A Comprehensive Study of Artificial Intelligence Applications for Soil Temperature Prediction

Hanifeh Imanian [1,*], Juan Hiedra Cobo [2], Pierre Payeur [3], Hamidreza Shirkhani [2] and Abdolmajid Mohammadian [1]

[1] Civil Engineering Department, University of Ottawa, Ottawa, Canada; himania3@uottawa.ca (H.I.); majid.mohammadian@uOttawa.ca (A.M.)

[2] National Research Council Canada, Ottawa, Canada; juan.hiedracobo@nrc-cnrc.gc.ca (J.C.); hamidreza.shirkhani@nrc-cnrc.gc.ca (H.S.)

[3] Computer Science Engineering Department, University of Ottawa, Ottawa, Canada; ppayeur@uottawa.ca

* Correspondence: himania3@uottawa.ca

Abstract: Soil temperature is a fundamental parameter in water resources and engineering. A cost-effective model which can forecast soil temperature accurately is extensively needed. Recently, many studies have applied artificial intelligence (AI) at both surface and underground levels for soil temperature prediction. However, there is no comprehensive and detailed assessment of the performance of different AI approaches in soil temperature estimation, and primarily limited atmospheric variables are used as input data for AI models. In the present study, great varieties of various land and atmospheric variables are applied to evaluate the performance of a wide range of AI methods on soil temperature prediction. Herein, thirteen approaches, from classic regressions to well-established methods of random forest and gradient boosting to advanced AI techniques like multi-layer perceptron and deep learning are taken into account. The results show that AI is a promising approach in climate parameter forecast and deep learning demonstrates the best performance among other models. It has the highest R-squared ranging from 0.957 to 0.980, the lowest NRMSE ranging from 2.237% to 3.287% and the lowest MAE, ranging from 0.510 to 0.743 in predicting soil temperature. The prediction is repeated for different sizes of data, and prediction outcomes confirm the conclusion mentioned above.

Keywords: artificial intelligence; climate forecast; deep learning; ensemble model; multi-layer perceptron; neural network; regression; soil temperature; stacking method

1. Introduction

Soil temperature is a pivotal parameter in geo-environmental and geotechnical engineering. Soil temperature prediction is significant for atmospheric models, numerical hydrological and land surface hydrological processes, as well as land-atmosphere interactions. In addition, in some other fields such as water resources and hydrologic engineering, soil temperature is an important factor.

In addition, the soil temperature is a catalyst for many biological processes. It influences soil moisture content, aeration and availability of plant nutrients, which are necessary for plant growth. Therefore, a precise and cost-effective model which can forecast soil temperature accurately is extensively needed.

There are two common ways of obtaining soil temperature: direct measurement and indirect prediction using numerical models. Since soil temperature is a stochastic parameter like other climatic parameters, researchers use the following approaches to calculate it: statistical models and machine learning methods. Statistical models use historical time series to estimate soil temperature in the future.

The commonly used method for time series forecasting is stochastic modeling such as auto-regressive moving average (ARMA) and the auto-regressive integrated moving average (ARIMA) [1]. The statistical methods assume that changes in the statistical

properties of soil temperature data series in the future would be similar to those in the past. This means that large amounts of data are required for long-term prediction.

Bonakdari et al. (2019) and Zeynoddin et al. (2020) proposed a linear stochastic method to model daily soil temperature with sufficient knowledge of time series structure [1,2]. Recently, the use of artificial intelligence (AI)-based techniques for predicting real-world problems has grown enormously. Many studies have applied AI models at both surface and under-ground levels for soil temperature prediction. George (2001) made use of a multi-layer neural network for the weekly mean soil temperature prediction during 1 year [3].

Monthly soil temperature was modeled using a 3-layer artificial neural network (ANN) constructed by Bilgili (2010) [4]. He used meteorological variables of atmospheric temperature, atmospheric pressure, relative humidity, wind speed, rainfall, global solar radiation and sunshine duration to make predictions at five depths below the ground level and compared them with linear and nonlinear regression results.

Ozturk et al. (2011) developed feed-forward artificial neural network models to estimate monthly mean soil temperature at five depths from 5 to 100 cm under the ground using meteorological data such as solar radiation, monthly sunshine duration and monthly mean air temperature [5].

Zare Abyaneh et al. (2016) used ANNs and co-active neuro-fuzzy inference system (CANFIS) for the estimation of daily soil temperatures at six depths from 5 to 100 cm underground using only mean air temperature data from a 14-year period as input data [6].

Adaptive neuro-fuzzy inference system (ANFIS), multiple linear regression (MLR) and ANN models were developed by Citakoglu (2017) to predict soil temperature data in monthly units at five depths from 5 to 100 cm below the soil surface using monthly air temperatures and monthly precipitation for at least 20 years [7].

Himika et al. (2018) made use of various existing regression and machine learning models to propose an ensemble approach to predict land temperature [8]. The chosen models were decision tree, variable ridge regression and conditional inference tree.

Delbari et al. (2019) evaluated the performance of a support vector regression (SVR)-based model in estimating daily soil temperature at 10, 30 and 100 cm depth at different climate conditions [9]. Climatic data used as inputs for the models were air temperature, solar radiation, relative humidity, dew point, and the atmospheric pressure. They compared the obtained results with classical MLR and found that SVR performed better in estimating soil temperature at deeper layers.

A study by Alizamir et al. (2020) compared four machine learning techniques, extreme learning machine (ELM), artificial neural networks (ANN), classification and regression trees (CART) and group method of data handling (GMDH) in estimating monthly soil temperatures [10]. They used monthly climatic data of air temperature, relative humidity, solar radiation, and windspeed at four different depths of 5 to 100 cm as model inputs. ELM was found to generally perform better than the others in estimating monthly soil temperatures.

Li et al. (2020) presented a novel scheme for forecasting the hourly soil temperature at five different soil depths [11]. They developed an integrated deep bidirectional long short-term memory network (BiLSTM) and fed their model with air temperature, wind speed, solar radiation, relative humidity, vapor pressure and dew point. Six benchmark algorithms were chosen to prove the relative advantages of the proposed method, namely, three deep learning methods, i.e., LSTM, BiLSTM and deep neural network (DNN), and three traditional machine learning methods: random forest (RF), SVR, and linear regression.

The proposed model of Penghui et al. (2020) is a hybridization of adaptive neuro-fuzzy inference system with optimization methods using mutation salp swarm algorithm and grasshopper optimization algorithm (ANFIS-mSG) [12]. The prediction of daily soil temperatures was conducted based on maximum, mean and minimum air temperature. The results were compared with seven models, including classical ANFIS, hybridized

ANFIS model with grasshopper optimization algorithm (GOA), salp swarm algorithm (SSA), grey wolf optimizer (GWO), particle swarm optimization (PSO), genetic algorithm (GA), and Dragonfly Algorithm (DA).

Shamshirband et al. (2020) modeled air temperature, relative humidity, sunshine hours and wind speed using multilayer perceptron (MLP) algorithm and SVM in hybrid form with the firefly optimization algorithm (FFA) to estimate soil temperature at 5, 10 and 20 cm depth [13].

In a study by Seifi et al. (2021), hourly soil temperatures at 5, 10, and 30 cm depth were predicted applying ANFIS, SVM, MLP and radial basis function neural network (RBFNN) with optimization algorithms of SSA, PSO, FFA and sunflower optimization (SFO) [14]. They used air temperature, relative humidity, wind speed and solar radiation as input information and found that wind speed did not have high coherence with soil temperature. Generalized likelihood uncertainty estimation (GLUE) approach was implemented to quantify model uncertainty and concluded that ANFIS-SFO produced the most accurate performance.

Hao et al. (2021) proposed a model called convolutional neural network based on ensemble empirical mode decomposition (EEMD-CNN) to predict soil temperatures at three depths of 5 to 30 cm [15]. They used Statistical properties of the maximum, mean, minimum and variance air temperature as the meteorological input information. The results were compared with four models: persistence forecast (PF), backpropagation neural network (BPNN), LSTM and EEMD-LSTM.

In a similar study, a convolutional 3D deep learning model with ensemble empirical mode decomposition (EEMD) was proposed by Yu et al. (2021) to predict soil temperatures over 1, 3 and 5 days at a depth of 7 cm underground [16].

The literature review shows that there are some gaps in the knowledge of AI application in prediction of soil temperature. First, there is an absence of a comprehensive and detailed assessment of the performance of different artificial intelligence approaches, from linear regression to complicated advanced techniques in soil temperature estimation. Second, in the context of atmospheric variables used as input data for AI models, previous studies have usually used limited atmospheric variables, while in the current investigation, a wide range of variables have been used. Although several researchers have developed codes equipped with some AI models, they have focused on limited meteorological parameters, mostly air temperature. There are many other climate data that affect soil temperature, directly or indirectly. Therefore, the impact of other land and atmospheric variables needs to be further studied.

The main purpose of this study is to evaluate the performance of a wide range of AI approaches on soil temperature prediction using various land and atmospheric variables. In this article, 13 methods, from classic regressions to well-established methods of random forest and gradient boosting to advanced AI techniques like ANFIS, ANN and deep learning are taken into account. Meanwhile, a broad selection of variables from a comprehensive reanalysis of ERA5 datasets have been chosen as input parameters for the developed prediction model to consider different aspects of the problem.

The rest of the paper is organized as follows: Section 2 describes the study area and databases, introduces the applied regression and artificial intelligence approaches. The evaluation metrics are also presented in this section. The subsequent section discusses the results and compares the performance of the methods. The last section presents the conclusions.

2. Materials and Methods

2.1. Study Area and Dataset

The climate data used in the present study are obtained from ERA5. ERA5 is the fifth-generation atmospheric reanalysis of the global climate covering the period from 1950 to present. It provides hourly estimates of a large number of atmospheric, land and oceanic climate variables in a gridded-base format with a regular latitude-longitude grid. The data coverage is global with horizontal resolution of 0.25° x 0.25° and resolves the atmosphere

using 137 levels from the surface up to a height of 80km. ERA5 includes information about uncertainties for all variables at reduced spatial and temporal resolutions. ERA5 combines vast amounts of historical observations into global estimates using advanced modelling and data assimilation systems.

The study area is Ottawa, the capital city of Canada (45.4° N, 75.7° W), located in the southeast of the country, in the province of Ontario. Figure 1 shows the geographical location of the considered site used in this study. Ottawa has a semi-continental climate with four distinct seasons. It has a warm, humid summer and a very cold winter.

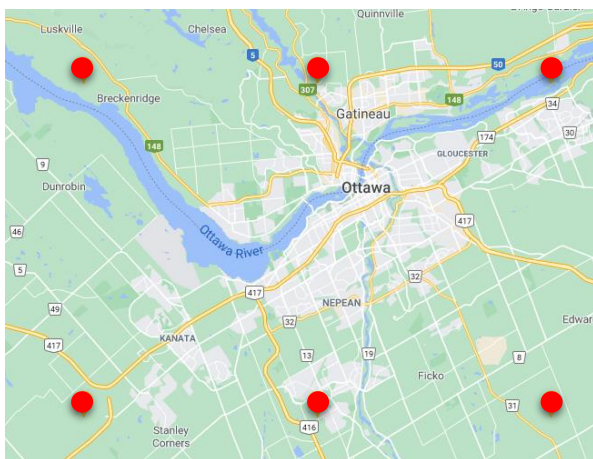
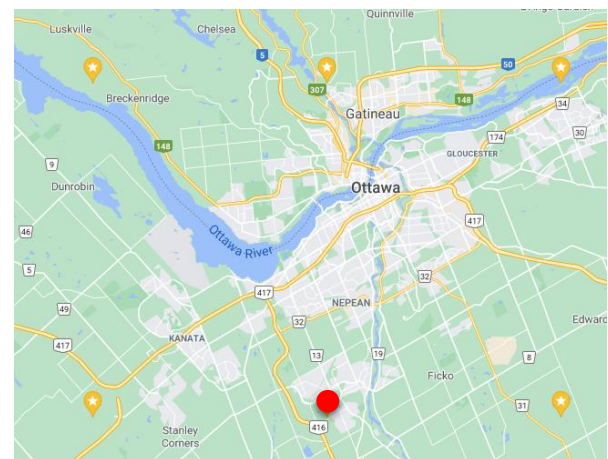
(a)

(b)   **(c)**

Figure 1. Location of the **a)** considered site used in the present study, **b)** six stations representing Ottawa, **c)** one selected station.

Six stations represent Ottawa covering an area of approximately 30 x 40 km. The longitude of these stations varied from 75.50° W to 76.00° W and their latitude varied from 45.25° N to 45.50° N.

The used variables were the hourly weather conditions including air temperature at 2 m above the surface (Kelvin), total precipitation (m), surface pressure (Pa), evaporation

(m of water), instantaneous wind gust at 10 m above the surface (m/s), dewpoint temperature 2 m above the surface (Kelvin), surface net solar radiation (J/m2) and surface net thermal radiation (J/m2). The valid data was from 1st July 2020 to 31st August 2020, a total of 92 days. Then, approximately 106,000 pieces of climatic information was gathered as the AI model's input. The output of each AI model predicted hourly soil temperature in Kelvin at the layer between 0 to 7 cm underground.

2.2. Descriptions of Artificial Intelligence Algorithms

A wide range of AI approaches are applied in the developed numerical model as described below.

2.2.1. Linear Regression, Ridge, Lasso and Elastic Net

Four different linear models were applied in the developed code including: Linear regression, Ridge, Lasso and Elastic Net. Linear regression is the most basic form of a linear model for minimizing the residual sum of squares. So, the objective function is as follows:

$$\sum (y - Xw)^2 \tag{1}$$

Ridge is a linear model that imposes a penalty on the sum of squared value of the weights, resulting in a group of weights that are more evenly distributed. The objective function becomes:

$$\sum (y - Xw)^2 + alpha \sum w^2 \tag{2}$$

where $alpha$ is a non-negative hyperparameter that controls the magnitude of the penalty.

For Lasso, a modification of linear regression is applied, in which the model is penalized for the sum of absolute values of the weights.

$$\frac{1}{2m} \sum (y - Xw)^2 + alpha \sum |w| \tag{3}$$

Elastic Net is a combination of the two last models such that both regularizations related to Ridge and Lasso models are exerted on the linear regression.

$$\frac{1}{2m} \sum (y - Xw)^2 + alpha * ratio * \sum |w| + \frac{alpha}{2} * (1 - ratio) * \sum w^2 \tag{4}$$

2.2.2. Nearest Neighbors Method

In the Nearest Neighbors method, learning is based on the fixed number of nearest neighbors of each query point, or on the neighbors within a fixed radius of the query point. It can be weighted uniformly or proportional to distance.

2.2.3. Decision Tree, Random Forest, Gradient Boosting and Extreme Gradient Boosting

There is another learning method called Decision Trees. This method's goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. In other words, a tree can be seen as a piecewise constant approximation.

Ensemble learning method is a technique that combines predictions from multiple machine learning algorithms in order to make a more accurate prediction than a single model. Some ensemble models are developed based on decision tree like Random Forest and Gradient Boosting.

Random Forest is a meta estimator that fits a number of decision trees on various subsets of the dataset. A Random Forest operates by constructing several decision trees during training time and outputting the mean of the classes as the prediction of all the trees. Several trees run in parallel with no interaction amongst them in this method.

Gradient Boosting fits a decision tree, repeatedly on the differentiable loss functions. This method builds one Decision Tree at a time, where each new tree helps to correct errors made by the previously trained tree.

Extreme Gradient Boosting (XGBoost) builds a model by a set of trees, reduces the errors, and builds a new model in subsequent iterations. Unlike Gradient Boosting method, XGBoost implements some regularization; therefore, it helps to reduce overfitting. Also, it is much faster compared to Gradient Boosting.

2.2.4. Support Vector Machine

A Support Vector Machine (SVM) constructs a decision surface for splitting target classes by making the largest distance to the nearest training data points. SVM uses a subset of training points in the decision function (support vectors) such that different Kernel functions can be specified for each of them. The radial basis function kernel is widely used for SVM models:

$$K(X, \acute{X}) = exp\left(-\gamma \|X - \acute{X}\|^2\right) \tag{5}$$

2.2.5. Stacking Method

Another group of methods is called stacking or Stacked Generalization. Stacking is an ensemble machine learning algorithm that learns how to best combine the predictions from multiple well-performing machine learning models. In this paradigm, the outputs of some aforementioned individual estimators are gathered and an additional regressor is used to compute the final prediction. Stacking often ends up with a model which is better than any individual intermediate model.

In the present study, different combinations of estimators are tried and eventually three methods, Random Forest, SVM and Ridge, computed the best outcome.

2.2.6. Multi-Layer Perceptron

Multi-Layer Perceptron (MLP), a class of feedforward ANN, is a non-linear function approximator in layers using back-propagation with no activation function in the output layer. It used rectified linear unit function as the activation function in the hidden layers:

$$g(z) = max(0, z) \tag{6}$$

MLP uses different loss functions depending on the problem type. For case of prediction, MLP uses the Square Error loss function; written as:

$$\frac{1}{2}\sum(y - Xw)^2 + \frac{alpha}{2}\sum w^2 \tag{7}$$

Starting from initial random weights, MLP minimizes the loss function by repeatedly updating these weights. After computing the loss, a backward pass propagates it from the output layer to the previous layers, providing each weight parameter with an update value meant to decrease the loss.

The algorithm stops when it reaches a pre-set maximum number of iterations, or when the improvement in loss is below a certain, small number.

2.2.7. Deep Learning

Deep Learning is a subset of ANN consisting of large neural networks with significant amounts of data. The word "deep" is referring to the depth of layers in a neural network. A neural network that consists of more than three layers, which would be inclusive of the inputs and the output, can be considered a deep learning algorithm.

2.2.8. Adaptive Neuro-Fuzzy Inference System

Adaptive Neuro-Fuzzy Inference System (ANFIS) models combine fuzzy systems and the learning ability of neural networks. ANFIS is considered an ANN model doing the preprocessing step by converting numeric values into fuzzy values.

2.3. Methodological Overview

The collected data has been split into two parts randomly. The first part, including 65% of the data, is used for the training phase, while the remaining 35% of the data is used as the testing set.

The air temperature, precipitation, surface pressure, evaporation, instantaneous wind speed, dewpoint temperature, solar radiation and thermal radiation are the atmospheric variables used as the inputs of the benchmark algorithm, and the soil temperature at a depth of 0-7 cm underground is the output of the model.

With the aim of making model training less sensitive to the scale of features and allowing our models to converge to better weights and, in turn, lead to a more accurate model, the data has been normalized in a way that removes the mean, and scales each variable to unit variance.

The overall flow of the simulation is illustrated in Figure 2.

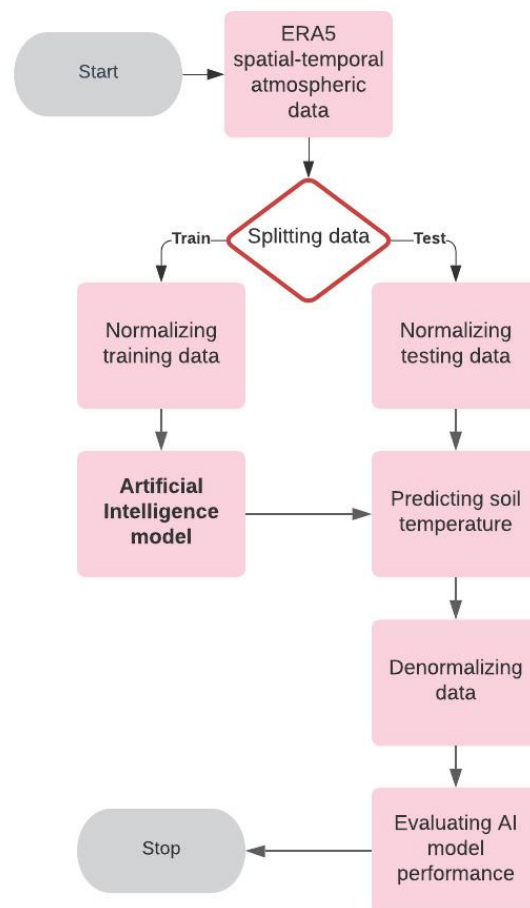



Figure 2. Algorithm of the developed numerical model in the present study.

To facilitate the assessment of the model performance, the outputs of applied AI models need to be compared. Several error indicators are employed to measure the quality of modeling including maximum error (MaxE), mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), normalized root mean square error (NRMSE) and coefficient of determination (R-squared). The evaluation metrics are defined as:

$$MaxE = Max(y_{obs} - y_{calc}) \qquad \text{optimal value: 0} \qquad (8)$$

$$MAE = \frac{\sum|y_{obs} - y_{calc}|}{n} \qquad \text{optimal value: 0} \qquad (9)$$

$$MSE = \frac{\sum(y_{obs} - y_{calc})^2}{n} \qquad \text{optimal value: 0} \qquad (10)$$

$$RMSE = \sqrt{\frac{\sum(y_{obs} - y_{calc})^2}{n}} \qquad \text{optimal value: 0} \qquad (11)$$

$$NRMSE = \frac{RMSE}{[Max(y_{obs}) - Min(v_{obs})]} \qquad \text{optimal value: 0} \qquad (12)$$

$$R^2 = 1 - \frac{\sum(y_{obs} - y_{calc})^2}{\sum(y_{obs} - \overline{y_{calc}})^2} \qquad\qquad \text{optimal value: 1} \qquad (13)$$

where y_{obs} is the observed value, y_{calc} is the predicted value by the AI mode, $\overline{y_{calc}}$ is the mean of calculated values and n is the number of data.

3. Results

Before applying the models to all stations representing Ottawa, one station was se­lected and AI models were applied. The location of the station was at the mid-south of Ottawa with coordinates of 45.25º N and 75.75º W, which is demonstrated in Figure 1-(c). As mentioned earlier, eight hourly climate parameters were used as input variables. The data was gathered from 1st to 31st of July 2020, a total of 31 days that leads to approxi­mately 6,000 pieces of climatic information.

According to what was stated, AI models were applied for two sets of data, first, on a confined database with a limited quantity of information, and second, on a big collection of data. An evaluation of each model's performance was carried out separately, then a comprehensive assessment was finally performed.

The primary step of modeling was splitting the data into two groups. Although the data was split for training and testing purposes randomly, the training and testing data was maintained for all models. So, all AI models were trained and tested with the same set of data.

After the model training procedure, the model was fed with testing data as input, and prediction results were obtained. The predicted outcomes and real values were sim­ultaneously reshaped into a 1-dimensional series, and the performance of the models was evaluated using the error metrics.

The developed model was executed each time, employing one of the 13 abovemen­tioned AI techniques, once on the limited database and once on the big dataset. Hence, 26 sets of predicted data were obtained. Figure 3 shows the time series of hourly soil temper­ature predicted by the various AI models explained in the previous section. Time series are presented for both limited dataset and big dataset.

(a)

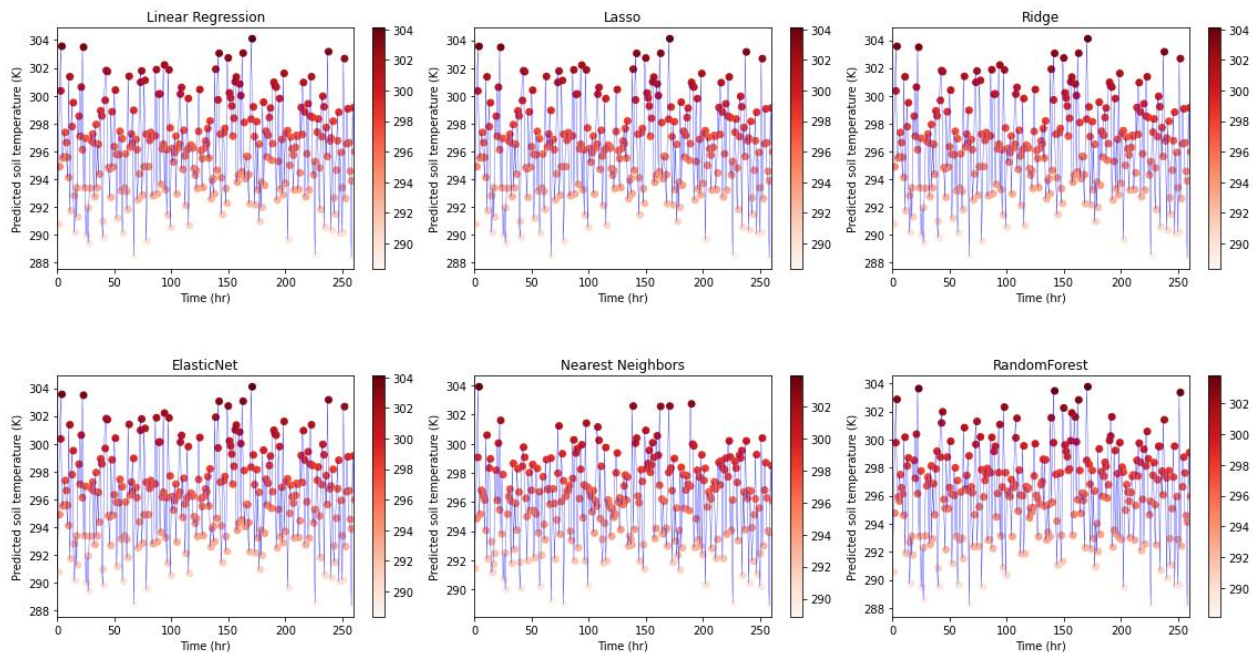
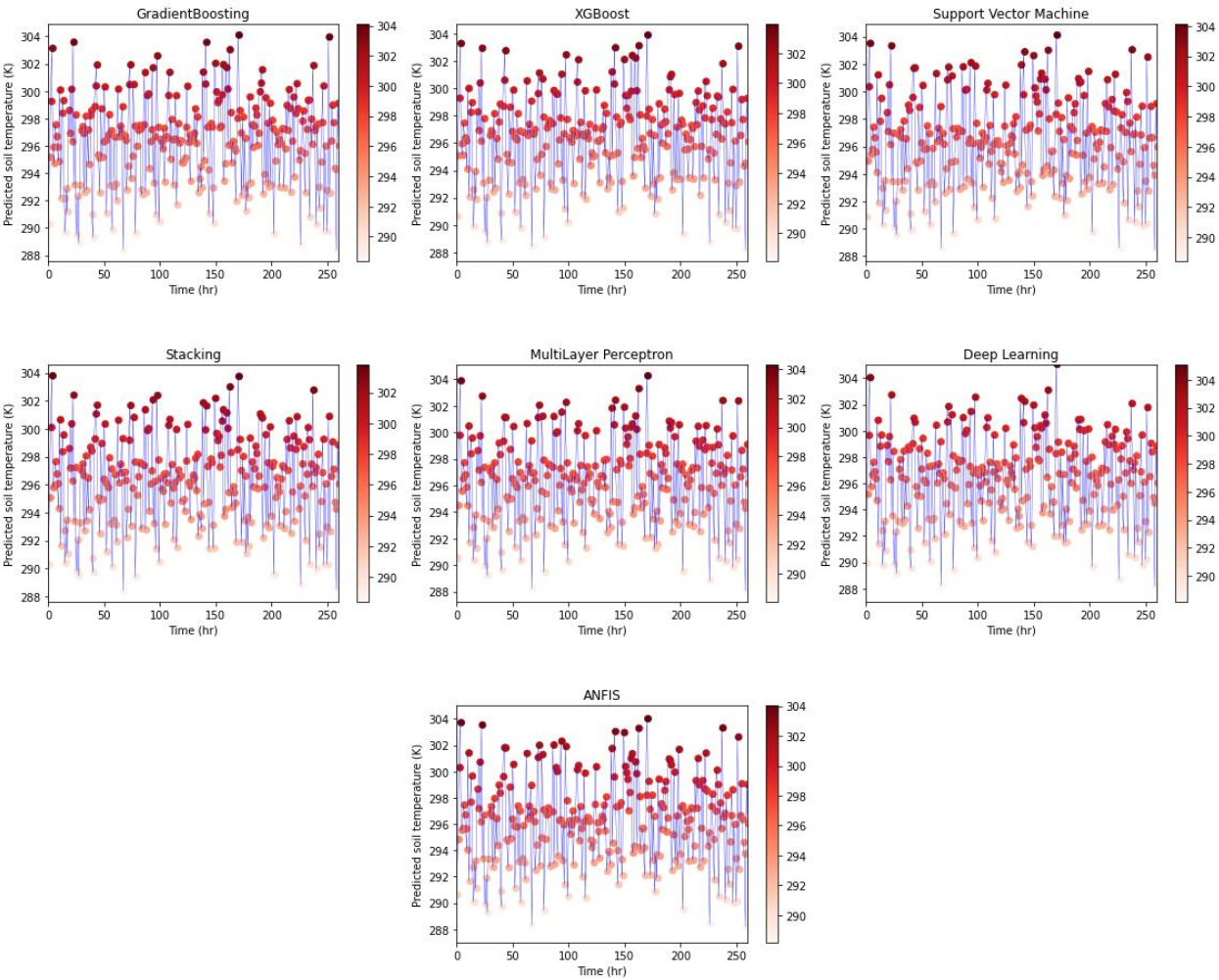



Figure 3. Time series of soil temperature predicted by different AI models for **a)** limited dataset **b)** big dataset.

(b)

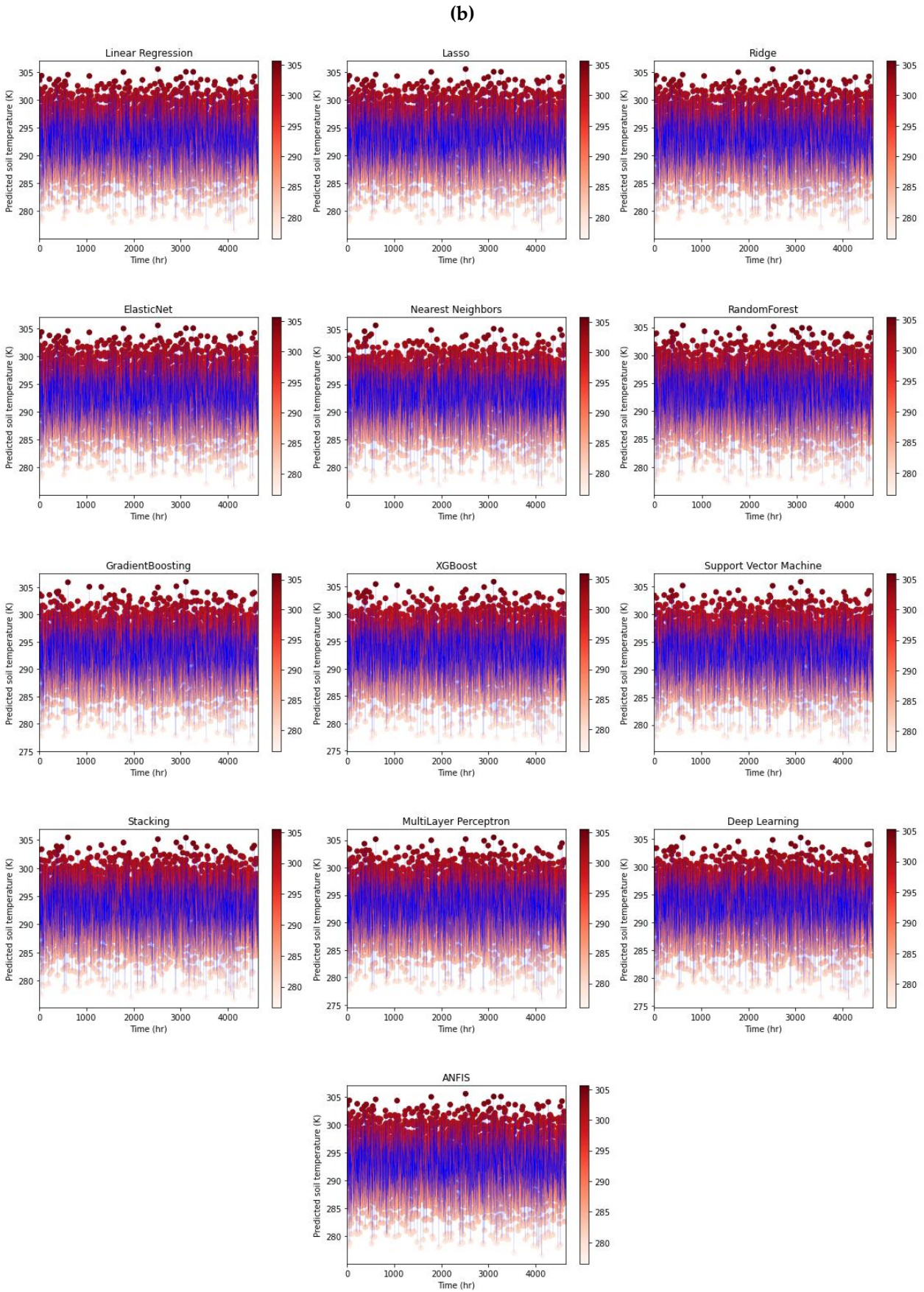



Figure 3. Time series of soil temperature predicted by different AI models for **a)** limited dataset **b)** big dataset (continue).

Figure 4 is a scatter plot of the predicted soil temperatures computed by different AI methods and reanalysis values, which demonstrates a good fit between the observed values and the models' predictions. The predicted soil temperatures show a very close match to the identity line in Figures 4(a) and 4(b). It was determined that all AI models can provide reliable soil temperature results for both limited dataset and big dataset.

Information presented in Figure 4 shows that the size of data in AI models plays a significant role in the correctness of results, and more data will lead to more robust and promising results. It can be seen that the correlation between actual and predicted data was 97% for the big dataset, whereas the R-squared value for the limited dataset was found to be 94.5%.

y=0.936x+18.959

R-squared=0.9452

y=0.968x+9.416

R-squared=0.9702

(a)   (b)

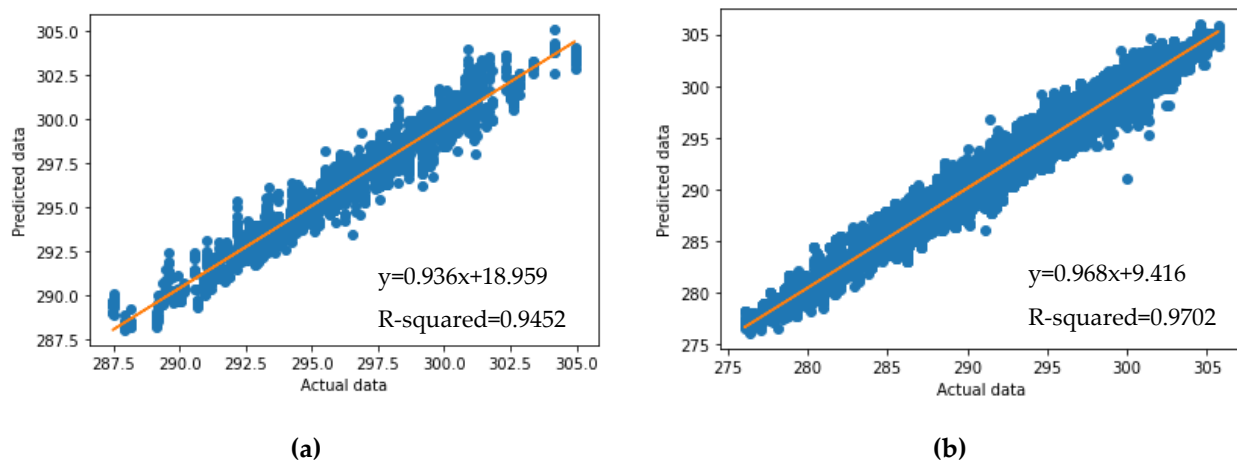



Figure 4. Scatter plots of predicted and observed soil temperature using different AI models **a)** limited dataset **b)** big dataset.

4. Discussion

In the current study, thirteen AI models, namely Linear regression, Lasso, Ridge, Elastic Net, K Neighbors, Random Forest Gradient Boosting, XG Boost, Support Vector Machine, Stacking, Multi Layer Perceptron, Deep Learning and Adaptive Neuro-Fuzzy Inference System, were employed to predict soil temperature. The mentioned models are applied to two sets of data with different quantities of information to assess the performance of the various AI models. Meanwhile, the effect of dataset size on the behaviour of AI models was evaluated.

To measure the quality of different AI models, the statistical indicators of Equations 8 to 13 were applied and the results of error analysis are presented in Tables 1 for both limited and big datasets.

As seen in Table 1, R-squared values are very near 1.00 for both limited and big datasets which shows a strong correlation between the results predicted by different AI models and soil temperature data. The average R-squared for limited and big datasets calculated equals 0.94 and 0.97, respectively. This issue is approved by scatter plots illustrated in Figure 3 and offers an overall acceptable performance for all AI methods. Meanwhile, the last column of Table 1 indicated that all AI models work better while increasing the quantity of information, leading to a more robust match between predicted results and soil temperatures.

An examination of the error values presented in Table 1 shows that the average NRMSE for the limited dataset was 4.5%, while this value equals 2.7% for the big dataset. This proves that using more data improves error measures extensively, and regardless of which AI method is applied, employing more data leads to better results.

Table 1. Error analysis of predicted soil temperature values using different AI models.

Size of database	AI Models	MaxE (K)	MAE (K)	MSE (K²)	RMSE (K)	NRMSE (-)	R² (-)
limited dataset	Linear Regression	2.31	0.63	0.65	0.81	4.6%	0.94
	Lasso	2.31	0.63	0.65	0.81	4.6%	0.94
	Ridge	2.31	0.63	0.65	0.81	4.6%	0.94
	Elastic Net	2.31	0.63	0.65	0.81	4.6%	0.94
	Nearest Neighbors	3.21	0.81	1.09	1.04	6.0%	0.91
	Random Forest	3.23	0.65	0.75	0.86	4.9%	0.94
	Gradient Boosting	3.08	0.58	0.62	0.79	4.5%	0.95
	XG Boost	2.84	0.61	0.66	0.81	4.7%	0.94
	SVM	2.42	0.64	0.65	0.81	4.6%	0.94
	Stacking	2.22	0.51	0.45	0.67	3.9%	0.96
	MLP	2.12	0.5	0.44	0.67	3.8%	0.96
	Deep Learning	2.04	0.46	0.39	0.62	3.6%	0.97
	ANFIS	2.2	0.62	0.65	0.81	4.6%	0.94
Big dataset	Linear Regression	3.89	0.74	0.95	0.97	3.3%	0.96
	Lasso	3.90	0.74	0.95	0.98	3.3%	0.96
	Ridge	3.89	0.74	0.95	0.97	3.3%	0.96
	Elastic Net	3.89	0.74	0.95	0.97	3.3%	0.96
	Nearest Neighbors	6.06	0.60	0.69	0.83	2.8%	0.97
	Random Forest	3.79	0.53	0.50	0.71	2.4%	0.98
	Gradient Boosting	3.53	0.53	0.48	0.69	2.3%	0.98
	XG Boost	3.14	0.52	0.46	0.68	2.3%	0.98
	SVM	4.79	0.53	0.51	0.71	2.4%	0.98
	Stacking	4.28	0.49	0.43	0.66	2.2%	0.98
	MLP	4.32	0.51	0.45	0.67	2.3%	0.98
	Deep Learning	3.44	0.51	0.44	0.66	2.2%	0.98
	ANFIS	8.99	0.74	0.97	0.98	3.3%	0.96

Moreover, it was found from Table 1 that four models: Linear Regression, Lasso, Ridge and Elastic Net showed the same error values for all evaluation metrics in both datasets, demonstrating they had very similar performance. As mentioned previously in the methodology section, the last three methods have linear bases and are refined versions of classic linear regression due to adding regularization terms. Relatively poor MAE, MSE and NRMSE obtained by linear regression results demonstrated that this method can not precisely predict soil temperature. At the same time, the same values of MAE, MSE and NRMSE obtained by Lasso, Ridge and Elastic Net models showed that linear regression modifications still were not appropriate tools to predict soil temperatures.

The Neighbors method had the worst performance among the investigated AI models for the limited database with the greatest MAE, RMSE and NRMSE. Although this method did not show the worst results for the big database, it was one of the weakest methods among all considered AI models with the greatest Maximum error and high RMSE. This finding goes back to the logic behind the Nearest Neighbors method, which does not work in the present study. In this method, a number of nearest neighbors are involved in the model, which does not apply for this prediction since it was more temporal rather than spatial.

A closer look at error values presented in Table 1 shows that three AI methods: stacking, MLP and Deep Learning had better performance than other models. On the limited dataset, the average MAE for these three AI models is less than 0.5 K, while other AI models had MAE of more than 0.65 K. The situation is the same for the big database. The average MAE for these three AI models is approximately 0.5 K, while other AI models had MAE values of approximately 0.65 K. The other error indicator, RMSE, showed a similar trend. On both datasets, the average RMSE for these three AI models was approximately 0.65 K, while other AI models had a RMSE of approximately 0.85 K.

Although deep learning was the best model, the stacking method, which is an ensemble of some not very advanced models, showed good performance and predicted the soil temperature with acceptable precision. This performance was better on the big dataset.

It is worth mentioning that the computation cost should be noted as an essential parameter in picking the best method. The execution time for the limited dataset was negligible, but it was significant for the big pieces of information. The average computation time for deep learning model was 17.5 s, while these values were 10.5 s and 20.5 s for MLP and stacking models, respectively. The stacking model suffers from insufficient execution speed despite showing adequate error metrics.

In statistical analysis, two concepts of confidence region and prediction bands are often used. The confidence band represents the uncertainty in an estimate of the regression on the data. The prediction band is the region that contains approximately 95% of the points. If another pair of actual value -calculated value is taken, there is a 95% chance it falls within the prediction band.

The 95% confidence region and 95% prediction band for these three models: stacking, MLP and Deep Learning, on both limited and big datasets are depicted in Figure 5, which supports the previously mentioned claim regarding the strength of these AI models.

(a)   (b)

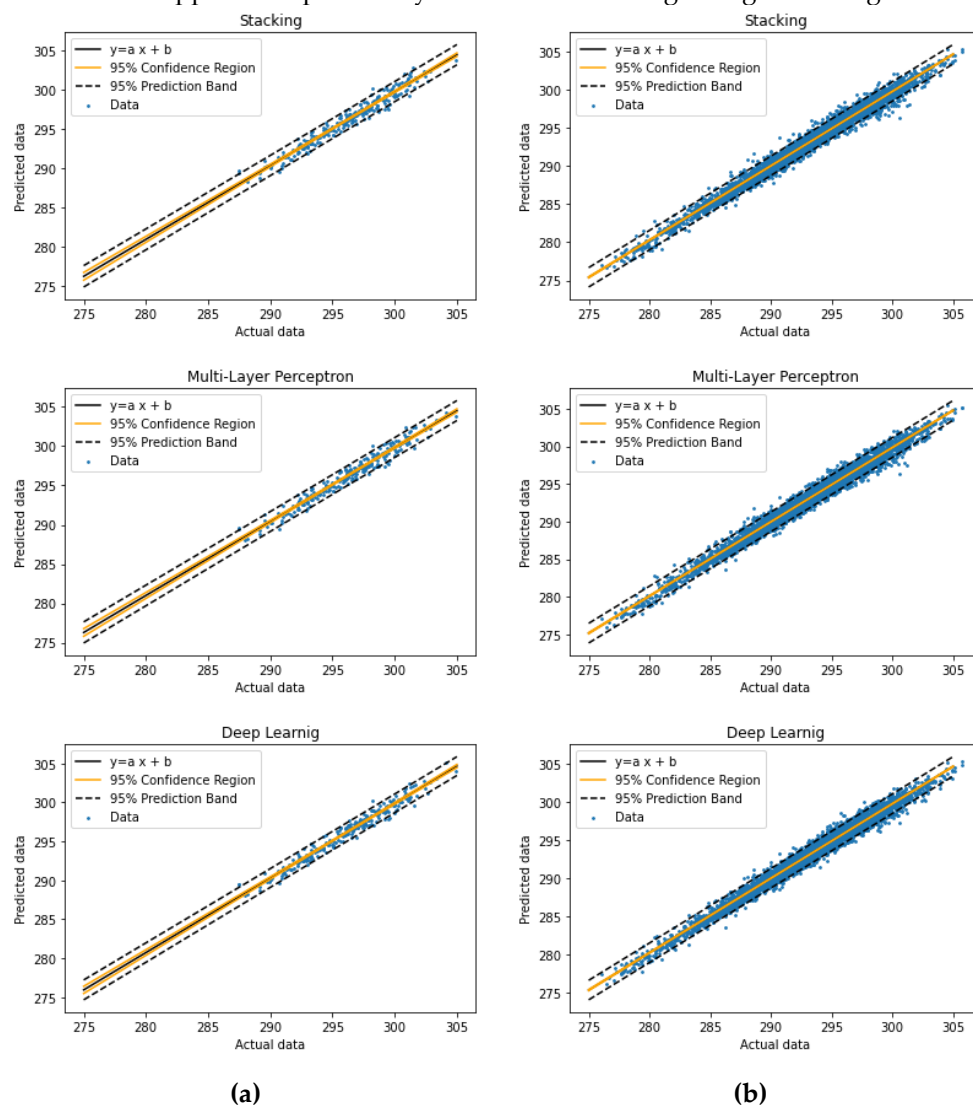



Figure 5. Confidence regions and prediction bands of different AI models **a)** limited dataset **b)** big dataset.

A sensitivity analysis has been done to find the importance level of atmospheric parameters used as AI models' input. In this regard, the three selected AI models of Stacking,

MLP and deep learning, which had the better performance, are executed several times. For each run, one of the input parameters is omitted. Then, the code was implemented with remained 7 variables out of 8 preliminary inputs of air temperature, precipitation, surface pressure, evaporation, wind, dewpoint temperature, solar radiation and thermal radiation. The error indicators are presented in Table 2.

Comparing the obtained errors with the original error values related to all input parameters presented in Table 1 shows that the most critical variable in soil temperature prediction is air temperature, followed by solar radiation. Also, it can be concluded from the sensitivity analysis that precipitation has a negligible effect on results. So, the precipitation does not play an important role in soil temperature forecast and can be omitted from the prediction models without decreasing precision. The mentioned importance level is the same for all three AI models and in both datasets.

5. Conclusions

A precise and cost-effective model for soil temperature forecasting, which enjoys the benefits of artificial intelligence techniques, is developed in the present research. Therefore, 13 AI models including Linear regression, Ridge, Lasso, Elastic Net, Nearest Neighbors, Random Forest, Gradient Boosting, XG Boosting, stacking method, SVM, MLP, Deep Learning and ANFIS are employed to generate a comprehensive and detailed assessment of the performance of different AI approaches in soil temperature estimation. In this regard, 8 hourly land and atmospheric variables of air temperature, precipitation, surface pressure, evaporation, wind gust, dewpoint temperature, solar radiation and thermal radiation were employed, and predictions were made using two limited and big datasets.

The key findings of this study are summarized as follows:

- AI is a promising approach in climate parameter forecast and developed AI models showed a reliable ability in soil temperature prediction.
- Applying AI models to more information led to better results even using the same method.
- Among all 13 AI models applied in the current study, deep learning showed the best performance in predicting soil temperature.
- Although deep learning was the best model, stacking method showed good performance with an acceptable precision in soil temperature prediction.
- Sensitivity analysis shows air temperature and solar radiation plays the most important roles in soil temperature prediction, while precipitation can be neglected in forecast AI models.

Table 2. Sensitivity analysis of input variables for AI models.

AI model	omitted variable	MaxE	MAE	MSE	RMSE	NRMSE	R²	Importance
Stacking, Limited data	precipitation	2.22	0.50	0.44	0.66	3.8%	0.96	8
	pressure	2.39	0.56	0.51	0.72	4.1%	0.96	
	evaporation	2.29	0.56	0.56	0.75	4.3%	0.95	
	wind	2.16	0.53	0.48	0.69	3.9%	0.96	
	dewpoint	2.22	0.52	0.48	0.69	4.0%	0.96	
	solar radiation	2.43	0.60	0.61	0.78	4.5%	0.95	2
	thermal radiation	2.12	0.52	0.46	0.68	3.9%	0.96	
	air temperature	7.31	1.39	3.24	1.80	10.3%	0.72	1

AI model	omitted variable	MaxE	MAE	MSE	RMSE	NRMSE	R²	Importance
MLP, Limited data	precipitation	1.98	0.49	0.42	0.65	3.7%	0.96	8
	pressure	2.17	0.53	0.46	0.68	3.9%	0.96	
	evaporation	2.32	0.54	0.52	0.72	4.1%	0.95	
	wind	2.12	0.53	0.47	0.69	3.9%	0.96	
	dewpoint	1.93	0.50	0.45	0.67	3.8%	0.96	
	solar radiation	2.21	0.58	0.59	0.77	4.4%	0.95	2
	thermal radiation	2.19	0.52	0.46	0.68	3.9%	0.96	
	air temperature	5.92	1.20	2.50	1.58	9.0%	0.79	1
Deep L, Limited data	precipitation	2.09	0.50	0.43	0.66	3.8%	0.96	
	pressure	2.18	0.52	0.45	0.67	3.9%	0.96	
	evaporation	2.16	0.52	0.49	0.70	4.0%	0.96	
	wind	2.01	0.53	0.50	0.70	4.0%	0.96	
	dewpoint	2.27	0.49	0.45	0.67	3.9%	0.96	
	solar radiation	2.51	0.59	0.61	0.78	4.5%	0.95	2
	thermal radiation	2.29	0.49	0.42	0.65	3.7%	0.96	8
	air temperature	5.69	1.33	2.93	1.71	9.8%	0.75	1
Stacking, Big data	precipitation	4.16	0.49	0.43	0.66	2.2%	0.98	8
	pressure	4.21	0.52	0.47	0.68	2.3%	0.98	
	evaporation	3.76	0.54	0.54	0.73	2.5%	0.98	
	wind	3.80	0.54	0.53	0.73	2.4%	0.98	
	dewpoint	3.99	0.53	0.51	0.71	2.4%	0.98	
	solar radiation	4.23	0.59	0.64	0.80	2.7%	0.97	2
	thermal radiation	4.26	0.52	0.49	0.70	2.4%	0.98	
	air temperature	10.24	1.19	2.78	1.67	5.6%	0.88	1
MLP, Big data	precipitation	3.82	0.53	0.49	0.70	2.3%	0.98	8
	pressure	4.27	0.53	0.49	0.70	2.4%	0.98	
	evaporation	4.10	0.60	0.64	0.80	2.7%	0.97	
	wind	3.88	0.58	0.59	0.77	2.6%	0.97	
	dewpoint	3.82	0.59	0.62	0.79	2.6%	0.97	
	solar radiation	4.34	0.65	0.74	0.86	2.9%	0.97	2
	thermal radiation	4.13	0.56	0.55	0.74	2.5%	0.98	
	air temperature	10.50	1.32	3.19	1.78	6.0%	0.86	1
Deep L, Big data	precipitation	3.94	0.52	0.46	0.68	2.3%	0.98	8
	pressure	4.33	0.52	0.46	0.68	2.3%	0.98	
	evaporation	3.74	0.57	0.58	0.76	2.6%	0.97	
	wind	3.71	0.58	0.58	0.76	2.6%	0.97	
	dewpoint	3.58	0.57	0.54	0.73	2.5%	0.98	
	solar radiation	4.43	0.61	0.65	0.81	2.7%	0.97	2
	thermal radiation	3.84	0.52	0.48	0.69	2.3%	0.98	
	air temperature	10.03	1.28	2.96	1.72	5.8%	0.87	1

Author Contributions: Conceptualization, J.C., A.M., H.S. and P.P.; methodology, H.I.; software, H.I.; validation, H.I. and P.P.; formal analysis, H.I. and P.P.; investigation, H.I.; resources, H.I.; data curation, H.I.; writing—original draft preparation, H.I. and A.M.; writing—review and editing J.C., P.P., H.S. and A.M.; visualization, H.I.; supervision, A.M. and J.C.; project administration, A.M., J.C.

and H.S.; funding acquisition, A.M., H.S. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Research Council Canada through the Artificial Intelligence for Logistics Supercluster Support Program, grant number AI4L-120.

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

References

1. Bonakdari H, Moeeni H, Ebtehaj I, Zeynoddin M, Mahoammadian A, Gharabaghi B. New insights into soil temperature time series modeling: linear or nonlinear? *Theoretical and Applied Climatology*. **2019**;135(3-4):1157-1177. doi:10.1007/s00704-018-2436-2
2. Zeynoddin M, Ebtehaj I, Bonakdari H. Development of a linear based stochastic model for daily soil temperature prediction: One step forward to sustainable agriculture. *Computers and Electronics in Agriculture*. **2020**;176. doi:10.1016/j.compag.2020.105636
3. George RK. Prediction of Soil Temperature by using Artificial Neural Network Algorithms. *Nonlinear Analysis*. **2001**;47:1737-1748.
4. Bilgili M. Prediction of soil temperature using regression and artificial neural network models. *Meteorology and Atmospheric Physics*. **2010**;110(1):59-70. doi:10.1007/s00703-010-0104-x
5. Ozturk M, Salman O, Koc M. Artificial neural network model for estimating the soil temperature. *Canadian Journal of Soil Science*. **2011**;91(4):551-562. doi:10.4141/cjss10073
6. Zare Abyaneh H, Bayat Varkeshi M, Golmohammadi G, Mohammadi K. Soil temperature estimation using an artificial neural network and co-active neuro-fuzzy inference system in two different climates. *Arabian Journal of Geosciences*. **2016**;9(5). doi:10.1007/s12517-016-2388-8
7. Citakoglu H. Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey. *Theoretical and Applied Climatology*. **2017**;130(1-2):545-556. doi:10.1007/s00704-016-1914-7
8. Himika, Kaur S, Randhawa S. Global Land Temperature Prediction by Machine Learning Combo Approach. In: 9th International Conference on Computing, Communication and Networking Technologies, Bengaluru, July 10-12, IEEE, **2018**:1-8.
9. Delbari M, Sharifazari S, Mohammadi E. Modeling daily soil temperature over diverse climate conditions in Iran—a comparison of multiple linear regression and support vector regression techniques. *Theoretical and Applied Climatology*. **2019**;135(3-4):991-1001. doi:10.1007/s00704-018-2370-3
10. Alizamir M, Kisi O, Ahmed AN, Mert C, Fai CM, Kim S, Kim NW, El-Shafie A. Advanced machine learning model for better prediction accuracy of soil temperature at different depths. *PLoS ONE*. **2020**;15(4). doi:10.1371/journal.pone.0231055
11. Li C, Zhang Y, Ren X. Modeling hourly soil temperature using deep BiLSTM neural network. *Algorithms*. **2020**;13(7). doi:10.3390/A13070173
12. Penghui L, Ewees AA, Beyaztas BH, Qi C, Salih SQ, Al-Ansari N, Bhagat SK, Yaseen ZM, Singh VP. Metaheuristic optimization algorithms hybridized with artificial intelligence model for soil temperature prediction: Novel model. *IEEE Access*. **2020**;8:51884-51904. doi:10.1109/ACCESS.2020.2979822
13. Shamshirband S, Esmaeilbeiki F, Zarehaghi D, Neyshabouri M, Samadianfard S, Ghorbani MA, Mosavi A, Nabipour N, Chau KW. 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*. **2020**;14(1):939-953. doi:10.1080/19942060.2020.1788644
14. Seifi A, Ehteram M, Nayebloei F, Soroush F, Gharabaghi B, Torabi Haghighi A. GLUE uncertainty analysis of hybrid models for predicting hourly soil temperature and application wavelet coherence analysis for correlation with meteorological variables. *Soft Computing*. **2021**;25(16):10723-10748. doi:10.1007/s00500-021-06009-4
15. Hao H, Yu F, Li Q. Soil Temperature Prediction Using Convolutional Neural Network Based on Ensemble Empirical Mode Decomposition. *IEEE Access*. **2021**;9:4084-4096. doi:10.1109/ACCESS.2020.3048028
16. Yu F, Hao H, Li Q. An Ensemble 3D Convolutional Neural Network for Spatiotemporal Soil Temperature Forecasting. *Sustainability*. **2021**;13(16):9174. doi:10.3390/su13169174