ORIGINAL PAPER

The performance comparison of the soft computing methods on the prediction of soil compaction parameters

Talas Fikret Kurnaz[1] · Yilmaz Kaya[2]

Abstract
The compaction parameters of soils known as the optimum moisture content (OMC) and maximum dry density (MDD) are necessary for the geotechnical engineering applications such as the fills, embankments, and dams. However, it takes a long time to determine the compaction parameters due to the laboratory test procedure. It was aimed to estimate the compaction parameters of soils with four soft computing methods and also to compare the performance of the methods in this study. For this purpose, a wide database consisting the index and standard proctor (SP) test results were used. Although all AI methods used in this study are successful on estimation of the MDD and OMC parameters, it was seen that the ELM method was the most successful method on the prediction of compaction parameters.

Keywords Compaction · Maximum dry density · Optimum moisture content · Soft computing

Introduction

Soil masses often do not have the properties required for engineering structure constructions. In such cases, in order to obtain the desired properties, the geotechnical properties of the soil can be improved. The selection of the soil improvement techniques to be applied depends on many factors such as the type of soil on the field, the condition of the soil and the economic factors. The aim in all improvement techniques is to increase the soil density and strength, and to reduce the permeability and settlements. One of them, the compaction method, is used to increase the density and bearing capacity of the soil and reduce the permeability.

The compaction is to increase the dry density by the applied energy of the soil where the water content (w) is changed. Air volume reduces while water and solid part do not compact, and the grains come closer together. If the soil is added with some water and compacted, the soil will have a certain dry unit weight (γd) and if the same soil is increased in water content and compacted in the same energy, the dry unit weight gradually increases. By increasing the water content, the dry unit weight reaches the highest value which is defined as maximum dry density (MDD) (γ_{dmax}), and after this limit value, the dry unit weight begins to decrease if the water is added to the soil. The water content providing the maximum dry density is defined as the optimum moisture content (OMC) (w_{opt}). The MDD and OMC are two significant parameters representing the compaction behavior of the soils which are determining with the standard proctor (SP) and modified proctor (MP) tests in the laboratory. These two parameters are determined on compaction curve that obtained from the laboratory tests and have an important role in compacted fillings which are inevitable for engineering structures such as highways, railways, and earth dams.

However, the laboratory tests for determination of the w_{opt} (OMC) and γ_{dmax} (MDD) have time consuming and laborious process and required an important effort. For this reason, many scientists and researchers have tried to determine the compaction parameters from index properties of soils by empirical correlations based on the regression analysis (Table 1). The effect of index properties on compacted the soils has been known for a long time. The grain size and grain distribution on coarse-grained soils and the consistency limits on fine-grained soils are the main factors. In addition, the tests to determine

Responsible Editor: Zeynal Abiddin Erguler

✉ Talas Fikret Kurnaz
  fkurnaz@mersin.edu.tr

[1] Technical Sciences Vocational School, Transportation Services, Mersin University, Mersin, Turkey

[2] Faculty of Engineering and Architecture, Computer Engineering, Siirt University, Siirt, Turkey

Table 1 Proposed empirical correlations for determining the compaction parameters

Equation	References
$\gamma_{dmax} = 0.89(LL - PL) - 1.26LL + 89.8Gs - 102.07$	Jeng and Strohm (1976)
$w_{opt} = 0.611LL - 0.42(LL - PL) + 2.14$	
$\gamma_{dmax} = (Gs/100)(45.6 - 1.28FM \, logD_{10} - 6.64 \times 10^2 \, FM \, PL + 1.43FM)$	Wang and Huang (1984)
$\gamma_{dmax} = (Gs/100)(45.9 + 7.5FM - 0.45logU - 7.54 \times 10^{-2}FM)$	
$w_{opt} = 0.01(26.14 + 12,7PL - 95FM^2 - 88.1log^2U$	
$w_{opt} = 0.01(1035 - 905logD_{50} + 0.22C^2 + 106FMlogD_{50}$	
$\gamma_{dmax} = 2.44 - 0.02 \, PL - 0.008 \, LL$ (Iraq)	Al-Khafaji (1993)
$\gamma_{dmax} = 2.27 - 0.019 \, PL - 0.003 \, LL$ (USA)	
$w_{opt} = 0.24 \, LL + 0,63 \, PL - 3.13$ (Iraq)	
$w_{opt} = 0.14 \, LL + 0.54 \, PL$ (USA)	
$\gamma_{dmax} = 1.6148 - 0.1329(- \, No.4) + 0.5576(+ \, No.200)$ (for SP)	Engin (2003)
$\gamma_{dmax} = 1.9300 - 0.1940(- \, No.4) + 0.4166(+ \, No.200)$ (for MP)	
$w_{opt} = 0.2108 + 0.0227(- \, No.4) - 0.1470(+ \, No.200)$ (for SP)	
$w_{opt} = 0.1589 + 0.0212(- \, No.4) - 0.1118(+ \, No.200)$ (for MP)	
$\gamma_{dmax} = 0.23(93.3 - PL)$	Sridharan and Nagaraj (2005)
$w_{opt} = 0.92PL$	
$\gamma_{dmax} = 0.223G + 0.411FC$ (for SP)	Olmez (2007)
$\gamma_{dmax} = 0.234G + 0.233S + 0.055FC$ (for MP)	
$w_{opt} = 0.618PL - 0.453PI$ (for SP)	
$w_{opt} = 0.284LL$ (for MP)	
$\gamma_{dmax} = 21.84 - 0.27 \, w_{opt}$ (for SP)	Sivrikaya (2008)
$w_{opt} = 0.92PL$ (for SP)	
$\gamma_{dmax} = - 0.1008 \, LL + 21.16$	Gunaydin (2009)
$\gamma_{dmax} = - 0.2283 \, PL + 21.88$	
$\gamma_{dmax} = - 0.078 \, LL - 0.062 \, PL$	
$w_{opt} = 0.8442 \, PL + 0.1076$	
$w_{opt} = 0.3802 \, LL + 2.4513$	
$w_{opt} = 0.323 \, LL + 0.157 \, PL$	
$\gamma_{dmax} = 0.253G + 0.236S + 0.218FC - 0.234LL + 0.161PI$ (for SP)	Sivrikaya et al. 2013
$\gamma_{dmax} = 23.673 - 0.357 \, w_{opt}$ (for SP)	
$\gamma_{dmax} = 0.265G + 0.278S + 0.127FC - 0.213LL + 0.166PI$ (for MP)	
$\gamma_{dmax} = 25.702 - 0.528 \, w_{opt}$ (for MP)	
$w_{opt} = - 0.032G - 0.009S + 0.046FC + 0.659LL - 0.473PI$ (for SP)	
$w_{opt} = 0.420LL$ (for SP)	
$w_{opt} = -0.005G - 0.007S + 0.141FC + 0.267LL - 0.073PI$ (for MP)	
$w_{opt} = 0.292LL$ (for MP)	

LL Liquid limit, *PL* plastic limit, *Gs* specific gravity, *FM* fineness modulus, D_{10} effective grain size, *U* uniformity coefficient, *C* clay content, D_{50} average grain size, *E* compaction energy, *No.4* passing No.4 sieve, *No.200* passing No.200 sieve, *G* gravel content, *S* sand content, *FC* fine content, *PI* plasticity index, *SP* standard proctor, *MP* modified proctor

the index properties have a fairly easy and inexpensive procedure compared with the compaction tests.

The proposed correlations related to relationships between the physical characteristics and the compaction parameters are generally based on multiple linear regression (MLR) analysis. The most important problem with the use of these correlations is that the correlations are usually developed for a particular locality or same geological origin soils. The use of these

Table 2 Statistical specifications of index and compaction parameters

	LL (%)	PL (%)	GC (%)	SC (%)	FC (%)	w_{opt} (%)	γ_{dmax} (kN/m³)
Min	19	10	0	1	1	6	12.70
Max	87.50	44	97	87.20	99	32	22.60
Mean	41.13	20.04	20.57	38.79	40.65	14.40	18.37
Std. Dev.	11.04	4.68	22.31	17.69	22.44	4.76	1.79

Table 3 Performance results of the GMDH model

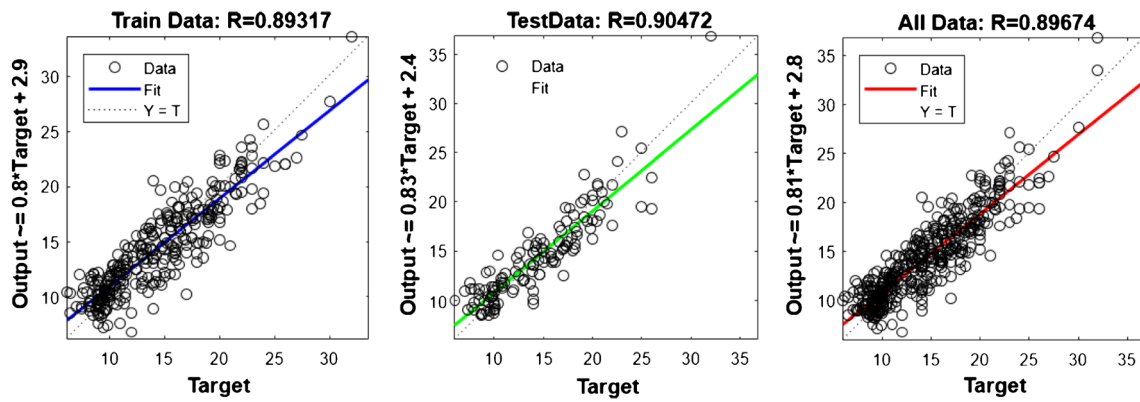| #Layer	OMC				MDD			
	Training		Testing		Training		Testing	
	MSE	R	MSE	R	MSE	R	MSE	R
2	0.619	0.8493	0.546	0.8793	0.809	0.8590	1.03	0.8373
3	0.603	0.8503	0.528	0.8798	0.804	0.8584	0.901	0.8657
4	0.671	0.8247	0.820	0.8297	1.17	0.7984	0.88	0.8476
5	0.635	0.8432	0.538	0.8834	0.834	0.8598	0.924	0.8401
6	0.821	0.8001	0.674	0.8325	0.925	0.8401	0.680	0.8943
7	0.587	0.8623	0.451	0.9047	0.628	0.8662	0.756	0.9174

Fig. 1 Determination coefficient performance values for OMC parameter

correlations for an area outside the grounds of the locality where the correlations are developed can cause significant differences between the expected and computed compaction parameters. Therefore, it is necessary to be cautious about the use of the compaction parameters determined by empirical correlations.

Recently, geotechnical engineering practices has begun to include many studies related with soft computing methods (Lee and Lee 1996; Najjar and Basheer 1996; Kiefa 1998; Juang and Chen 1999; Sakellariou and Ferentinou 2005; Wang et al. 2005; Kim and Kim 2006; Sinha and Wang 2008; Samui 2008; Abdel-Rahman 2008; Kuo et al. 2009; Gunaydin 2009; Nejad et al. 2009; Kalinli et al. 2011; Samui and Kothari 2011; Isik and Ozden 2013; Tenpe and Kaur 2015; Abdalla et al. 2015; Chenari et al. 2015; Suman et al. 2016).

In this paper, soft computing methods such as group method of data handling (GMDH)–type neural network, support vector machine (SVM), Bayesian regularization neural network (BRNN), and extreme learning machine (ELM) were used to predict the compaction parameters of the soils. The index properties of soil samples were used as an input parameter on estimation of compaction parameters for all models.

Database compilation

Database containing the index and standard proctor (SP) test results of 451 soil samples was used for this study. Approximately half of the dataset was provided from published studies (Gunaydin 2009; Olmez 2007), and the remainder part was provided from laboratories focused on different soil investigations such as compaction tests in various parts of Turkey. As mentioned in the introduction section, the grain size distribution on coarse-grained soils and the consistency limits on fine-grained soils are the most effective features. In this context, the database was formed by the percentages of liquid limit (LL), plastic limit (PL), fines content (FC), sand content (SC), and gravel content (GC) of the compacted soils. The soil classes of the soil samples has a wide scale such as CH, CI, CL, GC, GM, GP, GP-GC, GP-GM, GW, GW-GC, GW-GM, MH, MI, ML, SC, SP, SP-SC, and SW-SC. Statistical description of the data is given in Table 2.

Method

The GMDH is a more complex model, which is gradually evaluated on a set of multiple input, single output data pairs
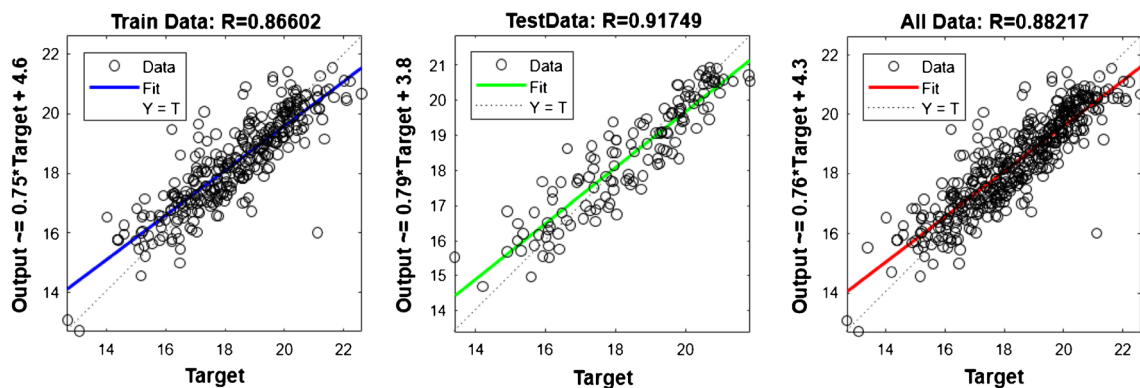

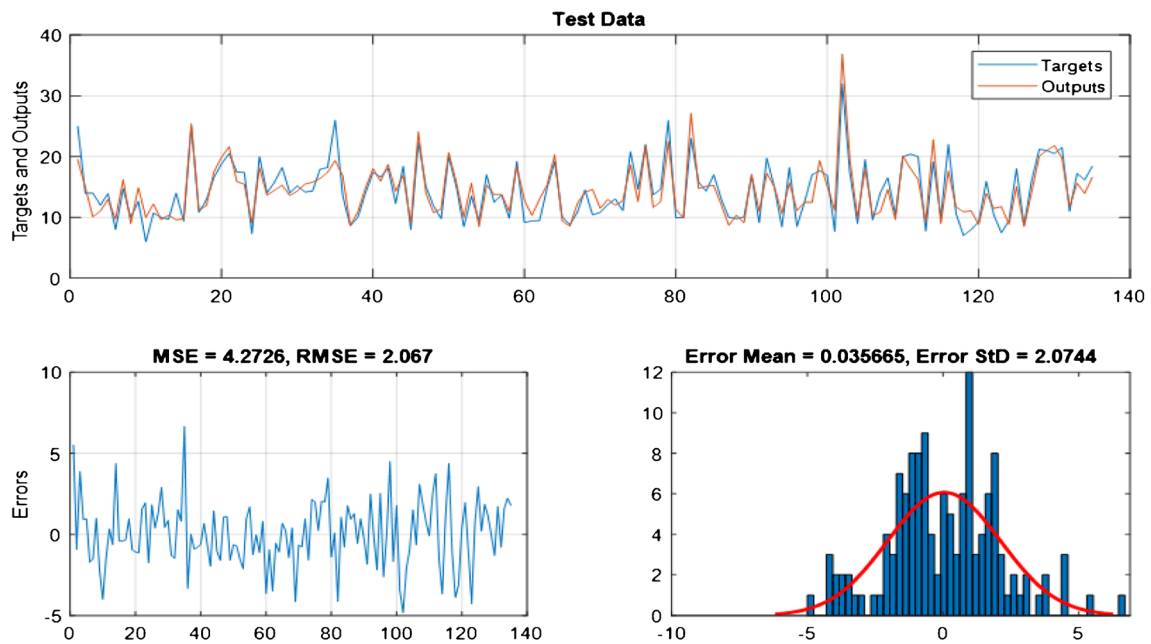**Fig. 2** Determination coefficient performance values for MDD parameter

Fig. 3 Distribution and error graphics related to the OMC parameter

(Vissikirsky et al. 2005). The GMDH architecture is a self-organizing polynomial neural network model with a flexible structure (Ghanadzadeh et al. 2012). Firstly, the data set is divided into training and test sets, and the calculations are made according to the following regression equation for both input variable pairs;

$$\text{Quadratic}: \widehat{y} = G(x_i, x_j) = w_0 + w_1 x_i + w_2 x_j + w_3 x_i x_j \quad (1)$$
$$+ w_4 x_i^2 + w_5 x_j^2$$

The weights in the above equation are obtained according to the least squares method. Then, new variables are obtained by using polynomial equations. The output of the polynomial equation of variables is carried to the next layer with the minimum error value. Thus, new variables are produced from the input variables that best describe the output variable.

$$E = \frac{\sum_{i=1}^{M} \left(y_i - G_i(x_i, x_j) \right)^2}{M} \rightarrow \text{minimum} \quad (2)$$
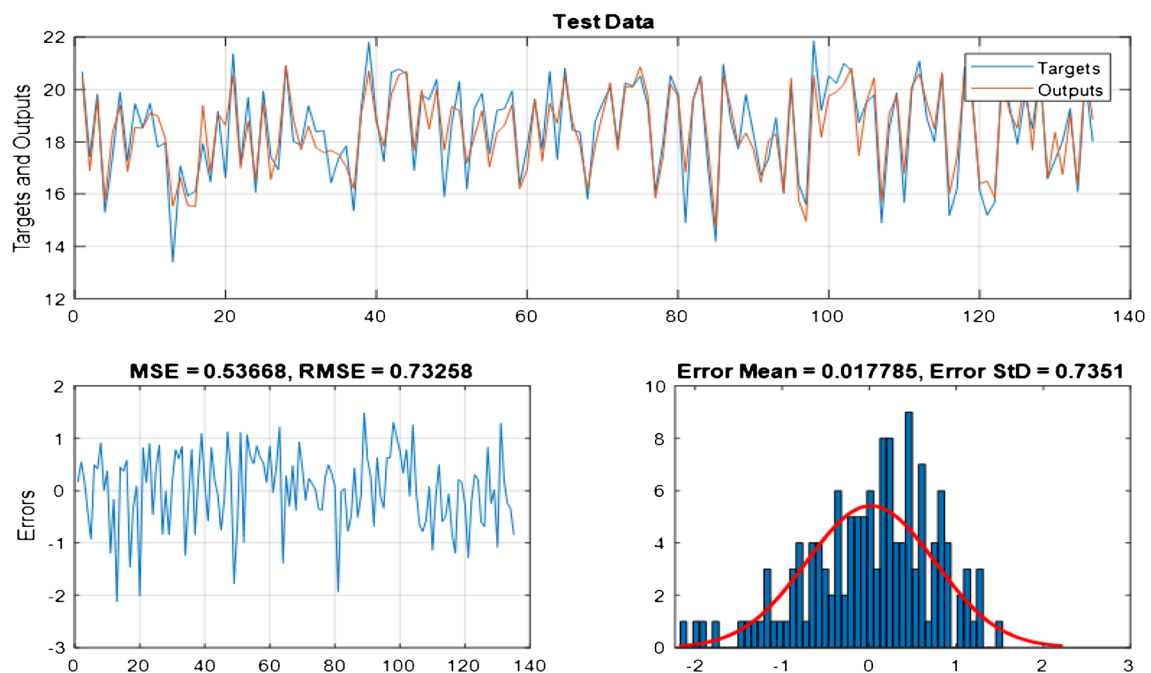

Fig. 4 Distribution and error graphics related to the MDD parameter

Table 4 Performance criteria in the case of output OMC and MDD

Data Sets	Train		Test	
	R	MSE	R	MSE
OMC	0.8678	0.3161	0.9358	0.2091
MDD	0.8502	0.3502	0.9521	0.5123

The GMDH model has been used with success in geotechnical practice in recent years (Jirdehi et al. 2014; Kordnaeij et al. 2015; Hassanlourad et al. 2017; Ardakani and Kordnaeij 2017).

ELM is basically similar to artificial neural networks with one hidden layer. Therefore, the working principle of ELM is to some extent the same as the working principles of artificial neural networks (ANN). However, in over-learning machines, the weights (W_i) in the hidden layer are randomly assigned, and these values are not changed (not updated) at a later stage of training. In contrast, the weights (β_i) between the output layer and the hidden layer are determined at once, analytically and quickly using a linear model. The basic ELM model is based on feedforward neural networks with a single hidden layer (Huang et al. 2006). The ELM method has begun to be used in most geotechnical problems and successful results were achieved (Muduli et al. 2013; Huang et al. 2017a; Huang et al. 2017b; Liu et al. 2014; Li et al. 2016; Liu et al. 2015).

Determining the number of hidden neurons is most of the biggest challenges in creating an ANN model. Overfitting can be occurs when the number of neurons is high, and there is a difficulty in training the network when there are few hidden nodes. In addition, an ANN model that is designed to be complex or simple will have poor predictive performance. To overcome this, BRNN model, which describes ANN's training function with a probabilistic approach, was proposed by MacKay (1991). BRNN is a widely used model for solving nonlinear problems (Bui et al. 2012; Okut 2016; Caballero and Fernández 2006). The BRNN method has been used on different geotechnical problems and obtained good results

(Nejad et al. 2009; Das et al. 2010; Muduli et al. 2014; Sabat 2015).

SVM model theory was first proposed by Cortes and Vapnik (1995). SVM is widely used for regression and classification purposes. An extended version of SVM, also known as support vector regression (SVR), has been developed for complex regression problems. SVM method has been widely used in different geotechnical problems in recent years (Kordjazi et al. 2014; Sabat 2015; Samui 2008; Samui et al. 2008; Samui and Kothari 2011). In these studies, the theory of the SVR model is also explained in detail.

Results

GMDH results

In this paper, different GMDH architectures were used to prediction of the compaction parameters. The results of these trials are given in Table 3. Since the rise in the number of hidden layers caused an increase in the calculation costs, a maximum of 7 trials were performed. However, when the number of hidden layers rises, it is seen that the MSE error values decrease and R determination coefficients increase and also the success of the model increases (Table 3). The regression graphics and curves belonging to the actual and predicted values are given on Figs. 1 and 2. The graphics and histograms related to errors and distributions are given in Figs. 3 and 4.

ELM results

The training and test sets were randomly generated in ELM models. The R and MSE performance values obtained for OMC and MDD parameters are given in Table 4.

In this study, different architectural structures of ELM have also been tried to estimate OMC and MDD parameters. The different architectural structures of the ELM model were carried out for the 70–30% training-test set, and the obtained results are given in Tables 5 and 6.

Table 5 Effect of different activation functions for OMC (70–30% training-test set)

Types of activation functions	R	MSE
Sigmoid-22	0.9304	0.2543
Sine-19	0.9336	0.2437
Hardlimit-6	0.8955	0.3441
Triangular basis-45	0.9068	0.2863
Radial basis-24	0.9369	0.2224

Table 6 Effect of different activation functions for MDD (70–30% training-test set)

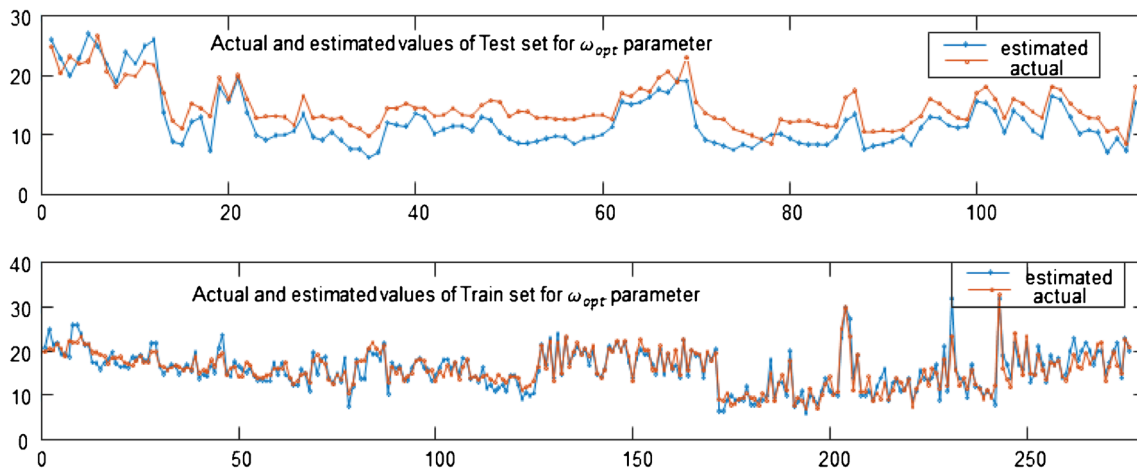| Types of activation functions	R	MSE
Sigmoid-6	0.9282	0.4840
Sine-6	0.9465	0.4673
Hardlimit-55	0.8893	0.5421
Triangular basis-5	0.9342	0.4385
Radial basis-52	0.9309	0.4402

Fig. 5 The estimated and the actual values for OMC

The best performance in predicting OMC was achieved with radial basis activation function (Table 5). The best performance in predicting MDD was obtained with sine activation function (Table 6). It is clear that the estimated and the actual values for the OMC and MDD were seen to be close when related graphics are examined (Fig. 5 and 6).

Many trials were performed for the estimation of the OMC and MDD parameters with ELM. The training set was used to train the model while the test set was used to verify the generalization ability of the model in the trials with ELM. The R and MSE performance values obtained in estimating the OMC and MDD parameters are given in Fig. 7. It is understood that the training period for the OMC parameter is very well performed (Fig. 7a). However, the same performance could not be demonstrated in the testing process. In trials where the number of neurons in the hidden layer is greater than 30, it is seen that the values of R and MSE are wavy for the test

samples; in addition, the MSE error values are increased. Accordingly, it can be said that there is no functional relationship between the R and MSE performance criteria and the number of neurons in the hidden layer. However, the performance of ELM's generalization ability depends on the number of neurons in the hidden layer. Similarly, when the number of neurons in the hidden layer of the ELM is greater than 30–40 for the estimation of the MDD parameter, the performance of the model was reduced. In other words, the MSE error values were increased, and the R values were decreased (Fig. 7b).

BRNN results

In this paper, the weights between hidden-output layers and input-hidden layers in neural networks were also calculated with Bayesian regularization model. To show the performance of success, BRNN method was compared with different training methods. The variables in the input
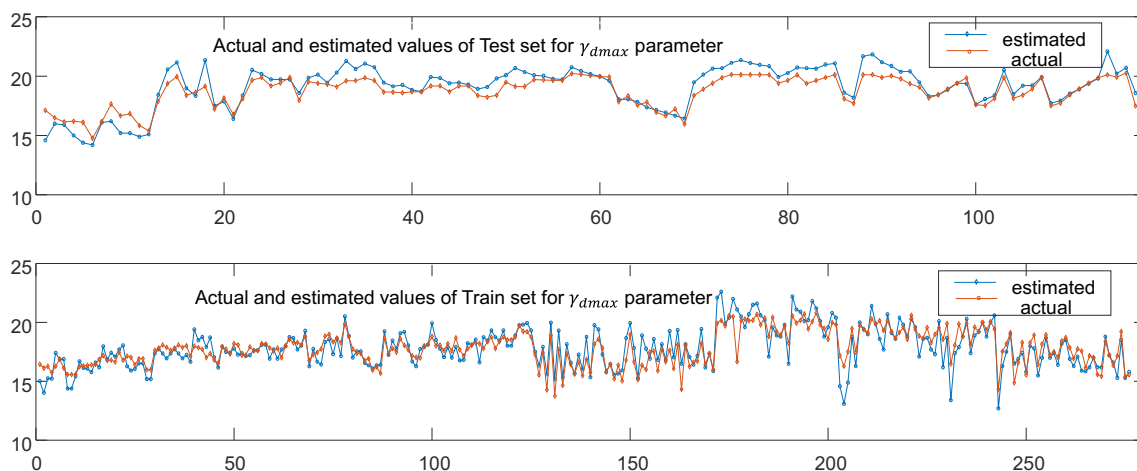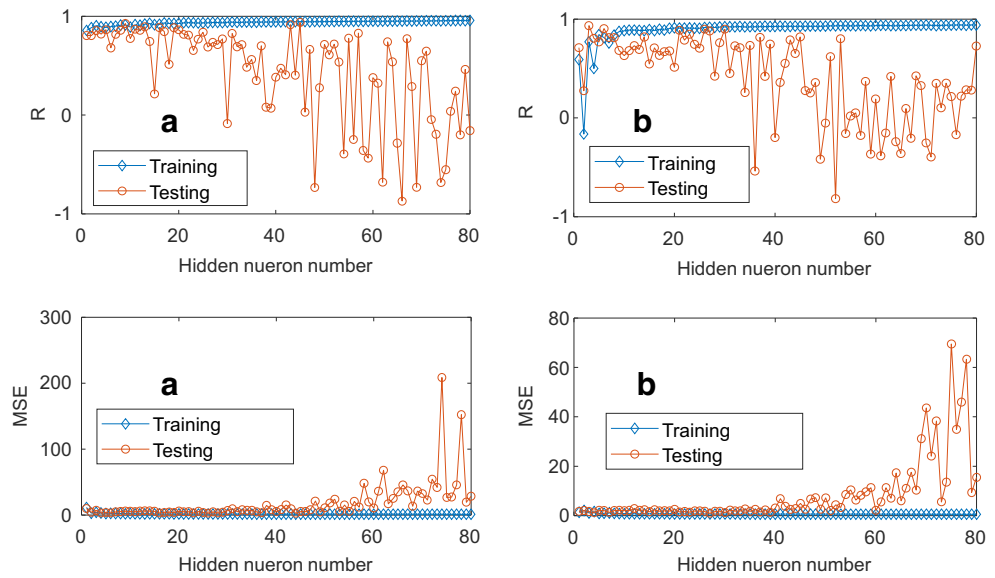

Fig. 6 The estimated and the actual values for MDD

Fig. 7 Predictive performance versus hidden neuron number (with radial basic activation function), **a** training and test performances for OMC parameter, **b** training and test performances for MDD parameter

layer are the index properties of soil such as LL, PL, FC, SC, GC, and the compaction parameters (OMC or MDD) are used in the output layer. Different activation functions and different neuron numbers for the hidden layer were used for the success of the BRNN model. Training and test sets were randomly generated in each attempt. Min–max conversion was applied to the data before the analysis. The best performance results were obtained by using log-sigmoid and linear activation functions in hidden and output layer, respectively. The performance results obtained with BRNN in the estimation of OMC and MDD parameters are given in Table 7. The best results obtained with BRNN as $R = 0.9191$ and MSE $= 0.0043$ for OMC and $R = 0.9219$ and MSE $= 0.0047$ for the MDD parameter, respectively. The regression graphs obtained for 70–30% training-test sets of OMC and MDD parameters are given in Figs. 8 and 9. The graphs show the regression models between measured and estimated values for training, testing, and all data.

To show the success of the BRNN method, other ANN training methods were also tested, and the results were compared with BRNN (Tables 8 and 9). All methods were used with the same activation functions and architectures.

As seen in the Tables 8 and 9, it has been observed that the best success for the estimation of OMC and MDD parameters is obtained with BRNN. The results obtained with BRNN for test data set as $R = 0.9191$ and MSE $= 0.0043$ for OMC (Table 8) and $R = 0.9219$ and MSE $= 0.0047$ for the MDD (Table 9), respectively. However, acceptable results have been obtained with other train methods.

SVM results

In this section, trials were performed with SVM models in different architectures for the estimation of OMC and MDD. Different kernel functions were used by investigating the effect of kernel function on SVM model performance. In addition to specifying the kernel function, the decision of the user-defined parameters C and ε is also important. The performance values for the estimation of OMC and MDD parameters are given in Tables 10 and 11.

The prediction achievements showed little change according to the variation of ε parameter. The best success in estimating the OMC parameter with SVM was $R = 0.8510$ and MSE $= 2.5247$ in the case of $\varepsilon = 0.010$, and the best success in estimating the MDD parameter was $R = 0.8483$ and MSE $= 0.9557$ in the case of $\varepsilon = 0.002$ (Tables 10 and 11).

The graphics of R values observed for different kernel functions and ε values are given in Fig. 10. The performances obtained by different kernel functions of OMC and MDD parameters are seen on Figs. 10 and 11. As shown in Figs. 10, 11, and 12, the most successful kernel function in estimating OMC and MDD parameters is observed as RBF

Table 7 Performance criteria in the case of output OMC and MDD

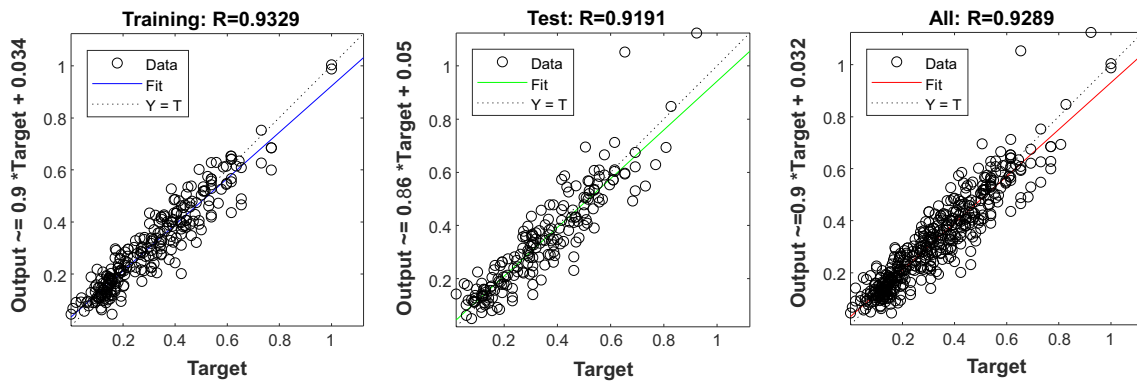| Data sets	Training set		Test set		All set	
	MSE	R	MSE	R	MSE	R
OMC	0.0039	0.9329	0.0043	0.9191	0.0041	0.9289
MDD	0.0045	0.9270	0.0047	0.9219	0.0046	0.9258

Fig. 8 Comparison between the estimated and measured values of OMC

Discussion

In the present study, many trials were performed to obtain best prediction performance of OMC and MDD with different training algorithms, different activations, or kernel functions in all models. Eventually, the performance results of the GMDH, ELM, BRNN, and SVM models are compared. Comparison of the prediction performances of all models are presented in Table 12.

Despite the overfitting problem, ANN models are widely preferred by many researchers in solving linear and nonlinear problems. The GMDH method, which has been gaining increasing popularity in recent years, is a successful model if there is a linear relationship between data (inputs and outputs). On the contrary, the ELM and BRNN methods are known as successful on solving nonlinear problems. SVM can be used for classification and regression purposes and has been successfully applied in the solution of nonlinear problems in recent years.

There are various research studies focused on to solve complex geotechnical problems by the types of AI techniques (Samui and Kothari 2011; Das and Basudhar 2007; Das et al. 2010; Liu et al. 2015; Huang et al. 2017a, b; Muduli

et al. 2013; Nejad et al. 2009; Sabat 2015; Samui 2008). It is clear that the AI techniques are found to be more efficient compared to statistical models in the aforementioned studies. When AI techniques are compared in themselves, the success results can vary according to the content of the study. Although ANN models seem to be more successful than SVM when the relationship between inputs and output is not linear, all AI techniques can be more successful in different data sets.

Conclusions

In this paper, prediction models were developed by using soft computing methods such as GMDH, ELM, BRNN and SVM for the compaction parameters of soils and compared the model performances. Totally 451 test data (index and Standard Proctor) belongs to the compacted soils were used for the study. Trials with GMDH were carried out with different architectures. The best performance success was obtained with 7 hidden layer (R = 0.9047, MSE = 0.451 for OMC; R = 0.9174, MSE = 0.756 for MDD) (Table 3). In trials with ELM, the best
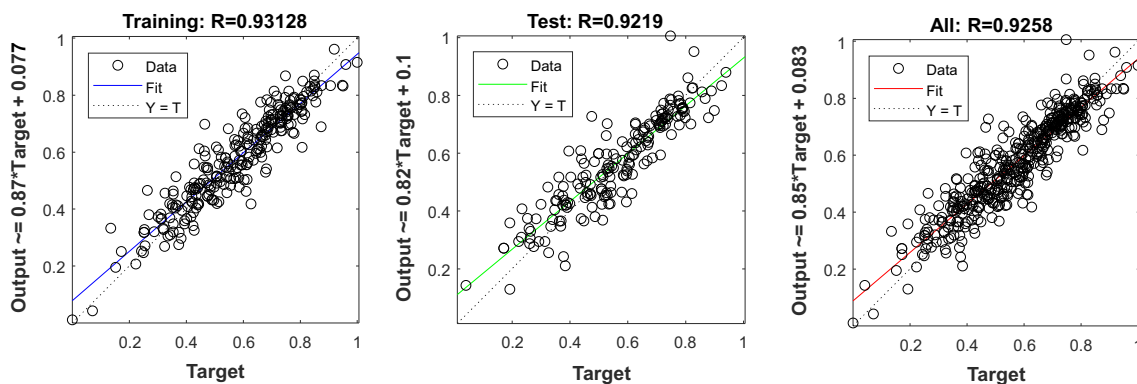

Fig. 9 Comparison between the estimated and measured values of MDD

Table 8 Performance criteria in the case of output OMC for 70–30% training-test sets

Training model	Training set		Test set		All set	
	MSE	R	MSE	R	MSE	R
Bayesian regularization	0.0039	0.9329	0.0043	0.9191	0.0041	0.9289
Levenberg-Marquardt	0.0099	0.8460	0.0091	0.8738	0.0093	0.8536
Conjugate gradient Polak-Ribiére	0.0036	0.9437	0.0050	0.8751	0.0047	0.9227
BFGS quasi-Newton	0.0033	0.9514	0.0042	0.8930	0.0040	0.9359
Conjugate gradient Powell-Beale	0.0038	0.9441	0.0046	0.8988	0.0044	0.9303
Gradient descent adaptive learning	0.0093	0.8419	0.0087	0.8915	0.0092	0.8613
Gradient descent with momentum	0.0159	0.7181	0.0165	0.7206	0.0163	0.7190
Gradient desc. with momentum and adaptive learning	0.0065	0.9033	0.0076	0.8229	0.0074	0.8804
Resilient	0.0033	0.9506	0.0042	0.9015	0.0044	0.9353
Scaled conjugate gradient	0.0032	0.9494	0.0044	0.8991	0.0041	0.9324

Table 9 Performance criteria in the case of output MDD for 70–30% training-test sets

Training model	Training set		Test set		All set	
	MSE	R	MSE	R	MSE	R
Bayesian regularization	0.0045	0.9270	0.0047	0.9219	0.0046	0.9258
Levenberg-Marquardt	0.0099	0.9506	0.0091	0.8477	0.0093	0.9193
Conjugate gradient Polak-Ribiére	0.0046	0.9214	0.0054	0.8993	0.0051	0.9137
BFGS quasi-Newton	0.0032	0.9494	0.0040	0.9071	0.0039	0.9372
Conjugate gradient Powell-Beale	0.0038	0.9403	0.0042	0.9179	0.0040	0.9332
Gradient descent adaptive learning	0.0072	0.8822	0.0080	0.8416	0.0078	0.8700
Gradient descent with momentum	0.0179	0.6936	0.0175	0.6934	0.0176	0.6885
Gradient desc. with momentum and adaptive learning	0.0051	0.9182	0.0085	0.7472	0.0080	0.8642
Resilient	0.0040	0.9366	0.0048	0.9018	0.0044	0.9250
Scaled conjugate gradient	0.0037	0.9386	0.0045	0.9076	0.0041	0.9283

Table 10 Performances observed with SVM in the case of output OMC

ε	Radial basis function		Polynomial function		Sigmoid function		Linear function	
	R	MSE	R	MSE	R	MSE	R	MSE
0.001	0.8504	2.5274	0.7875	3.6064	0.8292	2.9004	0.8391	2.7843
0.002	0.8503	2.5276	0.7875	3.6064	0.8294	2.9013	0.8383	2.7932
0.003	0.8503	2.5285	0.7873	3.6142	0.8300	2.902	0.8379	2.7957
0.004	0.8505	2.5264	0.7873	3.6142	0.8306	2.8981	0.8378	2.7936
0.005	0.8507	2.5256	0.7873	3.6142	0.8306	2.8981	0.8386	2.787
0.006	0.8505	2.5266	0.7873	3.6142	0.8306	2.8981	0.8402	2.7816
0.007	0.8506	2.5256	0.7873	3.6142	0.8288	2.8965	0.8402	2.7816
0.008	0.8505	2.5263	0.7873	3.6221	0.8294	2.9088	0.8407	2.7723
0.009	0.8507	2.5251	0.7874	3.6130	0.8277	2.9058	0.8396	2.7802
0.010	0.8510	2.5247	0.7874	3.6130	0.8277	2.9058	0.8396	2.7802

Table 11 Performances observed with SVM in the case of output MDD

| ε	Radial basis function		Polynomial function		Sigmoid function		Linear function	
	R	MSE	R	MSE	R	MSE	R	MSE
0.001	0.8482	0.956	0.7717	1.1914	0.8421	0.9697	0.8418	0.9708
0.002	0.8483	0.9557	0.7717	1.1914	0.8420	0.9701	0.8450	0.9611
0.003	0.8483	0.9557	0.7721	1.1909	0.8419	0.9704	0.8447	0.9619
0.004	0.8481	0.9563	0.7717	1.1915	0.8419	0.9704	0.8452	0.9606
0.005	0.8479	0.9564	0.7719	1.1912	0.8413	0.9718	0.8451	0.9608
0.006	0.8480	0.9564	0.7719	1.1912	0.8418	0.9705	0.8451	0.9611
0.007	0.8478	0.9565	0.7719	1.1912	0.8415	0.9718	0.8450	0.9612
0.008	0.8471	0.9588	0.7722	1.1911	0.8417	0.9712	0.8446	0.9624
0.009	0.8471	0.9595	0.7722	1.1911	0.8427	0.9687	0.8442	0.9641
0.010	0.8469	0.9599	0.7739	1.1891	0.8428	0.9684	0.8441	0.9638
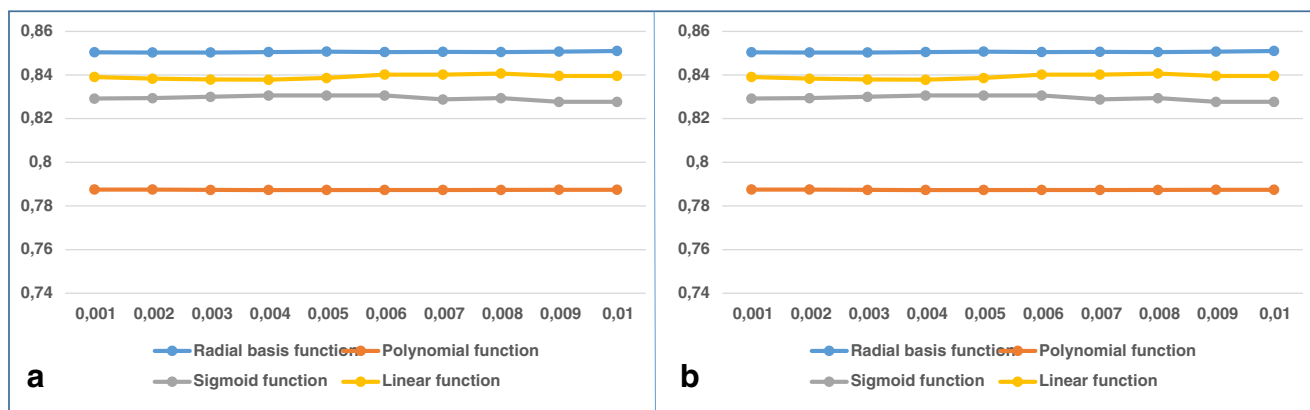
Fig. 10 . R values for different kernel functions. (A) R values for the OMC, (B) R values for the MDD

performance was obtained with Radial Basis activation function and 24 neurons in the hidden layer on the prediction of OMC (R = 0.9369 and MSE = 0.2224) (Table 5) while the best performance on the prediction of MDD was obtained with Sine activation function and 6 neurons in the hidden layer (MSE = 0.4673and R = 0.9465) (Table 6). Besides, in trials where the number of neurons in the hidden layer is greater than 30, it was seen that the MSE error values are increased (Fig. 7). Trials with BRNN were also performed different activation functions. The best performance results were obtained as R = 0.9191 and MSE = 0.0043 for OMC; R = 0.9219 and MSE = 0.0047 for the MDD (Table 7). In addition, the BRNN was compared to the other ANN training methods such as the LM, BFG, CGB, CGP, GDA, GDM, GDX, RP and SCG and found to be the most successful (Tables 8 and 9). In trials with SVM, the best performance was determined with RBF kernel function on the prediction of both OMC (R = 0.8510 and MSE = 2.5247) and MDD (R = 0.8483 and MSE = 0.9557) parameters (Tables 10 and 11).

As a result of this study, it is seen that the ELM method was the most successful method on the prediction of compaction parameters. And also, The ELM and BRNN methods are found to be more successful than GMDH due to the lack of a very linear relationship between the data. This has also been confirmed with compare of the performance results of all methods (Table 12). However, the SVM has been considered as the lowest-performing

Table 12 The comparison of the model performance results

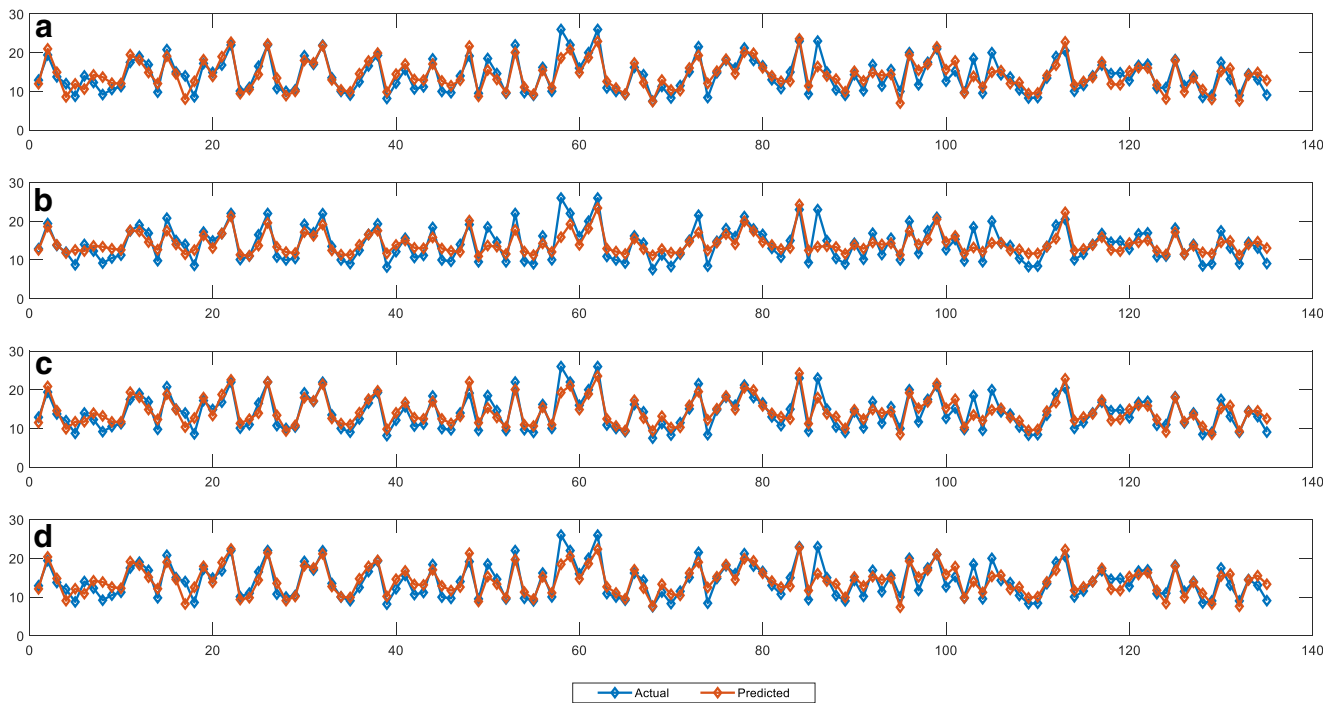| MODEL	OMC		MDD	
	R	MSE	R	MSE
ELM	0.9358	0.2091	0.9521	0.5123
BRNN	0.9191	0.271	0.9219	0.596
GMDH	0.9047	0.628	0.9174	0.756
SVM	0.8510	2.5200	0.8483	0.9557

Fig. 11 The actual and predicted with different kernel functions OMC values. **a** Predicted with LF, **b** predicted with PF, **c** predicted with RBF, **d** predicted with SF

method in this study. The results obtained with SVM method were less successful in all trials on the prediction of both OMC and MDD parameters comparing to the ANN-based models. It is believed that the limitation on achieving more successful results is due to the small number of data (451 test data) and it is thought that the success rates of different soft computing models will increase if the data set is expanded in the future.
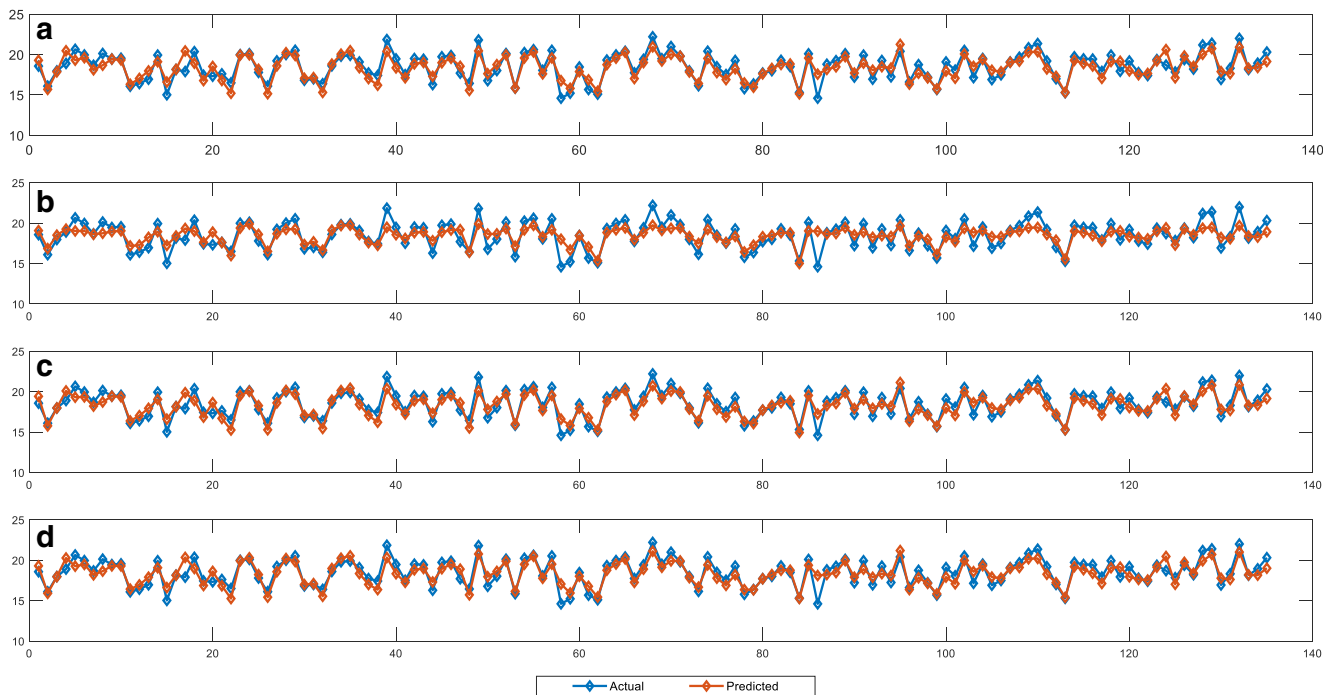

Fig. 12 The actual and predicted with different kernel functions MDD values. **a** Predicted with LF, **b** predicted with PF, **c** predicted with RBF, **d** predicted with SF

Compliance with ethical standards

Conflict of interest   The authors declare that they have no conflict of interest.

References

Abdalla JA, Attom MF, Hawileh R (2015) Prediction of minimum factor of safety against slope failure in clayey soils using artificial neural network. Environ Earth Sci 73(9):5463–5477

Abdel-Rahman AH (2008) Predicting compaction of cohesionless soils using ANN. In: Proceedings of the institution of civil engineers ground improvement, vol 161, pp 3–8

Al-Khafaji AN (1993) Estimation of soil compaction parameters by means of Atterberg limits. Q J Eng Geol Hydrogeol 26:359–368. https://doi.org/10.1144/GSL.QJEGH.1993.026.004.10

Ardakani A, Kordnaeij A (2017) Soil compaction parameters prediction using GMDH-type neural network and genetic algorithm. Eur J Environ Civ Eng, 1–14

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Landslide susceptibility assessment in the Hoa Binh province of Vietnam: a comparison of the Levenberg–Marquardt and Bayesian regularized neural networks. Geomorphology 171:12–29. https://doi.org/10.1016/j.geomorph.2012.04.023

Caballero J, Fernández M (2006) Linear and nonlinear modeling of antifungal activity of some heterocyclic ring derivatives using multiple linear regression and Bayesian-regularized neural networks. J Mol Model 12(2):168–181

Chenari RJ, Tizpa P, Rad MRG, Machado SL, Fard MK (2015) The use of index parameters to predict soil geotechnical properties. Arab J Geosci 8(7):4907–4919

Cortes C, Vapnik V (1995) Support vector networks. Mach Learn 20(3): 273–297

Das SK, Basudhar PK (2007) Prediction of hydraulic conductivity of clay liners using artificial neural network. Lowland Technol Int J 9(1): 50–58

Das SK, Samui P, Sabat AK, Sitharam TG (2010) Prediction of swelling pressure of soil using artificial intelligence techniques. Environ Earth Sci 61(2):393–403

Engin H (2003) A laboratory investigation on the correlations of standard and modified compaction test values. Master's thesis, Dokuz Eylul University, Turkey

Ghanadzadeh H, Ganji M, Fallahi S (2012) Mathematical model of liquid–liquid equilibrium for a ternary system using the GMDH-type neural network and genetic algorithm. Appl Math Model 36: 4096–4105

Gunaydin O (2009) Estimation of soil compaction parameters by using statistical analyses and artificial neural networks. Environ Geol 57: 203–215. https://doi.org/10.1007/s00254-008-1300-6

Hassanlourad M, Ardakani A, Kordnaeij A, Mola-Abasi H (2017) Dry unit weight of compacted soils prediction using GMDH-type neural network. Eur Phys J Plus 132:357

Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489e501

Huang F, Huang J, Jiang S, Zhou C (2017a) Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. Eng Geol 218:173–186

Huang F, Yin K, Huang J, Gui L, Wang P (2017b) Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. Eng Geol 223:11–22

Isik F, Ozden G (2013) Estimating compaction parameters of fine- and coarse-grained soils by means of artificial neural networks. Environ Earth Sci 69:2287–2297. https://doi.org/10.1007/s12665-012-2057-5

Jeng YS and Strohm WE (1976) Prediction of the sherar strength and compaction characteristics of compacted fine-grained cohesive soils. Final report, U.S. Army engineer waterways Experiment Station, soils and pavement laboratory, Vicksburg

Jirdehi RA, Mamoudan HT, Sarkaleh HH (2014) Applying GMDH-type neural network and particle warm optimization for prediction of liquefaction induced lateral displacements. Appl Appl Math 9(2): 528–540

Juang CH, Chen CJ (1999) CPT-based liquefaction evaluation using artificial neural networks. Comput Aided Civ Inf Eng 14(3):221–229

Kalinli A, Acar MC, Gunduz Z (2011) New approaches to determine the ultimate bearing capacity of shallow foundations based on artificial neural networks and ant colony optimization. Eng Geol 117(1–2): 29–38. https://doi.org/10.1016/j.enggeo.2010.10.002

Kiefa MAA (1998) General regression neural networks for driven piles in cohesionless soils. Geotech Geoenviron Eng 124(12):1177–1185

Kim YS, Kim BT (2006) Use of artificial neural networks in the prediction of liquefaction resistance of sands. J. Geotech. Geoenviron. Eng. ASCE 132(11):1502–1504. https://doi.org/10.1061/ASCE1090-02412006132:111502

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

Kordnaeij A, Kalantary F, Kordtabar B, Mola-Abasi H (2015) Prediction of recompression index using GMDH-type neural network based on geotechnical soil properties. Soils Found 55(6):1335–1345

Kuo YL, Jaksa MB, Lyamin AV, Kaggwa WS (2009) ANN-based model for predicting the bearing capacity of strip footing on multi-layered cohesive soil. Comput Geotech 36(3):503–516. https://doi.org/10.1016/j.compgeo.2008.07.002

Lee I, Lee J (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

Li AJ, Khoo S, Lyamin AV, Wang Y (2016) Rock slope stability analyses using extreme learning neural network and terminal steepest descent algorithm. Autom Constr 65:42–50

Liu Z, Shao J, Xu W, Chen H, Zhang Y (2014) An extreme learning machine approach for slope stability evaluation and prediction. Nat Hazards 73(2):787–804

Liu Z, Shao J, Xu W, Wu Q (2015) Indirect estimation of unconfined compressive strength of carbonate rocks using extreme learning machine. Acta Geotech 10(5):651–663

MacKay DJC (1991) Bayesian Methods for Adaptive Models. Ph.D. thesis, California Institute of Technology

Muduli PK, Das SK, Das MJ (2013) Prediction of lateral load capacity of piles using extreme learning machine. Int J Geotech Eng 7(4):388–394

Muduli PK, Das SK, Bhattacharya S (2014) CPT-based probabilistic evaluation of seismic soil liquefaction potential using multi-gene genetic programming. Georisk 8(1):14–28. https://doi.org/10.1080/17499518.2013.845720

Najjar YM, Basheer IA (1996) Utilizing computational neural networks for evaluating the permeability of compacted clay liners. Geotech Geol Eng 14:193–212

Nejad FP, 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

Okut H (2016) Bayesian regularized neural networks for small n big p data. In Artificial Neural Networks-Models and Applications. InTech

Olmez A (2007) Determination of compaction parameters by means of regression approaches. Master's thesis, Niğde Universty, Turkey (in Turkish)

Sabat AK (2015) Prediction of California bearing ratio of a stabilized expansive soil using artificial neural network and support vector machine. Electron J Geotech Eng 20(3):981–991

Sakellariou MG, Ferentinou M (2005) A study of slope stability prediction using neural networks. Geotech Geol Eng 24(3):419–445

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

Samui P, Kothari DP (2011) Utilization of a least square support vector machine (LSSVM) for slope stability analysis. Scientia Iranica 18(1):53–58. https://doi.org/10.1016/j.scient.2011.03.007

Samui P, Sitharam TG, Kurup PU (2008) OCR prediction using support vector machine based on piezocone data. J Geotech Geoenviron Eng ASCE 134(6):894–898. https://doi.org/10.1061/(ASCE)1090-0241(2008)134:6(894)

Sinha SK, Wang MC (2008) Artificial neural network prediction models for soil compaction and permeability. Geotech Geol Eng 26:47–64

Sivrikaya A (2008) Models of compacted fine-grained soils used as mineral liner for solid waste. Environ Geol 53:1585–1595

Sivrikaya A, Kayadelen C, Cecen E (2013) Prediction of the compaction parameters for coarse-grained soils with fines content by MLR and GEP Acta Geotechnica Slovenica, 2013/2

Sridharan A, Nagaraj HB (2005) Plastic limit and compaction characteristics of fine-grained soils. Ground Improv 9(1):17–22

Suman S, Mahamaya M, Das SK (2016) Prediction of maximum dry density and unconfined compressive strength of cement stabilised soil using artificial intelligence techniques. Int J Geosynth Ground Eng 2:11. https://doi.org/10.1007/s40891-016-0051-9

Tenpe A, Kaur S (2015) Artificial neural network modeling for predicting compaction parameters based on index properties of soil. International Journal of Science and Research (IJSR), Volume 4, Issue 7

Vissikirsky VA, Stepashko VS, Kalavrouziotis IK, Drakatos PA (2005) Growth dynamics of trees irrigated with wastewater: GMDH modeling, assessment, and control issues. Instrum Sci Technol 33(2):229–249

Wang MC and Huang CC (1984) Soil compaction and permeability prediction models. J Environ Eng ASCE, Vol 110. 6:1063-1083

Wang HB, Xu WY, Xu RC (2005) Slope stability evaluation using back propagation neural networks. Eng Geol 80:302–315