Proceedings of the XVIII ECSMGE 2024
GEOTECHNICAL ENGINEERING CHALLENGES
TO MEET CURRENT AND EMERGING NEEDS OF SOCIETY

ISBN 978-1-032-54816-6
DOI 10.1201/9781003431749-245

ECSMGE 24
XVIII EUROPEAN CONFERENCE ON SOIL MECHANICS AND GEOTECHNICAL ENGINEERING

Assessment of geotechnical properties using optimisation methods

Évaluation des propriétés géotechniques par des méthodes d'optimisation

M. Stãnciucu*, I. Mircea
University of Bucharest, Romania

**stanciucumihaela@yahoo*

ABSTRACT: The paper presents the results of the complex geotechnical evaluation of two sites situated on the alluvial plain of the Dâmbovita River in Bucharest, Romania, where geotechnical investigations were completed with the geophysical survey. Despite short distances between sites, the results illustrate significant variabilities of sedimentary layers and consequently of geotechnical properties. To obtain local representative correlations, we applied in the first instance several well-known formulas, but the degree of fitness was inferior. In the second stage, we developed several evolutionary polynomial regression algorithms (EPR), a method of solving optimization problems, when the objective function is nonlinear. The new predicted values were analyzed using statistical parameters and residual analysis. Thus, we obtain new formulas for each layer, allowing more precise correlations of geotechnical and geophysical parameters inside the sedimentary alluvial structure, by reporting to the relations proposed in the literature.

RÉSUMÉ: L'article présente les résultats de l'évaluation géotechnique complexe de deux sites situés sur la plaine alluviale de la rivière Dâmbovita à Bucarest, en Roumanie, où les études géotechniques ont été complétées avec l'étude géophysique. Malgré de courtes distances entre les sites, les résultats illustrent des variabilités significatives des couches sédimentaires et par conséquent des propriétés géotechniques. Pour obtenir des corrélations représentatives locales, nous appliquons en premier lieu plusieurs formules bien connues, mais le degré de fitness était très faible. Dans un second temps, nous avons développé plusieurs algorithmes de régression polynomiale évolutive (EPR) qui est une méthode de résolution de problèmes d'optimisation, lorsque la fonction objective est non linéaire. Les nouvelles valeurs prédites ont été analysées à l'aide de paramètres statistiques et d'analyses résiduelles. Ainsi, nous obtenons de nouvelles formules pour chaque couche, permettant des corrélations plus précises des paramètres géotechniques et géophysiques à l'intérieur de la structure alluviale sédimentaire, par rapport aux relations proposées dans la littérature.

Keywords: Evolutionary polynomial regression; alluvial deposits; lacustrine.

1 INTRODUCTION

The two sites we refer to (A and B, North and South) are located in the alluvial plain at small distances from the canalized course of the Dâmboviţa River the middle of Bucharest City.

In this research we considered three of the superficial layers, starting from the surface of the terrain, as:

(I) the upper cohesive (7-9m thickness) composed mainly on silty clays;

(II) the middle noncohesive, consisting of sands and fine gravel of 13-30m thickness with thin cohesive intercalations, and

(III) the deeper lacustrine clays, which were partially opened on 15m to 40m.

On each site geotechnical investigations (boreholes and laboratory analyses) were executed in tandem with geophysical survey works (cross-hole and down-hole).

2 CURRENT CORRELATIONS

Estimation of the density layers based on geotechnical usual practices is a very difficult task in case of under consolidate cohesive and noncohesive sedimentary deposits.

In consequence, realistic assessment of relative density through correlation with various in situ investigation results represent the only path to evaluate this important geotechnical parameter. One of the very few such correlations is given by Mayne and Schneider, (1999), which evaluates the relative density of soils ρ (g/cm^3) as a function of shear wave velocities Vs (m/s) and depth z (m):

$$\rho = 0.85\, log(V_s) - 0.16\, log(z) \qquad (1)$$

We apply this correlation for both sites A and B, but the fitting of results with measured data is poor considering the coefficient of correlation $0.2 \le R^2 \le 0.3$, fact that may be visually observed also in Figure 1.

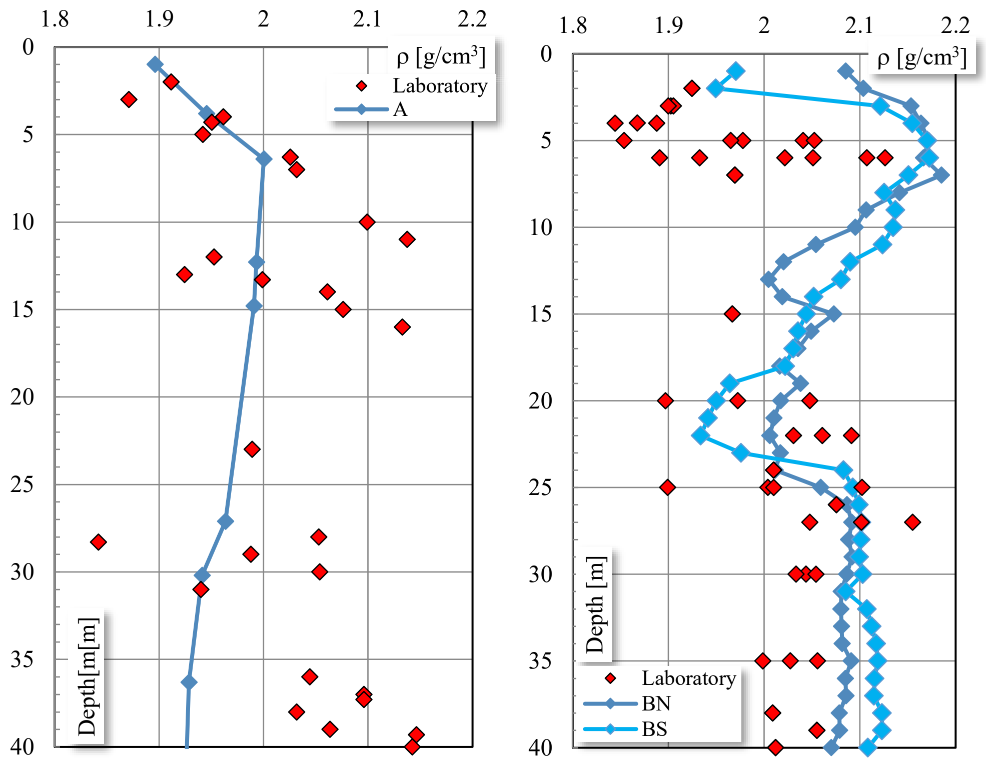

Figure 1. Variation in depth of measured and calculated values of relative density of soils, formula (1).

3 EPR MODEL CONSTRUCTION

3.1 Theoretical frame of EPR models

The mathematical model proposed for this paper is based on evolutionary polynomial regression (EPR) which is a method intensely used in geotechnics in order to find polynomial structures with an output-dependent variable Y and a set of independent variable X (Clegg et al, 2005; Keramati et al, 2014; Rezania et al, 2009).

The function can be written as:

$$Y_{N\times 1} = F(\alpha_{1xk}, X_{N\times m}) \quad (2)$$

where F is a function that will be determined using input – output data and $\alpha_{1\times k} = [a_0\ a_1\ a_2\ \dots\ a_n]$ is a vector with $k = n+1$ parameters. For a matrix of inputs considered as $X_{N\times m} = [X_1\ X_2\ \dots\ X_m]$ and a matrix of exponents whose elements can take values within user–defined bounds, $E_{n\times m}$, can be defined n vectors whose elements are products of independent inputs X as:

$$Z^i_{N\times 1} = X_1^{E(i,1)} \cdot X_2^{E(i,2)} \cdot \dots \cdot X_m^{E(i,m)}, \quad i = 1, \dots, n \quad (3)$$

Thus, a matrix equation results:

$$Y_{N\times 1}(\alpha, Z) = Z_{N\times k} \times \alpha^T_{k\times 1} \quad (4)$$

where $Y_{N\times 1}(\alpha, Z)$ is the least squares estimate vector of the N target values.

To successfully determine the optimal values of the exponents, the genetic algorithm is recommended. This algorithm represents a method of solving optimization problems, especially when the objective function is nonlinear. The process consists of generating a population of individuals who are used as parents to produce a new generation through mutations and crossover techniques. The algorithm is used to obtain an equation that ensure the best possible fit of the data. Then, adjustable parameters a_i, $i = 0, 1, \dots, n$, can be determined by the linear square's method. The estimated equation of regression will be evaluated using the determination coefficient expressed as:

$$R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2} \quad (5)$$

where $\hat{y}_i$ is the estimated value of the output of the process, y_i is the value of a dependent variable and $\bar{y}$ is the mean of the variable y. As R^2 it increases with the inclusion of several variables, the coefficient of determination is adjusted accordingly:

$$\bar{R}^2 = \frac{n-1}{n-p} R^2 - \frac{p-1}{n-p} \quad (6)$$

where n is the number of observations and p is the number of independent variables.

Another statistically important parameter is the residual standard error which is a measure used to assess the precision of the predictions. Regression testing consisted of studying the dispersion analysis, the *F - test* of global significance and the t-test. Thus, if *Significance F* has a value lower than the established significance threshold, then the null hypothesis of the statistic F is rejected. Also, for the *t - test*, a value of *P-value* lower than the significance threshold leads to the rejection of the nullity hypothesis (Clocotici, 2007). Residue analysis assumes the validity of an error normal distribution. This can be verified by studying the diagrams predicted values – residues, $(\hat{y}_i\ ,\ d_i)$, $i = 1, \dots, n$, where the normalized residuals d_i is given by (Montgomery et al., 2003; Pimpan and Suwattee, 2009):

$$d_i = \frac{e_i}{s(e_i)} \tag{7}$$

in which $e_i = y_i - \hat{y}_i$ and $s^2(e_i)$ is the dispersion of the residue e_i, $i = 1, \dots, n$.

If there is an observation i with a large standardized residual ($|d_i| > 3$) then that observation is a potential outlier and can be excluded from the data set or analyzed in another subject of interest.

3.2. EPR models used for geotechnical data

In the next step, several EPS models have been developed and analyzed according to the procedure described above for all three layers, on both sites.

For the upper cohesive layer (I), the EPR model is given in Eq. 8. Specific values of the coefficients $a_{i,i=0,1,2..6}$, as the statistical parameters of regression model are exposed in Table 1 (Figure 2).

$$\rho = a_0 + a_1 \frac{{V_s}^2}{z} + a_2 \frac{1}{z^2 {V_s}^2} + a_3 \frac{1}{z} + a_4 \frac{V_s}{z^2} + a_5 \frac{V_s}{z} + a_6 z V_s \tag{8}$$

Table 1. Regression coefficients and statistical parameters for layer (I), Eq. 8.

Coef-ficients	Value	Regression Statistics		Condi-tions
a_0	20.442344	MR	0.993737	2m<z<8m
a_1	-0.000650	R Sq.	0.987513	
a_2	16341302	Ad. R Sq	0.965661	
a_3	-268.4582	St.Err	0.012176	
a_4	0.108107	SS Res	0.000593	
a_5	0.693340	F	0.001202	
a_6	-0.002579			

For the second layer (II - the middle noncohesive), the regression equation is presented in Eq. 9, and Table 2 (Figure 3) contain the specific values of the coefficients and statistical parameters.

$$\rho = a_0 + a_1 \frac{1}{z {V_s}^2} + a_2 \frac{V_s}{z^2} + a_3 \frac{1}{z^2 V_s} + a_4 \frac{1}{\sqrt{z}} + a_5 z^2 + a_6 \frac{1}{z^2} + a_7 z^2 V_s \tag{9}$$

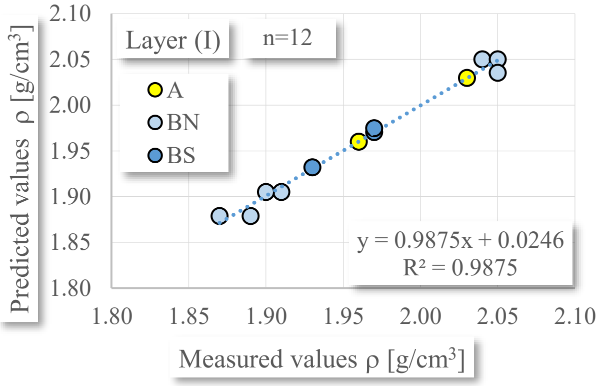

Figure 2. Results of EPR model of layer (I), Eq. 8.

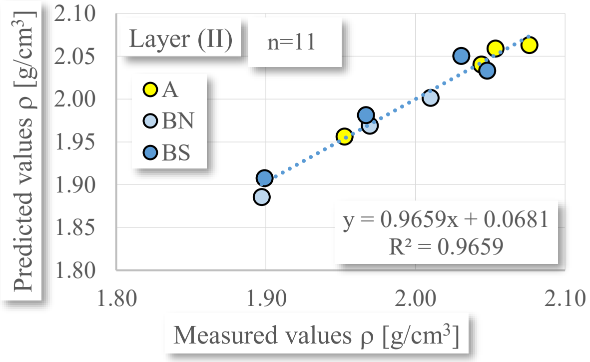

Figure 3. Results of EPR model of layer (II), Eq. 9.

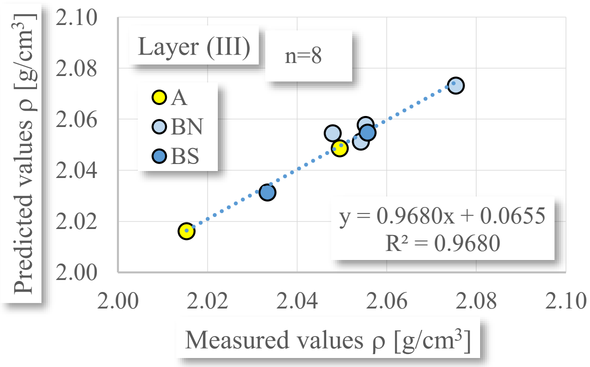

Figure 4. Results of EPR model of layer (III), Eq. 10.

Table 2. Regression coefficients and statistical parameters for layer (II), Eq. 9.

Coef-ficients	Value	Regression Statistics		Cond-itions
a_0	8.532624	MR	0.98278	5m<z<40m
a_1	17639312	R Sq.	0.96586	
a_2	-15.4535	Ad. R Sq	0.88619	
a_3	-3543866	St. Err	0.02099	
a_4	-55.7057	SS Res	0.00132	
a_5	-0.00728	F	0.03254	
a_6	15173.55			
a_7	0.000018			

Finally, for the deeper lacustrine clays (III) the regression equation is given in the following equation (Figure 4. Table 3):

$$\rho = a_0 + a_1 \frac{V}{z} + a_2 V^2 + a_3 \frac{1}{z} \tag{10}$$

Table 3. Regression coefficients and statistical parameters for layer (III), Eq. (10).

Coef-ficients	Value	Regression Statistics		Condi-tions
a_0	1.016192	MR	0.983889	25m<z<50m
a_1	-0.142700	R Sq.	0.968038	
a_2	0.000004	Ad. R Sq.	0.944066	
a_3	76.770632	St. Err.	0.004179	
		SS Res.	0.000070	
		F	0.001895	

4 CONCLUSIONS

Depositional and spatial variability of recent alluvial deposits of the rivers could often mislead the assessment of geotechnical and geophysical parameters based on general correlation equations available in the literature. For this reason, the establishment of such particular relationships at the local scale of a geological or geomorphological formation is an important task in order to obtain precise and reliable results of investigations. Regardless of the quality, the amount or the diversity of investigations, the accuracy and the representativeness of such correlations are strongly influenced by the mathematical tool used for this purpose.

In this paper we examined geotechnical and geophysical investigations of two sites situated on the alluvial plain of Dâmbovița River, Bucharest City. For the assessment of relative density based on shear wave velocity we elaborated an EPR model for every layer of the geological structure.

Using the genetic algorithm tool and the multidimensional linear regression method, equations were obtained that accurately describe the mathematical relationship of one of the most important geotechnical parameters (ρ) for every specific encountered layer. Finally, this may lead to a more precise approach in such regions where general relations are not relevant to the field situation

In these new EPR models statistical parameters extended between the limits exposed in Table 4, proved to be much appropriate than the well-known correlation formula ρ - Vs used in such situations.

Due to the fact that the accuracy of all mathematical models depends among other attributes, on the volume and density of the data used, it appears that it would be of scientific interest to elaborate a communitarian database associated with large geological and geomorphological units, in order to obtain a proper and realistic assessment of geotechnical parameters, which are the most important key parameters of serious and safe design of all civil or industrial projects.

Table 4. Limits of statistical parameters of the EPR models obtained for layers (I-III), Eqs. 8-9-10.

Regression Statistics	EPR models		Previous models
	Minimum value	Maximum value	
Multiple R	0.9828	0.9937	0.2-0.3
R Square	0.9659	0.9875	
Adjusted R Square	0.8862	0.9657	
Standard Error	0.0042	0.0210	
SS Residual	0.0001	0.0013	
Significance F	0.0012	0.0325	

REFERENCES

Clegg, J., Dawson, J.F., Porter, S.J. and Barley, M.H. (2005). The Use of a Genetic Algorithm to Optimize the Functional Form of a Multi-dimensional Polynomial Fit to Experimental Data, *Conference Evolutionary Computation. The 2005 IEEE Congress* on Volume 1. DOI: 10.1109/CEC.2005.1554782.

Clocotici, V. (2007). *Introducere in statistica multivariata*, Editura Universitatii Alexandru Ioan Cuza – Iasi (in Romanian). ISBN: 978-973-703-322-2.

Keramati, M., Reshad, S.K., Asgarpour, S. and Tutunchian, M.A. (2014). Predicting Shear Strength of Municipal Waste Material by Evolutionary Polynomial Regression (EPR). *Electronic Journal of Geotechnical Engineering*, 19, pp. 53-62.

Mayne, P.W. and Schneider, J.A. (1999). Small and large strain soil properties from seismic flat dilatometer tests. *Proceedings of the Pre-failure Deformation Characteristics of Geomaterials*, Jamiolkowski et al. editors, Torino, pp. 419-426. ISBN 90 5809 075 2.

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2003). *Introduction to Linear Regression Analysis*, 3rd ed.

New York: John Wiley & Sons, 672 pag. ISBN-10: 0471315656.
Pimpan, A. and Suwattee, P. (2009). A comparative study of outlier detection procedures in multiple linear regressions. *Proceedings of the International Multi-Conference of Engineers and Computer Scientists*. Vol. 1. IMECS 2009, March 18 - 20, Hong Kong. ISBN: 978-988-17012-2-0.
Rezania, M., Javadi, A.A. and Giustolisi, O. (2009). Evaluation of liquefaction potential based on CTP results using evolutionary polynomial regression, *Computers and Geotechnics* 37, pp. 82-92. DOI: 10.1016/j.compgeo.2009.07.006.