Water Practice & Technology

IWA PUBLISHING

Water Practice & Technology Vol 17 No 12, 2625 doi: 10.2166/wpt.2022.151

Estimation of hydraulic conductivity of porous media using data-driven techniques

Divya Thakur *, Abhishish Chandel and Vijay Shankar

National Institute of Technology, Hamirpur, India
*Corresponding author. E-mail: thakurdivya976@gmail.com

DT, 0000-0002-9578-8704

ABSTRACT

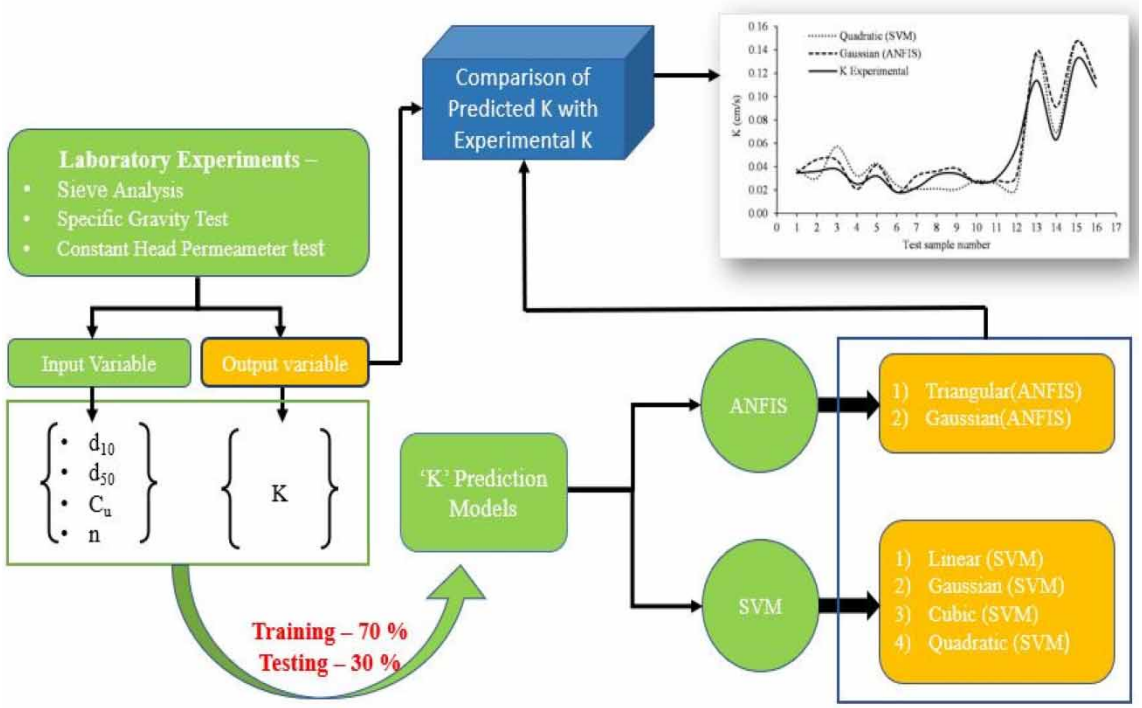
Knowledge of hydraulic conductivity (*K*) is inevitable for sub-surface flow and aquifer studies. Hydrologists and groundwater researchers are employing data-driven techniques to indirectly evaluate *K* using porous media characteristics as an alternative to direct measurement. The study examines the ability of the Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict the *K* of porous media using two membership functions (MFs), i.e., triangular and Gaussian, and support vector machine (SVM) via four kernel functions, i.e., linear, quadratic, cubic, and Gaussian. The techniques used easily measurable parameters namely effective and mean grain size, uniformity coefficient, and porosity as input variables. A 70 and 30% dataset is used for the training and testing of models, respectively. The correlation coefficient (*R*) and root mean square error (RMSE) were used to evaluate the models. The Gaussian MF-based ANFIS model outperformed the triangular model having *R* and RMSE values of 0.9661 & 0.0010 and 0.9532 & 0.0015, respectively, whereas the quadratic kernel-based SVM model with *R* and RMSE values of 0.9520 and 0.0015 performs better than the other SVM models. Based on the evaluation of ANFIS and SVM models, the study establishes the efficacy of the Gaussian MF-based ANFIS model in estimating the *K* of porous media.

Key words: grain size, hydraulic conductivity, porosity, SVM

HIGHLIGHTS

- The study focuses to develop a hydraulic conductivity model based on easily measurable grain-size parameters using the two data-driven techniques, i.e., ANFIS and support vector machine.
- The data-driven approaches result in the quick estimation of hydraulic conductivity, which is useful in determining the groundwater recharge with precise accuracy.

GRAPHICAL ABSTRACT



INTRODUCTION

Hydraulic conductivity (K) is a simple, yet critical porous media parameter that directs the flow of fluids through porous media using Darcy's law. The K of porous media is reliant on the physical properties of the flowing fluid and the transmission medium such as particle size, porosity, and pore connectivity (Chandel *et al.* 2021). Darcy's law directly connects the seepage of fluid to the K and its knowledge is important for groundwater recharge, landslide, and soil stability analysis. Soil's K values can be measured or predicted using direct and indirect approaches (Chapuis 2012; Elbisy 2015). Various efforts have been invested in estimating the K due to the complex geometry of the soil particles and the multiscale pore structure of porous media (Jougnot *et al.* 2021).

The direct approach consists of the K measurement in the laboratory or field. Laboratory methods include constant-head and falling-head tests, while the field methods include ring infiltrometer, instant profile, test basins (Williams & Ojuri 2021; Chandel *et al.* 2022a), auger hole, tension infiltrometer (Raoof *et al.* 2011), borehole permeameter, and pressure infiltrometer (Deb & Shukla 2012). A detailed review of various laboratory and field methods for determining the K of soil is covered by Amoozegar & Warrick (1986) and Wang *et al.* (2017). Direct measurement is expensive and time-consuming and becomes infeasible due to temporal and geographical variations. It also needs sophisticated instruments and competent operators (Arshad *et al.* 2013; Chandel & Shankar 2022). This resulted in the invention and widespread use of indirect techniques for estimating the K from more easily and inexpensively measured soil parameters such as porosity, specific gravity, and the percentage of sand, silt, and clay content (More *et al.* 2022). The indirect approaches involve the estimation of K using empirical equations and models developed using data-driven techniques based on correlation analysis, which relates the K to relevant contributing factors. The existing empirical methods have the significant benefit of being able to estimate the K value more quickly than the direct measurement (Williams & Ojuri 2021; Chandel *et al.* 2022b). As a result of their domain-specific development, empirical equations cannot be applied outside of those boundary conditions (Chandel & Shankar 2021).

Recent research shows the effective use of data-driven techniques in different fields of engineering (Das *et al.* 2012). Learning, modelling, and obtaining a pattern from an experimental approach are all complex tasks that data-driven techniques can perform with precise accuracy. The data-driven techniques can either be single such as artificial neural networks (ANNs), fuzzy, and support vector machine (SVM), or hybrids such as ANFIS and

genetic algorithm-ANN (Williams & Ojuri 2021). Particle swarm optimization and Monte Carlo analysis can also be used for model optimization and uncertainty analysis (Torabi Podeh *et al.* 2022). These techniques have been effectively practiced in water resource engineering challenges (Ghorbani *et al.* 2016; Kumar *et al.* 2020) and have a better prediction efficiency as compared to the direct approaches (Sihag *et al.* 2021). Ekhmaj (2010) developed the ANN and the multilinear regression (MLR) model to estimate the infiltration rate for different types of Libyan soil and suggested that ANN has better prediction efficiency than MLR. To estimate the K value for clay liners, Das *et al.* (2012) found that the SVM model was more effective than the neural network (NN) model. Yilmaz *et al.* (2012) used ANNs and ANFIS to estimate the K of coarse-grained soils and found the ANFIS model to be more reliable. Arshad *et al.* (2013) predicted the K using multi-layer perceptron neural networks (MLPNNs), MLR, ANFIS, and radial basis function neural network (RBFNN) models and found that ANFIS and RBFNN were efficient approaches for K prediction with precise accuracy. Elbisy (2015) studied SVM for estimating the K of sandy soil and found that SVM based on the radial basis function (RBF) model had a higher level of accuracy when compared to the linear and sigmoid-based models. Sihag (2018) developed a fuzzy logic and ANN-based model for estimating the K of soil, and the results showed that the ANN approach performed well. Sihag *et al.* (2018) used ANN, SVM, Gaussian process (GP), random forest (RF), M5P model tree, and two conventional models to estimate the infiltration rate of fly-ash-mixed soils, and the result showed that SVM with RBF kernel was the best-fit modelling technique among others. Boadu (2020) explored support vector regression (SVR) and MLR to predict K and found that the SVR models were more accurate and achieve better performance. Singh *et al.* (2021a) used SVM, RF, and MLR models to estimate the K of the soil and concluded that the SVM model has better prediction efficiency. Singh *et al.* (2021a) investigated ANN, MLR, RF, and M5P tree-based models to predict the infiltration rate. The results showed that all models have relatively good prediction capability, with RF-based models outperforming the others.

Several data-driven techniques were found to be useful in the literature review for estimating the K of porous media using easily measurable soil properties; however, ANFIS and SVM showed impressive prediction performance in comparison to other techniques in the respective studies. The Takagi–Sugeno fuzzy inference system is used in ANFIS, and it has been proven to be an effective tool for capturing non-linear relationships between inputs and outputs (Nawaz *et al.* 2015). SVMs reduce overfitting as they are based on the principle of structural risk optimization and statistical learning algorithm, and they have gained more popularity because of better generalization (Das *et al.* 2012; Elbisy 2015). To the best of our knowledge, the ANFIS and SVM techniques have not been used together to examine which technique is more effective in estimating the K of porous media. Based on the literature review and the potential of ANFIS and SVM, the current study is conducted to examine and compare the potency of both techniques in developing prediction models for estimating the K of porous media. The main objectives of the study are given as follows:

1. To estimate the hydraulic conductivity of porous media by developing the ANFIS and SVM models and evaluate their efficacy via statistical indicators, i.e., R and RMSE.
2. To compare the outcomes of the ANFIS and SVM models in terms of prediction performance capability.

Theoretical overview of ANFIS and SVM

Adaptive Neuro-Fuzzy Inference System

The fuzzy theory was developed by Zadeh in 1962, and ANFIS is the most convenient neuro-fuzzy model established by Jang in 1993 (Azad *et al.* 2018). ANFIS makes utilization of neural network learning and fuzzy logic reasoning capabilities. It produces a fuzzy inference system (FIS) based on the input–output dataset (Nawaz *et al.* 2015). Each fuzzy rule in the FIS specifies a local system behaviour, and the network structure implements the FIS and uses hybrid learning rules to train the model. ANFIS's five-layer architecture is depicted in Figure 1, and the layers and nodes are described below (Jalalkamali 2015; Nawaz *et al.* 2015).

Layer 1 – Every single node produces membership grades of an input variable. The shapes of the membership functions, such as generalized bell-shaped, Gaussian, triangle-shaped, and trapezoidal-shaped functions, are used to categorize the fuzzy set associated with each node.

Layer 2 – Every node is a fixed node labelled as $\prod$ and the output of every node is the product of the incoming signal, which represents the firing strength of a rule.

Layer 3 – Every node is a fixed node labelled as N and computes the normalized firing strengths.

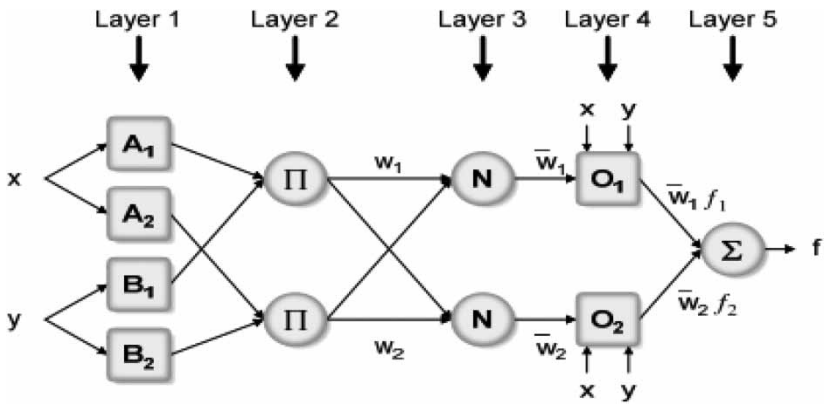


Figure 1 | Basic structure of ANFIS.

Layer 4 – Every node i is an adaptive node and computes the input of the ith rule for the model output with node function.

Layer 5 – Computes the overall output of the ANFIS by summing outputs of all incoming signals.

Support vector machine

SVM was developed by Vapnik, and is based on statistical learning theory. In support of vector regression, the input x is projected into a high-dimensional feature space using a kernel function, and then a linear model is developed in this feature space (Das *et al.* 2012).

Consider a training set (x_i, y_i), $(i = 1, 2, 3, \ldots, n)$. The linear model $f(x; w)$ in the feature space is given by

$$f(x) = (w \cdot \varphi(x)) + b \tag{1}$$

where x is the input vector and y is the output value

$\varphi(x)$ is the non-linear mapping function that transforms the original input space to a higher dimensional feature space, w is the weight vector of the regression function, b is the bias of the regression function.

The coefficients w and b are estimated by minimizing the regularized risk function (R_{reg}).

$$R_{reg} = \frac{1}{2}\|w\|^2 + \frac{C}{n}\sum_{i=1}^{n} L_\varepsilon(f(x), y) \tag{2}$$

where

n is the number of patterns that contain all the information necessary to solve the learning task at hand hereinafter refers to as support vectors.

C is a regularization constant that determines the trade-off between the training error and the generalization performance.

$\frac{1}{2}\|w\|^2$ = regularization term measures the flatness of the function $L_\varepsilon(f(x), y)$ (Elbisy 2015; Boadu 2020).

The quality of prediction is measured by the ε-insensitive loss function $L_\varepsilon(f(x), y)$.

$$L_\varepsilon(f(x), y) = \begin{Bmatrix} 0 & \text{if } |f(x) - y| \le \varepsilon \\ |f(x) - y| - \varepsilon & \text{if } |f(x) - y| > \varepsilon \end{Bmatrix} \tag{3}$$

where ε is a precision parameter representing the radius of the tube located around the regression function.

Using positive slack variables ξ and ξ^* in Equation (2), the optimization problem can be transformed into a quadratic programming problem. Finally, the regression function of the SVM can be expressed as follows:

$$f(x) = \sum_{i=1}^{m} y_i(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{4}$$

Subjected to$\{0 \le \alpha_i^* \le C\}$and $\{0 \le \alpha_i \le C\}$

where α^* and α are Lagrange multipliers and $K(x_i, x) = \varphi(x_i)\varphi(x)$ is called the kernel function (Elbisy 2015). The SVM technique was discussed in detail by Vapnik (1999), Samui *et al.* (2008), and Das *et al.* (2012).

MATERIALS AND METHODOLOGY

Experimental study

In the first phase of the study, experiments have been performed in the Hydraulic Laboratory located at NIT Hamirpur, which includes the sieve analysis, specific gravity, and constant-head permeameter test. The samples were collected from the Ravi riverbank (32°33′10.3″ N, 76°07′19.2″ E) nearby the Irrigation and Public Health Department Chamba of Himachal Pradesh in India. A total number of 56 porous media samples having different proportions of fines (diameter <0.075 mm), fine sand (0.075–0.425 mm), medium sand (0.425–2 mm), coarse sand (2–4.75 mm), and fine gravel (4.75–20 mm) were used for measurement. Soil samples were sieved according to the American Society for Testing and Materials standards (ASTM 2007). From the grain-size curve, the parameters namely d_{10}, d_{50}, and d_{60} (grain diameter at which 10, 50, and 60% particles are finer) and C_u (uniformity coefficient) were determined. To determine the porosity (n) of the sample, the specific gravity test was carried out using the pycnometer method (Chandel & Shankar 2021). Figure 2 depicts the usage of a constant-head permeameter with a diameter of 10.16 cm and total and test lengths of 106 and 46.5 cm, respectively, for the determination of K. To observe the head difference between the start and endpoint of the soil sample, four pressure taping points at an angle of 90° were provided along the circumference of the permeameter. The K was calculated by dividing the flow rate (cm^3/s) by the permeameter area (cm^2) times the hydraulic gradient (ASTM 2006; Alabi 2011). At the beginning and end of each test run, the water temperature was measured digitally using a thermometer.

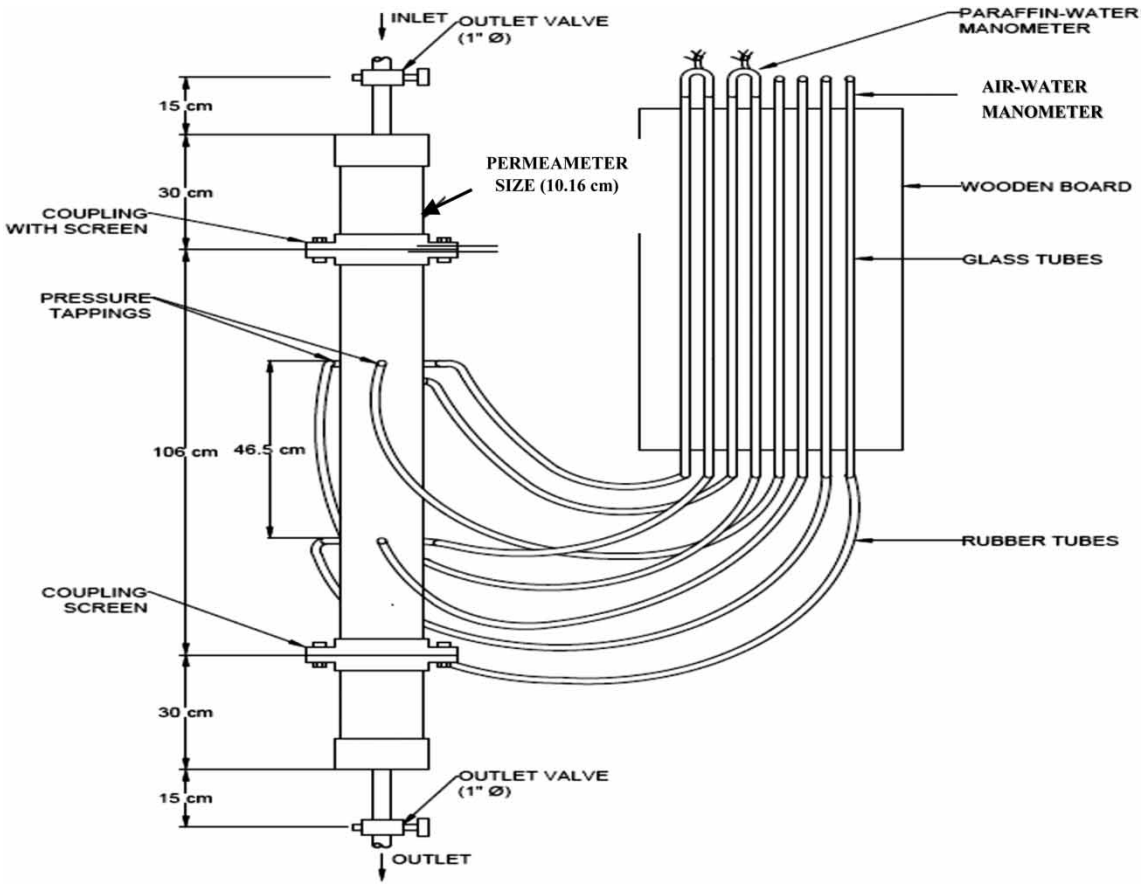


Figure 2 | Experimental set-up of constant-head permeameter.

Model development

The second phase of the study used the observations obtained from the experimental investigations for estimating the K of porous media using ANFIS and SVM techniques. The agreement between the predicted and observed K values was checked statistically by calculating R and RMSE. The modelling strategies used and performance evaluation parameters are summarized in the following section.

ANFIS modelling strategy

Developing the ANFIS model is a trial-and-error method. Estimation using the ANFIS model includes defining the number and shape of MF. In this study, two numbers of MF for each parameter and two shapes (triangular and Gaussian) of the MFs for input and output parameters were selected for generating the FIS. The training parameters were optimized using the hybrid optimization algorithm. The ANFIS modelling was performed using the neuro-fuzzy designer toolbox in MATLAB R2021a according to the procedure as shown in Figure 3.

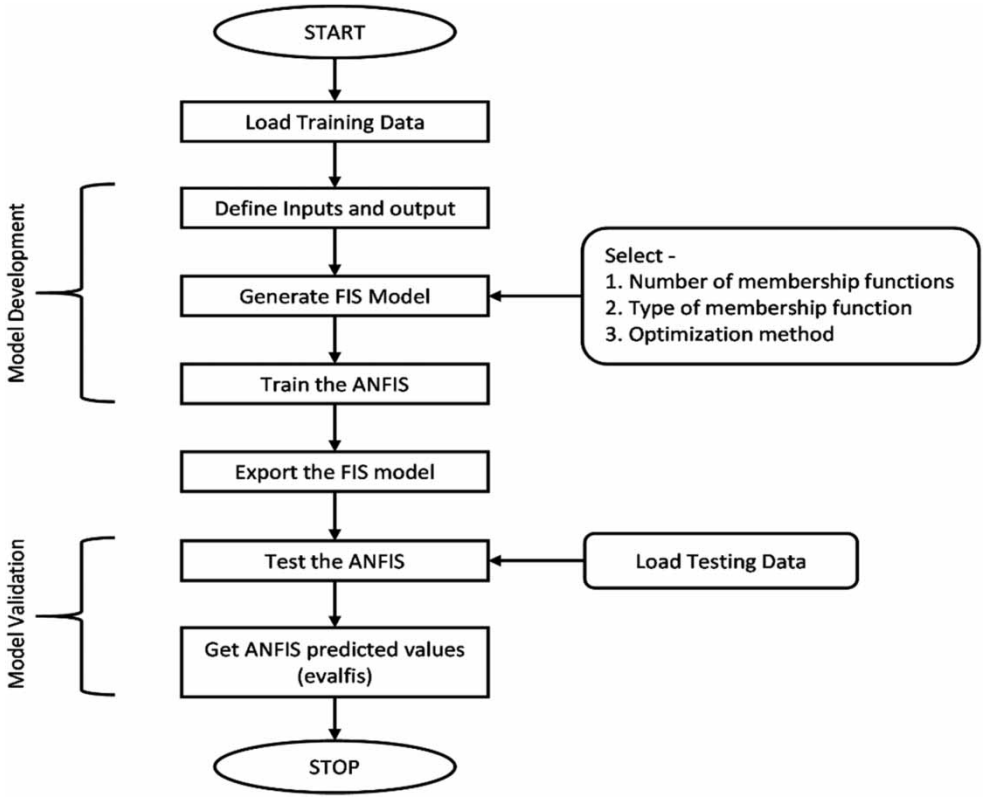


Figure 3 | Flowchart representing the ANFIS modelling approach.

SVM modelling strategy

The SVM regression models with linear, quadratic, cubic, and Gaussian kernel functions were developed and validated using the regression learner toolbox of MATLAB R2021a. The 5-fold cross-validation of the development dataset was done to protect the model from overfitting. The procedure implemented for SVM modelling is shown in Figure 4.

Model performance evaluation

To analyse the competence of ANFIS and SVM models in estimating the K of the porous media, R and RMSE values were calculated. Values closer to 1 for R and lower values close to 0 for RMSE indicate better agreement between the observed and predicted values (Chandel & Shankar 2021). The model performance evaluation

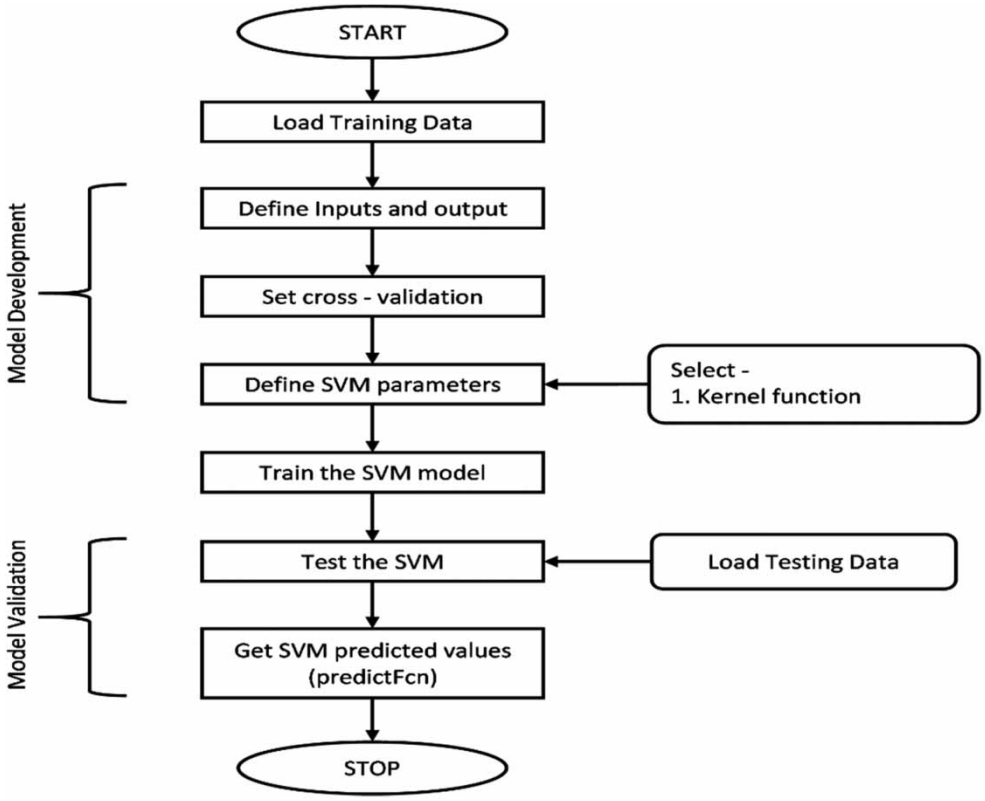


Figure 4 | Flowchart representing the SVM modelling approach.

parameters are defined as follows:

$$R = \frac{\sum_{i=1}^{N}(P_i - \bar{P})(E_i - \bar{E})}{\sqrt{\sum_{i=1}^{N}(P_i - \bar{P})^2 \sum_{i=1}^{N}(E_i - \bar{E})^2}} \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - E_i)^2} \tag{6}$$

where N is the number of samples;

E_i is the experimentally observed K value; P_i is the model-predicted K value;
$\bar{E}$ is the experimental observation mean value; $\bar{P}$ is the model prediction mean value (Elbisy 2015).

RESULTS AND DISCUSSION

The observations of the laboratory experiments and findings of the developed ANFIS and SVM models to estimate the K have been analysed and discussed in the following section. The first section describes the statistical analyses of the porous media dataset, and the second section presents the performance analysis of the developed models.

Statistical analysis

The porous media properties obtained in the laboratory include fines (%), fine sand (%), medium sand (%), coarse sand (%), fine gravel (%), d_{10}, d_{50}, d_{60}, C_u, n, and K. The statistical indices such as maximum, minimum, mean,

and standard deviation values of each porous media property are presented in Table 1. The observed K values for 56 porous media samples have maximum and minimum values of 0.315 and 0.006 cm/s, respectively. The correlation coefficient of d_{10}, d_{50}, C_u, and n with the K is presented in Table 2. From Table 2, it is evident that d_{10} and n have a more significant influence on K, whereas d_{50} and C_u have a moderate influence on K with correlation coefficients of 0.94 & 0.77 and 0.58 & 0.52, respectively.

Table 1 | Physical properties of the samples

Property	Maximum	Minimum	Mean	Standard deviation
Fines (%)	3.256	0.000	0.479	0.745
Fine sand (%)	91.234	11.354	44.662	22.302
Medium sand (%)	88.646	8.579	49.250	20.614
Coarse sand (%)	13.529	0.000	4.046	4.763
Fine gravel (%)	20.944	0.000	1.563	4.689
d_{10} (mm)	0.410	0.090	0.213	0.073
d_{50} (mm)	1.184	0.249	0.472	0.191
d_{60} (mm)	1.906	0.273	0.566	0.315
C_u[a]	5.580	1.793	2.883	0.806
n[a]	0.427	0.316	0.377	0.030
K (cm/s)	0.315	0.006	0.071	0.065

[a]The dimensionless properties.

Table 2 | Correlation coefficient of various soil parameters with K

	d_{10}	d_{50}	C_u	n	K
d_{10}	1.00				
d_{50}	0.52	1.00			
C_u	0.36	0.83	1.00		
n	0.77	0.51	0.33	1.00	
K	0.94	0.58	0.52	0.77	1

Dataset for ANFIS and SVM

The success of the model to estimate the K of the porous media depends upon the degree of the training dataset. Out of 56 sample datasets, 40 were chosen for training, while the remaining 16 were used for testing the model. Input variables include d_{10}, d_{50}, C_u, and n, whereas K was considered as the output variable. Table 3 shows the range of different input and output parameters used in ANFIS and SVM techniques.

Table 3 | Characteristics of training and testing data

	Training data			Testing data		
	Maximum	Minimum	Mean	Maximum	Minimum	Mean
Input variable						
d_{10} (mm)	0.410	0.090	0.218	0.296	0.136	0.200
d_{50} (mm)	1.184	0.249	0.503	0.871	0.300	0.395
C_u[a]	5.580	1.793	3.033	4.261	1.946	2.508
n[a]	0.427	0.316	0.381	0.415	0.321	0.367
Output variable						
K (cm/s)	0.315	0.006	0.080	0.132	0.018	0.050

[a]The dimensionless properties.

Estimation of *K* using ANFIS

For the study, the structure of the ANFIS model used with two numbers of MF is shown in Figure 5. The results of the performance evaluation parameters of the ANFIS model for both training and testing datasets are listed in Table 4. Notably, the triangular (ANFIS) and Gaussian (ANFIS) are associated with the results of the triangular MF and Gaussian MF-based ANFIS models, respectively. From Table 4, the Gaussian (ANFIS) model outperforms the triangular (ANFIS) model, with *R* and RMSE values of 0.9884 & 0.0002 and 0.9661 & 0.0010 for the training and testing dataset, respectively. Scatter plots of experimentally observed and model-predicted *K* values for training and testing datasets are shown in Figures 6 and 7 respectively. The performance of ANFIS was satisfactory during training and testing, which infers that the model constructed via ANFIS can effectively predict the *K* of porous media. Furthermore, the scatter plot of the testing dataset as shown in Figures 6(b)

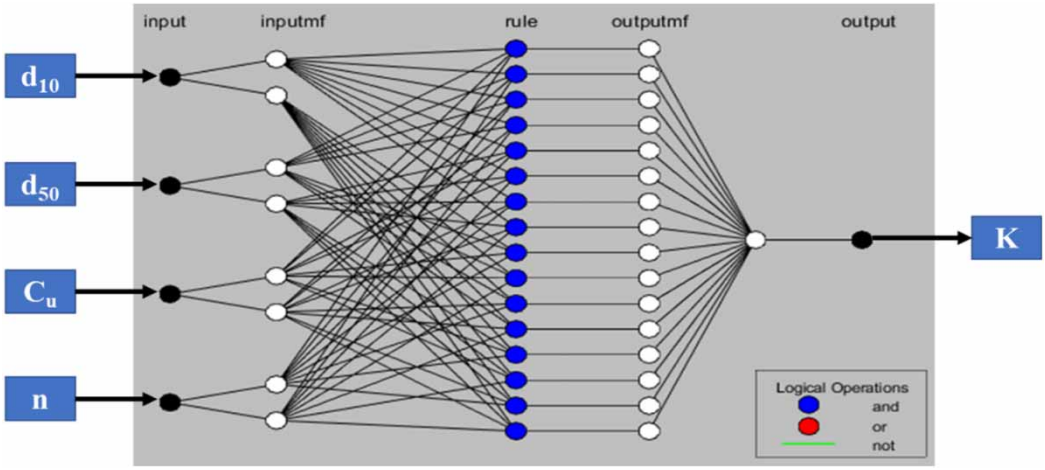


Figure 5 | Structure of the ANFIS MF2 model for *K* prediction.

Table 4 | Performance evaluation parameter of training and testing datasets for the ANFIS model

		Training		Testing	
Studied model	**No. of MF**	***R***	**RMSE**	***R***	**RMSE**
Triangular (ANFIS)	2	0.9726	0.0007	0.9532	0.0015
Gaussian (ANFIS)	2	0.9884	0.0002	0.9661	0.0010

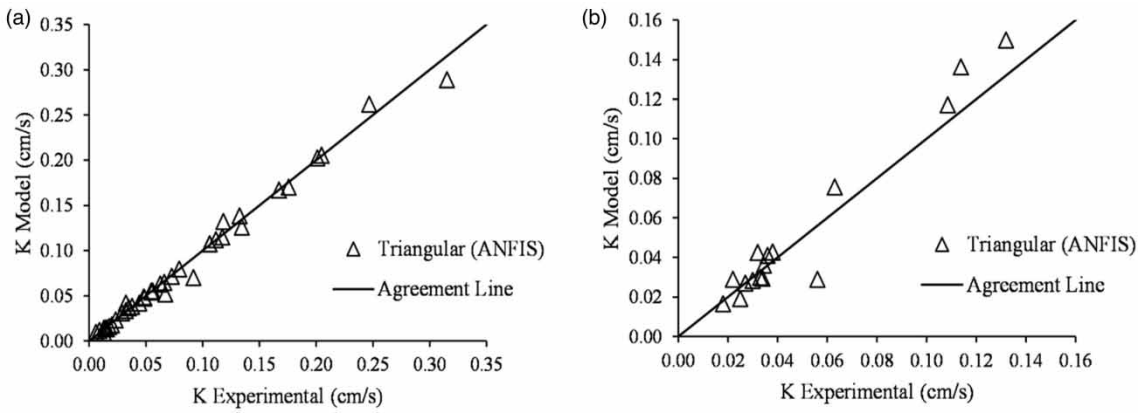


Figure 6 | Scatter plots of triangular (ANFIS) model for predicting *K* during (a) training and (b) testing stages.

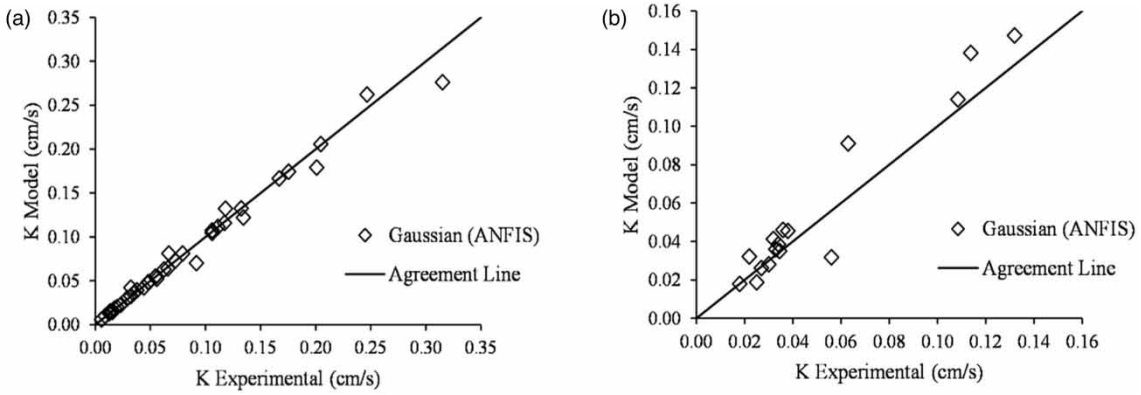


Figure 7 | Scatter plots of Gaussian (ANFIS) model for predicting *K* during (a) training and (b) testing stages.

and 7(b) shows that the points are concentrated close to the agreement line for lower values of *K*, which implies superior prediction performance, but for higher *K* values the prediction capability was relatively decent (More *et al.* 2022). ANFIS model's predictions are influenced by what they learn during training. The more the data in a given input range, the better the performance. More *K* values in the (0.006–0.07) range appeared in the training dataset for the ANFIS model, which may have contributed to the model's superior performance in the lower region during testing.

Estimation of *K* using SVM

Results of the performance evaluation parameters of the SVM model based on different kernel functions for training and testing datasets have been provided in Table 5. Notably, the linear (SVM), Gaussian (SVM), cubic (SVM), and quadratic (SVM) are associated with the results of the linear, Gaussian, cubic, and quadratic kernel function-based SVM models, respectively. The linear (SVM) and cubic (SVM) models performed well during training with *R* and RMSE values of 0.9741 & 0.0008 and 0.9724 & 0.0011, respectively, but performed poorly during testing with *R* and RMSE values of 0.9296 & 0.0021 and 0.9352 & 0.0028, respectively. The performance of the quadratic (SVM) and Gaussian (SVM) models was better during both stages. The comparison of the performance evaluation parameters shows that the quadratic (SVM) model has better prediction capability with *R* and RMSE values of 0.9520 and 0.0015, respectively, during testing. The analysis of the results shows that the non-linear kernel function has a better prediction efficiency of *K* than the linear kernel. Scatter plots of experimentally observed and model-predicted *K* values for training and testing datasets are shown in Figures 8–11. The figures show that all the kernel function-based models give significant prediction performances for lower values of *K* as the points were concentrated close to the agreement line, but, for higher *K* values, the prediction capability was relatively decent (More *et al.* 2022).

Table 5 | Performance evaluation parameter of training and testing datasets for the SVM model

	Training		Testing	
Studied model	***R***	**RMSE**	***R***	**RMSE**
Linear (SVM)	0.9741	0.0008	0.9296	0.0021
Gaussian (SVM)	0.9781	0.0006	0.9404	0.0016
Cubic (SVM)	0.9724	0.0011	0.9352	0.0028
Quadratic (SVM)	0.9831	0.0004	0.9520	0.0015

Comparison of models

The results of the ANFIS and SVM models were compared to determine the best-performing model for estimating the *K* of the porous media. Based on the statistical indicators, i.e., *R* and RMSE as shown in Tables 4 and 5, the

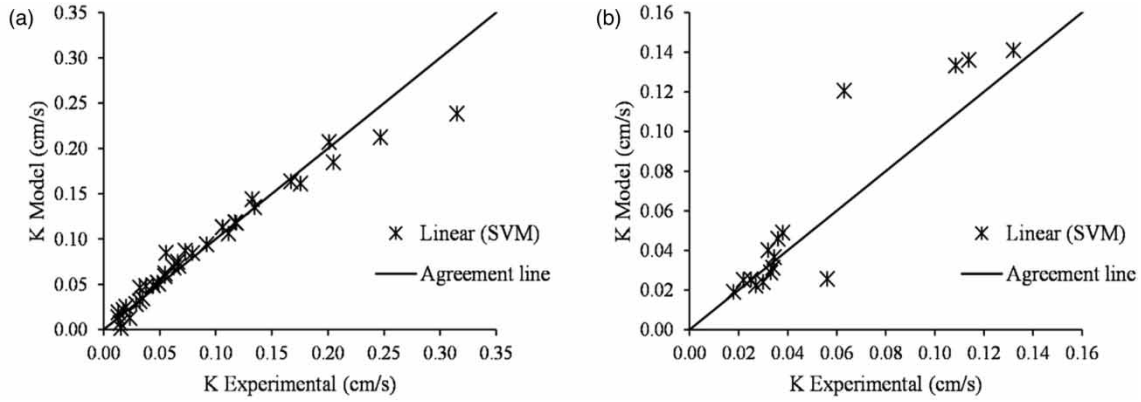


Figure 8 | Scatter plots of linear (SVM) model for predicting *K* during (a) training and (b) testing stages.

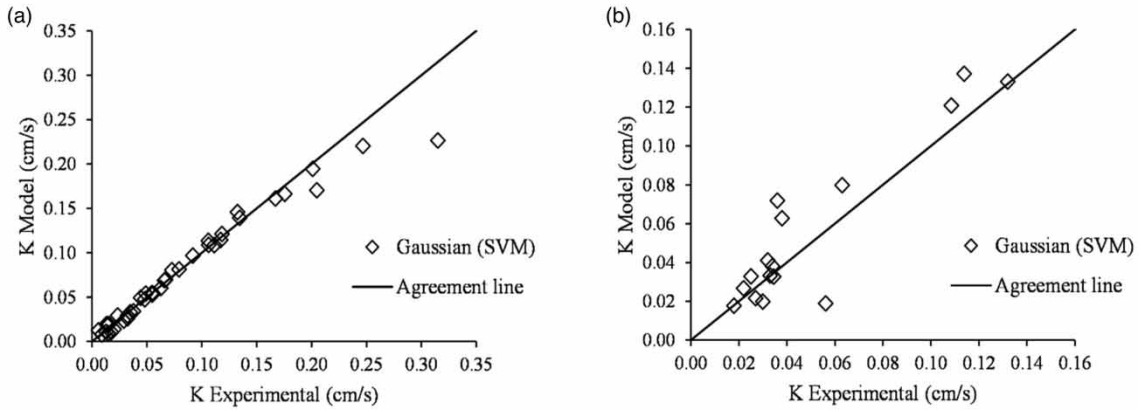


Figure 9 | Scatter plots of Gaussian (SVM) model for predicting *K* during (a) training and (b) testing stages.

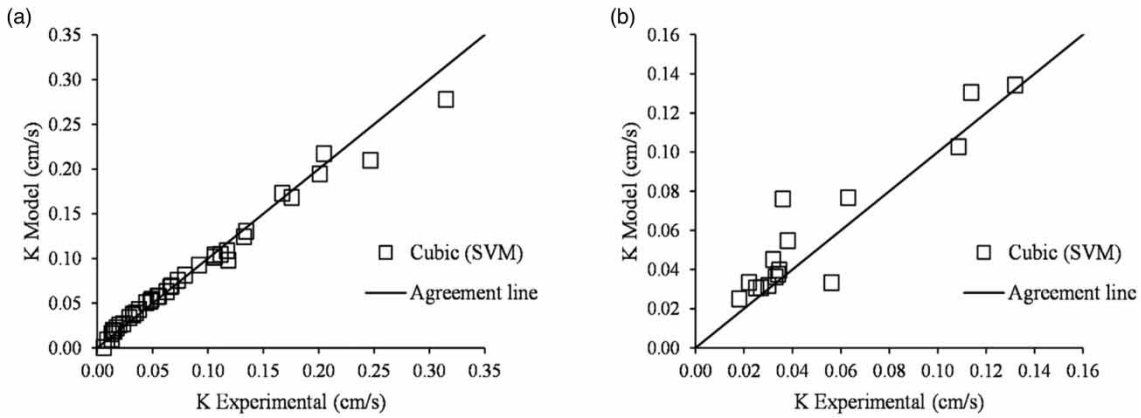


Figure 10 | Scatter plots of cubic (SVM) model for predicting *K* during (a) training and (b) testing stages.

prediction performance of the Gaussian (ANFIS) model is better than the quadratic (SVM) model. A graph between the number of test samples and *K* is plotted as shown in Figure 12 to indicate the variation in predicted *K* values obtained from Gaussian (ANFIS) and quadratic (SVM) models in comparison to the experimentally observed *K* values. It can be inferred from Figure 12 that predicted *K* values provided by ANFIS were close to the experimental *K* values and were found to follow the same pattern.

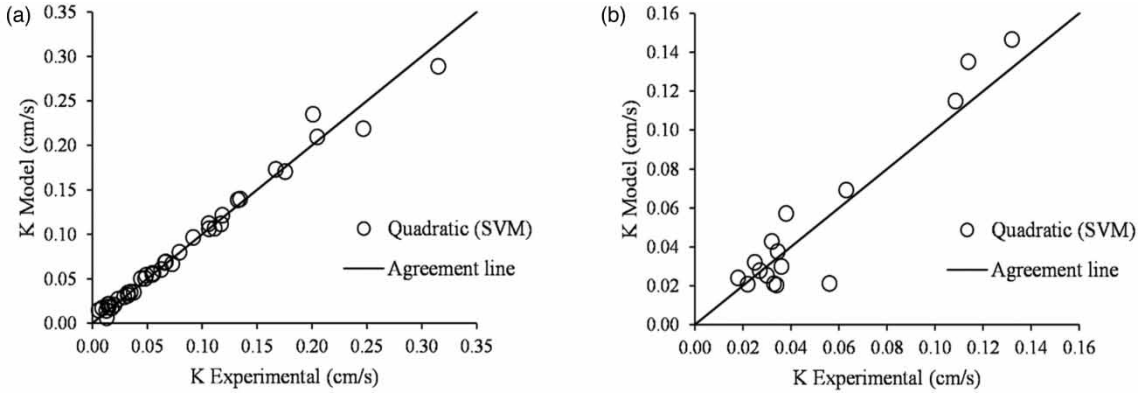


Figure 11 | Scatter plots of quadratic (SVM) model for predicting *K* during (a) training and (b) testing stages.

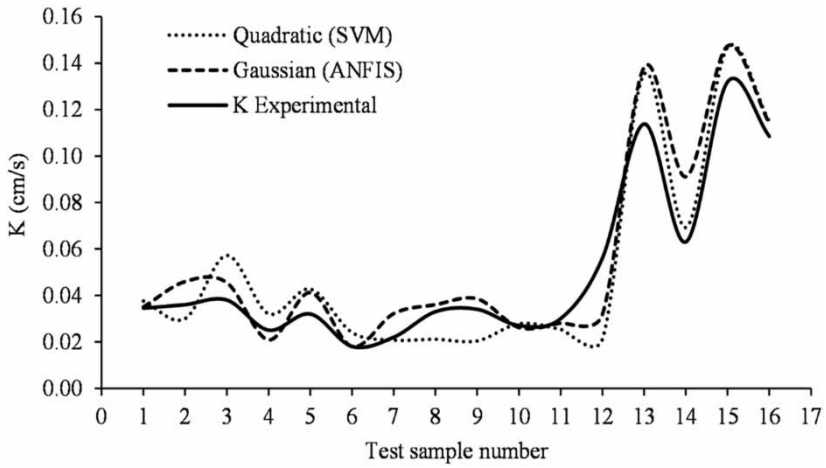


Figure 12 | Comparison of experimental *K* values with the predicted *K* values using Gaussian (ANFIS) and quadratic (SVM) models.

CONCLUSIONS

The study focuses on estimating the *K* of porous media using two data-driven techniques, i.e., ANFIS and SVM. Based on the correlation analysis, d_{10}, d_{50}, C_u, and n were found to be the influential parameters to predict the *K* of the porous media. The results of the study infer that both techniques are effective with the dataset in estimating the *K* values. The findings of the ANFIS model show that the Gaussian (ANFIS) model outperforms the triangular (ANFIS) model having *R* and RMSE values of 0.9661 and 0.0010 for the testing dataset, respectively, whereas the quadratic (SVM) model performs better than linear (SVM), Gaussian (SVM), and cubic (SVM) models with *R* and RMSE values of 0.9520 and 0.0015, respectively, for the testing dataset. The comparison of performance evaluation parameters signifies that the Gaussian (ANFIS) approach predicts better *K* values as compared to the SVM approaches. Furthermore, the prediction performance of these techniques can be checked with a greater number of datasets, and the application of these approaches in other hydrological studies could be the subject of in-depth study in the future.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Alabi, O. O. 2011 Validity of Darcy's law in laminar regime. *Electronic Journal of Geotechnical Engineering* **16**(A), 27–40.

Amoozegar, A. & Warrick, A. W. 1986 Hydraulic conductivity of saturated soils: field methods. *Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods* **5**, 735–770.

Arshad, R. R., Sayyad, G., Mosaddeghi, M. & Gharabaghi, B. 2013 Predicting saturated hydraulic conductivity by artificial intelligence and regression models. *International Scholarly Research Notices* **2013**, 1–8.

ASTM 2006 *Standard D2434 – Permeability of Granular Soils (Constant Head)*. ASTM International, West Conshohocken, PA, USA.

ASTM 2007 *Standard D422 – Particle-Size Analysis of Soils*. ASTM International, West Conshohocken, PA, USA.

Azad, A., Karami, H., Farzin, S., Saeedian, A., Kashi, H. & Sayyahi, F. 2018 Prediction of water quality parameters using ANFIS optimized by intelligence algorithms (case study: Gorganrood River). *KSCE Journal of Civil Engineering* **22**(7), 2206–2213.

Boadu, F. K. 2020 A support vector regression approach to predict geotechnical properties of soils from electrical spectra based on Jonscher parameterization. *Geophysics* **85**(3), EN39–EN48.

Chandel, A. & Shankar, V. 2021 Evaluation of empirical relationships to estimate the hydraulic conductivity of borehole soil samples. *ISH Journal of Hydraulic Engineering* 1–10.

Chandel, A. & Shankar, V. 2022 Grain-size model for hydraulic conductivity estimation of porous media. *Water Practice & Technology* **17**(9), 1836–1848.

Chandel, A., Shankar, V. & Alam, M. A. 2021 Experimental investigations for assessing the influence of fly ash on the flow through porous media in Darcy regime. *Water Science and Technology* **83**(5), 1028–1038.

Chandel, A., Fayaz, F. & Shankar, V. 2022a Assessment of column to particle diameter ratio on the hydraulic conductivity of porous media: wall effect in Darcy Regime. *ISH Journal of Hydraulic Engineering* 1–9.

Chandel, A., Sharma, S. & Shankar, V. 2022b Prediction of hydraulic conductivity of porous media using a statistical grain-size model. *Water Supply* **22**(4), 4176–4192.

Chapuis, R. P. 2012 Predicting the saturated hydraulic conductivity of soils: a review. *Bulletin of Engineering Geology and the Environment* **71**(3), 401–434.

Das, S. K., Samui, P. & Sabat, A. K. 2012 Prediction of field hydraulic conductivity of clay liners using an artificial neural network and support vector machine. *International Journal of Geomechanics* **12**(5), 606–611.

Deb, S. K. & Shukla, M. K. 2012 Variability of hydraulic conductivity due to multiple factors. *American Journal of Environmental Sciences* **8**(5), 489.

Ekhmaj, A. I. 2010 Predicting soil infiltration rate using artificial neural network. In *2010 International Conference on Environmental Engineering and Applications*. IEEE, pp. 117–121.

Elbisy, M. S. 2015 Support vector machine and regression analysis to predict the field hydraulic conductivity of sandy soil. *KSCE Journal of Civil Engineering* **19**(7), 2307–2316.

Ghorbani, M. A., Khatibi, R., Goel, A., FazeliFard, M. H. & Azani, A. 2016 Modeling river discharge time series using support vector machine and artificial neural networks. *Environmental Earth Sciences* **75**(8), 1–13.

Jalalkamali, A. 2015 Using of hybrid fuzzy models to predict spatiotemporal groundwater quality parameters. *Earth Science Informatics* **8**(4), 885–894.

Jougnot, D., Thanh, L. D., Van Do, P., Thuy, T. T. C., Hue, D. T. M. & Hung, N. M. 2021 Predicting water flow in fully and partially saturated porous media: a new fractal-based permeability model. *Hydrogeology Journal* **29**(6), 2017–2031.

Kumar, M., Tiwari, N. K. & Ranjan, S. 2020 Prediction of oxygen mass transfer of plunging hollow jets using regression models. *ISH Journal of Hydraulic Engineering* **26**(1), 23–30.

More, S. B., Deka, P. C., Patil, A. P. & Naganna, S. R. 2022 Machine learning-based modeling of saturated hydraulic conductivity in soils of tropical semi-arid zone of India. *Sādhanā* **47**(1), 1–15.

Nawaz, N., Harun, S. & Talei, A. 2015 Application of adaptive network-based fuzzy inference system (ANFIS) for river stage prediction in a tropical catchment. *Applied Mechanics and Materials* **735**, 195–199.

Raoof, M., Nazemi, A. H., Sadraddini, A. A., Marofi, S. & Pilpayeh, A. 2011 Measuring and estimating saturated and unsaturated hydraulic conductivity in steady and transient states on sloping lands. *World Applied Sciences Journal* **12**(11), 2023–2031.

Samui, P., Sitharam, T. G. & Kurup, P. U. 2008 OCR prediction using support vector machine based on piezocone data. *Journal of Geotechnical and GeoEnvironmental Engineering* **134**(6), 894–898.

Sihag, P. 2018 Prediction of unsaturated hydraulic conductivity using fuzzy logic and artificial neural network. *Modeling Earth Systems and Environment* **4**(1), 189–198.

Sihag, P., Singh, B., Gautam, S. & Debnath, S. 2018 Evaluation of the impact of fly ash on infiltration characteristics using different soft computing techniques. *Applied Water Science* **8**(6), 1–10.

Sihag, P., Kumar, M. & Singh, B. 2021 Assessment of infiltration models developed using soft computing techniques. *Geology, Ecology, and Landscapes* **5**(4), 241–251.

Singh, B., Sihag, P., Pandhiani, S. M., Debnath, S. & Gautam, S. 2021a Estimation of permeability of soil using easy measured soil parameters: assessing the artificial intelligence-based models. *ISH Journal of Hydraulic Engineering* **27**, 38–48.

Singh, B., Sihag, P., Parsaie, A. & Angelaki, A. 2021b Comparative analysis of artificial intelligence techniques for the prediction of infiltration process. *Geology, Ecology, and Landscapes* **5**(2), 109–118.

Torabi Podeh, H., Parsaie, A., Shahinejad, B., Arshia, A. & Shamsi, Z. 2022 Development and uncertainty analysis of infiltration models using PSO and Monte Carlo method. *Irrigation and Drainage*.

Vapnik, V. N. 1999 An overview of statistical learning theory. *IEEE Transactions on Neural Networks* **10**(5), 988–999.

Wang, J. P., François, B. & Lambert, P. 2017 Equations for hydraulic conductivity estimation from particle size distribution: a dimensional analysis. *Water Resources Research* **53**(9), 8127–8134.

Williams, C. G. & Ojuri, O. O. 2021 Predictive modelling of soils' hydraulic conductivity using artificial neural network and multiple linear regression. *SN Applied Sciences* **3**(2), 1–13.

Yilmaz, I., Marschalko, M., Bednarik, M., Kaynar, O. & Fojtova, L. 2012 Neural computing models for prediction of permeability coefficient of coarse-grained soils. *Neural Computing and Applications* **21**(5), 957–968.

First received 1 September 2022; accepted in revised form 18 November 2022. Available online 24 November 2022