Logistic regression approach for estimating the groutability of granular soils with cement-based grouts

E. Tekin
Dr., Gazi University, Ankara, Turkey, etekin@gazi.edu.tr

S. O. Akbas
Associate Professor, Gazi University, Ankara, Turkey, soakbas@gazi.edu.tr

KEYWORDS: Groutability, Logistic regression, Granular soil, Microfine cement

ABSTRACT: A reliable estimation of the groutability of the target geomaterial is an essential part of any grouting project. However, this is not a simple task due to the complex dependence of groutability on the properties of both the soil and the grouting agent. In this study, the groutability of granular soils with cement based grouts is investigated by logistic regression analysis, using a database of 88 laboratory results. The proposed logistic regression analysis uses the water-cement ratio of the grout, relative density of the soil, grouting pressure, and the ratio of the diameters of the sieves through which 15% of the soil particles and 85% of the grout pass as independent variables. The predictive ability of the logistic regression analysis was compared to those employing artificial neural networks and discriminant analysis, as well as with traditional empirical estimation methods. The results indicate that the proposed logistic regression-based groutability estimation methodology is superior to discriminant analysis and traditional empirical methods but is less successful than the artificial neural networks approach.

1 INTRODUCTION

Permeation grouting is a widely-used ground improvement technique that involves the injection of suitable particulate grouts and chemical solutions into soil and rock with the aim to reduce permeability and / or improve mechanical properties (Jessberger 1983, Krizek 1985, Welsh 1986, Welsh & Burke 1991, Gouvenot 1998, Dupla et al. 2004). Because of a number of adverse effects that arise from the use of chemical solutions, such as the toxicity and long term strength reduction, the properties and behavior of cement grouts have been a major research focus in recent years (e.g. Zebovitz et al. 1989, Akbulut 1999, Kim et al. 2009, Tekin & Akbas 2010, Tekin & Akbas 2011).

Grouting describes the process of filling voids of a granular soil with a grout by replacing water or air in voids. The success of grouting can be measured by groutability (N). The main problem in the utilization of cement-based grouts is the reliable estimation of groutability of the target geomaterial. The grouting process is based on the complex time-dependent transport process of cement grains through the soil matrix, and is a function of at least the grain size distribution of soil and grout, the concentration and viscosity of grout suspensions, the pore size and hydraulic conductivity of soil, as well as the injection pressure. As a result, although the subject of groutability of granular soils by cement-based grouts has been studied for many years, an agreement on a universal set of criteria or a methodology could not be reached.

As summarized in Table 1, early studies on the estimation of groutability of granular soils involved only a comparison of the grain size of the host soils with that of the cement grout.

Table 1. Examples of Empirical Methods for Estimation of Groutability

Method	Grouting
Burwell (1958) & De Beer (1970) k (cm/s) = 116 (0.7 + 0.034t) $(D_{10})^2$ t= temperature (°C)	$k>1x10^{-1}$ cm/s, grouting is possible by cement grouts $k>5x10^{-3}$ cm/s, grouting is possible by microfine cement grouts $k>1x10^{-4}$ cm/s, grouting is possible by solution grouts
Burwell (1958) & Mitchell (1981) $N_\alpha = (D_{15}) / (d_{85})$	$N_\alpha > 25$, grouting is consistently possible, $N_\alpha < 11$, grouting is possible, $11 < N_\alpha < 25$, in-situ tests should be performed, $N_\beta>11$, grouting is consistently possible, $N_\beta<5$, grouting is not possible. $5<N_\beta<11$, in-situ tests should be performed,
Incecik & Ceren (1995) $N_\gamma = (D_{10}) / (d_{90})$	$N_\gamma>10$, grouting is possible, $N_\gamma<10$, grouting is not possible,
Akbulut & Saglamer (2002) $N_\zeta = \frac{(D_{10})}{(d_{90})} + k_1 \frac{w/c}{FC} + k_2 \frac{P}{D_r}$ $k_1 = 0.5$, $k_2 = 0.01$ 1/kPa	$N_\zeta > 28$, grouting is possible, $N_\zeta < 28$, grouting is not possible.

D_{15} = the diameter of a sieve through which 15% of the soil passes and d_{85} = the diameter of a sieve through which 85% of the cement grout passes, w / c = water / cement ratio of the grout, FC = content of soil passing through a 0.6 mm sieve, P = grouting pressure in kPa, N_α, N_β, N_γ = dimensionless constants, and D_r = relative density of the host soil.

However, recent studies (e.g. Akbulut and Saglamer 2002; Tekin 2004; Kim et al. 2009) indicated the need for the consideration of additional parameters for a more reliable estimation of groutability.

Logistic regression (LR) is a novel approach for classifying observations among classes. It is different from multiple regression in the sense that in multiple regression the dependent variable is assumed to follow normal distribution, but in case of logistic regression the dependent variable follows Bernoulli distribution (if dichotomous) which means it will be only 0 or 1. In cases where the dependent variable can take any numerical value for a given set of independent variables multiple regression can be used instead. This approach has become popular in recent years in geotechnical engineering applications also, because it requires fewer assumptions than traditional discriminant analysis (DA) and therefore is considered more robust (Baecher & Christian 2003). LR have been used in the estimation of liquefaction potential (e.g. Juang et al. 2002), landslide susceptibility assessments (e.g. Ohlmacher & Davis 2003, Ayalew & Yamagishi 2005), slope stability analyses (e.g. Nandi & Shakoor 2008), and retaining wall design (e.g. Choi & 2010).

Within this context, in this paper an LR model is developed to estimate the groutability of granular soils with cement-based grouts. To achieve this, results from a database consisting of 88 laboratory grouting tests were used. The performance of the LR model is compared with those of the most commonly used existing groutability estimation methods (Table 1), newly developed discriminant analysis (Tekin & Akbas 2010) and artificial neural network methods (Tekin & Akbas 2011).

2 LOGISTIC REGRESSION

Consider functional relationship between failure rate (π) and parameters that have the form (Baecher & Christian 2003):

$$\pi = g(X) \tag{1}$$

in which π is probability, and $X = \{X_1, \ldots, X_n\}$ is a vector of parameters upon which the probability depends.

The first thing to note is that probabilities range from zero to one, while regression analysis fits a model to variables that vary from −∞ to +∞. Thus, the probabilities, the classification probabilities – are transformed by forming the log odds-ratio, usually called, the logit:

$$Q = \ln(\frac{\pi}{1-\pi}) \tag{2}$$

to give values that vary within (−∞, +∞). The logit or log odds-ratio is set equal to a linear regression of the parameters X

$$Q = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_m X_m \tag{3}$$

in which the parameters X ={X_1,…, X_n} are the factors influencing the regressed probability, and the scalar weights β ={β_0, β_1,…, β_n} are regression coefficients estimated from data. Thus

$$\ln(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_m X_m \tag{4}$$

From which it follows, by taking the exponential of each side, that

$$\pi = \frac{1}{1+\exp(-X^T \beta)} \tag{5}$$

Eq (5) is called logistic regression model.

3 CLASSIFICATION OF GROUTABILITY WITH LOGISTIC REGRESSION

The database presented in Tekin & Akbas (2010) was used to construct a logistic regression model for the estimation of the groutability of granular soils with cement-based grouts. The database consists of 88 experimental grouting records, 51 of which resulted in success. The success of the grouting depends on the finer content of the grout particles as well as on the size of the voids to be penetrated (Bell, 1993). The proposed logistic regression model uses W/C ratio of the grouting agent, relative density of the soil (D_r), grouting pressure (P), weight percentage of soil passing through a 0.6 mm sieve (FC), and the ratio of the diameter of a sieve through which 15% of the soil passes to the diameter of a sieve through which 85% of the cement grout passes (D_{15}/d_{85}) as input parameters. The statistical properties of the input parameters that are grouped as successful (1) and unsuccessful (0) in terms of groutability, are presented in Table 2.

Table 2. Statistical parameters of grouting samples

Variables		Minimum	Maximum	Mean	Standard deviation	Coefficient of variation
W/C	1[a]	0.8	6.0	1.90	1.55	0.816
	0	0.8	4.0	1.59	0.98	0.616
D_r (%)	1	27.0	80.0	51.53	20.87	0.405
	0	30.0	80.0	55.95	24.43	0.437
Pressure, P (kPa)	1	50.0	690.0	338.33	200.38	0.592
	0	100.0	690.0	209.73	175.74	0.838
Fine content, FC (%)	1	1.0	100.0	43.80	31.84	0.727
	0	5.0	100.0	39.65	36.18	0.912
D_{15}/d_{85}	1	19.0	762.0	99.74	118.66	1.190
	0	10.0	58.0	19.62	9.28	0.473

a - 1: Successful grouting, 0: grouting failed

For the model to be validated, the null hypothesis that is based on the maximum likelihood method should be accepted. -2logL statistics is used to test the null and the alternative hypotheses. For a case where the model represents the data in a perfect manner, -2logL statistics will be equal to zero and it will have a chi-square (χ^2) distribution with n-k degrees of freedom, where k represents the number of parameters in the proposed model. In addition, -2logL statistics enables the assessment of the contribution of independent variables that are included within the model. For the model developed in this study, two different -2logL statistics are calculated. The value of the -2logL statistics for the models with only the constant term and with categorical independent variables are calculated as 119.757 and 57.756, respectively. For one degree of freedom, the results express a difference of 62.001 between the -2logL statistics for these two models and indicate significance for 0% significance level (Table 3).

Table 3. Iteration history of logistic regression model

Iteration	-2 log likelihood	Coefficients[a,b,c,d]					
		Constant	W/C	D_r	P	FC	D_{15}/d_{85}
1	96.932	0.459	-0.218	0.006	-0.002	0.005	-0.007
2	80.224	1.535	-0.552	-0.003	0.001	0.006	-0.026
3	72.356	2.435	-0.757	-0.008	0.003	0.006	-0.045
4	65.016	3.444	-0.729	-0.012	0.003	0.001	-0.076
5	59.112	4.974	-0.570	-0.015	0.003	-0.007	-0.130
6	57.832	6.079	-0.571	-0.019	0.004	-0.009	-0.169
7	57.757	6.432	-0.581	-0.020	0.004	-0.009	-0.181
8	57.756	6.458	-0.582	-0.020	0.004	-0.009	-0.182
9	57.756	6.458	-0.582	-0.020	0.004	-0.009	-0.182

a. Method: Enter
b. Constant is included in the model.
c. Initial -2 Log Likelihood: 119.757
d. Estimation terminated at iteration number 9 because parameter estimates changed by less than 0.001.

For logistic regression, a widely-accepted measure to quantify the amount of variation in the response variables, such as the coefficient of determination (R^2) in regression analysis, does not exist. Thus, several R^2 statistics can be found in the literature concerning logistic regression analyses (Nagelkerke, 1991). The Cox and Snell statistics which resembles coefficient of multiple determination in terms of probability estimation is calculated as 0.506 in the 9th iteration. This ratio roughly indicates a 50% relationship between dependent and independent variables. Nagelkerke R^2 statistics, which is another method to assess the logistic regression model, estimates a 68% relationship between dependent and independent variables.

The parameter estimation results and other related statistics including standard errors of logistic regression analysis are presented in Table 4.

Table 4. Variables in the Equation of Logistic Regression Model

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
W/C	-0.582	0.517	1.267	1	0.260	0.559
D_r	-0.020	0.017	1.394	1	0.238	0.981
P	0.004	0.003	1.510	1	0.219	1.004
FC	-0.009	0.012	0.566	1	0.452	0.991
d_{15}/D_{85}	-0.182	0.050	13.342	1	0.000	0.834
Constant	6.458	1.871	11.915	1	0.001	637.764

The constant parameter of the proposed model is calculated to be 6.458 and the standard errors for the terms W/C, D_r, P, FC, D_{15}/d_{85} are determined to be 0.517, 0.017, 0.03, 0.012 and 0.050, respectively. Using the results presented in Tables 3 and 4, the proposed model to estimate the groutability of granular soils using cement-based grouts is given as follows:

$$\ln\left(\frac{\pi}{1-\pi}\right) = 6.458 - 0.582\left(\frac{W}{C}\right) - 0.02(D_r) + 0.04(P) - 0.09(FC) - 0.182\left(\frac{d_{15}}{D_{85}}\right) \quad (6)$$

Note that the cut-off value in Equation 6 is 0.5, i.e, if the result obtained through the application of the given equation is less than 0.5, the grouting operation is expected to be unsuccessful and vice-versa.

The overall comparison of the performance of the LR model with that of the commonly-used empirical methods for estimating groutability is shown in Table 5. As can be seen in Table 5, the highest ratio of successful predictions to the total number of case records considered (95.4%) is obtained by the artificial neural networks (ANN) model. LR model correctly predicted the groutability condition for 77 of the 88 cases, which puts it into the second place in terms of performance, among the prediction methods considered. Note that the number of cases that could be considered by empirical models varies, as no definite groutability prediction can be performed for borderline situations.

Table 5. Success rates of the methods

Method	Incorrect Prediction	Not assessed number of data	Correct Prediction	Success Rate (%)
De Beer (1970)	37	0	51	57.9
Burwell (1958) & Mitchell (1981)	5	25	58	65.9
Incecik & Ceren (1995)	23	1	64	72.7
Akbulut & Saglamer (2002)	9	2	77	87.5
DA, Tekin & Akbas (2010)	14	0	74	84.1
ANN, Tekin & Akbas (2011)	4	0	84	95.4
Logistic Regression Model	11	0	77	87.5
Experimental Study	0	0	88	100.0

4 CONCLUSIONS

In this study, a LR model was developed to predict the groutability of granular soils using cement-based grouts. A total of 88 laboratory case records were used to develop the model. The proposed model uses w/c, D_r, P, FC, and (D_{15}) base soil / (d_{85}) cement grout as input parameters.

The results obtained using the proposed logistic regression model compared very well with the groutability results determined experimentally. For 87.5% of the cases in the compiled database, the model correctly predicted the groutability result. Furthermore, a comparison of the performance of the LR model with the results using empirical methods from the literature and discriminant analysis method indicated the relative superiority of the proposed LR model. The only empirical model that performs equally successfully is that of Akbulut ve Saglamer (2002).

Although artificial neural networks model for the prediction of groutability resulted in a higher success rate than the proposed LR model, it should be pointed out that LR model has a higher versatility because it is well-defined by Equation 6, whereas the ANN model is a black box.

REFERENCES

Akbulut S. (1999). The improvement of geotechnical properties in granular soils by grouting. Ph.D. Thesis, Istanbul Technical University, Istanbul.

Akbulut, S., Saglamer, A. (2002). Estimating the groutability of granular soils: A new approach, Tunnelling and Underground Space Technology, 17:4, 371-380.

Ayalew, L., Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, Geomorphology, 65(1–2): 15–31.

Baecher, G.B., Christian, J.T. (2003). Reliability and statistics in geotechnical engineering, John Wiley & Sons Inc, Cornwall.

Bell, F.G. (1993). Engineering Treatment of Soils. E&FN Spon, London, 10-160.

Burwell, E.B. (1958). Cement and clay grouting of foundations: Practice of the Corps of Engineering, Journal Soil Mechanics and Foundation Division, ASCE 84: 1551/1-1551/22.

Choi, M., Lee, G., (2010). Decision tree for selecting retaining wall systems based on logistic regression analysis, Automation in Construction, 19(7): 917–928.

De Beer, E. E. (1970). Volume 2 of Grondmechanica, Standard Scientific Pulbishers.

Dupla J-C., Canou J., and Gouvenot D. (2004) An advanced experimental set-up for studying a monodirectional grout injection process. Ground Improvement 8(3): 91-99

Gouvenot D. (1998). State of the art in European grouting. Ground Improvement 2(2): 51-67

Incecik, M., Ceren, I. (1995). Cement grouting model tests, Bulletin of The technical University of Istanbul, 48:2, 305-317.

Jessberger H.L. (1983). Soil grouting general report: improvement of ground. Proc. 8th European Conference on Soil Mechanics and Foundation Engineering, Helsinki, pp 1069-1078

Juang, C.H., Jiang, T., Andrus, R.D, (2002). Assessing Probability-based Methods for Liquefaction Potential Evaluation, J. Geotech. Geoenviron. Engrg. 128(7): 580:589.

Kim J-S, Lee I-M, Jang, J-H, and Choi, H. (2009). Groutability of cement-based grout with consideration of viscosity and penetration phenomenon. Int Journal for Numerical and Analytical Methods in Geomechanics, 33(16): 1771-1797.

Krizek R.J. (1985). Chemical grouting in soils permeated by water. Journal of Geotechnical Engineering Division ASCE 111(7): 898-915

Lai, S.Y., Chang, W.C., Lin, P.S. (2006). Logistic Regression Model for Evaluating Soil Liquefaction Probability Using CPT Data, Journal of Geotechnical and Geoenvironmental Engineering, 132(6): 694,704.

Mitchell, J. K. (1981). Soil improvement: State-of-the-Art, 10th International Conference on Soil Mechanics and Foundation Engineering, Stockholm, 4: 509-565.

Nagelkerke, N.J.D. (1991). A note on a general definition of the coefficient of determination, Biometrika, 78:3, 691-692.

Nandi, A., Shakoor, A., (2008). Application of logistic regression model for slope instability prediction in Cuyahoga River Watershed, Ohio, USA, 2(1): 16-27.

Ohlmacher, G.C., Davis, J.C. (2003). Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA, Engineering Geology, 69(3-4): 331–343.

Tekin, E., Akbas, S.O. (2010). Estimation of the groutability of granular soils with cement-based grouts using discriminant analysis, Journal of The Faculty of Engineering and Architecture of Gazi University, 25:3, 625-633.

Tekin, E., Akbas, S.O. (2011). Artificial neural networks approach for estimating the groutability of granular soils with cement-based grouts, Bulleting Engineering Geology Environment 70:153–161.

Welsh J.P., (1986). In-situ testing for ground modification techniques. In ASCE Geotechnical Special Publication No. 6. New York, pp 322-335

Welsh J.P., Burke G.K. (1991). Jet grouting: uses for soil improvement. Proc. Geotech. Eng. Congress. University of Colorado at Boulder, pp 334-345

Zebovitz S., Krizek R.J., and Atmazidis D.K. (1989). Injection of fine sands with very fine cement grout. J Geotech Eng Div ASCE 115(12): 1717-1733