MUNDO：用于时间降尺度的多分辨率生成式框架

Aymeric Delefosse[1*]、**Manon Tavernier**[1*]、**Julie Keisler**[1]，
Anastase Charantonis[1] **和 Dominique Béréziat**[2]Inria，
ARCHES，法国巴黎[2]索邦大学，法国国家科学研究中心，LIP6，法国巴黎

arXiv:2611.30868v1 [cs.LG] 23 Nov 2026

摘要

动力系统通常以较粗的时间分辨率进行观测，观测形式为稀疏状态或时间聚合量，因此底层的精细尺度轨迹无法解析。要恢复这些轨迹，需要概率模型既能捕捉未解析的变异性，又能与观测保持一致。我们提出MUNDO（**M**多分辨率**UN**统一**DO**降尺度），这是一个适用于一般线性观测算子的时间降尺度生成式框架。MUNDO 引入了一种零空间参数化，只需进行一次网络评估即可确保观测精确一致，从而统一时间插值与解聚。为捕捉复杂的多尺度动力学，我们利用反向马尔可夫学习，将观测算子分解为由连续粗化操作构成的层级结构。每一层仅对相邻分辨率之间损失的自由度进行建模。我们在双尺度 Lorenz-96、四种二维物理系统以及欧洲多变量 ERA5 再分析数据上，针对时间插值和解聚任务评估 MUNDO。在不同数据集和任务中，MUNDO 的表现通常优于基于流的基线方法。单步 MUNDO 具有较强的预测能力，而递归 MUNDO 通常能提升时间真实性和集合校准效果，使离散度与预测能力更加吻合，并提高预测覆盖率。

1 引言

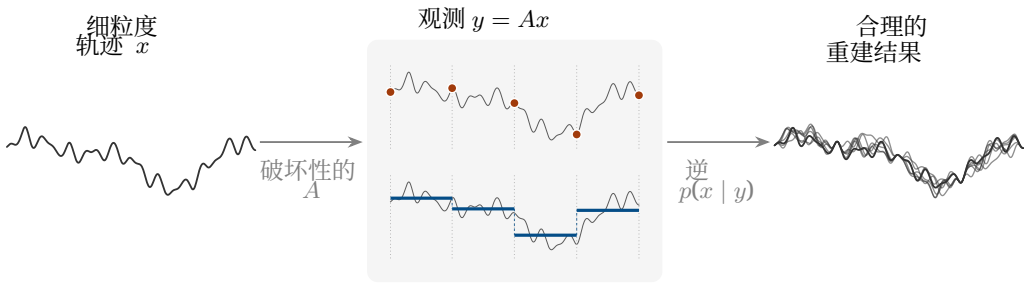


图1：**动力系统的粗尺度观测定义了一个一对多的逆问题**。子采样（橙色圆点）产生插值，而平均（蓝色条形）产生解聚。在这两种情况下，重建目标都是条件分布 $p(x \mid y)$。

动力系统很少会以其演化所处的时间分辨率进行观测。数值模型以精细的时间步长进行积分，但通常只存储状态的稀疏子集或聚合量，因为存储完整轨迹的成本高得难以承受。这在天气和气候领域尤为常见：气候模拟通常只保留亚小时尺度动力学的日均值或月平均值（Eyring 等人，2016年），而数据驱动天气预报模型通常以六小时间隔预测状态（Lam 等人，2023年；Kurth 等人，2023年）。其他领域也存在同样的细尺度时间信息损失，例如

*同等贡献。

模拟和观测到的物理系统，包括流体和波动动力学。时间降尺度以很低的成本重建缺失的信息，无需重新运行或训练模型（Brenowitz 等人，2025年）。它通常不加区分地涵盖插值（观测到稀疏状态）和解聚（给定聚合量）（图1）。两者都是同一个逆问题的实例：从不完整的观测中恢复精细轨迹。

对于混沌多尺度动力学，这个逆问题是多对一的：端点无法确定其间的轨迹，聚合量也无法确定产生它的细尺度变异。因此，目标是得到一个轨迹分布，其样本与观测一致。学习型时间降尺度器大多是确定性的，并针对逐点误差进行优化（Harilal 等人，2023年；Wang 等人，2025年；Leinonen 等人，2024年），因此只会给出一种重建结果，即条件均值，而无法提供与观测相容的各种可能结果。概率方法能够表示这些可能结果，但只能通过条件输入或惩罚项间接强制满足观测（Ingstad 等人，2026；Bassetti 等人，2024年），因此生成的样本未必满足观测。后验采样器接受任意线性算子，并通过引导或反复进行数据一致性更新来强制满足观测（Kim 等人，2025年；Aich 等人，2026）。在这些方法中，没有一种能在不进行迭代求解的情况下，针对两种算子都生成严格满足观测的样本。

本文介绍了 MUNDO（**M**多分辨率**UN**统一**DO**降尺度），这是一种在一般线性观测算子下进行时间降尺度的生成式框架。MUNDO 基于 engression（Shen 和 Meinshausen，2025年），该方法使用能量评分（Gneiting 和 Raftery，2007）这一严格适当评分规则来学习条件采样器。MUNDO 通过一次网络评估生成每个样本，将其表示为经典重建与观测算子零空间中的学习型残差之和。因此，对于插值和重分解，每个样本都严格满足观测约束。这样的采样器必须一次性恢复所有未解析尺度。然而，时间降尺度会自然地跨分辨率分解：较大的时间粗化操作可以表示为多个较小粗化操作的组合。我们通过反向马尔可夫学习（Shen 等人，2025年）利用这一结构，将该分解转化为条件采样器层级。每一层仅生成相邻分辨率之间丢失的自由度。零空间参数化确保每一层与其父节点保持严格一致，进而与观测保持一致。

我们的贡献如下：

- 一种适用于一般线性观测算子的零空间参数化方法，每次抽样只需进行一次网络评估，即可实现精确的观测一致性，适用于插值和细化分解；
- 对观测算子进行多分辨率因式分解，使每个层级仅生成相邻分辨率之间丢失的自由度，并与其父节点严格保持一致；
- 在层级结构自身的样本上进行端到端微调，从而消除训练时看到的真实父级与采样时接收的生成父级之间的不匹配；
- 一种与能量评分互补的空间相干项。

我们在混沌系统、模拟物理动力学和多变量 ERA5 再分析数据上对 MUNDO 进行了广泛评估，衡量其概率技巧、校准、观测一致性和时间结构。整个过程中，我们将递归 MUNDO 与其单步对应方法进行比较，以剖析层级结构带来的作用。

本文结构安排如下：第2 节回顾时间降尺度的相关工作。第3 节阐述逆问题和目标条件分布，然后介绍所提出的方法。第4节介绍系统、基线方法和评估方案，并报告结果。

2 相关工作

附录A 对下文介绍的方法作了更全面的回顾，并提供了对比表。

统计时间降尺度。Denton–Cholette（Denton，1971年；Cholette，1984年）重建与每个聚合量一致且最平滑的时间序列记录。CDF-t（Déqué，2007；Michelangeli 等人，2009）调整不同时间尺度上的边缘分布。两者都是时间分解的经典基线方法，但都无法对精细尺度轨迹的条件分布进行建模。

学习式时间降尺度。STint（Harilal 等人，2023年）、TemDeep（Wang 等人，2025年）和 ModAFNO（Leinonen 等人，2024年）将视频帧插值（Jiang 等人，2018年）方法应用于地球物理场。它们的逐点损失以条件均值为目标，因此会产生第3.3节所述的方差缺失。生成模型通过对条件分布进行采样来避免这种缺失，并且在如何落实观测约束方面各不相同。扩散模型以粗分辨率记录为条件输入，例如根据月平均值重建日尺度场（Bassetti 等人，2024年）。然而，这些样本无法保证重现条件输入中的聚合量，而且生成过程需要迭代去噪链。HourGlass（Ingstad 等人，2026）只需一次前向传播即可生成样本，但通过损失函数鼓励端点和窗口统计量的一致性，且仅处理六小时到小时插值。RainDisaggGAN（Scher & Peßenteiner，2021年）通过以日总量缩放时间 softmax，可在单次传播中实现精确结果，但仅适用于一个非负变量的解聚。

逆问题求解器。另一类互补研究在采样过程中将观测值施加于无条件生成式先验。DDNM（Wang 等人，2023年）在反向扩散期间保留由测量确定的分量，同时细化零空间内容；对于线性算子而言，这一过程是精确的。基于引导的求解器则向采样动力学中加入似然梯度更新，但这只能是近似的。FlowDPS（Kim 等人，2025年）将这一方法用于流匹配，Schmidt 等人（2025年）则利用它从六小时气候场生成小时集合；WIND（Aich 等人，2026年）使用一个在 ERA5 上预训练的扩散先验。

3 问题与方法

3.1 设置与观测算子

令 $X_i \in \mathbb{R}^d$ 表示系统的瞬时状态，令 δ 表示目标时间分辨率。一个窗口包含 N 个长度为 δ 的步长，其细尺度记录 $x = (x_1, \ldots, x_N) \in \mathcal{X} \subseteq \mathbb{R}^{Nd}$ 是从 μ 中抽取的样本；μ 是分辨率为 δ 的 N 个连续状态的联合分布，其密度未知，但我们持有从中获得的样本。我们用 x 同时表示随机记录及其某一个实现，具体含义由上下文确定。

细尺度记录从不直接被观测到，而只能通过一个会使其退化的观测算子

$$y = f(x), \qquad f : \mathbb{R}^{Nd} \to \mathbb{R}^m, \quad m < Nd, \tag{1}$$

其中 y 是粗分辨率记录。我们令 f 为线性算子，并将其与矩阵 $A \in \mathbb{R}^{m \times Nd}$, $f(x) = Ax$等同；不失一般性，假设其满行秩为 m。于是，f满射到 $\mathbb{R}^m$，且其右逆为 $A^\dagger = A^\top (AA^\top)^{-1}$，但它不是单射：其核的维数为 $\dim\ker A = Nd - m > 0$。给定 y，相容集是其原像

$$\mathcal{X}_y = f^{-1}(\{y\}) = \{ x \in \mathbb{R}^{Nd} : Ax = y \} = A^\dagger y + \ker A, \tag{2}$$

一个由 $\ker A$张成的仿射子空间。特别地，真实记录可分解为

$$x = A^\dagger y + r, \qquad r = (I - A^\dagger A)\, x \in \ker A, \tag{3}$$

一个由观测确定的分量，以及一个与之正交的残差，而观测对此残差完全无法确定：核恰好就是 f 所抹去的信息。公式（3）即为值域-零空间分解，$A^\dagger y$ 承载值域空间内容，而 r 承载零空间内容。它是逆问题中零空间学习的基础（Schwab 等人，2019年），并用于在基于扩散的图像复原中强制实现数据一致性（Wang 等人，2023年）。

在本研究中，f对记录进行时间粗化：仅以更粗的粗步长 $\Delta = k\delta$观测 x，其中 $k \in \mathbb{N}^*$可整除 N，因此保留 $n = N/k$ 个粗步和 $m = nd$，通过两个算子之一实现

$$(A_{\text{sub}}\, x)_j = x_{jk}, \qquad (A_{\text{avg}}\, x)_j = \frac{1}{k} \sum_{i \in B_j} x_i, \qquad j = 1, \ldots, n, \tag{4}$$

其中 $B_j = \{(j-1)k + 1, \ldots, jk\}$ 是 j 个精细步组成的第 k 个区块。从 A_{sub} 进行重建就是插值，而从 A_{avg} 进行重建则是解聚。每个粗粒度值都取自其自身区块，因此$AA^\top = c\, I_{nd}$，其中 $c = 1$ 用于子采样，$c = 1/k$ 用于分块平均，而 $A^\dagger = c^{-1} A^\top$。这些核函数，

$$\begin{aligned} \ker A_{\text{sub}} &= \{\, u : u_{jk} = 0,\ j = 1, \ldots, n \,\}, \\ \ker A_{\text{avg}} &= \left\{\, u : \textstyle\sum_{i \in B_j} u_i = 0,\ j = 1, \ldots, n \,\right\}, \end{aligned} \tag{5}$$

两者的维度均为 $(N-n)d$，代表算子所舍弃的子-Δ 变异性：前一种情况下，是一个情形中保留时刻之间的数值；后一种情况下，则是围绕区块均值的波动。

3.2 一个病态的逆问题

从 x 根据 y 进行恢复，在 Hadamard 意义下是病态的Hadamard (1902)。由满射性可知，对每个 $y \in \mathbb{R}^{nd}$都存在一个解，但解并不唯一，因为 $\mathcal{X}_y$ 是一个维数为 $(N-n)d$的仿射子空间，而非单点（Engl 等人，1996）。因此，确定性规则 $g:\ \mathbb{R}^{nd} \to \mathbb{R}^{Nd}$ 必须从 $\mathcal{X}_y$中选出一个代表。其核分量是 y的确定性函数，因此，给定 y 时，其输出的条件分布是狄拉克质量，无法表示与同一观测相容的各条轨迹之间的差异。

3.3 条件均值的方差缺失

最小化期望平方误差得到条件均值 $\mathbb{E}[x \mid y]$。由于 f 是线性的，条件均值仍与粗尺度观测完全一致，如下所示：

$$f(\mathbb{E}[x \mid y]) = \mathbb{E}[f(x) \mid y] = y. \tag{6}$$

然而，其方差存在系统性缺失。令 $\varphi:\ \mathbb{R}^{Nd} \to \mathbb{R}$为细尺度记录的线性泛函。全方差定律给出

$$\mathrm{Var}[\varphi(\mathbb{E}[x \mid y])] = \mathrm{Var}[\varphi(x)] - \mathbb{E}[\mathrm{Var}[\varphi(x) \mid y]], \tag{7}$$

其中，第二项是粗尺度观测未能解析的 $\varphi(x)$ 变异性。因此，条件均值恰好低估了 $\varphi(x)$ 的方差，低估量正是该项。这一缺失源于二次准则，而非任何特定的模型类别。当 $\varphi(x)$ 由 y 确定时，这一差额便消失；当 $\mathbb{E}[\varphi(x) \mid y]$ 为常数时，差额则达到 $\mathrm{Var}[\varphi(x)]$。

3.4 概率表述

确定性重建的局限性促使我们采用不同的目标。我们不再从 $\mathcal{X}_y$ 中选取单个成员，而是寻求给定粗尺度观测时精细记录的条件分布，该分布在 μ 下定义为 [1]

$$p(x \mid y) \;=\; \mu\left(x \,\middle|\, f(x) = y\right). \tag{8}$$

重建也就意味着采样 $p(\cdot \mid y)$：在固定 y 的条件下，生成多个相容记录，其离散度反映观测未能解析的变异性。

3.5 多分辨率采样

对式（8）中的条件分布进行采样，相当于生成式（3）中的残差。对于任意右逆 $A^\dagger$（对应 A），每个与 y 相容的记录都可以表示为基线重建加上 $\ker A$中的一个元素。因此，MUNDO 会抽取

$$\hat{x} = \underbrace{A^\dagger y}_{\text{classical reconstruction}} + \underbrace{\left(I - A^\dagger A\right) G_\theta(y, \varepsilon, c)}_{\text{residual} \in \ker A}, \tag{9}$$

其中 G_θ 是一个神经网络，c表示额外的条件输入；遵循 engression（Shen 和 Meinshausen，2025年）的做法，随机性通过 ε引入，每次抽样只需进行一次网络评估。由于 $AA^\dagger = I$，仿射映射 $P_y(u) = A^\dagger y + (I - A^\dagger A)u$ 是到 $\mathcal{X}_y$上的投影。因此，按构造，$\hat{x} = P_y(G_\theta(y, \varepsilon, c))$ 满足 $A\hat{x} = y$，无需惩罚项。残差头初始化为零，因此未训练的层级会返回基线重建 $A^\dagger y$。MUNDO 不规定 G_θ的架构；第4节说明了各个系统所采用的架构。

[1] 由于条件输入事件的概率为零，$\mu(\cdot \mid f(x) = y)$ 被理解为沿 f 对 μ 的解体，其对 $f_\#\mu$-几乎处处的 y 定义，并满足 $\mu = \int p(\cdot \mid y)\, \mathrm{d}(f_\#\mu)(y)$，其中 $p(\cdot \mid y)$ 的支撑集包含于 $\mathcal{X}_y$。

多分辨率因子分解。 遵循反向马尔可夫学习（Shen 等人，2025年），我们将单个条件采样器扩展到多个时间分辨率构成的层级结构。两个观测算子都可分解为一系列满行秩线性粗化，

$$A = A_R \circ \cdots \circ A_1, \qquad x^{(r)} = A_r\, x^{(r-1)}, \qquad x^{(0)} = x, \quad x^{(R)} = y. \tag{10}$$

每个 $x^{(r)}$ 因此都是中间分辨率下的一条记录，而每个 A_r 都定义了一个与方程（1）形式相同的逆问题，并具有各自的相容集和核。由于每个层级都能确定性地确定所有更粗尺度的层级，因此以 $x^{(r+1)}$ 为条件也会固定 $x^{(r+2)}$，……，$x^{(R)}$。因此，条件分布沿层级结构因式分解，

$$p\,(x \mid y) = \int \prod_{r=0}^{R-1} p\left(x^{(r)} \mid x^{(r+1)}\right) \mathrm{d}x^{(1)} \cdots \mathrm{d}x^{(R-1)}. \tag{11}$$

这种因式分解是精确的：一个跨越粗化比率 k 的单一条件问题，被替换为一系列形式相同、规模更小的条件问题。MUNDO 从粗到细对这些条件分布进行采样，在生成 $x^{(r-1)}$ 时，使用 A_r 通过方程（9） $x^{(r)}$。因此，每个生成的层级都能精确复现其父节点，并使一致性传递至原始粗尺度观测。近似仅来自学习型条件因子，而非因式分解本身。我们将 $R = 1$ 情形称为单步 MUNDO，将具有$R > 1$的情形称为递归 MUNDO。在递归 MUNDO 中，每个层级还可以在自身分辨率下，以前一窗口的最后一个状态作为条件（附录B.5）。

训练目标函数。 学习条件采样器需要一个总体最优解为目标分布的目标函数。我们使用能量评分，这是 engression 中采用的严格适当评分规则（Gneiting 和 Raftery，2007年；Shen 和 Meinshausen，2025年）。对于由 M 个相互独立的抽样 $u^{(1)}$，…，$u^{(M)}$ 以及目标 v组成的集合，其无偏集合估计量为

$$S_M\left(u^{(1:M)}, v\right) = \frac{1}{M}\sum_{m=1}^{M}\left\|u^{(m)} - v\right\| - \frac{1}{2M(M-1)}\sum_{\substack{m,m'=1 \\ m\neq m'}}^{M}\left\|u^{(m)} - u^{(m')}\right\|. \tag{12}$$

递归 MUNDO 在投影样本上计算能量评分，而单步 MUNDO 则在未投影样本上进行训练，并在训练后应用投影（附录 B.3）。每个层级都以其真实父节点为训练依据，因此学习的是（11）中的条件因子，而不是由较粗层级的学习误差传播而来的误差。

尽管能量评分是严格适当的，但在高维情况下，对于依赖关系设定错误的辨别能力可能较弱（Scheuerer & Hamill，2015年）。在我们的一些空间系统中，仅使用能量评分进行训练会高估细尺度变异（附录B.2）。当空间相干性未得到充分捕捉时，我们会在能量评分的基础上加入辅助惩罚项，

$$\mathcal{L} = S_M + \lambda \mathcal{R}_{\text{spatial}}. \tag{13}$$

一种选择是采用变差函数惩罚项（Scheuerer & Hamill，2015）。令 $\gamma_h^a(x) = |x_{a+h} - x_a|^p$ 表示单元格 a处的 p阶成对差异，以及空间滞后 h，其中 $p = \frac{1}{2}$。无偏 M次抽样估计量为

$$V_M(u_{1:M}, v) = \sum_h w_h \sum_a \frac{1}{M(M-1)}\sum_{\substack{m,m'=1 \\ m\neq m'}}^{M}\left(\gamma_h^a(v) - \gamma_h^a(u_m)\right)\left(\gamma_h^a(v) - \gamma_h^a(u_{m'})\right). \tag{14}$$

我们沿每个空间轴使用间隔为一个、两个和四个网格单元的滞后。另一种选择是采用谱 CRPS 惩罚项，这一做法借鉴了其在数据驱动天气预报中的应用（Bonev 等人，2025年）。该方法将集合成员和目标转换到谱域后评估概率技巧，从而惩罚空间尺度分布中的误差。我们将变异函数项和谱项都视为辅助空间惩罚项，而非 MUNDO 定义的一部分。使用空间惩罚项时，λ 最初设定为使能量项和空间项的量级相近。

端到端微调。教师强制使用真实父节点训练每个条件因子，而推理时接收的却是生成的父节点，因此误差可能沿反向链累积（Huang 等人，2026年）。因此，我们可以选择通过从粗尺度观测生成完整轨迹，并将最终损失反向传播至所有层级，从而对整个层级结构进行端到端微调。这样，各层级都能接触到推理时遇到的父级分布，并使较细层级能够补偿上游误差。在此阶段，我们保留能量评分目标，并以谱 CRPS 项替代变差函数惩罚项（附录 B.2）。只有当验证诊断显示递归 MUNDO 中存在误差传播或校准退化时，才会进行微调。

4 实验

这些实验探讨了三个问题：(i) 生成的轨迹是否满足粗尺度观测；(ii) 在什么情况下，多分辨率采样比一次性生成完整窗口效果更好；(iii) 随着可预测性降低，集合能否恢复未解析的条件变异性和精细尺度时间结构？

4.1 实验设置

该方案适用于表 1 中的三个层级。从精细分辨率记录出发，我们应用（4）中的一个观测算子，并根据由此得到的粗分辨率记录重建留出轨迹。插值窗口包含 $N+1$ 个状态，其两个端点 x_0 和 x_N 均为观测值。对于两个模拟层级，完整解析的轨迹均可用，且可系统地改变可预测性，因此我们能够研究性能与窗口长度 N 以及动力学内在去相关时间之间的关系。ERA5 提供了相应的基于档案数据的设置。附录C 给出了方程、数据划分和训练细节。

表 1：**实验系统**。 我们在三个状态维度和真实度逐渐提高的层级上评估 MUNDO。每个包含 N 步的窗口都粗化为一个值；因式分解列出了从粗到细依次应用的粗化操作，每个因子对应一个层级（2^6：六个因子均为二）。

Tier	系统	状态形状	窗口 N 和因式分解	控制 可预测性	骨干网络
Lorenz-96		36×1	64: 2^6; 128: 2^7	强迫项 $F \in \{10, 24, 32, 64\}$	循环卷积神经网络
AutoSim	平流-扩散 Gray–Scott 纳维–斯托克斯 Gross–Pitaevskii	$64^2 \times 1$ $64^2 \times 2$ $64^2 \times 3$ $64^2 \times 3$	8: 2^3; 16: 2^4; 24: 2, 2, 2, 3; 32: 2^5	机制和 N （通过 $N\delta/\tau$）	NATTEN （Hassani 等人，2023年）
ERA5		$149 \times 281 \times 7$	24: 2, 2, 2, 3	–	ArchesWeather （Couairon等人，2026年）

方法。我们将单步和递归 MUNDO 与经典基线方法及学习型基线进行比较。经典基线方法包括用于插值的线性方法和自然三次样条，以及用于解聚的 Denton–Cholette 和 CDF-t。学习型基线包括条件流匹配（Lipman 等人，2023年）、FlowDPS（Kim 等人，2025年），以及在二维系统上作为确定性基线的 SimVP（Gao等人，2022年）。在 Lorenz-96 和 AutoSim 上，MUNDO、条件流匹配和 FlowDPS 的参数数量保持一致；而 ERA5 则采用架构特定配置（附录 C.4 ）。插值和解聚分别训练独立模型。在所有三个基准测试中，采样后均使用方程（ 9 ）中的映射，将学习型基线投影到约束集上。在解聚任务中，MUNDO 和两种基于流的基线方法将相邻窗口的聚合量作为跨窗口上下文（附录 B.5）。

评估。 我们分别评估重建结果是否满足观测，以及是否恢复了未解析的条件分布。如果重建结果在插值时保留了保留状态，并在解聚时保留了施加的聚合量，则称其具有一致性。在表格中，一致性表示某种方法是否能在 10^{-4} 标准化单位的精度范围内，通过构造实现一致性（✓）、通过采样后投影实现一致性（$\checkmark^{\pi}$），或完全无法实现一致性（×）。随后，我们评估学习得到的条件

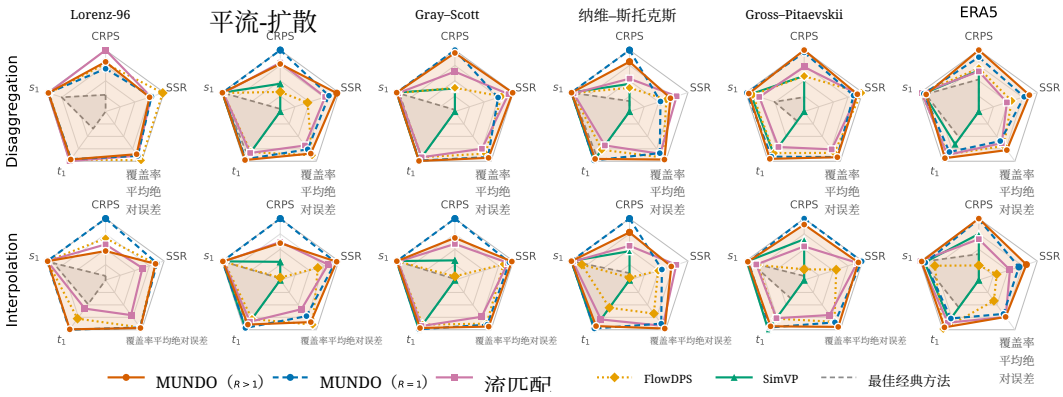


图2：**所有基准测试中的主要指标**，分别采用解聚（上）和插值（下）。每个坐标轴表示一个得分，范围为 $(0, 1]$，，其中 1 为理想值：多边形越大越好。连续排序概率得分相对于各面板中的最佳方法计算；其他坐标轴为绝对值。灰色虚线表示各面板中按连续排序概率得分评定的最佳经典基线。得分列于表 2、3 和 4。

分布，采用公平连续秩概率得分（Ferro，2014）、离散度-技巧比（Rasp 等，2024年，以及名义中心区间上经验覆盖率的平均绝对误差进行评估（Greenbury 等，2026年）。由于这些评分在时间上是边际的，我们还报告时间变差函数比 t_1，以衡量短时滞时间结构。附录 C.5 定义了所有指标。在 Lorenz-96 和 AutoSim 上，评分取三个训练随机种子的平均值，其中 $\pm$ 表示随机种子间的标准差。ERA5 使用单个随机种子。附录 C详细介绍了评估协议。

4.2 结果

图 2汇总了所有基准测试中的主要指标。在大多数面板中，MUNDO 覆盖的多边形面积最大，而最佳经典基线明显落后于所有学习方法。单步 MUNDO 在大多数面板中的连续排序概率得分最低，尤其是在插值情况下；递归则几乎在所有情况下都能改善校准。

Lorenz-96。 我们使用双尺度 Lorenz-96 系统（Lorenz，2006年）作为可控设置，通过强迫项直接调节可预测性 F。表 2 报告了 $N = 64$，$F = 32$下的代表性案例，而附录 D.1 报告了其余强迫项和窗口长度设置。在最终投影之前，尽管 FlowDPS 使用了一致性更新，但在 Lorenz-96 上，两个基于流的基线都不满足观测约束，且投影后它们的连续排序概率得分几乎没有变化。在解聚情况下，流匹配的连续排序概率得分低于递归 MUNDO，校准效果相近。在插值情况下，其连续排序概率得分介于两种 MUNDO 变体之间；其成员表现出过度的逐步噪声，短时滞变率几乎是真值的两倍，而递归 MUNDO 保持 t_1 接近 1。在整个参数扫描中，单步 MUNDO 在所有混沌设置下的插值连续排序概率得分都更低；而在解聚情况下，除一种设置外，递归 MUNDO 的得分在所有设置中都更低。随着可预测性降低，连续排序概率得分差距缩小，递归方法的优势转向校准和短时滞结构，而这两项指标在单步 MUNDO 中会退化。图 3 展示了连续四个窗口中一个慢变量的集合。在解聚情况下，与其他方法相比，递归 MUNDO 的成员能更好地保持窗口边界处的连续性。在插值情况下，其集合范围比单步 MUNDO 更宽，包含更多不同轨迹的真值；而流匹配的成员在相邻步之间明显带有噪声。

AutoSim。 我们使用 AutoSim 中的四个系统（Greenbury 等人，2026）：将平流-扩散作为近确定性对照，将 Gray–Scott 用于多模态图案形成，将条件纳维–斯托克斯用于湍流动力学，并将 Gross–Pitaevskii 用于色散波动力学。表4 报告了 $N = 32$处的基准层级，而附录 D.2 报告了其余的窗口长度和可预测性状态。在表 4中，单步 MUNDO 在所有系统和任务上的连续排序概率得分（CRPS）均已优于每个基线，并且在插值任务中仍是最强的变体。两种基于流的方法在每个系统和任务上都高估了短滞后变异性，而 t_1 更接近 1，用于

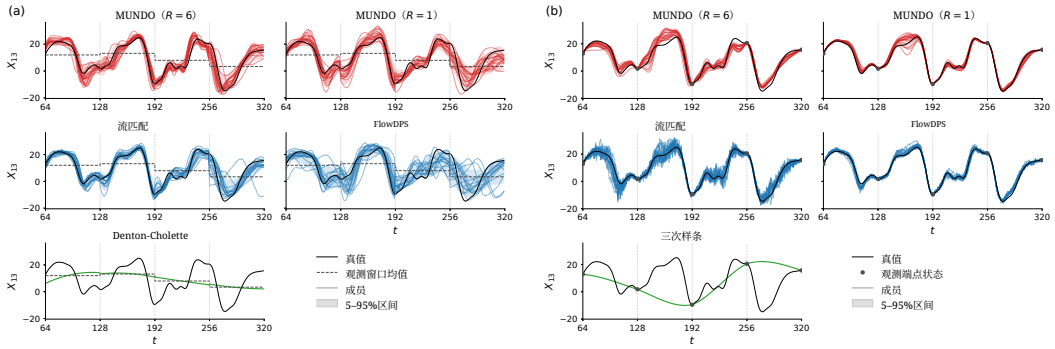


图 3： **Lorenz-96,** $F = 32$, $N = 64$**方差最大的慢变量**，涵盖图8和9中的四个窗口。（a）解聚，（b）插值。每个面板显示真值、观测、16 个成员（细线）以及完整集合的 5–95% 范围（带状区域）。经典基线方法显示其单次重构结果。

表 2：**Lorenz-96,** $N = 64$, $F = 32$。 †：因果链。粗体：最佳；下划线：各指标中的次优。

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 6$)†	✓	0.2255±0.0030	0.75	0.083		0.96
MUNDO ($R = 1$)	✓$^{\pi}$	0.2618±0.0015	0.79	0.063		0.96
流匹配	✓$^{\pi}$	**0.1823**±0.0046	0.78	0.065		**1.00**
FlowDPS	✓$^{\pi}$	0.2341±0.0019	**1.02**	**0.011**		0.96
Denton–Cholette	✓	0.6715	–	–		0.36
CDF-t	×	0.7850	–	–		0.48
插值						
MUNDO ($R = 6$)	✓	0.1112±0.0068	0.85	**0.021**		0.99
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0526**±0.0009	**0.88**	0.026		**1.00**
流匹配	✓$^{\pi}$	0.0906±0.0021	1.59	0.172		1.74
FlowDPS	✓$^{\pi}$	0.0774±0.0065	0.86	0.028		1.29
线性	✓	0.7816	–	–		0.49
三次样条	✓	0.8177	–	–		0.50

表3：**ERA5，日到小时**。粗体：最佳；下划线：各指标的次佳结果。

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 4$)	✓	**0.0961**	**0.79**	**0.147**		**1.01**
MUNDO ($R = 1$)	✓$^{\pi}$	0.1089	0.75	0.250		1.19
流匹配	✓$^{\pi}$	0.1484	0.48	0.224		0.84
FlowDPS	✓$^{\pi}$	0.1376	0.55	0.190		0.84
SimVP	✓$^{\pi}$	0.1438	–	–		0.64
Denton–Cholette	✓	0.1841	–	–		0.52
CDF-t	×	0.2165	–	–		0.78
插值						
MUNDO（$R = 4$）	✓	**0.0929**	**0.81**	0.151		1.04
MUNDO（$R = 1$）	✓$^{\pi}$	0.0937	0.68	0.186		1.27
流匹配	✓$^{\pi}$	0.1397	0.53	**0.150**		0.84
FlowDPS	✓$^{\pi}$	0.3918	0.31	0.342		**1.00**
SimVP	✓$^{\pi}$	0.1231	–	–		0.80
线性	✓	0.2204	–	–		0.55
三次样条	✓	0.2211	–	–		0.56

MUNDO。在更困难、可预测性更低的设置中，递归通常会改善集合校准；在可预测性最低的情况下，它还会降低解聚任务下的连续排序概率得分（CRPS），同时保留真实的细粒度时间结构。图4 和 5 分别展示了纳维–斯托克斯和 Gross–Pitaevskii 上解聚与插值任务的代表性重建结果。附录 D.2 还展示了平流-扩散和 Gray–Scott 上的重建结果（图10 和 11）。

ERA5。 我们在欧洲（Hersbach 等人，2020年）的 ERA5 再分析数据上，使用七个地表变量，在 149×281 网格上评估日到小时降尺度。降水为非负值，而线性约束并不保证这一点，因此我们将每种方法生成的降水样本投影到 $\{Ax = y,\ x \geq 0\}$（附录B.3）。表 3 显示，MUNDO 的两个变体在连续排序概率得分（CRPS）上都优于所有基线方法。在基线方法的参数数量下，单步 MUNDO 相较于表现更好的基于流的基线，在解聚任务中已将 CRPS 降低21%，在插值任务中降低33%；而基于流的集合则明显离散度不足。递归对解聚最为有用：它进一步将 CRPS 降低了 12%，改善了离散度 — 技巧一致性和覆盖率，并使 t_1 接近 1。在插值任务中，两个变体的 CRPS 几乎相同，但递归再次改善了校准和时间结构。附录D 报告了逐变量得分和其他诊断结果。图 6 展示了两种算子下具有代表性的 ERA5 总降水重建结果，其中不同成员从同一粗尺度观测中恢复出了合理的日内结构。

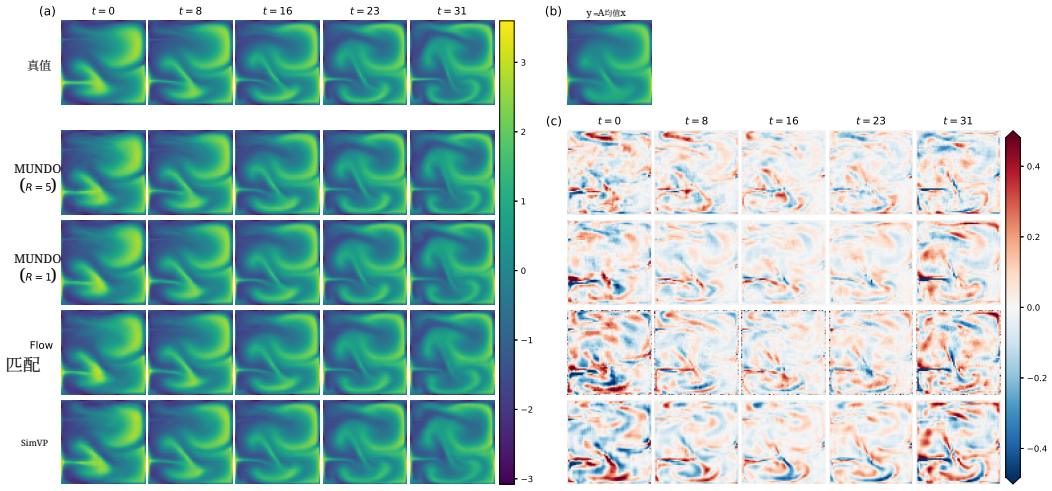


图 4：**纳维–斯托克斯方程解聚于** $N = 32$。(a) 所选时间点的真值和重建结果。(b) 粗尺度观测，即窗口内的时间平均。(c) 所选时间点的重建误差 $\hat{x} - x$；各行的方法顺序与 (a) 一致。

表 4：**AutoSim 于** $N = 32$。 ‡：端到端微调。粗体：最佳；下划线：各指标的次优。

(a) **平流-扩散**

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 5$)	✓	0.0026±0.0000	**0.97**	0.090	**1.02**	
MUNDO ($R = 1$)	✓$^\pi$	**0.0020**±0.0000	0.83	0.140	1.04	
流匹配	✓$^\pi$	0.0026±0.0000	0.77	0.181	1.21	
FlowDPS	✓$^\pi$	0.0062±0.0002	0.47	**0.070**	1.22	
SimVP	✓$^\pi$	0.0044	–	–	1.03	
Denton–Cholette	✓	0.0382	–	–	0.89	
CDF-t	×	0.0432	–	–	0.88	
插值						
MUNDO ($R = 5$)	✓	0.0022±0.0000	**1.04**	0.090	1.12	
MUNDO ($R = 1$)	✓$^\pi$	**0.0013**±0.0000	0.94	0.160	1.04	
流匹配	✓$^\pi$	0.0021±0.0000	0.82	0.242	1.19	
FlowDPS	✓$^\pi$	0.0326±0.0024	0.64	**0.060**	1.28	
SimVP	✓$^\pi$	0.0044	–	–	**1.03**	
线性	✓	0.0452	–	–	0.89	
三次样条	✓	0.0420	–	–	0.92	

(b) **Gray–Scott**

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 5$)	✓	0.0034±0.0000	**1.01**	**0.040**	**1.01**	
MUNDO ($R = 1$)	✓$^\pi$	**0.0032**±0.0000	1.35	0.048	**1.01**	
流匹配	✓$^\pi$	0.0049±0.0001	0.93	0.144	1.09	
FlowDPS	✓$^\pi$	0.0083±0.0002	0.80	0.089	1.09	
SimVP	✓$^\pi$	0.0085	–	–	**1.01**	
Denton–Cholette	✓	0.1079	–	–	0.93	
CDF-t	×	0.1440	–	–	0.90	
插值						
MUNDO ($R = 5$)	✓	0.0035±0.0000	**1.03**	**0.037**	1.04	
MUNDO ($R = 1$)	✓$^\pi$	**0.0024**±0.0000	0.93	0.060	1.01	
流匹配	✓$^\pi$	0.0041±0.0001	1.10	0.153	1.08	
FlowDPS	✓$^\pi$	0.0399±0.0008	1.20	0.072	1.12	
SimVP	✓$^\pi$	0.0076	–	–	**1.00**	
线性	✓	0.1563	–	–	0.91	
三次样条	✓	0.1361	–	–	0.95	

(c) **纳维–斯托克斯**

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 5$)	✓	0.0159±0.0002	0.70	**0.020**	1.05	
MUNDO ($R = 1$)	✓$^\pi$	**0.0129**±0.0001	0.53	0.095	**1.02**	
流匹配	✓$^\pi$	0.0236±0.0003	**1.24**	0.088	1.51	
FlowDPS	✓$^\pi$	0.0334±0.0007	0.63	0.044	1.30	
SimVP	✓$^\pi$	0.0286	–	–	**1.02**	
Denton–Cholette	✓	0.0892	–	–	0.89	
CDF-t	×	0.1665	–	–	0.92	
插值						
MUNDO ($R = 5$)	✓	0.0117±0.0000	0.71	**0.015**	1.08	
MUNDO ($R = 1$)	✓$^\pi$	**0.0091**±0.0000	0.55	0.072	1.03	
流匹配	✓$^\pi$	0.0163±0.0003	**1.25**	0.046	1.26	
FlowDPS	✓$^\pi$	0.2365±0.0050	0.49	0.191	1.80	
SimVP	✓$^\pi$	0.0193	–	–	**1.02**	
线性	✓	0.1041	–	–	0.92	
三次样条	✓	0.0933	–	–	0.94	

(d) **Gross–Pitaevskii**

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 5$)‡	✓	**0.1202**±0.0018	0.91	**0.047**	**1.05**	
MUNDO ($R = 1$)	✓$^\pi$	0.1249±0.0014	0.85	0.049	0.91	
流匹配	✓$^\pi$	0.1578±0.0009	0.90	0.138	1.39	
FlowDPS	✓$^\pi$	0.2085±0.0018	**1.02**	0.098	1.20	
SimVP	✓$^\pi$	0.2071	–	–	0.85	
Denton–Cholette	✓	0.5154	–	–	0.21	
CDF-t	×	0.6528	–	–	0.38	
插值						
MUNDO ($R = 5$)‡	✓	0.0450±0.0005	0.93	**0.034**	1.07	
MUNDO ($R = 1$)	✓$^\pi$	**0.0406**±0.0004	**0.98**	0.087	1.05	
流匹配	✓$^\pi$	0.0743±0.0012	0.91	0.173	1.30	
FlowDPS	✓$^\pi$	0.2332±0.0059	0.55	0.093	1.29	
SimVP	✓$^\pi$	0.0612	–	–	**1.00**	
线性	✓	0.6231	–	–	0.36	
三次样条	✓	0.6519	–	–	0.37	

5 结论

时间插值和解聚是同一个逆问题的两种实例：恢复与粗尺度观测相容的细粒度轨迹分布。MUNDO 是一个同时解决这两个问题的生成式框架。每个采样器都使用能量评分进行训练，并生成一个

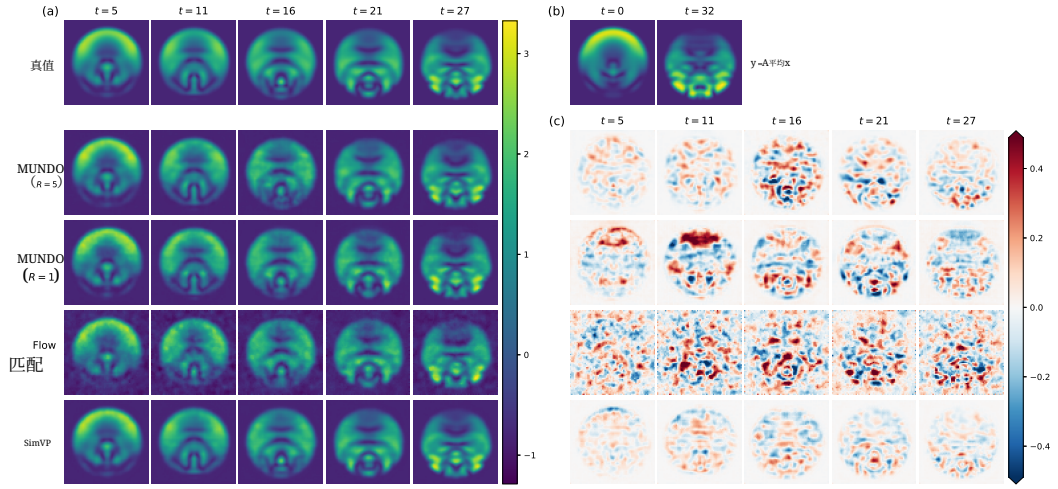


图5：**Gross–Pitaevskii 插值，时间为** $N = 32$。(a) 所选时刻的真实值和重建结果。(b) 观测端点状态。(c) 所选时刻的重建误差 $\hat{x} - x$；各行按 (a) 中的方法顺序排列。

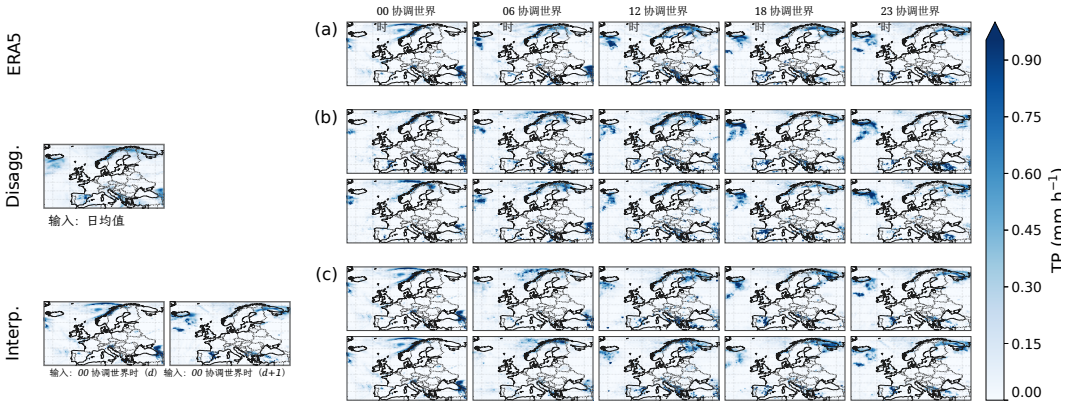


图6：**ERA5 总降水量（mm h^{-1}）的日到小时重建**。(a) 所选时刻的真实小时降水量。(b) 根据观测日均值进行解聚（最左侧），两行分别显示两个独立的递归 MUNDO 成员。(c) 在连续两天观测到的 00 协调世界时状态之间进行插值（最左侧），布局相同。

观测算子零空间中的残差，从而精确满足观测。单步 MUNDO 直接采样精细轨迹，而递归 MUNDO 将较大的由粗到细条件分解为一系列较小的条件，并按逆序采样。

我们在复杂度递增的三种场景中使用了相同的框架：Lorenz-96、四种模拟物理系统以及 ERA5 再分析，仅更换骨干网络。单步 MUNDO 已展现出较强的预测能力，并且在插值下通常能取得较低的连续排序概率得分。递归在保留精细尺度时间结构的同时提升了校准效果，并且在可预测性最低的场景下，通过解聚降低了连续排序概率得分。在包含七个耦合场的 ERA5 场景中，递归在两种算子下的所有指标上均优于单步 MUNDO。

将该构造扩展到联合时空观测算子以及更一般的观测算子（包括数据同化场景）是自然的下一步。第二个方向是将 MUNDO 应用于学习型预报和气候模型档案数据，包括没有配对高分辨率真值的场景，以生成气候预测中合理的细尺度时间实现。

人工智能使用声明

在本研究中，我们使用生成式人工智能工具协助软件实现、撰写、编辑和图表制作。我们未使用这些工具来提出假设、构建概念框架、设计实验或评估、生成或预处理数据，或解读结果。其余需要披露的事项不适用于本文。所有经人工智能辅助生成的代码均由作者审查，所有报告中的数值均来自我们的实验记录。我们对本研究的最终内容负责，包括借助生成式人工智能生成的文本、论断或成果。

可复现性声明

代码可在 https://github.com/dataymeric/mundo 获取。AutoSim 系统由以下工具生成：https://github.com/alan-turing-institute/autosim。ERA5 数据可从哥白尼气候数据存储库获取，网址为 https://doi.org/10.24381/cds.adbb2d47。附录 C 详细介绍了数据集、训练和评估。

致谢

作者感谢 Graham Clyne 富有见地的讨论。本研究得到法国国家研究署（ANR）和高等教育与研究部的支持。本项目获 GENCI 在 IDRIS 分配的高性能计算资源，分配编号为 AD011016443，使用的是 Jean-Zay 超级计算机。作者声明不存在任何利益冲突。

参考文献

Michael Aich、Andreas Fürst、Florian Sestak、Carlos Ruiz-Gonzalez、Niklas Boers 和 Johannes Brandstetter。WIND：用于零样本大气建模的天气逆扩散。在第*43*届国际机器学习会议论文集，第306卷，载于机器学习研究论文集，第1404–1433页。PMLR，2026年。网址：https://proceedings.mlr.press/v306/aich26a.html。

Pierre Ailliot、Denis Allard、Valérie Monbet 和 Philippe Naveau。随机天气生成器：天气类型模型综述。法国统计学会期刊，156(1)：101–113，2015年。

Seth Bassetti、Brian Hutchinson、Claudia Tebaldi 和 Ben Kravitz。DiffESM：利用三维扩散模型对地球系统模型中的温度和降水进行条件模拟。《地球系统建模进展》，16(10)：e2023MS004194，2024年。doi：10.1029/2023MS004194。

Yochai Blauand Tomer Michaeli。感知失真权衡。载于*2018*年 *IEEE/CVF* 会议 *o* *n* *Computer Vision and Pattern Recognition*（*CVPR*），第 6228–6237 页，2018年。doi: 10.1109/CVPR.2018.00652 .

Boris Bonev、Thorsten Kurth、Ankur Mahesh、Mauro Bisson、Jean Kossaifi、Karthik Kashinath、Anima Anand- kumar、William D. Collins、Michael S. Pritchard 和 Alexander Keller。FourCastNet 3：一种用于大规模概率机器学习天气预报的几何方法。*arXiv:2507.12144v2 [cs.LG]*, 2025。doi: 10.48550/arXiv.2507.12144。

Noah D. Brenowitz、Tao Ge、Akshay Subramaniam、Peter Manshausen、Aayush Gupta、David M. Hall、Morteza Mardani、Arash Vahdat、Karthik Kashinath 和 Michael S. Pritchard。瓶中气候：迈向公里尺度全球大气的生成式基础模型。*arXiv:* *2505 06474 2 [ ], .vphysics.ao-ph*。doi: . /arXiv. . 。
2025 10 48550 2505 06474

T. A. Buishand 和 T. Brandsma。通过最近邻重采样，对莱茵河流域的日降水和温度进行多站点模拟。《水资源研究》，37(11):2761–2776，2001年。doi: 10.1029/ 2001WR000291。

Pierre A.Cholette。将年内序列调整为年度基准值。《调查方法学》，10 49，1984年 1 35( ): – 。

Gregory C. Chow 和 An-loh Lin。通过相关序列对时间序列进行最佳线性无偏插值、分布和外推。《经济学与统计学评论》，53(4):372–375，1971年。doi: 10.2307/ 1928739。

Guillaume Couairon、Renu Singh、Anastase Charantonis、Christian Lessig 和 Claire Monteleoni
Arch- esWeatherGen：利用机器学习实现精准且计算高效的概率天气预报。科学进展，12(17):eadx2372，2026年。doi: 10.1126/sciadv.adx2372。

Estela Bee Dagum 和 Pierre A. Cholette。时间序列的基准评测、时间分布与协调方法，统计学讲义第186卷。Springer，2006年。doi: 10.1007/0-387-35439-5。Frank T. Denton。基于二次最小化调整月度或季度序列以匹配年度总量的方法。美国统计学会会刊，66(333):99–102，1971年。doi: 10.1080/01621459.1971.10482227。Michel Déqué。在人为影响情景下法国降水和极端温度的发生频率：模型结果及根据观测值进行的统计校正。*Global and Planetary Change*，57 (1-2):16–26，2007年。doi: 10.1016/j.gloplacha.2006.11.030。Heinz W. Engl、Martin Hanke 和 Andreas Neubauer。逆问题正则化，Mathematics and Its Applications 第375卷。Kluwer Academic Publishers，1996年。doi: 10.1007/978-94-009-1740-8。Veronika Eyring、Sandrine Bony、Gerald A. Meehl、Catherine A. Senior、Bjorn Stevens、Ronald J. Stouf- fer 和 Karl E. Taylor。第六阶段耦合模式比较计划（CMIP6）实验设计与组织概述。*Geoscientific Model Development*，9(5):1937–1958，2016年。doi: 10.5194/gmd-9-1937-2016。C. A. T. Ferro。集合预报的公平评分。*Quarterly Journal of the Royal Meteorological Society*，140 (683):1917–1923，2014年。doi: 10.1002/qj.2270。

Kristian Fösterling、Jonas Klesse、Fabian Wolter、Max Kanetzky 和 Ulrich Stemmer。科学开源气象观测时间序列解析工具（MELODIST v0.1.1）。地球科学模型开发，9(7):2315–2333，2016年。doi: 10.5194/gmd-9-2315-2016。

Zhangyang Gao、Cheng Tan、Lirong Wu 和 Stan Z. Li。SimVP：更简单却更出色的视频预测。发表于*2022*年 *IEEE/CVF*计算机视觉与模式识别会议（*CVPR*），第 3160–3170 页，2022年。doi: 10.1109/CVPR52688.2022.00317。Tilmann Gneiting 和 Adrian E. Raftery。严格适当评分规则、预测与估计。美国统计学会会刊，102(477):359–378，2007年。doi: 10.1198/016214506000001437。Sam F. Greenbury、Radka Jersakova、Paolo Conti、Marjan Famili、Christopher Iliffe Sprague、Edwin Brown 和 Jason D. McEwen。物理系统概率仿真的可靠性。*arXiv*:

2606 12997 1 .

[ ], 2026 10 48550 2606 12997

vcs.LG。doi: . /arXiv. . 。Jacques Hadamard。关于偏微分问题及其物理意义。普林斯顿大学公报，13:49–52，1902年。

Xue Han、Chunping Chang、Zhongling Guo、Shuiqing Yin、Jifeng Li、Huiru Li、Rende Wang 和 Qing Li。用于风蚀建模的确定性风速生成模型比较。《农业与森林气象学》，334:109438，2023年。doi: 10.1016/j.agrformet.2023.109438。

Paula Harder、Alex Hernandez-Garcia、Venkatesh Ramesh、Qidong Yang、Prasanna Sattigeri、Daniela Szwar- cman、Campbell Watson 和 David Rolnick。用于气候降尺度的硬约束深度学习。《机器学习研究杂志》，24(365):1–40，2023年。

Nidhin Harilal、Bri-Mathias Hodge、Aneesh Subramanian 和 Claire Monteleoni。STint：用于地理空间数据的自监督时间插值。*arXiv:2309.00059v1 [cs.CV]*, 2023。doi: 10.48550/arXiv.2309. 00059。

Lucy Harris、Andrew T. T. McRae、Matthew Chantry、Peter D. Dueben 和 Tim N. Palmer。《采用生成式深度学习对降水预报进行随机降尺度的方法》。《地球系统建模进展》，14(10):e2022MS003120，2022年。doi: 10.1029/2022MS003120。Ali Hassani、Steven Walton、Jiachen Li、Shen Li 和 Humphrey Shi。邻域注意力变换器。载于*2023*年 *IEEE/CVF*计算机视觉与模式识别会议（*CVPR*），第 6185–6194 页，2023年。doi: 10.1109/CVPR52729.2023.00599。

Hans Hersbach、Bill Bell、Paul Berrisford、Shoji Hirahara、András Horányi、Joaquín Muñoz-Sabater、Julien Nicolas、Carole Peubey、Raluca Radu、Dinand Schepers、Adrian Simmons、Cornel Soci、Saleh Abdalla、Xavier Abellan、Gianpaolo Balsamo、PeterBechtold、Gionata Biavati、Jean Bidlot、Massimo Bonavita、Gio- vanna De Chiara、Per Dahlgren、Dick Dee、Michail Diamantakis、Rossana Dragani、Johannes Flemming、Richard Forbes、Manuel Fuentes、Alan Geer、Leo Haimberger、Sean Healy、Robin J. Hogan、Elías Hólm、

Marta Janisková、Sarah Keeley、Patrick Laloyaux、Philippe Lopez、Cristina Lupu、Gabor Radnoti、Patri- cia De Rosnay、Iryna Rozum、Freja Vamborg、Sebastien Villaume 和 Jean-Noël Thépaut。《ERA5全球再分析》。皇家气象学会季刊，146(730):1999–2049，2020年。doi：10.1002/qj.3803。

Jiaqi Huang、Gongjun Xu 和 Ji Zhu . engression 和逆马尔可夫 engression 的理论分析 . *arXiv:2606.01002v1 [stat.ME]*, 2026。doi：10.48550/arXiv.2606.01002。

Zhewei Huang、Tianyuan Zhang、Wen Heng、Boxin Shi 和 Shuchang Zhou。用于视频帧插值的实时中间光流估计。载于欧洲计算机视觉会议 *(ECCV)*，第 624–642 页，2022年。doi：10.1007/978-3-031-19781-9_36。

Magnus Sikora Ingstad、Mariana C. A. Clare、Olav Ersland、Vera Gahlen、Håvard Homleid Haugen、Ophélia Miralles、Even M. Nordhagen、Thomas N. Nipen、Ivar A. Seierstad、John Bjørnar Bremnes、Michael Maier-Gerber、Zied Ben Bouallègue、Harrison Cook、Christian Lessig、Gert Mertes、Cathal O' Brien、Florian Pinault、Ana Prieto Nemesio 和 Matthew Chantry。HourGlass：用于全球和区域天气预报的概率数据驱动时间降尺度模型。*arXiv:2607.11457v1 [physics.ao-ph]*, 2026。doi：10.48550/arXiv.2607.11457。

Huaizu Jiang、Deqing Sun、Vanja Jampani、Ming-Hsuan Yang、Erik Learned-Miller 和 Jan Kautz。Super SloMo：面向视频插帧的多 Mini 帧高质量估计。载于*2018年IEEE/CVF*计算机视觉与模式识别会议（*CVPR*），第 9000–9008 页，2018年。doi：10.1109/ CVPR.2018.00938。

Tarun Kalluri、Deepak Pathak、Manmohan Chandraker 和 Du Tran。FLAVR：一种可流无关的视频表示方法，用于快速帧插值。载于*IEEE/CVF*冬季计算机视觉应用会议（*WACV*），第 2070–2081 页，2023年。doi：10.1109/WACV56688.2023.00211。

Jeongsol Kim、Bryan Sangwoo Kim 和 Jong Chul Ye。FlowDPS：用于逆问题的流驱动后验采样。载于*2025*年 12337，2025年 12328 10 1109 51701 2025 01146*IEEE/CVF*国际计算机视觉会议（*ICCV*），第 – 页。doi：。 /ICCV . . .

Thorsten Kurth、Shashank Subramanian、Peter Harrington、Jaideep Pathak、Morteza Mardani、David Hall、An- drea Miele、Karthik Kashinath 和 Anima Anandkumar。《FourCastNet：利用自适应傅里叶神经算子加速全球高分辨率天气预报》。载于 高级科学计算平台会议论文集，第 1–11 页，2023年。doi：10.1145/3592979.3593412。

Remi Lam、Alvaro Sanchez-Gonzalez、Matthew Willson、Peter Wirnsberger、Meire Fortunato、Ferran Alet、Suman Ravuri、Timo Ewalds、Zach Eaton-Rosen、Weihua Hu、Alexander Merose、Stephan Hoyer、George Holland、Oriol Vinyals、Jacklynn Stott、Alexander Pritzel、Shakir Mohamed 和 Peter Battaglia。学习高超的全球中期天气预报能力。《科学》，382(6677)：1416–1421，2023年。doi：10.1126/ science.adi2336。

Jussi Leinonen、Daniele Nerini 和 Alexis Berne。利用生成对抗网络对随时间演变的大气场进行随机超分辨率降尺度。《*IEEE*地球科学与遥感汇刊》，59(9):7211–7223，2021年。doi: 10.1109/TGRS.2020.3032790。

Jussi Leinonen、Boris Bonev 和 Thomas Hamilton。用于天气预报时间插值的调制自适应傅里叶神经算子。*arXiv:2410.18904v1 [physics.ao-ph]*, 2024。doi: 10.48550/ arXiv.2410.18904。

Xin Li、Ali Meshgi、Xuan Wang、Jingjie Zhang、Serene Hui Xin Tay、Gerard Pijcke、Nishtha Manocha、Matthias Ong、Manh Tuan Nguyen 和 Vladan Babovic。基于片段法的三种重采样方法，用于将日降水细化为亚日尺度。《国际气候学杂志》，38(S1):e1119–e1138，2018年。doi: 10.1002/joc.5438。

Yaron Lipman, Ricky T . Q. Chen、Heli Ben-Hamu、Maximilian Nickel 和 Matthew Le . 流匹配 f or ge nerative modeling. In国际学习表征会议, 2023年.

F. Lombardo、F. Volpe、D. Koutsoyiannis、and F. Serinaldi。用于间歇性降雨时间分解的理论一致随机级联。《水资源研究》，53(6):4586–4605，2017。doi: 10.1002/2017WR020529。

Edward N. Lorenz。天气与气候的可预测性，第40–58页。剑桥大学出版社，剑桥，2006年。doi: 10.1017/CBO9780511617652.004。

Ilya Loshchilov 和 Frank Hutter。发表于国际学习表征会议, 解耦权重衰减正则化 .

D. Maraun, F. Westerhall, A. Ireson, R. K. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. Rust, T. Sauter, M. Themeßl, V. K. C. Venema, K. P. Chun, C. M. Goodess, R. G. Jones, C. Onof, M. Vrac, 和 I. Thiele-Eich。气候变化背景下降水降尺度：弥合动力模型与最终用户之间差距的最新进展。《地球物理学评论》，48(3):RG3003，2010。doi: 10.1029/ 2009RG000314。

P.-A. Michelangeli, M. Vrac, 和 H. Loukos。概率降尺度方法：应用于风累积分布函数。《地球物理研究快报》，36(11):L11708，2009。doi: 10.1029/ 2009GL038401。

Jonas Olsson、Wai Ming Cheng、Rasmus A. Bergström 和 Yeong Bae Lee。《水文气候学中的降尺度模型》，《水文与地球系统科学》，2(1):19–30，1998年。doi: 10.5194/hess-2-19-1998。

Stephan Rasp、Stephan Hoyer、Alexander Merose、Ian Langmore、Peter Battaglia、Tyler Russell、Alvaro Sanchez-Gonzalez、Vivian Yang、Rob Carver、Shreya Agrawal、Matthew Chantry、Zied Ben Bouallegue、Peter Dueben、Carla Bromberg、Jared Sisk、Luke Barrington、Aaron Bell 和 Fei Sha。《WeatherBench 2：面向下一代数据驱动全球天气模型的基准测试》。《地球系统建模进展》，16(6):e2023MS004019，2024年。doi: 10.1029/2023MS004019。

C. W. Richardson. 日降水、温度和太阳辐射的随机模拟。*Water Re- sourcesResearch*，17(1):182–190，1981年。doi: 10.1029/WR017i001p00182. Chitwan Saharia、Jonathan Ho、William Chan、Tim Salimans、David J. Fleet 和 Mohammad Norouzi。通过迭代细化实现图像超分辨率。《*IEEE*模式分析与机器智能汇刊》，45(4):4713–4726，2023年。doi: 10.1109/TPAMI.2022.3204461. Sebastian Scher 和 Stefanie Peßenteiner。技术说明：利用生成对抗网络对空间降雨场进行时间分解。《水文与地球系统科学》，25(6):3207–3225，2021年。doi: 10.5194/hess-25-3207-2021. Daniel Schertzer 和 Shaun Lovejoy。利用各向异性标度乘法过程对雨和云进行物理建模与分析。《地球物理研究杂志：大气》，92(D8):9693–9714，1987年。doi: 10.1029/JD092iD08p09693. Michael Scheuerer 和 Thomas M. Hamill。针对多变量量概率预报的基于变异函数的适当评分规则。《月刊天气评论》，143(4):1321–1334，2015年。doi: 10.1175/MWR-D-14-00269.1. Maybritt Schillinger、Maxim Samarin、Xinwei Shen、Reto Knutti 和 Nicolai Meinshausen。EnScale：通过适当评分规则实现时间一致的多变量生成式降尺度。《地球系统建模进展》，18(8):e2025MS005593，2026年。doi: 10.1029/2025MS005593. Jonathan Schmidt、Luca Schmidt、Felix M. Strnad、Nicole Ludwig 和 Philipp Hennig。用于气候模拟概率降尺度的时空协调生成框架。*npj Climate and At- mosphericScience*，8(1):270，2025年。doi: 10.1038/s41612-025-01157-y. Johannes Schwab、Stephan Antholzer 和 Markus Haltmeier。用于逆问题的深度零空间学习：收敛性分析与收敛速度。逆问题，35(2):025008，2019年。doi: 10.1088/1361-6420/aaf14a. Xinwei Shen 和 Nicolai Meinshausen。Engression：从分布回归的视角进行外推。《皇家统计学会杂志*B*辑（统计方法学）》，87(3):653–677，2025年。doi: 10.1093/jrsssb/qkae108. Xinwei Shen、Nicolai Meinshausen 和 Tong Zhang。反向马尔可夫学习：用于复杂分布的多步生成模型。*arXiv:2502.13747v2 [cs.LG]*, 2025. doi: 10.48550/arXiv.2502.13747. Karen Stengel、Andrew Glaws、Dylan Hettinger 和 Ryan N. King。气候学风能与太阳能数据的对抗式超分辨率。《美国国家科学院院刊》，117(29):16805–16815，2020年。doi: 10.1073/pnas.1918964117. Hans von Storch。论统计降尺度中“膨胀”的使用。《气候杂志》，12 12 3505 3506, ( ): 1999 10 1175 1520 0442 1999 012<3505 >2 0 2 – 。doi: . / - ( ) :OTUOII . .CO; . Liwen Wang 、 Qian Li 、 Qi Lv 、 Xuan Peng 和 Wei You。TemDeep：用于以任意时间分辨率对大气场进行时间降尺度的自监督框架。*Geoscientific Model Development*，18(8): 2427–2442，2025年。doi: 10.5194/gmd-18-2427-2025. Yinhuai Wang、Jiwen Yu 和 Jian Zhang。使用去噪扩散零空间模型进行零样本图像复原。载于国际学习表征会议，2023年。

Daniel S. Wilks。利用随机天气生成器进行降水降尺度。*Wiley Interdisciplinary Reviews: ClimateChange*，1(6):898–907，2010年。doi: 10.1002/wcc.85.

Florian Zabel 和 Benjamin Poschlod。Teddy 工具 v1.1：用于气候影响分析的每日气候模型数据时间分解。*GeoscientificModelDevelopment*，16(18):5383–5399，2023年。doi: 10.5194/ gmd-16-5383-2023。

附录

A 扩展的相关工作 16

A.1 统计降尺度 .16A.2 神经网络降尺度 .18

B 方法论细节 19

B.1 评分规则与逆回归目标 .19B.2 空间相干性项 .20B.3 训练和采样期间的投影 .20B.4 层数 .21B.5 跨窗口条件化 .21

C 实验详情 25

C.1 Lorenz-96 .25C.2 AutoSim.25C.3 ERA5 .25C.4 训练配置和基线方法 .26C.5 指标 .27

D 补充结果 28

D.1 洛伦兹-96 .28D.2 AutoSim.32D.3 ERA5 .33D.4 AutoSim 完整表格 .37

A 扩展相关工作

本附录扩展了第2节。表5 根据各方法所处理的观测算子、是否返回 $p(x \mid y)$的样本、这些样本是否满足观测要求，以及推理是否需要迭代求解器，对这些方法进行了比较。

A.1 统计降尺度

动力降尺度利用粗尺度数据驱动高分辨率区域模型。在对流允许分辨率下，其成本限制了使用时长或区域范围（Maraun 等人，2010）。统计降尺度则学习粗尺度记录与细尺度记录之间的关系。以下所有方法均为统计方法。

确定性剖面。 MELODIST（Förster 等人，2016年）使用振幅由日最低值和最高值设定的余弦曲线，将日温度分配到一天之中；对于日风速，则使用经站点校准、并按日均值缩放的余弦曲线。 Han 等人（2023年）比较了这类用于风速的确定性模型。在这些模型中，固定剖面忽略了日内波动，因此预测的标准差过低（Förster 等人，2016年）。这些方法返回单条记录；对于温度，它们使用日最低值和最高值，而不是均值。

表5： **时间降尺度方法的定位**。 插值和解聚：该方法分别适用于插值和解聚。概率：返回 p（$x \mid y$）的抽样结果。精确：观测通过构造得以满足（✓）、通过惩罚项、学习型目标或引导近似满足（∼），或完全不满足（×）。无求解器：推理无需迭代式 ODE/SDE 求解器。–：不适用。 †在我们的基线方法中。在我们的实验中，每个学习型基线方法在采样后都会被投影到约束集上（第 4.1 节）。

方法	插值	解聚	概率	精确	无求解器	范围	
经典方法							
线性／三次样条†	✓	×	×	✓	–		
Denton–Cholette† （Denton，1971年；Cholette，1984年）	✓	✓	×	✓	–		最平滑的可接受的记录
CDF-t† （Michelangeli 等人，2009）	–	✓	×	×	–		仅边缘分布
随机级联（Lombardo 等，2017）	×	✓	✓	✓	–		非负间歇性变量
学习型，确定性							
Super SloMo(Jiang 等人，2018年)	✓	×	×	✓	✓		RGB视频
SimVP† （Gao等人，2022年）	✓	✓	×	×	✓		视频预测
STint (Harilal 等人，2023年)	✓	×	×	✓	✓		逐点损失
ModAFNO（Leinonen 等人，2024年）	✓	×	×	✓	✓		逐点损失
TemDeep（Wang 等人，2025年）	✓	×	×	✓	✓		逐点损失
学习型、生成式							
RainDisaggGAN (Scher & Peßenteiner, 2021年)	×	✓	✓	✓	✓		非负变量
Harder 等人 （2023年）	–	–	✓	–	✓		空间块 mean
条件流匹配† （Lipman 等人，2023年）	✓	✓	✓	×	×		条件模型
DiffESM（Bassetti 等人，2024年）	×	✓	✓	×	×		仅条件化
HourGlass（Ingstad 等人，2026）	✓	×	✓	∼	✓		仅从 6 小时到 1 小时
EnScale（Schillinger 等人，2026年）	–	–	✓	×	✓		空间，全球气候模型到 RCM
逆问题求解器							
DDNM（Wang 等人，2023年）	✓	✓	✓	✓	×		无条件先验
FlowDPS† (Kim 等人，2025年)	✓	✓	✓	∼	×		无条件先验
Schmidt 等人(2025年)	✓	✓	✓	∼	×		无条件先验
WIND（Aich 等人，2026）	✓	✓	✓	∼	×		预训练的， 1.5°，每 6 小时
Ours							
MUNDO，单步	✓	✓	✓	✓	✓		
MUNDO ，递归	✓	✓	✓	✓	✓		

约束优化。 计量经济学中的基准评测方法会在聚合约束下，将高频序列调整为低频总量（Denton，1971年；Cholette，1984年；Dagum & Cholette，2006年）。在没有指标序列的情况下，Denton–Cholette 返回一条一阶差分平方和最小、且其聚合量与观测相匹配的记录；对于经过子采样的序列，同一准则会简化为线性插值。Chow & Lin 的回归方法Chow &Lin （1971年）要求在目标分辨率下观测到相关序列，而我们的基准测试中没有任何方法提供此类序列。

分布校正。分位数映射根据模拟分布与目标分布之间的分位数—分位数关系校正每个值（Déqué，2007年），而 CDF-t（Michelangeli 等人，2009年）则通过让目标分布随模拟分布变化来扩展这一方法。逐位置应用时，两者都能校正边缘分布，但既不约束时间顺序，也不约束聚合量。对于缺乏方差的重建，常用的两种修正方法都无法恢复公式（7）中的条件方差项：膨胀只能放大已由 y （von Storch，1999年）解释的变异性，而随机化则用时间结构预先设定、而非通过学习获得的噪声来表示 y 尚未解释的变异性。

天气生成器与重采样。随机天气生成器根据拟合分布模拟日序列，并采用马尔可夫链或天气类型发生模型（Wilks，2010年；Richard-son，1981年；Ailliot 等人，2015年），或对档案数据进行重采样（Buishand & Brandsma，2001年）。片段法利用档案中具有相似日值的一天的日内分数来分解一天；按构造，此方法可以保留日总量（Li 等人，2018年）。Teddy（Zabel& Poschlod，2023年）将同一原理应用于七个变量。两种方法的日内形态都取自档案数据。

随机级联。 随机级联通过使用随机权重逐级将每个时间区间拆分为子时间区间，从而对降雨进行细化分解。微正则级联在每次拆分时严格守恒体积（Olsson，1998年），正则级联仅在平均意义上守恒（Schertzer & Lovejoy，1987），而 Lombardo 等人 （2017）提出的非乘法级联通过调整步骤恢复总量。级联使用预先设定的参数模型对一个非负间歇变量进行建模，而 MUNDO 的各层则学习给定父节点时子节点的分布，并能精确复现该分布。

A.2 神经网络降尺度

确定性插值。 视频帧插值利用光流（Jiang 等人，2018年；Huang 等人，2022年）或不利用光流（Kalluri 等人，2023年）重建两个 RGB 帧之间的帧。地球科学适配方法包括 STint（Harilal 等人，2023年），该模型通过循环一致性进行预训练；TemDeep（Wang 等人，2025年），该模型以连续时间偏移为条件；以及 ModAFNO（Leinonen 等人，2024年），该模型将六小时分辨率的 ERA5 状态插值为小时分辨率。它们的逐点损失会将输出拉向某个中心值 $p(x \mid y)$ （第3.3节），并将边界状态作为观测，因此都无法处理聚合量。我们的确定性基线是 SimVP（Gao 等人，2022年），一种视频预测模型。

生成模型。 逐点保真度与分布真实性之间存在权衡（Blau&Michaeli，2018年），这推动了生成式降尺度的发展。最初，它采用对抗网络（Stengel 等人，2020年；Leinonen 等人，2021年；Harris 等人，2022年）开发，随后又采用扩散（Sa- haria 等人，2023年）和流匹配（Lipman 等人，2023年）。在空间尺度上，EnScale（Schillinger 等人，2026）通过在能量评分上训练一个注入噪声的生成器来模拟区域气候模型，并采用由粗到细的阶段式训练；这些阶段以块平均目标为条件，但不强制拟合这些目标。在时间尺度上，DiffESM（Bassetti 等人，2024年）将地球系统模型输出从月尺度降尺度到日尺度，以粗分辨率记录为条件但不强制拟合，并对每次抽样进行迭代求解。与本研究同期的 HourGlass（Ingstad 等人，2026）采用近似公平 CRPS 训练单次生成器，以生成两个六小时预报状态之间的五个小时状态以及最终状态。对于预报变量，最终状态的残差目标为零，因此模型学习观测而非强制拟合观测；对窗口统计量的惩罚则有助于提升时间一致性。该方法仅处理六小时到小时的插值。我们采用这一类方法中的基线：条件流匹配。

硬约束网络。 RainDisaggGAN（Scher & Peßenteiner，2021年）通过对时间维度应用 Softmax 生成日降水总量的小时份额，并将其乘以总量，因此每日总和是精确的。这种构造要求变量非负，且观测为聚合量。 Harder 等人（2023年）添加约束层（加法、乘法或 Softmax），使空间超分辨率能够保持粗网格块均值不变。对于分块平均，加法层即为公式（9）中的投影。他们的时空模型使用光流插值一个中间时间步，并采用平方损失；其约束仅作用于空间均值。

逆问题求解器。 逆问题求解器训练无条件先验，并在采样期间施加观测约束。DDNM（Wang 等人，2023年）在每一步都将去噪估计的值域空间分量替换为观测所确定的分量；对于无噪声线性算子，这一做法是精确的。FlowDPS（Kim 等人，2025年）在流匹配采样中加入似然梯度步。 Schmidt 等人（2025年）在再分析数据上训练无条件的基于评分的先验。为生成小时集合，他们遵循基于评分的数据同化方法，通过观测模型以粗分辨率的六小时气候模拟为条件。WIND（Aich 等人，2026）在 ERA5 上训练一个视频扩散先验，涵盖 1.5° 以及 70 个变量的六小时时步，并通过引导而无需微调来处理预测、降尺度和重建。在其自身的日尺度到六小时尺度拆分实验中，特定任务模型取得了更低的均方根误差。其发布的检查点与我们的 ERA5 设置（0.25°，小时尺度）不匹配，而重新训练先验超出了我们的预算。所有这些求解器都需要对每次抽样进行迭代求解。我们采用该类方法中的 FlowDPS 作为基线，使用在我们的数据上训练的先验，并将其样本投影到约束集上。

B 方法论细节

B.1 S 评分规则与 engression 目标 TIVE

评分规则 $S(F,\ v)$ 用于评估预测分布 $F \in \mathcal{P}$ 与观测值 $v \in \mathbb{R}^D$。根据方程（12）中的损失约定，当满足以下条件时，它是 适当的：

$$\mathbb{E}_{v\sim Q}\big[S(F,v)\big] \geq \mathbb{E}_{v\sim Q}\big[S(Q,v)\big] \tag{15}$$

对于所有 $F,\ Q \in \mathcal{P}$，且当等号成立意味着 严格适当 $F = Q$ (Gneiting 和 Raftery，2007）。因此，真实分布唯一地使总体评分达到最小值。

指数为 $\beta \in (0,\ 2)$ 的能量评分为

$$S(P,v) = \mathbb{E}\,\|U-v\|_2^\beta - \frac{1}{2}\,\mathbb{E}\,\|U-U'\|_2^\beta\,, \tag{16}$$

其中 $U,\ U'$ 独立同分布$\sim P$。对于具有有限 β 阶矩的分布，它是严格适当的。沿用 engression，我们取 $\beta = 1$ （Gneiting 和 Raftery，2007；Shen 和 Meinshausen，2025年）。

条件化应用时，严格适当性意味着，对于 p_Y-几乎处处的 y，

$$\mathbb{E}_{X\sim p^*(\cdot|y)}\left[S\left(p^*(\cdot\mid y),X\right)\right] \leq \mathbb{E}_{X\sim p^*(\cdot|y)}\left[S\left(p_g(\cdot\mid y),X\right)\right], \tag{17}$$

且仅当 $p_g\left(\cdot\mid y\right)\ = p^*\left(\cdot\mid y\right)$ 时等号成立。对 Y 进行平均可得逆回归目标（Shen 和 Meinshausen，2025年），

$$g^* \in \arg\min_{g\in\mathcal{G}} \mathbb{E}\left[\|X-g(Y,\varepsilon)\|_2 - \frac{1}{2}\,\|g(Y,\varepsilon)-g(Y,\varepsilon')\|_2\right], \tag{18}$$

其中 ε和 ε' 是独立噪声样本。在可实现性条件下，其总体最优解满足$g^*\left(Y,\ \varepsilon\right)\ \mid Y = y \sim p^*\left(\cdot\mid y\right)$ 对于 p_Y- 几乎处处的 y 均成立。无需进行密度评估，因此同一目标函数可在（ 11 ）的每个层级使用。

局限性。 能量评分是严格适当的，但在高维情况下，其对空间依赖的有限样本判别能力可能较弱（Scheuerer & Hamill，2015）。式（12）的两项仅取决于整个场上的欧氏距离，因此即使局部结构或谱结构不同，不同的空间排列也可能获得相近的评分。严格适当性保证了在非限制分布上的正确总体最优解，但不能保证受限模型在有限样本优化下准确恢复空间依赖。

有限集合估计量。 训练期间使用的集合估计量也必须避免因成员数量较少而惩罚离散度。对于 $M = 2$，能量评分中的公平成对项简化为 $\frac{1}{2}\|u-u'\|$，因此，公式（12）已经是相应总体量的无偏估计量。式（14）中的变差函数估计量采用相同的跨成员构造。相比之下，集合均值变换增量的朴素平方误差会向上偏差 $_{1\ 2}\,\mathrm{Var}[\gamma_h]$，从而隐含地惩罚离散度。在交叉成员乘积中使用独立抽样，会使相应的中心化误差不相关，因此，这个额外的方差项在期望上消失。

表 6：**空间相干项消融**。 Gross–Pitaevskii 解聚，取值为 $N = 24$，使用三个随机种子，并采用跨窗口条件输入。能量表示能量评分，变差函数表示式（14）中的变差函数评分，谱表示谱 CRPS。阶段 1 为逐层训练目标，阶段 2 为端到端微调目标；– 表示不进行微调。 † 标记主实验中使用的配置。SSR 为离散度-技巧比， s_1 为单格空间变差函数比，最精细频带为最精细频带谱功率比；它们的目标值均为 1。

Arm	阶段 1	阶段 2	连续排序概率得分 ↓	SSR → 1	$s_1 \to 1$	最精细频带 → 1
MUNDO ($R = 4$)	能量	–	0.1047±0.0004	**0.85**	1.131	2.15±0.06
	能量 +变异函数	–	**0.1025**±0.0002	**0.85**	1.073	2.78±0.30
	能量 +谱	–	**0.1025**±0.0005	0.83	**1.055**	**1.96**±0.10
	能量 +变异函数	能量	0.0802±0.0008	0.92	1.089	2.75±0.19
	能量 +变异函数	能量 + 变异函数	**0.0788**±0.0007	0.92	**1.057**	2.21±0.19
	能量 +变异函数	能量 + 谱†	0.0799±0.0010	0.92	**1.057**	**1.17**±0.03
	能量	能量 + 谱	0.0801±0.0000	0.92	1.083	1.21±0.05
	能量 +谱	能量 +谱	**0.0795**±0.0003	**0.93**	**1.062**	**1.12**±0.04
MUNDO ($R = 1$)	能量	–	0.0866±0.0005	0.83	1.135	2.97±0.35
	能量 +变异函数†	–	**0.0828**±0.0009	0.83	**1.026**	**1.08**±0.08

B.2 空间相干性项

我们消融了式（13）中的辅助空间相干性惩罚项。在逐层训练期间，我们比较单独使用能量评分与分别加入变差函数评分或谱 CRPS 的效果。在端到端微调期间，我们比较同样的三种目标函数。单步 MUNDO 没有微调阶段，因为它唯一的父节点就是观测本身。谱 CRPS 分别针对每个二维傅里叶系数的实部和虚部独立计算（Bonev 等人，2025年）。

表 6报告了 $N = 24$上的 Gross–Pitaevskii 解聚对比结果。由于消融旨在评估对精细空间尺度的恢复效果，我们还报告了最精细频带谱功率比（表中记为 Finest），以及单元空间变差函数比 s_1；两者的定义见附录 C.5。对于单步 MUNDO，变差函数项十分重要：移除该项后，精细尺度谱功率会被大幅高估，且 s_1 会恶化。对于递归 MUNDO，逐层训练中的变差函数项影响较小；而采用谱 CRPS 进行端到端微调后，无论逐层训练采用哪种目标函数，最精细频带功率都接近真值。

B.3 训练与采样期间的投影

公式（9）定义了仿射投影 $P_y(u) = A^\dagger y + (I - A^\dagger A)\, u$ 到 $\mathcal{X}_y$。递归 MUNDO 在每个层级都对投影样本计算训练损失。单步 MUNDO 则基于未投影重建 $A^\dagger y + G_\theta(y,\ \varepsilon,\ c)$ 进行训练，并在训练后使用 P_y 投影生成样本。

假设存在有限一阶矩且满足可实现性，则当评分所评估的分布等于目标条件分布时，期望能量评分取得唯一最小值（附录B.1）。若在损失函数内部应用投影，则这一保证适用于投影后的分布。若在未投影样本上计算损失，则最优的未投影分布本身已等于目标分布，且经投影后保持不变。因此，两种形式都能在总体最优时恢复目标分布。然而，对于有限网络而言，训练目标的选择可能会影响生成分布。

表7 比较了基准制度下，AutoSim 解聚任务中单步 MUNDO 的两种训练选择，针对 $N = 32$。在损失函数中应用投影会使所有四个系统的连续排序概率得分上升，而其对校准的影响不一，对时间结构的影响则相对较小。无论何时应用投影，单步 MUNDO 在所有四个系统上的连续排序概率得分都低于基线方法。

表 7： **训练或采样期间投影的效果**。AutoSim 解聚、基准层级， $N = 32$。

	仅采样				在训练损失中			
系统	连续排序概率得分 ↓	SSR $\to 1$	覆盖率平均绝对误差 ↓	$t_1 \to 1$	连续排序概率得分 ↓	SSR $\to 1$	覆盖率平均绝对误差 ↓	$t_1 \to 1$
平流-扩散	0.0020	0.83	0.140	1.04	0.0023	0.95	0.135	1.01
Gray–Scott	0.0032	1.35	0.048	1.01	0.0034	1.76	0.058	1.00
纳维–斯托克斯	0.0129	0.53	0.095	1.02	0.0131	0.57	0.064	1.02
Gross–Pitaevskii	0.1249	0.85	0.049	0.91	0.1262	0.97	0.066	0.97

非负变量。投影 P_y 施加了 $Ax = y$，但不约束变量的符号。在 ERA5 总降水量中，MUNDO 生成的像素小时中有 28%（递归式）和 24%（单步式）为负值，而 Denton–Cholette 为 17%，流匹配为 20%。对于此变量，我们将 P_y 替换为 $\mathcal{X}_y \cap \{x \geq 0\}$上的欧几里得投影。在解聚任务中，该约束通过日均值 $\bar{y} \geq 0$将一个像素的 N 小时耦合起来，而一个数据块 $u_{1:N}$ 的投影为

$$x_t = \max(u_t - \lambda,\, 0), \tag{19}$$

其中 λ 满足 $\frac{1}{N}\sum_{t=1}^{N} \max(u_t - \lambda,\ 0) = \bar{y}$，并通过二分法求得。没有负值的区块保持不变（$\lambda = 0$）。在插值方法中，生成的小时值不受该约束耦合，且保留状态为非负，因此投影简化为截断至零。我们在采样后将其应用于每种方法。或者，也可以在对数（$x + \varepsilon$）空间中生成非负变量，再按观测到的聚合量缩放，从而避免源数据中出现负值。

B.4 层数

粗化比率 k 固定后，层数 R 决定相应质因数在层级结构中的分布方式。本文始终采用最精细的因式分解，即每个质因数对应一个层级。较粗的因式分解会将多个因子合并为一个步骤，而该步骤在层级结构中的位置可以变化。表 8考察了 R 对 $N = 32$纳维–斯托克斯方程解聚的影响，涵盖从单步 MUNDO（$R = 1$）到最精细因式分解（$R = 5$）的各类配置，以及 $R = 2$、3和4的所有可能因子排序。我们会针对不同配置调整骨干网络宽度，使参数数量控制在 模型的 $R = 5$ 10% 以内。

增加层数可以改善校准，但会牺牲连续排序概率得分（图 7）。随着 R从 1 增加到 5，连续排序概率得分从 0.0129 上升到 0.0159（+23%），而 SSR 从 0.53 增加到 0.70，覆盖误差则从 0.095 降至 0.020。值得注意的是， $R = 4$ 已经达到了与 $R = 5$相当的校准性能：最粗层因子为 2 的三个配置，其 SSR 和覆盖误差均与之相当，同时连续排序概率得分低 9%。在所有深度下， t_1都保持稳定，范围为 1.02 至 1.05。

除层数之外，因子排序也会影响性能，主要体现在最粗层上。对于 $R = 3$ 和 $R = 4$，最粗因子为 4 的配置相比同等深度下的其他排序方式，连续排序概率得分高出 4–7%，同时 SSR 更高、覆盖误差更低。相比之下，较精细层级中的因子排序影响很小：在 $R = 4$下，将因子 4 调整到这些层级中的不同位置，连续排序概率得分的变化小于 1%；而在 $R = 2$下，两种可能的排序方式也只会产生同样微小的差异。因此，中等深度配置引入了一种排序选择，在 $R = 4$下可能使连续排序概率得分相差最多 7%。最精细因式分解不合并任何因子，因此避免了这种歧义，这也是我们选择在全文中采用该方式的原因。

B.5 跨窗口条件化

条件输入 c 在（9）中可以包含来自相邻窗口的两类信息。跨窗口上下文对于解聚尤其有用，因为连续的重建窗口之间并不共享观测到的边界状态。

观测上下文。 第一种是观测上下文：前序窗口和后续窗口的聚合量，每个聚合量都配有一个存在标志，并在训练期间随机丢弃，因此同一个模型可以在有或没有这些信息的情况下运行。表 9 评估了相同训练模型在

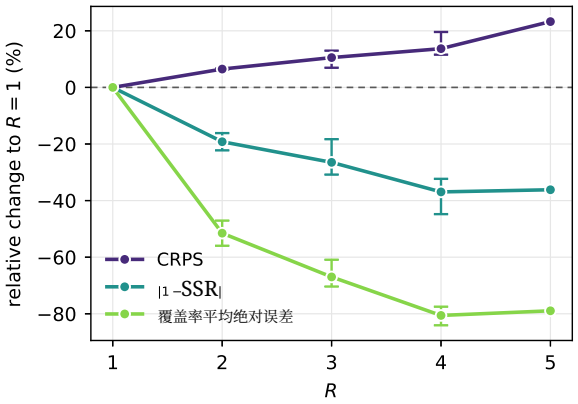


图 7：**层级数量的影响** R。纳维–斯托克斯方程解聚，$N = 32$。每条曲线表示相对于单步 MUNDO（$R = 1$）的评分变化百分比。对于这三个评分，数值越低越好，因此负值表示改进：$|1 - \text{SSR}|$ 表示离散度-技巧比与 1 的距离，覆盖率平均绝对误差表示覆盖误差。对于每个 R，标记表示表8中各链的均值，条形范围覆盖其最小值和最大值；$R = 1$ 和 $R = 5$ 各自只有一条链。

表 8：**层级配置和因子排序**。 $N = 32$ 下的纳维–斯托克斯方程解聚结果，取三个随机种子的平均值。对于连续排序概率得分，还报告了随机种子间的标准差。每条链均按从最粗到最精细的顺序列出层级因子。参数数量（参数）是所有层级的总和。

R	链	宽度	参数	连续排序概率得分 ↓	SSR → 1	覆盖率平均绝对误差 ↓	$t_1 \to 1$
1	[32]	85	12.6 M	**0.0129**±0.0001	0.53	0.095	**1.02**
2	[4,8]	60	13.7 M	0.0138±0.0001	0.61	0.050	**1.02**
	[8,4]	60	13.7 M	0.0137±0.0001	0.63	0.042	1.03
3	[2,4,4]	48	12.0 M	0.0138±0.0000	0.62	0.037	1.03
	[4,2,4]	48	12.0 百万	0.0144±0.0001	0.67	0.028	1.03
	[4,4,2]	48	12.0 百万	0.0146±0.0001	0.67	0.029	1.04
4	[2,2,2,4]	41	11.4 百万	0.0144±0.0001	0.68	0.021	1.03
	[2,2,4,2]	41	1140 万	0.0144±0.0002	0.69	0.019	1.04
	[2,4,2,2]	41	1140 万	0.0144±0.0001	0.70	0.018	1.04
	[4,2,2,2]	41	1140 万	0.0154±0.0002	**0.74**	**0.015**	1.04
5	[2,2,2,2,2]	37	12.7 M	0.0159±0.0002	0.70	0.020	1.05

在所有基准制度下的窗口长度中，均提供或屏蔽跨窗口上下文。短窗口的收益最大：在 $N = 8$ 时，跨四个系统，MUNDO 的连续排序概率得分降低了 25–87%，且跨窗口步长更接近真值；随着 N 增加，这两种效果都会减弱。校准变化相对较小，只有 Gray–Scott 例外，新增的上下文会显著改变集合离散度。即使流匹配保留了相邻聚合量，在基准制度下的 $N = 32$，屏蔽相邻聚合量也不会逆转单步比较结果：四个系统上的连续排序概率得分仍然更低（分别为 0.0023、0.0036、0.0153、0.1371，而对照值分别为 0.0026、0.0049、0.0236、0.1578；对应的系统依次为平流-扩散、Gray–Scott、Navier–Stokes 和 Gross–Pitaevskii）。

生成的上下文。第二种是生成的上下文：递归 MUNDO 还可以在每个层级上，根据该层级分辨率下前一窗口的最后一个状态进行条件设定。训练期间，这个状态来自真实轨迹。推理时，它来自同一集成成员生成的前一窗口。第 3.5 节的端到端微调解决了这种训练集与测试集之间的不匹配。

在给定观测值的条件下，生成的窗口构成一个马尔可夫链。不可用的相邻观测值会被屏蔽，并通过存在标志指示其是否可用。模型的分布

表9：**跨窗口条件消融**。在 AutoSim 基准测试情境下进行解聚，采用三个随机种子，并使用相同的训练模型进行评估，跨窗口上下文要么被屏蔽，要么作为输入。CRPS 增益为 1 –连续排序概率得分$_{\text{fed}}$/连续排序概率得分$_{\text{masked}}$。 r_{seam} 是跨窗口步长比，定义见（34），该比值为 1 表示与真值一致。

		MUNDO					MUNDO（$R=1$）				
		CRPS	SSR $\to 1$		$r_{\text{seam}} \to 1$		CRPS	SSR $\to 1$		$r_{\text{seam}} \to 1$	
系统	N	gain	掩码的	fed	掩码的	fed	gain	掩码的	fed	掩码的	fed
平流-扩散	8	87%	0.97	0.92	2.56	1.03	83%	0.97	0.96	2.22	1.05
	16	49%	0.95	0.94	2.06	1.16	48%	0.94	0.92	1.74	1.12
	24	36%	0.91	0.93	1.88	1.24	30%	0.86	0.84	1.61	1.20
	32	23%	0.94	0.97	1.79	1.39	14%	0.85	0.83	1.49	1.28
Gray–Scott	8	31%	1.31	1.41	1.02	1.01	9%	1.20	1.22	1.01	1.01
	16	21%	1.10	1.19	1.05	1.01	9%	1.27	1.26	1.03	1.02
	24	19%	1.00	1.10	1.09	1.03	9%	1.28	1.35	1.06	1.03
	32	19%	0.86	1.01	1.17	1.07	11%	1.25	1.35	1.13	1.05
纳维–斯托克斯	8	59%	0.89	0.93	1.47	1.03	48%	0.82	0.84	1.33	1.03
	16	46%	0.85	0.85	2.20	1.26	42%	0.74	0.74	1.87	1.19
	24	29%	0.76	0.76	2.87	1.84	27%	0.63	0.62	2.48	1.63
	32	17%	0.70	0.70	4.03	2.97	16%	0.53	0.53	3.29	2.45
Gross–Pitaevskii	8	25%	0.79	0.74	1.04	1.01	27%	0.87	0.92	1.02	1.01
	16	15%	0.83	0.79	1.42	1.26	19%	0.89	0.91	1.28	1.18
	24	10%	0.88	0.85	1.86	1.65	12%	0.82	0.83	1.60	1.46
	32	8%	0.91	0.90	2.39	2.13	9%	0.84	0.85	1.93	1.75

分布 q 分解为

$$q\left(x_{1:T} \mid y_{1:T}\right) = q\left(x_1 \mid y_{1:2}\right) \prod_{t=2}^{T} q\left(x_t \mid x_{t-1}, y_{t-1:t+1}\right). \tag{20}$$

每个窗口级因子均由多分辨率采样器定义。对于 $t \geq 2$，记 $b_{t-1}^{(r)}$ 为前一窗口在分辨率 r下的最后一个状态，则有

$$q\left(x_t \mid x_{t-1}, y_{t-1:t+1}\right) = \int \prod_{r=0}^{R-1} q_r\left(x_t^{(r)} \mid x_t^{(r+1)}, b_{t-1}^{(r)}, y_{t-1:t+1}\right) \mathrm{d}x_t^{(1)} \cdots \mathrm{d}x_t^{(R-1)}, \tag{21}$$

其中 $x_t^{(0)} = x_t$ 和 $x_t^{(R)} = y_t$。第一个窗口采用相同的层级结构，但会屏蔽生成的上下文。在每个层级限制为 $b_{t-1}^{(r)}$ 是一项建模假设。我们将生成的上下文留给递归 MUNDO 使用，因为它在每个中间分辨率下都可用。

表 10比较了以真实和生成的前一窗口上下文作为条件输入的效果。在大多数设置下，两者的连续排序概率得分相差约2% 以内，但在对流-扩散问题中， $N = 8$反馈生成的状态会降低连续排序概率得分和校准效果。

递归变体。因此，在解聚设置下，递归 MUNDO 有四种变体：是否使用生成的上下文，以及是否进行端到端微调，两者各有两种选择。在插值设置下，强制共享观测端点可使相邻窗口在交界处保持一致，因此我们不使用生成的跨窗口上下文，只改变微调设置。表中报告了普通变体或微调变体；针对每种 AutoSim 系统、可预测性机制和窗口长度，仅依据验证诊断进行选择，并在与其他所有行相同的测试窗口上评分。因果链以前一窗口的自身输出为条件，因此评分对象是链，即每条测试轨迹中除第一个窗口以外的所有窗口；表11 比较了这四种变体与单步 MUNDO 在这些窗口上的表现，针对基准测试和混沌机制下的 $N \in \{24, 32\}$ 。在这些窗口上，每种设置中，微调后的链都是四种变体中表现最佳的；在纳维–斯托克斯和 Gross–Pitaevskii 系统上，其连续排序概率得分比普通递归 MUNDO 低13–35% 。与单步 MUNDO 相比，它在 Gross–Pitaevskii 和混沌平流扩散上表现更好，在纳维–斯托克斯上持平，而在 Gray–Scott 和基准测试平流扩散上表现较差。对于 Lorenz-96，我们则预先固定变体：解聚采用因果链，插值采用普通采样器，两者都不进行微调。

表 10：**生成的跨窗口上下文**。 解聚，三个随机种子。无输入：递归 MUNDO 在不使用前一窗口输入的情况下训练。其余三列报告同一模型在使用该输入训练后，分别在评估时屏蔽该输入、填入真实前一状态，或填入同一集成成员生成的前一窗口（即部署时的做法）时的结果。每行中最低的连续排序概率得分以粗体显示；SSR $\to 1$； t_1 单步时间变差函数比， $\to 1$。

			无输入			掩码的			true			已生成		
系统	机制	N	连续排序概率得分	SSR	t_1	连续排序概率得分	SSR	t_1	连续排序概率得分	SSR	t_1	连续排序概率得分	SSR	t_1
对流-扩散基准测试		8	0.00044	0.89	1.02	0.00044	0.94	1.02	**0.00031**	0.90	1.02	0.00047	2.26	1.02
平流-扩散 温和		8	0.00395	1.23	1.04	0.00384	1.20	1.04	**0.00319**	1.13	1.03	0.02587	1.93	1.53
平流-扩散 混沌		8	0.01514	1.18	1.07	0.01516	1.20	1.07	**0.01411**	1.24	1.09	0.02774	1.76	1.41
平流-扩散 基准测试		16	0.00173	0.94	1.02	0.00161	0.96	1.02	**0.00145**	0.94	1.02	0.00146	0.95	1.02
平流-扩散 温和		16	0.00657	1.02	1.02	0.00651	1.01	1.02	**0.00610**	1.02	1.02	0.00617	1.03	1.02
对流-扩散混沌		16	0.02107	1.09	1.03	0.02131	1.17	1.05	**0.02094**	1.17	1.05	0.02111	1.19	1.06
Gray–Scott	基准测试	8	0.00049	0.89	1.01	0.00054	0.91	1.01	**0.00046**	0.87	1.01	0.00047	0.86	1.01
纳维–斯托克斯	基准测试	8	0.00173	0.94	1.04	0.00184	0.94	1.04	**0.00166**	0.95	1.04	0.00167	0.95	1.04

表 11：**AutoSim 上的递归 MUNDO 变体**。测试轨迹上的连续排序概率得分和覆盖率平均绝对误差，使用三个随机种子。解聚：在相同的链窗口（每条测试轨迹中除第一个以外的所有窗口）和成员上比较四种变体与单步 MUNDO；涵盖普通版和微调版，以及不使用和使用因果链生成的上下文的情况。插值：在标准评估窗口上比较普通版、微调版和单步 MUNDO。 ⋆：主表中报告的变体，普通版或微调版，根据验证轨迹选定。粗体：每行每项指标中的最低值，精度相同时并列。

			纯文本		微调		因果链		因果链，微调		单步	
系统	机制	N	连续排序概率得分 ↓	覆盖率平均绝对误差 ↓	连续排序概率得分 ↓	覆盖率平均绝对误差 ↓	连续排序概率得分 ↓	覆盖率平均绝对误差 ↓	连续排序概率得分 ↓	覆盖率平均绝对误差 ↓	连续排序概率得分 ↓	覆盖率平均绝对误差 ↓
解聚												
对流-扩散基准测试		24	0.0022⋆		**0.114**	0.0020	0.118	0.0020	0.117	0.0018	0.125	**0.0017**	0.158
对流-扩散基准测试		32	0.0027⋆		**0.093**	0.0026	0.100	0.0026	0.100	0.0025	0.108	**0.0021**	0.142
对流-扩散混沌		24	0.0247		0.014	0.0230⋆	**0.010**	0.0242	0.016	**0.0225**	0.014	0.0344	0.043
对流-扩散混沌		32	0.0281		0.026	0.0265⋆	**0.010**	0.0283	0.032	**0.0256**	0.013	0.0412	0.042
Gray–Scott	基准测试	24	0.0014⋆		**0.048**	0.0014	0.051	0.0014	**0.048**	**0.0013**	0.051	**0.0013**	0.052
Gray–Scott	基准测试	32	0.0019⋆		**0.044**	0.0019	0.049	0.0018	**0.044**	0.0018	0.049	**0.0016**	0.053
Gray–Scott	混沌	24	0.0100		0.061	0.0091⋆	0.068	0.0088	0.059	0.0082	0.069	**0.0070**	**0.043**
Gray–Scott	混沌	32	0.0183		0.054	0.0163⋆	0.058	0.0160	0.055	0.0146	0.061	**0.0115**	**0.037**
纳维–斯托克斯	基准测试	24	0.0084		**0.023**	0.0081⋆	0.024	0.0071	0.038	**0.0069**	0.038	0.0071	0.048
纳维–斯托克斯	基准测试	32	0.0153⋆		**0.019**	0.0147	**0.019**	0.0127	**0.019**	**0.0123**	0.021	0.0125	0.095
纳维–斯托克斯	混沌	24	0.0574		0.028	0.0541⋆	0.037	0.0503	0.027	**0.0484**	0.035	0.0519	**0.017**
纳维–斯托克斯	混沌	32	0.0869		0.027	0.0823⋆	0.034	0.0777	**0.020**	0.0754	0.024	**0.0746**	0.021
Gross–Pitaevskii	基准测试	24	0.1146		0.068	0.0899⋆	0.055	0.1002	0.074	**0.0748**	0.078	0.0938	**0.048**
Gross–Pitaevskii	基准测试	32	0.1523		0.072	0.1291⋆	**0.042**	0.1384	0.077	**0.1130**	0.062	0.1345	0.044
插值												
对流-扩散基准测试		24	0.0015⋆		**0.099**	0.0014	0.118	–	–	–	–	**0.0010**	0.172
对流-扩散基准测试		32	0.0022⋆		**0.090**	0.0020	0.106	–	–	–	–	**0.0013**	0.160
对流-扩散混沌		24	0.0222		**0.036**	0.0207⋆	**0.036**	–	–	–	–	**0.0185**	0.048
对流-扩散混沌		32	0.0285		0.041	0.0266⋆	**0.032**	–	–	–	–	**0.0213**	0.047
Gray–Scott	基准测试	24	0.0030⋆		0.044	0.0029	**0.039**	–	–	–	–	**0.0022**	0.059
Gray–Scott	基准测试	32	0.0035⋆		**0.037**	0.0034	0.039	–	–	–	–	**0.0024**	0.060
Gray–Scott	混沌	24	0.0115		**0.037**	0.0100⋆	0.041	–	–	–	–	**0.0074**	0.041
Gray–Scott	混沌	32	0.0203		**0.031**	0.0169⋆	0.040	–	–	–	–	**0.0112**	0.032
纳维–斯托克斯	基准测试	24	0.0074⋆		**0.017**	0.0071	0.020	–	–	–	–	**0.0063**	0.048
纳维–斯托克斯	基准测试	32	0.0117⋆		**0.015**	0.0112	0.022	–	–	–	–	**0.0091**	0.072
纳维–斯托克斯	混沌	24	0.0506		0.028	0.0427⋆	0.033	–	–	–	–	**0.0370**	**0.022**
纳维–斯托克斯	混沌	32	0.0842		**0.024**	0.0663⋆	0.036	–	–	–	–	**0.0650**	0.028
Gross–Pitaevskii	基准测试	24	0.0320		0.072	0.0253⋆	**0.055**	–	–	–	–	**0.0226**	0.082
Gross–Pitaevskii	基准测试	32	0.0660		0.056	0.0450⋆	**0.034**	–	–	–	–	**0.0406**	0.087

C 实验详情

C.1 Lorenz-96

双尺度 Lorenz-96 模型（Lorenz,2006年）将每个慢变量 X_k 与 J 快变量$Y_{j相耦合,\ k}$ 国产成人

$$\frac{d}{dt}X_k = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=1}^{J} Y_{j,k}, \tag{22}$$

$$\frac{d}{dt}Y_{j,k} = -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + \frac{hc}{b}X_k, \tag{23}$$

其中 $h = 1$、 $b = 10$和 $c = 10$。我们使用 $K = 36$ 和 $J = 32$。我们将强迫项在$F \in \{10, 24, 32, 64\}$上变化，以涵盖混沌性逐渐增强的机制，其中 $F = 10$为非混沌， $F = 64$为强混沌。系统采用四阶龙格–库塔方法进行积分，步长为 10^{-3} 个时间单位，在经过 100 个时间单位的自旋启动后运行 1000 个时间单位。仅观测和重构慢变量。快变量保持潜在状态，并作为慢动力学的未解析强迫项。

目标分辨率为 $\delta = 0.01$ 个时间单位，窗口长度为 $N \in \{64, 128\}$ 步。轨迹按顺序划分为 70% 训练集（时间单位 100–800）、15% 验证集（800–950）和 15% 测试集（950–1100），且窗口不会跨越划分边界。训练从训练区间采样 50 000 个窗口。评估采用 25 条由连续测试窗口组成的链，各窗口之间至少间隔 0.6 个时间单位，每个窗口包含 256 个集合成员。每条链在 $N = 64$ 时包含 8 个窗口，或在 $N = 128$ 时包含 4 个窗口。我们排除没有前序窗口的第一个窗口，因此分别剩下 175 个和 75 个评分窗口。

C.2 AUTOSIM

我们使用 AutoSim（ 64×64 生成四种物理系统（Greenbury 等人，2026），涵盖输运、反应扩散、湍流和非线性色散动力学。其控制方程为

平流-扩散

$$\partial_t \omega = \nu \nabla^2 \omega - \mu\, \mathbf{u} \cdot \nabla \omega, \tag{24}$$

Gray–Scott

$$\begin{aligned} \partial_t u &= \delta_u \nabla^2 u - uv^2 + F(1-u), \\ \partial_t v &= \delta_v \nabla^2 v + uv^2 - (F+k)\, v, \end{aligned} \tag{25}$$

条件化 Navier–Stokes

$$\begin{aligned} \partial_t s &= -\mathbf{u} \cdot \nabla s, \\ \partial_t \mathbf{u} &= -(\mathbf{u} \cdot \nabla)\mathbf{u} - \nabla p + \nu \nabla^2 \mathbf{u} + \beta s \mathbf{e}_y, \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned} \tag{26}$$

Gross–Pitaevskii

$$i\,\partial_t \psi = \left[-\tfrac{1}{2}\nabla^2 + V(\mathbf{x}, t) + g|\psi|^2\right] \psi. \tag{27}$$

重建的场包括：平流-扩散中的涡量、Gray–Scott 中的两种反应物、条件化 Navier–Stokes 中的被动标量和速度分量，以及 Gross–Pitaevskii 中的密度与 ψ 的实部和虚部。在平流-扩散中，$\mathbf{u}$是由 ω 通过其流函数诱导的速度。在 Gross–Pitaevskii 中， V是一个带有振荡障碍物的盒形势阱。

我们为每个系统生成 200 条训练轨迹以及 20 条验证集和测试集轨迹，并按轨迹划分数据集。我们评估长度为 $N \in \{8, 16, 24, 32\}$步的窗口，每个窗口包含 $M = 64$个集合成员；每个系统在测试轨迹上选取 200 个测试窗口（插值时为 180 个）。除基准配置外，我们还采用温和和混沌机制，其去相关时间逐渐缩短。对于 Gray–Scott，这些机制通过增大快照步长实现，而非改变底层动力学。表 12 给出了模拟参数。

C.3 ERA5

我们使用逐小时的 ERA5 再分析（Hersbach 等人，2020年），在覆盖欧洲（ $0.25°$ 的网格上，25°西–45°东、 35°北–72°北、 149×281 个网格点），包含七个地表场：10米u、v风、2米气温

表 12：**生成参数**。按（24）–（27）的顺序排列的系统。区间表示轨迹长度，δ 以及保存快照间隔，即目标分辨率，二者单位均为模拟单位。未列出的参数采用 AutoSim 随附配置中的取值。

系统	参数	基准测试	Mild	混沌
平流-扩散	对流 μ	$[0.5, 2]$	$[1, 4]$	$[2, 8]$
	扩散率 ν	$[10^{-4}, 10^{-2}]$	$[5, 500] \cdot 10^{-5}$	$[2.5, 250] \cdot 10^{-5}$
	快照步长 δ		0.25	
	span		80	
Gray–Scott	快照步长 δ	4	8	16
	span	1280	2560	5120
纳维–斯托克斯	浮力 β	$[0.2, 0.8]$	$[1.6, 6.4]$	$[3.2, 12.8]$
	快照步长 δ		0.261	
	span		85.9	
Gross–Pitaevskii	障碍物频率 ω	$[0.5, 2]$	$[1.5, 3]$	–
	相互作用 g	$[80, 200]$	$[160, 400]$	–
	快照步长 δ		0.05	
	span		17.55	

以及露点、平均海平面气压、整层水汽含量和总降水量。数据按年份划分：1979–2015年用于训练，2016–2018年用于验证，2019–2022年用于测试。MUNDO 和基于流的基线方法以相邻窗口观测和日历信息（一天中的小时和一年中的第几天）为条件，并结合静态地形、陆海掩膜、纬度和经度场，设置与 ArchesWeather（Couairon等人，2026年）相同。评估使用2019–2022年间均匀分布的96个不重叠测试窗口，包含 $M = 24$ 个集合成员。

C.4 训练配置与基线方法

对于每个层级，单步 MUNDO、递归 MUNDO、条件流匹配和 FlowDPS 使用相同的骨干网络，并在可行的情况下使参数数量相匹配；SimVP 则保留其原始架构。插值和解聚分别训练独立模型。所有学习型分支均使用 AdamW（Loshchilov 与 Hutter，2019年）。AutoSim 和 ERA5 使用余弦衰减，并进行 500 步预热。Lorenz-96 和 AutoSim 模型采用 FP32 训练，ERA5 模型则采用 BF16 混合精度。MUNDO 的训练损失根据 $M = 2$ 次抽样估算。表 13 总结了训练配置和采样成本。对于 AutoSim，采用 $N = 32$ 处的 Gross–Pitaevskii 作为代表性设置。

ERA5 上的模型容量。在 ERA5 上，递归 MUNDO 使用四个 57.4M 参数的 ArchesWeather 网络，而单步 MUNDO 和各个基于流的基线方法则各使用一个网络。层级结构使这一模型容量变得经济可行：每个层级均独立训练，因此四个网络可以并行训练，成本与一个单步网络相当（总计 90.1 图形处理器小时，而 22.8，见表13）。将共享骨干网络扩展到与四个层级总容量相当的规模，无论用于单步 MUNDO 还是基于流的基线方法，都超出了我们的计算预算。因此，递归模型的比较无法将层级结构的效果与其增加的模型容量区分开来；不过，单步 MUNDO 提供了参数数量相匹配的比较。

在全部三个基准测试中，所有学习型基线都在采样后，使用公式（9）中的映射投影到约束集上。在 AutoSim 和 ERA5 上，FlowDPS 也在每个采样步骤中使用相同的投影作为其一致性更新。其先验使用相同的流匹配损失进行训练，且不会接触窗口观测；窗口观测仅通过每一步的这一投影发挥作用。在 Lorenz-96 上，FlowDPS 遵循Kim 等人（2025年）的梯度更新。条件流匹配是其条件版本，不包含一致性步骤。每个生成成员所需的网络评估次数见表 13。

表 13：**训练配置与计算成本**。ES 表示能量评分，VS 表示变差函数惩罚项，SC 表示谱 CRPS，NFE 表示每个生成成员的网络评估次数。训练成本以 H100 图形处理器小时计。预算在 Lorenz-96 上以轮次（ep.）计，在其他数据集上以优化器步数计。MUNDO 的训练预算按采样器层级报告；端到端微调则联合更新整个层级结构。[a]采样在八个 CPU核心上测量。

		模型			训练				Cost		
Tier	Arm	骨干网络	宽度	参数	目标函数	预算 批次		lr	图形处理器小时	神经函数评估	毫秒/成员
Lorenz-96 $F=32, N=64$	MUNDO ($R=6$)	循环卷积神经网络	256	6.18 M	ES	60 个周期	256	10^{-3}	0.10	6	1.6[a]
	MUNDO ($R=1$)	循环卷积神经网络	620	6.18 M	ES	60 个训练周期	256	10^{-3}	0.06	1	1.4[a]
	流匹配	循环卷积神经网络	618	6.18 M	流匹配	60 个训练周期	256	10^{-3}	0.06	56	2.6
	FlowDPS	循环卷积神经网络	618	618 万	流匹配	60 个训练周期	256	10^{-3}	0.05	56	2.7
Gross–Pitaevskii $N=32$	MUNDO ($R=5$)	NAT U-Net	37	12.7 M	能量评分+变差函数惩罚 → 能量评分+谱音转换	40k → 余	16	$3{\cdot}10^{-4} \to 10^{-4}$	3.1 + 0.6	5	1.3
	MUNDO ($R=1$)	NAT U-Net	85	12.6 M	能量评分+变差函数惩罚	40k	16	$3{\cdot}10^{-4}$	1.4	1	0.7
	流匹配	NAT U-Net	85	12.4 M	流匹配	40k	16	$3{\cdot}10^{-4}$	0.72	56	39
	FlowDPS	NAT U-Net	85	12.4 M	流匹配	40k	16	$3{\cdot}10^{-4}$	0.70	56	40
	SimVP	SimVP-gSTA	–	970 万	MSE	40k	16	$3{\cdot}10^{-4}$	0.36	1	2.4
ERA5 $N=24$	MUNDO($R=4$)	ArchesWeather	96	4×57.4M	能量评分+变差函数惩罚	40k	4	$3{\cdot}10^{-4}$	90.1	4	507
	MUNDO($R=1$)	ArchesWeather	96	57.4 M	能量评分+变差函数惩罚	40k	4	$3{\cdot}10^{-4}$	22.8	1	127
	流匹配	ArchesWeather	96	57.4 M	流匹配	40k	4	$3{\cdot}10^{-4}$	11.4	25	3150
	FlowDPS	ArchesWeather	96	57.4 M	流匹配	40k	4	$3{\cdot}10^{-4}$	11.3	25	3160
	SimVP	SimVP-gSTA	–	7.3 M	MSE	40k	4	$3{\cdot}10^{-4}$	2.5	1	6

C.5 指标

设 $x=(x_1,\dots,x_N)$ 为测试窗口的隐藏精细记录，$y=Ax$ 为其观测，$u^{(1)},\dots,u^{(M)}$ 为某种方法根据该观测生成的集合成员。除非通过窗口汇总定义，以下各项均按窗口计算，并在窗口、变量和像素上取平均。

准确度。连续排序概率得分（CRPS）采用公平形式进行估计（Ferro，2014），

$$\mathrm{CRPS}\left(u^{1:M},v\right)=\frac{1}{M}\sum_{m=1}^{M}|u^{(m)}-v|-\frac{1}{2M\left(M-1\right)}\sum_{m\neq m'}|u^{(m)}-u^{(m')}|, \tag{28}$$

在每个位置 i，将其应用于 x_i的每个变量。

校准。对于经过校准的集合，x_i在 $u_i^{(1)},\dots,u_i^{(M)}$ 中的秩服从$\{1,\dots,M+1\}$上的均匀分布（按位置和变量分别计算）。我们通过中心成员区间的实际覆盖率，以及其在一组名义水平上的平均绝对误差来概括其直方图（Greenbury 等，2026），

$$\widehat{c}(\alpha)=\left\langle \mathbf{1}\left\{\widehat{q}_{(1-\alpha)/2}\left(u_i^{1:M}\right)\le x_i\le\widehat{q}_{(1+\alpha)/2}\left(u_i^{1:M}\right)\right\}\right\rangle,\qquad \text{cov. MAE}=\frac{1}{|\mathcal{A}|}\sum_{\alpha\in\mathcal{A}}\left|\widehat{c}(\alpha)-\alpha\right|, \tag{29}$$

with $\mathcal{A}=\{0.1,0.2,\dots,0.9,0.95,0.99\}$ 和 $\widehat{q}$韦布尔绘图位置处的成员分位数。

我们还报告了经过有限集合校正的离散度-技巧比（Rasp 等人，2024年），

$$\mathrm{SSR}=\sqrt{\frac{M+1}{M}}\,\frac{\sqrt{\left\langle \mathrm{var}\left(u_i^{1:M}\right)\right\rangle}}{\sqrt{\left\langle\left(\overline{u}_i-x_i\right)^2\right\rangle}}, \tag{30}$$

其中，var 为无偏成员方差，$\langle\cdot\rangle$ 为对各位置和窗口求均值。值接近 1 表示集合离散度与集合均值误差一致。值低于 1 表示相对于该误差而言离散度不足，值高于 1 则表示离散度过大。

一致性。当重建能够精确复现粗尺度观测时，即认为其具有一致性：插值后的保留状态，以及解聚后施加的聚合量。一致性误差为

$$\epsilon_{\mathrm{cons}}=\begin{cases}\max\limits_{j,m}\left\|(A_{\mathrm{avg}}u^{(m)})_j-y_j\right\|_\infty, & \text{disaggregation},\\ \max\limits_{j,m}\left\|u_{i_j}^{(m)}-y_j\right\|_\infty, & \text{interpolation},\end{cases} \tag{31}$$

其中 i_j 表示第 j个存储状态的位置。当相应误差低于10^{-4} 个标准化单位时，该方法即通过。对于标记为 ✓$^{\pi}$的方法， ϵ_{cons} 是在附录 C.4 的投影之后测得的。

时间相干性。 我们用阶数为p的时间差分之比来衡量时间相干性

$$t_h = \frac{\langle |u_{t+h} - u_t|^p \rangle}{\langle |x_{t+h} - x_t|^p \rangle}, \qquad p = \tfrac{1}{2}, \tag{32}$$

其中，平均值是在集合成员、有效时间对、窗口、变量和空间位置上计算的。我们报告 t_1，其中 1 表示与真值一致。大于 1 的值表示短滞后增量大于参考值，而小于 1 的值表示增量较小。

空间相干性。 利用阶数为p 的成对差异 γ_h^a （14），我们用其比值衡量空间相干性

$$s_h = \frac{\langle \gamma_h^a(u) \rangle}{\langle \gamma_h^a(x) \rangle}, \qquad p = \tfrac{1}{2}, \tag{33}$$

其中，平均值是在集合成员、帧、窗口、变量和空间位置上计算的。我们报告 s_1，其中 1 表示与真值一致。大于 1 的值表示网格尺度上的场过于粗糙，而小于 1 的值表示场过于平滑。

精细尺度谱功率。 对于二维系统，我们还报告生成功率与真值功率在最精细已解析空间频带中的比值。在周期性 AutoSim 区域上，谱沿一个周期轴计算且不使用窗口，最精细频带是已解析波数范围的上半部分， $[K/2, K]$，其中 K 为奈奎斯特波数。报告的比值将该频带上的功率求和，并在集合成员、帧、窗口和正交空间轴上取平均。比值为 1 表示重建在最精细已解析尺度上携带了正确的功率。该指标对功率所在的位置不敏感，因此仅作为附录 B.2 中空间相干性消融的补充诊断指标。

跨窗口连续性。 窗口内指标无法检测连续重建结果交界处的伪影，因此我们评估拼接后的窗口链。令 $\mathcal{B}$ 表示跨越窗口边界的时间步，令 $\tilde{u}$ 和 $\tilde{x}$ 分别表示拼接后的重建结果和真值。我们定义接缝比率

$$r_{\text{seam}} = \frac{\left\langle |\tilde{u}_{t+1} - \tilde{u}_t| \right\rangle_{t \in \mathcal{B}}}{\left\langle |\tilde{x}_{t+1} - \tilde{x}_t| \right\rangle_{t \in \mathcal{B}}}, \tag{34}$$

其中，平均值是对边界时间步、链、变量和空间位置求得的。该值为 1 表示跨窗口时间步的幅度正确，但方向或数值未必正确，因此该指标可能因误差相互抵消而得到改善。

D 附加结果

D.1 洛伦兹-96

表 14 和 15 将主要的 Lorenz-96 结果扩展到其余强迫项以及长度为$N = 128$步的窗口。每个面板报告两项指标，用于衡量持续时间为 $T = N\delta$的窗口内动力学的去相关程度：比率 $N\delta/\tau$，其中 τ是慢变量的 e倍去相关时间；以及 λT，其中 λ 是最大李雅普诺夫指数，因此 λT 表示窗口内小扰动的 e倍衰减次数。两者均随 F 和 N增加，从非混沌的 $F = 10(\lambda = 0$增长至约十三次 e倍衰减，达到 $F = 64$， $N = 128$。

在解聚任务中，流匹配在六种混沌设置下的连续排序概率得分最低；对于 $F \in \{32, 64\}$，FlowDPS 的得分也低于递归 MUNDO，具体为 $N = 128$ 。在 $F = 10$ 下，递归 MUNDO 的连续排序概率得分是所有方法中最低的。在 $N = 64$ 下，流匹配在每种强迫项下的覆盖误差也更低。在 $N = 128$ 下，递归 MUNDO 的离散度-技巧比更接近 1，且在四种强迫项中的三种下覆盖误差更低；而流匹配则高估了 $F \leq 32$ 的短滞后变异性。与单步 MUNDO 相比，递归在八种设置中的七种降低了连续排序概率得分

在八种设置中的七种，并且在单步 MUNDO 的时间变差函数比低于 1 的情况下，仍使该比值保持接近 1，不过在其中六种设置下，单步 MUNDO 的覆盖误差更低。在插值任务中，递归仅在非混沌区间改善了连续排序概率得分。对于 $F \geq 24$，单步 MUNDO 在两种窗口长度下的连续排序概率得分均较低，而递归 MUNDO 在所有设置下的覆盖误差仍较低。流匹配在每种强迫项下都高估了短滞后变异性，并且在低强迫项下严重过度离散；而递归 MUNDO 在所有设置下都使 t_1 接近 1。经典基线方法满足约束，但连续排序概率得分仍远逊于其他方法；其 t_1 低于 1，反映出平滑重建结果缺少短滞后变异性。

图 8和 9展示了每种方法在 $F = 32$、 $N = 64$这四个连续窗口中的一个成员，图 3 则展示了正文中一个慢变量对应的集合。递归式和单步 MUNDO 成员在时间上平滑，与参考轨迹相似；而两种基于流的方法在相邻步之间都存在噪声，其中插值下的流匹配尤为明显，这正是 t_1 所衡量的过量短滞后变异性。FlowDPS 仅通过最终投影满足观测。

表 14：**Lorenz-96 在** $N = 64$, $F \in \{10, 24, 64\}$。 †：因果链。粗体表示最佳结果；下划线表示各指标中的次优结果。

(a) $F = 10 (N\delta/\tau = 2.1, \lambda T = 0.0$；非混沌)。

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 6$)†	✓	**0.0085**±0.0011	1.17	0.062	**1.00**
MUNDO ($R = 1$)	✓$^\pi$	0.0093±0.0008	1.30	0.088	**1.00**
流匹配	✓$^\pi$	0.0137±0.0008	**1.15**	**0.036**	1.12
FlowDPS	✓$^\pi$	0.0139±0.0011	0.79	0.068	1.07
Denton–Cholette	✓	0.3733	–	–	0.76
CDF-t	×	0.6146	–	–	0.72
插值					
MUNDO ($R = 6$)	✓	**0.0073**±0.0001	**1.15**	0.053	**1.00**
MUNDO ($R = 1$)	✓$^\pi$	0.0075±0.0002	1.28	0.062	**1.00**
流匹配	✓$^\pi$	0.0260±0.0029	1.89	0.162	1.60
FlowDPS	✓$^\pi$	0.0271±0.0008	1.19	**0.050**	1.47
线性	✓	0.4084	–	–	0.85
三次样条	✓	0.3686	–	–	0.86

(b) $F = 24 (N\delta/\tau = 3.9, \lambda T = 1.8)$。

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 6$)†	✓	0.0857±0.0009	0.61	0.106	**0.99**
MUNDO ($R = 1$)	✓$^\pi$	0.1201±0.0008	0.63	0.098	0.98
流匹配	✓$^\pi$	**0.0803**±0.0017	0.73	0.043	1.02
FlowDPS	✓$^\pi$	0.1085±0.0019	**0.88**	**0.024**	1.02
Denton–Cholette	✓	0.6434	–	–	0.42
CDF-t	×	0.7404	–	–	0.54
插值					
MUNDO ($R = 6$)	✓	0.0356±0.0015	0.89	**0.008**	**1.00**
MUNDO ($R = 1$)	✓$^\pi$	**0.0261**±0.0002	1.09	0.033	**1.00**
流匹配	✓$^\pi$	0.0517±0.0028	1.32	0.107	1.42
FlowDPS	✓$^\pi$	0.0502±0.0010	**0.97**	0.029	1.32
线性	✓	0.7302	–	–	0.56
三次样条	✓	0.7709	–	–	0.57

(c) $F = 64$ ($N\delta/\tau = 8.3, \lambda T = 6.5$)。

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 6$)†	✓	0.4848±0.0012	0.90	0.049	**0.97**
MUNDO ($R = 1$)	✓$^\pi$	0.4886±0.0015	**1.04**	**0.037**	0.80
流匹配	✓$^\pi$	**0.4623**±0.0011	0.88	0.041	0.93
FlowDPS	✓$^\pi$	0.4856±0.0021	0.86	0.052	0.89
Denton–Cholette	✓	0.7205	–	–	0.25
CDF-t	×	0.8554	–	–	0.38
插值					
MUNDO ($R = 6$)	✓	0.4006±0.0161	0.91	**0.022**	0.92
MUNDO ($R = 1$)	✓$^\pi$	0.2772±0.0019	0.88	0.054	**0.96**
流匹配	✓$^\pi$	**0.2747**±0.0054	**0.92**	0.049	1.20
FlowDPS	✓$^\pi$	0.2786±0.0022	0.82	0.028	1.09
线性	✓	0.8711	–	–	0.38
三次样条	✓	0.9124	–	–	0.39

表 15：**Lorenz-96 的** $N = 128$, $F \in \{10, 24, 32, 64\}$。 †：因果链。粗体：最佳；下划线：各项指标中的第二佳。

(a) $F = 10 (N\delta/\tau = 4.2, \lambda T = 0.0$; 非混沌)。

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 7$) †	✓	**0.0150**±0.0021	**1.24**	0.062	**1.00**
MUNDO ($R = 1$)	✓$^\pi$	0.0181±0.0033	1.36	**0.055**	**1.00**
流匹配	✓$^\pi$	0.0658±0.0583	6.52	0.246	2.63
FlowDPS	✓$^\pi$	0.0308±0.0031	0.70	0.112	1.31
Denton–Cholette	✓	0.8645	–	–	0.32
CDF-t	×	1.0879	–	–	0.54
插值					
MUNDO ($R = 7$)	✓	**0.0130**±0.0012	**1.12**	**0.043**	**1.00**
MUNDO ($R = 1$)	✓$^\pi$	0.0157±0.0064	1.33	0.067	**1.00**
流匹配	✓$^\pi$	0.1458±0.0029	10.25	0.380	4.25
FlowDPS	✓$^\pi$	0.0635±0.0016	3.68	0.300	2.71
线性	✓	0.9766	–	–	0.55
三次样条	✓	1.0873	–	–	0.56

(b) $F = 24 (N\delta/\tau = 7.9, \lambda T = 3.7)$。

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 7$)†	✓	0.3300±0.0032	0.81	0.064	**0.95**
MUNDO ($R = 1$)	✓$^\pi$	0.3596±0.0025	0.87	0.038	0.92
流匹配	✓$^\pi$	**0.3038**±0.0013	0.62	0.130	1.32
FlowDPS	✓$^\pi$	0.3353±0.0053	**0.99**	**0.009**	1.11
Denton–Cholette	✓	0.7487	–	–	0.24
CDF-t	×	0.8999	–	–	0.38
插值					
MUNDO ($R = 7$)	✓	0.2376±0.0027	**0.96**	**0.015**	0.97
MUNDO ($R = 1$)	✓$^\pi$	**0.1508**±0.0008	0.78	0.025	**1.00**
流匹配	✓$^\pi$	0.2205±0.0013	2.37	0.290	3.66
FlowDPS	✓$^\pi$	0.1562±0.0025	1.22	0.170	2.43
线性	✓	0.8976	–	–	0.39
三次样条	✓	0.9434	–	–	0.40

(c) $F = 32$ ($N\delta/\tau = 9.8, \lambda T = 6.2$).

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 7$) †	✓	0.4928±0.0015	0.86	0.061	0.93
MUNDO ($R = 1$)	✓$^\pi$	0.4865±0.0017	**1.00**	**0.025**	0.83
流匹配	✓$^\pi$	**0.4608**±0.0008	0.86	0.049	1.16
FlowDPS	✓$^\pi$	0.4862±0.0002	0.92	0.028	**1.01**
Denton–Cholette	✓	0.7488	–	–	0.22
CDF-t	×	0.9005	–	–	0.35
插值					
MUNDO ($R = 7$)	✓	0.3823±0.0044	0.96	**0.006**	0.94
MUNDO ($R = 1$)	✓$^\pi$	0.2991±0.0039	0.93	0.058	**0.96**
流匹配	✓$^\pi$	0.3129±0.0014	1.55	0.210	3.27
FlowDPS	✓$^\pi$	**0.2696**±0.0016	**1.02**	0.104	2.18
线性	✓	0.9140	–	–	0.34
三次样条	✓	0.9578	–	–	0.35

(d) $F = 64$ ($N\delta/\tau = 16.6, \lambda T = 13.1$).

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 7$)†	✓	0.5374±0.0002	**1.02**	**0.026**	1.06
MUNDO ($R = 1$)	✓$^\pi$	0.5525±0.0041	1.04	0.093	0.90
流匹配	✓$^\pi$	**0.5363**±0.0008	0.87	0.047	**0.98**
FlowDPS	✓$^\pi$	0.5371±0.0008	0.86	0.050	0.93
Denton–Cholette	✓	0.7666	–	–	0.15
CDF-t	×	0.9491	–	–	0.27
插值					
MUNDO ($R = 7$)	✓	0.5041±0.0043	0.90	0.036	**0.97**
MUNDO ($R = 1$)	✓$^\pi$	0.4782±0.0052	**0.98**	0.059	0.79
流匹配	✓$^\pi$	0.4539±0.0026	1.19	0.098	2.40
FlowDPS	✓$^\pi$	**0.4528**±0.0064	0.86	**0.021**	1.62
线性	✓	0.9538	–	–	0.27
三次样条	✓	1.0006	–	–	0.28

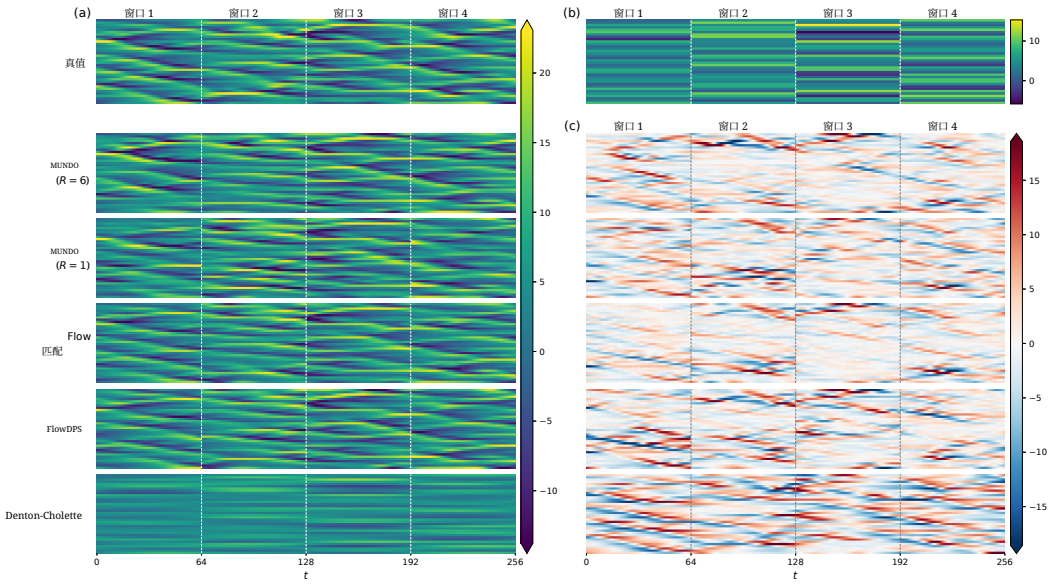


图 8：**Lorenz-96 在** $F = 32$**,** $N = 64$ 处的解聚，连续四个窗口，每种方法一个成员。(a) 真值和重建结果，慢变量随时间的变化。(b) 观测窗口均值。(c) 重建误差 $\hat{x} - x$；各行按 (a) 中的方法顺序排列。虚线标示窗口边界。

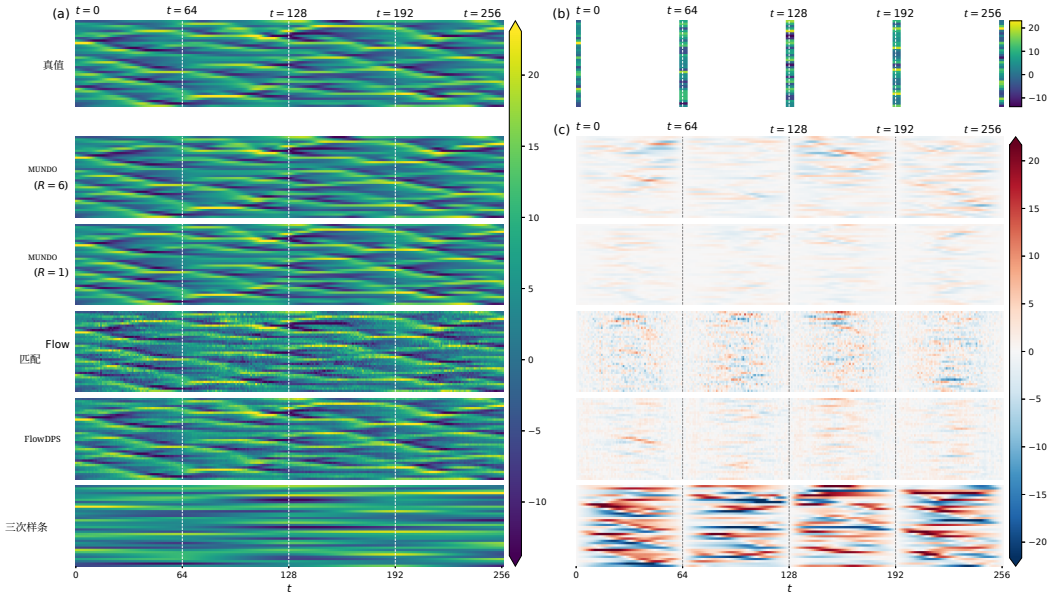


图 9：**Lorenz-96 在** $F = 32$**,** $N = 64$ 处的插值，连续四个窗口，每种方法一个成员。(a) 真值和重建结果，慢变量随时间的变化。(b) 为便于观察而扩展的观测端点状态。(c) 重建误差 $\hat{x} - x$；各行按 (a) 中的方法顺序排列。虚线标示观测端点。

D.2 AUTOSIM

表格 17–27 报告了其余 AutoSim 配置：基准层级为 $N \in \{8, 16, 24\}$，温和层级和混沌层级为 $N \in \{8, 16, 24, 32\}$。温和层级和混沌层级通过增强对流、浮力或障碍物强迫，或增大 Gray–Scott 的快照步长（表12），缩短动力学的去相关时间，因此固定长度的窗口涵盖了更多去相关时间。图注报告相应的 $N\delta/\tau$，其定义与 Lorenz-96 相同。在每张表中，递归式 MUNDO 变体（无论是否经过微调）均根据验证轨迹选定（附录 B.5）。Gross–Pitaevskii 没有混沌层级。图 10和 11展示了正文中未配图的两个系统的重建结果，并在表格行中加入了 FlowDPS 和 SimVP。

在插值任务中，单步 MUNDO 在大多数设置下取得最低的连续排序概率得分，而递归 MUNDO 通常取得较低的覆盖误差。在解聚任务中，单步 MUNDO 在大多数设置下仍保持较低的连续排序概率得分；在可预测性最低的情况下，递归会进一步降低该得分：在 $N \geq 16$ 的混沌平流扩散任务中降低 19–39%，在基准层级的 Gross–Pitaevskii 任务中降低 4–14%，在轻度级别的 Gross–Pitaevskii 任务中也降低 4–14%。递归模型是校准效果更好的采样器：在 44 种设置中的 37 种设置下，其离散度-技巧比更接近 1，在 23 种设置下覆盖误差更低；此外，当单步 MUNDO 的离散度出现问题时，递归模型能够予以纠正，具体而言，在基准层级和轻度级别的 Gray–Scott 任务中纠正过度离散，并在 $N \geq 16$ 的混沌平流扩散任务中纠正离散不足。在近确定性对照的平流-扩散任务中，流匹配在混沌层级 $N \geq 16$ 和轻度级别 $N \geq 24$ 的解聚任务中取得最低的连续排序概率得分，尽管其集合在这些情况下离散不足。递归 MUNDO 在纳维–斯托克斯插值任务中也略微高估了短滞后变异性，而单步 MUNDO 的 t_1 更接近 1。在基线方法中，FlowDPS 是插值任务中表现最弱的学习型方法。SimVP 满足观测，并在 Gray–Scott 和纳维–斯托克斯任务中匹配短时滞结构；但由于它是确定性的，其连续排序概率得分约为 MUNDO 的 1.2 至 10 倍。经典基线方法的平滑程度最高，Gross–Pitaevskii 任务尤其如此。

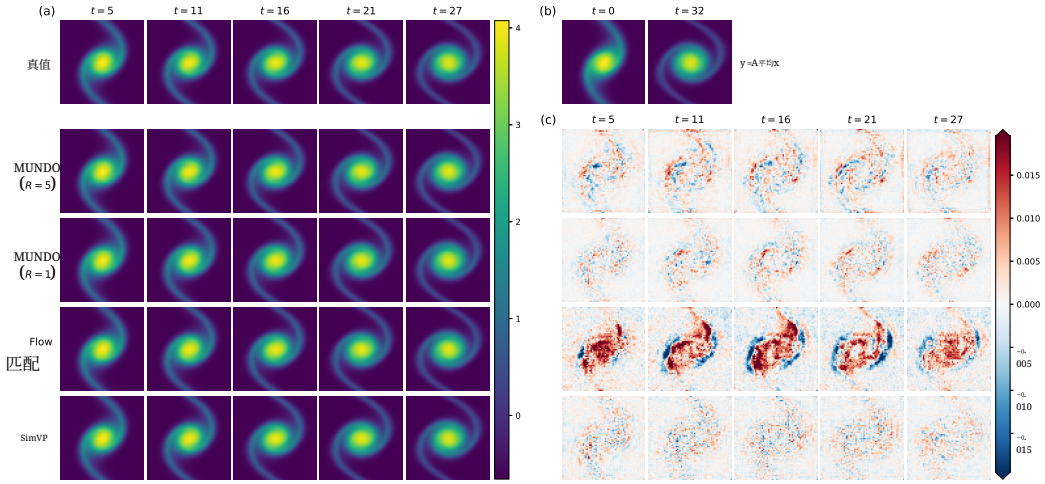


图 10：在 $N = 32$基准层级上的平流-扩散插值。(a) 所选时刻的真实值和重建结果。(b) 观测端点状态。(c) 所选时刻的重建误差 $\hat{x} - x$；各行按 (a) 中的方法顺序排列。

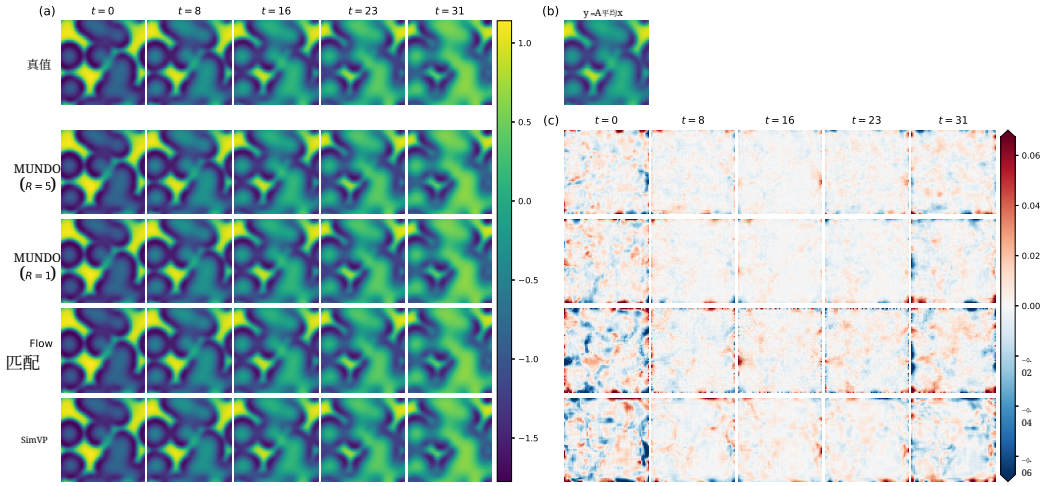


图11：**Gray–Scott 在** $N = 32$基准层级下的解聚。(a) 所选时刻的真实值和重建结果。(b) 粗尺度观测，即窗口内的时间平均。(c) 所选时刻的重建误差 $\hat{x} - x$；各行的方法顺序与 (a) 相同。

D.3 ERA5

D.3.1 降水的非负投影

表 16 对每种方法在方程（19）投影前后的降水通道进行评分。在解聚任务中，投影使递归和单步 MUNDO 的降水连续排序概率得分分别降低 5% 和 4%。对于递归模型，投影还使最精细频带的频带谱比从 1.56 降至 1.05，并使 t_1 从 1.39 降至 1.03；同时，SSR 从 0.95 降至 0.84，因为部分离散度来自负值。在插值任务中，将数值截断为零后，每种方法的降水连续排序概率得分变化均小于3%。该投影不会影响表 3中各方法的排名。

表 16：**ERA5 非负投影前后的总降水量**。Neg. 表示投影前生成的负像素小时所占比例。连续排序概率得分的单位为 mmh^{-1}；最精细频带表示最精细频带谱功率比。CDF-t 和线性插值不会产生负值，因此不作改动。

	负值（%）	连续排序概率得分 ↓		SSR → 1		最精细频带 → 1		$t_1 \to 1$	
		之前	之后	之前	之后	之前	之后	之前	之后
解聚									
MUNDO（$R = 4$）	27.9	0.1725	0.1634	0.95	0.84	1.56	1.05	1.39	1.03
MUNDO ($R = 1$)	23.9	0.1854	0.1774	0.88	0.84	0.77	0.75	1.44	1.15
流匹配	20.0	0.2324	0.2281	0.49	0.49	0.35	0.35	1.15	0.95
FlowDPS	20.4	0.2242	0.2204	0.58	0.56	0.42	0.41	1.09	0.90
SimVP	21.0	0.2456	0.2337	–	–	0.10	0.11	0.72	0.58
Denton–Cholette	16.6	0.3167	0.3007	–	–	0.11	0.11	0.44	0.35
插值									
MUNDO（$R = 4$）	20.1	0.1728	0.1715	0.88	0.87	0.92	0.91	1.26	1.16
MUNDO（$R = 1$）	21.1	0.1803	0.1770	0.70	0.70	0.47	0.46	1.50	1.36
流匹配	19.5	0.2354	0.2338	0.53	0.52	0.49	0.49	1.02	0.94
FlowDPS	23.1	0.2743	0.2726	0.35	0.35	0.07	0.07	0.95	0.85
SimVP	17.3	0.2517	0.2456	–	–	0.01	0.01	0.82	0.75
三次样条	20.5	0.3763	0.3666	–	–	0.50	0.50	0.43	0.38

D.3.2 各变量的逐小时评分

图 12 和 13 展示了按一天中的小时和变量划分、以物理单位表示的公平连续秩概率得分。在解聚条件下，表 3中的排序在几乎每个小时和变量上都成立；除降水外，误差在一天中变化很小，而降水的误差会在午后对流期间达到峰值。在插值条件下，观测到的状态处误差为零，而在窗口中间附近达到峰值；经典方法的插值器在此处表现最差。流匹配在此处与 MUNDO 的差距最大，尤其是在风和气压方面。

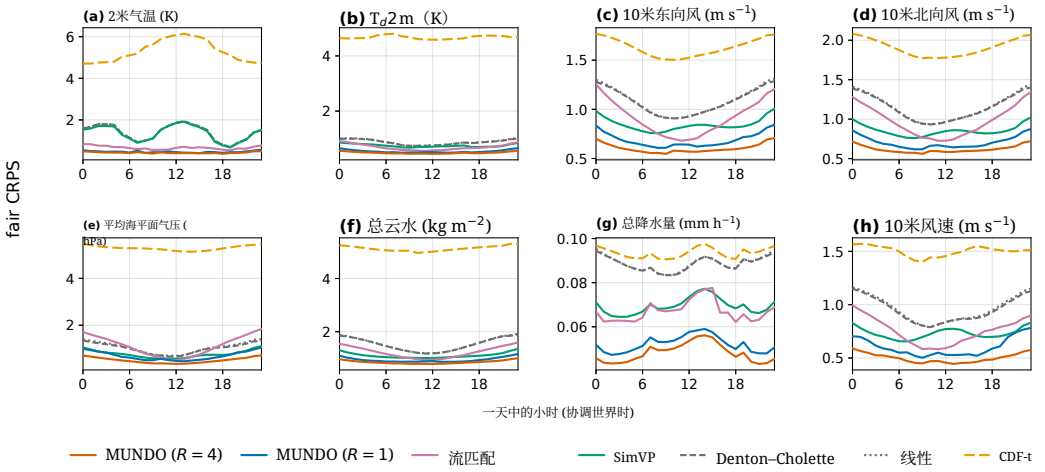


图12：**解聚条件下按一天中的小时划分的公平连续秩概率得分。** 各变量以物理单位表示，并在测试集窗口上取平均。风速（h）由两个风分量计算得出。

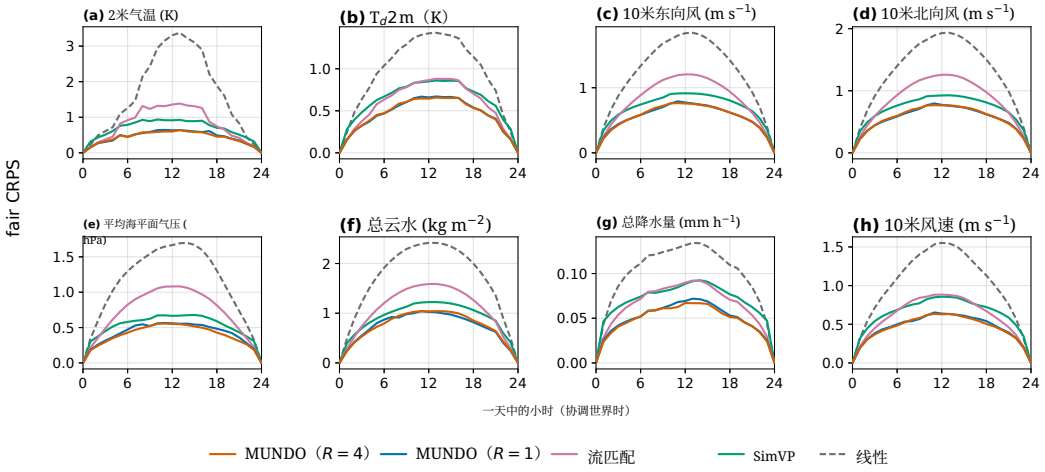


图13： **插值条件下按一天中的小时划分的公平连续秩概率得分。** 构造方式与图 12相同。

D.3.3 功率谱

空间功率谱衡量方差在不同空间尺度上的分布，因此，过度平滑或过量的小尺度噪声会直接表现为精细尺度功率的缺失或过量。图 14 和15 展示了相对于 ERA5 的功率谱比值。在解聚情况下，MUNDO 的两个变体在各变量上都保持接近 ERA5，而流匹配和 SimVP 则损失了精细尺度功率：风和露点的损失较为温和，降水的损失最为明显，其功率谱始终低于

参考值，且在所有尺度上均低于参考值。在插值情况下，SimVP 在各变量上都出现低估。对于降水，流匹配和 SimVP 在所有尺度上都低于 ERA5，而 MUNDO 在中等尺度上低于 ERA5，并在精细尺度上恢复到接近 ERA5 的水平。

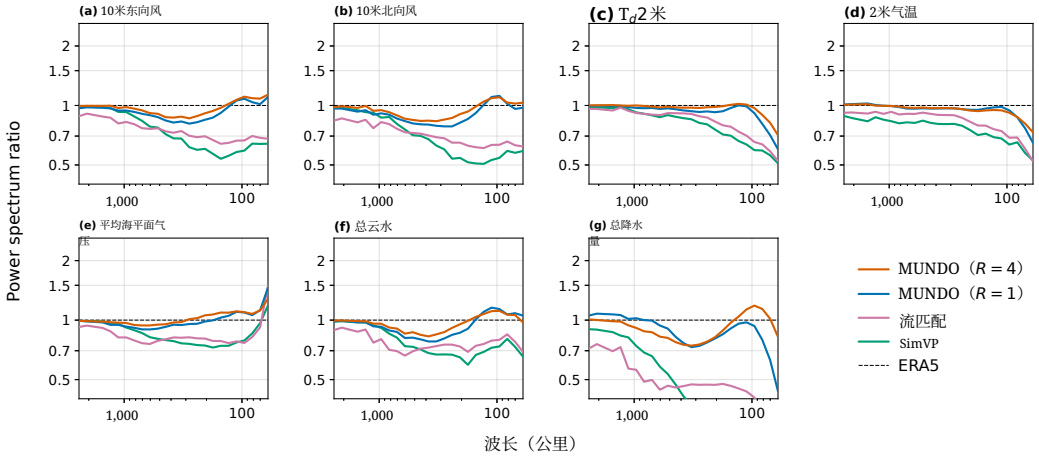


图14：**解聚下的空间功率谱比值**。各方法的功率谱密度与 ERA5 的功率谱密度之比，按变量分别计算，谱值为测试窗口和集合成员的平均值。虚线表示比值为一。

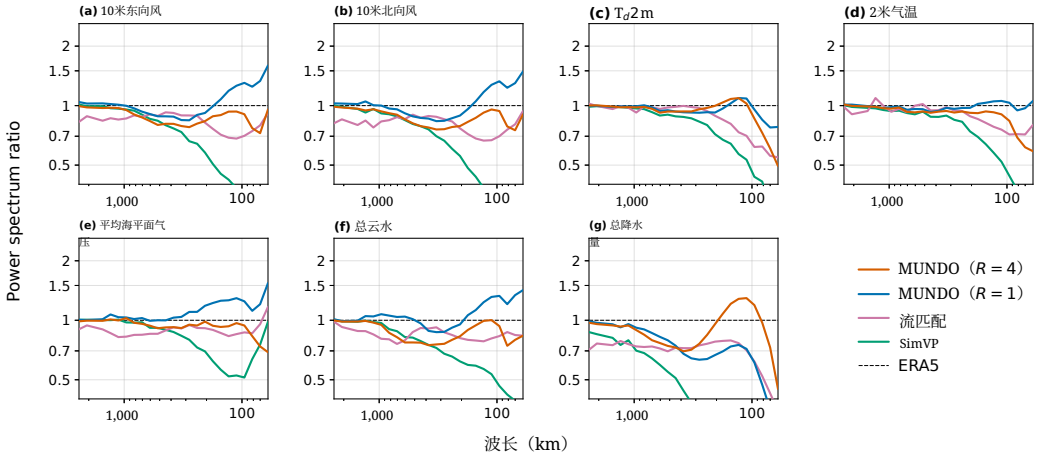


图15：**插值下的空间功率谱比值**。构造方式与图14相同。

D.3.4 日变化周期

日变化周期是地表温度的主导日内信号，在日均值中并不存在，因此恢复其振幅和相位可以检验生成的逐小时数据是否包含具有物理时间结构的信号，而非噪声。图 16 和 17 展示了大陆像元、海洋像元以及整个区域的平均 2米气温日变化周期，数值为相对于每天均值的距平。MUNDO 的两个变体在两项任务中都再现了均值周期的振幅和相位，包括几乎平坦的海洋性周期。

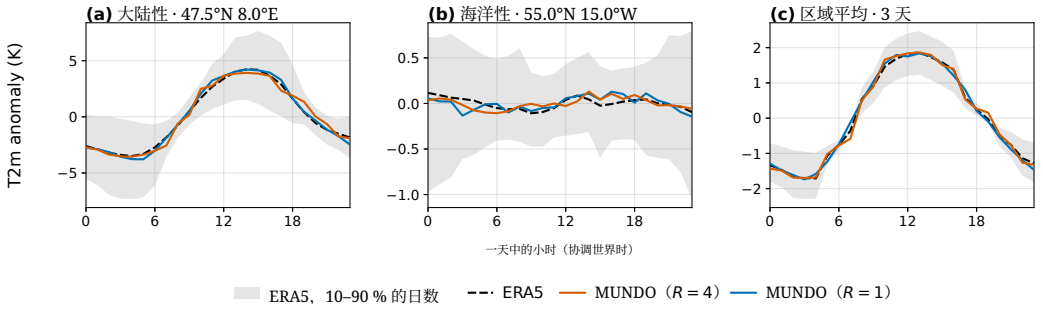
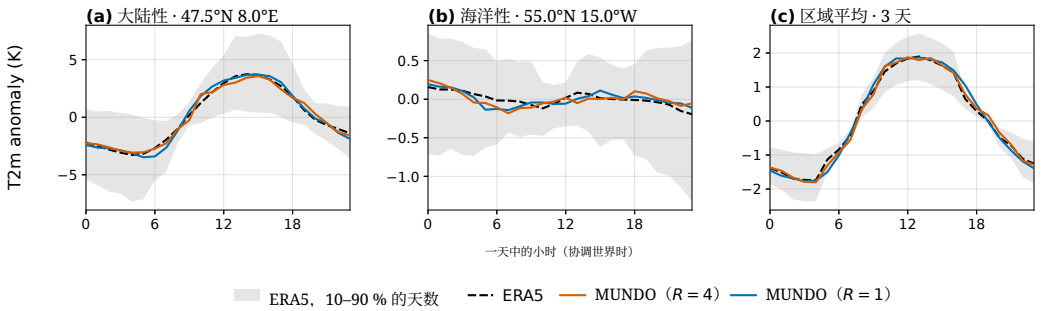


图16：**解聚条件下2米气温的日变化周期**。显示为相对于每天均值的距平，分别针对一个大陆像元、一个海洋像元，以及在整个区域上的平均值。曲线表示测试日的平均值，每天对应一个成员；阴影表示ERA5逐日变化周期的10–90%范围。



Fig图17：**插值条件下2米气温的日变化周期**。构造方式与图16**相同** .

D.4 FULLAUTOSIM 表格

在下方所有表格中，‡ 表示端到端微调模型；粗体条目为最佳结果，下划线条目为各指标的次佳结果。

表17： $N = 32$处各系统最混沌层级的 **AutoSim**。

(a)**平流-扩散**（$N\delta/\tau = 4.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 5$)‡	✓	0.0263±0.0008	**1.02**	**0.011**	0.94
MUNDO ($R = 1$)	✓π	0.0432±0.0063	0.69	0.038	0.71
流匹配	✓π	**0.0222**±0.0002	0.74	0.049	**0.95**
FlowDPS	✓π	0.0441±0.0018	0.54	0.159	0.77
SimVP	✓π	0.0331	–	–	0.78
Denton–Cholette	✓	0.0986	–	–	0.33
CDF-t	×	0.1196	–	–	0.43
插值					
MUNDO（$R = 5$）‡	✓	0.0266±0.0003	0.97	0.032	**1.00**
MUNDO（$R = 1$）	✓π	**0.0213**	**0.99**	0.047	0.87
流匹配	✓π	**0.0213**±0.0001	0.77	**0.025**	0.96
FlowDPS	✓π	0.0765±0.0027	0.67	0.092	0.76
SimVP	✓π	0.0353	–	–	0.81
线性	✓	0.1234	–	–	0.42
三次样条	✓	0.1281	–	–	0.43

(b)**Gray–Scott**（$N\delta/\tau = 3.2$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 5$）‡	✓	0.0199±0.0006	**1.03**	0.047	1.01
MUNDO（$R = 1$）	✓π	**0.0146**±0.0001	0.82	0.025	1.01
流匹配	✓π	0.0173±0.0003	0.89	0.082	1.02
FlowDPS	✓π	0.0467±0.0009	0.95	**0.015**	1.02
SimVP	✓π	0.0309	–	–	**1.00**
Denton–Cholette	✓	0.6247	–	–	0.56
CDF-t	×	0.6630	–	–	0.61
插值					
MUNDO ($R = 5$)‡	✓	0.0169±0.0001	**0.95**	0.040	1.02
MUNDO ($R = 1$)	✓π	**0.0112**±0.0002	0.77	**0.032**	1.01
流匹配	✓π	0.0142±0.0002	0.90	0.093	1.02
FlowDPS	✓π	0.2287±0.0095	0.90	0.052	0.97
SimVP	✓π	0.0214	–	–	**1.00**
线性	✓	0.6953	–	–	0.64
三次样条	✓	0.7244	–	–	0.68

(c)**纳维–斯托克斯 (**$N\delta/\tau = 1.5$**)**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 5$)‡	✓	0.1076±0.0008	**0.82**	**0.027**	1.24
MUNDO ($R = 1$)	✓π	0.0977±0.0038	0.76	**0.027**	**0.96**
流匹配	✓π	**0.0948**±0.0008	0.74	0.045	1.29
FlowDPS	✓π	0.1313±0.0018	0.49	0.081	1.06
SimVP	✓π	0.1352	–	–	0.94
Denton–Cholette	✓	0.2178	–	–	0.62
CDF-t	×	0.2831	–	–	0.68
插值					
MUNDO ($R = 5$)‡	✓	0.0663±0.0016	**0.79**	0.036	1.26
MUNDO ($R = 1$)	✓π	0.0650±0.0036	0.77	**0.028**	1.02
流匹配	✓π	**0.0600**±0.0002	0.71	0.051	1.19
FlowDPS	✓π	0.2544±0.0021	0.68	0.071	1.47
SimVP	✓π	0.0778	–	–	**1.00**
线性	✓	0.2610	–	–	0.69
三次样条	✓	0.2594	–	–	0.71

表 18： $N = 24$处的 **AutoSim**。

（a） **平流-扩散**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	0.0020±0.0000	**0.93**	0.114	**1.02**
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0016**±0.0000	0.84	0.158	1.05
流匹配	✓$^{\pi}$	0.0022±0.0000	0.81	0.184	1.19
FlowDPS	✓$^{\pi}$	0.0049±0.0001	0.48	**0.093**	1.20
SimVP	✓$^{\pi}$	0.0037	–	–	1.04
Denton–Cholette	✓	0.0274	–	–	0.93
CDF-t	×	0.0419	–	–	0.92
插值					
MUNDO（$R = 4$）	✓	0.0015±0.0000	0.99	0.099	1.09
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0010**±0.0000	**1.00**	0.172	**1.04**
流匹配	✓$^{\pi}$	0.0017±0.0000	0.87	0.245	1.17
FlowDPS	✓$^{\pi}$	0.0205±0.0013	0.65	**0.027**	1.22
SimVP	✓$^{\pi}$	0.0036	–	–	**1.04**
线性	✓	0.0338	–	–	0.92
三次样条	✓	0.0294	–	–	0.95

（b） **Gray–Scott**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 4$）	✓	**0.0028**±0.0000	1.10	**0.043**	**1.00**
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0028**±0.0000	1.35	0.048	1.01
流匹配	✓$^{\pi}$	0.0040±0.0000	**0.94**	0.128	1.06
FlowDPS	✓$^{\pi}$	0.0067±0.0000	0.79	0.089	1.07
SimVP	✓$^{\pi}$	0.0081	–	–	1.01
Denton–Cholette	✓	0.0805	–	–	0.92
CDF-t	×	0.1412	–	–	0.90
插值					
MUNDO ($R = 4$)	✓	0.0030±0.0001	1.04	**0.044**	1.02
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0022**±0.0000	**1.01**	0.059	1.01
流匹配	✓$^{\pi}$	0.0033±0.0000	**1.01**	0.141	1.06
FlowDPS	✓$^{\pi}$	0.0337±0.0005	1.15	0.067	1.09
SimVP	✓$^{\pi}$	0.0071	–	–	**1.00**
线性	✓	0.1191	–	–	0.93
三次样条	✓	0.0993	–	–	0.95

(c) **纳维–斯托克斯**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 4$）‡	✓	0.0088±0.0000	0.75	**0.020**	1.05
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0077**±0.0000	0.62	0.052	**1.02**
流匹配	✓$^{\pi}$	0.0116±0.0001	**0.85**	0.028	1.14
FlowDPS	✓$^{\pi}$	0.0160±0.0001	0.60	0.045	1.11
SimVP	✓$^{\pi}$	0.0200	–	–	**1.02**
Denton–Cholette	✓	0.0570	–	–	0.94
CDF-t	×	0.1509	–	–	0.95
插值					
MUNDO（$R = 4$）	✓	0.0074±0.0000	0.76	**0.017**	1.06
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0063**±0.0000	0.62	0.048	1.03
流匹配	✓$^{\pi}$	0.0092±0.0001	**0.78**	0.027	1.12
FlowDPS	✓$^{\pi}$	0.1420±0.0070	0.42	0.219	1.25
SimVP	✓$^{\pi}$	0.0139	–	–	**1.02**
线性	✓	0.0729	–	–	0.95
三次样条	✓	0.0583	–	–	0.97

(d) **Gross–Pitaevskii**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	**0.0799**±0.0010	**0.92**	0.065	1.06
MUNDO ($R = 1$)	✓$^{\pi}$	0.0828±0.0008	0.83	**0.057**	**0.98**
流匹配	✓$^{\pi}$	0.1001±0.0011	0.88	0.089	1.06
FlowDPS	✓$^{\pi}$	0.1402±0.0006	0.90	0.058	**0.98**
SimVP	✓$^{\pi}$	0.1536	–	–	0.91
Denton–Cholette	✓	0.4900	–	–	0.29
CDF-t	×	0.6140	–	–	0.45
插值					
MUNDO ($R = 4$)‡	✓	0.0253±0.0003	0.92	**0.055**	1.06
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0226**±0.0004	**0.95**	0.082	1.06
流匹配	✓$^{\pi}$	0.0305±0.0002	0.82	0.113	1.07
FlowDPS	✓$^{\pi}$	0.1385±0.0016	0.49	0.094	0.91
SimVP	✓$^{\pi}$	0.0388	–	–	**1.02**
线性	✓	0.5898	–	–	0.42
三次样条	✓	0.6169	–	–	0.43

表 19： $N = 16$处的**AutoSim**。

（a） **平流-扩散**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	0.0016±0.0000	**0.94**	**0.132**	**1.03**
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0012**±0.0000	0.92	0.175	1.05
流匹配	✓$^{\pi}$	0.0016±0.0000	0.79	0.199	1.16
FlowDPS	✓$^{\pi}$	0.0036±0.0001	0.47	**0.132**	1.17
SimVP	✓$^{\pi}$	0.0040	–	–	1.08
Denton–Cholette	✓	0.0127	–	–	0.94
CDF-t	×	0.0270	–	–	0.96
插值					
MUNDO ($R = 4$)	✓	0.0011±0.0000	**0.94**	0.117	1.08
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0006**±0.0000	**1.06**	0.199	1.04
流匹配	✓$^{\pi}$	0.0012±0.0000	0.91	0.247	1.15
FlowDPS	✓$^{\pi}$	0.0134±0.0008	0.64	**0.040**	1.18
SimVP	✓$^{\pi}$	0.0028	–	–	**1.03**
线性	✓	0.0176	–	–	0.95
三次样条	✓	0.0132	–	–	**0.97**

（b） **Gray–Scott**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	**0.0023**±0.0000	1.19	**0.053**	**1.00**
MUNDO ($R = 1$)	✓$^{\pi}$	0.0025±0.0000	1.26	0.057	1.01
流匹配	✓$^{\pi}$	0.0031±0.0001	**1.00**	0.138	1.05
FlowDPS	✓$^{\pi}$	0.0051±0.0001	0.75	0.106	1.06
SimVP	✓$^{\pi}$	0.0081	–	–	1.01
Denton–Cholette	✓	0.0536	–	–	0.89
CDF-t	×	0.1209	–	–	0.88
插值					
MUNDO ($R = 4$)	✓	0.0026±0.0000	1.10	**0.048**	1.02
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0020**±0.0000	1.03	0.064	1.01
流匹配	✓$^{\pi}$	0.0027±0.0000	**1.01**	0.139	1.04
FlowDPS	✓$^{\pi}$	0.0245±0.0003	1.18	0.090	1.08
SimVP	✓$^{\pi}$	0.0068	–	–	**1.00**
线性	✓	0.0920	–	–	0.95
三次样条	✓	0.0741	–	–	0.95

(c) **纳维–斯托克斯**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 4$）‡	✓	0.0043±0.0000	**0.85**	0.039	1.05
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0037**±0.0000	0.74	**0.022**	**1.03**
流匹配	✓$^{\pi}$	0.0056±0.0001	0.73	0.031	1.07
FlowDPS	✓$^{\pi}$	0.0076±0.0001	0.59	0.030	1.08
SimVP	✓$^{\pi}$	0.0127	–	–	**1.03**
Denton–Cholette	✓	0.0222	–	–	0.96
CDF-t	×	0.1264	–	–	1.00
插值					
MUNDO ($R = 4$)‡	✓	0.0043±0.0000	**0.79**	**0.022**	1.07
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0038**±0.0000	0.75	0.027	1.03
流匹配	✓$^{\pi}$	0.0051±0.0000	0.70	0.028	1.07
FlowDPS	✓$^{\pi}$	0.0627±0.0017	0.62	0.118	1.13
SimVP	✓$^{\pi}$	0.0092	–	–	1.02
线性	✓	0.0349	–	–	0.98
三次样条	✓	0.0224	–	–	**1.00**

（d） **Gross–Pitaevskii**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0509±0.0004	**0.96**	0.073	1.06
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0490**±0.0005	0.91	0.079	**1.03**
流匹配	✓$^{\pi}$	0.0577±0.0012	0.88	0.080	**1.03**
FlowDPS	✓$^{\pi}$	0.0898±0.0008	0.75	**0.049**	0.96
SimVP	✓$^{\pi}$	0.1023	–	–	0.95
Denton–Cholette	✓	0.4263	–	–	0.42
CDF-t	×	0.5121	–	–	0.56
插值					
MUNDO ($R = 4$)‡	✓	0.0149±0.0001	**0.94**	**0.060**	1.05
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0117**±0.0001	0.90	0.087	1.05
流匹配	✓$^{\pi}$	0.0151±0.0002	0.81	0.094	1.06
FlowDPS	✓$^{\pi}$	0.0781±0.0002	0.57	0.061	0.94
SimVP	✓$^{\pi}$	0.0242	–	–	**1.03**
线性	✓	0.5018	–	–	0.53
三次样条	✓	0.5229	–	–	0.55

表 20： $N = 8$处的**AutoSim**。

(a) **平流-扩散**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	**0.0004**±0.0000	0.89	**0.178**	**1.02**
MUNDO ($R = 1$)	✓π	0.0005±0.0000	**0.96**	0.219	1.05
流匹配	✓π	0.0009±0.0000	0.68	0.238	1.13
FlowDPS	✓π	0.0015±0.0000	0.53	0.192	1.14
SimVP	✓π	0.0027	–	–	1.05
Denton–Cholette	✓	0.0026	–	–	**0.98**
CDF-t	×	0.0162	–	–	1.03
插值					
MUNDO（$R = 3$）	✓	0.0005±0.0000	**1.02**	0.213	1.06
MUNDO（$R = 1$）	✓π	**0.0003**±0.0000	1.18	0.250	1.05
流匹配	✓π	0.0008±0.0000	0.78	0.224	1.11
FlowDPS	✓π	0.0078±0.0001	0.72	**0.117**	1.15
SimVP	✓π	0.0017	–	–	1.03
线性	✓	0.0057	–	–	0.98
三次样条	✓	0.0029	–	–	**0.99**

(b) **Gray–Scott**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	**0.0011**±0.0000	1.84	0.092	**1.00**
MUNDO ($R = 1$)	✓π	0.0012±0.0000	1.22	**0.091**	1.01
流匹配	✓π	0.0014±0.0000	**0.95**	0.183	1.03
FlowDPS	✓π	0.0019±0.0000	0.78	0.162	1.04
SimVP	✓π	0.0049	–	–	1.01
Denton–Cholette	✓	0.0204	–	–	0.95
CDF-t	×	0.0999	–	–	0.98
插值					
MUNDO（$R = 3$）	✓	0.0013±0.0000	1.12	**0.056**	1.01
MUNDO（$R = 1$）	✓π	**0.0011**±0.0000	1.12	0.122	1.01
流匹配	✓π	0.0014±0.0000	**0.99**	0.176	1.03
FlowDPS	✓π	0.0151±0.0001	1.15	0.104	1.05
SimVP	✓π	0.0047	–	–	**1.00**
线性	✓	0.0357	–	–	0.98
三次样条	✓	0.0275	–	–	0.99

(c) **纳维–斯托克斯**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)	✓	**0.0018**±0.0000	**0.93**	0.067	1.04
MUNDO ($R = 1$)	✓π	**0.0018**±0.0000	0.84	0.049	1.03
流匹配	✓π	0.0024±0.0000	0.69	0.050	1.05
FlowDPS	✓π	0.0026±0.0000	0.63	**0.043**	1.06
SimVP	✓π	0.0066	–	–	1.02
Denton–Cholette	✓	0.0041	–	–	**0.99**
CDF-t	×	0.1212	–	–	1.08
插值					
MUNDO（$R = 3$）	✓	0.0025±0.0000	0.90	0.046	1.05
MUNDO（$R = 1$）	✓π	**0.0023**±0.0000	0.88	0.069	1.03
流匹配	✓π	0.0026±0.0000	0.67	**0.033**	1.05
FlowDPS	✓π	0.0259±0.0001	**1.01**	0.041	1.05
SimVP	✓π	0.0053	–	–	1.02
线性	✓	0.0103	–	–	**1.00**
三次样条	✓	0.0043	–	–	1.01

(d) **Gross–Pitaevskii**

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	0.0169±0.0003	**0.96**	0.094	1.04
MUNDO ($R = 1$)	✓π	**0.0152**±0.0000	0.92	0.087	1.04
流匹配	✓π	0.0188±0.0002	0.88	0.089	1.04
FlowDPS	✓π	0.0407±0.0002	0.71	**0.048**	**1.01**
SimVP	✓π	0.0408	–	–	1.02
Denton–Cholette	✓	0.2263	–	–	0.69
CDF-t	×	0.2534	–	–	0.76
插值					
MUNDO ($R = 3$)‡	✓	0.0054±0.0001	**0.93**	0.079	1.04
MUNDO ($R = 1$)	✓π	**0.0041**±0.0000	0.88	0.084	1.04
流匹配	✓π	0.0057±0.0000	0.74	0.080	1.05
FlowDPS	✓π	0.0359±0.0001	0.81	**0.056**	**1.01**
SimVP	✓π	0.0135	–	–	1.03
线性	✓	0.2997	–	–	0.73
三次样条	✓	0.2851	–	–	0.77

表 21： $N = 32$ 处的 **AutoSim，各系统的轻度级别**。

(a)**平流-扩散**（$N\delta/\tau = 2.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 5$)	✓	0.0091±0.0001	**1.02**	**0.022**	1.02
MUNDO ($R = 1$)	✓π	0.0082±0.0001	0.54	0.049	0.98
流匹配	✓π	**0.0074**±0.0000	0.78	0.070	1.08
FlowDPS	✓π	0.0183±0.0006	0.42	0.054	**1.01**
SimVP	✓π	0.0127	–	–	0.96
Denton–Cholette	✓	0.0748	–	–	0.65
CDF-t	×	0.0828	–	–	0.69
插值					
MUNDO（$R = 5$）‡	✓	0.0079±0.0001	0.92	0.061	1.07
MUNDO（$R = 1$）	✓π	**0.0057**±0.0000	**0.95**	0.081	**1.00**
流匹配	✓π	0.0067±0.0000	0.78	0.109	1.08
FlowDPS	✓π	0.0440±0.0045	0.68	**0.037**	1.01
SimVP	✓π	0.0135	–	–	0.98
线性	✓	0.0884	–	–	0.68
三次样条	✓	0.0904	–	–	0.71

(b)**Gray–Scott**（$N\delta/\tau = 1.6$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 5$)	✓	0.0088±0.0002	0.79	0.043	1.01
MUNDO（$R = 1$）	✓π	**0.0066**±0.0001	1.17	**0.042**	1.01
流匹配	✓π	0.0088±0.0001	**0.91**	0.119	1.04
FlowDPS	✓π	0.0187±0.0003	1.10	0.049	1.04
SimVP	✓π	0.0147	–	–	**1.00**
Denton–Cholette	✓	0.2933	–	–	0.90
CDF-t	×	0.2885	–	–	0.86
插值					
MUNDO ($R = 5$)‡	✓	0.0061±0.0000	1.06	**0.048**	1.03
MUNDO（$R = 1$）	✓π	**0.0044**±0.0000	0.86	0.053	1.01
流匹配	✓π	0.0063±0.0001	**0.97**	0.130	1.04
FlowDPS	✓π	0.0692±0.0058	0.86	0.084	1.03
SimVP	✓π	0.0111	–	–	**1.00**
线性	✓	0.3509	–	–	0.84
三次样条	✓	0.3406	–	–	0.91

(c)**纳维–斯托克斯**（$N\delta/\tau = 1.3$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 5$）‡	✓	0.0757±0.0022	0.75	**0.021**	1.18
MUNDO（$R = 1$）	✓π	0.0718±0.0030	0.72	0.032	**0.99**
流匹配	✓π	**0.0703**±0.0005	**0.79**	0.051	1.29
FlowDPS	✓π	0.1049±0.0009	0.49	0.073	1.11
SimVP	✓π	0.1014	–	–	0.97
Denton–Cholette	✓	0.1989	–	–	0.69
CDF-t	×	0.2628	–	–	0.75
插值					
MUNDO ($R = 5$)‡	✓	0.0460±0.0010	0.68	0.023	1.16
MUNDO ($R = 1$)	✓π	**0.0359**±0.0008	0.63	**0.019**	1.02
流匹配	✓π	0.0447±0.0001	**0.73**	0.053	1.19
FlowDPS	✓π	0.2724±0.0166	0.65	0.098	1.56
SimVP	✓π	0.0583	–	–	**1.01**
线性	✓	0.2385	–	–	0.74
三次样条	✓	0.2366	–	–	0.76

(d)**Gross–Pitaevskii**（$N\delta/\tau = 4.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 5$)‡	✓	**0.1324**±0.0045	0.95	0.058	**1.02**
MUNDO ($R = 1$)	✓π	0.1532±0.0098	0.85	**0.052**	0.86
流匹配	✓π	0.1636±0.0003	0.93	0.127	1.33
FlowDPS	✓π	0.2612±0.0013	**0.97**	0.087	1.12
SimVP	✓π	0.2059	–	–	0.84
Denton–Cholette	✓	0.5340	–	–	0.16
CDF-t	×	0.6741	–	–	0.33
插值					
MUNDO ($R = 5$)‡	✓	**0.0583**±0.0020	0.93	**0.033**	1.07
MUNDO ($R = 1$)	✓π	0.0823±0.0110	1.06	0.084	**0.97**
流匹配	✓π	0.0950±0.0010	**0.95**	0.150	1.27
FlowDPS	✓π	0.2807±0.0068	0.52	0.088	1.21
SimVP	✓π	0.0791	–	–	0.95
线性	✓	0.6460	–	–	0.31
三次样条	✓	0.6766	–	–	0.32

表22： $N = 24$ 处各系统轻度级别的 **AutoSim**。

(a)**平流-扩散**（$N\delta/\tau = 1.5$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	0.0073±0.0001	**1.01**	**0.039**	1.02
MUNDO ($R = 1$)	✓$^{\pi}$	0.0063±0.0001	0.67	0.067	**1.00**
流匹配	✓$^{\pi}$	**0.0062**±0.0000	0.77	0.073	1.07
FlowDPS	✓$^{\pi}$	0.0162±0.0003	0.40	0.049	**1.00**
SimVP	✓$^{\pi}$	0.0114	–	–	0.98
Denton–Cholette	✓	0.0673	–	–	0.72
CDF-t	×	0.0739	–	–	0.73
插值					
MUNDO ($R = 4$)	✓	0.0058±0.0001	**1.03**	0.065	1.05
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0045**±0.0001	0.96	0.088	**1.01**
流匹配	✓$^{\pi}$	0.0054±0.0000	0.77	0.110	1.07
FlowDPS	✓$^{\pi}$	0.0359±0.0029	0.62	**0.032**	0.98
SimVP	✓$^{\pi}$	0.0119	–	–	**0.99**
线性	✓	0.0768	–	–	0.73
三次样条	✓	0.0778	–	–	0.77

(b)**Gray–Scott** ($N\delta/\tau = 1.2$)

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	0.0059±0.0001	1.12	**0.045**	**1.00**
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0056**±0.0001	1.56	0.046	1.01
流匹配	✓$^{\pi}$	0.0068±0.0001	0.91	0.102	1.03
FlowDPS	✓$^{\pi}$	0.0128±0.0003	**1.02**	0.053	1.03
SimVP	✓$^{\pi}$	0.0124	–	–	**1.00**
Denton–Cholette	✓	0.2044	–	–	0.95
CDF-t	×	0.2053	–	–	0.92
插值					
MUNDO ($R = 4$)	✓	0.0047±0.0001	**1.01**	**0.036**	1.01
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0035**±0.0000	0.93	0.057	1.01
流匹配	✓$^{\pi}$	0.0048±0.0000	0.96	0.119	1.03
FlowDPS	✓$^{\pi}$	0.0355±0.0007	1.11	0.100	1.04
SimVP	✓$^{\pi}$	0.0098	–	–	**1.00**
线性	✓	0.2643	–	–	0.88
三次样条	✓	0.2429	–	–	0.95

(c)**纳维–斯托克斯**（$N\delta/\tau = 1.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0508±0.0004	**0.74**	0.023	1.09
MUNDO ($R = 1$)	✓$^{\pi}$	0.0449±0.0004	0.70	0.032	**1.01**
流匹配	✓$^{\pi}$	**0.0436**±0.0005	0.71	**0.022**	1.08
FlowDPS	✓$^{\pi}$	0.0684±0.0002	0.57	0.081	1.02
SimVP	✓$^{\pi}$	0.0742	–	–	**0.99**
Denton–Cholette	✓	0.1731	–	–	0.76
CDF-t	×	0.2371	–	–	0.79
插值					
MUNDO ($R = 4$)‡	✓	0.0285±0.0004	**0.69**	0.024	1.08
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0238**±0.0001	0.63	**0.019**	**1.02**
流匹配	✓$^{\pi}$	0.0273±0.0001	0.65	0.025	1.08
FlowDPS	✓$^{\pi}$	0.2600±0.0071	0.55	0.151	1.37
SimVP	✓$^{\pi}$	0.0431	–	–	**1.02**
线性	✓	0.1974	–	–	0.80
三次样条	✓	0.1922	–	–	0.82

(d)**Gross–Pitaevskii**（$N\delta/\tau = 3.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 4$）‡	✓	**0.0939**±0.0014	0.92	0.063	**1.03**
MUNDO（$R = 1$）	✓$^{\pi}$	0.1000±0.0021	0.82	**0.050**	0.92
流匹配	✓$^{\pi}$	0.1156±0.0001	0.86	0.083	1.04
FlowDPS	✓$^{\pi}$	0.1839±0.0030	**1.00**	0.062	0.93
SimVP	✓$^{\pi}$	0.1631	–	–	0.87
Denton–Cholette	✓	0.5234	–	–	0.20
CDF-t	×	0.6548	–	–	0.40
插值					
MUNDO（$R = 4$）‡	✓	0.0343±0.0007	**0.97**	**0.042**	1.05
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0332**±0.0004	1.10	0.084	1.04
流匹配	✓$^{\pi}$	0.0476±0.0004	0.82	0.091	1.05
FlowDPS	✓$^{\pi}$	0.1743±0.0062	0.47	0.095	0.91
SimVP	✓$^{\pi}$	0.0529	–	–	**0.99**
线性	✓	0.6400	–	–	0.35
三次样条	✓	0.6683	–	–	0.37

表23： $N = 16$处的 **AutoSim**，各系统的轻度级别。

（a）**平流-扩散**（$N\delta/\tau = 1.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0056±0.0001	**0.92**	0.056	1.02
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0044**±0.0001	0.90	0.097	1.01
流匹配	✓$^{\pi}$	0.0047±0.0000	0.78	0.064	1.06
FlowDPS	✓$^{\pi}$	0.0126±0.0001	0.43	**0.042**	**1.00**
SimVP	✓$^{\pi}$	0.0110	–	–	1.02
Denton–Cholette	✓	0.0426	–	–	0.80
CDF-t	×	0.0523	–	–	0.82
插值					
MUNDO（$R = 4$）‡	✓	0.0038±0.0000	0.95	0.082	1.04
MUNDO（$R = 1$）	✓$^{\pi}$	**0.0028**±0.0000	**0.97**	0.108	1.02
流匹配	✓$^{\pi}$	0.0036±0.0000	0.79	0.124	1.06
FlowDPS	✓$^{\pi}$	0.0248±0.0012	0.55	**0.021**	0.97
SimVP	✓$^{\pi}$	0.0095	–	–	**1.01**
线性	✓	0.0513	–	–	0.80
三次样条	✓	0.0503	–	–	0.84

(b)**Gray–Scott**（$N\delta/\tau = 0.8$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	**0.0039**±0.0000	1.20	**0.042**	**1.00**
MUNDO（$R = 1$）	✓$^{\pi}$	0.0040±0.0001	1.36	0.047	**1.00**
流匹配	✓$^{\pi}$	0.0047±0.0001	**0.94**	0.094	1.02
FlowDPS	✓$^{\pi}$	0.0087±0.0001	0.89	0.061	1.02
SimVP	✓$^{\pi}$	0.0112	–	–	**1.00**
Denton–Cholette	✓	0.1227	–	–	0.93
CDF-t	×	0.1531	–	–	0.89
插值					
MUNDO ($R = 4$)	✓	0.0036±0.0000	1.11	**0.033**	1.01
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0029**±0.0000	0.98	0.055	**1.00**
流匹配	✓$^{\pi}$	0.0036±0.0000	**0.99**	0.106	1.02
FlowDPS	✓$^{\pi}$	0.0240±0.0007	1.18	0.106	1.03
SimVP	✓$^{\pi}$	0.0100	–	–	**1.00**
线性	✓	0.1871	–	–	0.91
三次样条	✓	0.1582	–	–	0.95

(c)**纳维–斯托克斯** ($N\delta/\tau = 0.6$)

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)	✓	0.0246±0.0004	**0.83**	0.040	1.04
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0197**±0.0002	0.71	**0.017**	**1.02**
流匹配	✓$^{\pi}$	0.0210±0.0003	0.73	0.018	1.04
FlowDPS	✓$^{\pi}$	0.0340±0.0003	0.57	0.062	1.03
SimVP	✓$^{\pi}$	0.0399	–	–	1.03
Denton–Cholette	✓	0.1092	–	–	0.84
CDF-t	×	0.1878	–	–	0.87
插值					
MUNDO ($R = 4$)‡	✓	0.0156±0.0002	**0.74**	0.035	1.06
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0130**±0.0002	0.67	0.023	**1.02**
流匹配	✓$^{\pi}$	0.0151±0.0001	0.64	**0.021**	1.04
FlowDPS	✓$^{\pi}$	0.1731±0.0051	0.44	0.173	1.24
SimVP	✓$^{\pi}$	0.0271	–	–	**1.02**
线性	✓	0.1317	–	–	0.87
三次样条	✓	0.1227	–	–	0.89

(d)**Gross–Pitaevskii**（$N\delta/\tau = 2.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO（$R = 4$）‡	✓	**0.0636**±0.0004	**0.95**	0.066	1.04
MUNDO（$R = 1$）	✓$^{\pi}$	0.0667±0.0004	0.78	**0.051**	0.97
流匹配	✓$^{\pi}$	0.0784±0.0006	0.85	0.063	**1.00**
FlowDPS	✓$^{\pi}$	0.1123±0.0009	0.71	0.054	0.91
SimVP	✓$^{\pi}$	0.1246	–	–	0.90
Denton–Cholette	✓	0.5085	–	–	0.28
CDF-t	×	0.6427	–	–	0.46
插值					
MUNDO ($R = 4$)‡	✓	0.0198±0.0000	0.96	**0.050**	1.04
MUNDO ($R = 1$)	✓$^{\pi}$	**0.0174**±0.0010	**0.97**	0.074	1.05
流匹配	✓$^{\pi}$	0.0219±0.0001	0.83	0.076	1.04
FlowDPS	✓$^{\pi}$	0.0938±0.0009	0.55	0.079	0.91
SimVP	✓$^{\pi}$	0.0330	–	–	**1.01**
线性	✓	0.6238	–	–	0.41
三次样条	✓	0.6552	–	–	0.43

表 24： $N = 8$ 时每个系统的温和级别 AutoSim。

(a)平流-扩散（$N\delta/\tau = 0.5$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	0.0031±0.0001	0.89	0.112	1.02
MUNDO ($R = 1$)	✓π	**0.0021**±0.0001	**0.99**	0.145	1.03
流匹配	✓π	0.0026±0.0000	0.78	0.108	1.05
FlowDPS	✓π	0.0062±0.0001	0.53	**0.068**	1.03
SimVP	✓π	0.0063	–	–	**1.01**
Denton–Cholette	✓	0.0169	–	–	0.90
CDF-t	×	0.0269	–	–	0.93
插值					
MUNDO ($R = 3$)‡	✓	0.0016±0.0000	0.97	0.142	1.03
MUNDO ($R = 1$)	✓π	**0.0011**±0.0000	**0.98**	0.164	**1.02**
流匹配	✓π	0.0017±0.0000	0.80	0.134	1.05
FlowDPS	✓π	0.0182±0.0025	0.57	**0.027**	0.97
SimVP	✓π	0.0064	–	–	**1.02**
线性	✓	0.0236	–	–	0.90
三次样条	✓	0.0206	–	–	0.93

(b)Gray–Scott（$N\delta/\tau = 0.4$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	**0.0018**±0.0000	1.89	0.063	**1.00**
MUNDO ($R = 1$)	✓π	0.0019±0.0001	1.43	**0.059**	**1.00**
流匹配	✓π	0.0020±0.0000	**1.01**	0.096	1.01
FlowDPS	✓π	0.0031±0.0000	0.79	0.065	1.02
SimVP	✓π	0.0068	–	–	**1.00**
Denton–Cholette	✓	0.0374	–	–	0.94
CDF-t	×	0.0906	–	–	0.96
插值					
MUNDO ($R = 3$)	✓	0.0018±0.0000	1.23	**0.036**	**1.00**
MUNDO ($R = 1$)	✓π	**0.0016**±0.0001	1.16	0.073	**1.00**
流匹配	✓π	0.0019±0.0000	**1.02**	0.089	1.01
FlowDPS	✓π	0.0147±0.0000	1.29	0.114	1.02
SimVP	✓π	0.0067	–	–	**1.00**
线性	✓	0.0721	–	–	0.97
三次样条	✓	0.0545	–	–	0.98

(c) 纳维–斯托克斯（$N\delta/\tau = 0.3$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)	✓	0.0064±0.0002	**0.89**	0.071	1.03
MUNDO ($R = 1$)	✓π	**0.0053**±0.0000	0.77	0.041	**1.02**
流匹配	✓π	0.0065±0.0000	0.73	**0.025**	1.03
FlowDPS	✓π	0.0098±0.0001	0.55	0.032	1.04
SimVP	✓π	0.0159	–	–	**1.02**
Denton–Cholette	✓	0.0330	–	–	0.95
CDF-t	×	0.1370	–	–	1.00
插值					
MUNDO ($R = 3$)‡	✓	0.0053±0.0000	0.78	0.041	1.04
MUNDO ($R = 1$)	✓π	**0.0047**±0.0000	0.74	0.042	1.03
流匹配	✓π	0.0057±0.0000	0.66	0.024	1.03
FlowDPS	✓π	0.0345±0.0001	**0.81**	**0.022**	1.03
SimVP	✓π	0.0123	–	–	**1.02**
线性	✓	0.0487	–	–	0.96
三次样条	✓	0.0378	–	–	**0.98**

(d)Gross–Pitaevskii（$N\delta/\tau = 1.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 3$)‡	✓	0.0216±0.0000	**0.98**	0.083	1.04
MUNDO ($R = 1$)	✓π	**0.0193**±0.0002	0.97	0.077	1.03
流匹配	✓π	0.0261±0.0003	0.88	0.073	**1.02**
FlowDPS	✓π	0.0535±0.0009	0.59	**0.059**	0.96
SimVP	✓π	0.0546	–	–	**0.98**
Denton–Cholette	✓	0.3224	–	–	0.61
CDF-t	×	0.3581	–	–	0.70
插值					
MUNDO ($R = 3$)‡	✓	0.0070±0.0002	**0.95**	0.070	1.03
MUNDO ($R = 1$)	✓π	**0.0052**±0.0001	0.93	0.072	1.03
流匹配	✓π	0.0068±0.0001	0.75	0.060	1.04
FlowDPS	✓π	0.0467±0.0006	0.67	**0.049**	0.96
SimVP	✓π	0.0180	–	–	**1.02**
线性	✓	0.4030	–	–	0.68
三次样条	✓	0.4003	–	–	0.72

表25： $N = 24$ 下 AutoSim 在各系统最混沌层级的表现。

(a)平流-扩散（$N\delta/\tau = 3.0$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0228±0.0003	**0.99**	**0.011**	**0.96**
MUNDO ($R = 1$)	✓π	0.0349±0.0015	0.58	0.040	0.77
流匹配	✓π	**0.0205**±0.0001	0.78	0.046	**0.96**
FlowDPS	✓π	0.0431±0.0004	0.45	0.173	0.76
SimVP	✓π	0.0314	–	–	0.82
Denton–Cholette	✓	0.0984	–	–	0.41
CDF-t	×	0.1123	–	–	0.49
插值					
MUNDO ($R = 4$)‡	✓	0.0207±0.0001	0.95	0.036	**0.99**
MUNDO ($R = 1$)	✓π	**0.0185**±0.0001	**0.99**	0.048	0.90
流匹配	✓π	0.0190±0.0000	0.81	**0.025**	0.97
FlowDPS	✓π	0.0763±0.0034	0.59	0.118	0.71
SimVP	✓π	0.0335	–	–	0.85
线性	✓	0.1203	–	–	0.47
三次样条	✓	0.1242	–	–	0.49

(b)Gray–Scott（$N\delta/\tau = 2.4$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0126±0.0001	0.79	0.054	1.01
MUNDO ($R = 1$)	✓π	**0.0100**±0.0000	0.76	0.031	**1.00**
流匹配	✓π	0.0118±0.0002	0.81	0.077	1.01
FlowDPS	✓π	0.0303±0.0003	**1.04**	**0.018**	1.01
SimVP	✓π	0.0221	–	–	**1.00**
Denton–Cholette	✓	0.5010	–	–	0.73
CDF-t	×	0.5081	–	–	0.74
插值					
MUNDO ($R = 4$)‡	✓	0.0100±0.0001	**0.95**	0.041	1.01
MUNDO ($R = 1$)	✓π	**0.0074**±0.0001	0.72	0.041	**1.00**
流匹配	✓π	0.0092±0.0001	0.87	0.086	1.01
FlowDPS	✓π	0.1219±0.0094	0.82	**0.018**	0.99
SimVP	✓π	0.0162	–	–	**1.00**
线性	✓	0.5477	–	–	0.76
三次样条	✓	0.5583	–	–	0.81

(c) 纳维–斯托克斯（$N\delta/\tau = 1.1$）

	一致性	连续排序概率得分 ↓	SSR → 1 覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚					
MUNDO ($R = 4$)‡	✓	0.0740±0.0008	0.77	0.027	1.13
MUNDO ($R = 1$)	✓π	0.0715±0.0028	**0.79**	**0.016**	**0.99**
流匹配	✓π	**0.0618**±0.0001	0.71	0.021	1.06
FlowDPS	✓π	0.0917±0.0005	0.58	0.087	0.98
SimVP	✓π	0.1002	–	–	0.96
Denton–Cholette	✓	0.1877	–	–	0.70
CDF-t	×	0.2497	–	–	0.75
插值					
MUNDO ($R = 4$)‡	✓	0.0427±0.0003	**0.73**	0.033	1.12
MUNDO ($R = 1$)	✓π	**0.0370**±0.0006	0.70	**0.022**	1.03
流匹配	✓π	0.0382±0.0002	0.64	0.025	1.07
FlowDPS	✓π	0.2572±0.0160	0.60	0.122	1.33
SimVP	✓π	0.0581	–	–	**1.01**
线性	✓	0.2210	–	–	0.75
三次样条	✓	0.2182	–	–	0.77

表 26： $N = 16$ 下各系统最混沌层级的 **AutoSim**。

(a)**平流-扩散**（$N\delta/\tau = 2.0$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 4$)‡	✓	0.0186±0.0001	**0.95**	**0.012**	0.95
MUNDO ($R = 1$)	✓π	0.0229±0.0035	0.66	0.020	0.86
流匹配	✓π	**0.0159**±0.0001	0.79	0.052	**0.97**
FlowDPS	✓π	0.0338±0.0004	0.42	0.160	0.79
SimVP	✓π	0.0292	–	–	0.85
Denton–Cholette	✓	0.0794	–	–	0.51
CDF-t	×	0.0911	–	–	0.58
插值						
MUNDO ($R = 4$)‡	✓	0.0153±0.0002	0.94	0.046	**1.01**
MUNDO ($R = 1$)	✓π	**0.0130**±0.0003	**0.97**	0.054	0.95
流匹配	✓π	0.0135±0.0000	0.82	**0.025**	**0.99**
FlowDPS	✓π	0.0532±0.0004	0.49	0.133	0.72
SimVP	✓π	0.0282	–	–	0.89
线性	✓	0.0981	–	–	0.55
三次样条	✓	0.1013	–	–	0.58

(b)**Gray–Scott**（$N\delta/\tau = 1.6$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 4$)‡	✓	0.0087±0.0000	**0.96**	0.053	**1.00**
MUNDO ($R = 1$)	✓π	**0.0070**±0.0002	0.83	0.038	**1.00**
流匹配	✓π	0.0080±0.0001	0.78	0.060	1.01
FlowDPS	✓π	0.0173±0.0002	**1.04**	**0.027**	1.01
SimVP	✓π	0.0168	–	–	**1.00**
Denton–Cholette	✓	0.3029	–	–	0.88
CDF-t	×	0.3004	–	–	0.86
插值						
MUNDO ($R = 4$) ‡	✓	0.0062±0.0001	**0.97**	**0.037**	1.01
MUNDO ($R = 1$)	✓π	**0.0050**±0.0001	0.74	0.045	**1.00**
流匹配	✓π	0.0055±0.0000	0.83	0.068	1.01
FlowDPS	✓π	0.0639±0.0035	0.54	0.076	0.98
SimVP	✓π	0.0123	–	–	**1.00**
线性	✓	0.3826	–	–	0.85
三次样条	✓	0.3576	–	–	0.91

(c)**纳维–斯托克斯**（$N\delta/\tau = 0.8$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 4$)‡	✓	0.0402±0.0007	**0.78**	0.048	1.09
MUNDO ($R = 1$)	✓π	0.0338±0.0001	0.68	0.018	1.02
流匹配	✓π	**0.0330**±0.0002	0.71	**0.017**	1.03
FlowDPS	✓π	0.0507±0.0005	0.59	0.066	**1.00**
SimVP	✓π	0.0594	–	–	1.02
Denton–Cholette	✓	0.1341	–	–	0.78
CDF-t	×	0.2058	–	–	0.82
插值						
MUNDO ($R = 4$)‡	✓	0.0239±0.0002	**0.72**	0.041	1.08
MUNDO ($R = 1$)	✓π	**0.0198**±0.0001	0.66	0.029	1.03
流匹配	✓π	0.0215±0.0002	0.62	**0.021**	1.04
FlowDPS	✓π	0.1936±0.0142	0.48	0.161	1.24
SimVP	✓π	0.0371	–	–	**1.02**
线性	✓	0.1607	–	–	0.82
三次样条	✓	0.1532	–	–	0.85

表27： $N = 8$ 下的 **AutoSim**，各系统的混沌层级。

（a）**平流-扩散**（$N\delta/\tau = 1.0$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 3$)‡	✓	0.0121±0.0002	**0.94**	**0.041**	**1.00**
MUNDO ($R = 1$)	✓π	**0.0098**±0.0002	0.87	0.053	0.98
流匹配	✓π	0.0099±0.0000	0.80	0.054	0.98
FlowDPS	✓π	0.0205±0.0003	0.51	0.109	0.87
SimVP	✓π	0.0213	–	–	0.88
Denton–Cholette	✓	0.0504	–	–	0.66
CDF-t	×	0.0579	–	–	0.72
插值						
MUNDO ($R = 3$)‡	✓	0.0079±0.0002	0.94	0.072	1.01
MUNDO ($R = 1$)	✓π	**0.0059**±0.0002	**0.98**	0.090	**1.00**
流匹配	✓π	0.0062±0.0001	0.81	**0.032**	**1.00**
FlowDPS	✓π	0.0401±0.0007	0.45	0.119	0.72
SimVP	✓π	0.0200	–	–	0.95
线性	✓	0.0646	–	–	0.69
三次样条	✓	0.0654	–	–	0.73

(b)**Gray–Scott**（$N\delta/\tau = 0.8$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 3$)‡	✓	0.0028±0.0000	1.48	0.046	**1.00**
MUNDO ($R = 1$)	✓π	**0.0027**±0.0000	1.35	0.043	**1.00**
流匹配	✓π	0.0028±0.0000	**0.92**	0.041	**1.00**
FlowDPS	✓π	0.0051±0.0001	0.85	**0.021**	**1.00**
SimVP	✓π	0.0081	–	–	**1.00**
Denton–Cholette	✓	0.0938	–	–	0.98
CDF-t	×	0.1184	–	–	0.97
插值						
MUNDO ($R = 3$)	✓	0.0024±0.0000	1.06	**0.032**	**1.00**
MUNDO ($R = 1$)	✓π	**0.0022**±0.0000	**1.05**	0.058	**1.00**
流匹配	✓π	0.0024±0.0000	0.92	0.048	**1.00**
FlowDPS	✓π	0.0158±0.0001	1.31	0.117	1.01
SimVP	✓π	0.0074	–	–	**1.00**
线性	✓	0.1500	–	–	0.95
三次样条	✓	0.1183	–	–	**1.00**

(c)**纳维–斯托克斯**（$N\delta/\tau = 0.4$）

	一致性	连续排序概率得分 ↓	SSR → 1	覆盖率	平均绝对误差 ↓	$t_1 \to 1$
解聚						
MUNDO ($R = 3$)‡	✓	0.0107±0.0002	**0.80**	0.079	1.04
MUNDO ($R = 1$)	✓π	**0.0088**±0.0001	0.72	0.042	1.03
流匹配	✓π	0.0099±0.0001	0.69	**0.019**	1.03
FlowDPS	✓π	0.0154±0.0002	0.52	0.046	1.03
SimVP	✓π	0.0224	–	–	**1.02**
Denton–Cholette	✓	0.0494	–	–	0.92
CDF-t	×	0.1449	–	–	0.98
插值						
MUNDO ($R = 3$)‡	✓	0.0076±0.0001	**0.76**	0.051	1.04
MUNDO ($R = 1$)	✓π	**0.0067**±0.0001	0.71	0.048	1.03
流匹配	✓π	0.0076±0.0000	0.64	0.021	1.03
FlowDPS	✓π	0.0404±0.0002	0.72	**0.014**	1.03
SimVP	✓π	0.0163	–	–	**1.02**
线性	✓	0.0661	–	–	0.93
三次样条	✓	0.0565	–	–	0.96