

融合多点地质统计与 Transformer 的分层自回归储层表征 框架

陈麒玉^{1,2,3,*}, 潘忠诚¹, 方洪峰⁴, 陈大颀¹, 刘刚^{1,2,3}

(1. 中国地质大学(武汉)计算机学院, 湖北 武汉, 430074; 2. 智能地学信息处理湖北省重点实验室, 中国地质大学(武汉), 湖北 武汉, 430074; 3. 自然资源信息管理与数字孪生工程软件教育部工程研究中心, 湖北 武汉 430074; 4. 中国地质大学(武汉)未来技术学院, 湖北 武汉 430074)

摘要: 为了解决深度生成模型在硬数据约束下易发生模式崩溃和伪影的问题, 本文提出一种分层自回归生成 (Stratified Autoregressive, SAG) 框架。该框架利用离线训练的 Transformer 架构作为条件分布估计器, 替代多点地质统计的在线搜索与计数过程; 采用三级由粗到细的分层策略, 先定义大尺度全局结构, 再向细尺度传播约束, 规避大网格上的二次计算复杂度。多组实验结果及多维尺度图与变差函数分析显示, 本文方法生成的实现具备多样性, 且准确再现了训练数据的全局统计特征与空间连续性; 直方图交叉量化评估证实了无伪影的高局部模式保真度; 不确定性评估显示不确定度由硬数据点向外逐渐增加, 收敛模式符合地质规律。本文提出的方法在不同数量的硬数据约束下, 其结果保持了空间连续性和样本多样性, 实现了复杂储层结构及物性的准确表征。

关键词: 多点地质统计; 深度生成模型; Transformer; 储层表征; 不确定性评估

中图分类号: P628+.2 收稿日期: 2025-12-15

Stratified Autoregressive (SAG) Framework for Reservoir Characterization: Bridging Multiple-Point Geostatistics and Transformer

CHEN Qiyu^{1,2,3,*}, PAN Zhongcheng¹, FANG Hongfeng⁴, CHEN Dajie¹, LIU Gang^{1,2,3}

(1. School of Computer Science, China University of Geosciences, Wuhan 430074, China; 2. Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China; 3. Engineering Research Center of Natural Resource Information Management and Digital Twin Engineering Software, Ministry of Education, Wuhan 430074, China; 4. School of Future Technology, China University of Geosciences, Wuhan 430074, China)

Abstract: To address the mode collapse and artifacts encountered in deep generative models under hard data constraints, this study aims to develop a robust reservoir characterization approach. A Stratified Autoregressive generation (SAG) method is proposed. This method utilizes an offline-trained Transformer architecture as a conditional distribution estimator to replace the computationally expensive online "search and count" process of MPS. A three-level, coarse-to-fine strategy is adopted to define global structures at large scales first and subsequently propagate constraints to finer scales, thereby avoiding the quadratic computational complexity on large grids. Multiple sets of experiments as well as

通信作者: 陈麒玉 (1990-), 男, 博士, 特任教授, 主要从事多点地质统计学、智能地质表征模拟、数字孪生等方面的教学和科研工作, Email: qiyu.chen@cug.edu.cn, ORCID: <https://orcid.org/0000-0003-3052-9223>

基金项目: 国家自然科学基金青年项目(42172333, 41902304)

multidimensional scaling and variogram analyses indicate that the generated realizations possess diversity and accurately reproduce the global statistics and spatial continuity of the training data. Quantitative assessment using histogram intersection further confirms high local pattern fidelity without artifacts. Uncertainty assessment reveals that uncertainty increases outward from hard data points, showing a convergence pattern consistent with geological laws. The results indicate that the proposed method maintains spatial continuity and realization diversity under varying amounts of hard data constraints, which achieves the accurate characterization of complex reservoir structures and properties.

Keywords: Multiple-point geostatistics; Deep generative model; Transformer; Reservoir characterization; Uncertainty assessment

0 引言

准确表征地下储层对油气与地热开发、地下水可持续管理以及二氧化碳地质封存等重大工程具有关键意义 (Strebelle, 2002; Arts et al., 2004; 袁哲等, 2025)。储层内部的岩相空间分布, 作为控制流体流动和采收率的关键因素, 其精确刻画是所有后续分析的基础 (Gelhar et al., 1992; 唐佳凡等, 2025; 陈麒玉等, 2025)。然而, 地下储层表征本质上是一个观测稀疏且多为间接响应 (如水力应答、地震时移等) 的高不确定性逆问题 (Garabedian et al., 1991; Santos et al., 2018)。任何单一、确定性的地质模型都只是对真实地下的欠约束近似; 基于错误的确定性模型做决策, 可能导致投资失误与风险误判 (Santos et al., 2018; Schiozer et al., 2019)。因此, 主流实践已从寻求单一“真值”转向随机建模: 在满足硬数据 (井点)、软数据 (趋势) 与合理先验的前提下, 生成等可能的模型集合, 并在集合层面开展流动模拟与风险评估, 从而量化不确定性并支撑稳健决策 (Haugen et al., 2006; Peters et al., 2010; Al-Ghattas et al., 2024)。

地质统计学为这一范式提供核心工具。两点地质统计通过变差函数刻画二阶结构, 并能高效条件化井点数据, 典型方法如序贯高斯模拟。但二阶统计难以复现曲流河道、三角洲分流等复杂高阶几何形态, 常出现“斑块或伪影”外观, 与实际连通性相悖 (Matheron, 1963; Deutsch, 2006)。为突破二阶统计的瓶颈, 多点地质统计 (Multiple-Point Statistics, MPS) 应运而生。MPS 引入训练图像 (Training Image, TI) 作为空间结构的先验模型, 通过扫描 TI 中的数据事件来近似高阶的联合概率分布, 并在模拟网格 (Simulation Grid, SG) 中逐点或逐块地复现这些模式 (Guardiano and Srivastava, 1993; Strebelle, 2002a)。MPS 的主要优势体现在: 能够捕捉几何复杂的高阶结构; 以 TI 作为显式先验模型刻画空间特征; 采用非参数统计避免预设分布假设。

在方法演进方面, 早期的方法如 ENESIM (Guardiano and Srivastava, 1993) 因需对每个待模拟点扫描整个 TI, 运行效率受限。Strebelle 和 Cavelius (2014) 提出两步策略以兼顾效率与质量: 首先以“搜索树”构建动态数据集存储从 TI 提取的空间模式; 随后在 SG 上逐点模拟, 按匹配到的模式赋值; 并引入多重网格以先重建长尺度结构、再细化小尺度纹理。此后改进大体分为两类: 一条思路聚焦于数据的组织方式, 如以列表代替树并通过 IMPALA 提升并行性与缓解内存压力 (Straubhaar et al., 2011), Gravey 和 Mariethoz (2020) 在将数据在频域当中利用快速傅里叶变换来高效计算“不匹配图”获得计算成本恒定模拟方式; 以及 (Zuo et al., 2022) 将最临近的模式进行聚类。另一条思路关注 TI 中的快速检索与匹配, 例如, 直接采样法 (Mariethoz et al., 2010) 通过在 TI 中在线搜索并将首个满足距离阈值的样式中心拷贝至 SG, 避免了庞大样式库的预构建; Cui 等人 (2021) 提出并实现了一个混合并行框架能够显著提高模拟效率。尽管 MPS 在复现高阶结构方面取得了成功, 但其发展与应用仍面临挑战: 绝大多数 MPS 算法默认 TI 的统计特征在空间上是平稳的。这难以处理真实地质环境中普遍存在的非平稳特征, 例如河道的宽度、走向、弯曲度等随空间位置发生系统性变化。模拟结果的

质量直接取决于 TI 的质量，在实际应用中，获取一个既能充分代表研究区地质特征，又与目标区域的局部数据（如井点数据）高度吻合的高质量三维 TI，本身就是一项极具挑战性的工作。

与之并行，近年的深度生成模型为学习复杂分布提供了强力工具。生成对抗网络（Generative Adversarial Networks, GANs）可生成外观逼真的地质实现，逐步发展出稳定性更好的变体（Goodfellow et al., 2020; Cui et al., 2024; 陈麒玉等, 2025）；变分自编码器（VAEs）学习到结构良好的潜空间，有利于低维参数化与历史同化（Chen et al., 2023）。然而，尽管深度生成模型在生成逼真地质模型方面表现出色，它们在整合稀疏硬数据时却面临一个普遍难题。强制模型精确匹配观测点，常导致模式坍塌、样本多样性受损，从而系统性地低估不确定性（Bhavsar et al., 2024; Cui et al., 2025b; Lee et al., 2025; Merzoug and Pyrcz, 2025）。此外，点误差主导的损失函数与地质连续性的内在矛盾，容易在观测点附近诱发局部不连续和伪影（Merzoug and Pyrcz, 2025）。为此，研究者们提出了一些针对性策略，例如通过扩展条件数据的影响范围以缓解伪影（Song et al., 2021），利用自表示学习减少生成结果中的伪影（Cui et al., 2025a），采用高斯混合模型来更好地捕捉后验分布的多样性（Bhavsar et al., 2024），或设计专门的映射模块以增强条件信息的传递（Lee et al., 2025）。

为融合二者之长，多种混合 MPS 与深度学习的策略被提出。这些混合策略大体可归为两类：其一为后处理修复，即利用深度学习模型对传统 MPS 生成的、与硬数据不符的局部区域进行修正（Bai and Tahmasebi, 2020）。这种补丁式方法中，深度学习仅作为修正工具，并未参与到核心的随机生成过程。其二为分层建模，即利用深度学习模型从稀疏或多模态数据中学习全局的宏观趋势，为后续的 MPS 提供初始模型或软约束（Hou et al., 2023; Liu et al., 2024）。该模式下，深度学习与 MPS 分别处理宏观与微观尺度，但二者是任务分离的串联关系，而非一个统一的生成框架。

这些策略在一定程度上缓解了硬数据条件化难题，但它们多是在现有深度生成模型的框架内进行修补，并未触及范式上的一系列固有局限。首先，绝大多数模型采用端到端方式一次性生成完整实现，缺乏对模型进行局部修改和迭代更新的灵活性。其次，当需要生成大量条件化实现以进行不确定性量化时，依赖逐案例反演或优化的策略依然计算成本高昂。更关键的是，其固有的“黑箱”特性使得模型难以显式地解释特定位置的赋值依据，这与地质建模中高度依赖邻域一致性检查和地质直觉的工作流程相悖。

这些挑战促使我们重新审视并回归 MPS 的核心思想，并探寻一条将其与深度学习相结合的道路。其关键在于，利用深度神经网络构建一个可学习的条件分布估计器，将传统 MPS 在推断阶段高昂的在线搜索与计数计算成本，摊销至模型的离线训练中。要实现这一目标，需要一个能够高效捕捉数据样板（data template）中复杂空间依赖关系的神经网络架构。近年来，以自注意力机制为核心的 Transformer 模型，在处理序列和图像等结构化数据时展现了强大的上下文理解能力（Parmar et al., 2018）。它能够动态地为邻域内的每个信息点分配权重，精确识别出对中心点预测起决定性作用的关键结构，这与 MPS 从邻域数据事件推断中心点值的内在逻辑契合，使其成为构建该估计器的理想选择。

为此，我们提出了分层自回归（Stratified Autoregressive, SAG）生成框架。该框架与 MPS 在思想上保持同构，但本质上是以一个经过充分训练的神经网络替代了传统的非参数统计过程。据我们所知，本文首次提出了一个将经典 MPS 的非参数思想与深度生成模型相融合的统一框架。通过构建一个可学习的条件分布估计器，将 MPS 高昂的在线搜索与计数的计算成本，摊销到模型的离线训练中。在训练阶段，该模型将整个数据集视为一个广义训练图像，通过自注意力机制自主学习从数据样板到中心值条件分布的复杂映射规则。而在推断时，该模型则转变为一个高效的条件分布预测器，仅需一次前向传播即可为待模拟位置输出高质量的条件概率分布（Conditional Probability Density Function, CPDF）。其关键创新在于：将传统 MPS 在推断阶段高昂的、基于数据计数的非参数 CPDF 估计过程，替换为一个由 Transformer 承担的、可学习的参数化建模问题，并通过将地质

属性场的表征任务分解为一系列从粗到细、在不同尺度上进行的、更易于处理的子任务减少其计算成本。

1 方法

1.1 直觉与动机

设研究域 $\Omega \subset R^d$ ($d = 2$ 或 3) 被离散化为一个规则网格 G 。定义在 G 上的空间属性场是一个随机场 $Z = \{Z_x: x \in G\}$ ，其任意一个实现 $z = \{z_x: x \in G\}$ 代表了一次地质建模的结果。我们的目标是，在给定一组硬数据约束 H 的条件下，从后验概率分布 $p(Z|H)$ 中进行高效且可靠的采样。直接从高维联合分布 $p(Z|H)$ 中采样在计算上是不可行的。联合概率可以被分解为一系列条件概率的乘积：

$$p(Z = z|H) = \prod_{i=1}^{|G|} p(z_{x_i} | z_{x_1}, \dots, z_{x_{i-1}}, H) \quad (1)$$

其中 $(x_1, \dots, x_{|G|})$ 是一条预定义的模拟路径。整个建模过程的核心任务，归结为在路径上的每一步，对未知节点 x 的局部条件概率分布进行准确估计：

$$p(z_x | \mathcal{D}(x)) \quad (2)$$

这里， $\mathcal{D}(x)$ 代表在 x 的局部邻域内，由所有先前已知信息共同构成的数据事件。这个条件分布 $p(z_x | \mathcal{D}(x))$ 编码了地质体在位置 x 处所有可用的局部空间依赖，是生成具有地质真实感模型的关键。经典 MPS 为此提供了非参数解：在 TI 中穷举扫描，搜寻与数据事件匹配的模式，以其中心值的经验频率近似上述 CPDF 并据此抽样赋值。

本文的核心思想是用深度神经网络来取代经典 MPS 的在线的训练图像扫描的过程。该网络直接学习从任意邻域数据事件到中心节点条件概率分布的复杂映射 $p_\theta(z_x | \mathcal{D}(x))$ 。通过这种方式，推断阶段高昂的搜索计算成本被转化为了一次性的离线训练过程。这种学习到的映射实现了一种更为灵活的软匹配，因为它能够从训练数据中学习潜在的地质模式分布规律并进行泛化，从而有效缓解了高维邻域中常见的数据稀疏和模式不匹配问题。然而，直接在原始高分辨率网格上应用 Transformer 这样的自回归模型，会面临巨大的计算挑战。Transformer 自注意力机制的计算复杂度与输入序列长度的平方成正比，像素级的序列长度将导致计算成本过高而不可行。为此，我们提出了 SAG 框架将地质属性场的表征任务分解为一系列从粗到细、在不同尺度上进行的、更易于处理的子任务。

1.2 分层自回归生成

SAG 框架的核心思想是将地质属性场的生成过程分解到多个尺度上。在每个尺度上，模型采用自回归过程，基于动态调度策略按顺序填充未知区域。该框架由三个主要部分组成：多尺度表示将高分辨率地质场分解为由粗到细的层级结构；条件分布估计器（本研究中实现为 Transformer 模块）对条件概率密度函数进行参数化；迭代解码调度通过多次迭代逐步去掩码并生成被掩盖位置的值。

多尺度表示将高分辨率地质属性场 Z 划分为一组层级网格 $\{G^{(k)}\}_{k=0}^L$ 。这里 k 代表任意一个层级， L 表示最高层级。每个层级 k 对应一个特定的地质观察尺度，最粗尺度 $G^{(0)}$ 关注的是宏观沉积格局，其网格单元 $c_i^{(0)}$ 对应的是大型的地质单元。相比之下，最精细尺度 $G^{(L)}$ 对应岩相尺度，其网格单元 z_x 直接代表具体的岩石类型。在给定尺度 k 下，网格为 $G^{(k)} \in R^{H_k \times W_k}$ ，定义本层的块尺寸 p_k （单位为像元），并将 $G^{(k)}$ 划分为不重叠的 $p_k \times p_k$ 小块。由此得到本层序列长度 $L_k = \frac{H_k}{p_k} \cdot \frac{W_k}{p_k}$ 。每个图像块 $c_i^{(k)}$ 被视为一个基本建模单元。

在任意一个地质尺度 k 上，建模过程并非一次性完成，而是通过一个迭代式的空间填充策略逐步构建，是

一个自回归的构建过程。建模开始时，该层级的整个研究区被视为一个待解释的未知域，由掩码 $M_0^{(k)}$ 标识。在随后的每一次迭代 $t = 1, 2, \dots, T_k$ 中，根据预设的调度策略 $S(t)$ ，在每一步选择一部分被掩码的位置进行预测。对于被选中的待建模单元， $k < L$ 时， $M^{(k)}$ （本文中为 Transformer 模型）输出的是传递到下一层的上下文潜在向量表征集合。这些潜在在表征编码了当前尺度的结构性上下文，并与上层传下的指令共同决定下层的生成倾向。当 $k = L$ 时，一个该尺度专属的条件分布估计器 $M^{(k)}$ 会综合所有已知信息，预测其最可能的地质属性。这些信息构成了预测的地质上下文，包括：周边已建模单元的地质属性 $\{f_i^{(k)} \mid c_i^{(k)} \notin M_t^{(k)}\}$ 确保局部空间连续性。来自上一层级 $k - 1$ 的、作为宏观趋势指导的地质指令 $C^{(k)}$ 。井点、岩心等直接观测数据 H ，作为硬约束条件。图 1 展示了 $L=2$ 时的网络架构和单次生成过程。

1.3 条件分布估计器

核心架构遵循标准的掩码建模范式（He et al., 2022; Chen et al., 2026）。具体而言，空间邻域首先被分块并线性嵌入为一系列 token，其中未知位置由一个可学习的 [MASK] token 表示。为了实现层级间的有效约束，来自上层 $k - 1$ 的地质属性向量作为条件前缀拼接到本层的输入序列。硬数据 H 则被保留作为已知值。通过自注意力机制，模型可以整合邻域信息、跨尺度指令与硬数据约束，共同预测被掩码位置的输出。

模型的输出头根据尺度进行配置：对于中间尺度 $k < L$ ，其目标是自回归地生成地质上下文潜在表征 $f_i^{(k)}$ ；而在最精细的岩相尺度 $k = L$ ，则输出 K 个岩相类别的概率分布。我们为三层级联架构 ($L = 2$) 设计的具体网络参数化方案，总结于表 1。

1.4 生成策略

在每个尺度 k 的迭代解码中，采用基于优先级的采样方法与调度式标记揭示相结合的策略，以在保证稳定性的同时提高生成效率与空间连续性。

该方法优先揭示具有最多可用条件信息的位置。在每次迭代 t 中，任何未生成候选位置 x 的优先级计算为其局部数据量 $|D(x)|$:

$$|D(x)| = \left| \left\{ y \in N(x) \mid y \notin M_t^{(k)} \right\} \right| + |H \cap N(x)| \quad (3)$$

其中 $N(x)$ 为 x 的邻域， H 为硬数据集合。本步需要揭示的位置数为 Δ_t （见下式），按 $|D(x)|$ 由高到低选取 Δ_t 个位置，并对其条件分布 $p_\theta^{(k)}(\cdot \mid D(x))$ 进行预测与赋值。

每一步揭示的标记数量 Δ_t 由余弦退火调度控制。设 $seqlen$ 为当前尺度的标记总数， \hat{T} 为最大计划步数。有效步数为 $T = \min(\hat{T}, seqlen)$ 。对于每一步 $t \in [0, T - 1]$ ，定义剩余掩码比例 $r_t = \cos\left(\frac{\pi}{2} \cdot \frac{t+1}{T}\right)$ ，则每次剩余的掩码数量 $m_t = \lfloor seqlen \cdot r_t \rfloor$ 。由此可得 $\Delta_t = \max\{0, m_{t-1} - m_t\}$ ，其中 $m_{-1} = seqlen$ 。

整个策略以自上而下的方式应用于所有尺度。较高尺度首先生成结构信息，作为较低尺度的约束。较低尺度相应地填充细节。

表 1 三层级联架构参数方案

Table 1 Parameter settings of the three-level cascaded architecture

层级	模型	输入单元尺寸	分块尺寸	序列长度	嵌入维度	网络结构
k=0	$M^{(0)}$	64×64	4×4	256	1024	32 层, 16 头
k=1	$M^{(1)}$	4×4	1×1	16	512	8 层, 8 头
k=2	$M^{(2)}$	单一特征向量	不适用	1	512	3 层, 4 头

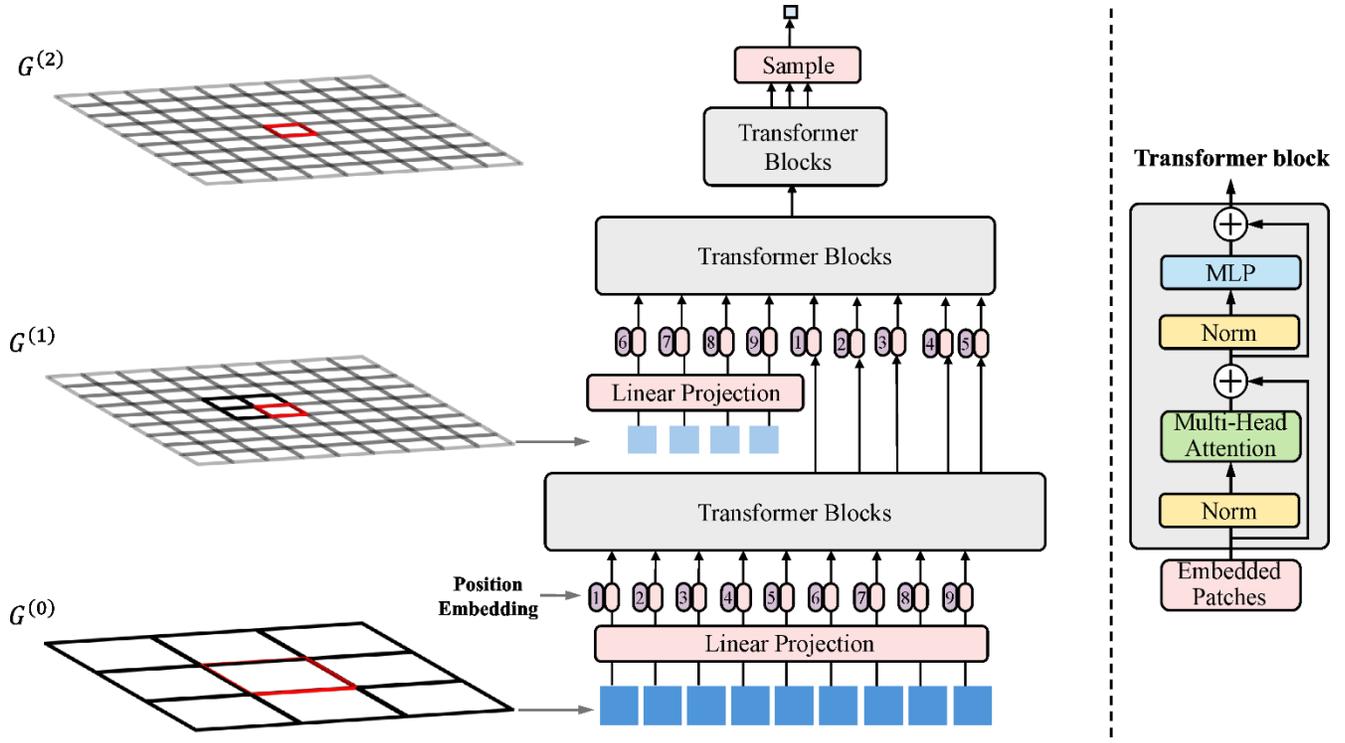


图 1 SAG 中的多尺度自回归生成。左图展示了将高分辨率地质属性场分解为由粗到细的层级结构 ($L=2$)。右图显示了所采用的 Transformer 模块的架构

Fig.1 Multi-scale autoregressive generation in SAG. The left panel illustrates the decomposition of a high-resolution geological attribute field into a coarse-to-fine hierarchy ($L=2$). The right panel displays the architecture of the employed Transformer block

1.5 训练目标

整个 SAG 框架被设计为一个统一的模型，并以端到端的方式进行联合优化。其训练目标是最大化在所有尺度上生成结果的对数似然。为此，我们设计了一个复合损失函数 \mathcal{L}_{SAG} ，它由最精细尺度上的岩相分类损失和所有中间尺度上的地质属性向量重建损失共同构成。

在训练阶段，一个来自训练图像的完整样本 z 被送入模型。该样本在每个尺度 k 上被相应地分块，为该层的模型 $M^{(k)}$ 提供学习目标。训练过程同样遵循一种分层迭代式的掩码预测范式，但在每一步，模型都需要对被掩码的位置进行预测，并与真实值计算损失。

在最精细的岩相尺度 ($k = L$)，模型的最终任务是为每个网格单元 x 预测其正确的岩相类别。因此，该层的损失 $\mathcal{L}_{\text{facies}}$ 定义为标准的交叉熵损失，用于度量预测的概率分布 $p_{\theta}^{(L)}(z_x | C^{(L)})$ 与独热编码的真实岩相标签 z_x^{true} 之间的差异：

$$\mathcal{L}_{\text{facies}} = -E_{x \sim M^{(L)}} \left[\sum_{j=1}^K (z_{x,j}^{\text{true}}) \log p_{\theta}^{(L)}(z_x = j | C^{(L)}) \right] \quad (4)$$

其中， $\mathcal{M}^{(L)}$ 是在该层级被掩码并用于计算损失的位置集合， K 是岩相总数。

对于中间尺度 ($k < L$)，模型 $M^{(k)}$ 的目标是生成能够有效指导下一层建模的地质上下文潜在表征 $f_i^{(k)}$ 。为了监督这些抽象向量的学习，我们在 $k+1$ 层引入了粗尺度的块级重建损失 $\mathcal{L}_{\text{recon}}^{(k)}$ 。在训练时，对任一被掩码块 ($c_i^{(k)}$)，从对应的真值抽取其块级类属指示向量 $y_i^{(k)} \in 0,1^K$ ，其中 K 为岩相属性数目。尺度 k 的条件分布估计器

输出该块的类属概率 $p_i^{(k)} \in \Delta^{K-1}$ 。层级重建项定义为被掩码块上的对数似然损失：

$$\mathcal{L}_{\text{recon}}^{(k)} = \mathbb{E}_{i \in \mathcal{M}^{(k)}} \left[- \langle y_i^{(k)}, \log p_i^{(k)} \rangle \right] \quad (5)$$

其中， $\langle \cdot, \cdot \rangle$ 表示向量内积。 $\mathcal{M}^{(k)}$ 是在尺度 k 上被掩码的位置集合。该项鼓励模型在块级别上复现上游约束所隐含的相属格局，从而为下一尺度的像元级细化提供稳定锚点。

最终的训练目标 \mathcal{L}_{SAG} 是上述所有尺度损失之和：

$$\mathcal{L}_{\text{SAG}} = \mathcal{L}_{\text{facies}} + \sum_{k=1}^{L-1} \mathcal{L}_{\text{recon}}^{(k)} \quad (6)$$

通过最小化这一复合损失，驱动整个级联模型学习从宏观到微观的地质模式，并确保各尺度间的生成过程是协调一致的，这使得模型在学习生成宏观结构的同时，也能精确地刻画微观细节。

2 实验设置与结果分析

2.1 数据集

本研究采用的训练与评估数据源于 (Song et al., 2021) 为其条件生成建模工作而开发的公开合成地质相数据集。该数据集包含了 35,640 个二维地质相模型，每个模型的分辨率均为 64×64 像素。这些模型通过过程法模拟生成，旨在复现复杂的曲流河沉积系统，其中包含了三种主要岩相：河道砂体、堤岸砂体与河道间泥岩，数据集示例见图 2。

我们遵循了原数据划分策略，从总数据集中，我们随机选择 32,640 个模型作为训练集，用于 SAG 模型的端到端学习。剩余的 3,000 个模型则构成独立的测试集，专门用于评估模型的泛化能力、生成样本的质量。

2.2 实现细节

我们提出的分层自回归 (SAG) 模型及所有实验均基于 PyTorch 深度学习框架实现。模型训练在一台配备 3 块 NVIDIA 3090 24GB GPU 的服务器上进行。我们采用 AdamW 优化器，并结合余弦退火学习率调度策略对模型进行优化。具体的模型参数配置已在表 1 中详述。

2.3 视觉评估与分布一致性分析

对 SAG 生成 1000 个样本随机采样，视觉评估表明结果在宏观和微观尺度上都很逼真 (图 3)。为了定量评估在不同数量条件下生成的样本质量，我们采用多维尺度 (Multidimensional Scaling, MDS) 分析来评估其分布。结果展示在图 4 中。在此分析中，我们首先将每个真实图像和生成图像转换为一维分类标签向量。然后我们计算所有样本之间的成对汉明距离，并使用 MDS 将生成的高维距离矩阵投影到二维空间进行可视化。测试集和生成集的点云分布显示出大量重叠，且生成集并未坍塌成少数几个簇。

2.4 岩相比例分析

我们首先检验模型在全局成分层面是否保持与真实数据一致的岩相配额。在测试抽取图片随机采样 N 个条件点，并且使用 SAG 进行模拟 1000 次，再对这 1000 个样本进行相比例计算。Reference 表示训练集总体的相比例。图 5 展示了在不同约束强度 N (此处 N 表示每个样本注入的井点，含 $N = 0$ 的无条件情形) 下的对比结果。可以看到，三类岩相的生成比例整体与参考集对齐，在增加约束 ($N = 1 \sim 16$) 时，堤岸砂体与河道砂体的配额保持稳定，泥岩在高 N 端出现轻微下调趋势，但幅度很小；这说明 SAG 在注入局部硬约束以保证可条件化采样的同时，不会牺牲全局配额一致性。

2.5 指示变差函数

为了评估空间连续性和各向异性的再现情况，我们检查了三种相的指示变差函数。该实验比较了不同条件数据密度 ($N \in \{0, 1, 2, 4, 8, 16\}$) 下的生成集和训练集。

结果以四分位距 (Interquartile Range, IQR) 重叠热图 (图 6) 和分面曲线图 (图 7) 的形式呈现。对于所有相和方向，生成变差函数曲线的均值始终落在参考集的 IQR 带内。在短滞后距处，曲线的初始斜率和形状与参考相匹配，这表明相边界的清晰度得以保留。在长滞后距处，基台值接近参考值，与由相比例确定的理论最大值一致。

通过 IQR 重叠百分比进行量化，定向和全向情况下的绝大多数网格点的值在 90% 到 98% 之间。特定组合观察到的最低重叠率约为 83%。未见系统性方向偏差或随 N 的退化趋势，但随着条件点的增加对岩相分布形成了一定约束使得重叠度变低。综合来看，SAG 在有和无条件约束下可稳定重现训练集的二阶统计与方向结构，并在不同条件点数量下保持一致的空间连续性刻画，这与前述视觉评估相吻合。

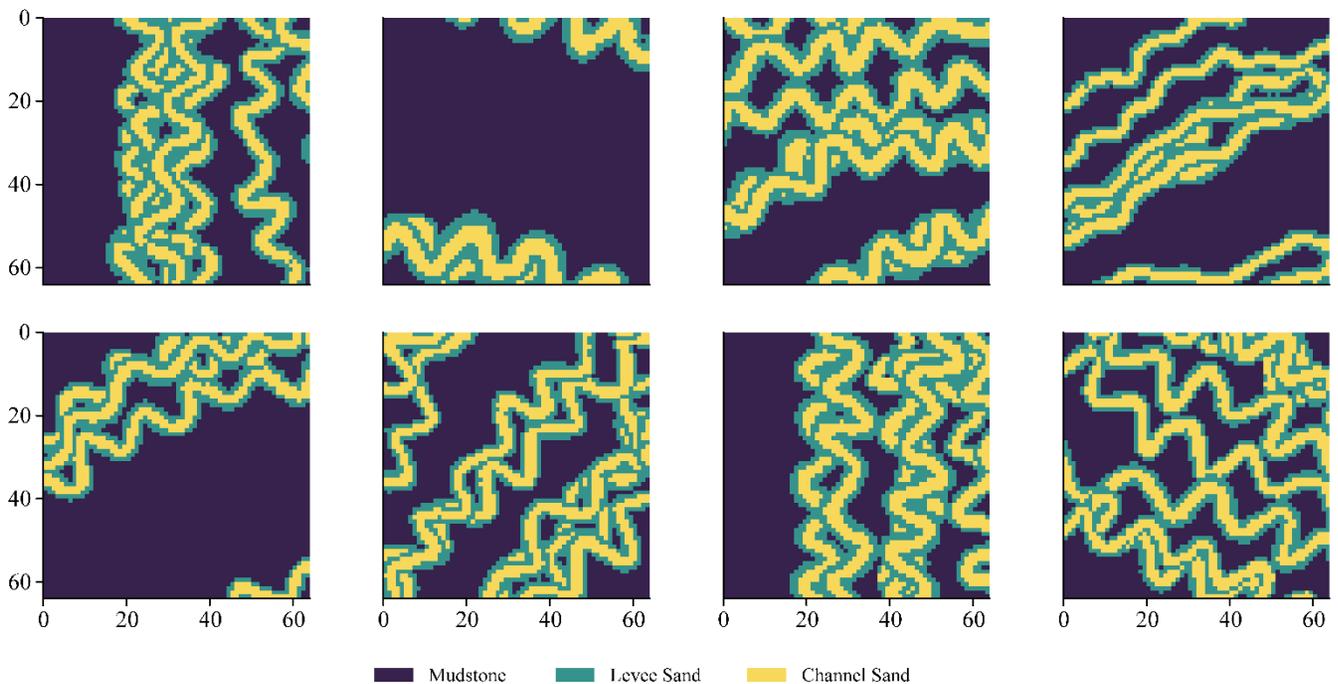


图 2 从训练集中随机选取的储层相样本。每个样本为 64×64 个单元，坐标轴一致；颜色表示河道砂（黄色）、天然堤砂（蓝绿色）和泥岩（紫色）

Fig.2 Random chosen reservoir facies samples from the training set. Each sample is 64×64 cells with consistent axes. Colors denote channel sand (yellow), levee sand (teal), and mudstone (purple)

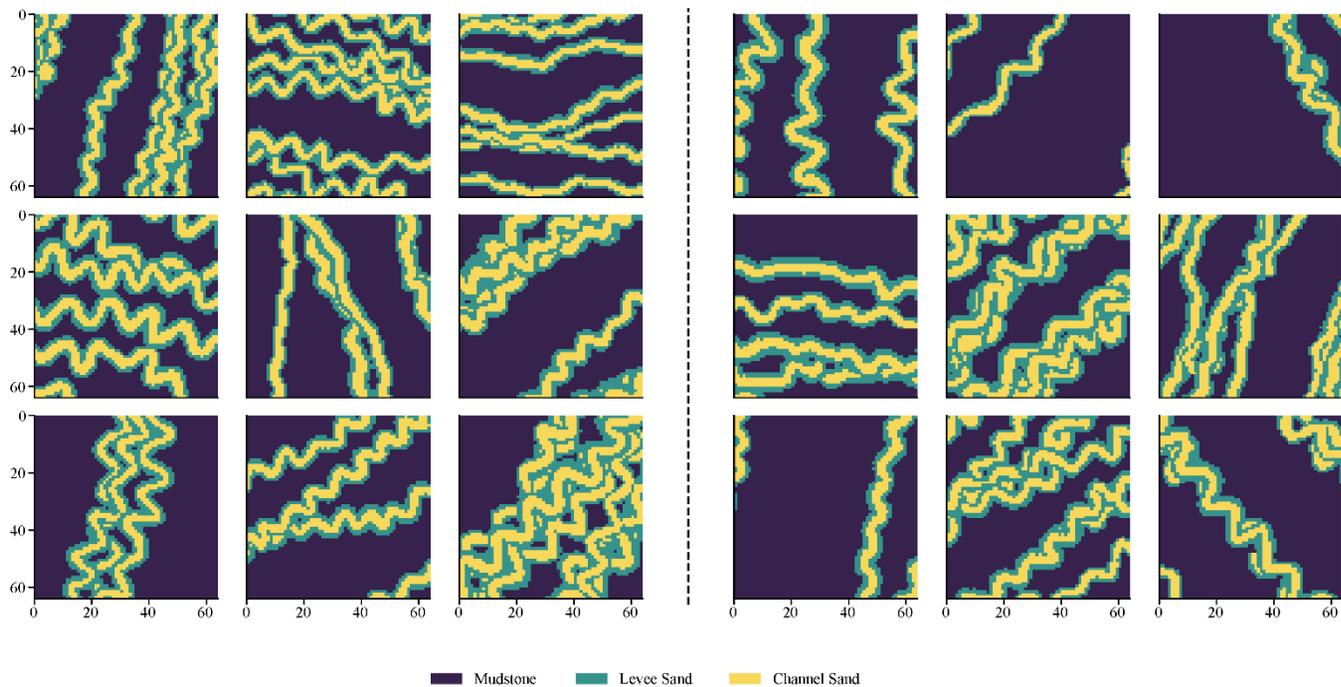


图 3 真实训练样本（左）与 SAG 模型生成的图像（右）的并排比较

Fig.3 Side-by-side comparison of real training samples (left) and images generated by the SAG model (right)

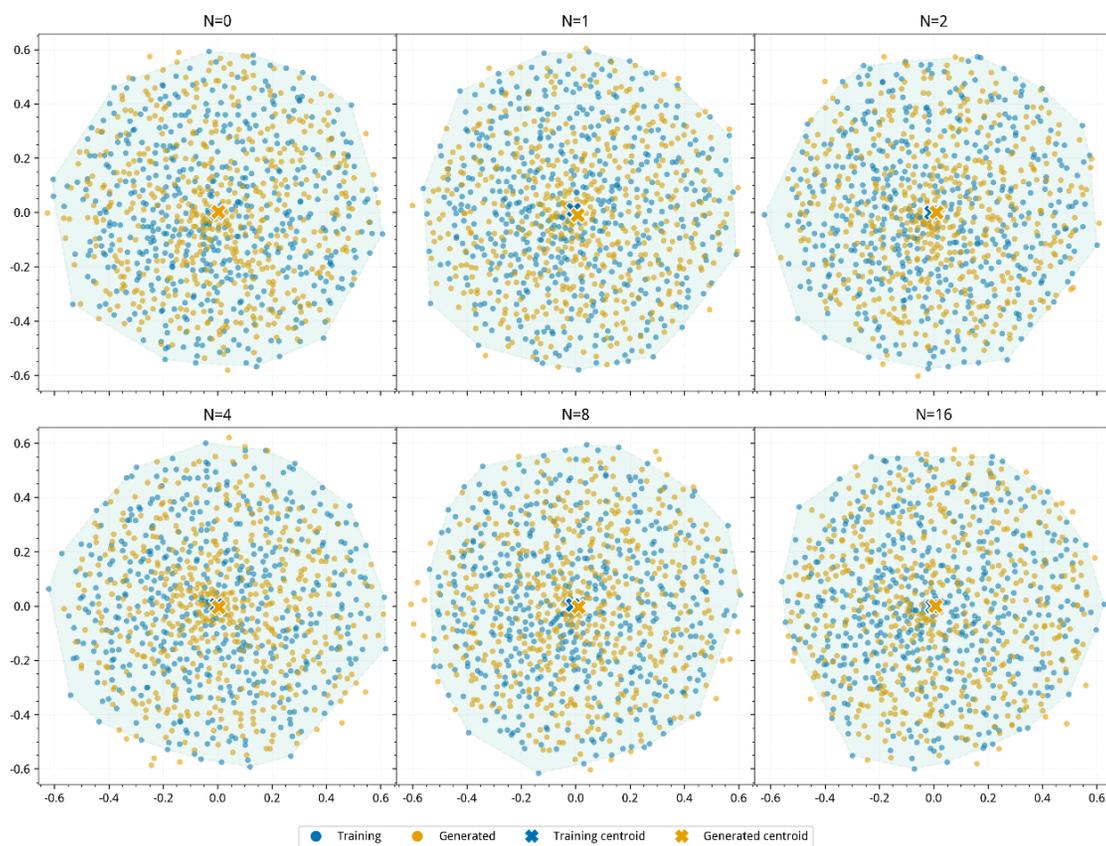


图 4 真实样本和生成样本分布的 MDS 可视化。N 表示条件数据的数量

Fig.4 MDS visualization of the distributions of real and generated samples. N denotes the number of conditioning data points

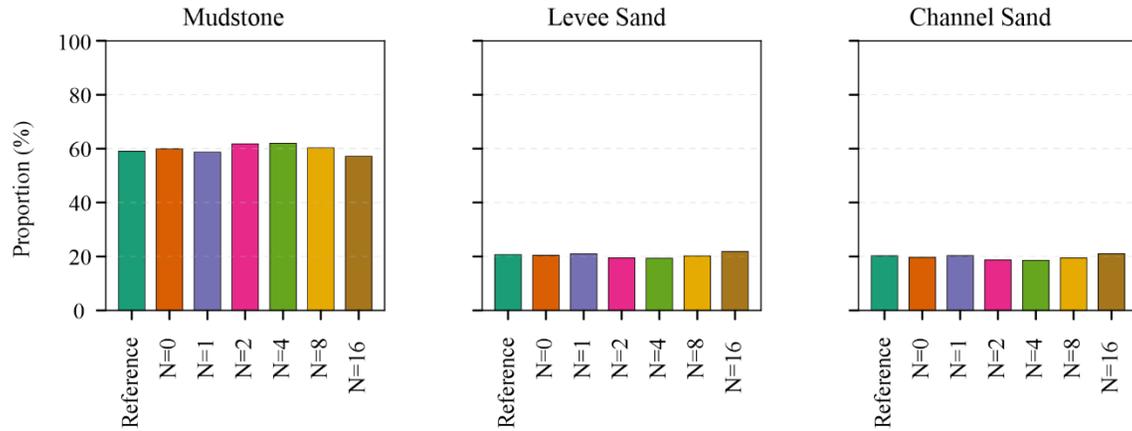


图 5 不同条件点数量 (N) 下生成样本与参考集的相比例对比。分析涵盖了无条件情况 ($N = 0$) 和 N 从 1 增加到 16 的条件模拟，证明了所有三种岩相的生成比例与参考值紧密一致

Fig.5 Facies proportions of generated samples versus the reference set across varying numbers of conditioning points (N). The analysis includes the unconditional case ($N = 0$) and conditional simulations with N from 1 to 16, showing close agreement with the reference proportions

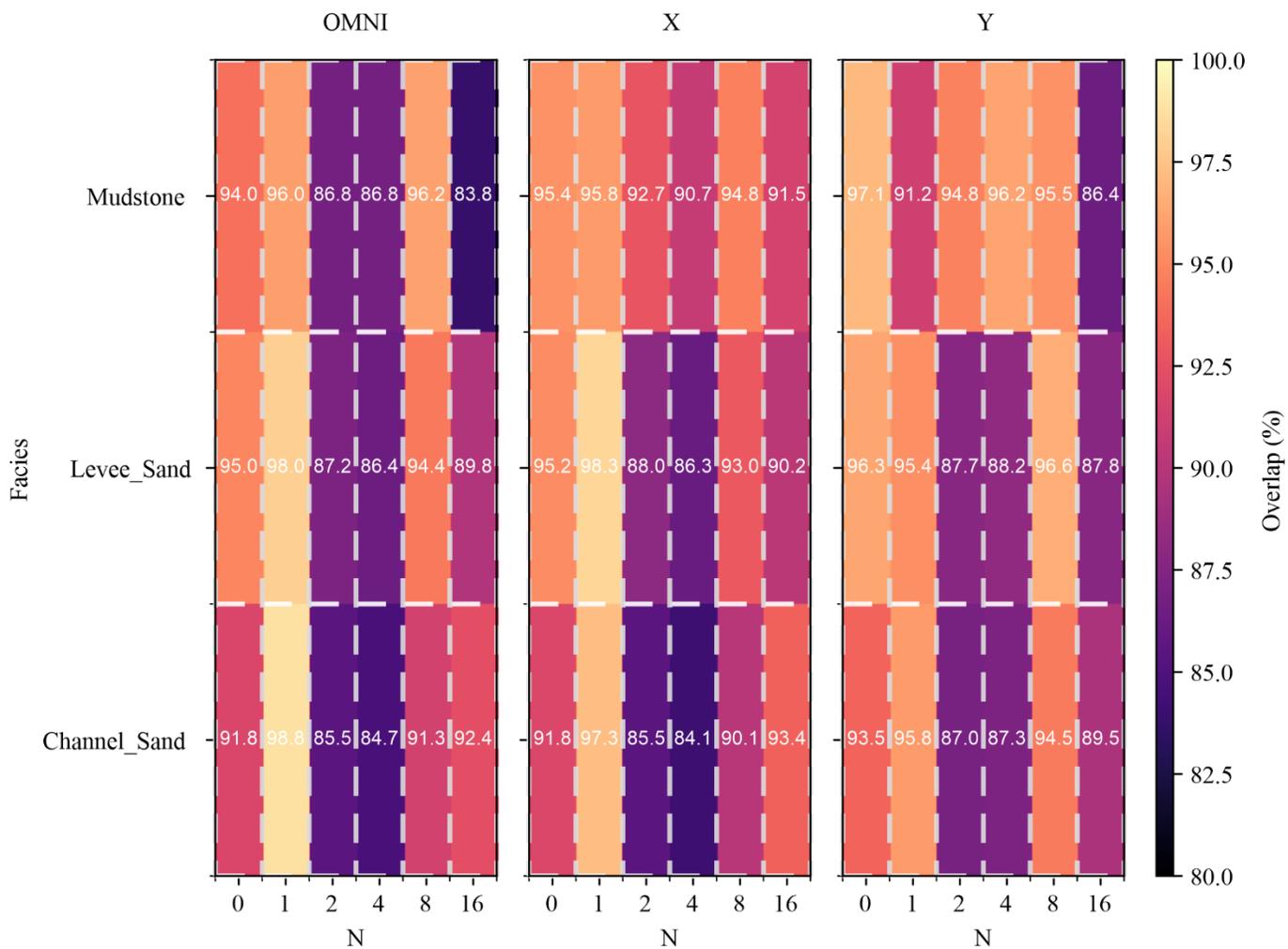


图 6 生成的和参考的指示变差函数之间的 IQR 重叠百分比热图

Fig.6 Heatmap of the IQR overlap percentage between generated and reference indicator variograms

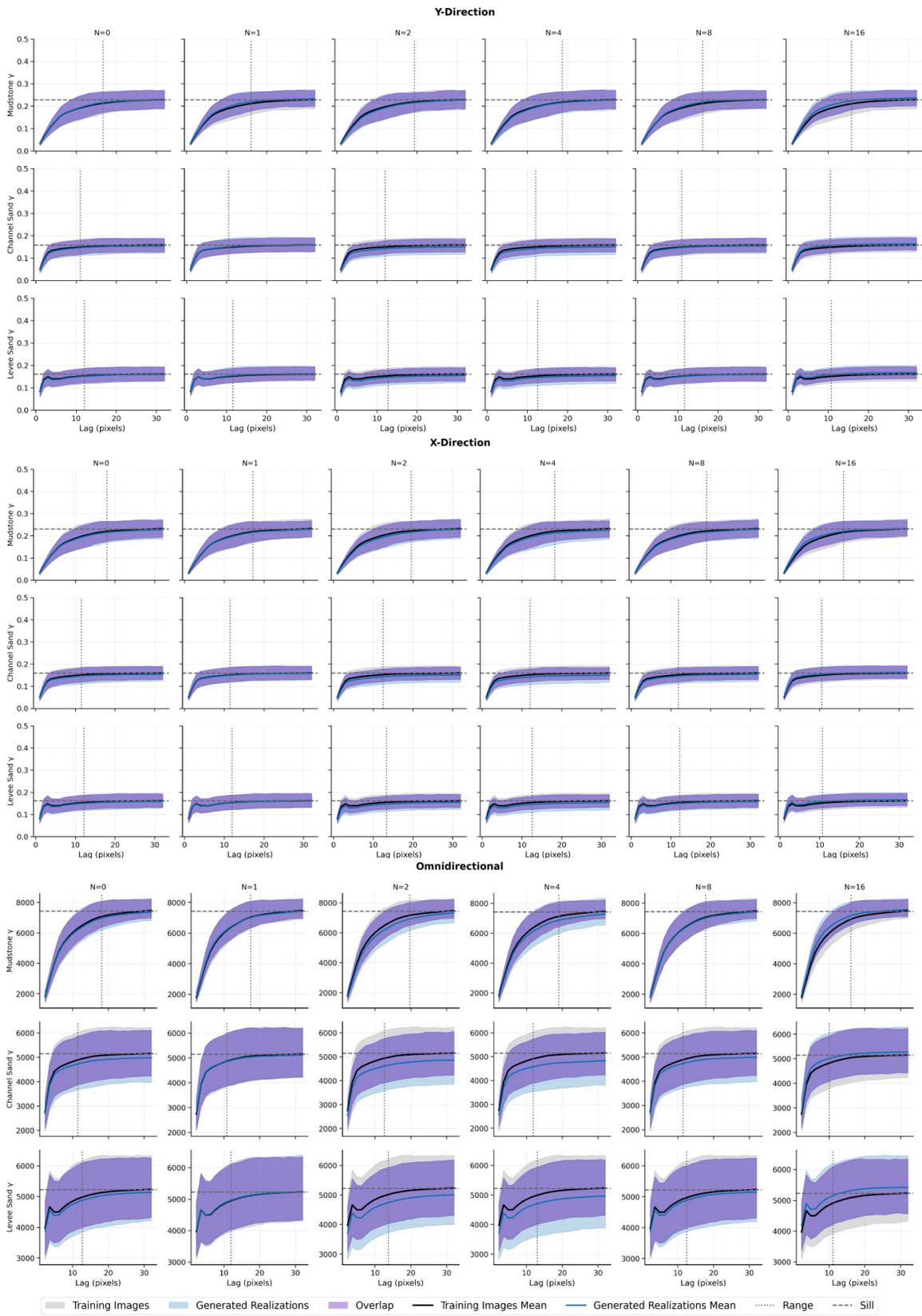


图 7 生成数据和训练集的指示变差函数曲线比较

Fig.7 Comparison of indicator variogram curves for generated and reference sets

2.6 地质模式一致性检查

为了客观且定量地评估我们生成的岩相模型对原始训练图像中地质模式的再现能力，我们参考了由 Boisvert 等人 (2007) 提出的 MPS 验证方法。核心思想是比较每个实现与原始训练图像之间的关键统计特征差异，差异越小，代表模型的还原度越高。具体做法是：对像每一类岩相 k ，在无条件模拟实现和训练图像中以包含中心点及其上下左右四个相邻点模板在整幅图像上滑动，有条件模拟仅采样以条件点为中心的十字形模板以确认判断条件点附近是否出现伪影，统计所有由该模板定义的局部模式出现的频率，得到训练图像的模式分布向量 $p^{(k)}$ 与生成样本的模式分布 $\hat{p}^{(k)}$ 。两者的一致性用直方图相交度 (Histogram Intersection, HI)。

$$HI(p, \hat{p}) = \sum_i \min(p_i, \hat{p}_i) \in [0,1] \quad (7)$$

量化值越接近 1 表示模式越一致。为兼顾统计稳健性与可视化可读性，我们将模式按训练图像的频度从高到低排序，报告模式频度曲线与 HI 的 Bootstrap 95% 置信区间。如图 8 (a-c) 所示，三类岩相的模式频度曲线与训练图像基线几乎重合：最常见的主模式占比匹配良好，低频复合模态的权重亦保持在合理区间。对应的 HI (均值及 95% 置信区间) 在所有条件点规模 $N \in \{0,1,2,4,8,16\}$ 下均接近 1.0 (图 8)，表明生成的样本的中以条件点为中心的模式与训练图像在模式组成趋于一致；随着 N 增大，HI 基本稳定，个别类别在最大 N 处有轻微回落但仍处于高位且置信区间重叠，说明条件点周围未产生伪影 (图 9)。

2.7 不确定性评估

从测试集中随机选取一幅 64×64 地质相图作为参考，并在其上随机采样 16 个像素作为条件点，在相同的条件点集合下，分别取 $N \in \{1,2,4,8,16\}$ 个条件点参与条件化采样；每个 N 下独立生成 100 次实现，用于统计逐像素的不确定性。为表征不确定性，我们对 100 次实现计算：Shannon 熵 $H = -\sum_k p_k \log p_k$ (三类岩相，单位比特)，以及逐像素标准差 (将类别标签视作数值以衡量类别波动)，对应的可视化见标准差网格图与熵网格图。条件点处不确定性近零，在所有 N 下，条件点像素的熵与标准差均降低 (图 10 中呈深色小斑)，表明硬约束被严格满足。随着 N 增大，低熵的影响区由条件点向周边扩展并逐渐连通，表明不确定性呈由点及面的衰减。这说明模型能够将局部约束通过学习到的空间依赖传播到更大尺度。远离条件点且位于相界的像素，熵接近三分类的上界，反映了几何结构与相界位置的不确定性主导。标准差与熵在条件点附近与沿连通通道的不确定性共同下降，而在未被约束的宽阔背景区仍然保持较高水平。从 $N = 1$ 到 $N = 16$ ，全图平均熵与平均标准差随 N 单调降低；尤其当条件点在同一空间子区密集时，该子区的不确定性显著塌缩，显示出模型对地质连通性先验的有效利用。总体而言，SAG 在就地条件化和先粗后细的生成范式下，能将有限点约束有效外推到结构尺度，呈现出符合地质直觉的不确定性收敛行为。

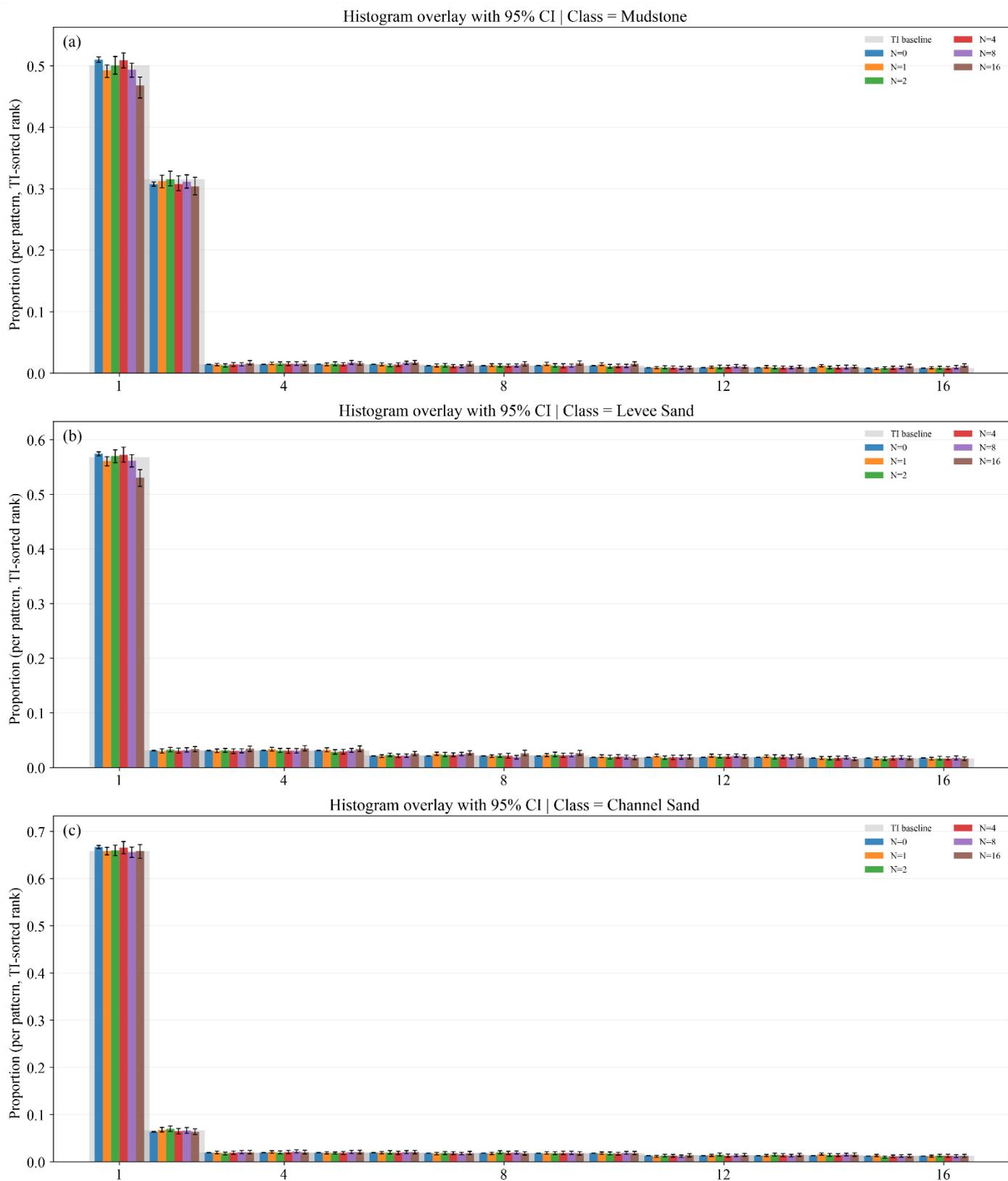


图 8 三种相类别的模板模式频率分布比较：河道砂体 (a)、堤岸砂体 (b) 与河道间泥岩 (c)。不同条件点数量 N 的模拟分布叠加在训练图像基线上。误差条代表 95% 置信区间

Fig.8 Comparison of template pattern frequency distributions for the three facies classes: Channel Sand (a), Levee Sand (b), and Mudstone (c). Distributions for simulations with different numbers of conditioning points

(N) are overlaid on the training-image baseline. Error bars represent 95% confidence intervals

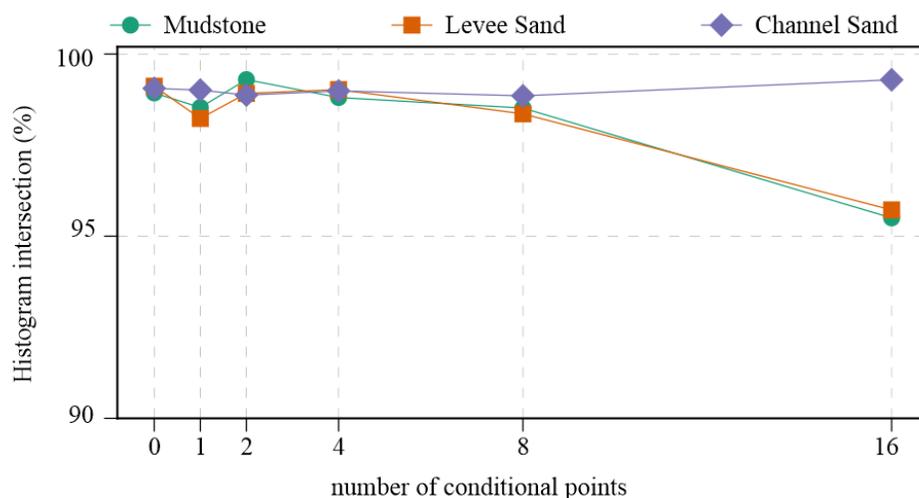


图 9 使用直方图交叉度量对每个相类别进行模式一致性定量评估。图表显示了不同条件点数量下模拟的平均直方图交叉值

Fig.9 Quantitative evaluation of pattern consistency using the histogram intersection metric for each facies class. The plots show the mean histogram intersection values for simulations with different numbers of conditioning points

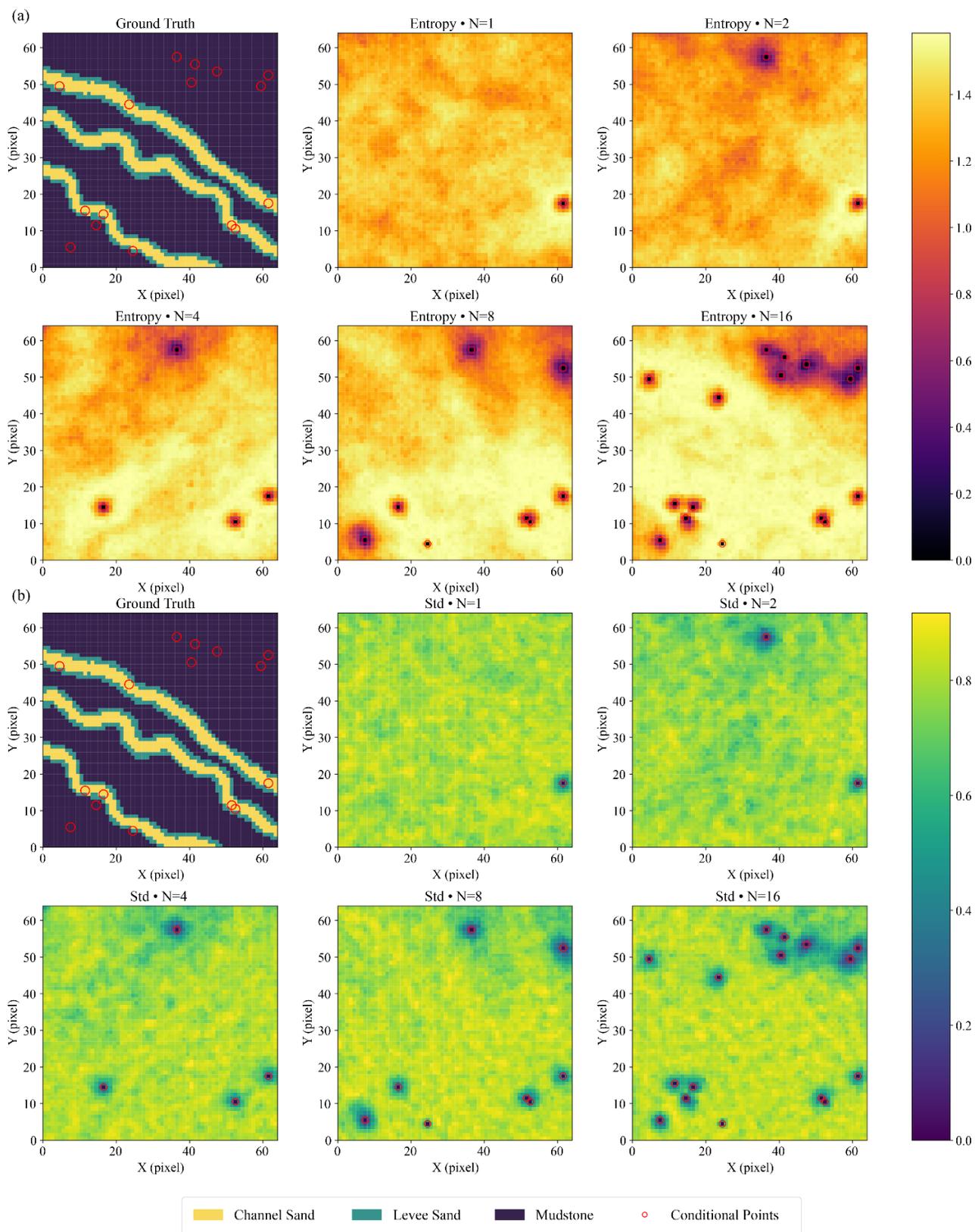


图 10 来自具有不同条件点数量 N 的 100 个独立实现的熵网格图 (a) 和标准差 (b)

Fig.10 Entropy maps (a) and standard deviation maps (b) from 100 independent realizations with varying numbers of conditioning points (N)

3 讨论

3.1 与现有技术的比较

经典 MPS 的核心思想在于：给定一个局部邻域的数据事件，通过在 TI 中进行穷举搜索与计数，来非参数地构造中心未知点的条件概率分布，并依据该分布进行抽样赋值（Guardiano and Srivastava, 1993; Strebelle, 2002b）。MPS 的概率估计严格受限于 TI 的有限尺寸，在处理高维复杂模式时常面临数据稀疏导致的统计失真。而 SAG 利用神经网络的泛化能力，通过学习到的软映射关系，能够在未见过的数据事件下依然输出鲁棒的概率预测，这本质上解决了 MPS 对大 TI 的依赖瓶颈。

现有的 MPS 与深度学习融合方法，仅在使用 MPS 之前用深度学习模型生成宏观趋势或利用深度学习进行后处理。这些方法中，深度学习仅作为 MPS 的辅助工具，传统 MPS 基于训练图像的“搜索与计数”核心算法并未发生改变，因此依然继承了搜索空间有限、计算效率随邻域指数增长以及对平稳性假设敏感等固有缺陷。

与近年来兴起的端到端深度生成模型（如 GANs、VAEs）相比，SAG 的差异体现在生成的粒度与可控性上。多数端到端模型倾向于一次性输出完整的地质实现，当需要强制匹配稀疏的硬数据约束（如井点）时，常因其全局优化的本质而引发模式坍塌、样本多样性降低或在约束点附近产生局部伪影。相反，SAG 通过其掩码自回归机制与先粗后细的级联生成流程，显式地构建了局部 CPDF，实现了逐点式构建。这不仅使得硬数据约束可以被无缝、精确地整合，更允许在任意尺度和任意子区域进行局部重采样与迭代修正。这种显式的概率建模还提供了像素级的不确定性度量，使得生成过程透明可控，而非 GAN 式的黑箱映射。

3.2 推断效率与复杂度

在直接的像素级 Transformer 中，自注意力的时间与显存复杂度通常与序列长度（ $seqlen$ ）的平方成正比。SAG 以两种机制缓解该问题：

（1）分块和降序列：在粗尺度以 $(p_k \times p_k)$ 块作为 *token*，序列长度由像素数降至块数；中尺度进一步以子块汇聚向量表达区域上下文，至细尺度才回落到像素级。

（2）余弦退火式揭示调度：每步只对一部分掩码位置进行预测，且早期小批量、高置信，后期平滑增加揭示规模，使得有效计算集中在后段。

若以每层嵌入维度 d 与头数 h 计，单步注意力的典型复杂度为 $O(seqlen^2 \cdot d/h)$ 。SAG 在粗尺度显著减少 $seqlen$ ，并通过有限步 $T_k \leq seqlen$ 的调度使总复杂度低于直接像素级建模。与无需训练的非参数 MPS 相比，SAG 需要一次性的训练成本与 GPU 资源，但在批量推断与局部重建情形下具有明显的时延优势：同一套已训练参数可快速产生大量等可能实现，适用于历史拟合、开发方案优选与风险对冲等需要蒙特卡罗样本的大规模工作流。

3.3 条件点密度对 SAG 生成机制的影响

SAG 框架在不同数量的硬数据约束下均能保持较高质量的生成结果。然而，我们也观察到，当条件点密度（ N ）增高时，IQR 重叠和直方图一致性出现了轻微的下落趋势。这一现象并非模型能力的衰减，而是其内部多重机制在强约束下协同作用的综合体现。

从贝叶斯推断的视角看，增加条件点的本质是利用更多观测数据来约束先验模型，从而使后验概率分布 $p(Z|H)$ 更加集中。这意味着，随着 N 的增加，等可能的地质实现集合会收窄，样本间的差异性（即不确定性）会合理地降低。

这种统计指标的微小变化，其根源可追溯至 SAG 所采用的先粗后细分层生成策略。粗尺度首先生成全局

结构骨架，作为细尺度生成的宏观趋势指令。硬数据作为不可违背的锚点，直接作用于最精细的岩相尺度。当 N 很大时，粗尺度生成的宏观趋势与密集的细尺度硬数据点所隐含的局部结构之间，出现不一致的可能性会增加。在这种情况下，细尺度模型在进行自回归填充时，必须在遵循上层宏观指令的同时，精确匹配所有硬数据点。为了修复这种潜在的局部不一致，模型可能会生成一些在无条件先验中频率较低的过渡模式。这解释了在高 N 时，模式一致性 HI 值会出现轻微回落。重要的是，由于 SAG 的序贯式生成特性，这种修复是局部的、渐进式的，并不会像端到端模型那样引发全局的模式坍塌或伪影。

关于岩相比例在高 N 出现的轻微波动，也可能是条件点采样偏置的影响。我们的条件点是从多个参考图像中随机抽取的。当 N 较小时，样本的代表性有限；但当 N 很大时，这组条件点的岩相比例会非常接近其所在的参考图像的真实比例，而这个比例未必与整个训练集的总体平均比例完全一致。因此，模型为了匹配这些条件点而调整全局相比例，是贝叶斯条件化的一个正确且预期的行为，而非模型失真的表现。

总的来看，在高条件点密度下观察到的部分指标轻微下降，是 SAG 框架在强约束下，其后验收缩、跨尺度协调以及对采样偏置等多种机制共同作用的综合体现。这恰恰证明了该框架能够有效地将硬数据信息整合到其生成过程中，并合理地收缩不确定性空间。

4 结论

本文提出了一种分层自回归生成框架，通过引入可学习的参数化条件分布估计器，对经典 MPS 的核心进行了重构。该框架将传统 MPS 在推断阶段高昂的在线搜索与计数成本成功摊销至模型的离线训练中。我们提出的 SAG 框架通过两个关键设计实现了这一重构。其一，以 Transformer 为核心的可学习条件分布估计器，它不仅通过泛化能力克服了传统 MPS 在高维邻域下面临的模式稀疏性难题，还将推断过程简化为高效的前向传播。其二，从粗到细的分层生成策略，该策略将单一的高维建模问题分解为计算上可行的多尺度子任务，有效规避了标准 Transformer 在处理大规模网格时的二次方复杂度瓶颈，同时模拟了地质体“由骨架到细节”的层级特征。

虽然 SAG 框架展示了潜力，但其目前的形式仍存在局限性，并为未来工作指明了方向。当前模型专注于 2D 离散相，将其扩展到 3D 环境和连续属性场将需要内存效率更高的架构。此外，作为一种数据驱动的方法，其泛化能力与训练数据的多样性相关，且其前期训练成本对于某些应用可能是一个实际约束。

利益冲突声明/ Statement of Conflict of Interest: 所有作者声明无利益冲突。

References

- [1] Al-Ghattas,O.,Bao,J.,Sanz-Alonso,D.,2024.Ensemble Kalman Filters with Resampling.*SIAM/ASA Journal on Uncertainty Quantification*,12(2): 411–441. doi: 10.1137/23M1594935
- [2] Arts,R.,Eiken,O.,Chadwick,A.,et al.,2004.Monitoring of CO2 Injected at Sleipner Using Time-Lapse Seismic Data.*Energy*,29(9): 1383–1392. doi: 10.1016/j.energy.2004.03.072
- [3] Bai,T.,Tahmasebi,P.,2020.Hybrid Geological Modeling: Combining Machine Learning and Multiple-Point Statistics.*Computers & Geosciences*,142: 104519. doi: 10.1016/j.cageo.2020.104519
- [4] Bhavsar,F.,Desassis,N.,Ors,F.,et al.,2024.A Stable Deep Adversarial Learning Approach for Geological Facies Generation.*Computers & Geosciences*,190: 105638. doi: 10.1016/j.cageo.2024.105638
- [5] Boisvert,J.B.,Pyrcz,M.J.,Deutsch,C.V.,2007.Multiple-Point Statistics for Training Image Selection.*Natural*

- Resources Research*,16(4): 313–321. doi: 10.1007/s11053-008-9058-9
- [6] Chen,D.,Chen,Q.,Cui,Z.,et al.,2023.SA-VAE: a novel approach for reservoir characterization based on variational auto-encoder and selective attention mechanism.*Earth Science Informatics*,16(4): 3283–3301. doi: 10.1007/s12145-023-01095-4
- [7] Chen,Q.,Zhou,R.,Chen,D.,et al.,2026.A Conditional Masked Autoencoder Network Based on Efficient Multiple-Head Self-Attention for Characterizing Heterogeneous Reservoirs.*Expert Systems with Applications*,296: 128973. doi: 10.1016/j.eswa.2025.128973
- [8] Chen,Q.,Zhang,R.,Cui,Z.,et al.,2025a. Multi-Scale Three-Dimensional Geological Model Reconstruction Method Based on Lap-SAGAN.*Earth Science Frontiers*,32: 1–30 (in Chinese with English abstract).
- [9] Chen,Q.,Xun,L.,Cui,Z.,et al.,2025b. Recent Progress and Development Trends of Three-Dimensional Geological Modeling.*Bulletin of Geological Science and Technology*,44(3): 373–387 (in Chinese with English abstract).
- [10] Chen,Z.,2006.Computational methods for multiphase flows in porous media.Society for Industrial and Applied Mathematics,Philadelphia.
- [11] Cui,Z.,Chen,Q.,Jiang,S.,et al.,2025a.Interpretable Deep Learning-Based Characterization of Hydrogeological Structures with Self-Representation Learning.*Journal of Hydrology*,662: 133943. doi: 10.1016/j.jhydrol.2025.133943
- [12] Cui,Z.,Chen,Q.,Liu,G.,et al.,2024.SA - RelayGANs: A Novel Framework for the Characterization of Complex Hydrological Structures Based on GANs and Self - Attention Mechanism.*Water Resources Research*,60(1): e2023WR035932. doi: 10.1029/2023WR035932
- [13] Cui,Z.,Chen,Q.,Liu,G.,et al.,2021.Hybrid Parallel Framework for Multiple-Point Geostatistics on Tianhe-2: A Robust Solution for Large-Scale Simulation.*Computers & Geosciences*,157: 104923. doi: 10.1016/j.cageo.2021.104923
- [14] Cui,Z.,Jiang,S.,Chen,Q.,et al.,2025b.Characterization of Reservoir Structures With Knowledge-Informed Neural Network.*SPE Journal*,30(08): 4469–4486. doi: 10.2118/228279-PA
- [15] Deutsch,C.V.,2006.A Sequential Indicator Simulation Program for Categorical Variables with Point and Block Data: BlockSIS.*Computers & Geosciences*,32(10): 1669–1681. doi: 10.1016/j.cageo.2006.03.005
- [16] Garabedian,S.P.,LeBlanc,D.R.,Gelhar,L.W.,et al.,1991.Large - scale Natural Gradient Tracer Test in Sand and Gravel, Cape Cod, Massachusetts: 2. Analysis of Spatial Moments for a Nonreactive Tracer.*Water Resources Research*,27(5): 911–924. doi: 10.1029/91WR00242
- [17] Gelhar,L.W.,Welty,C.,Rehfeldt,K.R.,1992.A Critical Review of Data on Field - scale Dispersion in Aquifers.*Water Resources Research*,28(7): 1955–1974. doi: 10.1029/92WR00607
- [18] Goodfellow,I.,Pouget-Abadie,J.,Mirza,M.,et al.,2020.Generative Adversarial Networks.*Communications of the ACM*,63(11): 139–144. doi: 10.1145/3422622
- [19] Gravey,M.,Mariethoz,G.,2020.QuickSampling v1.0: A Robust and Simplified Pixel-Based Multiple-Point Simulation Approach.*Geoscientific Model Development*,13(6): 2611–2630. doi: 10.5194/gmd-13-2611-2020
- [20] Guardiano,F.B.,Srivastava,R.M.,1993.Multivariate Geostatistics: Beyond Bivariate Moments.In:Soares,A.,ed.,Geostatistics Tróia '92: Volume 1.Springer Netherlands,Dordrecht,133–144.

- [21] Haugen, V., Natvik, L.-J., Evensen, G., et al., 2006. History Matching Using the Ensemble Kalman Filter on a North Sea Field Case. In: SPE Annual Technical Conference and Exhibition. OnePetro, SPE-102430-MS.
- [22] He, K., Chen, X., Xie, S., et al., 2022. Masked Autoencoders Are Scalable Vision Learners. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 15979–15988.
- [23] Hou, W., Chen, Y., Liu, H., et al., 2023. Reconstructing Three-Dimensional Geological Structures by the Multiple-Point Statistics Method Coupled with a Deep Neural Network: A Case Study of a Metro Station in Guangzhou, China. *Tunnelling and Underground Space Technology*, 136: 105089. doi: 10.1016/j.tust.2023.105089
- [24] Lee, D., Ovanger, O., Eidsvik, J., et al., 2025. Latent Diffusion Model for Conditional Reservoir Facies Generation. *Computers & Geosciences*, 194: 105750. doi: 10.1016/j.cageo.2024.105750
- [25] Liu, H., Xia, S., Fan, C., et al., 2024. Integrating Multimodal Deep Learning with Multipoint Statistics for 3D Crustal Modeling: A Case Study of the South China Sea. *Journal of Marine Science and Engineering*, 12(11). doi: 10.3390/jmse12111907
- [26] Mahjour, S.K., Santos, A.A.S., Correia, M.G., et al., 2021. Scenario Reduction Methodologies under Uncertainties for Reservoir Development Purposes: Distance-Based Clustering and Metaheuristic Algorithm. *Journal of Petroleum Exploration and Production Technology*, 11(7): 3079–3102. doi: 10.1007/s13202-021-01210-5
- [27] Matheron, G., 1963. Principles of Geostatistics. *Economic Geology*, 58(8): 1246–1266. doi: 10.2113/gsecongeo.58.8.1246
- [28] Merzoug, A., Pyrcz, M., 2025. Conditional Generative Adversarial Networks for Multivariate Gaussian Subsurface Modeling: How Good Are They? *Mathematical Geosciences*, 57(4): 733–757. doi: 10.1007/s11004-025-10176-7
- [29] Parmar, N., Vaswani, A., Uszkoreit, J., et al., 2018. Image Transformer. In: Proceedings of the 35th International Conference on Machine Learning. PMLR, 4055–4064.
- [30] Peters, E., Arts, R.J., Brouwer, G.K., et al., 2010. Results of the Brugge Benchmark Study for Flooding Optimization and History Matching. *SPE Reservoir Evaluation & Engineering*, 13(03): 391–405. doi: 10.2118/119094-PA
- [31] Santos, S.M.G., Gaspar, A.T.F.S., Schiozer, D.J., 2018. Managing Reservoir Uncertainty in Petroleum Field Development: Defining a Flexible Production Strategy from a Set of Rigid Candidate Strategies. *Journal of Petroleum Science and Engineering*, 171: 516–528. doi: 10.1016/j.petrol.2018.07.048
- [32] Schiozer, D.J., De Souza Dos Santos, A.A., De Graça Santos, S.M., et al., 2019. Model-Based Decision Analysis Applied to Petroleum Field Development and Management. *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, 74: 46. doi: 10.2516/ogst/2019019
- [33] Song, S., Mukerji, T., Hou, J., 2021. GANSim: Conditional Facies Simulation Using an Improved Progressive Growing of Generative Adversarial Networks (GANs). *Mathematical Geosciences*, 53(7): 1413–1444. doi: 10.1007/s11004-021-09934-0
- [34] Straubhaar, J., Renard, P., Mariethoz, G., et al., 2011. An Improved Parallel Multiple-Point Algorithm Using a List Approach. *Mathematical Geosciences*, 43(3): 305–328. doi: 10.1007/s11004-011-9328-7
- [35] Strebelle, S., 2002. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology*, 34(1): 1–21. doi: 10.1023/A:1014009426274
- [36] Strebelle, S., Cavelius, C., 2014. Solving Speed and Memory Issues in Multiple-Point Statistics Simulation Program SNESIM. *Mathematical Geosciences*, 46(2): 171–186. doi: 10.1007/s11004-013-9489-7

- [37] Tang,J.,Tang,M.,Lu,S.,et al.,2024. Three-Dimensional Modeling of Estuary Reservoir Based on Coupling Sedimentary Dynamics Simulation and Multipoint Geostatistics Method. *Earth Science*, 49(1): 174-188. (in Chinese with English abstract).
- [38] Yuan,Z.,Ren,P.,Liu,J.,et al.,2025. Research on Whole Strata Geological Modeling Technology for CO2 Geological Storage in Salt Water Layer. *Earth Science*, 50(5): 1987-1998. (in Chinese with English abstract).
- [39] Zuo,C.,Pan,Z.,Yin,Z.,et al.,2022.A Nearest Neighbor Multiple-Point Statistics Method for Fast Geological Modeling.*Computers & Geosciences*,167: 105208. doi: 10.1016/j.cageo.2022.105208

附中文参考文献

- [1] 陈麒玉, 张如甜, 崔哲思, 等, 2025a. 基于 Lap-SAGAN 的多尺度三维地质模型重构方法. *地学前缘* 32, 1 - 30.
- [2] 陈麒玉, 荀磊, 崔哲思, 等, 2025b. 三维地质建模技术的最新进展和发展趋势. *地质科技通报* 44, 373 - 387.
- [3] 唐佳凡, 唐明明, 卢双舫, 等, 2024. 基于耦合沉积动力学模拟与多点地质统计学方法的河口湾储层三维建模. *地球科学* 49, 174 - 188.
- [4] 袁哲, 任培罡, 刘金华, 等, 2025. 咸水层CO2地质封存全地层地质建模技术研究. *地球科学* 50, 1987 - 1998.