Proceedings of Indian Geotechnical Conference 2020
December 17-19, 2020, Andhra University, Visakhapatnam

Visakhapatnam Chapter

Multiple Regression (MR) and Artificial Neutral Networks (ANN) for Predicting Factor of Safety in Slope Stability Analysis

Aswathi M[1] and S. Chandrakaran[2]

[1] M Tech Scholar, National Institute of Technology Calicut, aswathimattummal@gmail.com
[2] Professor, National Institute of Technology Calicut, chandra@nitc.ac.in

Abstract. To simplify the problem of stability of slopes, predictive modelling techniques can be effectively used. Slope stability refers to the ability of a slope surface to resist movement in its existing condition. It is affected by various parameters which can be utilized for the prediction stability of slopes in the form of factor of safety. This paper presents results of a study to analyze the stability of slopes using Geo-Studio software, to observe the effect of various parameters on the factor of safety and to develop prediction models for analysis of stability of slopes. In this study factor of safety is considered as a function of slope height, slope angle, unit weight of soil, cohesion, angle of internal friction and pore water pressure coefficient. Development of different prediction models was done by using Multiple Regression (MR) and Artificial Neural Networks (ANN). Training is done with a set of data of 570 homogeneous finite slopes having different soil and slope parameters obtained from models developed in SLOPE/W using Morgenstern–Price method. Statistical analysis using new input data is to be done to confirm the practical applications of the developed models. A comparative study of the developed and already existing models is to be done. The slope prediction models will be useful for prediction of stability of slopes and hence for its safe design. The model developed will be helpful to make the slope stability problem less time consuming and easier than the traditional analysis methods.

Keywords: Slope Stability, Regression Analysis, Artificial Neural Networks.

1 Introduction

Analysis of slope stability is of great importance in geotechnical engineering. With the increase in developmental activities, more construction works are being done in sloped soil surfaces. Landslides caused by slope instability have become one of the major disasters in the world. Failure of natural slopes and man-made slopes has resulted in much death and destruction. Soil slope failures are generally of four types such as translational failure, rotational failure, wedge failure, compound failure. Slope stability is the resistance of inclined surface to failure by sliding or collapsing. Slope stability analysis is performed to assess the safe design of manmade or natural slopes.

The main objectives of slope stability analysis are investigation of potential failure mechanisms, determination of the slope sensitivity to different triggering mechanisms, designing of optimal slopes with regard to safety, reliability and economics, designing possible remedial measures such as barriers and soil stabilization.

The advent of electronic computers made it possible to more readily handle the iterative procedures and the use of slope stability software has simplified the analysis to a great extent. Both limit equilibrium method (LEM) and finite element method (FEM) based software are commonly used in geotechnical computations. SLIDE, SLOPE/W, GSLOPE, STABLE WV are examples of softwares used for slope stability analysis. The goal of this study is to develop prediction models for slope stability using multiple linear regression method and artificial neural networks. In statistical modelling, regression analysis is a set of statistical processes for estimating the relationships among variables. It includes many techniques for modelling and analysing the relationship between a dependent variable and one or more independent variables or predictors. With the modern intensive research activities which have occurred in the field of artificial intelligence, it has been possible to solve even badly structured problems. This may not be possible with the conventional programming techniques. One of the significant results of this research is the artificial neural networks (ANN) or simply the neural networks.

2 Regression Analysis

Regression is one of the most widely used statistical technique. It is used to estimates relationships among variables. Regression models provide a very flexible framework for describing and testing hypotheses about relationships between explanatory variables and a response variable. The basis of regression analysis is the linear model. Based on the number of predictor variables used, the linear regression methods can be categorized into two types as simple linear regression and multiple linear regression. Simple linear regression summarize and study relationships between two variables. One variable is regarded as the predictor or independent variable. The other variable is the response or dependent variable. Multiple linear regression is the statistical tool that allows you to examine how multiple independent variables are related to a dependent variable. Once we identified how these multiple variables relate to your dependent variable, we can take information about all of the independent variables and use it to make much more powerful and accurate predictions.

3 Artificial Neural Networks

An artificial neural network (ANN) attempts to simulate the biological structure of the human brain and nervous system. ANNs are very sophisticated modelling techniques. Neural networks consists of three or more layers which includes an input layer, an output layer and one or more hidden layers. Each layer consists of a number of interconnected processing elements, commonly known as neurons. Each neuron is connected to all the neurons in the next layer. Input layer is the one which data are presented to the network and output layer holds the responses of the network to the input. Hidden layers helps to represent and compute complicated association between inputs and outputs. The number of hidden layers used depends on the degree of the complexity of the problem. The neural network learns by modifying the weights of the neurons in response to the errors between the actual output values and target output values. Several learning algorithms have been developed. The back propagation learning algorithm is the most commonly used neural network algorithm and has been applied with great success to model many phenomena in the field of geotechnical engineering. Each hidden and output neuron processes its inputs by multiplying each input by its weight, summing the product and then processing the sum using a non-linear transfer function. The most commonly used transfer function is sigmoid function.

4 Slope Stability Analysis using Geo-Studio Software

SLOPE/W module of GEOSTUDIO 2019 is used to model all slopes in this study. Input is given as slope geometry in the form of slope height and slope angle. Material properties used are cohesion, angle of internal friction and bulk density. Pore-water pressure coefficient is used to represent the piezometric condition. Limit equilibrium methods like Janbu, Spencer, Morgenstern- Price, Bishop and Ordinary methods are available in the software for use. This study is done by opting the Morgenstern–Price method.

5 Parametric Study

Modelling of slope stability is done with different parameters using Geo-Studio Software. The slope stability analysis was carried out using Limit equilibrium analysis.
Slope stability is usually expressed in terms of an index, most commonly the factor of safety, which is usually defined as the ratio of the available shear strength to the shear stress acting on the plane.

A parametric study was done to analyze the behavior of each input parameter towards factor of safety. It was observed that cohesion and angle of internal friction are directly proportional to factor of safety (see Fig. 1&2). And among them cohesion has more influence. Also, when the slope angle and slope height was increased, factory of safety was decreased, leading to failure of slope tremendously (see Fig. 5&6). Factor of safety slightly decreased as the unit weight of soil and pore water pressure coefficient increased (see Fig. 3&4). Compared to shear strength parameters change in unit weight of soil and pore water pressure coefficient has less influence on the stability of slope, as the shear strength of soil depends more on shear parameters such as cohesion and angle of internal friction.

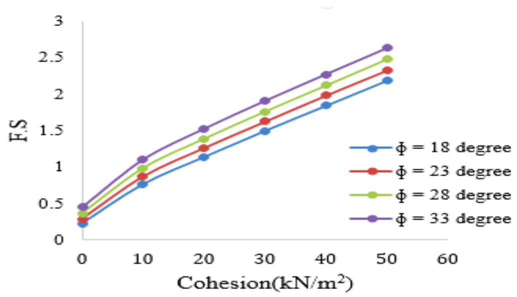

Fig. 1. Behaviour of Slope with respect to Cohesion

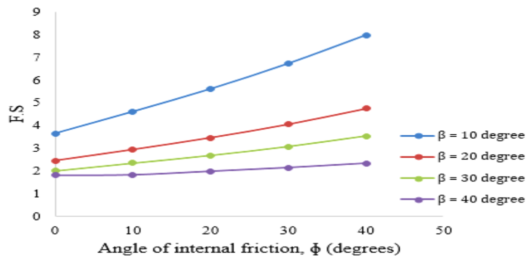

Fig. 2. Behaviour of Slope with respect to Angle of internal friction

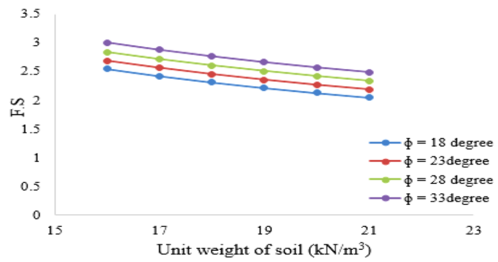

Fig. 3. Behaviour of Slope with respect to unit weight of Soil

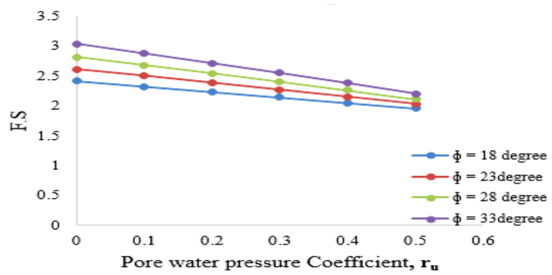

Fig. 4. Behaviour of slope with respect to pore water pressure coefficient

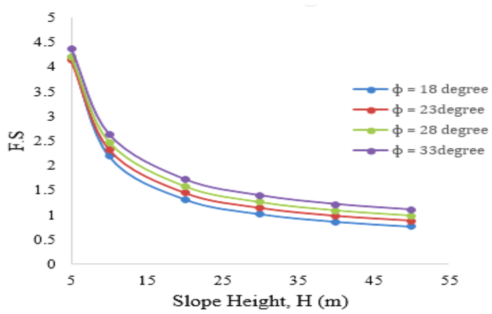

Fig. 5. Behaviour of Slope with respect to height of slope

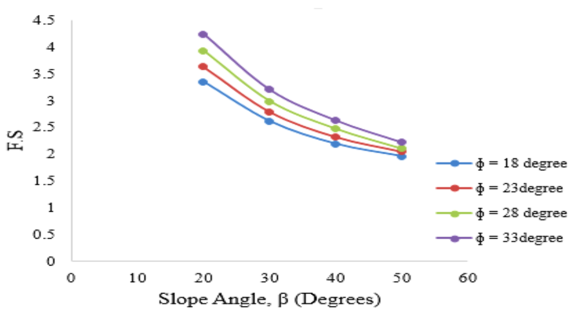

Fig. 6. Behaviour of Slope with respect to slope angle

6 Training Set of Data

On the basis of previous researches six important input parameters are selected for the study. Slope height and slope angle as slope geometry parameters, cohesion, angle of internal friction and bulk density as soil parameters and pore water pressure coefficient are used in this study to develop a slope stability prediction model. Slope stability analysis is done for each case and the corresponding factor of safety is determined using SLOPE/W 2019 software. Morgenstern Price method is used for the analysis. The range of values for each parameter chosen based on literature review is given in table 1. A set of 570 data is used for the training purpose.

Table 1. Range of values for the selected input parameters (Srdan Kostic et al., 2016)

Input Parameter	Minimum Value	Maximum Value
Slope Height, H (m)	6	10
Slope Angle, β (degrees)	25	70
Bulk density, ϒ (kN/m3)	16	20
Cohesion, c (kPa)	0	50
Angle of internal friction, ϕ (degrees)	10	50
Pore-water pressure Coefficient, r_u	0	0.5

7 Results and Discussions

The main objective of this study is to investigate the feasibility of two machine learning based methods, namely, multiple linear regression and artificial neural networks for slope stability determination. This work is carried out based on factor of safety in slope stability analysis. Once the slope stability factors for different geometrical and soil properties of the slope were calculated, the mathematical models were developed. In the present study six input parameters are selected for the prediction of stability of slopes as factor of safety.

7.1 Multiple linear regression models

Table 2. Performance of each Predictor with Output

Dependent Variable	Independent Variable	R^2	RMSE
Factor of Safety	H	0.015	1.36
	B	0.222	1.21
	ϒ	0.000385	1.37
	C	0.519	0.952
	ϕ	0.0429	1.34
	r_u	0.037	1.35

The performance of each predictor with the output is analyzed using regression analysis. Based on the R-squared value from table 2, it is clear that cohesion has more influence on factor of safety than other parameters. The decreasing order of effects of parameters on factor of safety is obtained as $c > \beta > \phi > r_u > H > \Upsilon$.

Regression analysis is carried out for different combinations of input parameters. Some of the results obtained are given in table 3. Based on the R-squared value from table 3, the best model combination is obtained by including all the six parameters together. Hence the selection of input parameters is confirmed.

Table 3. Performance of combinations of Predictors with Output

Combinations	Dependent Variable	Independent Variable	R^2	RMSE
1.	Factor of safety	H, β, ϒ, c, ϕ, r_u	0.864	0.509
2.		β, ϒ, c, ϕ, r_u	0.826	0.575
3.		H, β, ϒ, c, ϕ	0.82	0.584
4.		β, ϒ, c, ϕ	0.782	0.642
5.		H, β, ϒ, c	0.734	0.71
6.		ϒ, c, ϕ, r_u	0.642	0.824
7.		c, ϕ, r_u	0.633	0.833
8.		β, ϒ, c	0.696	0.758

9.	ϒ, c, ϕ	0.603	0.866
10.	H, β, ϒ	0.239	1.2

Training using multiple linear regression with least-squares fit method gives the correlation of determined critical factor of safety to the different input parameters with an R-squared value of 0.869. The model prediction scatter plot on training data set is plotted (see Fig. **1**7). And the model developed revealed the correlation as in Eq. (1).

$$F.S = 4.4969 - 0.17155H - 0.035118\beta - 0.073051\Upsilon + 0.056649c + 0.026885\phi - 1.544r_u \quad (1)$$

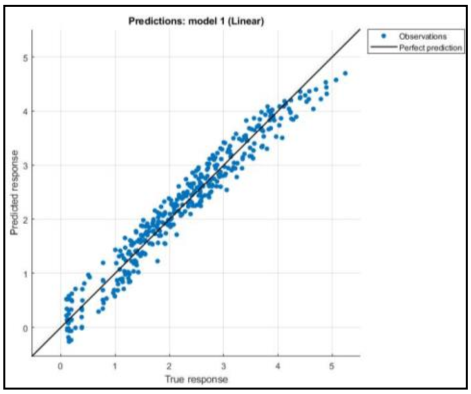

Fig. 7. Model prediction scatter plot on training data set for Model 1

7.2 Stepwise linear regression model

Stepwise regression essentially does multiple regression a number of times, each time removing the weakest correlated variable. At the end there will be only variables that explain distribution best. The model prediction scatter plot on training data set is plotted (see Fig. **1**8). Training using stepwise linear regression method revealed the correlation as in Eq (2).

$$\begin{aligned} F.S = &- 0.33313 - (0.073452*H) - (0.015626*\beta) + (0.027174*\Upsilon) + (0.18681*c) \\ &+ (0.092431*\phi) - (0.15619*r_u) + (0.0024077*H*\beta) - (0.0057242*H*c) \\ &- (0.0016322*H*\phi) - (0.0005021*\beta*c) - (0.00087872*\beta*\phi) \\ &+ (0.016745*\beta*r_u) - (0.0034823*\Upsilon*c) + (0.0001427*c*\phi) \\ &- (0.010601*c*r_u) - (0.061684*\phi*r_u) \end{aligned} \quad (2)$$

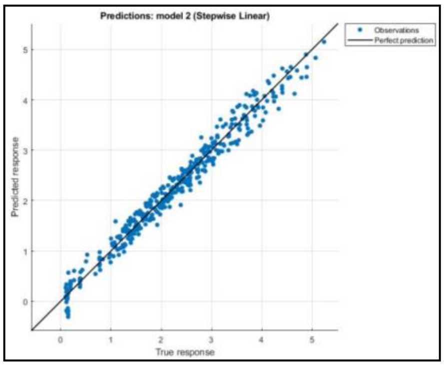

Fig. 8. Model prediction scatter plot on training data set for Model 2

7.3 Artificial Neural Network model

A feed-forward back propagation network is generated having 6 neurons in input layer and one neuron in output layer. The network has a hidden layer with 40 neurons (see Fig. 9). The activation function used for the network is tan-sigmoid. The main objective of training the ANN is to bring the regression coefficient for target and output very close to unity. Another objective is to reduce the mean square error between target and output. Training is done using this method and an overall R-squared value of 0.9983 is achieved (R- value = 0.99916) (see Fig. 10).

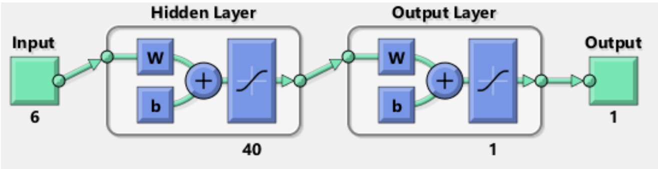

Fig. 9. Framework of Neural Network

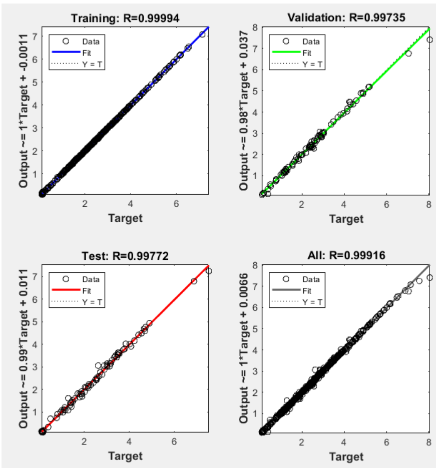

Fig. 10. Scatter diagram showing regression coefficient

7.4 Testing

To check the reliability of the proposed model using multiple linear regression method, its predictive power is tested by analyzing some examples of slopes reported in previous researches. A set of 25 data is used for the testing purpose as given in table 3. The comparison of results obtained by using the different models is shown in table 4.

Table 3. Testing set of data

Sl. No.	Reported by	H (m)	β (Degrees)	ϒ (kN/m^3)	C (kPa)	ϕ (Degrees)	r_u	Target F.S
1.	Sah et al	10	30	22.4	10	35	0	1.924
2.		10.67	22	20.41	24.9	13	0.35	1.521
3.		10.67	25	18.84	15.32	30	0.38	1.608
4.		8	33	22	0	40	0.35	0.652
5.		8	33	24	0	40	0.3	0.744
6.		8	20	20	0	24.5	0.35	0.758
7.	Hoek and Bray	12	40	21	20	40	0	1.84
8.		12	49	21	45	25	0.3	1.53
9.		12	40	21	30	35	0.4	1.49
10.		12	40	21	35	28	0.5	1.43
11.		6	34	20	10	29	0.3	1.34
12.		15	30	20	40	30	0.3	1.84
13.		14	25	18	45	25	0.3	2.09
14.		11	35	19	30	35	0.2	2
15.		10	40	20	40	40	0.2	2.31
16.	Recent studies in Kannur region, Kerala	6	25	18.9	0	32.6	0	1.374
17.		8	30	20.3	20	28.6	0	2.308
18.		10	35	29.1	20	16.4	0	1.126
19.		12	40	31.3	0	32.8	0	0.77
20.		10	45	26.2	30	27.7	0	1.636
21.		14	50	26.7	30	23.5	0.5	0.66
22.		7	55	19.3	50	23	0.5	3.06
23.		9	28	18.9	0	32.6	0.5	0.433
24.		11	32	20.3	20	28.6	0.5	1.192
25.		13	48	29.1	20	16.4	0.5	0.459

Table 4. Comparison of Results

Sl. No	Target F.S	F.S by developed models			Errors		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
1.	1.924	1.5989	1.9506	1.6703	0.3250	0.0267	0.2537
2.	1.521	1.6225	1.2164	1.6625	0.1015	0.3045	0.1415
3.	1.608	1.4999	1.3673	1.7634	0.1080	0.2406	0.1554
4.	0.652	0.8934	1.0872	0.7348	0.2414	0.4352	0.0828
5.	0.744	0.8245	1.2451	0.8310	0.0805	0.5011	0.0870
6.	0.758	1.0794	0.7431	0.7436	0.3214	0.0148	0.0143
7.	1.84	1.7078	2.0068	1.9476	0.1321	0.1668	0.1076
8.	1.53	1.9415	1.4224	1.8974	0.4115	0.1075	0.3674
9.	1.49	1.5223	1.3178	1.5696	0.0323	0.1721	0.0796
10.	1.43	1.4630	1.1615	1.5974	0.033	0.2684	0.1674
11.	1.34	1.6955	1.5139	1.1174	0.3555	0.1739	0.2226
12.	1.84	2.0184	1.2130	1.913	0.1784	0.6269	0.073
13.	2.09	2.6604	1.8182	2.2934	0.5704	0.2717	0.2034
14.	2	2.3243	2.2291	2.0543	0.3244	0.2291	0.0543
15.	2.31	2.9482	2.7985	2.379	0.6382	0.4885	0.069
16.	1.374	2.0854	1.6880	1.3123	0.7114	0.3140	0.0617
17.	2.308	2.4898	2.4429	2.3587	0.1818	0.1349	0.0507
18.	1.126	1.0003	1.0226	0.9839	0.1257	0.1033	0.1420
19.	0.77	0.3712	1.4030	0.7665	0.3988	0.6330	0.0035
20.	1.636	1.7313	1.6281	1.3771	0.0953	0.0078	0.2589
21.	0.66	0.052	0.0858	1.1783	0.608	0.5741	0.5183
22.	3.06	2.6334	2.6641	2.6679	0.4266	0.3958	0.3921
23.	0.433	0.6934	0.5717	0.4879	0.2604	0.1387	0.0549
24.	1.192	1.133	1.1089	1.4904	0.059	0.0830	0.2984
25.	0.459	0.7429	0.2292	0.8036	0.2839	0.2297	0.3446

The variations observed between the predicted and target factor of safety for the testing set of data is in the range of 0.033 to 0.7114 for model 1, 0.0078 to 0.63307 for model 2 and 0.0035 to 5183 for model 3. In case of model 1 more variations are observed and ANN model gives better results among the three.

8 Conclusions

Important factors affecting factor of safety in slope stability are slope height, slope angle, cohesion, angle of internal friction, unit weight of soil and pore water pressure coefficient. Among all the six parameters which affect factor of safety, cohesion and angle of internal friction has a positive influence and unit weight of soil, pore water pressure coefficient, slope height and slope angle shows a negative influence. Cohesion has more influence among all the factors. Slope height has a high negative influence on F.S. When the slope angle and slope height increases, factory of safety decreases, leading to failure of slope tremendously. Compared to shear strength parameters effect of unit weight of soil and pore water pressure coefficient is less on F.S, as the shear strength of soil depends more on shear parameters. The coefficient of determination (R-squared) value obtained is 0.869 in case of linear regression (model 1), 0.964 in case of step wise linear regression (model 2) and 0.9983 in case of ANN (model 3). The variations observed between the predicted and target factor of safety for the testing set of data is in the range of 0.033 to 0.7114 for model 1, 0.0078 to 0.63307 for model 2 and 0.0035 to 0.5183 for model 3. In case of multiple linear regression model more variations are observed and ANN model gives better results among the three. The higher variations are observed in cases when height of slope is more than 10m. Height of slope used for training is in the range of 6m to 10m. The variations may be able to minimize if the training set consists a wider range of value of height of slope.

References

1. Adrin Tohari, Makoto Nishigaki and Mitsuru Komatsu, "Laborotary Rainfall-Induced Slope Failure with Moisture Content Measurement", Journal of Geotechnical and Geoenvironmental Engineering, Vol.133, Issue 5, 2007.
2. Alireza Ahangar-Asr, Asaad Faramarzi and Akbar A. Javadi, "A New Approach for prediction of the Slope Stability of Soil and Rock Slopes", International Journal for Computer Aided Engineering and Software, Vol. 27, No 7, 878-893, 2010.
3. Ashutosh Kainthola, Dhananjai Verma, Rahul Thareja, Singh T N, "A Review on Numerical Slope Stability Analysis", International Journal of Science, Engineeringand Technology Research, Vol. 2, No 6, 2013.
4. Ataollah Shirzadi, Karim Solaimani, Mahmood Habibnejad Roshan, Ataollah Kavin, Kavian, Kamran Chapi, Himan Shahabi, Saskia Keesstra, Baharin Bin Ahmad, Dieu Tien Bui, "Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution", ELSEVIER, Vol. 178, 172-188, 2019.
5. Attom M F, Hawileh R and Abdalla J, "Development of prediction models for determining minimum factor of safety against slope failure in clayey soils", International Journal of Geotechnical Engineering, 7:2, 187-192, 2016.
6. Brian D. Collins and Dobroslav Znidarcic, "Stability Analyses of Rainfall Induced Landslides", Journal of Geotechnical and Geoenvironmental Engineering, Vol 130, No. 4, 362-372, 2004.

7. Sah N. K. et al., "Maximum Likelihood Estimation of Slope Stability", Int. J. Rock Mech. Min. Sci. & Geomech., Vol. 31, No. I, 47-53, 1994.
8. Srdan Kostic, Nebojsa Vasovic and Dusko Sunaric, "Slope stability analysis based on experimental design", International Journal of Geomechanics, Vol. 16, Issue 5, 2016.
9. Upasna Chandarana Kothari and Moe Momayez (2018), "Machine Learning: A Novel Approach to Predicting Slope Instabilities", International journal of Geophysics, Volume 2018, Article ID 4861254, 9 pages.
10. Usama R Mishal and Thair Sh. Khayyun (2018), "Stability analysis of an earth dam using GEO-SLOPE model under different soil conditions", Engineering and technology journal, Vol. 36, Part A, No.5.