

# Modeling of soil shear strength using multiple linear regression (MLR) at Penang, Malaysia

Bala Balarabe\*\*\* and Andy Anderson Bery\*

\*School of Physics, Universiti Sains Malaysia, 11800 Penang, Malaysia

\*\*Department of Physics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria

\*Corresponding Author: andersonbery@yahoo.com.my

**Submitted:** 30/10/2018

**Revised:** 07/11/2020

**Accepted:** 15/11/2020

## ABSTRACT

This paper presents multiple linear regression (MLR) soil shear strength models developed from electrical resistivity and seismic refraction tomography data. The MLR technique is used to estimate the value of dependent variables of soil shear strength based on the value of two independent variables, namely, resistivity and velocity. These parameters were regressed using regression statistics technique for generating MLR model. The results of MLR model, which is based on the estimation of model dependent parameters ( $\text{Log}_{10}$  resistivity and  $\text{Log}_{10}$  velocity), calculated for  $p$ -value, are less than 0.05 and VIF value less than 10 for cohesion and friction angle models. This result shows that there is a statistically significant relationship between cohesion and friction angle with geophysical parameters (independent variables). The estimation accuracy of the MLR models is also conducted for verification, and the result shows that *RMSE* value for predicted cohesion and predicted friction angle is  $0.77 \text{ kN/m}^2$  and  $1.73^\circ$  which is close to zero. Meanwhile, *MAPE* value was found to be 4.57 % and 7.61 %, indicating highly accurate estimation for the MLR models of predicted cohesion and predicted friction angle. Based on the application of near surface, the study area was successfully classified into two regions, namely, medium and hard clayey sand. Thus, it is concluded that MLR method is suitable in estimating the subsurface characterization that covered more regions compared to the traditional method (laboratory test).

**Keywords:** Multiple linear regression; Resistivity; Seismic refraction; Shear strength; Statistical.

## INTRODUCTION

Near surface imaging is one of the most challenging methods in geotechnical and geophysical studies. As emphasized in previous studies (Take and Bolton, 2004; Greenwood *et al.*, 2004; Smethurst *et al.*, 2006), the understanding of ground subsurface behavior is related to factors including seasonal slope behavior, vegetation contribution, and pore water pressure changes. In addition, for an effective slope monitoring study, the measurement of internal factors is important and required. Many other researches have used different approaches to investigate the subsurface behavior using slope geometry design (Alejano *et al.*, 2007), groundwater table position (Rahardjo *et al.*, 2010), rainfall infiltration study (Zhan *et al.*, 2007), pore pressure profile measurements (Huang *et al.*, 2012), and boundary effects (Ali *et al.*, 2014). Landslide prediction and slide motion analysis methods were used in Japan by Suwa *et al.* (2010). Near surface study has been a focus study to many researchers other than engineering, as subsurface investigation helps understand the nature of soil and its properties, as it influences the stability of civil structures (Roy and Bhalla, 2017) and also in construction industry (Hytiris *et al.*, 2014). These include geophysicists and geologists, where soil and rock are used as their targets, as these mediums played an important role for human sustainability. Besides soil and rock, slope stability is another concern matter in subsurface investigation since mass movement events such

as landslide and rockfall are related to human and building safety factors. The current slope stability assessment or monitoring requires skilled engineers for operation and installation, resulting in a high cost of study, and it does not provide large areal coverage as explained by Uchimura *et al.* (2015). Upon considering factors using sensors, this study considers the application of geophysical methods, as it can give large areal coverage. One of the well-known methods used for near surface investigation is electrical resistivity imaging (Bery, 2016; Maslinda *et al.*, 2017).

The work by Suzuki and Higashi (2001) used resistivity monitoring inversion data at landslide-slope area. This study is focused on groundwater flow after heavy rain, and laboratory experiments were carried out to estimate differences in the geological characteristics. Therefore, this method is used because it can provide more information related to subsurface compared to borehole record, where borehole only provides subsurface information at a limited point of location (Bery and Ismail, 2018). Meanwhile, electrical resistivity method is used for characterization of ground subsurface related to meteorite impact study (Kiu *et al.*, 2012).

Statistical methods have been previously used for predicting properties of soils. Elaoud *et al.* (2017) demonstrated how statistical methods can be effectively used for soil and tractor properties to estimate penetration resistance of soil. In their study, there were four different predictors used in the statistical analysis. The predictors were the tractor weight, the ground granulometry as its depth function, the humidity variation, and the pressure of wheels. Besalatpour *et al.* (2013) used a comparison study between multiple linear regression (MLR) models, artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), and generalized linear model (GLM) to predict soil wet aggregate stability. Meanwhile, Yilmaz and Kaynar (2011) used multiple regression (MR), artificial neuro-fuzzy inference system (ANFIS), and artificial neural network (ANN) models for prediction of swell potential of clayey soils.

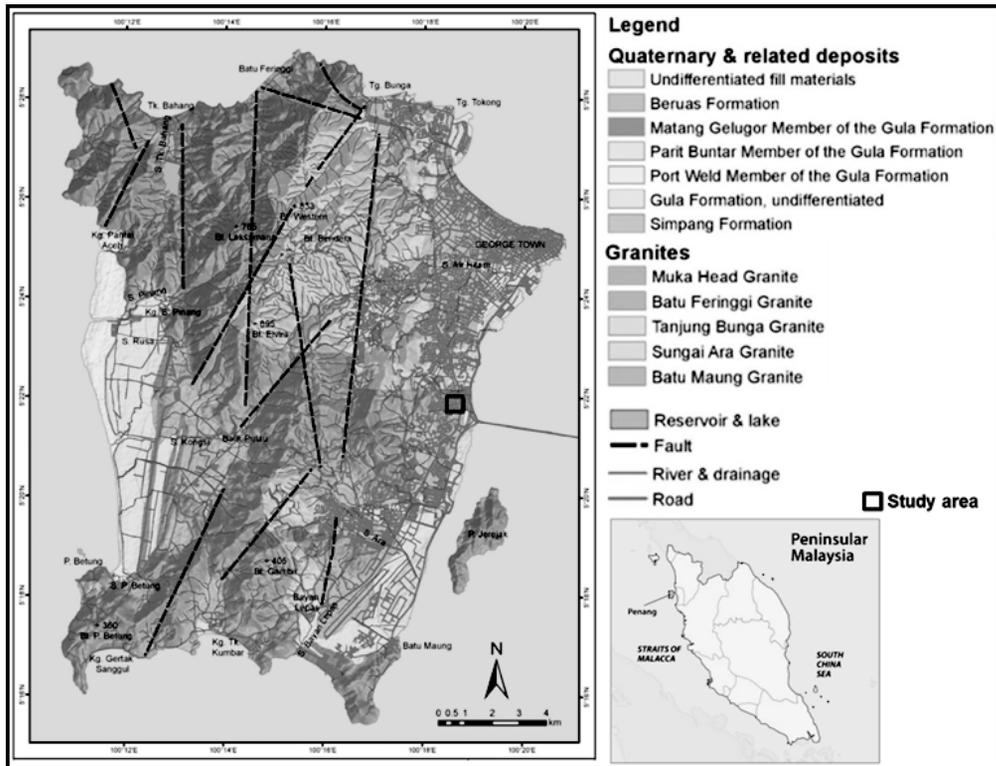
In this study, we proposed a combination of both geophysical and geotechnical methods for shear strength modeling using MLR model. The developed regression model is tested for validity and accuracy. The proposed MLR model is different from previous statistical approaches (Egbe *et al.*, 2017; Jung *et al.*, 2017), where MLR model is used, but prediction accuracy of MLR model was not assessed. Although Israil *et al.* (2006) and Chand *et al.* (2004) considered the application of geoelectrical parameters in their model, they excluded the significance of multiple factors. Thus, in this study, MLR model is checked for its *p*-value and variance inflation factor (*VIF*) for verification before the study stretches into final model development.

The proposed multiple regression model in this study depends simultaneously on two predictors, namely, electrical resistivity and seismic velocity. The developed multiple regression model can yield a more reliable soil cohesion estimation than the traditional method of direct shear test. In addition, we also carried out estimation accuracy evaluation on the MLR model and actual data obtained from direct shear test. The developed MLR model is also applicable to other areas with similar geological setting.

## **GEOLOGICAL SETTING**

This study is carried out in Penang Island, Malaysia. Penang Island is governed under Penang State authority. This state is one of the most rapidly developed areas, and it is located at north west Peninsula Malaysia (Kong, 1994). Penang Island enjoys an equatorial climate, which is similar to other parts of Malaysia. Its temperature is often higher than the mainland, with daytime temperatures ranging between 29 °C and 35 °C. This study area is made up of Penang Granitic rocks (Kong, 1994). Generally, Penang Island can be divided into two types of granitic rocks, namely, Type I, known as Bukit Bendera, and Type II, known as Sungai Ara (Kong, 1994).

The study area is located at Minden, Gelugor, and this area is covered by Type II Penang Granitic rock (Figure 1). The black box shown in Figure 1 is the location of the study area, Minden. Minden area is characterized by medium to coarse-grained biotite granites. The soil of the study area is mainly made up of clayey sand soil with sand domination due to weathering process of the granitic rock.



**Figure 1.** Geological map of Penang, Island, Malaysia (modified after Abdul Hamid *et al.*, 2019).

## MATERIALS AND METHODS

### 2D resistivity and seismic refraction methods

2D resistivity data acquisition is conducted at the Minden, Penang, using ABEM SAS4000 equipment. This equipment is a multielectrode resistivity system, suitable for 2D resistivity surveys. The survey line in this study area is selected in university vicinity due to its location away from civil structures, and it is used for the dependent variable for the multiple regression model developed in this study. In addition, this study area is suitable for slope monitoring purpose. Wenner–Schlumberger array is used as an electrode array because of its potential to resolve vertical and horizontal subsurface changes (Bery, 2018). The total survey length for this study is 20 m, with a minimum electrode spacing of 0.5 m. According to Okpoli (2013) and Bery (2016), small electrode spacing is capable of providing relatively good horizontal resolution compared to a larger electrode spacing. The data acquisition is carried out in the morning, because the subsurface is suitable for planting electrodes into the ground. And this approach can reduce error in the data set because of the dry subsurface conditions occurring in the afternoon or evening. The data is collected automatically using Electrode Selector (ES10-64) equipment. The maximum output of 20 mA and total data stacking of 2 were used for this study. The acquired field dataset of resistivity is processed and iteratively inverted using Res2dinv software through Least-square inversion process. This software was developed by Loke and Barker (1996). A damping factor of 0.05 with a minimum value of 0.01 was employed to increase the accuracy of the calculated apparent resistivity. Moreover, finer mesh type and four nodes between adjacent electrodes were incorporated to enhance the forward modelling subroutine accuracy. Bad data points in the data set will contribute to a large percentage error value of apparent resistivity and will usually be removed using *RMS* statistical error. The maximum cut-off error was set at 30%; hence, only data points with percentage error apparent resistivity of 30% and less were used in the final inversion process.

Seismic refraction method is one of the oldest geophysical methods and has been used for geophysical exploration. This method is governed by measuring the time arrival of refracted waves when the seismic wave is produced. The recorded waves are detected by a set of geophones planted at the ground surface. In this study, seismic refraction method used 15 shot points (including offset locations at  $\pm 4\text{m}$ ,  $\pm 8\text{m}$  and  $\pm 12\text{m}$ ) and 24 geophones with spacing of 1.0 m. Seismic data set is then processed using SeisOpt@2D, which is refraction velocity optimization software, and the input it requires is first arrival travel times and geometry of survey to derive information of subsurface velocity (Optim, 2006).

The electrical resistivity with topography results for survey area is shown in Figure 2. The electrical resistivity result has values ranging from 5 to 1300 ohm m, and *RMS* error for the topography model is 2.3% (Figure 2). The low *RMS* values indicate that the model is reliable, and there are small changes in resistivity distribution throughout the model results. It was observed from the electrical resistivity model that the subsurface characteristics are made up of clayey sand as reported by Bery (2016). The resistivity values of 5–700 ohm m are interpreted as region of medium clayey sand; meanwhile, resistivity values of 700–1300 ohm m are interpreted as hard clayey sand region. For model seismic refraction with topography, it shows that the study area also can be divided into two regions (Figure 3). The first region is interpreted as medium to hard soil with velocity values of 150–632 m/s, and the second region is interpreted as weathered parent material with velocity values of 632 to 1500 m/s.

### Direct shear test

The geotechnical laboratory test using direct shear is carried out for collected soil samples. This is based on forcing soil sample to fail along a plane, while it is being subjected to normal load. The shear stress obtained is defined as the shear resistance developed along known sample section area within sliding plane. The Mohr–Coulomb failure criterion, expressed in terms of effective stress, will be of the form

$$\tau_f = c' + \sigma' \tan \phi' \quad (1)$$

where  $c'$  = cohesion and  $\phi'$  = friction angle, based on effective stress. This approach is conducted to determine the shear strength parameters ( $c'$  and  $\phi'$ ) of soil from known location (distance and depth) along the geophysical survey line.

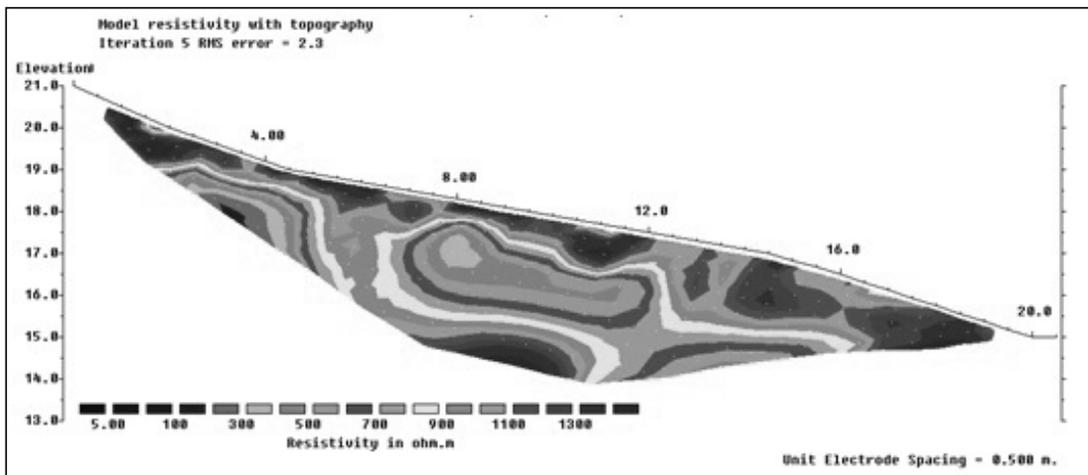
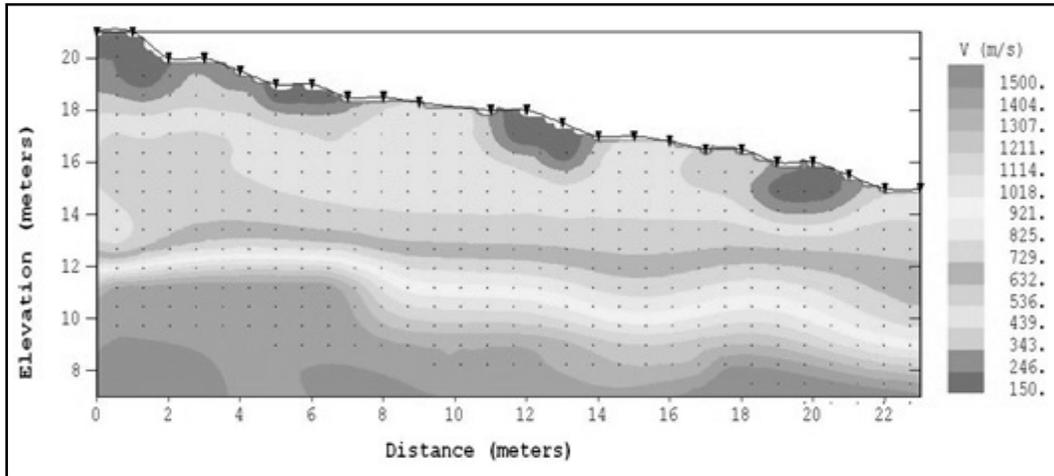


Figure 2. Model resistivity with topography of the study area.



**Figure 3.** Model seismic refraction with topography of the study area.

### Soil shear strength parameters and geophysical data correlations

The electrical resistivity (ohm m) and seismic velocity (m/s) are generated from the 2D imaging scheme. The actual values of soil shear strength parameters (cohesion and friction angle) from laboratory tests were correlated with true resistivity and seismic velocity values at the same location ( $X$ ) and elevation ( $Y$ ). This approach is employed to obtain reliable information about the subsurface. Based on the previous study conducted by Bery (2016), the relationship between resistivity (model resistivity with topography) and cohesion parameters is given by equations (2) and (3). Both equations were used as the actual data, as shear strength of soil parameters is derived from direct shear test in laboratory.

$$c' = 3.156 + 0.014(\rho) \quad (2)$$

$$\phi' = 53.80 - 0.041(\rho) \quad (3)$$

### Multiple linear regression (MLR) model

Indicator variables are used to represent qualitative parameters selected for regression models. It is important to understand the analysis of variance (ANOVA) using indicator variables, which are the models used to analyze geophysical parameters from field study. Consider the following generalized MLR model:

$$Y = \beta_0 + \beta_1(x_1) + \beta_2(x_2) + \dots + \beta_n(x_n) + \varepsilon_i \quad (4)$$

where  $\beta_0$  is the intercept;  $\beta_1$  and  $\beta_2$  are the slopes of the regression line with independent variables or predictors ( $x_1$  and  $x_2$ ), respectively;  $\varepsilon_i$  is the error term; and lastly,  $Y$  is the dependent variable or response as reported in Koutsoyiannis (2001). The MLR model developed in this study is used to construct novel models of soil cohesion and friction angle based on two geophysical parameters. These two geophysical parameters are set as independent variables, known as predictors. In this study, the ANOVA statistical analysis is carried out, and the MLR model used two factorial independent variables. Meanwhile, the dependent variable is set as soil's shear strength parameters (cohesion and friction angle). The statistical analysis conducted in this study was done using IBM SPSS software, and the significance level ( $\alpha$ ) is set at 0.05. In other words, the confidence level is 95%. The MLR models used in the present study are expressed as equations (5) and (6).

$$c' = \beta_0 + \beta_1(\text{Log}_{10}\rho) + \beta_2(\text{Log}_{10}V_p) \tag{5}$$

$$\phi' = \beta_0 + \beta_1(\text{Log}_{10}\rho) + \beta_2(\text{Log}_{10}V_p) \tag{6}$$

The electrical resistivity and velocity values are set in  $\text{Log}_{10}\rho$  and  $\text{Log}_{10}V_p$  terms. According to Loke (2014), the complexity of the Earth’s subsurface is nonlinear; thus, it requires the use of nonlinear equation for these two geophysical parameters. Therefore, resistivity and seismic velocity data sets obtained from subsurface imaging are transformed into nonlinear forms. The coefficients for  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  were determined through regression analysis using the ANOVA scheme.

## RESULTS AND DISCUSSION

### Multiple linear regression (MLR) of soil shear strength models

The quantitative statistical analysis is carried out using IBM SPSS software for the development of new shear strength models. This is agreed by Minichiello (1990), where data are collected measuring parameters and the results reported through statistical analyses. In this study, the first null hypothesis is that there is no statistically significant relationship between cohesion and geophysical parameters. Meanwhile, the second null hypothesis is that there is no statistically significant relationship between friction angle and geophysical parameters. Meanwhile, there are two alternative hypotheses used in this study. The first alternative hypothesis is that there is a statistically significant relationship between cohesion and geophysical parameters, whereas the second alternative hypothesis is that there is a statistically significant relationship between friction angle and geophysical parameters. From the statistical results, the  $p$ -value of predictors’ parameters is inspected of each of the independent variables. If the  $p$ -value is larger than 0.05, we cannot conclude that a significant difference exists (over 0.05, not significant). Therefore, there is no significance relationship, and thus, both null hypotheses can be accepted, and both alternative hypotheses can be rejected.

On the other hand, if the  $p$ -value of the predictors’ parameters is less than 0.05, there is a statistically significant relationship for both MLR models, and thus, both alternative hypotheses can be accepted, while null hypotheses can be rejected. The summary outcomes for Tables 1 and 2 using a sample size of 70 showed that the evaluated resistivity and velocity parameters have a statistically significant relationship with outcomes. This statistically significant relationship is supported by the  $p$ -value for each predictor that is found with less than 0.05 (5%). This implies that the significance of both selected predictors (resistivity and velocity parameters) at confidence level  $\geq 95\%$  can explain the cohesion model (equation 6) and the friction angle model (equation 7), respectively. From the model summary output (Tables 1 and 2), both coefficients of multiple determination are 0.826; therefore, about 82.6% of variation in cohesion and friction angle are explained by resistivity and velocity. The  $p$ -value (less than 0.05 (5%)) in the ANOVA tables indicates that there is a statistically significant relationship between shear strength parameters and geophysical parameters.

**Table 1.** Multiple linear regression results for cohesion parameter with geophysical data with cohesion as dependent variable.

	Coefficients	p-value	VIF
(Constant)	-7.502	0.008	
$\text{Log}_{10}\rho$	9.555	0.000	1.035
$\text{Log}_{10}V_p$	-1.963	0.025	1.035

R Square = 0.826; n = 70

**Table 2.** Multiple linear regression results for cohesion parameter with geophysical data with friction angle as dependent variable.

	Coefficients	<i>p</i> -value	<i>VIF</i>
(Constant)	83.650	0.000	
$\text{Log}_{10} \rho$	-26.754	0.000	1.035
$\text{Log}_{10} V_p$	5.497	0.025	1.035

R Square = 0.826; n = 70

In addition, the variance inflation factor (*VIF*) value is less than 10, and it indicates that there is no multicollinearity. As the rule of thumb, the value of *VIF* that exceeds 10 is often regarded as indicating multicollinearity. From the MLR analysis, the values obtained for equation (5) were  $\beta_0 = -7.502$ ,  $\beta_1 = 9.555$ , and  $\beta_2 = -1.963$ , and for equation (6), they were  $\beta_0 = 83.650$ ,  $\beta_1 = -26.754$ , and  $\beta_2 = 5.497$ . By substituting all the coefficients into equations (5) and (6), the new cohesion ( $c'$ ) and friction angle ( $\phi'$ ) models are obtained as follows:

$$c' = -7.502 + 9.555(\text{Log}_{10}\rho) - 1.963(\text{Log}_{10}V_p) \quad (7)$$

$$\phi' = 83.650 - 26.754(\text{Log}_{10}\rho) + 5.497(\text{Log}_{10}V_p) \quad (8)$$

Appendix A shows normal P-P plot of regression standard residual for dependent variables, and it shows that the residuals are normally distributed. Equations (7) and (8) are the new MLR equations having  $c'$  and  $\phi'$  as the dependent variables or outcomes with independent variables of resistivity and seismic velocity parameters. According to Israil *et al.* (2006), equations (7) and (8) are referred to as models. Therefore, equations (7) and (8) are considered as the new models of soil shear strength parameters with resistivity and seismic velocity as unique independent parameters for this study area.

### Estimation accuracy

In estimating numerical outcome with statistical model, two methods, namely, root mean square error (*RMSE*) and mean absolute percentage error (*MAPE*), are used to examine the strength of estimation between actual and prediction values determined. These approaches are important to understand how strong the estimation accuracy between prediction values (equations 7 and 8) and actual values (equations 2 and 3) of cohesion and friction angle parameters is. The first one, *RMSE*, is used as a measure of the predicted values as being really matching the actual values exactly (Equation 9), and it tells how concentrated the data is around the best fit line. *RMSE* has the same units as the quantity being estimated. *RMSE* close to zero are indicatives of high accuracy between actual and predicted values.

Meanwhile, for *MAPE*, it is the mean difference between the actual values and the predicted values (Equation 10). The smaller the value of *MAPE*, the better (accurate) the model in predicting values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (9)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (10)$$

where  $y_i$  is the actual value,  $\hat{y}_i$  is the predicted value, and  $n$  is the total number of observations.

Table 3 shows the results obtained from sample size of 105 data. The result shows that the cohesion model gives low *RMSE* value ( $\text{kN/m}^2$ ) compared to the *RMSE* value (Degree) of friction angle. The *RMSE* value of both models is found to be less than 2, which indicates that there is a small error between the two data sets; thus, it shows that the predicted values are close to the actual values. Then, the *MAPE* value of both models is found to be less than 10%, which indicates that it is highly accurate for predicted cohesion and friction angle values.

**Table 3.** Summary result for MLR of cohesion and friction angle models.

Model	<i>RMSE</i>	<i>MAPE</i>
Cohesion	0.77	4.57
Friction angle	1.73	7.61

Sample size,  $n = 105$

### Near surface investigation using the MLR models

The last stage in this study is the actual application of subsurface investigation using MLR models. Based on the estimation accuracy result (Table 3), the predicted shear strength parameters are acceptable for MLR model using both geophysical parameters.

Firstly, the geophysical data is processed until the final model is obtained. Then, a new excel file is created, which contained four parameters, namely, distance ( $X$ ), elevation ( $Y$ ), and resistivity ( $\rho$ ) and velocity ( $V_p$ ) data. Two more new columns are created for predicted cohesion and predicted friction angle parameters using MLR, which indicated the same distance and depth with geophysical data sets. Cohesion (equation 7) and friction angle (equation 8) models were used to estimate or predict the soil shear strength parameters distribution from the study area. Figure 4 represented the predicted cohesion and friction angle parameters using the multiple linear regression (MLR) models. It was observed from the final predicted models that the cohesion and friction angle are capable of imaging the subsurface with large coverage of parameter's distribution. The visualization interpretation shows that low to medium cohesion with a value range of  $13.8\text{--}15.6 \text{ kN/m}^2$  is interpreted as medium clayey sand (Figure 4). The medium clayey sand zones are located nearly at middle part of pseudosections, stretching from 4 to 14 m at depth of about 2 m from the surface. Meanwhile, medium to high cohesion with a value range of  $15.6\text{--}17.04 \text{ kN/m}^2$  is interpreted as hard clayey sand. This hard clayey sand soil is found to dominate near the surface of the study area. The interpretation for cohesion model is aligned as discussion made for the model resistivity with topography (Figure 2). The friction angle model result shows that the subsurface can be classified into two zones. The first region with low to medium friction angle with a range value of  $14.5\text{--}20$  degree is interpreted as hard clayey sand soil. Meanwhile, the second region medium to high friction angle with a range value of  $20\text{--}23.85$  degree is interpreted as medium clayey sand soil. The result shows that the greater soil cohesion indicated the lower friction angle between clayey sand particles. In addition, interpretation of this soil shear strength models can be used to derive cementation between sand grains and electrostatic attraction forces between clay particles; thus, direct shear strength increases with the increase of shear strain rate as explained by Bery (2018).

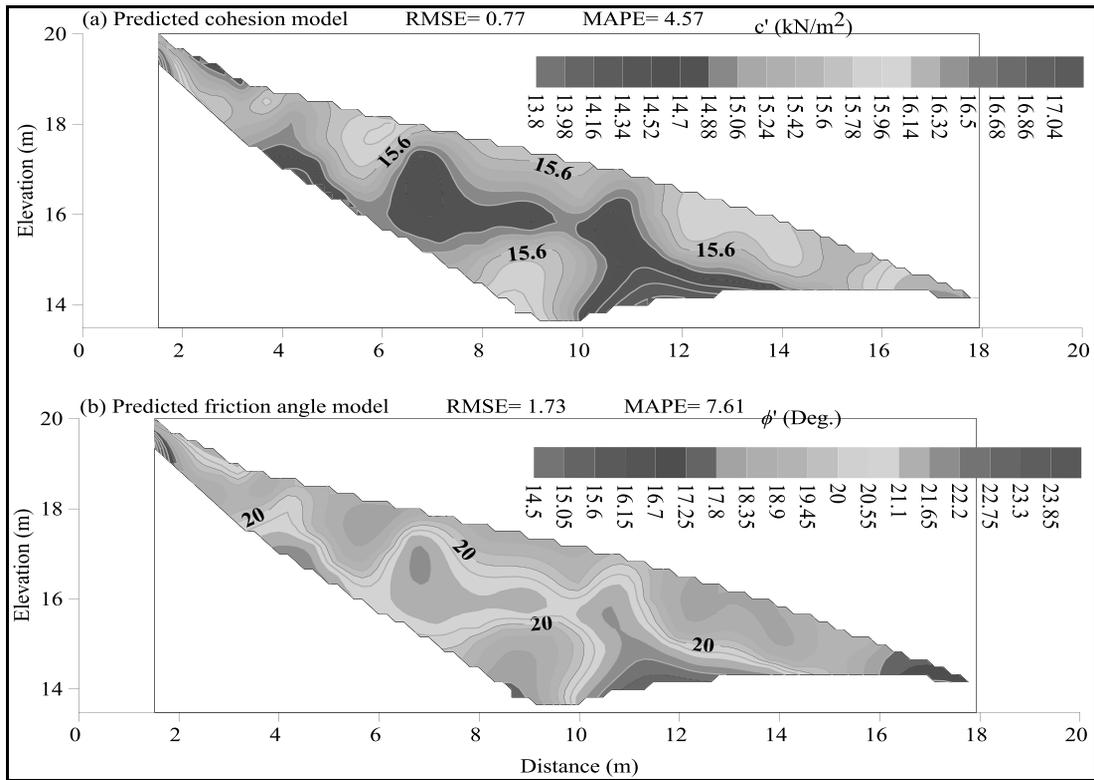


Figure 4. Shear strength models using MLR with cohesion (top) and friction angle (bottom).

High soil cohesion zone means that the soil has larger quantity of clay components than low soil cohesion zone (lesser clay quantity). Sand is defined as free running soil type, whose strength depends on frictions between particles (known as frictional soil). Hard clayey sand indicates that increase in clay element in this soil. Dominant sand grains in soil allow water to flow because of the empty space between sand particles unfilled by clay material. Therefore, increase in cohesion indicates increase in clay content and reduce in sand content. Meanwhile, increase in friction angle indicates increase in sand content and reduce in clay content. These results of soil shear strength models agreed with the results from the resistivity model and seismic refraction model (Table 4). It further indicated that low resistivity and high seismic velocity as functions of low cohesion and high friction angle values in clayey sand soil. Meanwhile, high resistivity and low seismic velocity as functions of high cohesion and low friction angle values in clayey sand soil at the study area.

High cohesion zones are attributed by soils with considerable amount of clay component in the soil compared to regions of low cohesion. This property of soil agrees with low amount of moisture content, even though the sand particles/grains present in the soil profile permit water passage through its pore voids/spaces. It is likely that the amount of sand grains interlocked by increased clay particles is in high expectation, since the soil cohesion changes with water content, grain size, and degree of compaction of soils. Therefore, low cohesion values are attributed to relative increase in the amount of moisture content as clay particles tend to separate with more moisture and number of sand grains, however, opposite behaviour in the friction angle across the model.

These regression-ships agreed by the developed MLR models for cohesion and friction angle (Equations 7 and 8) and supported with statistical tests. Hence, the developed soil cohesion model is a reliable tool for near surface investigation with the help of geophysical methods. In summary, this newly shear strength model developed from MLR is suitable to predict or estimate the subsurface geotechnical parameters at the study area.

**Table 4.** Regression-ships of shear strength models and geophysical models from study area.

Material	Resistivity	Seismic velocity	Cohesion	Friction angle
	$\rho$ (ohm m)	$V_p$ (m/s)	$c'$ (kN/m <sup>2</sup> )	$\phi'$ (Deg.)
Medium clayey sand	5 – 700	439 – 632	13.8 – 15.6	20 – 23.85
Hard clayey sand	700 – 1300	150 – 439	15.6 – 17.04	14.5 – 20

## CONCLUSION

Two shear strength models are developed using the MLR method, which are cohesion and friction angle. The newly proposed shear strength models were generated from geophysical parameters, namely, electrical resistivity and seismic velocity. In addition, estimation accuracy analysis is conducted to check the validity of the newly proposed shear strength models using *RMSE* and *MAPE* statistical methods before generating predicted shear strength models. The MLR method was used to predict or estimate the distribution of cohesion and friction angle parameters of the soil throughout the survey lines. These two models are reliable to image the subsurface of study area in two-dimensional form, which covered more regions compared to traditional method (laboratory). The approach in this study can provide a quick and cost-effective prediction of shear strength parameters by geophysical field studies. Further study on the estimation accuracy of shear strength models can be achieved by increasing the number of soil samples used for actual cohesion values in direct shear test (laboratory). In addition, the developed shear strength models using the MLR method might not be suitable for other study areas with different parent materials. This is because the soil condition is greatly influenced by various physical factors such as weathering process, rainfall amount, and weather condition.

## ACKNOWLEDGMENT

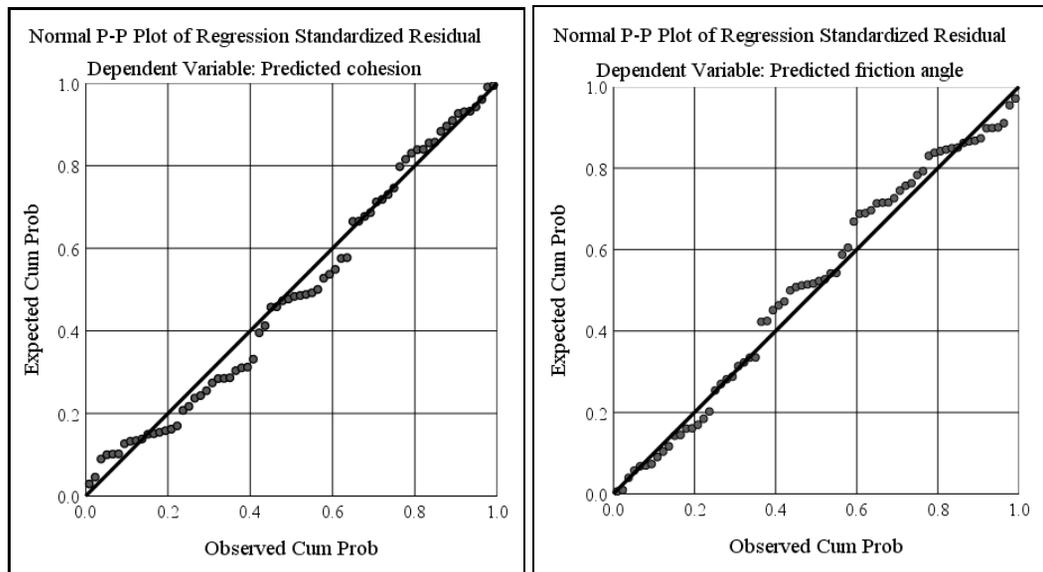
The authors greatly appreciate insightful comments from anonymous reviewers and language editor, which contributed enormously by improving the quality of the manuscript and increased its readability.

## REFERENCES

- Abdul Hamid, F.A.Z., Abu Bakar, A.F., Ng, T.F., Ghani A.A. & Mohamad Zulkifley, M.T. 2019.** Distribution and contamination assessment of potentially harmful element (As, Pb, Ni, Cd) in top soil of Penang, Malaysia. *Environmental Earth Sciences*, **78**(21): 616-627.
- Alejano, L.R., Pons, B., Bastante, F.G., Alonso, E. & Stockhausen, H.W. 2007.** Slope geometry design as a means for controlling rockfalls in quarries. *International Journal of Rock Mechanics and Mining Sciences*, **44**(6): 903-921.
- Ali, A., Huang J., Lyamin, A.V., Sloan, S.W. & Cassidy, M.J. 2014.** Boundary effects of rainfall-induced landslides. *Computers and Geotechnics*, **61**: 341-354.
- Bery, A.A. 2016.** Slope monitoring study using soil mechanics properties and 4-D electrical resistivity tomography methods. *Soil Mechanics and Foundation Engineering*, **53**(1): 24-29.
- Bery, A.A. 2018.** Development of soil electrical conductivity (SEC) modelling for slope assessment in Penang, Malaysia. *Proceedings of EAGE-HAGI 1<sup>st</sup> Asia Pacific Meeting on Near Surface Geoscience and Engineering*. Yogyakarta, Indonesia.
- Bery, A.A. & Ismail, N.E.H. 2018.** Empirical correlation between electrical resistivity and engineering properties of soils. *Soil Mechanics and Foundation Engineering*, **54**(6): 425-429.
- Besalatpour, A.A., Ayoubi, S., Hajabbasi, M.A., Mosaddeghi, M.R. & Schulin, R. 2013.** Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed. *CATENA* **111**: 72-79.
- Chand, R., Chandra, S., Rao, V.A., Singh, V.S. & Jain, S.C. 2004.** Estimation of natural recharge and its dependency on sub-surface geoelectric parameters. *Journal of Hydrology*, **299**(1): 67-83.
- Egbe, J.G., Ewa, D.E., Ubi, S.E., Ikwa, G.B. & Tumenayo, O. 2017.** Application of multilinear regression analysis in modeling of soil properties for geotechnical civil engineering works in Calabar South. *Nigerian Journal of Technology*, **36**(4): 1059-1065.

- Elaoud, A., Hassen, H.B., Salah, N.B., Masmoudi, A. & Chehaibi, S. 2017.** Modeling of soil penetration resistance using multiple linear regression (MLR). *Arabian Journal of Geoscience*, **10**(20): 442.
- Greenwood, J.R., Norris, J.E. & Wint, J. 2004.** Assessing the contribution of vegetation to slope stability. *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, 199-207.
- Huang, A.B., Lee, J.T., Ho, Y.T., Chiu, Y.F. & Cheng, S.H. 2012.** Stability monitoring of rainfall-induced deep landslides through pore pressure profile measurements. *Soils and Foundations*, **52**(4): 737-747.
- Hytiris, N., Stott, R. & McInnes, K. 2014.** The importance of site investigation in the construction industry: a lesson to be taught to every graduate civil and structural engineer. *World Transactions on Engineering and Technology Education*, **12**(3): 414-419.
- Israil, M., Al-hadithi, M., Singhal, D.C. & Kumar, B. 2006.** Groundwater-recharge estimation using a surface electrical resistivity method in the Himalayan foothill region, India. *Hydrogeology Journal*, **14**(1): 44-50.
- Jung, C., Lee, Y., Cho, Y. & Kim, S. 2017.** A study of spatial soil moisture estimation using a multiple linear regression model and MODIS land surface temperature data corrected by conditional merging. *Remote Sensing*, **9**(8): 870.
- Kiu, Y.C., Saad, R., Saidin, M., Nordiana, M.M. & Bery, A.A. 2012.** Characterization of Bukit Bunuh ground subsurface by 2D resistivity for meteorite impact study. *Electronic Journal of Geotechnical Engineering*, **17**(RS): 3575-3583.
- Kong, T.B. 1994.** Engineering properties of granitic soils and rocks of Penang Island, Malaysia”, *Geological Society of Malaysia*, **35**: 69-77.
- Koutsoyiannis, A. 2001.** *Theory of Econometrics*. Palgrave Macmillan Limited, New York.
- Loke, M.H. & Barker, R.D. 1996.** Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method. *Geophysical Prospecting*, **44**(1): 131-152.
- Loke, M.H. 2014.** *Inversion and interpretation of multi-dimensional resistivity surveys*. Geotomo Software, Malaysia.
- Maslinda, U., Nordiana, M.M. & Bery, A.A. 2017.** Porosity determination from 2-D resistivity method in studying the slope failures. *AIP Conference Proceedings*, **1861**: 030025.
- Minichiello, V. 1990.** *In-depth interviewing: Researching people*. Longman Cheshire.
- Okpoli, C.C. 2013.** Sensitivity and resolution capacity of electrode configurations. *International Journal of Geophysics*, **2013**: 608037.
- Optim. 2006.** User’s manual SeisOpt@2D version 5.0. Optim. Inc.
- Rahardjo, H., Nio, A.S., Leong, E.C. & Song, N.Y. 2010.** Effect of groundwater table position and soil properties on stability of slope during rainfall. *Journal of Geotechnical and Geoenvironmental Engineering*, **136**(11): 1555-1564.
- Roy, S. & Bhalla, S.K. 2017.** Role of geotechnical properties of soil on civil engineering structures. *Resources and Environment*, **7**(4): 103-109.
- Smethurst, J.A., Clarke, D. & Powrie, W. 2006.** Seasonal changes in pore water pressure in a grass covered cut slope in London Clay. *Geotechnique*, **56**(8): 523-537.
- Suwa, H., Mizuno, T. & Ishii, T. 2010.** Prediction of a landslide and analysis of slide motion with reference to the 2004 Ohto slide in Nara, Japan. *Geomorphology*, **124**(3-4): 157-163.
- Suzuki, K. & Higashi, S. 2001.** Groundwater flow after heavy rain in landslide-slope area from 2-D inversion of resistivity monitoring data. *GEOPHYSICS*, **66**(3): 733-743.
- Take, W.A. & Bolton, M.D. 2004.** Identification of seasonal slope behaviour mechanisms from centrifuge case studies. *Proceedings of Advances in Geotechnical Engineering: The Skempton Conference*, 992-1004.
- Uchimura, T., Towhata, I., Wang, L., Nishie, S., Yamaguchi, H., Seko, I. & Qiao, J. 2015.** Precaution and early warning of surface failure of slopes using tilt sensors. *Soils and Foundation*, **55**(5): 1086-1099.
- Yilmaz, I. & Kaynar, O. 2011.** Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Systems with Application*, **38**(5): 5958-5966.
- Zhan, T.L.T., Ng, C.W.W. & Fredlund, D.G. 2007.** Field study of rainfall infiltration into a grassed unsaturated expansive soil slope. *Canadian Geotechnical Journal*, **44**(4): 392-408.

## APENDIX A



A normal predicted probability (P-P) plot compares the observed cumulative distribution of the standardized residual to the expected cumulative distribution of the normal distribution. The residuals for both dependent variables (predicted cohesion and predicted friction angle) are normally distributed when data points cluster to the diagonal normality line.