Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.jzus.zju.edu.cn; www.springerlink.com
E-mail: jzus@zju.edu.cn

EPR-RCGA-based modelling of compression index and RMSE-AIC-BIC-based model selection for Chinese marine clays and their engineering application*

Ze-xiang WU[1], Hui JI[†‡1], Chuang YU[2], Cheng ZHOU[3]

[1]*State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean, and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Department of Civil Engineering, Wenzhou University, Wenzhou 325035, China*

[3]*State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource and Hydropower Engineering, Sichuan University, Chengdu 610065, China*

[†]E-mail: jihui_v@hotmail.com

Received Feb. 20, 2017; Revision accepted May 24, 2017; Crosschecked Jan. 31, 2018

Abstract: The compression index is a key parameter in the field of soft clay engineering. In this paper, we propose an improved method for correlating the compression index with the physical properties of intact Chinese marine clays that are involved in many construction projects in coastal regions in China. First, the compression index and some common physical properties of clays from 21 regions along the Chinese coast are extracted from the literature. Then, a basic regression analysis for the compression index using the natural water content and Atterberg limits is conducted. To improve the correlation performance, an evolutionary polynomial regression (EPR) and real coded genetic algorithm (RCGA) combined technique is adopted to formulate different equations involving different numbers of variables. An optimal correlation using only natural water content and liquid limit as input parameters is finally selected according to the root mean square error (RMSE), Akaike's information criterion (AIC), and Bayesian information criterion (BIC). The proposed correlation is evaluated and shown to perform better than existing empirical correlations in predicting the compression index for all selected Chinese marine clays. This correlation is validated to be reliable and applicable to engineering applications through the prediction of the properties of an embankment on the southeast coast of China using finite element method. All comparisons show that the EPR and RCGA combined technique is powerful for correlating the compression index with the physical properties of the clay, and that model selection by RMSE, AIC, and BIC is effective. The proposed correlation could be used to update current formulations, and is applicable to engineering design in coastal regions of China.

Key words: Clay; Compressibility; Correlation index; Atterberg limits; Finite element; Embankment; Soft clay
https://doi.org/10.1631/jzus.A1700089 **CLC number:** TU433

1 Introduction

Natural clayey deposits up to several hundred meters thick are widely distributed in the coastal regions of China (Hong et al., 2010; Liu et al., 2011; Zhu et al., 2013; Wu HN et al., 2015; Yin et al., 2015). Most marine sediments in these regions have been deposited continuously due to the Eustachy transgression–regression cycle and coastline changes over thousands of years (Liu et al., 2011; Wu HN et al., 2015). For clayey soils, compressibility is an important characteristic in evaluating the settlement of the soil layer for any construction (Yao and Sun, 2000; Yin and Hicher, 2008; Yao et al., 2009; Yin et al.,

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 41002091)
ORCID: Ze-xiang WU, https://orcid.org/0000-0002-1143-1274

2009, 2010, 2011, 2013, 2014, 2015; Karstunen and Yin, 2010; Yin and Wang, 2012; Shen et al., 2014, 2015). Using traditional methods, a test for obtaining the compression index would take around two weeks (three days for saturation, eight days for loading from 12.5 kPa to 1600 kPa with 24 h for each loading, and a couple of days for unloading to 12.5 kPa). If possible, a reliable correlation between the compression index and the basic physical properties of the clay could reduce the test cost and save time.

To obtain the compression index rapidly and conveniently, many attempts have been made to correlate the compression index with some physical properties of soil, such as initial water content, liquid limit, initial void ratio, and plastic index (Azzouz et al., 1976; Wroth and Wood, 1978; Bowles, 1979; Carrier, 1985; Nagaraj and Murthy, 1986; Burland, 1990; Sridharan and Nagaraj, 2000; Nath and DeDalal, 2004; Yoon et al., 2004; Sridharan and Gurtug, 2005; Tiwari and Ajmera, 2012). For Chinese marine clays, correlations between the compression index and soil physical properties have been evaluated for some clays, e.g. Dalian clay (Li et al., 2016), Lianyungang clay (Liu et al., 2011; Hong et al., 2012), Jiangsu marine clays (Miao et al., 2007), Nanjing clay (Huang et al., 2011), Shanghai clay (Wei and Hu, 1980; Gao et al., 1986; Ng et al., 2011; Chen et al., 2013; Wu et al., 2014), Kemen clay (Hong et al., 2012), and Hong Kong marine clay (Yin, 1999, 2002). Up to now, a comprehensive and systematic study evaluating the compression index and physical properties for Chinese marine clays has not been reported.

Recently, soft computing techniques, such as artificial neural networks (ANNs), genetic programming (GP), and evolutionary polynomial regression (EPR), have been increasingly developed. For assessing the compressibility of soil, ANNs have been applied successfully to calculate the compression index of some soils (Ozer et al., 2008; Park and Lee, 2011). However, both ANNs and GP have inherent drawbacks for modelling complex nonlinear problems (Giustolisi and Savic, 2006; Rezania et al., 2010). Compared with other data-driven and classical regression techniques, EPR is a hybrid data-driven technique that attempts to overcome some of those shortcomings. Recently, EPR has been frequently and successfully used in geotechnical engineering (Rezania et al., 2010; Faramarzi et al., 2014). Therefore, the application of EPR to estimating the correlation between soil compressibility and physical properties may be worthwhile and give better results. Furthermore, advanced real coded genetic algorithm (RCGA) (Jin et al., 2016a, 2016b, 2017a, 2017b) may enhance the EPR-based correlation, and indexes like root mean square error (RMSE), Akaike's information criterion (AIC) (Akaike, 1998), and Bayesian information criterion (BIC) (Schwarz, 1978) might help the final selection of correlated equations. If all these techniques are combined, the correlation could be much improved and made more applicable to engineering applications.

Therefore, in this study, the compression index and physical properties of clays from 21 regions along the Chinese coast are first sourced from the literature. Then, a basic regression analysis for the compression index using the natural water content, liquid limit, and plastic index is conducted. To improve the correlation performance, the EPR-RCGA combined technique is adopted to formulate equations involving different numbers of variables. An optimal correlation is selected according to RMSE, AIC, and BIC, which is then compared with existing empirical correlations for its accuracy in predicting the compression index for all selected Chinese marine clays. Finally, an application of the suggested correlation to the prediction of the compression index of an embankment on the southeast coast of China is performed to confirm its reliability and accuracy.

2 Database of Chinese marine clays

2.1 Description of database

To obtain a correlation between compressibility and soil properties which is valid for all Chinese marine clays, the database must cover a sufficiently wide range of clays. Along the Chinese coast, 21 regions distributed uniformly from north to south were selected for sampling of marine clays. These clays can be considered typical of marine clays in China, and their compressibility reflects that of all Chinese marine clays. Note that all the selected clay samples were natural intact clays. For each selected marine clay, the natural water content w_n, initial void ratio e_0, liquid limit w_L, plastic limit w_p, plastic index I_p, and compression index ($C_c=\Delta e/\Delta \log p'$, where Δe is the change of void ratio, and p' is the mean effective stress) in a semi-logarithmic plot were collected to establish the database. All the data used in this study were obtained from previous studies (Wei and Hu, 1980; Shen et al., 2005, 2014; Miao et al., 2007; Liu et al., 2011; Ng et al., 2011; Hong et al., 2012; Chen et al., 2013; Zhu et al., 2013; Wu et al., 2014; Li et al., 2016). To obtain reliable correlations, data from 108 samples of Chinese marine clays were selected. For some clays, samples were taken from different depths and thus had variable properties for the same location. For physical properties, the data presented in the literature were entered directly into the database. For the compression index, the data were obtained from the compression curves. All the data were selected from well-documented articles. Note that clays in the crust layer, usually with a high over-consolidated ratio due to drying-wetting or human activities, were not considered in this study, since the natural water content of these over-consolidated clays is much lower than that of clays below the crust layer. Thus, in practice, an accurate compression index of the crust layer should be obtained separately by conducting a compression test.

The selected clays were classified by plotting their plasticity (Fig. 1). To assess the adequacy of the database, the statistics of properties for all selected samples were calculated based on all the selected data. For each investigated variable, a histogram, a probability distribution curve, and maximum, minimum, mean μ, and standard deviation σ are shown in Fig. 2.

2.2 Basic regression analyses

The correlations between C_c and w_n, w_L, and I_p were first estimated. Note that for the Atterberg limits, selecting the liquid limit and plastic limit or liquid limit and plastic index is physically the same for estimating the correlation. To propose a correlation with good applicability in engineering practice for C_c, w_n

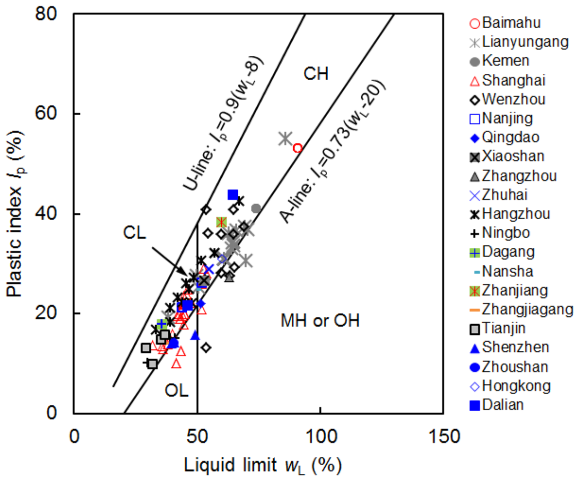

Fig. 1 Plasticity of the selected Chinese marine clays

CL: low plasticity inorganic clays; CH: high plasticity inorganic clays; OH: opaque clays; MH: silty clays; OL: low plasticity organic silty clays; U-line: upper bound for general soils; A-line: the boundary to separate the claylike materials from silty materials

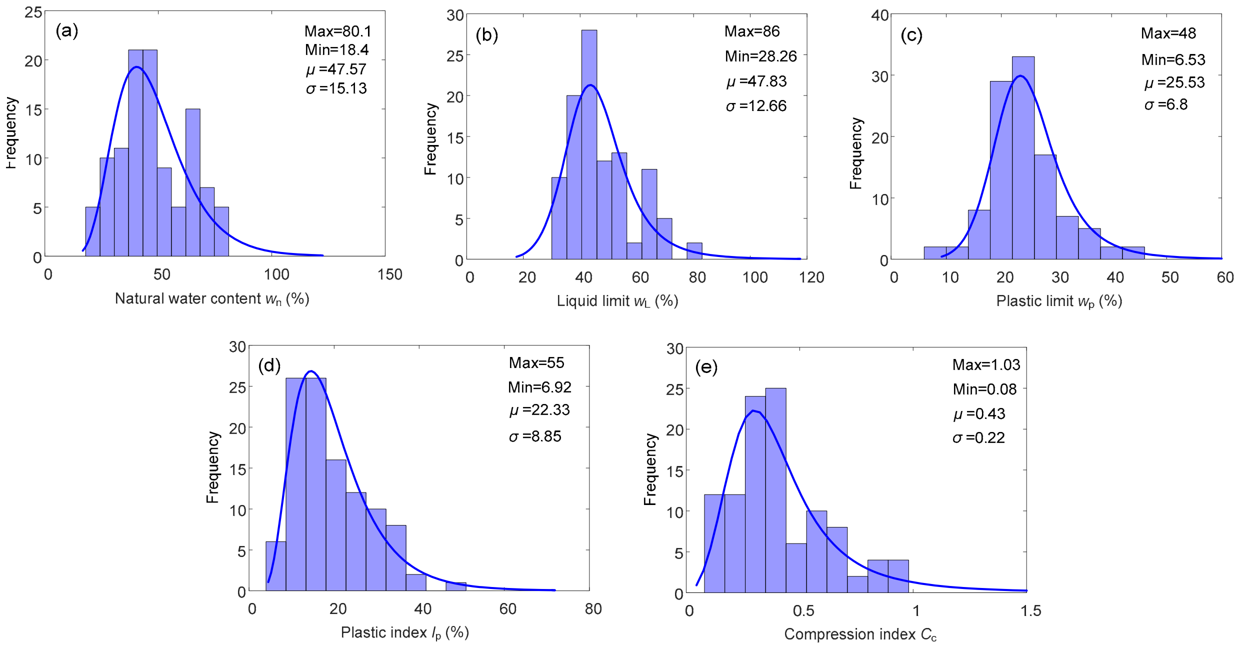

Fig. 2 Histograms of selected physical properties: (a) w_n; (b) w_L; (c) w_p; (d) I_p; (e) C_c

was selected to represent the in-situ deposit environment and geology. The initial void ratio e_0 was not selected as a correlated variable since e_0 is physically equivalent to w_n, but experimentally less convenient. For each parameter, the correlations between the compression index and each physical property are shown in Fig. 3. The compression index of natural intact samples shows a remarkable linear correlation with w_n, and a poor correlation with the Atterberg limits, as reported by Ng et al. (2011) and Wu CJ et al. (2014, 2015) for Shanghai clay. However, the correlation coefficient R^2 even for the best fit (R^2<0.74) still needs to be improved if the method is to be adopted in engineering practice.

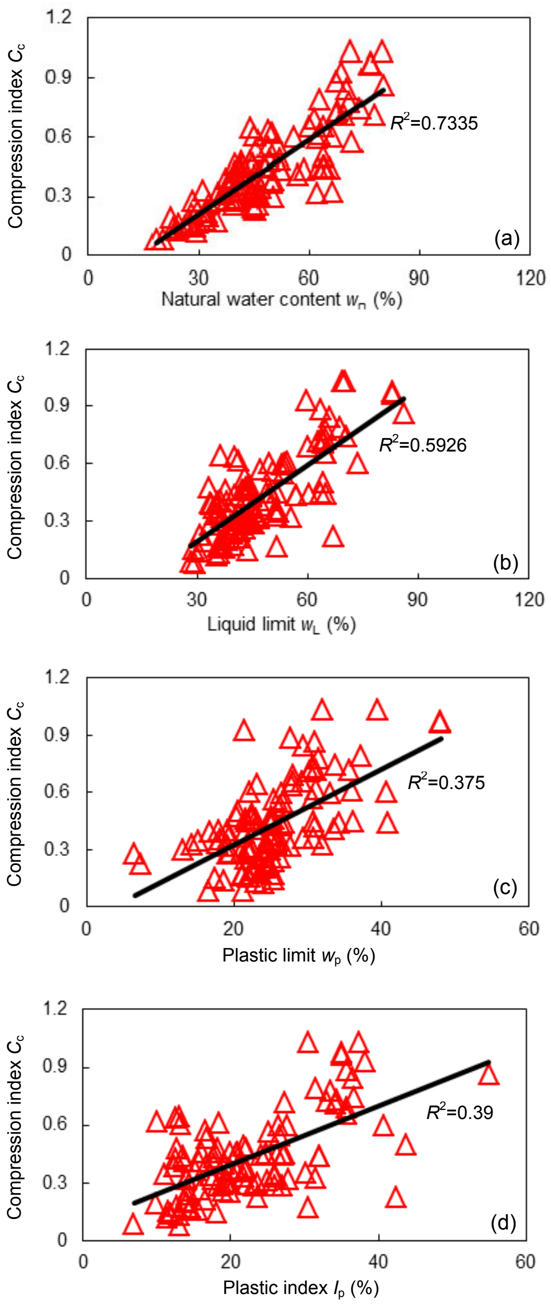

Fig. 3 Linear correlations between the compression index and each physical property of Chinese marine clays: (a) w_n; (b) w_L; (c) w_p; (d) I_p

For two or three parameters, the correlations with their R^2 values are summarized and plotted in Fig. 4. For two parameters, the correlation coefficient shows a slight increase to 0.75. The combination of w_L and I_p gives an R^2 value of <0.7, indicating the necessity for w_n.

Overall, all the correlations based on basic regression analysis were unsatisfactory. Therefore, the use of the EPR method using genetic optimization was investigated.

3 EPR-RCGA-based correlation

3.1 Evolutionary polynomial regression technique

The EPR is generally expressed as

$$\boldsymbol{y}=\sum_{j=1}^{m}F\left(\boldsymbol{X}, f(\boldsymbol{X}), \boldsymbol{a}_j\right)+a_0, \tag{1}$$

where $\boldsymbol{y}$ is the estimated vector of the output, a_0 and $\boldsymbol{a}$ are parameters, F is a given function, $\boldsymbol{X}$ is input variables, f is a given function, and m is the number. This equation can be transformed to

$$\begin{aligned}\boldsymbol{Y}_{N\times1}(\boldsymbol{\theta},\boldsymbol{Z})&=\left[\boldsymbol{I}_{N\times1}\quad \boldsymbol{Z}^j_{N\times m}\right]\times\left[a_0\quad a_1\quad \dots\quad a_m\right]^{\mathrm{T}}\\&=\boldsymbol{Z}_{N\times d}\times\boldsymbol{\theta}^{\mathrm{T}}_{d\times1},\end{aligned} \tag{2}$$

where $\boldsymbol{Y}_{N\times1}(\boldsymbol{\theta}, \boldsymbol{Z})$ is the least-squares vector, $\boldsymbol{\theta}_{d\times1}$ is the vector of d, and $\boldsymbol{Z}_{N\times d}$ is a matrix formed by $\boldsymbol{I}$, unitary column vector for bias a_0, and m vectors of variables $\boldsymbol{Z}^j$ that for a fixed j are a product of the independent predictor vectors of variables/inputs, $\boldsymbol{X}$=<$\boldsymbol{X}_1$, $\boldsymbol{X}_2$, …, $\boldsymbol{X}_k$>.

In this study, the EPR procedure combined with the RCGA-based optimization method was used to investigate the correlation between the compression index and soil physical properties of natural intact Chinese marine clays. A flow chart of the method is shown in Fig. 5.

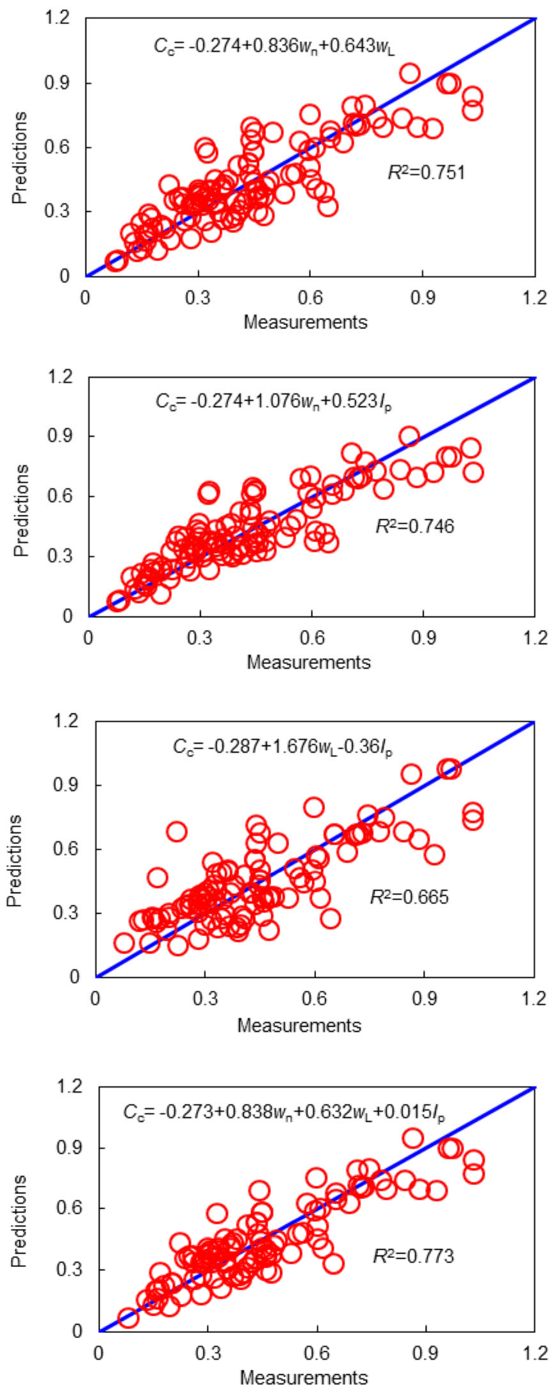

Fig. 4 Correlations between the compression index and two or three combined physical properties of Chinese marine clays

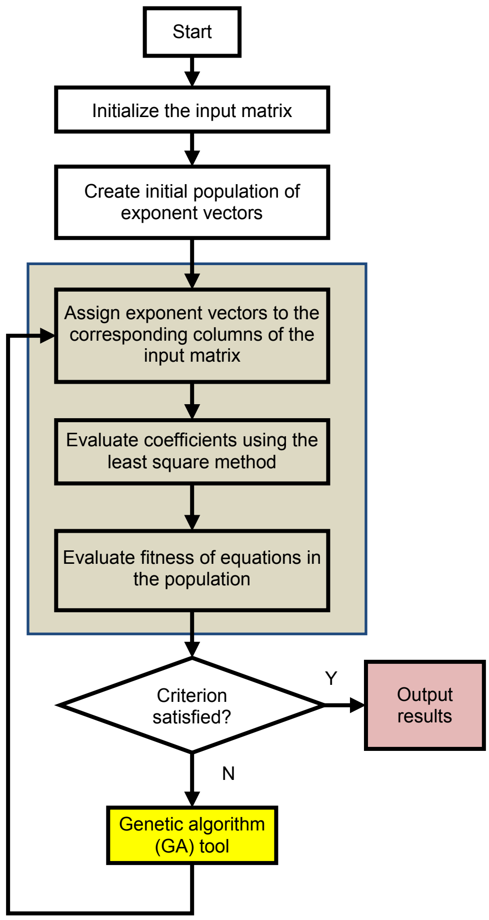

Fig. 5 Flow chart of the proposed optimization based EPR process

3.2 EPR-RCGA-based correlation construction

Based on the values of correlation coefficients shown in Figs. 2 and 3, the natural water content w_n, liquid limit w_L, and plastic index I_p were the most common properties, so they were selected as the correlated variables in the EPR-RCGA-based correlation. The general structure of the EPR-RCGA-based model can be expressed as

$$C_c = f(w_n, w_L, I_P) + a_0, \tag{3}$$

where C_c is the compression index measured from the e-logp' plane (void ratio e vs. mean effective stress p') for natural intact clays, and a_0 is a constant.

The whole database was used for training the EPR-RCGA models. Simply, all exponents were set to [−2, 2] with a maximum of five terms. During the evolutionary process, the level of accuracy was evaluated based on the coefficient of determination (COD) as the fitness function:

$$\mathrm{COD}=\left(1-\frac{\sum_{i=1}^{N}\left(\boldsymbol{Y}_{\mathrm{a}}-\boldsymbol{Y}_{\mathrm{p}}\right)^{2}}{\sum_{i=1}^{N}\left(\boldsymbol{Y}_{\mathrm{a}}-\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{Y}_{\mathrm{a}}\right)^{2}}\right)\times 100\%, \quad (4)$$

where N is the number of data points on which the COD is computed, $\boldsymbol{Y}_a$ is the actual output value, and $\boldsymbol{Y}_p$ is the predicted value. The initial population was set to 50 for all EPR procedures. The population was generated by the uniform random sequence generator SOBOL (Poles et al., 2009). Multiple runs were performed and the analysis was repeated with various combinations in GA (Rezania et al., 2010; Jin et al., 2016a, 2016b, 2017a, 2017b) to obtain the most suitable form for the EPR-RCGA model.

To obtain a robust and reliable correlation for the compressibility of natural intact Chinese marine clays, EPR-RCGA-based correlations involving different kinds and numbers of selected variables (physical properties) were constructed successively. The EPR-RCGA-based correlations were divided into two groups: (1) those involving two selected variables, and (2) those involving three selected variables. A total of three EPR-RCGA-based correlations involving two selected variables, and one involving three selected variables were constructed. The optimal EPR-RCGA-based correlations along with their correlation coefficients R^2 for different situations are summarized in Table 1 for comparison.

The performance of the optimized correlations improved slightly with an increasing number of variables. Moreover, the natural water content w_n and liquid limit w_L had a more significant effect on the compressibility of Chinese marine clays. Note that other physical properties (e.g. the void ratio at liquid limit e_L, the specific gravity G_s, the activity A, the plastic limit w_p, and the shrinkage index I_S) involved in Eq. (3) could theoretically improve the performance of the correlation. Since the database was lacking these physical properties, only w_n, w_L, and I_p were selected as variables for correlation.

3.3 Evaluation of ERP-RCGA-based correlations

3.3.1 Evaluation by statistical data

For EPR-RCGA-based correlations for predicting the compression index of natural intact Chinese marine clays, the following error function was proposed:

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{Y}_{\mathrm{a}}-\boldsymbol{Y}_{\mathrm{p}}\right)^{2}}. \quad (5)$$

At the same time, the mean μ and the standard deviation σ of $\boldsymbol{Y}_p/\boldsymbol{Y}_a$ were also calculated. A μ value greater than 1.0 indicates over-estimation and other values under-estimation. The best correlation was represented by a μ value close to 1.0 and σ value close to 0.

Table 2 presents the error index, μ and σ of $\boldsymbol{Y}_p/\boldsymbol{Y}_a$ of optimized correlations for natural intact Chinese marine clays. Considering the values of RMSE, μ, and σ, the optimized correlation involving w_n and w_L gave more accurate predictions than others. Similarly, the optimized correlation involving w_n, w_L, and I_p gave

Table 1 Optimal correlation equations with correlation coefficient values for Chinese marine clays

Number of variables	Combination of variables	Correlation expression	R^2
2	w_n, w_L	$C_c=-0.366-4.9\frac{w_n^2}{w_L}+5.53w_n^2w_L+10.2w_nw_L+1.36\frac{w_n^2}{w_L^2}-10.66w_nw_L^2$	0.872
2	w_n, I_p	$C_c=-0.13-12.1I_p^2+0.027\frac{w_n^2}{I_p^2}+10.85w_nI_p-1.48w_n^2+2.01\frac{I_p^2}{w_n}$	0.861
2	w_L, I_p	$C_c=-0.714-0.26\frac{w_L}{I_p}+0.33\frac{I_p}{w_L^2}+0.08\frac{w_L^2}{I_p^2}+9.11I_p-9.9\frac{I_p^2}{w_L}$	0.657
3	w_n, w_L, I_p	$C_c=-0.218+1.34\frac{w_nI_p^2}{w_L^2}+1.86w_L^2-8.67w_L^2I_p^2+0.028\frac{w_n^2}{w_L^2I_p}+6.92w_n^2I_p^2w_L$	0.876

Note: w_n, w_L, and I_p are real numbers, not percentages reliable predictions of the compression index with a small scatter (Fig. 6).

Table 2 Error index, μ and σ of Y_a/Y_p, AIC, and BIC when using different optimal correlations for Chinese marine clays

Variable	RMSE	μ	σ	AIC	BIC
w_n, w_L	0.0777	1.0393	0.2332	−48.0	−486.6
w_n, I_p	0.0809	1.0576	0.2590	−40.2	−478.8
w_L, I_p	0.1218	1.1481	0.5297	39.3	−399.3
w_n, w_L, I_p	0.0764	1.0481	0.2391	−49.2	−485.3

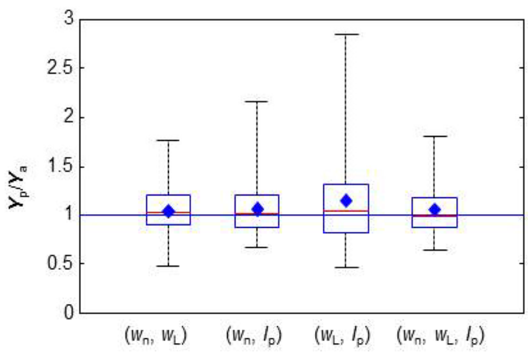

Fig. 6 Box plot of the ratio Y_p/Y_a of the compression index of different optimized correlations for Chinese marine clays

3.3.2 Evaluation based on two information criteria

Note that each EPR-RCGA-based correlation has a different number of parameters. The evaluation of EPR-RCGA-based correlations should consider the number of parameters in addition to their predictive performance. To assess how well a model explains the data, two widely used criteria can be adopted: (a) Akaike's information criterion (AIC) and (b) Schwartz's Bayesian information criterion (BIC).

Akaike (1998) proposed AIC as a measure of model quality which can be expressed as

$$\mathrm{AIC} = -2\log L(\hat{\theta}) + 2k, \tag{6}$$

where θ is the set of parameters of the model, $L(\hat{\theta})$ is the likelihood of the candidate model given the data when evaluated at the maximum likelihood estimate of θ, and k is the number of estimated parameters in the candidate model. The AIC can be expressed equivalently as

$$\mathrm{AIC} = n\cdot\log(\mathrm{RSS}) + 2k, \tag{7}$$

where RSS is the residual sum of squares, $\mathrm{RSS}=\sum_{i=1}^{n}(U_{\exp}^{i}-U_{\mathrm{num}}^{i})^2$, $U_{\exp}^{i}$ is the experimental value corresponding to point i, U_{num}^{i} is the numerical value corresponding to point i, and n is the number of values in the estimation dataset.

Similar to AIC, the BIC is computed according to Schwarz (1978) as

$$\mathrm{BIC} = n\cdot\log(\mathrm{RSS}/n) + k\cdot\log n. \tag{8}$$

The best model is the one that provides the minimum values of AIC and BIC. To select the most "appropriate" EPR-RCGA-based model, the values of AIC and BIC for each EPR-RCGA-based model were calculated (Table 2). Based on those values, we conclude that the model involving only w_n and w_L was the most "appropriate".

Therefore, based on the above evaluation, the optimized correlation involving only w_n and w_L was selected as the best correlation of compression index for natural intact Chinese marine clays, as expressed by

$$\begin{aligned} C_c &= -0.366 - 4.9\frac{w_n^2}{w_L} + 5.53w_n^2 w_L \\ &\quad + 10.2w_n w_L + 1.36\frac{w_n^2}{w_L^2} - 10.66w_n w_L^2. \end{aligned} \tag{9}$$

Fig. 7 shows a good agreement between measurements and predictions for the suggested optimized correlation.

3.3.3 Comparison with existing empirical correlations

To highlight the performance of EPR-RCGA-based correlations of the compression index of Chinese marine clays, the predictions given by the suggested EPR-RCGA-based correlation were compared with those predicted by existing empirical correlations. Due to the limitations of available data, correlations involving void ratio at liquid limit e_L, specific gravity G_s, activity A, and shrinkage index I_S were not considered in this study. For a fair comparison, empirical correlations involving the initial void ratio e_0 also were not compared in this study. Various correlations previously proposed involving w_n and the Atterberg limits (w_L and I_p) were selected for comparison (Table 3).

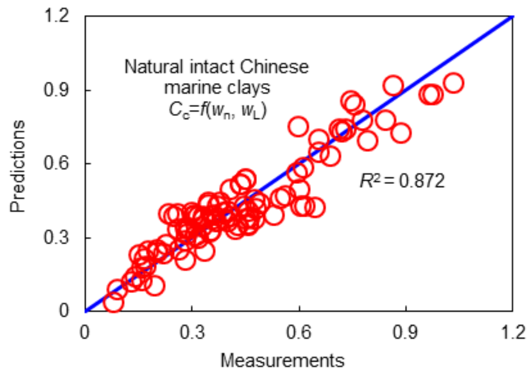

Fig. 7 Predictions compared to measurements for different suggested correlations

The correlation coefficient R^2, the value of the error index, μ and σ of $\boldsymbol{Y}_a/\boldsymbol{Y}_p$ of all selected correlations for natural intact Chinese marine clays were calculated and are plotted in Fig. 8. The results indicate that the EPR-RCGA-based correlation performed better for each indicator than empirical correlations. That is, the comparisons show that the EPR-RCGA-based correlation is more powerful than empirical methods for predicting compressibility.

Overall, the reliability and accuracy of the EPR-RCGA-based correlation in representing the compressibility of natural intact Chinese marine clays were better than those of empirical correlations.

4 Model validation

To evaluate the accuracy and reliability of the suggested EPR-RCGA-based correlation in engineering practice, an embankment constructed on soft clay deposits in a coastal area of southeast China was selected and simulated. Predictions of the vertical settlement, horizontal displacement, and the generation of excess pore pressure were compared with field measurements.

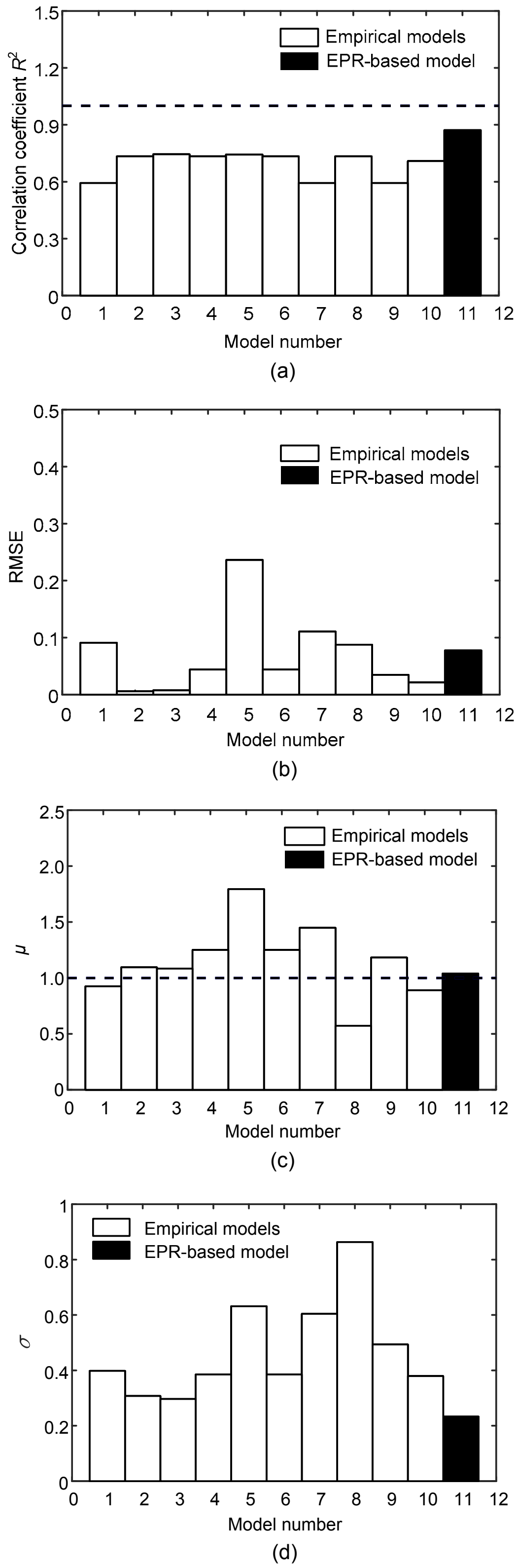

Fig. 8 Correlation coefficient R^2 (a), RMSE (b), mean value μ (c), and standard deviation σ (d) of Y_a/Y_p when using different correlations for Chinese marine clays

Table 3 Selected empirical correlations of the compression index for natural clays

Number	Equation	Applicability	Reference
1	$C_c=0.01(w_n-5)$	All clays	Azzouz et al., 1976
2	$C_c=0.009w_n+0.002w_L-0.1$	All clays	Azzouz et al., 1976
3	$C_c=0.01w_n$	All clays	Koppula, 1981
4	$C_c=0.009w_n+0.005w_L$	All clays	Koppula, 1981
5	$C_c=0.01(w_n-7.549)$	All clays	Herrero, 1983
6	$C_c=0.009(w_L-10)$	Normal consolidated clays	Terzaghi et al., 1996
7	$C_c=0.016(w_L-14)$	Liangyungang clay	Liu et al., 2011
8	$C_c=-0.704+0.022w_n$	Jiangsu marine clays	Miao et al., 2007
9	$C_c=0.018(w_L-14)$	Shanghai clay	Gao et al., 1986
10	$C_c=0.00426[\exp(0.0444w_n)-0.794]$	Shanghai clay	Wu et al., 2014a, 2014b
11	EPR-based correlation	Chinese marine clays	This study

4.1 Adopted embankment and finite element model

The adopted embankment was an instrumented test embankment on soft clay deposits, which had been the subject of previous studies (Shen et al., 2005). The subsoil was divided into six layers, and the geometry of the finite element analysis mesh is shown in Fig. 9. The physical properties of the soil in each layer are shown Fig. 10. The water table was at a depth of 1.5 m. The model was symmetrical, 120 m horizontally and 25 m vertically, and under plane strain condition. Two lateral boundaries were restrained in the horizontal direction, and the bottom boundary was restrained laterally and vertically. The mesh comprised 1720 four-node elements, resulting in 1815 nodes. The embankment fill was simulated using a linear elastic model with Young's modulus E=25 000 kPa, Poisson's ratio ν=0.25, and unit weight γ=20 kN/m^3. The loading was applied by increasing the unit weight of the elements of the embankment fill (4.66 m high) over about four months (Shen et al., 2005).

To simulate the behaviour of the subsoils, a modified cam-clay (MCC) model was employed. The parameters of the MCC model for layers 1–5 were the same as those of Shen et al. (2005), except for the compression index λ and the swelling index κ (in MCC, $\lambda=C_c/\ln 10$ and the swelling index κ was set to one-tenth of λ). The compression index of each soil layer was estimated by the suggested EPR-RCGA-based correlation using w_n and w_L (Fig. 12). Note that in the crust layer the correlated compression index

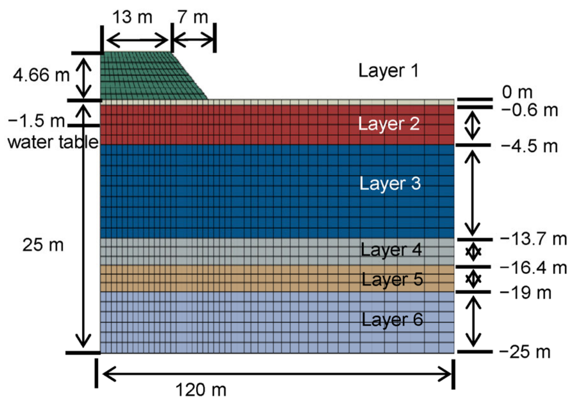

Fig. 9 Finite element model of the embankment

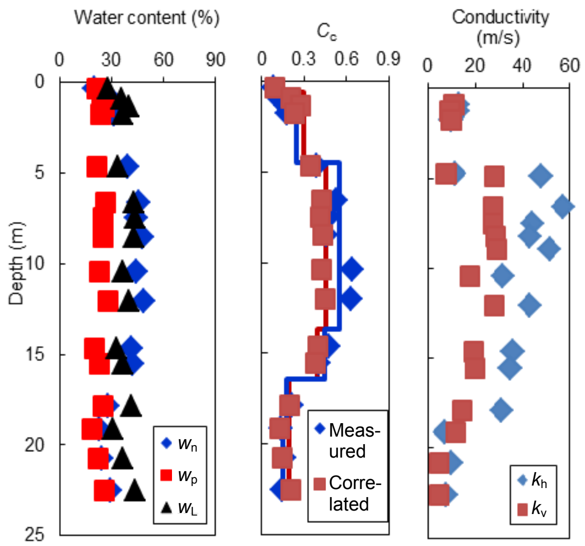

Fig. 10 Soil properties throughout the depth of the embankment

k_h: horizontal permeability; k_v: vertical permeability was directly adopted even though test data of the crust layer soils were not included during the correlation. This was unlikely to have influenced the calculated results since the crust layer was very thin and highly over-consolidated. The soil in layer 6 was simulated using the Mohr-Coulomb (MC) model with Young's modulus E=25 000 kPa, Poisson's ratio ν=0.25, friction angle ϕ=35°, and dilation angle ψ=0°, which is more reasonable than the linear elastic model of Shen et al. (2005). For hydraulic conductivity, the vertical and horizontal nonlinear hydraulic conductivities used by Shen et al. (2005) were adopted.

For comparison, the simulation was also conducted using the measured compression index according to the profile in Fig. 11, which was more precise than that of Shen et al. (2005). All parameters used for MCC and MC are summarized in Table 4.

4.2 Comparisons with field measurements

Fig. 12 presents the predicted and measured surface settlements under the centerline and shoulder of the embankment. Compared to field measurements, both predictions were acceptable, with only slight differences between predictions using measured and correlated C_c values. Fig. 13 shows the predicted and measured horizontal displacements at the end of construction and 160 d after construction. Predictions obtained using correlated C_c values were more satisfactory than those using measured C_c values. Fig. 14 shows the predicted and measured excess water pore pressure values at different locations under the centerline of the embankment. All the predictions captured well the generation and dissipation processes of excess water pore pressure. The difference between the two different predictions was slight.

Overall, the use of correlated C_c gave an acceptable performance in simulating the embankment on soft clay deposits, showing that the suggested EPR-RCGA-based correlation of C_c for Chinese marine clays is reliable.

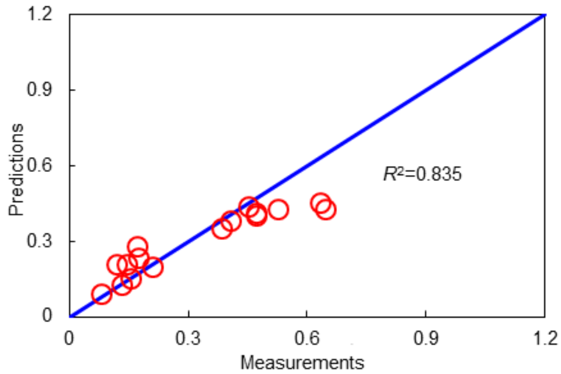

Fig. 11 Values of the compression index predicted by the optimal correlation compared with measured values

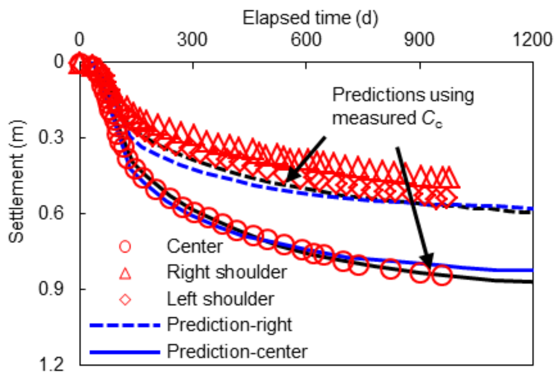

Fig. 12 Measured and calculated values of surface settlements of the embankment

Table 4 Model parameters for embankment fill and subsoils

Layer	ν	κ (me)	κ (co)	λ (me)	λ (co)	M	e_0	Γ	γ (kN/m^3)	k_h (×10^{-3} m/d)	k_v (×10^{-3} m/d)
Fill	0.25	–	–	–	–	–	–	–	20.0	193.50	193.50
1	0.20	0.004	0.008	0.044	0.089	1.0	0.81	0.98	19.3	2.72	2.72
2	0.25	0.011	0.015	0.107	0.130	1.0	1.07	1.22	18.5	0.56	0.22
3	0.25	0.024	0.020	0.238	0.197	0.8	1.36	2.22	17.3	2.54	1.69
4	0.20	0.019	0.017	0.192	0.172	0.8	1.10	1.98	17.9	2.06	1.01
5	0.20	0.080	0.084	0.076	0.084	1.0	0.81	0.99	19.5	0.39	0.18
6	0.25	–	–	–	–	–	–	–	–	25.90	25.90

me: measured value; co: correlated value predicted by EPR; M: slope of critical state line; Γ: specific volume

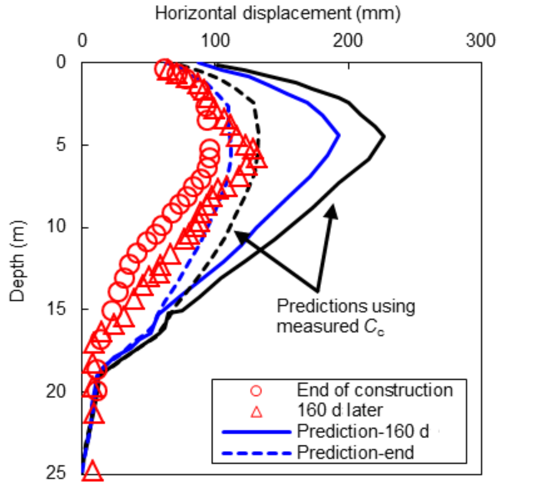

Fig. 13 Measured and calculated values of horizontal displacements at different time points

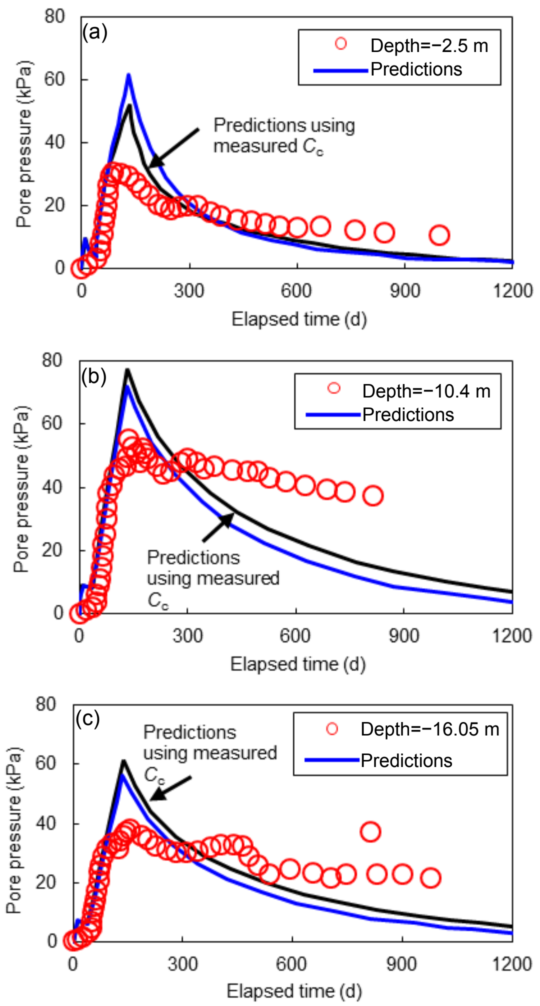

Fig. 14 Measured and calculated evolution of excess pore water pressures in subsoils: (a) depth=−2.5 m; (b) depth=−10.4 m; (c) depth=−16.05 m

5 Conclusions

The correlation of the compression index with the physical properties of Chinese marine clays was investigated through an EPR-RCGA-RMSE-AIC-BIC combined technique. A database containing the compression index and soil physical properties of Chinese marine clays from 21 regions along China coast was established.

First, C_c was correlated with physical properties (w_n, w_L, and I_p) using basic regression analyses. The C_c was found to be highly correlated with w_n as a single parameter. Then, for two parameters, the use of w_n and w_L gave better correlations. By changing from two to three parameters, the correlation was slightly enhanced, but was still not satisfactory.

The correlations between the compression index and different combinations of w_n, w_L, and I_p were then investigated using the EPR procedure. The results indicate that the performance of EPR-RCGA-based correlations is slightly improved by increasing the number of variables involved, and that correlations involving w_n and w_L may give a better performance than others.

For selecting the most "appropriate" EPR-RCGA-based correlation, the performance of each correlation was first evaluated using the RMSE index, the mean value and the standard deviation value of $\boldsymbol{Y}_p/\boldsymbol{Y}_a$. Then the performance was estimated using two evaluation criteria (AIC and BIC). Based on these, the EPR-RCGA-based correlation involving only w_n and w_L was selected as the most "appropriate" correlation of C_c for Chinese marine clays.

Then, the reliability of the suggested correlation was tested by comparing it with existing empirical correlations for Chinese marine clays. The correlated compression index was used to simulate an embankment on the southeast coast of China. Compared to field measurements and simulation results based on compression index measurements, good agreement was achieved for surface settlement of the embankment, horizontal displacement of subsoils, and excess water pore pressure in soils. This demonstrates that the suggested correlation of C_c for Chinese marine clays is reliable and applicable.

In the future, the proposed correlation of C_c for Chinese marine clays should be applied to real constructions. The results may help improve the Chinese design code.

References

Akaike H, 1998. Information theory and an extension of the maximum likelihood principle. Selected Papers of Hirotugu Akaike. Springer, p.199-213.
https://doi.org/10.1007/978-1-4612-1694-0_15

Azzouz AS, Krizek RJ, Corotis RB, 1976. Regression analysis of soil compressibility. *Soils and Foundations*, 16(2):19-29.
https://doi.org/10.3208/sandf1972.16.2_19

Bowles JE, 1979. Physical and Geotechnical Properties of Soils. McGraw Hill, New York, USA.

Burland J, 1990. On the compressibility and shear strength of natural clays. *Géotechnique*, 40(3):329-378.
https://doi.org/10.1680/geot.1990.40.3.329

Carrier WD, 1985. Consolidation parameters derived from index tests. *Géotechnique*, 35(2):211-213.
https://doi.org/10.1680/geot.1985.35.2.211

Chen B, Sun DA, Lu HB, 2013. Experimental study of compression behavior of marine soft clays. *Rock and Soil Mechanics*, 34(2):381-388 (in Chinese).

Faramarzi A, Alani AM, Javadi AA, 2014. An EPR-based self-learning approach to material modelling. *Computers & Structures*, 137:63-71.
https://doi.org/10.1016/j.compstruc.2013.06.012

Gao DZ, Wei DD, Hu ZX, 1986. Geotechnical properties of Shanghai soils and engineering applications. Marine Geotechnology and Nearshore/Offshore Structures. ASTM International, USA.

Giustolisi O, Savic DA, 2006. A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics*, 8(3):207-222.
https://doi.org/10.2166/hydro.2006.020

Herrero AA, 1983. End-product inhibition in anaerobic fermentations. *Trends in Biotechnology*, 1(2):49-53.
https://doi.org/10.1016/0167-7799(83)90069-0

Hong ZS, Yin J, Cui YJ, 2010. Compression behaviour of reconstituted soils at high initial water contents. *Géotechnique*, 60(9):691-700.
https://doi.org/10.1680/geot.09.P.059

Hong ZS, Zeng LL, Cui YJ, et al., 2012. Compression behaviour of natural and reconstituted clays. *Géotechnique*, 62(4):291-301.
https://doi.org/10.1680/geot.10.P.046

Huang MS, Liu YH, Sheng DC, 2010. Simulation of yielding and stress–stain behavior of Shanghai soft clay. *Computers and Geotechnics*, 38(3):341-353.
https://doi.org/10.1016/j.compgeo.2010.12.005

Jin YF, Yin ZY, Shen SL, et al., 2016a. Investigation into MOGA for identifying parameters of a critical state based sand model and parameters correlation by factor analysis. *Acta Geotechnica*, 11(5):1131-1145.
https://doi.org/10.1007/s11440-015-0425-5

Jin YF, Yin ZY, Shen SL, et al., 2016b. Selection of sand models and identification of parameters using an enhanced genetic algorithm. *International Journal for Numerical and Analytical Methods in Geomechanics*, 40(8): 1219-1240.
https://doi.org/10.1002/nag.2487

Jin YF, Yin ZY, Riou Y, et al., 2017a. Identifying creep and destructuration related soil parameters by optimization methods. *KSCE Journal of Civil Engineering*, 21(4): 1123-1134.
https://doi.org/10.1007/s12205-016-0378-8

Jin YF, Yin ZY, Shen SL, et al., 2017b. A new hybrid real-coded genetic algorithm and its application to parameters identification of soils. *Inverse Problems in Science and Engineering*, 25(9):1343-1366.
https://doi.org/10.1080/17415977.2016.1259315

Karstunen M, Yin ZY, 2010. Modelling time-dependent behaviour of Murro test embankment. *Géotechnique*, 60(10): 735-749.
https://doi.org/10.1680/geot.8.P.027

Koppula S, 1981. Statistical estimation of compression index. *Geotechnical Testing Journal*, 4(2):68-73.
https://doi.org/10.1520/GTJ10768J

Li G, Zhang JL, Yang Q, 2016. Geotechnical investigations at the Dalian offshore airport, China. *Marine Georesources & Geotechnology*, 34(8):747-758.
https://doi.org/10.1080/1064119X.2015.1087438

Liu S, Shao G, Du Y, et al., 2011. Depositional and geotechnical properties of marine clays in Lianyungang, China. *Engineering Geology*, 121(1-2):66-74.
https://doi.org/10.1016/j.enggeo.2011.04.014

Miao LC, Zhang JH, Chen YN, 2007. Study on compressibility of Jiangsu marine clay. *Chinese Journal of Geotechnical Engineering*, 29(11):1711-1714 (in Chinese).

Nagaraj T, Murthy BS, 1986. A critical reappraisal of compression index equations. *Géotechnique*, 36(1):27-32.
https://doi.org/10.1680/geot.1986.36.1.27

Nath A, DeDalal SS, 2004. The role of plasticity index in predicting compression behaviour of clays. *Electronic Journal of Geotechnical Engineering*, 9:1-7.

Ng CW, Li Q, Liu GB, 2011. Characteristics of one-dimensional compressibility of Shanghai clay. *Chinese Journal of Geotechnical Engineering*, 33(4):630-636 (in Chinese).

Ozer M, Isik NS, Orhan M, 2008. Statistical and neural network assessment of the compression index of clay-bearing soils. *Bulletin of Engineering Geology and the Environment*, 67(4):537-545.
https://doi.org/10.1007/s10064-008-0168-8

Park HI, Lee SR, 2011. Evaluation of the compression index of soils using an artificial neural network. *Computers and Geotechnics*, 38(4):472-481.
https://doi.org/10.1016/j.compgeo.2011.02.011

Poles S, Fu Y, Rigoni E, 2009. The effect of initial population sampling on the convergence of multi-objective genetic

algorithms. Multiobjective Programming and Goal Programming. Springer, p.123-133.
Rezania M, Javadi AA, Giustolisi O, 2010. Evaluation of liquefaction potential based on CPT results using evolutionary polynomial regression. *Computers and Geotechnics*, 37(1-2):82-92.
https://doi.org/10.1016/j.compgeo.2009.07.006
Schwarz G, 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461-464.
https://doi.org/10.1214/aos/1176344136
Shen SL, Chai JC, Hong ZS, et al., 2005. Analysis of field performance of embankments on soft clay deposit with and without PVD-improvement. *Geotextiles and Geomembranes*, 23(6):463-485.
https://doi.org/10.1016/j.geotexmem.2005.05.002
Shen SL, Wu HN, Cui YJ, et al., 2014. Long-term settlement behaviour of metro tunnels in the soft deposits of Shanghai. *Tunnelling and Underground Space Technology*, 40:309-323.
https://doi.org/10.1016/j.tust.2013.10.013
Shen SL, Wang JP, Wu HN, et al., 2015. Evaluation of hydraulic conductivity for both marine and deltaic deposits based on piezocone testing. *Ocean Engineering*, 110:174-182.
https://doi.org/10.1016/j.oceaneng.2015.10.011
Sridharan A, Nagaraj H, 2000. Compressibility behaviour of remoulded, fine-grained soils and correlation with index properties. *Canadian Geotechnical Journal*, 37(3):712-722.
https://doi.org/10.1139/t99-128
Sridharan A, Gurtug Y, 2005. Compressibility characteristics of soils. *Geotechnical & Geological Engineering*, 23(5):615-634.
https://doi.org/10.1007/s10706-004-9112-2
Terzaghi K, Peck RB, Mesri G, 1996. Soil Mechanics in Engineering Practice. John Wiley & Sons.
Tiwari B, Ajmera B, 2012. New correlation equations for compression index of remolded clays. *Journal of Geotechnical and Geoenvironmental Engineering*, 138(6):757-762.
https://doi.org/10.1061/(ASCE)GT.1943-5606.0000639
Wei DD, Hu ZX, 1980. Experimental study of preconsolidation pressure and compressibility parameters of Shanghai subsoil. *Chinese Journal of Geotechnical Engineering*, 2(4):13-22 (in Chinese).
Wroth C, Wood D, 1978. The correlation of index properties with some basic engineering properties of soils. *Canadian Geotechnical Journal*, 15(2):137-145.
https://doi.org/10.1139/t78-014
Wu CJ, Ye GL, Wang JH, 2014. Relationship between compression index and natural water content of Shanghai clay. *Rock and Soil Mechanics*, 35(11):3184-3190 (in Chinese).
Wu CJ, Ye GL, Zhang LL, et al., 2015. Depositional environment and geotechnical properties of Shanghai clay: a comparison with Ariake and Bangkok clays. *Bulletin of Engineering Geology and the Environment*, 74(3):717-732.
https://doi.org/10.1007/s10064-014-0670-0
Wu HN, Shen SL, Ma L, et al., 2015. Evaluation of the strength increase of marine clay under staged embankment loading: a case study. *Marine Georesources & Geotechnology*, 33(6):532-541.
https://doi.org/10.1080/1064119X.2014.954180
Yao YP, Sun DA, 2000. Application of Lade's criterion to Cam-clay model. *Journal of Engineering Mechanics*, 126(1):112-119.
https://doi.org/10.1061/(asce)0733-9399(2000)126:1(112)
Yao YP, Hou W, Zhou AN, 2009. UH model: three dimensional unified hardening model for overconsolidated clays. *Géotechnique*, 59(5):451-469.
https://doi.org/10.1680/geot.2007.00029
Yin JH, 1999. Properties and behaviour of Hong Kong marine deposits with different clay contents. *Canadian Geotechnical Journal*, 36(6):1085-1095.
https://doi.org/10.1139/t99-068
Yin JH, 2002. Stress-strain strength characteristics of a marine soil with different clay contents. *Geotechnical Testing Journal*, 25(4):459-462.
https://doi.org/10.1520/GTJ11298J
Yin ZY, Hicher PY, 2008. Identifying parameters controlling soil delayed behaviour from laboratory and in situ pressuremeter testing. *International Journal for Numerical and Analytical Methods in Geomechanics*, 32(12):1515-1535.
https://doi.org/10.1002/nag.684
Yin ZY, Wang JH, 2012. A one-dimensional strain-rate based model for soft structured clays. *Science China Technological Sciences*, 55(1):90-100.
https://doi.org/10.1007/s11431-011-4513-y
Yin ZY, Chang CS, Hicher PY, et al., 2009. Micromechanical analysis of kinematic hardening in natural clay. *International Journal of Plasticity*, 25(8):1413-1435.
https://doi.org/10.1016/j.ijplas.2008.11.009
Yin ZY, Chang CS, Karstunen M, et al., 2010. An anisotropic elastic-viscoplastic model for soft clays. *International Journal of Solids and Structures*, 47(5):665-677.
https://doi.org/10.1016/j.ijsolstr.2009.11.004
Yin ZY, Karstunen M, Chang CS, et al., 2011. Modeling time-dependent behavior of soft sensitive clay. *Journal of Geotechnical and Geoenvironmental Engineering*, 137(11):1103-1113.
https://doi.org/10.1061/(ASCE)GT.1943-5606.0000527
Yin ZY, Xu Q, Hicher PY, 2013. A simple critical-state-based double-yield-surface model for clay behavior under complex loading. *Acta Geotechnica*, 8(5):509-523.
https://doi.org/10.1007/s11440-013-0206-y
Yin ZY, Zhu QY, Yin JH, et al., 2014. Stress relaxation coefficient and formulation for soft soils. *Géotechnique Letters*, 4(1):45-51.
https://doi.org/10.1680/geolett.13.00070

Yin ZY, Yin JH, Huang HW, 2015. Rate-dependent and long-term yield stress and strength of soft Wenzhou marine clay: experiments and modelling. *Marine Georesources & Geotechnology*, 33(1):79-91. https://doi.org/10.1080/1064119X.2013.797060

Yoon GL, Kim BT, Jeon SS, 2004. Empirical correlations of compression index for marine clay from regression analysis. *Canadian Geotechnical Journal*, 41(6):1213-1221. https://doi.org/10.1139/t04-057

Zhu QY, Jin YF, Yin ZY, et al., 2013. Influence of natural deposition plane orientation on oedometric consolidation behavior of three typical clays from southeast coast of China. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, 14(11):767-777. https://doi.org/10.1631/jzus.A1300156

中文概要

题　目：中国海相黏土的压缩指数的 EPR-RCGA 回归模型和 RMSE-AIC-BIC 模型选择及其工程应用

目　的：压缩指数是软土工程领域的关键参数。本文旨在提出一个基于进化多项式回归和实编码遗传算法（EPR-RCGA）的回归分析方法，将压缩指数与物理特性建立相关关系并应用于工程实践。

创新点：结合 EPR 和 RCGA 方法，将中国沿海 21 个不同区域的黏土的压缩性指数与天然含水率和液塑限之间建立相关关系，并采用均方根误差（RMSE）、赤池信息量准则（AIC）和贝叶斯信息准则（BIC）对所建立的不同回归模型进行优选。

方　法：1. 从文献中收集中国沿海 21 个地区的黏土的压缩指数和常见的基本物理性质，并对数据进行整理和分类。2. 进行压缩指数和天然含水量及液塑限之间的 EPR 回归关系分析，并采用新近提出的 RCGA 优化方法来提高回归关系的精度。3. 采用 RMSE、AIC 和 BIC 对不同组合下的回归关系进行优选，并确定最佳回归关系。4. 将得到的关系式应用到有限元路堤计算来验证所得关系式的实用性和准确性。

结　论：1. 本文提出的压缩指数关系式比现有的经验公式更好，预测得到的压缩指数更为精确。2. 采用所提出的压缩指数回归模型预测了东南沿海一路堤下不同土层的压缩指数，并应用所得数据和有限元方法对路堤的沉降进行了模拟分析，验证了所提方法的可靠性。3. 所有结果表明，结合基于 EPR 和 RCGA 的回归分析方法以及基于 RMSE、AIC 和 BIC 的模型选择方法对分析压缩指数与黏土的物理性质的相关关系是切实可行的，可以更好地服务于中国沿海地区的工程设计。

关键词：黏土；压缩性；相关系数；液塑限；有限元；路堤；软黏土