Jordan Journal of Civil Engineering, Volume **12**, No. **3**, **2018**

Performance Evaluation of Empirical Equations for Compression Index Using Experimental Data and Statistical Analysis

***Mohammad Bahrami* [1)*] *and Seyed Morteza Marandi* [2)]**

[1)] PhD Student, Civil Engineering Department, Shahid Bahonar University of Kerman, Kerman 7616914111, Iran. E-Mail: m.bahrami@eng.uk.ac.ir
* Corresponding Author.
[2)] Professor, Civil Engineering Department, Shahid Bahonar University of Kerman, Kerman 7616914111, Iran. E-Mail: marandi@uk.ac.ir

ABSTRACT

A quick and economical solution to find the consolidation settlement is to use empirical equations to estimate the soil compression index. These equations may not be appropriate for soils in different regions. In this study, the performance of some empirical equations was determined by estimating the compression index of Kerman clayey soil, Iran. In order to evaluate the compression index, 122 standard odometer tests were performed on undisturbed specimens and the compression indices calculated. Four quantity criteria including error and ranking were selected. The computational measures included determination of the root mean square error (RMSE), ratio of the estimated compression index to the measured value in the laboratory (K), the ranking index (RI) and the ranking distance (RD). Ranking index and ranking distance are related to the accuracy and precision of the estimated results compared to real data. Finally, evaluation of performance of the empirical equations was carried out according to the RD, since it is a self-inclusive statistical parameter that gives accurate and precise estimation compared to other parameters. Results of the analysis indicated that the best estimate of compression index of Kerman soil with empirical equations is based on the initial void ratio compared to other soil properties.

KEYWORDS: Compression index, Experimental data, Empirical equations, Statistical analysis, Ranking distance, Ranking index.

INTRODUCTION

Correlation relationships have made important contributions to geotechnical studies. The importance of these relationships is due to extensive changes in soil deposits caused by geological conditions, including stress level, time and soil formation environment. The characteristics, derived experimentally with difficultly, can be obtained more easily through relationships associated with soil physical parameters (Onyejekwe et al., 2015). One of the characteristics required in estimating the settlement caused by the consolidation phenomenon is the compression index. The compression index is used to estimate the consolidation settlement of fine-grain soils under different foundations before the design and selection of the appropriate foundation system. The compression index is a dimensionless quantity, equal to the end straight-line gradient of the $e - log\sigma'$ curve in the consolidation test (Han, 2015).

In recent decades, researchers have tried, by

Received on 24/1/2018.
Accepted for Publication on 14/4/2018.

statistical analysis of the data obtained from the results of experimental and semi-experimental consolidation tests, to relate the compression index to the physical properties of soils and estimate it. Compressibility of a number of different types of clay was investigated by Skempton (1994), who found a correlational relationship for compression index. Many researchers presented correlation relations using the soil liquid limit (Terzaghi and Peck, 1967; Azzouz et al., 1976; Mayne, 1980; Bowles, 1979). Equations based on the initial void ratio (e_0) were reported by Azzouz et al. (1976), Hough (1957) and Ahadian et al. (2008). Other researchers reported equations using some soil properties with the help of the artificial neural network method (Kalantary and Kordnaeij, 2012; Park and Lee, 2011; Kordnaeij et al., 2015; Al-Taie et al., 2017). The natural water content of the soil was another independent quantity which was used in the relationships (Azzouzz et al., 1976; Kalantary and Kordnaeij, 2012; Rendon-Herrero, 1983b).

Another correlation relationship associated with liquid limit and specific gravity of solid soil was presented (Nagaraj and Murthy, 1986). A model combining the liquid limit and soil moisture content was reported by Azzouz et al. (1976). In their research, they reported a different relationship using e_0 and soil moisture content. Al-Khafaji and Andersland (1992) presented a correlation equation with compression index through the combination of e_0 and the soil liquid limit.

Initial void ratio and specific gravity of solid soil were used in empirical equations (Rendon-Herrero, 1983a; Azzouz et al., 1976). A correlation equation was presented by combining three quantities of e_0, *in situ* water content and soil liquid limit (Yoon and Kim, 2004). Mohammadzadeh et al. (2014) developed a Multi-Expression Program (MEP) by introducing a computational method for modeling and predicting the compression index. Tiwari and Ajmera (2012) presented a new correlation relationship for the compression index of reconstituted soils from Kaolinite, Ilite, Montmorillonite and Quartz mixtures. Past research has shown that only Nishida's relationship has theoretical basis among the presented relationships (Nishida, 1956).

In determining the correlation relations, researchers divided the physical properties of soils into three groups. The first group is based on the mineralogical combination of the soil composition such as the liquid limit, while the second group of equations is based on environmental characteristics such as moisture content, dry unit weight,… etc. and the third group is a combination of the characteristics of the first and second groups.

Generally, the number of relationships presented in the research shows that they cannot be considered valid for any soil conditions. But, at the same time, these relationships may be valid in a certain framework (Giasi et al., 2003). Therefore, the optimal use of these relationships with respect to cost and time should be tested for an area in order to verify their consistency with real data. Since the adoption of the best practical relationship is not always easy, in this study, an effort has been made to investigate the reliability of some of the empirical relationships related to the compression index using a statistical evaluation criterion for the Kerman-Iran region, where the water level gradually rose and caused a lot of problems in the structure.

Geological and Geotechnical Characteristics of the Study Region

The region is geologically located in the main sedimentary-structural zone of central Iran. The Kerman deposits are a mixture of clay, silt and sand, whose proportions vary in depth and surface. The thickness of alluvium in the plains is relatively high. So, it is estimated to be 350 m in the south-east of Kerman plain. The thickness of the alluvial plains ranges from 150 to 170 meters and in the west of the city from 150 to 200 meters. In the south-east of the plains, it ranges from 300 to 350 meters. The alluvium thickness in the central part of the plains is between 100 and 150 meters and is measured or estimated in the range from 100 to 130 meters in the old texture of the city. The results of geophysical studies on Kerman plains indicated that the depth of alluvial deposits is at about 300 meters and that there are fine grains above 100 meters. The total thickness fluctuation of alluvium in Kerman is between 50 and 350 meters in total. Alluvial limestone in the north of the plains is often composed of Cretaceous limestone. In the eastern margin, there is the Neogene conglomerate or lime and in the center of Kerman plains, there are Marine Neogene deposits (Aminizadeh et al., 2013). Geological map of the location of the site studied is shown in Fig. 1.

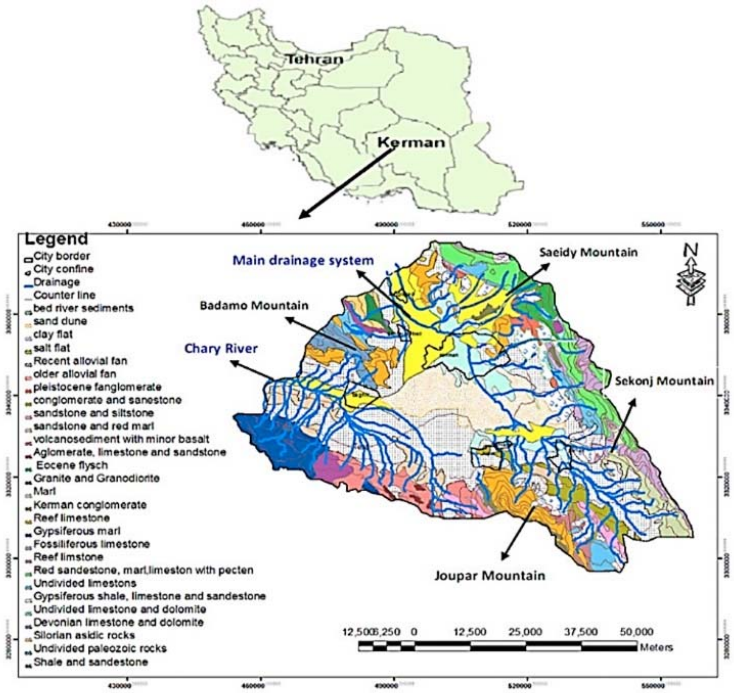


Figure (1): Geological map of sedimentary basin of Kerman city

In order to determine the soil compression index in the Kerman region, undisturbed samples are provided from the drilling boreholes in different parts of the city. In Table 1, some data is derived from geotechnical borehole log plotted at this site. Physical soil experiments were carried out to identify and determine the soil classification. The soil is generally classified into the CL category for a unified classification system of up to 30 meters depth. Soil grain diagrams at different depths are shown in Fig. 2.

Table 1. Data of drilled borehole log

Depth (m)	**Unified classification** (USCS)	***In situ* moisture content** (%)	**Plasticity index** (%)	***In situ* dry density** (kN/m^3)
2.0	CL	5.4	12	13.4
4.0	CL	7.9	11	13.1
7.5	CL	11.1	9	14.3
10.0	CL	12.8	10	14.9
12.0	CL	13.5	11	15.2
13.0	CL	14.2	9	
15.0	CL	15.9	10	

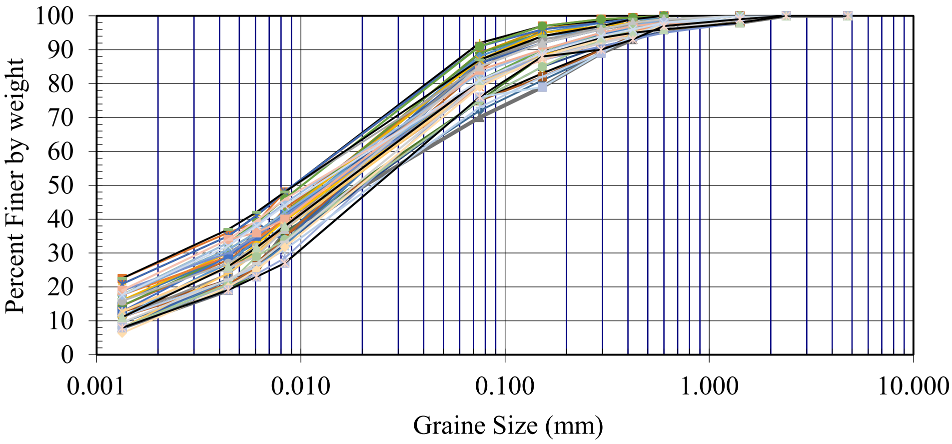


Figure (2): The results of grain size and soil hydrometric tests of the site under study

Data Preparation

Due to enormous changes in the type and structure of soils, variation in a structure's application along with increased loading as well as various levels of consolidation settlement of soils in different areas, it is impossible to provide a proper average value for this type of settling in a vast city. The estimation of the average amount of consolidation settlement requires extensive experiments and tests (Le et al., 2015). Therefore, 10 sites in the city that witnessed the highest number of constructions were selected. In order to obtain the required data from geotechnical basic information, 122 samples of 21 borehole logs with a maximum depth of 20 and 30 meters were taken. All samples were undistributed and prepared with a four inch-diameter sampler with disturbance of less than 10% (according to Das, 2010).

One-dimensional consolidation test was performed on all specimens according to ASTM D 2435-11 standard with applying 25, 50,100,200,400 and 800 kPa vertical stresses. Sieve analysis and hydrometric tests were conducted using the ASTM D 422-07 method, Atterberg limits were determined by ASTM D 4318, *in situ* water content and specific gravity of the solid soil were found by ASTM D 2216 and ASTM D 854-10, respectively. The data obtained included the percentage passing through 75 and 2 micron screens, liquid limit (LL) values, plastic index (PI), *in situ* water content (W_n), dry unit weight, specific gravity of solid soil (G_s), e_0 and saturation (S).

The data was extracted and recorded for various depths. The values of soil compression index were extracted and recorded by measuring the end straight line slope of the diagrams. Consolidation test graphs in saturation and natural moisture content conditions shown in Fig. 3 suggest that the soil in this area is of collapsible clayey type.

The data is presented as descriptive statistics of the variables in Table 2. Similarly, the data is shown in Fig. (4; a-e) as a graph of the frequency of data variables with normal distribution curve.

The error and ranking measurement method was used to determine the performance of empirical equations for compression index among other soil physical properties. The computational criteria included determination of root mean square error, ratio of predicted to measured data from laboratory, ranking index and ranking distance (Hamza et al., 2016).

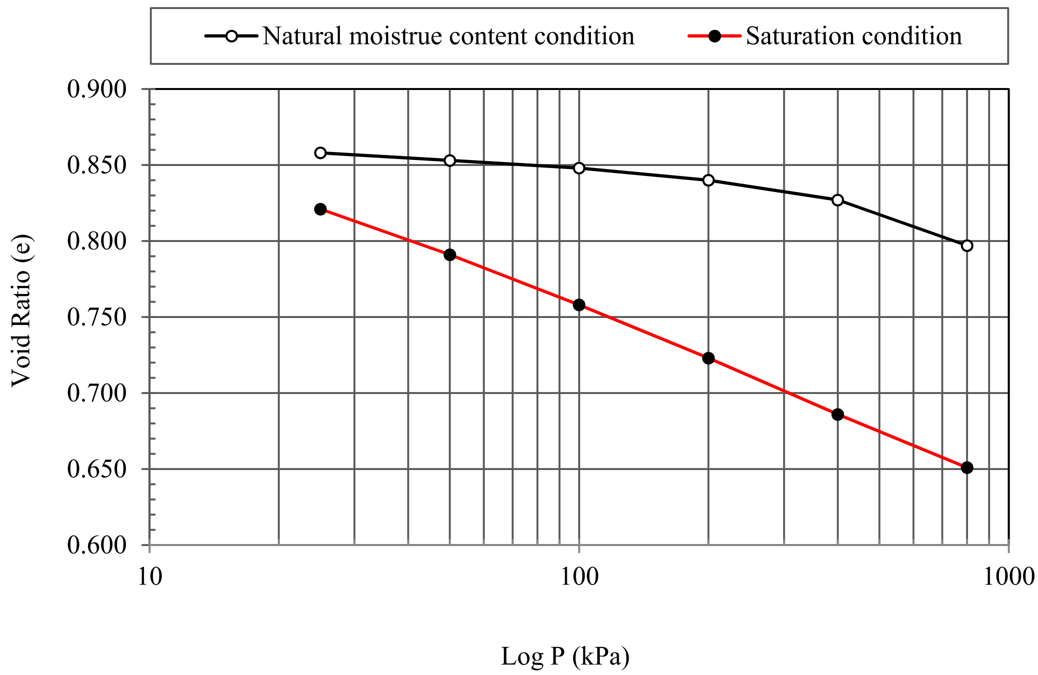


Figure (3): Variation of consolidation curves due to collapsibility phenomenon

Table 2. Descriptive statistics of data for some physical properties of soil

Soil properties	N	Mean	Minimum	Maximum
W_n (%)	122	21.10	8.20	33.10
γ_d (kN/m^3)	122	14.90	12.40	17.60
e_0	122	0.75	0.517	1.17
G_s	122	2.59	2.510	2.72
C_c	122	0.11	0.04	0.27
LL (%)	122	29.00	25.00	34.00
PI (%)	122	9.00	5.00	13.00

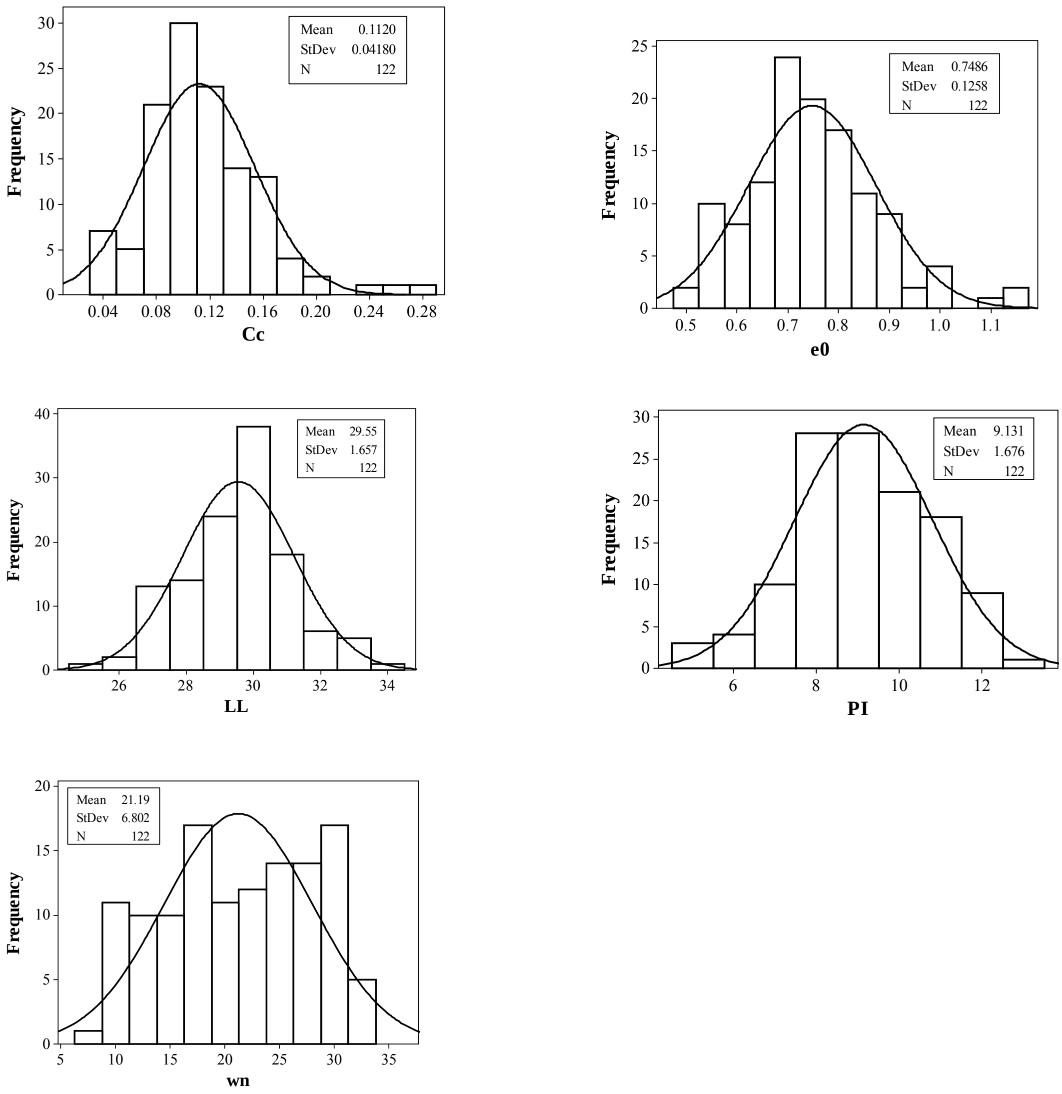


Figure (4; a-e): Frequency charts of some data variables

Empirical Relationships Used

The selected empirical relationships related to the compression index, whose functional ability has been investigated in this study, are presented in Table 3. The relationships have been selected based on the dependent variables and functionality.

Table 3. Empirical equations used in ranking evaluation

Independet Variable	No. of Model (M)	Equation	Reference	Region / Conditions of Applicability
W_n	1	$Cc = 0.01W_n - 0.05$	Azzouz et al. (1976)	All natural soils
	2	$Cc = 0.01W_n - 0.75$	Rendon-Herrero (1983b)	Soil systems of all complexities and types
	3	$Cc = 0.0074W_n - 0.007$	Kalantary and Kordnaeij (2012)	Clayey soils of Mazandaran, Iran
e_o	4	$Cc = 0.156(e_0 + 0.0107)$	Bowles (1979)	All clays
	5	$Cc = 0.4(e_0 - 0.25)$	Azzouz et al. (1976)	All natural soils
	6	$Cc = 0.287(e_0 - 0.015)$	Ahadiyan et al. (2008)	Over-consolidated clay with low to meduim plasticity
	7	$Cc = 0.3(e_0 - 0.27)$	Hough (1957)	Inorganic silty, sandy -silty clay
	8	$Cc = 0.3608\, e_0 - 0.0713$	Kalantary and Kordnaeij (2012)	Clayey soils of Mazandaran, Iran
LL	9	$Cc = 0.0046(LL - 9)$	Bowles (1979)	Brazilian clay, disturbed over-consolidated clay
	10	$Cc = 0.009(LL - 10)$	Terzaghi and Peck (1967)	Normally consolidated, moderately sensitive clays
	11	$Cc = 0.006(LL - 9)$	Azzouz (1976)	All natural soils
	12	$Cc = (LL - 13)/109$	Mayne (1980)	Low-plasticity clay
	13	$Cc = 0.007(LL - 10)$	Skempton (1944)	Remolded clays, normally consolidated and low-sensitivity clays
LL,G_s	14	$Cc = 0.2343(\frac{LL}{100})G_s$	Nagaraj and Murthy (1985)	-
LL, W_n	15	$Cc = 0.009W_n + .002LL - 0.1$	Azzouz (1976)	All natural soils
e_o , W_n	16	$Cc = 0.04\ (e_0 + 0.001W_n - 0.25$	Azzouz (1976)	All natural soils
LL, e_o	17	$Cc = 0.156 + 0.141e_0 + 0.00058\, LL$	Al-Khafji and Andersland (1992)	
e_o , G_s	18	$Cc = 0.141\, Gs^{1.2} (\frac{1 + e_0}{Gs})^{2.38}$	Rendon-Herrero(1983a)	Soil systems of all complexities and types
LL, e_o , W_n	19	$Cc = 0.37(e_0 + 0.003LL + 0.0004Wn - 0.34)$	Azzouz (1976)	All natural soils
	20	$Cc = -0.404 + 0.341e_0 + 0.004LL + 0.006Wn$	Yoon and Kim (2006)	Marine clay

As shown in Table 3, the compression index is related to various variables of soil physical properties. The relationships are presented in single-parameter or multi-parameter form in various combinations of independent variables.

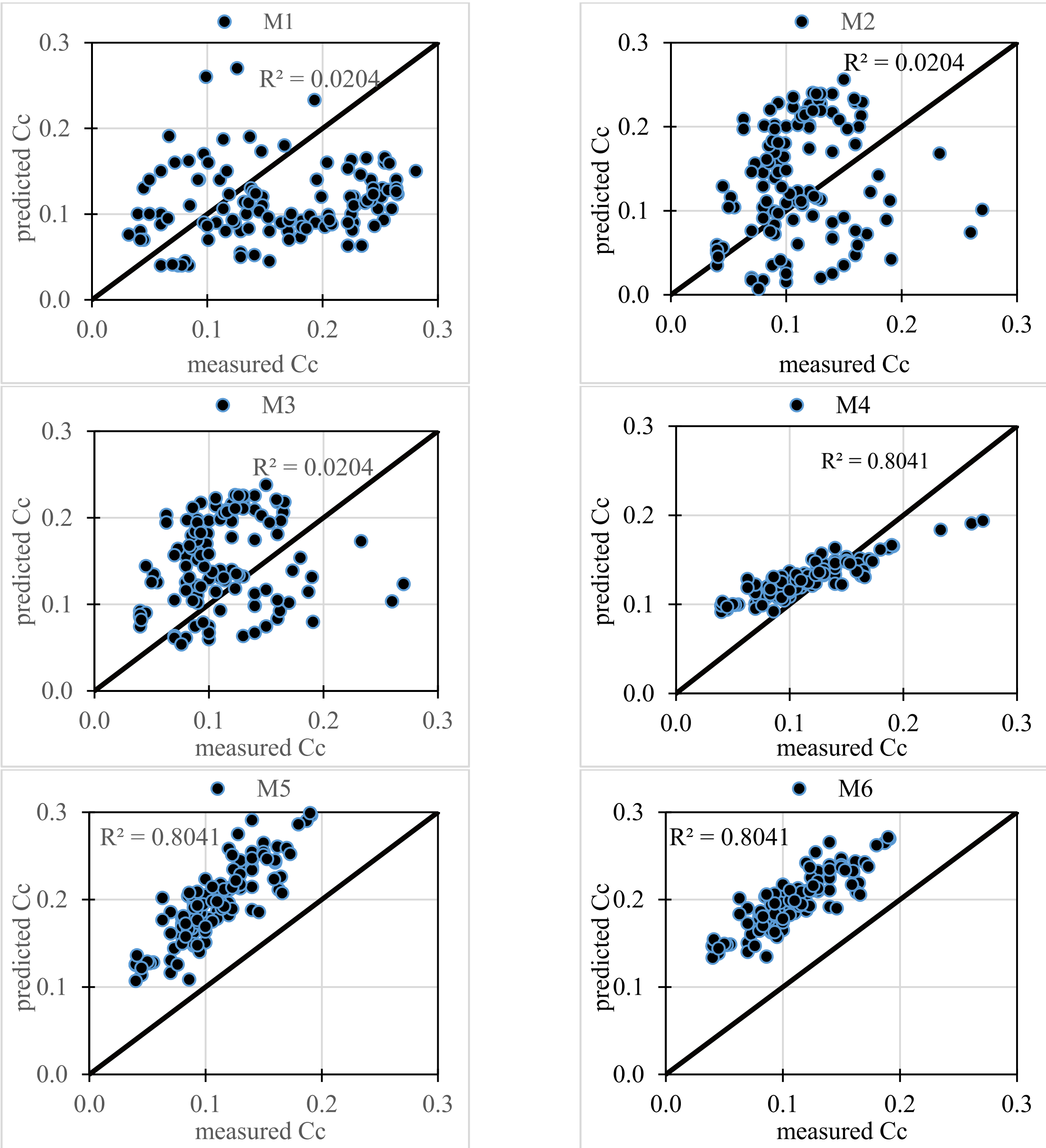
Most of the equations that use data with extensive resources can be used for fine-grained clayey soils. But, relationships derived from location-specific data are more appropriate for the same specific area. As can be seen from Table 3, most of the equations are based on single-variable linear relationships. Liquid limit is related to particle size and composition of soil minerals (Onyejekwe et al., 2015). The liquid limit can be considered an overcoming factor that shows the amount of water absorbed by soil particles and plays a fundamental role in determining the compression index from the equations (Giasi et al., 2003). On the other hand, liquid limit-based relationships have been evaluated by some researchers in terms of reliability. Some soils, despite the same liquid limit, showed different compressibilities (Al-Khafaji and Andersland, 1992; Giasi et al., 2003). This suggests that the liquid limit does not fully determine the soil's plastic properties for the prediction of compressibility.

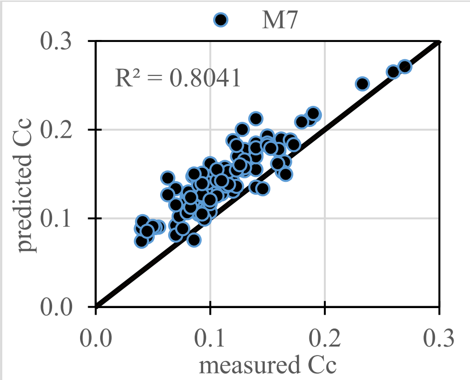
As shown in Table 3, in a significant number of equations, the compression index is gained using e_0 as a parameter. Settlement due to consolidation is usually determined by e_0. Here, it should be noted that the initial void ratio will be determined in the description of the soil structure through geological history, as well as the real expression of the state of the soil at the site. Hence, the equations do not show a history of stress in soils without considering e_0.

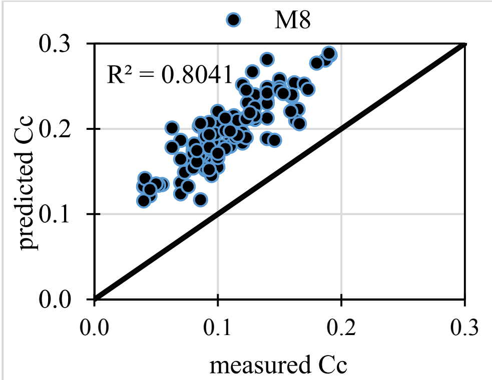
Considering these causes, to realize the behavior of specimens in undisturbed state, a multi-variable equation corresponding to the initial undistributed state of the soil (for example, the initial void ratio) along with other physical properties can be considered as beneficial for compressibility prediction (Hamza et al., 2016; Onyejekwe et al., 2015; Giasi et al., 2003). On the other hand, instead of using the quantities for an ideal relationship, it should be noted that the final decision on the best equation must be taken after a precise statistical analysis for predicting compression index with the help of experimental data.

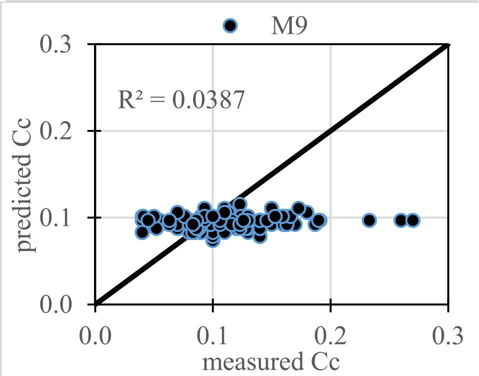
Empirical equations in Table 3 are selected from a vast scale of data sources, where special conditions of applicability for a specific site are not suitable for other sites. These relationships show different results between estimated and measured data due to their uncertainties and there is a lot of error in their use. Hence, the simple application of these equations without adequate basic evaluation can have a detrimental effect on the design of structures.

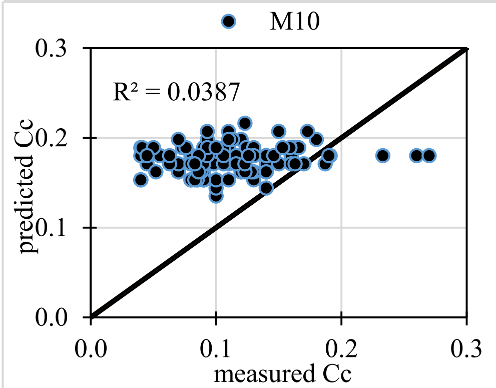
From this point of view, the compression index values associated with the predicted results (from the correlation equations) and the real results (from those measured in laboratory) for the selected equations are presented in Fig. 5. Models M1-M20 in Fig. 5 show that except for cases where the correlation coefficient is moderate to good (from 0.28 to 0.79), in most of the equations, there is a large difference between the estimated and experimental results in terms of compression index. This calls for the need for a precise statistical evaluation of the application of these equations. Therefore, using the performance criteria related to the method of statistical analysis for this research, these are the most applicable relationships proposed for Kerman's soil.

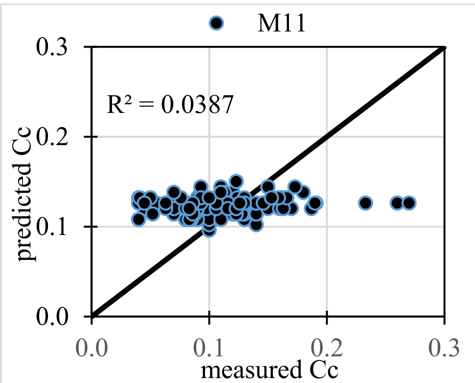

M1
R² = 0.0204
predicted Cc
measured Cc
0.0
0.1
0.2
0.3
M2
R² = 0.0204
predicted Cc
measured Cc
M3
R² = 0.0204
predicted Cc
measured Cc
M4
R² = 0.8041
predicted Cc
measured Cc
M5
R² = 0.8041
predicted Cc
measured Cc
M6
R² = 0.8041
predicted Cc
measured Cc

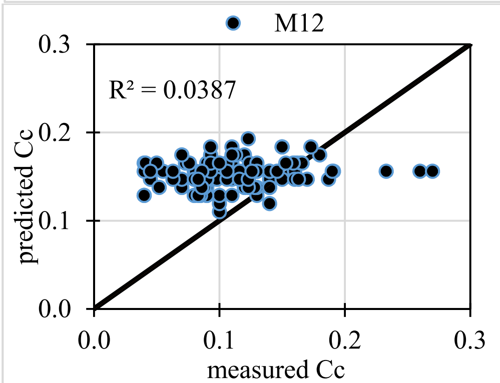

M7
R² = 0.8041
predicted Cc
measured Cc
0.0
0.1
0.2
0.3

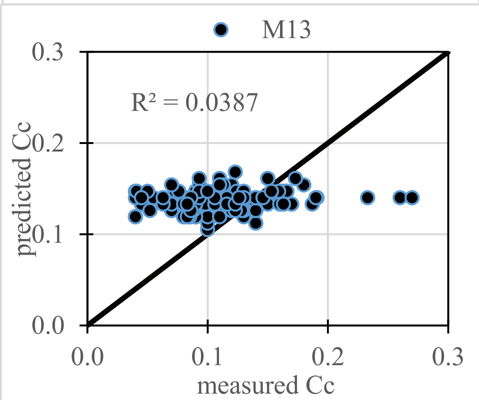

M8
R² = 0.8041
predicted Cc
measured Cc
0.0
0.1
0.2
0.3

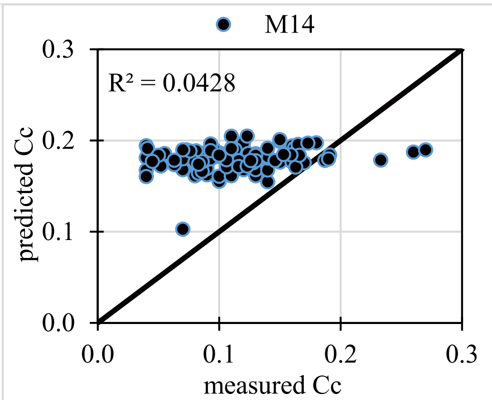

M9
R² = 0.0387
predicted Cc
measured Cc
0.0
0.1
0.2
0.3


M10
R² = 0.0387
predicted Cc
measured Cc
0.0
0.1
0.2
0.3


M11
R² = 0.0387
predicted Cc
measured Cc
0.0
0.1
0.2
0.3


M12
R² = 0.0387
predicted Cc
measured Cc
0.0
0.1
0.2
0.3


M13
R² = 0.0387
predicted Cc
measured Cc
0.0
0.1
0.2
0.3


M14
R² = 0.0428
predicted Cc
measured Cc
0.0
0.1
0.2
0.3

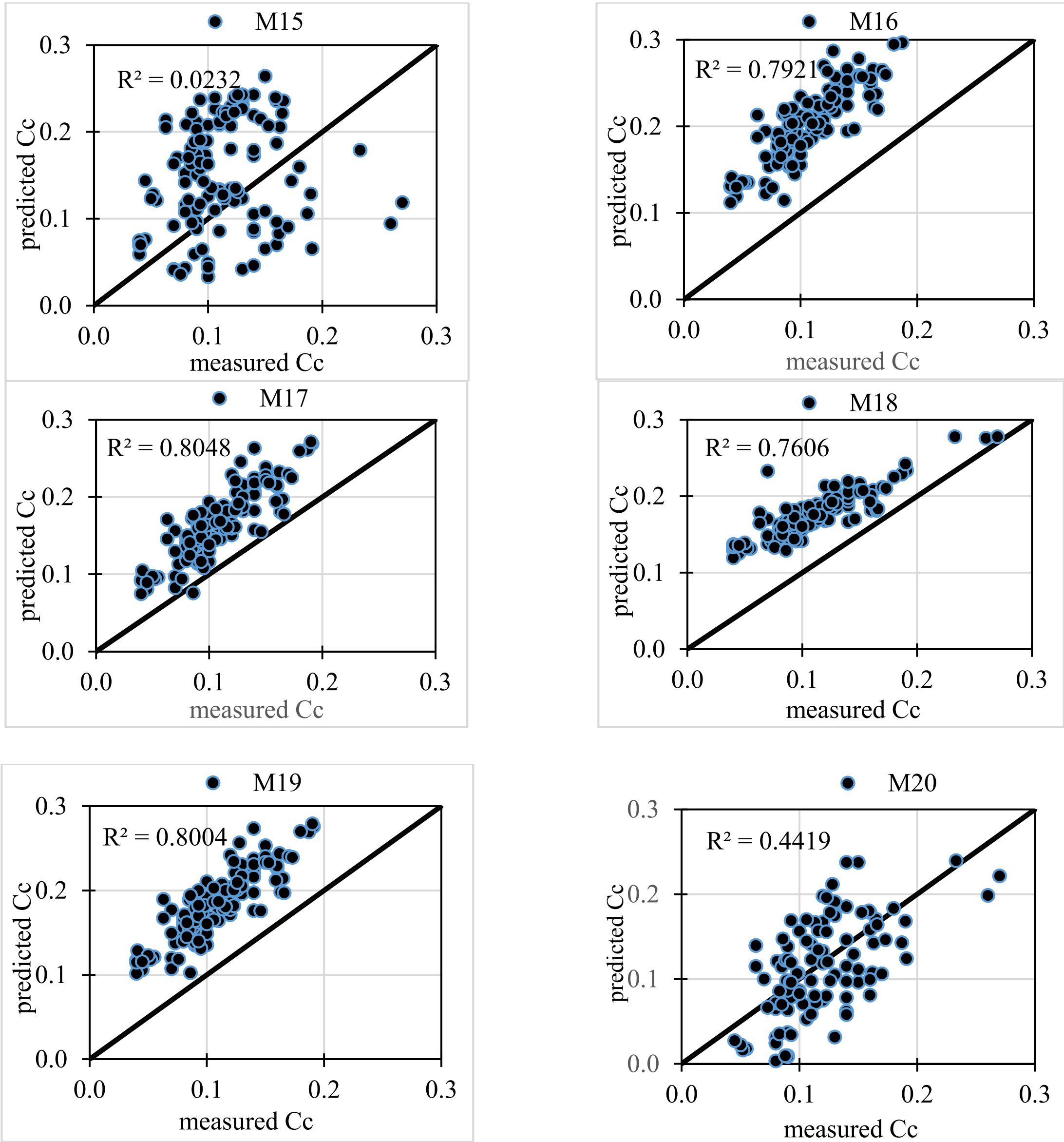


Figure (5; M1-M20): Compression index predicted from the correlation equations in terms of compression index measured from laboratory data

Statistical Analysis Method

The process of evaluation and validation of selected correlation equations is carried out using a statistical analysis method based on the criteria described below:

Root Mean Square Error (RMSE)

The RMSE indicator is very useful and relatively robust when major errors are detected. In terms of performance evaluation, the smaller the RMSE value, the better the performance of the model (Ozer, 2008). The RMSE index is calculated from Eq. 1.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Cc_{est} - Cc_{lab})^2}{N}} \quad (1)$$

where N is the number of data, $C_{C_{est}}$ is the compression index estimated (predicted) from the correlation equations and $C_{C_{lab}}$ is the compactness index measured in the laboratory.

Bias Factor (K)

The bias factor (K), which is defined as the ratio of the estimated compression index to the measured compression index in the laboratory, is calculated from Eq. 2.

$$K = \frac{C_{C_{est}}}{C_{C_{lab}}} \quad (2)$$

K is determined based on a probabilistic combination method and is useful for obtaining selected variables from a set of values of that index (Orr and Cherubini, 2003). It can be said that in empirical relations, for k<1, the data obtained is less than the specified amount and for k >1 it is higher than the specified amount. From the theoretical point of view, the value of K can vary from zero to infinity and its optimal value is one. This leads to asymmetrical K distribution around the mean value. In this distribution, the numbers of less-estimated and more-estimated values are not equal (Briaud and Tucker, 1998). For this reason, in order to compare different methods of estimating the compression index, the K values obtained from each method should be statistically analyzed.

Therefore, the appropriate statistical parameters are determined and evaluated to specify which method of estimation and prediction is more relevant to the experimental data (Giasi et al., 2003).

The mean value and standard deviation for K represent the importance of the accuracy and precision in the prediction statistical analysis method. The mean value (i.e., center) of K can be closely related to accuracy, whereas standard deviation (i.e., dispersion) of K can be useful for prediction precision (Orr and Cherubini, 2003). In an exact correlation relationship, an accurate and precise relationship provides a mean value multiplied by K equal to 1 and a standard deviation multiplied by K equal to 0. This means that the predicted result is equal to the laboratory data. This is an excellent result. However, in fact, it is better when standard deviation multiplied by K is closer to 0 and mean value multiplied by K is closer to 1 (Abu-Farsakh and Titi, 2004).

In past research related to the methods of estimating load bearing capacity and settlement, K value has been used as a satisfactory part for evaluating the performance of empirical equations. Specifically, the mean value and standard deviation of K values are used to process these equations, including some ranking measures, such as RI and RD (Hamza et al., 2016; Onyejekwe et al., 2015; Giasi et al., 2003; Sridhrarn and Nagaraj, 2000; Abu-Farskh and Titi, 2004).

Ranking Index (RI)

The ranking index is derived from Eq.3:

$$RI = \mu \times ln\left[\left(\frac{C_{c_{est}}}{C_{c_{lab}}}\right)\right] + S \times ln\left[\left(\frac{C_{c_{est}}}{C_{c_{lab}}}\right)\right] \quad (3)$$

where μ and s respectively indicate the mean value and standard deviation of all K values of the analyzed data sequence. RI is obtained by applying the log-normal distribution of K. This is a general indicator that simplifies the asymmetrical K distribution problem. In addition, it states a general judgment on the accuracy and precision of an empirical equation using the mean value (μ) and the standard deviation (s) of all the K data (Briaud and Tucker, 1988; Onyejekwe et al., 2015; Orr and Cherubini, 2003). To evaluate the empirical equation, one can say that the lower the RI, the better the performance of the equation (Hamza et al., 2016; Briaud and Tucker, 1998; Abu-Farsakh and Titi, 2004).

Ranking Distance (RD)

The ranking distance is calculated from Eq. 4:

$$RD = \sqrt{\left[1 - \mu_{\left(\frac{Cc_{set}}{Cc_{lab}}\right)}\right]^2 + \left[s_{\left(\frac{Cc_{set}}{Cc_{lab}}\right)}\right]^2} \quad (4)$$

In equation (4), the RD index is a general alternative parameter for determining the accuracy and precision of a prediction and is related to the mean value (μ) and standard deviation (s) associated with all the values of K in each series of the analyzed data. Considering the diagram of K mean value (for example μ) on the X axis and the standard deviation of K on the Y axis, the value of RD is equal to the distance between the point of estimation determined with the empirical equations and the optimal point determined by the mean (central) $K = 1$ (For example, $\mu = 1$) as well as the standard deviation $K = 0$ (dispersion; for example, $s = 0$). The RD index has the ability to predict empirical correlation with accuracy. For evaluation, it could be expressed that low RD values with empirical correlation predictions are consistent with high precision and high accuracy and when high RD values are relevant to predictions, these may be highly inaccurate or lacking precision (Hamza et al., 2016; Orr and Cherubini, 2003.

When rankings are compared in terms of RD index and RI value, it is determined that the RD index ranks the performance of the empirical relationship better than the RI value and considers the accuracy of the equation. It can be said that, in estimating methods that are very accurate or very precise, the RI rank is better than the RD index. Similarly, when ranking correlation equations with the same precision level with RI, the importance of the accuracy problem is greater than that of the accuracy problem when ranking with RD. Hence, the RD index seems a better index for comparing different prediction procedures. However, considering the results based on RD and RI indicators can be useful in evaluating the validity of empirical equations (Orr and Cherubini, 2003).

ANALYSIS AND RESULTS

Twenty equations were used in terms of performance evaluation of compression correlation equations. One hundred twenty two samples of fine-grained silty clay soil were studied and evaluated for prediction and estimation of the compression index. The uncertainty of the correlation equations was calculated and the best relations were determined by ranking the equations.

The results of the empirical relationships are summarized in Table 4, based on statistical estimates of the RMSE values K (percentages, mean values and standard deviations of K values), RD and RI. For RMSE, it is seen from Table 4 that the experimental correlation equation M4 with e_0 (RMSE = 0.029) among the single-variable equations, M17 with e_0 and LL (RMSE = 0.061) among the two-variable relations and M20 with e_0, W_n and LL (RMSE = 0.053) among the three-variable relationships had the best performance. The best performance for RMSE among all the equations is associated with the M4 equation with e_0 (RMSE= 0.029); $C_c = 0.156\ (e_0 + 0.0107)$ and the worst performance is related to the M16 relationship with W_n and LL (RMSE = 0.098). In general, M4, M7, M11 and M9 equations are ranked, respectively, according to RMSE values.

Regarding performance of the K values, with respect to their proportionality with K = 1 (for example, K <1 or K > 1), it was observed that equations M2–M8 considered the data to be overestimated (K> 1), showing that more than 85% of the K values are higher than 1. This result is also consistent with the work of Giasi et al. (2003). The mean K values range is from 0.36 to 5.17 and the standard deviation range between 0.114 and 1.707. Furthermore, a few equations, such as M20, M9 and M1, underestimate the K values.

Table 4. Correlation equations' performance with applying laboratory data and computing statistical indices

Independent variable	Number of equation	RMSE	K				RI	RD
			<1%	>1%	Mean	SD		
W_n	M1	0.089	25	75	1.59	0.736	1.073	0.941
	M2	0.078	10	90	3.42	1.701	6.302	2.960
	M3	0.071	1	99	3.73	1.297	6.607	3.019
e_0	M4	0.029	8	92	1.05	0.183	0.258	0.164
	M5	0.090	1	99	2.48	0.663	2.855	1.622
	M6	0.089	1	99	2.48	0.501	2.714	1.566
	M7	0.074	3	97	1.79	0.495	1.323	0.929
	M8	0.089	1	99	2.47	0.606	2.784	1.591
LL	M9	0.044	25	75	0.94	0.114	0.052	0.130
	M10	0.076	1	99	1.75	0.217	1.096	0.778
	M11	0.042	2	98	1.22	0.194	0.278	0.269
	M12	0.057	1	99	1.51	0.204	0.703	0.547
	M13	0.048	1	99	1.36	0.196	0.468	0.397
LL,G_s	M14	0.078	1	99	4.44	1.510	7.375	3.476
LL, W_n	M15	0.0700	2	98	3.73	1.573	6.980	3.150
e_0, W_n	M16	0.089	1	99	5.17	1.351	10.717	4.385
LL, e_0	M17	0.061	1	99	4.20	1.341	7.957	3.471
e_0 , G_s	M18	0.067	2	98	1.36	1.486	1.073	2.704
LL, e_0, W_n	M19	0.078	1	99	4.65	1.243	0.071	3.860
	M20	0.053	24	76	2.42	1.707	0.648	2.221

In terms of accuracy, it can be concluded from the mean values of K (the closest mean value K is 1) that the most accurate equations are M9 and M4 with K = 0.94 and 1.05, respectively. Also, M11 with parameter LL and K = 1.22 was acceptable. The performance of the equations with respect to the mean values of K is proportional to the analysis based on RMSE. This suggests that both methods consider the true error equations for ranking.

In terms of precision, considering K as standard deviation (with small K standard deviation), the most precise equations are M9 with LL (K = 0.114), M4 with e_0 (K = 0.183) and M11 with LL (K = 0.194). In examining the relationship performance with respect to RI, in which accuracy and precision are considered (with low ranking index), the best initial estimates of Cc are given by equations M9 with LL (K = 0.052), M4 with e_0 (K = 0.258) and M11 with LL (K = 0.278).

In the case of performance evaluation according to the RD value, which considers accuracy and precision (with small ranking distance), the optimum estimates of Cc are obtained from equations M9 with LL (K = 0.130), M4 with e_0 (K = 0.164) and M11 with LL (K = 0.269).

Taking RMSE, K, RI and RD indices into account, the best equations are summarized in Table 5. The equations for parameters e_0 (M4); $C_c = 0.156\ (e_0 + 0.0107)$ and LL (M9); $C_c = 0.0046\ (LL - 9)$ are compatible with all the assessment criteria (RMSE, K, RI and RD). These relationships are suitable for estimating Cc for the soil of Kerman, Iran, in terms of conformance with all the criteria.

From the best correlation equations in Table 5, we can see that the presence of the parameter e_0 in these relations indicates the importance of soil primitive conditions for compressibility, as was reported in past research (Onyejekwe et al., 2015; Giasi et al., 2003).

Table 5. The best correlations according to independent variables and indices among all equations

Independent variables	RMSE	K		RI	RD
		Mean	SD		
One variable W_n	M3	M1	M1	M1	M1
One variable e_0	M4	M4	M4	M4	M4
One variable LL	M11, M9	M9	M9	M9	M9
Two variables	M17	M18	M17	M18	M18
Three variables	M20	M20	M19	M19	M20
Total	M4	M4	M4	M4	M4

The results of RMSE, RI and RD analyses are presented in Fig. 6. When differences in RMSE, RI and RD values are taken into account in Fig. 6, it is observed that M2, M3 and M14 to M16 equations in the RD curve have a greater distance compared to other equations. In most of the equations, there is the same RMSE curve path with only higher values. We compared the values of RI and RD except for two cases of the mentioned equations in which excesses have occurred and their use is not recommended. In many cases, the RD curve draws a path slightly different from the RMSE curve.

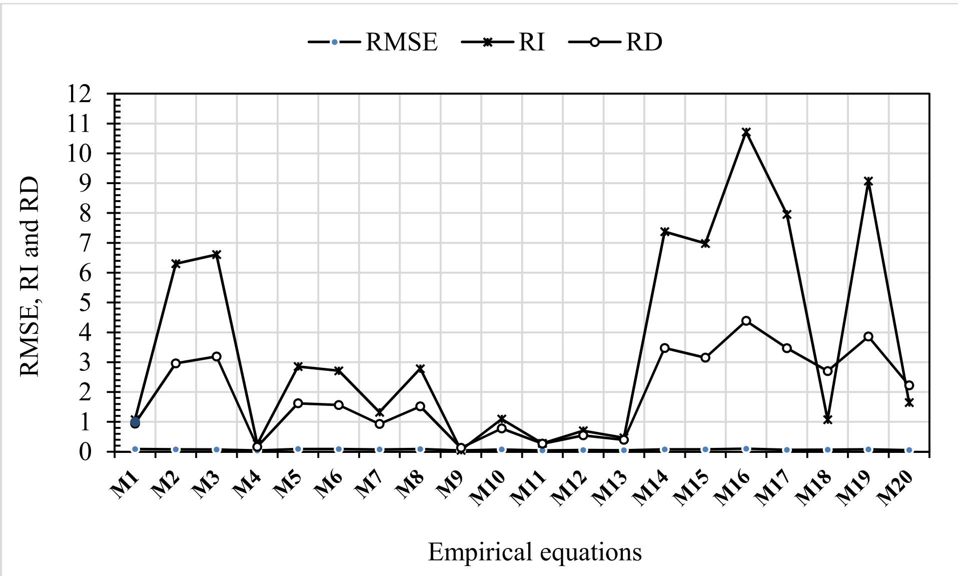


Figure (6): The results of the evaluation of empirical relationships according to RMSE, RD and RI indices

In the overall assessment, the greater the difference between the two criteria (RI and RD), the poorer is the accuracy of Cc relationships (for example, the mean value of K is not close to 1), and if there is a lower difference, there will be more accuracy in the estimates for Cc. High differences (bias) in equations of limited accuracy may be related to the fact that RI, while considering accuracy and precision, places the focus on accuracy due to its mathematical form (Onyejekwe et al., 2015; Giasi et al., 2003).

Regarding the studied area, M4, M9 and M11 equations provide the most appropriate estimates for Cc with respect to low distance while comparing RI and RD values (Table 5).

The current study showed that the initial void ratio was the most important parameter for providing the compression index for Kerman, Iran. Therefore, it is vital to determine the performance of empirical equations with experimental data of each location.

CONCLUSIONS

A number of compression index equations presented in past research have been studied to determine the most reliable relationship. Based on the data collection of 122 samples, the performance of existing relations for estimating Cc, which includes soil parameter variables, was studied for the silty clay soil of Kerman, Iran.

The performance of a number of existing correlation equations for determining Cc was evaluated using the statistical criteria (RMSE, RI and RD), error estimation and ranking. Final evaluation was calculated with prioritization of the RD index, which takes into account the precision and accuracy of the estimate alike. The reason for this was that the RMSE does not take into account precision of the estimates and the RI index does not express equal effect for accuracy and precision of estimation. According the analytical results, the following conclusions can be drawn regarding the estimation of Cc for the soil in Kerman, Iran:

- The single-variable equation with the parameter e_0, M4:$C_c = 0.156\ (e_0 + 0.0107)$ with RD=0.164, provides the best estimate of Cc for the region's data among all RD values. With respect to this, it can be derived that the most valid equation of Cc is related to the initial void ratio of the soil.
- The other best performing equations based on the parameters are: (i) M1 with W_n, (RD = 0.941) and equation $C_c = 0.0.1w_n - 0.05$ (single-variable relationship), (ii) M17 with e_0 and LL (RD=3.47) and equation $C_c = 0.156 + 0.141e_0 + 0.00058LL$, as well as M18 with e_0, Gs and (RD = 2.704) with equation $C_c = 0.141G_s^{1.2}(\frac{1+e_0}{G_s})^{2.38}$ (two-variable relationship) and (iii) M20 with e_0, LL, W_n and (RD = 2.221) with $C_c = -0.404 + 0.341e_0 + 0.004LL + 0.006w_n$ (three-variable relationship).
- Among the best-performing relationships, the correlation equation M4: $C_c = 0.156\ (e_0 + 0.0107)$ is appropriate for all evaluation criteria (RMSE, K, RI and RD) used for predicting Cc in the area under study.

REFERENCES

Abu-Farsakh, M. Y., and Titi, H. H. (2004). "Assessment of direct cone penetration test method for predicting the ultimate capacity of friction driven piles." Journal of Geotechnical Engineering and Geo-environmental Engineering, 130 (9), 935-944.

Ahadiyan, J., Ebne, J.R., and Bajestan, M.S. (2008). "Prediction determination of soil compression index Cc in Ahwaz Region." J. Faculty Eng., 35 (3), 75-80. (In Persian).

Al-Khafaji, A.W.N., and Andersland, O.B. (1992). "Equations for compression index approximation." J. Geotech. Eng., ASCE, 118 (1), 148-153.

Al-Taei, A. J., Al-Bayati, A. F., and Taki, Z.N.M. (2017). "Compression index and compression ratio prediction by artificial neural networks." Journal of Engineering, 23 (12), 96-106.

Aminizadeh, B.M.R., Lashkaripour, Gh., Ghafoori, M., Hafezi Moghadas, N., and Bahrami, M. (2013). "Assessing sedimentary basin effect on compressibility of city of Kerman sediments using stress sensitivity and standard penetration test." Journal of Applied Environmental and Biological Sciences, 3 (10), 132-141.

Azzouz, A.S., Krizek, R.J., and Corotis, R.B. (1976). "Regression analysis of soil compressibility." Soils and Foundations, 16 (2), 19-29.

Bowles, J.E. (1979). "Physical and geotechnical properties of soils". 2nd Edn., McGraw-Hill Book Company, Inc., New York, 88-89.

Briaud, J.L., and Tucker, L.M. (1998). "Measured and predicted axial response of 98 piles". Journal of Geotechnical Engineering, 114 (9), 984-1001.

Das, B.M. (2010). "Principles of geotechnical engineering". 7th Edn., Cengage Learning, Stamford, USA, 635-639.

Giasi, C.I., Cherubini, C., and Paccapelo, F. (2003). "Evaluation of compression index of remolded clays by means of Atterberg limits." Bulletin of Engineering Geology and the Environment, 62 (4), 333-340.

Hamza, G., Canakci, H., and Alhashemy, A. (2016). "Use of ranking measure for performance assessment of correlations for the compression index." European Journal of Environmental and Civil Engineering,http://www.tandfonline.com/doi/abs/10.1080/19648189.2016.1210036

Han, J. (2015). "Principles and practice of ground improvement". John Wiley and Sons, Inc., Hoboken, New Jersey.

Hough, B.K. (1957). "Basic soils engineering". The Ronald Press Company, New York.

Kalantary, F., and Kordnaeij, A. (2012). "Prediction of compression index using artificial neural networks". Scientific Research and Essays, 7 (31), 2835-2848.

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

Le, T.M., Fatahi, B., and Khabbaz, H. (2015). "Analyzing consolidation data to obtain elastic viscous plastic parameters of clay". Geo-mechanics and Engineering, An International Journal, 8 (4), 559-594.

Mayne, P. (1980). "Cam-clay predictions of undrained strength". J. Geotech. Eng. Div., ASCE, 106 (11), 1219-1242.

Mohammadzadeh, D.S., Bolouri Bazaz, J., and Alavi Amir, H. (2014). "An evolutionary computational approach for formulation of compression index of fine-grained soils." Engineering Applications of Artificial Intelligence, 33, 58-68.

Nagaraj, T.S., and Srinivas Murthy, B.R. (1986). "Critical reappraisal of compression index equation". Geotchnique, 36 (1), 27-32.

Nishida, Y. (1956). "A brief note on compression index of soil". Journal of Soil Mechanics and Foundation Engineering, 82 (3), 1-14.

Onyejekwe, S., Kang, X., and Ge, L. (2015). "Assessment of empirical equations for the compression index of fine-grained soils in Missouri." Bulletin of Engineering Geology and the Environment, 74 (3), 705-716.

Orr, T.L.L., and Cherubini, C. (2003). "Use of ranking distance as an index for assessing accuracy and precision of equations for the bearing capacity of piles and at-rest earth pressure coefficient". Canadian Geotechnical Journal, 40 (6), 1200-1207.

Ozer, M., Isik, N.S., and Orhan, M. (2008). "Statistical and network assessment of the compression index of clay-bearing soils". Bulletin of Engineering Geology and the Environment, 67 (4), 537-545.

Park, H., and Lee, S.R. (2011). "Evaluation of compression index of soils using an artificial neural network." Computers and Geotechniques, 38 (4), 472-481.

Rendon-Herrero, O. (1983a). "Universal compression index equation discussion." J. Geotech. Eng. Div., ASCE, 109 (10), 1179-1200.

Rendon-Herrero, O. (1983b). "Universal compression index equation closure." J. Geotech. Eng. Div., ASCE, 109 (5), 755-761.

Skempton, A. W. (1994). "Notes on the compressibility of clays." Quarterly Journal of the Geological Society of London, 100 (2), 119-135.

Sridhrarn, A., and Nagaraj, H.B. (2000). "Compressibility behavior of remolded, fine-grained soils and correlation with index properties". Canadian Geotechnical Journal, 37 (3), 712-722.

Terzaghi, K., and Peck, R. (1967). "Soil mechanics in engineering practice". John Wiley and Sons, New York.

Tiwari, B., and Ajmera. (2012). "New correlation equations for compression index of remolded clays". J. Geotech. Geoenviron. Eng., 138 (6), 757-762.

Yoon, G.L., Kim, B.T., and Jeon, S.S. (2004). "Empirical correlations of compression index for marine clay from regression analysis". Canadian Geotechnical Journal, 41 (6), 1213-1221.