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

Facies Classification and Porosity–Permeability Modeling Using Self-Organizing Maps in Hybrid Carbonate Rocks

Gabriel S. S. Ribeiro, Silvia L. B. Bermudez, Maira C. O. Lima

LRAP⁺, Civil Engineering Program, COPPE, Federal University of Rio de Janeiro
Moniz Aragão Avenue, 360 - BL4 Cidade Universitária, 21941594, Rio de Janeiro, Brazil
gabriel.ribeiro@petroleo.ufrj.br,silviabermudez@petroleo.ufrj.br,maira.lima@petroleo.ufrj.br

Abstract. Detailed studies on the hybrid carbonate rocks of the Morro do Chaves Formation are essential, as these outcrops serve as analogues to the Brazilian pre-salt carbonate reservoirs, allowing for the validation and application of advanced petrophysical characterization methods. However, the heterogeneity resulting from the complex mixture of carbonate and siliciclastic minerals significantly challenges facies classification and the prediction of critical properties such as porosity and permeability. Unsupervised techniques based on Self-Organizing Maps (SOM) have gained prominence in petrophysics for their ability to project complex multidimensional data onto two-dimensional maps, preserving original topology and revealing clusters and nonlinear patterns not perceptible through traditional approaches. In this context, four distinct SOM models were developed in the Interactive Petrophysics (IP) software, and the best-performing model was selected. This spherical-topology network, comprising 1,442 nodes, was trained over 50,000 epochs (η_0=0.1). To train the SOM, we employed well logs, 57 core-plug measurements of porosity and permeability, and a lithologic profile based on core descriptions. Classification was validated using contingency tables, with the best model achieving 72.59% accuracy, and was complemented by qualitative analysis through Fit Flag. Subsequently, petrophysical properties were predicted and statistically evaluated using the mean absolute error (MAE), root-mean-square error (RMSE), and uncertainty bands. These bands correspond to a ±1 σ standard deviation, weighted at each node, and are derived from continuous calibration data, and plotted alongside the predicted curves. Finally, regression fits between predicted and observed values yielded correlation coefficients (R^2) of 0.9398 for porosity and 0.8266 for permeability, corroborating the robustness and expanding the applicability of this approach to well-scale petrophysical characterization of hybrid carbonate systems.

Keywords: Self-Organizing Maps (SOM), Hybrid Carbonate Rocks, Facies Classification, Porosity and Permeability Prediction, Petrophysical Upscaling.

1 Introduction

Petrophysical characterization is essential for reservoir modeling, particularly in the complex carbonates of the Brazilian pre-salt. The heterogeneity arising from depositional and diagenetic processes creates dual-porosity systems with variable connectivity and preferential wettability, posing additional challenges to flow behavior prediction [1–2]. The Morro do Chaves Formation (Aptian–Albian), located in the Sergipe-Alagoas Basin, stands out as a representative analogue of such hybrid reservoirs due to its textural and mineralogical similarity. Composed of hybrid lacustrine coquinas — a mixture of carbonate and siliciclastic minerals — and featuring a dual-porosity system with vuggy, moldic, and interparticle pores, this formation serves as a natural laboratory for testing advanced petrophysical characterization methodologies and evaluating their effectiveness in capturing heterogeneities typical of carbonate reservoirs [3–5].

Despite its value as an analogue, characterizing these hybrid rocks remains challenging. The multiscale heterogeneity resulting from the coexistence of carbonate and siliciclastic minerals leads to abrupt variations in facies and pore connectivity, hindering both qualitative interpretation and quantitative prediction [1,4]. This

geological complexity is further compounded by data limitations: although a continuous core description by Favoreto [4] is available at 40 cm intervals, lithogeochemical profiles are lacking and only 57 plugs with routine porosity and permeability measurements are available across 130.6 m of section, complicating the upscaling process. These data gaps demand methods capable of integrating discrete measurements with well log data to provide continuous property estimation. Unsupervised learning techniques — particularly Self-Organizing Maps (SOM) — have gained prominence for projecting multidimensional datasets into intuitive topological representations, enabling electrofacies identification and petrophysical property prediction even under high lithological complexity and limited data availability [6–8]. In this context, the present work applies SOM to classify facies and estimate porosity and permeability in the hybrid carbonate rocks of the Morro do Chaves Formation, integrating well log data with routine core petrophysical measurements to reduce uncertainty and support petrophysical upscaling at the well scale.

2 Methodology

In this study, an integrated dataset was compiled from well logs acquired in the Morro do Chaves Formation, Sergipe–Alagoas Basin. The logs analyzed comprise gamma ray (GR), spontaneous potential (SP), shallow resistivity (RES16N), deep resistivity (RES64N), lateral resistivity (RESLAT), compressional sonic (BHC), and a density curve derived from acoustic velocity by Gardner's equation and subsequently calibrated through Bayesian inversion with a Markov-chain Monte Carlo (MCMC) algorithm, as detailed in Ribeiro et al. [9]. Direct calibration of the predicted petrophysical properties employed 57 discrete laboratory measurements of porosity and permeability obtained from core plugs. In addition, a high-resolution (40 cm) lithological profile, constructed from continuous petrographic core descriptions and published by Favoreto [4], was incorporated into the well layout as the facies reference. The workflow was organized into six main stages, described below:

2.1 Data Pre-processing

The dataset first underwent conditioning to identify and correct missing values and outliers. To ensure that every variable exerted equivalent influence on the predictive model, all geophysical logs were normalised to zero mean and unit variance prior to training.

2.2 SOM Model Training

The Self-Organizing Map (SOM) technique, originally proposed by Kohonen [10], is an unsupervised neural network that projects multidimensional data onto two- or three-dimensional maps while preserving their topological relationships. During training, neurons compete to represent input patterns closest to their weight vectors, iteratively updating those weights and revealing clusters and non-linear relationships among the input variables. In this study, a spherical-topology SOM (Figure 1) was adopted to mitigate edge effects. The network comprised 1,442 nodes and was trained for 50,000 epochs in the SOM module of Interactive Petrophysics (IP), using an initial learning rate $\eta_0 = 0.1$.

BMU selection. For each input vector x, the Euclidean distance to each node j is computed as Eq. (1):

$$D_j = \sum_{k=1}^{m}(x_k - w_{j,k})^2. \quad (1)$$

Where m is the number of input attributes, x_k is the k-th component of the input vector, and $w_{j,k}$ is the k-th component of the weight vector of node j. The BMU is defined as the node with the smallest D_j.

Weight update with neighbourhood. Once the BMU is identified, its weight vector and those of its neighbouring nodes on the spherical lattice are updated as:

$$w_i(t+1) = w_i(t) + \theta_i(t)L_t[x(t) - w_i(t)]. \quad (2)$$

where L_t is the learning rate at iteration t and the Gaussian neighbourhood gain is described as Eq. (3):

$$\theta_i(t) = exp\left(\frac{-dist(i,BMU)^2}{2\sigma_t^2}\right). \quad (3)$$

The dist(i,BMU) is the geodesic distance between node i and the BMU on the spherical lattice, and σ_t is the neighbourhood radius at iteration t.

Learning schedules. Both L_t and σ_t decay exponentially with iteration t according to Eq. (4) and Eq. (5), respectively:

$$L_t = L_0\, exp\left(\frac{-t}{\lambda}\right), \quad (4)$$

$$\sigma_t = \sigma_0\, exp\left(\frac{-t}{\lambda}\right), \quad (5)$$

with $L_0=\eta_0$ and σ_0 the initial radius (initialized in IP as a fraction of the map width), and λ the time constant governing the decay.

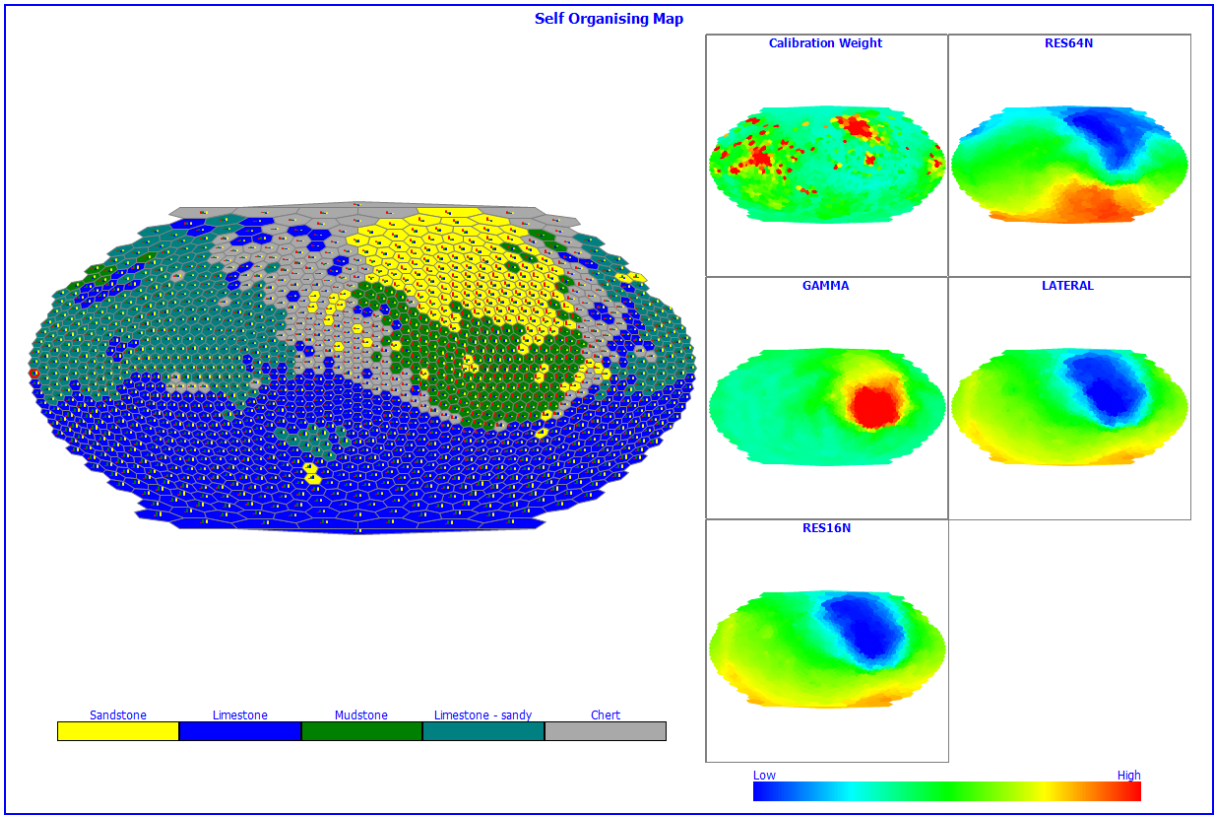


Figure 1 – Spherical geometry of the Self-Organizing Map (SOM) selected for facies classification. On the left, the lattice of 1,442 nodes displays the final distribution of classes: Sandstones (yellow), Rudstones (dark blue), Mudstones (dark green), Grainstones (light green), and Chert (gray). On the right, the minilogos (weight maps) show, from top to bottom, the relative influence of the calibration curve (Calibration Weight) and the input curves RES64N, GAMMA, RESLAT, and RES16N for each neuron, highlighting in a color scale (blue → red) the regions from low to high values for each attribute.

2.3 Calibration for Input Curves

Discrete calibration — facies classification. The continuous lithologic profile (Favoreto [4]) was presented as a calibration curve. For each calibration sample (x_i, c_i), node j receives a distance-based weight (inverse square of D_j); node classes are assigned by maximizing the accumulated weights per class. Each depth is then classified by its BMU with optional percentile display for confidence.

Continuous calibration — porosity and permeability. Plug measurements linked to each node define a calibrated node value y_j (arithmetic mean for porosity, geometric handling for permeability). Predictions at depth

d are obtained by projecting node values from the BMU and its neighbours via the Gaussian kernel above:

$$y(d) = \frac{\sum_j \theta_j(d) y_j}{\sum_j \theta_j(d)}. \quad (6)$$

2.4 Facies Validation

Validation of the SOM-derived facies classes combined quantitative and qualitative assessments. Quantitatively, contingency tables were built by cross-referencing the SOM predictions with the reference classification obtained from Favoreto's detailed petrographic core description [4], allowing the global accuracy of the model to be computed as the primary performance metric. Qualitatively, the Fit Flag method was employed to visually pinpoint depth intervals where the lithological reference curve and the SOM-classified facies exhibited the highest concordance, thereby reinforcing the consistency of the results.

2.5 Prediction and Assessment of Petrophysical Properties

Continuous porosity and permeability logs were estimated entirely within the SOM module of IP software. After training the spherical map, each neuron inherited the core-plug measurements located within its area of influence and was assigned their arithmetic mean (porosity) or geometric mean (permeability). Simple linear regressions were then applied between these neuronal mean values and the 57 laboratory measurements to remove residual bias and extrapolate the relationship to depths lacking direct data. The resulting coefficients (a and b) were applied to the SOM weight matrix to generate continuous curves along the well, obtained by inverse-distance weighting to the Best Matching Unit at each depth. Uncertainty was propagated by retaining, for every sample, an envelope of ±1 standard deviation (±1 σ) around the calibrated value. Model performance was summarized by comparing the predicted curves with plug data, using the Mean Absolute Error (MAE) and the Root-Mean-Square Error (RMSE) to quantify average accuracy and dispersion, respectively.

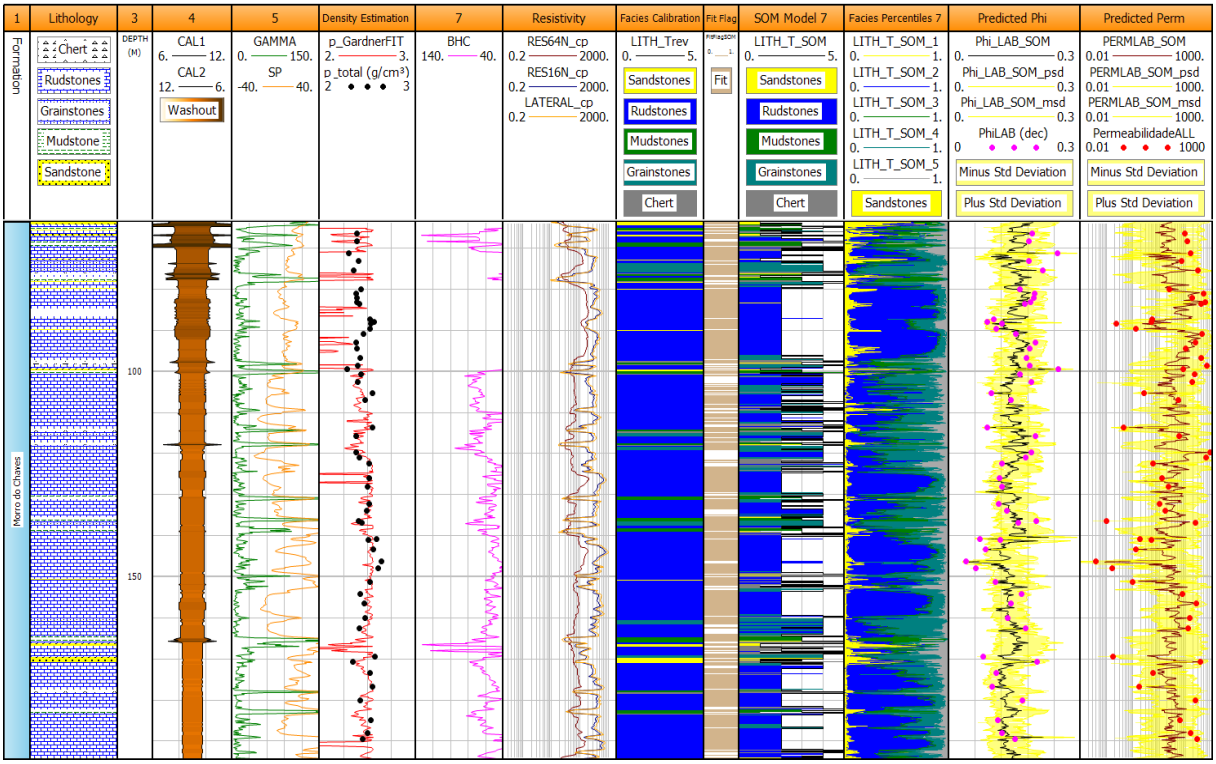
2.6 Integration and Visualization of Results

The results obtained were integrated and presented in comparative logs, containing calibrated versus classified facies curves, and laboratory-measured porosity and permeability curves versus SOM model predictions. Heatmaps were also used to provide a quick and intuitive visualization of the distribution of lithological classes and predicted petrophysical properties throughout the analyzed interval, assisting in the qualitative and integrated interpretation of the results.

3 Results and Discussions

Figure 2 presents the input curves (tracks 2 to 8) and the results obtained from facies classification using the spherical SOM model (track SOM Model7) applied to the well logs of the SMC-02-AL well in the Morro do Chaves Formation. The Facies Percentiles track displays one curve for each facies, indicating their relative weighting at each depth level and enabling the assessment of confidence in facies selection. It is observed that the facies with the highest weighting correspond to the classifications produced by the model. The Fit Flag curve, showing high agreement, reinforces the accuracy of the SOM model in recognizing facies. The initial qualitative analysis based on this track shows that the curve remains predominantly in the beige ("match") class throughout most of the evaluated section, indicating excellent agreement between the SOM-predicted facies and those defined by the calibrated lithologic profile (Facies Calibration track). Small intervals with occasional deviations are observed, represented by the absence of color ("mismatch"), but these do not significantly affect the overall classification quality. This strong qualitative agreement is further supported quantitatively by the contingency table associated with SOM Model 7, which shows an overall accuracy of 72.59%, considered excellent for hybrid facies sets with strong overlap in petrophysical properties. This remarkable value demonstrates the robustness and effectiveness of the SOM method in the automated identification of the various facies present in this complex lithologic context, representing a substantial improvement over traditional methods previously employed.

A detailed analysis of the contingency table results highlights the SOM model's ability to accurately classify Mudstone facies, achieving an accuracy of 90.8%. This high performance is primarily due to the highly distinctive and characteristic log signatures associated with Mudstones compared to other facies identified in the well. These rocks are characterized by a very fine-grained matrix, predominantly composed of clay and silt particles. Such fine particles produce a homogeneous and compact texture, often associated with low permeability and reduced fluid storage capacity. Due to this significant clay content, Mudstones typically exhibit high responses in the gamma ray (GR) log, resulting from the abundant presence of clay minerals containing radioactive elements in their structure. In addition, these rocks present reduced electrical resistivity, caused by the high water retention capacity of clays, as well as equally low porosity, consistent with their compact and poorly porous matrix. This unique petrophysical pattern greatly facilitates the accurate differentiation and identification of these facies by the applied SOM model, as confirmed by both the quantitative and qualitative results obtained in this study.



Contingency Table

Total Correct : 72.59%

Calibration Input	1	2	3	4	5	Total
1	86.1	3.1	6.2	2.6	2.1	100.
2	1.5	66.	1.3	23.5	7.7	100.
3	6.3	0.	90.8	0.7	2.1	100.
4	0.	1.5	0.	98.5	0.	100.
5	0.	0.	13.6	4.5	81.8	100.

SOM Result

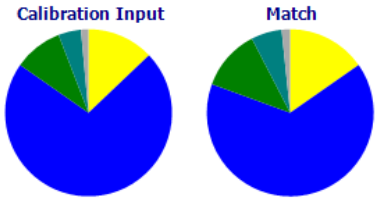


Figure 2 – Integrated layout of facies classification and petrophysical property prediction along the SMC-02-AL well in the Morro do Chaves Formation. From left to right: (i) lithologic description based on core observations (track 2); (ii) input curves used in the modeling (gamma ray, spontaneous potential, density estimated via Gardner's equation fit, compressional sonic, and resistivities RES64N, RES16N, and Lateral – tracks 5 to 8);

(iii) facies calibration curve (track 9); (iv) SOM classification results (track 11), indicating the facies assigned along the well; (v) facies percentile curve for evaluating classification confidence (track 12); and (vi) porosity and permeability predictions with uncertainty intervals (±1σ), highlighting agreement with laboratory data (tracks 13 and 14). The lower panel displays the contingency table comparing calibrated and classified facies, showing an overall accuracy of 72.59%, and pie charts illustrating (i) the percentage distribution of classes in the calibration set and (ii) the proportion of correct matches obtained by the model.

Another important point is the well-known difficulty in distinguishing between different carbonate facies, especially Rudstones and Grainstones. Texturally, Rudstones are characterized by coarse carbonate fragments supported by larger grains, whereas Grainstones consist of finer, well-sorted, and generally well-rounded grains, with no significant muddy matrix [11–12]. Despite these textural differences, the two facies have similar mineral compositions, which results in very similar response patterns in the well logs, making precise separation challenging. This difficulty is clearly reflected in the contingency table, which shows that approximately 98.5% of Grainstones were correctly classified by the SOM model, whereas only 66% of Rudstones were correctly identified. Moreover, about 23.5% of Rudstones were misclassified as Grainstones, demonstrating the method's limitation in reliably distinguishing between these two facies. Nevertheless, this is a well-known challenge, widely reported in the petrophysical literature [3,13]. If a single generalized class such as "carbonates" were considered, instead of the specific textural subdivision into Grainstones and Rudstones, the overall classification accuracy would certainly be higher, due to the reduced ambiguity in petrophysical responses.

After the successful validation of the facies classification, the same spherical SOM model was applied in the subsequent stage for the quantitative prediction of petrophysical properties—porosity and permeability. All previously used hyperparameters were maintained, with only the target curves for calibration being replaced. The network was directly calibrated with laboratory measurements from 57 core plugs along the studied interval, ensuring a solid basis for validating and adjusting the model's predictive capability. The results of this stage are shown in Figure 2, specifically in the Predicted Phi and Predicted Perm tracks. In these tracks, the continuous curves represent the estimates generated by the SOM model, while the dots correspond to the laboratory data used as reference. Additionally, yellow bands surrounding the predicted curves define the uncertainty interval corresponding to ±1σ (standard deviation), providing a robust assessment of the variability of the estimated properties throughout the depth profile. To objectively quantify the model's performance, linear regressions were performed between the laboratory measurements and the SOM predictions, yielding coefficients of determination (R^2) of 0.9398 for porosity and 0.8266 for permeability. These values clearly demonstrate the accuracy and consistency of the adopted method in capturing petrophysical variations from well logs, even in heterogeneous lithologic contexts such as the hybrid coquinas of the Morro do Chaves Formation.

Table1 – Descriptive statistics and prediction errors for porosity and permeability (Lab vs. SOM).

Properties	Minimum	Mean	Median	Mode	Maximum.	SD	MAE	RMSE
Phi_LAB	0.0400	0.1518	0.1630	0.0870	0.2490	0.0484	N/A	N/A
Phi_SOM	0.0660	0.1487	0.1457	0.1421	0.2398	0.0404	0.0100	0.0155
PermeabilidadeALL	0.04	160.72	70.77	0.04	853.53	213.15	N/A	N/A
Perm_SOM	0.12	102.56	30.39	0.12	723.85	151.03	61.79	126.71

The SOM errors indicate high performance for porosity and adequate performance for permeability (Table 1). For Phi, MAE = 0.010 and RMSE = 0.0155 demonstrate precise predictions with limited tails; the SD of Phi_SOM (0.0404) is slightly lower than Phi_LAB (0.0484), suggesting mild smoothing without bias. For permeability, MAE = 61.79 mD and RMSE = 126.71 mD reflect heterogeneity and the presence of high-mD outliers (RMSE > MAE); even so, the SOM reduces dispersion (SD = 151.03 mD vs. 213.15 mD in the laboratory) while preserving the central tendency. These indicators support the applicability of the SOM to well-scale upscaling, with clear gains for porosity and consistent performance for permeability in heterogeneous coquinas.

4 Conclusions

The application of Self-Organizing Maps (SOM) proved effective both in the accurate classification of facies and in the quantitative prediction of essential petrophysical properties—porosity and permeability—in hybrid

carbonate rocks of the Morro do Chaves Formation, characterized by high lithological heterogeneity. The careful use of a spherical-topology SOM model, calibrated with well logs and laboratory data from 57 core plugs, achieved an overall accuracy of 72.59% in facies classification relative to the reference lithologic profile, with notable recognition rates of 98.5% for Grainstones and 90.8% for Mudstones. This outstanding performance, combined with the careful selection of hyperparameters—determined through distortion analysis—and the reduced computational time, demonstrates the robustness and operational efficiency of the method. In the petrophysical property prediction stage, the model achieved high coefficients of determination (R^2 of 0.9398 for porosity and 0.8266 for permeability) and provided continuous estimates accompanied by explicit uncertainty bands ($\pm 1\sigma$), substantially increasing the reliability of the results. Therefore, the proposed approach stands out as a promising and practical tool to support petrophysical upscaling at the well scale, significantly contributing to the reduction of uncertainties and enhancing the technical characterization of complex carbonate reservoirs.

Acknowledgements. This research was carried out in association with the ongoing R&D project registered as ANP nº 23020-1, "Análise Experimental da Recuperação Avançada de Petróleo em Reservatórios Carbonáticos do Pré-Sal do Brasil Através de Injeção Alternada de CO2 e Água - Fase-II - Condições de Reservatório" (UFRJ/Shell Brasil/ANP), sponsored by Shell Brasil Petróleo Ltda under the ANP R&D levy as "Compromisso de Investimentos com Pesquisa e Desenvolvimento". The authors would like to acknowledge the support of Programa de Recursos Humanos (PRH-9.1) from Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – PRH-ANP, financed with resources from qualified oil and gas companies by R, D&I Clause ANP N. 50/2015, and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). The authors also acknowledge the partial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (CNPq) - through the research productivity grant 310291/2022-4.

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] Bruhn, C. H. L., Gomes, J. A. T., Lucchese Jr., C., & Johann, P. R. S. (2003). Campos Basin: reservoir characterization and management — Historical overview and future challenges. Offshore Technology Conference, paper OTC-15220-MS. https://doi.org/10.4043/15220-MS
[2] Gomes, J. A. T., et al. (2020). Carbonate reservoirs from Santos Basin pre-salt, Brazil: Understanding reservoir heterogeneities to improve oil recovery. AAPG Bulletin, 104(1), 89-120.
[3] Lucia, F. J. (2007). Carbonate Reservoir Characterization: an integrated approach (2ª ed.). Springer-Verlag, 336 p. DOI: 10.1007/978-3-540-72742-2
[4] Favoreto, J. (2020). Depositional controls on lacustrine coquinas from an early Cretaceous rift lake: Morro do Chaves Formation, Northeast Brazil. Tese de Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
[5] Santo, M.C.O.L. (2020). The effects of pore connectivity on the permeability of coquinas from the Morro do Chaves Formation. Tese de Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
[6] Kohonen, T. (2001). Self-Organizing Maps. Springer Series in Information Sciences. DOI: 10.1007/978-3-642-56927-2
[7] Roden, R., Smith, T., & Sacrey, D. (2015). Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps. Interpretation, 3(4), SAE59–SAE83. https://doi.org/10.1190/INT-2015-0037.1
[8] Martins, M. L. J. (2022). Electrofacies classification using unsupervised method and comparison with formation evaluation through NMR logs and facies association from Barra Velha Formation, Santos Basin. IX Simpósio Brasileiro de Geofísica, Curitiba, 04-06 Outubro 2022.
[9] Ribeiro, G. S. S., Bermudez, S. L. B., Lima, M. C. O., Alves, J. L. D., & Couto, P. (2025). Modeling and inversion of petrophysical properties using Monte Carlo method in coquinas of the Morro do Chaves Formation, Sergipe-Alagoas Basin. Brazilian Journal of Petroleum and Gas, 19(1), 1–15. https://doi.org/10.5419/bjpg2025-0011
[10] Kohonen, T. (1981b). Construction of similarity diagrams for phonemes by a self-organizing algorithm. Technical Report TKK-F-A463, Helsinki University of Technology, Espoo, Finland.
[11] Dunham, R. J. (1962). Classification of carbonate rocks according to depositional texture. In W. E. Ham (Ed.), Classification of Carbonate Rocks (AAPG Memoir 1, pp. 108–121). Tulsa, OK: American Association of Petroleum Geologists.
[12] Embry, A. F., & Klovan, J. E. (1971). A late Devonian reef tract on northeastern Banks Island, NWT. Bulletin of Canadian Petroleum Geology, 19(4), 730–781. DOI: 10.35767/gscpgbull.19.4.730
[13] Gomes, J. S., Távora, E. & Carvalho, M. S. S. (2008). Petrofísica de Rochas Carbonáticas. Rio de Janeiro: Interciência, Petrobras, 280 p.