Knowledge-based Engineering and Sciences (2023), **4**, (1), 55–77
doi:10.51526/kbes.2023.4.1.55-77

RESEARCH PAPER

Interpretation the Influence of Hydrometeorological Variables on Soil Temperature Prediction Using the Potential of Deep Learning Model

Salah Elsayed,[*,1] Meenu Gupta,[2] Gopal Chaudhary,[3] Soham Taneja,[3] Harshit Gaur,[3] Mohamed Gad,[4] Mohamed Hamdy Eid,[5,6] Attila Kovács,[5] Szűcs Péter,[5] Aissam Gaagai,[7] and Urs Schmidhalter[8]

[1]Agricultural Engineering, Evaluation of Natural Resources Department, Environmental Studies and Re-search Institute, University of Sadat City, Minufiya, 32897, Egypt
[2]Chandigarh University, Punjab, India
[3]Bharati Vidyapeeth's College of Engineering, New Delhi, India
[4]Hydrogeology, Evaluation of Natural Resources Department, Environmental Studies and Research Institute, University of Sadat City, Menoufia 32897, Egypt
[5]Institute of Environmental Management, Faculty of Earth Science, University of Miskolc, 3515 Miskolc, Hungary
[6]Geology Department, Faculty of Science, Beni-Suef University, Beni-Suef 65211, Egypt
[7]Scientific and Technical Research Center on Arid Regions (CRSTRA), Biskra 07000, Algeria
[8]Department of Life Science Systems, Chair of Plant Nutrition, Technical University ofMunich, Freising, Germany
[*]Corresponding author. Email: salah.emam@esri.usc.edu.eg

(Received 15 March 2023; revised 27 April 2023; accepted 29 April 2023; first published online 30 April 2023)

Abstract

The importance of soil temperature (ST) quantification can contribute to diverse ecological modelling processes as well as for agricultural activities. Over the literature, it was evident that soil supports more than 95% of living habitats and food production on earth, and this demand will increase to 500 years' times in expected consumption in 2060. This paper aims to analyses the contrastive approach to predict the ST of a certain region with the help of different machine learning models, including Random Forest (RF), Support Vector, Neural Network (NN), Linear Regression (LR) and Long Short-Term Memory Network (LSTM). The study was utilized the hourly humidity, dew point, rainfall, solar radiation, and barometer readings for the formulation of the models. Various performance criteria were employed to evaluate the prediction skills of the models and the results depicted that the promising ability belong to LSTM despite the acceptable prediction accuracy achieved by other models. The modelling outcomes revealed that LSTM model attained the lowest root mean square error (RMSE = 3.3255) decreased the average prediction error by 6% with regards to NN (RMSE = 3.4796), SVM (RMSE = 3.5766), and RF (RMSE = 3.8128), and improved the prediction accuracy of LR by 15%. The model is in compliance with the latest machine learning industry standards and allows low-cost experimental performances on low powered edge computing devices.

Keywords: Soil Temperature; Geo-science engineering; atmospheric data; deep learning.

1. Introduction

Soil temperature (ST) demonstrates a vital role in various biomes [1]. ST is a very scrutinizing tool that governs important variables and factors such as ecological and atmospheric variables, chemical, and terrestrial systems [2]. In simpler terms, it can be described as the function of heat

flux in the soil and the predominant cycles of heat exchanges between soil and the atmosphere [3]. Studying the ST at various depths provides vital information about different aspects of geothermal and extrinsic atmospheric processes of specific regions and aids in research parameters, such as agronomy, hydrology, the nitration cycle, and environmental sciences [4]. The prediction of these ecological parameters, such as wind speed, ST, and relative humidity, is a necessary aspect of agricultural sciences owing to the relationship of these parameters with solar energy [5], [6]. Variation in ST can cause significant changes in the properties of soil, leading to consequent environmental problems and an associated change in carbon footprints balance [7]. ST is connected to rainfall and air temperature in each cycle of natural process. Many researchers have tried to establish the relationship between ST and meteorological parameters at different surface soil depths (5, 10, 20, 20, 50, 100, and 150 cm). Other factors that have been studied include the moisture content of the soil, rate of evaporation from the soil surface, plant cover, and surface albedo [8]. Heat transfer mechanisms from the surface of the soil are important for the plantation of seasonal crops. ST has a strong effect on the rate of germination and early growth of vegetative plants [9]. The greatest seed germination cycles are only met when the optimum ST is available for growth [10]. Moreover, ST is crucial for studying the exchange of energy and setting up correlations between the atmosphere and the land. Thus, developing theoretical methods to accurately predict ST is important. Soil surface temperature at various depths is spatially difficult owing to the availability of no specific methods. ST has drastically changed under the influence of climate change, and it also impacts the green-up date and environment [11].

The research established by environmental science practitioners demonstrated the effect of temperatures in soil freezes and thaw states [12], deciding the seed harvesting dates [13], and moreover it influences infiltration [14], soil water systems [15]. In the past current trends, extensive use of machine learning techniques has been helping the meteorological and environmental sciences to set up widely optimized harvest systems harnessing artificial intelligence (AI) [16], [17], [26], [18]–[25]. Although ST is said to be an effective variable agricultural practices, only a few types of research have developed data-driven forecasting models for its prediction. Thus, establishing new and modern algorithm driven techniques is important to continuously evaluate the soil profile and enhance the agricultural industry. Recently several studies were established in different field of hydro environmental engineering using different AI-based models. However, a significant level of challenges can still be attributed to this model, such as generalization [22]–[30]. According to nor free lunch theorem there is no single model that proves to be the best in all types of data sets. In addition, several scholars have proved that even when using similar datasets, there are still variations in the performance for different models. Hence, it has become necessary to design a generally applicable programmed AI models which can be realized over different local scales [31], [32].

Various studies have outlined comparative techniques for predicting ST from data available to them, such as humidity, atmospheric pressure, heat escape from soil, and fusion of gases with the soil surface [33]. These discuss the methods to implement AI to estimate the daily ST in distributed regions [34]. For instance, Hanks et al. (1971) predicted ST as an inclusive function of time, and they considered the initial temperature of the soil as a function of depth. The difference in his study was created as the formula proposed could not be applied for two- or three-dimensional problems under the soil surface. Bilgili (2010) stated that ST is determined by various meteorological parameters, such as atmospheric temperature and pressure, wind speed, rainfall, relative humidity, and duration of average sunshine received. Rai and Varma (2010) executed an investigation at Zahedan and Ramsar stations with the help of multivariate regression. Considering the best and the worst-case scenarios, the correlation was 0.94 and 0.64, respectively. Bilgili (2011) performed experiments which overlanded a contrast scheme where he tested regression and artificial neural network models in predicting ST in the Adana province of Turkey. The authors used a new set of atmospheric data and monthly meteorological surveys, and he confined his result to the Artificial Neural Network

(ANN) where he exerted that an ANN is a more principal method for ST prediction.

Similarly, Tabari et al. (2011) investigated ANN for estimating ST for specific regions and approximated from the calculations that values of ANN models were better aggregates for enhanced observations than traditional regression methods. He could find accurate temperature through the model with depths up to 15 cm to 30 cm. Sulaiman et al. (2018) took the help of ANN models and coupled Gene Expression Programming (GEP) for estimating daily dew points. Dew points are effective parameters for data localisation in ST prediction but are limited to regional disparities. It was assumed dew points to be a sublimation for ST prediction. Tabari et al. (2014) carried out the propositional study where he used the ANN and multivariate regression (MLR) considering large atmospheric data points and estimated ST up to 100 cm of soil depth at five sites in Iran. The results were significant and explained the importance of relative humidity and mean temperature on soil average temperature. Kisi (2006) also investigated the use of generalized neural networks, MLP neural networks, and radial basis neural networks to estimate ST by harnessing meteorological surveyed data. However, in more recent study, Samadianfard et al. (2018) showed a different method where he proposed his study on ST by using wavelet neural networks (WNN) for profile analysis at depths. His study established the fact that with increasing soil depth, the accuracy for predicting ST appreciably reduces. Kazemi et al. (2018) demonstrated the supremacy of genetic neural network ensemble, commonly called as GNNE combined with Neighbourhood Cleaning Rule (NCL) and least mean squares (LMS) algorithm which outperformed other prediction methods of soil at different depths based on regression for the prediction of daily ST. Furthermore, the author proposed to include other fuzzy systems to make effective data decomposition algorithms that will help in constructing the components in GNNE.

Most recent studies conducted by Singhal et al. (2021) which developed a three-layer feed forward neural network models for the estimation of soil temperature in the Himalayan glaciated region. For this purpose, different input combinations was trained and tested temperatures using concurrent and antecedent air-soil temperature data. The results displayed the capability of ANN based models to provide promising result for soil-temperature prediction. Alizamir et al. (2021) compared the feasibility of novel deep echo state network (Deep ESN) model with three different traditional AI models for the prediction of ST at various depth. Several inputs were combined based on sensitivity analysis and the results were evaluated using statistical indices. The outcomes indicated that Deep ESN outperformed the other AI-based models in both training and testing phase. Cao et al. (2021) demonstrated the basis relationship between ST and mois-ture content at different depth level using an empirical modelling techniques. Yaseen (2021) studied an insight into machine learning for understanding the soil and water application using several AI based models, optimization algorithms, deep learning model, and hybrid models. Andrade et al. (2021) employed the capability of ANNs to simulate soil heat flux on land cover integrated with remote sensing and meteorological reanalysis. Empirical model was also employed for comparison, the obtained results indicated the robustness of AI-based models. However, most of the studies are performed on the data varying from the regions such as Qinghai-Tibetan Plateau [49], Turkish Regions Turkish State Meteorological Service [50]. They also reported a pattern of ST variation with increase in soil depth across 261 stations of Turkey while such ST changes were not seen in the regions. Insitu measurement of ST and soil moisture is successfully implemented by the Tibetan Plateau observatory [49], which have incorporated three regional scale regional networks equipped with satellite-based instruments and remote sensing sensors. But these can only be used to investigate the surface temperature reference and do not provide accurate ST at different depths across different regions. Many researchers have also investigated different machine learning techniques to compare the best suitable model for accurate prediction. One such investigation performed by [50] compared the three models for the same across regions in Turkey. The ANN, adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) models were trained, and the outcomes depicted that AFNIS outperformed

the other two models with RMSE score 1.99 and MAE 1.09 (indication for less errors thus more accuracy). This was a significant study as a more vali-dated model was proposed for prediction and the dependent variables resolved into simpler data points such as monthly minimum and maximum air temperatures, calendar month dates, precipitation, etc [51], [52]. Based on the reported literature, the Scopus database reported that there were 168 research articles starting from 2010 up to date adopted over the literature using the feasibility of machine learning models for the ST prediction. In Figure 1a, 262 keywords occurrence between those researches indicating the lack of deep learning implementation for ST simulation. Also, the popularity of this research domain explored for 48 regions all around the world majority for China, USA and India (Figure 1).

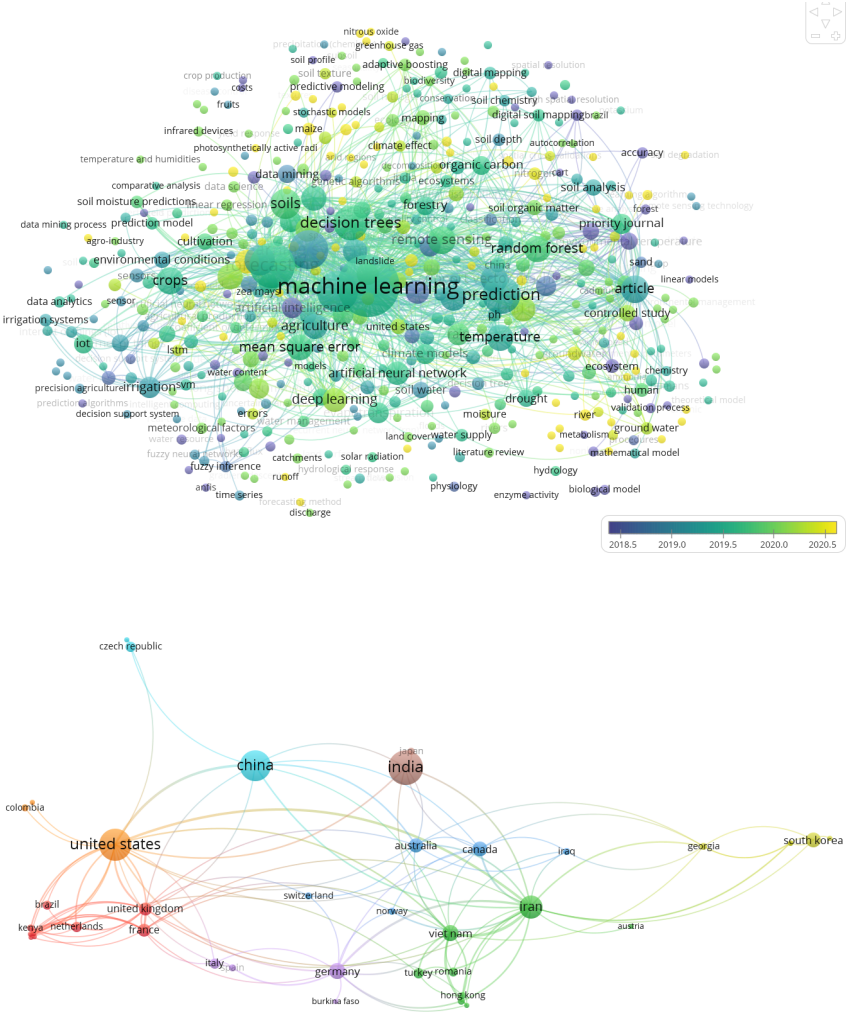


Figure 1: *a) the major keywords used over the literature on the ST simulation using machine learning models (2010-2021), b) the investigated region for the ST prediction.*

These studies have shown exemplary AI techniques for ST prediction, and each study has gradually progressed with more proofs of accuracy for observations and efficiency at increasing soil depths than the previous ones. However, these studies have only focused on an aspect of the problem, which is the use of better AI techniques each time. This research corresponds to the fact that, by targeting

modern deep learning models and easily fetchable data points for DSTs that can be calculated by low energy powered edge computing devices. This study mainly considers the efficiency of the computation of models and feeding the data acquired by low-cost devices across distributed regions such as hills, steeps, plateaus, etc. Estimation of long-term and average monthly ST is necessary to understand the properties of soil and acquire the numerical relationship of ecological variables. The present study emphasizes the use of Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Neural Network (NN), and Long Short-Term Memory Network (LSTM) and compares the accuracy score for better ST prediction based on atmospheric data points at a certain region.

The major problem with the ST prediction is the non-uniform nature of the variables that it depends upon. For a specific region, the variables can have different values and sometimes, during a major climatic change for a specific region, it becomes difficult to predict the ST based on these environmental factors. The previous models have not completely demonstrated a solution tha can address the problem of these rapidly changing parameters. Although, some researchers have made an effort to increase the accuracy over the large distribution of these factors using hydro-climatic parameters [53], there is no evidence on the effect of meteorological station over the ST prediction. Deep ESN model is a good approach to increase the RMSE for the 5-input and 1-input models, but concreted evidence over more distributed data could have given a proof for the single prediction method. This is the major problem with the prediction of ST as discussed earlier due to the non-uniform climatic nature of meteorological stations.

It is obvious that ST consists of complex processes, and it is complex behaviour are quite difficult to explain by simple process. This is owing to the fact that most of the traditional and mathematical methods were built on rough approximation that were known to be the conventional approach for determining the prediction skill of ST phenomena. Being the fundamental modelling methods, the traditional and physic-based methods explain the physical processes but still attributed with various weakness such as, time consuming, computational burden, failure to capture chaotic, and complex process. In contrast, it was known to overcome the above limitation but neglect the physical process, particularly when focus is on the accuracy and reliability of the estimation rather than understanding the simple physical process [54]–[57]. AI and traditional model were employed in our study. The proposed model is cost-effective since lightweight ML models can be used to perform predictions on edge devices with low computation power. This method can provide real-time temperature predictions, and future values can also be predicted by using the forecasted weather conditions. Enabling an integrated system for spontaneous ST prediction based on ecological data points. The proposed model can be fine-tuned to predict ST of a specific region. This helps in surveying in the depth of the region and enables soil profiling and modeling.

2. Materials and Methods

This section analyses the machine learning models formulation used in the prediction of ST with various steps and processes involved. Schematically, it involves feature extraction and normalization of the values. Normalization helps the algorithms to model the data correctly as the numeric values of columns are set to a more common scale without disturbing the difference in the ranges of values of information or losing values [58]–[61]. LR, RF, SVM, and LSTM models are used in independent combinations to process the data (see, Fig. 2). Equation 1 is used to scale and normalize the data set used in this study into a range of 0 to 1; this was done to ensure that the dependent and independent variables were equally treated, and that their dimensions were minimized equally. Prior to AI modeling, data normalization is commonly used to remove redundant data attributes and complex numerical errors.

$$y = 0.05 + \left(0.95 \times \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right)\right) \tag{1}$$

For the machine learning task, 3 models were chosen - SVM, LR and RF. Feature extraction was done on input features and data was normalized as shown in Fig. 2. The pre-processed data was fed to the respective ML models and predictions were inverse normalized to obtain the predicted temperature. One by one, all the three models were used and a comparative analysis was carried out. For the deep learning neural network (DNN) model, the input features are normalized to change mean to 0 and standard deviation to 1. The pre-processed features are then fed to a DNN and finally, a linear output layer is used for inference as shown in Fig. 3. Adam optimizer was used to perform back propagation.

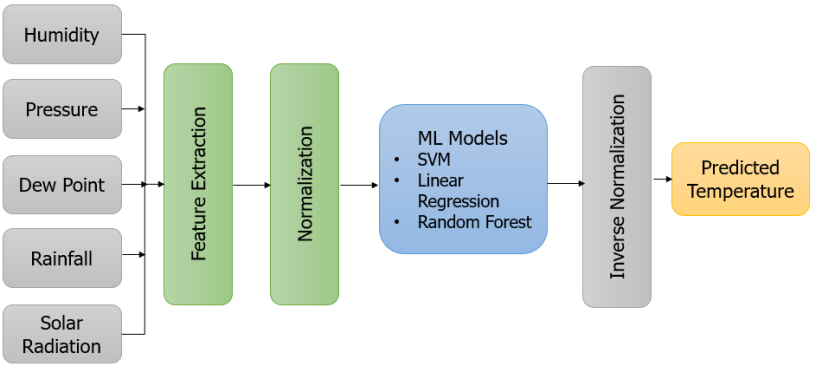


Figure 2: *The proposed ML Models.*

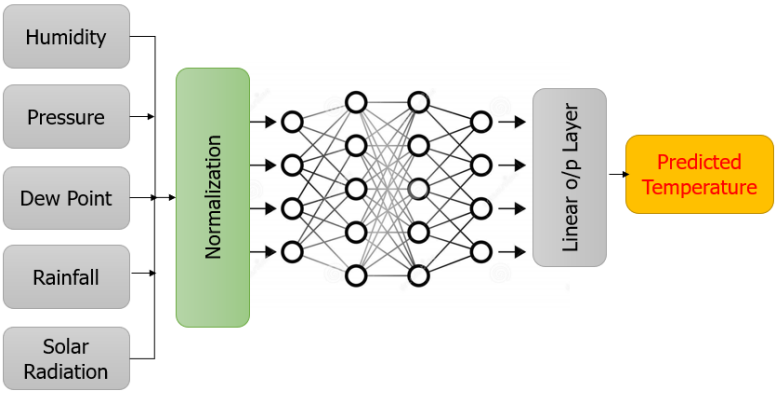


Figure 3: *Proposed DNN architecture.*

However, In LSTM architecture, the input features are normalized and fed to a LSTM cell, which consists of update, forget and output gates. The output of this cell is then fed to a dense neural layer. Finally, a single size dense layer is used as an inference layer to obtain the output. Adamax optimizer was used to perform back propagation as shown in Fig. 4. The dataset for this study has

been obtained from five locations in North Dakota which is located in the centre of North America. The five stations are climatically diverse, and the dataset obtained was recorded over 2010–2018. The data comprises readings of ST to accurate 10 cm of soil depth. It consists of hourly, daily and monthly ST patterns of the grand fork area. The study utilized the hourly humidity, dew point, rainfall, solar radiation and barometer readings for formulating our prediction model. Further, the data as well as target temperatures were normalized: z = x – u/sigma where, x is temperature, u = mean and sigma = standard deviation as stated above. A total of 43,824 samples were used to train our models. Validation split of was used to divide data. Training data size = 35059, Validation data size = 8765. Further, the daily and monthly averages can be derived from the hourly readings.

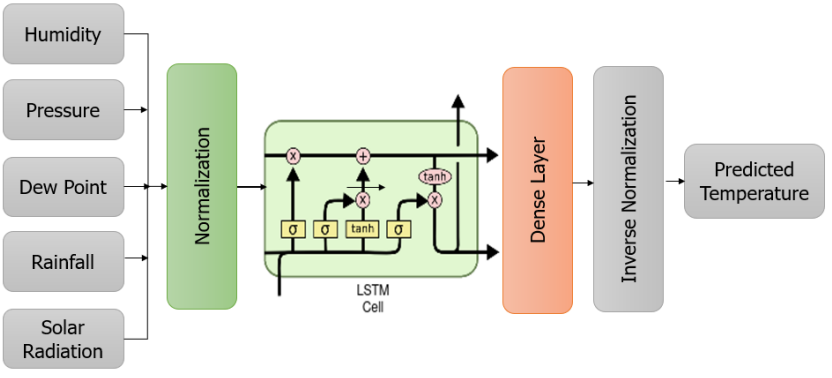


Figure 4: *Proposed DNN architecture.*

The fitness of the model within an acceptable dataset depending on the employed indicators is one of the key goals; this is to ensure that any unknown any machine learning or deep learning model can be used to simulate any unknown dataset and achieve trustworthy and resilient results. However, some concerns, such as local minima and overfitting demand the validation of the dataset since the training performance may not be sufficient, particularly when the study used a very small data set. Some of the available validation methods include cross-validation (i.e., k-fold), holdout, and leave one out. To ensure that the issue of overfitting is prevented in this work, the k-fold validation approach was used in this study [62].

2.1 Linear regression model (LR)

Linear regression is a simple technique that attempts to draw the relationship between two variables by fitting a linear equation [63]. The two variables are distinctly identified as the explanatory variable and the dependent variable [64]. The model is first examined for the suitability to perform the regression (having one relationship is mandatory). For the analysis of Linear Regression: while loss is not minimal: the output predictor variable Z is calculated as shown in Eq. (2):

$$Z = W \cdot X + b \tag{2}$$

where W is weights, b is bias, n is number of data points, alpha α is learning rate. The loss is defined as:

$$l = \frac{1}{n} \times \sum_{i=0}^{i=n} (Y_i - Z_i)^2 \tag{3}$$

Here, Y_i is actual value and Z_i is the predicted value, which is the difference between original and predicted values.

$$dW = -\frac{2}{n} \times \sum_{i=0}^{i=n} X(Y - Z) \tag{4}$$

And,

$$db = -\frac{2}{n} \times \sum_{i=0}^{i=n} (Y - Z) \tag{5}$$

To calculate the gradients for weights and bias respectively. $W = w - \alpha \times dw$ and $b = b - \alpha \times db$ are used to calculate the new values of weights and bias, thus completing one step of gradient descent algorithm. In this work, two parameters (i.e., X and Y) have been considered for analysis of ST. In this work, four different machine learning algorithm has been used. The array X consists of data points: Humidity, Dew Point, Pressure, Rainfall, Solar Radiation, which are used as features to predict the ST array Y. X = [Humidity, Dew Point, Pressure, Rainfall, Solar Radiation], Y = [Soil Temperature].

2.2 Support vector machine (SVM)

At the discovery of the support vector machine (SVM) concept, it was considered an observer-based learning method; however, SVM has been lately used successfully in several fields of chemistry, such as chromatography, synthesis, spectroscopy, and spectrometry [65]. SVM is generally divided into 2 types which are Support Vector Regression (SVR) that is mainly used for prediction works, and Support Vector Classifiers (SVC), mainly used for classification works. SVR can be denoted using:

$$f(x) = w \times \phi(x) + b \tag{6}$$

where, w, φ, and b are the vector weight (shown in the feature space), the transfer function, and the bias, respectively. The SVR function $f(x)$ can be represented by demonstrating the regression problems as follows:

Minimize

$$\frac{1}{2}|0w|0^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{7}$$

Subject to the condition:

$$\{y_i - f(x) \leq \varepsilon + \xi_i f(x) - y_i \leq \varepsilon + \xi_i^* \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, 3, \ldots, N \tag{8}$$

where C = the penalty parameter while ξ_i and ξ_i^* are two slack parameters. The non-linear regression function can be solved using the Lagrangian functions based on the optimization as follows:

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{9}$$

where; $K(x, x_i)$ is the kernel function and α_i and α_i^* are two variables; C, α_i and $\alpha_i^* > 0$. Numerous kernel functions are available, such as linear, radial basis function (RBF), sigmoid, and polynomial [66] but among these kernel functions, the most popular in the literature is the RBF kernel and this is the reason for its utilization in this work. The RBF kernel is defined as:

$$K(x, x_i) = \exp\left(-\gamma\,|0x_i - x|\,0^2\right) \tag{10}$$

where, γ is the kernel parameter, which means C, γ, and ϵ are three parameters that are responsible for SVR performance. A detailed description of SVR and SVM can be found in [67], [68].

2.3 Random forest (RF)

This is one of the effective supervised learning techniques that is mostly used for classification and regression problems in machine learning. It comprises a number of decision trees on various subsets of the dataset stream. It works on the concept of ensemble learning which aims to improve the accuracy of the dataset by combining multiple classifiers to solve a complex problem [67]. For the analysis of Random Forest (RF):

$$n_i = \text{ node importance} \tag{11}$$

The variables used to represent the tree are w (weighted samples reaching node) and C (impurity of node). To calculate the importance of each input feature are defined as:

$$\text{feature importance } = \frac{\text{summation}\,(n_i(j))}{\text{summation}\,(n_i)} \tag{12}$$

To calculate the norms of all the features importance are defined as:

$$\text{norm (feature importance)} = \frac{f_i(i)}{\text{summation }(f_i)} \tag{13}$$

To determine the final features used to form the RF tree is:

$$\text{Random Forest features } = \frac{\text{summation}\,(\text{norm}\,(f_i(i,j)))}{\text{number of trees}\,[i = \text{ node } i, j = \text{ tree } j]} \tag{14}$$

2.4 Neural network (NN)

Commonly called as the neural net, an Artificial Neural Network (ANN) or simply NN that resembles the biological method of human process information through computational models. It contains the basic unit of computation known as the neuron which is responsible for carrying the inputs to different nodes of decision and data processing [69], [70]. The nodes consist of Multi-Layer Perceptron (MLP) which is responsible to execute different functions on the received data. The best advantage of a neural net is if a neuron is not responding, the network can still detect the error and gener-ate the output. For the analysis of Neural Networks (NN):

$$Z_1 = W \cdot X + b_1 \tag{15}$$

The equation is used for 1st dense layer and activation function used for 1st dense layer is defined as $A_1 = \text{relu}\,(Z_1)$. The second dense layer is defined as $Z_2 = W \cdot A_1 + b_2$ with activation function as $A_2 = \text{relu}\,(Z_2)$. Finally, the final inference layer is defined as $Z_3 = W \cdot A_2 + b_3$. The gradient calculation of back-propagation for the 3rd dense layer is defined as $dZ_3 = Z_3 - Y$ and its weight is calculated as $dW_3 = (1/n) * dZ_3 \cdot A_2$. The gradient calculation for the bias of 3rd dense layer is defined as $db_3 = (1/n) *$ summation (dZ_3), and for 2nd dense layer is defined as $dZ_2 = W_2 \cdot dZ_3 * g'\,(Z_2)$. The gradient calculation for the weight of 2nd dense layer is defined as $dW_2 = (1/n) * dZ_2 * A_1$ and for the bias of 2nd dense layer is defined as $db_2 = (1/n) *$ summation (dZ_2). The gradient calculation for 1st dense layer is defined as $dZ_1 = W_1 \cdot dZ_2 * g'\,(Z_1)$, weight is denoted as $dW_1 = (1/n) * dZ_1 * X$ and bias of 1st layer is defined as $db_1 = (1/n) *$ summation (dZ_1).

2.5 Long short-term memory (LSTM)

LSTM was conceptually designed by Hochreiter and Schmidhuber, to overcome the problem of long-term dependencies of RNNs [71]. An RNN is not feasible in terms of accuracy of predictions when the gap length increases. It is not suitable for persistent information. On the other hand, LSTM can retain the information for a long period of time [72]. It is generally used for predicting and classifying on the basis of time series data. For the analysis of LSTM:

$$U(t) = \text{sigmoid}\left(W_u[A(t-1), X(t)] + b_u\right) \tag{16}$$

Equation (16) is defined as an update gate of LSTM. The Forget Gate is represented as:

$$F(t) = \text{sigmoid}\left(W_f[A(t-1), X(t)] + b_f\right) \tag{17}$$

And output gate is defined as:

$$O(t) = \text{sigmoid}\left(W_o[A(t-1), X(t)] + b_o\right) \tag{18}$$

The candidate cell and memory is defined as:

$$C \sim (t) = \tanh(Wc[A(t-1), X(t)] + bc) \text{ and } C(t) = U(t) * C \sim (t) + F(t) * C(t-1) \tag{19}$$

The activated output function is notated as:

$$A(t) = O(t) * \tanh(C(t)) \tag{20}$$

The dense layer output calculation is done on the basis of equation (21):

$$Z_1 = W_1 \cdot A(t) + b_2 \tag{21}$$

The activation function used on dense layer as:

$$A_1 = \text{relu}\left(Z1_1\right) \tag{22}$$

And output layer inference is defined as:

$$Z_2 = W_2 \cdot A_1 + b_2 \tag{23}$$

where W_u=Weights for update gate, b_u=bias unit for update gate, W_f=Weights for forget gate, b_f=bias unit for forget gate, W_o= eight for output gate, and b_o=bias unit for Output gate.

The primary purpose of any modelling is to fit the model to the given data based on the employed indicators to achieve reliable prediction on the unknown data set. Despite the promising ability of some AI-based models (e.g., ANN) in the forecasting of nonlinear systems, the issue of over fitting provides that there is a disparity between the achieved satisfactory training performance and the testing performance. Hence, combining multiple performance criteria is essential for good analysis. Different metrics were used to determine the performance accuracy; this was done by comparing the measured values with the predicted ones. The models were evaluated using R^2, MSE, and RMSE as the evaluation indices [73], where:

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left[(Y)_{obs,j} - (Y)_{com,j}\right]^2}{\sum_{j=1}^{N}\left[(Y)_{obs,j} - \underline{(Y)_{obs,j}}\right]^2} \tag{24}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(Y_{obsj} - Y_{comj}\right)^2 \tag{25}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(DO_{obs_j} - DO_{com_j}\right)^2}{N}} \tag{26}$$

where $N, Y_{obsf}, \underline{Y} and Y_{com\ j} are data number,$
observed data, the average value of the observed data and computed values, respectively.

2.6 Pseudo code

The pseudo code for proposed model gives a detailed explanation of how the model will execute the process and take different environmental factors into account. Dense layer setup is explained in Figure 5 which depicts the overall reducing the overfitting with the help of dropout regularization. Relu is used as the activation layer for the same. Optimizer helped in compiling the model. The model is compiled and fit into the validation data.

```
inp = Input(shape = (5))
x1 = Dense (512, activation = 'relu') (inp)
x1 = Dropout (0.25) (x1)
x2 = Dense (256, activation = 'relu') (inp)
out = Dense (1) (x2)
net = Model (inp, out)
net.compile('adamax','mean  squared  error')
net.fit (trx,np.array(trainy),validation_data = (tsx,np.array(testy)),epochs = 5)
```

```
inp = Input (shape = (5))
x1 = LSTM (64) (inp)
x2 = Dense (128, activation = 'relu') (x1)
x3 = Dropout (0.15) (x2)
out = Dense (1) (x3)
net = Model (inp, out)
net.compile ('adamax','mse')
net.fit (trx,np.array(trainy),validation_data = (tsx,np.array(testy)),epochs = 5)
```

Figure 5: *(a) Pseudo code for dense layer (b) Pseudo code for LSTM.*

2.7 Sample Data analysis

The brief analysis of the sample data collected for this experimental investigation is given In Table 1. The data is outsourced from different labelled data that provided a broad distribution of atmospheric pressures and its corresponding effect on ST. The given data has logged the soil bare and turf temperature, which are the target features. The predictor features consist of soil humidity, wind speed, solar radiation and atmospheric pressure. The data is evenly distributed as per the region it is collected from. It provides enough uniformity for the prediction of ST based upon these factors. The data points are gathered from some meteorological institutes and climate prediction centres from various regions and stations, as indicated on their official documented site on the internet.

Normalization was necessary in order to maintain a common scale for the numeric values in the columns which includes the attributes. The data, since it was gathered from different cited sources corresponded to various regions and had some irregularities such as missing values or non-optimized data. Normal distribution had to be created for this range set.

Table 1: *Sample data for the analysis of environmental factors.*

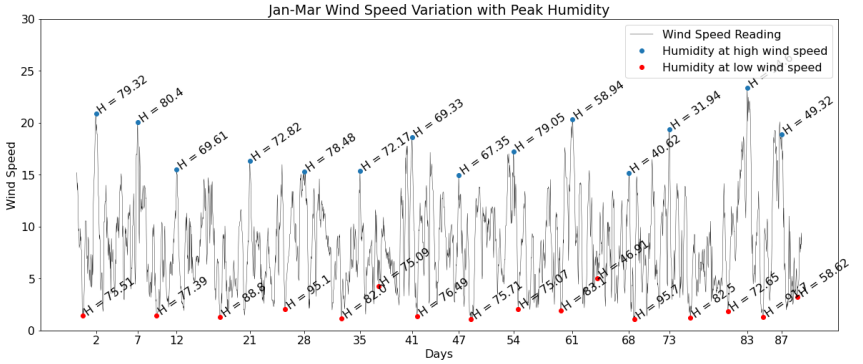
	Humidity (g.m-3)	Bare Temp (Degrees F)	Turf Temp (Degrees F)	Wind Speed (mph)	Solar Radiation (Lys)	Barometer	Dew Pt. (Degrees F)	Year
0	79.87	22.705	26.492	15.203	0.0	977.0	19.655	2015
1	87.60	23.005	26.688	15.037	0.0	976.0	22.288	2015
2	93.00	23.301	26.886	13.516	0.0	976.0	23.549	2015
3	92.10	23.612	27.070	14.224	0.0	976.0	24.410	2015
4	92.70	23.887	27.248	13.003	0.0	976.0	25.010	2015
...	...	...	...	...	...	...	...	...
43819	82.30	31.041	30.906	10.636	0.0	971.6	6.008	2019
43820	82.80	31.039	30.906	8.391	0.0	970.2	6.604	2019
43821	84.70	31.046	30.909	10.600	0.0	968.4	6.125	2019
43822	85.10	31.030	30.895	10.427	0.0	967.9	6.444	2019
43823	85.50	31.014	30.870	9.171	0.0	966.9	5.628	2019

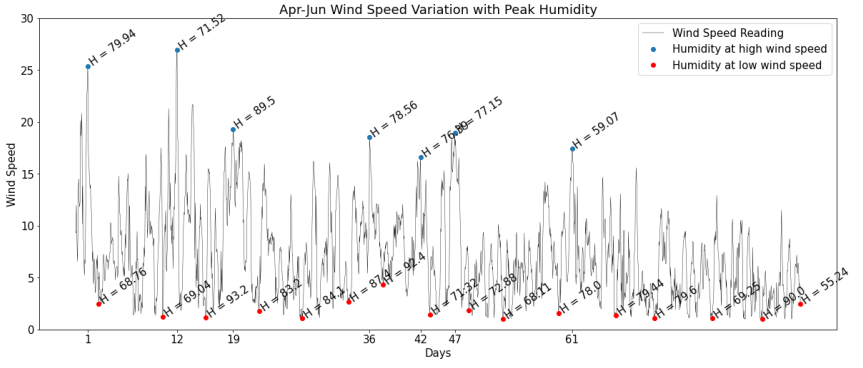
3. Application Results and Discussion

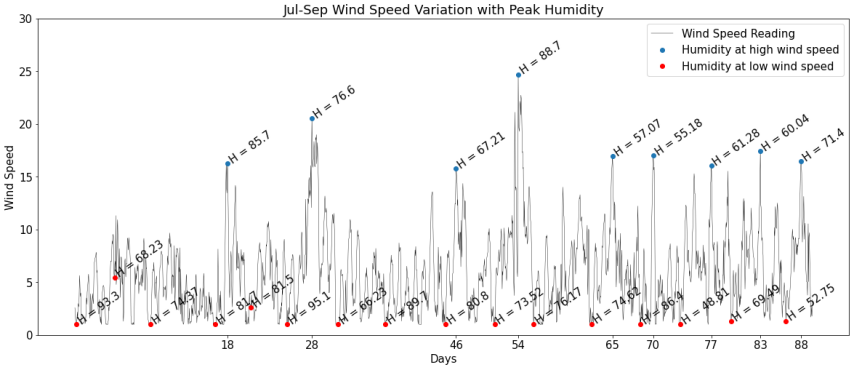
The major motivation and contribution of this research is to develop and to comparatively study a machine learning models (SVM, RF, NN, LSTM) and a linear model (LR) for the determination of ST. In this section, the results obtained are demonstrated both in visualized and quantitative format. A preanalysis on the raw data including maximum and minimum humidity levels is also marked on the graphs to indicate that on days with low wind speed. Prior to the pre-diction pre-processing of data was carried out using various methods including normalization, and standardization. For any time-series data, pre-analysis of the individual data, i.e., the individual inputs, is paramount because their ac-curacy can significantly contribute to determining the efficiency of the individual models. Hair et al. [74] claim that a dimension's variables are internally consistent if their Cronbach's alpha values are greater than 0.7. According to Dickey and Fuller, [75], the ADF test is used to assure more reliable and valid outcomes as well as the stationarity of all variables. The data utilized in this investigation confirmed that the unit root test met all of the stationarity requirements.

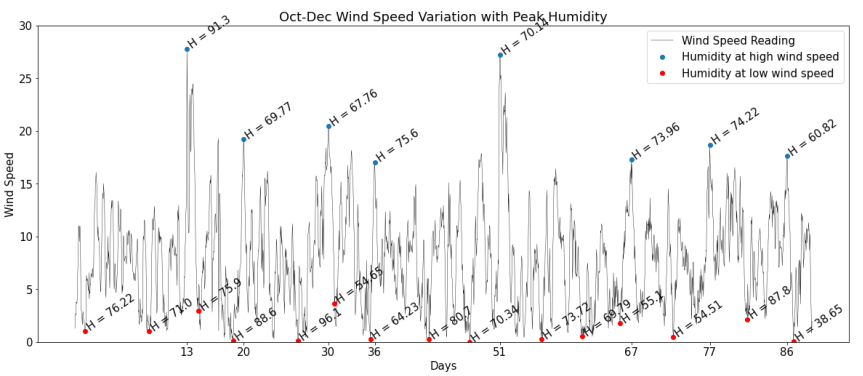
However, various graphs are obtained from the month-wise distribution of wind speed vs peak humidity at fixed geolocation points in the North Dakota region (see, Fig. 6. On various days, some trends were observed which helped in setting up the correlation between the variables such as humidity, dew point, wind speed, etc. These parameters helped in predicting the ST by calculating the average on that scale. The correlation analysis technique-based graph also determines the bearing and level of correlation between the variables, and this helps in the formation of the 3 distinct classes. The analysis further enhanced the understanding of the mechanism and science of the data by displaying the required input parameter with the strongest correlation with the output variable for the modelling process. This can also aid with experimental analysis, particularly model selection. As a result, correlation analysis is the basis for the formation of numerous models used to observe the practical influence of selected methods on the simulation process.

From the Fig. 6a it shows the daily wind speed variation for the 1st quarter of the year. Maximum and minimum humidity levels are also marked on the graphs to indicate that on days with low wind speed, humidity is generally high. Fig. 6b depicted that the daily wind speed variation for the 2nd


Jan-Mar Wind Speed Variation with Peak Humidity
Wind Speed Reading
Humidity at high wind speed
Humidity at low wind speed
Wind Speed
Days


Apr-Jun Wind Speed Variation with Peak Humidity
Wind Speed Reading
Humidity at high wind speed
Humidity at low wind speed
Wind Speed
Days


Jul-Sep Wind Speed Variation with Peak Humidity
Wind Speed Reading
Humidity at high wind speed
Humidity at low wind speed
Wind Speed
Days


Oct-Dec Wind Speed Variation with Peak Humidity
Wind Speed Reading
Humidity at high wind speed
Humidity at low wind speed
Wind Speed
Days

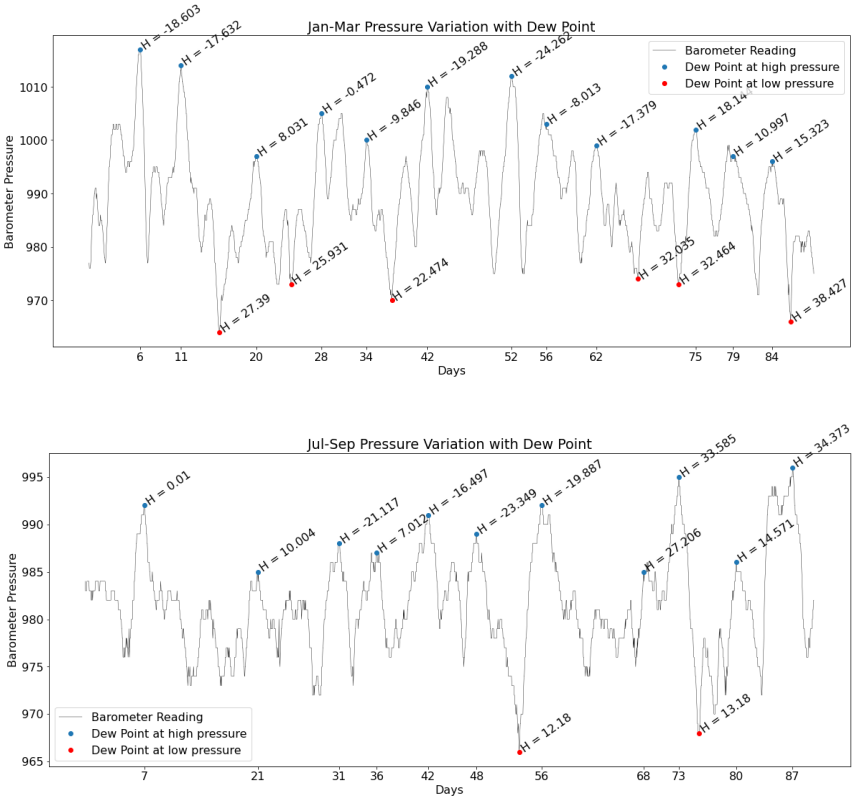


Figure 6: *Wind speed variation with Peak Humidity for (a) Jan-March (b) April-June (c) June -Sept (d) Oct-Dec, and with Pressure Variation with Dew Point (e) Jan-June (f) July-Dec.*

quarter of the year. Maximum and minimum humidity levels are also marked on the graphs to indicate that on days with low wind speed, humidity is generally high. Fig. 6c shows the daily wind speed variation for the 3rd quarter of the year. Maximum and minimum humidity levels are also marked on the graphs to indicate that on days with low wind speed, humidity is generally high. However, Fig. 6d indicated that the daily wind speed variation for the 4th quarter of the year. Maximum and minimum humidity levels are also marked on the graphs to indicate that on days with low wind speed, humidity is generally high. Whereby Fig. 6e shows the pressure variation on all days and for the first half of the year, dew point rises with decrease in pressure, and Fig. 6f showed that the pressure variation on all days and for the second half of the year, dew point drops with decrease in pressure.

In this section, the results of AI (LSTM, SVM, NN, and RF) and deterministic linear model (LR) have been discussed in relative terms of the data collected throughout the investigation. Various observations were recorded with the help of correlation between the dependent and independent variables across five major models used for prediction of ST points. The SVM, LSTM, NN, and RF models were implemented using MATLAB 9.3 (R2020a) while the LR model was implemented by using EViews (version 9.5). According to Abba et al.[76] determining suitable hidden nodes is the crucial aspect of any ANN modelling to avoid over-fitting caused by different factors. As reported in several studies in the field of science and engineering, there is no particular standard for determining the appropriate number of hidden neurons. It should be noted that the best hyper-parameter's structure was attained using trial and error for all four models. The Lavenberg-Marquartd algorithm

was used in training the ANN model due to its excellent performance in relevant studies; it is found to improve convergence speed and training effectiveness by resolving non-linear least squares problems. For the SVM, RF, and LSTM modelling, various types of parametric functions and iterations were explored using trial and error to identify the best structure.

The performance results of the four models are shown in Table 2. R^2 score describes the proportion of the variance in the dependent variable that can be easily predicted from the correlated independent variables. A 100% score indicates that the two variables are completely in correlation and a lower value would show a low level of correlation, indicating that the regression model is not valid but not limited to all possible cases. This is well defined as the total variance explained by the model/total variance. The MSE is defined as the taken average of the square of the errors. A larger value indicates a larger error in the estimated values since it describes the difference in the estimated values and what is estimated in the experiment. This is important for the study to propose a more accurate and near to the correct values of the estimated ST points via the developed model. RMSE defines the standard deviation from the estimated values in terms of the average of squares of errors. This is more relevant to this study as it consists of large values for the prediction made on the dataset. The higher the values indicate better performance of the compared model. It expresses more accuracy when comparing the performance of dependent and independent variables of different models. As it can be in Table 2, the best combinations were reported, and the results were presented in both calibration and verification phases.

Table 2: *Prediction accuracy results.*

	Calibration Phase				Verification Phase			
Techniques	R2	MSE	RMSE	R	R2	MSE	RMSE	R
LR-model	0.8050	49.3295	7.0310	0.8972	0.8000	49.2795	7.0199	0.8944
RF-Model	0.9425	14.5871	3.8193	0.9708	0.9375	14.5371	3.8128	0.9682
SVM-Model	0.9493	12.8423	3.5836	0.9743	0.9443	12.7923	3.5766	0.9718
NN-Model	0.9521	12.1573	3.4862	0.9758	0.9471	12.1073	3.4796	0.9732
LSTM-Model	0.9562	11.1087	3.3320	0.9779	0.9512	11.0587	3.3255	0.9753

Modelling ST from atmospheric data points using machine learning technique is paramount important as mentioned in section 1. The results suggest different levels of efficiency of the predictive approaches using the considered performance metrics. This is due to the fact that the level of robustness of each model differs according to the nature of capturing the data pattern between the input and target variables. In general, all the models (RF, LSTM, SVM, and NN) have shown a certain level of accepted accuracy including the LR. Among the 4 models LSTM served as the best model for the prediction of ST, this can be proved by considering the performance criteria (R^2, MSE, RMSE, and R).

The promising capability of the LSTM model is certainly not surprising, because it is an emerging non-linear deep learning model and has shown better predictive ability in various studies [77], [78]. Even though the models cannot be ranked based on the achieved accuracies, the best prediction accuracy was achieved by the LSTM and NN techniques in terms of ST behaviour modelling as the models exhibited > 95% performance efficiency in the calibration and verification stages. However, the overall accuracy of the models is satisfactory with regards to capturing nonlinear relationship between predictors and their corresponding targets.

From the numerical observation with regards to goodness-of-fit and absolute error indicated that LSTM (R^2=0.9512 and RMSE= 3.3255), NN (R^2=0.9471, and RMSE=3.4796), SVM (R^2=0.9443, and RMSE=3.5766), RF (R^2=0.9375, and RMSE=3.8128, and LR (R^2=0.800, and RMSE=7.0199). Based on the performance skill of the four model it could be observed that LSTM model reduced the

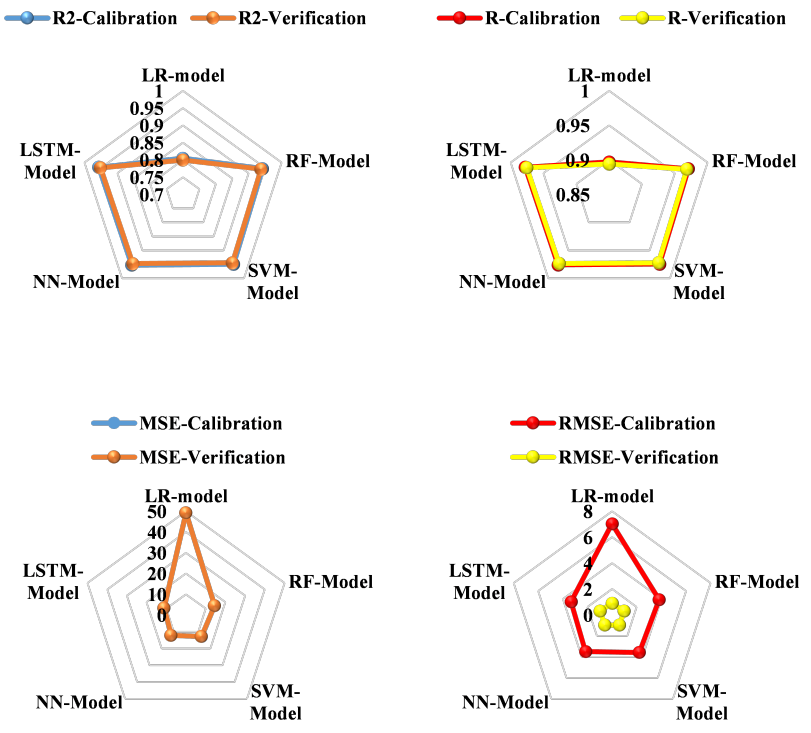


Figure 7: *Radar chart showing the discrepancies between the prediction skill of the models.*

prediction error averagely by 6% compared to NN, SVM and RF models and in-creases the prediction accuracy of SVM, NN and LR up to averagely 1% and 15%, respectively. The marginal accuracy attained by LR could be improved by developing the hybrid technique through the applications of the optimizations algorithms or ensemble learning techniques as demonstrated by [80]–[83]. Note that the promising estimations were observed at the calibration phase, which is generally used to correctly calibrate models using known input parameters and targets. The verification process, on the other hand, is critical in evaluating a model's performance since it checks the level of accuracy of the model based on unknown set values. The calibration set does not offer this advantage and as a result, it is expected that a reliable model should exhibit stable and balanced performances at the calibration and verification phases. Furthermore, the time series plot can help visualize the comparative study of the models (see, Fig. 8); it is a sophisticated graphical representation of data that provides a high-level overview, as well as a numerical summary of datasets. The plot sowed that the best model is the one with the nearest prediction to the experimental values. Based on the spread of the values achieved by the prediction and experimental models; the ranking of the models was as follows: LSTM>NN>SVM>LR, showing that LSTM is the best among all the models.

The blue highlighted curves indicate the plot of the original values of ST at the specific geolocation points and atmospheric conditions. The yellow highlighted curves depict the predicted ST points collected from the LR model created for this study. For the RF model, the correlation between the models has been majorly covered in the later part of the prediction. The initial instances were quite low for this model. All the specifications were similar from the previous model. RF is capable of handling missing values and is less impacted from the unnecessary noise due to the irregular variation in the dataset obtained. Moreover, SVM model schemes out a similar trend to RF, but it

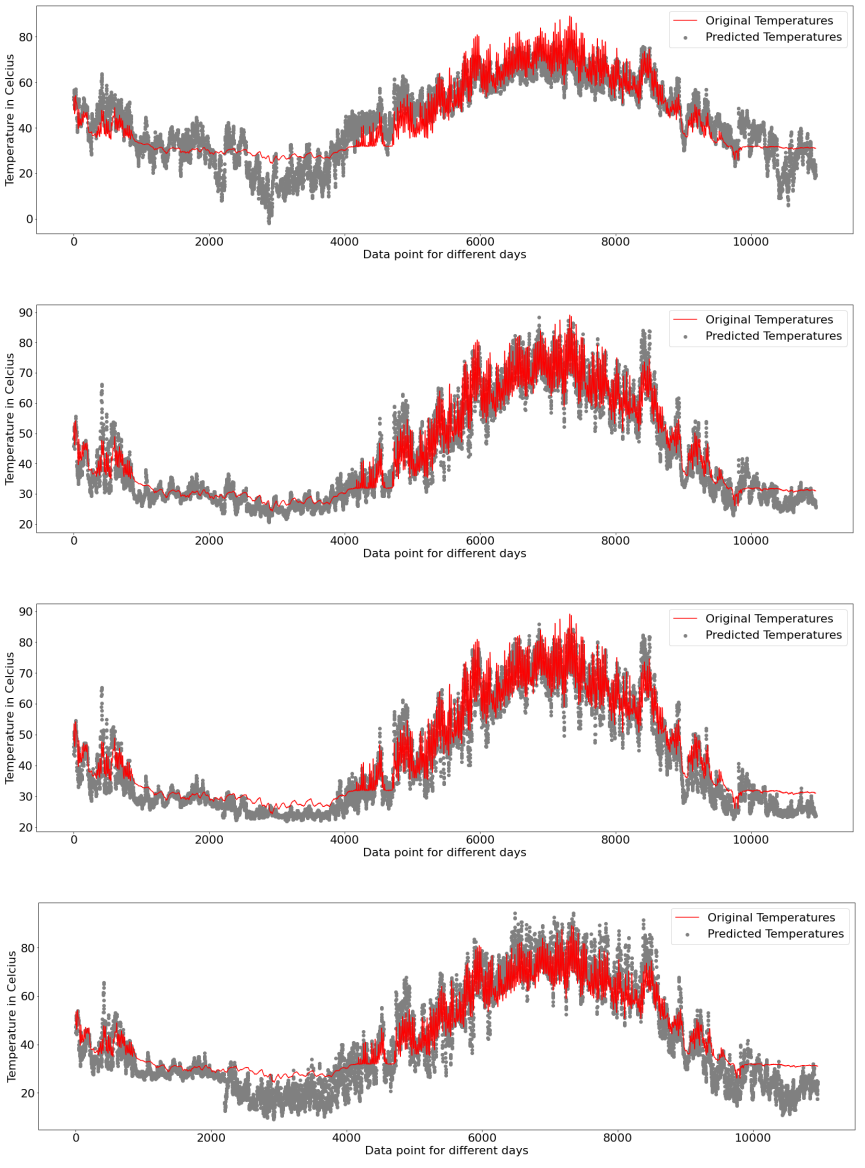


Figure 8: *Time series plot between the observed and predicted ST.*

was quite more stable in the performance as the dataset was sparse and irregular. It can be observed from the higher peak and the damping peak over the two models and comparing them. A little difference was there between the actual value the calculated values of ST. It can be seen LSTM model outperformed all ML models as well as the DNN. This is due to LSTM blocks have the capability to keep a track of the data by using the gates present. MSE of 11.1087 is well under the limits for predicting temperatures accurately. This study in line with one conducted by Singhal et al. [45] which study different architecture of ANNs for the prediction of soil temperature in the Himalayan glaciated region. The value of determination coefficient obtained was in the range of 0.8 to 0.9. Others studies also proved the capability of AI based models in handling nonlinear chaotic system such as ST modelling [7], [33], [84]–[86].

4. Conclusions

Soil temperature has great potential use in determining the distribution of plants and the prediction of the suitability of plant biodiversity. It affects many biological activities, humans, animals, and cycles of ecosystems such as water movement, nitrification, transpiration, etc. Moreover, studying the ST and soil profile helps in agricultural research to study about prolific growth and development of roots and vegetative parts. ST also provides appropriate tools to capture the spread of plant diseases and control methods of the specific bio-vegetative region. In the present research, we have drawn a comparative analysis of determining the model for predicting ST by harnessing the atmospheric data points. For the analysis, various machine learning techniques such as LR, RF, SVM, NN, and LSTM were used on the meteorological dataset of North Dakota to create deterministic models for collecting ST at various depths of soil surface. The data was collected from the atmospheric dataset that incorporated the readings from an experimental investigation based on precision instruments. the performance criteria. Various performance indicators including (MSE, RMSE, R^2, and R) were used to assess the performance evaluation of the models. Based on the performance skill of the four model it could be observed that LSTM model decreases the prediction error averagely by 6% compared to NN and SVM models and increases the prediction accuracy of SVM, NN and LR up to averagely 1% and 15%, respectively. The overall prediction results suggested that other optimization method for example, genetic algorithms, ensemble learning techniques, nature and bioinspired algorithms should be employed to improve the prediction results.

Funding: **There is no funds received.**

Data Availability Statement: **Data can be provided upon request from the corresponding author.**

Conflicts of Interest: **The authors declare no conflict of interest.**

Ethical Approval: **The manuscript is conducted within the ethical manner advised by the target-ed journal.**

Consent to Participate: **Not applicable**

Consent to Publish: **The research is scientifically consent to be published.**

Acknowledgement: **The authors acknowledge the data sources provider The United States Geological Survey (USGS).**

References

[1] J. S. Kemp, E. Paterson, S. M. Gammack, M. S. Cresser, and K. Killham, "Leaching of genetically modified Pseudomonas fluorescens through organic soils: Influence of temperature, soil pH, and roots," Biol. Fertil. Soils, vol. 13, no. 4, pp. 218–224, 1992.

[2] C. Huang, W. Chen, Y. Li, H. Shen, and X. Li, "Assimilating multi-source data into land surface model to simultaneously improve estimations of soil moisture, soil temperature, and surface turbulent fluxes in irrigated fields," Agric. For. Meteorol., vol. 230–231, pp. 142–156, 2016.

[3] M. E. Essington, J. E. Foss, and Y. Roh, "Reproduced from Soil Science Society of America Journal . Published by Soil Science Society of America . All copyrights reserved . The Soil Mineralogy of Lead at Horace ' s Villa," Scanning, no. 1963, pp. 2069–2077, 2004.

[4] T. Jackson, K. Mansfield, M. Saafi, T. Colman, and P. Romine, "Measuring soil temperature and moisture using wireless MEMS sensors," Meas. J. Int. Meas. Confed., vol. 41, no. 4, pp. 381–390, 2008.

[5] R. K. George, "Prediction of Soil Temperature by Using Artificial Neural Networks Algorithms," Nonlinear Anal. Theory, Methods & Appl., vol. 47, no. 3, pp. 1737–1748, 2001.

[6] R. Tur and S. Yontem, "A Comparison of Soft Computing Methods for the Prediction of Wave Height Parameters," Knowledge-Based Eng. Sci., vol. 2, no. 1, pp. 31–46, 2021.

[7] B. Qian, E. G. Gregorich, S. Gameda, D. W. Hopkins, and X. L. Wang, "Observed soil temperature trends associated with climate change in Canada," J. Geophys. Res. Atmos., vol. 116, no. 2, pp. 1–16, 2011.
[8] P. Sviličić, V. Vučetić, S. Filić, and A. Smolić, "Soil temperature regime and vulnerability due to extreme soil temperatures in Croatia," Theor. Appl. Climatol., vol. 126, no. 1–2, pp. 247–263, 2016.
[9] D. T. Coelho and R. F. Dale, " An Energy-Crop Growth Variable and Temperature Function for Predicting Corn Growth and Development: Planting to Silking 1 ," Agron. J., vol. 72, no. 3, pp. 503–510, 1980.
[10] A. Krishnan and G. G. S. N. Rao, "Soil temperature regime in the arid zone of India," Arch. für Meteorol. Geophys. und Bioklimatologie Ser. B, vol. 27, no. 1, pp. 15–22, 1979.
[11] M. Bayatvarkeshi et al., "Modeling soil temperature using air temperature features in diverse climatic conditions with complementary machine learning models," Comput. Electron. Agric., vol. 185, p. 106158, 2021.
[12] F. Shati, S. Prakash, H. Norouzi, and R. Blake, "Assessment of differences between near-surface air and soil temperatures for reliable detection of high-latitude freeze and thaw states," Cold Reg. Sci. Technol., vol. 145, no. October 2017, pp. 86–92, 2018.
[13] A. Araghi, M. Mousavi-Baygi, J. Adamowski, C. Martinez, and M. van der Ploeg, "Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network," Meteorol. Appl., vol. 24, no. 4, pp. 603–611, 2017.
[14] R. HUANG et al., "Soil temperature estimation at different depths, using remotely-sensed data," J. Integr. Agric., vol. 19, no. 1, pp. 277–290, 2020.
[15] B. Souffaché, P. Kessouri, P. Blanc, J. Thiesson, and A. Tabbagh, "First Investigations of In Situ Electrical Properties of Limestone Blocks of Ancient Monuments," Archaeometry, vol. 58, no. 5, pp. 705–721, 2016.
[16] Z. M. Yaseen, S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction," J. Hydrol., vol. 569, pp. 387–408, 2019.
[17] S. J. Hadi, S. I. Abba, S. S. H. Sammen, S. Q. Salih, N. Al-ansari, and Z. M. Yaseen, "Non-Linear Input Variable Selection Approach Integrated With Non-Tuned Data Intelligence Model for Streamflow Pattern Simulation," pp. 1–16, 2019.
[18] Z. M. Yaseen et al., "Rainfall Pattern Forecasting Using Novel Hybrid Intelligent Model Based ANFIS-FFA," Water Resour. Manag., vol. 32, no. 1, pp. 105–122, 2017.
[19] Z. M. Yaseen et al., "Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis," Water (Switzerland), 2019.
[20] Z. M. Yaseen, "An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions," Chemosphere, vol. 277, p. 130126, Aug. 2021.
[21] M. Hameed, S. S. Sharqi, Z. M. Yaseen, H. A. Afan, A. Hussain, and A. Elshafie, "Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia," Neural Comput. Appl., vol. 28, no. 1, pp. 893–905, 2017.
[22] Q. B. Pham et al., "Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall," Water Resour. Manag., vol. 33, no. 15, pp. 5067–5087, 2019.
[23] A. G. Usman, S. Işik, and S. I. Abba, "A Novel Multi-model Data-Driven Ensemble Technique for the Prediction of Retention Factor in HPLC Method Development."
[24] S. I. Haruna et al., "Compressive Strength of Self-Compacting Concrete Modified with Rice Husk Ash and Calcium Carbide Waste Modeling: A Feasibility of Emerging Emotional Intelligent Model (EANN) Versus Traditional FFNN," Arab. J. Sci. Eng., no. June, 2021.
[25] K. Mahmoud et al., "Prediction of the effects of environmental factors towards COVID-19

outbreak using AI-based models," IAES Int. J. Artif. Intell., vol. 10, no. 1, pp. 35–42, 2021.
[26] M. H. Ahmad, A. G. Usman, and S. I. Abba, "Comparative performance of extreme learning machine and Hammerstein–Weiner models for modelling the intestinal hyper-motility and secretory inhibitory effects of methanolic leaf extract of Combretumhypopilinum Diels (Combretaceae)," Silico Pharmacol., vol. 9, no. 1, 2021.
[27] H. U. Abdullahi, A. G. Usman, and S. I. Abba, "Modelling the Absorbance of a Bioactive Compound in HPLC Method using Artificial Neural Network and Multilinear Regression Methods," vol. 6, no. 2, pp. 362–371, 2020.
[28] S. Idris, M. A. A. Musa, S. I. Haruna, U. U. A. A. G. Usman, and M. I. A. Abba, "Implementation of soft-computing models for prediction of flexural strength of pervious concrete hybridized with rice husk ash and calcium carbide waste," Model. Earth Syst. Environ., 2021.
[29] S. I. Abba, A. G. Usman, and S. IŞIK, "Simulation for response surface in the HPLC optimization method development using artificial intelligence models: A data-driven approach," Chemom. Intell. Lab. Syst., vol. 201, no. April, 2020.
[30] Z. S. A. A.S. Mubarak, Parvaneh Esmaili, M. S. G. , R.A. Abdulkadir, M. Ozsoz, and S. I. A. , Gaurav Saini, "Metro-environmental data approach for the prediction of chemical oxygen demand in new Nicosia wastewater treatment plant," vol. 27049, pp. 1–10, 2021.
[31] S. Oleiwi, S. Jalal, S. Hamed, S. Ozgur, K. Zaher, and M. Yaseen, "Precipitation pattern modeling using cross-station perception: regional investigation," Environ. Earth Sci., vol. 0, no. 0, p. 0, 2018.
[32] H. Tao et al., "Artificial intelligence models for suspended river sediment prediction: state-of-the art, modeling framework appraisal, and proposed future research directions," Eng. Appl. Comput. Fluid Mech., vol. 15, no. 1, pp. 1585–1612, Jan. 2021.
[33] K. M. Hinkel, F. Paetzold, F. E. Nelson, and J. G. Bockheim, "Patterns of soil temperature and moisture in the active layer and upper permafrost at Barrow , Alaska: 1993 – 1999," pp. 1993–1999, 2001.
[34] G. Mihalakakou, "On estimating soil surface temperature profiles," Energy Build., vol. 34, no. 3, pp. 251–259, Mar. 2002.
[35] R. J. Hanks, D. D. Austin, and W. T. Ondrechen, "Soil temperature estimation by a numerical method," Soil Sci. Soc. Am. J., vol. 35, no. 5, pp. 665–667, 1971.
[36] M. Bilgili, "Prediction of soil temperature using regression and artificial neural network models," Meteorol. Atmos. Phys., vol. 110, no. 1–2, pp. 59–70, 2010.
[37] M. Rai and A. Varma, "Mycotoxins in food, feed and bioweapons," Mycotoxins Food, Feed Bioweapons, no. January, pp. 1–405, 2010.
[38] M. Bilgili, "The use of artificial neural networks for forecasting the monthly mean soil temperatures in Adana, Turkey," Turkish J. Agric. For., vol. 35, no. 1, pp. 83–93, 2011.
[39] H. Tabari, A. A. Sabziparvar, and M. Ahmadi, "Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region," Meteorol. Atmos. Phys., vol. 110, no. 3, pp. 135–142, 2011.
[40] S. O. Sulaiman, J. Shiri, H. Shiralizadeh, O. Kisi, and Z. M. Yaseen, "Precipitation pattern modeling using cross-station perception: regional investigation," Environ. Earth Sci., 2018.
[41] H. Tabari, P. Hosseinzadeh Talaee, and P. Willems, "Short-term forecasting of soil temperature using artificial neural network," Meteorol. Appl., vol. 22, no. 3, pp. 576–585, 2014.
[42] Ö. Kisi, "Generalized Regression Neural Networks for Evapotranspiration modelling," Hydrol. Sci. J., vol. 51, no. 6, pp. 1092–1105, 2006.
[43] S. Samadianfard et al., "Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths," Soil Tillage Res., vol. 175, pp. 37–50, 2018.
[44] S. M. R. Kazemi et al., "Novel genetic-based negative correlation learning for estimating

soil temperature," Eng. Appl. Comput. Fluid Mech., 2018.
[45] M. Singhal, A. C. Gairola, and N. Singh, "Artificial neural network-assisted glacier forefield soil temperature retrieval from temperature measurements," Theor. Appl. Climatol., vol. 143, no. 3–4, pp. 1157–1166, 2021.
[46] M. Alizamir, S. Kim, M. Zounemat-Kermani, S. Heddam, A. H. Shahrabadi, and B. Gharabaghi, "Modelling daily soil temperature by hydro-meteorological data at different depths using a novel data-intelligence model: deep echo state network model," Artif. Intell. Rev., vol. 54, no. 4, pp. 2863–2890, 2021.
[47] Z. Cao, S. Mu, L. Xu, M. Shao, and H. Qu, "Causal Research on Soil Temperature and Moisture Content at Different Depths," IEEE Access, vol. 9, pp. 39077–39088, 2021.
[48] B. C. C. de Andrade et al., "Artificial Neural Network Model of Soil Heat Flux over Multiple Land Covers in South America," Remote Sens., vol. 13, no. 12, p. 2337, 2021.
[49] Z. Su et al., "The tibetan plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products," Hydrol. Earth Syst. Sci., vol. 15, no. 7, pp. 2303–2316, 2011.
[50] H. Citakoglu, "Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey," Theor. Appl. Climatol., vol. 130, no. 1–2, pp. 545–556, 2017.
[51] B. B. Arikan, L. Jiechen, I. I. D. Sabbah, A. Ewees, R. Homsi, and S. O. Sulaiman, "Dew Point Time Series Forecasting at the North Dakota," Knowledge-Based Eng. Sci., vol. 2, no. 2, pp. 24–34, 2021.
[52] S. A. Faskari, G. Ojim, T. Falope, Y. B. Abdullahi, and S. I. Abba, "A Novel Machine Learning based Computing Algorithm in Modeling of Soiled Photovoltaic Module," Knowledge-Based Eng. Sci., vol. 3, no. 1, pp. 28–36, 2022.
[53] O. H. Kombo, S. Kumaran, Y. H. Sheikh, A. Bovim, and K. Jayavel, "Long-term groundwater level prediction model based on hybrid KNN-RF technique," Hydrology, vol. 7, no. 3, pp. 1–24, 2020.
[54] S. I. Abba, R. A. Abdulkadir, M. S. Gaya, M. A. Saleh, P. Esmaili, and M. B. Jibril, "Neuro-fuzzy ensemble techniques for the prediction of turbidity in water treatment plant," 2019 2nd Int. Conf. IEEE Niger. Comput. Chapter, Niger. 2019, pp. 1–6, 2019.
[55] S. I. Abba et al., "Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration," IEEE Access, vol. 8, no. September, pp. 157218–157237, 2020.
[56] V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence - An ensemble approach," Water Sci. Technol., vol. 78, no. 10, pp. 2064–2076, 2018.
[57] S. I. Abba, G. Elkiran, and V. Nourani, "Non-linear Ensemble Modeling for Multi-step Ahead Prediction of Treated COD in Wastewater Treatment Plant," vol. 2, pp. 683–689, 2020.
[58] M. S. Gaya, S. I. Abba, A. M. Abdu, and A. I. Tukur, "Estimation of water quality index using artificial intelligence approaches and multi-linear regression," vol. 9, no. 1, pp. 126–134, 2020.
[59] A. G. USMAN, S. IŞIK, S. I. ABBA, and F. MERİÇLİ, "Artificial intelligence–based models for the qualitative and quantitative prediction of a phytochemical compound using HPLC method," Turkish J. Chem., vol. 44, no. 5, pp. 1339–1351, 2020.
[60] Z. M. Yaseen et al., "Predicting compressive strength of lightweight foamed concrete using extreme learning machine model," Adv. Eng. Softw., vol. 115, pp. 112–125, Jan. 2018.
[61] S. I. Abba, S. J. Hadi, and J. Abdullahi, "River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques," Procedia Comput. Sci., vol. 120, pp. 75–82, 2017.
[62] S. I. Abba et al., "Results in Engineering Emerging Harris Hawks Optimization based load demand forecasting and optimal sizing of stand-alone hybrid renewable energy systems – A case

study of Kano and Abuja , Nigeria," Results Eng., vol. 12, no. July, p. 100260, 2021.
[63] O. O. Aalen, "A linear regression model for the analysis of life times," Stat. Med., 1989.
[64] S. S. Sonawane, S. S. Chhajed, S. S. Attar, and S. J. Kshirsagar, "An approach to select linear regression model in bioanalytical method validation," J. Anal. Sci. Technol., 2019.
[65] S. Raghavendra and P. C. Deka, "Support vector machine applications in the field of hydrology: A review," Appl. Soft Comput. J., vol. 19, pp. 372–386, 2014.
[66] Z. M. Yaseen et al., "Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq," J. Hydrol., vol. 542, pp. 603–614, 2016.
[67] L. Breiman, "Random Forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
[68] T. Zhou, F. Wang, and Z. Yang, "Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction," Water (Switzerland), vol. 9, no. 10, 2017.
[69] N. M. Gazzaz, M. K. Yusoff, A. Z. Aris, H. Juahir, and M. F. Ramli, "Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors," Mar. Pollut. Bull., vol. 64, no. 11, pp. 2409–2420, 2012.
[70] M. P. Darji, V. K. Dabhi, and H. B. Prajapati, "Rainfall forecasting using neural network: A survey," in Conference Proceeding - 2015 International Conference on Advances in Computer Engineering and Applications, ICACEA 2015, 2015.
[71] A. A. Ismail, T. Wood, and H. C. Bravo, "Improving Long-Horizon Forecasts with Expectation-Biased LSTM Networks," 2018.
[72] P. Liu, J. Wang, A. K. Sangaiah, Y. Xie, and X. Yin, "Analysis and prediction of water quality using LSTM deep neural networks in IoT environment," Sustain., vol. 11, no. 7, pp. 1–14, 2019.
[73] A. M. Araba, Z. A. Memon, M. Alhawat, M. Ali, and A. Milad, "Estimation at Completion in Civil Engineering Projects: Review of Regression and Soft Computing Models," Knowledge-Based Eng. Sci., vol. 2, no. 2, pp. 1–12, 2021.
[74] J. . Hair, W. . Black, B. . Babin, and R. Anderson, Multivariate Data Analysis, 7th ed. Upper Saddle River, NJ, USA.: Prentice-Hall, Inc., 2010.
[75] D. A. Dickey, W. A. Fuller, D. A. Dickey, and W. A. Fuller, "Journal of the American Statistical Association Distribution of the Estimators for Autoregressive Time Series with a Unit Root Distribution of the Estimators for Autoregressive Time Series With a Unit Root," Taylor, Publ., no. July 2015, pp. 37–41, 2012.
[76] S. I. Abba et al., "Evolutionary computational intelligence algorithm coupled with self-tuning predictive model for water quality index determination," J. Hydrol., vol. 587, p. 124974, Aug. 2020.
[77] F. Shahrin, L. Zahin, R. Rahman, A. S. M. J. Hossain, A. H. Kaf, and A. K. M. Abdul Malek Azad, "Agricultural analysis and crop yield prediction of habiganj using multispectral bands of satellite imagery with machine learning," in 11th International Conference on Electrical and Computer Engineering, ICECE 2020, 2020, pp. 21–24.
[78] U. Ashwini, K. Kalaivani, K. Ulagapriya, and A. Saritha, "Time Series Analysis based Tamilnadu Monsoon Rainfall Prediction using Seasonal ARIMA," in 6th International Conference on Inventive Computation Technologies, ICICT 2021, 2021, pp. 1293–1297.
[79] R. A. Abdulkadir, S. I. A. Ali, S. I. Abba, and P. Esmaili, "Forecasting of daily rainfall at Ercan Airport Northern Cyprus: a comparison of linear and non-linear models," Desalin. Water Treat., vol. 177, no. May 2019, pp. 297–305, 2020.
[80] S. Shamshirband et al., "Comparative analysis of hybrid models of firefly optimization algorithm with support vector machines and multilayer perceptron for predicting soil temperature at different depths," Eng. Appl. Comput. Fluid Mech., vol. 14, no. 1, pp. 939–953, 2020.
[81] J. Yu, C. H. Kim, and S. B. Rhee, "The Comparison of Lately Proposed Harris Hawks

Optimization and Jaya Optimization in Solving Directional Overcurrent Relays Coordination Problem," Complexity, vol. 2020, no. iii, 2020.
[82] A. Malik, Y. Tikhamarine, S. S. Sammen, S. I. Abba, and S. Shahid, "Prediction of meteorological drought by using hybrid support vector regression optimized with HHO versus PSO algorithms," Environ. Sci. Pollut. Res., Mar. 2021.
[83] S. I. Abba et al., "Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration," Hydrol. Sci. J., vol. 0, no. 0, 2021.
[84] O. Kisi, H. Sanikhani, and M. Cobaner, "Soil temperature modeling at different depths using neuro-fuzzy, neural network, and genetic programming techniques," Theor. Appl. Climatol., vol. 129, no. 3–4, pp. 833–848, May 2016.
[85] O. Kisi, M. Tombul, and M. Z. Kermani, "Modeling soil temperatures at different depths by using three different neural computing techniques," Theor. Appl. Climatol., vol. 121, no. 1–2, pp. 377–387, 2015.
[86] Himika, S. Kaur, and S. Randhawa, "Global Land Temperature Prediction by Machine Learning Combo Approach," 2018 9th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2018, pp. 1–8, 2018.