



# Estimation of Settlement of Pile Group in Clay Using Soft Computing Techniques

Jitendra Khatti · Hanan Samadi ·  
Kamaldeep Singh Grover

Received: 28 April 2023 / Accepted: 24 August 2023 / Published online: 18 September 2023  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

**Abstract** The present research introduces an optimum performance soft computing model by comparing deep (multi-layer perceptron neural network, support vector machine, least square support vector machine, support vector regression, Takagi Sugeno fuzzy model, radial basis function neural network, and feed-forward neural network) and hybrid (relevance vector machine) learning models for estimating the pile group settlement. Six kernel functions have been used to develop the RVM model. For the first time, the single (mentioned by SRVM) and dual (mentioned by DRVM) kernel function-based RVM models have been employed for the reliability analysis of settlement of pile group in clay, optimized by genetic and particle swarm optimization algorithms. For that purpose, a database has been collected from the published article. Sixteen performance metrics

have been implemented to record the model's performance. Based on the performance comparison and score analysis, models MS3, MS9, MS17, MS23, and MS25 have been recognized as the better-performing models. Furthermore, the regression error characteristics curve, Uncertainty analysis, cross-validation (k-fold=10), and Anderson–Darling test reveal that model MS23 is the best architectural model in reliability analysis of pile group settlement. The comparison of model MS23 with published models shows that model MS23 has outperformed with a performance index of 1.9997, a20-index of 100, an agreement index of 0.9971, and a scatter index of 0.0013. The compression index, void ratio, and density influence the pile group settlement prediction. Also, the problematic multicollinearity level (variance inflation for > 10) significantly affects the performance and accuracy of the deep learning model.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10706-023-02643-x>.

J. Khatti (✉) · K. S. Grover  
Department of Civil Engineering, Rajasthan Technical University, Kota, Rajasthan, India  
e-mail: jitendrakhatti197@gmail.com

K. S. Grover  
e-mail: ksgrover@rtu.ac.in

H. Samadi  
School of Geology, College of Science, University of Tehran, Tehran, Iran  
e-mail: samadi.hanan1996@ut.ac.ir

**Keywords** Hybrid learning · Multicollinearity · Pile group settlement · Reliability analysis

## Abbreviations

3D	Three-dimensional
a20	A20-index
ACP	Axial capacity for driving piles
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
ANOVA	Analysis of variance
BCP	Bearing capacity of pile
BF	Bias factor

BPNN	Back propagation neural network	PLC	Pile load capacity
CAM	Cosine amplitude method	PLT	In-situ pile load test
Cc	Compression index	PLTC	Pile load test using a calibration chamber
CPT	Cone penetration test	Pm	Mean value of cyclic load
D	Footing depth	PSO	Particle swarm optimization algorithm
df	Degree of freedom	Pu	Ultimate bearing capacity of pile
DRVM	Dual Kernel function-based RVM	R	Coefficient of correlation
e	Void ratio	R <sup>2</sup>	Coefficient of determination
ED	Embedment depth	RBFNN	Radial basis function neural network
F	F State value	RMSE	Root mean square error
F crit	F critical value	RNN	Recurrent neural network
FAP	Footing net applied pressure	ROC	Receiver operating characteristics
FE	Finite element	RSR	Root mean square error to observation's standard deviation ratio
FEM	Finite element method	RVM	Relevance vector machine
FFNN	Fee-forward neural network	SA	Actual settlement
FN	Functional network	SE	Standard error
FOS	Factor of safety	SPM	Supplementary materials
$\gamma$	Soil density	SPT	Standard penetration test
GA	Genetic algorithm	SRVM	Single Kernel function-based RVM
GMDH	Group method of data handling	SS	Sum of square
GP	Genetic programming	StDev	Standard deviation
GPR	Gaussian process regression	SVM	Support vector machine
H0	Null hypothesis	SVR	Support vector regression
HR	Research hypothesis	TSFL	Takagi–Sugeno fuzzy method
IOA	Index of agreement	UB	Upper bound
IOS	Index of scatter	UBC	Ultimate bearing capacity
L	Footing width	UCS	Unconfined compressive strength
L/W	Length-to-width ratio of footings	UL	Upper level
LB	Lower bound	VAF	Variance accounted for
LL	Lower level	VIF	Variance inflation factor
LLP	Laterally loaded piles	W	Footing length
LMI	Legate and McCabe's index	WBC	Width of confidence bound
Lp/D	Total length of pile/pile diameter	WMAPE	Weighted mean absolute percentage error
Ls/Lt	Length of pile in the soil layer/length of pile in the rock layer		
Ls-SVM	Least square support vector machine		
MAE	Mean absolute error		
MAPE	Mean absolute percentage error		
MBE	Mean bias error		
ME	Margin of error		
MOE	Mean of error		
MS	Model structure		
MS	Mean of squares		
NMBE	Normalized mean bias error		
NS	Nash–Sutcliffe efficiency		
$\phi$	Angle of internal friction		
Pcyc	Half amplitude of the cyclic load		
PDR	Pile driving records		
PI	Performance index		

## 1 Introduction

Piles are one of the most important deep structures that transfer axial building loads into appropriate bearing strata. Unlike shallow foundations, pile foundations can convey axial loads through low-strength soil layers to suitable bodies. Pile foundations are more widely used than shallow foundations as they provide the best efficiency and optimum cost, even in the case of shallow load-bearing soil layers (Smolczyk 2003). In recent decades, the use of piles has increased due to the reducing raft settlement without decreasing the safety and foundation's performance

(Poulos 2001). One of the essential aspects in the design of pile foundations is assessing and investigating the load-settlement performance of a single or group pile. Many parameters affect a pile's behaviour, including the physical and mechanical behaviour of the soil (nonlinear or linear), pile properties, and the pile's installation methods (Berardi and Bovolenta 2005). However, the elastic settlement is a key factor for the final settlement of piles in sandy soil (Murthy 2002). Vesic (1977) computed the immediate settlement of a pile using a semi-empirical method. The elastic settlement of piles is not the major part of the final settlement and contributes to consolidation settlement for saturated clay soils (Poulos and Davis 1980; Murthy 2002). Determining the final settlement of a single or group pile is time-consuming. Therefore, several researchers and scientists used different methods, i.e., theoretical, experimental, and computational, to assess pile settlement. The theoretical procedures for assessing pile settlement are traditional, cumbersome, and less accurate. The computational methods are more accurate and less time-consuming than the traditional methods. The computational methods consist of several techniques, i.e., ANN, SVM, GPR, DT, RF, etc., associated with a conventional, advanced machine, hybrid, and deep learning.

Ray et al. (2023) analyzed the settlement of the shallow foundation on clayey soil using ANFIS and FN computational techniques. The authors predicted shallow foundation settlement using void ratio, compression index, and unit weight as independent variables. The performance of models was evaluated by several metrics, and it was concluded that the FN model is more reliable than ANFIS in predicting shallow foundation settlement. FN model predicted settlement with  $R^2$  of 0.9821, RMSE of 0.0012, and VAF of 98.44, showing better curve fitness than ANFIS. The defective pile failure scenario, pile spacing, and pile group configuration play a significant role if the pile group is axially loaded and designed for sand strata (Alhashmi et al. 2023). Zheng et al. (2020) employed the MARS model to assess the liquefaction-induced settlement of shallow foundations. The authors used the finite difference method to create an artificial database utilizing ground motion, structures, and soil properties. The proposed model assessed the settlement with a coefficient of determination of 0.884. Kumar and Samui (2020) employed the Ls-SVM, GMDH, and GPR soft computing approaches

to analyze the reliability of the pile group settlement in clay. The authors concluded that the Ls-SVM, GMDH, and GPR soft computing approaches are highly capable of analyzing the settlement reliability for the pile group in clay. Armaghani et al. (2020) applied a neuro-swarm approach to assess pile settlement. The authors developed a particle swarm-optimized neural network model in the reported study. The input variables were the SPT N-value,  $P_u$ ,  $L_s/L_t$ ,  $L_p/D$ , and UCS. The authors concluded that the optimized model assessed settlement with a determination coefficient of 0.892 in the testing phase, higher than the conventional neural network model. Moreover, the authors also concluded that the optimization algorithm improves the prediction capabilities of the neural network model. In recent decades, the use of artificial neural network (ANN) algorithms has been increased to predict the settlement of group and single piles due to their high efficiency and accuracy (Teh et al. 1997; Alkroosh and Nikraz 2011a, 2011b; Tarawneh and Imam 2014). A model based on recurrent neural network (RNN) was used to assess the load-settlement of axially driven steel piles. The defined RNN model was developed and validated by 23 in situ, full-scale pile load tests and cone penetration test (CPT) datasets. The model's accuracy was evaluated by the root mean square error (RMSE) performance metric. The result of this loss function was determined as 74.5, indicating the defined algorithm's accuracy (Shahin 2014). Kordjazi et al. (2014) employed a support vector machine (SVM) network to predict the settlement of piles as a target factor using 108 collected datasets. The test RMSE was obtained as 318.85 for the SVM model. Nevertheless, limited works and algorithms were applied to estimate the load-settlement behaviour of piles (Abu-Kiefa 1998; Pooya Nejad and Jaksa 2017). Table 1 presents the details of models used in the published research for predicting the single or group pile settlement.

Interestingly, many researchers and scientists have predicted pile settlement using traditional methods. Afnan et al. (2023) presented a theoretical study for investigating the vertical settlement and horizontal displacement of piles through clayey and silty soil bodies using the commercial finite element method. Furthermore, Li and Deng (2023) introduced a method to study the pile group foundation settlement behaviour for linear viscoelastic soil.

**Table 1** Details of soft computing models available in the literature

S. No	Soft computing approaches	Input variables		R <sup>2</sup> Test	Loss function	Reference
		Database	Variables (Input(s) and Output)			
1	LSSVM, GMDH, <b>GPR</b>	80	Cc, e, $\gamma$	1.00	RMSE: 0.02	Kumar and Samui (2020)
2	ANN	500	CPT	0.99	RMSE: 3.21	Pooya Nejad and Jaksa (2017)
3	RNN	23	PLT, CPT	0.99	RMSE: 74.5	Shahin (2014)
4	SVM	108	ACP, CPT	0.97	RMSE: 318.85	Kordjazi et al. (2014)
5	<b>GP</b> , SVM	-	W, D, L/W, $\phi$ , $\gamma$ , UBC	<b>0.99</b> , 0.98	MARE: <b>7.68</b> , 9.45	Adarsh et al. (2012)
6	<b>SVM</b> , GP	94	PDR, LLP, PLC	<b>0.94</b> , 0.89	RMSE: <b>334.05</b> , 308.3	Pal and Deswal (2010)
7	ANN	174	ACP, CPT	0.97	MSE: 4.1	Shahin (2010)
8	ANN	1000	SPT	0.72	RMSE: 12.42	Pooya Nejad et al. (2009)
9	SVM	272	L, W, FAP, ED, SPT, SA	0.92	RMSE: 8–10	Samui (2008)
10	ANN	272	L, W, FAP, ED, SPT, SA	0.94	RMSE: 6.2	Shahin et al. (2000)
11	BPNN	58	BCP, PLTC, PLT	0.85	SSE: 40.7	Lee and Lee (1996)

\*Bold values present the best model

This evaluation was based on several parameters, including the Poisson ratio on pile and load ratio, elastic modulus and slenderness ratio, pile spacing, and pile-pile interaction factors. Bao et al. (2023) reported that the stress-bearing ratio and axial stress of corner and side piles decrease due to increased pile spacing. Still, the stress-bearing ratio increases with pile spacing for piles at the center. Hakro et al. (2022) utilized a two-dimensional finite element to investigate foundations' settlement and structural behavior on different soil bodies by considering various load combinations and geometry situations. In another study, Oteuil et al. (2022) developed an analytical and numerical framework for the axial and lateral capacity of bored piles based on two parameters, including a conventional and a cone penetration test. Voyagaki et al. (2022) assessed the settlement of the axially loaded piles using the DINGO database. The database consists of pile results obtained from over 500 test piles. For this aim, the authors employed two analytical models to assess the settlement of the foundation. The models predicted foundation settlement accurately, close to actual values. Model 1 was based on the strength design principles, and Model 2 was designed using depth-dependent soil stiffness, elastic-perfectly plastic curve, bilinear force–displacement, and soil yielding. Finally, the authors concluded that models 1

and 2 predicted settlement for 2.5 FOS without significant differences. It was also noted that the settlement prediction mainly depends on soil properties. Ponomaryov and Sychkina (2022) determined the impact of compaction around driven piles in clay. The authors also estimated pile settlement using the numerical method (modeling in Plaxis 2D) and analytical method (based on Russian standards). The published research was carried out for the clay and claystone soils zone. Satisfactory settlement results were obtained from both methods for both zones. Zhang et al. (2021) estimated pile settlement by introducing a novel soil-pile interaction model considering Biot's poroelastic continuum equations and soil pile model. The authors concluded that the final settlement at the failure of the pile is affected by Young's modulus. Gomes Filho and Moura (2021) proposed modifications to the t-z and q-z curves, introduced by Bohn et al. (2017), to analyze the load-settlement ratios in pile groups. The method was implemented for eight soil cases studied by Dai et al. (2012). The authors concluded that the predicted results agreed with the experimental data of the pile group due to soil heterogeneity. Chen et al. (2021a) conducted experimental research for pile cyclic settlement in silt soil. The saturated silt soil strata were considered in the reported work using pore water pressure, soil pressure on the pile

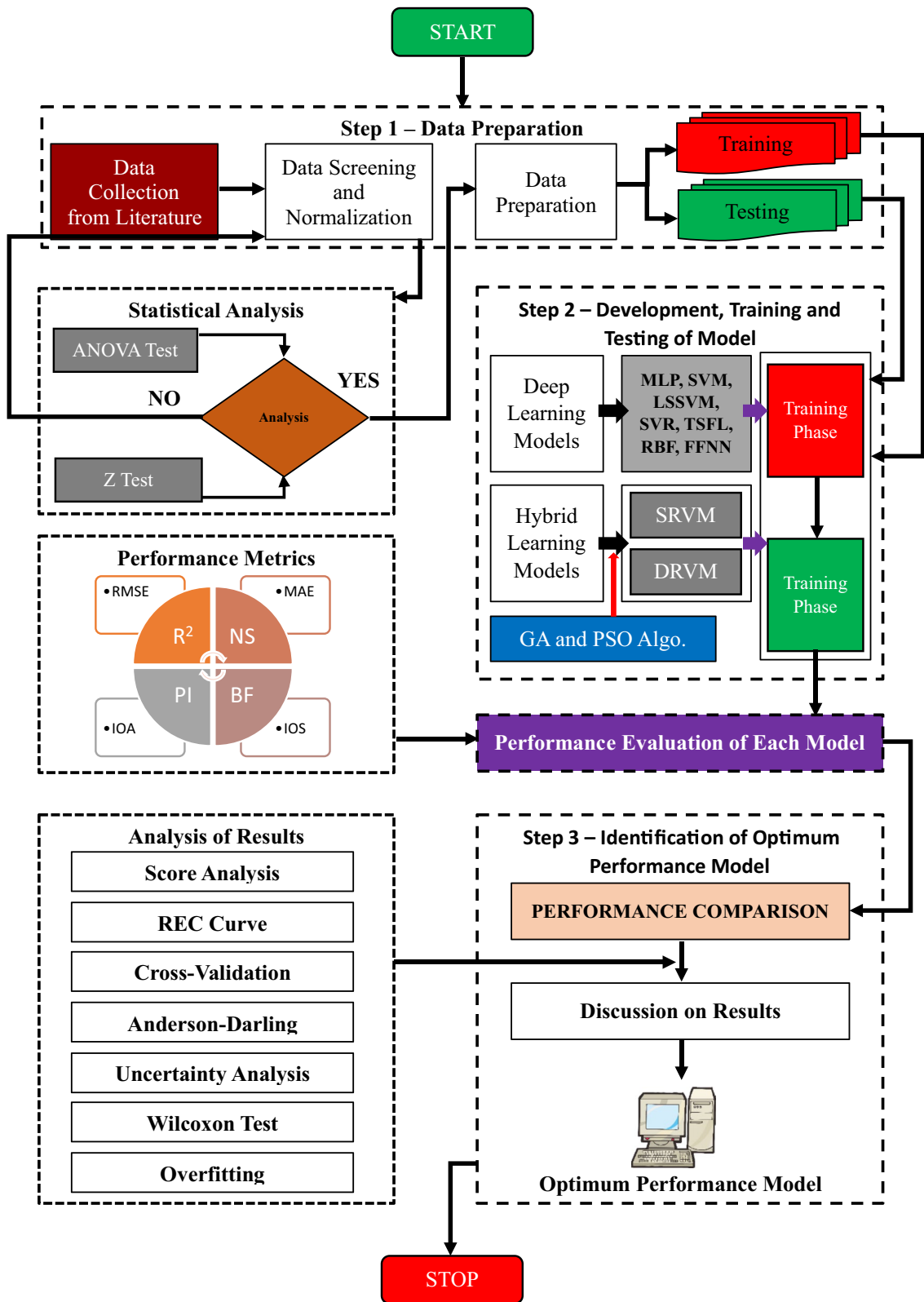
surface, and pile settlement. The authors concluded that the pile settlement develops due to  $P_m$ ,  $P_u$ , and  $P_{cyc}$  parameters. Bhartiya et al. (2021) studied the time-dependent settlements of pile raft foundations utilizing 3D FEM and critical state-based soil constitutive modified Cam-Clay (MCC) models. Lee et al. (2020) reported that the inside frictional resistance is a significant parameter for understanding the behaviour of steel pipe piles. Voyagaki et al. (2019) used linear elastic, linear elastic–plastic, and power-law nonlinear soil models and predicted pile settlement. The authors concluded that the adopted theoretical approaches agree well with the actual database. Sychkina (2019) implemented Koltunov’s creep kernel in the rheological deformation model to estimate the long-term settlement of piles. Li and Gong (2019) estimated the nonlinear settlement of pile groups in clay strata. The authors estimated settlement using an analytical approach based on the strength and stiffness of the soil. The authors assumed the soil strata elastically under loading. The proposed approach also considers the reinforcement effects of the adjacent pile and pile–pile interaction. The authors summarized that the predicted pile group settlement agrees well with the actual settlement. Cui et al. (2019) studied the long-term time-dependent load settlement for piles in clay. Crispin et al. (2019) evaluated the prediction error using the t-z curve. Also, the authors developed a nonlinear model (using pile shaft and base) compared to the linear elastic-perfectly plastic model developed by linear elastic stiffness. In addition, several researchers utilized the finite element method (including Defp1g and Napra) to investigate pile settlement (Katzenbach et al. 2000; Balakumar and Ilamparuthi 2007; Baziar et al. 2009; Sheil 2017). Moreover, the small-scale model, numerical, and analytical approaches were used to assess the settlement of single and group piles in clayey and sandy soil layers (Balakumar et al. 2005; Yamashita et al. 2013; Saha et al. 2015; Lai et al. 2016; Kumar et al. 2016). Wong (2002) performed numerical analyses and parametric studies to assess the responses of pavements and embankments supported by piles. Randolph (2003) addressed estimating pile capacity in clay and siliceous sands using analytical approaches. Mandolini et al. (2005) discussed the effect of the installation method on the bearing capacity and the load–settlement response

of a single pile. Poulos (2006) properly estimated the pile group settlements by focusing on pile–soil interaction. Furthermore, considering the multiphase model (soil–pile interaction), the settlement of a vertically loaded piled raft was analyzed and investigated (Bourgeois et al. 2012). Masani and Vanza (2018) investigated the impact of vertical load on the behavior of a group of piles by considering lateral load through laboratory experiments on aluminum pipe piles.

*Gap Identification in the Literature Survey* The literature study demonstrates that limited soft computing approaches have been employed to predict pile settlement by researchers and scientists. The quality and quantity of the database highly influence the performance of the soft computing models. Most researchers have employed models based on the RNN, SVM, ANN, Ls-SVM, GMDH, and GPR to assess the pile group settlement. It has also been observed that the models based on the Takagi–Sugeno Fuzzy (TSFL), Radial Basis Functional Neural Network (RBFNN), Feed-forward Neural Network (FFNN), and relevance vector machine (RVM) approaches have not been developed and employed for assessing the pile group settlement. Also, the hypothesis tests, such as ANOVA, Z, and Chi-tests, have not been carried out to select the research hypothesis. Furthermore, the impact of multicollinearity has not been investigated in the reliability analysis of pile group settlement. The Anderson–Darling test has not been carried out to study the behaviour of the predicted pile group settlement.

*Novelty of the Present Research* The present research develops, trains, tests, and analyzes the capabilities of deep and hybrid learning approaches. Considering the gap identified in the literature survey, the present research contains the following novelty.

- The present research checks the capabilities of deep (Multi-Layer Perceptron, Least Square Support Vector Machine, Support Vector Machine, Support Vector Regression, Takagi–Sugeno Fuzzy, Radial Basis Function Neural Networks, Feed-Forward Neural Network) and hybrid (Single Kernel Function-based and Dual Kernel Function-Based, which is optimized by genetic algorithm and particle swarm optimization algorithm) learning soft computing approaches in the reliability analysis of pile group settlement in clay.



◀**Fig. 1** Flow chart of the present research work

- This study also demonstrates the impact of the database multicollinearity on the performance, accuracy, and overfitting of the developed models.
- This work draws the hypothesis statement for reliability analysis and performs the ANOVA and Z tests to select the research hypothesis.
- This research analyzes the impact of input variables, compression index, void ratio, and density in predicting pile group settlement.
- The score analysis, uncertainty analysis, and regression error characteristics curve are carried out to find the optimum performance soft computing model for reliability analysis of pile group settlement. The Wilcoxon test has been performed to compare the actual and predicted settlement for the best architectural models in terms of confidence intervals.

## 2 Research Methodology

The present research has been performed to introduce the optimum performance model for reliability analysis of the settlement of pile group in clay. For this purpose, a database has been collected from the article of Kumar and Samui (2020). The database has been screened for training and testing, and missing data points have been removed. The min–max normalization function has been implemented to normalize the data points. The database contains the results of the compression index ( $C_c$ ), void ratio, density, and pile group settlement. The multicollinearity analysis has been performed to check the multicollinearity level of the data points and to determine the impact of multicollinearity on soft computing models. Also, the ANOVA and Z tests have been performed to identify the research hypothesis. In this research, the deep learning models, MLPNN, SVM, LS-SVM, SVR, TSFL, RBFNN, and FFNN, have been developed, trained, tested, and analyzed. On the other side, hybrid learning models have been developed using the RVM approach. The relevance vector machine is an advanced soft computing approach to support vector machine. The support and relevance vector machines are the kernel function-based approaches. The Gaussian, Exponential, Linear,

Laplacian, Sigmoid, and Polynomial kernel functions have been implemented in relevance vector machine models. Furthermore, each GA and PSO optimization algorithm has optimized the employed single kernel function-based (mentioned by SRVM) models. Thus, twelve RVM models (six GA-optimized SRVM and six PSO-optimized SRVM) have been employed. RMSE, MAE, R, MAPE, VAF, WMAPE, NS, PI, BF, NMBE, MBE, LMI, RSR, a20-index, IOA, and IOS performance metrics have been used to measure the performance of the employed models. Based on the performance comparison of six GA-optimized SRVM models, one kernel function has been selected as the best kernel function and mentioned by K1.

Similarly, one kernel function has been identified as the best kernel function by comparing six PSO-optimized SRVM models. The different kernel function combinations have been developed to employ the dual kernel function-based RVM models (mentioned by DRVM). Each GA and PSO-optimization algorithm has also optimized the DRVM models. Thus, ten DRVM models (five GA-optimized DRVM models and five PSO-optimized DRVM models) have been employed. The five GA-optimized DRVM and five PSO-optimized DRVM models have been individually compared. From the comparison, one GA-optimized DRVM and one PSO-optimized DRVM model has been recognized as the best architectural model. Thus, one best architectural model has been obtained from each deep learning, GA-optimized SRVM, PSO-optimized SRVM, GA-optimized DRVM, and PSO-optimized DRVM model. Finally, a performance comparison has been drawn among the five best architectural models to identify the optimum performance model. In addition, the results have been analyzed by performing score analysis, REC curve, cross-validation, Anderson–darling test, and uncertainty analysis. Moreover, the impact of multicollinearity on the overfitting of the best architectural model has been studied and analyzed. Based on the analysis, the obtained results have been discussed, and the optimum performance model has been identified. Figure 1 illustrates the execution of the present research.

## 3 Data Analysis and Computational Methods

This section presents the database source and analysis, followed by the computation methods adopted in

this study. A brief discussion has been drawn for each computational method. Furthermore, the performance metrics, sensitivity analysis, multicollinearity analysis, and hypothesis testing have been discussed in this section.

### 3.1 Data Analysis

A database of pile group settlement has been collected from the article of Kumar and Samui (2020). In the reported article, the authors have used the pile group of 9 piles embedded in normally consolidated clay. The authors have selected the following parameters in the reported study; (i) length of pile = 5 m, (ii) diameter of pile = 0.5 m, (iii) hard stratum at 7 m, (iv) pile spacing = 0.5 m, (v) dispersion angle = 30°, (vi) vertical load = 500 kN. The authors have also considered the compression index ( $C_c$ ), void ratio ( $e$ ), and density in the range of 0.26–0.85 (obtained from Ibrahim et al. 2012), 0.5–1, and 14.8–19 kN/m<sup>3</sup> (obtained from Das 2008). Eighty data points have been developed in the reported study. The database has been divided into training and testing databases by randomly picking up 80% and 20% of 80 data points, respectively. The frequency distribution of database variables has been drawn, as shown in Fig. 2.

Figure 2 demonstrates the frequency distribution of normalized data. The database consists of (i) 17, 15, 16, 13, 19 compression index data points, (ii) 20, 17, 18, 14, and 11 void ratio data points, (iii) 18, 18, 19, 19, and 6 density data points, and (iv) 14, 12, 12, 14, and 30 settlement data points in the range of 0 to 0.2, 0.2 to 0.4, 0.4 to 0.6, 0.6 to 0.8, and 0.8 to 1.0, respectively. Furthermore, the Pearson

product-moment correlation coefficient method has been used to determine the relationship among the  $C_c$ ,  $e$ , density, and settlement variables. The correlation coefficient  $\pm 0.0$  to  $\pm 0.20$ ,  $\pm 0.21$  to  $\pm 0.40$ ,  $\pm 0.41$  to  $\pm 0.60$ ,  $\pm 0.61$  to  $\pm 0.80$ , and  $\pm 0.81$  to  $\pm 1.0$  demonstrate very strong, strong, moderate, weak, and no relationship for database pairs (Hair et al. 2013; Khatti and Grover 2023a). Figure 3 demonstrates the relationship between variables in terms of the correlation coefficient.

Figure 3 illustrates that all variables very strongly correlate with each other. The correlation of  $C_c$  with  $e$  (=0.9832) and settlement (0.9846) presents the multicollinearity. Also, the settlement has multicollinearity with  $C_c$  (=0.9846) and density (=0.9784). The descriptive statistics of the overall, training and testing database are given in Table 2.

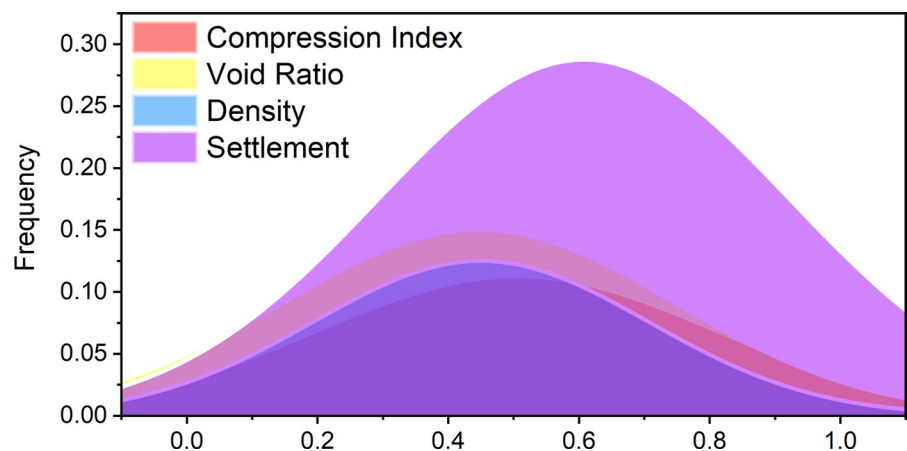
### 3.2 Computation Methods

Eight soft computing approaches have been used to develop the different configuration-based models to execute the present study. The theory of the adopted soft computing has been discussed in this section.

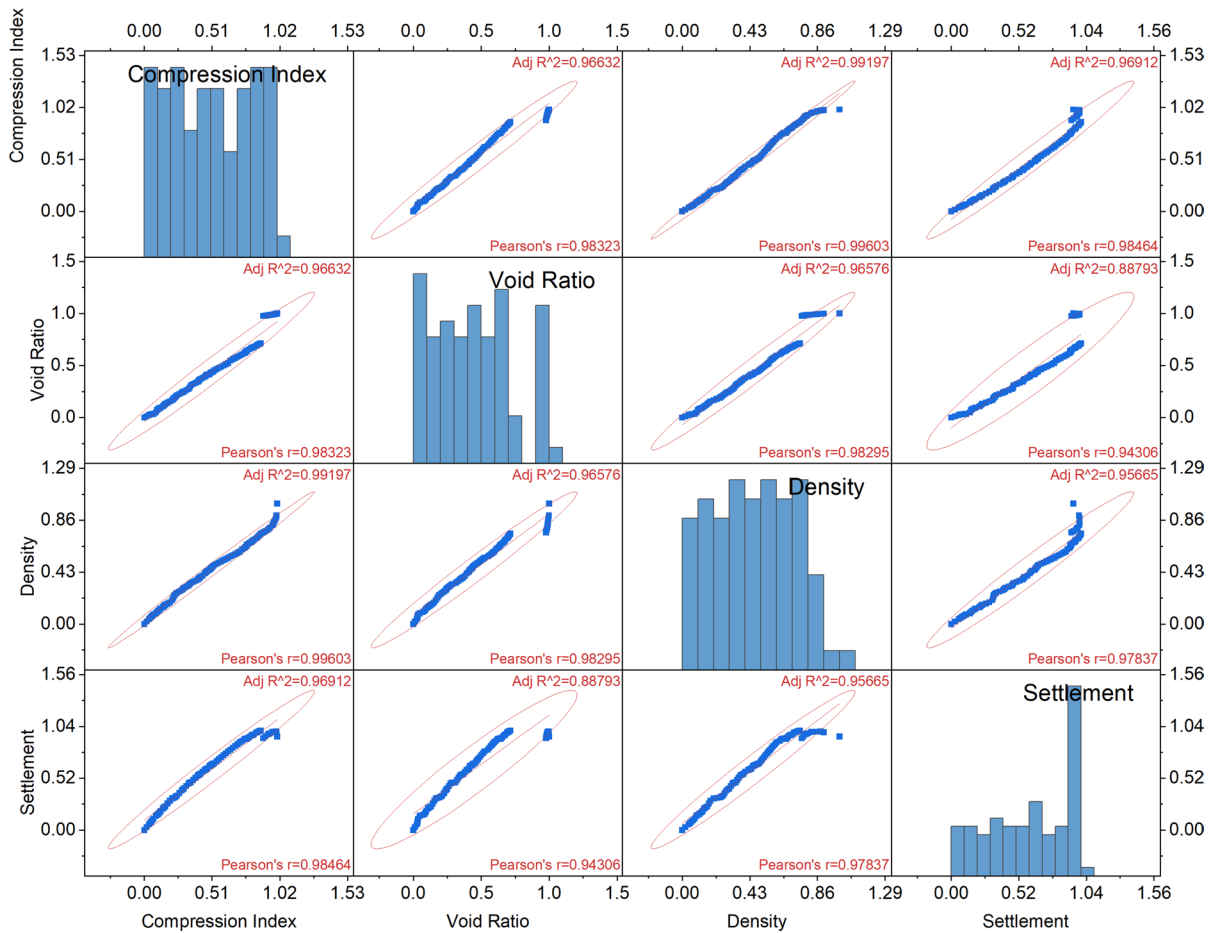
#### 3.2.1 Multi-layer Perceptron Neural Network (MLPNN)

An Artificial neural network (MLPNN) is a soft computing system composed of artificial neurons with inputs generating an output signal that could be submitted to several other neurons. A network of connections to which the weight is assigned modifies the input strength. The activation function such as

**Fig. 2** Illustration of the frequency distribution of variables







**Fig. 3** Demonstration of relationship among variables

Threshold, Sigmoid, Hyperbolic tangent, and ReLu are used in the network to convert input signals in input layers to the output signal in the output layer (Menhrotra et al. 1990).

The optimum architecture network assumed in this paper is 3-13-1, which means the ANN has three layers, including five neurons and five input parameters in this research. One hidden layer with thirteen neurons, followed by one neuron in the output layer, eventually generates the group pile settlement. The structure of MLPNN is presented in Fig. 4.

### 3.2.2 Support Vector Machine (SVM)

Support vector machine (SVM) is one of the established methods of supervised machine learning algorithms, applied to classify and predict with small

samples and nonlinearity by constructing a hyperplane or set of hyperplanes in a high or infinite-dimensional space. The SVM algorithm is a complicated nonlinear relationship between the target and input parameters, making it a desirable case for this analysis. SVM aims to determine the maximum possible margin between the classes (Cortes and Vapnik 1995). The SVM margin and its description are shown in Fig. 5.

### 3.2.3 Least Square Support Vector Machine (Ls-SVM)

Least-squares support-vector machines (Ls-SVM) are used for prediction and statistical modeling. Least-squares (Ls) is a version of support-vector machines (SVM), a set of related supervised learning methods

**Table 2** Descriptive statistics of overall, training, and testing databases

Variables	N	Mean	SE Mean	StDev	Variance	CoefVar	Sum	Mini	Q1	Median	Q3	Max	IQR	Mode	N for Mode	Skewness	Kurtosis	MSSD	
<i>Overall database</i>																			
Cc	80	0.55	0.02	0.18	0.03	31.98	26.45	0.26	0.39	0.54	0.71	0.83	0.32	0.7	2	0.03	-1.3	0	
e	80	16.68	0.14	1.26	1.59	7.56	22,375.21	14.80	15.61	16.58	17.54	19.00	1.93	*	0	0.39	-0.79	0.009	
Density	80	0.73	0.01	0.13	0.02	17.70	43.36	0.50	0.62	0.73	0.83	1.00	0.21	*	0	0.03	-1.04	0	
Settlement	80	0.20	0.00	0.04	0.00	19.76	3.33	0.12	0.17	0.20	0.24	0.25	0.07	*	0	-0.33	-1.25	0	
<i>Training database</i>																			
Cc	64	0.49	0.02	0.14	0.02	28.94	16.49	0.26	0.36	0.49	0.61	0.73	0.25	0.7	2	0.13	-1.2	0	
e	64	16.20	0.11	0.88	0.77	5.42	16,849.75	14.80	15.41	16.23	16.94	17.65	1.53	*	0	0.03	-1.25	0.001	
Density	64	0.68	0.01	0.10	0.01	14.99	30.33	0.50	0.59	0.68	0.77	0.85	0.18	*	0	-0.06	-1.19	0	
Settlement	64	0.19	0.00	0.04	0.00	19.05	2.36	0.12	0.16	0.19	0.22	0.24	0.06	*	0	-0.08	-1.19	0	
<i>Testing database</i>																			
Cc	16	0.79	0.01	0.03	0.00	4.00	9.96	0.74	0.76	0.79	0.82	0.83	0.06	*	0	-0.17	-1.51	0.00002	
e	16	18.58	0.14	0.58	0.33	3.10	5525.46	17.70	17.78	18.93	18.97	19.00	1.19	*	0	-0.89	-1.38	0.04	
Density	16	0.90	0.01	0.04	0.00	4.30	13.03	0.86	0.87	0.89	0.93	1.00	0.06	*	0	1.13	1.28	0.00011	
Settlement	16	0.25	0.00	0.00	0.00	1.13	0.97	0.24	0.24	0.25	0.25	0.25	0.00	*	0	-1.18	0.52	0.000005	

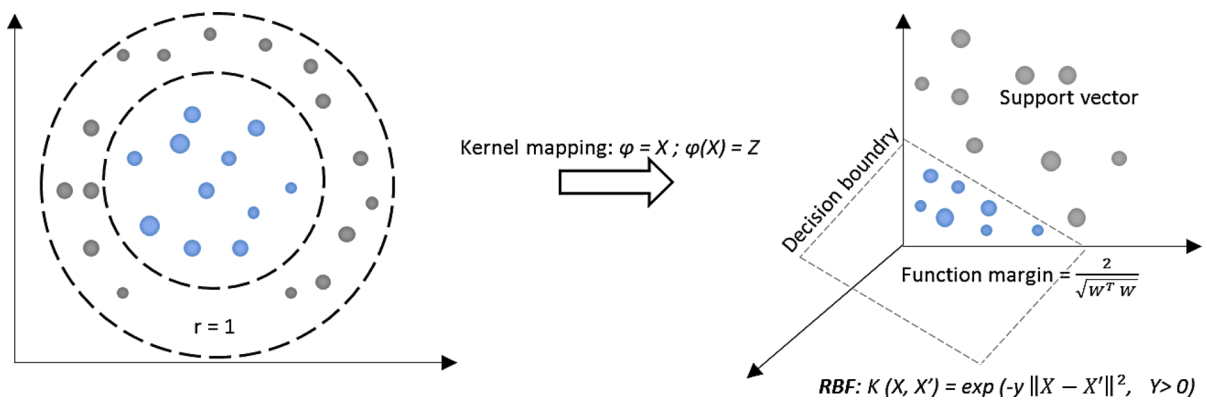
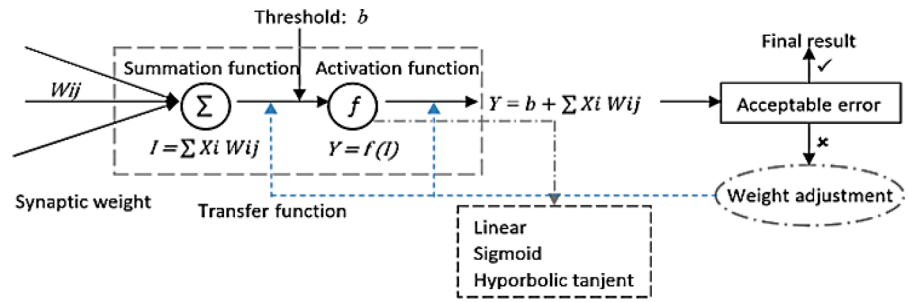
that analyze data and recognize patterns and are used for classification and regression analysis (Khatti and Grover 2023f). Least-squares SVM classifiers were proposed by Suykens and Vandewalle (1999), which are a class of kernel-based learning methods. There are three types of kernel-based functions, namely linear, polynomial, and Gaussian, which are different in making hyperplane decision boundaries between the data classes. Linear and polynomial kernels are

less time-consuming and provide less accuracy than the Gaussian kernels (Hongwei 2011), as presented in Fig. 6. Therefore, the Gaussians kernel-based Ls-SVM is performed in this research.

### 3.2.4 Support Vector Regression (SVR)

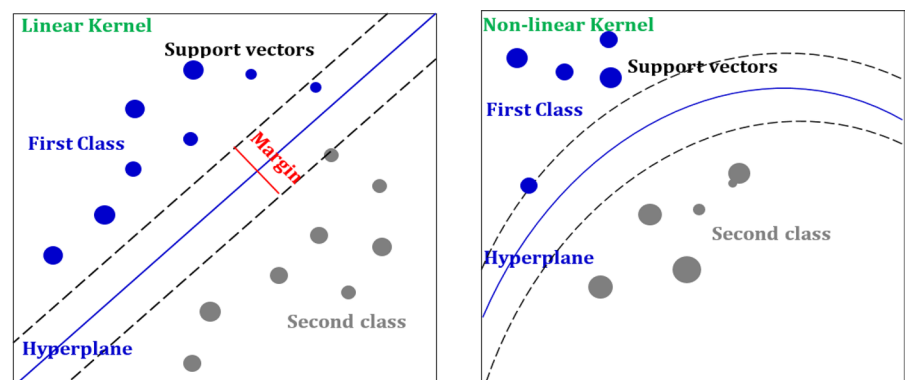
Support vector regression (SVR) is a widely used supervised learning network to estimate continuous

**Fig. 4** The structure of ANN based on the determined database (Samadi et al. 2021a)



**Fig. 5** Maximum margin hyperplane (MMH) and margins of an SVM (Samadi et al. 2021a)

**Fig. 6** The structure of linear and nonlinear kernel function



variables. SVR is similar to the support vector machine (SVM) algorithm principle. The best-fit line in this computing machinery model is the hyperplane with maximum marks (Smola and Scholkopf 2004).

### 3.2.5 Takagi–Sugeno Fuzzy Model (TSFL)

Takagi and Sugeno (1985) introduced the Takagi–Sugeno fuzzy model to develop a systematic fuzzy rule generation approach by a given input–output dataset. The typical TS fuzzy rule is defined as follows:

$$\text{If } X \text{ is } A \text{ and } Y \text{ is } B, \text{ then } Z = f(X, Y) \tag{1}$$

where A and B are fuzzy sets in the antecedent, while  $Z=f(X, Y)$  is a crisp function in the consequent. Usually,  $f(X, Y)$  is a polynomial consisting of input variables X and Y. Any function describing the appropriate output within the fuzzy region specified by the rule’s antecedent can be considered  $F(X, Y)$ . In this research, the TSFL model is used because the Sugeno fuzzy systems are well established using linear weighted mathematical expressions and are well suited to human input adaptive techniques (Takagi and Sugeno 1985), as illustrated in Fig. 7. The output is a weighted mean that is defined as follow:

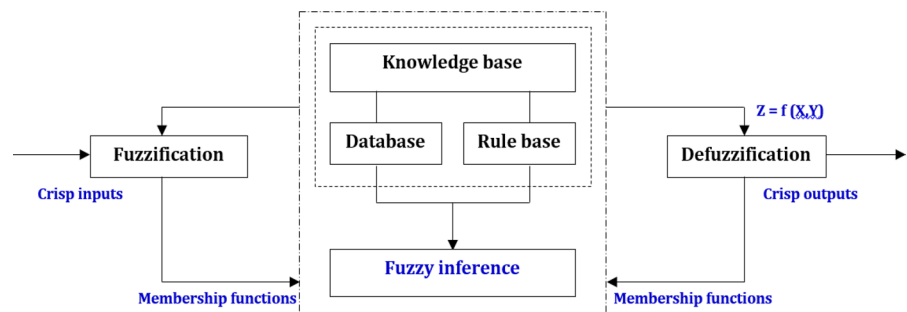
$$Z = \frac{\sum w_i f_i(X, Y)}{\sum w_i} \tag{2}$$

where  $W_i$  is the firing strength of the  $i^{\text{th}}$  output, transforming a crisp quantity into a fuzzy set is known as Fuzzification. This procedure accurately expresses the crisp input values into linguistic variables.

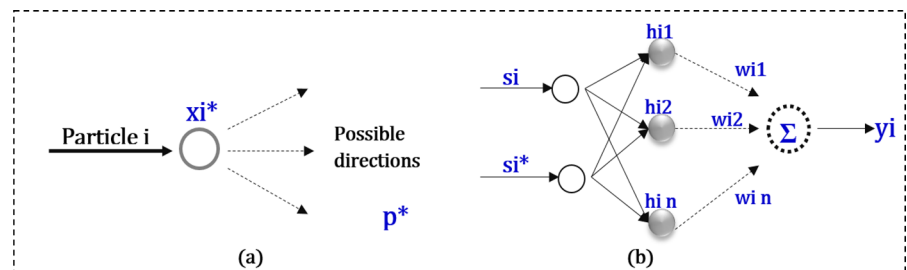
### 3.2.6 Radial Basis Function Neural Network (RBFNN)

The values of the RBF method depend on the distance between the inputs and fixed points. This function assigns an actual value to each input and the value predicted by the network, which cannot be negative. Broomhead and Lowe (1988) indicated that RBF is traditionally associated with radial functions in a single-layer network. The structure of this model is shown in Fig. 8. Three-layer were determined in the training stage, including input layers, hidden layers, and output layers. To increase the linear separability of the feature vector, the dimension of the feature vector must be increased. Due to the computational reliability of SVM, it can utilize a radial basis function (RBF) as the best Gaussian kernel function, which is defined as follows (Apostolopoulou et al. 2020):

**Fig. 7** Structure of Takagi–Sugeno fuzzy model (Fuzzification and Defuzzification)



**Fig. 8** A view of radial basis function structure (Samadi and Farrokh 2021b)



$$W = \sum_{i=0}^n a_i y_i x_i \tag{3}$$

$$K(x, y) = \exp(-y||x - z||), y > 0, y = \frac{1}{2\sigma^2} \tag{4}$$

In this function, W is assumed at the margin between the groups that the RBF algorithm attempts to maximize the W value.

### 3.2.7 Feed-Forward Neural Network (FFNN)

A feed-forward neural network (FNN) is one of the neural networks that move in only one direction (forward) without any loops in the network structure. The network considers the input nodes and moves through the hidden layer to estimate output factors (Andreas 1994).

### 3.2.8 Relevance Vector Machine (RVM)

The RVM is a machine learning method used in mathematics that employs Bayesian inference to produce parsimonious solutions for probabilistic classification and regression problems. The RVM offers probabilistic classification yet has a functional form equivalent to the support vector machine (Tipping 2001). In reality, it is similar to a Gaussian process model with a covariance function:

$$k(j, j') = \sum_{m=1}^N \frac{1}{\alpha_m} \varphi(j, j_m) \varphi(j', j_m) \tag{5}$$

where  $\varphi$  is the kernel function (default Gaussian),  $\alpha_m$  are the variances of the prior on the weight vector  $\omega \sim N(0, \alpha^{-1}I)$ , and  $j_1, \dots, j_N$  are the input parameters of the training dataset (Candela 2004). The Gaussian, Linear, Laplacian, Polynomial, Sigmoid, and Exponential kernel functions are used in RVM models (Khatti and Grover 2023g). The mathematical formulation of the kernel functions is

Linear

$$k(x_i, x_j) = m(x_i, x_j) + c \tag{6}$$

Polynomial Kernel

$$k(x_i, x_j) = (x_i \cdot x_j + 1)^d \tag{7}$$

Gaussian Kernel

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{8}$$

Exponential Kernel

$$k(x_i, x_j) = D^2 \exp\left(-\frac{\|x_i - x_j\|}{l}\right) \tag{9}$$

Laplace RBF Kernel

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|}{\sigma}\right) \tag{10}$$

where  $x_i, x_j$  are input and output parameters, c is constant, m is the gradient/slope of a line, d is a degree, D is the scale factor, l is the length-scale hyper-parameter, and  $\sigma$  is the standard deviation. The Bayesian formulation of the RVM eliminates the set of free parameters of the SVM compared to those of support vector machines (SVM). However, because RVMs employ a learning strategy similar to expectation maximization (EM), they are susceptible to local minima. In contrast, SVMs often use sequential minimal optimization (SMO)-based methods, ensuring a global optimum.

Several researchers have solved complex problems associated with geotechnical engineering using the RVM approach (Khatti and Grover 2021; Mittal et al. 2021; Chen et al. 2021; Fattahi and Hasani-panah 2021; Li et al. 2021; Yang et al. 2022; Raja et al. 2022). The published studies demonstrate the efficiency of the RVM approach in solving geotechnical problems. Therefore, relevance vector machine models have been developed using single (SRVM) and dual (parallel mentioned by DRVM) kernel functions. The Gaussian, Exponential, Linear, Laplacian, Sigmoid, and Polynomial kernel functions have been implemented to develop the proposed RVM models in this study. The hyperparameters of the developed RVM models are given in Table 3.

Thus, each GA and PSO algorithm has developed and optimized six single kernels (gaussian,

**Table 3** Details of hyperparameters selected for RVM models

Hyperparameters	SRVM	DRVM
<i>Parameter settings</i>		
Free basis	Enable	
Kernel functions	Gaussian, Exponential, Linear, Laplacian, Sigmoid, Polynomial	
Max. Iterations	1000	
Number of Kernels	Single	Two
<i>Optimizer settings</i>		
Methods	GA & PSO	
Target	Single Kernel	Two Kernels
lb	$2^{-6}$	$2^{-5}, 10^{-2}, 10^{-3}, 10^{-3}$
uB	$2^{-6}$	$2^{-5}, 10^0, 10^3, 10^3$
Num. variable	1	4
Max. iterations	100	
Kfolds	5	

exponential, linear, laplacian, sigmoid, and polynomial) based on SRVM models.

The performance of single kernel-based RVM models optimized by each GA and PSO has been compared to select the primary kernel for dual (parallel) kernel-based RVM model. However, six kernel functions have been used in this research. Therefore, a combination of the primary kernel (kernel 1) and the rest of the kernel (kernel 2, 3, 4, 5, 6) has developed the dual (parallel) kernel-based DRVM models. For example, GA optimized Gaussian kernel-based SRVM model has performed better than other GA-optimized SRVM models. Thus, kernel combinations are Gaussian + Linear, Gaussian + Laplacian, Gaussian + Polynomial, Gaussian + Sigmoid, and Gaussian + Exponential for GA-optimized DRVM models. The same procedure has been used to develop the PSO-optimized dual (parallel) kernel-based DRVM models.

### 3.3 Performance Metrics

The different statistical parameters are used to determine the performance of soft computing called performance metrics. These performance metrics are linear and nonlinear. In this study, sixteen performance metrics have been implemented to measure the performance of deep and hybrid learning models. The mathematical expression of the implemented performance

metrics is as follows (Kumar and Samui 2020; Asteris et al. 2021a, 2021b; Khatti and Grover 2023b, 2023c):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\alpha - \beta)^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |(\omega - \alpha)| \quad (12)$$

$$R^2 = \frac{\sum_{i=1}^r (\alpha - \beta)^2 - \sum_{i=1}^r (\alpha - \omega)^2}{\sum_{i=1}^r (\alpha - \beta)^2} \quad (13)$$

$$R = \frac{\sum (\alpha_i - \bar{\beta})(\omega_i - \bar{\omega})}{\sqrt{\sum (\alpha_i - \bar{\beta})^2 \sum (\omega_i - \bar{\omega})^2}} \quad (14)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\alpha - \omega}{\alpha} \right| * 100 \quad (15)$$

$$WMAPE = \frac{\sum_{i=1}^n \left| \frac{\alpha - \omega}{\alpha} \right| * \alpha}{\sum_{i=1}^n \alpha} \quad (16)$$

$$VAF = \left( 1 - \frac{\text{var}(\alpha - \omega)}{\text{var}(\alpha)} \right) * 100 \quad (17)$$

$$PI = R^2 + \left( VAF/100 \right) - RMSE \tag{18}$$

$$RSR = \frac{RMSE}{\sqrt{\frac{1}{N} \sum_{i=1}^N (\alpha - \beta)^2}} \tag{19}$$

$$NS = 1 - \frac{\sum_{i=1}^n (\alpha - \omega)^2}{\sum_{i=1}^n (\alpha - \beta)^2} \tag{20}$$

$$BF = \frac{1}{n} \sum_{i=1}^n \frac{\omega}{\alpha} \tag{21}$$

$$a\ 20\ index = \frac{m20}{H} \tag{22}$$

$$IOA = 1 - \frac{\sum_{i=1}^n (\omega - \alpha)}{2 \sum_{i=1}^n (\alpha - \beta)} \tag{23}$$

$$IOS = \frac{RMSE}{Avg.\ of\ Actual\ Values} \tag{24}$$

$$LMI = 1 - \left[ \frac{\sum_{i=1}^n |\alpha - \omega|}{\sum_{i=1}^n |\alpha - \beta|} \right] \tag{25}$$

$$NMBE = \frac{\frac{1}{N} \sum_{i=1}^n (\omega - \alpha)^2}{\frac{1}{N} \sum_{i=1}^n \alpha} \tag{26}$$

$$MBE = \frac{1}{N} \sum_{i=1}^n (\omega - \alpha) \tag{27}$$

where  $\alpha$  and  $\omega$  are the actual and predicted  $i$ th value,  $n$  presents the total number of data,  $\beta$  is the mean of the actual values,  $\bar{\omega}$  is the mean of the predicted value,  $k$  is the number of independent variables,  $m20$  is the ratio of experimental to the predicted value, varies between 0.8–1.2, and  $H$  is the total number of data samples.

The primary advantage of the a20-index is that the proposed model predicts values with a deviation of  $\pm 10\%$  compared to laboratory values. On the other hand, the index of agreement (mentioned in Eq. 23) is bounded by  $- 1.0$  and  $1.0$  (Willmott et al. 2012). Moreover, the least value of the scatter

index (mentioned in Eq. 20) presents a better prediction and accuracy (Mentaschi et al. 2013). The R-squared value higher than 0.95 ( $R=0.9747$ ) demonstrates that the developed model is highly reliable and accurate. Also, the value of  $R$  more than 0.8 ( $R^2=0.64$ ), between 0.2 ( $R^2=0.4$ ) to 0.8 ( $R^2=0.64$ ), and less than 0.2 ( $R^2=0.4$ ) presents the strong, good and weak correlation between pair of data (Smith 1986). A perfect predictive model always has performance indicators value equal to the ideal value, as given in Table 4.

### 3.4 Sensitivity Analysis

A sensitivity analysis has been performed to determine the most influencing input parameters in predicting soil compaction. The cosine amplitude method (CAM) is utilized in this study to compare the strength of the input parameters with the soil compaction parameters (Ardakani and Kordnaeij 2019). It aids in determining the degree of the correlation between input and output dimensions. In CAM, the data array  $X$  ( $n$  data samples in the same space) can be expressed as (Hasanzadehshoouili et al. 2012):

$$X = \{x_1, x_2, x_3 \dots x_m\} \tag{28}$$

Each predictor  $x_i$  of the data array,  $X$  is the vector (length  $m$ ) in Eq. 28 and is defined as:

$$X = \{x_{i1}, x_{i2}, x_{i3} \dots x_{im}\} \tag{29}$$

Thus, Eq. 29 estimates the strength between predictors ( $x_i$ ) and output ( $x_j$ ) variable (Ghorbani et al. 2020).

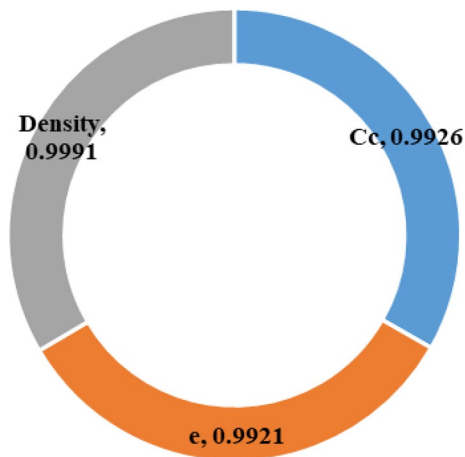
**Table 4** Ideal value of the different performance indicators

Indicators	Value	Indicators	Value
RMSE	0	NMBE	0
MAE	0	NS	1
R2	1	LMI	0
R	1	RSR	0
MAPE	0–100	a20-index	100
WMAPE	0	IOA	1
VAF	100	IOS	0
PI	2	BF	0

$$CAM = \frac{\sum_{k=1}^m x_{ik}x_{jk}}{\sqrt{\sum_{k=1}^m x_{ik}^2 \sum_{k=1}^m x_{jk}^2}} \quad (30)$$

The strength between predictors and output is measured between 0 and 1. The value of CAM near 1 demonstrates the higher strength between data points, and the zero value represents no strength. The sensitivity analysis has been performed in this research for the 50%, 60%, 70%, 80%, 90%, and 100% training databases. Figure 9 illustrates the strength of data points for the complete database used in this study.

Figure 9 illustrates that all input variables are highly sensitive to pile group settlement. It can be seen that the soil density (=0.9991) very strongly influences the prediction of pile group settlement, followed by the compression index (=0.9926) and void ratio (=0.9921).



**Fig. 9** Illustration of sensitivity of input variables

### 3.5 Multicollinearity Analysis

A relationship between variables resulting in their correlation is called multicollinearity. Multicollinear data are difficult to analyze since they are not independent (Khatti and Grover 2023a). The Pearson product moment-correlation coefficient and variance inflation factor methods determine the multicollinearity database. In this research, the variance inflation factor (VIF) method has been adopted to determine the database multicollinearity, and the results of VIF are presented in Table 5.

Table 4 illustrates that compression index, void ratio, and density have 31.65, 134.89, and 132.66 multicollinearities, respectively. All input variables have a VIF value of more than 10, showing the problematic multicollinearity level (Khatti and Grover 2023a).

### 3.6 Hypothesis Testing

Hypothesis testing is a process for determining the hypothesis type for a particular study. For this purpose, several parametric and nonparametric statistical tests are used. The parametric ANOVA and Z tests have been performed in this research to identify the hypothesis type. The present study has the following statements for the research hypothesis ( $H_R$ ):

- Soil compression index, void ratio, and density are essential for assessing pile group settlement.
- Soil compression index, void ratio, and density are highly related.
- Multicollinearity does not affect the performance and accuracy of deep and hybrid learning soft computing models.

**Table 5** VIF results for defining database multicollinearity level

Statistical parameters	Coeff	SE	t Stat	p-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	R <sup>2</sup>	VIF
Intercept	0.3786	0.0284	13.3428	0.0000	0.3221	0.4351	0.3221	0.4351	–	–
Cc	0.3798	0.0323	11.7436	0.0000	0.3154	0.4442	0.3154	0.4442	0.9684	31.65
e	–0.0238	0.0022	–10.9487	0.0000	–0.0282	–0.0195	–0.0282	–0.0195	0.9926	134.89
Density	0.0153	0.0438	0.3505	0.7269	–0.0719	0.1025	–0.0719	0.1025	0.9925	132.66



### 3.6.1 ANOVA Analysis

Analysis of variance (ANOVA) is a parametric statistical analysis performed to identify the research hypothesis. In the present research, the ANOVA analysis has been performed using the Data Analysis Tool of Microsoft Excell 2021. The results of the ANOVA analysis are given in Table 6.

Table 5 illustrates that the p-value for the Cc, e, and density is less than the significance level, i.e.,  $p < 0.05$ . Also, the F (F state) is higher than the F critical (F Crit) for each input variable. Finally, the ANOVA analysis rejects the present research’s null hypothesis ( $H_0$ ).

### 3.6.2 Z-Test

Z-test is another parametric statistical test performed to find the hypothesis type. The Z-test has been performed for the complete database used in this study.

Table 7 illustrates the results obtained from the Z-test. It can be seen that (i) Z (z state) is higher than Z critical one and two tail, (ii) Z critical one tail is less than z critical two tail, (iii) p-value is less than to significance p-value. After analyzing the results, it has been observed that Z-test accepts the research hypothesis for the study.

## 4 Results and Discussion

### 4.1 Simulation of Soft Computing Models

To execute the present research, the models based on the MLPNN, SVM, Ls-SVM, SVR, TSFL, RBF, FFNN, and RVM have been developed, trained, tested, and analyzed for performing reliability analysis of pile group settlement. Each GA and PSO algorithm has been used to develop and optimize the single and dual kernel function-based RVM

**Table 6** Results of the ANOVA analysis for a pile group settlement

Source of variation	SS	df	MS	F	p-value	F crit
<i>ANOVA for Cc and settlement</i>						
Between groups	4.84	1.00	4.84	299.68	0.00	3.90
Within groups	2.55	158.00	0.02			
Total	7.39	159.00				
<i>ANOVA for e and settlement</i>						
Between groups	10,859.25	1.00	10,859.25	13,647.94	0.00	3.90
Within groups	125.72	158.00	0.80			
Total	10,984.97	159.00				
<i>ANOVA for density and settlement</i>						
Between groups	11.02	1.00	11.02	1221.23	0.00	3.90
Within groups	1.43	158.00	0.01			
Total	12.44	159.00				

**Table 7** Results of the ANOVA analysis for a pile group settlement

Statistical parameters	Cc	e	Density	Settlement
Mean	0.5480	16.6769	0.7251	0.2002
Known variance	0.0307	1.5898	1.5898	0.0016
Observations	80	80	80	80
Hypothesized mean difference	0	0	0	–
z	17.3047	116.8222	3.7210	–
P(Z < = z) one-tail	0	0	9.92E-05	–
z Critical one-tail	1.6449	1.6449	1.6449	–
P(Z < = z) two-tail	0	0	0.000198	–
z Critical two-tail	1.9600	1.9600	1.9600	–

models. Sixteen performance metrics have been used to measure the training and testing performance of the employed models and reported in supplementary materials (SPM). The training and testing phase results have been discussed and analyzed below.

#### 4.1.1 Deep Learning Models

The deep learning models have been developed, trained, and tested using the MATLAB software. Table A (refer SPM) presents the training and testing performances of the deep learning models. The performance comparison of the MLPNN, RBFNN, and FFNN demonstrates that the FFNN model MS6 has attained over 99% accuracy in the training ( $R=0.9971$ ) and testing ( $R=0.9973$ ) phase. It has also been observed that model MS6 has outperformed the RBFNN and MLPNN models with the least prediction error in the training (RMSE=0.0030 m, MAE=0.0022 m, MAPE=1.1563%, WMAPE=0.0110 m, NMBE=0.0000) and testing (RMSE=0.0037 m, MAE=0.0032 m, MAPE=1.6706%, WMAPE=0.0156 m, NMBE=0.0001) phase. The a20, IOA, and IOS demonstrate the superiority of the FFNN model MS7 over RBFNN and MLPNN models. On the other side, the performance comparison of SVM, SVR, and Ls-SVM reveals that the Ls-SVM model MS3 has predicted pile group settlement with RMSE=0.0016 m, MAE=0.0015 m,  $R=0.9995$ , and MAPE=0.7644%, comparatively higher than SVM and SVR models. The TSFL model MS5 has predicted pile group settlement with the RMSE of 0.0080 m, MAE of 0.0079 m, and  $R$  of 0.9881.

Furthermore, the performance comparison of models MS3, MS5, and MS7 reveals that the least square support vector machine model MS3 has outperformed the other deep learning models with the higher training (RMSE=0.0039 m, MAE=0.0027 m,  $R=0.9952$ , MAPE=1.4675%, VAF=98.98, WMAPE=0.0134, NS=0.9895, PI=1.9763, BF=1.0032, NMBE=0.0001, MBE=0.0007, LMI=0.0814, RSR=0.1027, a20=100.00, IOA=0.9593, and IOS=0.0195) and testing (RMSE=0.0016 m, MAE=0.0015 m,  $R=0.9995$ , MAPE=0.7644%, VAF=99.90, WMAPE=0.0072, NS=0.9987, PI=1.9964, BF=0.9971, NMBE=0.0000, MBE=-0.0006, LMI=0.0378, RSR=0.0354, a20=100.00, IOA=0.9811, and

IOS=0.0075) performance. Figure 10a depicts the relationship between actual and predicted settlement for the Ls-SVM model MS3. The following observations have been mapped from the performance comparison of the deep learning models.

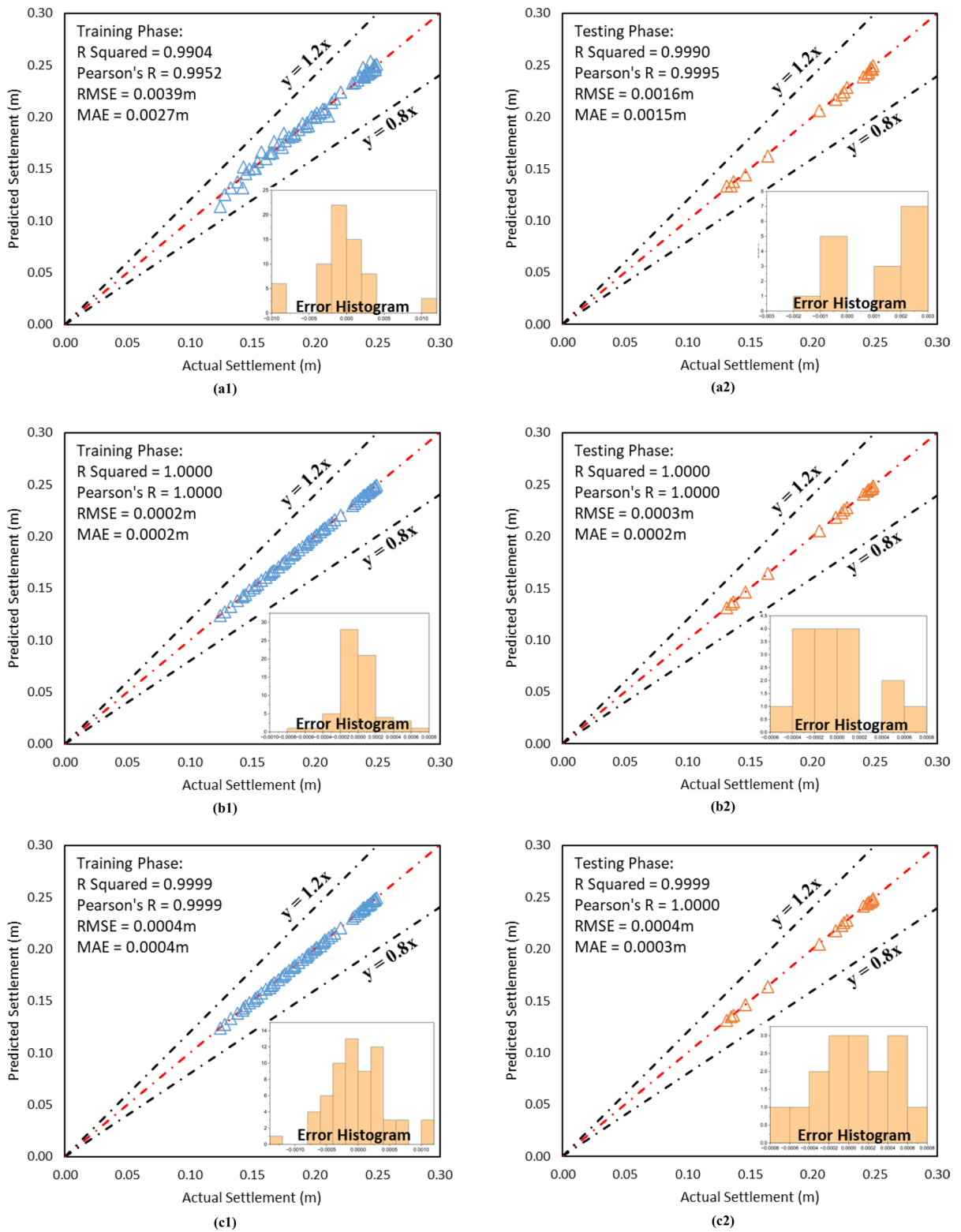
- The feed-forward neural network has performed better than RBFNN and MLPNN models because the problem is complex and requires more neurons. Also, the computational process runs in one direction only.
- The Ls-SVM model MS3 has attained higher performance because of the sum of the square error cost function.
- The Takagi–Sugeno Fuzzy Logics did not perform better because of the unavailability of proper membership functions and fuzzy rules from numerical data (Hong and Lee 1996).

#### 4.1.2 Hybrid Learning Models

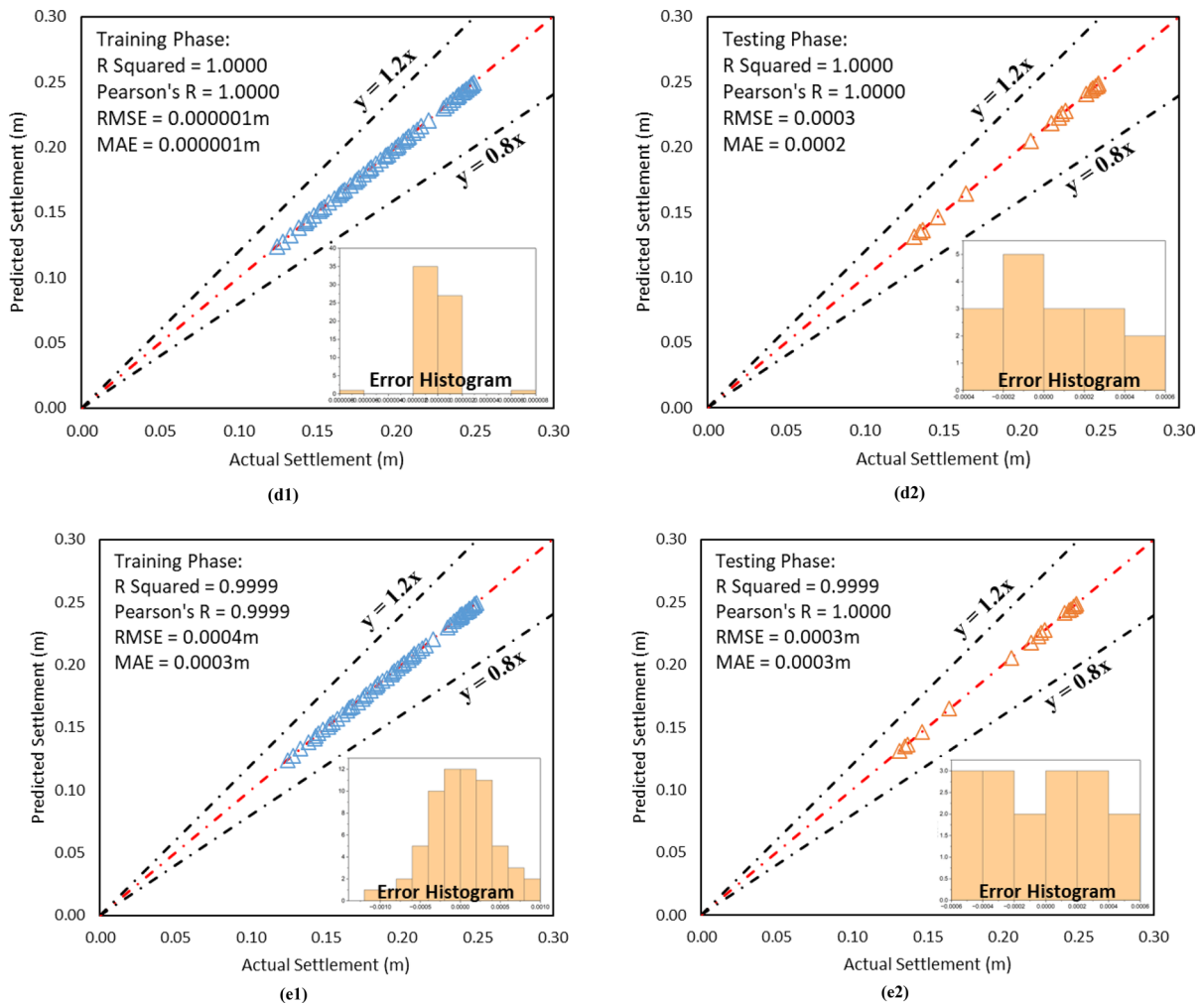
For employing the hybrid learning models, the relevance vector machine approach has been selected in this research. The Gaussian, exponential, linear, laplacian, sigmoid, and polynomial kernel functions have been implemented to develop 22 relevance vector machine (6GA-optimized SRVM, 6PSO-optimized SRVM, 5GA-optimized DRVM, and 5PSO-optimized DRVM) models. The training and testing performance of the RVM models has been summarized in Table A.

Table A (refer SPM) demonstrates that the GA-optimized exponential kernel function-based SRVM model MS9 has attained high performance in both phases (training  $R=1.0000$ , testing  $R=1.0000$ ). Model MS9 has predicted the pile group settlement with the least residuals in training (RMSE=0.0002 m, MAE=0.0002 m) and testing (RMSE=0.0003 m, MAE=0.0002 m) phase. The performance comparison of PSO-optimized SRVM shows that the PSO-optimized laplacian kernel function-based SRVM model has performed better than other PSO-optimized SRVM models. The PSO-optimized model MS17 has predicted pile group settlement with the RMSE of 0.0004 m, MAE of 0.0003 m, and  $R$  of 1.0000 in the testing phase.

Models MS9 and MS17 have been developed by implementing the exponential and laplacian kernel



**Fig. 10** Illustration of actual vs predicted plot for models **a** MS3, **b** MS9, **c** MS17, **d** MS23, and **e** MS25 in training (a1,b1,c1,d1,e1) and testing (a2, b2, c2, d2, e2) phase



**Fig. 10** (continued)

functions, and models have attained higher performance in the testing phase. Therefore, five GA-optimized DRVM models have been developed by implementing the different combinations of kernel functions, i.e., Exponential+Gaussian, Exponential+Linear, Exponential+Laplacian, Exponential+Sigmoid, and Exponential+Polynomial. On the other hand, five PSO-optimized DRVM models have been developed using the Laplacian+Gaussian, laplacian+Exponential, Laplacian+Linear, Laplacian+Sigmoid, and Laplacian+Polynomial combinations of kernel functions.

Table A (refer SPM) demonstrates that the GA-optimized exponential+sigmoid dual kernel function model MS23 has achieved higher training

( $R=1.0000$ ) and testing ( $R=1.0000$ ) performance than other GA-optimized DRVM models. Also, it has been observed that model MS23 has predicted the pile group settlement with the least prediction error in the testing phase, i.e.,  $RMSE=0.0003$  m,  $MAE=0.0002$  m. The performance comparison of PSO-optimized DRVM models reveals that the PSO-optimized Laplacian+Gaussian kernel function-based DRVM has outperformed the other PSO-optimized DRVM models with  $R$  of 1.0000,  $RMSE$  of 0.0003 m, and  $MAE$  of 0.0003 m, close to the ideal values.

Based on the performance comparison, it has been observed that the Ls-SVM model MS3, GA-optimized SRVM model MS9, PSO-optimized

SRVM model MS17, GA-optimized DRVM model MS23, and PSO-optimized DRVM model MS25 has attained higher performance ( $R > 0.95$ ) in the testing phase and recognized as the better-performing models. A relationship plot with an error histogram has been drawn for models MS9, MS7, MS23, and MS25, as shown in Fig. 10b–e.

The following observations have been drawn from the performance comparison of the GA-optimized SRVM, PSO-optimized SRVM, GA-optimized DRVM, and PSO-optimized DRVM models.

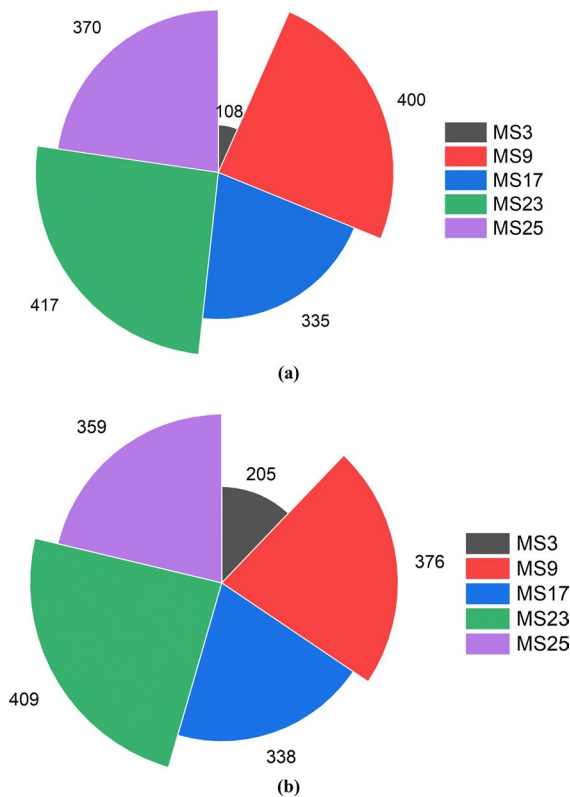
- The GA-optimized exponential kernel function-based SRVM model MS9 has performed better than other GA-optimized SRVM models because of its simplicity. Still, it decays the prediction error much more quickly.
- The PSO-optimized laplacian kernel function-based SRVM model MS17 has attained higher performance than other PSO-optimized SRVM models because it decays the error faster.
- The performance comparison of models MS9 and MS17 shows that model MS9 has predicted the pile group settlement, optimized by genetic algorithm. Hence, it can be stated that the genetic algorithm is the most suitable algorithm for single kernel function-based relevance vector machine models.
- Model MS23 has been employed using the combination of exponential and sigmoid kernel functions and optimized by the genetic algorithm. Model MS23 has performed better than other PSO-optimized DRVM models because it works as a two-layer perceptron neural network.
- The PSO-optimized Laplacian+Gaussian kernel function-based DRVM model MS25 has outperformed this study's other PSO-optimized DRVM model. The combination of both kernels demonstrates that the model MS25 is robust because the Gaussian kernel performs better if the prior information is not given, and Laplacian decays the error faster.
- The performance comparison of models MS23 and MS25 shows that the genetic algorithm is slightly better than the particle swarm optimization algorithm in predicting pile group settlement.

## 4.2 Analysis of Results

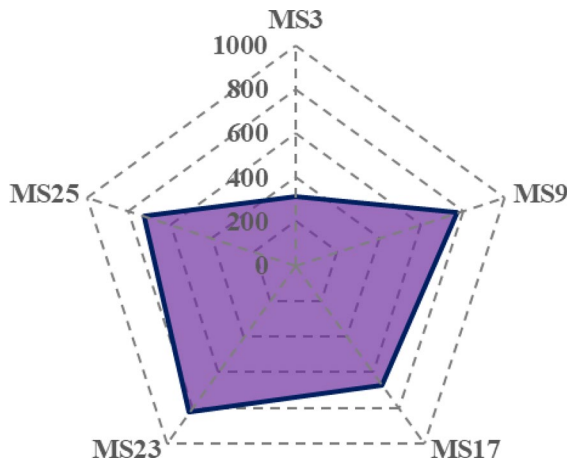
In this section, the performance results have been analyzed and presented by performing different tests and analyses. For this purpose, the score analysis, Anderson–darling test, uncertainty analysis, Wilcoxon test, overfitting, etc., has been performed and discussed.

### 4.2.1 Score Analysis

The score analysis compares the effectiveness of the best architectural models through statistical analysis. The model for choosing the optimal value for each performance indicator is given a score of  $n$  (in this study,  $n=29$ ; see best architectural models for soft computing that are taken into account in the analysis). The better and poorer training and testing examples for the models are shown by the higher and lower values of performance indicators in score analysis (Khatti and Grover 2023e). The next step is to add the performance indicator scores from each training and testing phase to determine the final model score. Finally, a model's overall score is computed by adding the training, testing, and validation score. The results obtained from the score analysis have been presented in Table B (refer SPM). Table B demonstrates that model MS23 has gained 417 and 409 scores in the training and testing phase, respectively, followed by models MS9, MS25, and MS17. Model MS3 has attained the lowest score in the training ( $=108$ ) and testing ( $=205$ ) phases. The graphical comparison of training and testing scores for the better-performing models is shown in Fig. 11 (a) and (b). Figure 11 presents the superiority of model MD23 in predicting pile group settlement, with a high score of 417 (in the training phase) and 409 (in the testing phase). The GA-optimized SRVM model MS9 scored higher than PSO optimized SRVM model MS17, i.e., 400 (in training) and 376 (in testing). Furthermore, the grand score has been calculated and presented in Fig. 12. Figure 12 depicts that model MS23 has attained an overall score of 826, which is higher than other models. Based on the overall score, GA-optimized exponential+sigmoid kernel function-based DRVM model MS23 has been recognized as the best architectural model in predicting the settlement of pile group in clay.



**Fig. 11** Illustration of score obtained by the better performing model in **a** training and **b** testing phase



**Fig. 12** Illustrates the overall score obtained by better-performing models

#### 4.2.2 Regression Error Characteristics (REC) Curve

The regression error characteristics curve is a regression version of the 2-D receiver operating characteristics (ROC) curve. The y-axis shows the percentage of predicted points that fall inside the tolerance, and the x-axis shows the error tolerance. The cumulative distribution function of the difference between experimental and expected values is calculated using this curve. The REC curve specifies the error amount as a squared residual or an absolute deviation. The area over the curve (AOC), referred to as the curve area, is a reliable measure of a regression model’s effectiveness. AOC should preferably be as small as feasible for a great regression model, and the curve should be positioned parallel to the y-axis. Figure 13 (a and b) represents the REC plot showing the error as "absolute deviation" for the pile group settlement for training and testing databases.

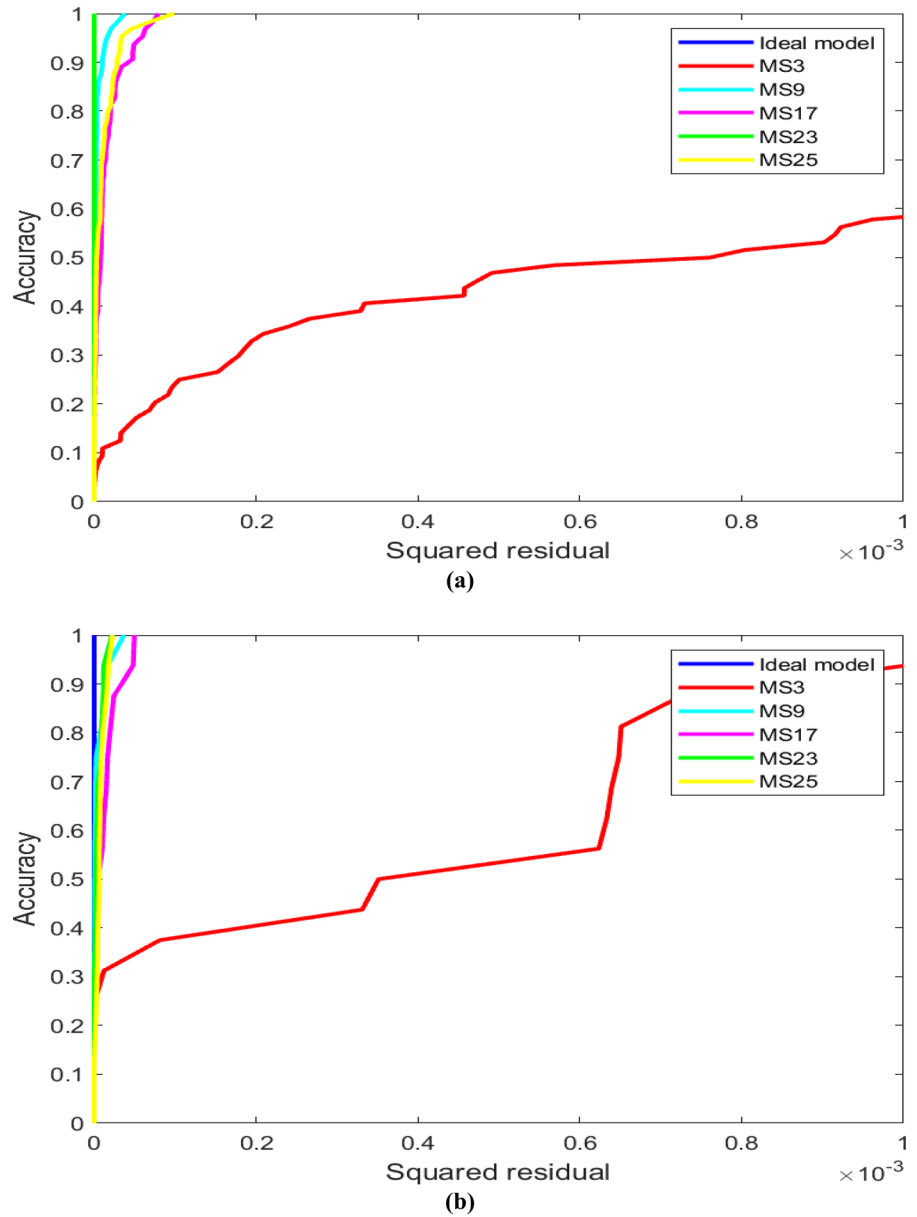
Figure 13a, b shows that model MS23 has attained the most negligible AOC value in the training and testing phase. The value of the AOC of the better-performing models is given in Table 8. It can be seen that the model MS23 has the smallest AOC (training = 6.67E-11 and testing = 4.02E-06) value compared to other models. Hence, model MS23 has been recognized as the best architectural model in assessing pile group settlement.

#### 4.2.3 Cross-Validation of Optimum Performance Model

In the present research, five models, MS3, MS9, MS17, MS23, and MS25, have been identified as the better-performing models in predicting pile group settlement in clay. For the cross-validation of the better-performing models, the computational cost analysis has been performed using k-fold = 10. The developed, trained, and tested models have been developed using k-fold = 5 (Khatti and Grover 2023d). Therefore, a comparison of the computational cost for models MS3, MS9, MS17, MS23, and MS25 has been drawn and presented in Table 9.

Table 9 demonstrates that model MS23 has attained the desired prediction with significantly less computational cost in k-fold 5 and 10. Therefore, it can be stated that model MS23 is the best architectural model for predicting pile group settlement.

**Fig. 13** Illustration of REC curve for the better performing models in **a** training and **b** testing phase



**Table 8** Details of AOC for REC curve

Phase	Actual	MS3	MS9	MS17	MS23	MS25
Train	0.00E+00	1.62E-03	3.22E-06	1.33E-05	6.67E-11	1.00E-05
Test	0.00E+00	3.91E-04	4.07E-06	1.17E-05	4.02E-06	6.86E-06

\*Bold values correspond to the best architectural model

**Table 9** Comparison of computational cost

k-fold value	MS3	MS9	MS17	MS23	MS25
k=5	0.0009	0.0024	0.0013	0.0030	0.0018
k=10	0.0011	0.0027	0.0015	0.0034	0.0020

#### 4.2.4 Anderson–Darling Test

For the hypothesis testing, parametric and nonparametric tests are performed. In this study, a non-parametric test, the "Anderson–Darling" test (AD),

has been performed to investigate the deviation of the outcomes. A statistical test called the AD test determines if a sample of data came from a population with a particular distribution. The AD test has been performed on the better-performing models (only for the validation phase) using the MiniTab Statistical Software.

Table 10 and Fig. 14a–f demonstrate the AD test results and probability plot. It can be seen that the p-value is 0.005 (Model: Actual), found to be less than the significance level, i.e., 0.05. Therefore, the AD test rejects the normality null hypothesis ( $H_0$ ). Also, the AD value of the model MS23 has been calculated closest to the actual value, confirming the superiority of the GA-optimized DRVM model MS23 over the other better-performing models.

#### 4.2.5 Uncertainty Analysis ( $UA_{95}$ )

Determining the reliability of any soft computing model is necessary to compute predictive targets accurately. In this research, uncertainly analysis (UA) has been performed to explain the quantitative prediction of error of the applied models in predicting the UCS of cohesive virgin soils. The UA test has been performed for the training and testing phases. Hence, the predictive outputs comparison with these actual data points is significant in predicting the reliability of the applied soft computing models, and UA is ideally suitable for this object. The UA test is performed by computing absolute error, MOE, StDev, SE, ME (@ 95% confidence level), WBC, UB, and LB, and presented in Table 11. A good model always contains a low value for the WCB (Bardhan et al. 2021).

Table 11 demonstrates that model MS23 has achieved the first rank in the training and testing phase in uncertainty analysis. Therefore, model MS23 has been recognized as the best architectural model for predicting the pile group settlement in clay.

#### 4.2.6 Wilcoxon Test

If a normal data distribution cannot be assumed, the one-sample Wilcoxon signed rank test is a nonparametric alternative to the one-sample t-test. It determines whether the sample's median equals a recognized standard value. In the present study, one sample Wilcoxon test has been performed for models MS3, MS9, MS17, MS23, and MS25. The results obtained for the soft computing models have been compared to the result obtained for actual pile group settlement data. Table 12 compares the Wilcoxon test for actual pile group settlement data and soft computing models.

Table 12 reveals that the model MS23 has predicted the pile group settlement with the upper and lower levels close to the actual data in the training and testing phase. Therefore, model MS23 has been recognized as the best architectural model in predicting pile group settlement.

#### 4.2.7 Overfitting of Soft Computing Models

The ratio of test RMSE to train RMSE is called overfitting of the soft computing model (Tenpe and Patel 2020; Khatti and Grover 2023c). Asteris et al. (2019) and Armaghani and Asteris (2021) stated that overfitting is a common problem in computational mechanics. It means the optimal model may predict the data used in training and development. Still, at the same time, for the other input variables than the one used in training and testing, the model may predict unusual values. For that purpose, experimental validation may be carried out by adopting different cases from the literature. In this study, the overfitting of the best architectural models MS3, MS9, MS17, MS23, and MS25 has been computed and graphically presented in Fig. 15.

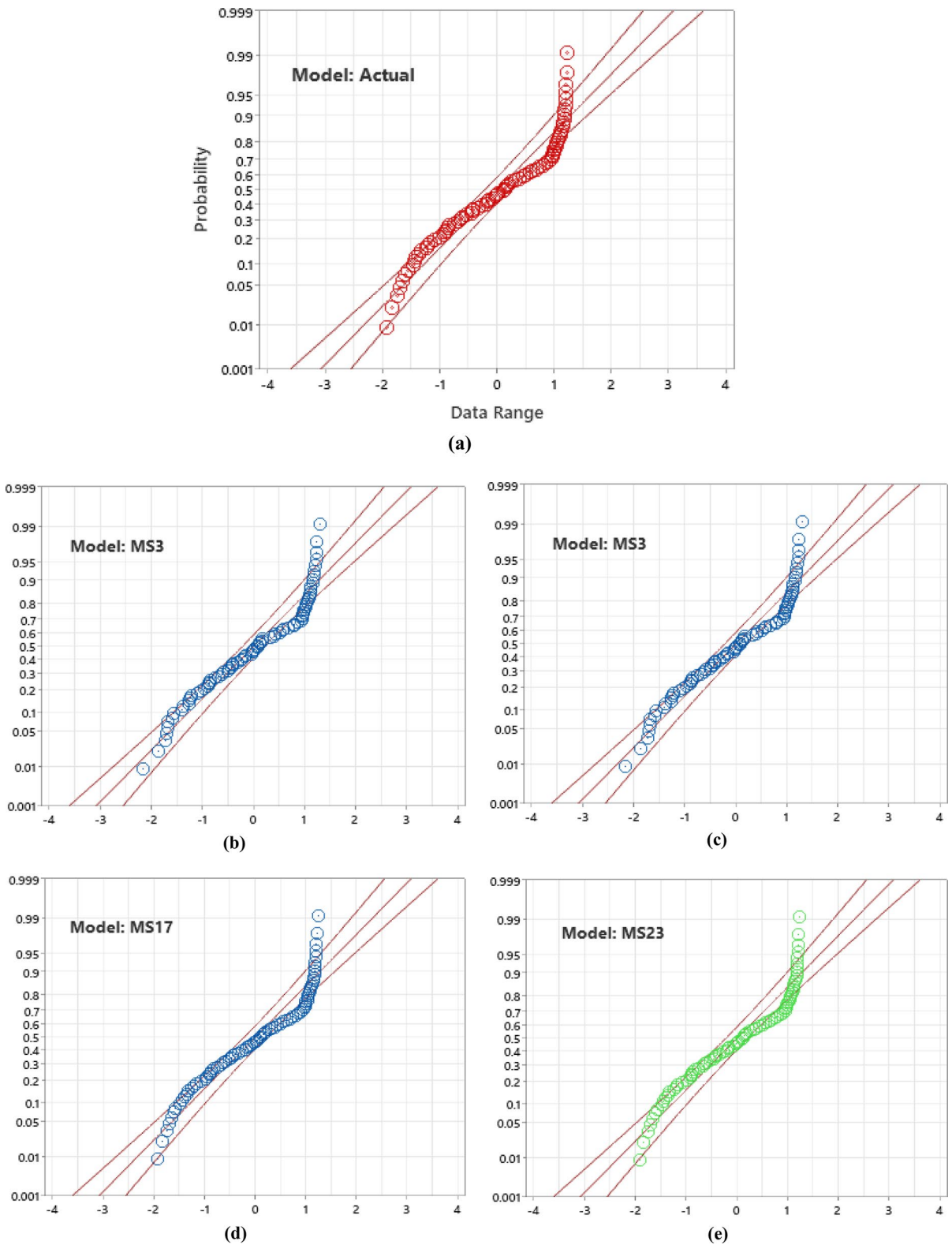
Figure 15 shows that model MS23 has attained the highest overfitting, i.e., 221.281 because all input variable has problematic multicollinearity ( $VIF > 10$ ). Still, model MS23 has performed well and predicted the pile group settlement with the least residuals in this research. Model MS3 has attained the least overfitting, i.e., 0.404. Based on the performance comparison, all

**Table 10** Details of AD normality test for the whole database

\*Bold values correspond to the best architecture model

Parameters	Actual	MS3	MS9	MS17	MS23	MS25
AD Value	2.139	1.914	2.147	2.143	<b>2.14</b>	2.142
p-value	<0.005					





**Fig. 14** Illustration of Anderson–Darling test for models **a** Actual, **b** MS3, **c** MS9, **d** MS17, **e** MS23, and **f** MS25

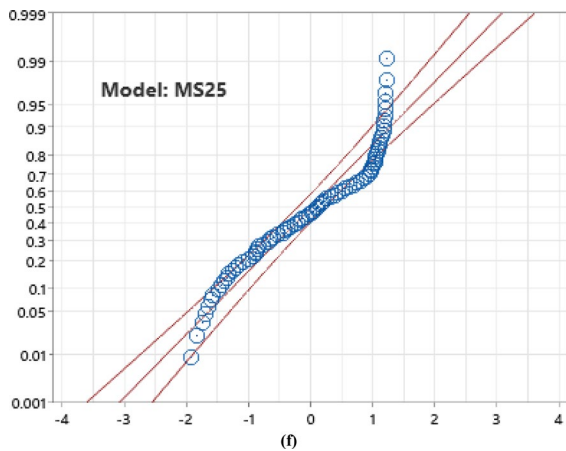


Fig. 14 (continued)

models have attained performance over 95%, presenting highly capable of predicting pile group settlement in clay.

### 4.3 Discussion of Results

In this research, MLPNN, SVM, Ls-SVM, SVR, TSFL, RBFNN, FFNN, and RVM models have been developed, trained, tested, and analyzed to determine the optimum performance model for predicting pile group settlement. Each GA and PSO algorithm has developed and optimized the single and dual kernel function-based RVM models. The database has been collected from the published article. The multicollinearity analysis has been performed for the database and found that the input variables contain the problematic

multicollinearity level. However, the ANOVA and Z tests reject the null hypothesis for the database used in this research. For performance analysis, sixteen performance metrics have been used and compared. Based on the performance comparison, it has been observed that (i) the FFNN model attains a higher performance than RBFNN and MLPNN models, (ii) the Ls-SVM model performance is better than SVM and SVR models, (iii) also, the Ls-SVM model gains higher performance than FFNN and TSFL models in the presence of multicollinearity. The performance comparison of GA and PSO-optimized SRVM models reveals that the GA-optimized SRVM model has predicted pile group settlement with the least error. Similarly, the GA-optimized DRVM model attains better performance and accuracy than the PSO-optimized DRVM model. The performance comparison also reveals that the GA-optimized DRVM model MS23 has outperformed the adopted soft computing models and is recognized as the best architectural model. Model MS23 has attained a higher score and rank in this research work. The analysis of results for model MS23 demonstrates that the combined kernel function-based RVM model, i.e., DRVM, is a robust model because both kernels strengthen the model’s performance.

The complex database has problematic multicollinearity, and overfitting is generated due to (i) if the model is simple and the data is complex or (ii) if the model is complex and the data is simple, or (iii) if both are complex. The present study presents that model MS23 is a complex model that combines exponential and sigmoid kernel functions. Also, it has

**Table 11** Results of uncertainty analysis for the better-performing models

Model ID	MOE	SD	SE	ME	LB	UB	WCB	Rank
Training phase								
MS3	0.0027	0.0028	0.0004	0.0007	0.0013	0.0066	0.0053	5
MS9	0.0002	0.0002	0.0000	0.0000	0.0007	0.0010	0.0003	2
MS17	0.0004	0.0003	0.0000	0.0001	0.0003	0.0007	0.0004	4
<b>MS23</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>1</b>
MS25	0.0003	0.0003	0.0000	0.0001	0.0001	0.0007	0.0006	3
Testing phase								
MS3	0.0041	0.0030	0.0007	0.0015	0.0018	0.0099	0.0081	5
MS9	0.0002	0.0002	0.0001	0.0001	0.0002	0.0010	0.0008	2
MS17	0.0003	0.0002	0.0001	0.0001	0.0001	0.0006	0.0005	4
<b>MS23</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>1</b>
MS25	0.0003	0.0002	0.0001	0.0001	0.0001	0.0007	0.0005	3

\*Bold values correspond to the best architectural model

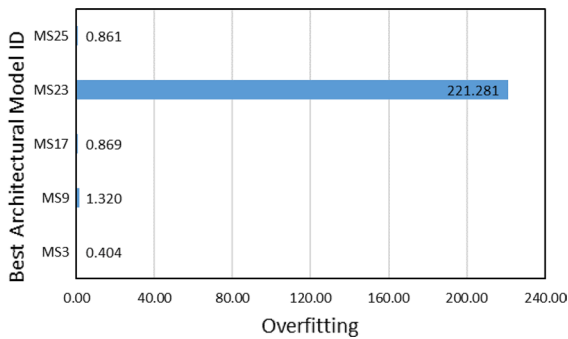
**Table 12** Results of the Wilcoxon test

Case	Phase	Confidence interval		Achieved confidence (%)
		LL	UL	
Actual	Train	0.1896	0.2098	95.03
	Test	0.1820	0.2355	94.75
MS3	Train	0.1900	0.2113	95.03
	Test	0.1811	0.2344	94.75
MS9	Train	0.1897	0.2097	95.03
	Test	0.1822	0.2354	94.75
MS17	Train	0.1897	0.2098	95.03
	Test	0.1821	0.2352	94.75
MS23	Train	0.1896	0.2098	95.30
	Test	0.1821	0.2354	94.75
MS25	Train	0.1895	0.2098	95.03
	Test	0.1821	0.2351	94.75

\*Bold values correspond to the best architectural model

**Table 13** Comparison of model MS23 and literature survey

Approach	LSSVM	GMDH	GPR	Model MS23
RMSE	0.0190	0.0200	0.0200	0.0003
MAE	0.0160	0.0160	0.0070	0.0002
R	1.0000	1.0000	1.0000	1.0000
MAPE	0.0000	0.0000	0.0000	0.1091
VAF	87.88	88.89	99.74	100.00
WMAPE	0.0860	0.0800	0.0920	0.0011
NS	1.0000	1.0000	1.0000	1.0000
PI	1.8700	1.8700	1.8300	1.9997
BF	1.0800	1.0800	1.0800	0.9998
NMBE	8.1300	8.0000	7.6600	0.0000
MBE	0.0000	0.0000	0.0000	-0.0001
LMI	1.0000	0.9990	1.0000	0.0059
RSR	0.0000	0.0000	0.0000	0.0061
a20	–	–	–	100.00
IOA	–	–	–	0.9971
IOS	–	–	–	0.0013



**Fig. 15** Illustration of overfitting for the best architectural models

been optimized by a genetic algorithm. Still, model MS23 has performed better than other models. Furthermore, the performance of model MS23 has been compared with the test performance of previously published models by Kumar and Samui (2020), as shown in Table 13.

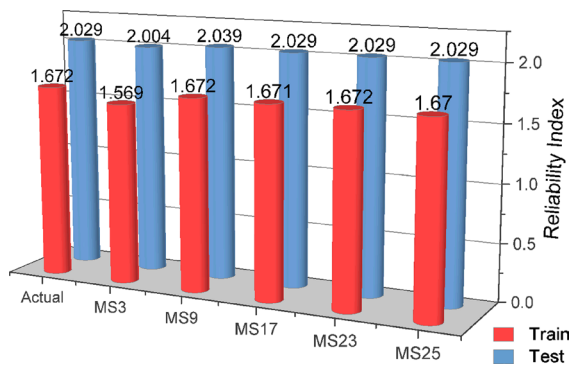
Based on the comparison presented in Table 13, model MS23 has been recognized as the optimum performance model for predicting the pile group settlement in clay.

#### 4.4 Reliability Analysis

The qualities of measurement scales and the components that make up the scales can be studied through

reliability analysis. In addition to providing data on the correlations between the scale’s constituent items, the reliability analysis technique creates various regularly used scale reliability measures. In this research, 80 data points have been used to perform the reliability analysis in predicting the pile group settlement using soft computing approaches. Still, it is questionable whether an amount of data points covers the full range of parameter values sufficient for reliability analysis. Cavaleri et al. (2017), Asteris et al. (2020), Lu et al. (2020), and Huang et al. (2020) reported that a sufficient amount of database is not essential for a high amount of data points, but rather data points that cover a wide range of combination of independent variable values, thus assisting in the ability of models to solve the problem. The requirement of a capable and reliable database is crucial in the case of laboratory databases. The reliability analysis has been performed for MS3, MS9, MS17, MS23, and MS25 models and compared with actual values. The results obtained from the reliability analysis are presented in Fig. 16.

Figure 16 depicts that models MS3, MS9, MS17, MS23, and MS25 can predict the settlement of the pile group. Still, model MS23 is the optimum performance model for predicting the pile group settlement. Model MS23 has attained exact reliability



**Fig. 16** Reliability index values for soft computing models

indexes in both phases, i.e., 1.671 (in training) and 2.029 (in testing).

## 5 Summary and Conclusions

The present study has been carried out to introduce an optimum performance soft computing model for predicting the settlement of pile group in clay. For that purpose, MLPNN, SVM, Ls-SVM, SVR, TSFL, RBFNN, FFNN, and RVM models have been developed, trained, tested, and analyzed using data available in the literature. Twenty-two RVM models (6 GA-optimized SRVM, 6 PSO-optimized SRVM, 5 GA-optimized DRVM, and 5 PSO-optimized DRVM) have been developed and analyzed. Sixteen performance metrics have been implemented in this research to identify the optimum performance model. The following conclusions are mapped through this study.

- *Capabilities of Models* (i) FFNN is more competent than RBFNN and MLPNN, (ii) Ls-SVM is better than SVM and SVR, (iii) Ls-SVM is significantly better than FFNN and TSFL, (iv) exponential kernel function is more accurate than other kernel function in case of GA-optimized SRVM model, (v) laplacian kernel function is potent than other kernel function in case of PSO-optimized SRVM models, (vi) the implementation of second kernel function enhance the performance and accuracy of SRVM model, (vii) exponential with sigmoid kernel function is better than other kernel combinations in case of GA-optimized DRVM

model, (viii) laplacian with gaussian kernel function is better than other kernel sets in case of PSO-optimized DRVM model, (ix) GA-optimized SRVM and DRVM models perform better than PSO-optimized SRVM and DRVM models.

- *Impact of Multicollinearity* (i) multicollinearity affects the performance of deep learning models, i.e., MLPNN, SVM, Ls-SVM, SVR, TSFL, RBFNN, and FFNN, than hybrid learning models, (ii) models based on the hybrid learning approach, i.e., RVM, attains the high overfitting in the testing phase, (iii) the hybrid learning models perform potentially in the presence of multicollinearity.
- As per the statistical clause for the hypothesis selection, the ANOVA and Z tests accept the research hypothesis for the present research.
- The results of score analysis, REC plot, computational cost, cross-validation, uncertainty analysis, AD test, Wilcoxon test, and comparison with available models introduce GA-optimized DRVM model MS23 as an optimum performance model for reliability analysis of the settlement of pile group in clay.

To sum up, the GA-optimized DRVM model is successfully employed in this research for reliability analysis of the pile group settlement in clay. The study demonstrates the high capabilities of the GA-optimized DRVM model, and it may be suggested that the GA-optimized DRVM model can be used to solve other geotechnical issues. One of the limitations of the employed machine learning models in this work is to determine the optimal structure using different analyses. Therefore, optimizing the coefficients/weights of the employed models using metaheuristic optimization algorithms is suggested. Different metaheuristics algorithms, such as squirrel search algorithm (SSA), improved squirrel search algorithm (ISSA), grey wolf optimizer (GWO) algorithm, random walk grey wolf optimizer (RW\_GWO) algorithm, sailfish optimizer (SAO) algorithm, may be implemented for models and comparison may be drawn with the present study. Also, a large database of pile group settlement can be utilized to develop another hybrid RVM model. The present research will help geotechnical engineers/designers determine the settlement of the pile group without performing experimental procedures. This research will save time and require fewer human resources. As per the

author's knowledge, the hybrid relevance vector machine models have been developed, trained, tested, and analyzed for reliability analysis of pile group settlement in clay for the first time.

## 6 Software Support

MATLAB R2020a: for employing soft computing models, analysis, evaluation, and prediction. Origin Lab 2022b: for graphical presentations and analysis. MiniTab Statistical Software: for statistical analysis.

**Author Contribution** JK: Main author, conceptualization, literature review, manuscript preparation, application of AI models, relevance vector machine model development, methodological development, statistical analysis, detailing, and overall analysis; HS: literature review, manuscript preparation, introduction and literature writing, theory of adopted methods, soft computing model development, detailing, overall analysis; KSG: conceptualization, comprehensive analysis, manuscript finalization, detailed review, and editing.

**Funding** No funding was received in assisting the preparation of this manuscript.

**Data Availability** All data, models, and code generated or used during the study appear in the submitted article. The database used in this research was collected from the literature.

### Declarations

**Conflict of interests** The authors declare no conflict of interest.

## References

- Abu-Kiefa MA (1998) General regression neural networks for driven piles in cohesionless soils. *J Geotech Geoenviron Eng ASCE* 124(12):1177–1185. [https://doi.org/10.1061/\(ASCE\)1090-0241\(1998\)124:12\(1177\)](https://doi.org/10.1061/(ASCE)1090-0241(1998)124:12(1177))
- Adarsh S, Dhanya R, Krishna G, Merlin R, Tina J (2012) Prediction of ultimate bearing capacity of cohesionless soils using soft computing techniques. *Int Schol Res Notices*
- Afnan HA, Abbas HO, Abd SH, Mohamed SM (2023) Behavior of pile group constructed on clayey soil under el-centro seismic. *Diyala J Eng Sci* 16(1):44–54. <https://doi.org/10.24237/djes.2023.16105>
- Alhashmi AE, El Naggat MH, Oudah F (2023) Performance of axially loaded defective pile groups in sand: capacity and serviceability evaluation. *Innov Infrastruct Solut.* <https://doi.org/10.1007/s41062-023-01086-w>
- Alkroosh I, Nikraz H (2011a) Correlation of pile axial capacity and CPT data using gene expression programming. *Geotech Geol Eng* 29:725–748. <https://doi.org/10.1007/s10706-011-9413-1>
- Alkroosh I, Nikraz H (2011b) Simulating pile load-settlement behavior from CPT data using intelligent computing. *Central Eur J Eng* 1(3):295–305. <https://doi.org/10.2478/s13531-011-0029-2>
- Andreas Z (1994) *Simulation neuronaler netze-simulation of neural networks*. Addison-Wesley
- Apostolopoulou M, Asteris PG, Armaghani DJ, Douvika MG, Lourenço PB, Cavaleri L, Bakolas A, Moropoulou A (2020) Mapping and holistic design of natural hydraulic lime mortars. *Cem Concr Res* 136:106167. <https://doi.org/10.1016/j.cemconres.2020.106167>
- Ardakani A, Kordnaej A (2019) Soil compaction parameters prediction using GMDH-type neural network and genetic algorithm. *Eur J Environ Civ Eng* 23(4):449–462. <https://doi.org/10.1080/19648189.2017.1304269>
- Armaghani DJ, Asteris PG (2021) A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. *Neural Comput Appl* 33(9):4501–4532. <https://doi.org/10.1007/s00521-020-05244-4>
- Armaghani DJ, Asteris PG, Fatemi SA, Hasanipanah M, Tarinejad R, Rashid ASA, Huynh VV (2020) On the use of neuro-swarm system to forecast the pile settlement. *Appl Sci* 10(6):1904. <https://doi.org/10.3390/app10061904>
- Asteris PG, Armaghani DJ, Hatzigeorgiou GD, Karayannis CG, Pilakoutas K (2019) Predicting the shear strength of reinforced concrete beams using Artificial Neural Networks. *Comput Concr* 24(5):469–488. <https://doi.org/10.12989/cac.2019.24.5.469>
- Asteris PG, Apostolopoulou M, Armaghani DJ, Cavaleri L, Chountalas AT, Guney D, Hajihassani M, Hasanipanah M, Khandelwal M, Karamani C, Koopialipoor M (2020) On the metaheuristic models for the prediction of cement-metakaolin mortars compressive strength. *Metaheuristic Comput Appl.* <https://doi.org/10.12989/mca.2020.1.1.063>
- Asteris PG, Koopialipoor M, Armaghani DJ, Kotsonis EA, Lourenço PB (2021a) Prediction of cement-based mortars compressive strength using machine learning techniques. *Neural Comput Appl* 33(19):13089–13121. <https://doi.org/10.1007/s00521-021-06004-8>
- Asteris PG, Lourenço PB, Hajihassani M, Adami CEN, Lemonis ME, Skentou AD, Marques R, Nguyen H, Rodrigues H, Varum H (2021b) Soft computing-based models for the prediction of masonry compressive strength. *Eng Struct* 248:113276. <https://doi.org/10.1016/j.engstruct.2021.113276>
- Balakumar V, Kalaiarasi V, Ilamparuthi K (2005) Experimental and analytical study on the behavior of circular piled raft on sand. In: *Proceedings of the 16th international conference on soil mechanics and geotechnical engineering*. Osaka, Japan. <https://doi.org/10.3233/978-1-61499-656-9-1943>
- Balakumar V, Ilamparuthi K (2007) Performance monitoring of a piled raft foundation of twelve storied building and analytical validation. *Indian Geotech J* 37(2):94–115
- Bao X, Cheng Z, Shen J, Zhang X, Chen X, Cui H (2023) Study on bearing capacity of reinforced composite pipe pile group in reclaimed stratum under vertical load. *J Mar Sci Eng.* <https://doi.org/10.3390/jmse11030597>

- Baziar MH, Ghorbani A, Katzenbach R (2009) Small-scale model test and three-dimensional analysis of pile-raft foundation on medium-dense sand. *Int J Civ Eng* 7(3):170–175
- Berardi R, Bovolenta R (2005) Pile settlement evaluation using field stiffness nonlinearity. *Proc Inst Civ Eng Geotech* 158:35–44. <https://doi.org/10.1680/geng.2005.158.1.35>
- Bhartiya P, Chakraborty T, Basu D (2021) Prediction of piled raft settlement using soil subgrade modulus in consolidating clays. *Pract Period Struct Des Constr* 26(4):04021037. [https://doi.org/10.1061/\(ASCE\)SC.1943-5576.0000608](https://doi.org/10.1061/(ASCE)SC.1943-5576.0000608)
- Bohn C, Lopes dos Santos A, Frank R (2017) Development of axial pile load transfer curves based on instrumented load tests. *J Geotech Geoenviron Eng ASCE* 143(1):04016081. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001579](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001579)
- Bourgeois E, Buhan P, Hassen G (2012) Settlement analysis of piled-raft foundations by means of a multiphase model accounting for soil-pile interactions. *Comput Geotech* 46:26–38. <https://doi.org/10.1016/j.compgeo.2012.05.015>
- Broomhead D, Lowe D (1988) Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal Signals and Radar Establishment Malvern, United Kingdom
- Candela JQ (2004) Learning with uncertainty—gaussian processes and relevance vector machines. Technical University of Denmark, Copenhagen, pp 1–152
- Cavaleri L, Chatzarakis GE, Di-Trapani F, Douvika MG, Roinos K, Vaxevanidis NM, Asteris PG (2017) Modeling of surface roughness in electro-discharge machining using artificial neural networks. *Adv Mater Res* 6(2):169. <https://doi.org/10.12989/amr.2017.6.2.169>
- Chen S, Gu C, Lin C, Zhang K, Zhu Y (2021) Multi-kernel optimized relevance vector machine for probabilistic prediction of concrete dam displacement. *Eng Comput* 37(3):1943–1959. <https://doi.org/10.1007/s00366-019-00924-9>
- Chen Y, Ma S, Ren Y, Chen R, Bian X (2021a) Experimental study on cyclic settlement of piles in silt soil and its application in high-speed railway design. *Transp Geotech* 27:100496. <https://doi.org/10.1016/j.trgeo.2020.100496>
- Cortes C, Vapnik V (1995) Support vector networks. *Mach Learn* 20:273–297. <https://doi.org/10.1007/BF00994018>
- Crispin JJ, Vardanega PJ, Mylonakis G (2019) Prediction of pile settlement using simplified models. *Proceedings of the XVII ECSMGE-2019-0388*, pp 1–9. <https://doi.org/10.32075/17ECSMGE-2019-0388>
- Cui J, Li J, Zhao G (2019) Long-term time-dependent load-settlement characteristics of a driven pile in clay. *Comput Geotech* 112:41–50. <https://doi.org/10.1016/j.compgeo.2019.04.007>
- Dai G, Salgado R, Gong W, Zhang Y (2012) Load tests on full-scale bored pile groups. *Can Geotech J* 49(11):1293–1308. <https://doi.org/10.1139/t2012-087>
- Das BM (2008) *Advanced soil mechanics*: Tylor & Frances
- Fattahi H, Hasanipanah M (2021) Prediction of blast-induced ground vibration in a mine using relevance vector regression optimized by metaheuristic algorithms. *Nat Resour Res* 30(2):1849–1863. <https://doi.org/10.1007/s11053-020-09764-7>
- Ghorbani B, Arulrajah A, Narsilio G, Horpibulsuk S, Bo MW (2020) Development of genetic-based models for predicting the resilient modulus of cohesive pavement subgrade soils. *Soils Found* 60(2):398–412. <https://doi.org/10.1016/j.sandf.2020.02.010>
- Gomes Filho FVC, Moura AS (2021) Proposal for considering the group effect in the prediction of settlements in pile groups through load transfer methods. *Soils Rocks*. <https://doi.org/10.28927/SR.2021.061320>
- Hair JF, Ortinau DJ, Harrison DE (2013) *Essentials of marketing research*. Oxford University Press, New York
- Hakro MR, Aneel K, Zaheer A, Mujahid A, Fahid A, Roman F, Sergey K, Alexander K, Linar S (2022) Numerical analysis of piled-raft foundations on multi-layer soil considering settlement and swelling. *Buildings* 12(3):356. <https://doi.org/10.3390/buildings12030356>
- Hasanzadehshooili H, Lakirouhani A, Medzvieckas J (2012) Superiority of artificial neural networks over statistical methods in prediction of the optimal length of rock bolts. *J Civ Eng Manag* 18(5):655–661. <https://doi.org/10.3846/13923730.2012.724029>
- Hong TP, Lee CY (1996) Induction of fuzzy rules and membership functions from training examples. *Fuzzy Sets Syst* 84(1):33–47. [https://doi.org/10.1016/0165-0114\(95\)00305-3](https://doi.org/10.1016/0165-0114(95)00305-3)
- Hongwei G (2011) A simple algorithm for fitting a Gaussian function. *IEEE Sign Proc Mag* 28(9):134–137. <https://doi.org/10.1109/MSP.2011.941846>
- Huang J, Asteris PG, Manafi-Khajeh-Pasha S, Mohammed AS, Hasanipanah M (2020) A new auto-tuning model for predicting the rock fragmentation: a cat swarm optimization algorithm. *Eng Comput* 38:1–12. <https://doi.org/10.1007/s00366-020-01207-4>
- Ibrahim NM, Rahim NL, Amat RC, Salehuddin S, Ariffin NA (2012) Determination of plasticity index and compression index of soil at Perlis. *APCBEE Proc* 4:94–98. <https://doi.org/10.1016/j.apcbee.2012.11.016>
- Katzenbach R, Arslan V, Moorman C (2000) Numerical stimulations of combined piled raft foundations for the new high rise building. Max in Frankfurt am main. In: *Proceedings of the 2nd international conference on soil structure interaction in urban civil engineering*, ETH, Zurich, Switzerland
- Khatti J, Grover KS (2021) Computation of permeability of soil using artificial intelligence approaches. *Int J Eng Adv Technol* 11(1):257–266. <https://doi.org/10.35940/ijeat.A3220.1011121>
- Khatti J, Grover KS (2023a) Prediction of compaction parameters for fine-grained soil: Critical comparison of the deep learning and standalone models. *J Rock Mech Geotech Eng* in Press. <https://doi.org/10.1016/j.jrmge.2022.12.034>
- Khatti J, Grover KS (2023b) CBR prediction of pavement materials in unsoaked condition using LSSVM, LSTM-RNN, and ANN approaches. *Int J Pavement Res Technol*. <https://doi.org/10.1007/s42947-022-00268-6>
- Khatti J, Grover KS (2023c) Assessment of fine-grained soil compaction parameters using advanced soft computing

- techniques. *Arab J Geosci* 16(3):208. <https://doi.org/10.1007/s12517-023-11268-6>
- Khatti J, Grover KS (2023d) Prediction of UCS of fine-grained soil based on machine learning part 2: comparison between hybrid relevance vector machine and Gaussian process regression. *Multiscale Multidiscip Model Exp Des*. <https://doi.org/10.1007/s41939-023-00191-8>
- Khatti J, Grover KS (2023e) Prediction of UCS of fine-grained soil based on machine learning part 1: multi-variable regression analysis, gaussian process regression, and gene expression programming. *Multiscale Multidiscip Model Exp Des*. <https://doi.org/10.1007/s41939-022-00137-6>
- Khatti J, Grover KS (2023f) Prediction of compaction parameters of compacted soil using LSSVM, LSTM, LSBoostRF, and ANN. *Innov Infrastruct Solut* 8(2):76. <https://doi.org/10.1007/s41062-023-01048-2>
- Khatti J, Grover KS (2023g) Prediction of soaked CBR of fine-grained soils using soft computing techniques. *Multiscale Multidiscip Model Exp Des* 6(1):97–121. <https://doi.org/10.1007/s41939-022-00131-y>
- Kordjazi A, Nejad FP, Jaksa MB (2014) Prediction of ultimate axial load-carrying capacity of piles using a support vector machine based on CPT data. *Comput Geotech* 55:91–102. <https://doi.org/10.1016/j.compgeo.2013.08.001>
- Kumar M, Samui P (2020) Reliability analysis of settlement of pile group in clay using LSSVM, GMDH. *GPR Geotech Geol Eng* 38:6717–6730. <https://doi.org/10.1007/s10706-020-01464-6>
- Kumar A, Houdhury D, Katzenbach R (2016) Effect of earthquake on combined pile-raft foundation. *Int J Geomech*. [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000637](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000637)
- Lai JX, Liu HQ, Qiu JL, Chen J (2016) Settlement analysis of saturated tailings dam treated by CFG pile composite foundation. *Adv Mater Sci Eng*. <https://doi.org/10.1155/2016/7383762>
- Lee IM, Lee JH (1996) Prediction of pile bearing capacity using artificial neural networks. *Comput Geotech* 18(3):189–200. [https://doi.org/10.1016/0266-352X\(95\)00027-8](https://doi.org/10.1016/0266-352X(95)00027-8)
- Lee K, Shin S, Lee H, Kim D (2020) Analysis of pile behaviors with friction resistance of skin of steel pipe pile in ground where settlement is predicted. *J Korean Geotech Soc* 36(11):107–117
- Li L, Deng Y (2023) Analysis of settlement of group pile foundation in linear viscoelastic soil. *Adv Civ Eng Mater*. <https://doi.org/10.1155/2023/3207304>
- Li L, Gong W (2019) Prediction of nonlinear vertical settlement of a pile group consisting of new and existing displacement piles in clay strata. *Soils Found* 59(5):1336–1348. <https://doi.org/10.1016/j.sandf.2019.06.001>
- Li TZ, Pan Q, Dias D (2021) Active learning relevant vector machine for reliability analysis. *Appl Math Model* 89:381–399. <https://doi.org/10.1016/j.apm.2020.07.034>
- Lu S, Koopialipoor M, Asteris PG, Bahri M, Armaghani DJ (2020) A novel feature selection approach based on tree models for evaluating the punching shear capacity of steel fiber-reinforced concrete flat slabs. *Materials* 13(17):3902. <https://doi.org/10.3390/ma13173902>
- Mandolini A, Russo G, Viggiani C (2005) Pile foundations: experimental investigations, analysis and design. In: *Proceedings of the 16th international conference on soil mechanics and geotechnical engineering, Osaka, vol 1, pp 177–213*
- Masani AK, Vanza MG (2018) Effect of vertical load on the lateral behavior of group of piles. *Int J Adv Res Innovat Ideas Educ* 4(2)
- Menhrotra K, Mohan C, Ranka S (1990) *Elements of artificial neural networks*. MIT Press, Cambridge
- Mentaschi L, Besio G, Cassola F, Mazzino A (2013) Problems in RMSE-based wave model validations. *Ocean Model* 72:53–58. <https://doi.org/10.1016/j.ocemod.2013.08.003>
- Mittal M, Satapathy SC, Pal V, Agarwal B, Goyal LM, Parwekar P (2021) Prediction of coefficient of consolidation in soil using machine learning techniques. *Microprocess Microsyst* 82:103830. <https://doi.org/10.1016/j.micpro.2021.103830>
- Murthy VNS (2002) *Principles and practices of soil mechanics and foundation engineering*. Marcel Dekker Inc
- Oteuil A, Oralbek A, Mukhamet T, Moon SW, Kim J, Tokbolat S, Satyanaga A (2022) Robust analysis and design of bored pile considering uncertain parameters. *Indian Geotech J* 52:720–734. <https://doi.org/10.1007/s40098-021-00588-7>
- Pal M, Deswal S (2010) Modelling pile capacity using Gaussian process regression. *Comput Geotech* 37(7–8):942–947. <https://doi.org/10.1016/j.compgeo.2010.07.012>
- Ponomaryov A, Sychkina E (2022) Effect of clay compaction around driven pile and prediction of pile settlement. In *E3S web of conferences, vol 363*. EDP Sciences. <https://doi.org/10.1051/e3sconf/202236302016>
- Pooya Nejad F, Jaksa MB (2017) Load-settlement behavior modeling of single piles using artificial neural networks and CPT data. *Comput Geotech* 89:9–21. <https://doi.org/10.1016/j.compgeo.2017.04.003>
- Pooya Nejad F, Jaksa MB, Kakhi M, McCabe BA (2009) Prediction of pile settlement using artificial neural networks based on standard penetration test data. *Comput Geotech* 36(7):1125–1133. <https://doi.org/10.1016/j.compgeo.2009.04.003>
- Poulos HG (2001) Piled raft foundations: design and applications. *Geotechnique* 51(2):95–113. <https://doi.org/10.1680/geot.2001.51.2.95>
- Poulos HG, Davis EH (1980) *Pile foundation analysis and design*. Wiley
- Poulos HG (2006) Pile group settlement estimation-research to practice. *Geo shanghai international conference*
- Raja MNA, Jaffar STA, Bardhan A, Shukla SK (2022) Predicting and validating the load-settlement behavior of large-scale geosynthetic-reinforced soil abutments using hybrid intelligent modeling. *J Rock Mech Geotech Eng* 15(3):773–788. <https://doi.org/10.1016/j.jrmge.2022.04.012>
- Randolph MF (2003) Science and empiricism in pile foundation design. 43rd Rankine Lecture. *Geotechnique* 53(10):847–875. <https://doi.org/10.1680/geot.2003.53.10.847>
- Ray R, Samui P, Roy LB (2023) Reliability analysis of a shallow foundation on clayey soil based on settlement criteria. *J Curr Sci Technol* 13(1):91–106
- Saha R, Dutta SC, Haldar S (2015) Seismic response of soil pile raft-structure system. *J Civ Eng Manag*

- 21(2):144–164. <https://doi.org/10.3846/13923730.2013.802716>
- Samadi H, Hassanpour J, Farrokh E (2021a) Maximum surface settlement prediction in EPB TBM tunneling using soft computing techniques. *J Phys Conf Ser.* <https://doi.org/10.1088/1742-6596/1973/1/012195>
- Samadi H, Farrokh E (2021b) Utilization of rock mass parameters for performance prediction of rock TBMs using machine learning algorithms. In 4th international conference on development of materials engineering technology, mining and geology, SJFST, vol 3, no 3, pp 1–9. <https://doi.org/10.47176/sjfst.3.3.1>
- Samui P (2008) Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Comput Geotech* 35(3):419–427. <https://doi.org/10.1016/j.compgeo.2007.06.014>
- Shahin MA (2010) Intelligent computing for modeling axial capacity of pile foundations. *Can Geotech J* 47(2):230–243. <https://doi.org/10.1139/T09-094>
- Shahin MA (2014) Load-settlement modeling of axially loaded steel driven piles using CPT-based recurrent neural networks. *Soils Found* 54(3):515–522. <https://doi.org/10.1016/j.sandf.2014.04.015>
- Shahin MA, Jaksa MB, Maier HR (2000) Predicting the settlement of shallow foundations on cohesion less soils using back-propagation neural networks. Department of Civil and Environmental Engineering, University of Adelaide, Australia
- Sheil B (2017) Numerical simulations of the reuse of piled raft foundations in clay. *Acta Geotech* 12:1047–1059. <https://doi.org/10.1007/s11440-017-0522-8>
- Smith GN (1986) Probability and statistics in civil engineering: an introduction. Collins, London
- Smola AJ, Scholkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14:199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Smolczyk U (2003) Geotechnical engineering handbook. Ernst & Sohn, Berlin
- Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. *Neural Process Lett* 9(3):293–300. <https://doi.org/10.1023/A:1018628609742>
- Sychkina EN (2019) Forecast of settlement of single pile based on hereditary creep. In IOP conference series: materials science and engineering, vol 687, no 4. IOP Publishing, pp 044033
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its application to modeling and control. *IEEE Trans Syst* 151:116–132. <https://doi.org/10.1109/TSMC.1985.6313399>
- Tarawneh B, Imam R (2014) Regression versus artificial neural networks: predicting pile setup from empirical data. *KSCE J Civ Eng* 18(4):1018–1027. <https://doi.org/10.1007/s12205-014-0072-7>
- Teh CI, Wong KS, Goh ATC, Jaritngam S (1997) Prediction of pile capacity using neural networks. *J Comput Civ Eng* 11(2):129–138
- Tenpe AR, Patel A (2020) Utilization of support vector models and gene expression programming for soil strength modeling. *Arab J Sci Eng* 45(5):4301–4319. <https://doi.org/10.1007/s13369-020-04441-6>
- Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res* 1:211–244
- Vesic AS (1977) Design of pile foundations. National Cooperative Highway research program, synthesis of practice, vol 42, Transportation Research Board. Washington DC
- Voyagaki E, Crispin J, Gilder C, Nowak P, O’Riordan N, Patel D, Vardanega PJ (2019) Analytical approaches to predict pile settlement in London Clay. In Sustainability issues for the deep foundations: proceedings of the 2nd GeoMEast international congress and exhibition on sustainable civil infrastructures, Egypt 2018—the official international congress of the soil-structure interaction group in Egypt (SSIGE). Springer International Publishing, pp 162–180. [https://doi.org/10.1007/978-3-030-01902-0\\_14](https://doi.org/10.1007/978-3-030-01902-0_14)
- Voyagaki E, Crispin JJ, Gilder CE, Ntassiou K, O’Riordan N, Nowak P, Sadek T, Patel D, Mylonakis G, Vardanega PJ (2022) The DINGO database of axial pile load tests for the UK: settlement prediction in fine-grained soils. *Georisk: assessment and Management of Risk for Engineered Systems and Geohazards*, vol 16, no 4, pp 640–661. <https://doi.org/10.1080/17499518.2021.1971249>
- Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. *Int J Climatol* 32(13):2088–2094. <https://doi.org/10.1002/joc.2419>
- Wong S (2002) Application of piles to pavement and embankment construction. Ph.D. thesis. Department of Civil Engineering. University of Sydney
- Yamashita K, Wakai S, Hamada J (2013) Large-scale piled raft with grid-form deep mixing walls on soft ground. In: Proceeding of the 18th ICSMGE, Takenaka Corporation, Chiba, Japan, pp 2637–2640
- Yang J, Cheng L, Ran L (2022) Research on slope reliability analysis using multi-kernel relevance vector machine and advanced first-order second-moment method. *Eng Comput* 38(4):3057–3068. <https://doi.org/10.1007/s00366-021-01331-9>
- Zhang Y, Wu W, Zhang H, El Naggar MH, Wang K, Jiang G, Mei G (2021) A novel soil-pile interaction model for vertical pile settlement prediction. *Appl Math Model* 99:478–496. <https://doi.org/10.1016/j.apm.2021.07.004>
- Zheng G, Zhang W, Zhou H, Yang P (2020) Multivariate adaptive regression splines model for prediction of the liquefaction-induced settlement of shallow foundations. *Soil Dyn Earthq Eng* 132:106097. <https://doi.org/10.1016/j.soildyn.2020.106097>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.