Prediction of Saturated Hydraulic Conductivity Using Artificial Neural Networks and Tree Boost Methods

Moussa S. Elbisy
Civil Engineering Department, College of Engineering and Architecture, Umm Al Qura University, Makkah, Saudi Arabia
mselbisy@uqu.edu.sa (corresponding author)

Abdullah S. Bostaji
Civil Engineering Department, College of Engineering and Architecture, Umm Al Qura University, Makkah, Saudi Arabia
asbostaji@uqu.edu.sa

Received: 10 March 2025 | Revised: 14 April 2025 | Accepted: 22 April 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | *DOI: https://doi.org/10.48084/etasr.10899*

ABSTRACT

The hydraulic conductivity of saturated soil is a critical parameter in the design of efficient drainage systems, reflecting the soil's ability to transmit water based on pore size and structure. This study explores the use of Artificial Neural Networks (ANNs) and tree boost models to predict the field saturated soil hydraulic conductivity (K_{field}) of sandy soils influenced by saline and alkaline conditions. Two ANN models were used, namely a Multilayer Perceptron (MPNN) and a General Regression Neural Network (GRNN). Soil samples were taken from El-Nubaria and Sinai in Egypt's western delta, and their physical and chemical properties were tested in the laboratory. Statistical analyses were performed to assess the relationships between these properties and K_{field}. Model training and evaluation were conducted using cross-validation and five evaluation metrics. The results show that the tree boost model outperformed the ANN model, demonstrating superior accuracy and predictive capability, indicating that tree boost algorithms could be very useful for estimating K_{field} in situations where data are limited or the soil is complicated.

Keywords-saturated soil hydraulic conductivity; soil properties; prediction; tree boost; general regression neural networks; multilayer perceptron

I. INTRODUCTION

Egypt is implementing large-scale subsurface drainage systems on newly reclaimed lands. Most of these lands have unstable sandy soil and need suitable drainage systems. Designing drainage systems requires information about the soil's properties, such as grain size distribution, texture and structure, plasticity index, hydraulic conductivity, mineralogy, and chemical composition. The saturated hydraulic conductivity (K_{sat}) of the soil is a key factor in the design and operation of subsurface drainage systems, as all drain spacing equations use this parameter. Therefore, to design or evaluate the needs of a drainage project, the determination of K_{sat} is required as accurately as possible. In [1], it was reported that K_{sat} is subject to variation in space and time, which means that one must adequately assess a representative value for K_{sat}. This study also mentioned that it is time-consuming and costly, so a balance has to be struck between budget limitations and desired accuracy.

Currently, there is no optimum surveying technique, and much depends on the skill of the person conducting the survey. The K_{sat} of the soil represents its average hydraulic conductivity, which depends mainly on the size, shape, and distribution of the pores. Moreover, it depends on the temperature of the soil and the viscosity and density of the water. In some soils (e.g. structureless sandy sediments), K_{sat} is the same in all directions; however, its value varies with the direction of flow. The vertical permeability of the soil or a soil layer is often different from its horizontal permeability due to vertical differences in texture, structure, and porosity. The measured K_{sat} of the soil often shows a log-normal distribution with wide variation [2]. In [3], the finite element method was used to determine the effect of the variation in K_{sat} on the drawing of the water table between the drains. This study concluded that the best estimate would be the average of the arithmetic and geometric means. The standard deviation of the observed K_{sat} depends on the method of determination.

Soil salinity usually has a positive influence on hydraulic conductivity, especially in clay soil [1]. Upon reclamation, saline soils can become less permeable. Sodic soils experience a dispersion of soil particles and deterioration in the structure, resulting in poor K_{sat}. Sodic soils are formed when sodium carbonates are present in the soil or are introduced with irrigation water [4]. The K_{sat} of soils can be determined using either correlation methods or hydraulic methods. Hydraulic methods can be either laboratory or in situ (field) methods.

The correlation methods for determining K_{sat} in drainage surveys are frequently based on relationships with one or more soil properties (i.e. soil texture, pore size distribution, grain size distribution, or soil mapping unit). In [5], a correlation was presented between the content of silt and clay for subsoil materials in Imperial Valley, California, USA, and the results of hydraulic laboratory tests. In [6], generalized tables with ranges of *K*-values were provided for certain soil textures. However, such tables should be handled with care [1]. In addition, a warning was highlighted in [6], underscoring that soils with identical texture may have quite different *K*-values due to differences in structure, and some heavy clay soils have well-developed structures and much higher *K*-values than those indicated in the table. The pore size distributions, the regularity of the pores, and their continuity have a significant influence on K_{sat}. However, the study and characterization of porosity to evaluate K_{sat} is not advanced enough to be practical on a large scale [1]. In [7], an example of the complexity of such a study was provided using micromorphometric data for clay soils. In [8], the pore size distribution was determined using the relationship between the soil water content and the metric head, providing another example. Applying Poiseuille's law to some fractions of the pF-curve can estimate K_{sat}. This method is mainly applicable to granular (sandy) soils without systematic continuous pores [1]. The grain size distribution determines the soil permeability in sandy soils that lack continuous systematic pores.

Over the past decade, Machine Learning (ML) methods have emerged as significant instruments for nonlinear modeling. ML models are particularly applicable when the use of conventional or phenomenological regression models is impractical or inconvenient. Recently, there has been a notable appearance of substantial publications dedicated to determining K_{sat} using ML approaches. In [9], Support Vector Machine (SVM) and Multiple Linear Regression (MLR) methods were employed to estimate the K_{sat} of sandy soil. This study focused on utilizing soil properties that are easily measurable, such as the clay/silt ratio, the liquid limit, and the chloride ion content. This study examined the effectiveness of SVM utilizing various kernel functions, including linear, radial basis, and sigmoid, and employed genetic algorithms to enhance model performance. The findings showed that the SVM model utilizing the RBF kernel outperformed the MLR models in predicting K_{sat}, indicating that SVM is an effective approach for predicting soil properties, providing a more accurate estimation of K_{sat} compared to MLR methods, particularly in the context of complex datasets such as sandy soils.

In [10], Artificial Neural Networks (ANNs) were used to improve the estimation of K_{sat} in smectitic soils, incorporating fractal parameters derived from soil particle and micro-aggregate size distributions. Pedotransfer Functions (PTFs) were used to estimate K_{sat} by integrating fractal parameters with traditional soil data, such as bulk density, clay content, and sand content. The findings showed that the incorporation of fractal parameters along with traditional soil data significantly improves the prediction of K_{sat}. ANN ensemble models, particularly those incorporating fractal parameters, demonstrated superior accuracy and reliability in predictions compared to conventional methods or individual ANN models. In [11], ANNs and MLR were used to predict soil hydraulic conductivity, concluding that ANNs are more effective for predicting K_{sat}, particularly across various soil types, due to their ability to model complex relationships. MLR models, while limited in their ability to capture non-linear dependencies, are still valuable in contexts where data relationships are relatively simple.

In [12], a Supervised Committee Machine Artificial Intelligence (SCMAI) model was employed to predict *K* based on the distribution of soil grain sizes. This model integrates three forms of artificial intelligence: Larsen Fuzzy Logic (LFL), Least Squares Support Vector Machine (LSSVM), and Wavelet-Artificial Neural Network (WANN). The SCMAI model effectively predicted hydraulic conductivity, outperforming conventional individual models. In [13], several PTFs were employed to examine the K_{sat} of agricultural soil. These functions determine K_{sat} by analyzing basic soil parameters such as particle size distribution, bulk density, and organic matter concentration. This study analyzed ten PTF models, which encompassed NN-based frameworks, such as Rosetta-SSC, in addition to ML algorithms including Random Forest (RF) and Boosted Regression Trees (BRTs). This study used data from a single agricultural field, which exhibited various tillage treatments and alterations in spatial and temporal parameters. The study concluded that ML-based PTF models typically fell short of measured K_{sat} values, indicating the need to improve these models for more precise predictions.

In [14], a novel model was used to estimate K_{sat}, highlighting the influence of particle sizes. This model enhanced the established Kozeny-Carman (KC) model by incorporating the concept of equivalent particle size and examining the influence of adsorbed water films on fine-grained soils. The method was designed to address the deficiencies of the KC model, which is less effective for fine-grained soils due to the intricate characteristics of the absorbed water and the pore structure of these materials. This model provided a more precise method to estimate K_{sat}, particularly in fine-grained soils, as it considered the influence of absorbed water and the pore structure. In [15], ML models, specifically Radial Basis Function Neural Networks (RBFNN), Multilayer Perceptron Neural Networks (MPNN), and hybrid models integrating Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) with NN (GA-NN and PSO-NN), were used to predict K_{sat}. This study used data from a field study conducted at the Bajgah Agricultural Experimental Station in Iran to evaluate the predictive accuracy of these models. Many soil parameters were evaluated, including bulk density, moisture content, and aggregate size. The PSO-NN model achieved the highest accuracy with a substantial correlation

coefficient of 0.958, surpassing RBFNN, MPNN, and GA-NN which exhibited commendable outcomes but with somewhat greater errors. In [16], three ML models, namely RF, SVM, and Least-Squares Support Vector Machine (LSSVM), were employed to investigate the prediction of K_{sat} using many soil properties, including bulk density, porosity, pH, texture, salinity, and sodium adsorption ratio. The results showed that RF was the best model to predict K_{sat}, outperforming both SVM and LSSVM based on several performance metrics.

II. MATERIALS AND METHODS

A. Study Area and Data

Soil sampling was performed in two designated areas located in the western delta of Egypt. These areas have sandy soils that exhibit varying physical and chemical properties. A total of 57 soil samples were collected from Nobaria, while 28 soil samples were obtained from Sinai. Soil samples exhibiting disturbances were collected from designated areas and sites, with hydraulic conductivity evaluated at each site using the auger-hole method [1]. The collected soil samples were analyzed in the laboratory to determine their physical and chemical characteristics, while a permeameter setup was employed to evaluate the laboratory hydraulic conductivity (K_{lab}) using Darcy's law. After analyzing the collected soil samples, they were categorized into two groups according to their chemical composition. The first group comprised samples from Nobaria and the other consisted of samples from Sinai.

All samples consisted of sandy soils, each exhibiting distinct physical and chemical properties. The Nobaria samples exhibited clay content ranging from 0.10 to 27.00%, while the Sinai samples demonstrated clay content ranging from 0.00 to 37.4%. The d_{90} values for the Nobaria samples ranged from 0.12 to 4.37 mm, while for the Sinai samples, they ranged from mm to 0.34 mm. Hydraulic conductivity in the Nobaria area ranged from 0.07 to 6.06 m/day, whereas in the Sinai samples, it ranged from 0.35 to 1.39 m/day. The Nobaria samples exhibited Electric Conductivity (EC) below 4.0 dS/m, probably due to their high salt content. The EC of the Sinai samples exceeded 37.0 dS/m, and the sodium absorption ratios exceeded 15.0, indicating that the soil was predominantly acidic. Hydraulic conductivity was measured in the laboratory utilizing a permeameter setup with disturbed soil samples obtained from two distinct areas. Tables I, II, III, and IV present summaries of laboratory analysis results for the soil samples.

The correlation coefficients between K_{field} and several soil characteristics were calculated. Intracorrelations among soil parameters were evaluated. Figures 1 and 2 illustrate the relationship between soil characteristics and K_{field} for Nobaria and Sinai. A complex link existed between K_{field} and other soil characteristics. This relationship is distinctive in each place because the soil comprises various constituents, exhibits diverse structures, and is uniquely influenced by its surroundings.

TABLE I. SUMMARY OF PHYSICAL PROPERTIES OF THE SOIL SAMPLES (NOBARIA AREA)

	K_{field} (M/d)	K_{lab} (M/d)	Sand (%)	Silt (%)	Clay (%)	d_{90} (mm)	LL (%)	PL (%)	PI (%)
Max	6.06	93.41	93.3	40.8	27.0	4.37	50.6	21.3	32.2
Min	0.07	1.90	37.2	2.2	0.1	0.12	19.3	12.2	1.7
Avg	1.05	17.76	67.7	17.4	14.9	1.35	27.9	17.1	11.0
STD	1.45	19.83	15.9	9.5	8.5	1.54	6.9	2.8	6.8

TABLE II. SUMMARY OF CHEMICAL PROPERTIES OF THE SOIL SAMPLES (NOBARIA AREA)

	pH	EC at 25°C d/m	Ca	Mg	Na	K	HCO_3	SO_4	Cl	Soil Sp (%)	Soil SAR	Soil ESP (%)	$CaCO_3$ (%)
			(meq/L)										
Max	8.1	3.4	17.0	11.9	20	0.7	4.6	34.7	11.4	74.0	8.8	11.7	54.9
Min	7.1	0.5	1.1	1.1	2.3	0.2	2.3	0.2	1.7	18.0	1.6	2.4	1.9
Avg	7.5	1.6	4.9	4.2	7.4	0.3	3.8	8.6	4.6	40.6	3.5	5.0	29.3
STD	0.3	0.9	3.4	3.3	4.5	0.1	0.6	7.8	2.9	15.2	1.7	2.2	16.6

TABLE III. SUMMARY OF PHYSICAL PROPERTIES OF THE SOIL SAMPLES (SINAI AREA)

	K_{field} (M/d)	K_{lab} (M/d)	Sand (%)	Silt (%)	Clay (%)	d_{90} (mm)	LL (%)	PL (%)	PI (%)
Max	1.4	87.9	88.0	19.3	37.4	0.3	24.5	16.5	8.0
Min	0.4	0.1	43.8	2.9	0.0	0.1	1.0	1.0	1.0
Avg	1.0	23.5	75.9	11.0	13.1	0.2	5.1	3.1	1.7
STD	0.4	24.5	11.7	6.0	11.3	0.1	8.4	5.3	1.9

TABLE IV. SUMMARY OF CHEMICAL PROPERTIES OF THE SOIL SAMPLES (SINAI AREA).

	pH	EC at 25C° d/m	Ca	Mg	Na	K	HCO_3	SO_4	Cl	Soil Sp (%)	Soil SAR	Soil ESP (%)	$CaCO_3$ (%)
			(meq/L)										
Max	7.5	122	260	543	1980	26	7.0	949	2058	43	197	74.7	13.1
Min	7.3	37	26	9	305	6	5.5	29	208	20	40	37.3	1.6
Avg	7.4	75	97	178	1033	15	6.2	212	975	31	91	54.2	3.7
STD	0.1	29	75	130	634	9	0.5	242	645	9	49	12.1	2.8

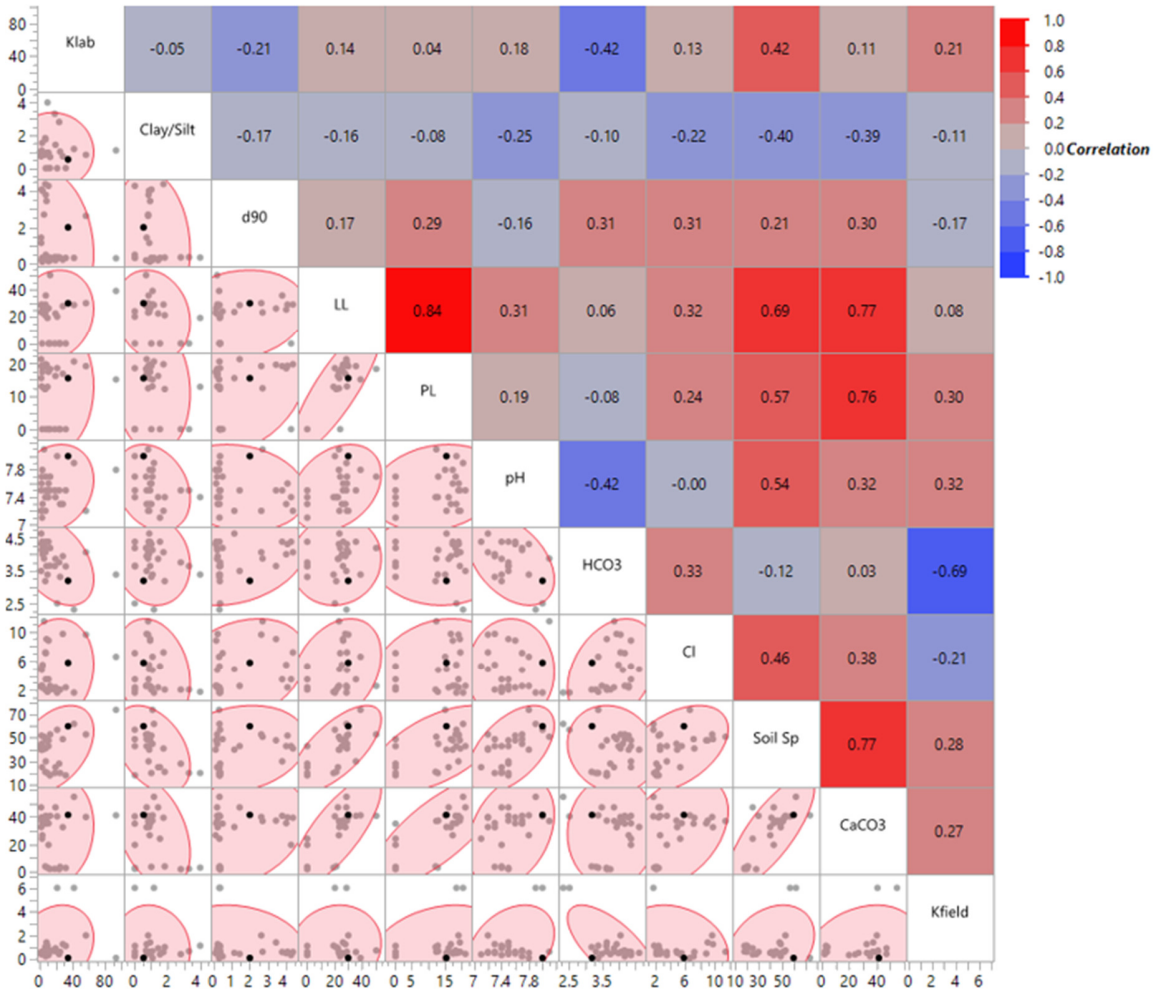

Fig. 1. Correlation matrix of soil parameters for Nobaria area.

B. Methods

ML approaches aim to identify accurate approximations that deliver a reliable, computationally efficient, and economical solution while minimizing computational time [17-19].

1) Artificial Neural Networks (ANNs)

The difficulty of the task, the size of training data, the number of predictor variables, the desired error level, and the network's purpose, specifically whether it involves pattern recognition or function approximation, are the primary factors that dictate the optimal ANN algorithm. This study utilized the MPNN and GRNN algorithms for K_{field} estimation through various network techniques.

The MPNN comprises an input layer, a minimum of one hidden layer, and an output layer, all fully interconnected. The network acquires a training dataset, calculates the errors based on the output, and iterates this procedure. The errors modify the weights and biases. This iterative process ultimately yields the optimal weight and bias values that most accurately replicate the real scenario. The conjugate gradient approach is employed in learning to adjust weight values by utilizing the gradient to propagate errors back through the network. Sigmoid and linear transfer functions were utilized as the activation functions for the hidden and output layers.

The GRNN method was introduced as a variant of the RBFNN. This model is an efficient artificial neural network that offers a viable solution for any approximation issue. Furthermore, it is a feed-forward ANN that relies on non-linear regression learning [20, 21]. The GRNN comprises four layers: input, hidden, pattern/summation, and decision layers. The learning process resembles the formation of a multidimensional surface in space that offers a statistically optimal approximation for the dataset. As the training set of the GRNN model increases in size, the prediction error converges towards zero. Furthermore, the Gaussian kernel function is employed, with the sigma value (σ) dictating the dispersion of the function.

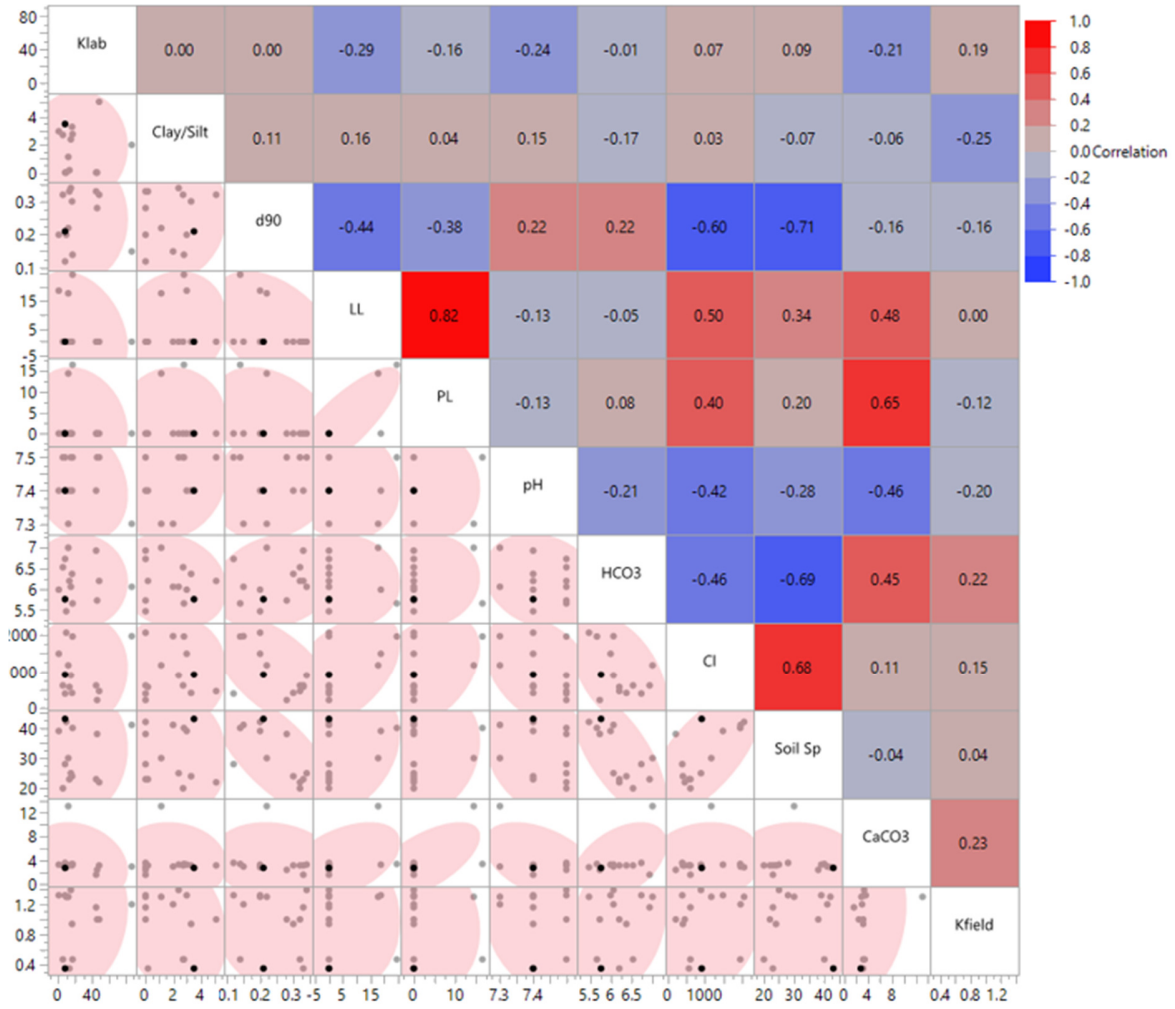

Fig. 2. Correlation matrix of soil parameters for Sinai area.

2) Tree Boost Method

The tree boost method belongs to the family of boosted regression tree modeling approaches. Tree boost improves the decision tree algorithm by incorporating a boosting mechanism. The core idea in a single robust consensus optimal model is to integrate a set of weak models to generate a resilient ensemble model. The tree boost technique sequentially builds new decision trees by minimizing the residuals of existing ones. The process of developing this sequential model constitutes a variant of the gradient descent function. This method can use both qualitative and quantitative variables in the regression analysis. Additionally, it can manage connected predictive factors and missing data. Moreover, it is deemed robust against the presence of outliers in the dataset and the inclusion of extraneous predictor factors. The tree boost model typically delineates three principal parameters: the learning rate (often referred to as the shrinkage parameter), the complexity of the tree, and the number of regression trees (tree size). Cross-validation techniques are utilized to assess the generalization of the tree boost model and reduce overfitting [22].

The training subset, utilized to determine the optimal model parameters, and the validation subset were both derived from the training data. Upon completion of the training, the models were evaluated to assess their ability to generalize knowledge to previously unexamined scenarios. Approximately 70% of the dataset was randomly allocated for model training, while the remaining 30% was designated for model testing.

III. RESULTS AND DISCUSSION

The careful consideration of input and output variables is a fundamental part of the development of an ML (ANN or tree boost) model. A subset of 10 parameters (K_{lab}, Clay/Silt ratio, d_{90}, LL, PL, pH, HCO3, Cl, Soil Sp (%), and $CaCO_3$) was chosen to develop the ML-based models to predict the K_{field} of sandy soil based on basic soil properties of saline and alkaline soil data. The ML-based models were evaluated using Mean Square Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), correlation coefficient (R), and Scatter Index (SI), as seen in Table IV.

A 10-fold cross-validation was used for all models, obtaining the mean results for each test dataset. Cross-

validation was used to ensure that all instances of the dataset were utilized for both training and testing the models. A GA was utilized to optimize the size of the MPNN model, designed based on input parameters including K_{lab}, Clay/Silt ratio, d_{90}, LL, PL, pH, HCO_3, Cl, Soil Sp (%), and $CaCO_3$.

TABLE V. STATISTIC METRICS

Features	Description
MSE	$MSE = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2$
MAE	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert P_i - O_i\rvert$
RMSE	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$
SI	$SI = \frac{RMSE}{\bar{O}}$
R	$R = \frac{\sum_{i=1}^{N}(P_i-\bar{P})(o_i-\bar{o})}{\sqrt{\sum_{i=1}^{N}(P_i-\bar{P})^2 \sum_{i=1}^{N}(o_i-\bar{o})^2}}$

O_i denotes the observed value, P_i is the predicted value, N is the number of observations, $\bar{O}$ is the mean value of the observations, and $\bar{P}$ is the mean value of the predictions.

The MPNN research revealed that a hidden-layer network with 18 neurons might yield the most optimal and steady performance. The statistical results for the test data on the projected K_{field} were MSE of 0.136, MAE of 0.257, RMSE of 0.368, and SI of 0.256. The RMSE and SI for the training data were 23.49% and 20.75% lower than those for the test data, respectively. Figures 3 and 4 show the correlations between the observed and predicted data using the different ML models for both the training and the test datasets. The observed K_{field} values were compared with the predictions made by the MPNN, GRNN, and tree boost models for both training and test datasets (Figures 5 and 6). For the GRNN model, the RMSE and SI values for the training data were 20.39% and 17.68% lower than those for the test data, respectively. The statistical results for the test data on the projected K_{field} were MSE of 0.094, MAE of 0.212, RMSE of 0.307, and SI of 0.213, demonstrating that the GRNN model is a dependable predictor of K_{field}.

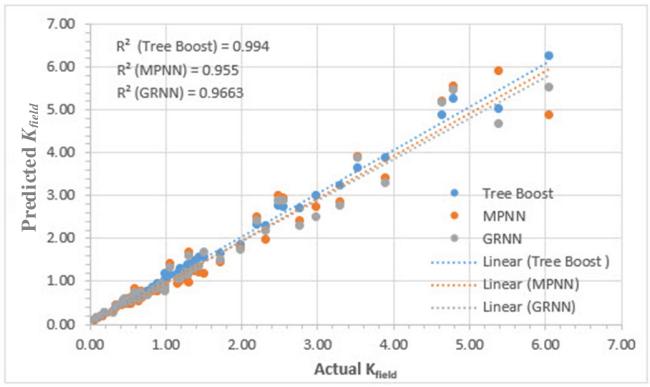

Fig. 3. Scatter plot of the measured and predicated K_{field} for MPNN, GRNN, and tree boost models for training data.

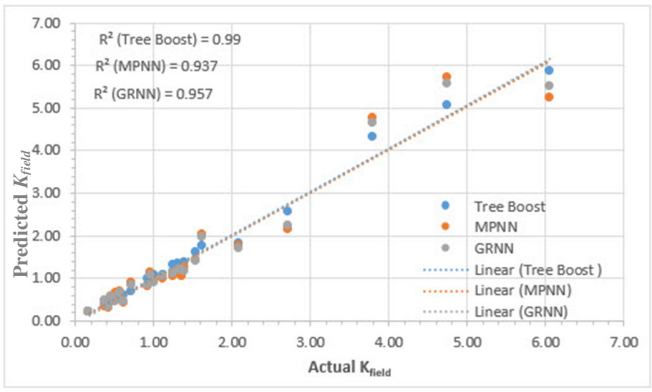

Fig. 4. Scatter plot of the measured and predicated K_{field} for MPNN, GRNN, and tree boost models for test data.

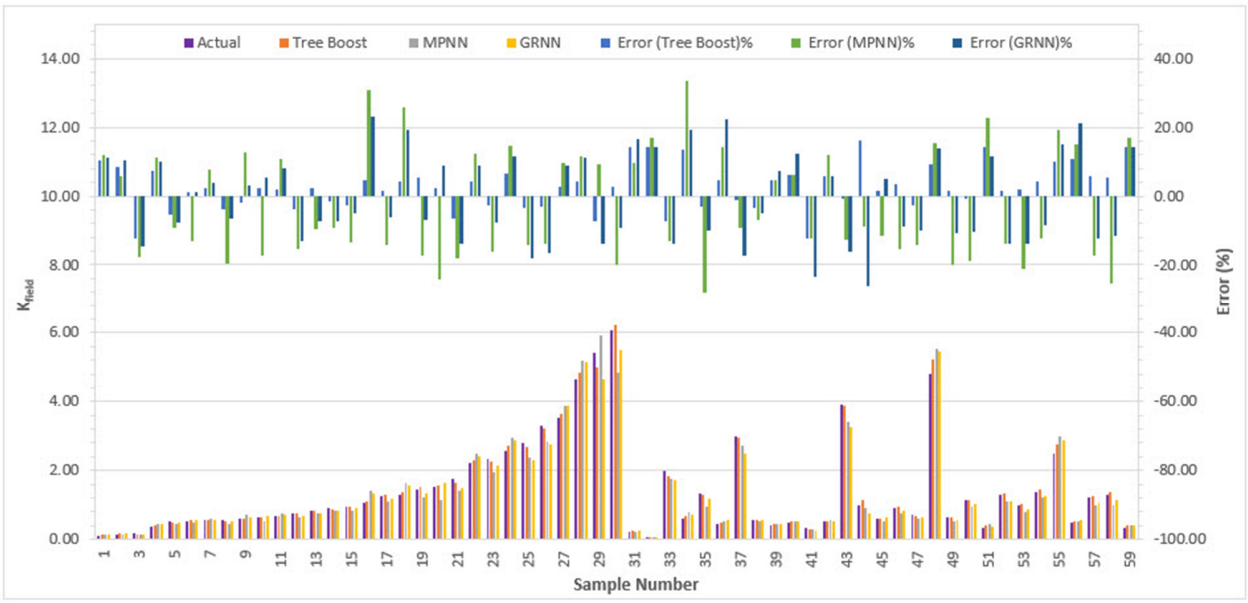

Fig. 5. Comparison of actual and predicted values of K_{field} for MPNN, GRNN, and tree boost models for training data.

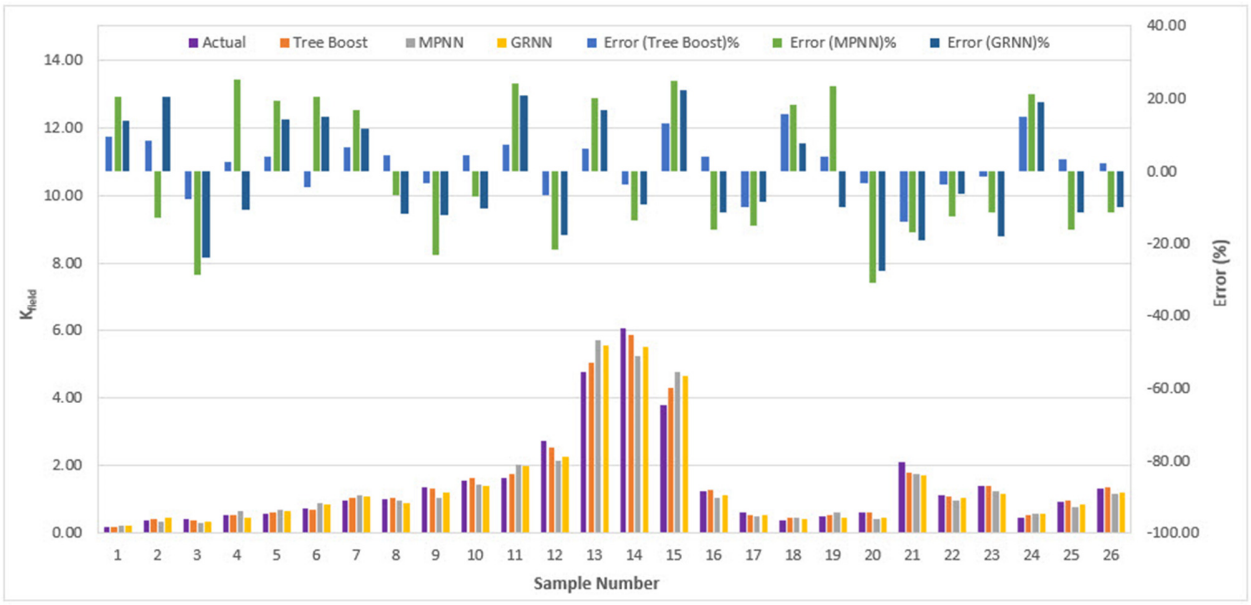

Fig. 6. Comparison of actual and predicted values of K_{field} for MPNN, GRNN, and tree boost models for test data.

The application of the tree boost model to the training data showed an MSE of 0.012, an MAE of 0.066, an RMSE of 0.107, and an SI of 0.076. The test data yielded an MSE of 0.021, an MAE of 0.091, an RMSE of 0.145, and an SI. The fluctuation of K_{field} exhibited a consistent tendency for both experimental data and the tree boost model. Overall, it can be inferred that the tree boost outperformed the other models.

The tree boost model considerably decreased overall predicting errors, outperforming the MPNN and GRNN models in predicting K_{field}. The tree boost model provided the most accurate predictions for K_{field}, surpassing the other models according to the five performance parameters and significantly diminishing overall error, as can be seen in Table VI.

To further validate the performance of the tree boost model, it was compared with an SVM model with an RBF, as proposed in [17]. The results of the RMSE and R values for the two models demonstrate the superior performance of the tree boost model, as shown in Figure 7.

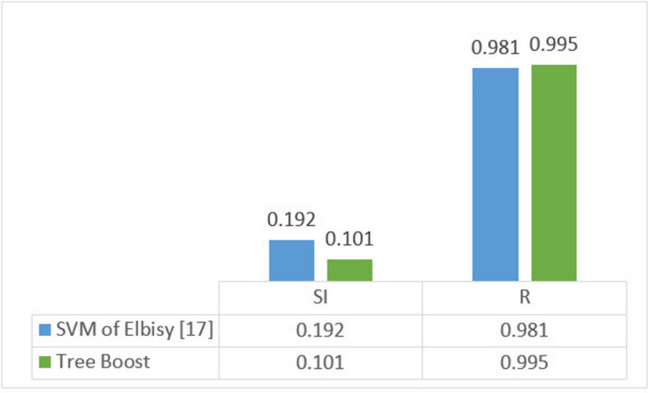

Fig. 7. Comparison of the tree boost model by SVM of Elbisy [17].

TABLE VI. ERROR STATISTICS FOR MPNN, GRNN, AND TREE BOOST MODELS.

Models		Error statistics				
		MSE	***MAE***	***RMSE***	***SI***	***R***
MPNN	Training	0.089	0.212	0.298	0.212	0.977
	Test	0.136	0.257	0.368	0.256	0.968
GRNN	Training	0.065	0.172	0.255	0.181	0.983
	Test	0.094	0.212	0.307	0.213	0.978
Tree Boost	Training	0.012	0.066	0.107	0.076	0.997
	Test	0.021	0.091	0.145	0.101	0.995

IV. CONCLUSIONS

This study developed an MPNN, a GRNN, and a tree boost model to predict the field-saturated soil hydraulic conductivity (K_{field}) of sandy soil, using basic soil properties that can be easily measured in the laboratory. The ANN and tree boost models were trained and validated using cross-validation, ensuring that all instances of the dataset were used. Using five performance measures, the predictive methods were comprehensively compared, drawing the following conclusions.

- The tree boost model demonstrated superior accuracy compared to the MPNN and GRNN models. The SI of the tree boost was 52.58% and 60.55% inferior to that of the GRNN and MPNN, respectively, for the test data.
- The tree boost had the highest predictive capability for K_{field}, succeeded by GRNN and MPNN. The proposed tree boost model is both efficient and accurate, making it a viable tool for K_{field} prediction.
- The tree boost model serves as a valuable adjunct to traditional forecasting methods due to its superior predictive capabilities and ease of configuration.

REFERENCES

[1] R. J. Oosterbaan and H. J. Nijland, "Determining the saturated hydraulic conductivity.," in *Drainage Principles and Applications*, Wageningen, The Netherlands: International Institute for Land Reclamation and Improvement (ILRI), 1994.

[2] P. J. Dielman and B. D. Trafford, "Drainage testing: Irrigation and drainage," *Food and Agriculture Organization (FAO)*, 1976.

[3] W. J. Bentley, R. W. Skaggs, and J. E. Parsons, *The effect of variation in hydraulic conductivity on water table drawdown*. North Carolina Agricultural Research Service, North Carolina State University, 1989.

[4] R. S. Ayers and D. W. Westcot, "Water quality for agriculture," *Food and Agriculture Organization (FAO)*, 1985.

[5] V. S. Aronovici, "The mechanical analysis as an index of subsoil permeability.," in *Proceedings - Soil Science Society of America*, 1946.

[6] D. W. Rycroft and L. K. Smedema, *Land drainage: planning and design of agricultural drainage systems*. London, UK: Batsford Academic & Educational, 1983.

[7] J. Bouma, A. Jongerius, and D. Schoonderbeek, "Calculation of Saturated Hydraulic Conductivity of Some Pedal Clay Soils Using Micromorphometric Data," *Soil Science Society of America Journal*, vol. 43, no. 2, pp. 261–264, 1979, https://doi.org/10.2136/sssaj1979.03615995004300020002x.

[8] T. J. Marshall, "Permeability and the Size Distribution of Pores," *Nature*, vol. 180, no. 4587, pp. 664–665, Sep. 1957, https://doi.org/10.1038/180664a0.

[9] M. S. Elbisy, "Support Vector Machine and regression analysis to predict the field hydraulic conductivity of sandy soil," *KSCE Journal of Civil Engineering*, vol. 19, no. 7, pp. 2307–2316, Nov. 2015, https://doi.org/10.1007/s12205-015-0210-x.

[10] A. Sedaghat, H. Bayat, and A. A. Safari Sinegani, "Estimation of soil saturated hydraulic conductivity by artificial neural networks ensemble in smectitic soils," *Eurasian Soil Science*, vol. 49, no. 3, pp. 347–357, Mar. 2016, https://doi.org/10.1134/S106422931603008X.

[11] C. G. Williams and O. O. Ojuri, "Predictive modelling of soils' hydraulic conductivity using artificial neural network and multiple linear regression," *SN Applied Sciences*, vol. 3, no. 2, Jan. 2021, Art. no. 152, https://doi.org/10.1007/s42452-020-03974-7.

[12] M. Khalili-Maleki, R. V. Poursorkhabi, A. A. Nadiri, and R. Dabiri, "Prediction of hydraulic conductivity based on the soil grain size using supervised committee machine artificial intelligence," *Earth Science Informatics*, vol. 15, no. 4, pp. 2571–2583, Dec. 2022, https://doi.org/10.1007/s12145-022-00848-x.

[13] K. Báťková *et al.*, "Prediction of saturated hydraulic conductivity Ks of agricultural soil using pedotransfer functions," *Soil and Water Research*, vol. 18, no. 1, pp. 25–32, 2023.

[14] J. Liu, X. Wang, and X. Ren, "Hydraulic conductivity and particle size of soils: modeling and experiment," *Acta Geotechnica*, Apr. 2025, https://doi.org/10.1007/s11440-025-02581-3.

[15] A. A. Moosavi, M. A. Nematollahi, and M. Omidifard, "Comparing machine learning approaches for estimating soil saturated hydraulic conductivity," *PLOS ONE*, vol. 19, no. 11, 2024, Art. no. e0310622, https://doi.org/10.1371/journal.pone.0310622.

[16] M. Farasati, M. Seyedian, and A. Fathaabadi, "Predicting soil hydraulic conductivity using random forest, SVM, and LSSVM models," *Natural Resource Modeling*, vol. 37, no. 4, 2024, Art. no. e12407, https://doi.org/10.1111/nrm.12407.

[17] M. S. Elbisy, "Predictive Modeling of Saturated Hydraulic Conductivity using Machine Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21348–21355, Apr. 2025, https://doi.org/10.48084/etasr.10225.

[18] S. Vimalkumar and R. Latha, "Advanced Soil Moisture Predictive Methodology in the Maize Cultivation Region," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19966–19970, Feb. 2025, https://doi.org/10.48084/etasr.9059.

[19] A. S. Kote and D. V. Wadkar, "Modeling of Chlorine and Coagulant Dose in a Water Treatment Plant by Artificial Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 9, no. 3, pp. 4176–4181, Jun. 2019, https://doi.org/10.48084/etasr.2725.

[20] D. F. Specht, "A general regression neural network," *IEEE transactions on neural networks*, vol. 2, no. 6, pp. 568–576, Jan. 1991, https://doi.org/10.1109/72.97934.

[21] J. Song, C. E. Romero, Z. Yao, and B. He, "A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler," *Knowledge-Based Systems*, vol. 118, pp. 4–14, Feb. 2017, https://doi.org/10.1016/j.knosys.2016.11.003.

[22] J. H. Friedman and J. J. Meulman, "Multiple additive regression trees with application in epidemiology," *Statistics in Medicine*, vol. 22, no. 9, pp. 1365–1381, 2003, https://doi.org/10.1002/sim.1501.