ORIGINAL PAPER

Probabilistic analysis of thermal conductivity of soil

Rashid Mustafa[1] · Kajol Kumari[1] · Sabnam Kumari[1] · Guddu Kumar[1] · Pritam Singh[1]

Abstract

Thermal conductivity is the unique thermal characteristic of soil that regulates the flow of heat energy. A significant impact on geothermal applications is caused by the heat conductivity of the soil. Generally thermal conductivity of soil depends on quartz content, degree of saturation, porosity, dry density, weather condition, and some topographical factors. In this study, four major factors are considered on which thermal conductivity of soil depends, viz., quartz content (Q_C), degree of saturation (S), porosity (η), and dry density (γ) of soil. In this study, three machine learning models, namely, adaptive neuro fuzzy inference system (ANFIS), extreme learning machine (ELM), and extreme gradient boosting (XGBoost) are used to predict thermal conductivity of soil more accurately and errorless. A total of 110 datasets have been used, where 70% (77 cases) of the dataset are used in the training phase and the rest 30% (33 cases) are used in the testing phase. Models' performances are judged using various performance parameters like R^2, a-20 index, VAF, WI, NS RMSE, MAE, SI, RSR, and WMAPE. Proposed models are also judged with the help of regression curve, error matrix, rank analysis, radar diagram, and William's plot. Reliability index (β) and failure probability (P_f) are computed with the help of FOSM (first-order second moment) approach. The overall performance of ANFIS model is superior as compared to the other models, and ELM performs worst. To know the influence of each input parameters on the output, sensitivity analysis is performed.

Keywords Reliability analysis · ANFIS · ELM · XGBoost · William's plot · Sensitivity analysis

Abbreviations

ML	Machine learning
Q_c	Quartz content
η	Porosity
ELM	Extreme learning machine
R^2	Coefficient of determination
WI	Willmott's index of agreement
RMSE	Root mean square error
SI	Scatter index
FOSM	First-order second moment method
P_f	Probability of failure
TMP	Trend measuring parameters
TR	Training
μ	Average value
E_e	Error for EMP
SOR	Strength of relation
ANFIS	Adaptive neuro fuzzy inference system
S	Degree of saturation
γ	Dry density
XGBoost	Extreme gradient boosting
VAF	Variance account factor
NS	Nash Sutcliffe efficiency
MAE	Mean absolute error
RSR	RMSE-observation standard deviation ratio
β	Reliability index
TC	Thermal conductivity
EMP	Error measuring parameters
TS	Testing
σ	Standard deviation
E_t	Error for TMP
I	Ideal value of EMP and TMP

Introduction

From the engineering point of view, soil is a very important and universal ancient construction and base material. Soil is very popular due to its versatile properties, such as moisture content, quartz content, porosity, mineral content, and

organic content. It is very popular in construction industries and geotechnical field. Furthermore, with the help of some soil properties, we can compute thermal conductivity of soil by various methods like conventional method that comes from ancient and another by using machine learning which uses a variety of machine learning algorithms. Thermal conductivity is world-widely very important from many points of view such as large and heavy construction, agriculture field, geothermal resources development, land surface processing, and underground construction. Thermal conductivity is a characteristic of soil that manages the thermal properties on various types of soil depending upon various types of parameters related to that soil. Among these parameters, some parameter like quartz content, water content and cross-section, and thermal conductivity is highly reliable. Thermal conductivity is basically saying about thermal energy transformation via soil to either atmosphere or building.

Many researchers have used experimental and laboratory methods to determine heat conductivity, including Sepaskhah and Boersma (1979) who have used a laboratory method to find the soil's thermal conductivity with heat probe method, and they have considered that soil's thermal conductivity depends upon water content and temperature. The measurement of thermal conductivity was done for loam, loamy sand, and soils with silty clay loam with the help of a cylinder-shaped heat sensor method at different water contents and temperatures. Abu-Hamdeh and Reeder (2000) have done a laboratory experiment to estimate the thermal conductivity of soil and have taken moisture content, bulk density salinity level, and organic substance as input parameters. The results showed that for the same salt type, thermal conductivity measurements for sand were greater than those for clay loam. Abu-Hamdeh et al. (2001) have done a comparison of soil heating and cooling techniques to find the soil's thermal conductivity for sand, loam, sandy loam, and clay. They have observed that based the thermal conductivity based on heating techniques is slightly more than cooling method. Abu-Hamdeh (2003) analyzed the thermal characteristics of soils that are influenced by water content and density. The result showed that the sandy soil had lesser volumetric heat capacity and specific heat in comparison to clay soil at equal soil density and moisture contents. Cosenza et al. (2003) have established the relationship between volumetric water content and soil's thermal conductivity. They have demonstrated through the use of a mathematical modelling approach that the relationship between volumetric water content and thermal conductivity of soil is affected by the microscopic structure of water. Nusier and Abu-Hamdeh (2003)Hamdeh (2003 used laboratory methods to find the thermal conductivity of soil such as sand and loam, and finally it has been observed that sand has a higher heat conductivity than loam. Malek et al. (2021) have done a laboratory work to find the

impact of different soil properties on thermal conductivity of soil. The output showed that the soil's water content has the greatest impact on its thermal conductivity. Dong et al. (2015) have reviewed thermal conductivity models for unsaturated soils for critical condition. The models that they have used are classified into three groups: empirical models, mixing models, and mathematical models. The typical thermal conductivity (k) can be arranged in decreasing order as $k_{mineral} > k_{saturated\ soil} > k_{water} > k_{dry\ soil} > k_{air}$. Tokoro et al. (2016) analyzed how soil thermal conductivity is affected by water and propose a new soil thermal conductivity model and empirical equation. Farouki (1981) used various methods to find the thermal conductivity of soil in cold region. They observed that Johansen's method gives better result over others. Hu et al. (2016) used the conduction–convection equation to estimate a Fourier series–based analytical approach in order to determine the soil's temperature and thermal characteristics. The conclusion obtained that the Fourier series model shows better observed field temperatures on comparing with the sine wave model. Tong et al. (2016) presented a simple empirical model for calculating soil thermal conductivity. In comparison to other complicated models, this one is more useful for massive computations of surface energy and water fluxes. Rubin and Ho (2021) give a semi-analytical way to evaluate the soil's thermal conductivity. It is one of the few models that currently exists that captures the physical geometry of different soil structure stages. Adrinek et al. (2022) gave the idea regarding the analytical computation of thermal conductivity based on sediment's physical characteristics.

Some of the researchers have also used machine learning model in the geotechnical field. TaeBang et al. (2020) used machine learning approach to anticipate a model to compute thermal conductivity for compacted bentonite soil. Rizvi et al. (2020) used soft computing approach such as thermal lattice element method (TLEM) and deep neural network (DNN) in which TELM shows best performance. Kardani et al. (2021) proposed new computational approaches along with hybridization of two metaheuristic optimization techniques to predict the thermal conductivity of unsaturated soil. The result shows that ELM-IFF gives the highest possible prediction accuracy. Da et al. (2022) used several machine learning algorithms that are broadly applicable to find thermal diffusivity and thermal conductivity of bentonite. Random forest, gradient boosting algorithms, and extra trees regressor have advantage in creation of thermal conductivity model. Li et al. (2022a, b) implemented six machine learning algorithms to predict soil's thermal conductivity. Outcome shows that adaptive boosting methods (AdaBoost) gives better predicted values. Li et al. (2022a, b) proposed a unified model based on artificial neural network algorithms, and the accuracy of the model is judged by evaluating it with current prediction models to compute

thermal conductivity of soil. Mustafa et al. (2022) created four hybrid ANFIS models to find the factor of safety against overturning. They observed that ANFIS-PSO shows superior prediction across all four models. Wang et al. (2022) analyzed that based on artificial intelligence (AI), support vector machine (SVM) is the best model to predict soil's thermal conductivity. Liu et al. (2023) made an experimental judgment between machine learning (ML) models and traditional parametric models. He observed that extreme gradient boosting (XGBoost) has better recreation accuracy and simplification ability. Mustafa et al. (2023) examined the gravity retaining wall's resistance to bearing failure with the help of machine learning and concluded that deep neural network (DNN) performed the best. In this study, we have used three machine learning (ML) models such as extreme learning machine (ELM), adaptive neuro fuzzy inference system (ANFIS), and extreme gradient boosting (XGBoost) to predict thermal conductivity of soil.

Details of present studies

Soil's thermal conductivity is an essential property which establishes how easily heat can flow through the soil. It is a key factor in understanding how energy is exchanged between the ground and the atmosphere. The soil's thermal conductivity plays a vital role in geothermal energy systems, where heat is extracted from or injected into the ground. Soils with higher mineral content tends to have higher thermal conductivity because minerals are better heat conductors than organic matter or water. In this study, we have used an empirical equation of thermal conductivity of soil (K_s) which was given by Johansen (1975) and can be expressed as follows:

$$TC_s = 7.7^{Q_c} . 2^{(1-Q_c)} \tag{1}$$

where Q_c is the quartz content and TC_S is soil's thermal conductivity.

Using Eqs. (2) and (3), we can compute the saturated thermal conductivity (TC_{sat}) and dry thermal conductivity (TC_{dry}). By using the values of TC_{sat}, TC_{dry}, and K_N, we can compute unsaturated thermal conductivity of soil (TC_{unsat}).

$$TC_{sat} = 0.57^\eta . TC_s^{(1-\eta)} \tag{2}$$

$$TC_{dry} = \frac{137\gamma + 64.7}{2700 - 947\gamma} \tag{3}$$

$$K_N = \left\{ 0.7 log\left(\frac{S}{100}\right) + 1.0 \right\} \tag{4}$$

where η is the porosity, γ is dry unit weight, S is the degree of saturation, and K_N is Kersten number. The unsaturated thermal conductivity of soil can be computed as:

$$TC_{unsat} = \left(TC_{sat} - TC_{dry}\right)K_N + TC_{dry} \tag{5}$$

In this study, we have considered four input parameters, namely, porosity (η), degree of saturation (S), dry density (γ), and Quartz content (Q_c) and one output that is soil's thermal conductivity in watts per meter Kelvin (W m^{-1} K^{-1}).

Proposed AI-based model

Adaptive neuro fuzzy inference system (ANFIS)

The ANFIS is based on Takagi–sugeno fuzzy inference system. The Takagi–sugeno fuzzy inference system serves as the foundation for ANFIS, which is the addition of artificial neural network (ANNs), and fuzzy logic includes membership function and fuzzy logic operators. The accuracy of any ANFIS model depends upon the suitability of the data provided.

The goal of ANFIS model is in order to reduce the error between the expected and actual output from the training datasets and provide integrated learning capacity using its five-layer system. Furthermore, a new study recently did ANFIS-PSO (adaptive neuro fuzzy inference system-particle swarm optimization) for accurate prediction by Mustafa et al. (2022).

The following fuzzy inference relies on a fuzzy rule that was established for two inputs (A and B) and one output (M).

Rule1: If A is X_1 and B is Y_1, then,

$$d_1 = e_1 A + f_1 B + K_1 \tag{6}$$

Rule 2: If A is X_2 and B is Y_2, then,

$$d_2 = e_2 A + f_2 B + K_2 \tag{7}$$

where A and B are the two inputs and X_1, X_2, Y_1, and Y_2 are membership function (MF) linked with inputs A and B connected to node function. The parameters e_1, f_1, K_1, and e_2, f_2, and K_2 are linked with output functions d_1 and d_2. A usual ANFIS illustration is shown in Fig. 1. It demonstrates that the ANFIS model has five layers. Every layers' functions is listed below:

Layer 1: The input layers have been turned into a fuzzy membership function (MF) through this layer. Premise parameters are referred to as layer 1 parameters. The node's output may be calculated as:

$$Q_j^1 = \mu_{Xj}(A) \text{ for } j = 1, 2, \quad Q_j^1 = \mu_{Yj-2}(B) \text{ for } j = 3, 4 \tag{8}$$

where both A and B represent the node j's inputs and Xj and Yj are fuzzy sets associated with node j in parameterized form.

Fig. 1 Adaptive neuro fuzzy inference system
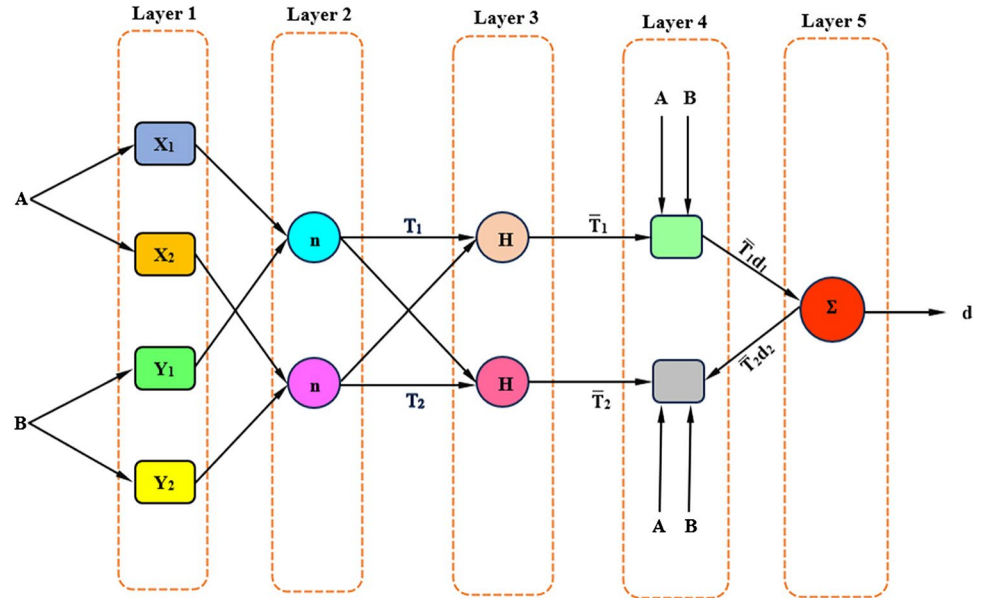
Layer 2: The nodes in this level are fixed, serve as multipliers, and are referred to as neural network layers.

$$Q_j^2 = p_j = \mu_{Xj}(A).\mu_{Yj}(B) \text{ for } j = 1, 2 \tag{9}$$

Layer 3: Each node in layer 3 has been fixed and classified by H. Fuzzy inference method firing intensities are standardized. Each node that may be accessed in this level computes its weight which is standardized. The term "outputs" refers to normalized firing intensities.

$$Q_j^3 = \overline{p_j} = \frac{p_j}{p_1 + p_2} \text{ for } j = 1, 2 \tag{10}$$

Layer 4: The nodes' parameters in this layer can be adjustable. A parameter collection is present in this fuzzy logic node $\{e_j, f_j, k_j\}$ and the output of previous layer $\overline{p_j}$. The node's output is indicated as follows:

$$Q_j^4 = \overline{p_j} d_j = \overline{p_j}(e_j A + f_j + k_j) \text{ for } j = 1, 2 \tag{11}$$

Layer 5: One particular node is fixed in this layer, and the output has been calculated as the sum of all entering signals. This node's output function is calculated as follows:

$$Q_j^5 = \sum \overline{p_j} d_j = \frac{\sum_j p_j d_j}{\sum_j p_j} Q_j^4 = \overline{p_j}(e_j A + f_j B + K_j) \text{ for } j = 1, 2 \tag{12}$$

Extreme learning machine (ELM)

Extreme learning machine is a type of computational algorithm which comes from feedforward neural networks. In many types of application, i.e., classification, regression, clustering, and feature extraction tec. It was introduced to improve the accuracy of single-hidden-layer feedforward network (SLFNs) (Huang et al. 2011). ELM is a new AI-based model developed in the recent decade (Huang et al. 2006). Application of ELM is found in many types of programming languages, namely, MATLAB, Python, and R. ELM is a feedforward neural network used to resolve regression, clustering, and featuring learning with the help of a single layer of hidden nodes with the help of above-described relationship; ELM tries to understand the complex relation between variables and gives a good prediction. A pictorial view of ELM is shown in Fig. 2.

For the formation of a classification model from a set of data ($W = \{m_h \in B^u\}$ (h = 1, 2, 3, …, t)) having parameters such as t samples and u input, it is simple to write the output of an ELM network with u inputs, d hidden neurons, and y output neurons.

$$Y_j(h) = M_j^N d(h) \tag{13}$$

where $M_j \in B^d$ (h ∈ { 1, 2, 3, … y}) is the weight vector; it touched with the hidden neurons to the jth output neurons; and output vector of hidden neurosis H(j) ∈ B^d, that is given by

$$H(h) = \left[f(x_1^N m(h) + r_1), f(x_2^N n(h) + r_2), \ldots, f(x_d^N(h) + r_d) \right] \tag{14}$$

where r_v ($v = 1, 2, …, d$) and $x_V \in B^u$ are the bias of the vth hidden neurons and weight factor of vth hidden neurons respectively, and $f(\bullet)$ indicates the initiation function (sigmoidal function). It was observed that the bias (r_k) and weight (x_v) vectors are produced in a casual means from which the hidden layer output matrix (T) can be formed.
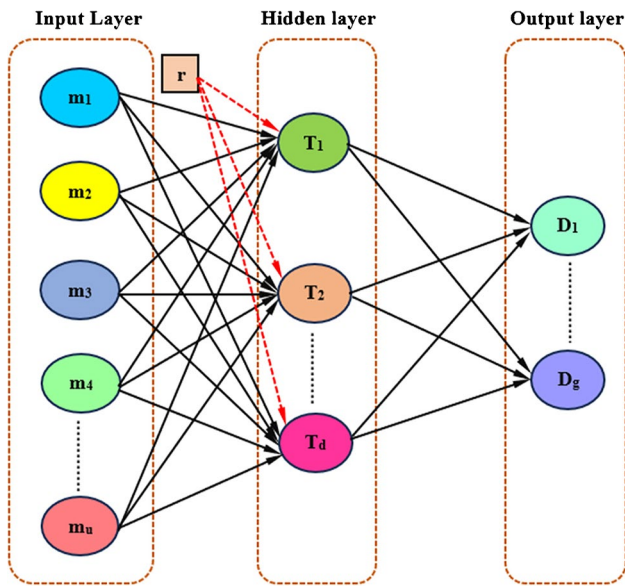
Fig. 2 Extreme learning machine

The weight matrix is then planned by the "Moore–Penrose pseudo inverse" method, which is given by:

$$X = \left(T \times T^N\right)^{-1} T \times Q^N \tag{15}$$

where $Q = [e(1), e(2) \ldots, e(t)]$ shows a matrix of $g \times t$ dimensions and its hth column is real largest vector $e(h) \in B^g$. The class label for new input parameter can be calculated as:

$$O = argmax\{D_j\}(j = 1, 2, \ldots, D) \tag{16}$$

where O represents predicted class label. Other detail working guidelines of ELM can be seen from Zhang et al. (2021) (Fig. 3).

Extreme gradient boosting (XGBoost)

XGBoost (extreme gradient boosting) is a powerful computational algorithm which is designed to improve the accuracy of a previous model, and it employs a complicated function to prevent overfitting. Applications of XGBoost are found in many types of programming language such as Python, R, Java, and Julia. Due to its high accuracy and speed, XGBoost is very popular in the field of data science (Chen and Guestrin 2016). It gives best performance in developing prediction models for DNA replications and DNA sequence in general (Do and Le., 2020), forecasts glucose concentration with the help of robust machine learning algorithm (Mekonnen et al. 2020), and resolves Walmart's sales problem using XGBoost algorithm and meticulous features engineering (Niu., 2020). The exact formula for the objective loss function of XGBoost approach is given below:
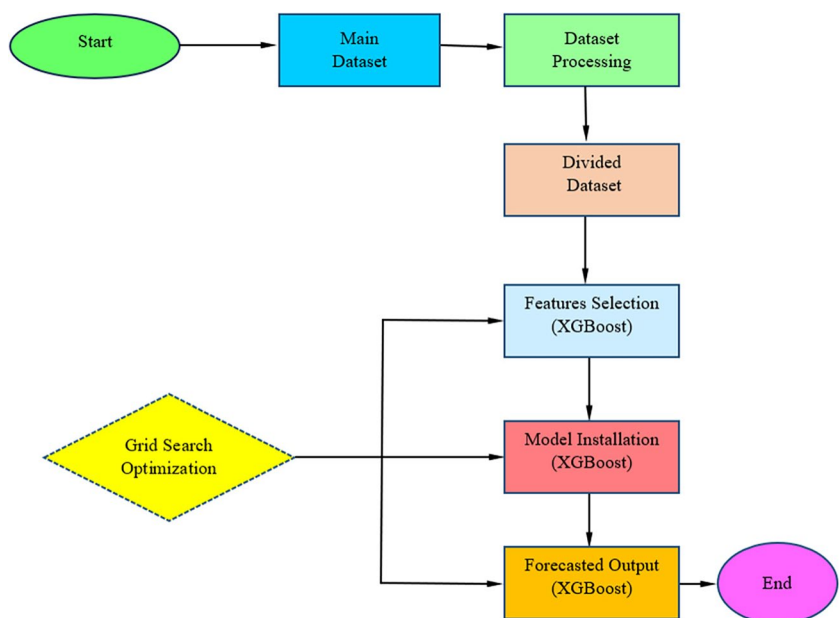
$$\text{Objective}(a) = \sum_{j=1}^{q} p\left(Y_j, Y_j^a\right) + \Omega(f_a) + As \tag{17}$$

$$Y_j^a = Y_j^{a-1} + f_a(z_j) \tag{18}$$

$$\Omega(f_a) = \rho X_a + \frac{1}{2} \lambda \|m\|^2 \tag{19}$$

where X_a indicates the number of leaves there are on the tree. A is constant, z_j is the input vector, and λ are ρ hyperparameters. Y_j stands for the mean the actual value and Y_j^a

Fig. 3 Extreme gradient boosting

stands for the forecast value of sales, respectively, similarly. $P(\bullet)$ and $f_a(\bullet)$ stands for square loss function and a regression tree, respectively. Based on Taylor's formula, function of an object ($\bullet$) is roughly stated in the following ways:

$$\text{Object}(a) \approx \sum_{j=1}^{q} [p(Y_j, Y_j^{a-1} + e_j f_a(z_a) + \frac{1}{2} k_j f_a^2(z_j)] + \Omega(f_a) + A \tag{20}$$

where e_j and $\frac{1}{2}k_j$ expressed as a first and quadratic term's coefficients in the Taylor expansion.

Dataset preparation

With the help of mean and standard deviation, input data are generated. Input variables, namely, quartz content (Q_C), degree of saturation (S), porosity (η), and dry density (γ) of coarse soil are considered to compute the output (thermal conductivity of soil). For this purpose, mean and standard deviation of coarse soil input variable were taken from past research's Tokoro et al. (2016). The statistical depictions of input parameters are given in Table 1.

A total of 110 dataset were taken, and after that, the input (Q_C, S, η, and γ) and output (thermal conductivity of soil) variables have been normalized. Before the consideration of the machine learning models, there is an important stage, i.e., normalization of datasets. Normalizations is conducted to remove dimensionality influence of the variable. That is why for the generation of any model, the output and input were normalized among 0 and 1 using the following equation:

$$M_{\text{normalized}} = \frac{M - M_{\min}}{M_{\max} - M_{\min}} \tag{21}$$

where $M_{\max}$ and $M_{\min}$ are the maximum and minimum values of the parameter (M) taken into account, respectively. After that the dataset obtained by normalization is split into two parts, i.e., training (TR) phase and testing (TS) phase. Training phase is done by 70% (77 cases) of the dataset of the normalized value and testing phase is done by 30% (33 cases) of the dataset of the normalized value. The methodology flowchart is represented in Fig. 4.

Statistical performance parameter

The prediction accuracy of proposed machine learning models such as extreme learning machine (ELM), adaptive neuro fuzzy inference system (ANFIS), and extreme gradient boosting (XGBoost) was examined by calculating numerous statistical performance parameters and evaluated graphically (i.e., scatter plot). We used various trend measuring statistical performance parameters like coefficient of determination (R^2), a-20 index, Willmott's index of agreement (WI), variance account factor (VAF), Nash–Sutcliffe efficiency (NS), and error measuring parameters like mean absolute error (MAE), root mean square error (RMSE), scatter index (SI), RMSE-observation standard deviation ratio (RSR), and weighted mean absolute percentage error (WMAPE).

Table 1 Statistical depiction of dataset

	Input parameters				Output
	Quartz content (Q_c)	Porosity (η)	Degree of saturation (S)	Dry density (γ) (In g/cc)	Thermal conductivity (TC) (in W m^{-1} K^{-1})
Mean (μ)	0.6125	0.4637	0.3321	1.4387	1.2396
Std. deviation (σ)	0.3052	0.0665	0.2173	0.1837	0.4080
1st quartile (Q1)	0.3863	0.4229	0.2482	1.3336	0.9772
2nd quartile (Q2)	0.5873	0.4675	0.3779	1.4434	1.1207
3rd quartile (Q3)	0.7582	0.5157	0.4655	1.5508	1.4546
Maximum	1.481	0.6059	1.0148	1.8744	2.6796
Minimum	0.0448	0.2876	0.1052	1.0403	0.5818
Skewness	0.3905	−0.1483	0.5639	0.0354	0.9632
Kurtosis	−0.1127	−0.3628	0.8786	−0.1795	1.0156
Std. error	0.0289	0.0063	0.0206	0.0174	0.1177
Geometric mean	0.5108	0.4618	0.3302	1.4360	1.1780
Harmonic mean	0.3905	0.4568	0.2872	1.4256	1.121
Coefficient of variation (%)	49.8286	14.3411	65.4321	12.7685	32.9151

Fig. 4 Flowchart of methodology
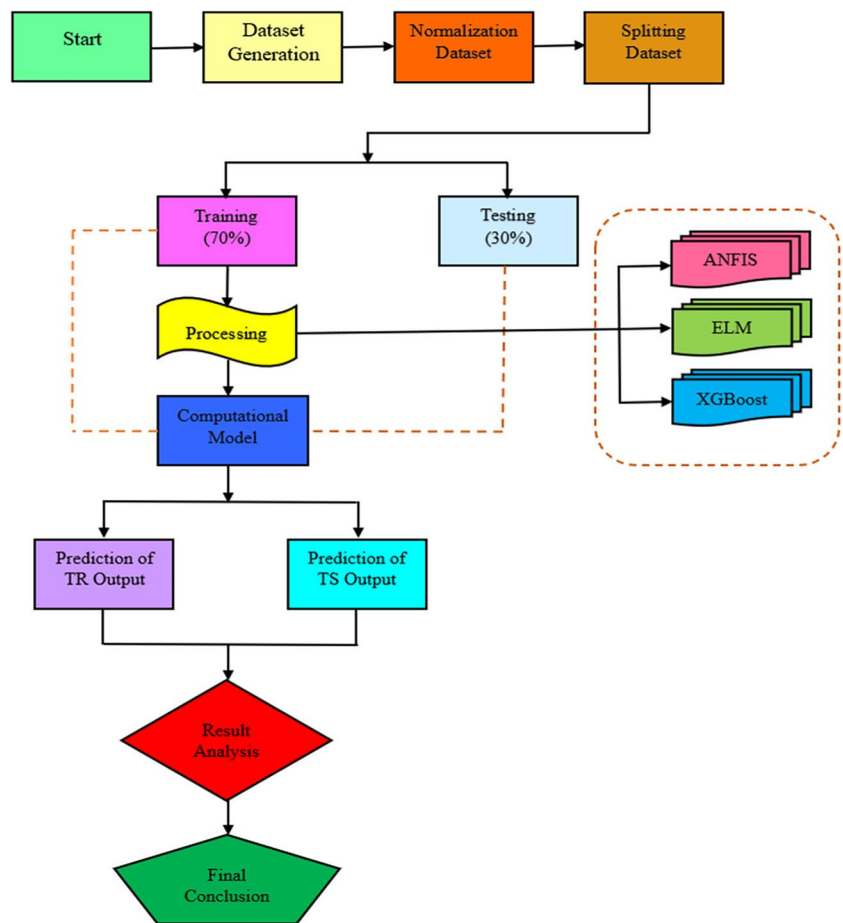
$$R^2 = \frac{\sum_{i=1}^{r}\left(TC_{obs,i} - \overline{TC_{obs}}\right)^2 - \sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2}{\sum_{i=1}^{r}\left(TC_{obs,i} - \overline{TC_{obs}}\right)^2}$$

$$(22)$$

$$a - 20\text{Index} = \frac{g20}{r} \qquad (23)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2}{\sum_{i=1}^{r}\left(\left|TC_{\text{pred}} - \overline{TC_{obs}}\right| + \left|TC_{obs,i} - \overline{TC_{obs}}\right|\right)^2}\right]$$

$$(24)$$

$$VAF = \left(1 - \frac{var\left(TC_{obs,i} - TC_{\text{pred},i}\right)}{var\left(TC_{obs,i}\right)}\right) \times 100 \qquad (25)$$

$$NS = 1 - \frac{\sum_{i=1}^{r}\left(TC_{obs} - TC_{\text{pred}}\right)^2}{\sum_{i=1}^{r}\left(TC_{obs} - \overline{TC_{obs}}\right)^2}$$

$$(26)$$

$$MAE = \frac{1}{r}\sum_{i=1}^{r}\left|\left(TC_{\text{pred},i} - TC_{obs,i}\right)\right| \qquad (27)$$

$$RMSE = \sqrt{\frac{1}{r}\sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2} \qquad (28)$$

$$SI = \frac{\sqrt{\frac{1}{r}\sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2}}{\overline{TC_{obs}}} = \frac{\text{RMSE}}{\overline{TC_{obs}}} \qquad (29)$$

$$RSR = \frac{\sqrt{\frac{1}{r}\sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2}}{\sqrt{\frac{1}{r}\sum_{i=1}^{r}\left(TC_{obs,i} - \overline{TC_{obs}}\right)^2}} = \frac{\text{RMSE}}{\sqrt{\frac{1}{r}\sum_{i=1}^{r}\left(TC_{obs,i} - TC_{\text{pred},i}\right)^2}}$$

$$(30)$$

$$\text{WMAPE} = \frac{\sum_{i=1}^{r}\left|\frac{TC_{obs,i} - TC_{\text{pred},i}}{TC_{obs,i}}\right| \times TC_{obs,i}}{\sum_{i=1}^{r} TC_{obs,i}} \qquad (31)$$

where $TC_{obs,i}$ and $TC_{\text{pred},i}$ are the actual and predicted thermal conductivity of ith values, respectively, $\overline{TC}_{obs}$ is the mean thermal conductivity of actual value, r is the total

number of training and testing samples, g20 is the number of data that has both observed and predicted value ratios that lie between 0.80 and 1.20. With the help of the above equations, the model having the least value of RMSE, SI, MAE, RSR, and WMAPE and higher value of R^2, a-20 Index, VAF, WI, and NS have superior prediction ability.

Results and analysis

Prediction power

In this portion, calculation of various performance parameters is done, which will be used in deciding the best predictive model. In this study, three machine learning models are utilized to predict thermal conductivity (unsaturated condition) of soil. The three machine learning models, namely, ANFIS, ELM, and XGBoost, are used to predict thermal conductivity of soil. Numerous performance parameters were computed to assess these models' performance. Performance parameter is further divided into two forms, that is, the trend measuring parameters (R^2, a-20 index, VAF, WI and NS) and error measuring parameters (RMSE, MAE, SI, RSR, and WMAPE). The results of the performance parameters are listed in Table 2 (TR dataset) and Table 3 (TS dataset). Based on the performance parameters, ANFIS has been found to have more accurate prediction power throughout the training with higher value of $R^2 = 0.9913$, a-20 index $= 0.9480$, VAF $= 99.1554$, WI $= 0.9978$, and NS $= 0.9913$ and lower value of RMSE $= 0.0171$, MAE $= 0.0050$, SI $= 0.0538$, RSR $= 0.0931$, and WMAPE $= 0.0157$, whereas the same is decreased in the testing phase ($R^2 = 0.9762$, a-20 index $= 0.7879$, VAF $= 97.7848$, WI $= 0.9941$, NS $= 0.9762$, RMSE $= 0.0334$, MAE $= 0.0204$, SI $= 0.1104$, RSR $= 0.1543$, WMAPE $= 0.0674$). The performance of XGBoost model ($R^2 = 0.9804$, a-20 index $= 0.9351$, VAF $= 98.0956$, WI $= 0.9952$, NS $= 0.9804$,

Table 2 Comparison of model prediction power for TR dataset

Parameters		Ideal value	ANFIS	ELM	XGBoost
R^2	Trend measuring Parameter (TMP)	1	0.9913	0.9896	0.9804
a-20 index		1	0.9480	0.9351	0.9351
VAF		100	99.1554	98.9626	98.0956
WI		1	0.9978	0.9973	0.9952
NS		1	0.9913	0.9896	0.9804
RMSE	Error measuring Parameter (EMP)	0	0.0171	0.0186	0.0257
MAE		0	0.0050	0.0129	0.0132
SI		0.1	0.0538	0.0586	0.0808
RSR		0	0.0931	0.1013	0.1398
WMAPE		0	0.0158	0.0391	0.0409

Table 3 Comparison of model prediction power for TS dataset

Parameters		Ideal value	ANFIS	ELM	XGBoost
R^2	Trend measuring Parameter (TMP)	1	0.9762	0.9553	0.9901
a-20 index		1	0.7879	0.8182	0.8788
VAF		100	97.7848	95.6701	99.0228
WI		1	0.9941	0.9882	0.9975
NS		1	0.9762	0.9553	0.9901
RMSE	Error measuring Parameter (EMP)	0	0.0334	0.0457	0.0215
MAE		0	0.0204	0.0203	0.0120
SI		0.1	0.1104	0.1512	0.0712
RSR		0	0.1543	0.2114	0.0995
WMAPE		0	0.0674	0.0670	0.0032

RMSE $= 0.0257$, MAE $= 0.0132$, SI $= 0.0808$, RSR $= 0.1398$, WMAPE $= 0.0409$) is lower than ANFIS model in the training phase, while in the testing phase, it has higher predicting power ($R^2 = 0.9901$, a-20 Index $= 0.8788$, VAF $= 99.0228$, WI $= 0.9975$, NS $= 0.9901$, RMSE $= 0.0215$, MAE $= 0.0120$, SI $= 0.0712$, RSR $= 0.0995$, WMAPE $= 0.0032$). Based on these performance parameters, we may conclude that XGBoost performs better during testing phase, while ANFIS performs better during training phase. In training stage, ANFIS model outperformed the other models by far, while the XGBoost performed better in the testing phase. The trend is reversed due to lesser data in testing phase, as we know that in machine learning model, the more data you provide to the ML system, the faster that model can learn and improve.

Rank analysis

In this section, ranking of the proposed models (ANFIS, ELM, and XGBoost) is done according to the performance indicators. The value which is closer to the ideal value of the performance parameters is ranked 1 and the farthest value ranked 3, as we have used three models. To get the overall rank of the model, ranks of the training phase and testing phase are added. The models which have the lower overall rank will be the best model and model of higher overall rank is worst model. Ranking of the models is shown in Table 4. In the rank analysis, we have found that ANFIS gives best results in training phase (ANFIS$_{TR} = 12$) and XGBoost in the testing phase (XGBoost$_{TS} = 11$). Overall ranks of the models are as follows: ANFIS $= 34$, ELM $= 47$, XGBoost $= 38$. So, it can be clearly seen that ANFIS model gives the better prediction in comparison of ELM and XGBoost, as ANFIS has the lower rank among the three models that are used in this study. The graphical representation of rank analysis is also represented by radar diagram which is shown in Fig. 5.
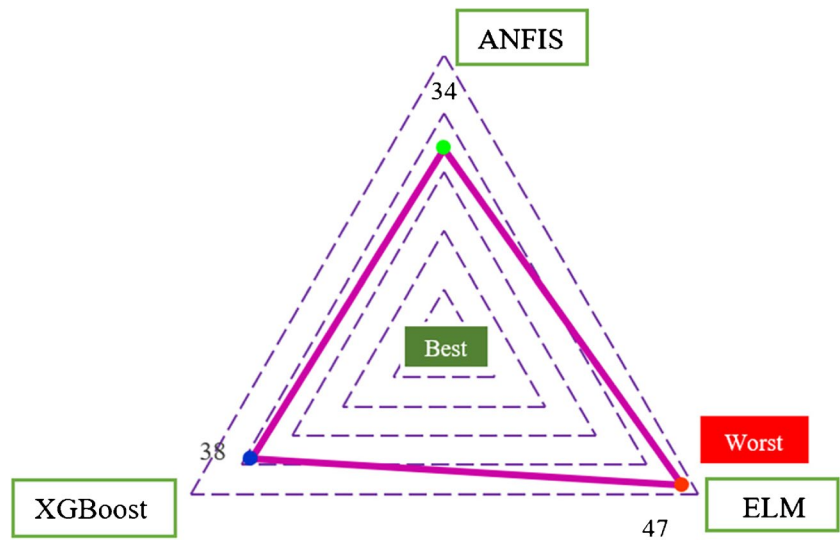
Table 4 Rank analysis of the models (training and testing)

Models	Phase	R^2	a-20 index	VAF	WI	NS	RMSE	MAE	SI	RSR	WMAPE	Total rank	Overall rank
ANFIS	TR	1	1	1	1	1	1	1	3	1	1	12	34
	TS	2	3	2	2	2	2	3	1	2	3	22	
ELM	TR	2	2	2	2	2	2	2	2	2	2	20	47
	TS	3	2	3	3	3	3	2	3	3	2	27	
XGBoost	TR	3	2	3	3	3	3	3	1	3	3	27	38
	TS	1	1	1	1	1	1	1	2	1	1	11	

Fig. 5 Radar diagram

Reliability analysis

Thermal conductivity of soil cannot be considered directly as accurate, as numerous uncertainties can possibly complicate it. This inaccuracy can be related to the soil properties or analytical parameter used. To identify unpredictability and to get a dependable method for predicting the thermal conductivity, reliability analysis has been done. To evaluate the reliability index (β), the first-order second moment (FOSM) approach has been used. In FOSM method, reliability index is computed with the help of average value (μ) and standard deviation (σ) of the performance function and can be computed as:

$$\beta = \frac{\mu - 1}{\sigma} \tag{32}$$

where μ and σ is the average value and standard deviation of thermal conductivity, respectively. The probability of failure (P_f) is directly dependent on the reliability index (β) on the assumption that all the normal variables are distributed normally. The probability of failure can be determined as:

$$P_f = 1 - \Phi(\beta) \tag{33}$$

Table 5 Ranking of models according to the reliability index (β)

Models	Existing β	Existing P_f	Model's β	Model's P_f	Rank
ANFIS			0.5871	0.2786	1
ELM	0.5848	0.2793	0.4351	0.3317	3
XGBoost			0.5605	0.2876	2

where $\Phi(\beta)$ denotes the standard normal cumulative probability. By using FOSM technique, the reliability index (β) of the models has been calculated (using Eq. 32), and comparison is done with the reliability index of the model with the actual value. A greater value of β shows a superior performance of the model. The greater value of reliability index and lesser value of the probability of failure (P_f) displays that ANFIS performs best among all three models (as shown in Table 5). Hence, the models are ranked on the basis of β and P_f. The model with greater value of β and lesser value of P_f will be ranked 1 as it shows better performance, and the model with lesser value of β and greater value of P_f will be ranked 3 as it shows worst performance. Also it can be observed that the ANFIS model's reliability index and probability of failure values closely match the existing value. As

a result, the ANFIS model may be trusted to be reliable for analyzing the thermal conductivity of soil.

Error matrix

An error matrix, also known as a confusion matrix, is a table used to evaluate the performance of the models that are used in this study. It is commonly used in machine learning. It provides a summary of the predictions made by the model with the observed outcomes. By analyzing the error matrix, various evaluation matrices can be derived (accuracy, precision) to check the performance of the models. These matrices can help in understanding the strength and weaknesses of the model. Additionally, it conveys the idea of the maximum and minimum error of the predicted model. The error matrix is generated according to the ideal values of the performance parameters. In this study, ten performance parameters are used, viz., R^2, a-20 index, VAF, WI, NS, RMSE, MAE, SI, RSR, and WMAPE, where R^2, a-20 index, VAF, WI, and NS are trend measuring parameters and RMSE, MAE, SI, RSR, and WMAPE are error measuring parameters.

$$E_e = |(1 - |P|)| \times 100 \tag{34}$$

$$E_t = \frac{(I - |P|)}{I} \times 100 \tag{35}$$

where E_e and E_t represent the error in error measuring parameters (EMPs) and error in trend measuring parameter (TMP), respectively; I is the ideal value of trend and error measuring parameter; and P is the value of performance parameters estimated for error measuring parameter and trend measuring parameter. Calculation of error for error measuring and trend measuring parameter estimated by using Eqs. 34–35. Tables 6, 7 and 8 represent the calculation of error for trend measuring parameter and Table 7, 8 and 9 show the calculation of error for error measuring parameters.

In this section, error is calculated for trend measuring parameter (TMP) and error measuring parameter (EMP). Error matrix for trend and error measuring parameter for both TR and TS dataset is shown in Figs. 6 and 7, respectively. TMP are R^2, a-20 index, VAF, WI, and NS and EMP are RMSE, MAE, SI, RSR, and WMAPE considered in this study. The lowest error has been shown by green color, moderate error by yellow color, and highest error by red color. As per the error matrix of the thermal conductivity of soil of trend measuring and error measuring parameter, ANFIS model gives less error than the other two models used in this study in the training phase. In the TS phase, XGBoost

Table 6 Calculation of error for trend measuring parameter (TR dataset)

TMP	Ideal value	ANFIS	Error (%)	ELM	Error (%)	XGBoost	Error (%)
R^2	1	0.9913	0.8661	0.9896	1.0404	0.9804	1.9549
a-20 index	1	0.9480	5.1948	0.9351	6.4935	0.9351	6.4935
VAF	100	99.1554	0.8446	98.9626	1.0374	98.0956	1.9044
WI	1	0.9978	0.2167	0.9973	0.2669	0.9952	0.4757
NS	1	0.9913	0.8661	0.9896	1.0404	0.9804	1.9549

Table 7 Calculation of error for error measuring parameter (TR dataset)

EMP	Ideal value	ANFIS	Error (%)	ELM	Error (%)	XGBoost	Error (%)
RMSE	0	0.0171	1.7134	0.0186	1.8658	0.0257	2.5741
MAE	0	0.0050	0.5040	0.0129	1.2884	0.0132	1.3164
SI	0.1	0.0538	4.6179	0.0586	4.1391	0.0809	1.9141
RSR	0	0.0931	9.3065	0.1013	10.1345	0.1398	13.9817
WMAPE	0	0.0158	1.5771	0.0391	3.9092	0.0409	4.0921

Table 8 Calculation of error for trend measuring parameter (TS dataset)

TMP	Ideal value	ANFIS	Error (%)	ELM	Error (%)	XGBoost	Error (%)
R^2	1	0.9762	2.3826	0.9553	4.4692	0.9901	0.9898
a-20 index	1	0.7879	21.2121	0.8182	18.1818	0.8788	12.1212
VAF	100	97.7848	2.21524	95.6701	4.3299	99.0228	0.9772
WI	1	0.9941	0.5876	0.9882	1.1795	0.9975	0.2510
NS	1	0.9762	2.3826	0.9553	4.4692	0.9901	0.9898

Table 9 Calculation of error for error measuring parameter (TS dataset)

EMP	Ideal value	ANFIS	Error (%)	ELM	Error (%)	XGBoost	Error (%)
RMSE	0	0.0334	3.3405	0.0457	4.5751	0.0215	2.1531
MAE	0	0.0204	2.0387	0.0203	2.0288	0.0120	1.2025
SI	0.1	0.1104	1.0423	0.1512	5.1233	0.0712	2.8828
RSR	0	0.1543	15.4357	0.2114	21.1404	0.0995	9.9488
WMAPE	0	0.0674	6.7390	0.0671	6.7063	0.0032	0.3255

Fig. 6 Error matrix for trend measuring parameters (TR and TS dataset)

Fig. 7 Error matrix for error measuring parameters (TR and TS dataset)

performs best as it has less error on comparing with other two models.

Regression curve

Regression curve is drawn among the predicted value and observed value of soil's thermal conductivity. Regression curve is also known as R-curve. It provides the calculated R-value that is displayed in Table 3 and 4. Observed thermal conductivity of soil is shown on x-axis, and predicted soil's thermal conductivity is shown on y-axis. Both the datasets of TR and TS phase is shown in the same plot. In Fig. 8, dotted lines indicate ± 10% deviation between the actual line and the projected data. It can be observed from the curve that the observed value and predicted value overlaps in the case of ANFIS and XGBoost model, but slight deviation can be seen in the ELM model.

William's plot

William's plot is drawn between standardized residual and leverage, in which standardized residual is plotted on the ordinate and leverage on the abscissa. In order to determine whether an AI model is capable of producing accurate

Fig. 8 Scatter plot for models **a** ANFIS; **b** ELM; and **c** XGBoost
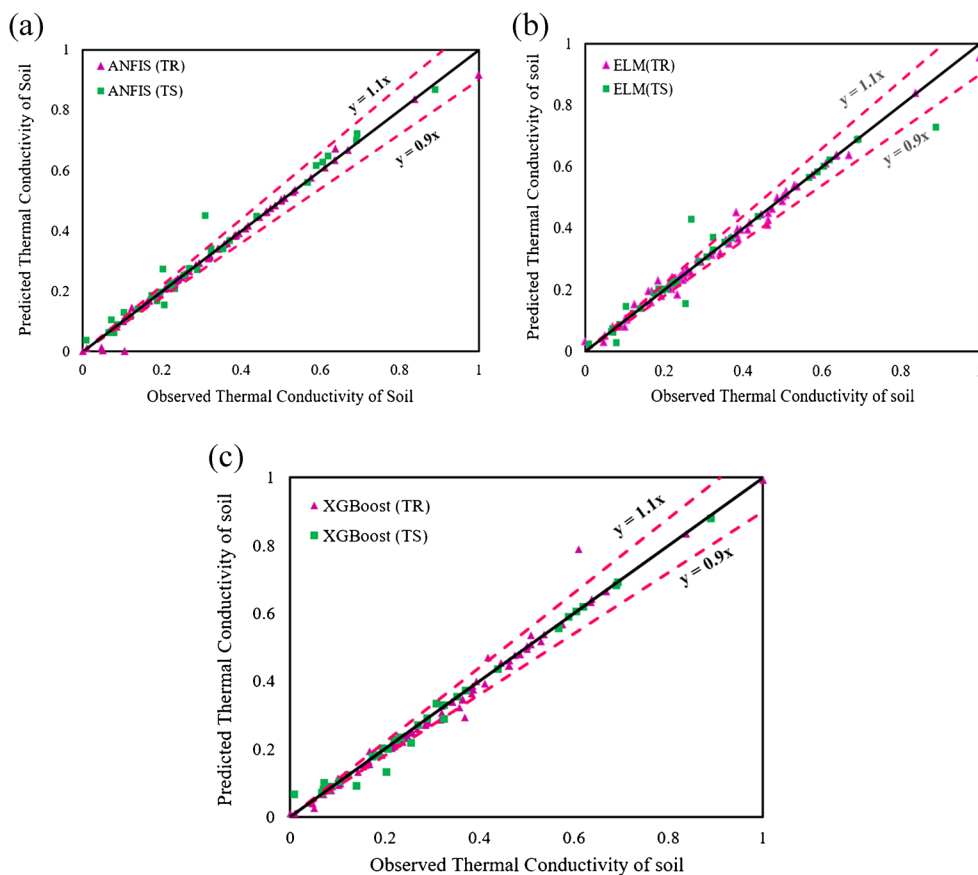
predictions, it is important to evaluate the applicability domain of three AI models that are considered in this study. By outlining the leverage (h) values of the TR and TS datasets, the three separate AI models' applicability areas were identified. The William's plot for both TR and TS dataset is shown in Fig. 9. A squared zone with a leverage threshold h* and a range of ± 3 standard deviations surround the applicability of this domain. The leverage threshold (h*) is normally fixed at 3 p/m, where m is the total number of training compound and p is the number of model parameters plus one (m + 1). The value of leverage threshold is taken as 0.195. From the plot, it can be seen that all the compounds had leverage lower than the threshold leverage of the TR and TS datasets, but despite that the leverage is less than threshold leverage, there is some value in the plot which lies outside the square boundary, i.e., two training and two testing in the case of ANFIS, one training and one testing for XGBoost, and for the ELM four testing and three training datasets.

Sensitivity analysis

The impact on the output (thermal conductivity of soil) due to each input parameters (Q_c, η, S, γ) is calculated for all the three models (namely, ANFIS, ELM, and XGBoost). A parameter called strength of relation is calculated to know the impact on thermal conductivity of soil. It can be calculated as:

$$SOR = \frac{\sum_{i=1}^{r} R_{m,i} S_{n,i}}{\sqrt{\sum_{i=1}^{r} (R_{m,i})^2 \sum_{i=1}^{r} (S_{n,i})^2}} \quad (36)$$

where $R_{m,i}$ denotes the ith value of mth independent variable, r denotes the total number of observation dataset, and m denotes the total number of input parameters considered, and $S_{n,i}$ denotes the ith value of nth dependent variable. Here, to know the impact of each input parameters, ($m = 4$, $n = 1$, $r = 110$) has been considered. The greater value of strength shows greater influence of input parameters on the output. The strength of relation of numerous input parameters will be different according to their impact on output parameter as shown in Fig. 10. From the figure, it can be observed that the dry density of soil (γ) is the most influential parameter and quartz content (Q_c) is the least influential parameter.
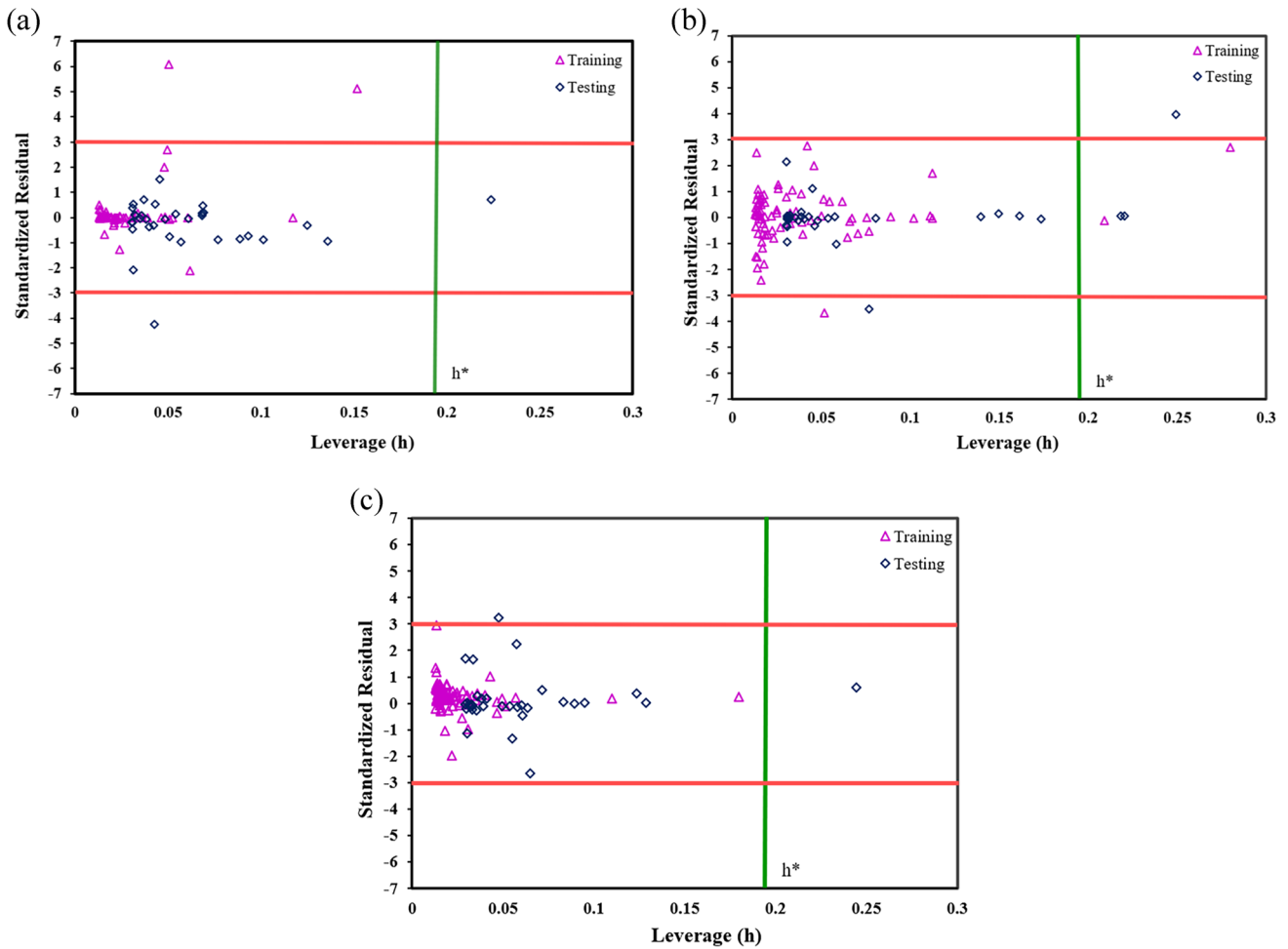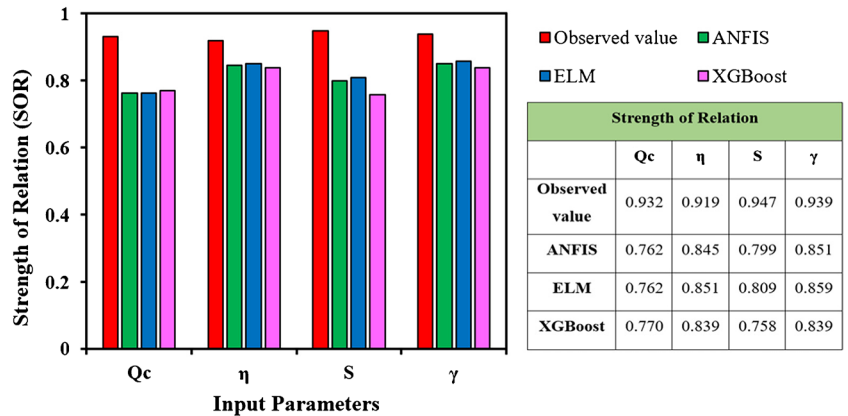
Fig. 9 William's plot for models **a** ANFIS; **b** ELM; and **c** XGBoost

Fig. 10 Strength of relation of input parameters

Strength of Relation				
	Qc	**η**	**S**	**γ**
Observed value	0.932	0.919	0.947	0.939
ANFIS	0.762	0.845	0.799	0.851
ELM	0.762	0.851	0.809	0.859
XGBoost	0.770	0.839	0.758	0.839

Conclusion

In this project, three predictive models were developed to compute the thermal conductivity of soil, namely, ANFIS, ELM, and XGBoost, using four input parameters of soil

(Q_c, η, S, γ). We compared all three models with several parameters like R-curve, rank analysis, and error matrix, and as a result, it can be seen that ANFIS shows better outcome amongst all three models (ANFIS, ELM, and XGBoost) in the training stage, and XGBoost shows the best prediction ability in the testing stage to determine

thermal conductivity of soil amongst all three models. The total rank achieved by ANFIS model in the testing stage is 22 and did better than the two models by a long shot, while the total rank achieved by ELM model is 27 and XGBoost model is 11 in testing phase. But, when evaluation is done according to overall rank (training stage rank + testing stage rank) gained by all models, the uppermost predictive exactness is displayed by ANFIS with the final rank of 34, which is traced by ANFIS, XGBoost, and ELM. Reliability index (β) and probability of failure (P_f) were calculated too for all three models and matched with existing value. A greater value of β and lesser value of P_f is given by ANFIS model amongst all the three models. Other than this, the radar diagram has also drawn to show the appropriateness of model. Sensitivity analysis is also performed to know the influence of each input parameters on the output. For this, strength of relation is calculated to know the impact of each input parameters on the output (thermal conductivity of soil). From sensitivity analysis it can be observed that dry density of soil (γ) is the most dominant parameter among all four input parameters, followed by porosity (η), degree of saturation (S), and quartz content (Q_c).

Author contributions RM: Conceptualization, formal analysis, investigation, software, validation, visualization. KK: Writing—original draft. SK: Formal analysis, data curation. GK: Methodology. PS: Writing—review and editing.

Data availability The data presented in this study are available on request from the corresponding author.

Declarations

Ethics approval Not applicable.

Informed consent Not applicable.

Conflict of interest The authors declare no competing interests.

References

Abu-Hamdeh NH, Reeder RC (2000) Soil thermal conductivity: effects of density, moisture content, salt concentration and organic matter. Soil Sci Soc Am J 64:1285–1290. https://doi.org/10.2136/sssaj2000.6441285x4

Abu-Hamdeh NH, Khdair AI, Reeder RC (2001) A comparison of two methods used to evaluate thermal conductivity for some soils. Int J Heat Mass Transf 44:1073–1078. https://doi.org/10.1016/S0017-9310(00)00144-7

Abu-Hamdeh NH (2003) Thermal properties of soils as affected by density and water content. Biosys Eng 86(1):97–102. https://doi.org/10.1016/S1537-5110(03)00112-0

Adrinek S, Singh RM, Janza M, Zerun M, Ryzynski G (2022) Evaluation of thermal conductivity estimation models with laboratory-measured thermal conductivities of sediments. Environ Earth Sci 81(15). https://doi.org/10.1007/s12665-022-10505-7

Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. The 22nd ACM SIGKDD international conference. https://doi.org/10.1145/2939672.2939785

Cosenza P, Guerin R, Tabbagh A (2003) Relationship between thermal conductivity and water content of soils using numerical modelling. Eur J Soil Sci 54:581–587. https://doi.org/10.1046/j.1365-2389.2003.00539.x

Da TX, Chen T, He WK, Elshaikh T, Ma Y, Tong ZF (2022) Applying machine learning methods to estimate the thermal conductivity of bentonite for a high-level radioactive waste repository. Nucl Eng Des 392. https://doi.org/10.1016/j.nucengdes.2022.111765

Dong Y, McCartney JS, Lu N (2015) Critical review of thermal conductivity models for unsaturated soils. Geotech Geol Eng 33:207–221. https://doi.org/10.1007/s10706-015-9843-2

Do DT, Le NQK (2020) Using extreme gradient boosting to identify origin of replication in Saccharomyces cerevisiae via hybrid features. Genomics 112(3):2445–2451. https://doi.org/10.1016/j.ygeno.2020.01.017

Farouki QT (1981) Thermal properties of Soils in Cold Region. Cold Reg Sci Technol 5:67–75. https://doi.org/10.1016/0165-232X(81)90041-0

Huang GB, Zhi QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501. https://doi.org/10.1016/j.neucom.2005.12.126

Huang GB, Wang DH, Lan Y (2011) Extreme learning machine: a survey. Int J Mach Learn Cybern 2:107–122

Hu G, Zhao L, Wu X, Li R, Wu T, Xie C, Qiao Y, Shi J, Li W, Cheng G (2016) New Fourier-series-based analytical solution to the conduction-convection equation to calculate soil temperature, determine soil thermal properties, or estimate water flux. Int J Heat Mass Transf 95:815–823. https://doi.org/10.1016/j.ijheatmasstransfer.2015.11.078

Johansen O (1975) Thermal conductivity of soils. Ph.D thesis, Trondheim, Norway. (CRREL Draft Translation 637, 1977). ADA 044002.

Kardani N, Bardhan A, Samui P, Nazem M, Zhou A, Armaghani DJ (2021) A novel technique based on the improved firefly algorithm coupled with extreme learning machine (ELM-IFF) for predicting the thermal conductivity of soil. Eng Comput 38:3321–3340

Li KQ, Kang Q, Nie JY, Huang XW (2022a) Artificial neural network for predicting the thermal conductivity of soils based on a systematic database. Geothermics. https://doi.org/10.1016/j.geothermics.2022.102416

Li KQ, Liu Y, Kang Q (2022) Estimating the thermal conductivity of soils using six machine learning algorithms. International Communications in Heat and Mass Transfer 136. https://doi.org/10.1016/j.icheatmasstransfer.2022.106139

Liu W, Li R, Wu T, Shi X, Zhao L, Wu X, Hu G, Yao J, Xiao Y, Ma J, Jiao Y, Wang S (2023) Simulation of soil thermal conductivity based on difference schemes: an empirical composition of 13 models. Int J Therm Sci. https://doi.org/10.1016/j.ijthermalsci.2023.108301

Malek K, Malek K, Khan MF (2021) Response of soil thermal conductivity to various soil properties. Int Commun Heat Mass Transfer 127:1–8. https://doi.org/10.1016/j.icheatmasstransfer.2021.105516

Mekonnen BK, Yang W, Hsieh TN, Liaw SH, Yang FL (2020) Accurate prediction of glucose concentration and identification of major contributing features from hardly distinguishable near-infrared spectroscopy. Biomed Signal Process Control. https://doi.org/10.1016/j.bspc.2020.101923

Mustafa R, Samui P, Kumari S (2022) Reliability analysis of gravity retaining wall using hybrid ANFIS. Infrastructures 7(9):121. https://doi.org/10.3390/infrastructures7090121

Mustafa R, Samui P, Kumari S, Mohamad ET, Bhatawdekar RM (2023) Probabilistic analysis of gravity retaining wall against

bearing failure. Asian J Civil Eng 24(8):3099–3119. https://doi.org/10.1007/s42107-023-00697-z

Niu Y (2020) Walmart sales forecasting using XGBoost algorithm and feature engineering. Inst Elect Electron Eng. https://doi.org/10.1109/ICBASE51474.2020.00103

Nusier OK, Abu-Hamdeh NH (2003) Laboratory techniques to evaluate thermal conductivity for some soils. Heat Mass Transfer 39:119–123. https://doi.org/10.1007/s00231-002-0295-x

Rizvi ZH, Zaidi HH, Akhtar SJ, Sattari AS, Wuttke F (2020) Soft and hard computation methods for estimation of the effective thermal conductivity of sands. Heat Mass Transf 56:1947–1959

Rubin AJ, Ho CL (2021) Soil thermal conductivity estimated using a semi-analytical approach. Geothermics. https://doi.org/10.1016/j.geothermics.2021.102051

Sepaskhah AR, Boersma L (1979) Thermal conductivity of soils as a function of temperature and water content. Soil Sci Soc Am J 43:439–444. https://doi.org/10.2136/sssaj1979.03615995004300030003x

TaeBang H, Yoon S, Jeon H (2020) Application of machine learning methods to predict a thermal conductivity model for compacted bentonite. Annals of Nuclear Energy 142 https://doi.org/10.1016/j.anucene.2020.107395

Tokoro T, Ishikawa T, Shirai S, Nakamura T (2016) Estimation method for thermal conductivity of sandy soil with electrical characteristics. Soils Found 56(5):927–936. https://doi.org/10.1016/j.sandf.2016.08.016

Tong B, Gao Z, Horton R, Li Y, Wang L (2016) An empirical model for estimating soil thermal conductivity from soil water content and porosity. American Meteorological Society 601–613 https://doi.org/10.1175/JHM-D-15-0119.1

Wang C, Cai G, Liu X, Wu M (2022) Prediction of soil thermal conductivity based on intelligent computing model. Heat Mass Transfer 58:1695–1708. https://doi.org/10.1007/s00231-022-03209-y

Zhang R, Wu C, Goh ATC, Bohlke T, Zhang W (2021) Estimation of diaphragm wall deflections for deep braced excavation in anisotropic clays using ensemble learning. Geosci Front 12(1):365–373. https://doi.org/10.1016/j.gsf.2020.03.003