SinFusion-based Geological Model Augmentation and Well Data Integration

Scan for Abstract

Eunsil Park[1)] · Junil Yoon[2)] · Hong Lee[2)] · Honggeun Jo[1)*]
1) Inha University, 2) Korea Gas Corporation
contact: *honggeun.jo@inha.ac.kr

Introduction

Gap

- **Multiple-Point Statistics (MPS)**: Enables well conditioning but is computationally expensive and limited in reproducing complex geological structures (Pyrcz and Deutsch, 2014)
- **Stratigraphic Forward Modeling (SFM)**: Provides realistic process-based depositional models but is computationally expensive and difficult to condition with well data (Jiménez et al., 2024)
- **GenAI**: Reproduces complex geological patterns efficiently (Park et al., 2025; Jo et al., 2020) but requires large datasets and lacks generalized conditioning methods (e.g., retraining)
- Existing methods struggle to balance geological realism, well conditioning, and computational efficiency

Opportunity

- **Single-image diffusion (SinFusion)** learns from a single geological model
- Combining SFM with SinFusion enables (1) **high efficiency**, (2) **flexible data conditioning**, and (3) generation from **a single geological model**

Objectives

- Develop a **SinFusion-based SFM modeling** framework
- Enable rapid generation of multiple realistic realizations with well conditioning
- Demonstrate applicability to a **3D deepwater submarine fan system**

Workflow

1. Data Preprocessing

- Convert SFM (SLB's GPM) outputs to Python, validate facies distributions, and interpolate undefined regions to ensure continuous, training-ready data

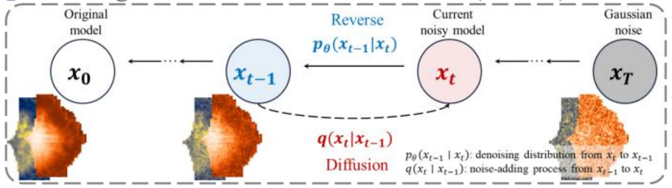
2. SinFusion Implementation

- Generate training patches from a single geological image, train **SinFusion** to **learn structural variability**, and produce realizations with **progressive well conditioning** during reverse diffusion

3. Performance Validation

- Evaluate the generated realizations using low-dimensional embedding and variograms to assess structural fidelity and diversity

Denoising Diffusion Probabilistic Models (DDPM)



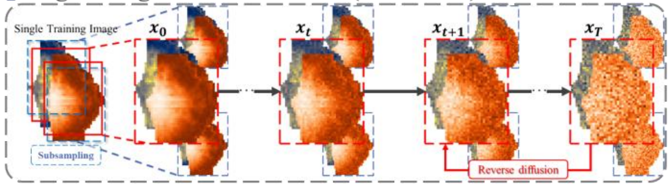
Forward diffusion and reverse generation mechanism (modified from Ho *et al.*, 2020)

① **Diffusion**: add Gaussian noise to the original image over timesteps t to obtain x_T
② **Reverse**: A neural network performs the reverse process by progressively removing noise
→ Starting from pure noise x_T, the model reconstructs x_0 similar to the original image

U-Net-based Neural Network for Denoising (Reverse Diffusion)

- Input: a noisy image at timestep t, $x_t = \sqrt{\alpha_t}x_{t-1} + \sqrt{1-\alpha_t}\varepsilon_t$
- Output: **noise predicted by the U-Net**, $\hat{\varepsilon}_t = \mathrm{UNet}(x_{t'}\, t)$
- Using $\hat{\varepsilon}_t$, the model **removes noise** from x_t and **reconstructs the previous state** x_{t-1}
- α_t: fraction of previous image preserved, ε_t: random Gaussian noise, $\hat{\varepsilon}_t$: predicted noise

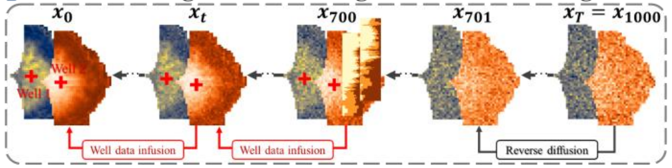
Single Image Diffusion Model (SinFusion)



Principle of SinFusion training data construction (modified from Nikankin et al., 2022)

- DDPM requires large-scale training data, limiting its use for single-image-based problems
- A method proposed to learn diversity and generalization **from a single image**
- We construct the input x_0 using random crops of about 95% of the original image size

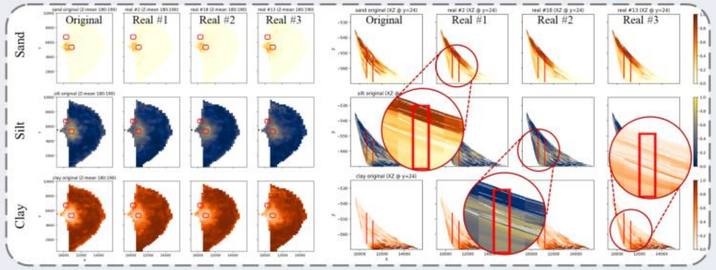
Well Data Integration and Progressive Conditioning



SinFusion well data integration workflow

- The pretrained model directly incorporates **well data without additional training**
- In the initial stage, unconditional diffusion is used for free generation
- From a **chosen timestep onward, well data are injected** at well locations ($t \le 700$)
- Enables **flexible integration regardless of well number or arrangement**

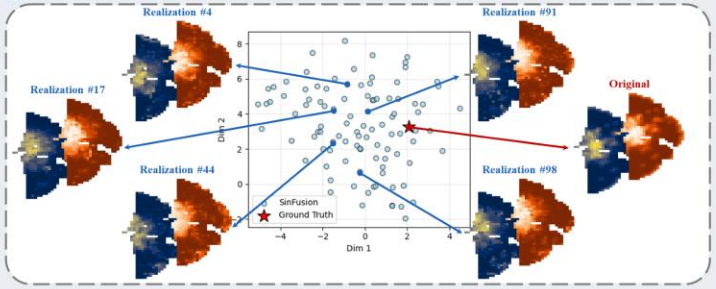
SinFusion Results: Top and Cross-Sectional Views



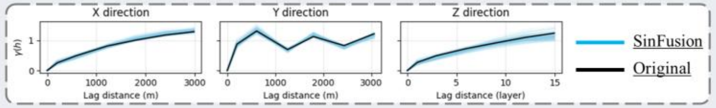
SinFusion results: Sand, Silt, and Clay in the XY plane and XZ section

- Preserves spatial **connectivity and boundary variability**, reflecting complexity
- Maintains **vertical continuity** and reproduces **depositional structures and stratigraphic trends**
- Well data are **honored while maintaining continuity** with surrounding areas

SinFusion Performance Evaluation



Low-dimensional embedding of SinFusion realizations



Directional Variograms Analysis

- Assessing **structural similarity** and **diversity** in low-dimensional space, realizations **cluster near the reference**, indicating **realistic reproduction** of the **patterns**, while also spreading across the space, demonstrating diversity in the generated results
- **Variogram-based evaluation** showing **stable agreement** with the **reference spatial continuity** in both horizontal and vertical directions

Computational Time Comparison for Geological Modeling

SinFusion training	**SinFusion generation**	MPS	SFM
261 min	**13 sec**	9 min	17 min

Conclusion

- We propose a **SinFusion-based geological modeling framework** that achieves **high efficiency**, **flexible data conditioning**, and realistic model generation using only a **single geological model**.
- The framework enables **rapid generation** of realistic models and **immediate updates with newly acquired well data**, effectively enforcing **local constraints near wells** while preserving the **global geological structure** without additional training.
- The method consistently reproduces **spatial connectivity** and **geological complexity**, while **low-dimensional analysis** and **variograms** confirm structural fidelity and diversity.
- Future work will extend the framework from **facies modeling** to properties such as **porosity and permeability**, and integrate **seismic data** for multi-data conditional modeling.

References

- M. J. Pyrcz and C. V Deutsch, Geostatistical reservoir modeling. Oxford university press, 2014.
- G. Jiménez, A. H. A. Latiff, W. Ben Habel, and M. Poppelreiter, "Effective application of geological process modeling for unravelling carbonate build-up complexity: A case study from the EX-Carbonate build-up in central Luconia Province, Malaysia," Mar. Pet. Geol., vol. 170, p. 107117, 2024.
- E. Park, H. Kim, H. Shin, and H. Jo, "Deep Learning-Assisted THM-Integrated InSAR Modeling for Co2 Storage Characterization and Surface Deformation Forecasting," Available at SSRN 5226149, 2025.
- J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," Adv. Neural Inf. Process. Syst., vol. 33, pp. 6840–6851, 2020.
- Y. Nikankin, N. Haim, and M. Irani, "SinFusion: Training diffusion models on a single image or video," arXiv preprint arXiv:2211.11743, 2022.

Acknowledgement

- This work was supported by Korea Gas Corporation (RD2025-0071)

Scan for an interactive 3D view

Sharing is encouraged

SinFusion-based Geological Model Augmentation and Well Data Integration

EGU26 (ESSI1.4: Deep Learning in Geosciences)

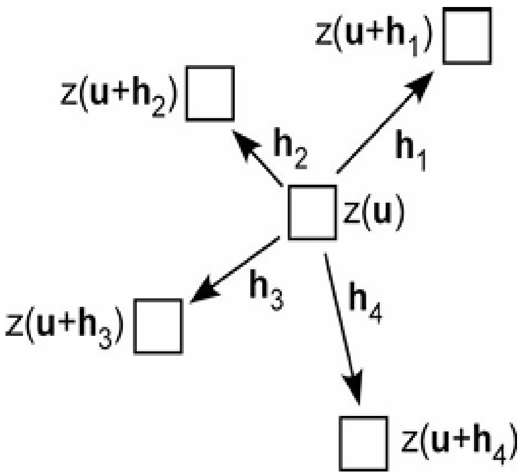
Eunsil Park[1], Hong Lee[2], Junil Yoon[2], Honggeun Jo[1]*

1) Inha University, Department of Energy Resources Engineering, Republic of Korea
2) Korea Gas Corporation, KOGAS Research Institute, Republic of Korea

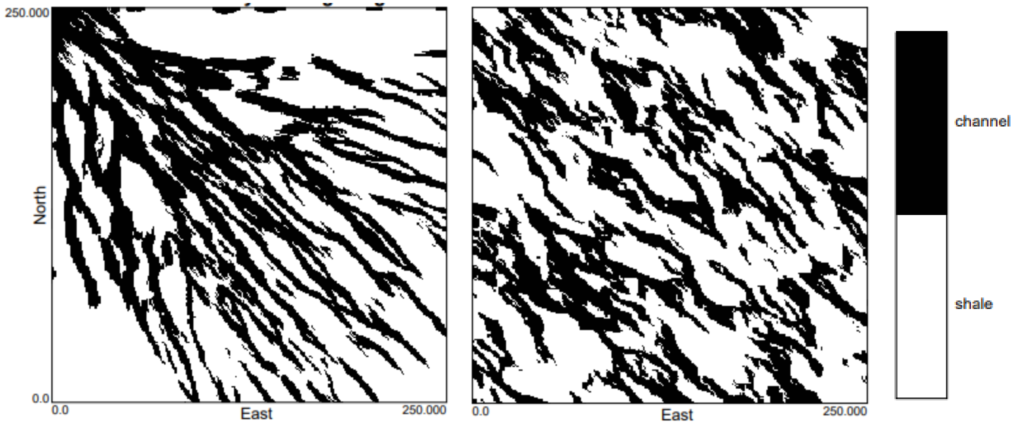
Limitations of Geological Modeling

Conventional statistics-based geological modeling methods

- **Sequential Gaussian Simulation (SGS), Sequential Indicator Simulation (SIS)**
  - Variogram-based geostatistical methods such as SGS and SIS can reproduce local continuity, but have limitations in representing complex geological structures
- **Multiple Point Statistics (MPS)**
  - MPS uses a training image to simultaneously account for **spatial relationships among multiple locations**
  - **Complex spatial patterns** contained in the training image (TI) are encoded into a search-tree data structure to generate realizations (e.g., SNESIM)



Multi-point spatial relationship concept (Pyrcz *et al.*, 2014)

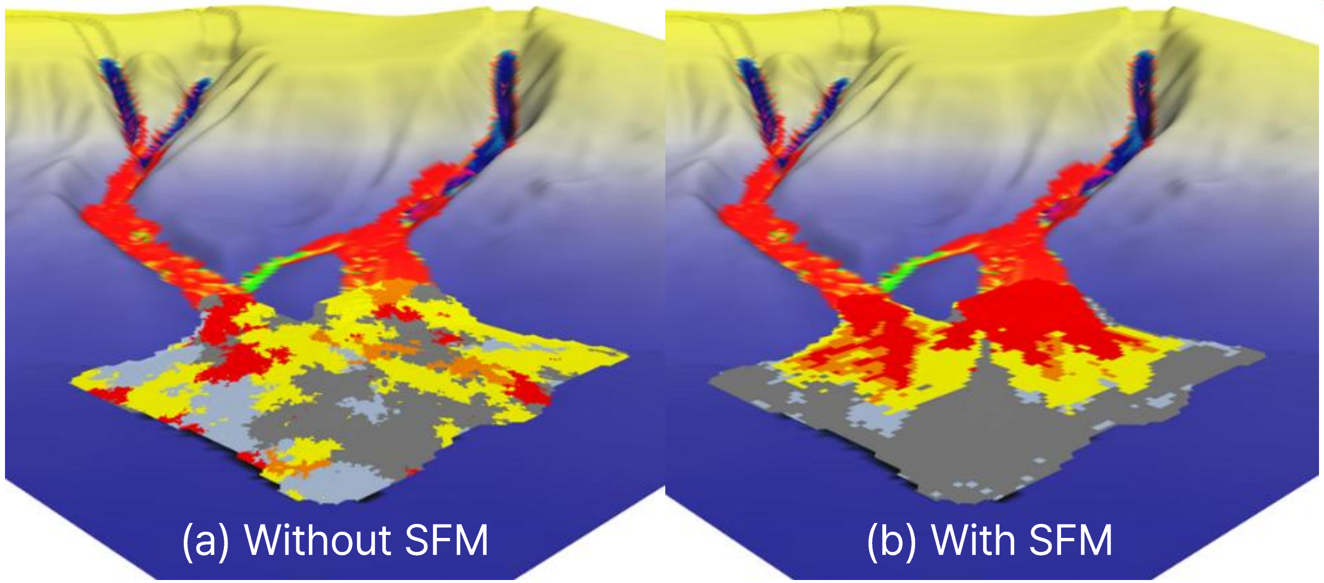


MPS results for non-stationary patterns (Liu, 2006)

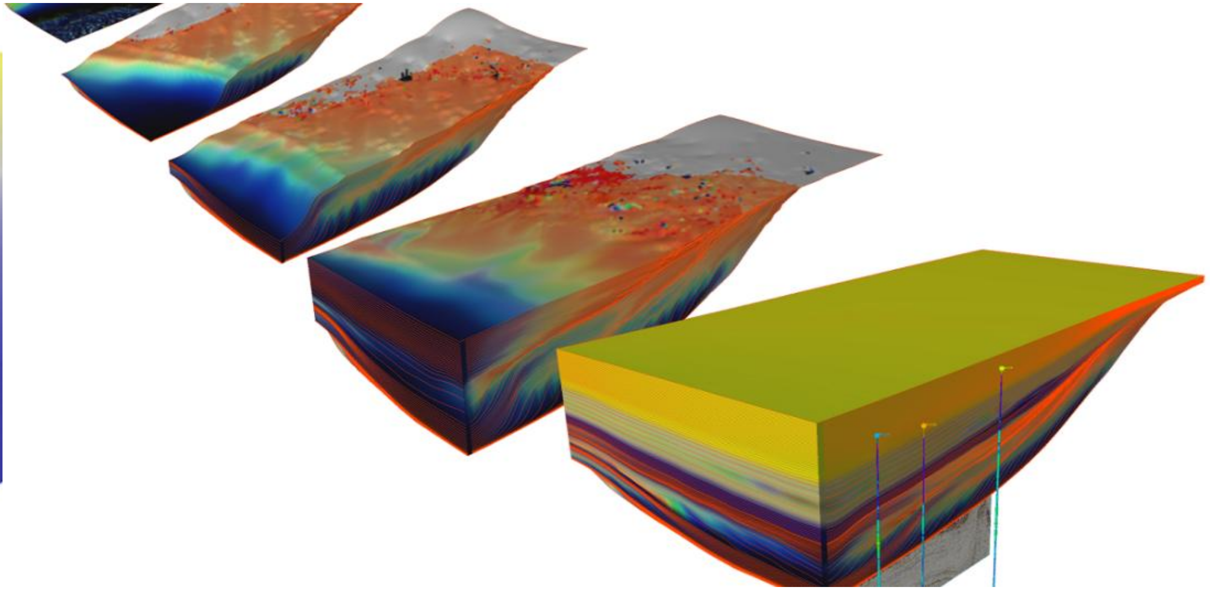
Limitations of Geological Modeling

Stratigraphic forward modeling (SFM)

- A physics-based simulator that models **depositional and stratigraphic processes** to **predict stratigraphic architecture** and **reservoir property distributions**
- It captures geologically realistic stratigraphic architectures by simulating sedimentary processes and their controls
- Limitation: **Conditioning the model to precisely honor local hard data (e.g., well constraints) remains challenging**



With vs Without SFM: Facies Modeling Comparison (Source: SLB)

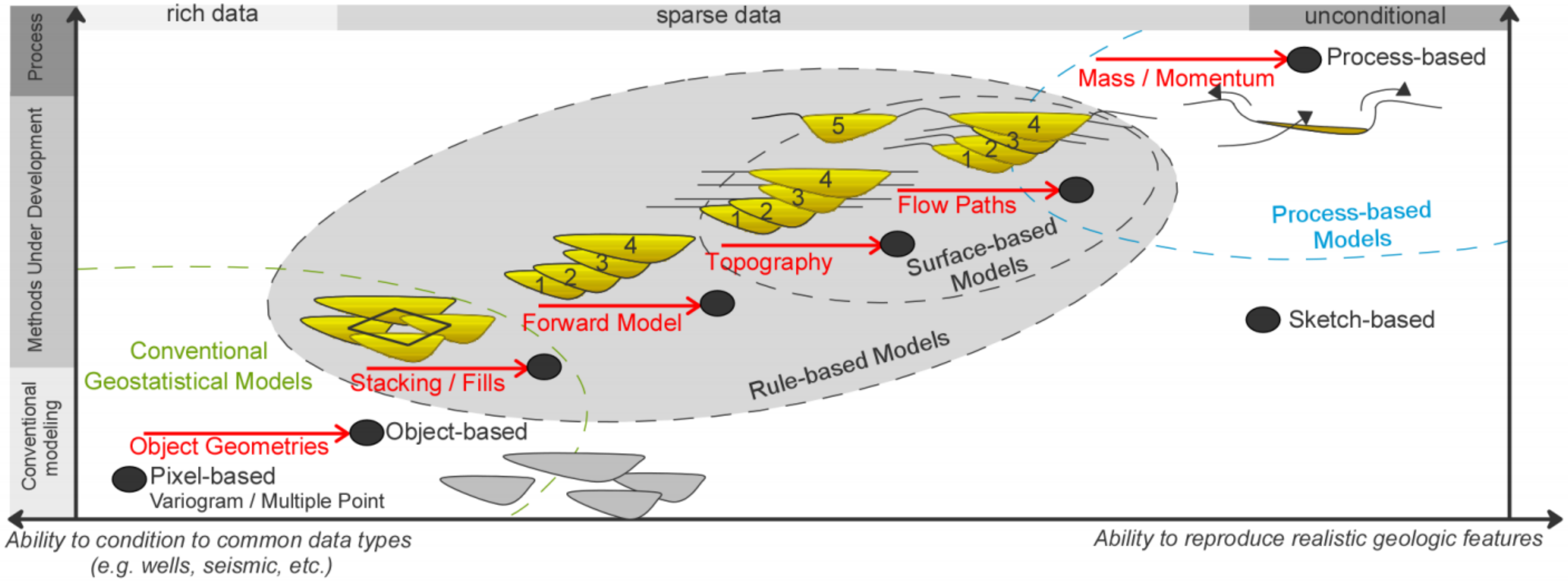


Depositional Development Over Simulation Time (Source: SLB)

Research Objective

Generative AI-based Geological Model Augmentation and Well Data Integration

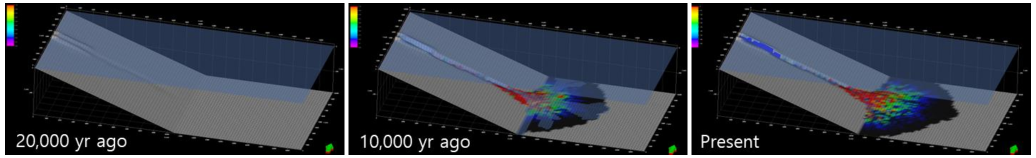
- **MPS** methods require **substantial computation time** during model generation and **have limitations in realistically reproducing complex geological structures**
- **SFM** approaches can represent **geologically plausible architectures**, but **direct integration of well data is challenging** and computational costs are high
- To address these limitations, we develop **a generative AI–based geological modeling** method that achieves both **efficient generation** and **improved structural realism**



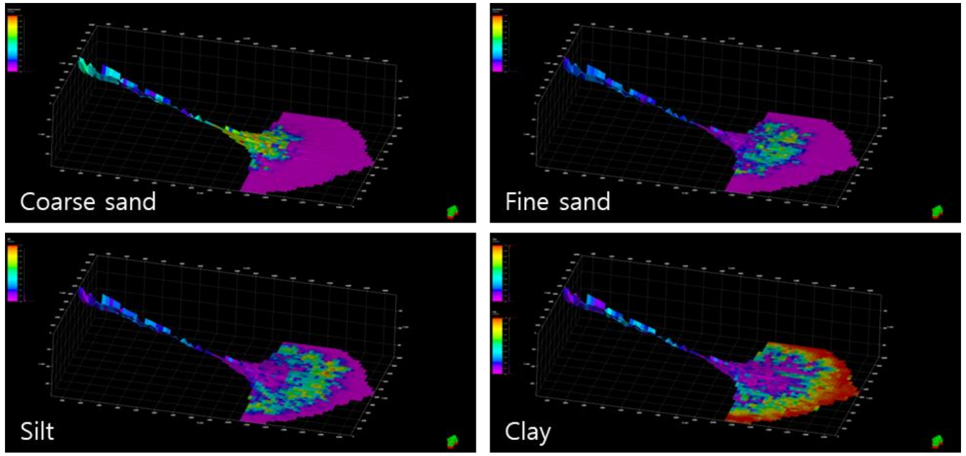
Comparative Performance Analysis of Geological Modeling Methods (Pyrcz, 2015)

Building Training Data for Generative AI

Stratigraphic forward modeling results



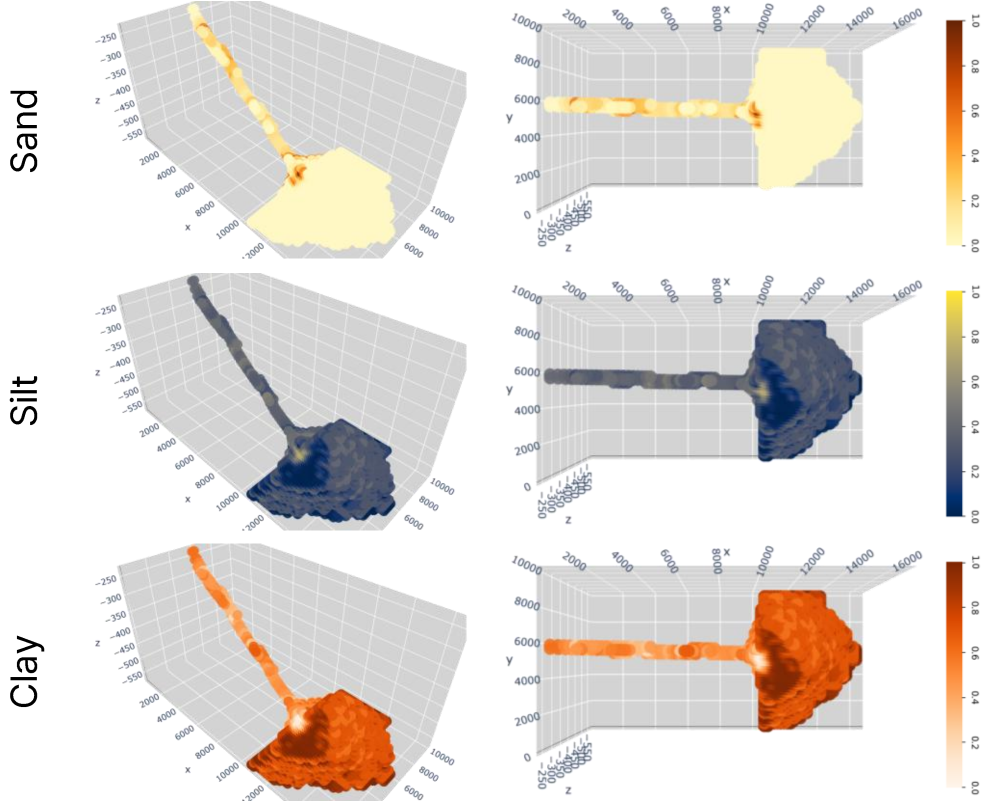
Time-lapse depositional simulation



Spatial distributions by facies

Petrel outputs → Python-based facies visualization

Perspective view | Top View

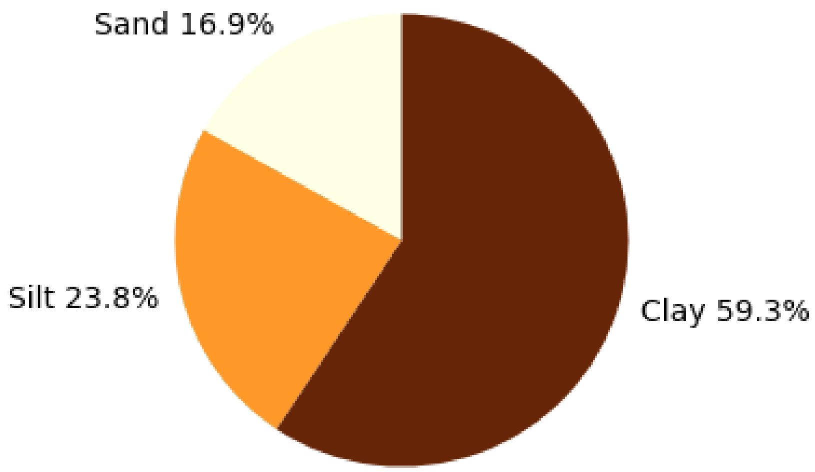


- Export SFM results from Petrel in GSLIB format
- Load and visualize in Python to validate **Petrel–Python I/O compatibility**

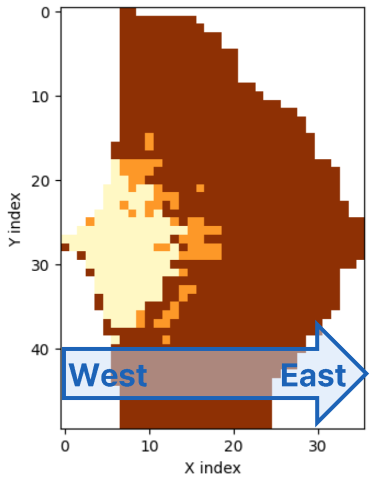
Building Training Data for Generative AI

Facies Statistics Analysis

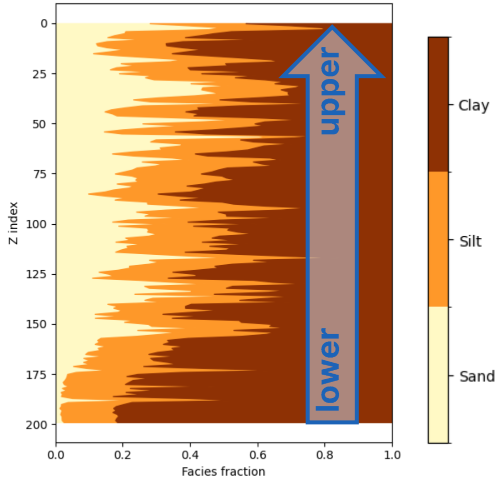
- From SFM models with over one million cells, we focus on the **continental shelf to deep-sea domain** with high oil and gas potential, excluding the continental slope
- Facies proportions: **Sand 16.9%, Silt 23.8%, Clay 59.3%**
- Facies Majority: for each cell, the facies that occurs most frequently along the z-direction
  - **Sand dominates in the west, while Clay dominates in the east**
- Facies proportions by z-slice:
  - **Sand is dominant in upper layers, while Clay is dominant in lower layers**



Facies Proportions



Facies Majority



Facies Proportions by Z-slice

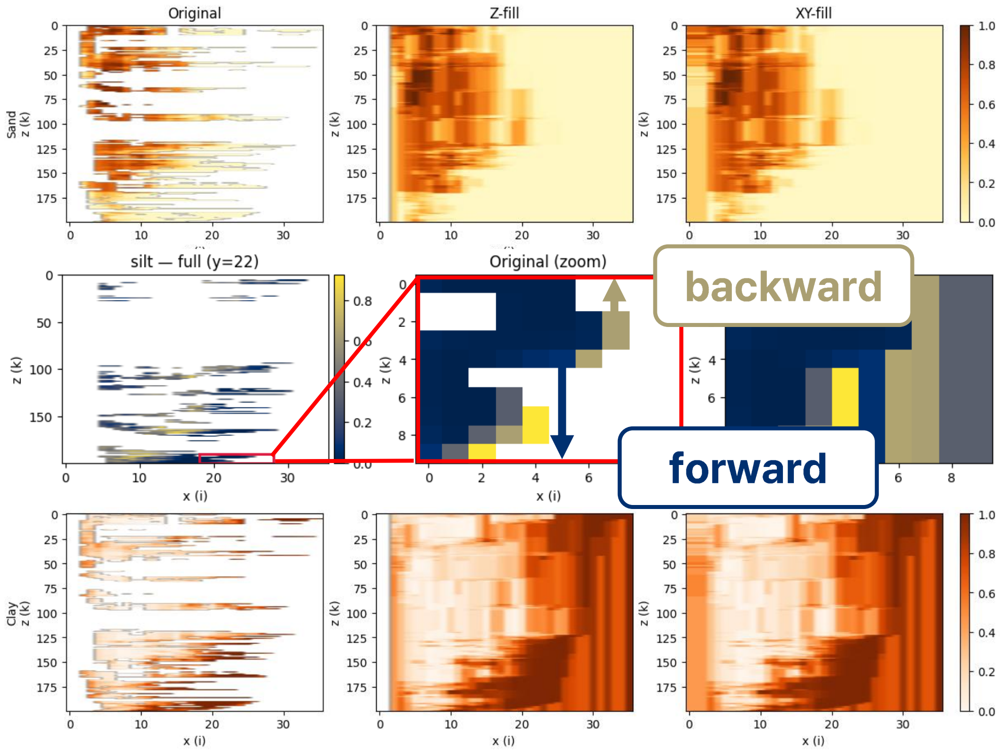
Building Training Data for Generative AI

Missing Value Handling: In SFM simulations, areas not reached by sediments contain NaNs

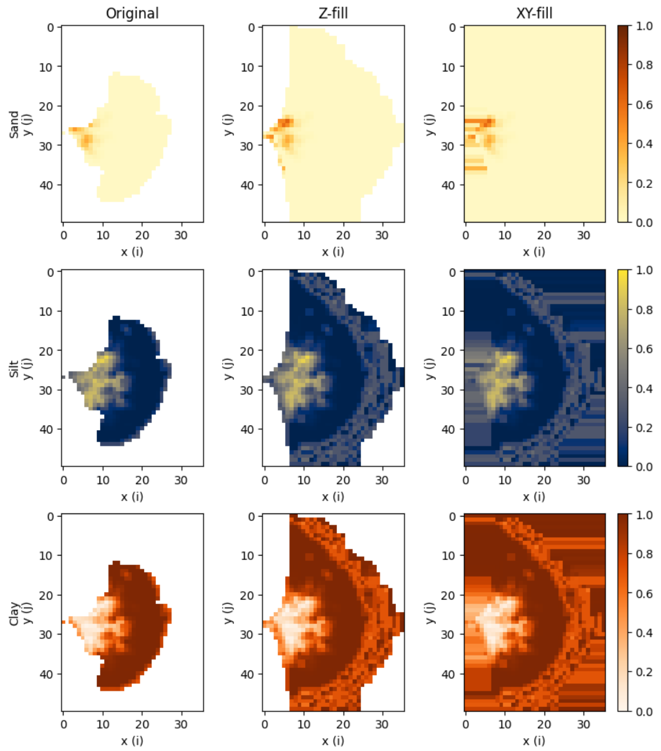
- **Z direction (depositional)**:
  - Fill bottom→top(**Backward**), then top→bottom (**Forward**)

  → Use the nearest valid value to preserve continuity and boundaries

- **XY plane (horizontal)**:
  - Fill left→right (**Backward**)
  - then right→left (**Forward**)



XY section: original → Z fill → XY fill

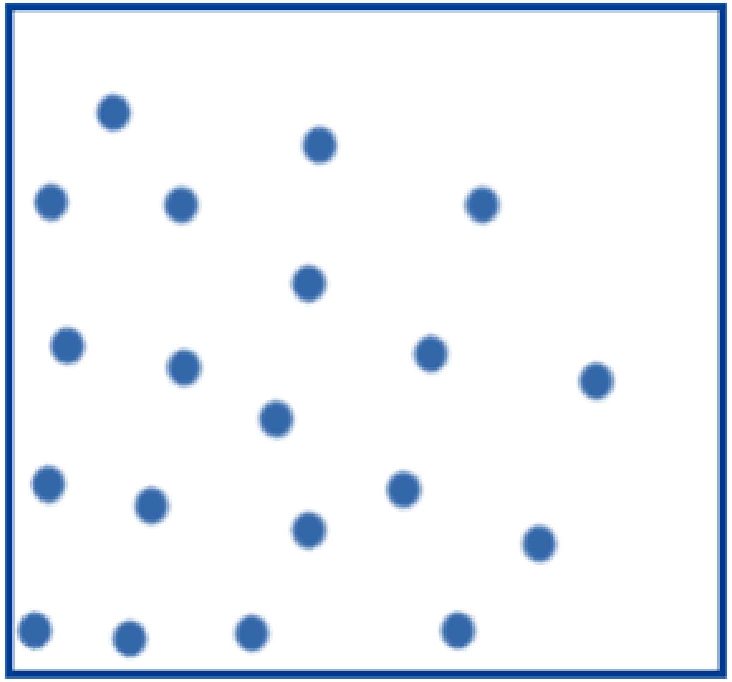


XZ section: original → Z fill → XY fill

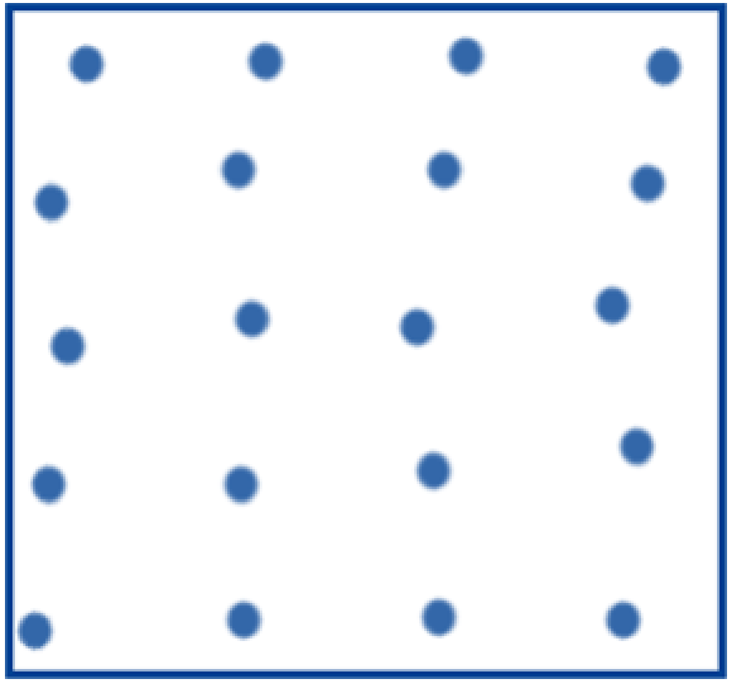
Basic Principle of Diffusion Models

Example of physical diffusion

① **Diffusion**: particles start at a point and gradually spread over time, becoming uniformly distributed in space



Partially diffused intermediate state, $t = T/2$



Fully diffused uniform distribution, $t = T$

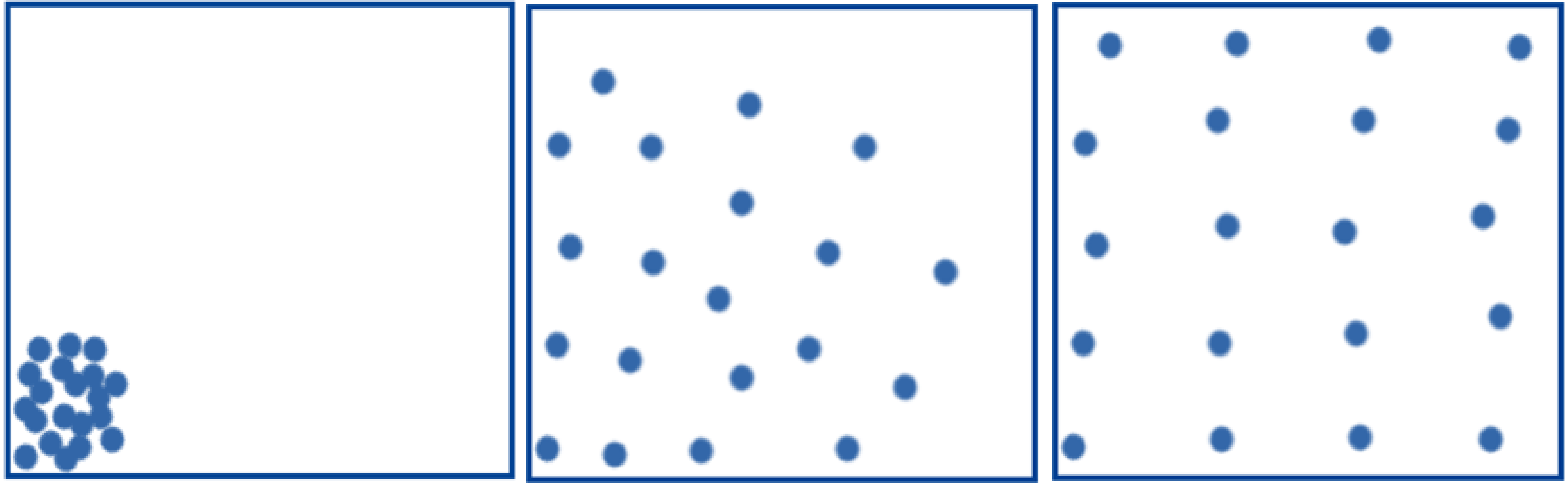
Basic Principle of Diffusion Models

Example of physical diffusion

① **Diffusion**: particles start at a point and gradually spread over time, becoming uniformly distributed in space

② **Reverse**: start from a fully diffused distribution and stochastically recover the initial state

→ Based on this diffusion and reverse-denoising concept, **generative AI diffusion models** have been proposed



Partially diffused intermediate state, $t = T/2$ Fully diffused uniform distribution, $t = T$

Diffusion model

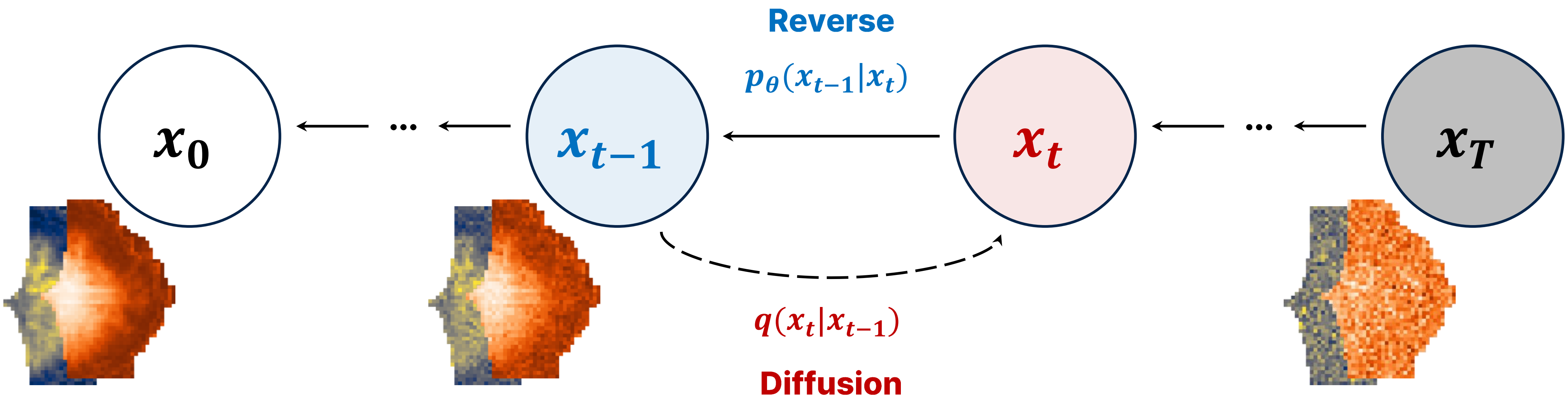
Concept of Diffusion Models

① **Diffusion**: gradually add Gaussian noise to the original image over timesteps t to obtain $x_T \sim \mathcal{N}(0, I)$

② **Reverse**: A noise-predicting neural network progressively removes noise in the reverse process

→ Starting from pure noise x_T ,the model generates a realistic sample x_0

→ The process uses 1,000 timesteps, where noise is incrementally added **(Diffusion)** and removed **(Reverse)**



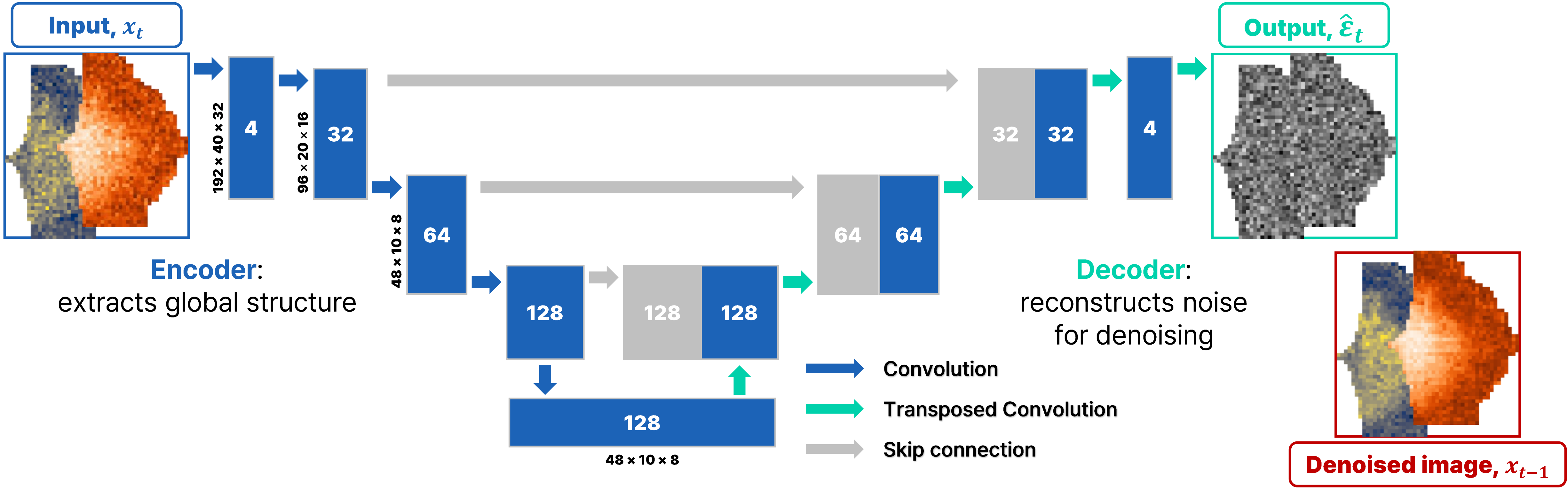
Forward diffusion and reverse generation mechanism of diffusion models (modified from Ho *et al.*, 2020)

Neural Network for Denoising (Reverse Diffusion)

Noise Prediction: Role of the U-Net

- Input: a noisy image at timestep t, $x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1-\alpha_t}\varepsilon_t$
- Output: **noise predicted by the U-Net**, $\hat{\varepsilon}_t = \mathrm{UNet}(x_t, t)$
- Using $\hat{\varepsilon}_t$, the model **removes noise** from x_t and **reconstructs the previous state** x_{t-1}

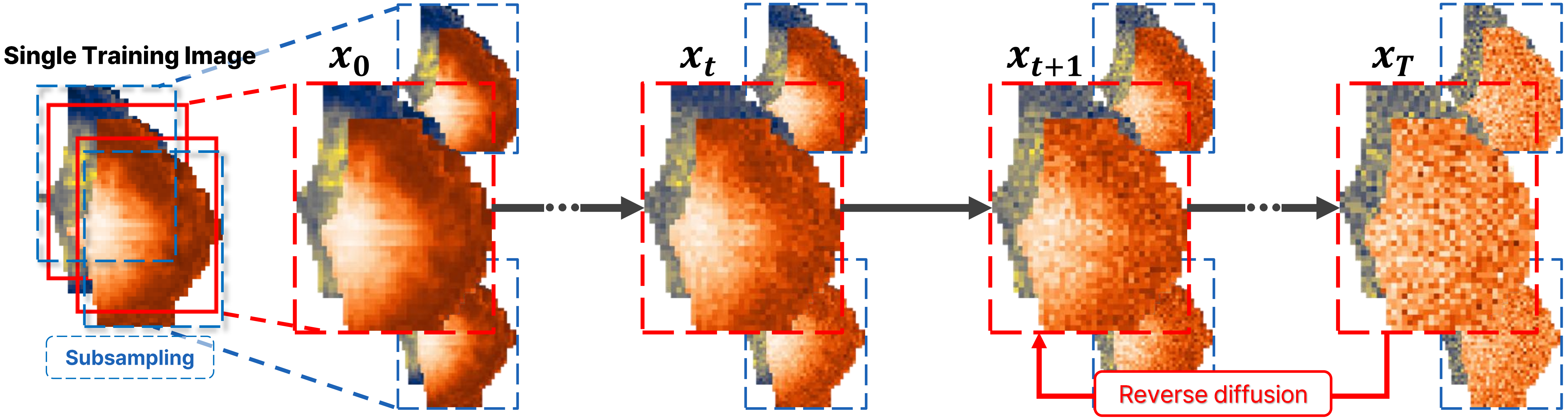
α_t: fraction that preserves information from the previous image
ε_t: random Gaussian noise
$\hat{\varepsilon}_t$: noise component predicted by the U-Net



Diffusion Model Trained on a Single Training Image

Single Image Diffusion Model (SinFusion)

- A method proposed to learn diversity and generalization from a single image **without large datasets**
- **With one geological interpretation**, we apply SinFusion **to capture structural variability and diversity**
- We construct the input x_0)subsample) using random crops of about 95% of the original image size

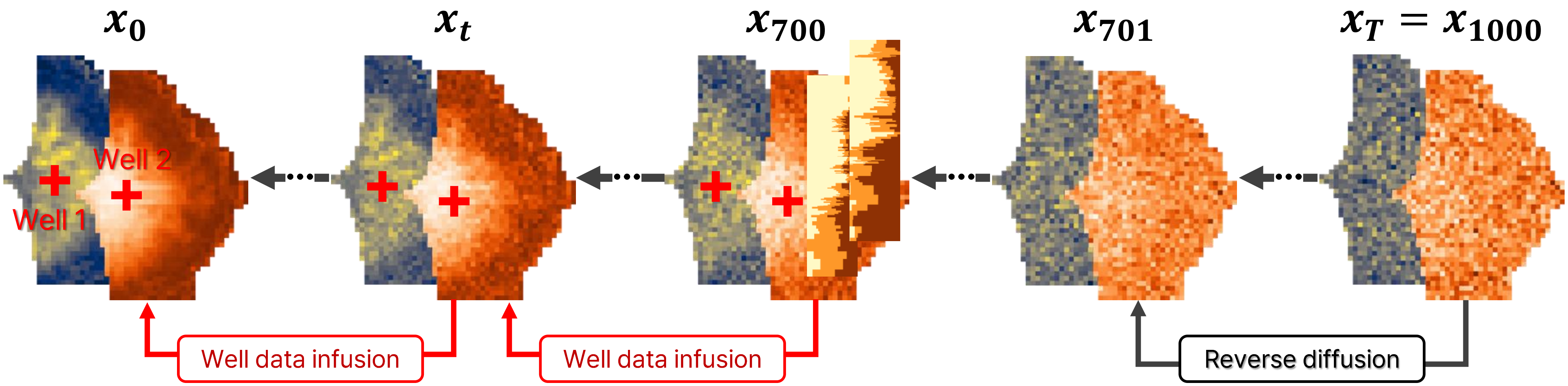


Principle of SinFusion training data construction (modified from Nikankin *et al.*, 2022)

SinFusion for Well Data Integration

SinFusion-based well data generation and progressive conditioning

- Using a pretrained model, **well locations and values can be incorporated immediately**
- In the initial stage, unconditional diffusion is used for free generation (total timesteps: $T = 1000$).
- From a **chosen timestep onward, well data are injected** at well locations (conditioning interval: $t \leq 700$).
- This enables **flexible integration without constraints on the number or spatial arrangement of wells**



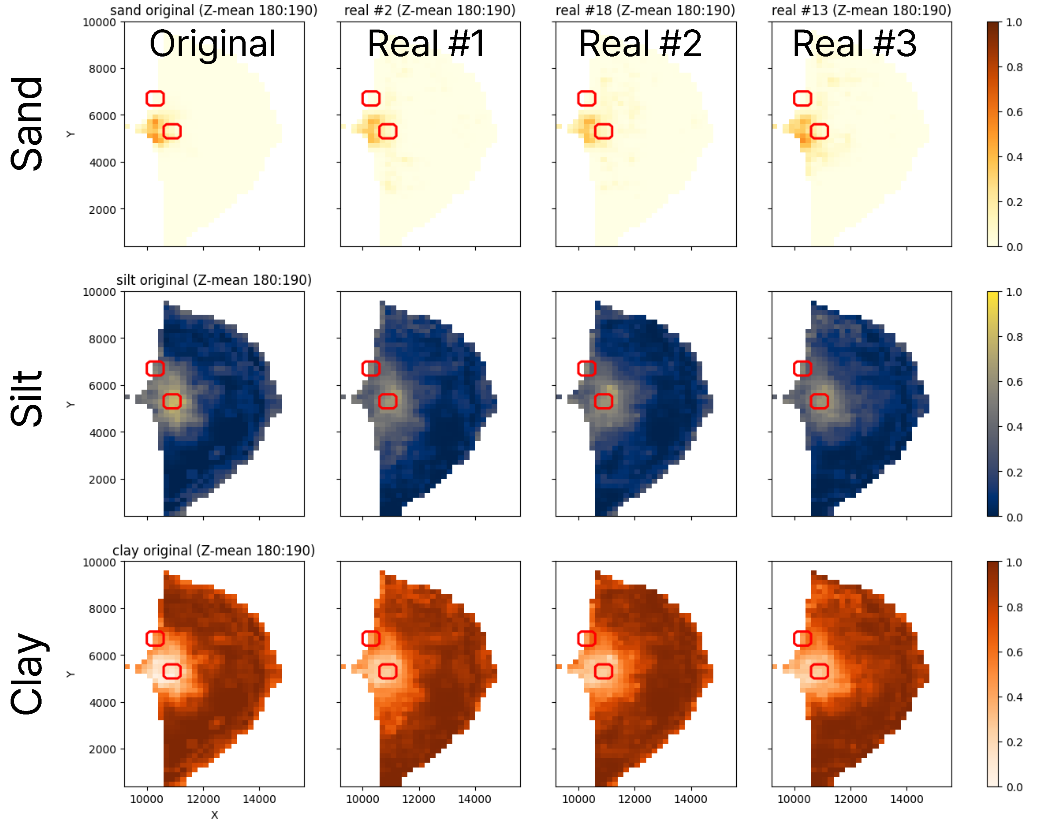
SinFusion well data integration workflow

SinFusion Results: Cross-Section Visualization

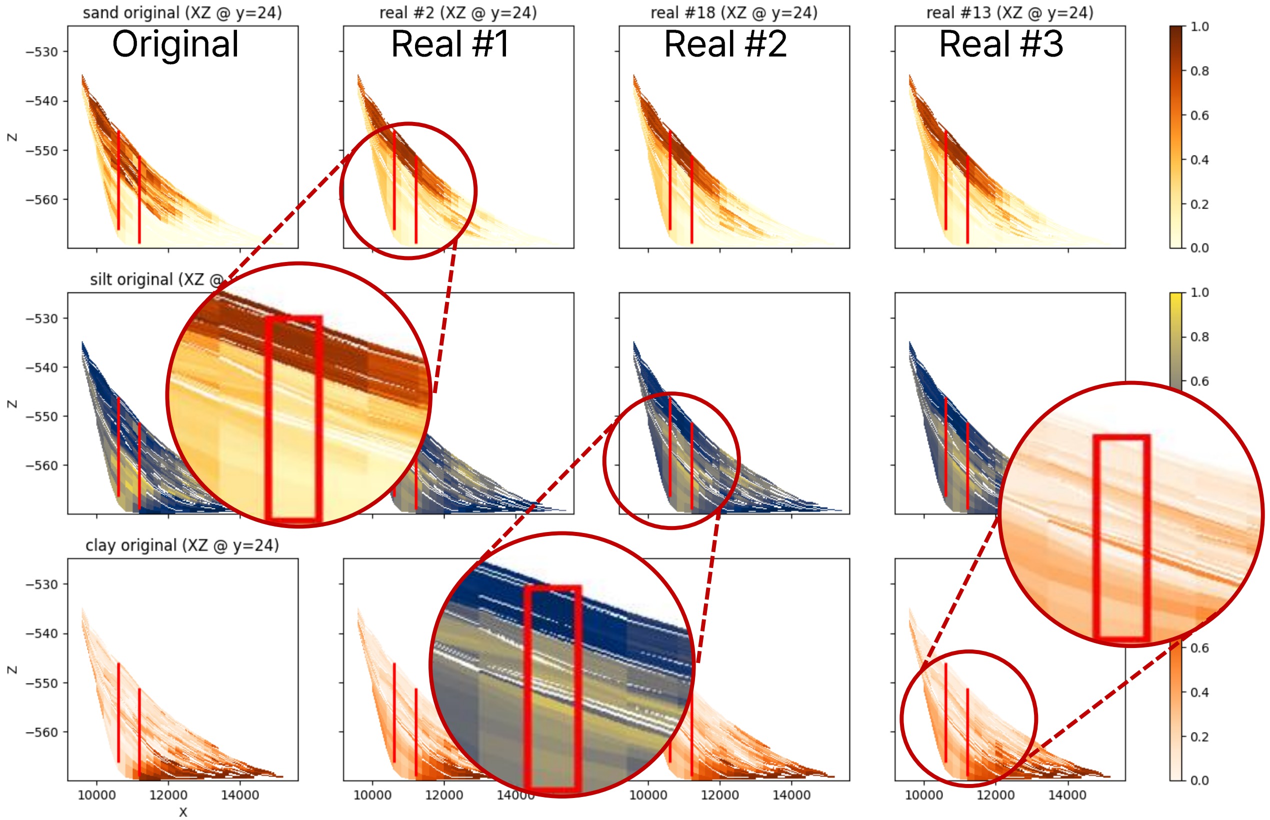
Well location

Spatial distribution reproduction of SinFusion trained on a single SFM realization

- Preserves spatial **connectivity and boundary variability**, reflecting geological complexity
- Maintains **vertical continuity** and reproduces **depositional structures and stratigraphic trends**



SinFusion results: Sand, Silty, and Clay on the XY plane

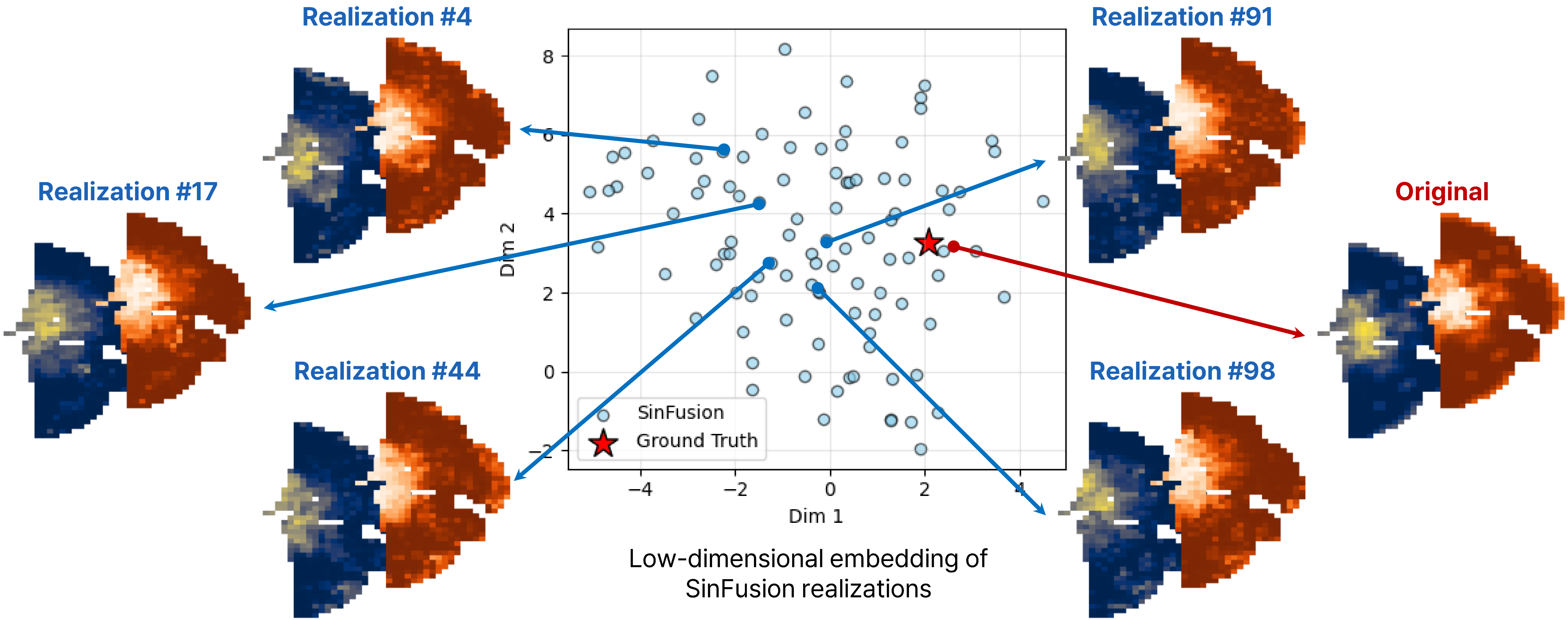


SinFusion results: Sand, Silty, and Clay on the XZ section

SinFusion Performance Evaluation

Assessing structural similarity and diversity in a low-dimensional space

- Realizations cluster **near the reference data**, indicating faithful **reproduction** of the overall data distribution
- They also **spread broadly across** the low-dimensional space, demonstrating **diversity** in the generated results



Low-dimensional embedding of SinFusion realizations

SinFusion Performance Evaluation

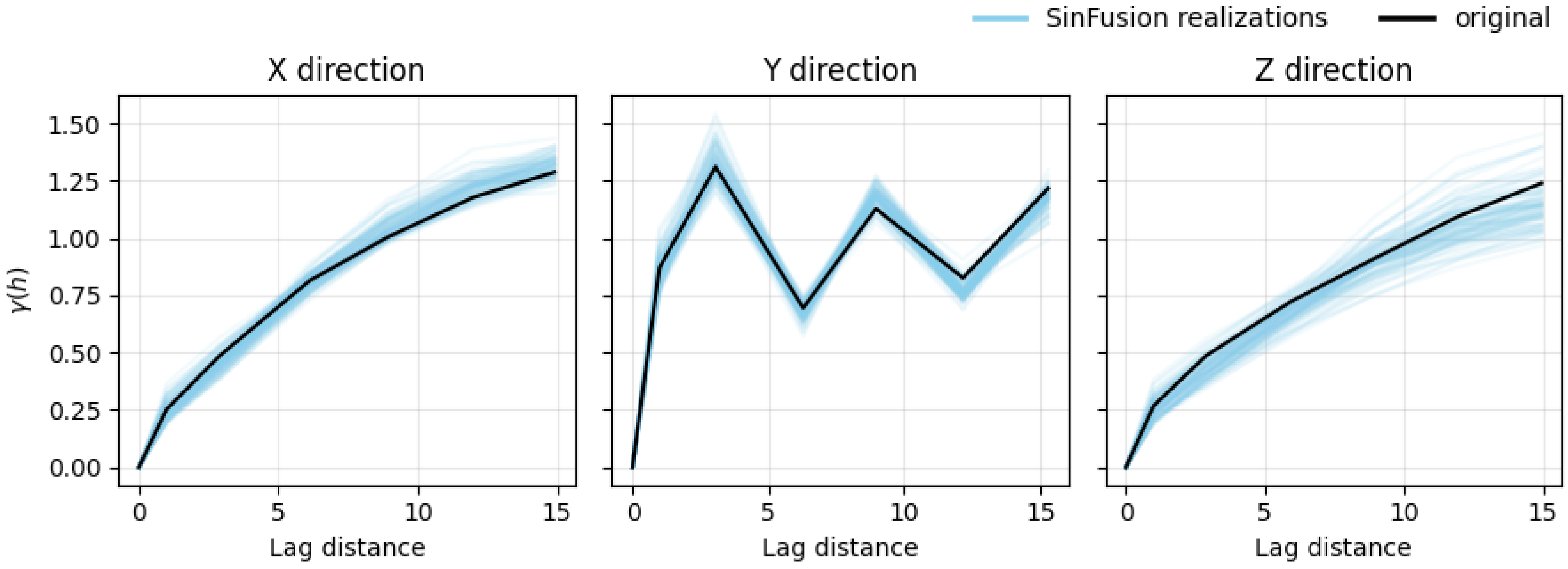
Assessment of Generated Results with Variograms and Quantitative Metrics

1. Variogram Analysis

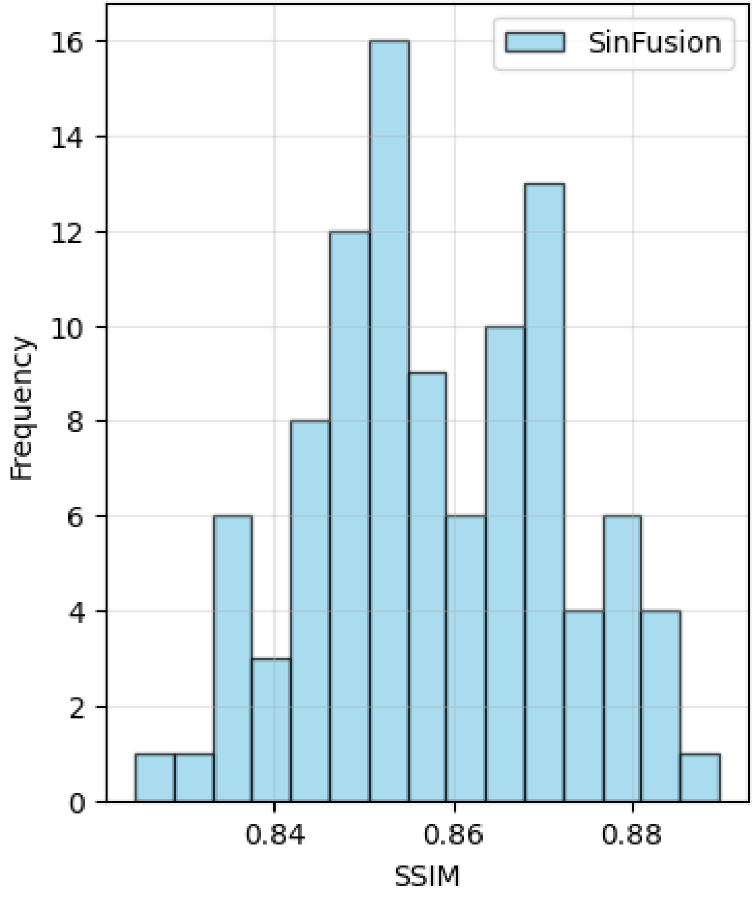
- **Stable agreement** with the reference variogram in all directions

2. Quantitative Metrics

- SSIM shows a broad distribution with generally **high values**, indicating **diversity**



Directional Variograms (X/Y/Z)

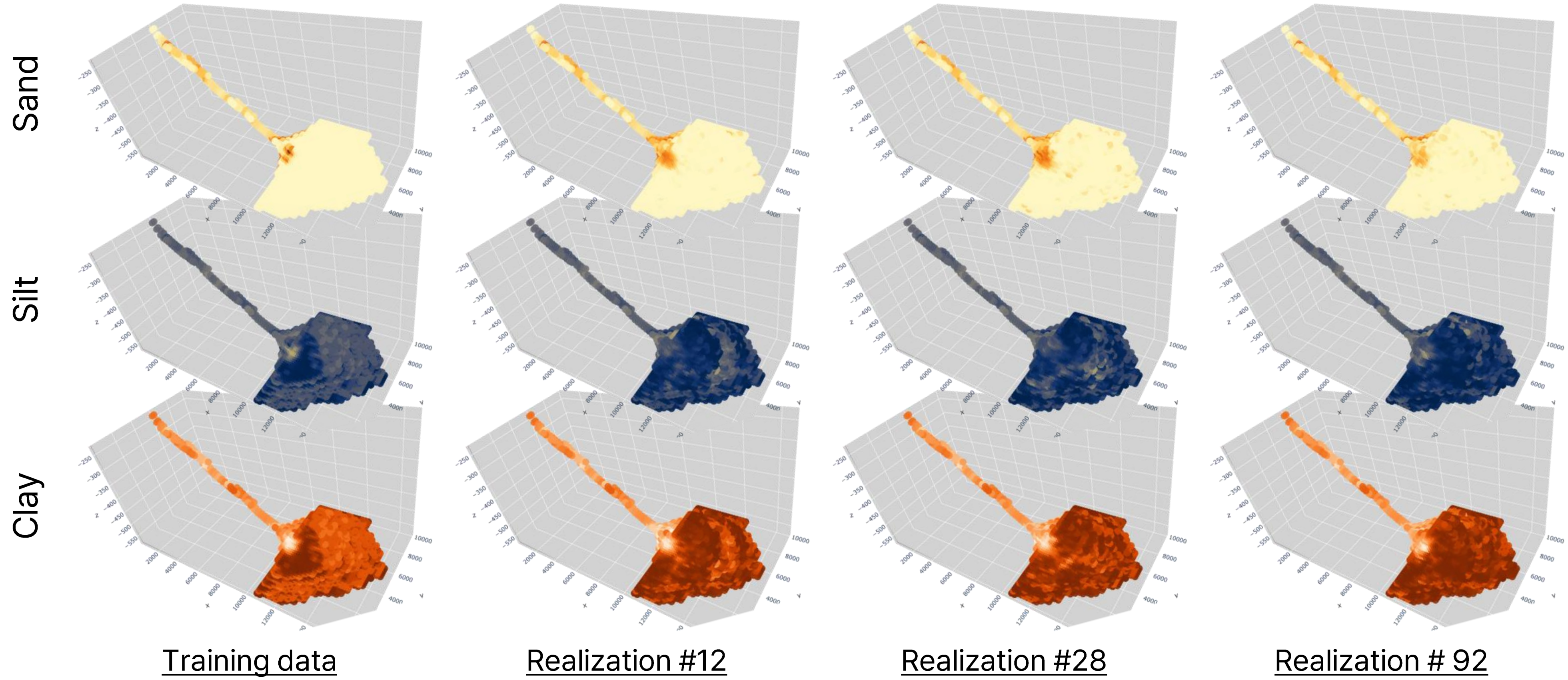


Structural similarity index measure

SinFusion Performance Evaluation

Generation Speed and 3D Geological Visualization

- **training** takes **~261 min**, and post-training **generation** takes **13 s per realization**
- **High** visual **similarity** to the reference data in terms of distribution **continuity** and **spatial smoothness**



Generative AI–Based Geological Model Augmentation and Well Data Integration

Significance

- We propose a **generative AI–based geological modeling** method that retains the strengths of MPS and SFM while addressing their limitations
- It enables **rapid** generation of **many realistic models** and is robust to changes in well number and location
- **Newly acquired well data** can be incorporated immediately to **update models without additional training**

Key Results

- **Spatial connectivity** and overall **geological complexity** are reproduced consistently
- Low-dimensional analysis and variograms show that samples remain **within the reference structural range while preserving diversity**
- **Well conditioning** enables local constraint enforcement near wells while maintaining the global structure

Future Work

- This study focuses on facies modeling; future work will extend to continuous properties such as porosity and permeability
- We will also integrate soft data (e.g., seismic) to develop a multi-data conditional geological modeling framework

References

- Liu, Y. (2006). Using the Snesim program for multiple-point statistical simulation. *Computers & geosciences*, *32*(10), 1544-1563.
- Pyrcz, M. J., & Deutsch, C. V. (2014). Geostatistical reservoir modeling. Oxford university press.
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, *33*, 6840-6851.
- Nikankin, Y., Haim, N., & Irani, M. (2022). SinFusion: Training diffusion models on a single image or video. arXiv preprint arXiv:2211.11743.
- Kozlowski, E., & Davies, A. (2023, June). Recent Advances on the Application of Stratigraphic Forward Models. In *84th EAGE Annual Conference & Exhibition* (Vol. 2023, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.
- Hu, L. Y., Liu, Y., Scheepens, C., Shultz, A. W., & Thompson, R. D. (2014). Multiple-point simulation with an existing reservoir model as training image. *Mathematical Geosciences*, *46*(2), 227-240.
- Zhang, T., McCormick, D., Hurley, N., & Signer, C. (2007, September). Applying multiple-point geostatistics to reservoir modeling–a practical perspective. In *EAGE Conference on Petroleum Geostatistics* (pp. cp-32). European Association of Geoscientists & Engineers.

Thank you