 0.76, while the LSTM and BILSTM models had R^2 values of 0.80 and 0.82, respectively. The BILSTM model



(c)

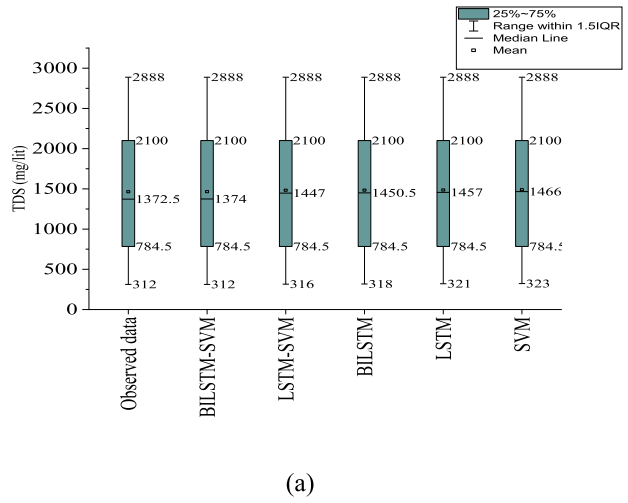
Fig. 3. (*continued*).

outperformed the LSTM and SVM models.

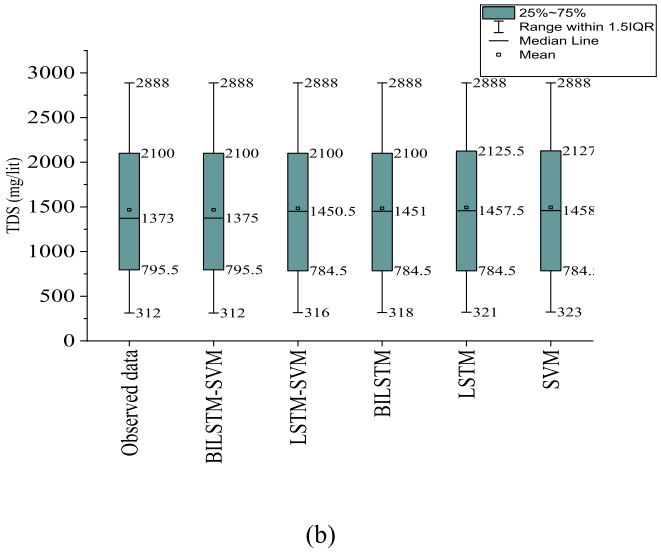
The evaluation of EC prediction accuracy at a one-day prediction horizon is shown in Fig. 3a. The key findings are observed as (1) The BILSTM-SVM, and LSTM-SVM models exhibited superior performance, with R^2 values of 0.94 and 0.92, respectively, at the testing level. In contrast, the BILSTM, LSTM, and SVM models had R^2 values of 0.88, 0.86, and 0.86, respectively. (2) The BILSTM-SVM model demonstrated the highest accuracy with training and testing AIs of 0.96 and 0.94, respectively. The LSTM-SVM model had training and testing AIs of 0.93 and 0.91, respectively. Thus, the BILSTM-SVM model outperformed the LSTM-SVM model. (3) MAPE of the BILSTM model was 1 and 4% lower than that of the LSTM and SVM models. The SVM model exhibited the poorest performance with training and testing MAPEs of 16% and 18%, respectively.

Fig. 3b illustrates the evaluation of the accuracy in the EC predictions at the two-day prediction horizon. The key finding related to the evaluations are (1) The MAPE value of the BILSTM-SVM model was 4, 10, 13, and 17% lower than that of the LSTM-SVM, BILSTM, LSTM, and SVM model at the testing level. The BILSTM-SVM models had the best precision among other models. (2) The R^2 values of the BILSTM, LSTM, and SVM models were 0.85, 0.82, and 0.80, respectively, at the testing level. Notably, the SVM model exhibited lower accuracy than the other models. (3) Training and Testing AIs of the BILSTM-SVM model were 0.93 and 0.92, respectively. BILSTM-SVM model outperformed the LSTM-SVM model.

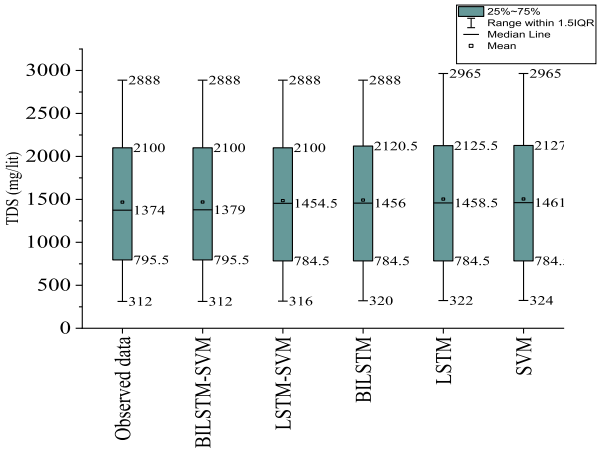
The evaluation of EC prediction accuracy at a three-day prediction horizon is shown in Fig. 3c. According to this evaluation, we observe that (1) At the training level, the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models had MAPEs of 3%, 7%, 12%, 17%, and 21%, respectively. Notably, the BILSTM-SVM models decreased the MAPEs of



(a)

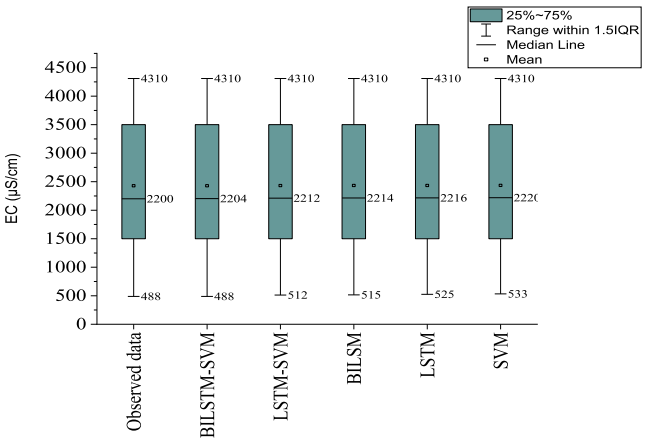


(b)

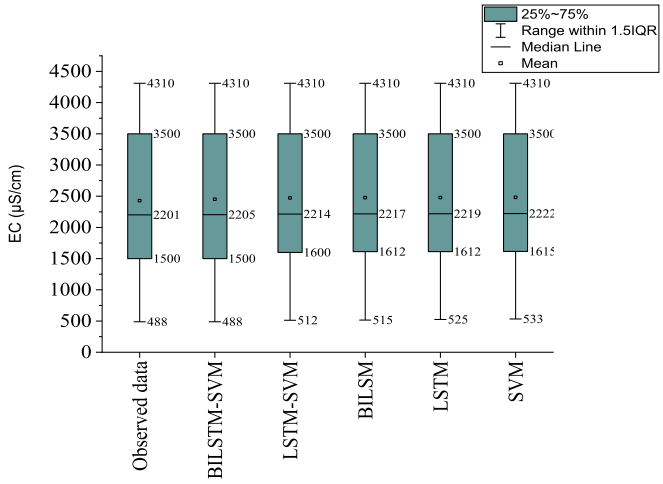


(c)

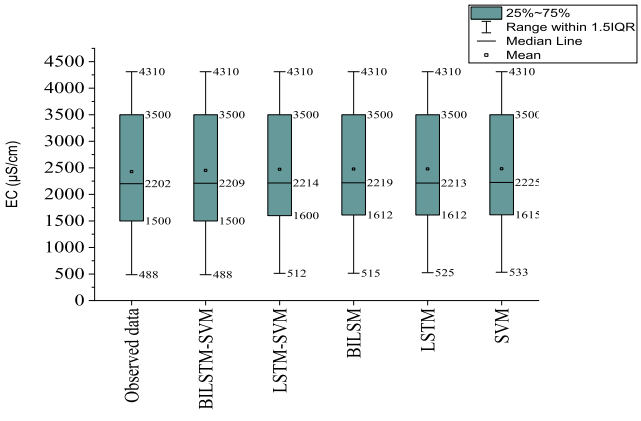
Fig. 4. Boxplots of models using TDS data at the one, two, and three-day prediction horizons. (a) one-day, (b) two-day, and (3) three-day prediction horizons.



(a)

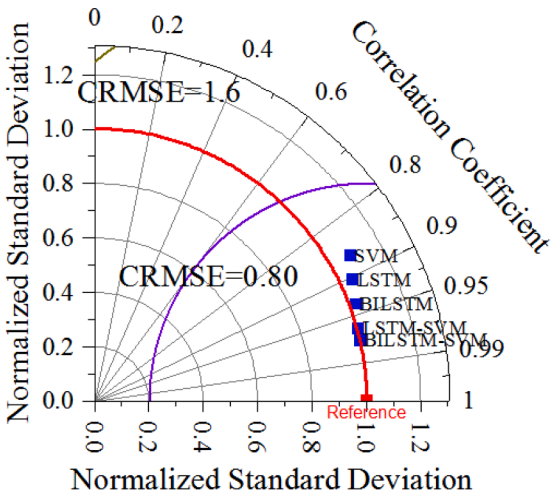


(b)

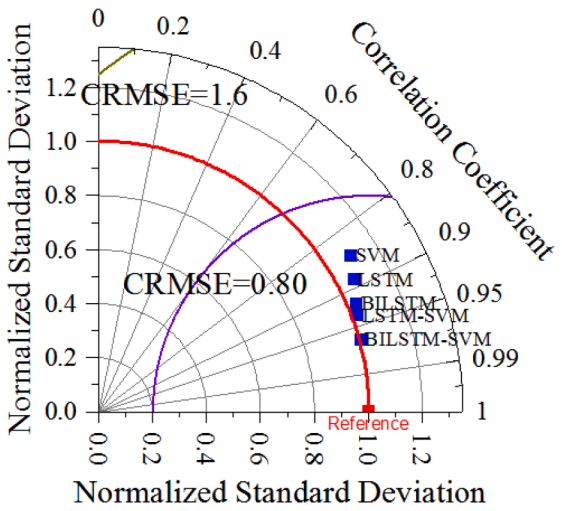


(c)

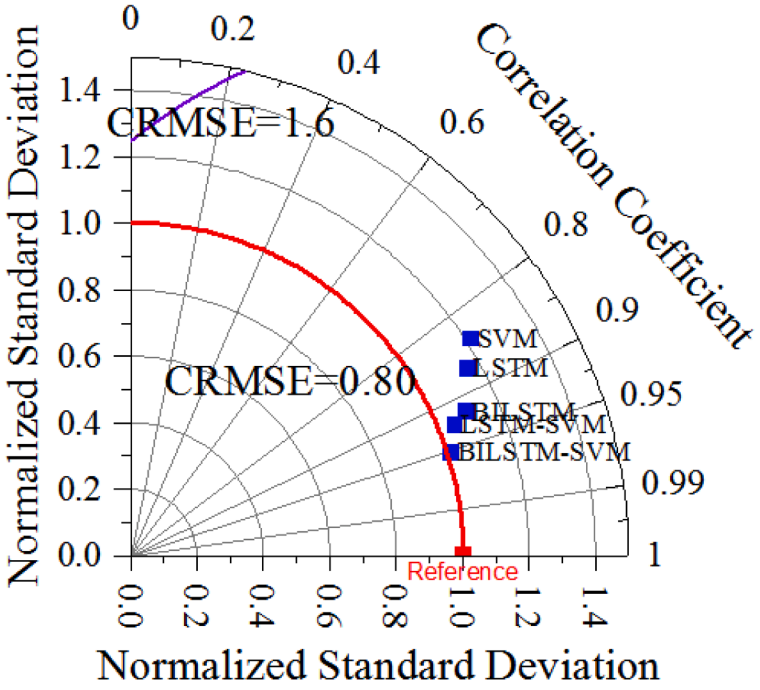
Fig. 5. Boxplots of models using EC data at the one, two, and three-day prediction horizons. (a) one-day, (b) two-day, and (3) three-day prediction horizons.



One –day ahead



two –day ahead



Three- day ahead

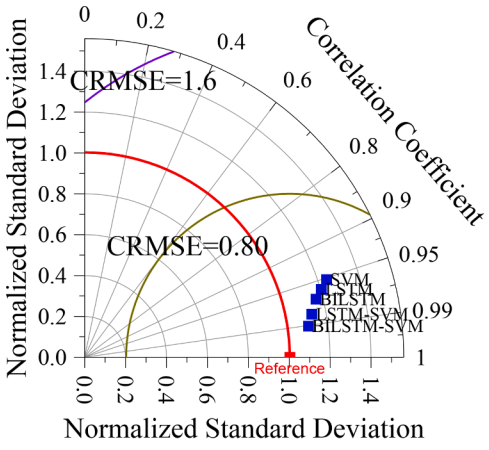
(a)

Fig. 6. Taylor diagram for assessing the models' accuracy using TDS and EC values. (a) TDS values, and (TDS) values.

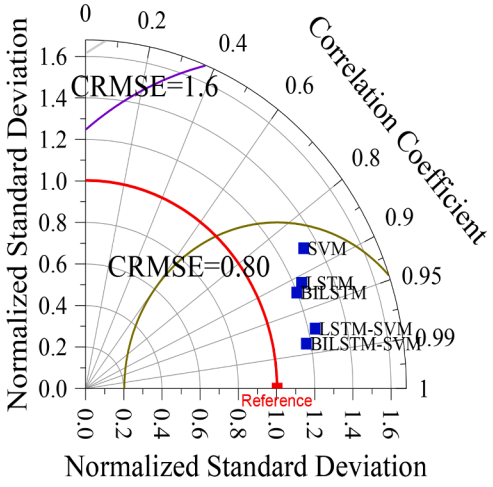
the LSTM-SVM, BILSTM, LSTM, and SVM models by 5%, 9%, 13%, and 18%, respectively. (2) At the testing level, the AI values of the BILSTM, LSTM, and SVM models were 0.88, 0.77, and 0.71. The BILSTM model performed better than the LSTM and SVM models. (3) Training and Testing R^2 values of BILSTM-SVM models were 0.93 and 0.91, respectively. Training and Testing R^2 values of LSTM-SVM models were 0.90 and 0.89, respectively.

Overall, the findings demonstrate that the accuracy of models decreased with an extension of the prediction time. The following results confirm this issue.

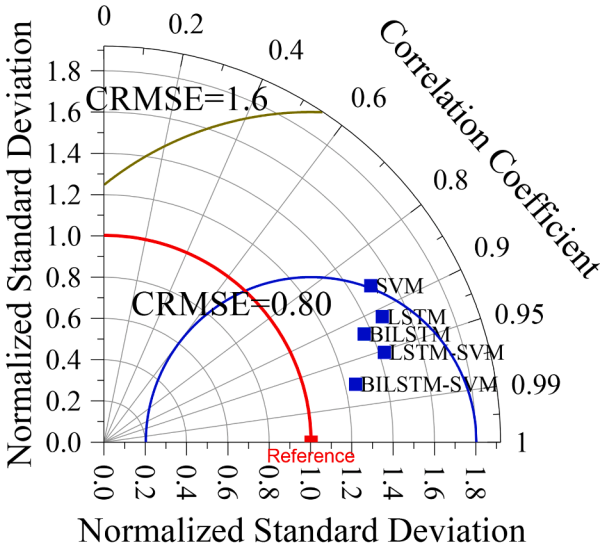
- **TDS prediction**



One- day ahead



Two-day ahead



Three –day ahead

(b)

Fig. 6. (*continued*).

The BILSTM-SVM model exhibited MAPE values of 4%, 5%, and 7% at one, two, and three-day prediction horizons, respectively. Furthermore, R^2 values for the BILSTM-SVM model were 0.96, 0.93, and 0.92 at the same prediction horizons. Additionally, the AI values for the BILSTM-SVM model were 0.95, 0.94, and 0.92 at one, two, and three-day prediction horizons, respectively.

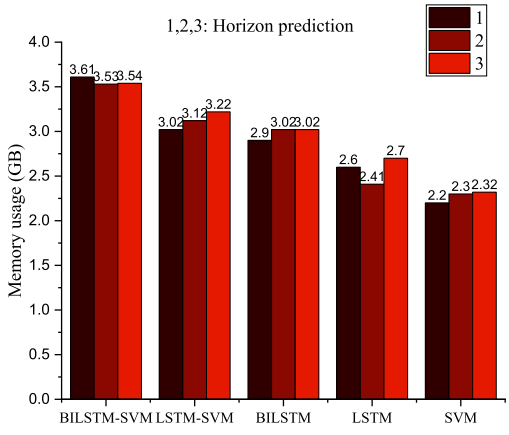
- **EC prediction**

At the one, two, and three-day prediction horizons, the MAPE values for the BILSTM-SVM model were 3%, 4%, and 6%, respectively. The R^2 values for the same prediction horizons were 0.94, 0.92, and 0.91. Additionally, the AI values for the BILSTM-SVM model were 0.94, 0.92, and 0.90, at one, two, and three-day prediction horizons, respectively.

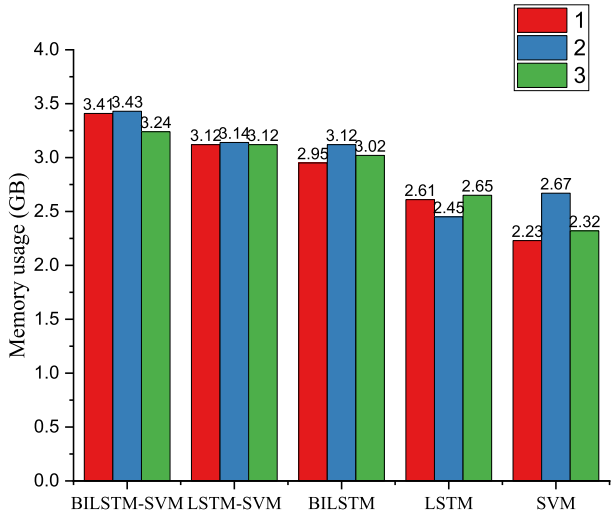
Boxplots at the one, two, and three-day prediction horizons are presented in Fig. 4a, 4b, and 4c, respectively. As shown in Fig. 4a, the median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 1372.5 mg/lit, 1374.0 mg/lit,

Table 5
Comparison of the accuracy of the BILSTM-SVM models with the previous models.

Author	Model	Parameter	Accuracy
Ravansalar and Rajaee [30]	Wavelet ANNANN	EC	R^2 (Wavelet-ANN) = 0.94 R^2(ANN) = 0.38
Ahmadianfar et al. [31]	Wavelet-support vector machine	EC	R^2 = 0.95
Banadkooki et al. [32]	Adaptive neuro fuzzy interface system- moth lame optimization	TDS	AI = 0.94
Jamei et al. [33]	wavelet pre-processing with multigene genetic programming (W-MGGP)	TDS	AI = 0.86
Ahmadpour et al. [34]	Seasonal autoregressive integrated moving average (SARIMA)- ANN	EC	R^2-0.86



(a)



(b)

Fig. 7. Memory usage for predicting a: EC and b: TDS.

Table 6
GPI values and Rank of models.

		TDS			EC		
		Horizon prediction			Horizon prediction		
Models	Rank	1	2	3	1	2	3
BILSTM-SVM	1	0.482	0.412	0.408	0.467	0.412	0.401
LSTM-SVM	2	0.308	0.307	0.305	0.323	0.331	0.320
BILSTM	3	0.112	0.110	0.104	0.298	0.234	0.222
LSTM	4	0.109	0.107	0.102	0.267	0.230	0.228
SVM	5	0.106	0.104	0.098	0.116	0.108	0.106

1447 mg/lit, 1450.5 mg/lit, 1457 mg/lit, and 1466 mg/lit. Fig. 4b illustrates that the median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 1373.0 mg/lit, 1375 mg/lit, 1450.5 mg/lit, 1451.0 mg/lit, 1457.5 mg/lit, and 1458 mg/lit, respectively. As shown in Fig. 4c, the median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 1374.0 mg/lit, 1379.0 mg/lit, 1454.5 mg/lit, 1456.0 mg/lit, 1458.5 mg/lit, and 1461 mg/lit, respectively.

Fig. 5 shows boxplots of observed and predicted EC values. Fig. 5a shows boxplots at the one-day prediction horizon. The median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 2200 mg/lit, 2204 mg/lit, 2212 mg/lit, 2214 mg/lit, 2216 mg/lit, and 2220 mg/lit. The minimum value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 488 mg/lit, 488 mg/lit, 512 mg/lit, 515 mg/lit, 525 mg/lit, and 533 mg/lit, respectively. Fig. 5b shows boxplots at the two-day prediction horizon. The median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 2201 mg/lit, 2205 mg/lit, 2214 mg/lit, 2217 mg/lit, 2219 mg/lit, and 2222 mg/lit. Fig. 5c shows boxplots at the three-day prediction horizon. The median value of the observed TDS data, BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models was 2202 mg/lit, 2209 mg/lit, 2214 mg/lit, 2217 mg/lit, 2219 mg/lit, and 2222 mg/lit.

A Taylor diagram is a visual tool for assessing the accuracy of a model. It measures the central root mean square error (CRMSE), standard deviation, and correlation coefficient to evaluate the accuracy of the models. An ideal model will be closer to the reference point (observed data). Fig. 6a evaluates the accuracy of TDS predictions. At the 1- day prediction horizon, the CRMSE of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, SVM model was 0.2, 0.24, 0.32, 0.40, and 0.48. The correlation coefficient (CC) of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, SVM model was 0.96, 0.93, 0.92, 0.88, and 0.85 at the 2-day prediction horizon. The CC of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, SVM model was 0.95, 0.92, 0.91, 0.87, and 0.84 at the 3-day prediction horizon. Thus, BILSTM-SVM and LSTM-SVM model had the best performance. Fig. 6b depicts the assessment of EC prediction accuracy. At the one-day prediction horizon, the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models had a CRMSE of 0.15, 0.21, 0.27, 0.32, and 0.37, respectively. At the two-day prediction horizon, the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models had a CRMSE of 0.23, 0.31, 0.42, 0.47, and 0.61, respectively. The CC of the BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, SVM model was 0.97, 0.95, 0.92, 0.91, and 0.86 at the 3- day prediction horizon.

5. Discussions

In this article, hybrid models are used to predict EC and TDS. Specifically, the study combines BILSTM and LSTM models with SVMs to improve the prediction accuracy. The BILSTM model was employed to extract both linear and nonlinear patterns, which were then sent to an SVM model. Combining a BILSTM model with an SVM model improved the efficiency of the SVM model. Notably, the BILSTM-SVM model, which processed information in both directions, demonstrated superior performance compared to LSTM. Furthermore, this research study

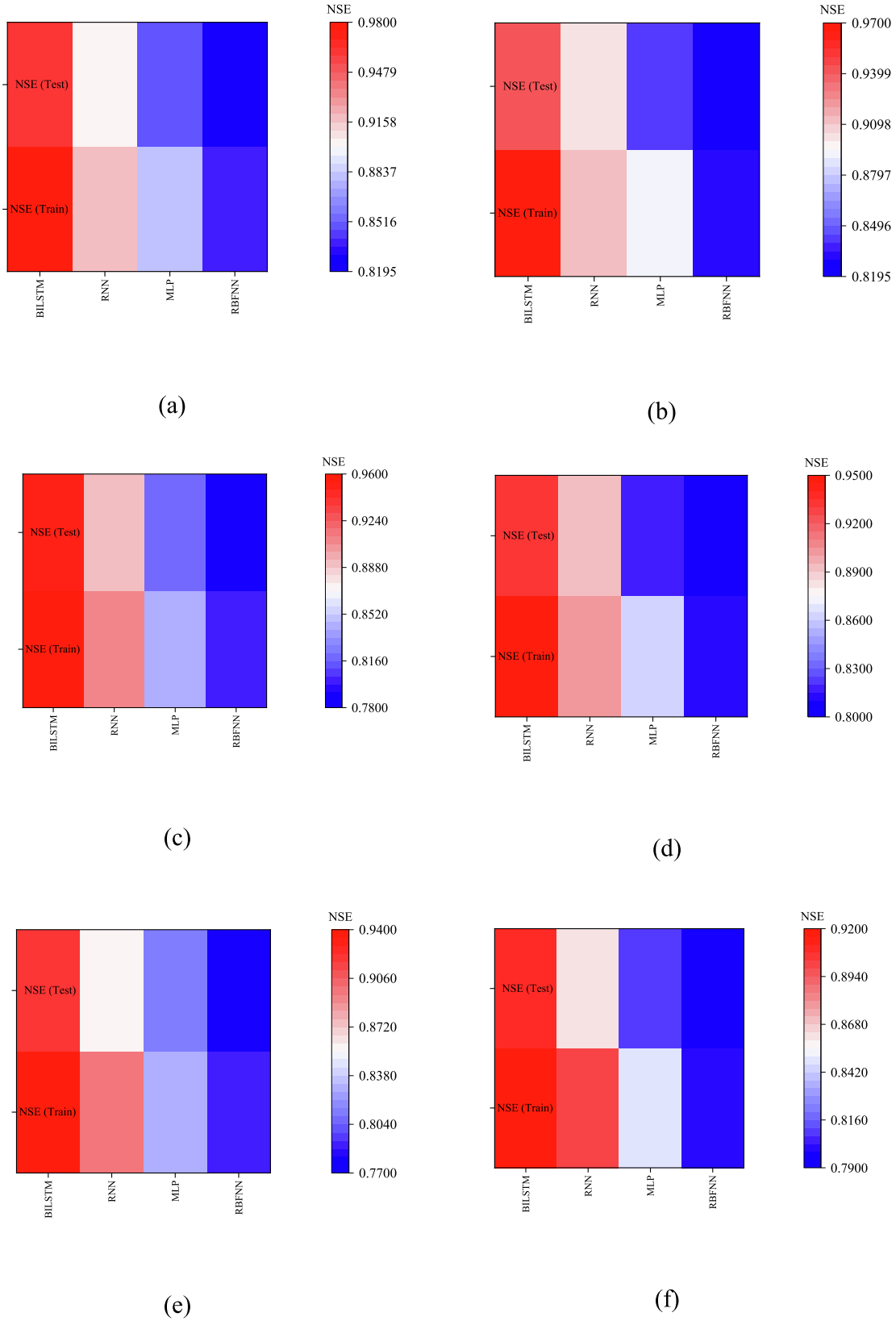


Fig. 8. TDS prediction (one-day ahead), b: EC prediction (one-day ahead), c: TDS prediction (two-day ahead), d: EC prediction (two-day ahead), e: TDS prediction (three-day ahead), and f: EC prediction (three-day ahead).

employed the binary version of ABFT to accurately extract optimal features from time series data. As our approach automatically selects input combinations, it is more accurate than random selection. ABFT-BILSTM-SVM was proposed as a powerful tool for data preprocessing and prediction. The findings revealed that the accuracy of models decreased with an increase in prediction time. However, the time series data may have more complex patterns over two and three days, which should be considered for future research. While our models successfully predicted data points, they were unable to predict time intervals. Future studies can combine models with Bayesian approaches to predict time intervals. Table 5 provides a comparative analysis of the current study with previous research efforts. Ravansalar and Rajaee (2015) used wavelet ANN and ANN models to predict EC. R^2 vales of wavelet ANN and ANN were 0.94 and 0.38. R^2 value of the BILSTM-SVM model was 0.97. Ahmadianfar et al. (2020) used the wavelet-SVM model to predict EC. R^2 of the wavelet-SVM was 0.95. Thus, BILSTM-SVM model outperformed the wavelet- SVM model. Banadkooki et al. (2020) used an adaptive neuro fuzzy interface system-moth flam optimization (ANFIS-MFO) to predict TDS. The AI value of the ANFIS-MFO was 0.94. The AI value of the BILSTM-SVM model was 0.97.

The developed models can serve as early warning systems and warn authorities when the values of water quality parameters exceed the permissible limit. In addition, the BILSTM SVM model can predict past and future values of water quality parameters. These capabilities can significantly enhance the efficacy and efficiency of water quality management efforts. For an unbiased comparison, we should consider the complexity and computational costs of hybrid models. Fig. 7a and 7b display the memory usage of all models for estimating EC and TDS. Compared to other models, hybrid models, such as the BILSTM-SVM and the LSTM-SVM, used more memory. SVM models use less memory because they are simple models. The hybrid model used more memory than the SVM model. The SVM model decreased memory usage of the BILSTM-SVM model, LSTM-SVM, BILSTM, and LSVM model by 28%, 26%, 23%, and 12% for predicting EC at one-day prediction horizon. Similarly, the SVM model reduced the memory usage of all models used to predict EC and TDS at other horizon predictions. Hybrid models have higher accuracy than standalone models, but their memory usage is greater than hybrid models. Now, we should choose a model as the best model. In order to determine the best model, we use a global performance index (GPI) to determine the rank of models based on MAPE, R^2, AI, correlation coefficient, standard deviation, memory usage, and CRMSE.

$$GPI_i = \sum_{j=1}^{n} \phi_j \left(X_{ij} - \bar{X}_j\right) \tag{25}$$

where n: number of inputs, X_{ij}: The scaled value of index j for model i, $\bar{X}_j$: The median of scaled values of index and ϕ_j of NSE, R^2, correlation coefficient: 1, and ϕ_j of other indices:-1. Table 6 determines the best rank of models and GPI values. The best model has the highest value of GPI. As can be seen from this table, the hybrid models have a better rank than standalone models.

In Fig. 8a, the NSE values for one-day-ahead Total Dissolved Solids (TDS) prediction are presented, showcasing the performance of the best model, BILSTM-SVM, in comparison to benchmark models. Notably, during training, the BILSTM-SVM achieved an NSE of 0.98, while the RNN, MLP, and RBFNN attained NSE scores of 0.92, 0.88, and 0.84, respectively. It is evident that the BILSTM-SVM model outperformed the benchmark models during testing, with improvements of 6.66, 9.10, and 10.24 in NSE compared to RNN, MLP, and RBFNN, respectively. Fig. 8b presents NSE values for one-day-ahead Electrical Conductivity (EC) prediction. Here, the BILSTM-SVM model again demonstrated its superiority by enhancing testing NSE scores over the RNN, MLP, and RBFNN by 6.02, 10.60, and 11.24, respectively. Moving on to Fig. 8c, which showcases NSE values for two-day-ahead TDS prediction, the BILSTM-SVM model displayed robust performance with training and testing NSE scores of 0.96 and 0.95, respectively. Fig. 8d provides NSE values for two-day-ahead EC prediction, revealing that the BILSTM-SVM model consistently outperformed the benchmark models. The training and testing NSE scores for the BILSTM-SVM model were 0.95 and 0.94, respectively. In Fig. 8e, NSE values for three-day-ahead Total Dissolved Carbon (TDC) prediction are depicted. Notably, the BILSTM-SVM model substantially improved both training and testing NSE scores, surpassing the RBFNN model by 15% and 16%, respectively. Finally, Fig. 8f presents NSE values for three-day-ahead EC prediction. Here, the training NSE of the BILSTM model surpassed that of the other models, further highlighting its effectiveness.

6. Conclusions

Accurately predicting water quality is necessary for sustainable freshwater production. This study proposes a novel hybrid deep learning model for predicting water quality parameters. Specifically, a BILSTM model was coupled with an SVM model to enhance the efficiency of the SVMs in extracting nonlinear patterns. We utilized the ABFT-BILSTM-SVM, LSTM-SVM, BILSTM, LSTM, and SVM models to predict EC and TDS, and the binary version of ABFT was used to determine the optimal input combinations. Notably, the BILSTM-SVM model outperformed the other models in predicting TDS at different prediction horizons. The training MAPE values of the models ranged from 2% to 21%. The BILSTM-SVM model achieved the lowest MAPE values. Moreover, the BILSTM-SVM model successfully predicted EC values. The BILSTM-SVM model ranged from 3% to 21%. The ABFT-BILSTM-SVM model has distinct advantages over other models, including determining the best input scenarios for predicting outputs, successfully extracting important features, and improving the accuracy of the SVM, BILSTM, and LSTM models.

This study contributes to water quality measurement technologies and advanced feature selection models. This study demonstrates the effectiveness of the ABFT for selecting optimal inputs, especially when modelers need a large number of inputs. Additionally, our results show that the BILSTM is highly proficient in extracting essential features from the input data. As a bidirectional analysis tool, BILSTM can be particularly suitable for hydrological simulations. The novel approaches presented in this study have the potential to advance the field of water quality measurements and contribute to improved water resource management practices.

Subsequent studies can utilize the BILSTM-SVM model to accurately forecast the temporal and spatial fluctuations of water quality parameters. Furthermore, this method may be integrated with other approaches to predict interval times, enhancing the accuracy of predictions.

Declaration of Competing Interest

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

References

[1] Shah MI, Javed MF, Alqahtani A, Aldrees A. Environmental assessment based surface water quality prediction using hyper-parameter optimized machine learning models based on consistent big data. Process Saf Environ Prot 2021;151:324–40.
[2] Wang F, Wang Y, Zhang K, Hu M, Weng Q, Zhang H. Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. Environ Res 2021;202:111660.
[3] Kisi O, Parmar KS. Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. J Hydrol 2016;534:104–12.
[4] Ji X, Shang Xu, Dahlgren RA, Zhang M. Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: a case study of Wen-Rui Tang River. Environ Sci Pollut Res 2017;24(19):16062–76.

[5] Kamyab-Talesh F, Mousavi SF, Khaledian M, Yousefi-Falakdehi O, Norouzi-Masir M. Prediction of water quality index by support vector machine: a case study in the Sefidrud Basin, Northern Iran. Water Resour 2019;46:112–6.
[6] Najafzadeh M, Niazmardi S. A novel multiple-kernel support vector regression algorithm for estimation of water quality parameters. Nat Resour Res 2021;30(5):3761–75.
[7] Deng T, Chau K-W, Duan H-F. Machine learning based marine water quality prediction for coastal hydro-environment management. J Environ Manage 2021;284:112051.
[8] Nasir N, Kansal A, Alshaltone O, Barneih F, Sameer M, Shanableh A, et al. Water quality classification using machine learning algorithms. J Water Process Eng 2022;48:102920.
[9] Koklu M, Unlersen MF, Ozkan IA, Aslan MF, Sabanci K. A CNN-SVM study based on selected deep features for grapevine leaves classification. Measurement 2022;188:110425.
[10] Agarwal C, Kaur I, Yadav S. Hybrid CNN-SVM model for face mask detector to protect from COVID-19. In: Artificial Intelligence on Medical Data. Singapore: Springer; 2023. p. 419–26.
[11] Wan H, Xu R, Zhang M, Cai Y, Li J, Shen X. A novel model for water quality prediction caused by non-point sources pollution based on deep learning and feature extraction methods. J Hydrol 2022;612:128081.
[12] Wu J, Wang Z. A hybrid model for water quality prediction based on an artificial neural network, wavelet transform, and long short-term memory. Water (Switzerland) 2022;14(4):610.
[13] Valadkhan D, Moghaddasi R, Mohammadinejad A. Groundwater quality prediction based on LSTM RNN: An Iranian experience. Int J Environ Sci Technol 2022;19(11):11397–408.
[14] Zhang Q, Wang R, Qi Y, Wen F. A watershed water quality prediction model based on attention mechanism and Bi-LSTM. Environ Sci Pollut Res 2022;29(50):75664–80.
[15] Bi J, Zhang L, Yuan H, Zhang J. Multi-indicator water quality prediction with attention-assisted bidirectional LSTM and encoder-decoder. Inf Sci 2023;625:65–80.
[16] Braik M, Ryalat MH, Al-Zoubi H. A novel meta-heuristic algorithm for solving numerical optimization problems: Ali Baba and the forty thieves. Neural Comput & Applic 2022;34(1):409–55.
[17] Vapnik V. The nature of statistical learning theory, Springer, New York; 1995.
[18] Jaseena KU, Kovoor BC. Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. Energ Conver Manage 2021;234:113944.
[19] Joseph RV, Mohanty A, Tyagi S, Mishra S, Satapathy SK, Mohanty SN. A hybrid deep learning framework with CNN and Bi-directional LSTM for store item demand forecasting. Comput Electr Eng 2022;103:108358.
[20] Bazrafshan O, Ehteram M, Dashti Latif S, Feng Huang Y, Yenn Teo F, Najah Ahmed A, et al. Predicting crop yields using a new robust Bayesian averaging model based on multiple hybrid ANFIS and MLP models: Predicting crop yields using a new robust Bayesian averaging model. Ain Shams Eng J 2022;13(5):101724.
[21] Bazrafshan O, Ehteram M, Moshizi ZG, Jamshidi S. Evaluation and uncertainty assessment of wheat yield prediction by multilayer perceptron model with bayesian and copula bayesian approaches. Agric Water Manag 2022;273:107881.
[22] Sammen SS, Ehteram M, Abba SI, Abdulkadir RA, Ahmed AN, El-Shafie A. A new soft computing model for daily streamflow forecasting. Stoch Env Res Risk A 2021;35(12):2479–91.
[23] Ghanbari-Adivi E, Ehteram M, Farrokhi A, Sheikh Khozani Z. Combining radial basis function neural network models and inclusive multiple models for predicting suspended sediment loads. Water Resour Manag 2022;36(11):4313–42.
[24] Banadkooki FB, Ehteram M, Ahmed AN, Teo FY, Fai CM, Afan HA, et al. Enhancement of groundwater-level prediction using an integrated machine learning model optimized by whale algorithm. Nat Resour Res 2020;29(5):3233–52.
[25] Darabi H, Mohamadi S, Karimidastenaei Z, Kisi O, Ehteram M, ELShafie A, et al. Prediction of daily suspended sediment load (SSL) using new optimization algorithms and soft computing models. Soft Comput 2021;25(11):7609–26.
[26] Mohamadi S, Ehteram M, El-Shafie A. Accuracy enhancement for monthly evaporation predicting model utilizing evolutionary machine learning methods. Int J Environ Sci Technol 2020;17(7):3373–96.
[27] Wu Z, Rincon D, Christofides PD. Process structure-based recurrent neural network modeling for model predictive control of nonlinear processes. J Process Control 2020;89:74–84.
[28] Weerakody PB, Wong KW, Wang G, Ela W. A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing 2021;441:161–78.
[29] Mosavi A, Hosseini FS, Choubin B, Abdolshahnejad M, Gharechaee H, Lahijanzadeh A, et al. Susceptibility prediction of groundwater hardness using ensemble machine learning models. Water 2020;12(10):2770.
[30] Ravansalar M, Rajaee T. Evaluation of wavelet performance via an ANN-based electrical conductivity prediction model. Environ Monit Assess 2015;187(6). https://doi.org/10.1007/s10661-015-4590-7.
[31] Ahmadianfar I, Jamei M, Chu X. A novel hybrid wavelet-locally weighted linear regression (W-LWLR) model for electrical conductivity (EC) prediction in surface water. J Contam Hydrol 2020;232:103641.
[32] Banadkooki FB, Ehteram M, Panahi F, Sh. Sammen S, Othman FB, EL-Shafie A. Estimation of total dissolved solids (TDS) using new hybrid machine learning models. J Hydrol 2020;587:124989.
[33] Jamei M, Ahmadianfar I, Chu X, Yaseen ZM. Prediction of surface water total dissolved solids using hybridized wavelet-multigene genetic programming: New approach. J Hydrol 2020;589:125335.
[34] Ahmadpour A, Mirhashemi S, Panahi M. Comparative evaluation of classic and seasonal time series hybrid models in predicting electrical conductivity of Maroun river. Iran Sustainable Water Resources Management 2022;8(5):1–8.