

# Support vector regression-based model for prediction of behavior stone column parameters in soft clay under highway embankment

Qasim A. Aljanabi<sup>1,2</sup> · Zamri Chik<sup>2</sup> · Mohammed Falah Allawi<sup>2</sup> · Amr H. El-Shafie<sup>3</sup> · Ali N. Ahmed<sup>4</sup> · Ahmed El-Shafie<sup>5</sup>

Received: 31 July 2016 / Accepted: 18 December 2016 / Published online: 9 January 2017  
© The Natural Computing Applications Forum 2017

**Abstract** In order to have a proper design and analysis for the column of stone in the soft clay soil, it is essential to develop an accurate prediction model for the settlement behavior of the stone column. In the current research, to predict the behavior in the settlement of stone column a support vector machine (SVM) method is developed and examined. In addition, the proposed model has been compared with the existing reference settlement prediction model that using the monitored field data. As SVM mathematical procedure has resilient and robust generalization aptitude and ensures searching for global minima for particular training data as well. Therefore, the potential that support vector regression might perform efficiently to predict the ground soft clay settlement is relatively valuable. As a result, in this study, comparison of two different developed types of SVM method is carried out. Generally, significant reduction in the relative error (RE%) and root mean square error has been achieved. Utilizing nu-SVM-type model through tenfold cross-validation procedure could achieve outstanding performance accuracy level with

RE% less than 2% and CR = 0.9987. The study demonstrates high potential for applying SVM in detecting the settlement behavior of SC prediction and ascertains that SVM could be effectively used for settlement stone columns analysis.

**Keywords** Settlement prediction · Soft clay · Stone columns · Embankment · Support vector machines

## 1 Introduction

Soft soil deposits, which are spread all over the world, are a limiting factor for civil engineering constructions. Low shear strength of soft soil deposits causes their excessive settlement, which can lead to circular or sliding failure. Large embankments constructed on such base can be seriously damaged. The structural engineer may be confronted with inconvenient consolidation and displacement, caused by soft soil porosity. That is the reason why geotechnical engineers must prepare the ground improvement schemes. The properties of soft soil deposits are very important for the geotechnical engineering. The soft soil improvement method by using encased stone columns is very effective for reduction in compression time [1].

For calculating settlement and forecast its evolution, the geotechnical engineers use approximate methods for simplifying assumptions and complex methods (finite element method), which are applying the elasticity and plasticity theory. Aboshi et al [2], Barksdale and Bachus [3] and Barksdale and Takefumi [4] described a simple method for calculating the reduction settlement for the soft soil improvement by using encased stone columns. According to Lo et al. [5], the usage of finite element method gives highly accurate results, which were verified by comparing

---

✉ Mohammed Falah Allawi  
mohmmd.falah@gmail.com

<sup>1</sup> Civil Engineering Department, Faculty of Engineering, Diyala University, Baqubah, Iraq

<sup>2</sup> Department Civil and Structural Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, Malaysia

<sup>3</sup> Civil Engineering Department, Giza High Institute for Engineering and Technology, Giza, Egypt

<sup>4</sup> Department of Civil Engineering, College of Engineering, University Tenaga Nasional, Kajang, Malaysia

<sup>5</sup> Civil Engineering Department, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

the numerical results with field measurements. The conclusions exposed that the stone columns could improve the clay bearing capacity and enrich the drainage behavior as well. In addition, the stone columns might dissipate the excess pore water pressure [6].

The artificial intelligence (AI) technique has been recently proved to be capable to represent any nonlinear processes by giving a wide range of complexity of the network. That is the reason for using this technique in wide areas, in different scientific research domains and in industry. Many researchers have applied AI methods with several engineering topics such as [7–11]. The artificial intelligence (AI) technique is also used for prediction over the soil settlement under a large structure [12–14]; the predictions help specialists to avoid future failure problems. Machine learning is a subfield of the artificial intelligence which deals with the methods development for allowing the computer processor to calibrate the natural behavior and learn its nonlinearity process. The development of the AI algorithms would provide the machine to learn and accomplish task that being required. Several techniques and approaches were developed in order for modeling engineering applications and natural behavior utilizing machine learning over particular time period [15]. Boser et al. [16] were the first who introduced support vector machine and the formulation of SVM embody the principle of structural risk minimization which seem to be superior than the traditional principle of empirical risk minimization (ERM) [17] which used by the conventional neural networks. The reason behind this is that the SRM tend to minimize the expected risk upper boundary; however, ERM tend to minimize the error on the training data. By this, the SVM can be equipped with the ability to generalize, the statistical learning goal. Initially, the SVMs were developed to solve the problem of classification; however, the SVM has recently been extended to the domain of regression problem [17, 18].

The SVM had been applied in solving the problem of geotechnical since SVM able to explore data between the several inputs and target variable [19, 20]. Tinoco et al. [21] predicting the jet grouting columns uniaxial compressive strength by using support vector regression. SVM model for predicting liquefaction has been developed by using cone resistance ( $q_c$ ) [22]. In addition, the SVM approach has also been applied in pattern recognition, function approximation and time series [23, 24]. Sun [25] used SVM model to predict deformation value due to deep foundation pit occurred in soft soil area. The results of this proposed SVM model, the fuzzifications of geotechnical engineering problems could be solved in outstanding manner with relatively high matching with the monitored data.

## 2 Methodology

### 2.1 Case study

In this research, the pilot study area is located in Rawang–Ipoh double-tracking project. The project is implementing electrified train and is mainly designed in order to improve the public transportation between Ipoh and Rawang and also to enhance the socioeconomic standard in the area between. In fact, this project is considered as one the part of Trans-Asia Railway line which is aiming to link China with Singapore. The distance of the project area is round 150 km. The designed alignment of the new line is a double track and attached closely to the existing track which is single line; in several locations, both tracks will be a shared. Figure 1 show the whole alignment for the study area of the project.

In this project, vibro-replacement is used with stone columns adopted as ground treatment. It is a use 1 m diameter with spacing 2 m c/c between columns. The presence of silt–clayey in nature—to depths of 9.5 m experienced low shear strength values which introduce serious problems of stability and long-term settlements. The highway embankments in the project have heights ranging from 2 m. The top of the embankment has a minimum width of 14.9 m. The side slopes of the embankments have gradients of 1 V:2H.

### 2.2 Regression in support vector machines

Support vectors are the training points that are the nearest to the separating hyperplane, and the basic concept of SVM is illustrated in Fig. 2. There are decision functions that are accountable, for example, hyperplanes that are able to delineate the positive and negative data that have marked the maximum margins. This shows the range from the nearest positive sample to a hyperplane, and the range between the nearest negative sample and the hyperplane shall be maximized.

$$y(x) = w^T \phi(x) + c \quad (1)$$

where  $\phi(x)$  denotes the high-dimensional of spaces characteristics, which is experienced nonlinearly mapping from the input space  $x$ .  $w$  and  $c$  are the coefficients which are estimated in order to minimize the regularized function:

$$R(C) = C \frac{1}{N} \sum_{i=1}^N L_c(d_i, y_i) + \frac{1}{2} w^2 \quad (2)$$

where

$$L_c(d_i, y_i) = \begin{cases} |d_i - y_i| - \varepsilon & \text{if } |d_i - y_i| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

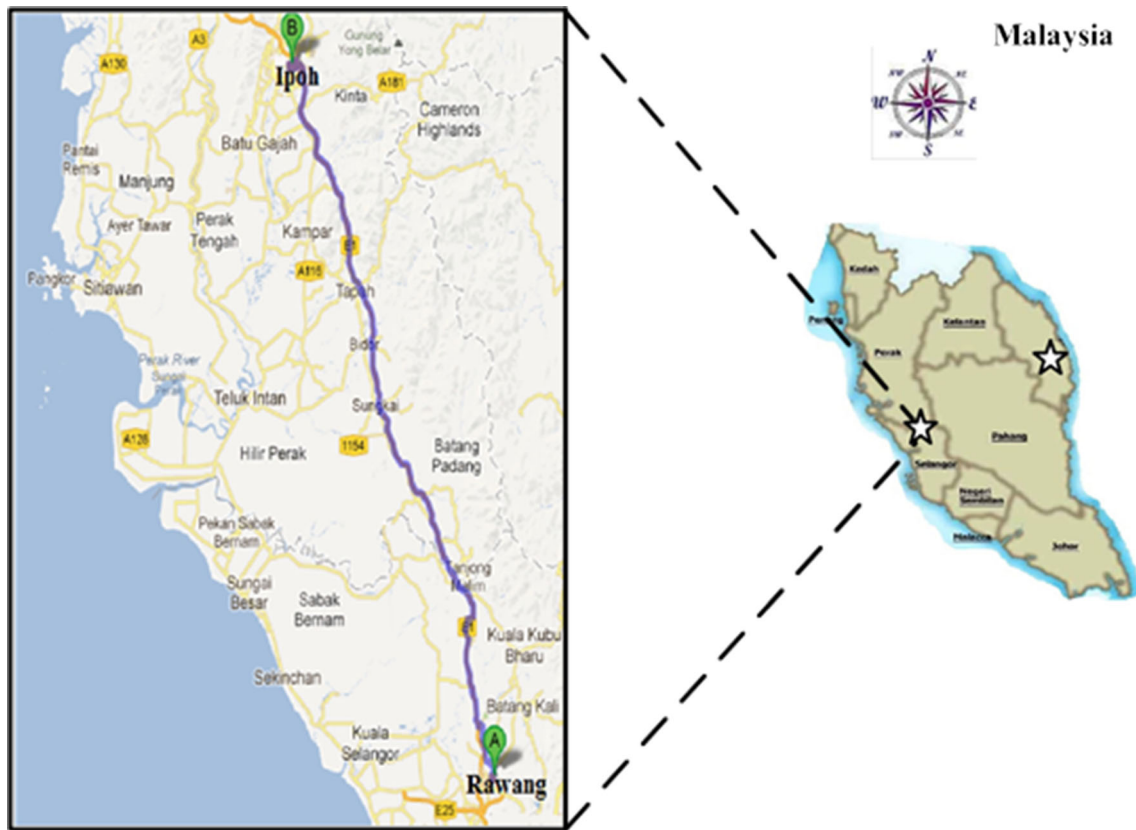


Fig. 1 Location of study sites in Malaysia

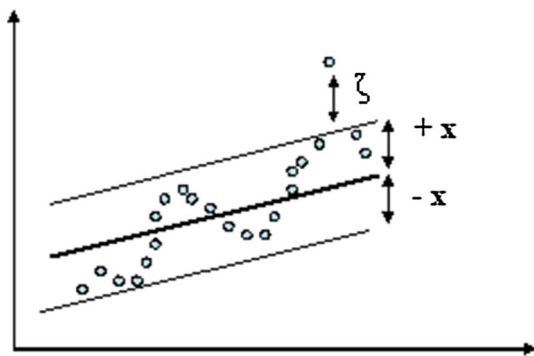


Fig. 2 Hyperplane and the basic concept of SVM

In order to attain the possible estimation values of  $w$  and  $b$ , Eq. (2) is transmuted to the original function which is given by Eq. (4) by compensating the positive slack variables  $\xi_i$  and  $\xi_i^*$  as follows:

$$\begin{aligned} &\text{Minimize } R(C) = C \left( \sum_{a=1}^u (X_a + X_a^*) + \frac{1}{2} E^2 \right) \\ &\text{Subject to } \begin{cases} d_i - w\phi(y_a) - l_a - \leq X + X_a, & a = 1, 2, 3, \dots, u \\ w\phi(y_a) + l_a - r_a \leq X + X_a^*, & a = 1, 2, 3, \dots, u \\ X_a, X_a^* \geq 0 & a = 1, 2, 3, \dots, u \end{cases} \end{aligned} \tag{4}$$

The first term  $(1/2\|E\|^2)$  is the weights vector norm,  $d_i$  is the targeted value and  $C$  is referred as constant value that regularized by defining the trade-off between the regularized terms and the empirical errors.  $x$  is the SVM's tube size as shown in Fig. 2.

It is similar to the equation accuracy that is related in the training data. Karush–Kuhn–Tucker (KKT) condition optimality as mentioned in [26] the usual multipliers will be zero. The multipliers that are nonzero are known as support vectors. This is where the variable slacks might be brought into the study.  $X_a$  and  $X_a^*$  are suggested by introducing this, and by exploiting constraint optimality, the function decision by Eq. (1) gives the following:

$$U(y) = \sum_{a=1}^u (\alpha_a - \alpha_a^*) K(y, y_a) + c \tag{5}$$

In Eq. (3),  $\alpha_a$  and  $\alpha_a^*$  are, respectively, the multipliers for Lagrange function. They fulfill the likenesses  $\alpha_a \times \alpha_a^* = 0$ ,  $\alpha_a \geq 0$  and  $\alpha_a^* \geq 0$  where  $a = 1, 2, 3, \dots, u$  and are found out by maximizing the twofold functions of Eq. (4) which has the following form:

$$R(\alpha_a, \alpha_a^*) = \sum_{a=1}^u d_i(\alpha_a - \alpha_a^*) - \varepsilon \sum_{a=1}^u (\alpha_a + \alpha_a^*) - \frac{1}{2} \sum_{a=1}^u \sum_{e=1}^u d_a(\alpha_a - \alpha_a^*)(\alpha_e - \alpha_e^*)K(y_a, y_e) \tag{6}$$

With the constraints,

$$\sum_{a=1}^u (\alpha_a) = \sum_{a=1}^u (\alpha_a^*) \quad \text{and} \quad \alpha_a, \alpha_a^* \in [0, C], \tag{7}$$

$a = 1, 2, 3, \dots, u$

$K(y_a, y_e)$  is known as the kernel function. The kernel value is equal to the internal value of two vectors  $y_a$  and  $y_e$  in the characteristics space  $\phi(y_a)$  and  $\phi(y_e)$ , that is,  $K(y_a, y_e) = \phi(y_a) \times \phi(y_e)$ . Several kernel function types of SVM, in this study four kernel functions, are introduced as follows:

Linear Kernel:  $k(y_a, y_e) = y_a^T \cdot y_e$  (8)

Polynomial Kernel:  $k(y_a, y_e) = (\gamma y_a^T \cdot y_e + r)^d, \quad \gamma > 0$  (9)

Radial Basis Kernel:  $k(y_a, y_e) = \exp(-\gamma \|y_a - y_e\|^2), \quad \gamma > 0$  (10)

Sigmoid Kernel:  $k(y_a, y_e) = \tanh(\gamma y_a^T \cdot y_e + r)$  (11)

Here  $\gamma, r$  and  $d$  are kernel parameters. There are two types of SVM regression: The first type of SVM regression is known as type 1 or epsilon and the second type of regression is known as nu.

### 2.3 Performance evaluation of SVM model

The proposed model has been developed in three different phases. The first phase is the training session which is performed in order to adjust the model parameters, then switching to the validation phase utilizing unseen data in the training session to make sure that model is successfully accomplished. The aim of the validation session is to guarantee the generalization of the model to be valid for untrained input data and just memorizing the given limit range of the input–output interrelationships that are experienced in the training data session [27].

Evaluation of the performance of ANN models can be carried out by:

1. The coefficient of determination ( $R^2$ ), which is used to measure the relation between the observed and predicted data.

Or

$$R^2 = \frac{\sum_{k=1}^n (A_k - D)^2}{(P_k - D)^2} \tag{12}$$

$$R^2 = \frac{(n \sum_1^n S_m S_p - (\sum_1^n S_m)(\sum_1^n S_p))^2}{(n(\sum_1^n S_m^2) - (\sum_1^n S_m)^2)(n(\sum_1^n S_p^2) - (\sum_1^n S_p)^2)} \tag{13}$$

where  $A_k$ : actual output value,  $P_k$ : predicted output,  $D$ : mean of the desired output,  $n$ : number of data.

Root mean square error (RMSE): it is furthermost widely index to calculate the bias

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (P_k - A_k)^2} \tag{14}$$

2. The mean absolute percentage error (MAPE).  
The mean absolute percentage error (MAPE) is a commonly utilized indicator in pattern recognition process in order for examining the level of accuracy to most of time series prediction model. Once the value of MAPE is near to zero, it gives the indicator the model is performing better in fitting the data. Actually, the MAPE value is the summation of the difference values between the model prediction values and the actual data during a particular session and then divided by the number of the records  $n$ . The following formula is mathematical expression for calculated the MAPE; notice that the value is in percentage form.

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{A_k - P_k}{P_k} \right| \tag{15}$$

3. The mean absolute error (MAE).

$$MAE = \frac{1}{n} \sum_{k=1}^n |P_k - A_k| \tag{16}$$

All these statistical analysis are used to check the validity and test the robustness of the SVM models.

### 2.4 Tenfold cross-validation of SVM model

One of the major methods used for not only quantifying the performance of prediction or classification models but also guaranteeing the generalization of the model is cross-

validation, which is usually used before switching the model from the training session to testing session which is considered as a sub-tree for the whole tree. Regularly, this method used for models applied for time series data. In fact, the modeling procedure could be similar to the tree concept. The tree perception is as growing a binary tree in order for expecting a desire value or category  $Y$  for input pattern  $V_1, V_2, \dots, V_p$ . The tree is consistent of leaves which is considered as incurable nodes and characterize the cells of the partition that keeps the model valid for precise cell. The examination achieved from the input  $V_i$  could be functionalized for every node at the tree level, and the results could occur around surrounding the tree sub-branch. At the leaf node where the forecasting is done, certain validation model procedure at the root node in the tree could be established. There are several method of cross-validation such as  $V$ -fold, Monte Carlo cross-validation and leave-one-out. Several research efforts have been addressed that  $V$ -fold range could be very effective in developing these kinds of models. Those researches reported that a range for  $V$  between 7 and 20 showed outstanding performance and results over  $V$  values lower than 7. Feng and Derynck [28] utilized statistical index testing and other indicators for prediction error to measure up the competing models.

Predominantly, in case the model developed utilizing SVM method, for example, if  $V$  equals to 10, it means that the data would split into different 10 subsets of identical size by the model and the process of the training will be repeated 10 times. For each training process, the model is running utilizing subsets for training and leaves one for examining the model error. On the other hand, although the leave-one-out cross-validation method experienced providing respectable results and relatively high accuracy, it encounters a serious drawback of over-fitting difficulty for the data. One of the major steps in the model selection in terms of the model internal parameters is how to restructure the data for training data set and the desired output data set to assure betterment of the prediction procedure. In this research, several type and length of training, cross-validation and testing structure were evaluated in order to achieve the optimal results.

### 3 Application and analysis

#### 3.1 Parameter characteristics of SVM

In this study, the proposed SVM model has been developed and evaluation in order to be to get the perfect model, which gives the most accurate results by using various functions as polynomial, sigmoid, linear and radial basis function [16, 21]. These models are common prior to use in other engineering applications and have been compared to

the final results with the results obtained from tests for the study of precipitation field common ways that (kernel function) the most common models used are known by researchers for their accuracy and their ability models, the four offered by previous studies, were examined and evaluated the use of the most powerful and most accurate model of which (RBF) for the purposes of its use in this study, as shown in Table 1.

Applications of searching procedure have been carried out to find out the finest kernel function type while using cross-validation-based tenfold process. The total partitioning training data are calibrated by applying the SVMs with different kernels in order to create the final models architecture. Table 2 shows a comparative analysis for the performance of the prediction skills for the  $S_{total}$  using the proposed kernel functions through SVM method. It could be observed that with respect to correlation coefficient value ( $R^2$ ) could be achieved while using radial basis function as the kernel function with best level equal to 0.98 during the test data session. On the other hand, the  $R^2$  is equal to 0.95, 0.94 and 0.91 utilizing linear, polynomial and sigmoid kernel function

Actually, searching for the best structure of SVM model for particular application, there are two vital parameters that supposed to be selected, namely capacity parameter  $C$  and  $\epsilon$ . The selection of  $C$  is very sensitive to the accuracy of the prediction, as the small value of  $C$  could tend the model to underestimate the target value during the training data; this is due to the fact that using relatively small weight in the training data would reflect a larger values of the predictor while examining the model in the testing data set and vice versa. On the top of that, when  $C$  is large, the weight will lose its significance in detecting the mapping between the input and the output.

Alternatively, the large value of  $C$  could reflect a wide range of support vector's values; accordingly, supplementary data records could be chosen for optimizing the support vectors. Furthermore, large value of  $\epsilon$  could tend for less number of the support vector to be achieved, and then,

**Table 1** Compared total settlement values ( $S_{total}$ ) in field with result values by using four-kernel-function SVM model

$S_{total}$ (measured in field)	$S_{total}$ by used kernel-type function			
	Linear	Polynomial	RBF	Sigmoid
0.1294	0.0653	0.2093	0.0754	-1.86
0.3778	0.2436	0.2250	0.2767	-0.702
0.4627	0.2516	0.2261	0.4524	-0.571
0.6113	0.4291	0.3097	0.6361	0.5029
0.8465	0.6364	0.5836	0.8510	1.6534
1.0000	0.8094	0.8306	0.1155	4.0183
0.1358	0.0130	0.1752	0.3096	-0.837



**Table 2** Performance of  $S_{\text{total}}$  SVM model with four kernel functions

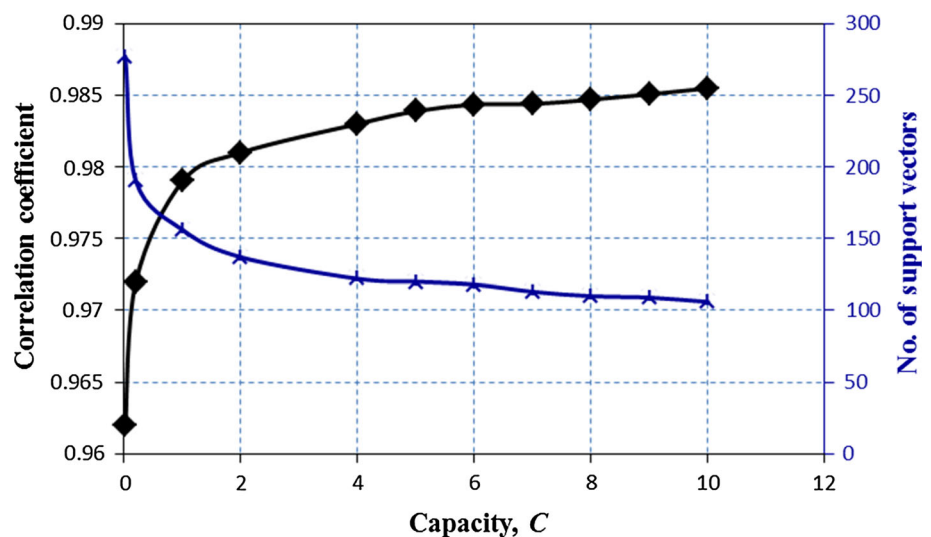
Kernel function	Mean squared error Testing	Correlation coefficient Testing
Linear	0.006183	0.956483
Polynomial	0.008233	0.941468
RBF	<b>0.002766</b>	<b>0.981186</b>
Sigmoid	3.506396	0.0115714

Bold values represent the best results

the expected illustration of the proposed elucidation is insufficient. Additionally, too large value of  $\varepsilon$  could lead to deterioration of the level of accuracy during the training. In this context, the selection of the optimal values of both  $C$  and  $\varepsilon$  should be achieved via several trial and error procedures. Once the optimal values of these two parameters are identified, there is a high potential to achieve high level of accuracy for predicting the desired data.

Radial basis function (RBF) kernel is conceptually used as the stepwise searching methodology to evaluate the performance of SVM [25]. In current study, with the purpose of determining the appropriate values of parameter  $C$  and  $\varepsilon$ , the replication processes with several parameters sceneries were also seen the sights of SVM  $S_{\text{total}}$  as major step of behavior prediction model. At the beginning, a consideration of constant value of  $\varepsilon$  to be 0.1, and changeable values of  $C$  to be ranged between 0 and 10 in order to build the proposed model during the training session of the input–output data. Consequently, a calculation of the prediction error as RMSE and  $R^2$  is carried out with identifying the number of the number of the support vectors. It could be depicted in Fig. 3 that slight reduction in the number of the support vectors and the value of RMSE is achieved while the used value of  $C$  increased; on the other hand, the value of the  $R^2$  increases. In addition,

**Fig. 3** Result of various capacity parameter  $C$ , of SVM model where  $\varepsilon = 0.1$  and  $\gamma = 0.2$



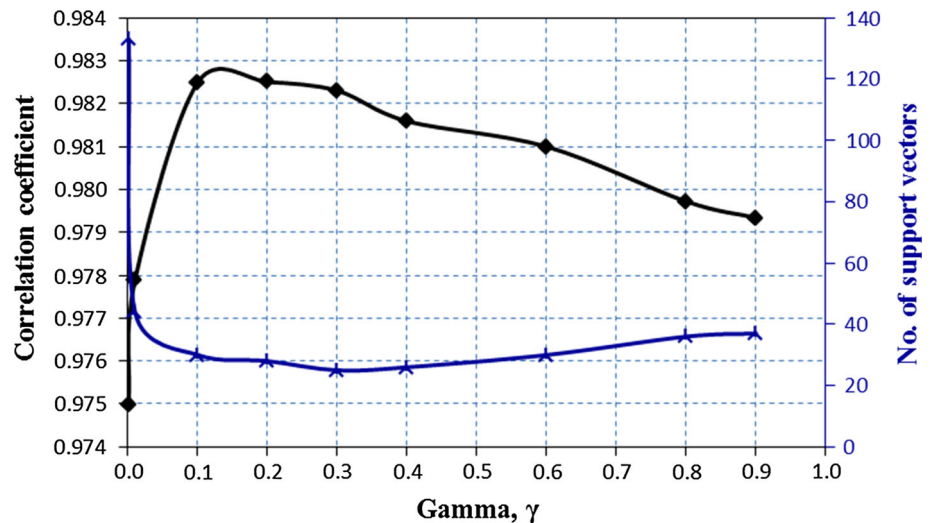
with focusing on the parameter  $C$ , it could be observed that the lowest value of RMSE point (0.00114) and one high correlation coefficient value (0.987588). RMSE first decreases marginally with the increase in parameter  $C$  and then grows again after the optimal point. Consequently, it is better to select parameter  $C$  to be 8.00.

Actually, it is necessary to figure out the appropriate values of the hyper-parameters  $C$  and  $\gamma$  as a major step while implementing any SVM model. Therefore, there is a need to carry out trial and error procedure. In this context, estimation of the generalized accuracy utilizing various value of kernel hyper has been done. With respect the limitation of parameter  $\gamma$ , it was decided to search its optimal value to be within the ranges of [0.001, 0.9] at increment of 0.1 for  $\gamma$  with being fixed to each  $C = 8$  and  $\varepsilon = 0.2$ . Consequently, the optimal selection value of  $\gamma$  is found out using tenfold cross-validation with repetitive ten times in order to enhance the reliability of the model results. The final architecture of the model is established when the minimum root mean square error is achieved during the validation session. Figure 4 shows the relationship between correlation coefficient and  $\gamma$ , where it began in the value of correlation coefficient rise with increasing  $\gamma$  until it reaches its peak value at 0.2 and after that value starts to decline along with a gradual decrease in the number of support vectors. The value of parameters that provides the minimum generalization error is then selected. The best result for the  $S_{\text{total}}$  model in training and forecasting phase when selecting  $\gamma = 0.2$  with an acceptable value of number of support vectors 28 (Fig. 5).

### 3.2 SVM model development

Generally, selecting the optimal number of parameters inputs of any particular SVM model is an essential step; however, up to date it is difficult to find certain theory that

**Fig. 4** SVM model performance using different values of  $\gamma$  where  $\varepsilon = 0.2$  and  $C = 8$



could be used to attendant achieving this step. While performing the training and testing session of SVM model, the same input arrangements of the data set of stone column parameters that were used have been mentioned in the previous sections of this paper. In fact, searching for the model parameters plays a crucial role in achieving a good performance of the SVM. Setting the hyper-parameters  $C$ ,  $\gamma$  and the kernel parameters (epsilon and nu) is considered as vital step in influencing the SVM generalization performance (estimation accuracy).

This section will be focused on the use of the two types of RBF kernel for its good performance and advantages in SC forecasting problem. The following explains optimized and selected kernel parameter values for each of the two models.

### 3.2.1 Epsilon-RBF model

Epsilon-RBF model is used for predicting  $S_{total}$ ; we fix  $C$  to be eight and  $\gamma = 0.2$ , and set  $\varepsilon$  as various values between 0.001 and 0.9. The results of model for training and testing are shown in Table 3; it can be observed that the RMSE increases with increasing the value of  $\varepsilon$ , while each of the  $R^2$  and the number of support vectors decreased. Finally, the  $\varepsilon$  value (0.2) that yields the minimum generalization error with acceptable number of support vector (28) is then chosen.

### 3.2.2 Nu-RBF model

In this model, we used optimal values for each gamma = 0.2, capacity = 8 SVM parameters. As nu model value increases, gradual increasing in the accuracy of the prediction for the test data is observed until the value of NU reach 0.4, and then, the values of the remaining are kept unchanged as shown in Table 4.

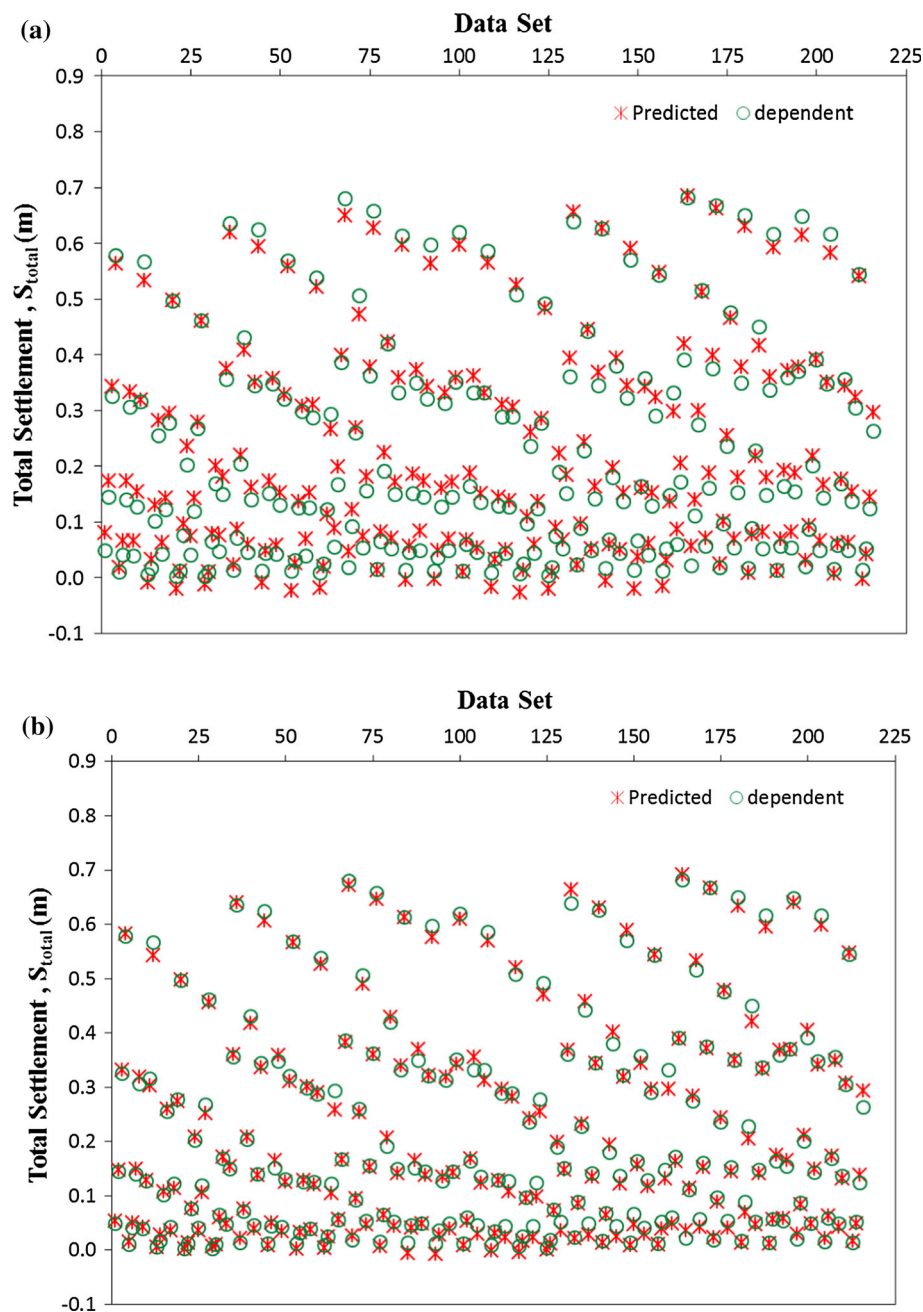
The cross-validation process is considered as one of the widely used procedures for evaluating the model architecture parameter values. In order to use the  $V$ -fold cross-validation, the training data set is unsystematically split into a set number of  $V$ -fold ( $V_1, V_2, \dots, V_n$ ). Then the selected type of SVM model is performed sequentially to the dataset that taken place to the  $V - 1$ -fold. The process then could achieve the results of the acting architecture with particular parameters on the sample  $V$  (the sample or fold that was unseen while training the SVM model; i.e., this is the testing sample) in order to figure out the error defined by one of the statistical indices. The major advantage of this process is that the average accuracy for the  $V$  times could lead to consistent measure model error and its stability, i.e., the validity of the model for testing session unseen data.

Table 5 illustrates that the MAE and RMSE values achieved utilizing tenfold cross-validation show best goodness fitting and outstanding performance if compared while utilizing 15-fold cross-validation. In addition, comparable results were attained for the MAPE values. It should be reported here that the major challenging of employing the cross-validation procedure in the current research is the choice of the size of the utilized data set. It seems essential for the choice to be representing for the feature for both the training the model and testing session.

However, after the optimal kernel parameters are found, and nu-RBF model is selected as optimal model, the whole training data of behavior SC parameters are trained using the optimal nu-RBF model and then tested. Figures 6 and 7 exhibit the prediction accuracy with percentage of error for the test data of behavior settlement of SC.

Another model to predict lateral displacement of stone column developed is shown in Fig. 7. The number of the inaccurate lateral bulging of stone column (Uh) prediction decreased significantly with the SVM method. The

**Fig. 5** Comparison of actual versus predicted behavior total settlement: **a** Nu-SVM-type model, **b** Epsilon-SVM-type model



application of physics-based distributed process complex computer software programs is often problematic, due to the need for massive amounts of detailed spatial and temporal environmental data which are not available [29]. In particular, that is the application of AI to the behavior settlement of stone column data estimation is limited in the literature. Lateral bulging of stone column ( $U_h$ ), model in figure below, illustrates the comparison between the predicted  $U_h$  and the measured  $U_h$  using 100% agreement line ( $45^\circ$ ) of graph and two deviation lines from the agreement line for both validation and testing data sets for the developed models. It was obvious that SVM model could

predict the  $U_h$  with relatively good level of accuracy, whereby the error for majority of the records did not reach 16%.

#### 4 Conclusion

In this study, the proposed SVM model and support vector regression have confirmed their attainment predicting and analyze settlement and statistical learning. Nevertheless, few works have been done for predicting the settlement behavior of stone column. In this paper, evaluation of the



**Table 3** Result of various epsilon parameter  $\epsilon$ , of SVM model where  $C = 8$  and  $\gamma = 0.2$  of  $S_{total}$  behavior prediction model

Epsilon, $\epsilon$	RMSE Testing	Correlation coefficient Testing	No. of support vectors
0.001	0.00085	0.99021	222
0.1	0.00116	0.98722	110
<b>*0.2</b>	<b>0.00177</b>	<b>0.98251</b>	<b>28</b>
0.3	0.00365	0.98158	16
0.4	0.00658	0.98111	10
0.5	0.01202	0.97709	10
0.6	0.01695	0.95826	8
0.7	0.02296	0.9597	7
0.8	0.03014	0.96015	6
0.9	0.0383	0.94957	4

Bold values represent the best results

**Table 4** Nu-RBF performance utilizing different values of nu with fixed ( $\gamma = 0.2$ , capacity = 8) of  $S_{total}$  behavior prediction model

Nu	RMSE Testing	Correlation coefficient Testing	No. of support vectors
0.001	0.049	0.933	8
0.01	0.009	0.972	16
0.1	0.004	0.985	62
0.2	0.003	0.988	118
0.3	0.002	0.989	150
<b>0.4</b>	<b>0.002</b>	<b>0.989</b>	<b>175</b>
0.5	0.003	0.989	188
0.6	0.003	0.987	201
0.7	0.003	0.987	201
0.8	0.002	0.989	203
0.9	0.002	0.988	202
1	0.002	0.990	202

Bold values represent the best results

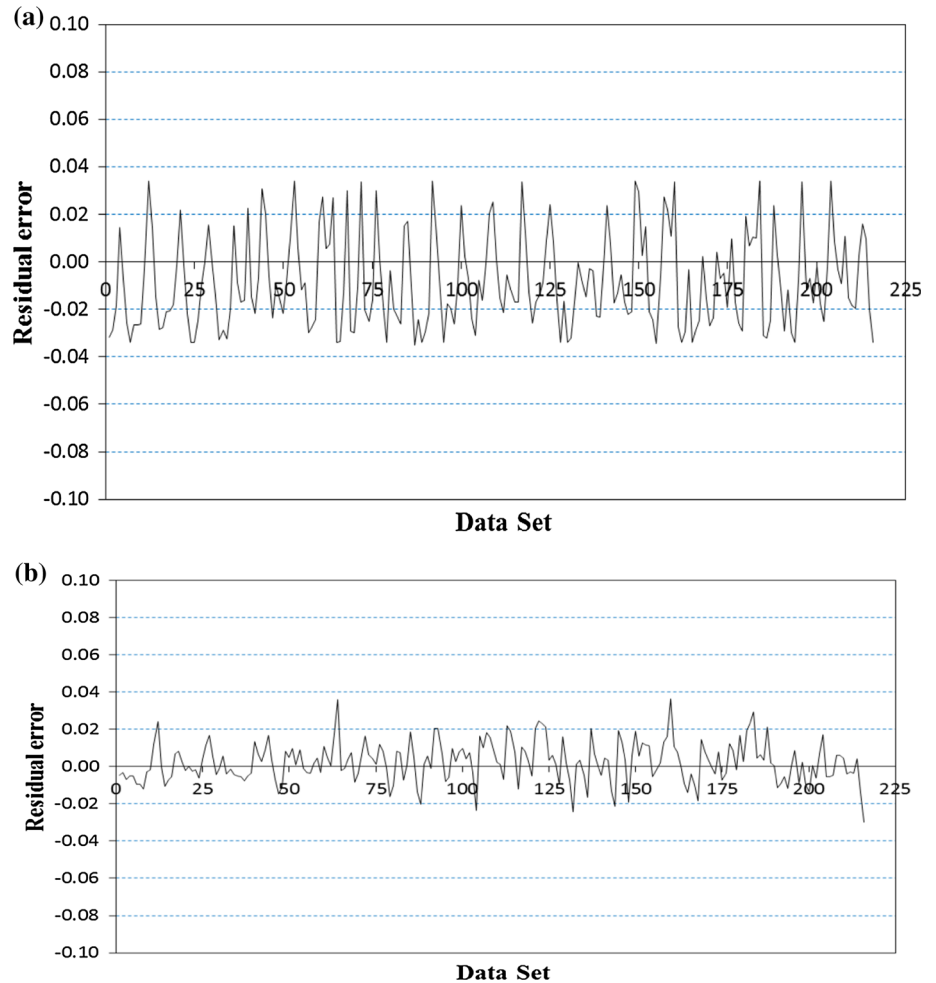
**Table 5** Statistical evaluation Stotal using 3-, 5-, 7-, 10- and 15-fold cross-validation for epsilon-RBF and nu-RBF models

Statistical evaluation	V-fold				
	3	5	7	10	15
<i><math>\epsilon</math>-RBF model</i>					
RMSE	0.003281	0.002807	0.002875	0.002686	0.002875
MAPE	0.177323	0.161928	0.16452	0.159176	0.16452
MAE	0.045329	0.041692	0.040612	0.040379	0.040612
CR	0.992476	0.997101	0.996374	0.997327	0.996374
<i>Nu-RBF model</i>					
RMSE	0.003278	0.002686	0.002531	<b>0.002398</b>	0.002531
MAPE	0.177171	0.159176	0.164146	<b>0.153674</b>	0.154886
MAE	0.04529	0.041682	0.04052	<b>0.037099</b>	0.038168
CR	0.993328	0.997327	0.998637	<b>0.9987</b>	0.998637

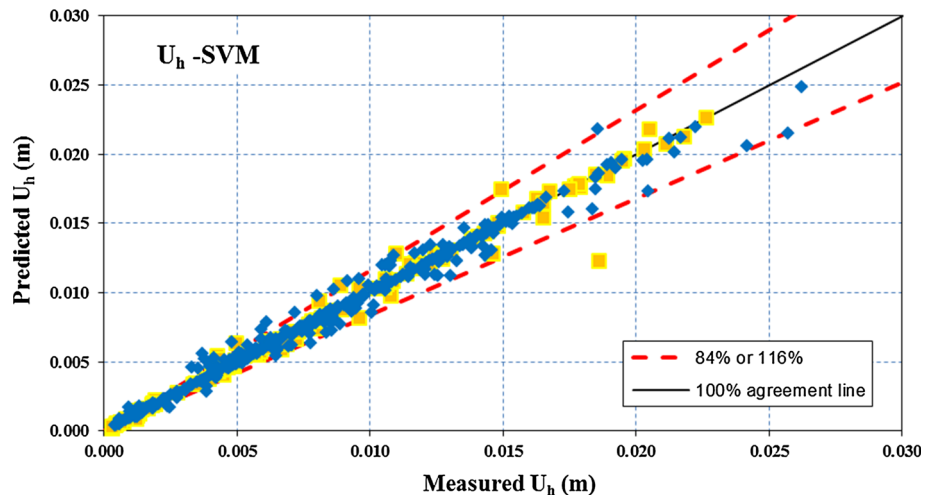
Bold values represent the best results

feasibility of employing support vector regression in mimicking the settlement behavior of stone column prediction is reported. Afterward various experiments propose set of SVM parameters that could be used to predict behavior settlement of SC with relatively high level of accuracy. In addition, the results depict that the proposed SVM model predictor meaningfully outperforms the other baseline prediction model. This evidences the applicability of support vector regression in settlement of stone column data analysis. SVM technique with varying fold cross-validation utilized to prediction of settlement behavior of stone column embedded in soft clay soils under embankment. In order for achieving the best result of SVM regression selection, several V-fold cross-validation procedures were examined. Type nu-SVM provides CR = 0.9987, which is higher than the 0.9973 by the type epsilon-SVM. On the other hand, tenfold cross-validation showed better performance over the other higher V-fold cross-validation.

**Fig. 6** Comparison of residual error for behavior total settlement: **a** nu-SVM-type model, **b** Epsilon-SVM-type model



**Fig. 7** Predict bulging of stone columns by using SVM model



In brief, utilizing the nu-SVM regression with tenfold cross-validation could achieve better prediction accuracy with maximum error equal to 2% and CR equal to 0.9987. Although the results appeared to be good, the application of the soil settlement is very sensitive application to the level of error. 2% as a maximum error

is relatively high in such application, and then, there is a need to enhance it.

**Acknowledgements** The authors would like to thank University Kebangsaan Malaysia (UKM) and Department of Civil and Structural Engineering, for the use of computer laboratory in this work.

## Compliance with ethical standards

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

## References

- Indraratna B, Basack S, Rujikiatkamjorn C (2013) Numerical solution of stone column-improved soft soil considering arching, clogging, and smear effects. *J Geotech Geoenviron* 139(3):377–394. doi:[10.1061/\(ASCE\)GT.1943-5606.0000789](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000789)
- Aboshi H, Ichimoto E, Enoki M, Harada K (1979) The compozer—a method to improve characteristics of soft clays by inclusion of large diameter sand columns. In: International conference on soil reinforcement: reinforced earth and other techniques, Paris, pp 211–216
- Barksdale RD, Bachus RC (1983) Design and construction of stone columns, vol I, and Vol. II. FHWA/RD-83/026, Federal Highway Administration, Washington, DC
- Barksdale RD, Takefumi T (1990) Design, construction and testing of sand compaction piles, symposium on deep foundation improvements: design. Construction and testing. ASTM Publications, Las Vegas
- Lo SR, Zhang R, Mak J (2010) Geosynthetic-encased stone columns in soft clay: a numerical study. *Geotext Geomembr* 28(3):292–302. doi:[10.1016/j.geotextmem.2009.09.015](https://doi.org/10.1016/j.geotextmem.2009.09.015)
- Bo MW, Choa V (2004) Reclamation and ground improvement. Thomson Learning, Singapore
- Allawi MF, El-Shafie A (2016) Utilizing RBF-NN and ANFIS methods for multi-lead ahead prediction model of evaporation from reservoir. *Water Resour Manag* 30(13):4773–4788
- Ahmed AN, Noor CM, Allawi MF, El-Shafie A (2016) RBF-NN-based model for prediction of weld bead geometry in Shielded Metal Arc Welding (SMAW). *Neural Comput Appl*. doi:[10.1007/s00521-016-2496-0](https://doi.org/10.1007/s00521-016-2496-0)
- Elzwayie A, El-Shafie A, Yaseen ZM, Afan HA, Allawi MF (2016) RBFNN-based model for heavy metal prediction for different climatic and pollution conditions. *Neural Comput Appl*. doi:[10.1007/s00521-015-2174-7](https://doi.org/10.1007/s00521-015-2174-7)
- Pelletier G, Mailhot A, Villeneuve JP (2003) Modeling water pipe breaks—three case studies. *J Water Resour Plan Manag* 129(2):115–123
- Hipni A, El-shafie A, Najah A, Karim OA, Hussain A, Mukhlisin M (2013) Daily forecasting of dam water levels: comparing a support vector machine (SVM) model with adaptive neuro fuzzy inference system (ANFIS). *Water Resour Manag* 27(10):3803–3823
- Shahin MA, Indraratna B (2006) Modeling the mechanical behavior of railway ballast using artificial neural networks. *Can Geotech J* 43(11):1144–1152. doi:[10.1139/t06-077](https://doi.org/10.1139/t06-077)
- Choobbasti AJ, Farrokhzad F, Barari A (2009) Prediction of slope stability using artificial neural network (case study: Noabad, Mazandaran, Iran). *Arab J Geosci* 2(4):311–319. doi:[10.1007/s12517-009-0035-3](https://doi.org/10.1007/s12517-009-0035-3)
- Soyupak S, Karaer F, Gurbuz H, Kivrak E, Senturk E, Yazici A (2003) A neural networkbased approach for calculating dissolved oxygen profiles in reservoirs. *Neural Comput Appl* 12(1):166–172
- Cristianini N, Shawe-Taylor John (2000) An introduction to support vector machines: and other kernel-based learning methods. Cambridge University Press, Cambridge
- Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: 5th Annual ACM workshop on COLT. ACM Press, Pittsburgh, PA
- Juran I, Guermazi A (1988) Settlement response of soft soils reinforced by compacted sand columns. *J Geotech Geoenviron Eng ASCE* 114(8):903–943
- Vapnik V, Golowich S, Smola A (1997) Support vector method for function approximation, regression estimation, and signal processing. In: Mozer M, Jordan M, Petsche T (eds) Advances in neural information processing systems 9. MIT Press, Cambridge
- Goh ATC, Goh SH (2007) Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data. *Comput Geotech* 34(5):410–421. doi:[10.1016/j.compgeo.2007.06.001](https://doi.org/10.1016/j.compgeo.2007.06.001)
- Tinoco J, Correia AG, Cortez P (2011) Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time. *Constr Build Mater* 25(3):1257–1262. doi:[10.1016/j.conbuildmat.2010.09.027](https://doi.org/10.1016/j.conbuildmat.2010.09.027)
- Tinoco J, Correia A, Cortez P (2012) Jet grouting mechanicals properties prediction using data mining techniques. In: Proceedings of the fourth international conference on grouting and deep mixing, pp 2082–2091
- Samui Pijush (2013) Liquefaction prediction using support vector machine model based on cone penetration data. *Front Struct Civ Eng* 7(1):72–82
- Pijush Samui (2008) Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Comput Geotech* 35(3):419–427
- Man C, Wang S, Wang W, Zhao J (2011) Subgrade settlement prediction based on support vector machine. Paper presented at the 6th international forum on strategic technology, Harbin, Heilongjiang, China
- Sun F (2010) SVM in predicting the deformation of deep foundation pit in soft soil area. Paper presented at the 2010 international conference on machine vision and human-machine interface, Kaifeng, China
- Fletcher R (1987) Practical methods of optimization. Wiley, New York
- Shahin MA, Maier HR, Jaksza MB (2002) Predicting settlement of shallow foundations using neural networks. *J Geotech Geoenviron* 128(9):785–793. doi:[10.1061/ASCE/1090-0241/2002/128/9/785](https://doi.org/10.1061/ASCE/1090-0241/2002/128/9/785)
- Feng XH, Derynck R (2005) Specificity and versatility in TGF- $\beta$  signaling through Smads. *Annu Rev Cell Dev Biol* 21:659–693
- Cigizoglu HK (2004) Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Adv Water Resour* 27(2):185–195