buildings

MDPI

Article

Developing Nomographs for the Unit Weight of Soils

Jonathan Dungca * and Joenel Galupino

Department of Civil Engineering, De La Salle University, Manila 1004, Philippines; joenel.galupino@dlsu.edu.ph
* Correspondence: jonathan.dungca@dlsu.edu.ph

Abstract: Engineers have created increasingly complex correlations based on laboratory and field tests. Over time, geotechnical engineering modeling techniques have evolved from simple analytical methods to complex numerical modeling techniques. Nomographs are traditional computational tools that have been widely employed in engineering. Combining nomographs with computational tools such as numerical models and machine learning algorithms can lead to better outcomes. Thus, this study aimed to develop a nomograph for geotechnical engineering that incorporates machine learning, specifically for the unit weight of soil. Four calibrated models were developed to determine the unit weight of soil: the moist unit weight of coarse-grained soil, the saturated unit weight of coarse-grained soil, the moist unit weight of fine-grained soil, and the saturated unit weight of fine-grained soil. An uncertainty test was conducted for the data used. Our results indicated a strong positive relationship to most of the models. The generated nomographs were tested in Malabon, a city in Metro Manila, where a low unit weight of soil was determined. This low unit weight was validated by the predominance of alluvial deposits and the shallow groundwater table, which soften and weaken the soil.

Keywords: nomograph; unit weight; machine learning; Philippines; estimate

1. Introduction

In the eighteenth century, advancements in basic soil testing techniques, such as visual inspection and manual testing, were developed. Significant geotechnical engineering advancements have occurred in earthworks and foundations, wherein they developed methods for constructing foundations in difficult soils [1]. Engineers also developed increasingly complex correlations based on laboratory and field tests. For example, based on laboratory testing, Terzaghi developed correlations between soil strength and parameters such as moisture content, void ratio, and over-consolidation ratio. These correlations were derived from historical laboratory or field data [2] and are typically processed by a model that can efficiently analyze and process this massive amount of data to identify trends, patterns, and statistical relationships. Over time, modeling techniques in geotechnical engineering evolved from simple analytical methods to complex numerical modeling methods [3]; as an outcome of this progress, modern geotechnical engineers have begun to use machine learning and artificial intelligence techniques to analyze large datasets and identify patterns and correlations that may not have been recognizable using traditional analysis techniques [4].

Nomographs are traditional computational tools that have been widely used in engineering, and they were invented in 1880 by Philbert Maurice d'Ocagne [5]. Nomographs were commonly used in numerous academic disciplines, including various fields in engineering.

Nomographs have applications in numerous engineering disciplines. In mechanical engineering, they aid in estimating crucial design parameters such as gear ratios and power transmission efficiency [6]. The visual representations of nomographs aid engineers in comprehending the complex variable interactions in mechanical systems, resulting in improved design decisions. Likewise, nomographs are utilized in chemical engineering for tasks such as material balance calculations, distillation, and heat exchanger design [7].

Citation: Dungca, J.; Galupino, J. Developing Nomographs for the Unit Weight of Soils. *Buildings* **2023**, *13*, 2315. https://doi.org/10.3390/buildings13092315

Academic Editors: Suraparb Keawsawasvong and Mengmeng Lu

Received: 18 July 2023
Revised: 9 August 2023
Accepted: 11 September 2023
Published: 12 September 2023

These tools enable engineers to rapidly estimate flow rates, temperatures, and chemical compositions, which is crucial for optimizing chemical processes. The graphical nature of nomographs facilitates the comprehension of the relationships between reactants, products, and process conditions.

Although nomographs have declined with the development of more advanced computational tools, they remain helpful in providing quick and accurate graphical calculations in specific situations where a portable calculator is required [8]. In this manner, nomography is still considered a valuable computational tool in academic settings since it provides a better understanding of complex formulas, how their variables are associated, and the sensitivity of the results to changes in those variables [9].

Many studies in civil engineering have utilized nomographs in construction [10–13], hydraulics and water resources [14–20], transportation engineering [21–23], and structural engineering [24,25]. In the construction industry, nomographs can be utilized for tasks such as determining concrete mix proportions, estimating material quantities based on project parameters, analyzing construction costs versus various variables, and optimizing resource allocation. Nomographs aid in designing water distribution systems, stormwater management plans, and flood control measures by assessing the relationships between variables such as flow rates, pipe sizes, and velocities in hydraulics and water resources. Nomographs are also valuable for transportation engineering for simplifying complex calculations and facilitating decision-making. These graphical charts allow designers to promptly assess relationships between various transportation parameters, such as vehicle speed, travel time, distance, and fuel consumption. Nomographs help optimize route planning, estimate travel costs, analyze traffic flow characteristics, and determine vehicle performance based on grade and load. In structural engineering, nomographs are helpful tools that simplify complex calculations and facilitate decision-making. Through using these graphical charts, engineers can evaluate relationships between various structural parameters, such as load, span length, material properties, and deflection. Nomographs can be used to determine suitable beam sizes, column capacities, and foundation dimensions, making them useful for preliminary design and quick calculations.

However, there have been limited studies on using nomographs in geotechnical engineering, creating a significant opportunity for their application in this field.

Despite being considered traditional in geotechnical engineering, nomographs are still relevant in spite of the advent of advanced computational tools such as numerical models and machine learning algorithms. These modern tools offer more precise solutions and additional insights into soil behavior. Recent studies have provided valuable insights into the application of machine learning in geotechnical engineering [26–30]. These contributions contribute to a deeper awareness of incorporating machine learning techniques in geotechnical engineering.

Combining nomographs with modern computational tools such as numerical models and machine learning algorithms can lead to even better results by allowing for a more comprehensive approach to solving complex problems. For this paper, a nomograph was developed for geotechnical engineering, specifically for the parameter of unit weight, in conjunction with machine learning. The developed nomograph was deployed in Metro Manila, Philippines, to provide valuable insights into soil behavior in this area.

2. Methodology

2.1. Data Collection and Cleansing

A systematic search was conducted to collect empirical data pertinent to the topic from previously published research papers to support the development of the geotechnical engineering nomograph. This entailed digitizing and organizing data from multiple sources to understand existing data and comprehensively improve future research [31]. To ensure the precision and applicability of the data, a calibration procedure was conducted to validate the collected data [32]. The required information included the author(s)' names, a description of the soil, the location of the study, the number of data points, the unit weight, and the SPT N-Value. These empirical data were then used as inputs for the machine learning model and source data for the programming code.

The unit weight of soil is an essential factor in determining its bearing capacity. This weight is affected by several variables, including moisture content, particle composition, and degree of compaction [3]. Commonly, the soil is categorized as coarse-grained or fine-grained [3,33] (a sample is shown in Table 1), as well as moist unit weight and saturated unit weight [4,34]. Unit weight, expressed in kilonewtons per cubic meter (kN/m^3), calculates total and effective stresses. However, unit weight information is typically not included in exhaustive geotechnical investigation data. Consequently, correlations are commonly employed to estimate the unit weight of the soil in each layer.

Table 1. Correlation of unit weight and SPT N-Value [33].

SPT N-Value	Unit Weight (kN/m^3)
0–4	11.00–15.71
4–10	14.14–18.07
10–30	17.28–20.42
30–50	17.28–22.00
>50	20.42–23.56

In determining the unit weight of soils, the general soil group to which the soil belongs is an essential factor. Soils are typically divided into two broad categories based on their characteristics: coarse-grained soils and fine-grained soils. Based on their properties, coarse-grained soils are further classified into several subcategories. Among these are GW (well-graded gravel), GP (poorly graded gravel), GM (silty gravel), GC (clayey gravel), SW (well-graded sand), SP (poorly graded sand), SM (silty sand), and SC (clayey gravel) (clayey sand). Similarly, fine-grained soils are subdivided into ML (silt), CL (clayey silt), OL (organic silt), MH (silty clay), CH (clayey sand), and OH (organic clay).

The drive cylinder method is a common technique for calculating the unit weight. This technique entails collecting an undisturbed soil sample on site and measuring its volume with a cylindrical instrument, typically a drive cylinder. The soil sample is then weighed to determine its moisture content. The bulk density of the soil is determined by dividing the soil's moist mass by the sample's volume. After determining the soil's bulk density, the sample is dried in an oven to remove all moisture. After weighing the soil's dry mass, the dry density is determined by dividing the dry mass by the sample volume, shown in Equation (1). Bulk and dry density are crucial parameters for calculating the unit weight of soil. Unit weight is simply the product of bulk density and gravitational acceleration, shown in Equation (2). Dry density and dry unit weight can also be determined using Equations (3) and (4).

$$\rho_t = \frac{M_t}{V} \tag{1}$$

$$\gamma_t = \rho_t * g \tag{2}$$

$$\rho_d = \frac{\rho_t}{1+w} \tag{3}$$

$$\gamma_d = \rho_d * g \tag{4}$$

where V is the volume of the soil sample in cm^3, ρ_t is bulk density in g/cm^3, M_t is the mass of the soil specimen in g, $\gamma_d\gamma_t$ is the dry unit weight, g is the acceleration due to gravity, ρ_d is the dry density g/cm^3, and γ_d is the dry unit weight.

Another essential in situ test is the Standard Penetration Test (SPT), which provides valuable information for site investigation. The test involves driving a standard sampler with a 51 mm outer diameter and a 63.5 kg weight that falls freely from a 760 mm height into the soil. The first 150 mm of soil are discarded, and the following 300 mm are used for the test. N-value refers to the number of blows required to push the sampler through 300 mm of soil. When an SPT N-value reaches 50, a rock layer is encountered during drilling. Thus, a rock coring method may be required.

Based on the findings of previous studies, correlations between SPT N-value and soil unit weight were determined. These correlations were used to create geotechnical engineering-useful nomographs. The relationship between the SPT N-value and the unit weight of soil with the various classifications used for modeling is presented in Table 2.

Table 2. Correlations of SPT N-value and the unit weight of soil.

Soil Layer Location	Reference Parameter	Correlating Parameter
Above Groundwater Table	SPT N-Value	Moist unit weight of coarse-grained soils Moist unit weight of fine-grained soils
Below Groundwater Table		Saturated unit weight of coarse-grained soils Saturated unit weight of fine-grained soils

2.2. Data Cleansing and Addressing Uncertainty

This study assumes the implementation of strict quality control measures during SPT field operations. This includes ensuring that the testing equipment is calibrated correctly, using standardized procedures, and adhering to established standards. Thus, it is assumed that the SPT N-value collected on site is reliable. However, data cleansing was performed. Data cleansing is an essential step in data management that involves the removal of inaccurate, corrupted, improperly formatted, duplicate, or incomplete data from a dataset. Multiple data sources can result in issues such as identical or incorrectly labeled data, making it essential to clean the data before storing or processing it. To ensure the integrity and acceptability of the dataset, the data must be processed and verified thoroughly. The data were checked on their consistencies in the reports, including identifying and rectifying errors, standardizing formats and labels, and eliminating irrelevant or incomplete data. Also included in the collected data are the equipment specifications, testing procedures, data collection process, and any identified limitations or uncertainties associated with each SPT test. This ensures transparency and allows other researchers or practitioners to comprehend and evaluate the SPT data's reliability and implications. In addition, as stated in the local code, conducting multiple SPTs at each location or area can aid in addressing the problem of low repeatability. The test results' variability and identifying any outliers or inconsistencies can be evaluated by taking multiple measurements. This results in a more accurate depiction of the subsurface conditions.

2.3. Modeling

Machine learning is a powerful data processing technique that automates model creation. Artificial intelligence is a subfield that enables systems to learn from data, recognize patterns, and make decisions with minimal human intervention. The significance of machine learning lies in its capacity to analyze large, complex datasets and generate more accurate, faster results on a massive scale, even when working with huge amounts of data.

In this study, regression models were utilized to determine the relationship between unit weight and SPT N-values, which required the manipulation of numerical data. The study employed numerous regression models, including tree, linear, quadratic, ensemble, and neural network models. Each of these models brought distinctive advantages, enabling us to better understand the underlying relationships between the parameters.

A tree model labels, records, and assigns variables to discrete classes; the tree model used in this study was constructed through a process known as binary recursive partitioning, which employs an iterative process of dividing the data into partitions, followed by further partitioning to achieve the desired accuracy rate. A neural network is a collection of algorithms that attempts to identify the underlying relationships in a dataset by simulating the human brain. Each input parameter is referred to as a "neuron", and each neuron receives input values with assigned weights. The weighted inputs are then summed within the node, and an activation function is applied to obtain the results. In addition, the relationships in linear regression were modeled with linear predictor functions whose unknown model parameters were estimated from the data in linear and quadratic modeling. Lastly, for ensemble modeling, this model combines the predictions of multiple base estimators built with a given learning algorithm to improve generalizability/robustness compared to a single estimator.

Improving the performance of a model necessitates a comprehensive understanding of the used datasets. Researchers often divide their datasets into three portions: 70% for training, 15% for validation, and 15% for testing. This enables them to train the model on a sufficiently large dataset while reserving sufficient data for validation and testing.

Once the model has been trained on these data, it can be optimized via hyperparameter tuning. This involves adjusting various model parameters to find the optimal combination that yields the best results. By changing the model in this manner, higher levels of accuracy were achieved.

Validation datasets are an essential instrument for assessing the performance characteristics of a model. Coefficient of determination (R^2) and root mean square error (RMSE) are frequently employed metrics for evaluating a model's performance.

The coefficient of determination quantifies a statistical model's ability to predict an outcome and is frequently used to evaluate the strength of the association between two variables. In a positive relationship, one variable tends to increase when the other variable increases. Higher values indicate a stronger relationship between the variables, shown in Table 3.

Table 3. Interpretation of the Coefficient of Determination (R^2).

Coefficient of Determination (R^2)	Interpretation
>0.70	Very Strong Positive Relationship
0.40–0.69	Strong Positive Relationship
0.30–0.39	Moderate Positive Relationship
0.20–0.29	Weak Positive Relationship
0.01–0.19	Negligible Relationship
0	No Relationship

In contrast, the root mean square error uses Euclidean distance to measure the distance between the model's predictions and the actual values. Lower values of the root mean square error indicate a better fit, indicating that the model can accurately predict the measured outcome.

In defining the architecture of a machine learning model, hyperparameters play a crucial role, and the process of determining the optimal set of hyperparameters is known as hyperparameter tuning. The primary goal of hyperparameter tuning is to determine the optimal values for these parameters, which can help to reduce error and enhance the overall performance of a trained model.

Several key hyperparameters were adjusted in this study to achieve optimal results. These included the number of layers for the neural networks, the polynomial degrees for linear models, and the number of models participating in ensembles. The winning model was identified by testing various hyperparameter values and comparing the resulting performance of the models.

MATLAB R2022b was this study's primary programming language and numerical computing environment. As a proprietary programming language, MATLAB offers a variety of tools and features designed to facilitate data analysis, modeling, and visualization.

Data preprocessing tasks involving data manipulation, transformation, cleaning, and normalization, which are fundamental steps in preparing data for machine learning models, were performed using MATLAB. Due to its comprehensive collection of machine learning algorithms and regression-focused libraries, it was also utilized in model development. This allowed for the development and testing of various model constructions and training algorithms. Standard metrics such as RMSE and R^2 were used to evaluate the performance of machine learning models utilizing MATLAB.

MATLAB also played a role in hyperparameter tuning, which optimizes algorithm performance by adjusting hyperparameters. Last but not least, the platform participated in deployment by integrating the chosen model into the existing syntax for automation purposes.

2.4. Addressing Uncertainty

Several steps are required to conduct a statistical uncertainty test and determine a confidence interval. Initially, it is essential to define the SPT dataset, ensuring that the variables of interest are specified precisely. Next, it is necessary to select the 95% confidence level, representing the desired certainty level for the interval.

The *t*-test is then followed by calculating relevant test statistics, such as comparing means and proportions.

The confidence interval can then be computed by adding and subtracting the margin of error from the parameter's point estimate. This interval provides an estimate of the plausible range in which the actual parameter of the population exists with the desired level of confidence. Finally, context must be considered when interpreting the confidence interval.

2.5. Nomographs

Nomographs were created to represent the summary of calibrated empirical models established for unit weight and SPT N-Value. A nomogram is a collection of n scales, one for every variable in an equation. A standard nomograph depicts the relationship between variables and their respective scales; many of the earliest published nomographs were created empirically. The dependent and independent scales of delineated nomographs are shown in Table 4.

Table 4. Dependent and independent scales of the delineated nomographs.

Independent Scale	Dependent Scale
SPT N-Value	Moist unit weight (coarse-grained soils) Moist unit weight (fine-grained soils) Saturated unit weight (coarse-grained soils) Saturated unit weight (fine-grained soils)

Nomographs are graphical devices that can be generated differently to solve mathematical equations. The first method involves Python programming and using code to create a nomograph. The second method is a computer-aided design (CAD) approach in which the nomogram is manually created using specialized software. The nomograph used in this study was created manually using AutoCAD 2023 software. The method required a graduated scale to precisely calibrate the divisions for both the dependent and independent parameters. This method ensured the precise alignment of these parameters on the nomograph, allowing for a trustworthy visual evaluation of their interrelationships.

Given the values of certain known variables, it is possible to determine the value of an unknown variable by aligning a straightedge (also known as an isopleth) across the known values on the scales and then reading the value of the unknown variable where the straightedge intersects the scale.

Please note that the findings and interpretations presented in this study are based on the assumption that Standard Penetration Test (SPT) field operations implemented strict quality control measures. These measures included correctly calibrating testing equipment, standardized procedures, and adherence to predetermined standards. Assuming that the SPT N-values collected on site are accurate, data cleansing was performed as a necessary step in data management. In addition to the collected SPT data, detailed information regarding equipment specifications, testing procedures, the process of data collection, and any identified limitations or uncertainties associated with each SPT test was included. Including such information is intended to increase transparency and permit other researchers or practitioners to comprehend and assess the validity and implications of the SPT data.

2.6. Case Study

This research was conducted in Metro Manila, the National Capital Region (NCR), which encompasses 619.57 square kilometers, as shown in Figure 1. Metro Manila consists of sixteen cities, which are further subdivided into 1690 barangays, and one municipality.

Figure 1. Map of Metro Manila, Philippines.

As the capital region of the Philippines, Metro Manila is the economic, cultural, and political hub.

In this study, the SPT N-value of selected locations and layers around Metro Manila were gathered and processed for the unit weight using the generated nomographs.

3. Results and Discussion

3.1. Data Collected

This study aims to estimate the unit weight of soil by analyzing data from multiple previously conducted studies. Several critical pieces of information must be included to ensure the collected data are organized and easily understood. These include the name(s) of the author(s), a detailed description of the soil according to the original research, the location of the study, the Standard Penetration Test (SPT) N-value, and the unit weight.

The empirical data consists of 458 data points, shown in Table 5, which can be categorized into four groups: moist-coarse, moist-fine, saturated-coarse, and saturated-fines.

Table 5. Collected data.

No. of Data	Description
158	Moist-coarse
120	Moist-fine
94	Saturated-coarse
86	Saturated-fine

The dataset is classified according to soil type and water table depth. The first category consists of 158 data points and is named moist-coarse. These data points represent measurements from coarse-grained soils above the groundwater table, such as sand and gravel. The second category, moist-fine, consists of 120 data points. These data points represent measurements from fine-grained soils above the groundwater table, such as silt and clay. The third category, saturated-coarse, consists of 94 data points. Coarse-grained soils from below the groundwater table, such as sand and gravel, were sampled for these measurements. The fourth and final category is saturated fines, which contains 86 data points. These measurements were taken from fine-grained soils below the ground water table, such as silt and clay.

3.2. Models

Four (4) calibrated models were created to determine the unit weight of soil, which depends on the amount of water in the ground, soil type, and SPT N-value. These are the moist unit weight of coarse-grained soil (kN/m^3), saturated unit weight of coarse-grained soil (kN/m^3), moist unit weight of fine-grained soil (kN/m^3), and saturated unit weight of fine-grained soil (kN/m^3). Figures 2–5 provide an overview of the machine learning models used in this study, each designed for a particular modeling approach. The first model, ML-ANN, uses an artificial neural network architecture to capture complex data relationships. The second algorithm, ML-Tree, uses decision tree models. The third method, ML-Linear Linear, uses linear modeling. The fourth, ML-Linear Quadratic, is a quadratic model-specific algorithm. Lastly, ML-Boosted uses ensemble learning, in particular, boosted tree models.

Above the groundwater table, the soil's natural in situ unit weight is equivalent to its moist unit weight. Various data on moist unit weight were used and trained using machine learning techniques to develop models for predicting unit weight. Figure 2 demonstrates that, among competing algorithms, the neural network regression model is the most effective for coarse-grained soils, exhibiting a very strong positive relationship with an R^2 value of 0.70.

Using parametric analysis, it was determined that there is an instantaneous increase in unit weight between N-values 0 and 15, indicating a range from extremely loose to loose. In addition, the unit weight increases gradually and consistently from N-values 15 to 50.

In this study, a new unit weight trend emerged. This trend deviated from the typical approaches of previous studies [4,31,34], which have primarily focused on the correlation of SPT N-values with the moist unit weight, which has increased steadily over time.

The new trend acknowledges a critical limitation in the models of previous studies, which fail to account for the case of an SPT value of "0". When the SPT N-value is zero, previous models [4,34] demonstrated an intercept of 12.25 kN/m^3 and 16 kN/m^3, respectively. This method is deemed completely inaccurate because it must account for an SPT value's absence.

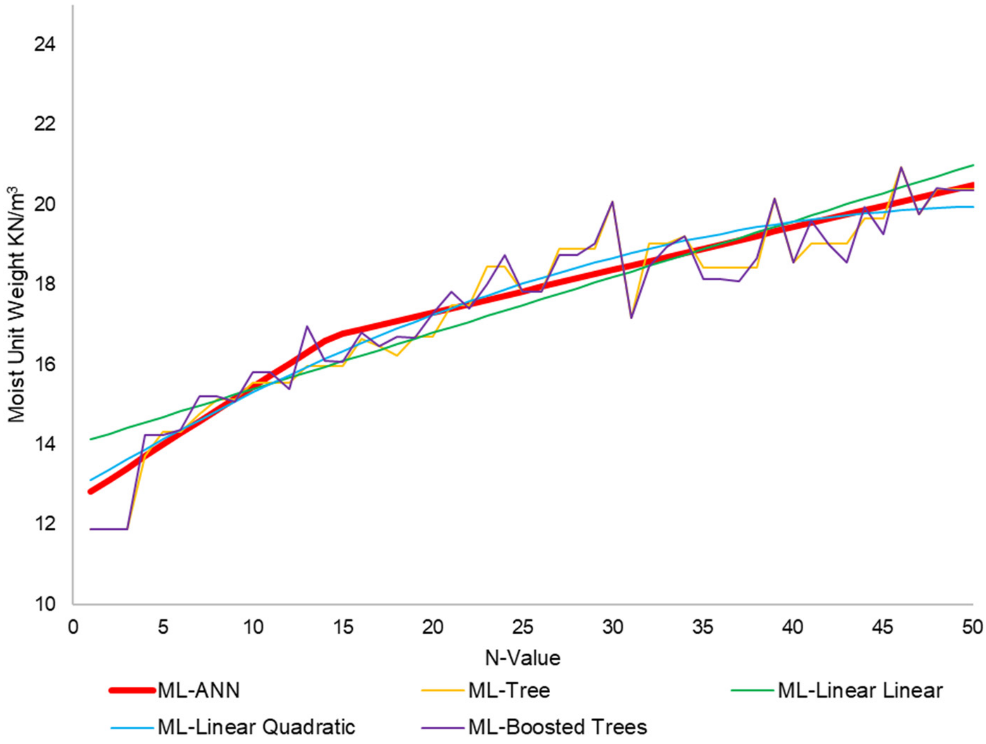

Figure 2. Unit weight of coarse-grained soils above the groundwater table.

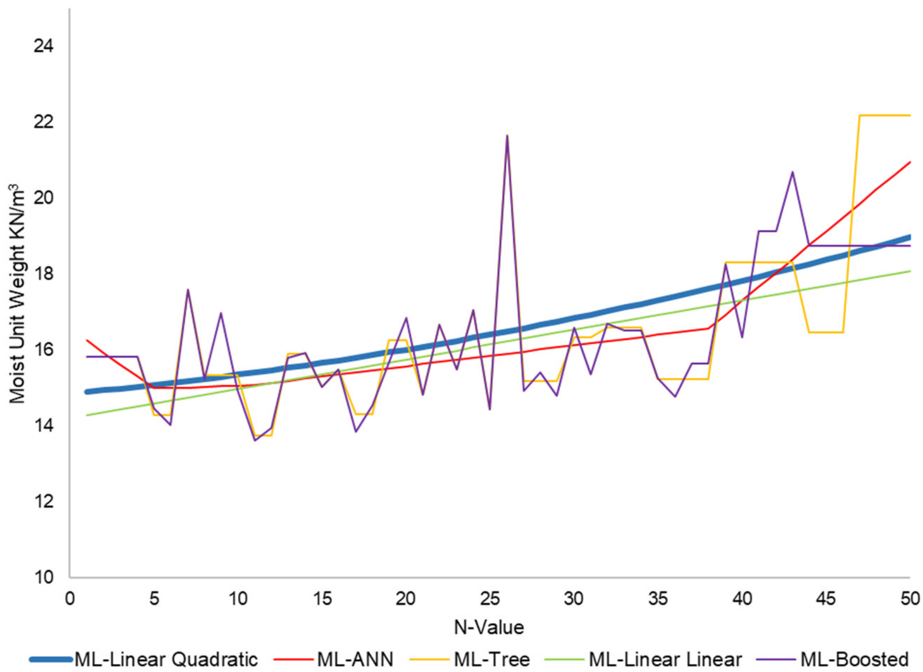

Figure 3. Unit weight of fine-grained soils above the groundwater table.

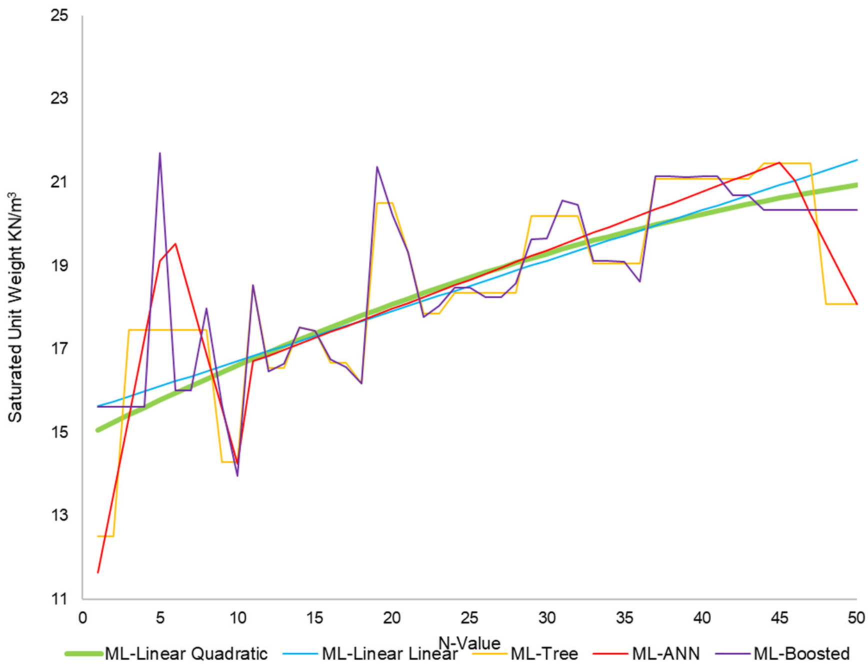

Figure 4. Unit weight of coarse-grained soils below the groundwater table.

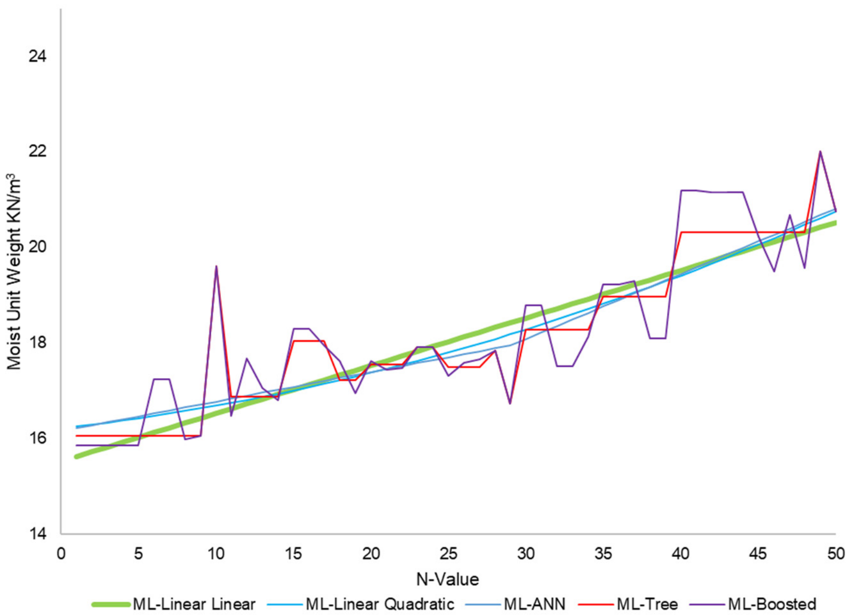

Figure 5. Unit weight of fine-grained soils below the groundwater table.

The quadratic model won the competition among various models for determining the moist unit weight of fine-grained soils, as shown in Figure 3. The resulting model had a weak relationship (R^2 = 0.16) due to the insufficient data collected. Even though the linear model was an option, its performance was inferior to the quadratic model's. The proposed model adhered to the trend [4,33], which indicates a direct correlation between the SPT N-Value and unit weight. This relationship was attributed to the compaction of soil particles and the presence of water in the soil layer above the groundwater table [3]. As the N-Value increased, the relative density of the soil particles increased, leading to increased compaction and, ultimately, a greater unit weight.

The saturated unit weight parameter describes soil layers below the groundwater table in geotechnical engineering. In contrast, the unit weight parameter describes soil particles above the groundwater table. The saturated unit weight accounts for the added weight of the water caused by saturation, resulting in greater values for these soil layers. This contradicts the expectation that unsaturated soil layers would have a higher unit weight due to their lack of water saturation. According to trained models, most saturated soil layers below the water table are more compacted, resulting in greater unit weight. The minimum and maximum saturated unit weight values are 15.63 kN/m^3 and 21.53 kN/m^3, respectively.

The winning models for predicting the saturated unit weight were linear for the coarse-grained and fine-grained soil layers. The model for coarse-grained soil demonstrated a moderately positive relationship with an R^2 value of 0.30. In contrast, the model for fine-grained soil showed a strong positive relationship with an R^2 value of 0.61. The lower R^2 value for the coarse-grained soil model compared to that of the fine-grained soil model suggests that the data are less sparse. Figures 4 and 5 represent the models for predicting the saturated unit weight of coarse-grained and fine-grained soils using parametric analysis. Based on the SPT N-value, the saturated unit weight of a soil layer can be estimated using the linear relationship between the SPT N-value and the saturated unit weight [4,34].

Certain limitations apply to the models used to calculate the unit weight of soil layers. Specifically, these models only apply to SPT N-values ranging from 1 to 50. If the SPT N-value exceeds 50, this indicates refusal, which suggests the presence of rocks and the inapplicability of the models. Regarding Figures 2–5, generally, ML models, such as tree, boosted, and neural network models, usually have strong relationship indications. However, these models were highly susceptible to overfitting because they followed the detail and noise in the training data to such an extent that it negatively impacted model performance. These overfitting models could not make accurate predictions on new data because they could not generalize from the training set to the test set.

3.3. Uncertainty Analysis

Using a T-test, this uncertainty analysis compares the mean raw unit weight to the mean predicted unit weight. The objective is to determine whether a statistically significant difference exists between the two variables. The mean raw unit weight was 17.95, while the predicted unit weight was 17.68. The calculated t-value, which measures the difference between the means in relation to the variance within the groups, was 1.607. The *t*-test's *p*-value of 0.108027 indicated statistically significant evidence of similarity between the raw and predicted unit weights.

The raw unit weight had a standard deviation of 2.32, whereas the predicted unit weight had a standard deviation of 1.81. Additionally, the F-ratio was given as 1.66. As indicated by the *p*-value, based on the results of this uncertainty analysis, there was a statistically significant similarity between the raw and predicted unit weights, addressing the uncertainty.

Moreover, the confidence intervals and measures of spread and variability were analyzed to address the uncertainty in the data provided. Shown in Table 6, the confidence intervals encompass the mean SPT N-Value (between 27.96 and 28.25) and the mean Unit Weight (between 17.93 and 17.96), providing a range within which the true population means are likely to fall with a specified degree of confidence. These intervals account for sampling variability and measure the estimated values' uncertainty. In addition, the variance (382.90) and standard deviation (19.57) of the SPT N-Value illustrate the variation and spread of the dataset. A more significant variance and standard deviation indicate that the SPT N-Value measurements are more uncertain or variable. A value near zero indicates approximate symmetry. Lastly, the large sample size of 69,723 allows for more precise estimates and narrower confidence intervals by reducing uncertainty. Nonetheless, carefully considering these metrics is essential when interpreting the results, as they provide insights into the data's range, variability, and representativeness, thereby effectively addressing the uncertainty underlying the analysis.

Table 6. Confidence interval analysis.

Variable	Mean	Conf.	Conf.	Min	Max	Var	Std. Dev.	Conf. SD	Conf. SD	Coef. Var.	Skewness
SPT N-Value	28.11	27.96	28.25	1.00	50.00	382.90	19.57	19.47	19.67	69.62	−0.04
Unit Weight	17.95	17.93	17.96	13.10	20.93	5.37	2.31674	2.30	2.33	12.91	0.00

3.4. Nomographs

Nomographs are a form of graphical representation created to summarize calibrated empirical models that were previously generated. The nomograph delineates both dependent and independent scales for these models. It is possible to determine several significant parameters by using only the SPT N-Value of a given soil layer and a nomograph. In utilizing the nomograph, it is assumed that the SPT N-value in the soil investigation report is reliable and has no uncertainty. Nomographs can provide a quick and easy visualization of on the soil strata's strength, providing helpful information. Please note that because the utmost care has been taken to address uncertainties and ensure data reliability, it is essential to acknowledge that inherent uncertainties may still exist in the SPT results.

A nomograph can determine the moist and saturated unit weights of coarse-grained and fine-grained soils. Nomographs are especially useful because they can provide these values without on-site computing tools. These essential soil parameters can be determined quickly and easily by simply reading the appropriate values from the nomograph based on the given SPT N-Value. The delineated nomograph is shown in Figure 6.

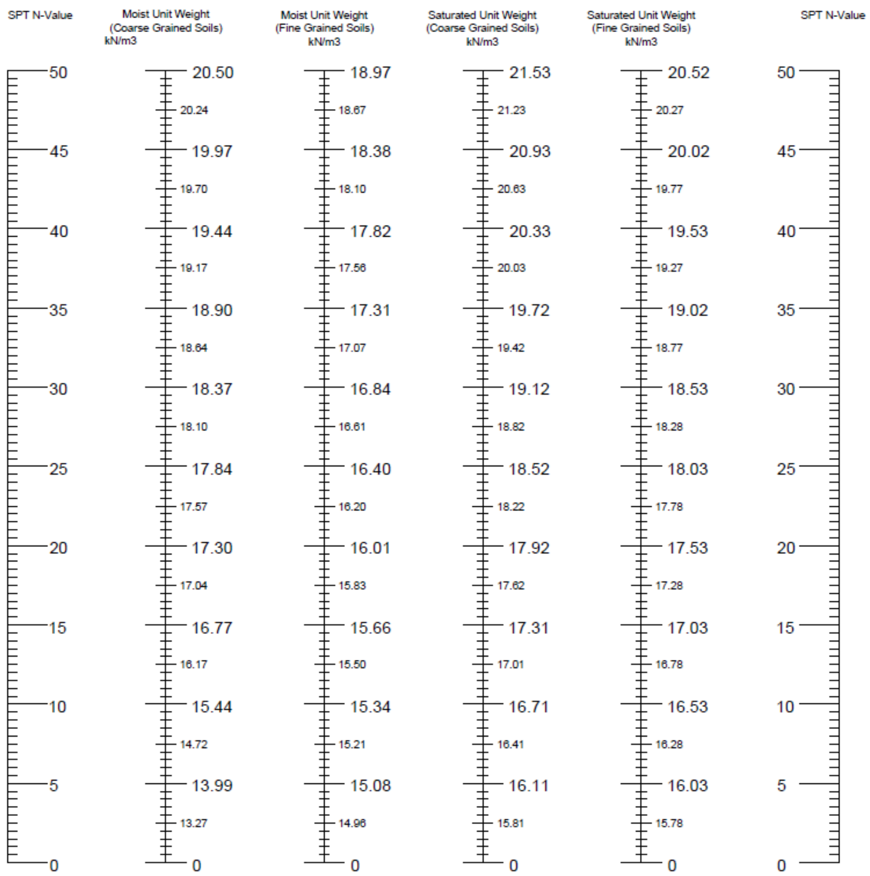

Figure 6. Delineated nomograph for unit weight of soils.

3.5. Deploying the Generated Nomograph to Metro Manila, Philippines

Metro Manila, also known as the National Capital Region (NCR), is situated on the western coast of Luzon. It comprises sixteen cities and one municipality and is located on a floodplain exposed to various geohazards.

We considered a soil layer located 3 m below ground level in the city of Malabon. This layer consists of poorly graded sand (SP) and has an N-value of 10 on the Standard Penetration Test (SPT). The water table was determined to be 1 m below ground level. Figure 7 was used to determine the unit weight of this soil layer.

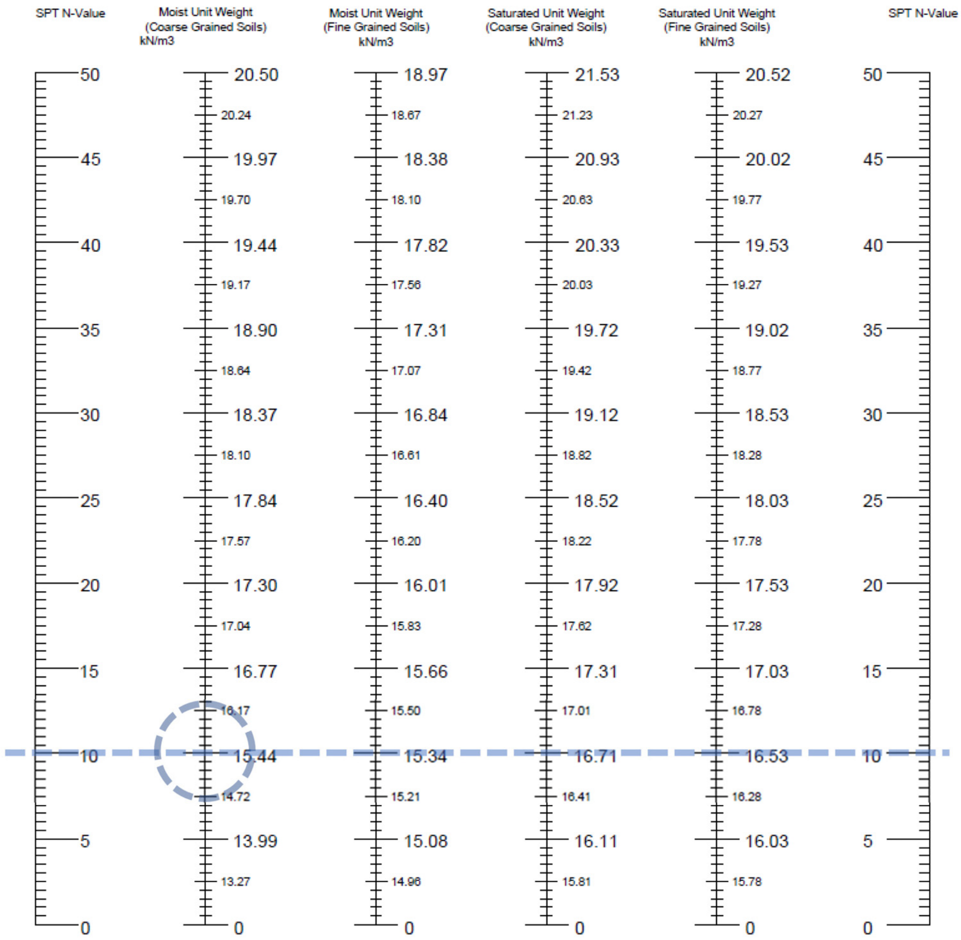

Figure 7. Determination of unit weight using the generated nomograph.

Due to its location beneath the groundwater table, the soil layer beneath the water table was considered to be the saturated soil layer. Additionally, the soil type is poorly graded sand, which belongs to the group of coarse-grained soils. Therefore, the nomograph that should be utilized in this scenario is the one for saturated coarse-grained soils. The unit weight of the soil was determined to be 15.44 kN/m^3 using the generated nomograph.

For validation, the region's soil is predominantly composed of alluvial deposits, sedimentary material transported and deposited by rivers and other bodies of water. Alluvial deposits consist of a mixture of clay, silt, sand, and gravel, and their characteristics vary based on their proximity to the sediment source. Additionally, some regions contain deposits of volcanic ash. Due to the region's shallow water table, Metro Manila's soil characteristics are influenced by alluvial deposits and the presence of water. The soil is typically soft, with low strength and rigidity.

The unit weight of soil is an essential design parameter in geotechnical engineering, as it is used to calculate the soil's bearing capacity, settlement, and lateral earth pressures.

It is used in various geotechnical engineering designs, such as shallow and deep foundation designs, to determine the foundation's bearing capacity, settlement, and soil pressure. In addition, it is utilized in the design of retaining walls to assess the lateral earth pressure exerted on the wall. The unit weight of the soil can be used to determine the importance of soil on a slope, which plays a crucial role in determining the stability of the slope. The unit weight of soil is also utilized when designing embankments, calculating the required slope angle and stability, and choosing the dynamic response of soil to seismic loads. In addition, it is used to determine the effort of compaction needed and density for a particular soil type. In conclusion, the unit weight of soil is an essential parameter that is used in geotechnical engineering designs alongside other soil parameters to determine the behavior of the ground under different loading and environmental conditions.

3.6. Further Validation

To determine the efficacy of the nomograph, unused data was examined. For example, the coordinates 14.683126 and 120.942258 for a location in Dampalit, Malabon City, were used. The obtained accuracy rate of 84% indicates that the estimated parameters and collected data are in good agreement. However, specific layers require improved accuracy. Nevertheless, this shows that the model and the predicted parameters have a reliable relationship. Given these findings, the nomograph can be a reference for future applications.

4. Conclusions and Recommendations

This study aimed to develop a geotechnical engineering-useful nomograph that incorporates machine learning, specifically for the unit weight parameter. Four calibrated models were developed to determine the soil unit weight based on various variables, including the soil's water content, soil type, and SPT N-value. Included in these models were the moist unit weight of coarse-grained soil, the saturated unit weight of coarse-grained soil, the moist unit weight of fine-grained soil, and the saturated unit weight of fine-grained soil. Our results indicated a very strong positive relationship for the moist unit weight of coarse-grained soil, a weak positive relationship for the moist unit weight of fine-grained soil, a moderately positive relationship for the saturated unit weight of coarse-grained soil, and a very strong positive relationship for the saturated unit weight of fine-grained soil.

Nomographs were created to visually represent the empirical models for unit weight and SPT N-Value that were calibrated for the study. By aligning a straightedge across the scales and reading the unit weight values where it intersects, using these nomographs, it is possible to determine the unit weight values based on the SPT N-values. It should be noted, however, that these models and nomographs only apply to SPT N-values between 1 and 50. If the SPT N-value is greater than 50, it indicates refusal, which suggests the presence of rocks and renders the models/nomographs inapplicable.

The generated nomographs were tested in Malabon, a city in Metro Manila, where a low unit weight of soil of 15.44 kN/m^3 was determined. This low unit weight is due to the predominance of alluvial deposits and the shallow groundwater table, which soften and weaken the soil.

According to this study's findings, further research suggests that the delineation of subsurface parameters using SPT N-value can be considered. These parameters include the angle of internal friction and cohesion, which can be derived from the SPT N-value. Other geotechnical parameters, such as the expansion and compression indices, can also be specified. It should be noted, however, that these parameters require nomographs with other independent variables instead of relying solely on the SPT N-value. Finally, we recommend that, in the absence of MATLAB, other free programming languages, such as Python, should be utilized. Python is a flexible and widely-used programming language that offers a variety of libraries and tools for scientific computing and data analysis, making it a viable alternative for geotechnical modeling and analysis.

Author Contributions: Conceptualization, J.D. and J.G.; methodology, J.D. and J.G.; software, J.G.; validation, J.D. and J.G.; formal analysis, J.D. and J.G.; investigation, J.D. and J.G.; resources, J.D. and J.G.; data curation, J.D. and J.G.; writing—original draft preparation, J.G; writing—review and editing, J.D.; visualization, J.G.; supervision, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research study was funded by the Department of Science and Technology Grants-in-Aid (DOST-GIA) and Engineering Research and Development for Technology (ERDT).

Data Availability Statement: Not applicable.

Acknowledgments: Thank you to DLSU, GEMMMS Project, and DOST.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Peck, R.; Hanson, W.; Thornburn, T. *Foundation Engineering*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 1974; Volume 2.
2. Ameratunga, J.; Sivakugan, N.; Das, B.M. *Correlations of Soil and Rock Properties in Geotechnical Engineering*; Springer: Berlin/Heidelberg, Germany, 2016.
3. Terzaghi, K.; Peck, R.; Mesr, G. *Soil Mechanics in Engineering Practice*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1996.
4. Puri, N.; Prasad, H.D.; Jain, A. Prediction of Geotechnical Parameters Using Machine Learning Techniques. *Procedia Comput. Sci.* **2018**, *125*, 509–517. [CrossRef]
5. Pynomo. Introduction—pyNomo Documentation 0.3.2.1 Documentation. Available online: http://pynomo.org/wiki/index.php/Software_documentation (accessed on 1 July 2022).
6. Esmail, E.L. Nomographs for the synthesis of epicyclic-type automatic transmissions. *Meccanica* **2013**, *48*, 2037–2049. [CrossRef]
7. McMillen, E.L. A Versatile Nomograph for Chemical Engineering Calculations. *Ind. Eng. Chem.* **1938**, *30*, 71–79. [CrossRef]
8. Levens, A. *Nomography*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1948.
9. Glasser, L.; Doerfler, R. A brief introduction to nomography: Graphical representation of mathematical relationships. *Int. J. Math. Educ. Sci. Technol.* **2018**, *50*, 1273–1284. [CrossRef]
10. Douglas, J.; Danciu, L. Nomogram to help explain probabilistic seismic hazard. *J. Seismol.* **2020**, *24*, 221–228. [CrossRef]
11. Coker, A.K. Cost estimation and economic evaluation. *Ludwig's Appl. Process Des. Chem. Petrochem. Plants* **2017**, *1*, 69–102.
12. Skibniewski, M.J.; Nof, S.Y. A framework for programmable and flexible construction systems. *Robot. Auton. Syst.* **1989**, *5*, 135–150. [CrossRef]
13. Zhao, Z.; Guan, X.; Xiao, F.; Xie, Z.; Xia, P.; Zhou, Q. Applications of asphalt concrete overlay on Portland cement concrete pavement. *Constr. Build. Mater.* **2020**, *264*, 120045. [CrossRef]
14. Mahpour, A.; El-Diraby, T. Incorporating Climate Change in Pavement Maintenance Policies: Application to Temperature Rise in the Isfahan County, Iran. *Sustain. Cities Soc.* **2021**, *71*, 102960. [CrossRef]
15. Mandare, A.B.; Ambast, S.K.; Tyagi, N.K.; Singh, J. On-farm water management in saline groundwater area under scarce canal water supply condition in the Northwest India. *Agric. Water Manag.* **2008**, *95*, 516–526. [CrossRef]
16. Moatar, F.; Person, G.; Meybeck, M.; Coynel, A.; Etcheber, H.; Crouzet, P. The influence of contrasting suspended particulate matter transport regimes on the bias and precision of flux estimates. *Sci. Total Environ.* **2006**, *370*, 515–531. [CrossRef] [PubMed]
17. Barker, G. Pipe sizing and pressure drop calculations. In *The Engineer's Guide to Plant Layout and Piping Design for the Oil and Gas Industries*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 411–472.
18. Chien, S.F. Laminar flow pressure loss and flow pattern transition of Bingham plastics in pipes and annuli. *Int. J. Rock Mech. Min. Sci.* **1970**, *7*, 339–356. [CrossRef]
19. Gamage, S.H.P.W.; Hewa, G.A.; Beecham, S. Modelling hydrological losses for varying rainfall and moisture conditions in South Australian catchments. *J. Hydrol. Reg. Stud.* **2015**, *4*, 1–21. [CrossRef]
20. Haan, C.T.; Barfield, B.J.; Hayes, J.C. Hydraulics of Structures. In *Design Hydrology and Sedimentology for Small Catchment*; Elsevier: Amsterdam, The Netherlands, 1994; pp. 144–181.
21. Srivastava, J.B.; Gupta, A.K.; Khanna, S.K. Modelling of Highway Traffic Pollution. *IFAC Proc. Vol.* **1994**, *27*, 889–891. [CrossRef]
22. Burke, C.M.; Scott, D.M. The space race: A framework to evaluate the potential travel-time impacts of reallocating road space to bicycle facilities. *J. Transp. Geogr.* **2016**, *56*, 110–119. [CrossRef]
23. Fricke, L.B. Traffic management and collision investigation. *Accid. Anal. Prev.* **1982**, *14*, 486–487. [CrossRef]
24. Martinelli, P.; Colombo, M.; Ravasini, S.; Belletti, B. Application of an analytical method for the design for robustness of RC flat slab buildings. *Eng. Struct.* **2022**, *258*, 114117. [CrossRef]
25. Minami, F.; Ohata, M.; Shimanuki, H.; Handa, T.; Igi, S.; Kurihara, M.; Kawabata, T.; Yamashita, Y.; Tagawa, T.; Hagihara, Y. Method of constraint loss correction of CTOD fracture toughness for fracture assessment of steel components. *Eng. Fract. Mech.* **2006**, *73*, 1996–2020. [CrossRef]
26. Chala, A.T.; Ray, R.P. Machine Learning Techniques for Soil Characterization Using Cone Penetration Test Data. *Appl. Sci.* **2023**, *13*, 8286. [CrossRef]
27. Daghistani, F.; Abuel-Naga, H. Evaluating the Influence of Sand Particle Morphology on Shear Strength: A Comparison of Experimental and Machine Learning Approaches. *Appl. Sci.* **2023**, *13*, 8160. [CrossRef]
28. Cheng, H.; Zhang, H.; Liu, Z.; Wu, Y. Prediction of Undrained Bearing Capacity of Skirted Foundation in Spatially Variable Soils Based on Convolutional Neural Network. *Appl. Sci.* **2023**, *13*, 6624. [CrossRef]
29. Chala, A.T.; Ray, R. Assessing the Performance of Machine Learning Algorithms for Soil Classification Using Cone Penetration Test Data. *Appl. Sci.* **2023**, *13*, 5758. [CrossRef]

30. Lee, S.; Kang, J.; Kim, J. Prediction Modeling of Ground Subsidence Risk Based on Machine Learning Using the Attribute Information of Underground Utilities in Urban Areas in Korea. *Appl. Sci.* **2023**, *13*, 5566. [CrossRef]
31. Dufour, J.C.; Mancini, J.; Fieschi, M. Searching for evidence-based data. *J. Chir.* **2009**, *146*, 355–367. [CrossRef] [PubMed]
32. Safadi, M.; Ma, J.; Wickramasuriya, R.; Daly, D.; Perez, P.; Kokogiannakis, G. Mapping for the Future: Business Intelligence Tool to Map Regional Housing Stock. *Procedia Eng.* **2017**, *180*, 1684–1694. [CrossRef]
33. Bowles, J.E. Foundation Analysis and Design. In *Civil Engineering Materials*; The McGraw-Hill Companies, Inc.: New York, NY, USA, 1997.
34. Rahman, M. Foundation Design using Standard Penetration Test (SPT) N-value. *Researchgate* **2019**, *5*, 1–39.