



A HYBRID SUPPORT VECTOR REGRESSION WITH ANT COLONY OPTIMIZATION ALGORITHM IN ESTIMATION OF SAFETY FACTOR FOR CIRCULAR FAILURE SLOPE

H. Fattahi^{*,†}

Department of Mining Engineering, Arak University of Technology, Arak, Iran

ABSTRACT

Slope stability is one of the most complex and essential issues for civil and geotechnical engineers, mainly due to life and high economical losses resulting from these failures. In this paper, a new approach is presented for estimating the Safety Factor (SF) for circular failure slope using hybrid support vector regression (SVR) and Ant Colony Optimization (ACO). The ACO is combined with the SVR for determining the optimal value of its user-defined parameters. The optimization implementation by the ACO significantly improves the generalization ability of the SVR. In this research, the input data for the SF estimation consists of the values of geometrical and geotechnical input parameters. As an output, the model estimates the SF that can be modeled as a function approximation problem. A data set that includes 46 data points is applied in current study, while 32 data points are used for constructing the model, and the remainder data points (14 data points) are used for assessment of the degree of accuracy and robustness. The results obtained show that the hybrid SVR with ACO model can be used successfully for estimation of the SF.

Keywords: safety factor; support vector regression; circular failure slope; ant colony optimization.

Received: 2 July 2015; Accepted: 30 September 2015

1. INTRODUCTION

Failures of slope are complex natural phenomena that constitute a serious natural hazard in many countries. Engineering assessment of the stability of slope is commonly performed using different computational methods such as limit equilibrium method (LEM) [1], finite element method (FEM) [2], boundary element method (BEM) [3] and finite difference

*Corresponding author: Department of Mining Engineering, Arak University of Technology, Arak, Iran

†E-mail address: H.fattahi@arakut.ac.ir (H. Fattahi)

method (FDM) in determining its susceptibility to failure in terms of the safety factor (SF). Out of the above different methods, the LEM is the most widely utilized approach for the stability of slope analysis. A detailed review of LEMs of slope stability analysis is presented by Duncan [4]. In recent years, in field of the stability of slope modeling, with the development of cheaper personal computer, soft computing approaches has been increasingly used in the stability of slope analysis such as: stability of slope prediction using artificial neural networks (ANNs) [5], slope stability prediction using fuzzy logic (FL) [6], finding the critical SF in slope stability analysis using simple genetic algorithm (GA) [7], using ant colony optimization algorithm [8] and using particle swarm optimization (PSO) algorithm [9]. The advantage of a soft computing approach in the analysis of slope stability problems over traditional LEMs is that no assumption needs to be made in advance about the shape or location of the failure surface, slice side forces and their directions.

In field of slope stability modeling, although previous efforts are valuable and revealed the better performance of intelligence-based models in preference to scaling equation, the pursuit of the novel model introduction with more accurate results is always ongoing. The SVR is a potent data mining model, which was developed by Vapnik [10] and co-workers based on statistical learning theory for solving problems encountered in petroleum industry. Although this method is a powerful methodology for modeling of different phenomena, it suffers from some shortcomings, which limit its application. In every SVR modeling, a series of user-defined parameters exist that required to be chosen by user precisely. Incorrect input of aforementioned parameters by user can lead to erroneous and even deceptive results. Hence, it is crucial to employ a potent optimization algorithm for searching the proper value of these parameters [11]. By now, there have been several optimization algorithms, such as genetic algorithm (GA) inspired by the Darwinian law of survival of the fittest [12] and biogeography-based optimization (BBO) inspired by the migration behavior of island species. Also, recently new optimization algorithms are developed consisting of charged system search (CSS) [13], ray optimization (RO) [14], democratic particle swarm optimization (DPSO) [15], colliding bodies optimization (CBO) [16] and enhanced colliding bodies optimization (ECBO).

In the present paper, in order to achieve the above goal, ant colony optimization (ACO) is applied as the searching strategy for finding the optimal value of user-defined parameters. ACO is capable of improving the performance of SVR through determining their free parameters. Integration of SVR model and ACO method produced a model, which can estimate the amount of the SF for circular failure slope with good precision.

2. HYBRID SUPPORT VECTOR REGRESSION WITH ANT COLONY OPTIMIZATION

2.1 Support vector regression

Support vector machines (SVMs) has been first proposed by Vapnik [10]. There are two main categories for SVMs: support vector regression (SVR) and support vector classification (SVC). SVM is a learning system using a high dimensional feature space. It yields prediction functions that are expanded on a subset of support vectors. SVM can generalize complicated gray level structures with only a very few support vectors and thus

provides a new mechanism for image compression. A version of a SVM for regression has been introduced in 1997 by Vapnik, Steven Golowich, and Alex Smola [17]. SVR is the most common application form of SVMs. An overview of the basic ideas underlying SVMs for regression and function estimation has been given in [18].

Let the training samples be denoted as $XY = \{(x, y) | (x_1, y_1), \dots, (x_n, y_n)\}$, where n is the number of training samples. In SVR, the ultimate goal is to find linear relation between n -dimensional input vectors $x \in R^n$, and output variables $y \in R$ as follow:

$$f(x) = w^T x + b \tag{1}$$

Where, b and w are offset of the regression line and the slope respectively. For determining the values of b and w , it is necessary to minimize following equation;

$$R = \frac{1}{2} \|w\|^2 + \frac{C}{l} \sum_{i=1}^l |y_i - f(x_i)|_\epsilon \tag{2}$$

Loss function, utilized in SVR is ϵ -insensitive which has been proposed by Vapnik [10] as below;

$$|y_i - f(x_i)|_\epsilon = \begin{cases} 0 & \text{if } |y_i - f(x_i)| \leq \epsilon \\ |y_i - f(x_i)| - \epsilon & \text{Otherwise} \end{cases} \tag{3}$$

This problem can be reformulated in a dual space by;

$$\text{Maximize } L_p(\alpha_i, \alpha_i^*) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j - \epsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i \tag{4}$$

$$\text{Subject to } \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l \\ 0 \leq \alpha_i^* \leq C, \quad i = 1, \dots, l \end{cases} \tag{5}$$

Where, $\alpha_i, \alpha_i^* \geq 0$ are positive Lagrange multipliers. C is regulated positive parameter which determines trade-off between the weight vector norm $\|w\|$ and approximation error. After calculation of Lagrange multipliers α_i and α_i^* , training data points, those meeting the conditions $\alpha_i - \alpha_i^* \neq 0$, will be applied to construct the decision function. Hence, the best linear hyper surface regression is given by;

$$f(x) = w_0^T x + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i^T x + b \tag{6}$$

Which desired weight vector of the regression hyper plane is given by;

$$w_0 = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i \quad (7)$$

In nonlinear regression, Kernel function is applied for mapping input data onto higher dimensional feature space in order to generate a linear regression hyper plane. In the case of the nonlinear regression, the learning problem is again formulated in the same way as in a linear case, i.e., the nonlinear hyperplane regression function becomes;

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (8)$$

where, $K(x_i, x)$ is kernel function which is defined as follow;

$$K(x_i, x_j) = \Phi^T(x_i) \Phi(x_j) \quad i, j = 1, \dots, l \quad (9)$$

Where, $\Phi(x_i)$ and $\Phi(x_j)$ are projection of the x_i and x_j in feature space respectively.

One may choose any arbitrary kernel functions, e.g., Radial Basis Function (RBF) $K(x_i, x_j) = \exp(-\|x_i - x_j\|/2\sigma^2)$, $\sigma > 0$, linear kernel function $K(x_i, x_j) = (x_i, x_j)$, polynomial kernel function $K(x_i, x_j) = ((x_i, x_j) + 1)^d$, $d > 0$, etc. In highly non-linear spaces, RBF kernel usually yields more promising results in comparison with other mentioned kernels [19]. Thus, we use only RBF kernel functions in this study.

2.2 Parameters optimization of the support vector regression based on ant colony optimization

The generalization ability of the SVR is extremely dependent upon its learning parameters, i.e., the RBF kernel parameter $\sigma \in [2^{-5}, 2^2]$, the error margin $\varepsilon \in [0.01, 0.6]$, and the regularization parameter $C \in [2^{-5}, 2^{15}]$, to be set correctly. Finding the best combination of hyper-parameters is often troublesome due to the highly non-linear space of the model performance with respect to these parameters. Although an exhaustive search method could be utilized to tune these hyper-parameters, it suffers from the main drawbacks of being very time-consuming and lacking a guarantee of convergence to the globally optimal solution. For example, the real-value GA was employed to determine the optimal parameters of SVR, which were then applied to construct the SVR model, referred to as SVR-GA [20]. The PSO has also been used to select the model parameters of SVR by several researchers [21]. Recently, the Harmony Search (HS) has also been utilized to select the model parameters of SVR [22].

In this paper, we have adopted the ACO for optimal parameter selection of SVR in order to improve runtime efficiency of learning procedure of SVR–ACO.

The ACO algorithm is a kind of algorithm inspired by real ants. This algorithm was first

proposed by Dorigo and his colleagues as a novel nature-inspired method for the solution of Combinatorial Optimization (CO) problems in the early 1990s [23]. From then on, researchers have successfully applied the ACO to many optimization problems such as continuous optimization problems [24], global optimum function [25], feature selection [26]. The principle of the method is based on the way ants search for food and find their way back to the nest. Ants can find the shortest path to food by laying a pheromone (chemical) trail as they walk. Other ants follow the pheromone trail to food. Ants that happen to pick the shorter path will create a strong trail of pheromone faster than the ones choosing a longer path. Since stronger pheromone attracts ants better, more and more ants choose the shorter path until eventually all ants have found the shortest path [27].

In the ACO, artificial ants find solutions starting from a start node and moving to feasible neighbor nodes in the process of building the solutions. Each ant builds a tour by frequently applying a stochastic greedy rule, which is called the state transition rule;

$$P(r,u) = \begin{cases} \arg \max_{u \in J(r)} \{ [\tau(r,u)]^\alpha \cdot [\eta(r,u)]^\beta \}, & \text{if } q \leq q_0 \\ S & , \text{ otherwise} \end{cases} \quad (10)$$

where (r,u) represents an edge between point r and u , and $\tau(r,u)$ stands for the pheromone on edge (r,u) . $\eta(r,u)$ is the desirability of edge (r,u) , which is usually defined as the inverse of the length of edge (r,u) . b is the parameter controlling the relative importance of the desirability, q_0 is a user-defined parameter with $0 \leq q_0 \leq 1$, q is a random number uniformly distributed in $[0, 1]$. $J(r)$ is the set of edges available at decision point r . S is a random variable selected according to the probability distribution given below;

$$P(r,s) = \begin{cases} \frac{[\tau(r,u)]^\alpha \cdot [\eta(r,u)]^\beta}{\sum_{u \in J(r)} [\tau(r,u)]^\alpha \cdot [\eta(r,u)]^\beta}, & \text{if } s \in J(r) \\ 0 & , \text{ otherwise} \end{cases} \quad (11)$$

While constructing its tour, an ant will modify the amount of the pheromone on the passed edges by applying the local updating rule;

$$\tau(r,s) \leftarrow (1-\rho) \cdot \tau(r,s) + \rho \tau_0 \quad (12)$$

Where ρ is the coefficient representing pheromone evaporation, $0 < \rho < 1$.

Once all ants have arrived at their destination, the amount of pheromone on the edge is modified again by applying the global updating rule;

$$\tau(r,s) \leftarrow (1-\delta) \cdot \tau(r,s) + \Delta \tau(r,s) \quad (13)$$

where;

$$\Delta\tau(r,s) = \begin{cases} 1/L_{gb}, & \text{if } (r,s) \in \text{global best tour} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

L_{gb} is the length of the globally best tour from the beginning of the trial, the δ is the global pheromone decay parameter, $0 < \delta < 1$ and the $\Delta\tau(r,s)$ is used to increase the pheromone on the path of solution [28].

In this paper, the ant is composed of the parameters C , σ and ε . Fig. 1 shows the algorithm process of the selection of the SVR model parameters based on ACO. With respect to the performance criteria, the detailed information is given in the next section. The detailed explanation of the ACO implementation [29] was provided elsewhere. For the sake of brevity, they are omitted in this study.

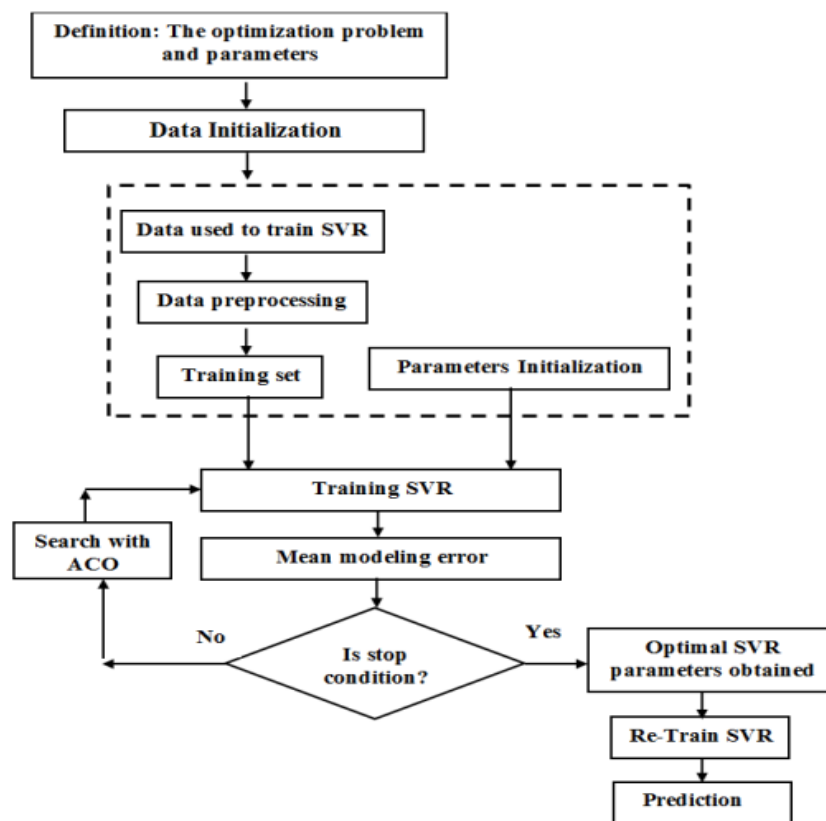


Figure 1. The process of optimizing the SVR parameters with the ACO

3. INPUT/ OUTPUT DATA SPACE

The main aim of this study is to implement the above method in the problem of the SF estimation for circular failure slope. Dataset applied in this study for determining the relationship among the set of input and output variables are gathered from open source

literature [30]. The input parameters that have been selected are related to the geometry and the geotechnical properties of each slope. The parameters utilized for circular failure (Fig. 2) were height (H), cohesion (c), slope angle (β), unit weight (γ), pore water pressure (r_u), and angle of internal friction (ϕ). Partial dataset used for training and testing model are presents in Table 1. Also, Table 2 shows statistical description of datasets used in this study.

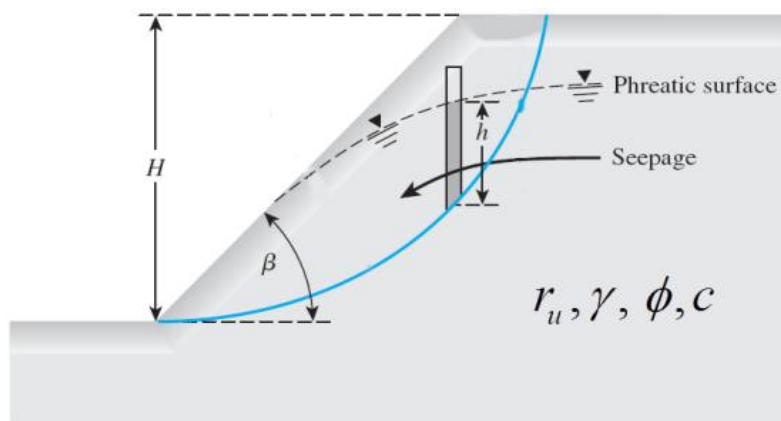


Figure 2. The parameters utilized for circular failure

Table 1: Partial dataset used for training and testing model [30]

Case No.	Input parameters					Output parameter	
	γ (KN/m ³)	C (KPa)	ρ ($^{\circ}$)	β ($^{\circ}$)	H (m)	r_u	SF
1	18.68	26.34	15	35	8.23	0	1.11
2	16.5	11.49	0	30	3.66	0	1
3	18.84	14.36	25	20	30.5	0	1.875
4	18.84	57.46	20	20	30.5	0	2.045
5	28.44	29.42	35	35	100	0	1.78
6	28.44	39.23	38	35	100	0	1.99
7	20.6	16.28	26.5	30	40	0	1.25
8	14.8	0	17	20	50	0	1.13
9	14	11.97	26	30	88	0	1.02
10	25	120	45	53	120	0	1.3

Table 2: Statistical description of dataset utilized in this study

Parameter	Min	Max	Average
γ (KN/m ³)	12	28.44	19.71
C (KPa)	0	150.05	20.66
ρ ($^{\circ}$)	0	45	27.46
β ($^{\circ}$)	16	53	33.02
H (m)	3.66	214	43.91
r_u	0	0.5	0.1746
SF	0.62	2.05	1.2446

4. ESTIMATING THE SAFETY FACTOR FOR CIRCULAR FAILURE SLOPE

In this paper, a hybrid SVR with ACO was proposed to estimate the SF, using MATLAB environment. In SVR-ACO model, height (H), cohesion (c), slope angle (β), unit weight (γ), pore water pressure (r_u), and angle of internal friction (φ) were defined as input parameters and the SF as output.

Furthermore, as shown in section 2.1, the generalization ability of SVR is highly dependent upon its learning parameters, i.e., $\{C, \sigma, \varepsilon\}$. Consequently, the ACO was used to manipulate these parameters and to form hybrid SVR- ACO. 10-fold cross-validation performance measure was applied to training dataset along with SVR- ACO to achieve reliable results. Related to the purpose, the ACO algorithm with number of ants=80, evaporation coefficient (ρ)=0.5, pheromone intensity=100, the initial value of the pheromone trail (τ)=100 and the algorithm was executed for 100 iterations for selecting optimal parameters. The adjusted parameters $\{C, \sigma, \varepsilon\}$ with maximal accuracy are selected as the most appropriate parameters. Then, the optimal parameters are used to train the SVR model. The optimal parameters of the SVR estimated by the ACO are presented in Table 3.

Table 3: The optimal parameters of the SVR estimated by the ACO

	Optimal value of σ parameter	Optimal value of C parameter	Optimal value of ε parameter
SVR-ACO model	0.721	1777	0.030

A comparison between estimated values of SF by the SVR-ACO model and measured values for 46 data sets at training and testing phases is shown in Figs. 3 and 4. It should be noted that the predicted and measured SF (Figs. 3 and 4) represent normalized values that was calculated using the following equation:

$$x_M = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (15)$$

Where x is the original value of the dataset, x_M is the mapped value, and x_{\max} (x_{\min}) denotes the maximum (minimum) raw input values, respectively. As shown in Figs. 3 and 4, the results of the SVR-ACO modeling compared with actual data show a good precision of the SVR-ACO model (see Table 5). Also, performance prediction of the predictive model proposed was evaluated, using Coefficient of Determination (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and variance account for (VAF) (Table 4) where, N is the number of samples, var denotes the variance, y and y' are the measured and predicted values, respectively.

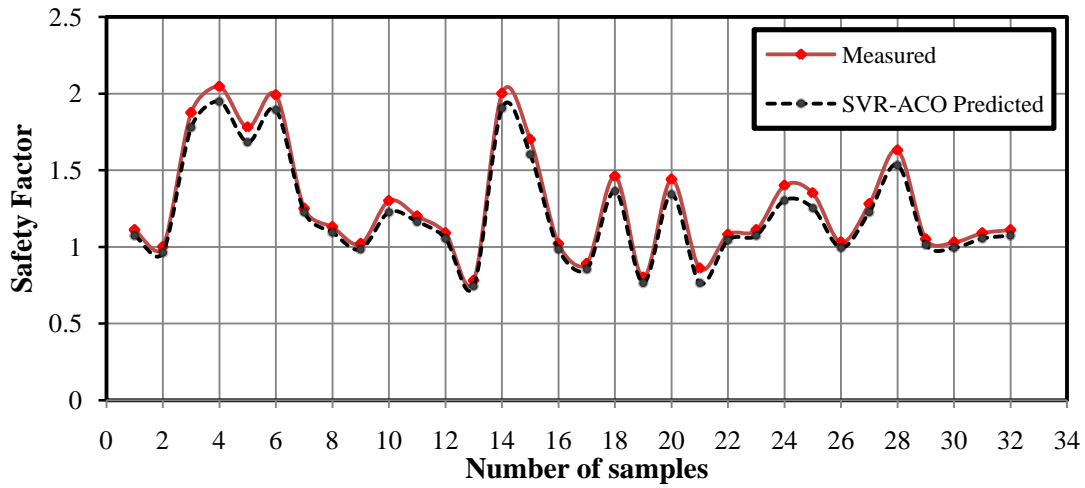


Figure 3. Comparison between measured and estimated safety factors for training data points

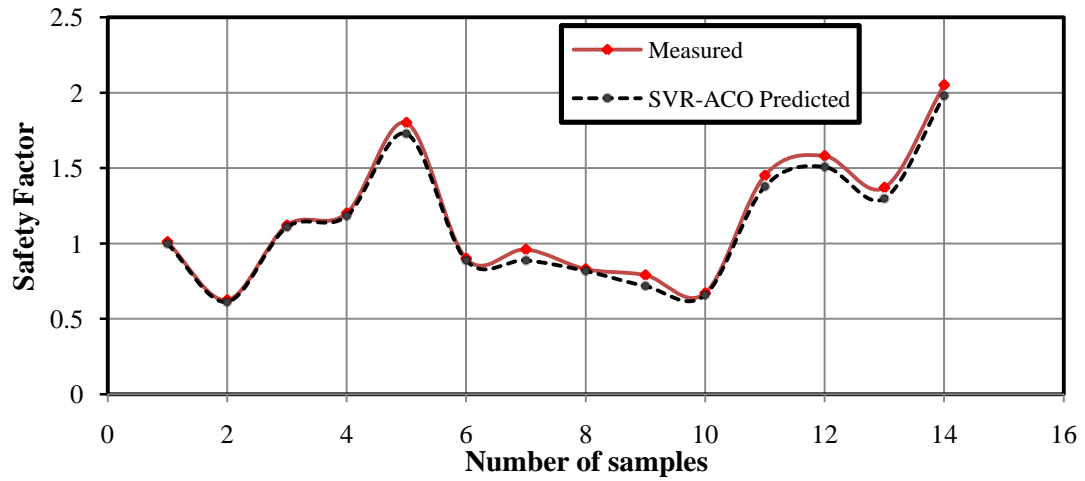


Figure 4. Comparison between measured and estimated safety factors for testing data points

Table 4: Statistical indicators

Statistical indicator	Equation
Mean squared error	$MSE = \frac{1}{N} \sum_{i=1}^N (y - y')^2$
Root mean square error	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - y')^2}$
Mean absolute percentage error	$MAPE = \frac{1}{N} \sum_{i=1}^N \left \frac{y - y'}{y} \right \times 100$
Variance account for	$VAF = \left(1 - \frac{\text{var}(y - y')}{\text{var}(y)} \right)$

Performance analysis of the SVR-ACO model for estimating SF is shown in Table 5. The performance indices obtained in Table 5 indicate the high performance of the SVR-ACO model that can be used successfully to the estimation of the SF for circular failure slope.

Table 5: Performance of the model for estimating the safety factor

Description	R^2	MSE	RMSE	VAF	MAPE
Training	0.9969	0.0025	0.049	99.35	2.98
Testing	0.9967	0.0027	0.052	99.49	3.51

Furthermore, the correlation between measured and estimated values of safety factor for training and testing phases are shown in Figs. 5 and 6.

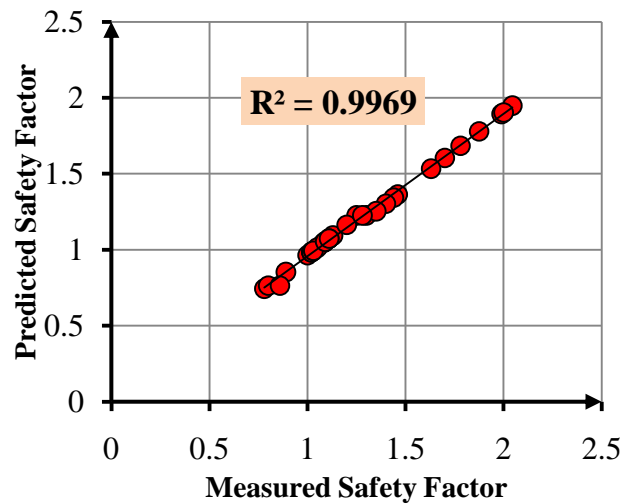


Figure 5. Correlation between measured and estimated values of safety factor for training data points

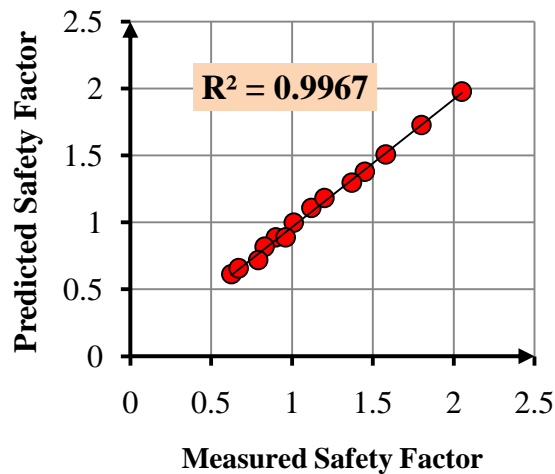


Figure 6. Correlation between measured and estimated values of safety factors for testing data points

Eventually, relative error (error percentage) for data point (training and testing samples) is assessed and revealed in Fig. 7. Relative error for most data points is located in range of $[-7\% 0\%]$, which is an acceptable value.

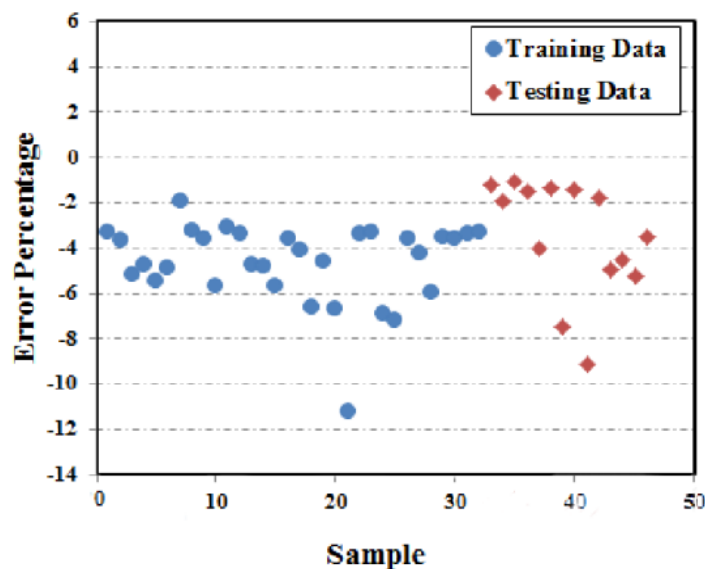


Figure 7. Relative error (error percentage) of SVR-ACO model in estimating the safety factor

5. CONCLUSION

The slope stability analysis is usually performed by engineers and researchers to estimate the stability of retaining walls, river training works, excavations, road embankments and embankment dams. In this paper, a new approach namely support vector regression optimized by ACO is proposed for estimating the safety factor for circular failure slope. In our methodology, ACO is applied as optimization tool for determining the optimal value of user defined parameters existing in formulation of SVR. The optimization implementation increases the performance of SVR model. Moreover, this method requires less time for setting optimal value in comparison to grid search, genetic algorithm and pattern search which are usually used for finding these values. Examination of the error analysis shows that the presented strategy produces results with satisfactory accuracy. This study indicates that SVR combined with ACO can be applied as a powerful tool for modeling of the non-linear regression problems involved in civil and mining engineering.

REFERENCES

1. Cheng Y, Lansivaara T, Wei W. Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods, *Comput Geotech* **34**(2007) 137-50.

2. Griffiths D, Fenton GA. Probabilistic slope stability analysis by finite elements, *J Geotech Geoenvironm Eng* **130**(2004) 507-18.
3. Martel SJ, Muller JR. A two-dimensional boundary element method for calculating elastic gravitational stresses in slopes, *Pure Appl Geophysic* **157**(2000) 989-1007.
4. Duncan JM. Soil slope stability analysis. *Landslides, investigation and mitigation US National Research Council Transportation Research Board Special Report*, **247**(1996) 337-71.
5. Das SK, Biswal RK, Sivakugan N, Das B. Classification of slopes and prediction of factor of safety using differential evolution neural networks, *Environm Earth Sci* **64**(2011) 201-10.
6. Mohamed T, Kasa A, Taha MR. Fuzzy logic system for slope stability prediction, *Int J Adv Sci, Eng Inform Technol* **2**(2012) 38-42.
7. Zolfaghari AR, Heath AC, McCombie PF. Simple genetic algorithm search for critical non-circular failure surface in slope stability analysis, *Comput Geotech* **32**(2005) 139-52.
8. Kahatadeniya KS, Nanakorn P, Neaupane KM. Determination of the critical failure surface for slope stability analysis using ant colony optimization, *Eng Geology* **108**(2009) 133-41.
9. Cheng Y, Li L, Chi SC, Wei W. Particle swarm optimization algorithm for the location of the critical non-circular failure surface in two-dimensional slope stability analysis, *Comput Geotech* **34**(2007) 92-103.
10. Vapnik V. *The Nature of Statistical Learning Theory*, Springer, 1999.
11. Üstün B, Melssen W, Oudenhuijzen M, Buydens L. Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization, *Analyt Chimica Acta*, **544**(2005) 292-305.
12. Holland JH. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press, 1975.
13. Kaveh A, Talatahari S. A novel heuristic optimization method: charged system search, *Acta Mech* **213**(2010) 267-89.
14. Kaveh A, Khayatazad M. A new meta-heuristic method: ray optimization, *Comput Struct* **112**(2012) 283-94.
15. Kaveh A, Zolghadr A. Democratic PSO for truss layout and size optimization with frequency constraints, *Comput Struct* **130**(2014) 10-21.
16. Kaveh A, Mahdavi VR. Colliding bodies optimization: a novel meta-heuristic method, *Comput Struct* **139**(2014) 18-27.
17. Vapnik V, Golowich SE, Smola A. Support vector method for function approximation, regression estimation, and signal processing, *Adv Neural Inform Proc Sys* (1997) 281-7.
18. Smola AJ, Schölkopf B. A tutorial on support vector regression, *Statist Comput* **14**(2004)199-222.
19. Geem ZW. *Music-inspired Harmony Search Algorithm: Theory and Applications*, Springer Verlag, 2009.
20. Hong WC, Dong Y, Chen LY, Wei SY. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Appl Soft Comput* **11**(2011) 1881-90.

21. Tsai HH, Jhuang YJ, Lai YS. An SVD-based image watermarking in wavelet domain using SVR and PSO, *Appl Soft Comput* **12**(2012) 2442-53.
22. Fattahi H, Gholami A, Amiribakhtiar MS, Moradi S. Estimation of asphaltene precipitation from titration data: a hybrid support vector regression with harmony search, *Neural Comput Applic* **26**(2014) 789-98.
23. Dorigo M, Blum C. Ant colony optimization theory: A survey, *Theoret Comput Sci* **344**(2005) 243-78.
24. Socha K, Dorigo M. Ant colony optimization for continuous domains, *Europ J Oper Res* **185**(2008) 1155-73.
25. Toksari MD. Ant colony optimization for finding the global minimum, *Appl Math Comput* **176**(2006) 308-16.
26. Kanan HR, Faez K, Taheri SM. Feature selection using ant colony optimization (ACO): a new method and comparative study in the application of face recognition system, *Advances in Data Mining Theoretical Aspects and Applications*, Springer, 2007, pp. 63-76.
27. Xu D, Fu L. Ant Colony Optimization and its Applications, Intelligent Systems and Applications, 2009 ISA 2009 International Workshop on: IEEE, 2009. pp. 1-4.
28. Zeng DH, Liu Y, Jiang LB, Li L, Xu G. A New Approach to Cutting Temperature Prediction Using Support Vector Regression and Ant Colony Optimization, *Advanced Engineering Forum*, Trans Tech Publ, 2012. pp. 145-52.
29. Dorigo M, Birattari M. Ant colony optimization, *Encyclopedia of Machine Learning*, Springer, 2010, pp. 36-39.
30. Sah N, Sheorey P, Upadhyaya L. Maximum likelihood estimation of slope stability, *Int J Rock Mech Mining Sci Geomech Abstracts*, **31**(1994) 47-53.