A Conditional GAN Framework for Joint Generative Modeling of Geological Facies and Acoustic Impedance

Israel Efraim De Oliveira
Department of Informatics and Statistics
Universidade Federal de Santa Catarina
Brazil
israel.oliveira@posgrad.ufsc.br

Rafael De Santiago
Department of Informatics and Statistics
Universidade Federal de Santa Catarina
Brazil
r.santiago@ufsc.br

Yineth Viviana Camacho-De Angulo
Universidad Tecnológica de Bolívar
Brazil
camachoy@utb.edu.co

Tiago Mazzutti
Department of Informatics
Instituto Federal Catarinense
Brazil
tiago.mazzutti@ifc.edu.br

Bruno Barbosa Rodrigues
PETROBRAS
Brazil
bbrodrigues@petrobras.com.br

Mauro Roisenberg
Department of Informatics and Statistics
Universidade Federal de Santa Catarina
Brazil
mauro.roisenberg@ufsc.br

***Abstract*—Facies modeling and the generation of associated rock properties are fundamental tasks in subsurface reservoir characterization, particularly in settings where direct observations are sparse and training data are limited. Traditional geostatistical workflows typically model categorical facies and continuous petrophysical properties in separate steps, while recent deep generative approaches often require large datasets and focus on single-variable synthesis. In this work, we propose a conditional generative adversarial network (cGAN) framework for the simultaneous generation of geological facies and acoustic impedance, explicitly conditioned on hard data from wells. The methodology builds upon a multi-scale SinGAN architecture, employing a hierarchical sequence of generators and discriminators that progressively refine joint facies–property realizations from coarse to fine resolutions. By extending the generator to a multi-channel output, the proposed approach captures the spatial dependencies between discrete facies distributions and continuous rock properties within an unified model. The method is validated using a limited set of two-dimensional realizations derived from the Stanford Earth Science dataset. Quantitative evaluation based on distributional metrics and conditioning accuracy demonstrates that the model effectively honors well data while producing geologically plausible and diverse joint realizations. The results indicate that the proposed framework offers a data-efficient and flexible solution for integrated facies and rock property modeling under constrained data availability.**

***Index Terms*—Acoustic impedance, Conditional data, Conditional generative adversarial network, Geostatistical simulations, Reservoir characterization, Seismic inversion**

I. Introduction

The accurate characterization of subsurface rock properties is fundamental in natural resource exploration (e.g., hydrocarbons, groundwater) and environmental management (e.g., geological CO_2 storage) [1]. However, obtaining direct subsurface data, such as well logs and cores, is an expensive process that results in sparse and spatially limited information [2].

This data scarcity makes representing the geological complexity of the reservoir a challenge, hindering large-scale analyses of properties and their distribution [3]. Therefore, the generation of realistic geological models that honor existing data and represent the inherent uncertainty is a central problem in geoscience [4].

Geostatistical methods are widely used to fill the gaps between sparse data points [5]. Traditional approaches based on two-point statistics, like the variogram, are efficient but often fail to capture complex and curvilinear geological patterns [6]. To overcome this limitation, Multiple-Point Statistics (MPS) emerged as a robust alternative. Unlike variogram-based methods, MPS extracts complex, high-order spatial patterns directly from a training image (TI), which is a conceptual model representing prior geological knowledge about the study area [7]. Although powerful, the application of MPS can be computationally expensive, and its effectiveness depends on the availability of a representative TI [8].

Recently, Generative Adversarial Networks (GANs), a type of Deep Generative Model (DGM), have emerged as a promising alternative for facies modeling, offering advantages over traditional geostatistical methods [9]. GANs consist of two neural networks, a generator and a discriminator, that compete against each other [10]. The generator learns to create synthetic realizations (e.g., facies maps) that are indistinguishable from training images, while the discriminator learns to differentiate between real and generated realizations. Conditional GANs (cGANs) are especially valuable in geological applications because they can generate data constrained by prior information, such as well logs. This ensures the resulting models are consistent with direct observations [11].

This project is sponsored by Petrobras under CT number 0050.0122221.22.9

Advanced cGAN implementations use a hierarchical architecture that operates at progressively higher resolutions [12], [13]. The process begins with a low-resolution generator conditioned on well data, refining the output in subsequent stages by adding finer details. Each generator is paired with a discriminator, ensuring consistency with conditioning data at all scales. This progressive strategy captures both large-scale structures and small-scale heterogeneities [13].

While existing works rely on large datasets or multi-stage workflows, we propose an application-driven extension of a SinGAN-based architecture to jointly generate geological facies and acoustic impedance (IP). Specifically, this work adapts a multi-scale conditional generative adversarial network (cGAN) to a **unified, multi-channel architecture**, bypassing sequential modeling workflows. Our framework operates efficiently under limited data by progressively refining joint facies-property realizations while injecting well log conditioning at every scale. Through quantitative evaluation—utilizing distributional metrics and conditioning accuracy—we demonstrate that this adapted multi-channel approach yields geologically plausible joint realizations and honors hard data, offering a **practical solution for data-constrained reservoir characterization**.

The remainder of this paper is organized as follows. Section 2 reviews related work on geostatistical modeling and deep generative approaches for facies and petrophysical property simulation. Section 3 presents the source data, the proposed cGAN architecture, and the conditioning strategy adopted in this study. Section 4 reports the experimental results and quantitative evaluation of the generated realizations. Finally, Section 5 concludes the paper and discusses potential extensions of the proposed framework. The code used to implement the proposed methodology is publicly available on GitHub[1].

II. Related Works

Recent advances in generative modeling have significantly expanded the range of tools available for subsurface characterization, particularly for the simulation of geological facies and associated petrophysical properties. Existing approaches can be broadly categorized along two main axes: (i) methods that jointly model multiple subsurface variables to preserve cross-property consistency, and (ii) methods specifically designed to operate under limited training data and strict conditioning constraints. While substantial progress has been made along each of these directions independently, combining joint multivariate generation with data-efficient, conditioning-aware architectures remains an open challenge. The following works review representative contributions across these categories and motivate the need for an integrated solution.

[9] integrate a GAN with the Spatially Adaptive Normalization (SPADE) technique. SPADE allows local probability maps to directly condition the facies distribution during generation. This approach is combined with co-DSS to model rock properties conditioned on facies, and both models (facies and continuous properties) are iteratively updated to minimize the mismatch with seismic data.

To achieve this, the method utilizes a two-step approach. First, facies realizations are generated using a GAN combined with SPADE to capture complex geological patterns. Subsequently, for each facies realization, multiple equiprobable acoustic impedance (IP) realizations are generated using direct sequential co-simulation (co-DSS). Co-DSS is a geostatistical technique that simulates a secondary variable (in this case, IP) conditioned on a primary variable (facies), ensuring that the histogram, variogram, and crucially, the joint probability distribution between the two properties are honored [14]. This allows modeling the spatial uncertainty of the continuous property (IP) consistently with each facies scenario.

Additionally, [15] proposed two generative models based on Variational Autoencoder (VAE) and Generative Adversarial Network (GAN) architectures to simultaneously predict the facies distribution and the associated acoustic impedance (IP), explicitly incorporating the spatial uncertainty of the elastic property. This approach uses variational Bayesian inference via Inverse Autoregressive Flows (IAF)—a method known as neural transport—to infer the solution of the inverse problem from the latent space of the pre-trained models. The authors demonstrated that while the GAN generates spatially sharper patterns and is statistically more accurate relative to the training images, the VAE performed better at approximating the posterior distribution in the latent space during seismic inversion, owing to the more continuous and regular nature of its mapping.

Addressing the challenge of training data scarcity and the need to capture complex geological patterns without extensive pre-existing databases, both [12] and [13] adopted architectures based on the SinGAN model. Both methods employ a hierarchical pyramid structure of fully convolutional generators and discriminators that operate at multiple scales. This approach allows the network to progressively learn the statistical distribution of image patches, from the coarsest to the finest level, capturing global structures down to detailed textures from a single training image (TI) or a limited dataset.

Despite these advances, a methodological gap remains between approaches that jointly generate facies and petrophysical properties and those designed to operate under limited data availability. Existing joint-generation frameworks typically rely on large training datasets or multi-stage workflows that decouple facies and property modeling, while data-efficient approaches based on SinGAN-style architectures have largely focused on facies generation alone. To the best of our knowledge, no prior work combines conditional, multi-scale generative modeling with joint facies–property synthesis in settings characterized by sparse training data and hard conditioning constraints. This study addresses this gap by extending a data-efficient cGAN architecture to simultaneously generate facies and acoustic impedance within a unified framework, enabling coherent joint modeling under data-constrained conditions.

[1]https://github.com/isreolv/reservoir-gan-inversion

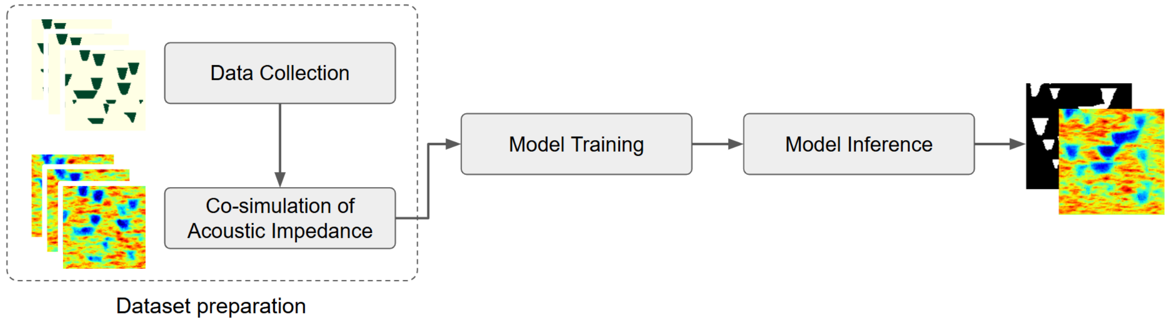


Fig. 1. Overview of the proposed workflow. Facies data are collected and paired with co-simulated acoustic impedance fields during dataset preparation. The joint facies–impedance realizations are used to train the conditional multi-scale GAN, which is then applied during inference to generate new conditioned facies and acoustic impedance realizations.

III. Method

The proposed methodology builds upon the conditional, multi-scale FaciesGAN framework introduced by [12]. As summarized in Fig. 1, the workflow comprises a dataset preparation stage in which facies realizations are paired with co-simulated acoustic impedance fields, followed by model training and inference using a unified, multi-channel generative architecture. While preserving the original SinGAN-based hierarchical training strategy and well-data conditioning mechanism, this work extends the architecture to jointly generate geological facies and associated acoustic impedance. The following subsections describe the data preparation, network architecture, and conditioning strategy adopted in this study.

A. Dataset Preparation

The training dataset consists of paired 2D realizations of facies and acoustic impedance (IP). The facies models, which serve as the primary geological framework, were sourced from the Stanford Earth Science Data dataset (Stanford-VI-E). The original 3D reservoir model was processed by [12] to generate 200 vertical slices from the top layer (80 m thick). Each 2D slice was resampled to a resolution of 256×256 pixels, where the original multi-class facies (floodplain, point bar, channel, boundary) were reclassified into a binary representation: reservoir (channel, labeled as 1) and non-reservoir (all other facies, labeled as 0).

For each binary facies realization, the associated acoustic impedance field was generated using Direct Sequential Co-Simulation (co-DSS) with multi-local distribution functions [14]. This stochastic simulation method ensures that the generated IP values respect the spatial statistics and the specific statistical distributions assigned to each facies type. For this study, Facies 0 (non-reservoir) and Facies 1 (reservoir) were assigned distinct IP distributions (Mean/Standard Deviation: 3500/300 and 2500/200, respectively), with the co-simulation process conditioning the IP field on the facies layout to ensure geological consistency.

B. cGAN Architecture

The proposed generative model adopts a conditional, multi-scale architecture based on the SinGAN framework [16], following the FaciesGAN methodology of [12]. As illustrated in Fig. 2, the model is structured as a pyramid of generators and discriminators operating at progressively increasing spatial resolutions. Training proceeds from the coarsest scale to the finest, enabling the network to first capture large-scale geological structure and subsequently refine local details. Conditioning information derived from well data is injected at every scale, ensuring consistency with hard constraints throughout the generation process. This hierarchical strategy allows the model to learn meaningful spatial variability from limited training data while maintaining stability during adversarial training.

Generator: The generator is implemented as a sequence of fully convolutional networks $\{G_0, \ldots, G_N\}$, each corresponding to a specific scale in the pyramid. At the coarsest scale, generation is purely stochastic, with the generator mapping spatial noise to an initial low-resolution realization. At finer scales, each generator G_n takes as input the upsampled output from the previous coarser scale, concatenated with a noise map and the downsampled conditioning mask representing known well data. The generator learns a residual correction that is added to the upsampled input, progressively enriching the realization with higher-resolution details. In contrast to the original FaciesGAN formulation, the generator output is extended to two channels, jointly producing a discrete facies map and a continuous acoustic impedance field, thereby learning their coupled spatial relationships.

Discriminator: Each generator stage is paired with a corresponding discriminator $\{D_0, \ldots, D_N\}$, forming an adversarial training loop at every scale. The discriminators are implemented as fully convolutional PatchGAN classifiers that operate on local image patches rather than producing a single global decision. At each scale, the discriminator receives either a real or generated two-channel realization and evaluates its realism with respect to the reference data at that resolution. This patch-based formulation encourages the preservation of local geological textures and facies–property relationships, while the multi-scale setup enforces consistency across spatial resolutions. Training is stabilized using the Wasserstein GAN with Gradient Penalty (WGAN-GP), which promotes smooth optimization and mitigates common issues such as mode

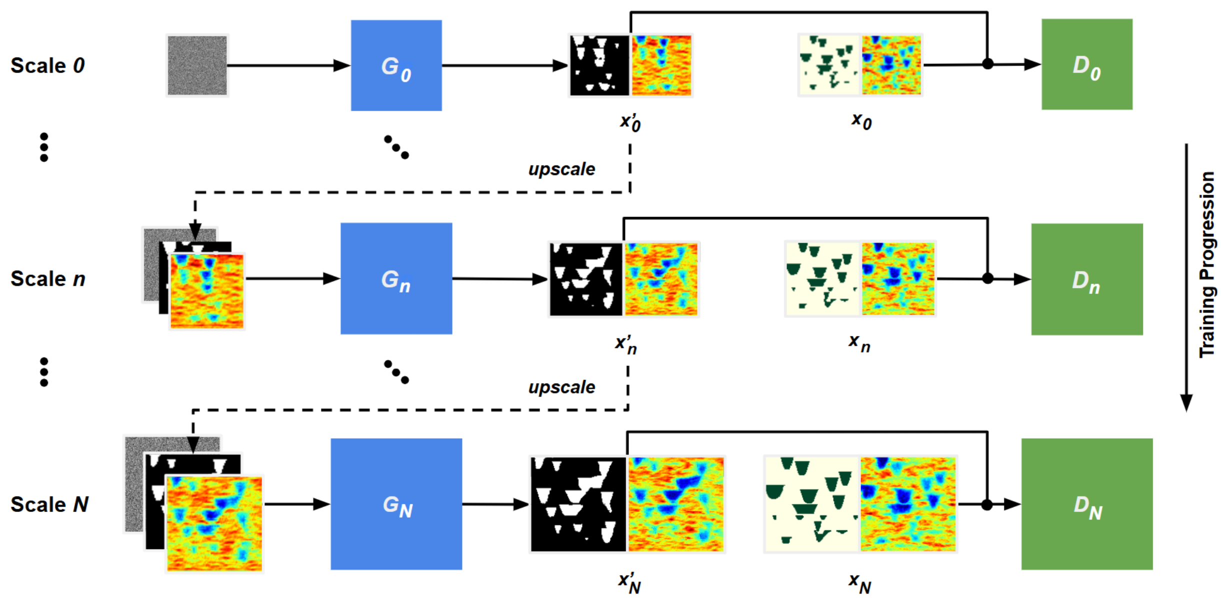


Fig. 2. **Hierarchical architecture of the proposed method.** The framework learns the internal patch distribution of a single image across multiple scales. Each level n captures features at a specific resolution, where the generator G_n produces the residual details required to refine the upsampled output x'_{n-1} from the preceding coarser scale $n-1$. A patch-discriminator D_n at each level ensures that the generated local statistics remain consistent with the original image. Facies maps are displayed using distinct color palettes for visualization purposes only.

collapse.

Convolutional Blocks: Both the generator and discriminator are constructed from repeated convolutional blocks. Each block comprises a 2D convolutional layer and a Leaky ReLU activation function, with batch normalization applied in the generator to stabilize training. This design facilitates gradient flow while preserving stable adversarial optimization under the WGAN-GP formulation.

C. Conditioning and Loss Functions

The model is optimized to honor hard well data through explicit input conditioning and a composite loss formulation.

Input Conditioning: Well information (known facies labels and acoustic impedance) is encoded as a mask. By concatenating this mask with the noise vector at every generator scale, the model strictly guides the joint generation to match direct observations.

Loss Function: The generator minimizes a composite objective combining the WGAN-GP adversarial loss ($\mathcal{L}_{adv}$) with reconstruction penalties:

$$\mathcal{L}_G = \mathcal{L}_{adv} + \alpha \cdot \mathcal{L}_{rec} + \lambda \cdot \mathcal{L}_{mask} \tag{1}$$

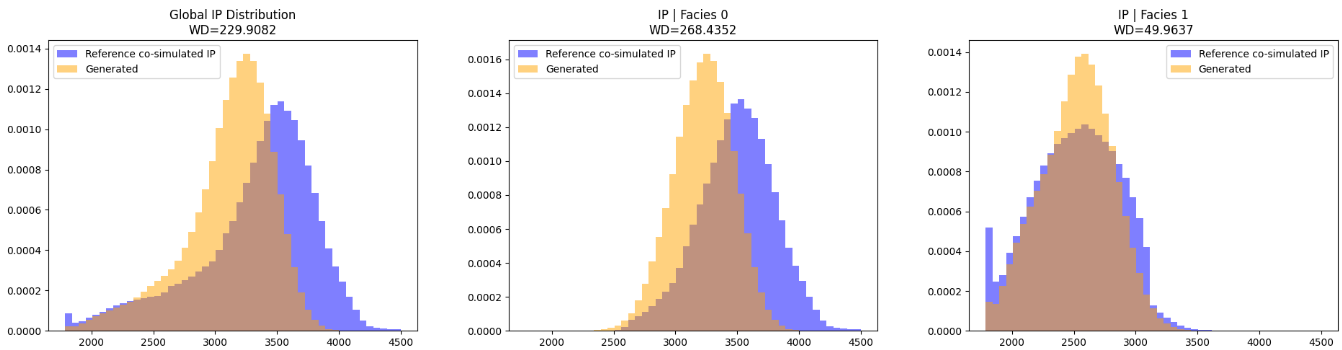
where $\mathcal{L}_{rec}$ calculates the Mean Squared Error (MSE) at well locations, and $\mathcal{L}_{mask}$ enforces strict pixel-wise adherence to the hard data. The coefficients α and λ weight the influence of reconstruction fidelity and hard conditioning, respectively.

The model was trained for 200 epochs at each of the 10 progressive scales. Optimization was performed using the Adam algorithm (learning rate of 5×10^{-5} and $\beta_1 = 0.5$) with a gradient penalty weight of $\lambda = 0.1$. All 2D convolutional filters employed 3×3 kernels with a stride of 1×1. Across the coarser scales, the spatial resolution progressively ranged from 16×16 to 128×128 pixels.

IV. Results

The quality of the generated realizations was assessed using distributional and conditioning-based metrics, focusing on marginal realism and adherence to hard data rather than second-order spatial statistics. Histogram-based comparisons were performed on 1,000 generated realizations, aggregating acoustic impedance (IP) values across all pixels and realizations to estimate global and facies-conditioned distributions, as shown in 3. Generated IP distributions were compared against *reference co-simulated IP* realizations produced using Direct Sequential Co-Simulation (co-DSS). Wasserstein-1 distance was used to quantify discrepancies between reference and generated distributions, as it provides an interpretable measure of distributional mismatch in physical units. In addition, conditioning quality was evaluated by computing the root-mean-square error (RMSE) between generated and reference impedance values at well locations, using vertical well log masks.

As illustrated in Fig.3a the global generated impedance distribution exhibits a moderate deviation from the reference data, with a Wasserstein distance of approximately 230. Visual inspection of the histograms indicates a systematic shift of the generated distribution toward lower impedance values, accompanied by reduced variance relative to the reference distribution. When conditioning on facies, a marked contrast in performance is observed (Fig.3b and Fig.3c). For reservoir facies (Facies 1), the generated impedance distribution closely



(a) Global IP distribution. (b) Non-reservoir (Facies 0). (c) Reservoir (Facies 1).

Fig. 3. Histogram-based comparison between real and generated acoustic impedance (IP) distributions aggregated over 1,000 realizations. (a) Global distribution across all facies. (b) Distribution conditioned on non-reservoir facies. (c) Distribution conditioned on reservoir facies. Wasserstein-1 distances are reported for each case to quantify distributional mismatch.

matches the reference distribution, with well-aligned modes and comparable spread, resulting in a low Wasserstein distance of approximately 50. In contrast, non-reservoir facies (Facies 0) show a more pronounced mismatch, with the generated distribution exhibiting both a leftward shift and reduced variability, reflected in a higher Wasserstein distance of approximately 268. This indicates that while the GAN effectively captures impedance characteristics within reservoir facies, it tends to underestimate impedance values and compress variability in non-reservoir regions.

Conditioning fidelity was evaluated by comparing generated impedance values against observed well log data. Across 454 well samples, the mean RMSE is approximately 419, with a standard deviation of 57 across realizations. Given the overall impedance range of the dataset (1888–4475), this corresponds to a relative error on the order of 16%, indicating that the model honors hard data reasonably well while maintaining stochastic variability across realizations. The non-zero spread in RMSE values suggests that the generator does not trivially overfit well locations, but instead balances conditioning constraints with global realism. Taken together, these results indicate that the proposed cGAN framework is capable of producing geologically plausible, well-conditioned impedance realizations, though with a facies-dependent bias that preferentially affects non-reservoir impedance distributions.

Figure 4 presents three random joint realizations of facies and acoustic impedance generated by the proposed cGAN for three randomly selected wells (Wells 30, 125, and 120). For each well, the reference facies and impedance maps are shown alongside three independent generated samples, with the well trajectory indicated by a vertical line. Visual examination shows that the generated facies realizations strictly honor the conditioning data at the well locations, while exhibiting controlled variability away from the wells that reflects the multi-scale patterns learned by the GAN pyramid. Differences between realizations primarily arise in regions not constrained by hard data, where the network introduces stochastic variations consistent with the training examples. The associated impedance fields remain spatially coherent with the generated facies, with lower impedance values predominantly aligned with reservoir facies and higher values in non-reservoir regions. Across samples, the impedance realizations satisfy the well constraints while preserving variability learned during training, illustrating the model's ability to jointly generate conditioned facies and petrophysical properties under limited data availability.

V. Conclusion

This work addresses the challenge of jointly modeling geological facies and associated rock properties in scenarios characterized by sparse subsurface data and limited availability of training examples. Traditional geostatistical approaches, while robust and well established, often require carefully designed training images and separate modeling steps for categorical and continuous variables. In contrast, recent deep generative models offer a flexible alternative but typically demand large training datasets and focus on single-variable generation. To bridge this gap, this work proposed a conditional GAN framework based on a multi-scale SinGAN architecture that simultaneously generates facies and acoustic impedance fields while explicitly conditioning on hard data from wells. By extending the generator to produce multi-channel outputs and embedding conditioning information at every scale, the method enables coherent joint simulation of discrete and continuous geological properties from relatively few training realizations.

The proposed methodology was evaluated using distributional and conditioning-based metrics on 1,000 generated realizations. Histogram and Wasserstein-1 distance analyses demonstrate that the model effectively reproduces the impedance distribution within reservoir facies, while exhibiting a systematic but interpretable bias toward lower impedance values and reduced variability in non-reservoir facies. Conditioning performance at well locations, quantified through RMSE, indicates that the network honors hard data with reasonable accuracy while preserving stochastic variability

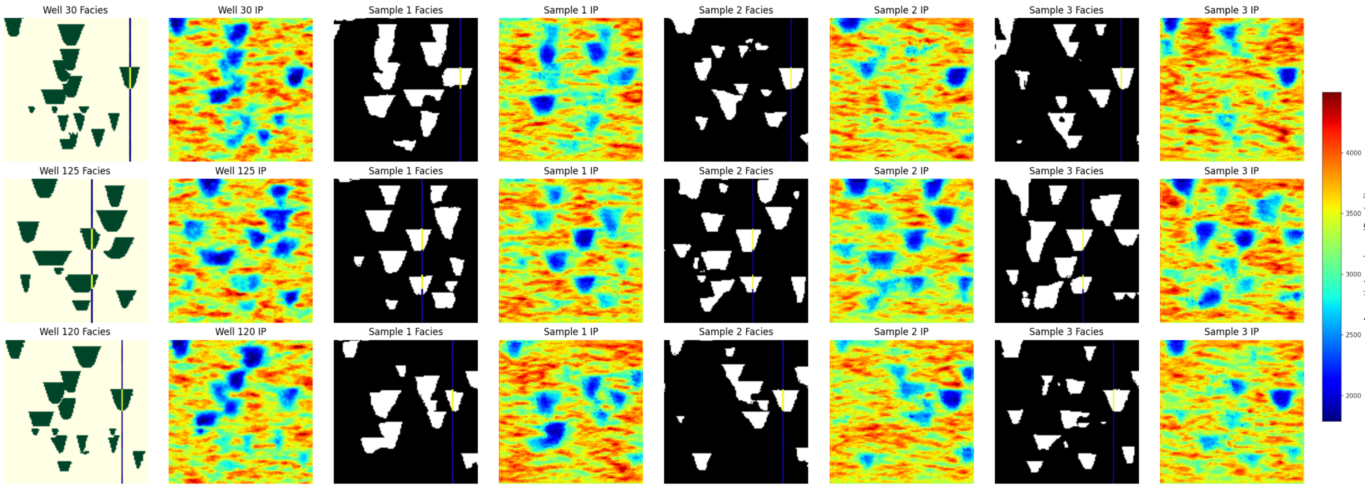


Fig. 4. Reference facies and acoustic impedance maps for three randomly selected wells (left), alongside three independent joint realizations generated by the proposed cGAN (right). Generated samples honor hard conditioning at well locations while exhibiting controlled variability away from the wells. Facies maps are displayed using distinct color palettes for visualization purposes only.

across realizations. These results highlight the ability of the model to balance data conditioning and generative diversity, a key requirement for uncertainty-aware geological modeling in data-constrained settings.

Overall, this study contributes a practical and extensible methodology for the joint generation of facies and petrophysical properties using conditional deep generative models trained on limited data. While the current implementation prioritizes marginal realism and conditioning fidelity over explicit reproduction of second-order spatial statistics, the framework provides a strong foundation for further extensions. Future work may include incorporating additional statistical constraints, alternative loss formulations, or extending the conditioning scheme to integrate seismic data as an additional conditioning source, enabling tighter coupling between subsurface models and geophysical observations.

References

[1] P. Doyen, *Seismic reservoir characterization: an earth modelling perspective*, ser. Education tour series. EAGE publications, 2007.

[2] M. Liu, P. Nivlet, R. Smith, N. BenHasan, and D. Grana, "Chapter 4 - recurrent neural network for seismic reservoir characterization," in *Advances in Subsurface Data Analytics*, S. Bhattacharya and H. Di, Eds. Elsevier, 2022, pp. 95–116.

[3] R. M. Slatt, "Chapter 6 - geologic controls on reservoir quality," in *Stratigraphic Reservoir Characterization for Petroleum Geologists, Geophysicists, and Engineers*, ser. Developments in Petroleum Science, R. M. Slatt, Ed. Elsevier, 2013, vol. 61, pp. 229–281.

[4] D. Grana, T. Mukerji, and P. Doyen, *Seismic Reservoir Modeling: Theory, Examples, and Algorithms*. Wiley, 2021. [Online]. Available: https://books.google.com.br/books?id=XIrfrQEACAAJ

[5] C. V. Deutsch and A. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, ser. Applied geostatistics series. Oxford University Press, 1998. [Online]. Available: https://books.google.com.br/books?id=CNd6QgAACAAJ

[6] A. G. Journel, "Nonparametric estimation of spatial distributions," *Journal of the International Association for Mathematical Geology*, vol. 15, no. 3, p. 445–468, Jun. 1983. [Online]. Available: http://dx.doi.org/10.1007/BF01031292

[7] G. Mariethoz, P. Renard, and J. Straubhaar, "The direct sampling method to perform multiple-point geostatistical simulations," *Water Resources Research*, vol. 46, no. 11, Nov. 2010. [Online]. Available: http://dx.doi.org/10.1029/2008WR007621

[8] E. F. G. lez, T. Mukerji, and G. Mavko, "Seismic inversion combining rock physics and multiple-point geostatistics," *Geophysics*, vol. 73, no. 1, pp. R11–R21, 2008.

[9] R. Miele, L. Azevedo, and D. Grana, "Iterative stochastic seismic inversion method using gan with spatiallyadaptive normalization," in *Third EAGE Conference on Seismic Inversion*, vol. 2024. European Association of Geoscientists & Engineers, 2024, pp. 1–5.

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[11] L. Mosser, W. Kimman, J. Dramsch, S. Purves, A. De la Fuente Briceño, and G. Ganssle, "Rapid seismic domain transfer: Seismic velocity inversion and modeling using deep generative neural networks," in *80th eage conference and exhibition 2018*, vol. 2018. European Association of Geoscientists & Engineers, 2018, pp. 1–5.

[12] Y. V. C.-D. Angulo, T. Mazzutti, B. B. Rodrigues, and M. Roisenberg, "Faciesgan: A conditional gan framework for realistic facies scenario generation as an efficient alternative to multiple-point statistics," *Journal of Seismic Exploration*, vol. 34, no. 6, p. 45, Nov. 2025. [Online]. Available: http://dx.doi.org/10.36922/JSE025370069

[13] R. Feng, D. Grana, and K. Mosegaard, "Geostatistical facies simulation based on training image using generative networks and gradual deformation," *Mathematical Geosciences*, vol. 57, no. 6, p. 1021–1044, Jan. 2025. [Online]. Available: http://dx.doi.org/10.1007/s11004-024-10169-y

[14] R. Nunes, A. Soares, L. Azevedo, and P. Pereira, "Geostatistical seismic inversion with direct sequential simulation and co-simulation with multi-local distribution functions," *Mathematical Geosciences*, vol. 49, no. 5, p. 583–601, Aug. 2016. [Online]. Available: http://dx.doi.org/10.1007/s11004-016-9651-0

[15] R. Miele, S. Levy, N. Linde, A. Soares, and L. Azevedo, "Deep generative networks for multivariate fullstack seismic data inversion using inverse autoregressive flows," *Computers &; Geosciences*, vol. 188, p. 105622, Jun. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.cageo.2024.105622

[16] T. R. Shaham, T. Dekel, and T. Michaeli, "Singan: Learning a generative model from a single natural image," 2019. [Online]. Available: https://arxiv.org/abs/1905.01164