XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGENEERING

Application of a Bayesian Approach for Soil Classification in Brazillian Marine Soil: Benchmarking Against Robertson Chart and Fuzzy Techniques

João V. G. Fernandes[1*], Christiano Várady[1], Joyce Tenório[1], Eduardo T. L. Junior[1], João P. L. Santos[1], Rafael Dias[2]

[1]*Dept. of Engineering, Federal University of Alagoas*
Av. Lourival Melo Mota, S/n - Tabuleiro do Martins, Maceió - AL, 57072-970, Alagoas, Brazil
joao.fernandes@ctec.ufal.br, christiano_varady@lccv.ufal.br, joyce.tenorio@lccv.ufal.br, limajunior@lccv.ufal.br, jpls@lccv.ufal.br
[2]*PETROBRAS*
Av. Horácio Macedo, 950 - Cidade Universitária, Rio de Janeiro - RJ, 21941-915
rafael_dias@petrobras.com.br

Abstract. This paper reproduces a Bayesian approach for classifying soil stratigraphy using CPTu data, with explicit incorporation of uncertainties. The approach estimates the most probable number of subsurface layers, their thicknesses, and soil types by combining the Robertson classification chart with a Bayesian inference. The approach consists of two main steps: model class selection and system identification. Unlike traditional methods, this approach integrates prior knowledge with CPTu data from the Campos Basin, Southeastern Brazil. The results cohere with both Robertson and fuzzy-based classifications, while providing uncertainty quantification. The model adapts to diverse subsurface conditions, enabling finer stratification with increasing model complexity. Particularly useful in offshore petroleum settings, the method enhances site characterization, offering a robust and flexible framework that supports more informed and reliable geotechnical design decisions.

Keywords: Soil classification, CPTu data, Stratigraphy, Bayesian approach.

1 Introduction

The accurate characterization of soil stratigraphy is a fundamental procedure in geotechnical engineering, especially in complex offshore projects, where safety, cost optmization and attendance to evironmental regulations are very important. Traditionally, soil classification based on the piezocone penetration test (CPTu) data relies on deterministic methods, such as the Robertson [1] chart, which do not quantify the uncertainties inherent in measurements and transition zones between layers. Probabilistic approaches have gained increasing prominence, to overcome these limitations.

In this matter, the present work compares advanced classification techniques. Specifically, we evaluate a robust Bayesian framework for soil classification, assessing its performance against two widely recognized methods: the conventional Robertson chart, enhanced with adjustments from a more recent approach, and a classification system based on fuzzy logic. By applying these three distinct approaches to real CPTu data from the Campos basin, study highlights the similarities approaches and the assessing of uncertainties. The main objective is to demonstrate how the Bayesian methodology not only provides soil classification, but also formalizes the stratigraphic delineation process by quantitatively evaluating the most probable number of layers and their boundaries, thus offering a more rigorous and less subjective alternative to traditional interpretation.

2 Methodology

The methodology integrates three distinct approaches: the well-known empirical chart by Robertson [1], incorporating Sodré [2] contributions to numerically determine the chart boundaries; a fuzzy logic-based classification technique proposed by Abu-Farsakh et al. [3] and a Bayesian probabilistic approach proposed by Wang et al. [4].

Analysis of classification data consists on comparison between all approaches and evaluation of the probabilistic technique in identifying stratigraphy. Figure 1 below presents a flowchart explaining the methodology described.

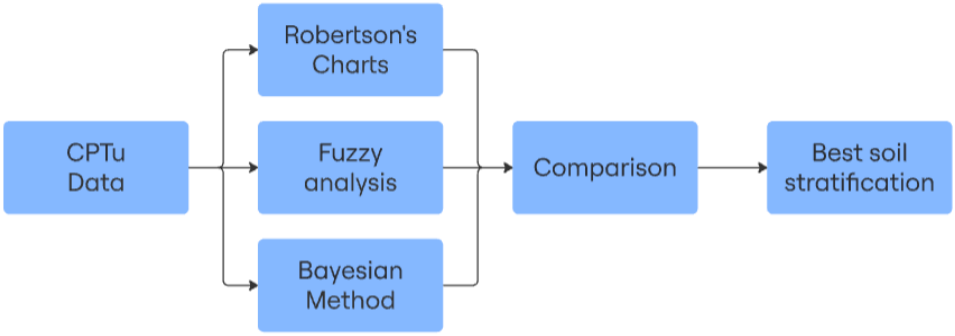


Figure 1. Flowchart of the applied methodology.

The following section summarizes the three classifications techniques.

2.1 Sodre's application of Robertson's Classification Chart

The first reference approach is the soil classification method proposed by Robertson [1]. This empirical method categorize the soil into nine Soil Behavior Types (SBT). The application requires the calculation of the two input parameters of the chart: the normalized tip resistance (Q_{tn}) and the normalized friction ratio (FR). The friction ratio relates the measured sleeve friction (fs) to the effective tip resistance ($q_t - \sigma_{v0}$). In turn, the normalized tip resistance Q_{tn} adjusts the effective tip resistance for the in-situ vertical effective stress (σ'_{v0}), using atmospheric pressure (Pa) as a reference.

This chart serves as a benchmark for calibrating and comparing new methodologies. The computational application of the Robertson chart, which is inherently visual, with numerical indices is used to allow translating the classification zone boundaries into mathematical equations. Previous indices, based on concentric circles proposed by Robertson and Wride [5], showed significant discrepancies compared to the original curved boundaries. In other aspect, Sodré [2] proposed a different boundary modeling defined in the Robertson chart are modeled using a numerical approximation for calculating the boundaries used in the Soil Behavior Index (Ics). According to the formulation proposed by Sodré [2], the calculation of Ics involved the following expression:

$$I_{cs} = e^{\frac{\ln\left(\sqrt{f^2+q^2}\right) - 0.0884\tan^{-1}\left(\frac{f}{q}\right)}{1-0.075\tan^{-1}\left(\frac{f}{q}\right)}} \tag{1}$$

Where:

$$f = 1.5 + 1.5\log F_r \tag{2}$$

$$q = 4 - \log Q_{tn}. \tag{3}$$

$$n = 0.3316 I_{cs} + 0.05\left(\frac{\sigma'_{vo}}{P_a}\right) - 0.2319. \tag{4}$$

The stress normalization exponent (n) is a crucial factor that reflects soil behavior and can be estimated through empirical correlations that depend on a soil behavior index. An iterative process is performed between I_{cs}

and n until n converges; afterward, the I_{cs} value is calculated using the final n value obtained. The table below presents the soil classification by type, serving as a reference to assist in interpreting the stratigraphic profiles generated by the Sodré and Wang methods.

Table 1. Soil classification by type

Type	Soil description
1	Sensitive, fine-grained
2	Organic soils (peats)
3	Clays (clay to silty clay)
4	Silt mixtures (clayey silt to silty clay)
5	Sand mixtures (silty sand to sandy silt)
6	Sands (clean sand to silty sand)
7	Gravelly sand to sand
8	Very stiff sand to clayey sand
9	Very stiff, fine-grained

2.2 Fuzzy Logic-Based Technique

The second reference approach employs fuzzy logic to classify soil profiles, based on the work of Abu-Farsakh et al. [3]. Unlike the rigid boundaries of the Robertson chart, the fuzzy technique handles uncertainty and the gradual transitions between soil types. Instead of a binary classification, the method assigns "degrees of membership" (membership functions) that indicate the likelihood of soil behavior, defining categories such as Highly Probable Clayey Soil (HPC), Highly Probable Mixed Soil (HPM), and Highly Probable Sandy Soil (HPS). This approach reflects more realistically the transition zones that frequently occur in natural deposits. This approach results a depth-wise profile of the accumulated probabilities of clay, mixture and sand. Main downside of this approach is its lower amount of possible classifications.

2.3 Bayesian Classification According to Wang (2013)

This approach explicitly models the uncertainty inherent in the Robertson classification chart. Given a set of CPTu data, the probability of a soil belonging to one of the nine SBT types is formulated using the Principle of Maximum Entropy, resulting in a joint Gaussian Probability Density Function (PDF) distributed over the Robertson chart. The proposed Bayesian method is carried out in two sequential steps:

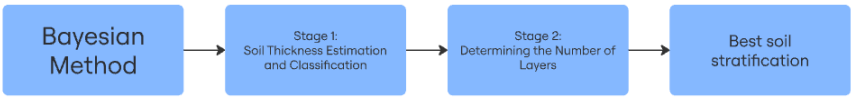


Figure 2. Flowchart of the Bayesian methodology.

This approach is structured in two main steps: Bayesian System Identification, which estimates layer thicknesses and soil classifications for a fixed number of layers, and Class Model Selection, which determines the optimal number of layers by balancing model complexity and data fit. Together, these steps provide a robust probabilistic framework for objectively identifying soil stratigraphy, quantifying uncertainty and noise inherent in traditional methods.

Remark 1: Step 1: Bayesian System Identification (Layer Thickness Estimation and Soil Classification). The first step performs supposing a predetermined k number of layers exists in the data. For each k amount of layers, the technique estimate simultaneously, which includes the corresponding soil classification and the thicknesses of each layer. The optimization process seeks the parameter values that maximize the posterior probability density

function obtained by combining prior information with the likelihood of the observed CPTu data by minimizing an objective function defined as the negative logarithm of that probability. This optimization process identifies the boundary positions that best explain the data, filtering out noise and small-scale variations in CPTu measurements that do not represent actual changes in stratigraphy. As a result, for a given number k of layers, this step provides a complete stratigraphic profile with the most probable boundaries and soil classification for each layer, along with a quantification of the associated uncertainty.

Remark 2: Step 2: Class Model Selection (Determination of the Number of Layers). This step determines which model (i.e. which amount of layers) best represents the actual soil stratigraphy. This selection is made by calculating the Bayesian evidence for each class model k, the evidence is a measure of how well the model, as a whole, predicts the observed data, the class model with the highest evidence value is then selected as the most probable, as it offers the best balance between data fit and model simplicity, a model with more layers (more complex) may fit the data better, but the Bayesian framework naturally penalizes excessive complexity. The final result is the most probable number of layers and its corresponding detailed stratigraphic profile obtained in Step 1.

3 Results

The following charts illustrate the identified soil stratigraphy for the reference CPTu borehole, designated hole A, located in Campos basin.

A visual analysis of the profiles allows for a direct comparison between the Robertson chart approach (using the I_{cs} index), the fuzzy logic method and the Bayesian approach. One of the main findings observed across all methods was the consistent identification of fine-grained soil layers. In particular, notable strata include those classified as Soil Behavior Type (SBT) 3, corresponding to ”Clays (clay to silty clay)” and SBT 4, corresponding to ”Silty mixtures (clayey silt to silty clay)”, according to the Robertson classification.

The generated profiles highlight how each method interprets the transitions between these layers, which will be the focus of a detailed comparative analysis at the conclusion of this study.

The distribution of these points across the Robertson chart further illustrates their classification within the defined soil behavior zones, is illustrated in Figure 3.

The application of the Robertson method, through the I_{cs} index (Figure 4), results in a point-by-point classification which, for borehole A, predominantly identifies the soil as Type 3 (Clays to silty clays) up to approximately 20 meters depth, with a transition to Type 4 (Silt mixtures) in the last 10 meters of the profile. Although accurate at each point, the results exhibit small-scale noise that requires subjective interpretation to group points and define boundaries.

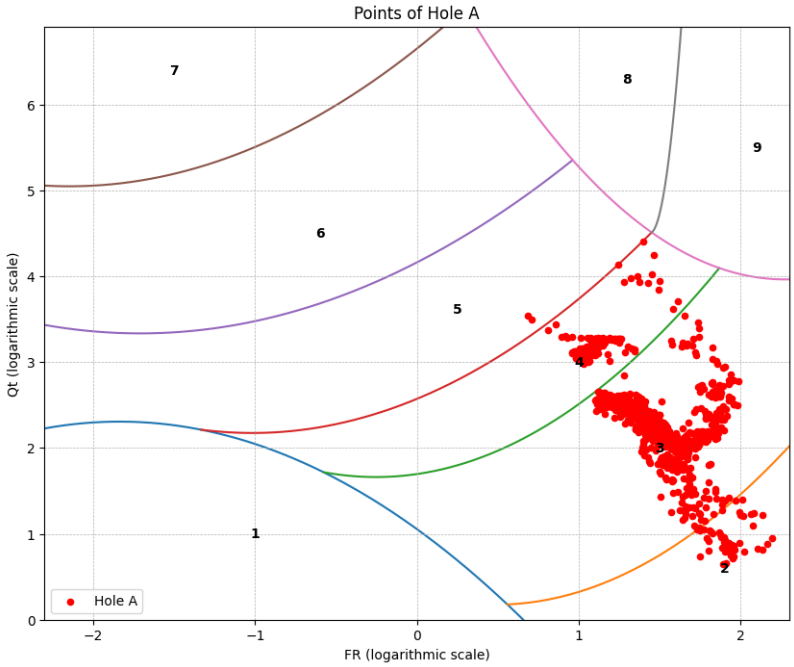


Figure 3. Points from hole A distributed on the Robertson chart.

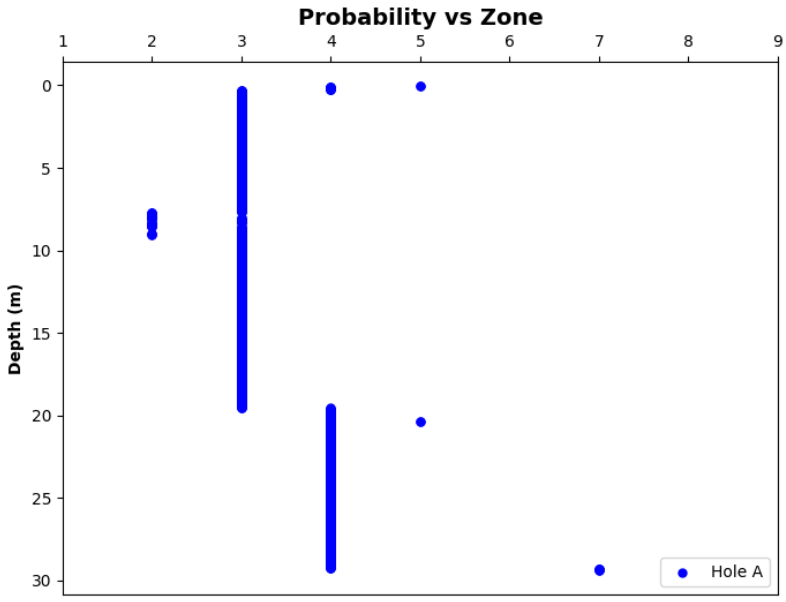


Figure 4. Classification by Sodré Method for hole A.

The fuzzy logic approach (Figure 5), in turn, effectively captures the gradual nature of transitions, displaying a continuous profile of accumulated probabilities. For hole A, a high probability of the soil being classified as "Clay" is observed in the first 16 meters, with a gradual increase in the probability of "Mixture" and "Sand" in the last 14 meters. However, this technique does not objectively define the layer boundaries, requiring interpretation for stratigraphic delimitation.

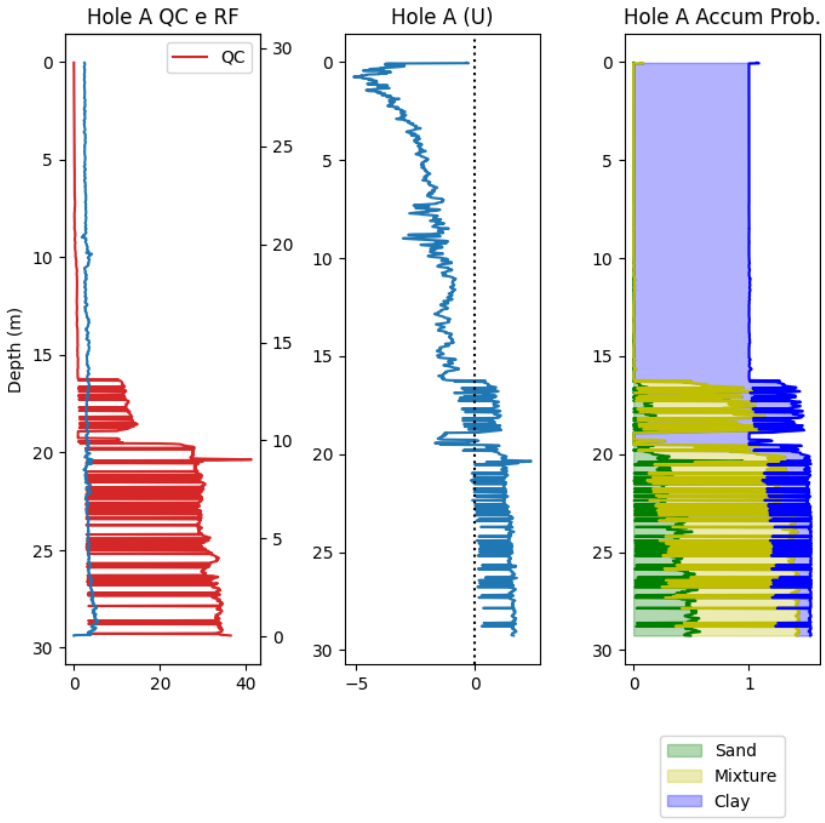


Figure 5. Classification by Abu-Farsakh Method for hole A.

In contrast, the Bayesian approach (Figure 6) provides a direct solution for stratigraphy. For the analyzed hole A, the model selection process indicated that a two-layer profile is the most probable. The model identified an upper layer extending to 21.87 meters, classified as Type 3 (Clays to silty clays), and a lower layer of 7.39 meters classified as Type 4 (Silt mixtures). The probability-by-depth graph shows that, although the lower layer is classified as Type 4, there is a significant competing probability (approximately 0.8) that the soil is of Type 3. This explicitly quantifies the uncertainty, defining not only the most probable classification but also the plausibility of alternatives.

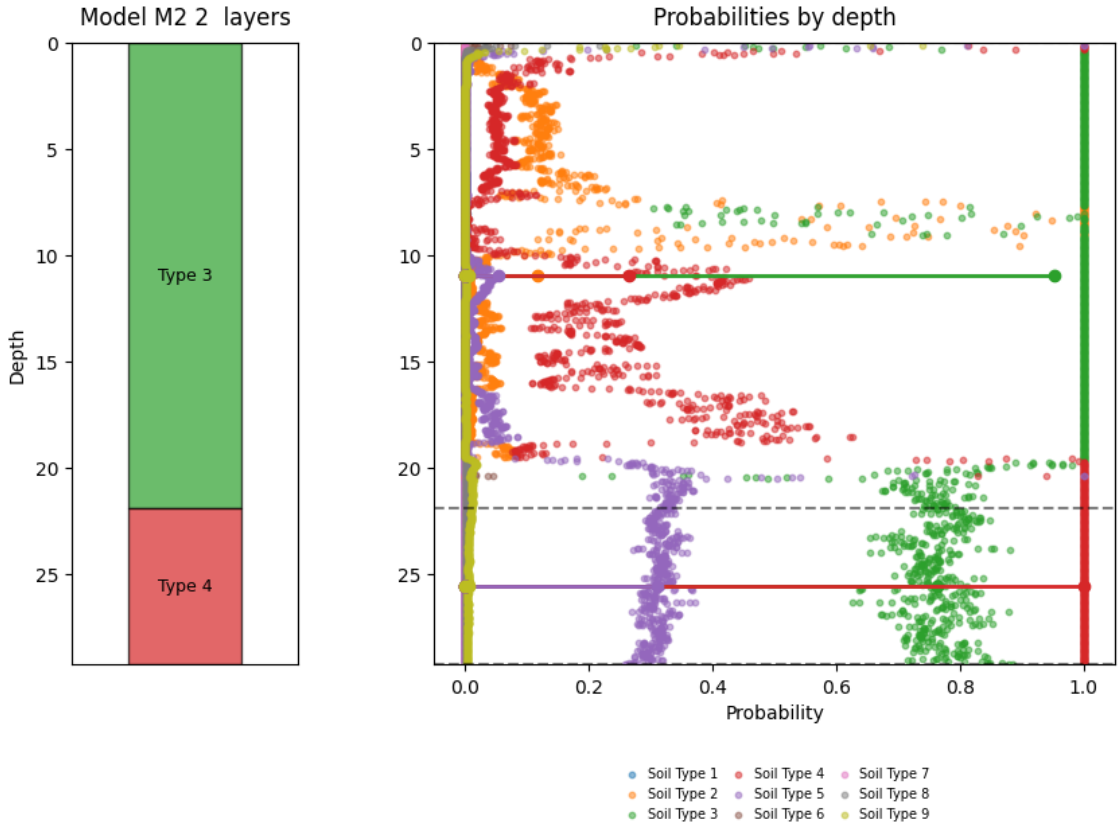


Figure 6. Classification by Wang Method for hole A.

Table 3 summarizes the results of this step, presenting the most likely thicknesses for each layer (h*) within the candidate models, ranging from one (M1) to six (M6) layers.

Table 2. Thickness of each model class

Model class	**h1***	**h2***	**h3***	**h4***	**h5***	**h6***
M1	29.26	–	–	–	–	–
M2	21.87	7.39	–	–	–	–
M3	9.35	10.36	9.55	–	–	–
M4	10.95	8.74	4.63	4.93	–	–
M5	4.79	8.07	6.84	4.76	4.80	–
M6	4.30	7.78	2.75	4.88	5.69	3.86

To consolidate the comparative analysis, the following table summarizes the stratigraphic profiles defined by each of the three approaches. The results are compared with the geotechnical data obtained from a report ceded by the partner oil and gas operator.

Table 3. Approach comparison by layer and soil type

Approach	h1	h2	Soil Type h1	Soil Type h2
Sodré	20.00	10.00	Clay	Silt Mixture
Abu-Farsakh	16.00	14.00	Clay	Silt Mixture
Wang	21.87	7.39	Clay	Silt Mixture
Reference data	22.00	7.00	Clay	Silt Mixture

The integrated analysis of the results reveals a clear overall similarity among the approaches, though each methodology has its own particularities. While the Robertson and fuzzy logic methods provide classifications with some noise, the Bayesian approach advances by delivering an objective and complete stratigraphic model. This ability to translate raw data into a well-defined layered profile, along with the explicit quantification of uncertainty, lays the foundation for the final conclusions regarding the effectiveness of this method.

4 Conclusions

This study demonstrated the robustness and advantages of a Bayesian approach for soil profile characterization by comparing it with well-established methods. Although all methods consistently identified the predominance of fine-grained soils Type (SBT) 3, corresponding to "Clays (clay to silty clay)," and SBT 4, corresponding to "Silty mixtures (clayey silt to silty clay)" in hole A, the Bayesian approach has been shown to reduce uncertainties in identifying stratigraphy. Unlike classical techniques, which produce noisy or probabilistic profiles that require significant engineering judgment to define layer boundaries.

The Bayesian method autonomously and quantitatively determines the most probable number of layers and their respective thicknesses by formalizing a process that is often subjective, filtering data noise, and identifying the most statistically relevant boundaries. Its ability to directly incorporate uncertainties into the analysis enhances repeatability and depth of interpretation, making it an especially valuable tool for critical applications such as offshore projects, where accurate subsurface characterization is essential for safe and optimized design decisions.

Acknowledgements. The author thanks the Laboratory of Scientific Computing and Visualization (LCCV) and PETROBRAS for providing the infrastructure and resources that made this work possible. I also thank Dr. Christiano Varady, for the insightful mentorship and support.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] P. K. Robertson. Soil classification using the cone penetration test. *Canadian Geotechnical Journal*, vol. 27, n. 1, pp. 151–158, 1990.

[2] D. J. R. Sodré. Cone penetration index for soil behaviour type prediction. *Scientific Reports*, vol. 12, n. 1, 2022.

[3] M. Y. Abu-Farsakh, Z. Zhang, M. Tumay, and M. Morvant. Computerized cone penetration test for soil classification: Development of ms-windows software. *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2053, pp. 47–64, 2008.

[4] Y. Wang, K. Huang, and Z. Cao. Probabilistic identification of underground soil stratification using cone penetration tests. *Canadian Geotechnical Journal*, vol. 50, pp. 766–776, 2013.

[5] P. K. Robertson and C. E. Wride. Evaluating cyclic liquefaction potential using the cone penetration test. *Canadian Geotechnical Journal*, vol. 35, n. 3, pp. 442–459, 1998.