Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Reconstruction of 3D multiphase microstructures using denoising diffusion models

Ali Aouf

ali.aouf@sckcen.be

Université Catholique de Louvain

Eric Laloy

Belgian Nuclear Research Centre

Bart Rogiers

Belgian Nuclear Research Centre

Christophe De Vleeschower

Université Catholique de Louvain

Research Article

Keywords: 3D Microstructure Generation, Denoising Diffusion Models, Generative Adversarial Networks, Multiphase Materials, Computational Imaging, Generative AI

Posted Date: October 20th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7448759/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Computational Geosciences on June 6th, 2026. See the published version at https://doi.org/10.1007/s10596-026-10437-2.

Cover Letter

Reconstruction of 3D multiphase microstructures using denoising diffusion models

Ali Aouf, Eric Laloy, Bart Rogiers, Christophe De Vleeschouwer

Dear Editors-in-Chief,

please find the enclosed manuscript "Reconstruction of 3D multiphase microstructures using denoising diffusion models" which we are submitting for exclusive consideration for publication in Computational Geosciences. We confirm that the submission follows all the requirements and includes all the items of the submission checklist.

This manuscript considers the generation of multi-phase microstructures. It applies denoising diffusion probabilistic models (DDPMs) to the generation of heterogeneous materials such as clay and cement paste, and demonstrates that this class of deep-learning generative models is particularly well suited to the task, yielding high-quality, multi-phase samples with consistent structural descriptors. We further benchmark DDPMs against GAN variants and find that DDPMs provide a more stable training process and are significantly less prone to artifact generation.
We provide the source codes in a public repository with details listed in the section "Code availability".

Thanks for your consideration.

Sincerely,

Ali Aouf (1), Eric Laloy (2), Bart Rogiers (3), Christophe De Vleeschouwer (4)
1, 2, 3: {ali.aouf, eric.laloy, bart.rogiers}@sckcen.be, 4: christophe.devleeschouwer@uclouvain.be

Highlights

Reconstruction of 3D multiphase microstructures using denoising diffusion models

Ali Aouf, Eric Laloy, Bart Rogiers, Christophe De Vleeschouwer

- First application of DDPMs to 3D multiphase microstructure generation.
- Comparison of DDPMs with GAN-based approaches (WGAN-GP and iWGAN) for illite clay, Boom Clay, and concrete.
- DDPMs outperform GANs in capturing complex spatial statistics and phase distributions.

Reconstruction of 3D multiphase microstructures using denoising diffusion models

Ali Aouf[a], Eric Laloy[b], Bart Rogiers[c] and Christophe De Vleeschouwer[d]

[a] *ali.aouf@sckcen.be — ali.aouf@uclouvain.be*

ARTICLE INFO

Keywords:
3D Microstructure Generation
Denoising Diffusion Models
Generative Adversarial Networks
Multiphase Materials
Computational Imaging
Generative AI

ABSTRACT

The reconstruction of three-dimensional (3D) multiphase microstructures is essential for understanding the physical properties of porous materials. In this study, we evaluate the performance of Denoising Diffusion Probabilistic Models (DDPMs) and Generative Adversarial Networks (GANs), specifically WGAN-GP and iWGAN, in generating 3D representations of multiphase materials with varying degrees of heterogeneity: homogeneous illite clay, heterogeneous Boom Clay, and highly complex concrete. Our findings demonstrate that DDPMs outperform GANs in capturing the intricate spatial statistics and multiphase structures of these materials. While GAN-generated samples exhibit mode collapse and structures that do not resemble the ground truth, DDPMs produce microstructures that better preserve phase distributions and morphological characteristics. These results highlight the potential of diffusion models for realistic 3D microstructure synthesis, paving the way for improved simulations in subsurface and construction material applications.

CRediT authorship contribution statement

Ali Aouf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing - original draft. **Eric Laloy:** Conceptualization, Methodology, Formal analysis, Writing - review and editing. **Bart Rogiers:** Conceptualization, Methodology, Formal analysis, Writing - review and editing. **Christophe De Vleeschouwer:** Conceptualization, Methodology, Formal analysis, Writing - review and editing.

1. Introduction

The accurate reconstruction of three-dimensional (3D) multiphase microstructures is fundamental to understanding and predicting the physical properties of heterogeneous materials such as clays and concrete. These microstructures influence key material behaviors, including permeability and gas diffusivity, making them critical in various fields, including geoscience and materials science. While experimental imaging techniques such as Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) provide high-resolution images of these structures, they remain costly, time-consuming, and constrained by sample preparation limitations (Gaboreau et al. (2016); Sezen (2016)). Consequently, computational methods for synthetic microstructure generation would provide a valuable alternative to obtain high-fidelity 3D representations with controllable properties (Luo et al. (2024), Feng et al. (2019a)).

Some early efforts in microstructure generation relied on simulated annealing, to iteratively optimize spatial arrangements (Lemmens et al. (2019a); Cherkasov et al. (2021); Chen et al. (2022a); Karsanina and Gerke (2023)). Other works such as Chen et al. (2022b) applied simulated annealing to generate three-phase sandstone microstructures, successfully reconstructing basic spatial patterns but with limited applicability to heterogeneous materials (Chen

ORCID(s): 0009-0004-1820-9707 (.A. Aouf)

et al. (2022b); Lemmens et al. (2019b)).

Deep generative models (Mosser et al. (2017); Feng et al. (2019b); Henkes and Wessels (2022); Feng et al. (2020); Jung et al. (2020); Luo et al. (2024)) have been successfully applied to binary microstructures, effectively learning spatial correlations and pore structures. Convolutional GANs for instance were employed to generate 2D multi-phase images, achieving promising results in simple and homogeneous materials such as silica and Li-ion cathodes/SOFC anodes (Feng et al. (2019a), Gayon-Lombardo et al. (2020)). Despite these advancements, GAN-based methods remain constrained by challenges such as mode collapse, unstable training dynamics, and limited generalization to highly heterogeneous 3D structures.

An alternative generative approach that rapidly became the state-of-the-art image generative model is Denoising Diffusion Probabilistic Models (DDPMs). Unlike Generative Adversarial Networks (GANs), which employ an adversarial learning process where a generator competes against a discriminator, DDPMs adopt a probabilistic denoising approach. Instead of directly building a direct mapping between a latent space and the image space, a DDPM starts with pure Gaussian noise and gradually refines it through a series of denoising steps, ultimately producing a sample in the image space (Ho et al. (2020a); Nichol and Dhariwal (2021); Dhariwal and Nichol (2021)). This iterative refinement process allows DDPMs to avoid common pitfalls of GANs. In our previous work (Aouf et al. (2025)) we showed that DDPM outperforms classical GANs for the generation of a binary clay microstructure. However, the capability of DDPMs for multi-phase 3D microstructure generation—especially for highly heterogeneous materials—remains largely unexamined.

A particularly relevant development in diffusion-based multi-phase microstructure generation is the Latent Diffusion Model (LDM) proposed by Baishnab et al. (2025). Their work introduced conditioning on phase fraction, allowing the model to generate structures with controlled phase distributions (Baishnab et al. (2024)). While this represents an important advancement, their approach has two key limitations: (1) it was not tested on heterogeneous microstructures, and (2) it relied on an autoencoder, meaning that the generated quality depends on both the autoencoder's ability to compress and reconstruct latent data and the diffusion model's ability to learn meaningful representations.

In this study, we evaluate the performance of DDPMs (Ho et al. (2020a)) and two popular GAN-based approaches, namely WGAN-GP (Gulrajani et al. (2017)) and iWGAN (Chen et al. (2021)), in generating clay and cement paste 3D multiphase microstructures. We focus on three materials with increasing structural complexity: compacted illite clay, which is relatively homogeneous; Boom Clay, which exhibits a higher level of heterogeneity; and cement paste, which shows the highest level of heterogeneity. By comparing these models across structural fidelity, phase distribution accuracy, and training stability, this study provides new insights into the applicability of generative models for realistic multi-phase 3D material reconstruction.

The remainder of this paper is structured as follows. Section 2 describes the methodology and implementation

aspects, detailing the datasets, preprocessing steps, model architectures, and training procedures. Section 3 presents a comparative evaluation of the GAN and diffusion-based generative models based on qualitative and quantitative analyses. Section 4 discusses the results, including key observations, limitations, and potential applications of the proposed approach. Finally, Section 5 concludes the paper by summarizing the key findings and outlining directions for future research.

2. Methodology and Implementation Aspects

3. Methodology

To compare GANs and DDPMs, we implement two GAN variants and a U-Net-based DDPM. All models are trained using identical datasets and preprocessing pipelines to ensure fair comparison.

3.0.1. Generative Adversarial Networks (GANs)

Generative adversarial networks (GANs) are a class of deep generative models originally introduced by Goodfellow et al. (Goodfellow et al. (2014)). They consist of two competing neural networks: a generator that learns to produce synthetic samples and a discriminator that attempts to distinguish real samples from generated ones. Through this adversarial process, both networks continuously improve, leading to the generation of increasingly realistic outputs. GANs have been widely adopted for image synthesis and microstructure generation.

In this study, we adapt two advanced GAN variants—Wasserstein GAN with Gradient Penalty (WGAN-GP) and Inferential Wasserstein GAN (iWGAN)—to work with 3D volumetric data for multi-phase microstructure generation. These architectures were chosen for their improved training stability compared to a regular GAN.

3.0.2. Wasserstein GAN with Gradient Penalty (WGAN-GP)

Instead of a traditional discriminator, WGAN employs a critic Gulrajani et al. (2017), which assigns a realness score to generated samples rather than classifying them as “real” or “fake”. Specifically, the critic must be 1-Lipschitz—i.e. satisfy

$$\|\nabla_x D(x)\|_2 \leq 1 \quad \forall\, x,$$

in order to meet the Kantorovich–Rubinstein dual requirement for the Wasserstein-1 distance.Weight clipping was originally applied to enforce this by bounding each critic weight within a compact interval, but at the cost of reduced capacity and training pathologies.The gradient penalty in WGAN-GP instead softly penalizes deviations of $\|\nabla_{\hat{x}} D(\hat{x})\|_2$ from unity, leading to more stable critic behavior and better sample diversity Gulrajani et al. (2017).

For WGAN-GP, the critic loss function is defined as:

$$L_D = \mathbb{E}_{x\sim P_r}[D(x)] \; - \; \mathbb{E}_{z\sim P_z}[D(G(z))] \; + \; \lambda\, \mathbb{E}_{\hat{x}\sim P_{\hat{x}}}\Big(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1\Big)^2 \tag{1}$$

where $D(x)$ is the critic output, $G(z)$ is the generator output, P_r represents the real data distribution, P_z is the prior over the latent space, and λ is a penalty coefficient applied to reduce the gradient related to the training of the discriminator when its magnitude becomes larger than 1. This situation (large discriminator gradient) is likely to occur switching from training the generator to training the discriminator.

3.0.3. Inferential Wasserstein GAN (iWGAN)

While WGAN-GP improves training stability, mode collapse (where the generator produces only a limited subset of possible structures) remains a concern. To address this, Inferential Wasserstein GAN (iWGAN) (Chen et al. (2021)) extends the WGAN framework by incorporating an encoder network, which maps samples to latent representations. This mapping helps improve latent space organization. The iWGAN process can be seen in figure 1.

Unlike traditional GANs, where the generator maps latent vectors to synthetic samples, iWGAN introduces an encoder Q that estimates the latent code z from real data samples x, ensuring that the generator can faithfully reconstruct training distributions. The iWGAN training objective consists of three loss components:

1. Critic Loss (Wasserstein Distance):

$$L_D = \mathbb{E}_{x\sim P_r}[D(x)] - \mathbb{E}_{z\sim P_z}[D(G(z))] \tag{2}$$

2. Generator Loss (Minimizing Critic Score for Generated Samples):

$$L_G = -\mathbb{E}_{z\sim P_z}[D(G(z))] \tag{3}$$

3. Encoder Loss (Ensuring Consistency in the Latent Space):

$$L_Q = \mathbb{E}_{x\sim P_r}[\|x - G(Q(x))\|^2] \tag{4}$$

where P_r denotes the real–data distribution and P_z denotes the prior distribution over latent vectors z.

The total iWGAN loss function combines these components:

$$L_{\text{iWGAN}} = L_D + \lambda_G L_G + \lambda_Q L_Q \tag{5}$$

where λ_G and λ_Q are weighting factors that balance the contributions of each term. The encoder network is only used during training to enhance the quality of the learned latent space; inference remains identical to a standard GAN.

3.0.4. Denoising Diffusion Probabilistic Models (DDPMs)

Denoising Diffusion Probabilistic Models (DDPMs), introduced by Ho et al. (2020b), provide a powerful framework for generating high-fidelity synthetic data by learning a gradual denoising process. Unlike adversarial approaches such as GANs, which generate samples in a single pass, DDPMs iteratively refine an initial Gaussian noise sample through a series of denoising steps, progressively reconstructing complex structures. The diffusion process is formulated as a Markov chain of T steps, where Gaussian noise is incrementally added to a real data sample until the original structure is entirely obscured. This forward process is adopted to generate the training samples that will be considered to learn the denoising model, i.e. a model that is able to generate novel samples from arbitrary noise. The forward process works as follows:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \tag{6}$$

where $\mathbf{x}_t$ is the noised image at timestep t, $\mathbf{x}_{t-1}$ is the previous timestep, and β_t controls the amount of noise added. The sequence $\{\beta_t\}_{t=1}^{T}$ determines the rate of information loss throughout diffusion. The schedule refers to the specific way in which the noise levels β_t are arranged over time, shaping the difficulty and dynamics of both the forward and reverse processes. The two commonly used schedules are:

- **Linear schedule**: $\beta_t = \beta_0 + t(\beta_T - \beta_0)/T$
- **Cosine schedule**:

$$\overline{\alpha}_t = \frac{f(t)}{f(0)} \quad \text{where} \quad f(t) = \cos\left(\frac{t/T + s}{1+s} * \frac{\pi}{2}\right)^2 \tag{7}$$

$$\beta_t = 1 - \frac{\overline{\alpha}_t}{\overline{\alpha}_t - 1} \tag{8}$$

where t is the current timestep and T is the total number of timesteps (Nichol and Dhariwal (2021)). These schedules determine how rapidly the real structure is corrupted during the diffusion process.

In the reverse diffusion process, a deep neural network is trained to denoise the corrupted data, progressively reconstructing the original microstructure. We employ a 3D U-Net-based architecture (Ronneberger et al. (2015)) that

predicts the noise component ϵ, to learn the inverse function $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$, formulated as:

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t)) \tag{9}$$

where μ_θ and Σ_θ are the predicted mean and variance of the noise distribution. The network is trained using a mean squared error (MSE) loss between the predicted noise and the actual noise added during the forward diffusion process:

$$\mathcal{L} = \mathbb{E}_{\mathbf{x}_0,\epsilon,t}\left[||\epsilon - \epsilon_\theta(\mathbf{x}_t, t)||^2\right] \tag{10}$$

where $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ represents Gaussian noise.

The DDPM training and sampling pipelines are described in Algorithms 1 and 2. The training phase consists of learning to predict the added noise, while the sampling phase generates 3D microstructures by iteratively reversing the noise addition process. Figure 2 illustrates the noising and denoising processes.

Algorithm 1 Training a Diffusion Model

Require: Training dataset X, noise schedule $\{\beta_t\}_{t=1}^T$

1: **for** each training iteration **do**
2: Sample a mini-batch $\mathbf{x}_0 \sim X$
3: Sample $t \sim$ Uniform($\{1, \dots, T\}$)
4: Sample noise $\epsilon \sim \mathcal{N}(0, \mathbf{I})$
5: Compute noised data $\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\epsilon$, where $\alpha_t = 1 - \beta_t$
6: Update model parameters by optimizing the loss:

$$\mathcal{L} = \mathbb{E}_{\mathbf{x}_0,\epsilon,t}\left[||\epsilon - \epsilon_\theta(\mathbf{x}_t, t)||^2\right]$$

7: **end for**

Algorithm 2 Sampling from a Diffusion Model

Require: Trained model ϵ_θ, noise schedule $\{\beta_t\}_{t=1}^T$

1: Initialize $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$
2: **for** $t = T, T-1, \dots, 1$ **do**
3: Sample noise $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ if $t > 1$, else $\epsilon = 0$
4: Compute:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(\mathbf{x}_t, t)\right) + \sqrt{\beta_t}\epsilon$$

where $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$

5: **end forreturn** $\mathbf{x}_0$

3.0.5. Known limitations of GANs and DDPMs

Both GANs and DDPMs come with inherent challenges that impact their effectiveness.

Despite their ability to generate high-resolution samples efficiently, GAN-based models face several critical issues:

- **Training instability**: The adversarial nature of GANs can lead to highly unstable training, requiring careful hyperparameter tuning to prevent oscillating loss dynamics or generator-discriminator imbalances.
- **Mode collapse**: GANs sometimes fail to represent the full diversity of the training data, generating only a limited subset of possible structures. While iWGAN partially mitigates this, underrepresentation of certain phases remains an issue in multi-phase materials.

These challenges motivated the need for alternative generative approaches, such as DDPMs, which offer a more structured, progressive generation process.

DDPMs provide several advantages over GANs, including training stability and better long-range consistency. However, their iterative nature introduces several drawbacks:

- **Computational Cost**: Unlike GANs, which generate images in a single step, DDPMs require iterative denoising over thousands of steps, making them significantly slower and more computationally expensive.
- **Sampling Latency**: The multistep sampling process makes DDPMs impractical for real-time applications, especially for high-resolution 3D microstructures, though, this is not important in our application.

In summary, GANs offer faster sample generation but suffer from training instability and artifacts presence (Noguchi and Inoue (2021); Borji (2021)), while DDPMs provide an easier training process and cleaner generated images at the cost of higher computational demands and longer sampling times.

4. Experimental validation

4.1. Dataset and Preprocessing

To train and evaluate the considered generative models, we use three datasets with increasing levels of heterogeneity. The datasets consist of one 3D scan per material acquired using Focused Ion Beam Scanning Electron Microscopy (as thoroughly explained Sezen (2016)), which captures 2D slices of the samples that can be stacked to obtain the 3D image. The considered images are:

- **Compacted Illite Clay**: This image is of 1096x1095x182 dimensions and voxel size of 15nm x 15nm x 5nm. This microstructure is made of three phases (pore, clay matrix, minerals) and is relatively homogeneous and isotropic, with a subvolumes having very close phase proportions (Gaboreau et al. (2016)) (See figure 3a).

- **Boom Clay**: Boom Clay is a natural geological formation with a complex pore network (Honty et al. (2021); Zeelmaekers et al. (2015); Yu et al. (2013); Hemes et al. (2013)). The image dimensions are 670x670x670 with a voxel size of 5nm x 5nm x 5nm, and it also consists of three phases (pore, clay matrix, minerals) (See figure 3b).

- **Cement paste**: The image of size 887x1114x1519 and voxel size of 5nm x 5nm x 5nm and exhibits complex structure and artifacts from imaging, with four phases: pore, hydrated cement, unhydrated cement, and minerals. Figure 3c shows the full image. It is worth noting that due to the very minimal presence of the minerals phase, we decided to omit it since it negatively impacted the quality of the generated images.

Each image undergoes the same preprocessing pipeline to construct the training and validation datasets. However, due to the limited number of available horizontal 2D slices per 3D image and the fact that their sizes are below the Representative Elementary Volume (REV), both training and validation samples are extracted from the same original image. This validation dataset allows us to evaluate the models' ability to faithfully reconstruct the microstructural patterns observed in the provided images, without implying performance on other scans. Validation against external scans will be pursued in future work. The preprocessing pipeline is as follows:

1. **Segmentation:** Multi-threshold segmentation was applied to the SEM images to distinguish between multiple phases. By selecting representative points from each phase, the method enabled phase-specific clustering based on grayscale intensity.
2. **Downsampling**: The images are downsampled by a factor of two using averaging. The downampling is done to reduce voxel sparcity and include more structures in the extracted blocks.
3. **Multi-phase Encoding**: The multi-phase data is represented as a one-hot encoded tensor, allowing generative models to learn distinct phase distributions. This process augments the memory requirements, since a three-phase image is then represented as two cubes.
4. **3D Cropping**: To ensure coverage over the full-size scans, we extract random 3D patches of size 64^3. We avoided using rotation, stretching, and skewing to preserve the microstructures and their orientation.

4.2. Evaluation Methods

To assess the quality of the generated 3D multi-phase microstructures and compare the performance of different generative models, we employ a combination of visual analysis, statistical descriptors, phase-fraction comparison, Minkowski fonctionals, and deep-learning-based similarity measures. The evaluation methodology are explained as follows:

4.2.1. Visual Inspection

A first qualitative assessment is conducted by visualizing the generated samples in multiple formats, allowing for direct comparison of their spatial organization, phase distribution, and overall structural coherence, among the tested generative methods.

- A series of renders of 3D generated samples.
- A series of 2D slices cut through the generated volumes.

4.2.2. Quantitative metrics

To comprehensively evaluate generation quality, a set of different quantitative criteria are used to complement visual inspection.

Structural Descriptors We implemented three key statistical descriptors commonly used in microstructure characterization (see Torquato and Haslach (2002)):

- Two-Point Probability Function $S_2^{(i)}(r)$: This function quantifies the probability that two points, separated by a distance r, belong to the same phase i.
- Lineal Path Function $L^{(i)}(r)$: This function measures the probability that a linear segment of length r remains entirely within phase i, capturing phase connectivity and continuity.
- Connectivity Function $C_k(r)$: This function calculates the probability that two points, separated by a distance r, belong to the same connected cluster, providing insights into phases' cluster sizes.

These metrics are computed for randomly selected slices from the generated and ground truth samples, on all orientations (XY, XZ, and YZ) to capture anisotropic features. A lag distance of 32 voxels is used to analyze phase correlations over spatial scales. To optimize computation time, a 1D vector approximation method is applied, where we flatten any selected 2D slice into a 1D vector.

Minkowski Functional Analysis We employ Minkowski functionals, which are robust, additive, and translation-invariant morphological descriptors (See Mecke (2000); Ohser and Mücklich (2000); Michielsen and de Raedt (2001)). We implemented four functionals:

- **Volume**: Quantifies the total amount of space occupied by a given phase. It provides a basic measure of material composition.
- **Surface Area**: Measures the total interfacial area between different phases. It is particularly important for understanding surface-driven phenomena such as diffusion, reaction kinetics, or phase interactions.

- **Integrated Mean Curvature (Mean Breadth)**: Captures the average curvature and roughness of the phase interfaces. This functional is sensitive to the morphology of the microstructure, such as the presence of rounded or sharp features.

- **Euler Characteristic**: Describes the topological complexity of the structure, including the number of isolated regions, tunnels, and voids. It provides insight into the connectivity and percolation behavior of phases.

A single threshold measurement is insufficient for capturing the multiscale nature of complex microstructures. Therefore, we perform a *morphological sweep* by dilating the binary image with a spherical structuring element (ball) of increasing radius r. This generates a family of Minkowski functional curves,
$F(r)$ (with $F \in \{\text{Volume, Surface Area, Mean Breadth, Euler Characteristic}\}$),
which reveal:

- **Scale-Dependent Geometry** or how the microstructural features evolve with r.

- **Connectivity Transitions** as reflected by changes in the Euler characteristic indicating percolation or isolation.

- **Interface Complexity** as shown by variations in surface area and curvature that distinguish smooth from rough boundaries.

Phase Fraction Analysis The phase fractions of the generated and real microstructures are compared using ternary diagrams. Ternary diagrams, also known as triangular plots, are graphical representations that display the proportions of three components in a system that sum to a constant value, typically 100% or 1.0. Each point within the triangle represents a specific compositional mixture of the three phases, with the distance from each side proportional to the fraction of the corresponding component. These plots are valuable for visualizing and comparing the multi-phase compositional distributions between ground truth and generated microstructures.

Deep Learning-Based Similarity Metrics To further measure the similarity between real and generated microstructures, we employ deep-learning-based feature embeddings extracted using InceptionV3 (Szegedy et al. (2015)), a convolutional neural network trained on the ImageNet (Deng et al. (2009)) dataset. These embeddings capture high-level descriptions of structural characteristics as observed by a neural-network, enabling statistical comparisons between real and synthetic distributions. We used two methods to calculate the similarity between the generated samples and the ground truth:

- The Fréchet Inception Distance (FID) score quantifies the distance between the feature distributions of real and

generated images (Heusel et al. (2017)):

$$\text{FID} = \|\mu_r - \mu_g\|^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{1/2}), \tag{11}$$

where μ_r, Σ_r are the mean and covariance of feature embeddings from real microstructures, and μ_g, Σ_g are the same statistics computed for generated samples. Lower FID scores indicate a closer resemblance between real and synthetic data.

- The cosine similarity between two feature vectors **A** and **B** is given by:

$$\text{Cosine Similarity} = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}. \tag{12}$$

This metric ranges from -1 to 1, where 1 indicates identical feature orientations, 0 indicates orthogonality, and -1 indicates completely dissimilar structures. By constructing a Cosine Similarity Matrix, we analyze the pairwise relationships between real and generated microstructures.

Whereas the FID score captures both mean (first-order) differences and covariance (second-order) differences in the feature space, cosine similarity directly evaluates the alignment between individual feature vectors, focusing on the direction of feature embeddings rather than their magnitude or distribution. By combining both metrics, we should obtain a comprehensive evaluation of the generated samples quality.

5. Results and Discussion

This section presents and analyzes the results of the multi-phase 3D microstructure generation using our considered GANs (WGAN-GP, iWGAN) and DDPMs. As written above, the performance of each model is evaluated using visual inspection together with a series of qualitative metrics related to structural fidelity, phase distribution accuracy and statistical similarities between real and generated samples.

5.1. Visual Inspection

Figures 4, 5, and 6 illustrate a 3D side-by-side comparison between the real and generated samples for compacted illite, Boom Clay, and cement paste, respectively. Each column presents a set of samples produced in the same way (either by cropping the real sample, or by sampling using one specific generative model with different input noise z). As a consequence, the similarity should be envisioned across set of samples. From visual assessement, it is observed that DDPM samples more closely resemble the ground truth sample across the three tested microstructures, compared to the GAN based models. WGAN-GP samples though looking noisy, tend to also show some resemblance to the

ground truth structures, albeit with some samples being total noise. While the iWGAN realizations look less noisy, their structures resemble the ground truth structures the less among the tested models.

Figures 7, 8, and 9, compare ground truth illite, Boom Clay, and cement paste image slices to generated ones. The DDPM-generated slices (b) across the three tested lithologies capture much of the large-scale structure and connectivity, though we observe a less frequent appearance of medium to small size pores and minerals in the boom clay samples compared to the ground truth. In contrast, WGAN-GP (c) often introduces higher noise levels, yielding speckle-like artifacts and fragmentations that obscure finer features, yet, its slices resemble the ground truth better than those produced by the iWGAN. Indeed, the iWGAN (d) results display low fidelity across all microstructures, with phases forming smaller pores or subtle grains in illite and boom clay samples, and unrealistic structures for cement paste.

5.2. Structural descriptors analysis

The structural descriptors provide quantitative information about spatial organization and phase connectivity observed in the qualitative analysis. Figures 10, 11, and 12 present the two-point probability function, lineal path function, and connectivity function for each considered material.

The results indicate that DDPMs structural descriptors align with that of the ground truth in most cases, indicating a preserved phase continuity across all materials. The two-point probability function analysis reveals that DDPM-generated microstructures honors phase spatial correlation better than their GAN-generated counterparts. The lineal path function shows that continuous pathways are more frequently preserved in DDPM samples, especially for the pore phase. WGAN-GP and iWGAN samples, on the other hand, display a decent alignment in some cases but with some inconsistencies and divergences for compacted illite and cement paste, where lineal paths were occasionally interrupted by artifacts in phase transitions. Furthermore, we notice that iWGAN consistently shows lower alignment with the ground truth.

5.3. Minkowski functionals analysis

Overall, the Minkowski functional curves (Figures 13, 14, 15) reveal that each generative model reproduces different aspects of the microstructure with varying degrees of fidelity, but the DDPM appears to capture bulk morphological characteristics (such as overall porosity and average pore size) more consistently across most dilation radii. This is evident from the “Volume vs. Radius” and “Mean Breadth vs. Radius” plots, where DDPM’s curves remain closer to the ground truth for both the pore phase and solid phases, indicating a more accurate replication of large-scale features and connectivity. The WGAN and IWGAN methods, while often competitive, show more pronounced deviations in specific radius ranges, suggesting either over or under-estimation of fine-scale pore throats or interface roughness.

5.4. Phase Fraction analysis

A key requirement for generating realistic microstructures is accurately reproducing phase fractions. Figure 16 uses ternary diagrams to compare phase distributions for compacted illite, Boom Clay, and cement paste across different generative models. The ternary plots reveal distinct performance patterns for each approach. DDPM consistently generates phase compositions that cluster closest to the ground truth across all three materials. While maintaining similar density and distribution patterns, DDPM exhibits increased variance compared to ground truth, particularly noticeable in compacted illite and Boom Clay samples. Despite this deviation, DDPM demonstrates the strongest compositional fidelity overall. WGAN-GP shows variable performance across materials. For compacted illite, it achieves decent overlap with ground truth data, though with noticeably wider scatter. In Boom Clay samples, some WGAN-GP points fall within the ground truth cluster, but the center of mass shifts away from the reference distribution. Cement paste results of were dispersed extensively across the ternary space, indicating poor compositional control. iWGAN displays a strong bias toward clay matrix generation in clay-based materials, failing to capture the compositional diversity present in ground truth samples. For cement paste, iWGAN produces distributions nearly as scattered as WGAN-GP, indicating similar difficulties in maintaining realistic phase relationships. These ternary comparisons demonstrate that, while accurately replicating the characteristics of multi-phase materials remains challenging for all generative approaches, DDPMs clearly outperform GAN-based methods in reproducing realistic phase distribution.

5.5. Similarity assessment using deep learning embeddings

A feature-based similarity assessment was conducted using deep-learning embeddings from InceptionV3. Table 1 presents the Fréchet Inception Distance (FID) and Cosine Similarity scores for each model.

The results indicate that DDPMs consistently achieved the lowest FID scores across all material strucutres, confirming that the generated microstructures closely align with the real data distribution. Lower FID scores suggest that the overall feature statistics of DDPM-generated samples resemble those of the real microstructures more closely than those of GAN-based models. Cosine similarity analysis further supports this observation, with DDPMs achieving the highest similarity scores to real samples, followed by iWGAN and WGAN-GP.

The relatively higher FID scores of GANs indicate that they introduce greater variation in structural features compared to DDPMs. This variation may stem from the artifacts in the generated samples or misalignment in their structures compared to the ground truth data.

6. Conclusions

This study evaluats the performance of generative models(DDPMs, WGAN-GP, and iWGAN) in reconstructing 3D multi-phase microstructures of compacted illite, Boom Clay, and cement paste. Our results demonstrate that, for

Table 1
Comparison of DDPM, WGAN-GP, and iWGAN on FID Score and Cosine Similarity across different materials. Lower FID scores indicate better structural similarity, while higher Cosine Similarity scores (closer to 1) suggest greater feature alignment.

Material	Model	FID Score ↓	Cosine Similarity ↑
Compacted Illite	DDPM	**2.8**	**0.998**
	WGAN-GP	4.7	0.975
	iWGAN	5.9	0.978
Boom Clay	DDPM	**12.9**	**0.99**
	WGAN-GP	15.7	0.99
	iWGAN	21.6	0.89
Cement Paste	DDPM	**2.09**	**0.998**
	WGAN-GP	11.7	0.984
	iWGAN	4.4	0.994

the considered materials, DDPMs consistently outperform GANs in preserving structural fidelity, phase connectivity, and statistical similarity to real microstructures. The diffusion-based approach effectively maintains long-range spatial correlations and accurately reproduces phase fractions, making it particularly suitable for complex heterogeneous materials.

While the tested GANs, particularly WGAN-GP, produced sharp phase boundaries, they also exhibited phase discontinuities and inconsistencies in highly heterogeneous systems.

Overall, DDPMs appears to be the most reliable option for generating realistic 3D multi-phase microstructures, particularly for applications requiring detailed structural preservation. Future research could explore hybrid approaches that integrate the accuracy of diffusion and the flexibility of adversarial training to further enhance generation quality and efficiency.

Code availability section

All models are trained on an NVIDIA A100 Tensor Core GPU, the software packages used are:

- Source code: https://github.com/40uf411/3DMPDDPM
- Python 3.10 as a programming language.
- Training time of about 24h.
- PyTorch for models implementation, training, and sampling.
- Numpy for samples representation.
- scikit-image and OpenCV for image processing.
- CherryMSA for samples analysis and visualization.

Statements and Declarations Christophe De Vleeschouwer is a Research Director of the Fonds de la Recherche Scientifique – FNRS.

References

Aouf, A., Laloy, E., Rogiers, B., De Vleeschouwer, C., 2025. 3d clay microstructure synthesis using denoising diffusion probabilistic models. Available at SSRN: https://ssrn.com/abstract=4814452 or http://dx.doi.org/10.2139/ssrn.4814452.

Baishnab, N., Herron, E., Balu, A., Sarkar, S., Krishnamurthy, A., Ganapathysubramanian, B., 2024. 3d multiphase heterogeneous microstructure generation using conditional latent diffusion models, in: AI4Mat-2024: NeurIPS 2024 Workshop on AI for Accelerated Materials Design, pp. –.

Baishnab, N., Herron, E., Balu, A., Sarkar, S., Krishnamurthy, A., Ganapathysubramanian, B., 2025. 3d multiphase heterogeneous microstructure generation using conditional latent diffusion models. URL: https://arxiv.org/abs/2503.10711, arXiv:2503.10711.

Borji, A., 2021. Pros and cons of GAN evaluation measures: New developments. CoRR abs/2103.09396. URL: https://arxiv.org/abs/2103.09396, arXiv:2103.09396.

Chen, D., Xu, Z., Wang, X., He, H., Du, Z., Nan, J., 2022a. Fast reconstruction of multiphase microstructures based on statistical descriptors. Phys. Rev. E 105, 055301. URL: https://link.aps.org/doi/10.1103/PhysRevE.105.055301, doi:10.1103/PhysRevE.105.055301.

Chen, D., Xu, Z., Wang, X., He, H., Du, Z., Nan, J., 2022b. Fast reconstruction of multiphase microstructures based on statistical descriptors. Phys. Rev. E 105, 055301. URL: https://link.aps.org/doi/10.1103/PhysRevE.105.055301, doi:10.1103/PhysRevE.105.055301.

Chen, Y., Gao, Q., Wang, X., 2021. Inferential wasserstein generative adversarial networks. URL: https://arxiv.org/abs/2109.05652, arXiv:2109.05652.

Cherkasov, A., Ananev, A., Karsanina, M., Khlyupin, A., Gerke, K., 2021. Adaptive phase-retrieval stochastic reconstruction with correlation functions: Three-dimensional images from two-dimensional cuts. Phys. Rev. E 104, 035304. URL: https://link.aps.org/doi/10.1103/PhysRevE.104.035304, doi:10.1103/PhysRevE.104.035304.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. doi:10.1109/CVPR.2009.5206848.

Dhariwal, P., Nichol, A., 2021. Diffusion models beat gans on image synthesis. Advances in neural information processing systems 34, 8780–8794.

Feng, J., He, X., Teng, Q., Ren, C., Chen, H., Li, Y., 2019a. Reconstruction of porous media from extremely limited information using conditional generative adversarial networks. Phys. Rev. E 100, 033308. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.033308, doi:10.1103/PhysRevE.100.033308.

Feng, J., He, X., Teng, Q., Ren, C., Chen, H., Li, Y., 2019b. Reconstruction of porous media from extremely limited information using conditional generative adversarial networks. Phys. Rev. E 100, 033308. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.033308, doi:10.1103/PhysRevE.100.033308.

Feng, J., Teng, Q., Li, B., He, X., Chen, H., Li, Y., 2020. An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning. Computer Methods in Applied Mechanics and Engineering 368, 113043. URL: https://www.sciencedirect.com/science/article/pii/S0045782520302279, doi:https://doi.org/10.1016/j.cma.2020.113043.

Gaboreau, S., Robinet, J.C., Prêt, D., 2016. Optimization of pore-network characterization of a compacted clay material by tem and fib/sem imaging. Microporous and Mesoporous Materials 224, 116–128. URL: https://www.sciencedirect.com/science/article/pii/S138718111500640X, doi:https://doi.org/10.1016/j.micromeso.2015.11.035.

Gayon-Lombardo, A., Mosser, L., Brandon, N.P., Cooper, S.J., 2020. Pores for thought: generative adversarial networks for stochastic reconstruction of 3D multi-phase electrode microstructures with periodic boundaries. npj Computational Materials 6, 82. URL: https://doi.org/10.1038/s41524-020-0340-7, doi:10.1038/s41524-020-0340-7.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial networks. arXiv:1406.2661.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C., 2017. Improved training of wasserstein gans. CoRR abs/1704.00028. URL: http://arxiv.org/abs/1704.00028, arXiv:1704.00028.

Hemes, S., Desbois, G., Urai, J., De Craen, M., Honty, M., 2013. Variations in the morphology of porosity in the boom clay formation: insights from 2d high resolution bib-sem imaging and mercury injection porosimetry. Netherlands Journal of Geosciences - Geologie en Mijnbouw 92, 275–300. doi:10.1017/S0016774600000214.

Henkes, A., Wessels, H., 2022. Three-dimensional microstructure generation using generative adversarial neural networks in the context of continuum micromechanics. Computer Methods in Applied Mechanics and Engineering 400, 115497. URL: http://dx.doi.org/10.1016/j.cma.2022.115497, doi:10.1016/j.cma.2022.115497.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Klambauer, G., Hochreiter, S., 2017. Gans trained by a two time-scale update rule converge to a nash equilibrium. CoRR abs/1706.08500. URL: http://arxiv.org/abs/1706.08500, arXiv:1706.08500.

Ho, J., Jain, A., Abbeel, P., 2020a. Denoising diffusion probabilistic models. CoRR abs/2006.11239. URL: https://arxiv.org/abs/2006.11239, arXiv:2006.11239.

Ho, J., Jain, A., Abbeel, P., 2020b. Denoising diffusion probabilistic models. CoRR abs/2006.11239. URL: https://arxiv.org/abs/2006.11239, arXiv:2006.11239.

Honty, M., Frederickx, L., Wang, L., De, C., Thomas, P., Moors, H., Jacops, E., 2021. Boom clay pore water geochemistry at the mol site: Experimental data as determined by in situ sampling of the piezometers. Applied Geochemistry 136, 105156. doi:10.1016/j.apgeochem.2021.105156.

Jung, J., Yoon, J.I., Park, H.K., Jo, H., Kim, H.S., 2020. Microstructure design using machine learning generated low dimensional and continuous design space. Materialia 11, 100690. URL: https://www.sciencedirect.com/science/article/pii/S2589152920301071, doi:https://doi.org/10.1016/j.mtla.2020.100690.

Karsanina, M.V., Gerke, K.M., 2023. Stochastic (re)constructions of non-stationary material structures: Using ensemble averaged correlation functions and non-uniform phase distributions. Physica A: Statistical Mechanics and its Applications 611, 128417. URL: https://www.sciencedirect.com/science/article/pii/S037843712200975X, doi:https://doi.org/10.1016/j.physa.2022.128417.

Lemmens, L., Rogiers, B., Jacques, D., Huysmans, M., Swennen, R., Urai, J.L., Desbois, G., Laloy, E., 2019a. Nested multiresolution hierarchical simulated annealing algorithm for porous media reconstruction. Phys. Rev. E 100, 053316. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.053316, doi:10.1103/PhysRevE.100.053316.

Lemmens, L., Rogiers, B., Jacques, D., Huysmans, M., Swennen, R., Urai, J.L., Desbois, G., Laloy, E., 2019b. Nested multiresolution hierarchical simulated annealing algorithm for porous media reconstruction. Phys. Rev. E 100, 053316. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.053316, doi:10.1103/PhysRevE.100.053316.

Luo, X., Sun, J., Zhang, R., Chi, P., Cui, R., 2024. A multi-condition denoising diffusion probabilistic model controls the reconstruction of 3d digital rocks. Computers and Geosciences 184, 105541. URL: https://www.sciencedirect.com/science/article/pii/S0098300424000244, doi:https://doi.org/10.1016/j.cageo.2024.105541.

Mecke, K.R., 2000. Additivity, convexity, and beyond: Applications of minkowski functionals in statistical physics, in: Mecke, K.R., Stoyan, D. (Eds.), Statistical Physics and Spatial Statistics, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 111–184.

Michielsen, K., de Raedt, H., 2001. Integral-geometry morphological image analysis. Physics Reports 347, 462 – 538. doi:10.1016/S0370-1573(00)00106-X. journal JUL 450EE PHYS REP-REV SECT PHYS LETT.

Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. E 96, 043309. URL: https://link.aps.org/doi/10.1103/PhysRevE.96.043309, doi:10.1103/PhysRevE.96.043309.

Nichol, A.Q., Dhariwal, P., 2021. Improved denoising diffusion probabilistic models, in: International Conference on Machine Learning, PMLR. pp. 8162–8171.

Noguchi, S., Inoue, J., 2021. Stochastic characterization and reconstruction of material microstructures for establishment of process-structure-property linkage using the deep generative model. Phys. Rev. E 104, 025302. URL: https://link.aps.org/doi/10.1103/PhysRevE.104.025302, doi:10.1103/PhysRevE.104.025302.

Ohser, J., Mücklich, F., 2000. Statistical analysis of microstructures in materials science. Chichester : Wiley. URL: http://lib.ugent.be/catalog/rug01:000818718.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597. URL: http://arxiv.org/abs/1505.04597, arXiv:1505.04597.

Sezen, M., 2016. Focused ion beams (fib) — novel methodologies and recent applications for multidisciplinary sciences, in: Janecek, M., Kral, R. (Eds.), Modern Electron Microscopy in Physical and Life Sciences. IntechOpen, Rijeka. chapter 6, pp. –. URL: https://doi.org/10.5772/61634, doi:10.5772/61634.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2015. Rethinking the inception architecture for computer vision. CoRR abs/1512.00567. URL: http://arxiv.org/abs/1512.00567, arXiv:1512.00567.

Torquato, Haslach, 2002. Random Heterogeneous Materials: Microstructure and Macroscopic Properties. Applied Mechanics Reviews 55, B62–B63. URL: https://doi.org/10.1115/1.1483342, doi:10.1115/1.1483342, arXiv:https://asmedigitalcollection.asme.org/appliedmechanicsreviews/article-pdf/55/4/B62/5439035/b61_1.pdf.

Yu, L., Rogiers, B., Gedeon, M., Marivoet, J., Craen, M.D., Mallants, D., 2013. A critical review of laboratory and in-situ hydraulic conductivity measurements for the boom clay in belgium. Applied Clay Science 75-76, 1–12. URL: https://www.sciencedirect.com/science/article/pii/S0169131713000446, doi:https://doi.org/10.1016/j.clay.2013.02.018.

Zeelmaekers, E., Honty, M., Derkowski, A., Środoń, J., De Craen, M., Vandenberghe, N., Adriaens, R., Ufer, K., Wouters, L., 2015. Qualitative and quantitative mineralogical composition of the rupelian boom clay in belgium. Clay Minerals 50, 249–272. doi:10.1180/claymin.2015.050.2.08.

List of Figures

1 This figure depicts the iWGAN framework, where an encoder maps ground truth data to an embedding, a generator reconstructs it, and a discriminator evaluates the realism score to improve generation quality. 21
2 Illustration of the DDPM process. The forward diffusion gradually corrupts an image by adding noise at each step, while the reverse process reconstructs the original structure by denoising in multiple stages. 22
3 The original 3D samples of compacted illite (Left sub-figure) and Boom Clay (Middle sub-figure) and Cement paste (Right sub-figure). 23
4 Visual comparison of ground truth compacted illite samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores. . . . 24
5 Visual comparison of ground truth Boom Clay samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores. 25
6 Visual comparison of ground truth cement paste samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores. 26
7 Comparison of multi-phase grid-slices of ground truth compacted illite to DDPM, WGAN-GP, and iWGAN generated ones. 27
8 Comparison of multi-phase grid-slices of ground truth Boom Clay to DDPM, WGAN-GP, and iWGAN generated ones. 28
9 Comparison of multi-phase grid-slices of ground truth cement paste to DDPM, WGAN-GP, and iWGAN generated ones. 29
10 Comparison of structural descriptor functions for the three phases of multi-phase compacted illite. Each row shows one phase; top to bottom: **Pores**, **Clay**, **Mineral**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability). 30
11 Comparison of structural descriptor functions for the three phases of multi-phase Boom Clay. Each row shows one phase; top to bottom: **Pores**, **Clay**, **Mineral**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability). 31
12 Comparison of structural descriptor functions for the three phases of multi-phase cement paste. Each row shows one phase; top to bottom: **Pores**, **Unhydrated cement**, **Hydrated cement**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability). 32
13 Multiscale Minkowski Functional Analysis of compacted illite. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink). 33
14 Multiscale Minkowski Functional Analysis of Boom Clay. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink). 34

15 Multiscale Minkowski Functional Analysis of cement paste. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink). 35

16 Ternary diagrams of materials composition. Each row shows one material; top to bottom: Compacted Illite, Boom Clay, Cement Paste. Within each ternary diagram, the points represent compositional distributions of the three phases, with colors indicating different data sources: ground truth samples (green), and microstructures generated by DDPM (orange), WGAN-GP (pink), and iWGAN (blue). The position of each point within the triangle represents the relative proportions of the three phases, with clustering patterns revealing how accurately each generative model reproduces the compositional constraints of the reference materials. 36

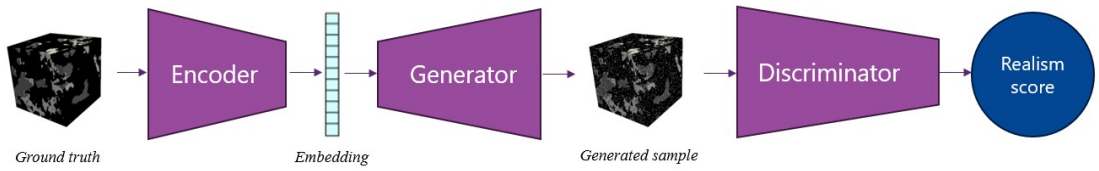


Figure 1: This figure depicts the iWGAN framework, where an encoder maps ground truth data to an embedding, a generator reconstructs it, and a discriminator evaluates the realism score to improve generation quality.

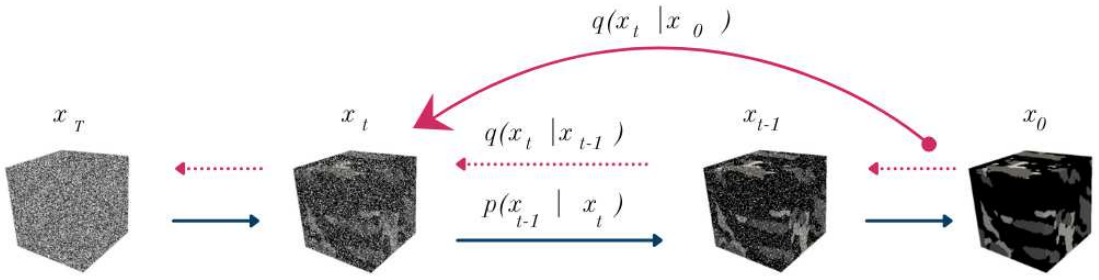
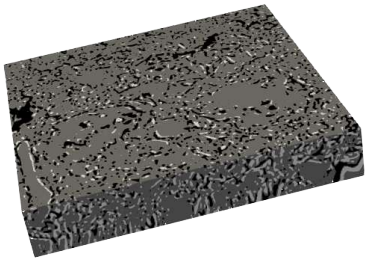


Figure 2: Illustration of the DDPM process. The forward diffusion gradually corrupts an image by adding noise at each step, while the reverse process reconstructs the original structure by denoising in multiple stages.

(a) Original 3D sample of compacted illite



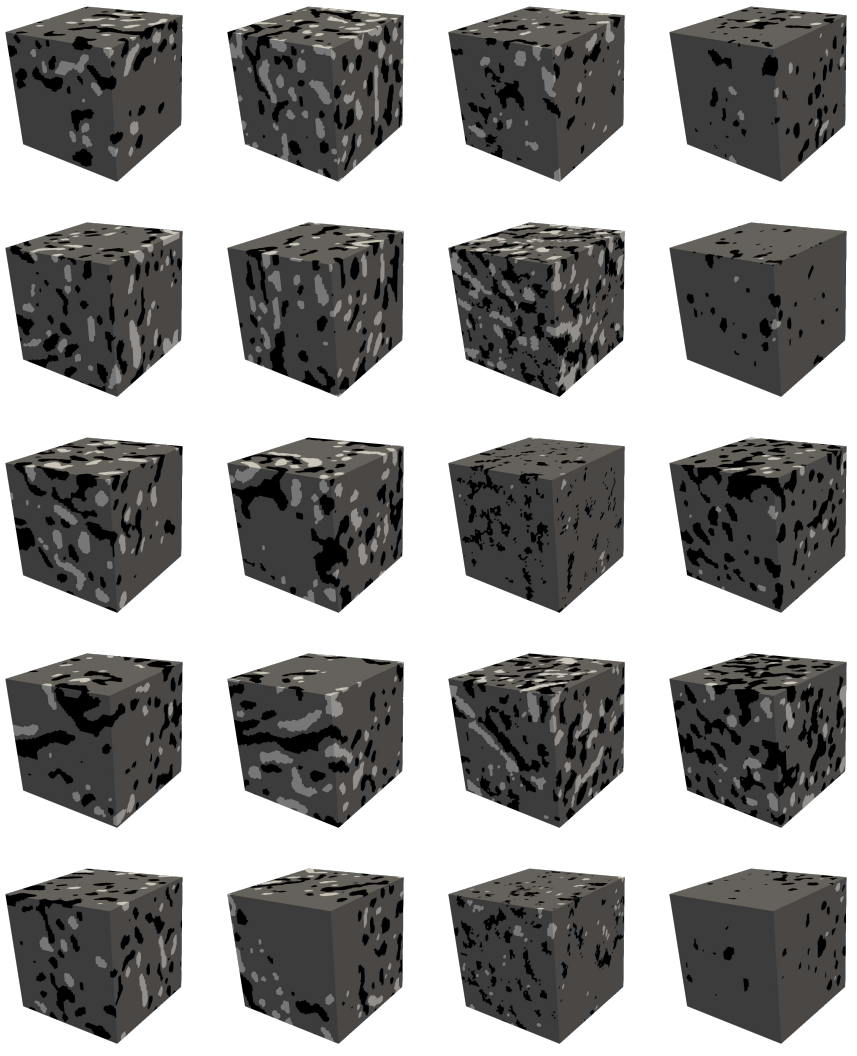
(b) Original 3D sample of Boom Clay



(c) Original 3D sample of Cement paste

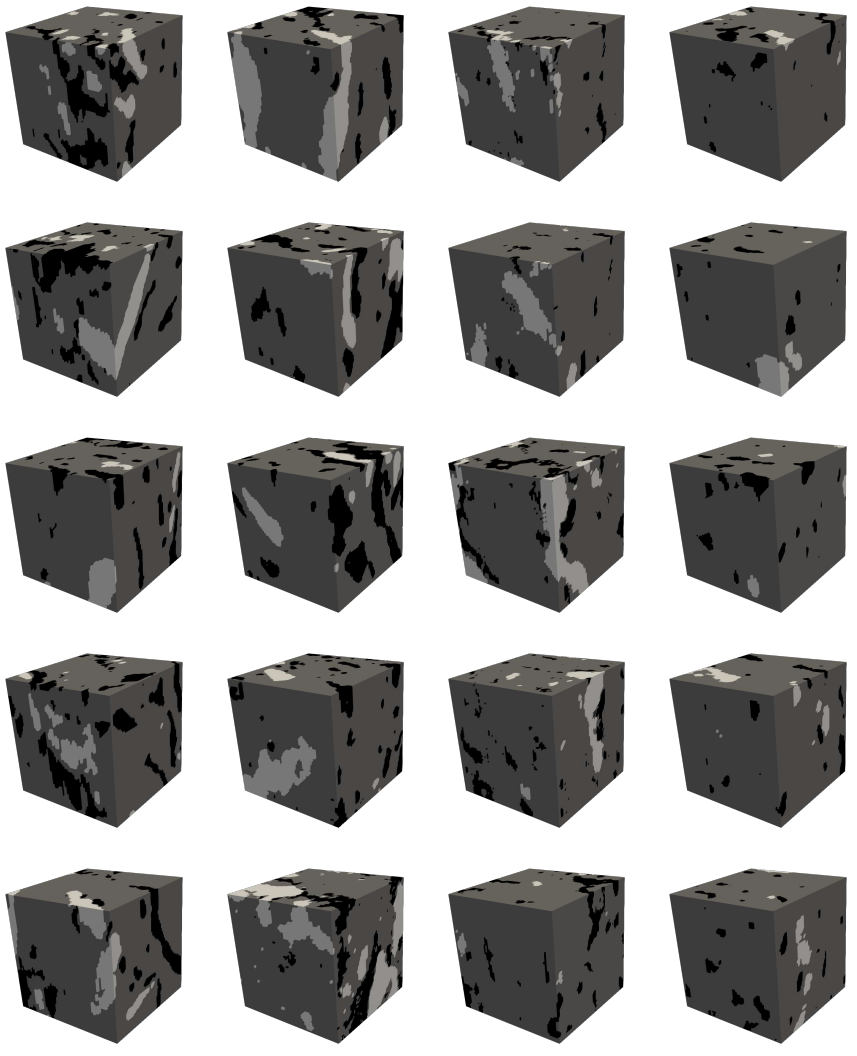


Figure 3: The original 3D samples of compacted illite (Left sub-figure) and Boom Clay (Middle sub-figure) and Cement paste (Right sub-figure).



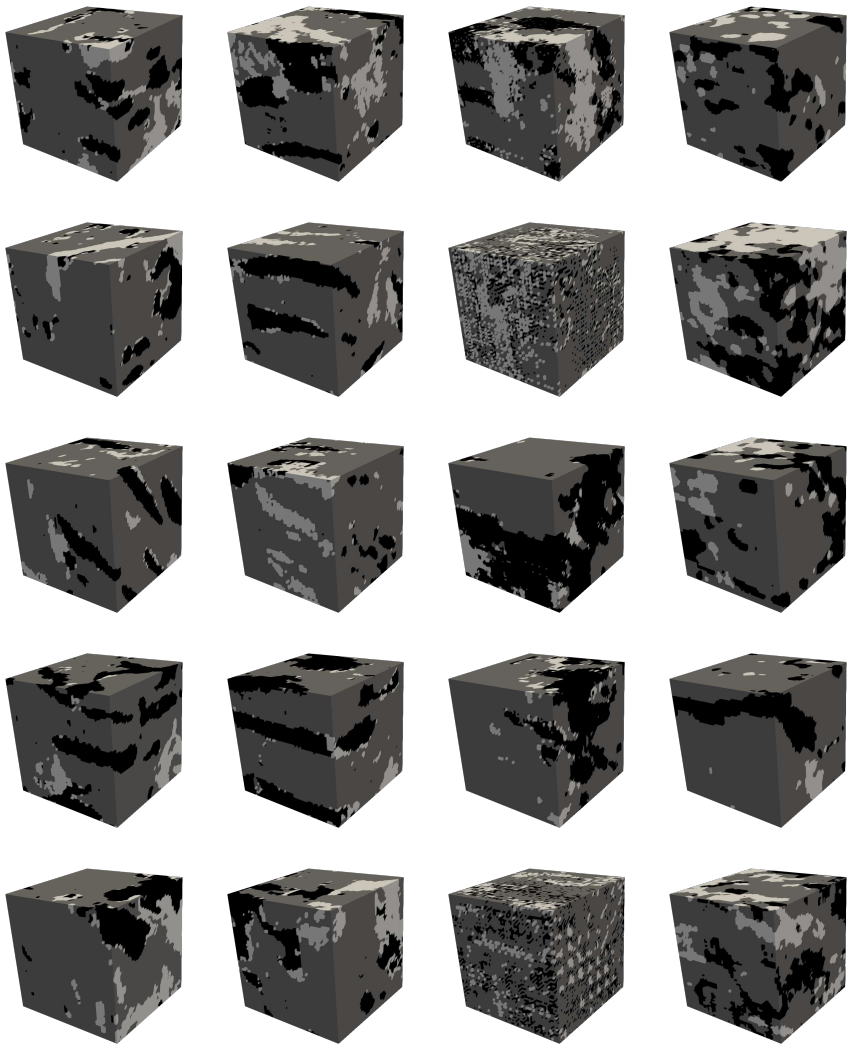
(a) Ground Truth (b) DDPM (c) WGAN-GP (d) iWGAN

Figure 4: Visual comparison of ground truth compacted illite samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores.



(a) Ground Truth (b) DDPM (c) WGAN-GP (d) iWGAN

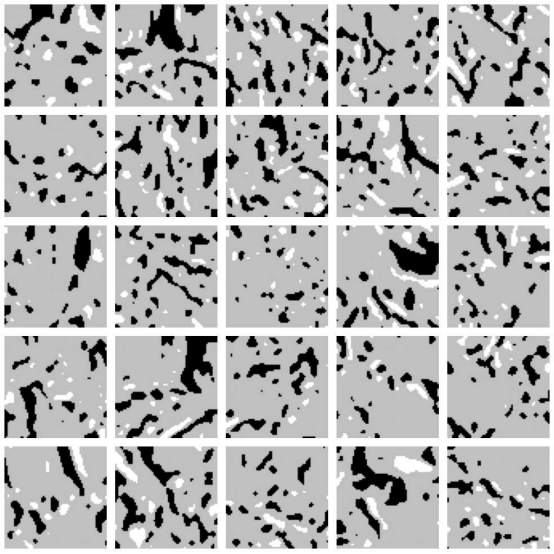
Figure 5: Visual comparison of ground truth Boom Clay samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores.



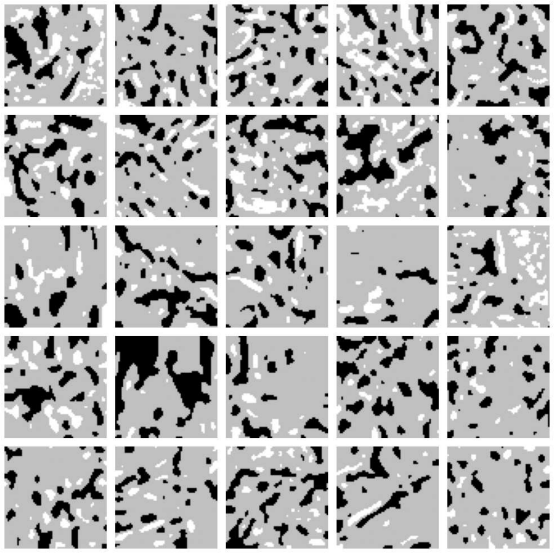
(a) Ground Truth (b) DDPM (c) WGAN-GP (d) iWGAN

Figure 6: Visual comparison of ground truth cement paste samples to DDPM, WGAN-GP, and iWGAN generated samples. Black represents the play matrix, white are the minerals, and gray are the pores.

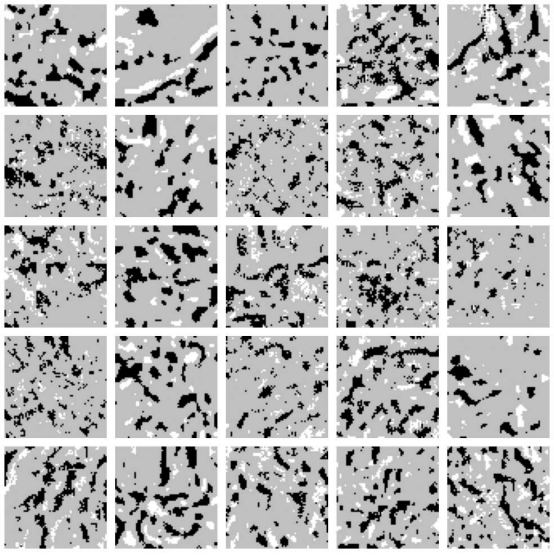
(a) Ground truth multi-phase slices of compacted illite



(b) Multi-phase slices of DDPM-generated compacted illite



(c) Multi-phase slices of WGAN-GP-generated compacted illite



(d) Multi-phase slices of iWGAN-generated compacted illite

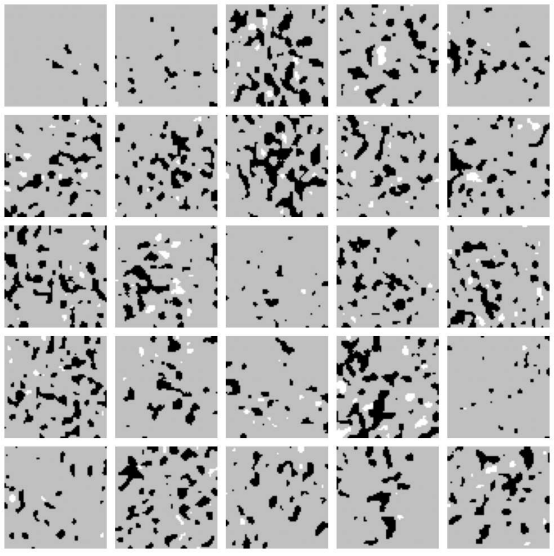
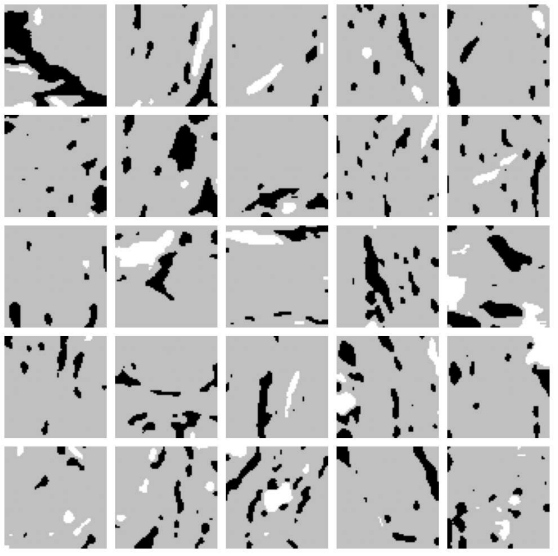
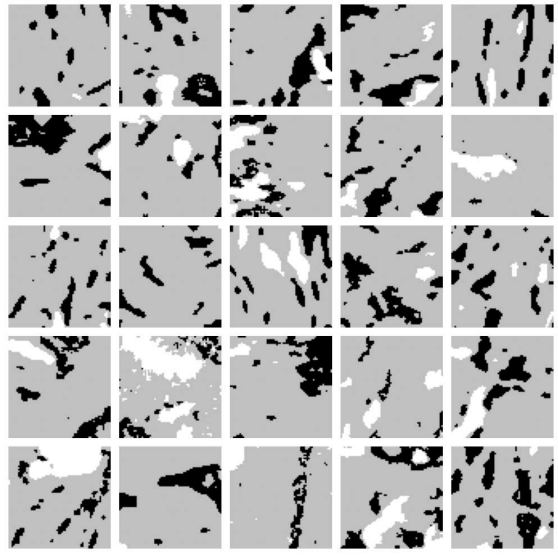


Figure 7: Comparison of multi-phase grid-slices of ground truth compacted illite to DDPM, WGAN-GP, and iWGAN generated ones.

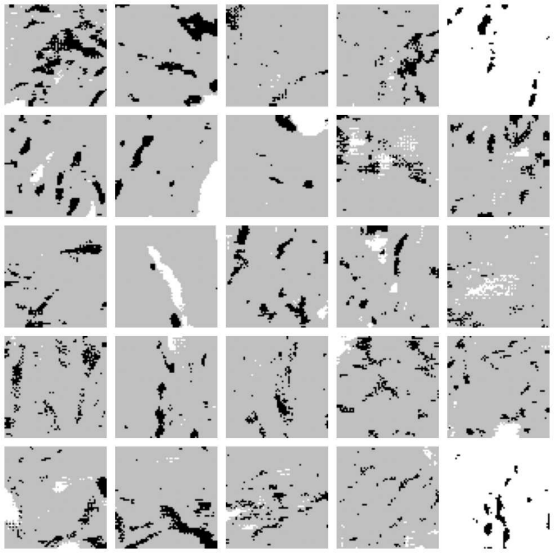
(a) Ground truth multi-phase slices of Boom Clay



(b) Multi-phase slices of DDPM-generated Boom Clay



(c) Multi-phase slices of WGAN-GP-generated Boom Clay



(d) Multi-phase slices of iWGAN-generated Boom Clay

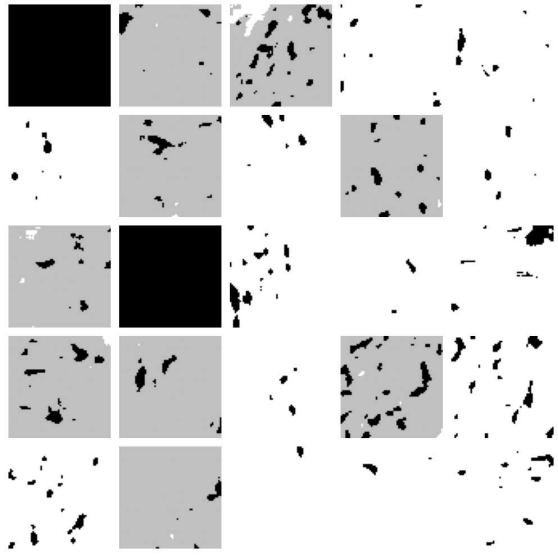
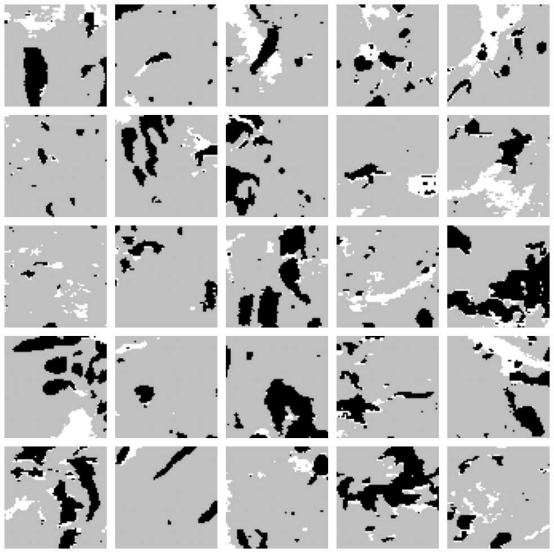
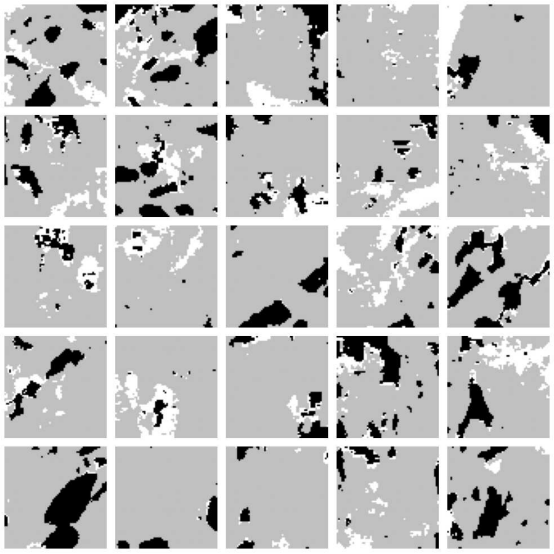


Figure 8: Comparison of multi-phase grid-slices of ground truth Boom Clay to DDPM, WGAN-GP, and iWGAN generated ones.

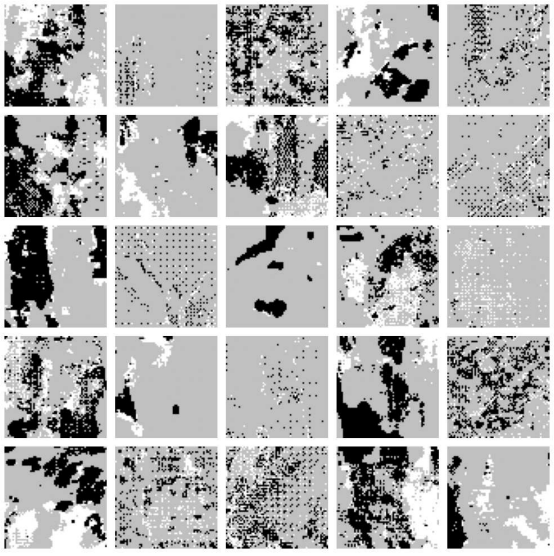
(a) Ground truth multi-phase slices of cement paste



(b) Multi-phase slices of DDPM-generated cement paste



(c) Multi-phase slices of WGAN-GP-generated cement paste



(d) Multi-phase slices of iWGAN-generated cement paste

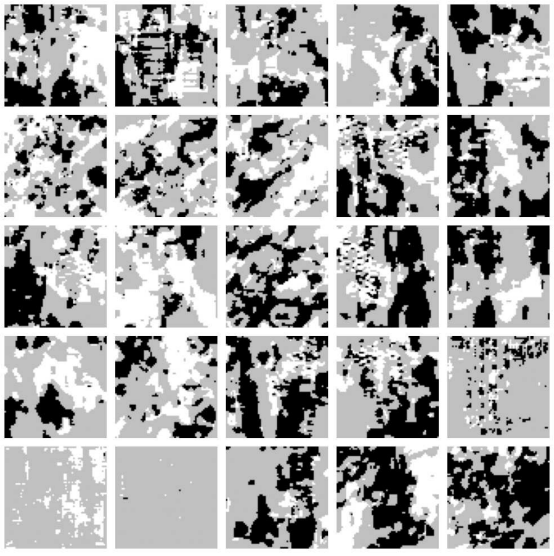


Figure 9: Comparison of multi-phase grid-slices of ground truth cement paste to DDPM, WGAN-GP, and iWGAN generated ones.

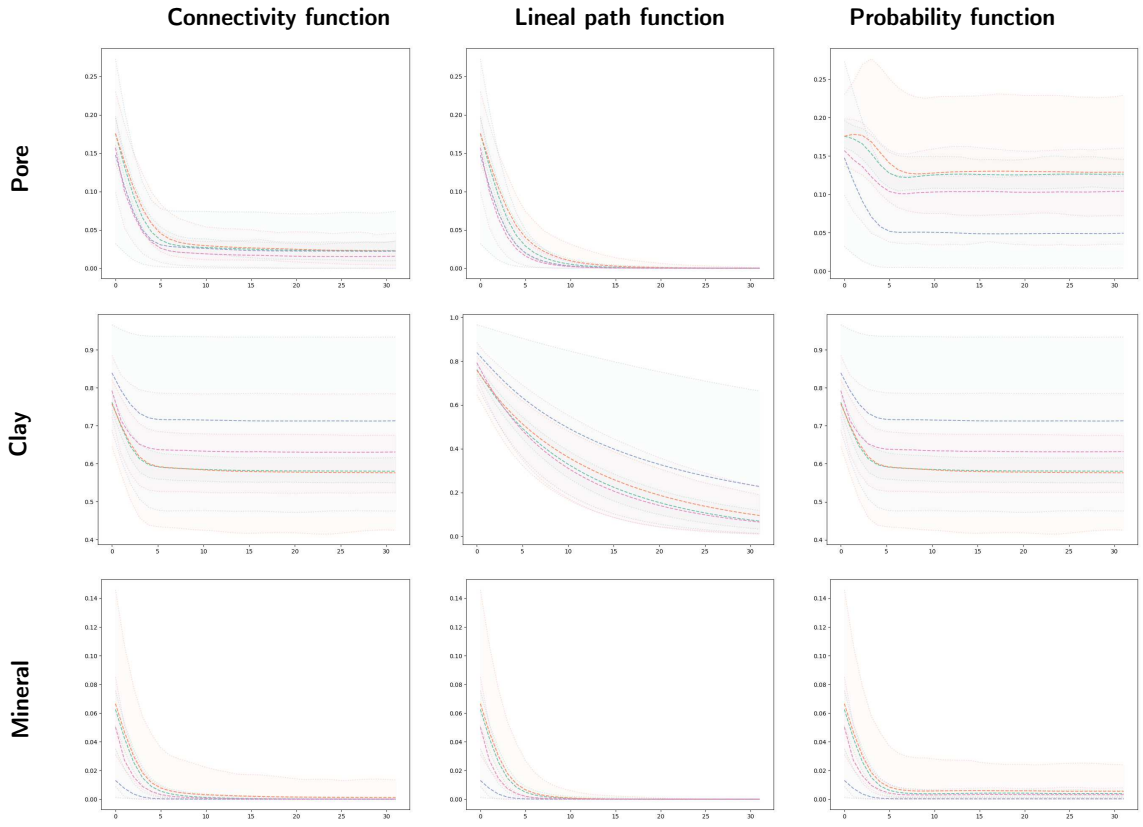


Figure 10: Comparison of structural descriptor functions for the three phases of multi-phase compacted illite. Each row shows one phase; top to bottom: **Pores**, **Clay**, **Mineral**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability).

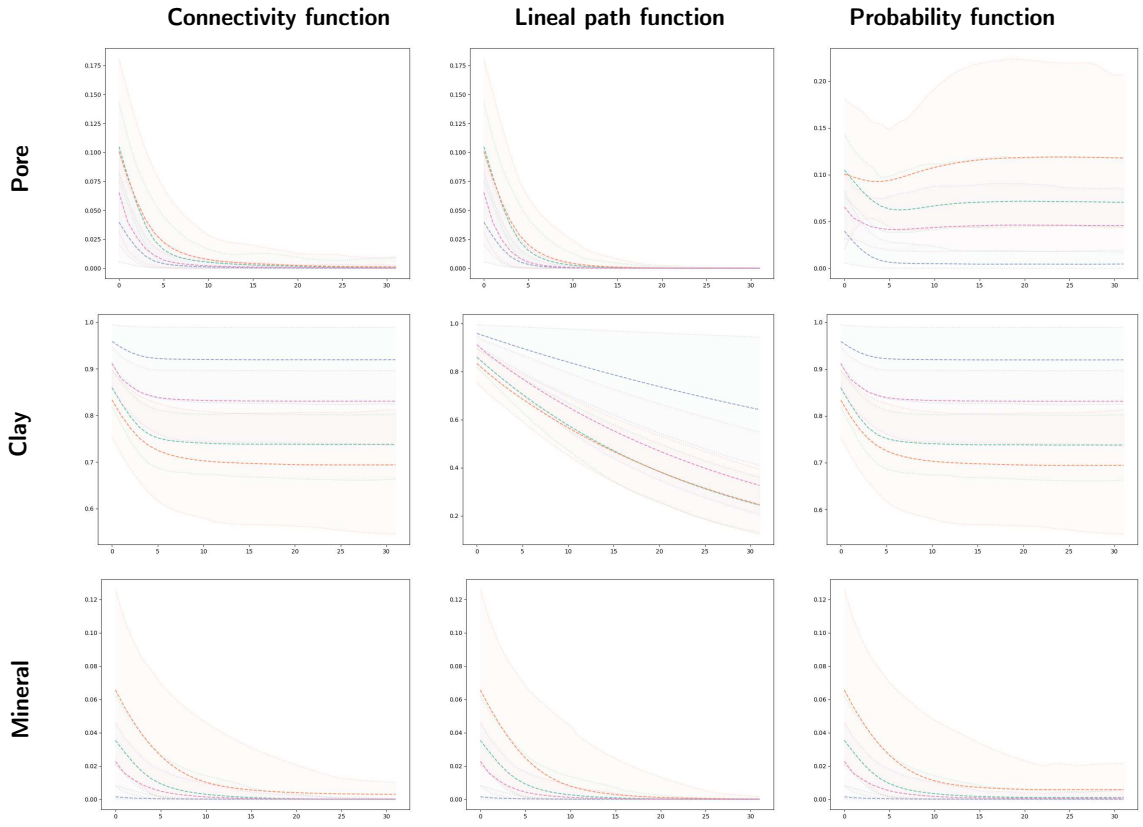


Figure 11: Comparison of structural descriptor functions for the three phases of multi-phase Boom Clay. Each row shows one phase; top to bottom: **Pores**, **Clay**, **Mineral**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability).

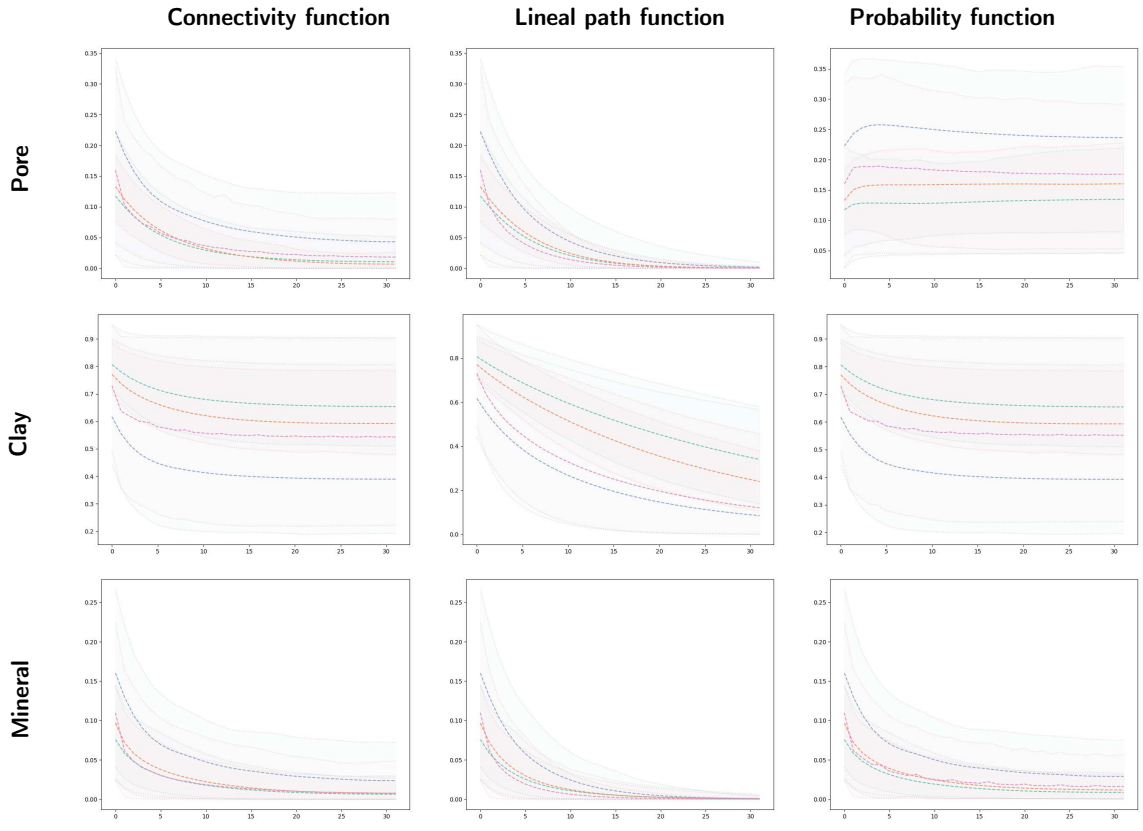
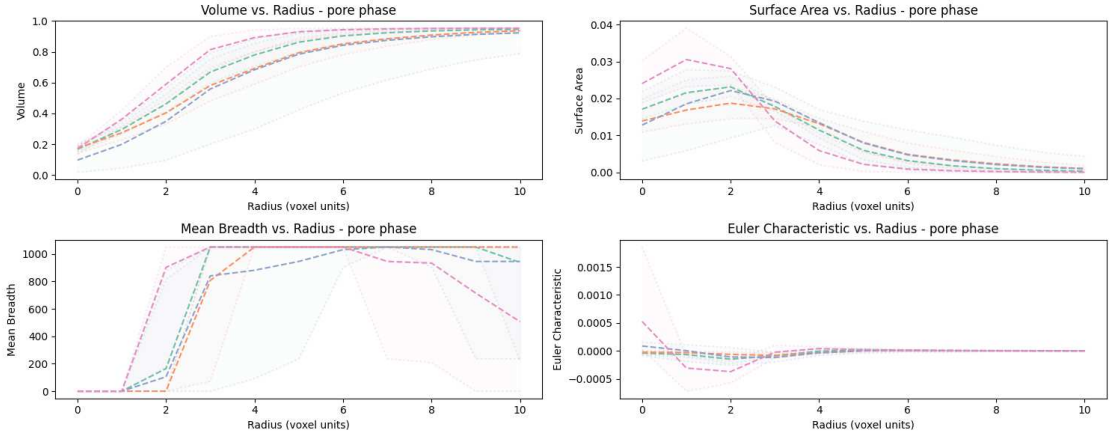
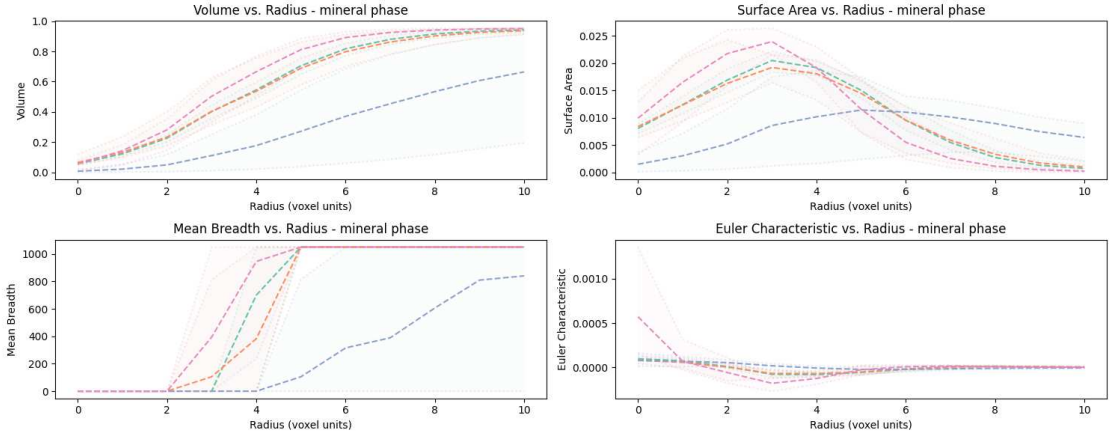


Figure 12: Comparison of structural descriptor functions for the three phases of multi-phase cement paste. Each row shows one phase; top to bottom: **Pores**, **Unhydrated cement**, **Hydrated cement**; and each column one descriptor; left to right: **Connectivity**, **Lineal path**, **Probability**. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink), plotted the frequencies (y-axis) as a function of spatial lag distance (x-axis). The dashed lines show the means, while the dotted lines and shaded areas represent the spread (variability).

(a) Minkowski functional curves for pores phase.



(b) Minkowski functional curves for minerals phase.



(c) Minkowski functional curves for clay matrix.

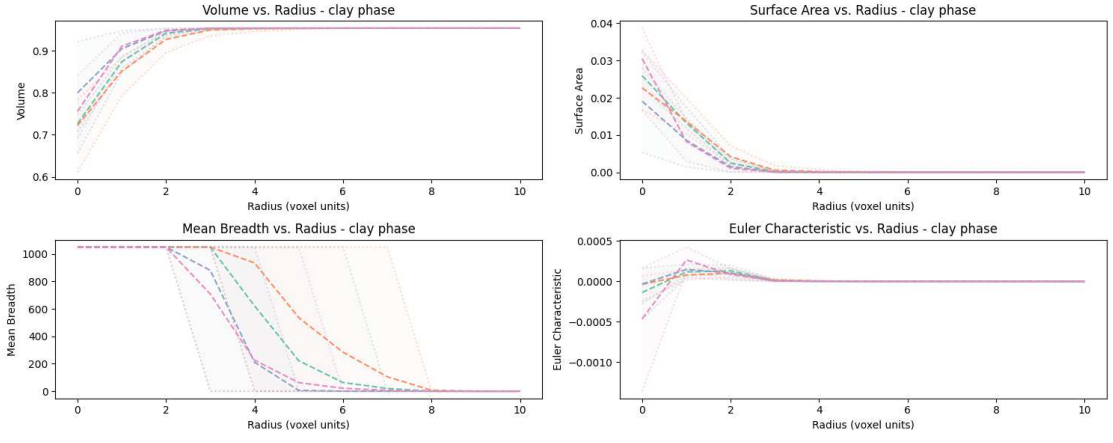
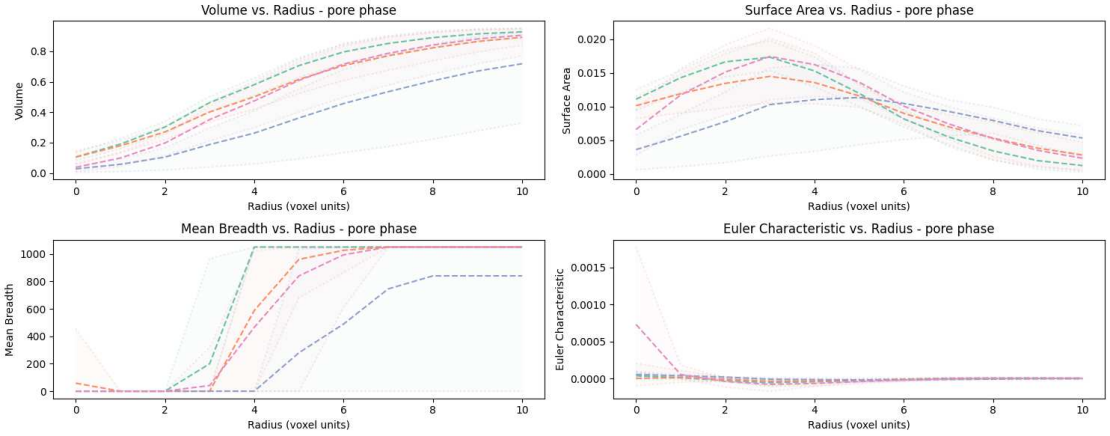
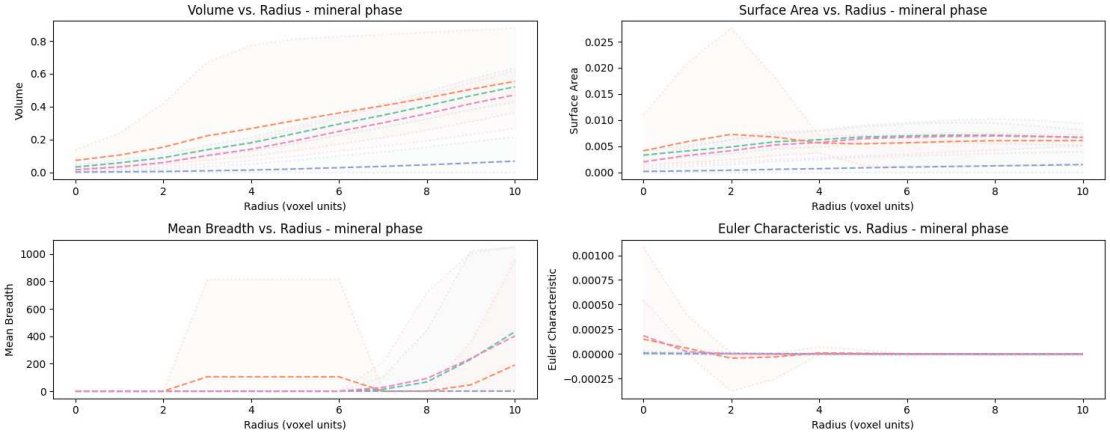


Figure 13: Multiscale Minkowski Functional Analysis of compacted illite. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink).

(a) Minkowski functional curves for pores phase.



(b) Minkowski functional curves for minerals phase.



(c) Minkowski functional curves for clay matrix.

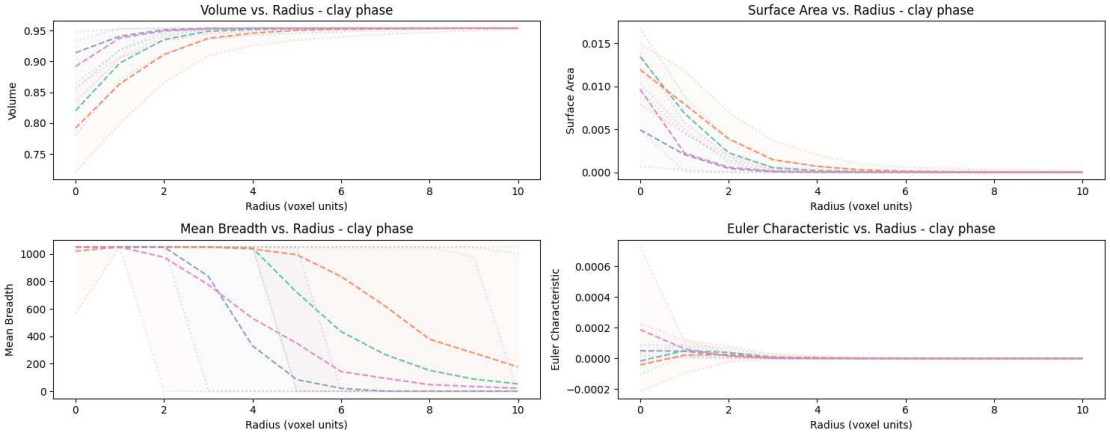
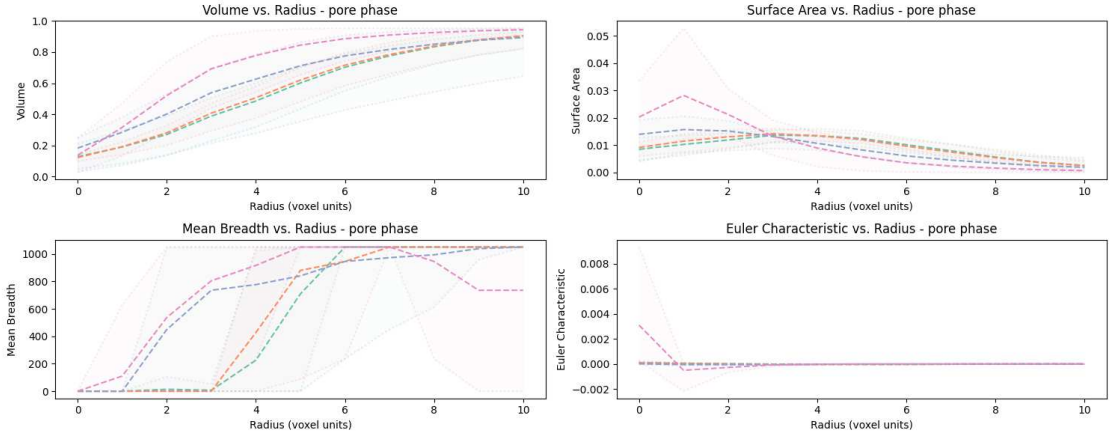
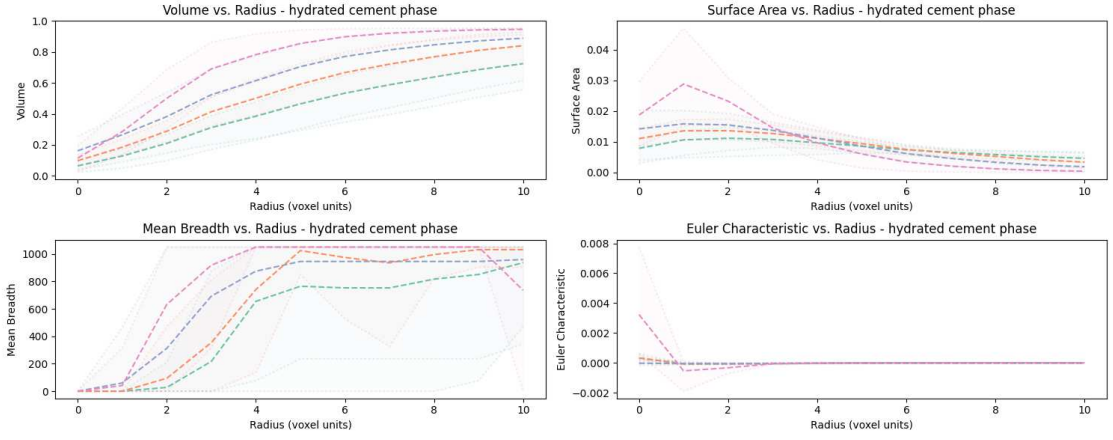


Figure 14: Multiscale Minkowski Functional Analysis of Boom Clay. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink).

(a) Minkowski functional curves for pores phase.



(b) Minkowski functional curves for hydrated phase.



(c) Minkowski functional curves for unhydrated matrix.

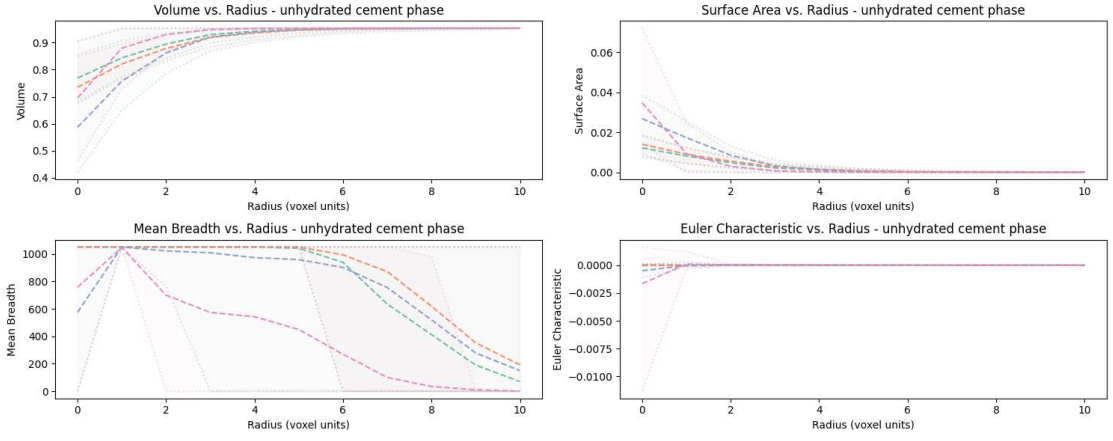


Figure 15: Multiscale Minkowski Functional Analysis of cement paste. The figure presents the evolution of four Minkowski functionals—volume, surface area, integrated mean curvature (mean breadth), and Euler characteristic—as a function of the dilation radius applied to binary microstructure data. The curves are obtained by performing a morphological sweep (using spherical dilation) on each 3D sample (averaged over 150 samples per source) and computed separately for individual phases. Within each subplot, the curves represent the *ground truth* sample (green), and realizations generated by *DDPM* (orange), *iWGAN* (blue), and *WGAN-GP* (pink).

Figure 16: Ternary diagrams of materials composition. Each row shows one material; top to bottom: Compacted Illite, Boom Clay, Cement Paste. Within each ternary diagram, the points represent compositional distributions of the three phases, with colors indicating different data sources: ground truth samples (green), and microstructures generated by DDPM (orange), WGAN-GP (pink), and iWGAN (blue). The position of each point within the triangle represents the relative proportions of the three phases, with clustering patterns revealing how accurately each generative model reproduces the compositional constraints of the reference materials.