Estimating Ground Elevation using Borehole Information: A Case of Metro Manila, Philippines

Galupino, Joenel
Department of Civil Engineering, De La Salle University

Dungca, Jonathan

https://doi.org/10.5109/5909058

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 8, pp.25-30, 2022-10-20. Interdisciplinary Graduate School of Engineering Sciences, Kyushu University
バージョン：

Estimating Ground Elevation using Borehole Information: A Case of Metro Manila, Philippines

Joenel Galupino[1*], Jonathan Dungca[2]
[1]Department of Civil Engineering, De La Salle University, Manila, Philippines
*Corresponding author email: joenel.galupino@dlsu.edu.ph

Abstract: *The ground elevation is an important site-specific parameter for civil engineering. LiDAR is used to determine Ground Elevation at the dawn of new technology; however, there are disadvantages to using LiDAR. The Ground Elevation is also measured during a Geotechnical Site Investigation, however, these data were only collected at the project's location, leaving unknown at other locations. In order to address this issue, machine learning models were used to generate the ground elevation for selected locations, in this case, Metro Manila Philippines. Models of machine learning were trained: Linear Regression Model, Quadratic Regression Model, Tree Regression Model, Boosted Trees Model, and Artificial Neural Network. The Tree Regression Model is the winning model, and its hyperparameter was optimized. To validate, the generated ground elevation was compared to the collected Metro Manila Digital Terrain Model (DTM).*

Keywords: Machine Learning, Metro Manila, Ground Elevation, Geospatial Intelligence

1. INTRODUCTION

The Philippines' seismic activity results from its unusual geographic location, which results in regular volcanic and tectonic activity. This is caused by the country's diverse soil types and irregular ground elevation. The region's soils and rocks have been significantly altered because of the region's dynamic geological formation phase.

Therefore, modern Civil Engineers face a challenge, as they require accurate site information to plan and design construction projects [1]. Due to the project's limited budget and timeline, only limited site information is collected. As a consequence, data can only be collected at the project site, with data at other locations remaining unknown [2].

Numerous efforts are being made at the local level to bridge this gap through the unification of soil parameter quantification, soil property maps, and hazard assessment maps [3-8]. Nonetheless, these studies have limitations, and there are still areas where information is uncertain or unreliable. Many of these studies used "inverse distance weighting" or "Kriging" to perform interpolation in Geographical Information Systems (GIS); however, if the distribution of sample data points is uneven, the quality of the interpolation result will decrease [9].

The ground elevation is a significant site-specific parameter. In Civil Engineering, ground elevation serves as a reference for ground water elevation, pipe embedment depth, foundation embedment depth, and watershed delineation, among other applications.

LiDAR (Light Detection and Ranging) is used to determine Ground Elevation at the dawn of new technology. LiDAR is a remote sensing technique that uses laser light pulses to measure Earth distances. However, there are disadvantages of using LiDAR, such as:

1. High operating costs,
2. Ineffective during heavy precipitation or low cloud cover
3. At high sun angles and reflections, visibility is diminished.
4. Unreliable with regards to water depth and turbulent crashing waves.
5. Errors in elevation due to the inability to penetrate extremely dense forests.
6. Incapable of penetrating dense vegetation.

During Geotechnical Site Investigation, Ground Elevation is determined. Locally, at each construction site, a Geotechnical Site investigation must be conducted and a Professional Report must be submitted. For structures with two or more storeys, an exhaustive geotechnical study is required to evaluate in-situ soil parameters for the design and analysis of foundations. There is a minimum number of required boreholes per structure, which is based on footprint area [10] of the structure. Thus, it is advantageous to collect these Geotechnical Site Investigations and process the data. However, these data were only collected at the project's site, leaving data at other locations unknown [2]. To fill this gap, site investigations are required, however, these are expensive. To address this issue, machine learning played a significant role in the development of cost-reduction models. Machine learning can automate models for processing ground elevation data by learning from the data, recognizing its patterns, and making decisions with minimal human input.

This study's objective is to apply Machine Learning Modeling to generate the ground elevation for specific locations, in this case, Metro Manila, Philippines. Locally, traditional regression techniques are still utilized [9–22], but Machine Learning models have proven effective [23–25] in estimating parameters.

2. METHODOLOGY

2.1 Research Locale

The study's scope of the study is Metro Manila, Philippines. Metro Manila, also known as the National Capital Region (NCR), has a total area of 619.57 km^2. Metro Manila is the capital of the Philippines and one of the country's three metropolitan areas. It consists of sixteen cities and one municipality: Manila, Quezon City, Caloocan, Las Pias, Makati, Malabon, Mandaluyong, Marikina, Muntinlupa, Navotas, Paranaque, Pasay, Pasig, San Juan, Taguig, and Valenzuela, shown in Fig. 1. It consists of 1,690 Barangays in total.

This region serves as the cultural, economic, educational, and political epicenter of the Philippines. The region, which has been designated a global power city, has a significant local and international impact on commerce, finance, media, art, fashion, research, technology, education, and entertainment.

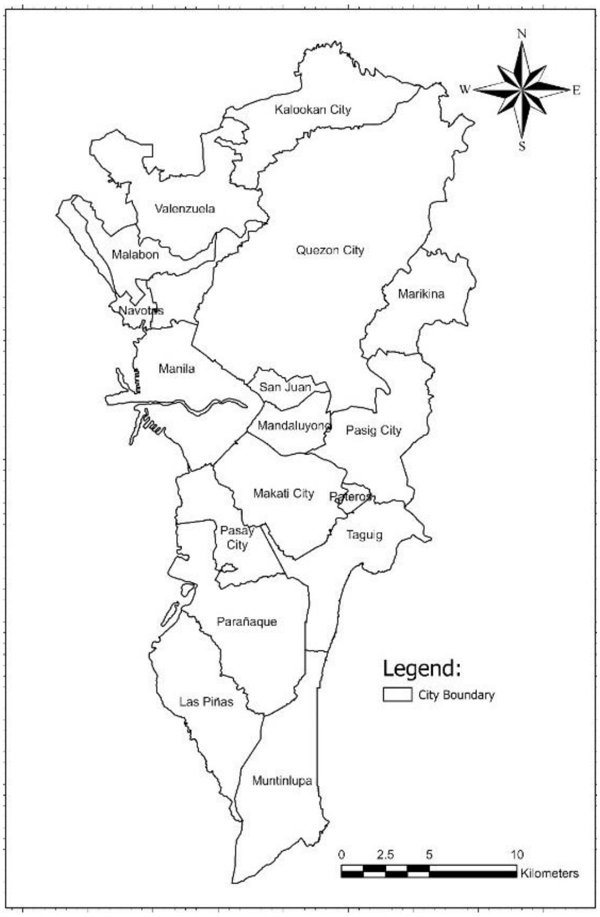

Fig. 1: Map of Metro Manila

2.2 Data

Important borehole information parameters, such as Ground Elevation, from the Borehole Log that was extracted and formatted to MS Excel and to be used in the Machine Learning modelling. The extracted data from the borehole logs will represent its location:

- Latitude, Longitude
- Ground Elevation

The United States Geological Survey USGS provided Metro Manila's Digital Terrain Model (DTM) for elevation.

The density of boreholes per city was checked. One Borehole/km^2 was followed [6]. However, there are instances that usable boreholes are not sufficient. Accuracy rates, Coefficient of Determination (R^2), and Root Mean Square Error (RMSE) were checked to determine the sufficiency of the number of usable borehole data.

Using Geographic Information System (GIS) software, the area of each zone was determined. The predicted properties will represent an area of less than 1 km^2, should there be a Barangay with Area greater than 1 km^2, the said Barangay was divided into smaller zones. Using Geographic Information System (GIS) software, the centroid, in Latitude/Longitude Format, of each zones.

2.3 Machine Learning Models

Before determining the Ground Elevation of a target location, Machine Learning Models were trained. Traditional Regression Models were used since the output data is a numerical value:

- Linear Regression Model [27],
- Quadratic Regression Model [28],
- Tree Regression Model [29],
- Boosted Trees Model [30], and
- Artificial Neural Network [31-32]

In modelling, the independent variables and dependent variables are:

Table 1. Independent and Dependent Variables of the Study

Independent Variables	Dependent Variables
• Latitude • Longitude	• Ground Elevation

The flow of modelling is shown in Fig. 2.

Once the models were trained, their R^2 [33] and RMSE [34] were compared, they are computed using Eq. 1 and Eq. 2. The model with the least RMSE and/or highest R^2 was the winning model.

$$R^2 = \left\{ \frac{\sum[(X - X_m) * (Y - Y_m)]}{\sqrt{\sum(X - X_m)^2 * \sum(Y - Y_m)^2}} \right\}^2 \qquad (1)$$

Where:

X – Data points in Data set X

Y – Data points in Data set Y

X_m – Mean of Data set X

Y_m – Mean of Data set Y

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} ||y(i) - \hat{y}(i)||^2}{N}} \qquad (2)$$

Where:

N - number of data points

$y(i)$ - the i-th measurement

$\hat{y}(i)$ - corresponding prediction.

Table 1 displays the parameters utilized by the Matlab Regression Learner program. The program was used as a foundation for training and validating regression models. After training competing models, compare their R^2 [33] and RMSE [34] side-by-side in order to select the best model. This is a typical procedure for training regression models within the Regression Learner program:

1. Data Selection and Validation;
2. Machine Learning Regression Model Tuning (Hyperparameter Tuning); and
3. Machine Learning Regression Model Training.
4. Implement the selected ML Regression Model.

The hyperparameters of the chosen model were adjusted to determine whether the R2 and RMSE might be enhanced.

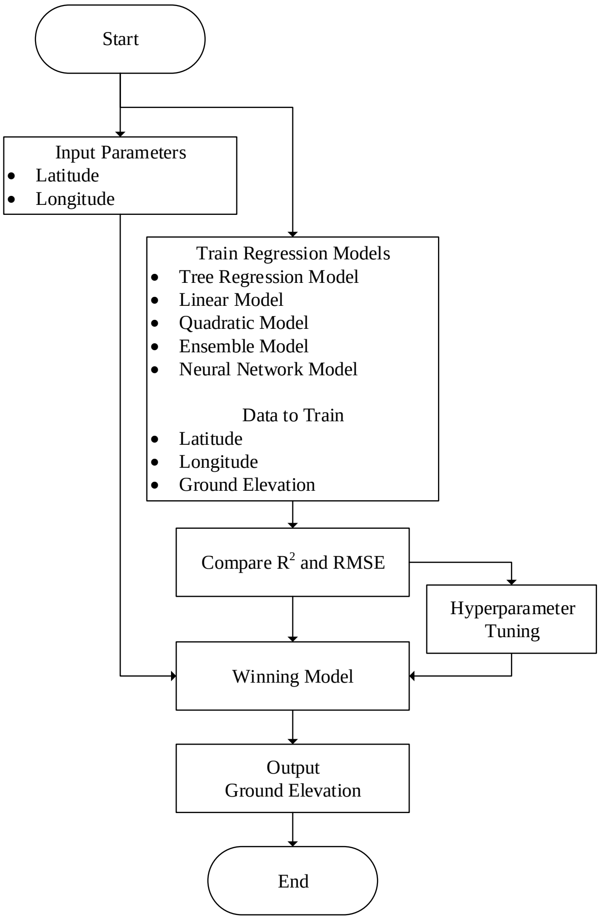

Fig. 2: Process of Modelling the Ground Elevation

2.4 Maps and Profiles

Once the required data has been imported into a GIS platform, it may be integrated with other data layers to generate many unique maps such as the elevation profile of Metro Manila.

2.5 Validation

The United States Geological Survey (USGS) provided topographic maps and geographic information system (GIS) data for elevation. Digital Terrain Model (DTM) of Metro Manila was extracted and compared for validation.

3. RESULTS AND DISCUSSIONS

3.1 Collected Data and Conversion to Smaller Zones

A total of 1,656 data were collected within and outside of Metro Manila, shown in Fig. 3.

Using Geographic Information System (GIS) software, the area of each barangay was calculated. a Barangay with Area greater than 1 km^2, the said Barangay was divided into smaller zones. Using Geographic Information System (GIS) software, the centroid, in Latitude/Longitude Format, of each zones. From the original total of 1,690 Barangays (Zones), it is further increased to 2,036 zones. The new centroids for deployment are shown in Fig. 4. By increasing the number of zones, the elevation can be modeled to further zones.

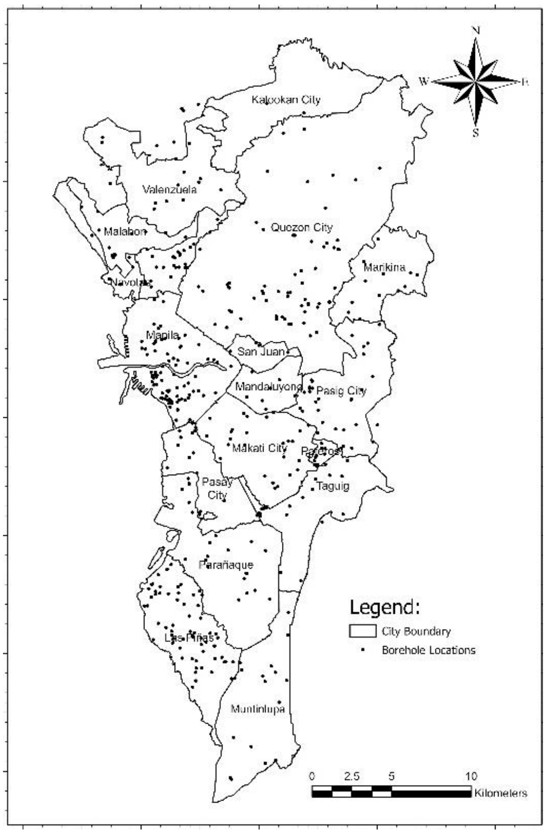

Fig. 3: Borehole Locations collected in Metro Manila

3.2 Coefficient of Determination (R^2) and Root Mean Square Error (RMSE) of Machine Learning Models

In the study, the R^2 and the RMSE were computed to determine the sufficiency of the number of usable borehole data, shown in Fig. 5.

Among the models, the Tree Regression Model has the greatest R^2 and the RMSE, with values of 1 and 2.73x10^{14}, respectively, for the 2.67 BH/km^2 density, thus it is the winning model.

The hyperparameter, Number of Leaves of Tree Regression Model, was adjusted to determine whether the R^2 and RMSE might be enhanced, shown in Fig. 6.However, upon tuning the number of leaves of the Tree Regression Model, the R^2 further decreased and the RMSE further increased, thus, the original hyperparameter was considered, which is One (1) Leaf.

3.3 Maps and Profiles

Maps and profiles were generated from the deployed model. Elevation from 2 meters to 89 meters above mean sea level were gathered, shown in Fig. 7.

The Coastal Area is located on the west side of the Plateau and faces the South China Sea. The Plateau Area is the middle Lengthwise of Metro Manila. The Plains Area are situated between the Plateau and the Province of Rizal and facing Laguna de Bay. The Marikina West Valley Fault is the border between the Plains and the

Plateau. The elevation agrees with the study of JICA (2001) [35].

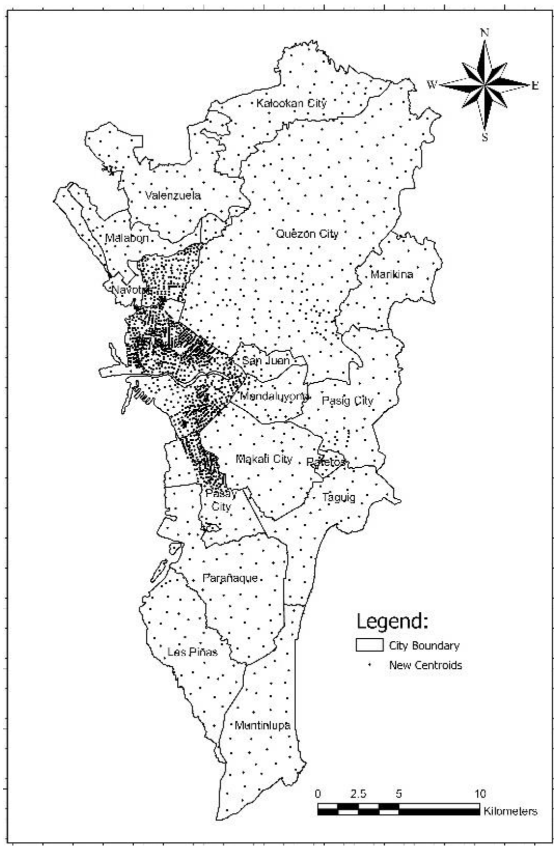

Fig. 4: Centroids of the Converted Smaller Zones

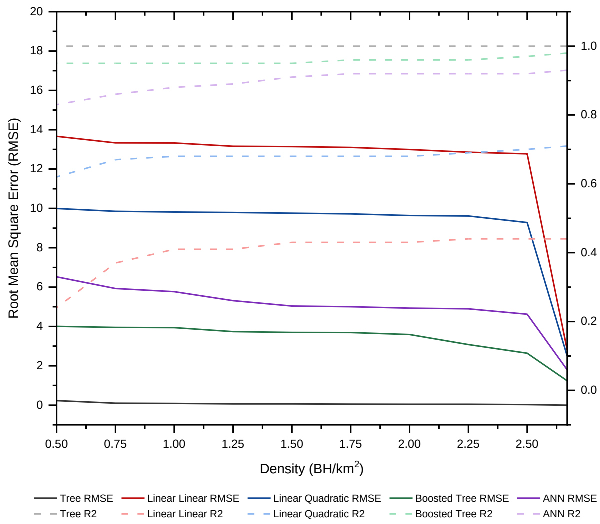

Fig. 5: The R² and the RMSE of the ML Models with Varying Borehole Densities

The Plateau is primarily residential and contains densely populated metropolitan Manila towns such as San Juan, Makati, Quezon City, parts of Paranaque and Muntinlupa. The ground's elevation ranges from nearly 19m to 37m and gradually declines westward. In the northwest region, the elevation varies from 72m to more than 89m, shown in Fig. 8.

The Coastal is a low, level plain that faces Manila Bay. The location of the City of Manila as well as its suburbs. The elevation varies from 2m on Manila Bay to 19m on the west side of Malabon, Navotas, and Paranaque Cities.

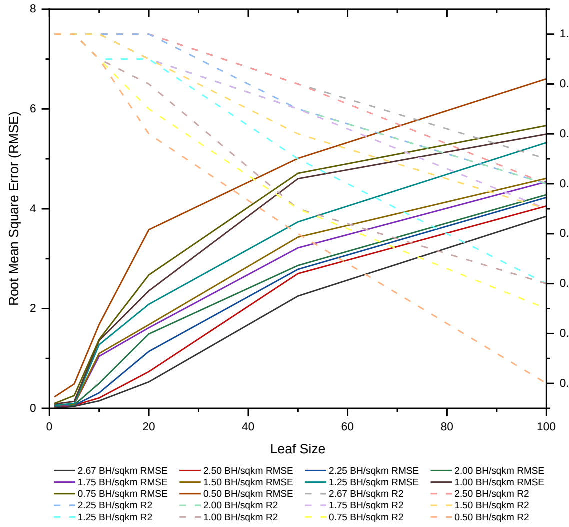

Fig. 6: Hyperparameter Tuning of Tree Regression Model

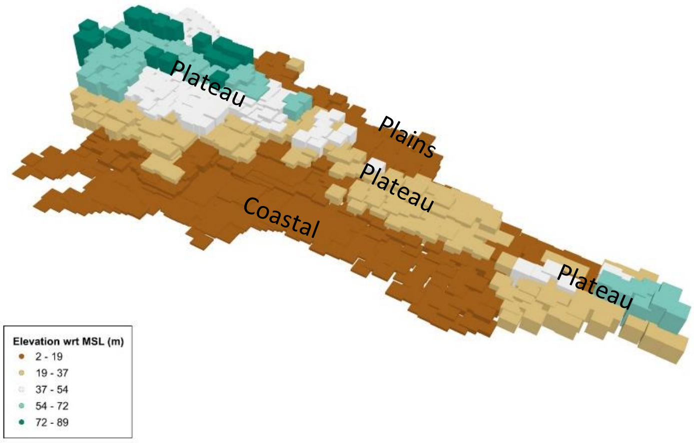

Fig. 7 Generated Ground Elevation for Metro Manila

Fig. 8 Cross Section of Metro Manila

The Plains are made up of floodplains along the Marikina River and a delta along Laguna de Bay. Its elevation ranges from two 2m on the Laguna de Bay side to nineteen 19m at Rizal on the north side. It is bordered by the Plateau and the Rizal Mountains.

3.4 Validation

The USGS provided topographic maps and GIS data for elevation. DTM of Metro Manila was extracted and compared for validation. Using a GIS software, data samples from the centroids in Fig. 4 were extracted on the USGS DTM and compared, shown in Fig. 9:

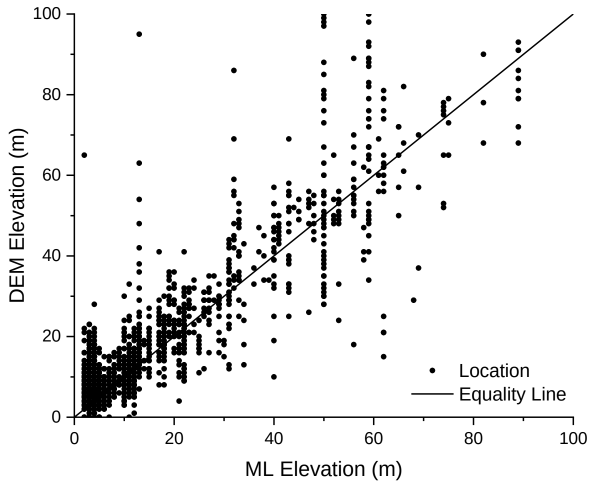

Fig. 9: USGS DTM vs. Machine Learning Generated Ground Elevation

The elevation 2m to 30m, the data are near the equality line, from 31m to 89m, the points are sparse, these may be the result of some gradual changes in the elevation and the availability of training data from these locations.

4. CONCLUSIONS

The Philippines' seismic activity results from its unusual geographic location, which results in regular volcanic and tectonic activity. This is caused by the country's diverse soil types and irregular ground elevation. The goal of this study is to apply Machine Learning Modeling Competition to generate the ground elevation to target locations, in the case of this study, Metro Manila, Philippines. From the original total of 1,690 Barangays (Zones), it is further increased to 2,036 zones.

Machine Learning Models were trained. Regression Models were used since the output data is a numerical value. Among the models, the Tree Regression Model has the greatest R^2 and RMSE, with values of 1 and 2.73x1014, respectively, for the 2.67 BH/km^2 density. However, upon tuning the number of leaves of the Tree Regression Model, the R^2 further decreased and the RMSE further increased, thus, the original hyperparameter was considered, which is One (1) Leaf.

Maps and profiles were generated from the deployed model. Elevation from 2 meters to 89 meters above mean sea level. The Coastal Area is located on the west side of the Plateau and faces the South China Sea. The Plateau Area is the middle Lengthwise of Metro Manila. The Plains Area are situated between the Plateau and the Province of Rizal and facing Laguna de Bay. The Marikina West Valley Fault is the border between the Plains and the Plateau.

The USGS provided topographic maps and GIS data for elevation. DTM of Metro Manila was extracted and compared for validation. The elevation 2m to 30m, the data are near the equality line, from 31m to 89m, the points are sparse, these may be the result of some gradual changes in the elevation and the availability of training data from these locations.

Using Machine Learning Models to predict ground elevation is incredibly useful. It is suggested to apply it to other fields as well [36-37].

5. REFERENCES

[1] Wang, Z. (2019). Site Selection, Power Load, and Power Generation Procedures. Design of Solar Thermal Power Plants, 417–424.

[2] Anbazhagan, P., Uday, A., Moustafa, S. S. R., & Al-Arifi, N. S. N. (2016). Correlation of densities with shear wave velocities and SPT N values. Journal of Geophysics and Engineering, 13(3), 320–341.

[3] Dungca, J., & Chua, R. (2016). Development of a Probabilistic Liquefaction Potential Map of Metro Manila. International Journal of GEOMATE, 2016(April), 1804–1809.

[4] Dungca, J. R. (2020). A reference for the allowable soil bearing capacities in Quezon city, Philippines. International Journal of GEOMATE, 19(71), 42–47.

[5] Galupino, J. G., Garciano, L. E. O., Paringit, M. C. R., & Dungca, J. R. (2017). Location based prioritization of surigao municipalities using probabilistic seismic hazard analysis (PSHA) and geographic information systems (GIS). Paper presented at the HNICEM 2017 - 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, 2018-January 1-7. doi:10.1109/HNICEM.2017.8269521

[6] Galupino, J. G., & Dungca, J. R. (2019). Quezon City soil profile reference. International Journal of GEOMATE, 16(58).

[7] Dungca, J. R., Concepcion, I., Limyuen, M., See, T., & Vicencio, M. (2017). Soil bearing capacity reference for Metro Manila, Philippines. International Journal of GEOMATE, 12(32), 5-11.

[8] Uy, E. E. S., Paringit, M. C. R., Cutora, M. D. L., Galupino, J. G., Garciano, L. E. O., & Dungca, J. R. (2020). Characterization of cebu province municipalities using probabilistic seismic hazard assessment (PSHA) and geographic information system (GIS). Paper presented at the IOP Conference Series: Earth and Environmental Science, 479(1)

[9] Galupino, J. G., & Dungca, J. R. (2015). Permeability characteristics of soil-fly ash mix. ARPN Journal of Engineering and Applied Sciences, 10(15), 6440-6447.

[10] Dungca, J. R., & Galupino, J. G. (2016). Modelling of permeability characteristics of soil-fly ash-bentonite cut-off wall using response surface method. International Journal of GEOMATE, 10(4), 2018-2024.

[11] Dungca, J., Galupino, J., Sy, C., & Chiu, A. S. F. (2018). Linear optimization of soil mixes in the design of vertical cut-off walls. International Journal of GEOMATE, 14(44), 159-165.

[12] R., Galupino, J. G., Alday, J. C., Barretto, M. A. F., Bauzon, M. K. G., & Tolentino, A. N. (2018). Hydraulic conductivity characteristics of road base materials blended with fly ash and bottom ash. International Journal of GEOMATE, 14(44), 121-127.
[13] [9] Adajar, M., Galupino, J., Frianeza, C., Faye Aguilon, J., Sy, J. B., & Tan, P. A. (2020). Compressive strength and durability of concrete with coconut shell ash as cement replacement. International Journal of GEOMATE, 17, 183-190.
[14] Dungca, J. R., & Dychangco, L. F. T. (2016). Strength properties of road base materials blended with waste limestones. International Journal of GEOMATE, 11(3), 2493-2498.
[15] Adajar, M. A. Q., Aquino, C. J. P., dela Cruz, J. D.,II, Martin, C. P. H., & Urieta, D. K. G. (2019). Investigating the effectiveness of rice husk ash as stabilizing agent of expansive soil. International Journal of GEOMATE, 16(58), 33-40.
[16] Adajar, M. A. Q., & Cutora, M. D. L. (2018). The effect of void ratio, moisture content and vertical pressure on the hydrocompression settlement of copper mine tailing. International Journal of GEOMATE, 14(44), 82-89.
[17] Adajar, M. A. Q., de Guzman, E., Ho, R., Palma, C. Jr.III, & Sindico, D. (2017). Utilization of aggregate quarry waste in construction industry. International Journal of GEOMATE, 12(31), 16-22.
[18] Uy, E. E. S., & Adajar, M. A. Q. (2017). Assessment of critical-state shear strength properties of copper tailings. International Journal of GEOMATE, 12(32), 12-18.
[19] Uy, E. E. S., & Dungca, J. R. (2018). A comparative settlement prediction of limestone blended materials using asaoka and hyperbolic method. International Journal of GEOMATE, 14(43), 63-69.
[20] Ubay, I. O., Alfaro, M., Alfaro, M., & Blatz, J. (2020). Stability assessment of an aging earth fill dam considering anisotropic behaviour of clay. International Journal of GEOMATE, 18(66), 84-91.
[21] Adajar, J. B., Ubay, I. O., Alfaro, M., & Chen, Y. (2020). Discrete element model parameters to simulate slope movements. International Journal of GEOMATE, 18(65), 192-199.
[22] Adajar, J. B., Ubay, I. O., Alfaro, M., & Chen, Y. (2020). Discrete element modelling of undrained consolidated triaxial test on cohesive soils. Paper presented at the Geotechnical Special Publication, 2020-February (GSP 317) 172-182.
[23] Elevado, K. J. T., Galupino, J. G., & Gallardo, R. S. (2019). Compressive strength optimization of concrete mixed with waste ceramics and fly ash. International Journal of GEOMATE, 16(53), 135-140.
[24] Elevado, K. J. T., Galupino, J. G., & Gallardo, R. S. (2018). Artificial neural network (ANN) modelling of concrete mixed with waste ceramic tiles and fly ash. International Journal of GEOMATE, 15(51), 154-159.
[25] Dungca, J. R., & Galupino, J. G. (2017). Artificial neural network permeability modeling of soil blended with fly ash. International Journal of GEOMATE, 12(31), 76-83.
[26] G.R. Mettam, L.B. Adams, How to prepare an electronic version of your article, in: B.S. Jones, R.Z. Smith (Eds.), Introduction to the Electronic Age, E-Publishing Inc., New York, 2009, pp. 281–304.
[27] Rath, S., Tripathy, A., & Tripathy, A. R. (2020). Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. Diabetes and Metabolic Syndrome: Clinical Research and Reviews, 14(5), 1467–1474.
[28] Sun, Y., & Wang, Q. (2020). Function-on-function quadratic regression models. Computational Statistics and Data Analysis, 142.
[29] Pekel, E. (2020). Estimation of soil moisture using decision tree regression. Theoretical and Applied Climatology, 139(3–4), 1111–1119.
[30] Elith, J., Leathwick, J. R., & Hastie, T. (2008, July). A working guide to boosted regression trees. Journal of Animal Ecology.
[31] Hopfield, J.J., Artificial neural networks. IEEE Circuits and Devices Magazine, 1988. 4(5): p. 3-10.
[32] Bhardwaj, A. and A. Tiwari, Breast cancer diagnosis using genetically optimized neural network model. Expert Systems with Applications, 2015. 42(10): p. 4611-4620.
[33] Cheng, C. L., Shalabh, & Garg, G. (2014). Coefficient of determination for multiple measurement error models. Journal of Multivariate Analysis, 126, 137–152.
[34] Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30(1), 79–82.
[35] JICA. (2001). Part Ii Metro Manila and Its 200Km Radius Sphere.
[36] Shaqour, A., & Farzaneh, H. (2021). Analyzing the Impacts of a Deep-Learning Based Day-Ahead Residential Demand Response Model on The Jordanian Power Sector in Winter Season. Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES), 7, 247–254.
[37] Liping, Z. (2011). The Application of Neural Network for Sneak Circuit. 0–4.