International Journal of GEOMATE, July, 2022, Vol.23, Issue 95, pp.57-64
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2022.95.3029
Geotechnique, Construction Materials and Environment

MACHINE LEARNING MODELS TO GENERATE A SUBSURFACE SOIL PROFILE: A CASE OF MAKATI CITY, PHILIPPINES

*Joenel Galupino[1] and Jonathan Dungca[2]

[1,2]Faculty, Department of Civil Engineering, De La Salle University, Philippines

*Corresponding Author, Received: 11 Dec. 2021, Revised: 25 April 2022, Accepted: 28 May 2022

ABSTRACT: Soils and rocks are natural geomaterials, and a variety of spatially varying factors influence their properties during their complex geological formation phase. As a result, geomaterials can have different properties at different points on a given site. It is advantageous to be aware of the soil profile of a target location in advance to avoid duplication of tests, determining the borehole depth, and sampling methodology. This can result in more economical borehole testing. The goal of this study is to apply Machine Learning Modeling Competition to generate the soil profile, in the case of this study, Makati City, Philippines. The models competing include Tree Model, Discriminant Model, Naïve Bayes Model, k-Nearest neighbor, and Artificial Neural Network. Among the models, k-Nearest Neighbor Model resulted in the highest accuracy rate, for validation. As an additional output, the generated data was transformed into a soil profile delineation that was represented by soils that are grouped into various classes.

Keywords: Geotechnology, Geospatial intelligence, Machine learning, K-nearest neighbor, Borehole data

1. INTRODUCTION

Soils and rocks are natural geomaterials, and a variety of spatially varying factors influence their properties during their complex geological formation phase. As a result, geomaterials can have different properties at different points on a given site. This has presented a challenge for geotechnical engineers, as they need accurate soil and rock information for the planning and design of their geotechnical construction project [1].

However, borehole logs are typically conducted infrequently for a given project due to the project's limited budget and schedule constraints. As a result, geological and geotechnical information can be probed only at the project's site. Subsurface information at other locations must be inferred based on available data from archived or proposed site investigations, or, more often than not, empirical correlations are used to collect additional information on a site to minimize cost, time spent, and effort [2].

In the Philippines, numerous efforts are being made to bridge this gap by integrating soil parameter quantification, soil property mapping, and hazard assessment mapping [3-8]. These studies, however, have limitations; there are still areas where data is uncertain or unreliable, and the majority of them are not publicly accessible. Geotechnical engineers can accurately predict the soil parameters they will encounter at their target locations by developing models that generate soil parameters for unknown sites.

Local codes specify the depth and number of boreholes required for each project. However, the depth is determined through trial and error. It is advantageous to be aware of the soil profile of a target location in advance to avoid duplication of tests, determining the borehole depth, and sampling methodology. This can result in more economical borehole testing.

Machine learning played a significant role in developing models to further reduce costs. Machine learning can automate soil data processing models by learning from the soil data, recognizing its patterns, and making decisions with minimal human intervention. The goal of this study is to apply Machine Learning Modeling Competition to generate the soil profile, in the case of this study, Makati City, Philippines. Traditional regression technique is still practiced locally [9-22], however, there are Machine Learning models that have been successful [23-25] in estimating geotechnical parameters.

Machine Learning models predict class labels for a given example of input data, which is beneficial for output which is in a format of strings. These models include Tree Model, Discriminant Model, Naïve Bayes Model, k-Nearest neighbor, and Artificial Neural Network.

One of the output is a soil profile that is represented by soils that are grouped into classes, which has similar physical properties and general characteristics in terms of behaviors. The grouping system is usually related to its physical properties inherent in the soil and not for a particular use. With only the soil type available it is not sufficient for design purposes but it will give the engineer an indication of the behavior of soil when used as a component in construction [26].

2. RESEARCH SIGNIFICANCE

Engineers now have a better idea of what to expect in the subsurface as a result of this study. They would know the depth to drill in order to prevent incurring excessive drilling costs. Given that this study has generated the soil profile through the competition of Machine Learning Models, the type of soil in the shallow layers was determined, this would assist Local Government Units in disaster risk reduction and management of problematic soils.

3. METHODOLOGY

3.1 Research Locale

Makati is a first-class, highly urbanized city located in the Philippines' National Capital Region. Makati City continues to be the richest local government unit (LGU) in the Philippines in terms of revenue from domestic sources and per capita income. Makati is the Philippines' financial center; it is home to the country's highest concentration of multinational and local corporations. Makati is home to numerous banks, corporations, department stores, and foreign embassies. As a result, structures are being built to house these developments, and as previously stated, soil explorations are a prerequisite for these projects.

3.2 Data

Soil borehole logs located within and around Makati City were collected. A density of one borehole log per square kilometer was used to describe the soil profile. The distribution was visually inspected and the areas that needed more data were determined. Borehole logs that seemed erroneous were also removed and disregarded [6]. Boreholes were plotted based on their locations on their official reports, shown in Fig. 1.

The independent and dependent variables are shown in Table 1, as patterned in the previous study [6]. The output soil data is a categorical parameter, which is very suitable for a Machine Learning Model.

For the output variable, Soil type, the Unified Soil System (USCS) is primarily concerned with classifying soils according to their textural and plasticity characteristics and behavior [27].

The soils are grouped into classes, which have similar physical properties and general characteristics in terms of behaviors. The grouping system is usually related to its physical properties inherent in the soil and not for a particular use.

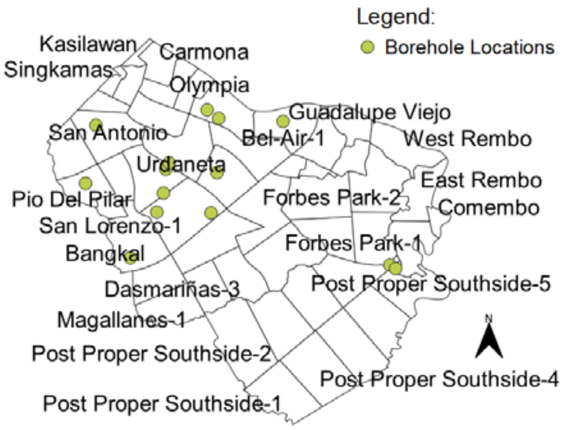

Fig. 1 Collected Boreholes for Makati City

Table 1 Dependent and Independent Variables of the Study

Independent Variable(s)	Dependent Variable(s)
1. Longitude 2. Latitude 3. Elevation	1. Soil Type

With only the soil type available it is not sufficient for design purposes but it will give the engineer an indication of the behavior of soil when used as a component in construction [27]. This study is mainly focused on the main groups used in the study of Galupino and Dungca [6] such as gravel, sand, silt, clay, and tuff.

3.3 Competing Machine Learning Models

Machine learning automated soil data processing models by learning from the soil data, recognizing its patterns, and making decisions with minimal human intervention. This study focused on applying Machine Learning Modeling to Generate the Soil Profile of Makati City, Philippines such as the Tree Model [28], Discriminant Model [29], Naïve Bayes Model [30], k-Nearest neighbor [6,31], and Artificial Neural Network [32-33].

3.4 Data Split Ratio, Accuracy Rates, Hyperparameter Tuning, and Deployment

3.4.1 Data Split Ratio

There are studies that provide detailed analyses of the training and testing process in machine learning [34]. Meaningful data has become more challenging as a result of the proliferation of data. To make sense of the data, a field called data mining has developed. Data mining is a collection of procedures that are used to extract meaning from data [35]. To detect the actions of a machine learning model, we must use observations that were not used during the training phase. Otherwise, the model's assessment will be skewed. The datasets are (1) Training Dataset, (2) Validation Dataset, and (3) Test Dataset. The Dataset Split Ratio is 70% Training Data, 15% Validation Data, and 15% Testing Data.

3.4.2 Accuracy Rates

The accuracy rates for each model were gathered. Among the five (5) competing Machine Learning Models, the highest Accuracy Rate model was used for deployment, this is also known as a Competition. The methodology of Fitri et al. [36] was followed in the study. Accuracy is the ratio of the number of correct predictions to the total number of input samples, shown in Eq. (1):

$$Acc.(\%) = \frac{No.of\ Correct\ Predictions}{Total\ No.of\ Predictions} \quad (1)$$

Parameters shown on Table 1 were used in the Matlab Regression Learner application. The application was used as a basis for training and validating regression models. After training different models, compare their validation errors side by side to see which model is the best. The following is a typical approach for training regression models in the Regression Learner application:

a) Data Selection and Validation
b) Tuning of ML Regression Models (Hyperparameter Tuning)
c) Train ML Regression Model
d) Evaluate the ML Regression Model using the Accuracy Rates (Competition)
e) Deploy selected ML Regression Model

The program utilized the model with the highest accuracy rate to determine the soil stratification of an unknown location. Additionally, it is fine-tuned to determine if the accuracy rates can be increased further as explained in the following section.

3.4.3 Hyperparameter Tuning

The hyperparameters of the selected model were tuned to check if the accuracy rates may be improved. Hyperparameter tuning is the problem of selecting the appropriate collection of hyperparameters for a machine learning system [37]. An example of a hyperparameter is the number of neighbors (k) for the k-NN model.

3.4.4 Deployment

The Machine Learning model with the highest Accuracy Rate was deployed at Barangay/Zone Level. To be consistent with the density of 1 borehole per square kilometer, since some of the Barangays are more than 1 km^2, these barangays were divided into zones, as shown in Fig. 2.

The centroids, in Latitude and Longitude format, of each Barangay were determined thru Geographic Information System (GIS).

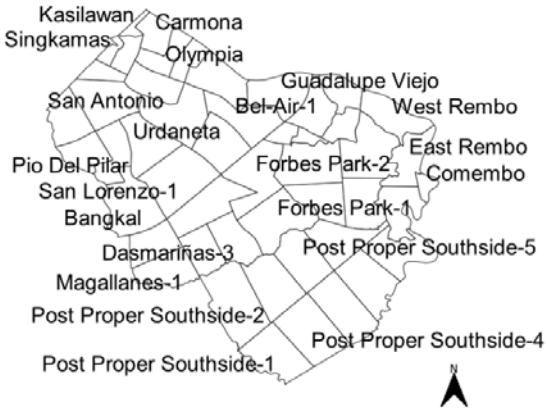

Fig. 2 Generated Map of Makati City Zones

3.5 Generated Soil Profile and Geographic Information System (GIS) Maps

Once all necessary data is entered into a GIS framework, it can be combined to create several unique maps, depending on the data layers used. The isometric soil profile delineation is also provided.

4. RESULTS AND DISCUSSIONS

4.1 Accuracy Rates of ML Models

Accuracy is the ratio of the number of correct predictions to the total number of input samples [36], it is computed using Eq. (1). Multiple Machine Learning Models were used to train the data and these models generated different accuracy rates shown in Table 2, with the highest accuracy rate being used for deployment. The hyperparameters used in the initial testing are the simplest forms. It was further tuned if the machine learning model is the highest on the accuracy rate. A total of 31,021 data points were utilized in the study.

Table 2 Accuracy Rates of the ML Models

Models	Acc. Rate (%)
Decision Tree Model	70.2
Discriminant Analysis	42.8
Naive Bayes Algorithm	53.2
k-Nearest Neighbor Model	93.9
Artificial Neural Network	56.3

Among the models, the k-Nearest Neighbor Model garnered the highest accuracy rate. For k-NN, most of its neighbors vote to classify an object, with the object being assigned to the most prevalent class among its k nearest neighbors. k-NN is very simple and is often used as a benchmark for more complex classifiers.

k-NN has the highest accuracy rate of validation since it used distance to classify an unknown parameter. The k-NN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other [39].

Since the data are location-based, Discriminant Analysis is expected to provide a lower accuracy rate. It is used to represent group distinctions wherein patterns of correlations between variables are assumed to be equivalent from one group to the next, in the study, data vary from each location thus resulting in a lower accuracy rate.

The k-Nearest Neighbor Model garnered the highest accuracy rate, it was further tuned to verify if the accuracy rate will increase if the number of neighbors *k* is also increased, the results are shown in Fig. 3.

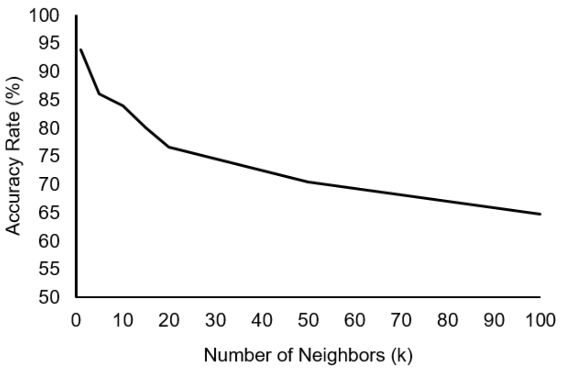

Fig. 3 Hyperparameter Tuning of k-NN Model

When the value of *k* is reduced, the anticipated value becomes less stable. In contrast, increasing the *k* value stabilizes the projected value while increasing the inaccuracy. As a result, the *k* value in the middle should be preferred, and it was also suggested that an odd number should be used as a tie-breaker. As a result, a *k* of five (5) was used for deployment in the study [39].

4.2 Generated Shallow Soil Type of Makati City

The lowest level of soil profile elevation was generated at Barangay Pio del Pilar and San Isidro, with a ground elevation of 3 meters above mean sea level (msl); the highest level of soil profile elevation is located in Forbes Park and Post Proper Southside, with a ground elevation of 37 meters above msl, as shown in Fig. 4.

Barangay Pio del Pilar and San Isidro are located in the Makati West district, which is part of Metro Manila's Coastal Lowland, a flat and low plain bordering Manila Bay. Forbes Park and Post Proper Southside are located on the Central Plateau, where the ground elevation varies between nearly 20 and 40 meters and gradually decreases toward the west.

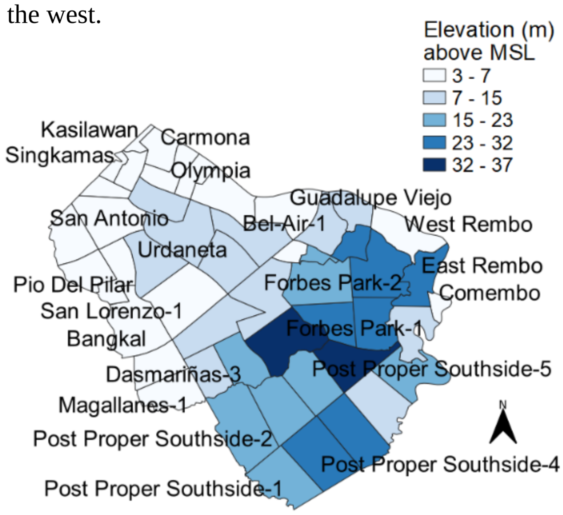

Fig. 4 Generated Elevation of Makati City

Makati City's surface soil type is dominated by clay and sand, with a little amount of silt in Barangay Pembo, as illustrated in Fig. 5. The western surface sand layers are hauled in from Metro Manila's Coastal Lowland, which is plentiful in sand. Along bodies of water, the northern and central barangays have surface sand layers. Pembo with a silt layer on the surface is also situated along the body of water that originates from the Sierra Madre. The meander of Pembo is very susceptible to building up fine-grained materials. Fine-grained sediment such as silt, tends to build up on meandering rivers [40]. This explains the prevalent presence of silt in its surface layer.

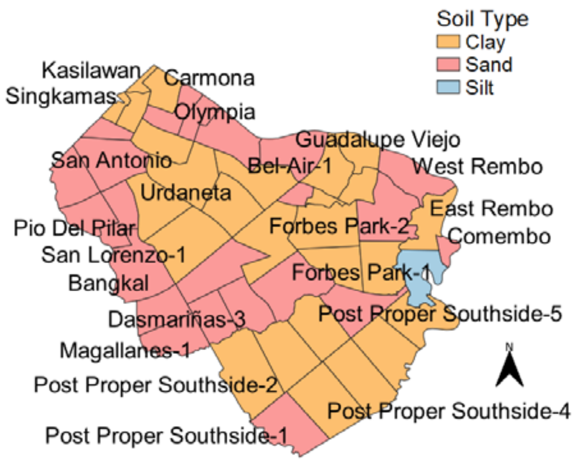

Fig. 5 Surface Layer/Ground Level Soil Type of Makati City

The clay and sand layers around Makati City, are reworked deposits of primary tuff and pyroclastic-flow deposits. Studies indicate that the pyroclastic flow unit overlies a succession of partially marine epiclastic sediments [38]. These marine epiclastic sediments are in relation to the connection of Manila Bay to Laguna Lake 3,000 years ago. However, it was separated from Manila Bay by repeated episodic uplifts along the West Marikina Valley Fault [41-42]. At the moment, the bay-lake interaction between Manila Bay and Laguna Lake occurs via the Pasig River, the lake's sole outlet, where Makati City is located.

Silt strata are encountered at depths of 1 meter, 2 meters, and 3 meters below the earth in the region of Bel-Air, Urdaneta, and Forbes Park, as illustrated in Fig. 6, Fig. 7, and Fig. 8. This is an example of Fine-grained sediments isolated during the geological process of Makati City. However, as one descends farther into Pembo, a continuous sand layer becomes visible.

An alternating layer of sand and clay can be seen in the surface layer; these are marine epiclastic sediments and eroded soil materials from the Sierra Madre mountains.

At a depth of 5 meters and lower, Makati City exhibits more prevalent sand and tuff layers, shown in Fig. 9. As a result of their position, the Northwest Barangays are dominated by sand layers. These barangays are located near the capital city of Manila, where predominant sand layers exist. The Southeast Barangay is dominated by tuff strata, as seen by the Central Plateau's widespread primary tuff and pyroclastic flow deposits. It can be noticed that Dasmariñas is dominated by clay layers, which is a result of the eroded soil materials from the Sierra Madre; nevertheless, when one descends farther, Dasmariñas, Cembo, Guadalupe Nuevo, and Pembo are dominated by tuff layers.

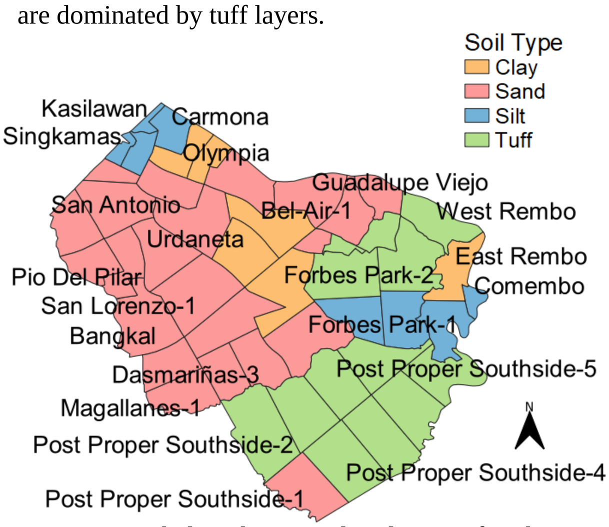

Fig. 6 1m below the ground Soil Type of Makati

These shallow clay and sand layers are critical indicators of the presence of potentially problematic soils in the area. It is advised that more testing and analyses for liquefaction and expansive soils be conducted.

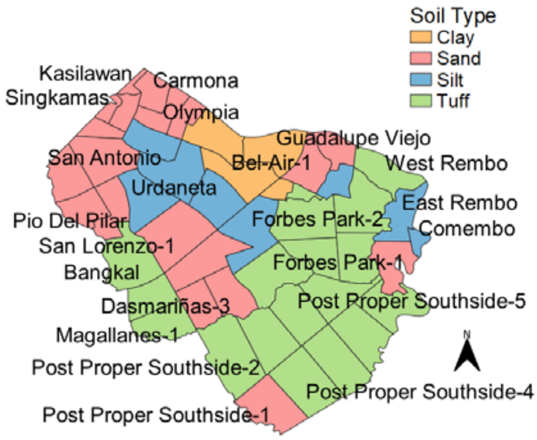

Fig. 7 2m below the ground Soil Type of Makati

4.3 Subsurface Isometric Soil Profile

The soils are usually grouped into classes, groups are based on similar physical properties and general characteristics in terms of behaviors.

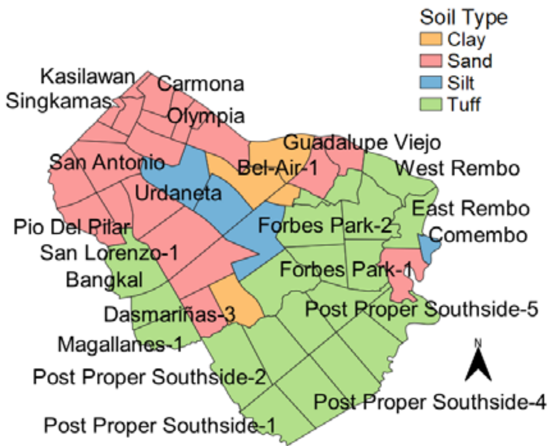

Fig. 8 3m below the ground Soil Type of Makati

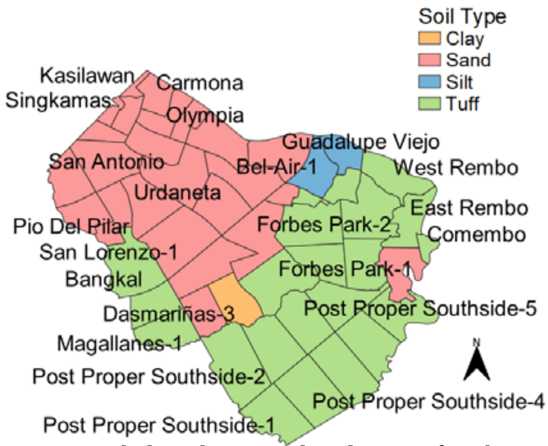

Fig. 9 4m below the ground Soil Type of Makati

As generated by the Machine Learning Model, k-Nearest Neighbor, Tuff is very prevalent in Makati City, just several meters below the ground, it would require from Soil Penetration Test (SPT) to

Rock Quality Designation (RQD) to accommodate the tuff lying beneath the clay and sand surface.

The generated soil profile is shown in Fig. 10 in terms of Azimuth and Elevation.

The Northwest Barangays are dominated by sand layers as a result of their location. These barangays are located near the City of Manila, an area with a high concentration of sand layers. As seen by the Central Plateau's vast main tuff and pyroclastic-flow deposits, the Southeast Barangay is dominated by tuff layers.

A total of 1,714 data points were generated by the k-Nearest Neighbor Model for the 49 Barangay/Zones of Makati City. The percentages of Sand, Gravel, Clay, Silt, and Tuff are 34.71%, 0.00%, 3.21%, 2.28%, and 59.81%, respectively.

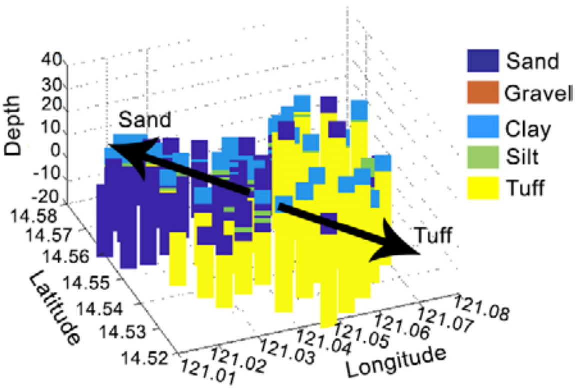

Fig. 10 Isometric Soil Profile of Makati City with -22 Azimuth-49 Elevation view

The preponderance of Tuff in the city's shallow surfaces is extremely cost-effective, as the foundations can be developed affordably due to tuff's ability to support a huge amount of weight.

There are clay, silt, and sand layers in the surface layers but several meters below the ground, early refusal levels may be expected. This is very beneficial for Engineers since they would know what to expect in the subsurface. They would have an idea of the depth to reach to avoid incurring unnecessary drilling costs.

Given that this study has generated the soil profile of Makati City through Machine Learning Models, the type of soil in the shallow layers was determined, as another form of validation, the Shear Wave Velocity Map, shown in Fig. 11, of the Philippine Earthquake Model created by the Philippine Institute of Volcanology and Seismology (PHIVOLCS) was compared.

As illustrated in Figure 10, the southeastern side of Makati City experiences a higher shear wave velocity than the northwestern side. This has the same configuration as the k-NN model-generated soil profile illustrated in Fig. 10. Tuff is considerably stronger in shear than sand. The PHIVOLCS Shear Wave Velocity Map was used to validate the k-Nearest Neighbor Model-generated subsurface soil profile.

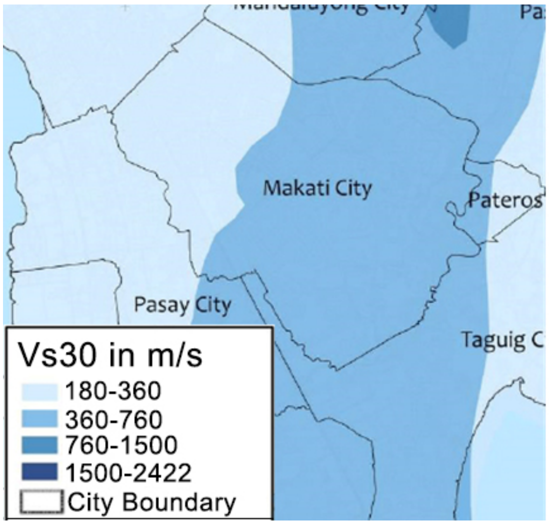

Fig. 11 Shear Wave Velocity Map showing Makati City [43]

5. CONCLUSION

Soils and rocks are natural geomaterials, and their properties are influenced by a range of spatially changing causes during their complex geological creation phase. As a result, geomaterials can exhibit a range of qualities depending on their location inside a specific site. It is beneficial to be knowledgeable of the soil profile of a target location in advance to minimize redundancy of tests, borehole depth determination, and sample methodology.

The goal of this study is to apply Machine Learning Modeling Competition to generate the soil profile, in the case of this study, Makati City, Philippines. Tree Model, Discriminant Model, Naive Bayes Model, k-Nearest Neighbor, and Artificial Neural Network are among the models. k-NN has the highest accuracy rate since it relies on distance to categorize an unknown parameter. The k-NN method presupposes the existence of comparable objects in close proximity. The k-NN hyperparameter is further tuned to determine if the accuracy rates can be improved. As a result, the study deployed a k of five (5).

In the research locale, the Northwest Barangays of Makati City is dominated by sand layers as a result of their location. These barangays are located near the City of Manila, an area with a high concentration of sand layers. As seen by the Central Plateau's vast main tuff and pyroclastic-flow deposits, the Southeast Barangay of Makati City is dominated by tuff layers. It was further validated by the Shear Wave Velocity Map of the

Philippine Earthquake Model created by PHIVOLCS.

6. REFERENCES

[1] Hu, Y., Zhao, T., & Wang, Y. (2020). Bayesian Learning of Site-Specific Spatial Variability Using Sparse Geotechnical Data. 553–558.

[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. https://doi.org/10.21660/2020.71.9203

[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). https://doi.org/10.21660/2019.58.8129

[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) doi:10.1088/1755-1315/479/1/012001

[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. doi:10.21660/2016.22.5173

[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. doi:10.21660/2018.44.7146

[12] Dungca, J. 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. doi:10.21660/2018.44.7145

[13] 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. doi:10.21660/2019.58.8123

[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. doi:10.21660/2018.44.7108

[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. doi:10.21660/2017.31.6511

[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. doi:10.21660/2017.32.6565

[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. doi:10.21660/2018.43.7170

[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. doi:10.21660/2020.66.9462

[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. doi:10.21660/2020.65.9424

[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. doi:10.21660/2019.53.14268

[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. doi:10.21660/2018.51.58567

[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. doi:10.21660/2017.31.6549

[26] Air Field Manuals. (1999, October 27). Materials Testing.

[27] USCS. (1977). The Unified Soil System. Technical Report Archive & Image Library, 1–28.

[28] De'Ath, G., & Fabricius, K. E. (2000). and regression trees: A powerful yet simple technique for ecological data analysis. Ecology, 81(11), 3178–3192.

[29] Li, X., Zhang, L., & You, J. (2019). Locally weighted discriminant analysis for hyperspectral image . Remote Sensing, 11(2).

[30] Dimitoglou, G., Adams, J. A., & Jim, C. M.(2012). Comparison of the C4. 5 and a Naïve Bayes classifier for the prediction of lung cancer survivability. ArXiv preprint arXiv:1206.1121.

[31] Statsoft. (2018, February). k-Nearest Neighbors. Retrieved from Statsoft: http://www.statsoft.com/Textbook/k-Nearest-Neighbors

[32] Hopfield, J.J., Artificial neural networks. IEEE Circuits and Devices Magazine, 1988. 4(5): p. 3-10.

[33] Bhardwaj, A. and A. Tiwari, Breast cancer diagnosis using genetically optimized neural network model. Expert Systems with Applications, 2015. 42(10): p. 4611-4620.

[34] Uçar, M. K., Nour, M., Sindi, H., & Polat, K. (2020). The Effect of Training and Testing Process on Machine Learning in Biomedical Datasets. Mathematical Problems in Engineering, 2020.

[35] I. H. Witten, E. Frank, and M. A. Hall, Data Mining-Practical Machine Learning Tools and Techniques, Morgan Kaufmann, Burlington, MA, USA, 2013.

[36] Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. Jurnal Transformatika, 18(1), 71.

[37] Ghawi, R., & Pfeffer, J. (2019). Efficient Hyperparameter Tuning with Grid Search for Text Categorization using kNN Approach with BM25 Similarity. Open Computer Science, 9(1), 160–180.

[38] JICA. General Profile of Metro Manila. Open Jica Report. JICA, https://openjicareport.jica.go.jp/pdf/12001491_02.pdf. Accessed 12 Oct 2021.

[39] Harrison, O. (2018). Machine Learning Basics with the K-Nearest Neighbors Algorithm. Retrieved September 11, 2018, from https://towardsdatascience.com/machine-learning-basics-the-k-nearest-neighbors-algorithm-6a6e71d01761.

[40] Ma, H., Nittrouer, J. A., Wu, B., Lamb, M. P., Zhang, Y., Mohrig, D., … Parker, G. (2020). Universal relation with regime transition for sediment transport in fine-grained rivers. Proceedings of the National Academy of Sciences of the United States of America, 117(1), 171–176.

[41] Siringan, F.P., Ringor, C.L., 1997. Predominant nearshore sediment dispersal patterns in Manila Bay. Science Diliman, 9(1 and 2), 29-40.

[42] Siringan, F.P., Ringor, C.L., 1998. Changes in bathymetry and their implications to sediment dispersal and rates of sedimentation in Manila Bay. Science Diliman, 10(2), 12-26.

[43] PHIVOLCS. (2017). The Philippine earthquake model.