Soils and Rocks
An International Journal of Geotechnical and Geoenvironmental Engineering
www.soilsandrocks.com
ISSN 1980-9743
ISSN-e 2675-5475

# A multiple model machine learning approach for soil classification from cone penetration test data

Lucas O. Carvalho[1] Dimas B. Ribeiro[1]#

Article

**Keywords**
Robertson
Charts
Tropical soil
Pore
Pressure
Budget constraints

**Abstract**

The most popular methods for soil classification from cone penetration test (CPT) data are based on examining two-dimensional charts. In the last years, several authors have dedicated efforts on replicating and discussing these methods using machine learning techniques. Nonetheless, most of them apply few techniques, include only one dataset and do not explore more than three input features. This work circumvents these issues by: (i) comparing five different machine learning techniques, which are also combined in an ensemble; (ii) using three distinct CPT datasets, one composed of 111 soundings from different countries, one composed of 38 soundings with information of soil age and the third composed of 64 soundings taken from the city of São Paulo, Brazil; and (iii) testing combinations of five input features. Results show that, in most cases, the ensemble of multiple models achieves better predictive performance than any technique isolated. Accuracies close to the maximum were obtained in some cases without the need of pore pressure information, which is costly to measure in geotechnical practice.

## 1. Introduction

The classical approach for soil classification from CPT data is based on examining two-dimensional charts, with pioneer studies pursuing to predict the soil granulometrical distribution from two raw CPT measurements (Begemann, 1965). Later work stated that predicting soil behavior would be more useful for real engineering projects than predicting soil granulometry (Douglas & Olsen, 1981). As a result, the well-known Robertson classification methods were proposed, using two charts obtained from three raw CPT measurements (Robertson et al.,1986; Robertson,1990). These charts became particularly popular due to the proposed input transformations, capable of better separating soil classes. Nonetheless, further investigations exposed limitations in those methods (Jefferies & Davies, 1991), associated with overconsolidated clays with dilative behavior. Although these methods evolved to minimize these problems (Robertson, 1991), other studies have shown that similar limitations remained (Ramsey, 2002; Schneider et al., 2008). To overcome these limitations two new charts were proposed (Schneider et al., 2008, 2012). In recent work, these charts were modified to create a full behavior-based classification method (Robertson, 2016).

Many recent works from the literature have also applied machine learning (ML) techniques to different geotechnical problems and most of them use artificial neural networks (ANN) to predict soil characteristics (Goh, 1995, 1996; Schaap et al., 1998; Juang & Chen, 1999; Kumar et al., 2000; Juang et al., 2002; Juang et al., 2003; Hanna et. al., 2007). On the other hand, Livingston et al. (2008) used decision trees (DT) models, Kohestani et al. (2015) employed random forests (RF), whilst Goh & Goh (2007) induced support vector machine (SVM) models. In addition, most studies dedicated to soil classification from CPT data seek for new soil classes using data clustering (Hegazy & Mayne, 2002; Facciorusso & Uzielli, 2004; Liao & Mayne, 2007; Das & Basudhar, 2009; Rogiers et al., 2017; Carvalho & Ribeiro, 2020). But another possible approach, which is relatively unexplored in the literature, is using ML techniques to replicate predefined soil classification systems, like classical soil classification methods based on charts (Arel, 2012). Most work adopting this approach use only ANN models (Kurup & Griffin, 2006; Arel, 2012; Reale et al., 2018) and, when more ML techniques are used, applications are restricted to small CPT datasets, with all soundings taken at the same location (Bhattacharya & Solomtine, 2006). Recent work has explored the additional potentialities of ML techniques to prospect and discuss alternative geotechnical aspects of soil classification, using the k-nearest neighbor ML technique (Carvalho & Ribeiro, 2019). Expanding this study with a larger

#Corresponding author. E-mail address: dimasbetioli@gmail.com
[1]Instituto Tecnológico de Aeronáutica, Divisão de Engenharia Civil, São José dos Campos, SP, Brasil.

Submitted on May 26, 2021; Final Acceptance on August 24, 2021; Discussion open until February 28, 2022.
https://doi.org/10.28927/SR.2021.072121

and more diverse dataset, comparing more ML techniques and investigating different combinations of input features are the main objectives of this work.

Herewith, several ML techniques are trained to classify soil from CPT data, aiming to replicate classification systems generated with a student version of the CPeT-IT v2.0.2.5 software. First, CPeT-IT is used to classify all examples from three datasets: one composed of 111 CPT soundings taken from different countries; one composed of 38 soundings including soil age information; and the third composed of 64 CPT soundings taken from the city of São Paulo, Brazil. The authors believe that using more diverse data samples is important to reveal general properties of the problem and to assess the competence of the ML models more properly. Next, the collected soil samples are used to train the following ML techniques: distance-weighted nearest neighbors (DWNN), boosted DT, RF, ANN, SVM and a multiple model predictor (MMP), which is a combination of the previous models, aka a heterogeneous ensemble of classification techniques. In addition, the combination of different input features is tested, including the original inputs required by CPeT-IT. This allows to investigate and discuss novel geotechnical aspects related to soil classification. As a result, this work has achieved the following original contributions:

- This is a first attempt to apply and compare multiple ML techniques of distinct biases (namely, DWNN, DT, RF, ANN, SVM) in a geotechnical application. In addition, their outputs are combined in an ensemble (MMP), resulting in higher predictive accuracies for soil classification;
- Discussing the utility and application of Robertson charts for classifying tropical soil, as their usage is more common in the analysis of soil data from temperate countries;
- Making possible to approximate Robertson soil classes without the need of pore pressure information, which is costly to measure in geotechnical practice. This is particularly important for the analysis of data from developing countries, which usually have severe budget constraints imposed on the engineering practice.

Although the results that sustain the last contribution, presented in Section 5.4, are not enough to dismiss measuring pore pressure in real engineering projects, they are important to motivate discussions concerning novel methods for soil classification that may be especially appealing for underdeveloped and developing countries.

## 2. Classification methods used in CPeT-IT

This section describes the two soil classification methods replicated in this work using ML techniques. For both cases, class 0 denotes a misclassified soil.

### 2.1. Method influenced by soil granulometry (ISG)

One of the chart-based classification methods replicated in this work was proposed by Robertson (1991), which is referred as ISG throughout this text. In this reference, the author intended to include soil behavior within the classification system, nonetheless the defined classes refer to granulometrical soil composition only. Furthermore, borehole samples were used to make soil classes compatible with real soil types. The ISG soil classes are:

- Sensitive, fine grained.
- Organic soils - peats.
- Clays - clay to silty clay.
- Silt mixtures - clayey silt to silty clay.
- Sand mixtures - silty sand to sandy silt.
- Sands - clean sand to silty sand.
- Gravelly sand to sand.
- Very stiff sand to clayey sand.
- Very stiff, fine grained.

The four basic parameters measured in CPT are depth ($z$), uncorrected cone resistance ($q_c$), lateral friction ($f_s$) and pore pressure in a disturbed state ($u_2$), usually measured behind the cone tip. In the method proposed by Robertson (1991), these parameters are combined to obtain normalized versions.

First, $q_c$ is corrected to discount the water pressure aiding cone penetration, resulting the total cone resistance $q_t$. Next, the equilibrium pore pressure $u_0$ is needed to calculate the excess pore pressure $u_2 - u_0$. The $u_0$ value can be obtained by drawing a straight line through the $u_2$ value in the graphic.

The effective $\sigma_{v0}^{'}$ and total $\sigma_{v0} = \sigma_{v0}^{'} + u_0$ overburden stresses are then obtained, enabling to calculate the net cone resistance $q_n = q_t - \sigma_{v0}$. In order to eliminate correlations, Robertson (1990) proposed that $q_n$ should be divided by $\sigma_{v0}^{'}$ to discount overburden and that $f_s$ and $u_2 - u_0$ should be divided by $q_n$, resulting in the normalizations presented in Equations 1 to 3:

$$Q_{t1} = \frac{q_n}{\sigma_{v0}^{'}} \tag{1}$$

$$F_r = \frac{f_s}{q_n} \tag{2}$$

$$B_q = \frac{u_2 - u_0}{q_n} \tag{3}$$

Later work (Robertson & Wride, 1998) found that the exponent $n$ of $\sigma_{v0}^{'}$ in the $Q_{t1}$ expression should be 1 only for pure sands, 0.5 only for pure clays and intermediary for mixtures of them. The result is presented in Equation 4:

$$Q_{tn} = \left(\frac{q_n}{pa}\right)\left(\frac{pa}{\sigma'_{v0}}\right)^n \tag{4}$$

where $pa$ is a reference pressure of $0.1$ MPa. The exponent can be obtained with the Equation 5:

$$n = 0.381 I_c + 0.05\left(\frac{\sigma'_{v0}}{pa}\right) - 0.15 \tag{5}$$

The parameter $I_c$ can be calculated as presented in Equation 6 (Robertson, 2009):

$$I_c = \left[\left(3.47 - log\, Q_{tn}\right)^2 + \left(log\, F_r + 1.22\right)^2\right]^{0.5} \tag{6}$$

Based on the previous equations, two charts are proposed by Robertson (1991) for soil classification. After obtaining raw CPT values and performing all procedures defined previously, a point can be placed in these charts, resulting in an attribution to each soil example. That is, the area to which the point belongs gives the class of the corresponding collected soil. If the obtained point is located outside the ranges defined within these charts, the soil is considered misclassified, receiving class $0$.

### 2.2. Method focused on soil behavior (FSB)

The second soil classification method replicated in this work was proposed by Robertson (2016) and is referred as FSB throughout this text. It includes, as a new application, a method to identify if soil contains microstructure. In this method, considered fully behavioral in the literature, soil classes are divided into three main blocks: clay-like, sand-like and transitional. One advantage of this division is that the behavior of sands and clays is clearly separable. Sands usually present high strength, low compressibility and high permeability, while clays usually present low strength, high compressibility and low permeability. Each soil group is subdivided as pursuing dilative or contractive behavior, according to the consolidation state. A separate class was created for contractive clays that are sensitive to disturbance. The FSB classes are:

- CCS: Clay-like - Contractive – Sensitive.
- CC: Clay-like – Contractive.
- CD: Clay-like – Dilative.
- TC: Transitional – Contractive.
- TD: Transitional – Dilative.
- SC: Sand-like – Contractive.
- SD: Sand-like – Dilative.

One problem of the ISG method, described in the previous section, is that $B_q$ has strong negative correlation with $Q_{tn}$, which makes highly overconsolidated clays indistinguishable from very dense sands (Schneider et al., 2008). To solve this problem, a new normalized excess pore pressure was proposed (Robertson, 2016) as:

$$U_2 = B_q Q_{t1} = \frac{u_2 - u_0}{\sigma'_{v0}} \tag{7}$$

The FSB method then employs two charts, one using $F_r$ and $Q_{tn}$ and the other using $U_2$ and $Q_{tn}$. The first is similar to the chart proposed in Schneider et al. (2008), while the second uses the hyperbolic curves presented in Schneider et al. (2012). New curves are also added to the $F_r \times Q_{tn}$ chart to separate dilative and contractive behaviors, as well as for separating the contractive sensitive behavior.

The values obtained for $Q_{tn}$, $F_r$ and $U_2$ enable obtaining one point in each of the charts. If classes given in both charts do not agree, the soil is considered misclassified (class $0$). In addition to that, a soil sample is attributed to class $0$ if the point is located outside the ranges of $Q_{tn}$, $F_r$ and $U_2$ of the charts and if a modified normalized small-strain rigidity index is greater than $330$.

Robertson (2016) highlights that the FSB method is inaccurate for aged or cemented soils, which contain microstructure.

## 3. Machine learning (ML) techniques employed

In this work, six ML techniques of distinct biases are used to replicate the soil classification methods described in Section 2. In this Section, a brief theoretical description is given for DWNN, DT, RF, ANN and SVM. In the MMP model, all previous five ML models have their outputs combined in the classification of new samples by a majority voting strategy. Table 1 presents the main advantages and disadvantages experienced by the authors, applying these ML techniques to soil classification problems.

### 3.1. Distance-weighted nearest neighbors (DWNN)

The DWNN technique (Dudani, 1976) is a distance-based technique, meaning that it uses distances to evaluate if two objects $x$ and $y$ are similar. In this work, the Euclidean distance is used, which can be written as:

$$d(x, y) = \sqrt{\Sigma\left|x_i - y_i\right|^2} \tag{8}$$

In DWNN, all known examples (composing the training dataset) can be regarded as a cloud of points within the input space. A new point can be classified according to its proximity to the known examples. For instance, it can be classified into the same class of its nearest neighbor. Or a

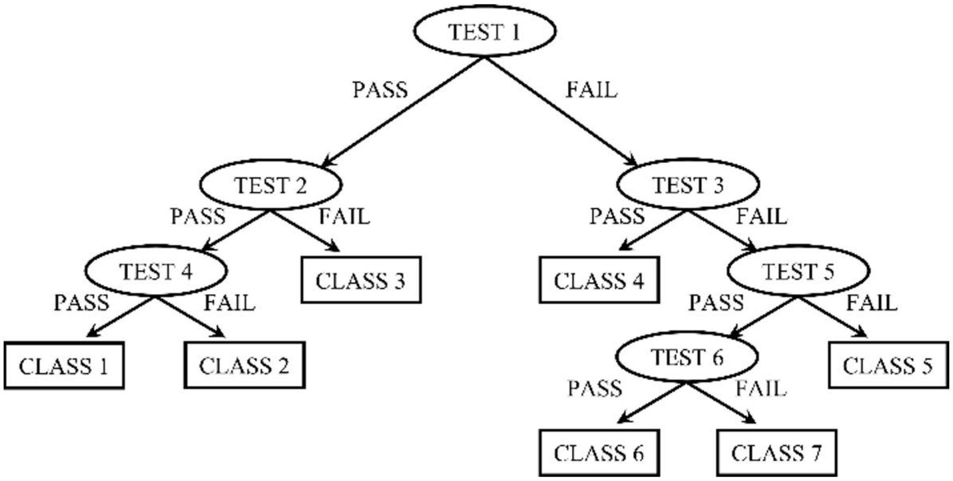

**Figure 1.** Example of DT.

**Table 1.** Advantages and disadvantages of each technique.

| Technique | Advantages | Disadvantages |
|---|---|---|
| DWNN | Flexible and easy to program | Sensitive to outliers, not so accurate |
| DT | Can lead to an interpretable model | Tends to overfit to training data |
| RF | Accurate in most cases | The model becomes too complex to be interpreted |
| ANN | Can be replicated with simple spreadsheets | Not interpretable and difficult to calibrate |
| SVM | Leads to a globally optimal solution | Difficult to tune the hyper-parameter values |
| MMP | In general, more accurate than the isolated techniques | Can combine disadvantages of isolated techniques |

majority voting of the classes of the *k* nearest neighbors can be employed instead. Weights can also be assigned to the votes of the nearest neighbors, proportional to the inverse of their distance to the new data point. This results in the DWNN technique. A Gaussian DWNN weighting is used in this work, which is given by:

$$w\left(d\left(x,y\right)\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}d\left(x,y\right)^{2}} \qquad (9)$$

where $d\left(x,y\right)$ is the Euclidean distance between two data items expressed in Equation 8. A recent work has shown that Gaussian weighting leads to better predictive performance in soil classification than attributing the same weights to all nearest neighbors (Carvalho & Ribeiro, 2019).

### 3.2. Decision trees (DT) and random forest (RF)

A DT can be defined as a graph with a tree structure, containing decision and leaf nodes (Quinlan, 1986). The decision nodes perform tests on the feature values of the data points, whilst leaf nodes output a class. Starting from the root node, the feature values of an example are used to decide to each branch of the tree the example will proceed until a leaf node is reached, giving the final classification of the object. Figure 1 illustrates a DT with six decision nodes (tests) and seven leaf nodes (classes).

The test performed by each decision node is usually chosen to maximize a goodness of split criterion, that is, the ability of distinguishing the classes. One problem of DTs is that they tend to overfit if they are induced to classify all training points correctly, meaning that the obtained solution can achieve good results only when applied to the same dataset that was used for its training. Overfitting can be avoided by DTs in multiple ways. One of them is pruning branches of the DT. Other strategy, employed in this work, is to join multiple trees trained using bootstrapping samples from the original dataset. From this point of this text, DT associated with the bootstrapping method is referred simply as DT. RF is another ensemble of tree-based models (Ho, 1995) which also randomly chooses subsets of input features from the original dataset in the bootstrapping procedure.

### 3.3. Artificial neural networks (ANN)

ANN are based on the brain structure and processing. Their fundamental units, the neurons, communicate to each other using weighted signals that usually belong to the [0,1] interval. The output of a neuron can be an input of another neuron, so that multiple layers of neurons can be combined. The neuron model presented in Figure 2 is called McCulloch

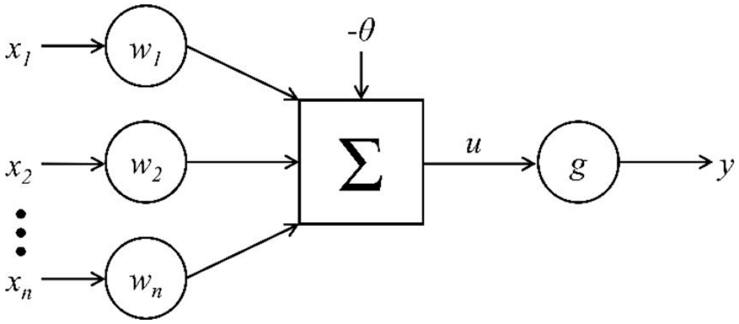

**Figure 2.** ANN neuron. Adapted from Carvalho et al. (2019).

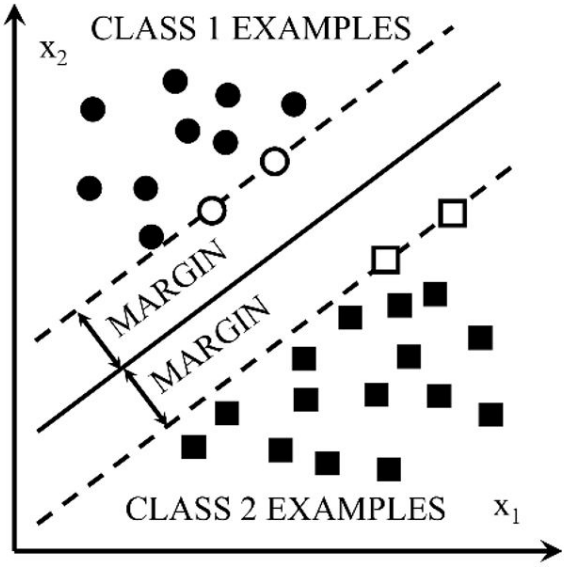

**Figure 3.** Hyperplane dividing the input space.

& Pitts (MCP) model and is used in the perceptron ANN (McCulloch & Pitts, 1943).

The MCP neuron receives input signals $x_i$, which are multiplied by weights $w_i$ and summed up. After an excitation threshold $\theta$ is discounted, a signal is produced. This signal is input to an activation function $g$, generating an output signal $y$. In the original MCP model, the activation function is a stepwise or signal function. Alternative functions, including non-linear functions, can provide more representative power to the ANN models.

If many artificial neurons are combined in layers, the model is called multi-layer perceptron neural network (Rumelhart et al., 1986). In this work, ANN architectures using up to two hidden layers were tested. The output layer has one neuron representing each class. The neuron outputting the highest value defines the final classification.

One can demonstrate that a network with a single hidden layer of neurons with non-linear activation functions can reproduce any continuous function, and that a network with two hidden layers of such neurons can reproduce any function (Hornik et al., 1989). Considering that a limit must be imposed to select among infinite possible architectures, in this work networks with three or more hidden layers are not tested.

### 3.4. Support vector machines (SVM)

In its simplest version, the SVM technique divides the input space with a hyperplane and assigns one class to each side. The optimal hyperplane seeks to maximize the margin of separation between both classes, as illustrated in Figure 3.

The support vectors correspond to examples that are placed over the margin limits after the hyperplane is defined. In Figure 3, for example, four support vectors are represented, two white circles and two white squares. In this work a soft-margin version of SVM is used, being possible that points remain within the margins or even on the wrong side of the decision border.

One limitation of this version of SVM is that it admits only linear separations between the classes. One way of extending the SVM to solve non-linear classification problems is by mapping the original input space into a higher dimension space, using a function called kernel. After preliminary tests, the polynomial kernel was chosen here due to its better predictive performance compared to other types of kernel functions. Considering $x$ and $y$ two points in input space, the polynomial kernel can be written as:

$$k\left(x,y\right)=\left(\delta\left(x\bullet y\right)+\kappa\right)^{\alpha} \tag{10}$$

where $\delta$, $\kappa$ and $\alpha$ are calibration parameters.

Although the described version of SVM is defined only for separating two classes, it is possible to extend it to multi-class problems by simply combining two or more binary classifiers. In this procedure, all classes must be evaluated in pairs, generating $\left(c\,2\right)$ classifiers for $c$ classes.

## 4. Data analysis

The analysis performed in this paper use the following parameters from CPT soundings:

- $B_q$: Dimensionless pore pressure normalization used by Robertson (1991).
- $F_r$: Dimensionless lateral friction normalization used by Robertson (2016).
- $f_s$: Lateral friction, measured in kPa.
- $q_c$: Uncorrected cone resistance, measured in MPa.
- $q_t$: Total cone resistance, calculated in MPa.
- $Q_{t1}$: Dimensionless cone resistance normalization used by Robertson (1990).
- $Q_{tn}$: Dimensionless cone resistance normalization used by Robertson (2016).
- $SA$: Soil age, represented by a dimensionless discrete number related to the geological epoch when the soil was deposited.
- $u_2$: Pore pressure in a disturbed state, measured in kPa.
- $U_2$: Dimensionless pore pressure normalization used by Robertson (2016).
- $z$ : Depth measured from the surface in m.

### 4.1. Description of the used datasets

Professor P. K. Robertson provided the 38 soundings described in Table 2 and Professor P. W. Mayne provided the 73 soundings described in Table 3. The information given by these 111 soundings compose the dataset used in the main studies of this work; therefore, it is hereafter named Main dataset.

A second dataset, here named Geological dataset, is gathered to investigate the influence of soil age within soil classification. The motivation for its usage is the difficulty reported in the literature for classifying aged soil (Robertson, 2016). A variable called soil age ($SA$) is then proposed, which is represented by a number related to the geological age when the soil was deposited. The Geological dataset, which is described in Table 4, uses information only from the 38 soundings provided by Robertson because no information about soil age was available for the other soundings.

The third dataset used in this work is composed of 64 CPT soundings from the metropolitan area of São Paulo, Brazil, being here named Tropical dataset. Measurements were taken at each 2 cm of depth and included more than forty thousand soil examples. These soundings were provided by the São Paulo Metropolitan Company under a confidentiality term, so most information about it cannot be exposed here.

Robertson charts were produced using samples taken from temperate regions, which can lead to uncertainty when applied to tropical soil. To discuss this issue, in section 5.2 the Tropical dataset is used to test if the performance of the ML techniques remains accurate. The study is divided in two parts, in the first the Main dataset is used for training the ML techniques and the Tropical dataset is used for testing. The objective of this first part is discussing if Brazilian soil can be accurately classified using soil information from other countries. In the second part, the Tropical dataset is used for both training and testing, aiming to observe if accuracy raises when compared to the first part. Figure 4 presents data of one of the CPT soundings to illustrate the used data.

**Table 2.** Dataset from P. K. Robertson. Adapted from Carvalho & Ribeiro (2019).

| Soil type | Location | Soundings |
|---|---|---|
| Mixed Soils | Canada | 3 |
| | Italy | 1 |
| | USA | 6 |
| | Switzerland | 1 |
| Soft Clay | UK | 1 |
| | Australia | 1 |
| | Norway | 1 |
| | USA | 3 |
| | Canada | 2 |
| | Sweden | 2 |
| | North Sea | 1 |
| | Very soft offshore | 1 |
| Soft Rock | USA | 4 |
| Stiff Clay | UK | 3 |
| | USA | 4 |
| | Italy | 1 |
| | France | 1 |
| | Ireland | 1 |
| | Alaska (USA) | 1 |
| Total | | 38 |

**Table 3.** Dataset from P. W. Mayne. Adapted from Carvalho & Ribeiro (2019).

| Location in USA | Soundings |
|---|---|
| Gosnell, Arkansas | 1 |
| Lenox, Tennessee | 1 |
| Memphis, Tennessee | 16 |
| Dexter, Missouri | 6 |
| Mooring, Tennessee | 6 |
| Marked Tree, Arkansas | 19 |
| Collierville, Tennessee | 1 |
| Meramec, Missouri | 4 |
| Opelika, Alabama | 4 |
| Wilson, Arkansas | 4 |
| Wolf, Wyoming | 7 |
| Wyatt, Missouri | 4 |
| Total | 73 |

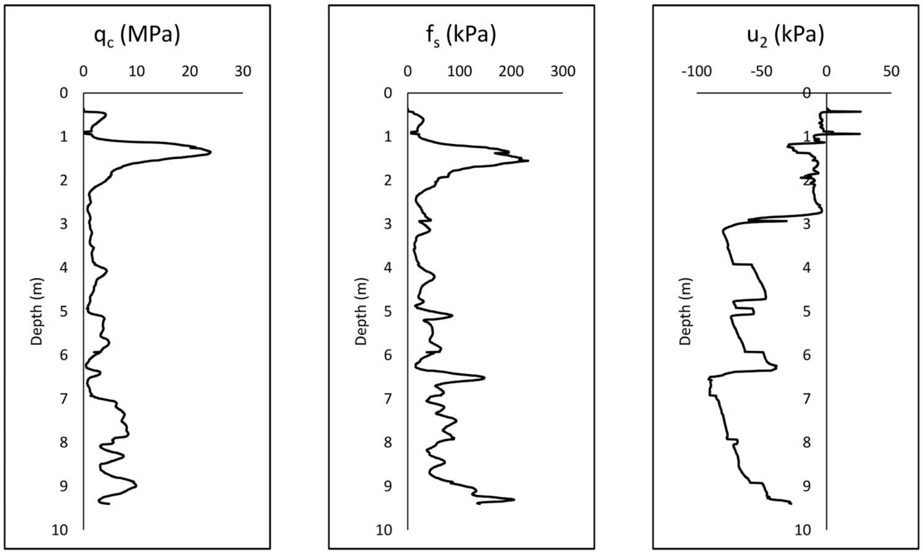

**Figure 4.** Illustration of the data recovered from one of the CPT soundings.

**Table 4.** Geological dataset. Adapted from Carvalho & Ribeiro (2019).

| Soil type | Identification | Geological age | SA |
|---|---|---|---|
| Mixed Soils | UBC, Canada | Holocene | 2 |
| | Venetian Lagoon, Italy | Holocene | 2 |
| | Ford Center, USA | Pleistocene | 4 |
| | San Francisco, USA | Late Pleistocene | 3 |
| | Tailings, USA | Recent | 1 |
| | UBC KIDD, Canada | Holocene | 2 |
| | UBC KIDD, Canada (2) | Holocene | 2 |
| Soft Clay | Bothkennar, RU | Holocene | 2 |
| | Burswood, Perth, Australia | Holocene | 2 |
| | Onsoy, Norway | Holocene | 2 |
| | Amherst, USA | Late Pleistocene | 3 |
| | San Francisco Bay, USA | Holocene | 2 |
| | San Francisco Bay, USA (2) | Holocene | 2 |
| Soft Rock | Newport Beach, USA | Miocene | 5 |
| | LA Downtown, USA | Miocene | 5 |
| | Newport Beach, USA (2) | Miocene | 5 |
| Stiff Clay | Madingley, UK | Cretaceous | 6 |
| | Houston, USA | Pleistocene | 4 |

### 4.2. Data preprocessing

As data-driven techniques, ensuring data quality is important when ML techniques are concerned. The identification and treatment of outliers, which are inputs with discrepant values, is one of the important steps for a proper data cleansing. One way of automatically detecting potential outliers is by the use of boxplots. Nonetheless, preliminary tests have shown that removing all potential outliers severely reduces accuracy. In this work, this problem is avoided by applying the Edit Nearest Neighbor technique (Wilson, 1972). It compares the classes of the potential outlier and its nearest neighbors, removing it only if their labels do not match.

Another problem is an imbalance within classes, which can bias the ML techniques towards the majority class in detriment of classes with less examples. An evaluation based on histograms allowed identifying some issues, solved as listed next:

1) There were too few ISG class 0 examples, therefore they were completely removed from the datasets. FSB class 0 examples were maintained;
2) ISG classes were very imbalanced within the Geological dataset, therefore all analysis with this dataset were restricted to the FSB method;
3) Random sampling was applied to reduce majority classes, considering that CPT data contains several redundancies due to many measurements taken within each soil layer;
4) Minority classes were incremented applying the SMOTE oversampling technique (Chawla et al., 2002).

After procedures 3 and 4, all classes have the same number of examples. A second data transformation is applied for the ANN, SVM and MMP analyses, imposing a logarithmic scale to each input feature. This procedure was adopted because the original charts from Robertson use logarithmic scale and preliminary tests showed that better performance is achieved with this transformation. Figure 5 shows an example of the logarithmic scale effect.

### 4.3. General methodology

The 10-fold cross-validation procedure is applied for each dataset and input combination. In this process, the original dataset is divided into 10 partitions called folds, in which the class proportion is kept the same as in the original dataset. Among these 10 folds, one is used for testing, one is used for validation and the remaining compose the training set. The training set is the only one subject to all preprocessing procedures and is used as a reference for all predictions. The validation fold is used to calibrate the parameters of each technique and the testing fold is used to measure predictive performance for new data points previously unseen by the ML techniques. At each step of the 10-step procedure a different testing fold is selected, and the final predictive performance is given by the average and standard deviation of the ten values obtained.

The most common performance metric adopted in multi-class problems is accuracy, which is given by the total number of correct predictions divided by the total number of objects. Nevertheless, majority classes can bias this measurement once the testing and validation folds are not balanced. To solve this problem, the predictive performance measure used in this work is obtained by calculating accuracy for each class separately and then calculating their mean value. This value would be the accuracy if the classes were balanced and had the same number of objects. For simplicity, this performance measure is called accuracy here, although it is commonly referred as balanced accuracy in the ML literature.

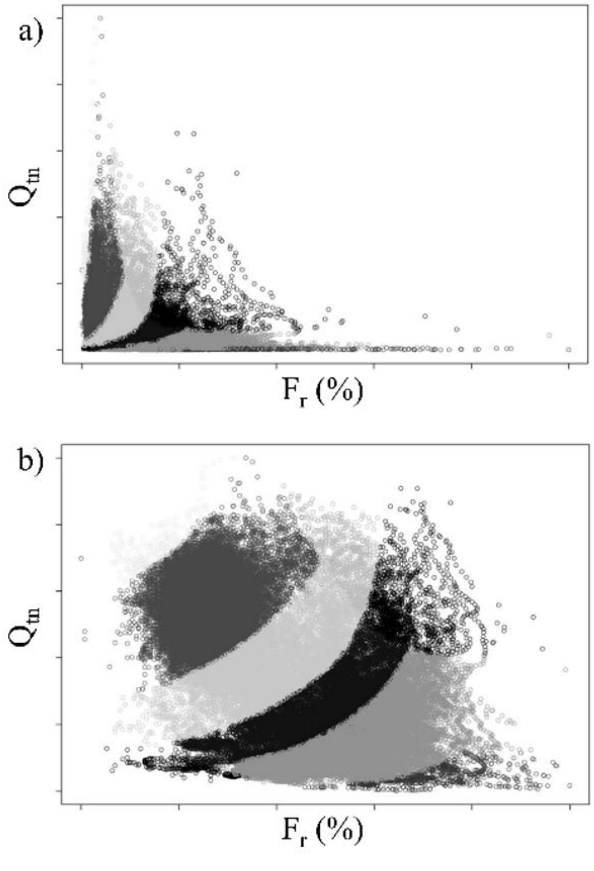

**Figure 5.** Example of logarithmic scale effect: (a) without logarithmic scale; (b) with logarithmic scale.

The calibration process performed for each technique is described in Section 3.

### 4.4. Comments about the inputs

Many variables mentioned in previous sections can be used as inputs for the ML techniques. Specific combinations are selected here considering previous work from the authors (Carvalho & Ribeiro, 2019; Carvalho et al., 2019) and the objectives of the present study. These combinations are:

1) $z$, $q_t$, $f_s$ and $u_2$: Raw CPT measurements, except for the correction of the cone tip resistance from $q_c$ to $q_t$;
2) $z$, $Q_{t1}$, $F_r$ and $B_q$: Depth plus normalizations proposed by Robertson (1990);
3) $z$, $Q_{tn}$, $F_r$ and $U_2$: Depth plus normalizations proposed by Robertson (2016);
4) $Q_{tn}$, $F_r$: Inputs used by the ISG method;
5) $Q_{tn}$, $F_r$ and $U_2$: Inputs used by the FSB method;
6) $z$, $Q_{t1}$, $F_r$, $B_q$ and $SA$: Depth plus normalizations proposed by Robertson (1990) plus soil age;
7) $z$, $Q_{tn}$, $F_r$, $U_2$ and $SA$: Depth plus normalizations proposed by Robertson (2016) plus soil age;
8) $z$, $q_c$ and $f_s$: Raw CPT measurements, excluding $u_2$ and not correcting $q_c$ to $q_t$.

The use of combination 1 has the objective of evaluating how accurately ISG and FSB can be replicated without using the normalizations proposed by Robertson. Combinations 2 and 3 aim to test predictive performance when such normalizations are combined to depth. The original input combinations 4 and 5 are used as a reference, while combinations 6 and 7 aim to evaluate if soil age improves predictive performance. The last combination 8 refers to CPT equipment which cannot measure pore pressure, making impossible to correct $q_c$ to $q_t$.

## 5. Results and discussion

### 5.1. General performance for replicating ISG and FSB

Results in this section refer to the general performance of the ML techniques when applied to the Main and Geological datasets. These results are summarized in Table 5, where each

**Table 5.** MA results obtained with all techniques (%).

| Input | Output | DWNN | DT | RF | ANN | SVM | MMP |
|---|---|---|---|---|---|---|---|
| $z\ q_t f_s\ u_2$ | ISG | 90.23 | 91.71 | 91.53 | 91.57 | 92.63 | 93.17 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 89.40 | 95.81 | 96.09 | 93.59 | 96.47 | 96.64 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 93.13 | 97.60 | 97.44 | 96.56 | 97.97 | 98.25 |
| $Q_{tn}\ F_r$ | | 96.58 | 96.97 | 97.31 | 96.48 | 98.03 | 97.95 |
| $z\ q_t f_s\ u_2$ | FSB | 90.28 | 91.32 | 91.43 | 85.88 | 87.57 | 91.82 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 88.82 | 96.40 | 96.38 | 84.39 | 91.24 | 96.15 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 93.77 | 97.31 | 97.27 | 86.18 | 92.77 | 97.08 |
| $Q_{tn}\ F_r\ U_2$ | | 93.06 | 94.69 | 94.63 | 82.85 | 89.24 | 94.66 |
| $z\ q_t f_s\ u_2\ SA$ | | 91.03 | 91.66 | 91.78 | 87.79 | 89.75 | 93.23 |
| $z\ Q_{tn}\ F_r\ U_2\ SA$ | | 94.73 | 97.01 | 97.31 | 89.96 | 94.44 | 97.42 |

line represents a 10-fold cross validation test (see Section 4.3). The first column presents the used inputs, and the second column represents the replicated method, ISG (Section 2.1) or FSB (Section 2.2). Considering that 10 tests (one for each fold) are made for each line, resulting in 10 separated accuracy measurements, other columns represent their mean value (MA stands for mean accuracy) for each technique. One can calculate MA from the individual accuracies $Ac_i$ using the expression:

$$MA = \frac{\sum_{i=1}^{10} Ac_i}{10} \tag{11}$$

One can observe that MA is above 91% in all lines for MMP, which can be considered a good predictive performance for soil profiling. In most cases MMP presents best performance, in others it presents a performance close to the best one. Results obtained with $z$, $q_t$, $f_s$ and $u_2$ show that accurate soil classification is possible without the data transformations proposed by Robertson. As expected, high accuracies are obtained when the original inputs are used for each method, $Q_{tn}$ and $F_r$ for ISG and $Q_{tn}$, $F_r$ and $U_2$ for FSB. Nonetheless, the highest accuracy for ISG was obtained when $z$, $Q_{tn}$, $F_r$ and $U_2$ were used as inputs for MMP and the highest accuracy for FSB was achieved when $z$, $Q_{tn}$, $F_r$, $U_2$ and $SA$ were used for MMP. This suggests that including depth as an input brings relevant information to soil classification. Reasonable accuracy was obtained for ANN and SVM only after applying logarithmic scale, as presented in Figure 5.

**Table 6.** Results obtained with MMP without the FSB class $0$ (%).

| Input | Output | MA | SD |
|---|---|---|---|
| $z\ q_t f_s\ u_2$ | FSB | 92.62 | 0.39 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 98.55 | 0.18 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 99.41 | 0.15 |
| $Q_{tn}\ F_r\ U_2$ | | 99.60 | 0.12 |
| $z\ q_t f_s\ u_2\ SA$ | | 92.37 | 1.19 |
| $z\ Q_{tn}\ F_r\ U_2\ SA$ | | 98.35 | 0.63 |

Preliminary tests have shown that objects assigned to class $0$ in FSB prejudice the predictive performance of ANN and SVM. In order to quantify this influence, additional experiments were performed removing these objects from the training and test sets, resulting the values presented in Table 6. SD stands for standard deviation and, for a sake of conciseness, only results for MMP are presented. As the proposal is to focus on the FSB method, results from the ISG method are omitted. One can observe that a higher MA is achieved for most of the cases, including values close to

100%. This suggests that objects assigned to the class 0 of the FSB method do not form a homogeneous region within input space, making the classification problem harder.

In order to complement the application of ML techniques for soil profiling, the MMP was employed to determine the soil profile according to the ISG method for a sounding taken in Vancouver, Canada and provided by Professor Renato da Cunha (Cunha, 1994). The Main dataset was used for training. Comparing the result obtained with CPeT-IT v2.0.2.5 to the one obtained with the MMP they are almost the same, with an accuracy of 95.4%.

### 5.2. Study with the Tropical dataset

Once the DWNN technique did not present good performance its results are omitted, as well as some input combinations tested in Section 5.2, to avoid redundancy.

Results from the first part of the study are shown in Table 7. One can observe that, even though the multiple model is not the best performing technique for all testing combinations, its performance is in general close to the best one. This shows that MMP is stable, while larger variations can be observed for the other techniques. Comparing Table 7 to Table 5, one can observe that accuracy drops in all cases.

Results from the second part of the study are presented in Table 8. The general behavior of the MMP is maintained, presenting stability and good performance when compared to other techniques. In some cases, accuracies close to 100% were obtained, showing that the information of the Tropical dataset is substantially different from the information of the Main dataset. This suggests that it is justifiable to develop new soil classification methods specific for tropical soil.

### 5.3. Soil classification without measuring the pore pressure

Once not all CPT equipment available in the market measure the pore pressure $u_2$, one could question if this variable

**Table 7.** MA results for the first part (%).

| Input | Output | DT | RF | ANN | SVM | MMP |
|---|---|---|---|---|---|---|
| $z\ q_t f_s\ u_2$ | ISG | 68.70 | 62.92 | 74.68 | 77.20 | 71.29 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 88.03 | 88.27 | 85.55 | 85.84 | 87.94 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 89.94 | 89.80 | 92.98 | 89.86 | 91.50 |
| $z\ q_t f_s\ u_2$ | FSB | 79.50 | 79.49 | 82.82 | 82.68 | 82.01 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 92.98 | 92.76 | 88.68 | 92.86 | 92.99 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 95.87 | 95.78 | 92.41 | 95.78 | 95.87 |

**Table 8.** MA results for the second part (%).

| Input | Output | DT | RF | ANN | SVM | MMP |
|---|---|---|---|---|---|---|
| $z\ q_t f_s\ u_2$ | ISG | 86.82 | 87.94 | 84.64 | 85.36 | 89.42 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 95.00 | 95.17 | 76.83 | 92.97 | 95.34 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 97.02 | 96.78 | 84.88 | 96.08 | 96.66 |
| $z\ q_t f_s\ u_2$ | FSB | 89.42 | 89.71 | 80.57 | 84.52 | 90.67 |
| $z\ Q_{t1}\ F_r\ B_q$ | | 96.86 | 96.85 | 27.61 | 69.89 | 90.47 |
| $z\ Q_{tn}\ F_r\ U_2$ | | 98.67 | 98.60 | 40.43 | 86.24 | 96.70 |

**Table 9.** Results using $z$ , $q_c$ and $f_s$ as inputs (%).

| Technique | ISG | | FSB | | FSB (no 0) | |
|---|---|---|---|---|---|---|
| | MA | SD | MA | SD | MA | SD |
| DWNN | 90.30 | 0.65 | 86.79 | 0.47 | 88.12 | 0.72 |
| DT | 90.13 | 0.56 | 87.35 | 0.45 | 88.99 | 0.40 |
| RF | 90.31 | 0.60 | 87.84 | 0.35 | 89.37 | 0.32 |
| ANN | 90.18 | 0.85 | 82.86 | 0.64 | 85.82 | 1.08 |
| SVM | 91.04 | 0.70 | 82.44 | 0.48 | 85.58 | 0.82 |
| MMP | 91.83 | 0.56 | 87.58 | 0.32 | 89.15 | 0.30 |

is really needed for soil classification. Consulting Section 2 one quickly concludes that, without $u_2$, classifying soil within the original ISG and FSB methods is not possible. Pore pressure $u_2$ plays a fundamental role throughout the methodology proposed, not only for correcting cone resistance but also for calculating stresses and obtaining the final normalizations. Therefore, since the approach presented here simply replicates those charts, one should not conclude from this study that measuring $u_2$ could be neglected for soil classification in real engineering projects. Nonetheless, the aim here is to start a discussion in this direction, possibly leading to further studies with conclusions that are more consistent.

In this context, additional experiments were performed to verify if the friction penetrometer without the pore pressure filter could provide enough information for obtaining a rough approximation of the soil classes. Therefore, all techniques plus the MMP were tested with the Main dataset using only $z$, $q_c$ and $f_s$ as inputs, resulting the values presented in Table 9. This study was replicated for the ISG method, for the FSB method with class 0 objects and for the FSB method without class 0 objects.

One can notice that all techniques achieved accuracy higher than 90% for the ISG method, which can be considered reasonable for soil profiling. Although lower accuracies were obtained for the FSB method, the accuracy values can also be considered practicable, especially when objects assigned to the class 0 are removed. These results show that, for this specific dataset, soil can be classified within reasonable accuracy with CPT data that do not include pore pressure filter measurements.

## 6. Conclusions and recommendations

A general methodology for the application of ML techniques for soil classification from CPT data is presented in this paper, including six ML techniques of distinct biases: DWNN, DT, RF, ANN, SVM and MMP, which is a combination of the previous techniques. MMP joins the predictions of the multiple individual models by majority voting, producing a heterogeneous ensemble of classifiers. All techniques are applied initially to a dataset composed of 111 CPT soundings, testing different input combinations within a 10-fold cross-validation procedure. Training data is also subject to a preprocessing procedure within each 10-fold cross-validation step for improving data quality, including data transformation, cleaning and balancing. Tests are also performed with two other datasets, one containing soil age information and the other with tropical soil information. The original CPT measurements included within the analysis are depth $z$, cone resistance $q_c$ and corrected cone resistance $q_t$, lateral friction $f_s$ and pore pressure $u_2$. Included normalizations are the cone resistances $Q_{t1}$ and $Q_{tn}$, the lateral friction $F_r$ and the pore pressures $B_q$ and $U_2$. A soil age *SA* parameter was also included, representing the geological age when the soil was deposited.

The machine learning techniques were successfully compared and combined in an ensemble that produces more accurate results that any isolated technique. MMP can be also considered the most stable technique, with accuracies above 93% in most cases. The predictive results in the classification of soil samples from tropical areas are in general inferior to those recorded for soil from temperate areas, especially when the models built from temperate areas are employed in the classification of soil from tropical areas. This indicates the need to develop classification methods specific for tropical soil, which the authors suggest as future work. Another important observation is that accuracy remains reasonable for all techniques even if pore pressure information is omitted during training. These results can encourage future work pursuing soil classification methods that do not use pore pressure information, which can be costly to measure and requires specialized equipment. The results do not allow concluding that pore pressure measurements can be dismissed in real engineering projects, but that soil classes can be roughly approximated without this information. This can become an alternative for initial geotechnical studies in underdeveloped and developing countries, where budget constrains limit engineering practice.

It is important to notice that none of these discussions would be possible by using the original Robertson charts alone, once these methods do not allow changing inputs or using incomplete data.

## Acknowledgements

To Peter K. Robertson, Paul W. Mayne, Renato da Cunha and São Paulo Metropolitan Company for making available the CPT soundings used in this work. This research received no external funding.

## Declaration of interest

The authors declare no conflict of interest.

## Authors' contributions

Lucas Orbolato Carvalho: formal analysis, methodology, writing – original draft, validation. Dimas Betioli Ribeiro: investigation, project administration, supervision, writing – review & editing.

## List of symbols

| | |
|---|---|
| ANN | Artificial neural networks. |
| CPT | Cone penetration test. |
| DT | Decision trees. |
| DWNN | Distance-weighted nearest neighbors. |
| FSB | Focused on soil behavior. |
| ISG | Influenced by soil granulometry. |
| MA | Mean accuracy. |
| ML | Machine learning. |
| MMP | Multiple model predictor. |
| MCP | McCulloch & Pitts model. |
| RF | Random forests. |
| SA | Soil age. |
| SD | Standard deviation. |
| SMOTE | Synthetic minority over-sampling technique. |
| SPT | Standard penetration test. |
| SVM | Support vector machines. |
| $z$ | Depth. |
| $q_c$ | Uncorrected cone resistance. |
| $f_s$ | Lateral friction. |
| $u_2$ | Pore pressure in a disturbed state. |
| $q_t$ | Total cone resistance. |
| $u_0$ | Equilibrium pore pressure. |
| $\sigma'_{v0}$ | Effective overburden stress. |
| $\sigma_{v0}$ | Total overburden stress. |
| $q_n$ | Net cone resistance. |
| $Q_{t1}$, $F_r$, $B_q$ | Normalizations proposed by Robertson (1990). |
| $Q_{tn}$, $U_2$ | Normalizations used by Robertson (2016). |
| $n$ | Exponent used to calculate $Q_{tn}$. |
| $pa$ | Reference pressure of 0.1 MPa. |
| $I_c$ | Parameter used to calculate $n$. |
| $x$, $y$ | Objects at the input space. |
| $d$ | Distance between two points. |
| $w$ | Gaussian weighting. |
| $w_i$, $\theta$, $u$, $g$, $y$ | Parameters used in the perceptron neuron. |
| $\delta$, , $\alpha$ | Calibration parameters of the polynomial kernel. |
| $c$ | Number of classes. |
| $SA$ | Soil age. |

## References

Arel, E. (2012). Predicting the spatial distribution of soil profile in Adapazari/Turkey by artificial neural networks using CPT data. *Computers & Geosciences*, 43, 90-100. http://dx.doi.org/10.1016/j.cageo.2012.01.021.

Begemann, H.K.S.P. (1965). The friction jacket cone as an aid in determining the soil profile. In *Proceedings of the 6th International Conference on Soil Mechanics and Foundation Engineering* (pp. 17-20).

Bhattacharya, B., & Solomtine, D.P. (2006). Machine learning in soil classification. *Neural Networks*, 19(2), 186-195. http://dx.doi.org/10.1016/j.neunet.2006.01.005.

Carvalho, L.O., & Ribeiro, D.B. (2019). Soil classification system from cone penetration test data applying distance-based machine learning algorithms. *Soils and Rocks*, 42(2), 167-178. http://dx.doi.org/10.28927/SR.422167.

Carvalho, L.O., & Ribeiro, D.B. (2020). Application of kernel k-means and kernel x-means clustering to obtain soil classes from cone penetration test data. *Soils and Rocks*, 43(4), 607-618. http://dx.doi.org/10.28927/SR.434607.

Carvalho, L.O., Ribeiro, D.B., & Monteiro, F.A.C. (2019). Comparing artificial neural networks with support vector machines for soil classification. In *Proceedings of the XL Ibero-latin american congress on computational methods in engineering* (pp. 11-14). Natal/RN, Brazil.

Chawla, N.V., Bowyer, K.W., Hall, L.O., & Kegelmeyer, W.P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357. http://dx.doi.org/10.1613/jair.953.

Cunha, R.P. (1994). *Interpretation of selfboring pressuremeter tests in sand* [Doctoral thesis, University of British Columbia]. University of British Columbia's repository. http://dx.doi.org/10.14288/1.0050418.

Das, S.K., & Basudhar, P.K. (2009). Utilization of self-organizing map and fuzzy clustering for site characterization using piezocone data. *Computers and Geotechnics*, 36(1-2), 241-248. http://dx.doi.org/10.1016/j.compgeo.2008.02.005.

Douglas, B.J., & Olsen, R.S. (1981). Soil classification using electric cone penetrometer. In *Proceedings of the Symposium on Cone Penetration Testing and Experience* (pp. 209-227). Reston: ASCE.

Dudani, S.A. (1976). The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(4), 325-327. http://dx.doi.org/10.1109/TSMC.1976.5408784.

Facciorusso, J., & Uzielli, M. (2004). Stratigraphic profiling by cluster analysis and fuzzy soil classification from mechanical cone penetration tests. In *Proceedings of the ISC-2 on Geotechnical and Geophysical Site Characterization* (pp. 905-912). Porto.

Goh, A.T., & Goh, S. (2007). Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data. *Computers and Geotechnics*, 34(5), 410-421. http://dx.doi.org/10.1016/j.compgeo.2007.06.001.

Goh, A.T.C. (1995). Modeling soil correlations using neural networks. *Journal of Computing in Civil Engineering*, 9(4), 275-278. http://dx.doi.org/10.1061/(ASCE)0887-3801(1995)9:4(275).

Goh, A.T.C. (1996). Neural-network modeling of CPT seismic liquefaction data. *Journal of Geotechnical Engineering*, 122(1), 70-73. http://dx.doi.org/10.1061/(ASCE)0733-9410(1996)122:1(70).

Hanna, A.M., Ural, D., & Saygili, G. (2007). Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data. *Soil Dynamics and Earthquake Engineering*, 27(6), 521-540. http://dx.doi.org/10.1016/j.soildyn.2006.11.001.

Hegazy, Y.A., & Mayne, P.W. (2002). Objective site characterization using clustering of piezocone data. *Journal of Geotechnical and Geoenvironmental Engineering*, 128(12), 986-996. http://dx.doi.org/10.1061/(ASCE)1090-0241(2002)128:12(986).

Ho, T.K. (1995). Random decision forests. In *Proceedings of the 3rd International Conference on Document Analysis and Recognition* (pp. 278-282). USA: IEEE. http://dx.doi.org/10.1109/ICDAR.1995.598929.

Hornik, K., Stinchombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366. http://dx.doi.org/10.1016/0893-6080(89)90020-8.

Jefferies, M.G., & Davies, M.P. (1991). Soil classification by the cone penetration test. *Canadian Geotechnical Journal*, 28(1), 173-176. http://dx.doi.org/10.1139/t91-023.

Juang, C.H., & Chen, C.J. (1999). CPT-Based liquefaction evaluation using neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 14(3), 221-229. http://dx.doi.org/10.1111/0885-9507.00143.

Juang, C.H., Lu, P.C., & Chen, C.J. (2002). Predicting geotechnical parameters of sands from CPT measurements using neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 17(1), 31-42. http://dx.doi.org/10.1111/1467-8667.00250.

Juang, C.H., Yuan, H., Lee, D.H., & Lin, P.S. (2003). Simplified cone penetration test-based method for evaluating liquefaction resistance of soils. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(1), 66-80. http://dx.doi.org/10.1061/(ASCE)1090-0241(2003)129:1(66).

Kohestani, V.R., Hassanlourad, M., & Ardakani, A. (2015). Evaluation of liquefaction potential based on CPT data using random forest. *Natural Hazards*, 79, 1079-1089. http://dx.doi.org/10.1007/s11069-015-1893-5.

Kumar, J.K., Konno, M., & Yasuda, N. (2000). Subsurface soil-geology interpolation using fuzzy neural network. *Journal of Geotechnical and Geoenvironmental Engineering*, 126(7), 632-639. http://dx.doi.org/10.1061/(ASCE)1090-0241(2000)126:7(632).

Kurup, P.U., & Griffin, E.P. (2006). Prediction of soil composition from CPT data using general regression neural network. *Journal of Computing in Civil Engineering*, 20(4), 281-289. http://dx.doi.org/10.1061/(ASCE)0887-3801(2006)20:4(281).

Liao, T., & Mayne, P.W. (2007). Stratigraphic delineation by three-dimensional clustering of piezocone data. *Georisk*, 1(2), 102-119. http://dx.doi.org/10.1080/17499510701345175.

Livingston, G., Piantedosi, M., Kurup, P., & Sitharam, T.G. (2008). Using decision-tree learning to assess liquefaction potential from CPT and $V_s$. *Geotechnical Earthquake Engineering and Soil Dynamics IV Congress*, 2008, 1-10. http://dx.doi.org/10.1061/40975(318)76.

McCulloch, W., & Pitts, W. (1943). A Logical calculus of ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5, 115-133. http://dx.doi.org/10.1007/BF02478259.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, 1, 81-106. http://dx.doi.org/10.1007/BF00116251.

Ramsey, N. (2002). A calibrated model for the interpretation of cone penetration tests (CPTs) in north sea quaternary soils. In *Proceedings of the Offshore site investigation and geotechnics 'Diversity and Sustainability' conference* (pp. 341-356). London. SUT.

Reale, C., Gavin, K., Libric, L., & Juric-Kacunic, D. (2018). Automatic classification of fine-grained soils using CPT measurements and Artificial Neural Networks. *Advanced Engineering Informatics*, 36, 207-215. http://dx.doi.org/10.1016/j.aei.2018.04.003.

Robertson, P.K. (1990). Soil classification using the cone penetration test. *Canadian Geotechnical Journal*, 27(1), 151-158. http://dx.doi.org/10.1139/t90-014.

Robertson, P.K. (1991). Soil classification using the cone penetration test. *Canadian Geotechnical Journal*, 28(1), 176-178. http://dx.doi.org/10.1139/t91-024.

Robertson, P.K. (2009). Interpretation of cone penetration tests – a unified approach. *Canadian Geotechnical*

*Journal*, 46(11), 1337-1355. http://dx.doi.org/10.1139/T09-065.

Robertson, P.K. (2016). Cone penetration test (CPT)-based soil behaviour type (SBT) classification system – an update. *Canadian Geotechnical Journal*, 53(12), 1910-1927. http://dx.doi.org/10.1139/cgj-2016-0044.

Robertson, P.K., & Wride, C.E. (1998). Evaluating cyclic liquefaction potential using the cone penetration test. *Canadian Geotechnical Journal*, 35(3), 442-459. http://dx.doi.org/10.1139/t98-017.

Robertson, P.K., Campanella, R.G., Gillespie, D., & Greig, J. (1986). Use of piezometer cone data. In *Proceedings of the SITU' 86 Use of In-Situ Testing in Geotechnical Engineering, ASCE Specialty Conference* (pp. 1263-1280).

Rogiers, B., Mallants, D., Batelaan, O., Gedeon, M., Huysmans, M., & Dassargues, A. (2017). Model-based classification of CPT data and automated lithostratigraphic mapping for high-resolution characterization of a heterogeneous sedimentary aquifer. *PLoS One*, 12, e0176656. http://dx.doi.org/10.1371/journal.pone.0176656.

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning representations by back-propagation errors. *Nature*, 323, 533-536. http://dx.doi.org/10.1038/323533a0.

Schaap, M.G., Leij, F.J., & van Genuchten, M.T. (1998). Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, 62(4), 847-855. http://dx.doi.org/10.2136/sssaj1998.03615995006200040001x.

Schneider, J.A., Hotstream, J.N., Mayne, P.W., & Randolph, M.F. (2012). Comparing CPTU *Q–F* and Q–$\Delta u_2/\sigma_{v0}'$ soil classification charts. *Géotechnique Letters*, 2(4), 209-215. http://dx.doi.org/10.1680/geolett.12.00044.

Schneider, J.A., Randolph, M.F., Wayne, P.W., & Ramsey, N.R. (2008). Analysis of factors influencing soil classification using normalized piezocone tip resistance and pore pressure parameters. *Journal of Geotechnical and Geoenvironmental Engineering*, 134(11), 1569-1586. http://dx.doi.org/10.1061/(ASCE)1090-0241(2008)134:11(1569).

Wilson, D.L. (1972). Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-2(3), 408-421. http://dx.doi.org/10.1109/TSMC.1972.4309137.