Machine Learning Approaches for One-Day Ahead Soil Temperature Forecasting

Mehmet BILGILI[a] ⓘD, **Şaban ÜNAL**[b*] ⓘD, **Aliihsan ŞEKERTEKIN**[c] ⓘD, **Cahit GÜRLEK**[d] ⓘD

[a]*Department of Mechanical Engineering, Ceyhan Engineering Faculty, Cukurova University, Adana 01950, TURKEY*

[b]*Department of Mechanical Engineering, Engineering Faculty, Osmaniye Korkut Ata University, Osmaniye 80000, TURKEY*

[c]*Igdir University, Vocational School of Higher Education for Technical Sciences, Department of Architecture and Town Planning, Igdir, TURKEY*

[d]*Department of Mechanical Engineering, Engineering Faculty, Sivas Cumhuriyet University, Sivas 58140, TURKEY*

ABSTRACT

Present study investigates the capabilities of six distinct machine learning techniques such as ANFIS network with fuzzy c-means (ANFIS-FCM), grid partition (ANFIS-GP), subtractive clustering (ANFIS-SC), feed-forward neural network (FNN), Elman neural network (ENN), and long short-term memory (LSTM) neural network in one-day ahead soil temperature (ST) forecasting. For this aim, daily ST data gathered at three different depths of 5 cm, 50 cm, and 100 cm from the Sivas meteorological observation station in the Central Anatolia Region of Turkey was used as training and testing datasets. Forecasting values of the machine learning models were compared with actual data by assessing with respect to four statistic metrics such as the mean absolute error, root mean square error (RMSE), Nash−Sutcliffe efficiency coefficient, and correlation coefficient (R). The results showed that the ANFIS-FCM, ANFIS-GP, ANFIS-SC, ENN, FNN and LSTM models presented satisfactory performance in modeling daily ST at all depths, with RMSE values ranging 0.0637-1.3276, 0.0634-1.3809, 0.0643-1.3280, 0.0635-1.3186, 0.0635-1.3281, and 0.0983-1.3256 °C, and R values ranging 0.9910-0.9999, 0.9903-0.9999, 0.9910-0.9999, 0.9911-0.9999, 0.9910-0.9999 and 0.9910-0.9998 °C, respectively.

Keywords: ANFIS, Daily soil temperature, LSTM, Elman neural network (ENN), Feed-forward neural network

1. Introduction

Annual, monthly, daily, and hourly meteorological data are among the most critical atmospheric parameters for many engineering systems and agricultural activities. As one of these meteorological parameters, soil temperature (ST) has crucial importance in distinct disciplines, including soil science, meteorology, agronomy, environmental studies, atmospheric, hydrological, and agricultural numerical models, ecological applications, and agricultural management (Mehdizadeh et al. 2017). Also, ST is a significant meteorological factor for agricultural activity, solar energy technologies, geothermal energy systems, ground source heat pumps, etc. The chemical structure of the soil and organic components are highly affected by ST. The soil heats inwards from the surface and cools by losing heat from inside to outside (Feng et al. 2019). Therefore, daily and seasonal temperature changes are high, although not as high as air. These changes decrease towards the depths, and the temperature remains constant after a certain level. Although the effects of the surface, in general, affect up to 10 m depth, temperature changes are negligible at depths more than 1.5~2 m. For these reasons, many studies have focused on ST forecasting and modeling (Araghi et al. 2017; Shahabi et al. 2021).

Meteorological parameters are measured at meteorological stations located at specific points in many parts of the world. ST measurements are generally made with soil thermometers and soil thermographs at depths of 5, 10, 20, 50, and 100 cm. Many meteorological and atmospheric variables are more easily measured than ST, and therefore more widely accessible. Measuring the ST of a specific location when needed is not as easy as measuring the air temperature of that point. Therefore, estimating ST based on various meteorological parameters, which can be measured much more quickly, has facilitated many engineering problems (Xing et al. 2018).

The thermal changes and energy balances between the soil and ground surface at a certain depth are highly affected by the ST (Araghi et al. 2019). Accurate ST forecasting is recognized as crucial information and foresight for this reason (Zeynoddin et al. 2019). Various studies have recently been conducted on short and mid-term ST forecasting (Penghui et al. 2020). In the first category, statistical approaches such as numerical weather prediction (NWP) methods are used, assuming that future

variations in the statistical characteristics of the ST data set will be similar to those in the past. However, these approaches often require much data that may not be available for long-term forecasts. In the second category, artificial intelligence (AI) and machine learning models such as an artificial neural network (ANN) (Citakoglu 2017; Singhal et al. 2021; Zhou et al. 2020), support vector machine (Xing et al. 2018), gene expression programming (GEP) (Mehdizadeh et al. 2017), genetic programming (Gill & Singh 2015; Stajkowski et al. 2020), adaptive neuro-fuzzy inference system (ANFIS) (Mehdizadeh et al. 2020a) and hybrid models (Sattari et al. 2020; Shamshirband et al. 2020) are used. Various studies have modeled ST as a non-linear physical aspect (Li et al. 2020; Xu et al. 2020; Zeynoddin et al. 2020; Hao et al. 2021).

There are many approaches to predict ST using numerical, analytical, and data-driven models based on the literature. In the early 18th century, Fourier suggested a one-harmonic analytical method that accepts the STs as a function of depths and year's date (Xing et al. 2018). This model is derived according to one-dimensional heat conduction equations, considering the ST on the surface as the boundary conditions. Analytical models can be applied for any desired location, but many parameters such as the soil's heat conductivity, density, and thermal capacity must be known. A particular study is needed to obtain this information correctly. Therefore, analytical models cannot be easily adapted wherever desired. Numerical methods are also utilized to predict the ST. While only conduction heat transfer is considered in analytical methods, convection heat transfer and mass transfer can be included in the mathematical model and conduction heat transfer in numerical methods. However, developing the mathematical model in numerical methods is complex, and at the same time, the model calculation time is comparatively long-time. Naranjo-Mendoza et al. (2018) investigated analytical and numerical methods for forecasting ST, and they presented that the sinusoidal approach was not appropriate for estimating the short-term temperature variations. However, the finite difference method was the most appropriate approach for long- and short-term temperature forecasting. Kayaci & Demir (2018) have presented a numerical model of transient ST distribution for a horizontal ground source heat pump. They obtained steady periodic ST and investigated the impacts of distinct parameters on the ST profile.

Different types of ST prediction models based on the correlation between the ST and meteorological parameters using the data-driven statistical methods can be listed as linear regression (LR), non-linear regression (NLR), ANFIS, ANN, wavelet neural network (WNN), and deep learning. George (2001) presented a study to estimate ST using ANN algorithms. Wind speed, relative humidity and atmospheric temperature data were utilized to predict ST. Gang et al. (2014) estimated the efficiency of a hybrid ground source heat pump unit with the ANN predictive control method. Yan et al. (2016) used data monitoring and mining techniques for predicting the long-term performance of a ground heat pump system using short-term data. Chen et al. (2018) developed an ANN model to estimate the vertical ground heat exchangers. A data-driven model was used by Xing et al. (2018) for daily ST estimations. They estimated the daily or monthly ST of a single site with great accuracy considering solar radiant, air temperature, and time as inputs. Samadianfard et al. (2018) applied the GEP, and WNN approaches to estimate short-term ST at distinct depths. Zeynoddin et al. (2020) used the linear-based stochastic model for ST estimation.

As seen recently, machine-learning approaches have been efficiently used to estimate ST. Considering previous studies, most of the studies modeling or forecasting STs use many environmental and atmospheric variables as inputs. However, the availability of these models may be limited in countries where atmospheric and environmental data are scarce. The STs estimation for any future short or long-term without the need for any other meteorological or geographical data by establishing a relationship between previous STs and future ST can have a significant advantage. Therefore, the main scope of this study is to predict the next day's ST based solely on the previously measured ST data. For this purpose, a univariate procedure was used. Creating such a univariate time series information is very important for areas where meteorological data are limited and can contribute to many previously mentioned applications. In this respect, in this work, six distinct machine learning techniques such as ANFIS-SC, ANFIS-GP, ANFIS-FCM, ENN, FNN, and LSTM were used in one-day ahead ST forecasting.

2. Material and Methods

This section describes the methods used for soil temperature prediction, and provides information about the study area and materials.

2.1. Study area and data

In this study, daily STs at 5, 50, and 100 cm depths were obtained from the Turkish Meteorological Service's station, located in Sivas province, Turkey. The climate of the study area, illustrated in Figure 1, is a continental climate with warm and dry summers and cold and snowy winters. The average air temperatures in wintertime and summertime are recorded as -1.7 °C and 18.7 °C, respectively (www.sivas.climatemps.com). Total annual precipitation is 427.4 mm. The ST data used in this study cover 10-year daily records from 2010 to 2020. In the machine learning applications detailed above, the first 80% of the dataset was used for training and 20% of the dataset was utilized for testing.
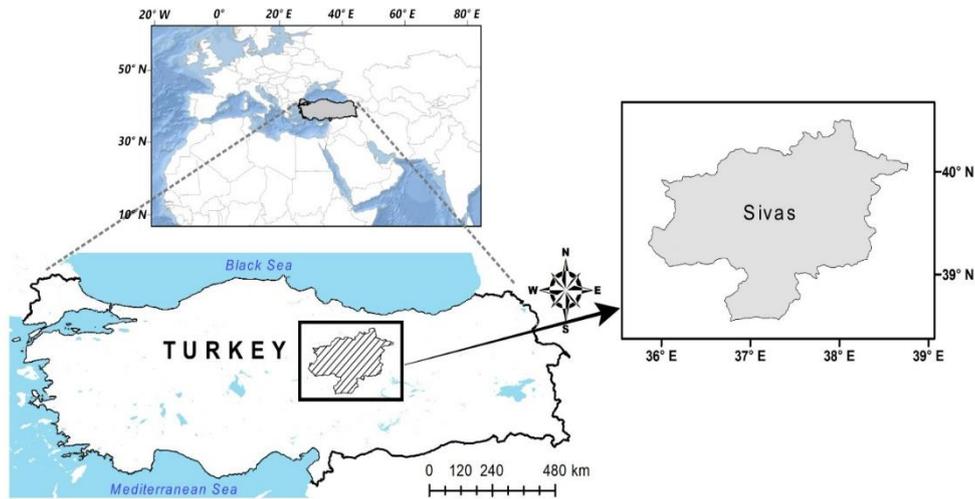
Figure 1- Location map of the study area, showing the place of the ST measurement station.

Table 1 presents descriptive statistics for the daily ST data used in training and testing phases. In Table 1, the skewness is a measure of the symmetry condition of distribution, and zero skewness stands for a completely normal (Gaussian) distribution. In this study, the ST data also indicate an almost normal distribution since the skewness values of both training and testing data sets are close to zero. Concerning the maximum and minimum ST values at different depths, Table 1 reveals that maximum values are higher and minimum values are lower in the training phase than in the testing phase. The standard deviation values show that the ST variation decreases as the depth decreases. Figure 2 illustrates the variation in 10-year ST data sets used in this study at 5, 50, and 100 cm depths. The figure also reveals the training and testing data sets, covering 80% and 20% of all data, respectively. Besides, it is clear from Figure 2 that the ST variation at different depths shows the almost normal distribution for each year, and the more the depth increases, the lower the ST variation is observed.

Table 1- The descriptive statistics for the daily ST data used in training and testing phases.

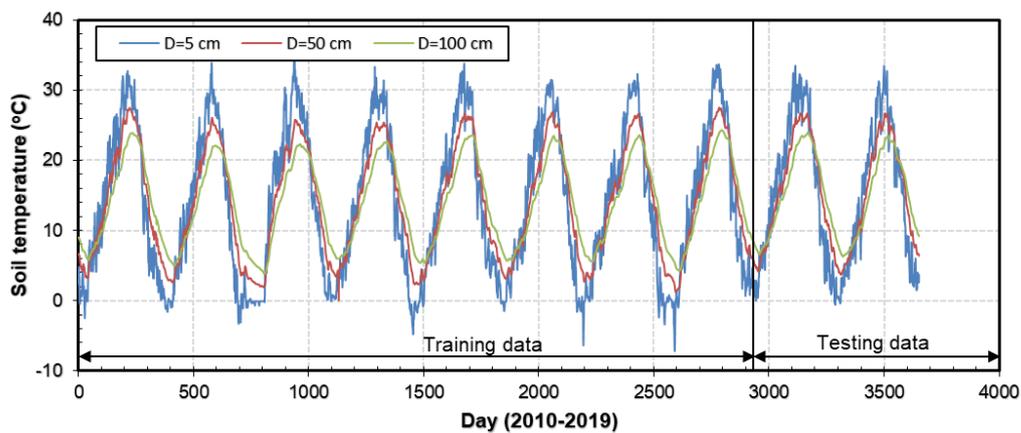| Data	Depth	Unit	Min.	Max.	Avg.	Std Dev.	Skewness
Training Data	5 cm		-7.2	34.8	14.0	10.6	0.13
	50 cm	ºC	1.2	27.5	14.0	7.9	0.09
	100 cm		3.9	24.3	13.9	6.1	0.11
Testing Data	5 cm		-0.6	33.5	15.0	9.9	0.14
	50 cm	ºC	3.7	26.7	15.2	7.5	0.08
	100 cm		6.3	24.0	14.9	5.8	0.08

Figure 2- 10 year ST data sets at 5, 50 and 100 cm depths used in the study. Training and testing data cover 80% and 20% of all data, respectively.

2.2. Artificial neural networks (ANNs)

ANNs are strongly applied in many scientific disciplines by using the dispersed storage features and large-scale parallel local processing techniques available in the human brain. Recently, they have become more effective tools, particularly for non-linear modeling processes, which are difficult to express with statistically or physically defined equations (Inyurt & Sekertekin 2019).

Figure 3 presents the structure of a primary neuron with R inputs. Each input p is identified by a weight value W. The sum of bias and weighted inputs constitutes the transfer function f. The different transfer functions f can generate neuron outputs (Mathworks 2020a).
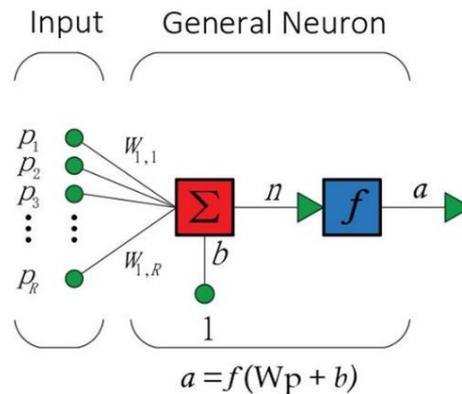

Figure 3- Structure of a basic neuron with R inputs (Mathworks 2020a)

Different ANNs with different algorithms and structures can be formed in simulating and modeling linear or non-linear parameters. Usually, ANNs can be classified into two categories: (a) feed-forward neural networks (FNNs) and (b) recurrent networks. The FNNs do not have feedback or delay elements. In recurrent networks, the network's current and previous inputs or outputs affect the output. The recurrent networks, unlike the FNNs, can use internal memory to process handling random input data series and exhibit transient behavior. This capability makes them applicable to time series forecasting with satisfactory predictive results. As a particular recurrent neural network, Elman neural network (ENN) has been widely and successfully used in time series estimation and forecasting (Mehdizadeh et al. 2017).

2.2.1. Feed-Forward Neural Network (FNN)

The simplest and first of artificial neural networks is FNN. In the structure of the FNN, information flows forward, moving in one direction from input nodes to output nodes. Feed-forward backpropagation ANN covers supervised learning, but there are no feedback loops or connections in the network. In the multi-layer FNNs, several neuron layers are connected in a forward direction. As shown in Figure 4, each node unit in one layer connects directly to neurons in the next layer (Mathworks 2020a).
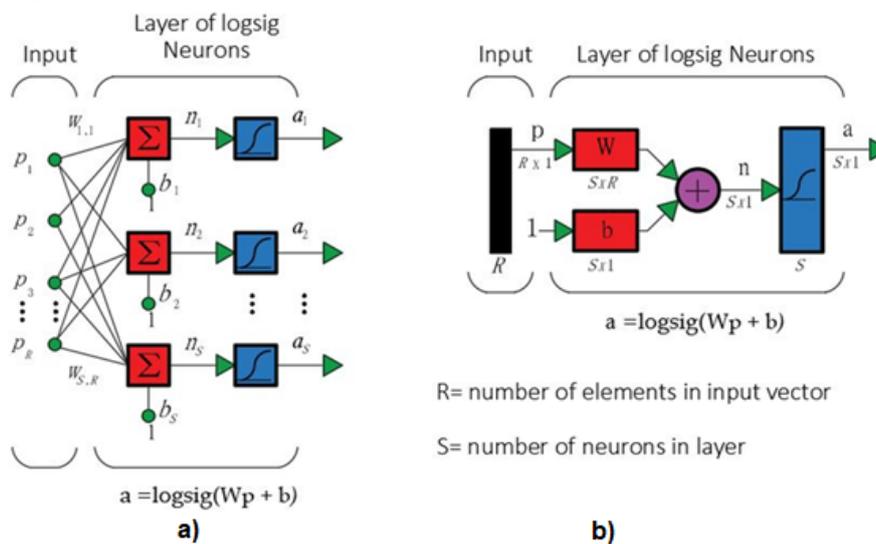

Figure 4- A single-layer network of S logsig neurons with R inputs (Mathworks 2020a)

FNNs usually consist of hidden layers followed by an output layer. Multiple neuron layers provide the non-linear relationships between the input and output vectors of the network with non-linear transfer functions. The linear output layer is usually applied for non-linear regression or function fitting cases. If it is desired to limit the outputs of a network between 0 and 1, the sigmoid transfer function should be preferred in the output layer. With this situation, a decision is made by the network, or the network's pattern recognition problem is solved. The two-layer logsig/purelin network is presented in Figure 5 (Mathworks 2020a).
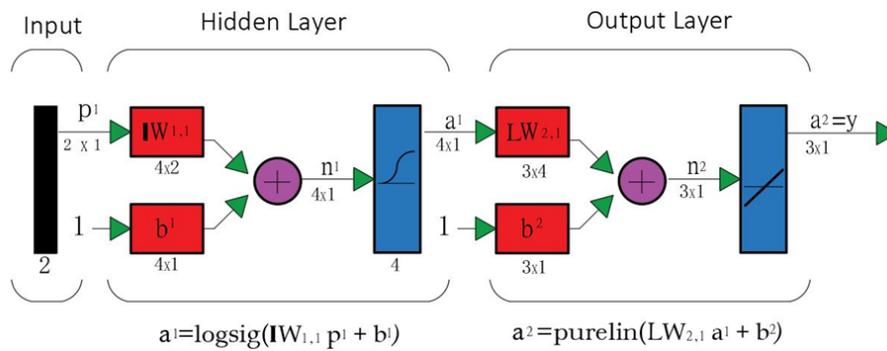
Figure 5- The two-layer logsig/purelin network (Mathworks 2020a)

2.2.2. Elman neural networks (ENN)

ENN consists of several layers and the additional set of connection units in an elementary FNN. A fixed weight value connects the hidden layer units and these context units. Both context units and input nodes activate the nodes in the hidden layer. As shown in Figure 6, content units are enabled to be activated with the feedback of hidden units. At time step, $t+1$, context units record a copy of the hidden unit values obtained at step t. The behavior of the context layer achieves the improvement of the dynamic information processing capacity of the network as local feedback (Mathworks 2020b).
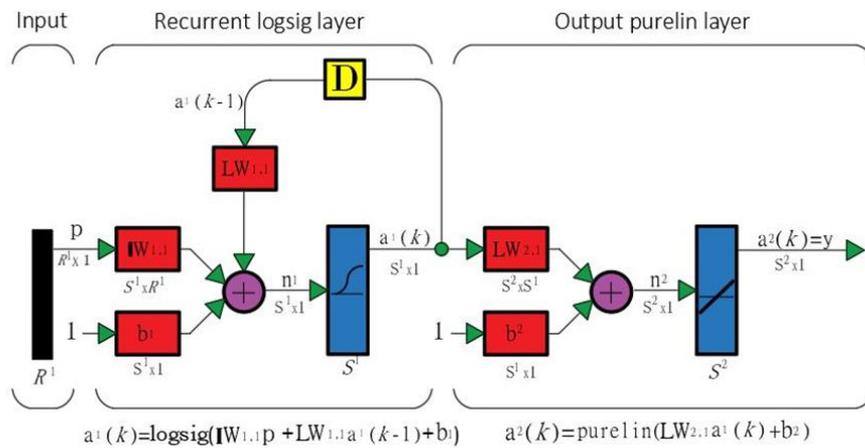

Figure 6- Structure of a two-layer Elman neural networks (Mathworks, 2020b)

Feedback is usually provided from the first layer output to the first layer input in a two-layer ENN. Thanks to this recurrent connection, the Elman network both detects and produces patterns that change over time. Logsig neurons and purelin neurons can be utilized in the recurrent or hidden and output layers of the ENN, respectively. This approach is special because two-layer networks with these transfer functions can estimate any function with random accuracy. However, the number of hidden layers must have sufficient neurons. The greater the complexity of the function available, the more hidden neurons are required. ENN differs from traditional two-layer networks in that the first layer has a recurrent connection. The delay in this connection reserves the values from the previous time step available in the current time step. Therefore, even if the same inputs are supplied to two ENN with the same biases and weights at a given time step, their outputs may differ due to different feedback situations. The network can learn spatial patterns and temporal patterns to store knowledge for future reference (Mathworks 2020b).

2.3. Adaptive neuro fuzzy inference system (ANFIS)

An ANFIS model combines two statistical systems: The Fuzzy Inference System (FIS) and the ANN. Figure 7 shows Type-3 fuzzy reasoning and corresponding equivalent ANFIS architecture (type-3 ANFIS), respectively. A circle describes a fixed node in the structure, while a square denotes an adaptive node. As a simple structure, x and y inputs and f output can be considered. The Sugeno model type is the most extensively applied fuzzy model found in the related works (Jang 1993; Karakuş et al. 2017; Tabari et al. 2012).
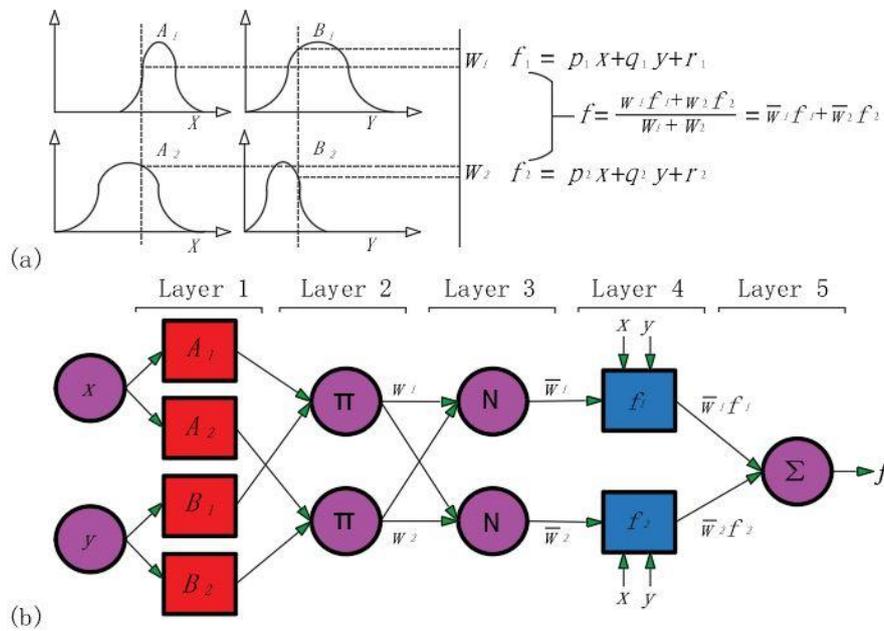
Figure 7- (a) First-order Takagi-Sugeno fuzzy model (b) equivalent ANFIS (Type-3 ANFIS) (Jang 1993)

In general, an ANFIS model contains two phases: construction and training. Membership functions' types and numbers are determined during the construction phase, and the input and output data are divided into rule patches. Therefore, clustering methods are employed for understanding and classifying inputs facilitating the training phase with the help of the ANFIS model. Three clustering approaches comprising the methods of subtractive clustering (SC), fuzzy c-means (FCM), and grid partitioning (GP) are used for this purpose. Fuzzy c-means (FCM) is a clustering method allowing each data point to have multiple clusters and belong to different degrees of membership. The Subtractive Clustering (SC) algorithm considers each data point a candidate cluster center, and the potential of each data point is calculated by measuring the density of the data point surrounding the cluster center. The algorithm uses an iterative process, assuming each point is potentially a cluster center considering their location for other data points (Benmouiza & Cheknane 2019). Grid partitioning (GP) algorithm divides the input data space into a rectangular subspace with the help of an axis-paralleled partition. Each input is split into identically shaped membership functions. The grid is created without using any physical meaning or data density repartition. According to system input-output training data, fuzzy rules are generated using each grid part, thus achieving rapid learning and optimized calculation time. However, the size of the inputs and grid considerably affects the method's performance. A finer grid usually yields higher performance, as expected. The size and location of the fuzzy grid regions can be optimized using adaptive grid partitioning (Benmouiza & Cheknane 2019). Further information regarding ANFIS may be obtained from the work of Jang (1993).

2.4. Long short-term memory (LSTM) neural network

LSTM neural network is a type of Recurrent Neural Network that addresses problems by adding memory cells with persistent errors. This way, errors can be regenerated without disappearing gradients. Three different gates are present in the LSTM neural network. An input gate learns to preserve the persistent error flow in the memory cell from irrelevant inputs. An output gate learns to protect other units from unrelated memory content saved in the memory cell. A forget gate teaches how long the value is in the memory cell (Hochreiter & Schmidhuber 1997; Piotrowski et al. 2015; Salman et al. 2018; Zahroh et al. 2019; Cai et al. 2020; Cho et al. 2020).

Figure 8 shows the LSTM layer architecture, indicating the flow of an X time series with S-length C properties (channels) across an LSTM layer. In this architecture diagram, h_t is also known as the hidden element and is the output. c_t is the cell state at time step t. (Mathworks 2020c).
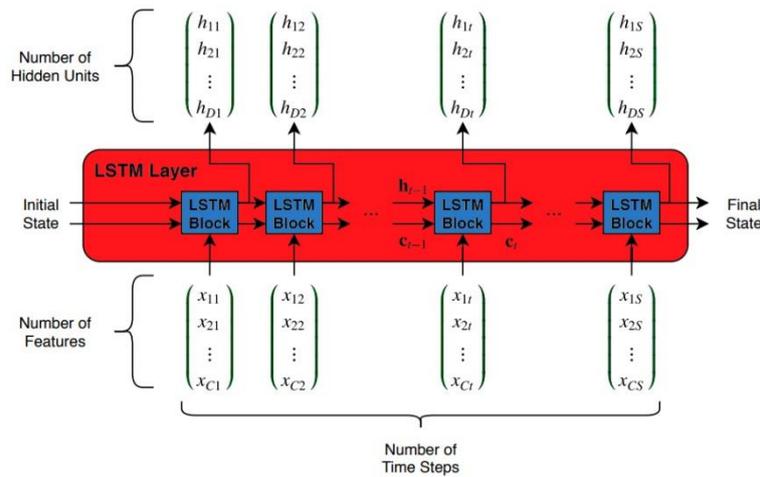
Figure 8- LSTM layer architecture (Mathworks 2020c)

Figure 9 presents the flow of data at time step *t*. This figure indicates how the gates forget, update, and output the cell and the hidden states. Some components in the LSTM layer architecture are used to control the cell state and the hidden state of the layer. For example, input gate (*i*) and output gate (*o*) control the cell state update and level of cell state added to the hidden state, respectively. Besides, the forget gate (*f*) checks the level of cell state reset (forget). On the other hand, cell candidate (*g*) adds the information to the cell state. Further information regarding LSTM can be obtained from Mathworks (2020c) study.
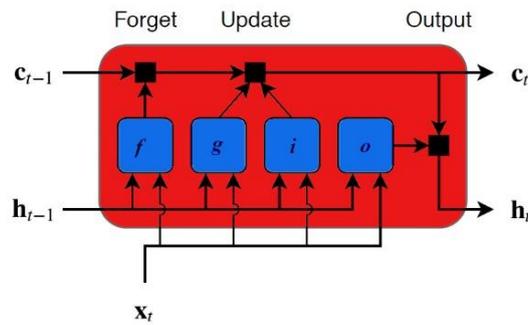

Figure 9- The flow of data at time step *t* (Mathworks 2020c)

2.5. Statistical parameters

In this study, four statistical error parameters such as mean absolute error (MAE), root mean square error (RMSE), Nash−Sutcliffe efficiency coefficient (NSE), and correlation coefficient (R) are used for the assessment of the accuracy of the models in forecasting the observed output variable. MAE, RMSE, NSE, and R parameters, respectively, are expressed as follows (Başakın et al. 2021; Citakoglu 2021):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|p(i) - o(i)| \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[p(i) - o(i)]^2} \tag{2}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}[o(i) - p(i)]^2}{\sum_{i=1}^{N}[o(i) - \bar{o}]^2} \tag{3}$$

$$R = \left(\sum_{i=1}^{N}[p(i) - \bar{p}][o(i) - \bar{o}]\right) / \left(\sqrt{\sum_{i=1}^{N}[p(i) - \bar{p}]^2}\sqrt{\sum_{i=1}^{N}[o(i) - \bar{o}]^2}\right) \tag{4}$$

Where: *p(i)* and *o(i)* are the predicted value and observed value at the time *i*, respectively; $\bar{p}$ and $\bar{o}$ are the means of the predicted values and observed values, respectively, and the total number of data is represented by N.

3. Results and Discussion

In this study, a time-series analysis was applied to predict one-day ahead ST. This is a technique for using time series data values to predict future values based on our historical data points. The proposed time series method is the univariate modeling based on time series data for the modeled variable. The most important advantage of univariate modeling is that there is no need to obtain independent variables. It is known that time-series tools may capture the stochastic component of the time series data, besides machine learning tools may forecast the determinative part of the time series data. In this respect, the used ANFIS-FCM, ANFIS-GP, ANFIS-SC, ENN, FNN, and LSTM techniques were developed based on past observations of the ST parameter as an input to train the model and to predict future values.

The performances of six machine learning techniques, namely, FNN, ENN, LSTM, ANFIS-SC, ANFIS-FCM, and ANFIS-GP, were investigated based on one-day ahead ST forecasting at three different depths of 5 cm, 50 cm, and 100 cm. The trial and error method determined the optimal parameters in each method. The results of the error criteria for determining the optimal parameters in each method by trial and error method are given in Table 2. In order to deal with over-training, a known disadvantage in machine learning models, the data sets were categorized into training and testing. All models were calibrated based on the training data, and then their performances were evaluated using the testing data, which did not take part in model training. Figure 10 represents measured and forecasted ST variations by six machine learning methods at the testing phase. It is clear from Figure 10 that daily ST variations at 5 cm depth show abrupt changes. The general overview of the model results in the central figure of Figure 10 reveals that almost all models overlap with the measured data; however, the zoomed images explain the differences more clearly. The zoomed images show that ANFIS-GP (blue dashed line) explicitly differs from the other methods.
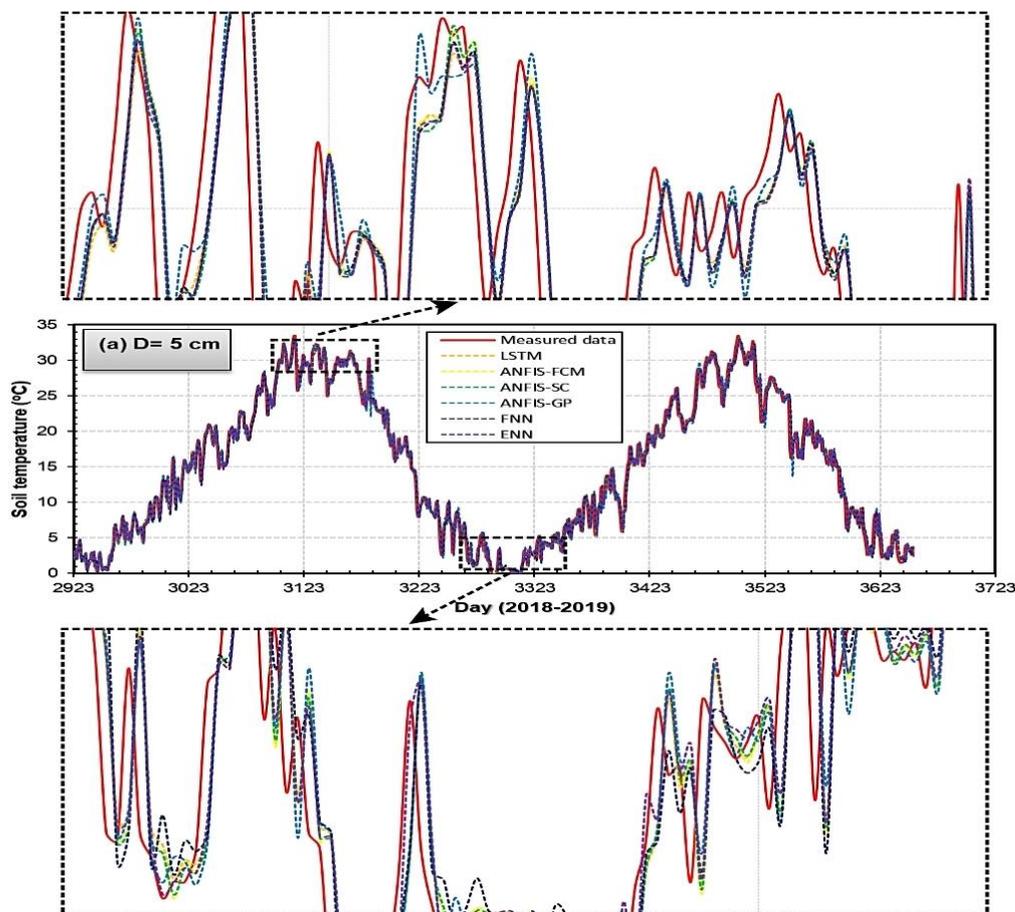

Figure 10- Illustration of the measured and forecasted ST by machine learning models at 5 cm depth

Table 2- The results of the error criteria for determining the optimal parameters in each method by trial and error method

Method	Depth (cm)	HLN*	MAE (ºC)	RMSE (ºC)	R	Method	Depth (cm)	HLN*	MAE (ºC)	RMSE (ºC)	R
LSTM	5	5	1.0145	1.3391	0.9908	FNN	5	6	1.0088	1.3308	0.9909
		10	1.0012	1.3292	0.9910			7	1.0078	1.3308	0.9910
		15	**0.9978**	**1.3256**	**0.9910**			**8**	**1.0065**	**1.3281**	**0.9910**
		20	1.0039	1.3292	0.9910			9	1.0136	1.3340	0.9909
		25	1.0039	1.3310	0.9910			10	1.0151	1.3366	0.9909
	50	**5**	**0.1783**	**0.2251**	**0.9995**		50	6	0.1087	0.1395	0.9998
		10	0.1838	0.2331	0.9995			7	0.1090	0.1396	0.9998
		15	0.1900	0.2400	0.9995			**8**	**0.1081**	**0.1389**	**0.9998**
		20	0.2065	0.2645	0.9994			9	0.1085	0.1391	0.9998
		25	0.2009	0.2532	0.9994			10	0.1105	0.1418	0.9998
	100	5	0.0809	0.1016	0.9998		100	6	0.0523	0.0633	0.9999
		10	0.0838	0.1060	0.9998			7	0.0521	0.0635	0.9999
		15	**0.0773**	**0.0983**	**0.9998**			**8**	**0.0519**	**0.0635**	**0.9999**
		20	0.0885	0.1123	0.9998			9	0.0523	0.0635	0.9999
		25	0.0887	0.1103	0.9998			10	0.0520	0.0638	0.9999
ANFIS-FCM	5	**2**	**1.0036**	**1.3276**	**0.9910**	ENN	5	6	1.0087	1.3300	0.9909
		4	1.0111	1.3334	0.9909			7	1.0074	1.3241	0.9910
		6	1.0161	1.3377	0.9909			**8**	**0.9967**	**1.3186**	**0.9911**
		8	1.0097	1.3365	0.9909			9	1.0368	1.3481	0.9907
		10	1.0132	1.3476	0.9907			10	1.0105	1.3305	0.9909
	50	**5**	**0.1087**	**0.1388**	**0.9998**		50	6	0.1167	0.1471	0.9998
		10	0.1097	0.1403	0.9998			7	0.1101	0.1406	0.9998
		15	0.1108	0.1415	0.9998			**8**	**0.1092**	**0.1409**	**0.9998**
		20	0.1106	0.1416	0.9998			9	0.1097	0.1405	0.9998
		25	0.1113	0.1431	0.9998			10	0.1099	0.1402	0.9998
	100	**5**	**0.0525**	**0.0637**	**0.9999**		100	6	0.0526	0.0636	0.9999
		10	0.0528	0.0642	0.9999			7	0.0525	0.0635	0.9999
		15	0.0530	0.0647	0.9999			**8**	**0.0521**	**0.0635**	**0.9999**
		20	0.0531	0.0649	0.9999			9	0.0523	0.0636	0.9999
		25	0.0529	0.0649	0.9999			10	0.0529	0.0641	0.9999
ANFIS-SC	5	0.1	1.0124	1.3331	0.9909	ANFIS-GP	5	**2**	**1.0407**	**1.3809**	**0.9903**
		0.3	1.0132	1.3354	0.9909			3	1.1600	1.6971	0.9853
		0.5	1.0101	1.3327	0.9909		50	**2**	**0.1106**	**0.1408**	**0.9998**
		0.7	1.0045	1.3294	0.9910			3	0.1120	0.1438	0.9998
		0.9	**1.0036**	**1.3280**	**0.9910**		100	**2**	**0.0517**	**0.0634**	**0.9999**
	50	0.1	0.1135	0.1470	0.9998			3	0.0520	0.0638	0.9999
		0.3	0.1109	0.1417	0.9998						
		0.5	0.1096	0.1398	0.9998						
		0.7	**0.1083**	**0.1387**	**0.9998**						
		0.9	0.1087	0.1389	0.9998						
	100	0.1	0.0539	0.0661	0.9999						
		0.3	0.0528	0.0648	0.9999						
		0.5	0.0528	0.0644	0.9999						
		0.7	0.0529	0.0645	0.9999						
		0.9	**0.0528**	**0.0643**	**0.9999**						

* HLN: Hidden layer number / Number of MFs / Influence radius / Number of neurons in the hidden layer

Concerning the visual interpretation of the daily ST forecasting at 50 cm depth (Figure 11), the temporal variations in ST at 50 cm depth are not as dynamic as the ST at 5 cm depth. As observed in Figure 10, the general overview of the model results in the central figure of Figure 11, showing that almost all models are in good agreement with the measured data; however, the zoomed images reveal the differences between the models. The zoomed images show that LSTM (orange dashed line) explicitly differs from the other methods. On the other hand, all other methods present the same results with minor differences at breaking points. As in the 50 cm ST results, similar trends were observed in ST at 100 cm depth (Figure 12), presenting that the LSTM results were different from all other methods, while the other methods acted almost in the same manner. These results also indicate that the temporal variations in ST decrease with the increasing depth, enabling the methods to provide higher accuracy with better modeling capacities.
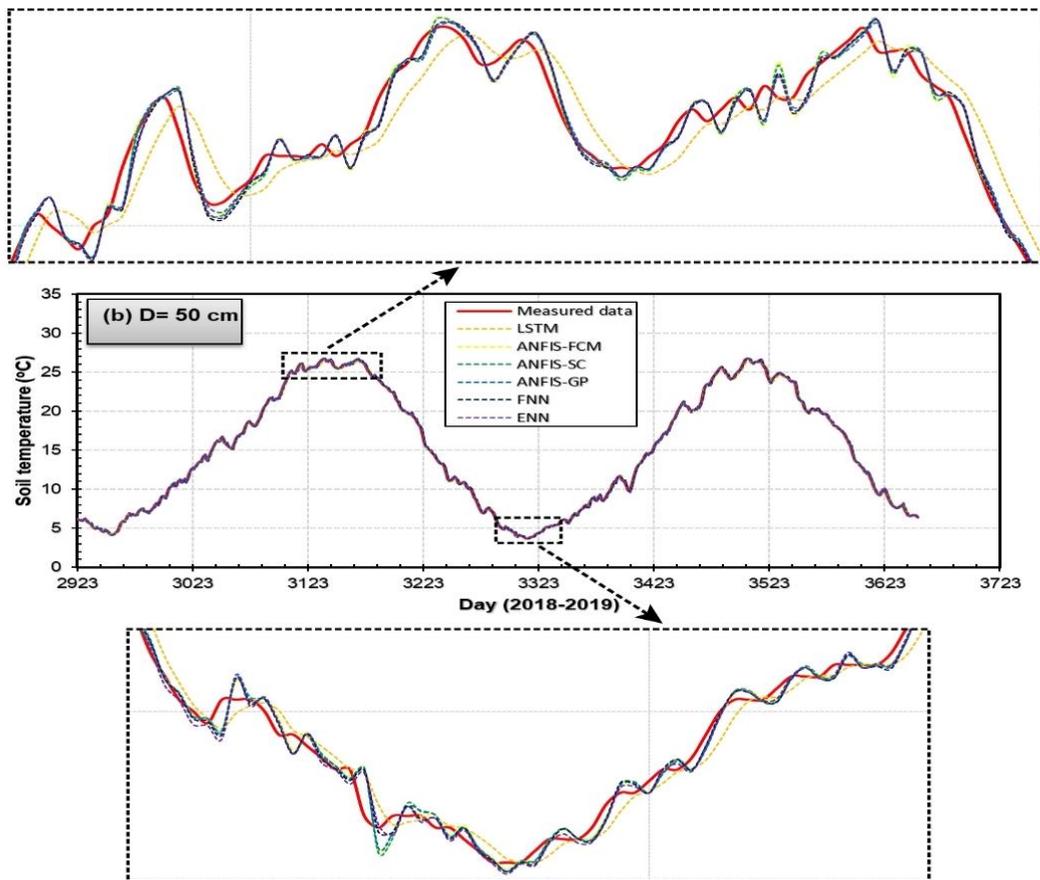
Figure 11- Illustration of the measured and forecasted ST by machine learning models at 50 cm depth.
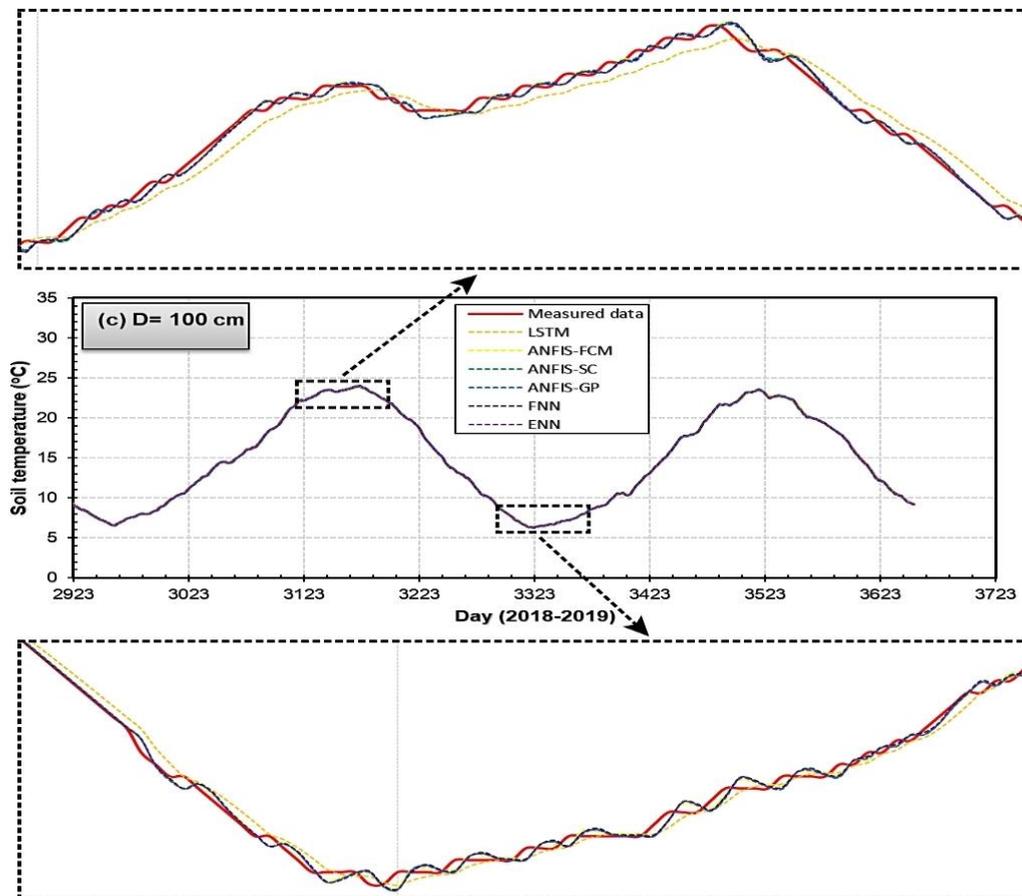

Figure 12- Illustration of the measured and forecasted ST by machine learning models at 100 cm depth.

In addition to the visual results above for forecasting the ST, Figures13-15 represent the regression analyses of the measured and forecasted ST data at 5 cm, 50 cm, and 100 cm, respectively. In the figures, X-axis presents measured ST values, while Y-axis corresponds to forecasted ST in °C. These figures provide the distribution of the measured and forecasted values. Concerning R for the ST forecasting at 5 cm depth (Figure 13 and Table 3), LSTM, ANFIS-FCM, ANFIS-SC, ANFIS-GP, FNN, and ENN presented 0.9910, 0.9910, 0.9910, 0.9903, 0.9910, and 0.9911. Although all results are satisfactory, ENN provided slightly better results than the other 5 cm ST forecasting methods. It is clear from Figures 14 and 15 that the regression plots show that all methods forecasted the ST at 50 cm and 100 cm depth at a higher level than the ST at 5 cm depth due to high temporal variations in ST. The highest R values for the ST forecasting at 50 cm and 100 cm depth were provided by the FNN and ANFIS-GP with the values of 0.9998 and 0.9999, respectively (Figures 14-15 and Table 3).

Table 3- Statistical accuracy results of the ST forecasting with six machine learning methods at various depths.

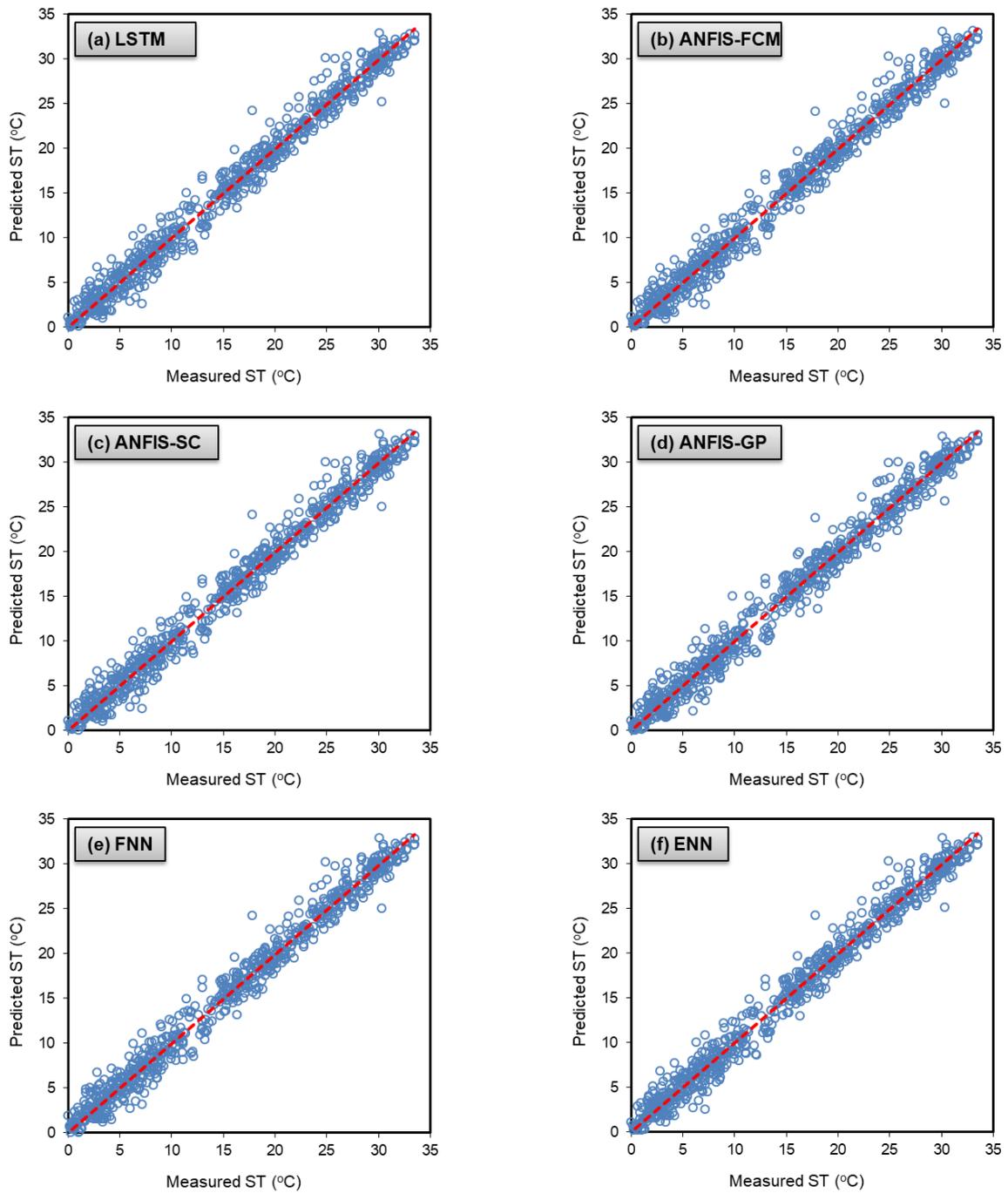| Depth (cm)	Method	MAE (ºC)	RMSE (ºC)	R	NSE
5	LSTM	0.9978	1.3256	0.9910	0.9822
	ANFIS-SC	1.0036	1.3280	0.9910	0.9821
	ANFIS-FCM	1.0036	1.3276	0.9910	0.9821
	ANFIS-GP	1.0407	1.3809	0.9903	0.9807
	FNN	1.0065	1.3281	0.9910	0.9821
	ENN	**0.9967**	**1.3186**	**0.9911**	**0.9824**
50	LSTM	0.1783	0.2251	0.9995	0.9991
	ANFIS-SC	0.1083	0.1387	0.9998	0.9997
	ANFIS-FCM	**0.1087**	**0.1388**	**0.9998**	**0.9996**
	ANFIS-GP	0.1106	0.1408	0.9998	0.9996
	FNN	0.1081	0.1389	0.9998	0.9996
	ENN	0.1092	0.1409	0.9998	0.9996
100	LSTM	0.0773	0.0981	0.9998	0.9997
	ANFIS-SC	0.0528	0.0643	0.9999	0.9999
	ANFIS-FCM	0.0525	0.0637	0.9999	0.9999
	ANFIS-GP	**0.0517**	**0.0634**	**0.9999**	**0.9999**
	FNN	0.0519	0.0635	0.9999	0.9999
	ENN	0.0521	0.0635	0.9999	0.9999

Figure 13- Regression analyses of the measured and predicted data of ST at 5 cm depth with various methods: a) LSTM, b) ANFIS-FCM, c) ANFIS-SC, d) ANFIS-GP, e) FNN, f) ENN methods.
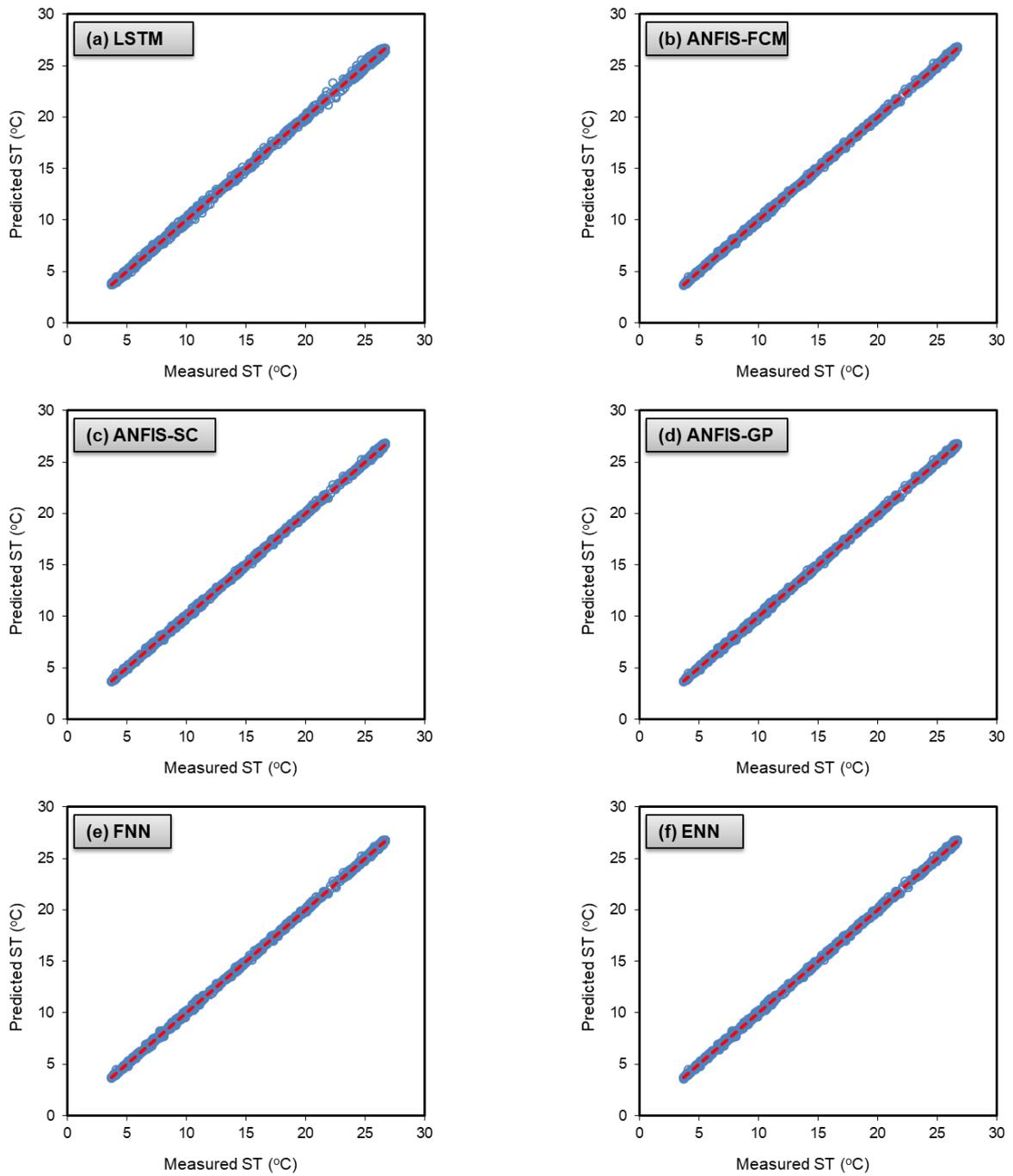
Figure 14- Regression analyses of the measured and predicted data of ST at 50 cm depth with various methods: a) LSTM, b) ANFIS-FCM, c) ANFIS-SC, d) ANFIS-GP, e) FNN, f) ENN methods.
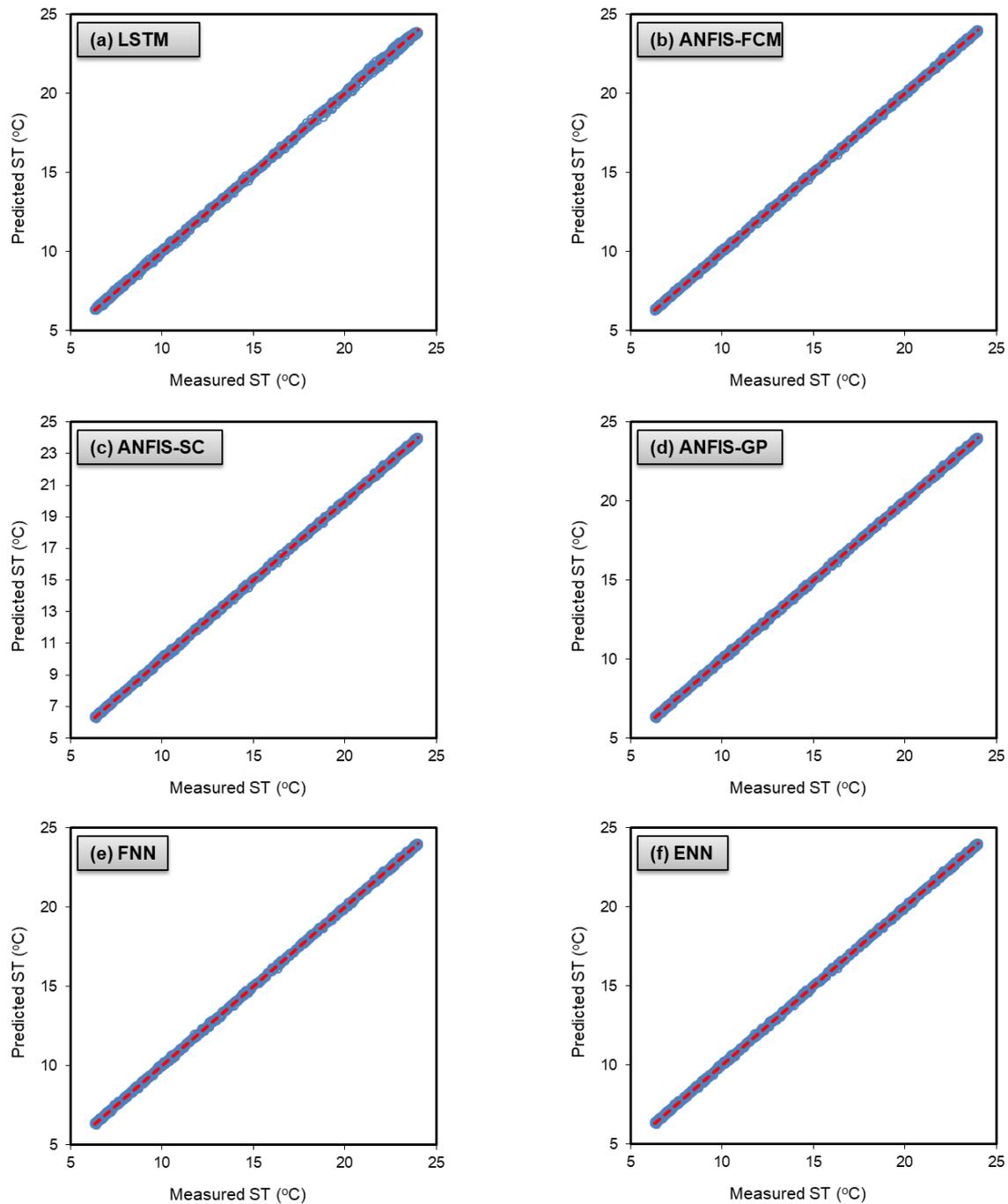
Figure 15- Regression analyses of the measured and predicted data of ST at 100 cm depth with various methods: a) LSTM, b) ANFIS-FCM, c) ANFIS-SC, d) ANFIS-GP, e) FNN, f) ENN methods.

In summary, Table 3 highlights the statistical accuracy results of the ST forecasting with six machine learning methods at various depths. It is clear from the table that every method provided satisfactory results at each depth compared to each other. However, we would like to demonstrate which method will give the best results, even with small differences. Considering the ST forecasting at 5 cm depth, ENN provided the best results with the statistical metrics of 0.9967 °C MAE, 1.3186 °C RMSE and 0.9824 NSE. On the other hand, the second-best method was the LSTM with 0.9978 °C MAE, 1.3256 °C RMSE, and 0.9822 NSE, and also the ANFIS-GP offered the worst results with 1.0407 °C MAE, 1.3809 °C RMSE, and 0.9807 NSE. For ST forecasting at 50 cm depth, the performance of the methods from the best to the worst can be listed as FNN, ANFIS-SC, ANFIS-FCM, ENN, ANFIS-GP, and LSTM, respectively. While the FNN presented the best results with 0.1081 °C MAE, 0.1389 °C RMSE, and 0.9996 NSE, the LSTM was the worst method with 0.1783 °C MAE, 0.2251 °C RMSE, and 0.9991 NSE in forecasting the ST at 50 cm depth. Concerning the ST forecasting at 100 cm depth, the effectiveness of the methods from the best to the worst, respectively, were ANFIS-GP, FNN, ENN, ANFIS-FCM, ANFIS-SC, and LSTM. The ANFIS-GP provided the best accuracy results as 0.0517 °C MAE, 0.0634 °C RMSE, and 0.9999 NSE, whereas the worst results obtained from the LSTM were 0.0773 °C MAE, 0.0983 °C RMSE, and 0.9997 NSE. Although the LSTM was effective in forecasting the ST at 5 cm depth, this performance was not the same as in forecasting the ST at 50 cm and 100 cm depths, which reveals it works better

in complex environments. Another important point of the performance results is that the optimal parameters of each method were determined by trial and error, which may affect the accuracy results, but the study was already organized under this assumption. Even though we listed the best and the worst methods based on the statistical accuracy metrics for every depth, in general, all methods provided identical and satisfactory results with slight differences. The RMSE results at 5 cm, 50 cm, and 100 cm depths varied from 1.3186 °C to 1.3809 °C, from 0.1388 °C to 0.2251 °C, and from 0.0634 °C to 0.0983 °C, respectively.

Taylor diagram based on statistical analysis was utilized while comparing six models in order to assess the consistency of the predicted data from the measured data. Thus, further comparisons of all models were provided using the Taylor diagram. Figure 16 shows the Taylor diagrams for the ST variable of LSTM, ANFIS-FCM, ANFIS-SC, ANFIS-GP, FNN, and ENN methods. As can be seen from the figure, it is understood that the 6 models used in the modeling of ST data at 5, 50, and 100 cm depths are similar to each other, the data lines overlap and it is difficult to separate from each other. However, it is clear that the closest results to the measurement values are obtained at a depth of 100 cm.
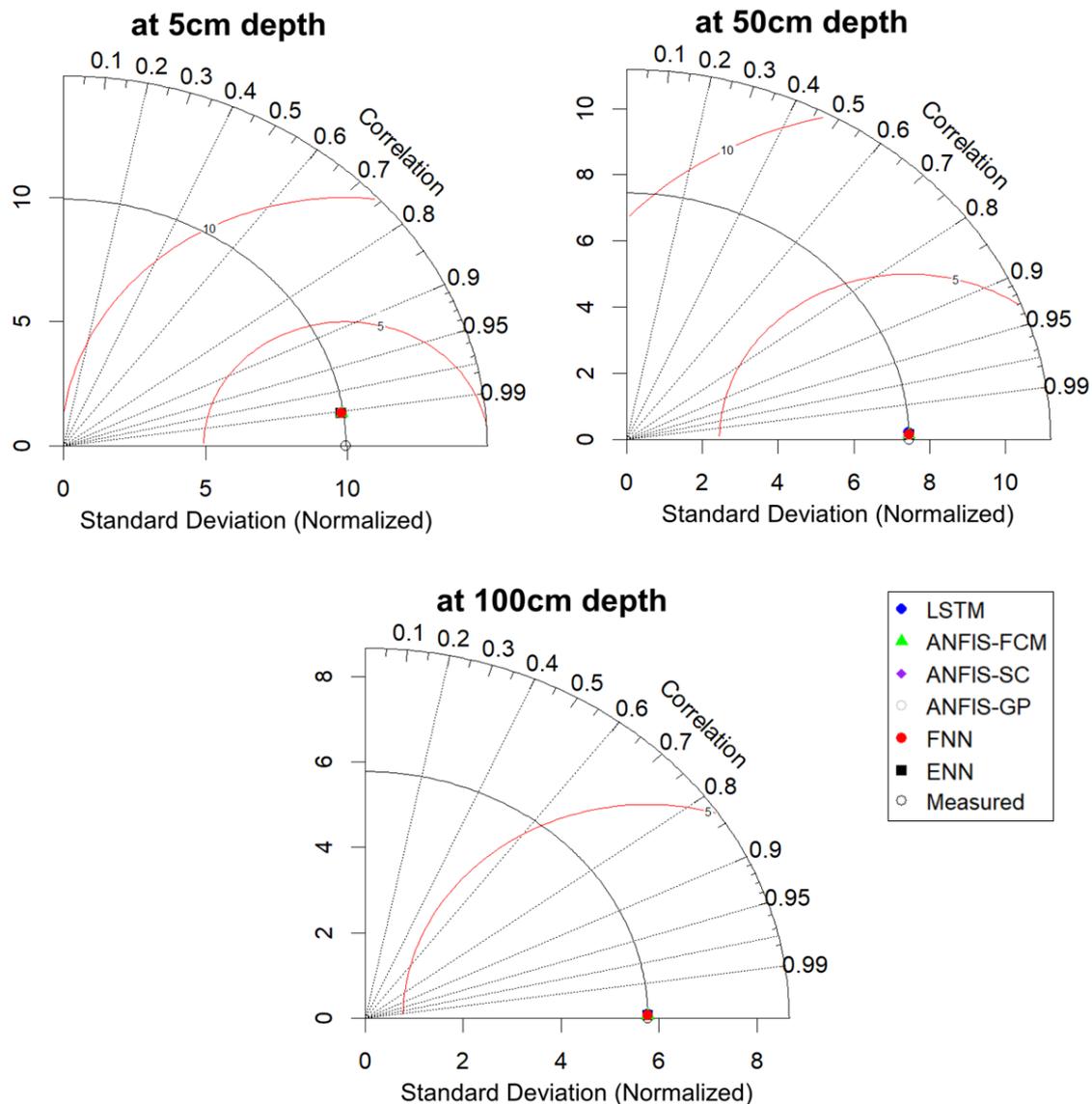

Figure 16- Taylor diagrams for ST variable of LSTM, ANFIS-FCM, ANFIS-SC, ANFIS-GP, FNN, and ENN methods.

Additionally, a comparison with previous studies was carried out to verify the accuracy of our results and models. The literature review showed that previously published studies mainly applied machine learning models using some meteorological parameters to predict ST. However, few studies have been reported in the literature for ST prediction using time series-based models. According to this, the prediction results of the proposed models are compared with the results from some published studies based on daily ST estimates in Table 4. The results showed that the models proposed in this study predicts daily ST with close accuracy to the models in other studies, and even more accurately than most. In summary, the results in this table show that the proposed time series forecasting models have been successfully applied for one-day ahead forecasting of ST.

Table 4- Summary of typical studies on daily ST forecasting

Reference	Country	Data	Method	Depth (cm)	R	RMSE (ºC)
Singhal et al. (2021)	India	Daily (2016 - 2017)	ANN	10	0.9560	0.5520
				20	0.9840	0.2950
				45	0.9740	0.3980
Hao et al. (2021)	Switzerland	Daily (2004 - 2014)	EEMD-CNN	5	0.9970	0.4660
				10	0.9980	0.3750
				25	0.9990	0.2760
Zeynoddin et al. (2019)	Iran	Daily (2013 - 2015)	Linear stochastic method	5	0.9953	1.3300
				10	0.9971	0.9800
				20	0.9983	0.7000
Yu et al. (2021)	China	Daily (2012 - 2020)	EEMD-Conv3d	0-7	0.9946	1.3096
Mehdizadeh et al. (2020b)	Iran	Daily (1998 - 2017)	GEP-FARIMA	5	-	0.2400
				10	-	0.1300
				50	-	0.0500
				100	-	0.1600
This study	Turkey	Daily (2010 - 2019)	ENN	5	0.9900	1.3186
			ANFIS-FCM	50	0.9988	0.1388
			ANFIS-GP	100	0.9999	0.0634

4. Conclusions

In this study, the performances of the six machine learning methods, namely, LSTM, ANFIS-SC, ANFIS-FCM, ANFIS-GP, FNN, and ENN, were evaluated based on one-day ahead ST forecasting at three different depths (5 cm, 50 cm, and 100 cm). The ST data at all depths cover the 10 years, and for all methods, the training and testing data sets were split into 80% and 20%, respectively, based on the whole data set. Concerning the data structure at different depths, it was observed that the daily temporal variation in ST increases with the decreasing depth. In other words, the daily ST at 5 cm depth shows more changes compared to the ST at 50 cm depth, and the variations in daily ST at 50 cm depth are higher than the ST at 100 cm depth. Considering the responses of the machine learning methods to these variations at different depths, it was proved that the methods provided high accuracies when the depth of the daily ST increased, such as from 5 cm to 100 cm. The visual and statistical results revealed that all methods presented satisfactory forecasting results at each depth, with only slight differences. The best performances in forecasting the daily ST at 5 cm, 50 cm, and 100 cm depths were obtained from ENN, FNN, and ANFIS-GP. In addition, the worst accuracies for the same depths were provided by the ANFIS-GP, LSTM, and LSTM, respectively. Even though LSTM was the second-best method in forecasting the ST at 5 cm depth, it did not respond well at 50 cm and 100 cm depths, which shows that it is good at forecasting the ST in complex environments. Overall, the obtained results indicate that all methods used in this study can be performed for the daily ST forecasting at different depths.

Acknowledgments

References

Araghi A, Mousavi-Baygi M, Adamowski J, Martinez C, Van der Ploeg M (2017). Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network. Met Apps 24:603-611. https://doi.org/10.1002/met.1661

Araghi A & Adamowski J, Martinez CJ, Olesen JE (2019). Projections of future soil temperature in northeast Iran. Geoderma 349:11-24. https://doi.org/10.1016/j.geoderma.2019.04.034

Başakın E E, Ekmekcioğlu Ö, Çıtakoğlu H & Özger M (2021). A new insight to the wind speed forecasting: robust multi-stage ensemble soft computing approach based on pre-processing uncertainty assessment. Neural Comput & Applic. https://doi.org/10.1007/s00521-021-06424-6

Benmouiza K & Cheknane A (2019). Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. Theor Appl Climatol 137:31-43. https://doi.org/10.1007/s00704-018-2576-4

Cai Q, Yan B, Su B, Liu S, Xiang M, Wen Y, Cheng Y & Feng N (2020). Short-term load forecasting method based on deep neural network with sample weights. Int Trans Electr Energy Syst 30:e12340. https://doi.org/10.1002/2050-7038.12340

Chen S, Mao J, Chen F, Hou P & Li Y (2018). Development of ANN model for depth prediction of vertical ground heat exchanger. *International Journal of Heat and Mass Transfer* 117:617-626. https://doi.org/10.1016/j.ijheatmasstransfer.2017.10.006

Cho M Y, Chang J M & Huang C C (2020). Application of parallel Elman neural network to hourly area solar PV plant generation estimation. Int Trans Electr Energy Syst 30:e12470. https://doi.org/10.1002/2050-7038.12470

Citakoglu H (2017). Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey. Theor Appl Climatol 130:545-556. https://doi.org/10.1007/s00704-016-1914-7

Citakoglu H (2021). Comparison of multiple learning artificial intelligence models for estimation of long-term monthly temperatures in Turkey. *Arabian Journal of Geosciences* 14:2131. https://doi.org/10.1007/s12517-021-08484-3

Feng Y, Cui N, Hao W, Gao L & Gong D (2019). Estimation of soil temperature from meteorological data using different machine learning models. Geoderma 338:67–77. https://doi.org/10.1016/j.geoderma.2018.11.044

Gang W, Wang J & Wang S (2014). Performance analysis of hybrid ground source heat pump systems based on ANN predictive control. Applied Energy 136:1138-1144. https://doi.org/10.1016/j.apenergy.2014.04.005

George R K (2001). Prediction of soil temperature by using artificial neural networks algorithms. Non-linear Analysis: Theory, Methods & Applications 47:1737-1748. https://doi.org/10.1016/S0362-546X(01)00306-6

Gill J & Singh S (2015). An efficient neural networks based genetic algorithm model for soil temperature prediction. *International Journal of Emerging Technologies in Engineering Research* (IJETER) 3:1-5

Hao H, Yu F & Li Q (2021). Soil temperature prediction using convolutional neural network based on ensemble empirical mode decomposition. IEEE Access 9:4084-4096. https://doi.org/10.1109/ACCESS.2020.3048028

Hochreiter S & Schmidhuber J (1997). Long short-term memory. Neural Computation. 9:1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Inyurt S & Sekertekin A (2019). Modeling and predicting seasonal ionospheric variations in Turkey using artificial neural network (ANN). Astrophys Space Sci. https://doi.org/10.1007/s10509-019-3545-9

Jang J S R (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665-685. https://doi.org/10.1109/21.256541.

Karakuş O, Kuruoğlu E E & Altınkaya M A (2017). One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renew Power Gener 11:1430-1439. https://doi.org/10.1049/iet-rpg.2016.0972

Kayaci N, Demir H (2018). Numerical modelling of transient soil temperature distribution for horizontal ground heat exchanger of ground source heat pump. Geothermics 73:33-47. https://doi.org/10.1016/j.geothermics.2018.01.009

Li C, Zhang Y & Ren X (2020). Modeling hourly soil temperature using deep BiLSTM neural network. Algorithms. https://doi.org/10.3390/a13070173

Mathworks (2020a). Multilayer Shallow Neural Network Architecture. https://www.mathworks.com/help/deeplearning/ug/multilayer-neural-network-architecture.html. Accessed 17 May 2020

Mathworks (2020b). Elman Networks. http://matlab.izmiran.ru/help/toolbox/nnet/recur94.html. Accessed 17 May 2020

Mathworks (2020c). Long Short-Term Memory Networks. Accessed. https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html. Accessed 17 May 2020

Mehdizadeh S, Behmanesh J & Khalili K (2017). Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data. Environ Earth Sci 76:325. https://doi.org/10.1007/s12665-017-6607-8.

Mehdizadeh S, Ahmadi F & Sales AK (2020a). Modelling daily soil temperature at different depths via the classical and hybrid models. Meteorol Appl 27:e1941. https://doi.org/10.1002/met.1941

Mehdizadeh S, Fathian F, Safari MJS & Khosrav A (2020b). Developing novel hybrid models for estimation of daily soil temperature at various depths. Soil & Tillage Research 197 (2020). 104513. https://doi.org/10.1016/j.still.2019.104513

Naranjo-Mendoza C, Wright A J, Oyinlola M A & Greenough R M (2018). A comparison of analytical and numerical model predictions of shallow soil temperature variation with experimental measurements. Geothermics 76:38-49. https://doi.org/10.1016/j.geothermics.2018.06.003

Penghui L, Ewees A A, Beyaztas B H, Qi C, Salih S Q, Al-Ansari N, Bhagat S K, Yaseen Z M & Singh V P (2020). Metaheuristic optimization algorithms hybridized with artificial intelligence model for soil temperature prediction: Novel Model. IEEE Access 8:51884-51904. https://doi.org/10.1109/ACCESS.2020.2979822

Piotrowski A P, Napiorkowski M J, Napiorkowski J J & Osuch M (2015). Comparing various artificial neural network types for water temperature prediction in rivers. J Hydrol 529:302-315. https://doi.org/10.1016/j.jhydrol.2015.07.044

Salman A G, Heryadi Y, Abdurahman E & Suparta W (2018). Single layer & multi-layer long short-term memory (LTSM) model with intermediate variables for weather forecasting. Procedia Computer Sci 135:89–98. https://doi.org/10.1016/j.procs.2018.08.153

Samadianfard S, Asadi E, Jarhan S, Kazemi H, Kheshtgar S, Kisi O, Sajjadi S & Manaf A A (2018). Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths. Soil and Tillage Research 175:37-50. https://doi.org/10.1016/j.still.2017.08.012

Sattari M T, Avram A, Apaydin H & Matei O (2020). Soil temperature estimation with meteorological parameters by using tree-based hybrid data mining models. Mathematics. https://doi.org/10.3390/math8091407

Shahabi M, Khojastehpour M & Sadrnia H (2021). Production and Evaluation of Agricultural Biodegradable Mulch through Heat and Moisture Distribution in Soil. *Journal of Agricultural Sciences*

Shamshirband S, Esmaeilbeiki F, Zarehaghi D, Neyshabouri M, Samadianfard S, Ghorbani MA, Mosavi A, Nabipour N & Chau KW (2020). Comparative analysis of hybrid models of firefly optimization algorithm with support vector machines and multilayer perceptron for predicting soil temperature at different depths. Engineering Applications of Computational Fluid Mechanics 14:939-953. https://doi.org/10.1080/19942060.2020.1788644

Singhal M, Gairola AC & Singh N (2021). Artificial neural network-assisted glacier forefield soil temperature retrieval from temperature measurements. Theoretical and Applied Climatology 143:1157-1166. https://doi.org/10.1007/s00704-020-03498-5

Stajkowski S, Kumar D, Samui P, Bonakdari H & Gharabaghi B (2020). Genetic-algorithm-optimized sequential model for water temperature prediction. Sustainability 12:5374. https://doi.org/10.3390/su12135374

Tabari H, Kisi O, Ezani A & Talaee P H (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. J Hydrol 444-445: 78-89. https://doi.org/10.1016/j.jhydrol.2012.04.007

Xing L, Li L, Gong J, Ren C, Liu J & Chen H (2018). Daily soil temperatures predictions for various climates in United States using data-driven model. Energy 160:430-440. https://doi.org/10.1016/j.energy.2018.07.004

Xu C, Qu JJ, Hao X, Zhu Z & Gutenberg L (2020). Surface soil temperature seasonal variation estimation in a forested area using combined satellite observations and in-situ measurements. *Int. J. Appl. Earth Obs. Geoinformation* 91:102156. https://doi.org/10.1016/j.jag.2020.102156

Yan L, Hu P, Li C, Yao Y, Xing L, Lei F & Zhu N (2016). The performance prediction of ground source heat pump system based on monitoring data and data mining technology. Energy and Buildings 127:1085-1095. https://doi.org/10.1016/j.enbuild.2016.06.055

Yu F, Hao H & Li Q (2021). An ensemble 3D convolutional neural Network for spatiotemporal soil temperature forecasting. Sustainability 13:9174. https://doi.org/10.3390/su13169174

Zahroh S, Hidayat Y & Pontoh RS (2019). Modeling and forecasting daily temperature in Bandung. In: Paper Presented at Proc. Int. Conf. Ind. Eng. Oper. Manag., Riyadh, Saudi Arabia

Zeynoddin M, Bonakdari H, Ebtehaj I, Esmaeilbeiki F, Gharabaghi B & Haghi DZ (2019) A reliable linear stochastic daily soil temperature forecast model. Soil & Tillage Research 189:73-87. https://doi.org/10.1016/j.still.2018.12.023

Zeynoddin M, Ebtehaj I & Bonakdari H (2020). Development of a linear based stochastic model for daily soil temperature prediction: One step forward to sustainable agriculture. Computers and Electronics in Agriculture 176:105636. https://doi.org/10.1016/j.compag.2020.105636

Zhou S, Chu X, Cao S, Liu X & Zhou Y (2020). Prediction of the ground temperature with ANN, LS-SVM and fuzzy LS-SVM for GSHP application. Geothermics 84:101757. https://doi.org/10.1016/j.geothermics.2019.101757