

GEOSTATISTICAL MODELING OF SPT DATA TO DETERMINE THE MINIMUM NUMBER AND DEPTH OF FUTURE BOREHOLE INVESTIGATIONS

*Andre C. Cruz¹, Ron Gabriel O. Navarro^{1,2}, Jedidiah Joel C. Aguirre^{1,3}, Roel Nikko T Magno¹ and Julius Jose M. Derecho¹

¹Department of Civil Engineering, University of the Philippines, Los Baños, Philippines

²Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands

³Swiss Seismological Services, Department of Earth Sciences, ETH Zurich, Switzerland

*Corresponding Author, Received: 21 Nov. 2022, Revised: 12 Dec. 2022, Accepted: 02 April 2023.

ABSTRACT: The standard of determining the minimum required number and depth of boreholes is currently governed by professional judgment and is not scientifically justified in the local structural code. In this study, previous Standard Penetration Test (SPT) reports within the University of the Philippines – Los Baños Science and Technology Park Complex were used to perform geostatistical modeling using the simple kriging interpolation of GIS software to predict the SPT N-values, uncertainties, and soil type in 22.5x 22.5m cell sizes, or blocks, within the study area with depths of down to 10 meters. Statistical measures such as mean error, mean average error, root mean square error and coefficient of determination were used to validate the accuracy of the models. A superimposed raster map of the maximum SPT N standard errors of 1-10m and 6-10m layers were created, and the blocks with at most ± 5 N-values as their 95% confidence interval were identified as 0-blocks, which suggests that no borehole exploration will be necessary. Guidelines to determine the minimum required number and depth of boreholes in a specific area using these interpretations were then developed. Finally, this study established a method that requires fewer boreholes and explorations with lesser depths compared to the current code and produced a total of 401 0-blocks that may save up to 20.30 hectares of land that needs geotechnical investigations.

Keywords: *Geotechnical, Kriging, Geostatistics, SPT, GIS*

1. INTRODUCTION

Geotechnical investigation is a necessary step for every construction to provide geotechnical reports that can be used in designing appropriate footing for the structures. The most commonly used test for geotechnical investigations is the Standard Penetration Test [1]. Currently, the number of boreholes necessary to have a sensible judgment for the soil-bearing capacity to be used in designing the foundations is determined by the footprint area of the structure [2], the geostructure to be used in the building [3], or the spacing per type of structure [1, 4]. These values are only rough estimates, have no experimental or theoretical basis by the authors, and can only be considered as products of professional experience.

Geotechnical investigations can be waived by a building official if past data from adjacent areas indicates that investigation is not necessary [5]. Past boreholes from projects near the proposed structure can be used as a reference but there is a space, distance, and depth variability that needs to be considered since soil layering varies. Also, there is no assurance that the data from the nearby project sites are analogous with the soil profile in the site,

hence determining the depth of boreholes should always be maximized.

Geostatistics is a division of statistics that focuses on geographically referenced data through analysis and interpretation [6]. It is used to interpolate values for locations where samples are not taken yet. Geostatistical tools also provide measures of uncertainty for the computed values important for informed decision-making [7].

Introducing geostatistical modeling to the concept of determining the minimum required number and depth of boreholes in geotechnical investigation can offer a new method or to strengthen the decision-making process that will be backed by data and statistics.

The University of the Philippines – Los Baños Science and Technology Park Complex (UPLB-STP Complex) buildings and nearby institutions have available borehole reports from previous investigations using Standard Penetration Tests (SPT) which can be used for geostatistical mapping. The study site has considerable vacant areas for building new structures which may benefit from this study.

Hence, the study will use the geostatistical modeling of ArcGIS Pro to determine the number of boreholes and the corresponding depth required

in a study area using SPT data from previous geotechnical investigations near the site.

2. RESEARCH SIGNIFICANCE

This study will provide a new method for determining the minimum number and depth of boreholes for geotechnical investigation, specifically for low-rise buildings which are usually constructed on shallow foundations. It will set a baseline supported by statistics in determining these objectives. The method of this study may offer a more reasonable and economical approach than the methods used in the current codes which are governed by experience only [3-5]. It will also map the summary of predicted values in the study area, which is within the Science and Technology Park Complex, UPLB, for future construction projects.

3. REVIEW OF RELATED LITERATURE

3.1 Determining the Minimum Required Number and Depth of Boreholes

In the Philippines, the National Structural Code of the Philippines (NSCP) 2015 is referred to for building requirements. Table 1 below shows the code guideline for the minimum required number of boreholes.

Table 1 Minimum required number of boreholes.

Footprint Area of Structure (m ²)	Minimum Required Number of Boreholes
$A \leq 50$	1
$50 < A \leq 500$	2
$A > 500$	2+ (A/1000) (rounded up to nearest integer)

The depth considered in this study was 10 meters below the surface. According to Budhu, the stress influence of a footing is felt up to 6% at approximately 3 times its width (B), hence this was considered in determining the study's depth while assuming a maximum footing width of three meters for low-rise structures.

Table 2 Minimum borehole depth based on the type of structure.

Structure	Minimum Depth
Shallow foundation for buildings	5m or 1B to 3B, where B is the foundation width
Deep foundations for buildings	5m or 1B to 3B, where B is the foundation width
Bridge	25-30m; if bedrock is encountered, drill 3m into it

The NSCP 2015 [2] also recommends that the borehole depth should be at least 5m into the hard strata, or until a suitable bearing layer (i.e., SPT N-Value is more than 50, refusal, or coring/rock layers) is reached. Budhu [3] presented an approach to determine the borehole depth as multiples of the foundation width as presented in Table 2.

3.2 Standard Penetration Test

The Standard Penetration Test (SPT) is the most popular and widely used *in situ* test for geotechnical investigation [1]. It was developed around 1927 and is currently standardized by ASTM D1586/D1586M. The number of blows to drive the sampler up to the last two 150 mm distances of the bore is counted to obtain the SPT N number, which is commonly known as the blow count.

3.3 Geostatistical Modeling

Geostatistical interpolation started in the 1950s from the mining industry in the search for ore reserves. They used the concept to estimate the probability of ore quantity in a certain area. This idea propagated to other fields after French mathematician G. Matheron derived the formulas that founded linear geostatistics [8]. Hengl [9] stated that the standard in interpolation is the kriging and its standard version is called ordinary kriging (OK). The ArcGIS Pro mapping software extends a variation of kriging, known as simple kriging, that has normal score transformation for data that are not normally distributed. Furthermore, geostatistics can assess the uncertainty associated with spatial variability of SPT results by generating multiple probable realizations using sequential Gaussian simulation [10].

4. RESEARCH METHOD

4.1 Research Area

The proposed University of the Philippines – Los Baños Science and Technology Park Complex is located in Bay, Laguna. Borehole data of the four existing buildings inside the research area are collected as well as the recorded borehole data of nearby institutions, to serve as boundaries of the models. The geostatistical mapping area has 8.31 km² while the study area has 1.14 km² of land and a perimeter of 4.48 km.

4.2 Data Gathering and Data Preparation

Borehole reports were obtained from previous SPTs conducted inside the Science and Technology Park Complex, UPLB, and nearby areas. The N-values and the soil type, classified as clay or sand,

were recorded in 1.5m depth and also every 1m depth intervals from soil strata of depths 1m-10m.

The raw data for SPT N-values were corrected and the box and whisker plots were used to remove outliers. A total of 42 boreholes were included in the geostatistical mapping as shown in Fig. 1.

4.3 Producing the Geostatistical Model

Geostatistical mapping area was used to produce the geostatistical models while the bounded study area was the specific area that was considered in the analysis. A map area that covers the UPLB Science and Technology Park Complex was then created to streamline the boundaries of the study area. Fig. 1 presents the locations of each borehole of the final dataset, the study area that was considered in the analysis, and the geostatistical mapping area.

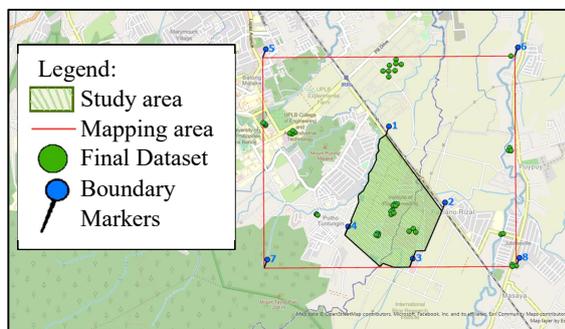


Fig. 1 Mapping areas with the final data set.

The next step was building the geostatistical models for each layer, known as GA Layers, in the GIS software. The workflow used was sourced from the ArcGIS Pro Manual. The default and optimized settings on the modeling of the software were applied for this study.

The dataset was run in three candidate models to determine the best-fit model that will be used in the study; one ordinary kriging run, and two varying simple kriging runs where an optimized setting and an adjusted variation of the optimized setting were used. The statistics of the three candidate models were computed and the best-fit model was selected as paralleled from the study conducted in assessing the accuracy of interpolation [11], specifically, the root mean square error (RMSE) and the average standard error. The optimized model of the simple kriging was used for the SPT N-values due to its higher RMSE and lower average standard error among the candidates. This best-fit model was examined using its Mean Error (ME), Mean Absolute Error (MAE), RMSE, and coefficient of determination, R^2 .

On the other hand, since the soil type is categorical data, the geostatistical mapping tool used was Empirical Bayesian Kriging (EBK) with

inputs that mimic the nearest neighbor interpolation [12].

The software has a built-in cross-validation section at the end of the modeling. In general, the idea is to remove one measured data at a time and predict its value using the rest of the measured data and compute the statistics. As the R^2 value gets closer to 1, a better fit between the data and the model will be achieved [13]. This study considered the $R^2 \geq 0.85$ criteria to have satisfactory accurate maps to be used in the decision-making [9].

4.4 Preparing the Maps for Analysis

For this study, a cell size of 22.5x 22.5m was determined based on the max footprint area of 500m² for a minimum of 2 boreholes as stated in Table 1.

After producing the GA Layers for each depth for SPT N value and soil type, raster maps were produced with a 22.5m x 22.5m cell size. The produced standard error (SE) maps were then superimposed considering the maximum values as their cell values. A superimposed map was produced using the layers 1m to 10m for the minimum required number of boreholes to produce a reference map for decision making, and another superimposed map with the layers 6m to 10m for the minimum required depth of boreholes.

For the analysis of the minimum required number of boreholes, the maximum standard error for the complete 1m-10m was considered to establish the minimum accuracy of the cell for the whole predicted exploration. Meanwhile, 6m-10m was considered in determining the minimum depth of exploration since the references recommend the minimum depth at 5m for any case.

4.5 Correlating the Standard Errors of SPT N to the Minimum Required Number and Depth of Boreholes

4.5.1 The Minimum Required Number of Boreholes

For the minimum required number of boreholes, this study considered the 1m-10m superimposed maximum standard error map. Cell locations with a 95% confidence interval (95% CI) with ± 5 N-value or lower are considered to be satisfactory enough not to conduct a borehole exploration at the grid area, and to assume the predicted values mapped in this study. It has 95% assurance that the true value of the site lies within this interval considering all the limitations of the mapping procedure. However, the engineer at the site may decide to have at least one borehole for verification. The desired range has points of at most 2.551 for their standard error as computed in the formula of 95% CI in Eq. (1) which x is defined as the predicted value in the specific cell. This equation uses the assumption that

the uncertainty in the SPT N-value follows a Gaussian distribution.

$$95\% CI = x \pm 1.96SE \quad (1)$$

The areas that were included having a value of at most 2.551 standard error were called the 0-blocks since these areas were recommended to adopt the values mapped in this study and save the necessary borehole explorations. Meanwhile, higher standard error areas were called 2-blocks since it was the standard in the code at 500m² [2].

4.5.2 The Depth of Borehole Explorations

For the depth of borehole exploration, this study considered the 6m-10m superimposed maximum standard error map. The study also used a 95% confidence interval (95% CI) with ± 5 N-value or lower, or 2.551 in standard error value, in identifying the limiting factor for the depth.

The analysis starts at the bottom part, at the 10m layer model, working upwards. The first layer from the bottom with a standard error higher than 2.551

value will be the depth of exploration since that layer would be considered as the deepest layer with an unacceptable accuracy for the prediction. Hence, it needs a borehole exploration up to that specific depth. Meanwhile, the deeper layers with acceptable accuracies were recommended to adopt the values mapped in this study.

4.6 Finalizing the Reference Maps and Guidelines

The flow of the process developed to determine the minimum required number and depth of boreholes using this study were summarized in a flowchart in Fig. 2.

The 1-block is introduced as a 2-block but with a boundary of the area of investigation that only covers 50% or less than half of the cell containing the 2-block. Meanwhile, block ratio refers to the ratio of 0-blocks to 2-blocks in the area of investigation. Sample building layouts were then assessed to determine the comparison of the method in this study to the NSCP 2015 standard.

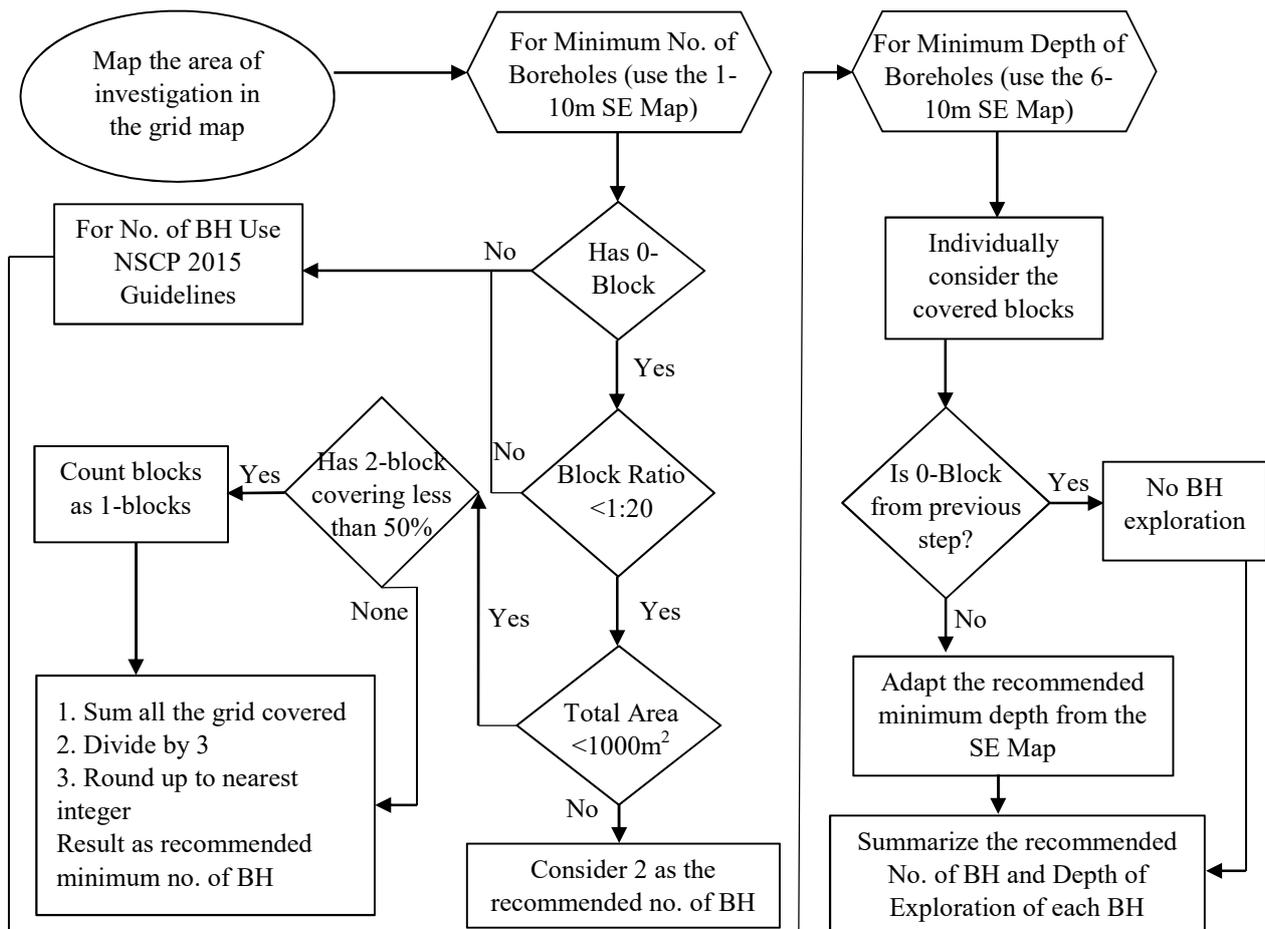


Fig. 2 Process for determining the number and depth of borehole explorations.

5. RESULTS AND DISCUSSION

5.1 Examining the Statistical Measures

The model used was examined using its Mean Error (ME), Mean Absolute Error (MAE), RMSE, and coefficient of determination (R^2). The summary is shown in Figures 3, 4, 5, and 6.

The ME has negative values in 7 out of 11 models which signifies that the model was mostly under predicting the SPT N-values. This is better than over predicting since lower values would mean more conservative predictions of SPT N-values. Additionally, the MAE showed low magnitudes of

the absolute error ranging from 0.942 of layer 1m up to 2.071 of layer 7m.

Meanwhile, the RMSE has extremes of 1.382 at layer 1m and 3.354 at layer 10m, respectively. It showed that the frequency distribution of error magnitudes of the models is relatively low and therefore, an indication of its good accuracy. Finally, all the R^2 values are above the standard set value of 0.85. Since 10 out of 11 layers have a higher than 0.90 R^2 -value, the model produced from the data set is highly accurate in predicting the values of the model. It indicates that as much as around 90% of the variance of the predicted value can be explained by the data set provided.

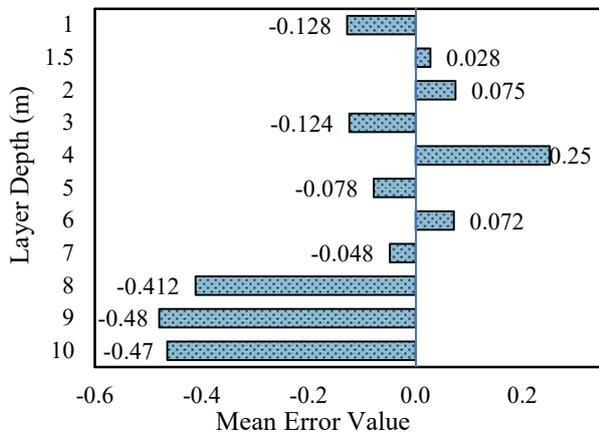


Fig. 3 ME Values of the resulting 11 models

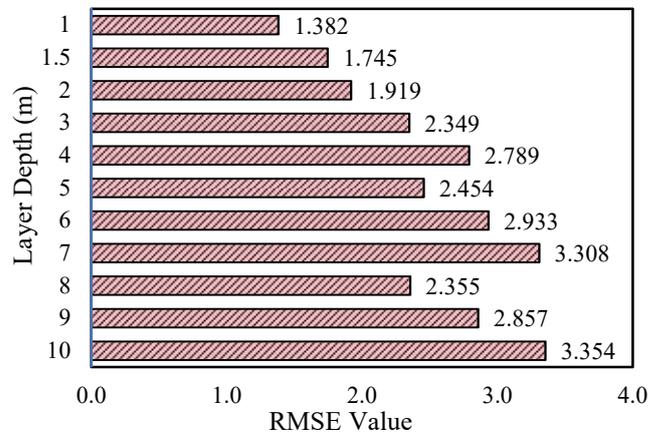


Fig. 5 RSME Values of the resulting 11 models

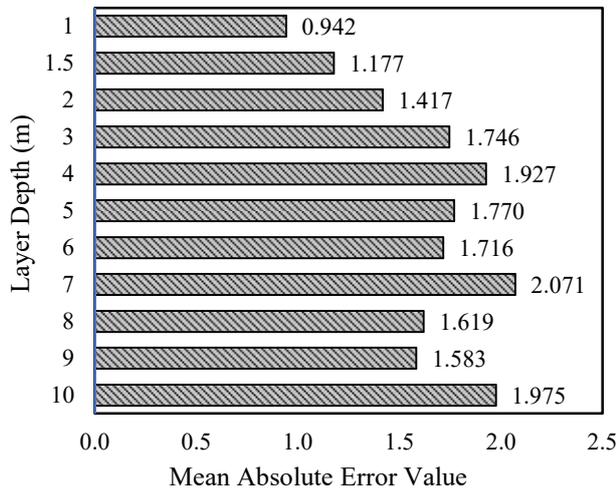


Fig. 4 MAE Values of the resulting 11 models

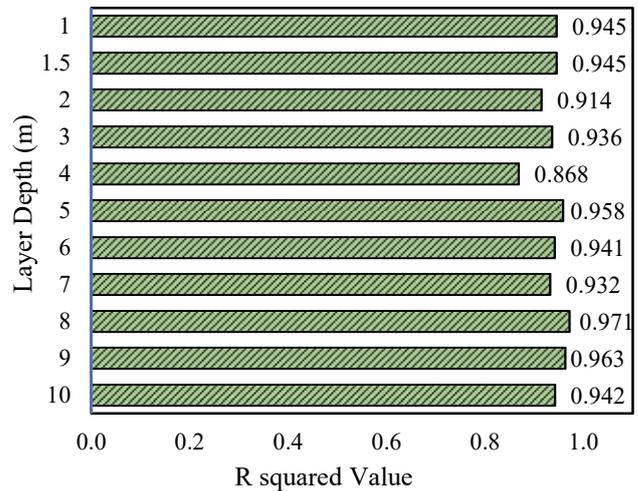


Fig. 6 R squared Values of the resulting 11 models

5.2 Prediction Value Maps

The prediction value (PV) and standard error maps of SPT N data were collected as shown in Fig. 7. On the other hand, the PV of the soil types is collected in Fig. 8.

It can be observed that in all layers or depths except layer 3m, the N-values were mostly below

25. This indicates loose to medium-dense sand and soft to stiff clays in the study area. In the deeper layers, it can be observed that the higher SPT N-values were skewed on the right side of the study area. These can be contributed to the high SPT N-values at the bottom right of the geostatistical mapping area which is mostly sand.

Since the soil type is categorical data and the Empirical Bayesian Kriging was used, the predicted values are highly dependent on the nearest measured points. The predicted value maps show that for the first two meters of the soil depth, clayey soil is the dominant composition in the study area. Meanwhile, for the 3-4 m layers, sand started to govern, and clay was left mostly on the lower right side.

For the 5-10 m except for the 7 m, clay occupies most of the top part of the study area while sand dominates the bottom part. The soil type information collected is necessary for foundation design, because governing soil bearing capacity equations are dependent on the soil type. Hence, the purpose of mapping the soil type alongside the SPT-N Values.

5.3 Reference Map in Determining the Minimum Required Number of Boreholes

The standard error maps of the SPT N-values were superimposed from 1m-10m and the maximum value for each block and the blocks with at most ± 5 N-values as their 95% confidence interval were identified as 0-blocks, which suggest that no borehole exploration will be necessary. These blocks are shown in Fig. 9 with blue shades.

It created 401 blocks that can be considered as 0-blocks, where 347 blocks have between ± 3 to ± 5 standard errors and 54 blocks had less than ± 3 standard errors of 95% confidence. The blue regions consisted of 17% of the total study area or 20.30 hectares of land area.

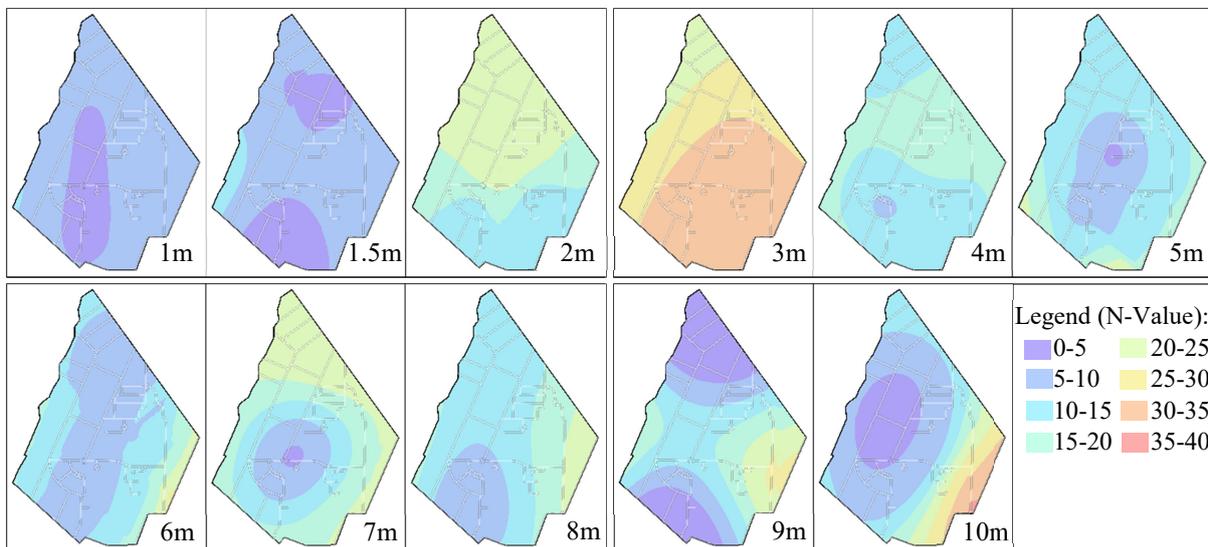


Fig. 7 Raster maps of SPT-N Values per layer

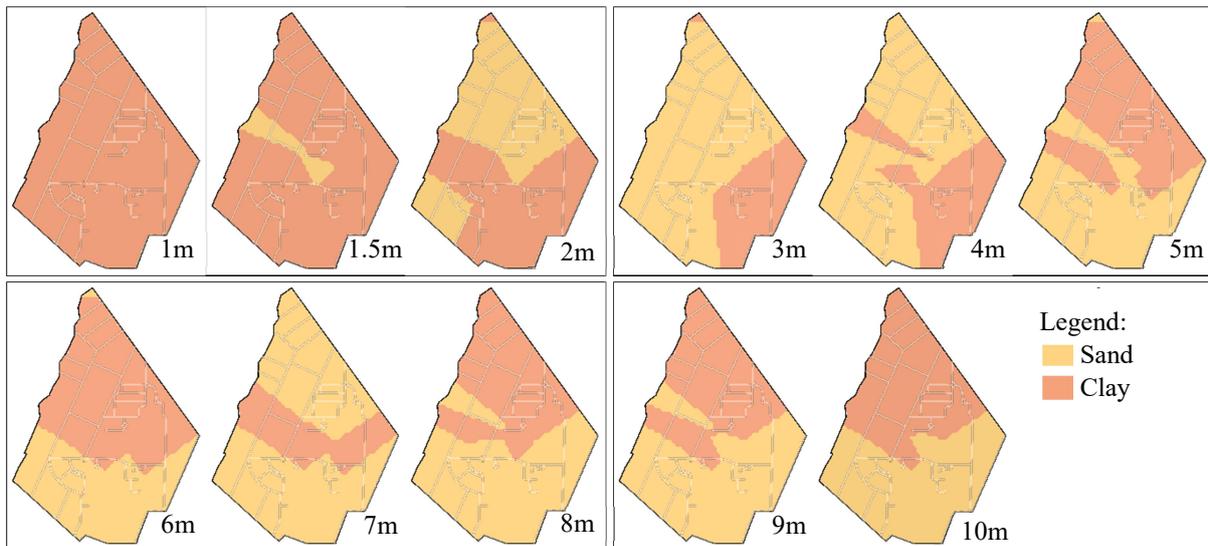


Fig. 8 Raster maps of Soil Types per layer

5.4 Reference Map in Determining the Minimum Required Depth of Boreholes

The standard error maps of the SPT N-values were superimposed from 6m-10m, and the first layer from the bottom with at most ± 5 N-values as their 95% confidence interval will be the recommended depth of borehole for that specific area. The summary map as shown in Fig. 10 verifies that deeper exploration is needed as the area of investigation goes farther from the known boreholes.

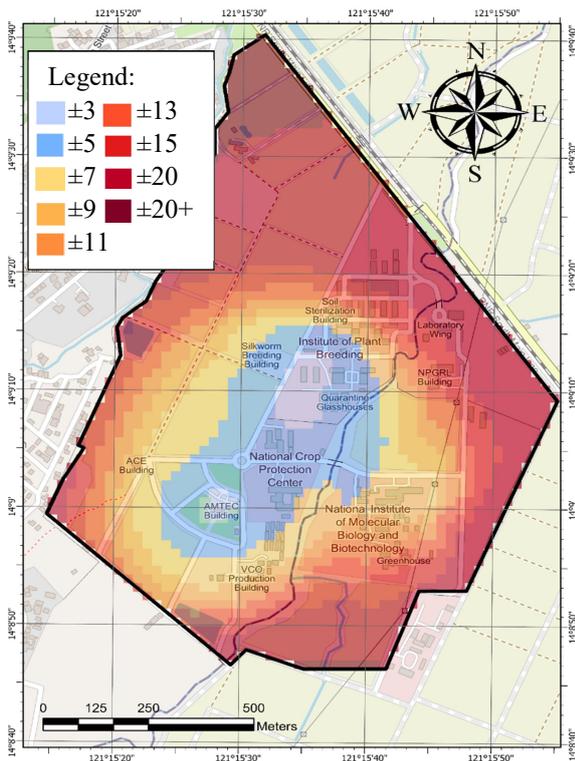


Fig. 9 Superimposed raster map for maximum standard errors for 1m-10m.

The map showed 470 cells that recommended a minimum depth of exploration of 5m, while 44, 3, 343, and 1390 cells recommended 7m, 8m, 9m, and 10m respectively. It should be noted that there is no recommendation for a 6m layer. This shows that the cells that have not been captured by the deeper layers yet at the 6-meter layer have acceptable accuracies. Consequently, the recommended depth for these cells leads directly to the next upper layer, which is the minimum depth of 5m, hence, the reason why there is no recommendation of 6m depth in the map.

5.5 Discussion on the Assessment of the Results

Three arbitrary building layouts were placed in the study area and the minimum required number

and corresponding depths of boreholes were computed based on the steps in Fig. 2 and by using the NSCP 2015 code for a 3m foundation width.

5.5.1 The assessment of building layouts for the minimum required number of boreholes

The site layout and the corresponding SE map (1m-10m) of the arbitrary building layouts is shown in Fig. 11 while the comparison of computed minimum required number of boreholes are shown in Table 3.

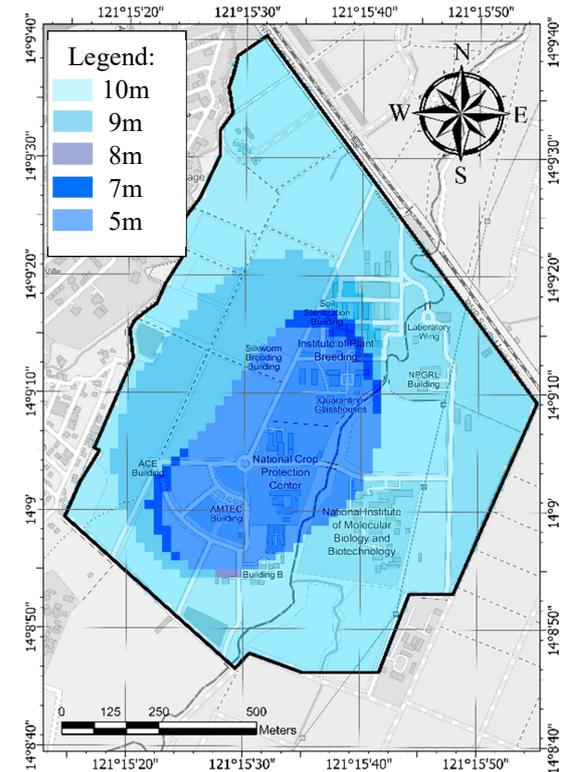


Fig. 10 Superimposed raster map for maximum standard errors for 6m-10m.

Table 3 Comparison of the computed minimum required number of boreholes.

Bldg	Area (m ²)	NSCP code	Recommended No. of Boreholes
1	2278	5	1
2	3037	6	2
3	5062	8	6

For building area 1, most of the area was covered with 0-blocks. Based on the geometry of the layout, it is proposed to conduct the borehole exploration at the top left most of the area. Additional borehole exploration can also be done to verify the accuracy of the 0-blocks or include another borehole on the 2-block to validate the SPT with two samples in that area.

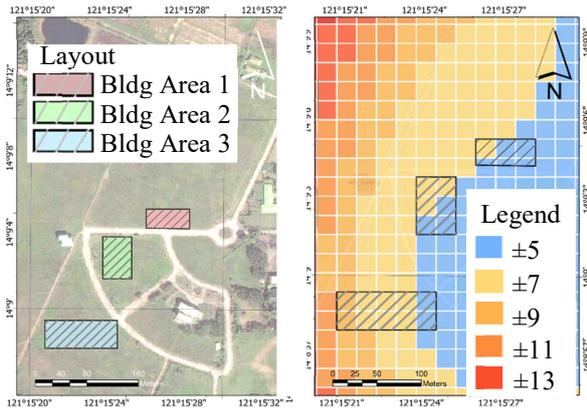


Fig. 11 Site layout and the corresponding SE map (1m-10m) of the arbitrary building layouts.

For building area 2, there are equal numbers of 0-blocks and 2-blocks. It is proposed to conduct the borehole exploration at the middle top area at an equal distance. Adding a borehole would lay out the borehole points to be well-spaced in the 2-blocks.

For building area 3, only a fifth of the area covers 0-block It is recommended to place the six borehole explorations equally spaced leaning to the right of the blocks of the 2-block areas.

It can be observed that for all the building layouts, the recommended number of boreholes obtained from this study is lower than that of NSCP 2015.

5.5.2 Assessment of building layout No. 3 for the minimum required depth of borehole exploration

As an illustration, building area 3 was used to assess the comparison of recommended minimum depths. Table 4 shows the recommended depth of boreholes for Building No. 3, while the corresponding SE map (6m-10m) of the arbitrary building layout No. 3 is shown in Fig 12.

Table 4 Comparison of computed minimum depths of exploration for Bldg. 3.

Grid code	Rec. Min Depth (m)	NSCP code (m)
1695	9	10
1696	9	10
1697	7	10
1698	5	10
1699	0-block	10
1741	9	10
1742	9	10
1743	7	10
1744	7	10
1745	0-block	10

The building layout No. 3 has two 0-blocks based on the previous assessment, and therefore is

recommended not to conduct a borehole exploration in these areas.

Meanwhile, the recommended minimum depth of borehole exploration ranges from 5m to 9m on the other cells. The values of the depths are relatively higher as the locations are seen farther from the known points. This supports the concept that accuracy diminishes with distance.

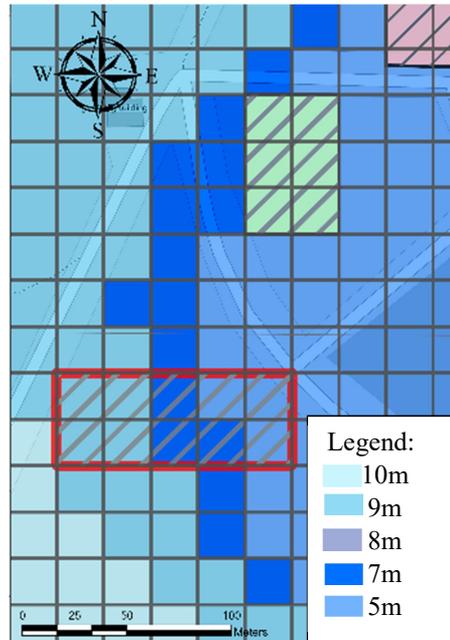


Fig. 12 The corresponding SE map (6m-10m) of the arbitrary building layout No. 3.

It can be observed also that since there is a reference in determining the depth of exploration, it is always lower than the recommended depth of NSCP 2015 which is 10m in the area based on the assumed maximum foundation width of 3m for shallow foundations of low-rise structures.

6. SUMMARY AND CONCLUSIONS

Standard Penetration Test (SPT) data were collected in the University of the Philippines – Los Baños Science and Technology Park Complex buildings and its nearby area. Geostatistical modeling of three candidates was then performed for 11 layers (1.5m and 1m-10m with 1 m intervals) for the SPT N value. The models for each layer with the lowest average standard error and root mean square error (RMSE) were selected as the best-fit model and were evaluated.

The maps were analyzed statistically; by using their mean, absolute mean, RMSE, and R^2 , which had a conclusion that they are satisfactory models that can be used for decision-making. Visually, the layers had generally low SPT N-values aside from the layers 2m, 3m, and 4m. Meanwhile, the

standard errors of the layers were superimposed in 1m-10m and 6m-10m to produce maximum standard error maps that will be used for the decision-making on determining the minimum required number and depth of boreholes, respectively. Guidelines for determining the minimum required number and depth of boreholes using the reference maps were then developed.

After the methods for determining the minimum required number of boreholes, there are a total of 401 blocks that can be considered as known points with at most ± 5 range of values, where 347 blocks have between ± 3 to ± 5 intervals and 54 blocks have less than ± 3 intervals. This study resulted to save as much as 20.30 hectares of land area on borehole explorations in the future in the study area.

Meanwhile, for determining the minimum required depth of exploration for necessary boreholes, the produced reference map showed 470 cells that recommended 5m depth or the minimum depth of exploration, while 44, 3, 343, and 1390 cells recommended 7m, 8m, 9m, and 10m respectively. All recommended depths are always lower than or equal to the recommended depth of the NSCP 2015 for low-rise structures assuming a maximum footing width of 3 meters.

Finally, this study shows that the minimum required number and depth of boreholes for future geotechnical investigations in an area can be determined by geostatistical mapping of existing SPT data and the results show that the number and depth of boreholes will be less than the NSCP Code. This will highly benefit locations with abundant and nearby borehole data recorded from prior geotechnical investigations, such as cities, industrial parks, subdivisions, communities, and business centers.

7. RECOMMENDATIONS

This study recommends modeling the area considering high-rise buildings; however borehole data up to 20 to 30m depth must be considered. Also, since only three candidate models were compared, it is recommended to conduct more candidate models to improve the accuracy. A deeper study into the kriging interpolation and its parameters to optimize the models may also be considered for further studies.

The study assumed that the area is homogenous in the interpolation of the software and have no erratic subsurface condition such as sinkholes or unusual soil profiles (10). If the area is identified to have an erratic stratification then this method is not recommended.

Likewise, the objective of the study which was determining through geostatistical mapping the minimum required number and depth of borehole exploration for future geotechnical investigations is

highly recommended to be done in other locations and study areas to verify and utilize the concept. Lastly, an SPT N Value with a soil type database is recommended to be made so that users can identify the SPT N Value and soil type at their specific grid code or location of interest.

8. REFERENCES

- [1] Das, B.M. and Sobhan K., Principles of Geotechnical Engineering, 9th Edition, 2018, pp. 740-750.
- [2] ASEP, National Structural Code of the Philippines 2015, 7th Ed., 2015, pp. 3-8.
- [3] Budhu M., Soil Mechanics Fundamentals, John Wiley & Sons, Ltd Registered, 2015, pp.76-155.
- [4] Look B.G., Handbook of Geotechnical Investigation and Design Tables. Handbook of Geotechnical Investigation and Design Tables, 2014, 2nd Ed., pp. 2-38.
- [5] ICC, 2021 International Building Code, 2020, pp. 488-515.
- [6] Goovaerts P., Geostatistics for Natural Resources Evaluation (Applied Geostatistics), Oxford University Press, Ney York, 1997, pp. 496.
- [7] ESRI ArcGIS Pro, The geostatistical workflow, <https://pro.arcgis.com/en/proapp/latest/help/analysis/geostatistical-analyst/the-geostatistical-workflow.htm>, 2020.
- [8] Cressie N. A. C., The origins of kriging, *Mathematical Geology*, Vol. 22, Issue 3, 1990, pp. 239-252.
- [9] Hengl T., A Practical Guide to Geostatistical Mapping in Government Publications Review, Vol. 13, Issue 6, 2009, pp. 15-30.
- [10] Basarir, H., Kumral, M., Karpuz, C., and Tutluoglu, L., Geostatistical modelling of spatial variability of SPT data for a borax stockpile site. *Engineering Geology*, Vol. 114, Issues 3-4, 2010, pp. 154-163.
- [11] Zirakbash T., Admiraal R., Boronina A., Anda M., and Bahri P.A, Assessing Interpolation Methods for Accuracy of Design Groundwater Levels for Civil Projects. *Journal of Hydrologic Engineering*, Vol. 25, Issue 9, 2020, pp. 4.
- [12] Mackey M., Explore Subsurface Geology Using Voxel Layers in Three Easy Steps, ESRI Canada, 2021, <https://resources.esri.ca/getting-technical/explore-subsurface-geology-using-voxel-layers-in-three-easy-steps>.
- [13] Barrera-Gonzalez J., Lavado Contador J.F., and Pulido Fernandez, M., Mapping Soil Properties at a Regional Scale vs. Geostatistical Interpolation Methods at Different Soil Depths. *Sustainability*, Vol. 14, Issue 16, 2022, pp. 1-20.