Journal of Soft Computing in Civil Engineering 4-3 (2020) 01-16

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Prediction of Compression Index of Saturated Clays Using Robust Optimization Model

Y. Erzin[1*], H. MolaAbasi[2], A. Kordnaeij[3], S. Erzin[4]

1. Professor, Department of Civil Engineering, Faculty of Engineering, Manisa Celal Bayar University, Manisa, Turkey
2. Assistant Professor, Department of Civil Engineering, Gonbad Kavous University, Gonbad, Iran
3. Department of Civil Engineering, Imam Khomeini International University, Iran
4. Instructor, Department of Physics, Faculty of Science, Dokuz Eylul University, Izmir, Turkey

Corresponding author: *yusuf.erzin@cbu.edu.tr*

https://doi.org/10.22115/SCCE.2020.233075.1226

ARTICLE INFO

Article history:
Received: 28 May 2020
Revised: 14 July 2020
Accepted: 23 July 2020

Keywords:
Compression index;
Saturated clays;
Consolidation settlement;
Robust optimization;
Second order cone.

ABSTRACT

Compression index (Cc) of normally consolidated (NC) clays determined by the oedometer experiments is utilized for calculating the consolidation settlement of shallow foundations. The determination of the Cc from the tests takes much more time and so empirical correlations based on clay properties can be a suitable alternative for the prediction of settlement. However, uncertainty in the measurements of input parameters has always been a major concern. Robust optimization is very popular due to its computational tractability for many classes of uncertainty sets and problem types. Therefore, in this research, an innovative method based on robust optimization has been used to investigate the effect of such uncertainties. To achieve these, the results of 433 oedometer tests taken from geotechnical investigation locations in Mazandaran province of Iran have been used. Based on Frobenius norm of the data points, uncertainty definition is presented and examined against the correlation coefficients for several empirical models for predicting Cc value and thus optimum values are determined. The results in compare with previous models indicate the robust method is a better pattern recognition tool for datasets with degrees of uncertainty. The variation of the Cc values with soil properties, namely, water content (ω_n), initial void ratio (e_o), and liquid limit (LL), by considering different value of uncertainties (0, 5 and 10%) was considered and indicated that the effect of e_o is more than other two physical parameters (ω_n and LL). The best model performance during in deterministic valuation and considering uncertainty is further shown.

How to cite this article: Erzin Y, MolaAbasi H, Kordnaeij A, Erzin S. Prediction of compression index of saturated clays using robust optimization model. J Soft Comput Civ Eng 2020;4(3):01–16. https://doi.org/10.22115/SCCE.2020.233075.1226.

1. Introduction

Stress increases in soil layers owing to structures are accompanied by settlements. Settlement-based design is an essential and important criterion for shallow foundations [1,2] The volume change characteristics of clays can be defined by oedometer tests using Terzaghi's one-dimensional consolidation theory [3]. Compression index (C_c), a compressibility parameter, presents the slope of the curve of the void ratio (e) versus the logarithm of the effective pressure and is obtained by these tests. Additionally, the C_c value, the main indicator of the compressibility of soil, is generally utilized for calculating the settlement of soil layers [1,4]. C_c is an important parameter used for the estimation of the primary consolidation settlement of clays [5]. High value of C_c indicates large settlement. The calculation of C_c value from oedemeter tests is relatively time consuming and expensive [6–9]. Performing oedemeter tests also requires precision, precautions and expertise and so is cumbersome [5]. Even a small disturbance can cause overestimation or underestimation of the C_c value and so the settlement [5]. Thus, if correlations between the C_c value and easily determined soil properties are developed, the determination of the C_c value may save a lot of time and laboratory expenditure [5]. Therefore, many investigators have performed various investigations to predict the C_c value from the soil properties determined more easily. They suggested single or multi-variable equations for the prediction of the C_c value, which is presented in Table 1. In this table, LL is the liquid limit; ω_n is natural water content; e_o is the in-situ void ratio; G_s is the specific gravity, and γ_d is the dry unit weight. All estimations based on the empirical methods need assessment of input parameters which incline to uncertainties and inaccuracies. Thus, a formula that is capable of resolving the uncertainties and inaccuracies in input parameters requires to be suggested.

Data uncertainty is observed generally in many optimization problems [10]. Measurement errors can lead to uncertainty in the parameters of an optimization model. Recently, robust optimization (RO) approach, a methodology to solve mathematical optimization problems with uncertain data, is developed to address such uncertainties [11] The objective of the RO is to find solutions that are immune to all perturbations of the data in a so-called uncertainty set. RO is popular because it is a computationally tractable methodology and has a wide range of applications in practice [12].

Keeping this in view, a RO approach was proposed in this paper in order to quantify the effect of uncertainties on evaluation of the input parameters. The RO model is the robust counterpart of the least square model and is a second order cone program (*SOCP*) in which, possible uncertainties can be reasonably adjusted. SOCP has widely been used in operation research [13]. In geotechnical engineering field, Kalantary et al. [14] and MolaAbasi et al. [15] developed robust optimization model to predict the liquefaction induced lateral displacement and shear wave velocity, respectively. It is shown that if uncertainty (%) is set to zero, the method reduces to ordinary regression method. Also, Kalantary et al. [14] have shown that logarithmic lateral spreading correlations perform better in deterministic valuation, whereas by considering uncertainty, they give similar degrees of accuracy to new linear model. Thus, this method can be used in empirical correlation of C_c as a step forward in comparison with statistical approaches that considers the variation of uncertainty.

Table 1
Some widely used compression index equations.

	Independent variable	Equation	References	Equation no.
Single variable equations	$\boldsymbol{\omega_n}$	$C_c = 0.01\omega_n - 0.05$	Azzouz et al. [16]	(1)
		$C_c = 0.01\omega_n$	Koppula [17]	(2)
		$C_c = 0.01\omega_n - 0.075$	Herrero [18]	(3)
		$C_c = 0.013\omega_n - 0.115$	Park and Lee [19]	(4)
	$\boldsymbol{e_0}$	$C_c = 0.49e_0 - 0.11$	Park and Lee [19]	(5)
		$C_c = 0.4(e_0 - 0.25)$	Azzouz et al. [16]	(6)
		$C_c = 0.287e_0 - 0.015$	Ahadiyan et al. [20]	(7)
		$C_c = 1.02 - 0.95e_0$	Gunduz and Arman [21]	(8)
		$C_c = 0.287e_0 - 0.015$	Farzi [22]	(9)
	$\boldsymbol{LL}$	$C_c = 0.006(LL - 9)$	Azzouz et al. [16]	(10)
		$C_c = (LL - 13)/109$	Mayne [23]	(11)
		$C_c = 0.009(LL - 10)$	Terzaghi and Peck [24]	(12)
		$C_c = 0.014LL - 0.168$	Park and Lee [19]	(13)
	$\boldsymbol{PI}$	$C_c = 0.0082PI + 0.0475$	Jain et al. [25]	(14)
Multi variable equations	$\boldsymbol{LL, G_s}$	$C_c = 0.2926\left(\frac{LL}{100}\right).G_s$	Park and Lee [19]	(15)
	$\boldsymbol{\omega_n, LL}$	$C_c = 0.009\omega_n + 0.005LL$	Koppula [17]	(16)
		$\mathrm{C_c = 0.009\omega_n + 0.002LL - 0.1}$	Azzouz et al. [16]	(17)
	$\boldsymbol{e_0, \omega_n}$	$C_c = 0.4(e_0 + 0.001\omega_n - 0.25)$	Azzouz et al. [16]	(18)
	$\boldsymbol{e_0, LL}$	$C_c = -0.156 + 0.411e_0 + 0.00058LL$	Al-Khafaji and Andersland [26]	(19)
		$C_c = -0.023 + 0.271e_0 + 0.001L$	Ahadiyan et al. [20]	(20)
	$\boldsymbol{e_0, \omega_n, LL}$	$C_c = 0.37(e_0 + 0.003LL + 0.0004\omega_n - 0.34)$	Azzouz et al. [16]	(21)
		$C_c = -0.404 + 0.341e_0 + 0.006\omega_n + 0.004LL$	Yoon and Kim [27]	(22)
	$\boldsymbol{G_s, e_0}$	$C_c = 0.141G_s^{1.2}[(1 + e_0)/G_s]^{2.38}$	Herrero [18]	(23)
	$\boldsymbol{\omega_n, LL, e_0, \gamma_d}$	$C_c = 0.1597(\omega_n^{-0.0187})(1 + e_0)^{1.592}(LL^{-0.0638})(\gamma_d^{-0.8276})$	Ozer et al. [28]	(24)
		$C_c = 0.151 + 0.001225\omega_n + 0.193e_0 - 0.000258LL - 0.0699\gamma_d$	Ozer et al. [28]	(25)

The study presented here aims both to predict the Cc values from the easily determined soil parameters by using RO model without performing consolidation tests and to evaluate the effects of uncertainty of each parameter (wn, eo, Gs and LL) independently on the analysis outcome Cc. To achieve these, the results of 433 laboratory tests (oedometer) on clay at 115 projects in north of Iran (Mazandaran province) were used while developing RO model. The measured C_c values from the experiments were compared with the predicted C_c values with different level of uncertainties.

The rest of the paper is arranged as follows. A review of RO model is given in Section 2. While the database compilation is presented in Section 3, the development of the ten different RO models is given in Section 4. The results of ten RO models and the discussion the RO results with the previous models in the literature given in Section 5.

2. Review of robust optimization (RO) model

In mathematical optimization models, it is generally assumed that the input data is precise and the effect of uncertainties on the optimality and feasibility of the models are ignored. It is thus possible that when the data is different from the assumed nominal values, the generated optimal solution may violate critical constraints and make it poor from an objective function point of view. These explanations provoke the need for methodologies in mathematical optimization models that explain solutions immune to data uncertainty [29,30]. For instance, inaccuracies enter in field measurements of C_c in the case histories alike all other natural phenomenon measurements. Such inaccuracies exist in other influencing parameters and can cause deviation. If such deviations are given as boundaries of the central point of the data (Figure 1) and the true data point could exist at any point within this boundary, the robust optimization (RO) investigates to minimize the maximum error with any particular level of uncertainty. In the particular case of zero uncertainty, RO approach decreases to normal regression analysis.

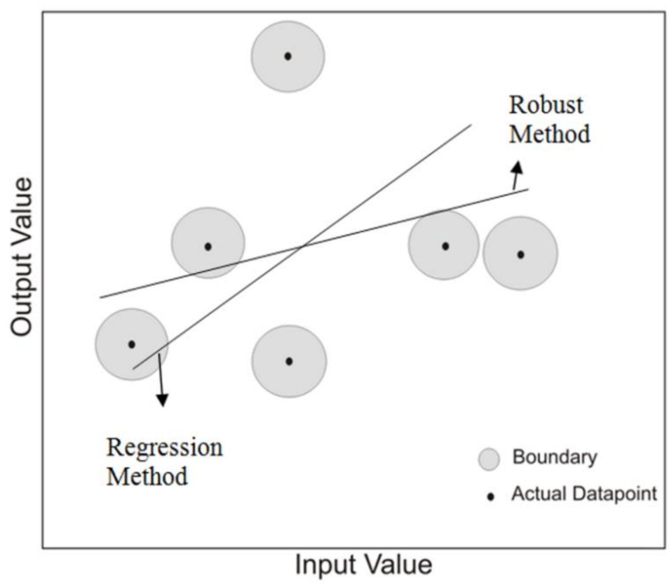

Fig. 1. Sample view of uncertainties.

In the past decade, there were noticeable developments in the theory of robust convex optimization. In this paper, a RO model for the least squares method was considered. It was firstly assumed that the level of uncertainty of a database is unknown and equal to ρ. Then, the RO model takes into account this level of uncertainty in the database and minimizes the worst case residual i.e.

$$\min_x \max_{\|[E,r]\| \le \rho} \| (A+E)x-(b+r) \| \tag{26}$$

Where E and r are uncertainties in A and b, respectively, and the matrix norm, $\| \quad \|$ is the Frobenius norm. Definitely, Eq. (26) cannot be solved by utilizing classical optimization algorithms. However, it can be rewritten in the following second order cone program (SOCP) form [31] (Eq. 26) and solved using an effective software such as SeDuMi [26], which is an interior point based software for solving SOCP and semi definite optimization.

$$\min(t + \rho s) \tag{27}$$

with the constrains, $\|Ax - b\| \leq t$, and $\sqrt{1 + \|x\|^2} \leq s$

First we rewrite problem (26) in the dual form of SeDuMi's input format namely

$$\max b^T y$$

$$c - A^T y \ \in K \tag{28}$$

where

$$c=\begin{pmatrix} 0 \\ -b \\ 0 \\ 1 \\ 0_{n\times 1} \end{pmatrix},\ A^t = \begin{pmatrix} -1 & 0 & 0_{1\times n} \\ 0_{m\times 1} & 0_{m\times 1} & -A \\ 0 & -1 & 0_{1\times n} \\ 0 & 0 & 0_{1\times n} \\ 0_{n\times 1} & 0_{n\times 1} & -I_{n\times n} \end{pmatrix},\ b = \begin{pmatrix} -1 \\ -\rho \\ 0_{n\times 1} \end{pmatrix},\ Y = \begin{pmatrix} t \\ s \\ x \end{pmatrix} \tag{29}$$

$$K = Q_{m+1} \times Q_{n+2} \tag{30}$$

and Q_k denotes the second order cone in R^k.

This kind of SOCP problem can be easily solved by means of SeDuMi which is one the most efficient interior-point methods based software packages for solving SOCP for four different values of uncertainty parameter. Thereafter, SeDuMi called by the following MATLAB command was used to solve it four different values of ρ in order to evaluate uncertainty:

```
[x,y]=sedumi(At,b,c,K)
```

Where At denotes transpose of the matrix A^t in Eq. (29); b and c are taken from Eq. (29); and K also is obtained from Eq. (30). The x and y outputs denote the solutions of Eq. (30) and its dual problem. Then, uncertainty is calculated from the following equation:

$$\text{Uncertanity } (\%) = \frac{\rho}{\|\text{DATA}\|_{\text{fro}}} \times 100 \tag{31}$$

where $\|\text{DATA}\|_{\text{fro}}$ is the Frobenius norm of maximum error of data matrix.

3. Database compilation

Databases have compiled the data from 433 oedometer tests for clays at 115 building sites in north of Iran (Mazandaran province), which is taken from the Geotechnical Report of Site investigations performed at Technical and Soil Laboratory of Mazandaran Province between 2000-2010 years. Considering the past investigations in the literature, in this paper, the C_c value of the clays was expected to be influences by the initial void ratio (e_o), liquid limit (LL), natural water content (ω_n), plastic index (PI), and specific gravity (G_s). The locations of collected database contain 433 data summarized in Figure 2. The samples were all collected and tested according to standard ASTM D2435-96. The database covers a wide spectrum of fine grained soils and soil parameters, including e_o, ω_n, LL, G_s, and C_c. The details of soil parameters used while developing the RO models are given in Figure 3 and in Table 2.

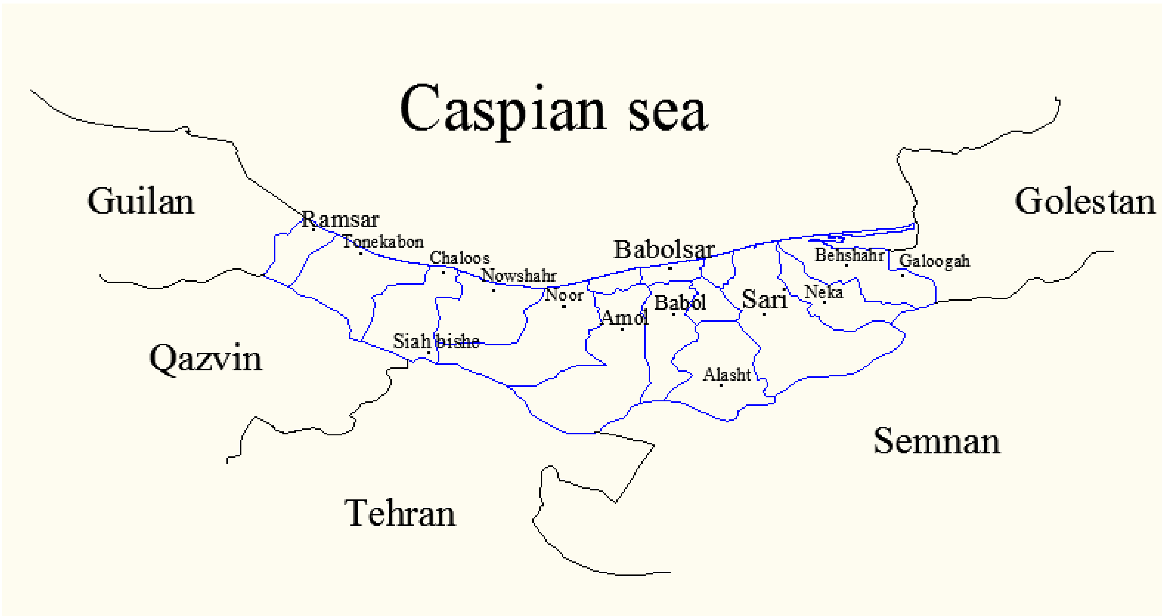

Fig. 2. Descriptive data collection location.

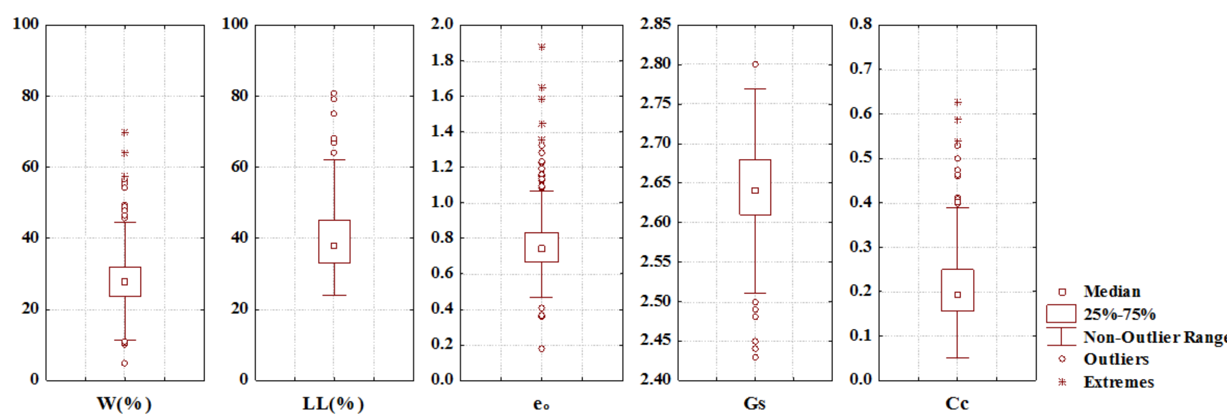

Fig. 3. Descriptive statistics of variables.

Table 2
The details of the parameters used in the RO models developed

Parameters used	Minimum	Maximum	Mean	Std. Deviation
Input parameters				
w (%)	5.00	70.00	28.6395	7.8835
γ_d (kN/m^3)	9.30	22.4	15.048	1.3942
LL (%)	24	81	39.8819	9.7299
G_s	2.43	2.80	2.6385	0.0603
e	0.18	1.88	0.7693	0.1796
Output parameter				
C_c	0.05	0.63	0.2084	0.0823

4. Development of robust optimization (RO) models

In order to investigate the effect of uncertainties in predicting the *Cc*, value ten different Robust Optimization (RO) models were developed by considering the previous studies given in Table 1. Description of the RO models used in this study is given in Table 3. Evaluating a_i coefficients in the ten different RO models given in Table 4 by means of robust optimization method constitutes the main goal of this paper.

Table 3
Description of different RO models used in this study.

Model	Model description
RO1	$C_c = a_1\omega_n + a_2$
RO2	$C_c = a_1 e_0 + a_2$
RO3	$C_c = a_1 LL + a_2$
RO4	$C_c = a_1 LL^{a_2} Gs^{a_3}$
RO5	$C_c = a_1\omega_n + a_2 LL + a_3$
RO6	$C_c = a_1 e_0 + a_2 LL + a_3$
RO7	$C_c = a_1 e_0 + a_2 LL + a_3\omega_n + a_4$
RO8	$C_c = a_1 G_s^{a_2}[(1 + e_0)/G_s]^{a_3}$
RO9	$C_c = a_1(\omega_n^{a_2})(1 + e_0)^{a_3}(LL^{a_4})(\gamma_d^{a_5})$
RO10	$C_c = a_1 + a_2\omega_n + a_3 LL + a_4 e_0 + a_5\gamma_d$

Table 4
The determined coefficients for ten different RO models.

Model	a_1	a_2	a_3	a_4	a_5
RO1	0.0931	0.1700	-	-	-
RO2	0.146	0.0904	-	-	-
RO3	0.0879	0.161	-	-	-
RO4	0.65726368	0.40297	-0.40743	-	-
RO5	0.0848	0.0805	0.143	-	-
RO6	0.133	0.0506	0.0794	-	-
RO7	0.125	0.0479	0.0479	0.0744	-
RO8	0.621015	-0.46271	0.21201	-	-
RO9	0.714573	0.46383	-0.16917	0.32257	-0.15777
RO10	0.051894	0.036503	0.035216	0.094971	0.034755

Therefore, these ten RO models can be formulated of the following form

$$Ax = b, A \in R^{mxn}, b \in R^{nx1} \quad (32)$$

which shows over-determined set of equations in the case of (m>n). Initially, all the RO models are examined using the compiled database and simple regression analyses, containing least square approaches. The classical method for solving least squares problems, e.g.:

$$min_{x \in R^n} \|Ax - b\| \quad (33)$$

In order to evaluate the prediction performance of the RO models developed in this study, different performance indices were used [32–34]. Four performance indices, namely, absolute fraction of variance (R^2), given by Eq. (34), root mean squared error (*RMSE*), given by Eq. (35), mean absolute percent error (*MAPE*), given by Eq. (36), and mean absolute deviation (*MAD*), given by Eq.(37), were calculated for examining of the prediction capacity of each RO model.

$$R^2 = 1 - \left[\frac{\sum_1^M (C_{mi} - C_{pi})^2}{\sum_1^M (C_{mi})^2}\right] \quad (34)$$

$$RMSE = \sqrt{\frac{1}{M}\sum_1^M (C_{mi} - C_{pi})^2} \quad (35)$$

$$MAPE = \frac{\sum_1^M |C_{mi} - C_{pi}|}{\sum_1^M C_{mi}} \times 100 \quad (36)$$

$$MAD = \frac{\sum_1^M |C_{mi} - C_{pi}|}{M} \quad (37)$$

where M is the number of data; and C_{mi} and C_{pi} are observed and calculated values, respectively. The better model performance will be obtained by lower *RMSE*, *MAPE* and *MAD* values. Under ideal conditions an accurate and precise method gives R^2 of 1.0, *RMSE*, *MAPE*, and *MAD* of zero.

5. Results and discussions

As mentioned earlier, to evaluate the uncertainty, variations of the coefficients (a_1, a_2, a_3, a_4, and a_5) presented in Table 3 have been evaluated against various levels of uncertainty for ten RO models. The results are summarized in Figures 4 and 5.These figures show that by increasing the uncertainty, the variability approaches a nearly stable value as observed by Kalantary et al. [14]. It must be pointed out at this stage that if uncertainty (%) is set to zero, the method reduces to an ordinary multiple linear regression (MLR) technique. The coefficients determined from Figs. 4 and 5 are given in Table 4 for ten RO models. As known, the physical parameters of soils have an important influence on soil's compressibility parameters. The variation of the C_c values with soil properties (w_n, e_o, and LL) by considering different value of uncertainties 0, 5 and 10% is also given in Figure 6. When compared a_1 coefficients given in Table 4 for the RO1, RO2 and

RO3 models developed using single variable, the RO3 model yielded the highest a_1 coefficient. This result indicates that the effect of e_0 is more than other two physical parameters (ω_n and LL). The effect of e_0 was also observed in the models RO6, RO7 and R10 developed using two, three and four variables, respectively.

In order to determine the precision of each RO model, the performance indices (R^2, $RMSE$, $MAPE$, and MAD) calculated by using Eqs. (34) to (37), respectively, were calculated for each RO model with different value of uncertainties, which is shown in Figure 7. As mentioned earlier, if uncertainty (%) is set to zero, the method reduces to an ordinary multiple linear regressions (MLR) technique. The calculated performance indices (R^2, $RMSE$, $MAPE$, and MAD) for the uncertainty of 100% are also given in Table 5 for ten RO models. The models RO2, RO6, and RO7 yielded lower $RMSE$, $MAPE$, and MAD values and higher R^2 values than the other RO models (Table 5).

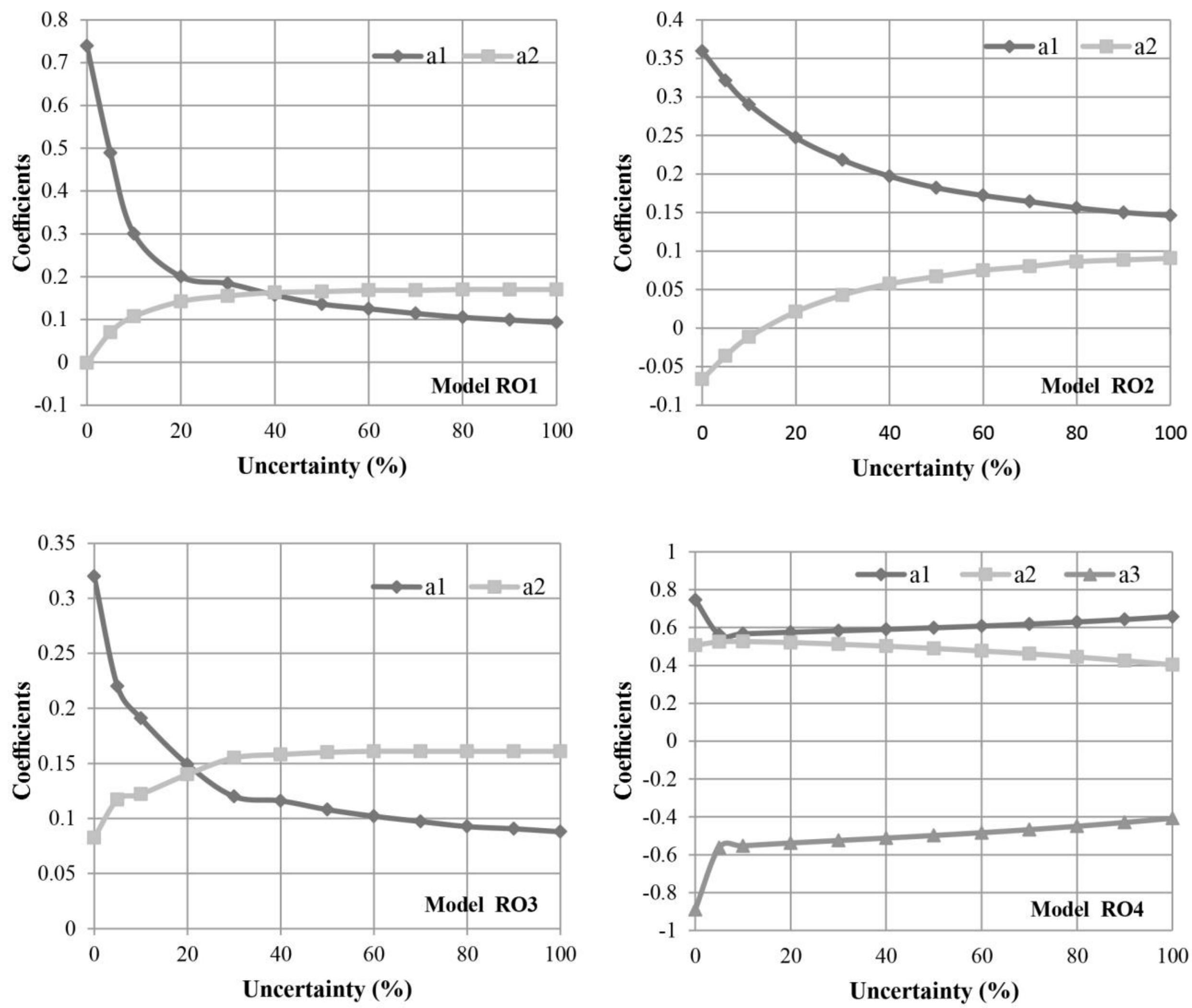

Fig. 4. The variation of constant coefficients (a_i) versus uncertainties for RO1, RO2, RO3, and RO4 models.

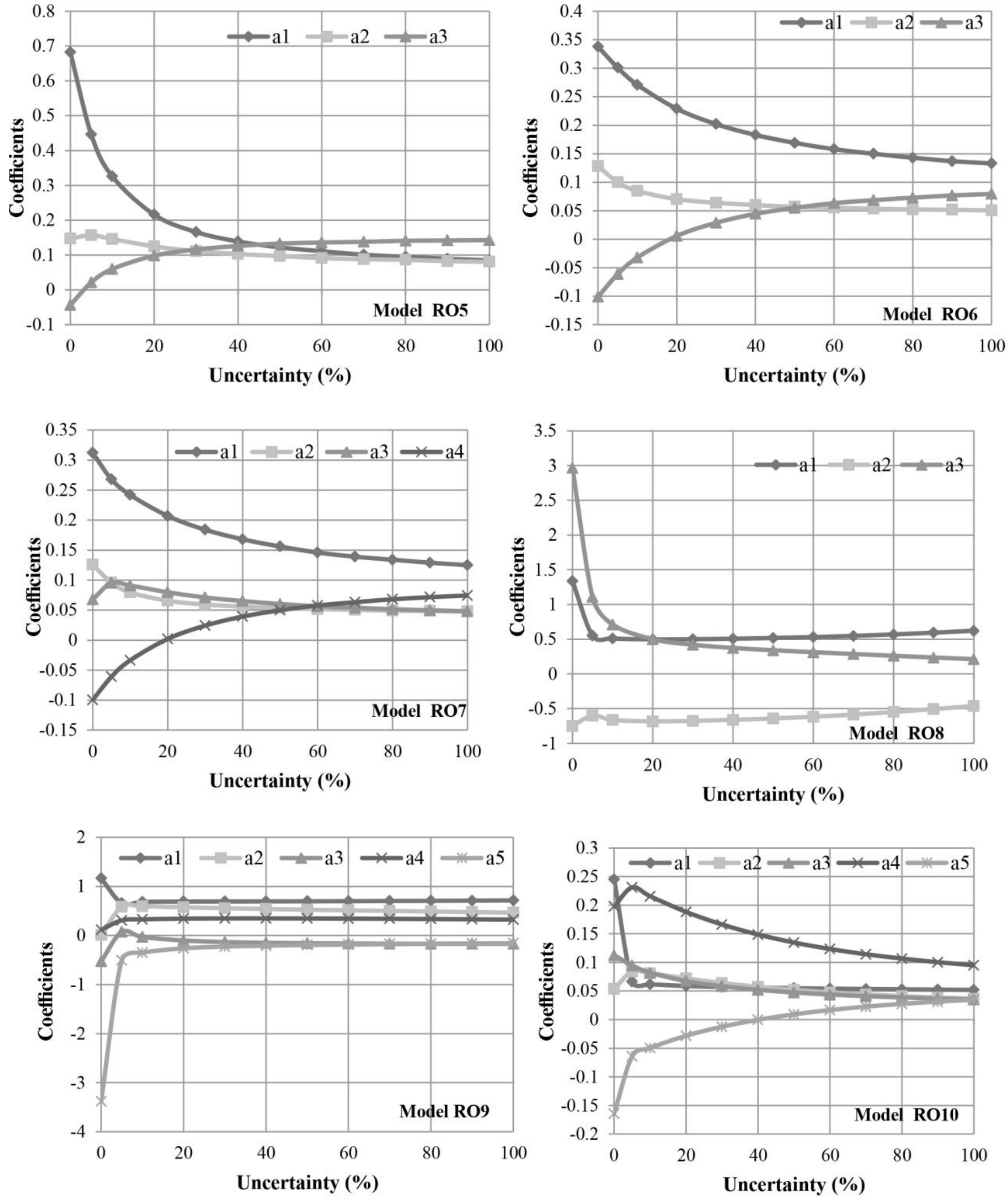

Fig. 5. The variation of constant coefficients (a_i) versus uncertainties for RO5, RO6, RO7, RO8, RO9, and RO10 models.

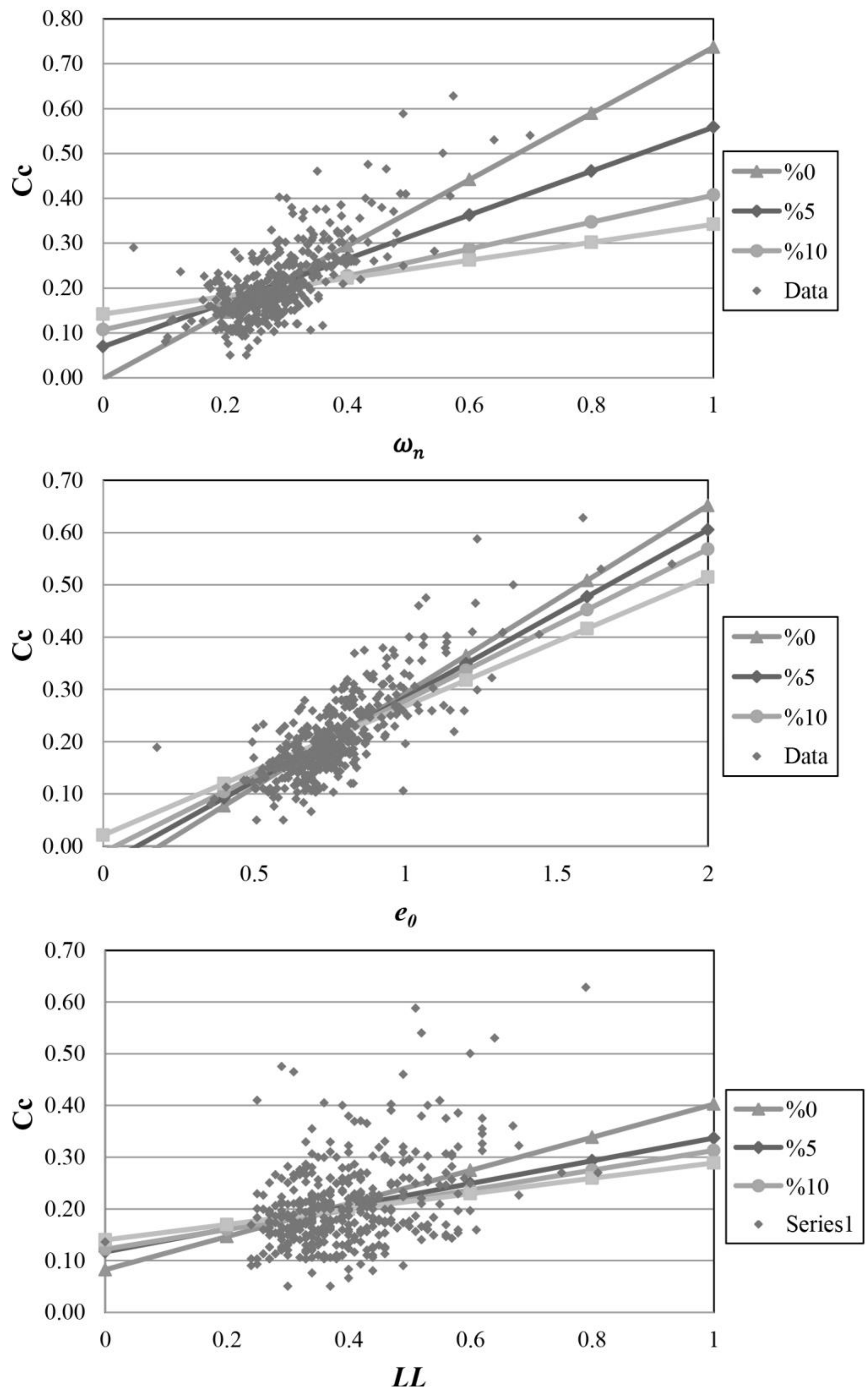

Fig. 6. The variation of C_c outputs with soil properties (w_n, e_o, and LL) by considering different value of uncertainties of 0, 5 and 20%.

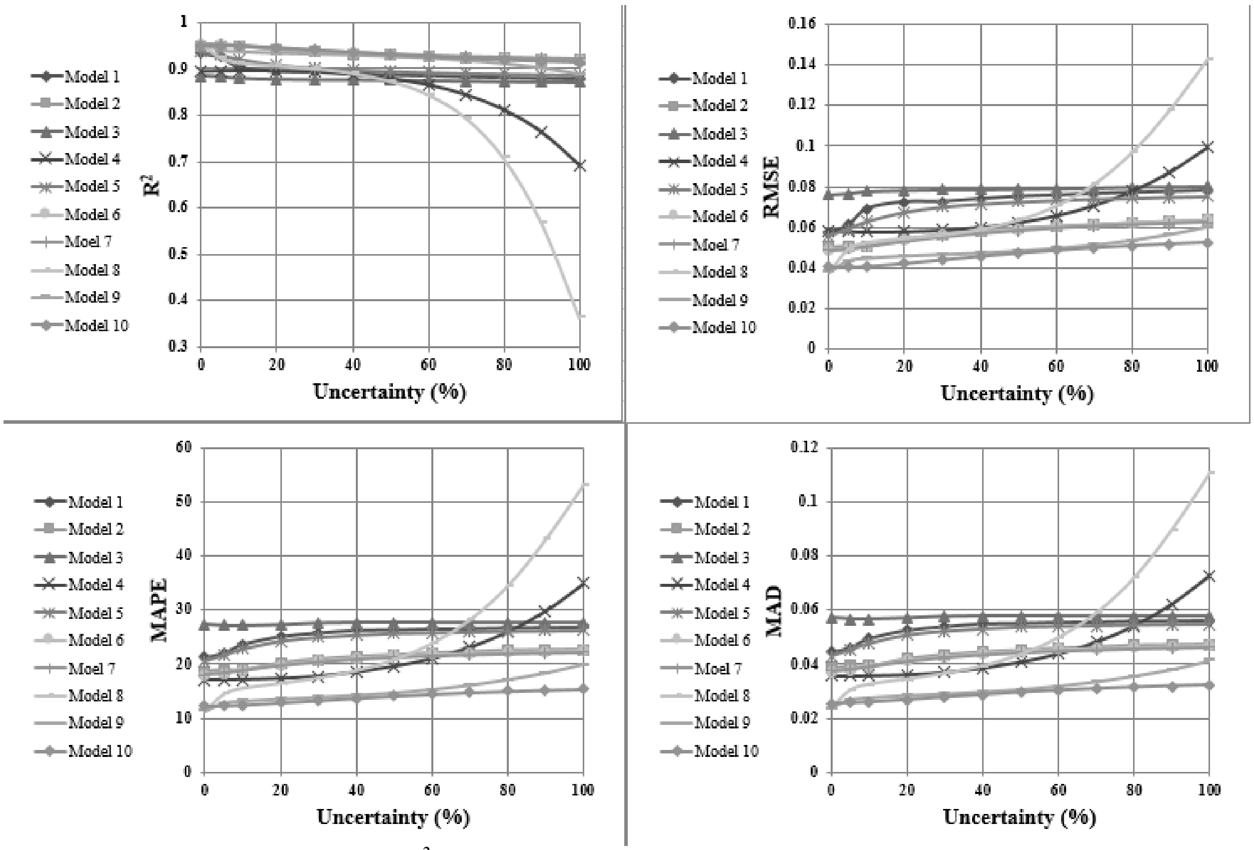

Fig. 7. The performance indices (R^2, *RMSE*, *MAPE* and *MAD*) versus uncertainties in predicting C_c value for different ten RO models.

Table 5
The performance indices calculated for all RO models.

Model	R^2	*RMSE*	*MAPE*	*MAD*
RO1	0.8788	0.0780	26.9115	0.0561
RO2	**0.9191**	**0.0637**	**22.6891**	**0.0473**
RO3	0.8717	0.0803	27.7060	0.0577
RO4	0.7019	0.1223	51.9701	0.1083
RO5	0.8865	0.0755	26.3032	0.0548
RO6	**0.9194**	**0.0636**	**22.5631**	**0.0470**
RO7	**0.9226**	**0.0623**	**22.2321**	**0.0463**
RO8	0.4031	0.1731	77.4982	0.1615
RO9	0.8900	0.0743	29.8022	0.0621
RO10	0.9019	0.0702	24.7954	0.0517

In order to compare the performance of RO models with the correlations suggested by previous researchers and given in Table 1, the performance indices (R^2, *RMSE*, *MAPE*, and *MAD*) were calculated for these correlations by using the test data in this study which is depicted in Table 5. When compared the performance indices of the correlations developed by previous researchers using single variable of ω_n (i.e. Eqs. 1 to 4) with those of the RO1 model developed by using the same variable, the correlations developed by Azzouz et al. [16] and Herrero et al. [18] represented by Eqs. (2) and (4), respectively, yielded higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values. When compared the performance indices of the correlations developed by previous researchers using single variable of e_0 (i.e. Eqs. 5 to 8) with those of the RO2 model developed by using the same variable, higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values were obtained for the correlations developed by Azzouz et al. [16] and Ahadiyan et al. [20], represented by Eqs. (6) and (7), respectively, and the RO2 model. When the performance indices of the correlations developed by previous researchers using single variable of *LL* (i.e. Eqs. 10 to 13) were compared with those of the RO3 model developed by using the same variable, the correlation developed by Azzouz et al. [16], represented by Eq. (10), and the RO3 model produced higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values. It can be noted from Tables 5 and 6 that the highest performance was obtained for the correlation suggested by Azzouz et al. [16], (Eq. 6), by using single variable of e_0, which indicates the higher effect of e_0 on the compression index (C_c) value as mentioned earlier.

Table 6
The performance indices calculated for the best used compression index equations.

	Independent variable	References	Equation no	R^2	*RMSE*	*MAPE*	*MAD*
Single variable equations	$\boldsymbol{\omega_n}$	Azzouz et al. [16]	(1)	0.9105	0.0670	26.0071	0.0542
		Herrero [18]	(3)	0.9259	0.0610	22.6792	0.0473
	$\boldsymbol{e_0}$	Azzouz et al. [16]	(6)	0.9490	0.0506	18.7712	0.0391
		Ahadiyan et al. [20]	(7)	0.9461	0.0136	19.0337	0.0397
	$\boldsymbol{LL}$	Azzouz et al. [16]	(10)	0.8577	0.0845	30.8510	0.0643
Multi variable equations	$\boldsymbol{\omega_n, LL}$	Azzouz et al. [16]	(17)	0.9141	0.0657	25.6741	0.0535
	$\boldsymbol{e_0, \omega_n}$	Azzouz et al. [16]	(18)	0.9460	0.0521	19.5910	0.0408
	$\boldsymbol{e_0, LL}$	Al-Khafaji and Andersland [26]	(19)	0.9060	0.0687	26.1167	0.0544
		Ahadiyan et al. [20]	(20)	0.9344	0.0574	19.9235	0.0415
	$\boldsymbol{e_0, \omega_n, LL}$	Azzouz et al. [16]	(21)	0.9520	0.0491	18.0270	0.0376
	$\boldsymbol{G_s, e_0}$	Herrero [18]	(23)	0.9274	0.0604	21.0635	0.0439
	$\boldsymbol{\omega_n, LL, e_0, \gamma_d}$	Ozer et al. [28]	(24)	0.9482	0.0510	19.3152	0.0402
		Ozer et al. [28]	(25)	0.9441	0.0530	20.2145	0.0421

When compared the performance indices of the correlation developed by Park and Lee [19](Eq. 4) using two variables of *LL* and G_s with those of the RO4 model developed by using the same variables, both the correlation and the RO4 model yielded lower R^2 values and higher *RMSE*, *MAPE*, and *MAD* values, which was not acceptable. When the performance indices of the correlations developed by past researchers using two variables of ω_n and *LL*, Eqs. (16) and (17), were contrasted with those of the RO5 model developed by using the same variables, the correlation developed by Azzouz et al. [16], represented by Eq. (17), yielded higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values. When compared the performance indices of the correlations developed by past researchers using two variables of e_0 and *LL*, Eqs. (19) and (20), with those of the RO6 model developed by using the same variables, the correlation developed by Ahadiyan et al. [20], represented by Eq. (20), yielded higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values. When compared the performance indices of the correlation developed by Herrero et al. [18] using two variables of e_0 and G_s, Eq. (23), with those of the RO8 model developed by using the same variables, both the correlation yielded higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values. It can be noted from Tables 5 and 6 that when compared the prediction performance of the correlations and RO models developed by using two variables, the highest performance obtained for the correlation suggested by Azzouz et al. [16], (Eq. 18), by using two variables of e_0 and ω_n, which indicates the higher effect of e_0 on the compression index (C_c) value as mentioned earlier.

When compared the performance indices of the correlations developed by past researchers using four variables of e_0, ω_n, γ_d, and LL, Eqs. (24) and (25), with those of the models RO9 and RO10 developed by using the same variables, the model RO10 and both correlations developed by Ozer et al. [28] yielded higher R^2 values and lower *RMSE*, *MAPE*, and *MAD* values.

When compared the performance of ten different Robust Optimization models, developed for the examination of the influences of uncertainty of each soil parameter independently on the analysis outcome *Cc*, the robust method is a better pattern recognition tool for datasets with degrees of uncertainty. The linear model contains *LL* (RO3) are found to be more effective than the available empirical formulas. The variation of the C_c values with soil properties (w_n, e_o, and *LL*) by considering different value of uncertainties (0, 5 and 10%) was evaluated and indicated that the effect of e_0 is found to be more than other two physical parameters (ω_n and *LL*).

6. Conclusions

The determination of the Compression index (Cc) of normally consolidated clays utilized to assess the consolidation settlement of shallow foundations takes much more time. Therefore, empirical correlations between Cc value and soil properties can be a suitable alternative for the prediction of Cc value. However, uncertainty in the measurements of input parameters has always been a major concern. Robust optimization is widely used due to its computational tractability for many classes of uncertainty sets and problem types. Therefore, in this research, an innovative method based on robust optimization has been used to investigate the effect of such uncertainties in the prediction of the Cc value. To achieve this, 433 data points have been taken from geotechnical investigation locations in Mazandaran province of Iran. The identification procedure used is based on the robust model of the least square which is a second order cone problem and is capably solved by the interior point technique. Based on Frobenius norm of the data points, uncertainty definition is presented and examined against the correlation coefficients for several empirical models and thus optimum values are determined. The comparison of results of the RO models with those of previous models indicate the robust method is a better pattern recognition tool for datasets with degrees of uncertainty. The variation of the Cc values with soil properties, namely, water content (ω_n), initial void ratio (e_o), and liquid limit (LL), by considering different value of uncertainties (0, 5 and 10%) was considered and indicated that the effect of e_o is more than other two physical parameters (ω_n and LL). Also, the linear robust model developed by using LL values is more efficient than the available empirical formulas. The best model performance during in deterministic valuation and considering uncertainty is further shown.

It should be noted that the robust models to predict *Cc* were obtained using limited amount of database. Again, more studies are required to check the validity of the equations that have been derived for other database.

Acknowledgment

This research database was supported by Technical and Soil Laboratory of Mazandaran Province of Iran (Sari City), A. Khatami, M. Taheri and H. Shirsavar cooperation. These supports are gratefully acknowledged.

References

[1] MolaAbasi H, Shooshpasha I, Ebrahimi A. Prediction of Compression Index of Saturated Clays (Cc) Using polynomial models. Sci Iran 2016;23:500–7. doi:10.24200/sci.2016.2134.

[2] Al-Taie AJ, Al-Bayati AF, Taki ZNM. Compression index and compression ratio prediction by artificial neural networks. J Eng 2017;23:96–106.

[3] Dagdeviren U, Demir AS, Kurnaz TF. Evaluation of the Compressibility Parameters of Soils Using Soft Computing Methods. Soil Mech Found Eng 2018;55:173–80. doi:10.1007/s11204-018-9522-4.

[4] Mohammadzadeh S D, Bolouri Bazaz J, Vafaee Jani Yazd SH, Alavi AH. Deriving an intelligent model for soil compression index utilizing multi-gene genetic programming. Environ Earth Sci 2016;75:262. doi:10.1007/s12665-015-4889-2.

[5] Kumar R, Jain PK, Dwivedi P. Prediction of compression index (Cc) of fine grained remolded soils from basic soil properties. Int J Appl Eng Res 2016;11:592–8.

[6] Polidori E. On the intrinsic compressibility of common clayey soils. Eur J Environ Civ Eng 2015;19:27–47. doi:10.1080/19648189.2014.926295.

[7] Nesamatha R, Arumairaj PD. Numerical modeling for prediction of compression index from soil index properties. Electron J Geotech Eng 2015;20:4369–78.

[8] Kurnaz TF, Dagdeviren U, Yildiz M, Ozkan O. Prediction of compressibility parameters of the soils using artificial neural network. Springerplus 2016;5:1801. doi:10.1186/s40064-016-3494-5.

[9] Habibbeygi F. Determination of The Compression Index of Reconstituted Clays Using Intrinsic Concept And Normalized Void Ratio. Int J GEOMATE 2017;13. doi:10.21660/2017.39.98271.

[10] Buchheim C, Kurtz J. Min-max-min robustness: a new approach to combinatorial optimization under uncertainty based on multiple solutions. Electron Notes Discret Math 2016;52:45–52. doi:10.1016/j.endm.2016.03.007.

[11] Gorissen BL, Yanıkoğlu İ, den Hertog D. A practical guide to robust optimization. Omega 2015;53:124–37. doi:10.1016/j.omega.2014.12.006.

[12] Yanıkoğlu İ, Gorissen BL, den Hertog D. A survey of adjustable robust optimization. Eur J Oper Res 2019;277:799–813. doi:10.1016/j.ejor.2018.08.031.

[13] Alizadeh F, Goldfarb D. Second-order cone programming. Math Program 2003;95:3–51.

[14] Kalantary F, MolaAbasi H, Salahi M, Veiskarami M. Prediction of liquefaction induced lateral displacements using robust optimization model. Sci Iran 2013;20:242–50.

[15] MolaAbasi H, Kalantary F, Salahi M. Uncertainty in Shear Wave Velocity Based on Standard Penetration Test by Using Error Least Square Model. J Eng Geol 2013;6:1559–76.

[16] Azzouz AS, Krizek RJ, Corotis RB. Regression Analysis of Soil Compressibility. Soils Found 1976;16:19–29. doi:10.3208/sandf1972.16.2_19.

[17] Koppula S. Statistical Estimation of Compression Index. Geotech Test J 1981;4:68. doi:10.1520/GTJ10768J.

[18] Rendon-Herrero O. Closure to " Universal Compression Index Equation " by Oswald Rendon-Herrero (November, 1980). J Geotech Eng 1983;109:755–61. doi:10.1061/(ASCE)0733-9410(1983)109:5(755).

[19] Park H Il, Lee SR. Evaluation of the compression index of soils using an artificial neural network. Comput Geotech 2011;38:472–81. doi:10.1016/j.compgeo.2011.02.011.

[20] Ahadiyan J, Ebne JR, Bajestan MS. Prediction determination of soil compression index, Cc. Ahwaz Reg (In Persian) J Fac Eng 2008;35:75–80.

[21] Gunduz Z, Arman H. Possible relationships between compression and recompression indices of a low-plasticity clayey soil. Arab J Sci Eng 2007;32:179.

[22] Farzi M. Suggesting a Formula to Calculate the Compression Index in Ahvaz. Indian J Sci Technol 2017;10:1–8. doi:10.17485/ijst/2017/v10i32/106232.

[23] Mayne PW. Cam-Clay Predications of Undrained Strength. J Geotech Eng Div 1980;106:1219–42.

[24] Terzaghi K, Peck RB, Mesri G. Soil Mechanics in Engineering Practice. Wiley; 1996.

[25] Jain VK, Dixit M, Chitra R. Correlation of plasticity index and compression index of soil. Int J Innov Eng Technol 2015;5:263–70.

[26] Al-Khafaji AWN, Andersland OB. Equations for Compression Index Approximation. J Geotech Eng 1992;118:148–53. doi:10.1061/(ASCE)0733-9410(1992)118:1(148).

[27] Yoon GL, Kim BT. Regression analysis of compression index for Kwangyang marine clay. KSCE J Civ Eng 2006;10:415–8. doi:10.1007/BF02823980.

[28] Ozer M, Isik NS, Orhan M. Statistical and neural network assessment of the compression index of clay-bearing soils. Bull Eng Geol Environ 2008;67:537–45. doi:10.1007/s10064-008-0168-8.

[29] Sturm JF. Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones. Optim Methods Softw 1999;11:625–53. doi:10.1080/10556789908805766.

[30] Wang B. Implementation of interior point methods for second order conic optimization, MSc thesis, MacMaster University 2003.

[31] Liu S, Cai G, Puppala AJ, Tu Q. Prediction of embankment settlements over marine clay using piezocone penetration tests. Bull Eng Geol Environ 2011;70:401–9. doi:10.1007/s10064-010-0329-4.

[32] Dutta RK, Rani R, Rao TG. Prediction of ultimate bearing capacity of skirted footing resting on sand using artificial neural networks. J Soft Comput Civ Eng 2018;2:34–46.

[33] Salahudeen AB, Ijimdiya TS, Eberemu AO, Osinubi KJ. Artificial neural networks prediction of compaction characteristics of black cotton soil stabilized with cement kiln dust. Soft Comput Civ Eng 2018;2:50–71. doi:10.22115/SCCE.2018.128634.1059.

[34] Srikanth S, Mehar A. Development of MLR, ANN and ANFIS Models for estimation of PCUs at different levels of service. J Soft Comput Civ Eng 2018;2:18–35. doi:10.22115/SCCE.2018.50036.