LEARNING MEANINGFUL GEOLOGICAL FIELDS REPRESENTATIONS BY SELF-SUPERVISED METHODS

Anonymous authors
Paper under double-blind review

ABSTRACT

Geological fields are challenging to model as they reflect complex multi-scale sedimentation processes during the whole history of Earth's existence. The geological objects that can be found in such fields have very different nature and scale and are often represented by statistical means. In this work we investigate the possibility to obtain meaningful representations for geological fields of different kinds with the help of self-supervised learning (SSL). Specifically, we adapt a self-supervised vision transformer architecture DINOv2 to a dataset of more than twenty thousands synthetic 3D geological cubes describing four sedimentary environments to learn their representations without labels. We show that learned embeddings cluster by sedimentary environment type and have high correlation with important geostatistical properties, such as variograms. When used as features for classification and regression, they match or exceed the performance of supervised CNNs trained from scratch. Our results demonstrate that SSL can capture meaningful geological structure in 3D geological data and serve as a strong foundation for downstream tasks, reducing the need for expensive labeled datasets in geological modeling.

1 INTRODUCTION

Geological fields are difficult to model because they reflect complex multi-scale sedimentation processes accumulated over geological time. Objects in such fields vary in nature and scale, and they are often characterized through statistical descriptors such as variograms, correlation lengths, and anisotropy. Traditional machine learning approaches in geoscience have relied heavily on supervised convolutional neural networks (CNNs) trained for facies classification, seismic interpretation, or regression of petrophysical properties. While effective, these methods demand large labeled datasets, which are costly to obtain.

Self-supervised learning (SSL) provides an alternative that bypasses the need for manual labels by exploiting invariances and data augmentations. Recent breakthroughs in SSL—contrastive learning (SimCLR (Chen et al., 2020), MoCo (He et al., 2020), SwAV (Caron et al., 2020)), non-contrastive predictive methods (BYOL (Grill et al., 2020)), and self-distillation frameworks (DINO (Caron et al., 2021), DINOv2 (Oquab et al., 2023))—have shown remarkable success in computer vision. Vision transformers (ViTs) (Dosovitskiy et al., 2021) trained with SSL objectives now deliver state-of-the-art results on downstream tasks and generalize well in low-label regimes. Masked autoencoders (MAE) (He et al., 2022) further extended SSL to reconstruction-based objectives.

Despite this progress, SSL has rarely been applied to geoscience, where data are volumetric, multi-channel, and governed by physical laws. Early works have explored supervised CNNs for facies classification (Liu, 2020), Bayesian CNNs for uncertainty quantification (Feng, 2021), and generative adversarial networks (GANs) or variational autoencoders (VAEs) for stochastic modeling of porous media and sedimentary environments (Mosser, 2017; Laloy, 2018). More recently, there have been first attempts to apply SSL to seismic data (Li, 2025), but systematic investigation of SSL embeddings for geological cubes and their geostatistical properties remains absent.

In this work, we investigate whether SSL can produce meaningful representations of geological fields. Specifically, we adapt the DINOv2 framework (Oquab et al., 2023) to a dataset of more than 20,000 synthetic 3D geological cubes describing four sedimentary environments. We show that the resulting embeddings cluster by environment type and correlate with key geostatistical descriptors.

These findings indicate that SSL can capture geologically meaningful structures in 3D data and provide a strong foundation for downstream modeling.

Our main contributions:

- Applied SSL to representation learning of geological fields
- We found that learned representation states is well aligned with important geological statistics

2 Related Work

2.1 Geological data classification

Depositional environments can be systematically classified into a particular reservoir architecture type according to their modelling approach of sand-body that forms heterogeneity of different types. Each type of heterogeneity influences fluid flow characteristics and hence sweep and recovery efficiencies (Caers, 2005).

There are three main types of depositional environments depending on their location within the sedimentary basin system. The information about them can be found in A.1.

Each environment exhibits characteristic depositional facies, energy regimes, and sediment textures that affect porosity and permeability distribution within reservoir rocks. The classification also includes process-based subdivisions, such as wave-dominated, tide-dominated and river-dominated that controls facies architecture and stacking patterns (Galloway, 1975; Bhattacharya & Walker, 1992).

The present phase of research focuses only on some of the depositional environments that are common for petroleum industry and forms relatively simple reservoir architecture. For better understanding the authors of present paper refer to sedimentological description of each environment that provided in Typical Facies Atlas (Tugarova & Zhukovskaya, 2019).

2.2 Overview of Geological Modeling Methods Using Geostatistics

Geostatistical modeling has become a foundational approach in geological reservoir characterization, offering quantitative tools for describing spatial variability and uncertainty in subsurface properties. Unlike deterministic methods, geostatistical techniques rely on statistical representations of spatial relationships—typically defined through variograms or covariance functions—to generate multiple equiprobable realizations of geological variables such as facies, porosity, and permeability (Journel & Huijbregts, 1978; Deutsch et al., 1992). Commonly applied methods include Sequential Indicator Simulation (SIS), which models categorical variables (e.g., facies types) by assigning class values based on conditional probabilities derived from indicator transforms. Sequential Gaussian Simulation (SGS) generates continuous property models by drawing values from a multivariate Gaussian distribution conditioned on neighboring data. Truncated Gaussian Simulation creates realistic geological patterns by applying thresholds to one or more Gaussian fields to define facies distributions with spatial continuity (Caers, 2005).

These methods are based on variogram analysis. The variogram quantifies spatial continuity and the degree of similarity between sampled data as a function of distance and direction. Variogram models are defined for each property and depositional environment and are critical for preserving geological realism in stochastic simulations. Key variogram parameters include the sill (total variance), range (distance of spatial correlation), and nugget (microscale variability or measurement error).

The strength of geostatistical methods lies in their ability to capture heterogeneity and assess uncertainty through stochastic realizations. By generating ensembles of models that honor both hard data (e.g., well logs) and soft constraints (e.g., trends or facies proportions), geostatistics supports probabilistic forecasting and risk-based decision-making in reservoir development (Chiles & Delfiner, 2012). Object-based modeling places predefined geological objects (e.g., channels, lobes) according to geological rules and spatial statistics. Multiple-point geostatistics (MPS) uses training images to replicate complex geological patterns by conditioning simulation on multiple neighboring points rather than just pairwise correlations (Strebelle, 2002). More recently, machine learning-assisted

workflows have been applied to learn spatial features directly from data and enhance the realism and efficiency of model generation (Zhou et al., 2024).

Despite their broad application, geostatistical modeling techniques have several important limitations that must be acknowledged during model construction. These methods rely heavily on variogram analysis, which assumes spatial stationarity. However, many sedimentary environments are non-stationary and heterogeneous, making it difficult to define a single representative variogram (Journel & Huijbregts, 1978). Sequential simulation methods (e.g., SIS, SGS) often neglect geological architecture and may produce noisy or unrealistic facies distributions. In addition, Truncated Gaussian Simulation is highly sensitive to threshold and correlation parameters, which may result in non-geological facies shapes (Galli et al., 1994). While Multiple-Point Statistics (MPS) improves pattern realism, it depends on training images that must be geologically representative and consistent with conditioning data. These simulations are also computationally demanding (Strebelle, 2002). Object-based models require detailed assumptions about facies geometry, while machine learning-based methods need large training datasets and often operate as "black boxes," which limits interpretability (Zhou et al., 2024).

2.3 Representation learning

Representation learning has become a cornerstone of modern machine learning. Classical supervised deep models such as CNNs demonstrated strong performance in vision tasks but required large annotated datasets. To reduce label dependence, SSL methods have emerged. Contrastive approaches like SimCLR (Chen et al., 2020), MoCo (He et al., 2020), and SwAV (Caron et al., 2020) maximize agreement between differently augmented views of the same sample. Non-contrastive predictive methods such as BYOL (Grill et al., 2020) avoid negative pairs and learn stable representations.

A parallel line of work explores distillation-based SSL. DINO (Caron et al., 2021) introduced a teacher–student framework with no labels, enabling vision transformers (ViTs) (Dosovitskiy et al., 2021) to learn high-quality features. Its successor, DINOv2 (Oquab et al., 2023), scaled this recipe with improved architectures and data pipelines, producing foundation models for vision. Reconstruction-based SSL has also shown promise; Masked Autoencoders (MAE) (He et al., 2022) reconstruct missing patches and have proven effective for large-scale pretraining. Beyond 2D vision, SSL has been extended to scientific and geospatial domains. In medical imaging, SSL methods such as contrastive 3D patch learning have improved performance in low-label MRI and CT tasks (Zhou et al., 2021; Azizi et al., 2021). In geospatial representation learning, domain-specific pretext tasks for remote sensing imagery have been proposed (Ayush et al., 2021). Despite these advances, applications of SSL to geological volumetric data remain rare.

3 Methodology

3.1 Data generation

In order to generate representative datasets for various depositional environments corresponding to particular cells of a reservoir architecture matrix, a multi-stage modeling workflow was implemented. This approach enables the assessment of geological uncertainty associated with sedimentary heterogeneity and spatial distribution of reservoir properties.

The modeling process began with the compilation of a comprehensive database of geological and geophysical parameters, collectively referred to as the Geological Context. This context includes detailed descriptions of facies sequences and lithological characteristics, as well as supporting data such as analog outcrop images, core descriptions, well logs, and stratigraphic motifs. These analog datasets served as the basis for defining variable ranges and training the geostatistical models.

To simulate subsurface heterogeneity in 3D, synthetic data sets were constructed and loaded into Petrel, a Schlumberger software platform widely used for reservoir characterization. Petrel enables integration of geological, petrophysical, and geostatistical modeling techniques at different scales. To learn more about the dataset acquisition, go to A.2.

Each depositional model was simulated in 5001 realizations, allowing for comprehensive assessment of uncertainty arising from spatial configuration of geological bodies, connectivity, and prop-

erty distribution. The parameters were varied systematically across realizations using probabilistic ranges informed by analog data. Each realization includes property grids for facies, porosity, and permeability—variables most sensitive to geological heterogeneity.

In total, our dataset comprises 20,004 synthetic 3D geological cubes with dimensions 150x144x80 voxels. Each cube has three channels: facies, porosity, and permeability. Such cubes can be treated as multi-channel volumetric images analogous to 3D medical, where SSL has already shown some potential (et al., 2021).

3.2 MODEL

We adapt DINOv2 (Caron et al., 2021; Oquab et al., 2023) to 3D volumetric inputs. DINOv2 is a self-distillation framework in which a student network is trained to predict the embeddings of a momentum-updated teacher given different augmented views of the same input. Specifically, we encode 3D image into 2D image using CNN. 2D image then goes into ViT. We also design volumetric augmentations, including random cropping, rotation, and noise. We also employ multi-crop training, combining global and local 3D views, following the original DINO recipe. We compare DINOv2 against alternative I-JEPA (Assran, 2023) framework. This ensures that our conclusions are not specific to a single SSL paradigm.

We implemented and trained:

- Supervised CNNs (3D ResNet) trained on labeled cubes.
- SSL models based on DINOv2 and I-JEPA

3.3 EVALUATION CRITERIA

We evaluate learned embeddings in three categories:

- Unsupervised analysis. We apply dimensionality reduction (t-SNE) to visualize clusters with respect to sedimentary environment labels. To test whether embeddings encode geological realism, we calculate correlations with geostatistical descriptors.
- Strength of correlation between embeddings and geostatistical descriptors.

For supervised classification and regression we use frozen embeddings as input features to lightweight classifiers/regressors (logistic/Ridge regression) for predicting environment type and geostatistical properties. We benchmark against 3D CNNs trained end-to-end.

4 RESULTS

4.1 REPRESENTATIONS

In this section, we describe representations learned by Supervised CNNs (representations comes from the last layer before the prediction head) and SSL models based on DINOv2 and I-JEPA. We analyze their meaningfullness as predicted

4.1.1 T-SNE DECOMPOSITION OF THE REPRESENTATIONS

We first reduce the representations dimensionality to 2 using T-SNE (Maaten & Hinton, 2008). Then we visualize them as presented in Figure 1.

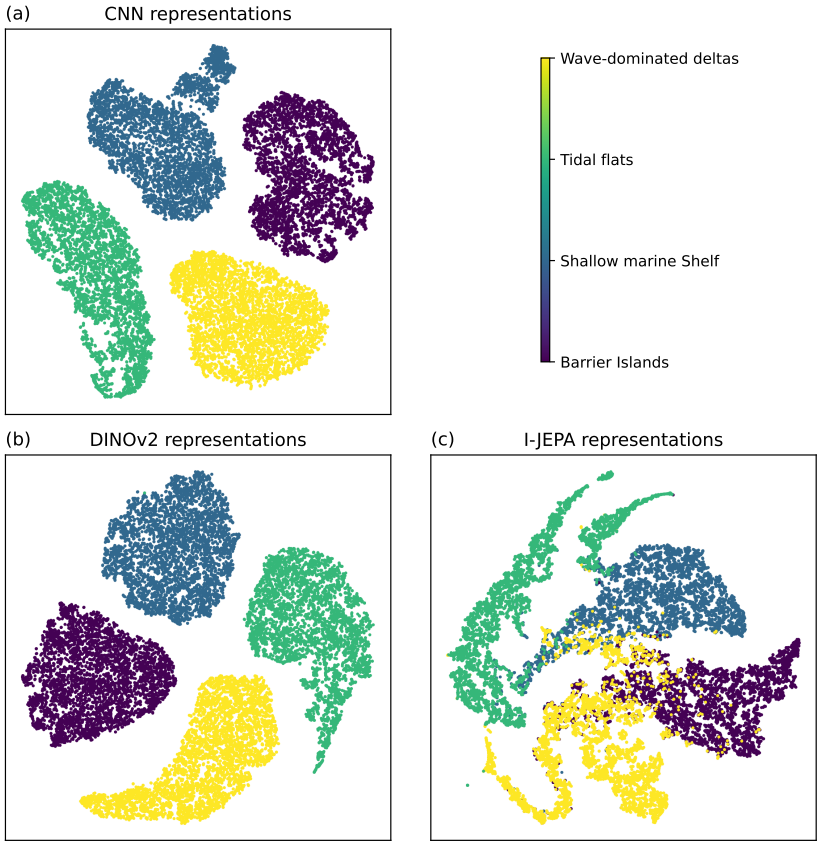

Figure 1: T-SNE decomposition of representations. The points are distributed into 4 clusters, each corresponding to the different geological depositional environment. (a) corresponds to CNN representations, (b) to DINOv2 representations, (c) to I-JEPA representations.

4.1.2 GEOLOGICAL STATISTICS PREDICTION

There are total 20 geological statistics computed in this study for each 3D geological cube (see Section B.1. In this section, we present the results of the Ridge regression, which take representation as input and predict the statistics as output.

As described in the Section B.1, there are 4 different groups of the geological statistics:

1. Facies statistics (total 4 statistics) - described in B.1.1, presented in Figure 2.
2. Lorenz coefficient (total 5 statistics) - described in B.1.2, presented in Figure 3.
3. Average derivatives (total 6 statistics) - described in B.1.3, presented in Figure 4.

4. Dykstra-Pasrsons coefficient (total 5 statistics) - described in B.1.4, presented in Figure 5.

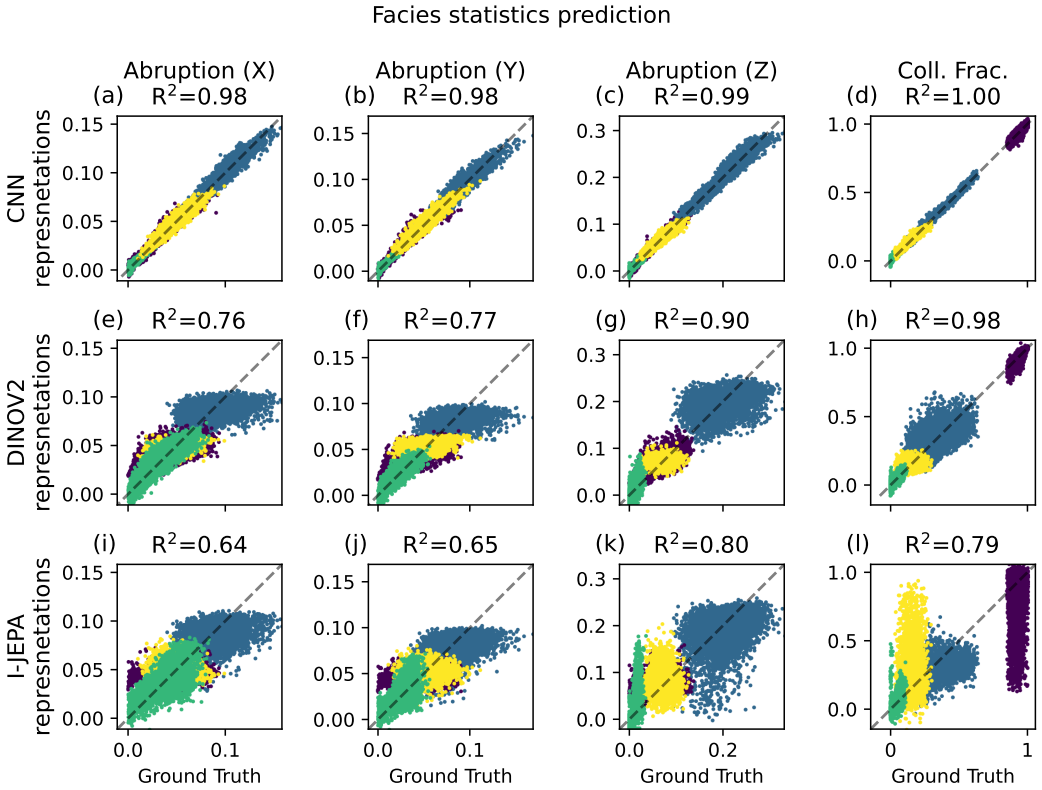

Figure 2: Prediction of the facies statistics from representations. Each points represents a geological statistics from single 3D cube. X axis corresponds to the ground truth value computed as described in the Section B.1.1. Y axis is the value of such statistics predicted from representations. The color represents depositional environment as described in Figure 1.

The average scores for each representation method are presented in the Table 1.

Table 1: Prediction scores for the Ridge regression tasks and depositional environment classification score (see Section 4.1.1).

Representations	CNN	DINOv2	I-JEPA
Depositional environment classification score	1.0	1.0	0.92
Facies statistics regression score	0.99	0.85	0.72
Lorenz coefficient regression score	0.94	0.76	0.65
Derivatives coefficient regression score	0.99	0.91	0.77
Dykstra-Parsons coefficient regression score	0.93	0.86	0.73

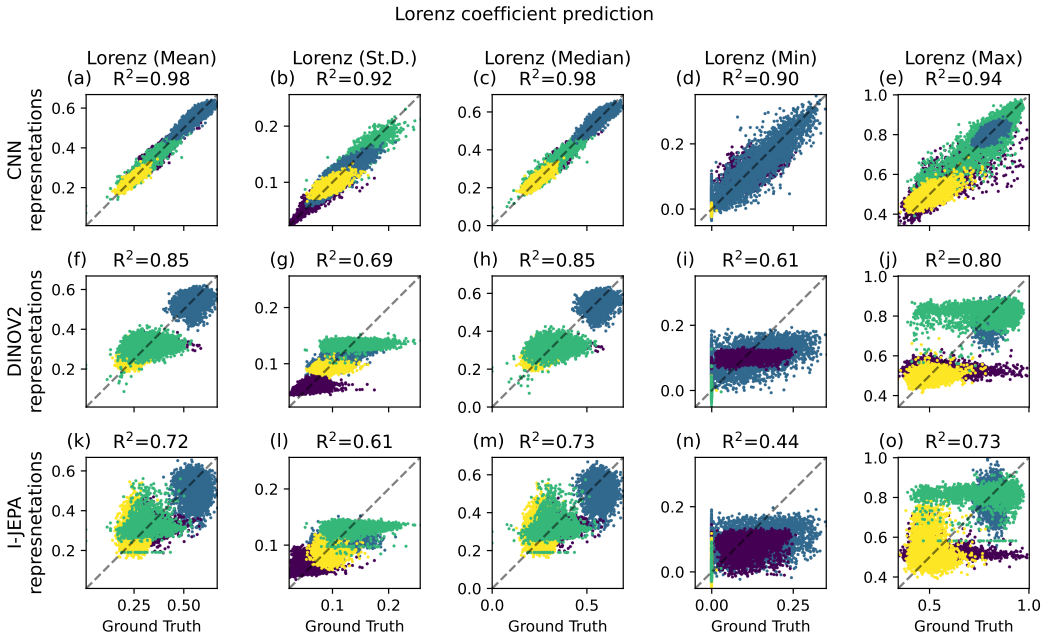

Figure 3: Prediction of the Lorenz coefficient from representations (see Section B.1.2).

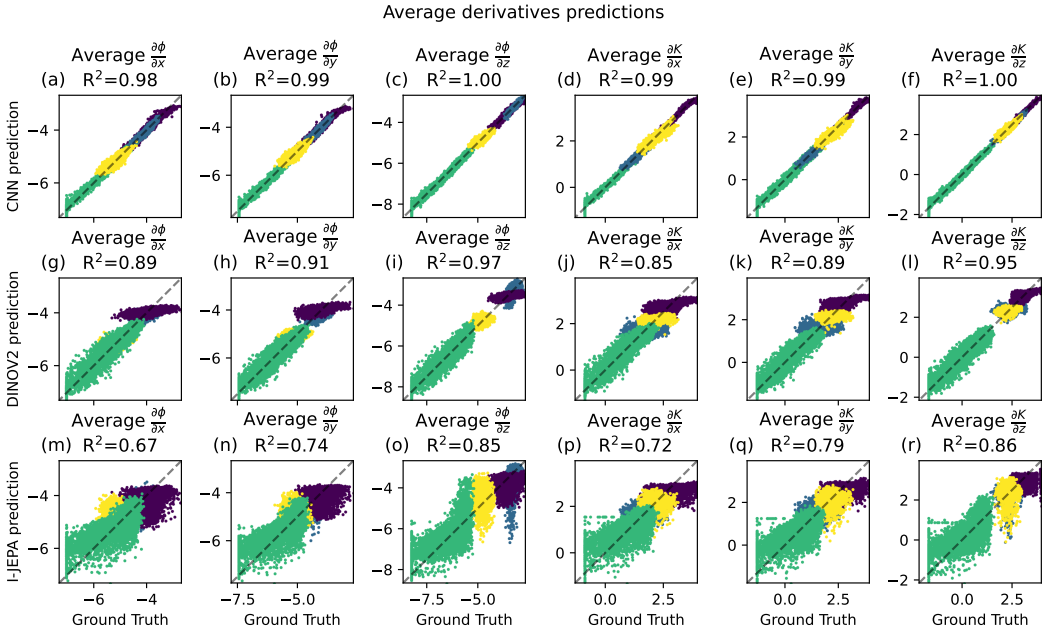

Figure 4: Prediction of the average derivatives from represenetaions (see Section B.1.3).

5 Conclusion and Discussion

The results show that self-supervised methods can extract non-trivial structure from synthetic geological realizations.

First, SSL representations separate depositional environments in latent space without direct supervision (see Figure 1). T-SNE visualizations display clear clusters aligned with depositional environments. This indicates that the DINOv2 framework captures depositional signatures.

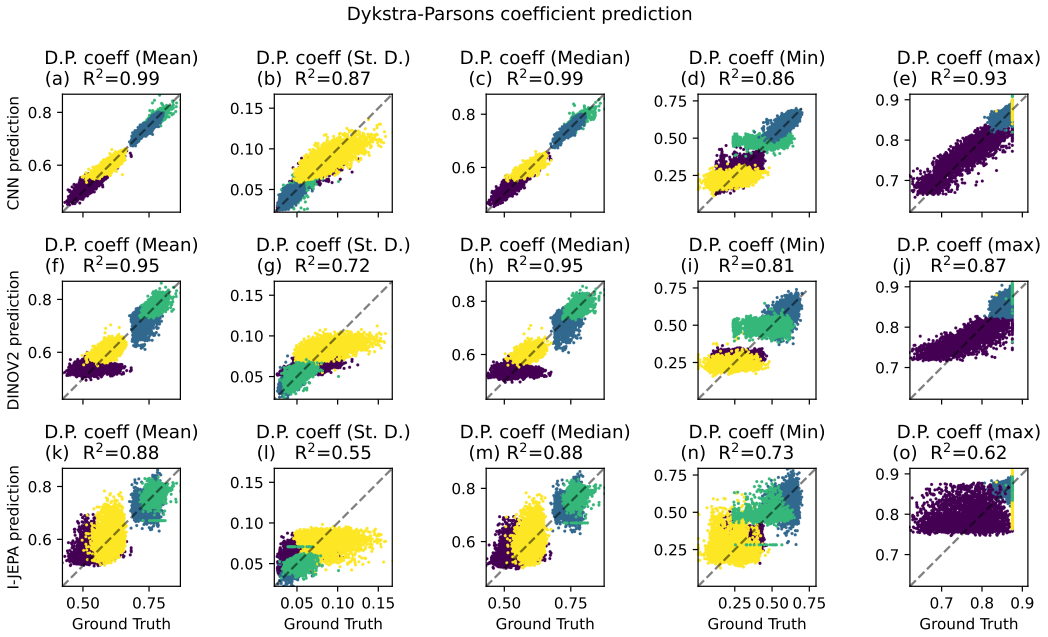

Figure 5: Prediction of the Dykstra-Parsons coefficient from representations (see Section B.1.4).

Second, the embeddings correlate strongly with geostatistical descriptors (see Figures 2, 3, 4, 5). Ridge regression on SSL features in DINOv2 case yields good R^2 scores without any supervision. I-JEPA model also works here, but worse then DINO approach. This confirms that self-distillation mechanisms preserve geological realism, not only categorical information. The supervised fine-tuning of these models can improve the results.

Comparison of paradigms shows that DINOv2 generally outperforms I-JEPA in classification and correlation tasks and in some cases (see Figure 4 and Table 1) can compete with CNN supervised features.

It is worth noting that we used the most simple scenario to adapt 3D data to a 2D model. The application of 3D ViT can improve the results, but will require much more memory and computation. Nevertheless, adapting image-based SSL to volumetric geological data is effective, despite the non-natural and high-dimensional character of the inputs. Conceptually, this shows that geological heterogeneity, traditionally represented by statistical descriptors, can also be encoded in learned embeddings that are transferable across tasks.

Limitations remain. Model size and training cost are also significant compared to classical geostatistical methods.

In addition, the trained represnetation models can be used for training autoencoders and variational autoencoders. As it is known from the literature Hou et al. (2017), VAE trained in such a way results in more realistic and less blurry images reconstruction. Thus, VAE trained on the geological 3D cubes with DINOv2 perceptual loss will likely preserve geological statistics even without explicit access to these statistics. This is the future direction of this study.

Overall, the study demonstrates that self-supervised vision transformers, when adapted to 3D geological cubes, provide representations that align with both geological categories and statistics. This positions SSL as a promising direction for geological modeling under uncertainty.

REFERENCES

J.R.L. Allen. *Physical Processes of Sedimentation*. Springer, 1997. ISBN 9780444196507.

M. Assran. Self-supervised learning from images with a joint-embedding predictive architecture. *arxiv:2301.08243*, 2023.

Kumar Ayush, Burak Uzkent, Chenlin Meng, Katyal Tanmay, Marshall Burke, David Lobell, and Stefano Ermon. Geography-aware self-supervised learning. In *ICCV*, 2021.

Shekoofeh Azizi, Basil Mustafa, Fiona Ryan, Zachary Beaver, Jesse Freyberg, Jeremy Deaton, Simon Kornblith, Ting Chen, William Li, Xiaohua Liu, et al. Big self-supervised models advance medical image classification. In *ICCV*, 2021.

J. P. Bhattacharya and Roger G Walker. Facies model: response to sea level change. *Geol. Asso. Canada*, 409:157–177, 1992.

Jef Caers. Petroleum geostatistics. *Society of Petroleum Engineers.*, 2005.

M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin. Unsupervised learning of visual features by contrasting cluster assignments. In *NeurIPS*, 2020.

M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin. Emerging properties in self-supervised vision transformers. In *ICCV*, 2021.

T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020.

Jean-Paul Chiles and Pierre Delfiner. *Geostatistics: modeling spatial uncertainty*. John Wiley & Sons, 2012.

Clayton V Deutsch, Andre G Journel, et al. Geostatistical software library and user's guide. *New York*, 119(147):578, 1992.

A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2021.

Yamel A. et al. Self-supervised learning for 3d medical image analysis using 3d simclr and monte carlo dropout. *arXiv:2109.14288*, 2021.

R. et al. Feng. Bayesian convolutional neural networks for seismic facies classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

A Galli, H Beucher, G Le Loc'h, B Doligez, and Heresim Group. The pros and cons of the truncated gaussian method. In *Geostatistical Simulations: Proceedings of the Geostatistical Simulation Workshop, Fontainebleau, France, 27–28 May 1993*, pp. 217–233. Springer, 1994.

William E Galloway. Process framework for describing the morphologic and stratigraphic evolution of deltaic depositional systems. *Houston Geological Society*, 1975.

J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. Pires, Z. Guo, M. Gheshlaghi Azar, et al. Bootstrap your own latent: A new approach to self-supervised learning. In *NeurIPS*, 2020.

K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick. Momentum contrast for unsupervised visual representation learning. In *CVPR*, 2020.

K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick. Masked autoencoders are scalable vision learners. In *CVPR*, 2022.

Xianxu Hou, Linlin Shen, Ke Sun, and Guoping Qiu. Deep feature consistent variational autoencoder. In *2017 IEEE winter conference on applications of computer vision (WACV)*, pp. 1133–1141. IEEE, 2017.

A. G. Journel and C. J. Huijbregts. *Mining geostatistics*. Academic Press, 1978.

E. et al. Laloy. Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, 2018.

M. et al. Li. A self-supervised deep learning framework for seismic facies segmentation. *Expert Systems with Applications*, 2025.

M. et al. Liu. Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks. *Geophysics*, 2020.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

Andrew D Miall. *The geology of fluvial deposits: sedimentary facies, basin analysis, and petroleum geology*. Springer, 1996.

L. et al. Mosser. Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Phys. Rev. E*, 2017.

VS Muromtsev. Electrometric geology of sand bodies–lithological traps of oil and gas. *L.: Nedra*, 260, 1984.

Emiliano Mutti and William R Normark. Comparing examples of modern and ancient turbidite systems: problems and concepts. *Marine clastic sedimentology: Concepts and case studies*, pp. 1–38, 1987.

M. Oquab, T. Darcet, T. Moutakanni, H. Vo, M. Szafraniec, V. Khalidov, Y. Chen, J. Xie, I. Misra, A. Joulin, P. Bojanowski, and M. Caron. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Michael J Pyrcz and Clayton V Deutsch. *Geostatistical reservoir modeling*. Oxford university press, 2014.

Sebastien Strebelle. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology*, 34:1–21, 2002.

M. A. Tugarova and E. A. Zhukovskaya. Atlas of typical facies. *Moscow: VNIGNI*, 2019.

KJ Webber and LC Van Geuns. Framework for constructing clastic reservoir simulation models. *Journal of Petroleum Technology*, 42(10):1248–1297, 1990.

KJ Weber. How heterogeneity affects oil recovery. *Reservoir characterization*, 487:544, 1986.

Wensheng Zhou, Chen Liu, Yuandong Liu, Zenghua Zhang, Peng Chen, and Lei Jiang. Machine learning in reservoir engineering: A review. *Processes*, 12(6):1219, 2024.

Zongwei Zhou, Vinay Sodha, Md Mahfuzur Rahman Siddiquee, Ruibin Feng, Nima Tajbakhsh, Michael Gotway, and Jianming Liang. Models genesis: Generic autodidactic models for 3d medical image analysis. *Medical Image Analysis*, 2021.

A Appendix A. Dataset generation details

The selected depositional environments are characterized by varying geological contexts, including the number of facies, SP log behavior according to the Muromtsev approach (Muromtsev, 1984), structural and textural features, facies reservoir properties, and the spatial distribution of facies.

Barrier islands form from regressive bars and show upward grain-size coarsening, with lens-shaped bodies aligned along the shoreline. They are associated with low-energy lagoonal settings are represented by shale sediments and non-reservoir.

Shallow marine shelf influenced by wave and tidal processes, display upward-coarsening trends from distal to nearshore sublittoral zones.

Transitional environments represented by Tidal Flat Environments and Wave-Dominated Delta (Distal Deltaic Systems).

A.1 Classification of Depositional Environments

Authors of ref. (Webber & Van Geuns, 1990) proposed a foundational framework for classifying clastic depositional environments based on their internal reservoir architecture, focusing on the geometric and connectivity attributes of sedimentary bodies that directly influence fluid flow. This approach moves beyond purely lithofacies-based interpretations, integrating sedimentological processes with reservoir simulation needs.

The classification identifies three principal architectural types, each associated with distinct depositional systems and flow characteristics:

Layer-cake Architectures. These are extensive, laterally continuous units with relatively uniform thickness, commonly associated with shoreface, shallow marine, and distal deltaic environments. Their homogeneity and lateral connectivity typically promote efficient sweep and recovery (Webber & Van Geuns, 1990; Allen, 1997).

Jigsaw Architectures. Found in fluvial and tidal channel systems, these are narrow, elongate bodies with high aspect ratios and strong directional anisotropy. These reservoirs are often heterogeneous, with variable connectivity and frequent internal baffles or barriers (Miall, 1996; Weber, 1986).

Labyrinth Architectures. These architectures form in depositional systems such as deltaic mouth bars, turbidite lobes, or fan deltas. They typically display radial or fan-shaped geometries, vertical stacking, and significant lateral facies transitions, resulting in complex reservoir compartmentalization (Mutti & Normark, 1987; Webber & Van Geuns, 1990).

The classification highlights the importance of understanding depositional architecture as a control on dynamic behavior, not just static heterogeneity. It also emphasizes the need to tailor simulation grid designs and flow unit definitions to match the dominant architectural style of each depositional system. This classification was adopted as the baseline framework for assigning depositional environments in the current study:

Barrier islands. The geomorphology of these bodies is sheet-like and elongate, oriented parallel to the shoreline, covering tens to hundreds of square kilometers (1 cell). The example is presented in Figure 6.

Shallow marine shelf. The shallow marine shelf is also characterized by laterally continuous bodies, but the number of facies in this model increases, thus shifting to cell 4. Vertical heterogeneity becomes more pronounced, while lateral heterogeneity remains low. The example is presented in Figure 7.

Tidal flats. This environment is characterized by discontinuous bodies due to the presence of channels. Channels often appear as sinuous, ribbon-like bodies in the subsurface, with thicknesses up to 20 meters and widths up to 500 meters (cell 2). The example is presented in Figure 8.

Wave-dominated deltas. This environment exhibits a jigsaw-like architecture, where incising channels cut through laterally continuous facies, fragmenting them into discontinuous bodies (cell 5). The example is presented in Figure 9.

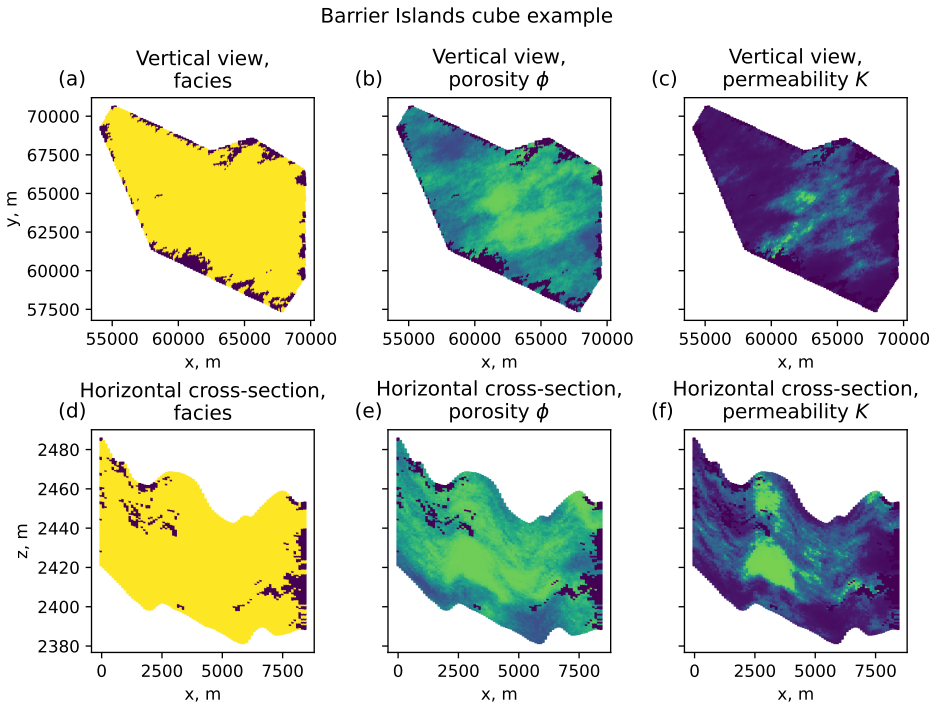

Figure 6: The example 3D cube from Barrier Island depositional environment.

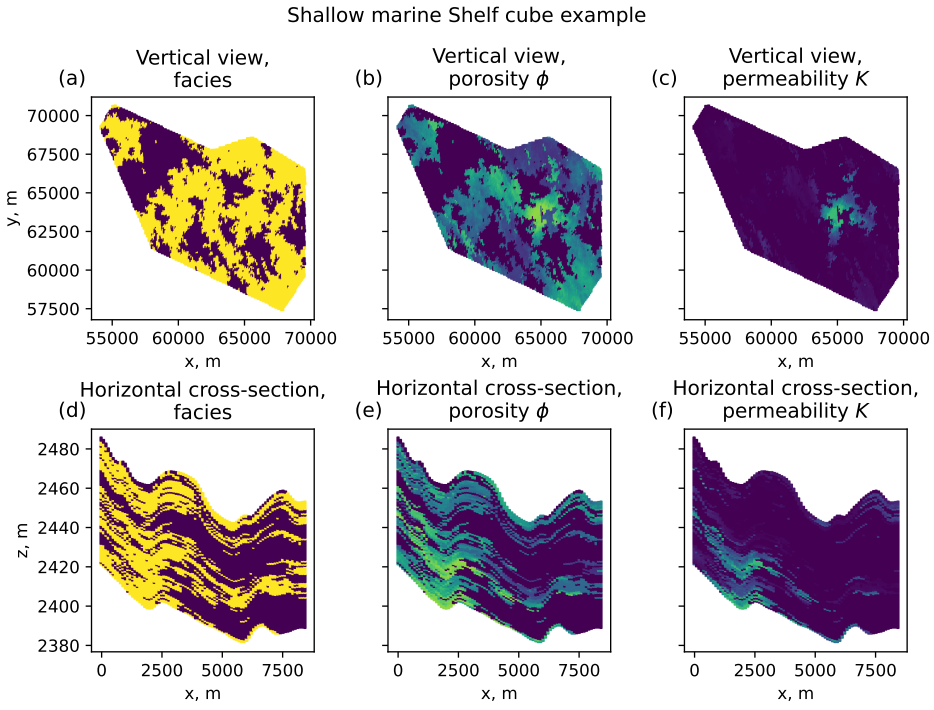

Figure 7: The example 3D cube from Shallow marine shelf depositional environment.

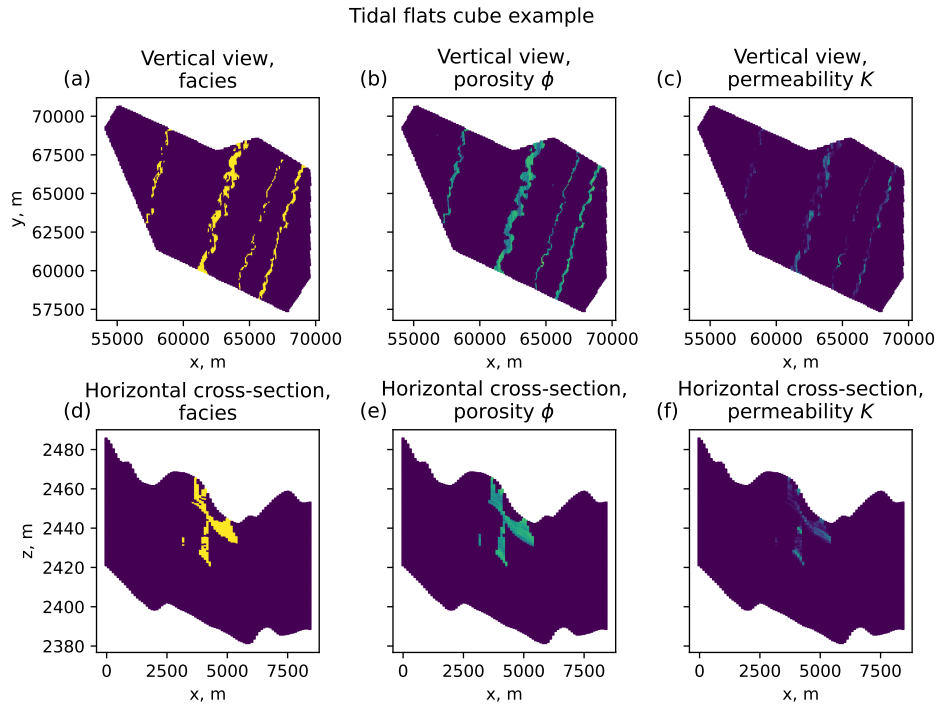

Figure 8: The example 3D cube from Tidal flats depositional environment.

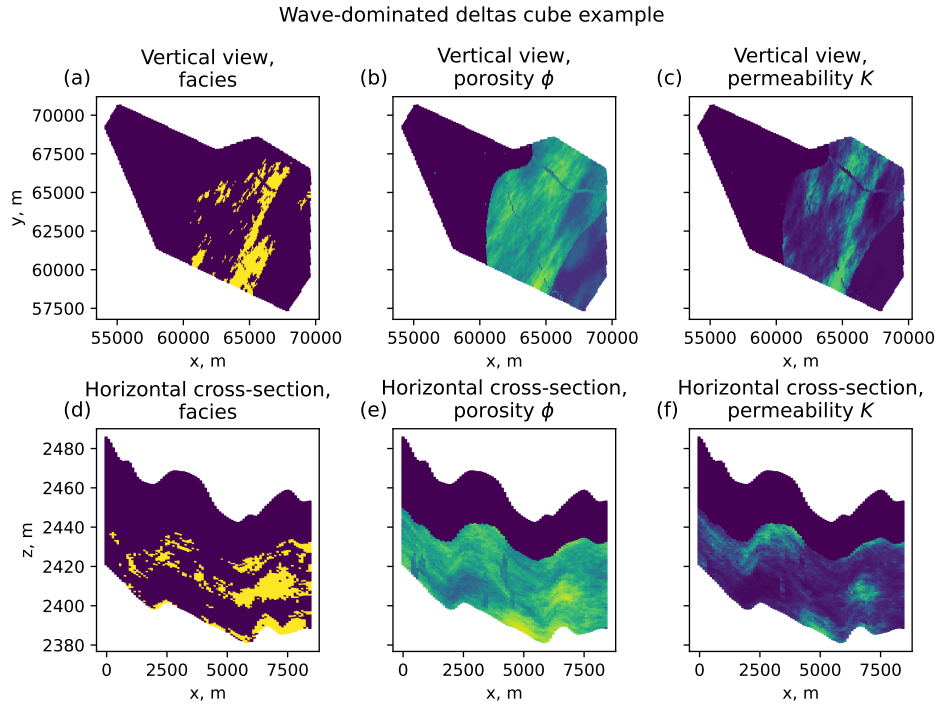

Figure 9: The example 3D cube from Wave-dominated deltas depositional environment.

A.2 Methods used

Facies modeling used methods of Sequential Indicator Simulation (SIS), Truncated Gaussian Simulation (TGS), and Object-based modeling.

Sequential Indicator Simulation (SIS) was used to model marine environments such as barrier island and shallow marine shelf. This method incorporates both horizontal and vertical trends, along with variogram parameter variation to reflect geological heterogeneity. The horizontal trend is closely tied to the geometry of sedimentary bodies and the prevailing depositional regime; its shape varies depending on the specific depositional environment, resulting in distinct model realizations. In this case, one horizontal trend and nine vertical trends were applied to capture the vertical variation in facies. The vertical trends were derived from analysis of lithological columns and spontaneous potential (SP) logs, reflecting the variation in sandstone content, which ranged from 65% to 100%. These trends represent upward increases or decreases in sand proportion typical of the modeled environment. In addition, variogram range parameters were varied to control the size and continuity of geological bodies, with major ranges from 400 to 10,000 meters and minor ranges from 100 to 5,000 meters, depending on the facies and depositional context.

Object-based modeling method was used for tidal flat environment, while combination of Truncated Gaussian Simulation (TGS) and Object-based modeling was used for wave-dominated delta.

Object-based modeling simulates geological bodies as discrete geometrical objects, such as channels, lobes, or bars. For tidal flat, there is only one object for modeling, which are tidal channels. Based on analysis of literature and analogue fields, boundary conditions for the channel facies were defined. The key input variables included: the proportion of reservoir to non-reservoir facies, the number of channels, their orientation, as well as amplitude, wavelength, width, and depth. Tidal flat shales were modeled as background facies as a low-permeable and non-reservoir matrix surrounding high-permeability features.

Truncated Gaussian simulation (TGS) is a geostatistical method used to model categorical variables, such as facies, by transforming continuous. For shelf transition zones from proximal to distal part, the TGS method was used. The core concept of TGS involves generating one or more Gaussian random fields with predefined spatial correlation structures (variograms). These continuous variables are then truncated based on a set of threshold values or lithotype rules, which assign each range of Gaussian values to a specific facies type. This allows the model to reflect both the spatial continuity of sedimentary bodies and the relative proportions and juxtapositions of facies. The result grid was used as background to model channels with a help of object modeling.

Porosity and permeability modeling was performed using upscaled synthetic porosity logs and the Sequential Gaussian Simulation (SGS) method, applied only to facies identified as reservoir rocks for all depositional environments. For non-reservoir facies, porosity values were assigned as zero. Each facies was assigned individual variogram ranges, consistent with those used in the facies model.

In addition, the input data was transformed into a normal distribution. The transformation parameters—mean, standard deviation (std), and minimum/maximum values—were calibrated based on well data, with the corresponding statistics on analogs fields. The modeling was carried out individually for each facies to ensure geological consistency.

For permeability modeling additional method co-located cokriging was used. It is a geostatistical technique used to model permeability by leveraging a secondary variable (e.g. porosity) as a spatial trend. The method estimates permeability at unsampled locations by combining sparse primary data (measured permeability) with a collocated, more densely sampled secondary variable (porosity), assuming a spatial correlation between the two. This allows for improved resolution and geological realism in permeability models, especially where direct permeability measurements are limited. applied with porosity as a secondary variable.

B Appendix B. Geological context

B.1 Global geological context variables

In this section we describe 20 total statistics that we have computed for each 3D cube. They have been used during supervised CNN training. They have also been used to estimate how meaningful are different representation in the Section 4.1.2.

B.1.1 FACIES STATISTICS

The facies fields are presented in Figures 6ad, 7ad, 8ad, 9ad. We compute 4 different statistics related to facies:

1. Abruption (X) - mean derivative of the facies along X axis.
2. Abruption (Y) - mean derivative of the facies along Y axis.
3. Abruption (Z) - mean derivative of the facies along Z axis.
4. Collectior Fraction - fraction of facies 1 in the cube.

B.1.2 LORENZ COEFFICIENT

Lorenz coefficient describes vertical reservoir heterogeneity (Pyrcz & Deutsch, 2014). We first compute it as 2D map for each cube. Then we compute the following 5 statistics over the obtained map:

1. Lorenz (Mean) - mean Lorenz coefficient for given 3D cube.
2. Lorenz (St.D.) - standard deviation of the Lorenz coefficient for given 3D cube.
3. Lorenz (Median) - median Lorenz coefficient for given 3D cube.
4. Lorenz (Min) - minimal Lorenz coefficient for given 3D cube.
5. Lorenz (Max) - maximal Lorenz coefficient for given 3D cube.

B.1.3 AVERAGE DERIVATIVES

To estimate horizontal heterogeneity, we compute average derivatives of the porosity ϕ and permeability K (see Figures 6, 7, 8, 9) over X, Y and Z axis, resulting in total 6 statistics:

1. Average $\frac{\partial \phi}{\partial x}$ - average derivative of porosity over X axis.
2. Average $\frac{\partial \phi}{\partial y}$ - average derivative of porosity over Y axis.
3. Average $\frac{\partial \phi}{\partial z}$ - average derivative of porosity over Z axis.
4. Average $\frac{\partial K}{\partial x}$ - average derivative of permeability over X axis.
5. Average $\frac{\partial K}{\partial y}$ - average derivative of permeability over Y axis.
6. Average $\frac{\partial K}{\partial z}$ - average derivative of permeability over Z axis.

In addition, the logarithm function is applied to these derivatives.

B.1.4 DYKSTRA-PARSONS COEFFICIENT

Just like Lorenz coefficient, Dykstra-Parsons coefficient also describes vertical reservoir heterogeneity (Pyrcz & Deutsch, 2014). It is also first computed as 2D map for the reservoir. We then compute another 5 statistics for each 3D cube.

1. D.P. (Mean) - mean Dykstra-Parsons coefficient for given 3D cube.
2. D.P. (St.D.) - standard deviation of the Dykstra-Parsons coefficient for given 3D cube.
3. D.P. (Median) - median Dykstra-Parsons coefficient for given 3D cube.
4. D.P. (Min) - minimal Dykstra-Parsons coefficient for given 3D cube.
5. D.P. (Max) - maximal Dykstra-Parsons coefficient for given 3D cube.