免责声明：使用美国数学会官方 AMS LaTeX 模板 v6.1 的双栏版式生成。本作品尚未经过同行评审，由投稿作者提供，旨在以非商业方式及时传播学术和技术成果。作品的版权及其所有相关权利均由作者或其他版权所有者保留。所有复制此信息的人员均应遵守各作者版权所规定的条款和限制。未经版权所有者明确许可，不得转载本作品。本作品的版权可能在不另行通知的情况下转让，届时本版本可能无法继续访问。

arXiv:2305.01822v1 [cs.LG] 2 May 2023

使用扩散桥进行流体流动的非配对降尺度

Tobias Bischoff &KatherineDeck

气候建模联盟，加州理工学院，美国加利福尼亚州帕萨迪纳，邮编 *91125*

摘要：我们提出一种方法，使用基于扩散映射的生成模型对理想化的地球物理流体模拟进行降尺度。通过分析取自不同数据分布的图像的傅里叶谱，我们展示了如何将两个独立的条件扩散模型串联起来，用于域转换。由此得到的变换是低分辨率数据集与高分辨率数据集之间的扩散桥，使人们能够根据特定的低分辨率特征生成新的高分辨率图像样本。生成新样本的能力使我们无需额外校准或训练，即可计算任何感兴趣的统计量。我们的无监督设置还旨在无需配对训练数据即可对图像进行降尺度；这种灵活性使我们能够在无需额外训练的情况下组合多个源域和目标域。我们展示了该方法能够提高分辨率并校正地球物理流体模拟中依赖于上下文的偏差，包括极端事件中的偏差。我们预计，同样的方法可用于对气候模拟的输出进行降尺度，包括温度场和降水场，无需针对每种应用训练新模型，并可显著节省计算成本。

研究意义：本研究旨在运用生成式机器学习技术的最新进展，以较低成本获取更高分辨率的地球物理流体动力学模型输出，相比直接模拟，同时保留高分辨率数据的重要统计特性。这一点十分重要，因为许多应用都需要高分辨率气候模型输出，而获取这类输出的计算成本也很高。

1. 引言

气候模拟是预测和分析气候变化情景的有力工具，但往往受计算资源限制，因此空间和时间分辨率也受到制约。因此，模拟既可能缺少许多应用所需的高分辨率细节，也可能因未能表示会反馈影响较大尺度的小尺度动力过程而存在固有偏差。例如，O（10 – 100 公里）的水平分辨率仍不足以准确模拟对流降水、热带气旋动力学，以及地形和土地覆盖的局地影响等重要现象，因此这类模拟在区域和次区域尺度上进行预测时用途有限。尤其是，低分辨率场在极端气温和降水率方面存在偏差，进而可能降低较小空间尺度气候灾害预测的准确性，例如 Abatzoglou 和 Brown（2012）；Gutmann 等（2014）；Hwang 和 Graham（2014）。

人们已开发出多种方法，以纠正流体和气候模拟中的偏差并提高分辨率；这一过程称为流体流动的“降尺度”（Fowler 等，2007；Salathé Jr 等，2007；Maurer 和 Hidalgo，2008）。引导技术通过约束动力系统的解，使其遵循大尺度信息，并在模拟过程中应用，属于动力降尺度。另一方面，统计降尺度指的是使用数据驱动模型，在流体模拟运行完成后对其结果或计算所得统计量进行校正的方法；这类方法的计算开销通常较低。通常，这些模型的训练成本可通过后处理中进行多次评估来分摊；相比之下，每当需要输出结果时，都要对高分辨率流体系统求解（但无需训练时间）。

气候变量偏差校正和分辨率增强的主流统计方法是偏差校正空间解聚（BCSD）算法（Panofsky 等，1958；Wood 等，2002、2004；Thrasher 等，2012）。BCSD 使用分位数映射来校正偏差，并利用傅里叶变换提高分辨率。然而，BCSD 算法存在局限性，包括无法纳入辅助数据集，以及缺乏多变量处理能力。分位数映射还可能对大尺度特征产生不利影响，例如均值的演变（Hagemann 等，2011；Pierce 等，2013；Maurer 和 Pierce，2014；Ballard 和 Erinjippurath，2022）。构造类比及其变体是一种用于降尺度的多变量方法，研究表明，其性能优于其他方法，这可能是因为它们考虑了变量之间的相关性（Pierce 等，2014；Abatzoglou 和 Brown，2012）。然而，这些方法确实……

通讯作者：tobias@caltech.edu 作者按字母顺序排列，且均为共同第一作者。所有作者对本研究的贡献相同。

并不能真正实现从某个分布中采样高分辨率数据。

与此同时，生成式机器学习模型，例如生成对抗网络（GAN；如 Good-fellow 等人（2020））、变分自编码器（VAE；如 Kingma 等人（2019））和归一化流（NF；如 Papamakarios 等人（2021）），已被证明在超分辨率和域转换方面效果显著（如 Wang 等人（2021）；Zhu 等人（2017）），如今也开始应用于降尺度任务。在这些应用中，模型是深度卷积神经网络，可将具有多个通道的图像（例如低分辨率的气候模型变量场）映射为具有多个通道的输出图像（例如高分辨率的气候模型变量场）。数据集之间的统计关系通过“监督式”或“无监督式”方式隐式学习。在监督式方法中，数据点以对齐的方式从低分辨率映射到高分辨率：训练集中的每个低分辨率数据点都对应一个训练集中的高分辨率数据点。另一种情况是，当无法获得配对数据点时，降尺度任务就变为对条件数据分布进行建模，并从这些分布中生成样本。

在本研究中，我们提出使用基于扩散映射的生成模型，通过未配对训练数据生成降尺度后的流体流动。扩散模型在从各种已学习的高维概率分布（例如图像、音频和视频，Dhariwal 和 Nichol（2021）；Kong 等人（2021）；Ho 等人（2022a,b））生成逼真样本方面表现出极大的灵活性。对于域转换和降尺度，即将源分布中的样本转换为目标分布中的样本，生成式扩散模型相较于经典方法具有明显优势，例如。

- 生成模型允许从高维概率分布中进行采样。根据这些样本，可以计算任何统计量；
- 基于扩散的模型可以使用非配对数据进行训练，因此无需针对每一对源域/目标域重新训练，即可用于多种域转换任务（Su 等，2023）；
- 预训练的基于扩散的模型可以通过引导采样进行“重新利用”，从而从域的特定部分进行采样，例如（Ho 和 Salimans，2022；Dhariwal 和 Nichol，2021）。

这些特点使基于扩散的模型有别于经典方法，并且在某些情况下甚至有别于基于 GAN 的方法，因为后者中的许多方法需要配对数据或配对的源域/目标域。这表明，基于扩散的模型是一个很有前景的候选方案，非常适合应用于流体动力学和气候科学领域，因为

- 由于涉及高维数据点（例如完整的气候场），机器学习模型可能需要付出高昂的训练成本，因此频繁重新训练并不可取；
- 对于降尺度任务，由于确定性混沌以及小尺度运动对大尺度的反馈，高分辨率和低分辨率气候模拟的配对数据集并不存在；
- 具有偏差尾部概率的极端事件可能在气候变量和空间位置之间存在相关性，而针对所有感兴趣的统计量校准降尺度方法具有挑战性。因此，生成样本的能力可能非常理想。

在本研究中，我们展示了如何使用扩散模型在低分辨率和高分辨率流体模拟之间进行域转换，而无需针对所考虑的特定转换任务进行定制。我们重点关注如何根据低分辨率输入生成一致的高分辨率信息，以及如何校正重要的统计偏差，例如均值偏移、未解析的空间尺度以及尾部事件估计不足。

a). 相关工作

1) 气候数据降尺度

此前的研究已成功将基于机器学习的方法应用于气候模拟的降尺度和偏差校正。Pan 等人（2021）使用生成对抗网络对美国本土的气候模拟数据进行偏差校正，重点是匹配相应观测数据的各种统计量。Harris 等人（2022）利用配对的高分辨率雷达测量数据、ECMWF 模拟数据及其他上下文信息，使用生成对抗网络对模拟的低分辨率降水场进行降尺度。他们发现，该模型的表现优于许多传统方法，在极端降雨事件方面亦是如此。同样，Price 和 Rasp（2022）表明，可以使用条件生成对抗网络，直接利用高分辨率真实雷达观测数据，对低分辨率降水预报进行偏差校正和降尺度。Ballard 和 Erinjippurath（2022）使用对比式转换生成对抗网络和高分辨率观测数据，对耦合模式比较计划的低分辨率模拟数据中的日最高气温和降水进行降尺度。这种特殊的生成对抗网络变体支持以无监督、非配对的方式进行训练（Park 等人，2020）。同样，作者发现其性能与现有方法相当或更优。类似地，Groenke 等人（2021）使用非配对数据集，结合归一化流与类似 CycleGAN 一致性损失函数的损失函数，学习从低分辨率模拟数据到高分辨率无偏数据的域转换映射（Zhu

等人，2017）。基于 CycleGAN 的方法通常要求模型在训练时同时使用源数据和目标数据，因此每项转换任务都必须创建一个新模型。

2）扩散建模

扩散是一种动力学过程，会在较长时间尺度上抹去初始条件。将观测数据（即未知数据分布中的样本）用作初始条件，我们可以在扩散模型下向前积分轨迹，并选择该模型使得当 $t \to \infty$ 时，系统的长期稳态对应于来自已知分布（如高斯分布（先验分布））的样本。将来自先验分布的样本用作初始条件，求解反向扩散模型即可生成来自无条件数据分布的样本。多个作者已经证明，扩散过程可以用作生成模型，通过扩散过程将来自已知先验分布的样本转换为来自未知数据分布的样本，例如 Sohl-Dickstein 等人（2015）；Song and Ermon（2019）；Ho等人（2020）。事实上，最近 Song 等人（2021b）指出，生成扩散模型创建的图像可以理解为“反向扩散”随机微分方程的数值解，其初始条件等于来自先验分布的样本。这依赖于这样一个事实：如果已知得分函数（与数据分布的梯度相关），就可以对（正向）扩散过程进行反转（Anderson 1982）。此外，尽管直接近似无条件数据分布颇具挑战性，Hyvärinen 和 Dayan（2005）；Vincent（2011）已展示如何利用神经网络和对可处理损失函数进行梯度下降来近似分布的得分函数。

如上所述，针对无条件分布的数学结果可以扩展到条件分布，从而实现条件采样（Song 等人 2021b；Batzolis 等人 2021）。许多条件扩散模型需要配对输入数据点，而许多重要的条件生成任务都会提供这类数据（例如超分辨率、图像修复、着色，以及其他插补任务，包括时间序列任务——Tashiro 等人（2021）；Saharia 等人（2023）；Giannone 等人（2022）；Ho等人（2022a）；Saharia 等人（2022））。另一种方法是，Meng等（2022）展示了如何通过选择合适的起始点（初始条件）和反向扩散轨迹的起始时间，根据细节有限的简单笔触画生成逼真的图像。正如我们将要讨论的，这可以理解为根据输入笔触画生成条件高分辨率图像。关键在于，这一过程使用的模型仅在高分辨率数据上训练，并未接触笔触画。也就是说，从条件分布中生成样本并不需要配对数据。

与 CycleGAN（Zhu 等人 2017）相比，Su 等人（2023）展示了如何利用扩散模型进行域转换，使模型只需针对每个数据域训练一次，而不必针对每种感兴趣的转换分别训练（StarGAN 也是一种解决方案，但代价是复杂度增加；Choi 等人（2018））。扩散模型的这一特性源于两个域的扩散模型具有易于关联的先验分布，其优势在于，同一个模型可以用于多种转换任务。转换过程是先将一个域中的图像完全扩散，然后将最终状态转化为另一个域的先验分布中的样本，再使用该域的模型执行反向扩散。本文将展示的方法结合了 Meng等（2022）和 Su 等人（2023）的思路，由一系列扩散模型充当不同数据域之间的桥梁，不过这一总体思路早已存在（Chetrite 等人 2021）。

b). 我们的贡献

我们提出并测试了一种用于流体模拟降尺度的无监督方法。该方法基于将基于扩散的生成模型串联起来。它依赖于这样一个事实：粗分辨率和高分辨率的气候模拟在小尺度和中间空间尺度上存在差异，但在最大尺度上基本一致。由于我们在此采用的扩散过程会首先抹除最小尺度上的信息，因此我们可以从源域（低分辨率）的样本开始，对其进行扩散，直到小尺度信息丢失；随后，使用针对目标域（高分辨率）预训练的扩散模型对其进行反向扩散。所得图像在大尺度上与源图像的匹配程度，以及其同时包含与高分辨率数据统计量相匹配的精细尺度特征的程度，取决于我们停止前向加噪过程并开始反向扩散过程的时间。遵循Meng等（2022）和Su等（2023）的做法，我们将这种源域到目标域的扩散模型称为扩散桥。如前所述，与现有降尺度方法相比，该方法具有优势，因为它支持样本生成、可用于非配对数据，并且训练好的模型可以重复使用，但其尚未针对这一应用进行测试。

此外，我们还对扩散模型中使用的神经网络进行了架构改进。这些改进虽然并非我们的总体目标的重点，但能够提升性能指标并缩短训练时间。众所周知，基于得分的扩散模型容易出现“色偏移”问题：生成图像可能具有正确的空间特征，但相对于训练数据，其平均颜色会发生偏移。对于尺寸较大的图像，这种误差会进一步增大。一种改善这一问题的方法是对模型使用指数移动平均（EMA）。

参数，通常采用指数移动平均所隐含的超长记忆（Song and Ermon 2020）。因此，需要大量训练迭代次数才能达到良好性能。虽然可以通过其他技术缓解色偏移，例如 Choi 等人（2022），但我们通过在神经网络架构中引入旁路层来减轻这一影响。即使对于较大的图像尺寸（例如空间分辨率为 512x512 像素），这也消除了使用指数移动平均的必要。

2. 数据与模拟

我们使用一种二维平流-凝结模型生成本研究所用的流体模拟数据，该模型与 O'Gorman and Schneider (2006) 提出的模型相似。该模型旨在近似描述类地行星热带外大气中水分在等熵面上的动力学（O'Gorman and Schneider 2006）。因此，它是一个理想化的玩具模型。尽管如此，与完整的气候模型相比，该模型能够以较低的计算成本，详细研究其涡度和水分场的谱特性与分布特性。在本研究中，我们重点关注两个统计特性近乎各向同性的量：涡度，以及一种平流示踪剂，它代表类地大气中的过饱和度 $q' = q - q_s$，其中 q 是水相对于湿空气的质量分数（比湿），而 q_s 是空气达到饱和时的质量分数（饱和比湿）。

涡度根据具有随机强迫和线性耗散的二维欧拉方程演化。过饱和度由涡度场所隐含的流场输运。它受到空间均匀源 e 的强迫，该源可解释为蒸发场，为流场补充水分；同时，还受到空间变化的凝结作用，在过饱和度条件下减少水分 $q > q_s (q' > 0)$。因此，凝结代表了过饱和度示踪剂分布的尾部，而极端凝结事件则相应地位于更远的尾部。有关理想化平流-凝结模型的更多数学细节，请参见附录 5a。

为了模拟地球大气中沿等熵面 q_s 的饱和比湿经向衰减，我们假设其剖面呈线性衰减，并受空间周期扰动调制。空间周期扰动很有用，因为它可用于在不同长度尺度上对过饱和度示踪物统计量施加空间非均匀性。宽泛地说，这些非均匀性可被视为地形影响饱和比湿场的一种高度理想化版本。

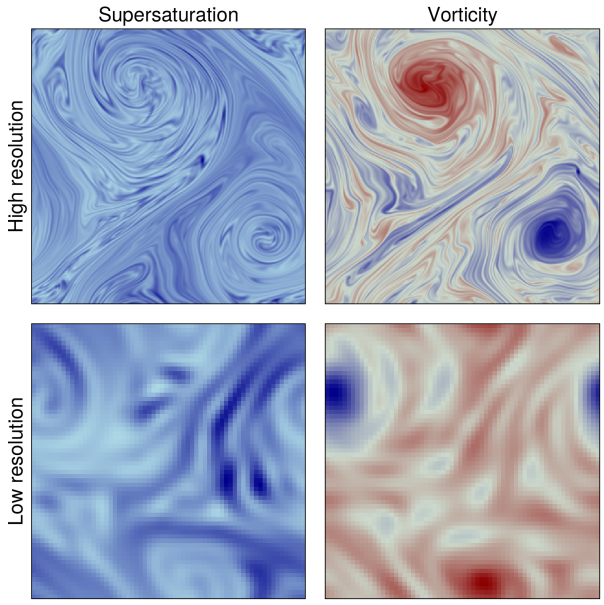

图 1：来自二维流体动力学模型的随机快照。顶行显示高分辨率快照，底行显示低分辨率快照。左列显示**过饱和度示踪物** 场，右列显示**涡度** 场。与颜色对应的数值对于本文所提出的方法并不重要，但对于过饱和场，正值（蓝色）对应于模拟模型中凝结项处于激活状态的区域（例如，发生理想化降雨事件）。过饱和场中的白色区域表示饱和亏缺区域（未发生理想化降雨事件）。对于涡度场，红色表示正值（气旋性涡度），而蓝色表示负值（反气旋性涡度）。

q_s 的时间无关形式由下列表达式给出

$$q_s(x, y) = \gamma y + A \sin\left(\frac{2\pi k_x x}{L}\right) \sin\left(\frac{2\pi k_y y}{L}\right), \quad (1)$$

其中 γ 表示背景饱和比湿梯度，A 是调制振幅，L 是 x 和 y 方向上的域大小，$k_x = k_y$ 表示波数，其取值为 $k_{x,\ y} \in \{1、2、4、8、16\}$，从而允许生成不同的大尺度饱和比湿廓线。因此，我们可以生成包含不同理想化地形或过饱和示踪物强迫的过饱和示踪场数据集。我们的目标是了解扩散模型在降尺度时能否充分利用上下文信息，因为地形、地表覆盖和其他场中的高频变化会影响 the

在更真实的气候模拟中模拟大气流动。这些特定地点的信息通常包含在生成模型中，包括第 1 节所述的模型。原则上，使用上下文训练模型将带来更好的性能和更强的泛化能力。

为了生成用于训练去噪扩散模型的数据，我们针对不同的分辨率和不同的饱和比湿廓线进行了一系列模拟，仅改变饱和比湿调制的波数 k_x，k_y。我们生成了一组高分辨率快照（512×512 像素），其背景饱和比湿廓线各不相同；同时生成了一个低分辨率数据集（64×64 像素），其中背景饱和比湿场固定且未经过调制，即$A = 0$。模拟中使用的参数见表 A1。达到稳态后，我们保存了这些模拟的快照，并将其用作扩散模型的训练数据。我们使用完整的高分辨率数据集训练扩散模型，其中包含所有上下文波数。图 1 展示了二维流体动力学模型的示例。顶排为高分辨率快照，底排为低分辨率快照。左列和右列分别展示过饱和示踪物场和涡度场。最后，我们使用最近邻加权将低分辨率64×64图像调整至 512×512分辨率，并通过应用低通滤波器去除高频混叠。我们确保真实 64×64 图像与调整后的图像之间的频谱信息保持不变。

3. 使用扩散桥进行降尺度

a). 扩散模型

我们对基于评分的生成模型的实现遵循 Song 等人（2021b）的方法。前向扩散（“加噪”）过程涉及向每个像素添加高斯噪声的独立样本，其中所添加噪声的均值为零，方差以预先规定的方式依赖于时间。具体而言，给定从数据分布中抽取的初始条件$\mathbf{x}(t = 0) \sim p_{\text{data}}(\mathbf{x})$，加噪过程由随机微分方程（随机微分方程）定义。

$$\mathrm{d}\mathbf{x} = g(t)\mathrm{d}\mathbf{W}, \tag{2}$$

其中，$g(t)$ 是预先设定的非负时间函数，$\mathbf{dW}$ 表示维纳过程。在任意时刻 t，该随机微分方程的解是“加噪”的图像$\mathbf{x}(t)$，其服从正态分布

$$\mathbf{x}(t) \sim \mathcal{N}(\mathbf{x}(0), \sigma^2(t)) = p(\mathbf{x}(t)|\mathbf{x}(0)), \tag{3}$$

其中，$\sigma(t)^2$ 是定义为以下形式的方差

$$\sigma^2(t) = \int_0^t g^2(t')\,\mathrm{d}t'. \tag{4}$$

这里，我们选择 $g(t)$，使得在 $t = 1$ 处，方差$\sigma^2(t = 1)$的量级远大于原始像素值，因此所有初始条件的信息都已丢失，即，

$$p(\mathbf{x}(1)|\mathbf{x}(0)) \approx \mathcal{N}(\mathbf{0}, \sigma^2(1)). \tag{5}$$

在这种观点下，扩散是一种将源图像嵌入潜在空间的过程，使得潜在空间中的样本 $\mathbf{x(1)}$ 近似地从一个已知分布中抽取，而该分布独立于源数据。

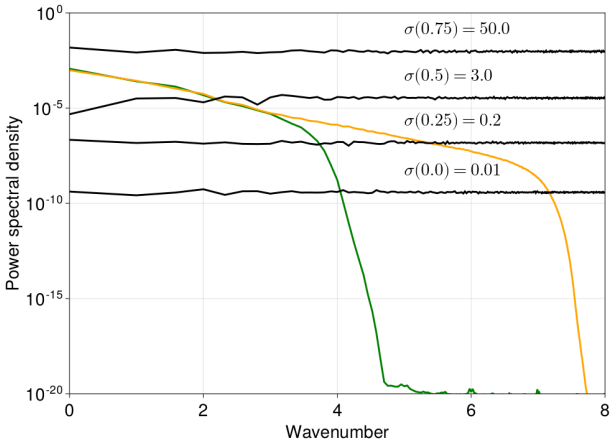

图 2：低分辨率和高分辨率数据集的涡度场功率谱密度，以及不同方差的高斯白噪声的功率谱密度 $\sigma(t)^2$。随着在较大的扩散时间（$t \to 1$）加入振幅更大的高斯噪声，加噪过程将首先抹去最小尺度上的信息；另请参见 Rissanen 等人（2023）。

为了近似地从数据分布中采样，我们反转这一过程。首先，我们从潜空间先验分布中采样，得到 $\mathbf{x(1)}$。这是反向扩散方程的初始条件，该方程从 $t = 1$ 求解至 $t = 0$。Anderson（1982）给出了反转方程（2）的方程：

$$\mathrm{d}\mathbf{x} = -g(t)^2\mathbf{s}(\mathbf{x}, t)\mathrm{d}t + g(t)\mathrm{d}\mathbf{W}, \tag{6}$$

其中，$\mathbf{s}(\mathbf{x}, t)$ 是数据分布的得分函数，

$$\mathbf{s}(\mathbf{x}, t) \equiv \nabla_x \log p_{\text{data}}(\mathbf{x}). \tag{7}$$

扩散建模所采用的训练过程的目标是确定得分函数的参数化表示

$\mathbf{s}_\theta(\mathbf{x},\ t) \approx s(\mathbf{x},\ t)$，通过对适当选择的损失函数进行梯度下降来实现。也就是说，我们对反向随机微分方程（SDE）中出现的时间导数进行参数化，并利用数据对其进行学习。通过这一反向模拟得到的最终图像$\mathbf{x}(t=0)$，结合经过训练的得分函数模型，即为新的数据样本。

b). 利用频谱信息的扩散桥

使用扩散桥进行域转换，需要将源数据集的前向模型与目标数据集的反向模型串联起来（例如，Meng等（2022）；Su等（2023））。我们的降尺度过程基于这一思路；下文将通过建立其与空间尺度差异化加噪之间的联系，明确说明这种方法为何有效。

根据公式（3），加噪后的图像$\mathbf{x}(t)$是$\mathbf{x}(0)$与高斯噪声之和，因此我们可以将其傅里叶变换写成$\mathbf{x}(0)$的傅里叶变换与高斯噪声的傅里叶变换之和（因为高斯噪声的傅里叶变换仍为（复值）高斯噪声）。由于二者不相关，我们可以近似地将功率谱密度写为 PSD $_{\mathbf{x}(t)}(k)$，其中 $k=\sqrt{k_x^2+k_y^2}$ 为波数：

$$\mathrm{PSD}_{\mathbf{x}(t)}(k) \approx \mathrm{PSD}_{\mathbf{x}(0)}(k) + \mathrm{PSD}_{\boldsymbol{\eta}(t)}(k) \tag{8a}$$

$$\eta(\mathbf{t}) \sim \mathcal{N}(\mathbf{0}, \sigma^2(t)). \tag{8b}$$

白噪声的功率谱密度与波数无关：$\mathrm{PSD}_{\boldsymbol{\eta}(t)}(k) = \sigma(t)^2/N^2$，其中 N 是图像大小。作为参考，功率谱密度将二维傅里叶变换压缩为一个与方向无关的一维信号。更多详情见附录 C4。流体流动的快照通常呈现出功率随 k增加而衰减的现象。因此，随着扩散时间从 $t=0$ 增加到 $t=1$，且 $\sigma(t)$增大，最小尺度（最大 k）会首先被噪声淹没，参见图2（另见 Choi 等人（2022）；Rissanen 等人（2023））。

我们假设存在一个空间尺度 $\lambda^\star$，高于该尺度时低分辨率数据无偏；经过低通滤波器处理的高分辨率模拟结果，在 $\lambda > \lambda^\star$处应与低分辨率模拟结果一致。$\lambda^\star$ 的存在意味着，低分辨率数据和高分辨率数据的期望功率谱密度 PSD(k) 在波数 $k < k^\star$处一致，其中 $k^\star = 2\pi/\lambda^\star$。因此，给定 $k^\star$ 的值，可以估计扩散时间 $t^\star$；在该时间点，所有小于 $\lambda^\star$ 的空间尺度上的信号信噪比均为 $\lesssim 1$，

$$t^\star = \sigma^{-1}\left(\sqrt{N^2\mathrm{PSD}(k^\star)}\right), \tag{9}$$

因为 $\sigma(t)$ 是已知的解析函数。

我们希望将源域 $\mathbf{x}_\mathcal{S} \in \mathrm{S}$ 中的图像转换为目标域 $\mathbf{x}_\mathcal{T} \in \mathrm{T}$ 中的图像。

更具体地说，我们来自 T 的样本是通过求解高分辨率流体模拟生成的 512×512 图像，而我们来自 S 的样本是通过求解分辨率为 64×64 的流体模拟，再进行上采样和低通滤波生成的 512×512 图像。所有数据$\mathbf{x}$ 均位于 $\mathbb{R}^{512\mathrm{x}512}$中；我们所说的“目标”域和“源”域 S 和 T，指的是我们假设数据所在的 $\mathbb{R}^{512\mathrm{x}512}$内的低维流形[1]。降尺度（域转换）算法定义了一个从 S $\in \mathbb{R}^{512\mathrm{x}512}$ 到 T $\in \mathbb{R}^{512\mathrm{x}512}$的映射函数。我们使用一个样本$\mathbf{x}_\mathcal{S}$ 作为初始条件，并求解源域的前向加噪模型，直至时间 $t^\star$。随后，我们将$\mathbf{x}(t^\star)$作为初始条件，并求解目标域的反向去噪模型，直至 $t=0$。由此得到的图像$\mathbf{x}(0)$即为从 T 生成的图像。该输运映射具有概率性，因为不同的计算结果会产生来自目标域的不同样本。此过程近似于从条件分布 $p(\mathbf{x}_\mathcal{T}|\mathbf{x}_\mathcal{S})$中采样，这一点在此处尚未得到证明，而且结果可能并不精确，但这一过程具有直观合理性：在此采样过程中，两幅图像之间的大尺度特征保持不变。降尺度算法的精确定义见算法 1。

算法 1 降尺度算法。仅步骤 1–3 执行一次，而步骤（4–7）则针对每个降尺度后的图像执行。

1：计算源域和目标域的期望功率谱密度 PSD$_\mathcal{S}(k)$ 和 PSD$_\mathcal{T}(k)$。2：求解 $k^\star$，使得 PSD$_\mathcal{S}(k^\star)$ = PSD$_\mathcal{T}(k^\star)$ ≡PSD*。3：计算 $t^\star$(PSD*)，见方程（9）。4：采样 $\mathrm{x}_\mathcal{S} \sim p_{\mathrm{data},\mathcal{S}}(\mathrm{x})$。5：通过求解方程（2），从 $t=0$ 到$t=t^\star$，并以 $\mathrm{x}_\mathcal{S}$ 为初始条件，得到 x($t^\star$)。6：通过求解方程（6），得到 x(0)，其中得分函数为 $\mathrm{s}_\mathcal{T}(\mathrm{x},t)$，求解区间为 $t=t^\star$ 到 $t=0$，初始条件为 x($t^\star$)。7：返回 x(0) $=\mathrm{x}_\mathcal{T} \sim p(\mathrm{x}_\mathcal{T}|\mathrm{x}_\mathcal{S})$。

图3展示了不同 $t^\star$取值下，降尺度过程生成的图像。当$t^\star \lesssim 0.5$时，扩散桥保留了低分辨率图像中的大尺度特征，但只填补了高分辨率图像中最精细尺度的特征，中间尺度的特征则缺失。当 $t^\star \gtrsim 0.5$时，前向加噪过程已经抹去了我们希望保留的部分或全部大尺度特征。当 $t^\star = 1$时，我们从$p_{\mathrm{data},\mathcal{T}}$采样得到一幅高分辨率图像，其中没有保留来自源图像的任何信息。就降尺度而言，最优值为

[1]鉴于流体模拟生成的轨迹受守恒量以及支配流场的偏微分方程约束，它们位于低维流形上似乎是合理的；许多高维数据集也被观察到位于低维流形上，例如 Brown 等人（2023）。

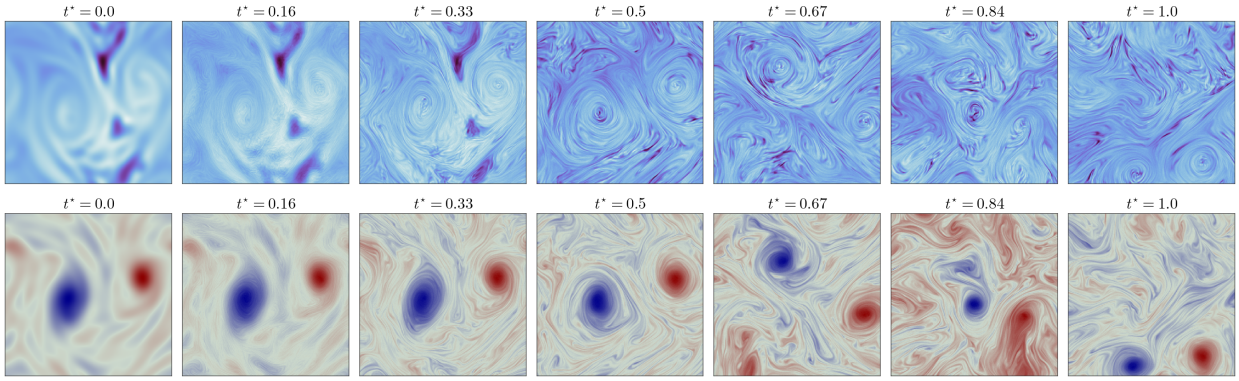

图3： $t^\star$ 对生成的**过饱和度示踪物**（上排）和**涡度**（下排）降尺度图像的影响。该值过小时，无法生成外观逼真的高分辨率图像；而该值过大时，虽然生成的高分辨率图像看起来逼真，但其选择是随机的。 $t^\star \lesssim 0.5$ 这一最优值能够生成与低分辨率源图像（$t = 0$）相符的逼真降尺度版本。对于过饱和场，正值（蓝色区域）对应模拟模型中凝结项处于激活状态的区域（例如，发生理想化降雨事件）。过饱和场中的白色区域表示饱和亏缺区域（未发生理想化降雨事件）。对于涡度场，红色表示正值（气旋性涡度），而蓝色表示负值（反气旋性涡度）。

约为 $t^\star = 0.5$，如图2所示。

在 Meng等（2022）中也探讨了寻找 $t^\star$ 最优值的核心思路，但未解释其与空间尺度之间的联系。在那项工作中，他们在忠实于“引导”图像（相当于我们的低分辨率图像）与目标图像的逼真度之间进行了权衡，并采用了核Inception评分和L2范数。选择最优的 $t^\star$ 取决于特定问题所关注的指标，而许多方法都可能奏效。

需要注意的是，该过程要求感兴趣图像的功率谱密度随波数增加而降低。此外，我们当前的实现依赖于图像不同通道具有相似谱密度形状这一事实，因此单个 $t^\star$ 值即可适用于每个通道。更通用的模型则允许不同通道采用不同的调度函数$\sigma(t)$。

c). 上下文信息

对于真实的气候模拟和数据，传输映射也可能取决于上下文信息，例如地表属性、地形等。为了研究这一点，我们采用了一个上下文场，我们的数据生成模型——平流-凝结模型——依赖于该上下文场（这种依赖关系将在第 2 节和附录 5a 中解释）。我们将这些上下文信息统称为 $\mathbf{x}_C$，它与流体状态变量场（记为 $\mathbf{x}$）相对齐。因此，我们能够从分布 $(\mathbf{x}, \mathbf{x}_C) \sim p(\mathbf{x}, \mathbf{x}_C)$ 中采样成对的数据。对于低分辨率在各次运行中，上下文被视为均匀的。在高分辨率下，上下文的分辨率与流体状态变量场相同，且随空间变化。在本研究中，我们将上下文作为额外通道，输入到用于建模得分函数的卷积神经网络中。我们不会对这些通道进行任何“扩散”。更多详情见附录 B2。

加入上下文信息不会改变上文对扩散桥的说明。我们假设存在一个空间尺度，在该尺度以上，低分辨率模拟没有偏差，且上下文信息尚未影响流场。对源（低分辨率）域中的一个数据点施加高斯噪声，直到小尺度信息消失，而最大尺度的信息仍大致保持不变。随后应用反向扩散，将加噪图像映射至目标（高分辨率）域。在反向扩散过程中，上下文信息才会介入并发挥作用。这意味着，目标域中最终到达的不同区域取决于上下文信息。这表明，上下文信息可用于引导生成过程。

该过程如图4所示，其中我们使用相同的源图像生成具有不同上下文的降尺度图像（我们仅展示过饱和场；在我们的设置中，涡度不受上下文影响）。对于$k_x = k_y = (8, 16)$情况（前两行），可以使用相同的 $t^\star$值，因为高分辨率和低分辨率功率谱密度仅在大于上下文扰动空间尺度的空间尺度上相一致。另一方面，对于 $k_x = k_y = 2$情况（底行），我们

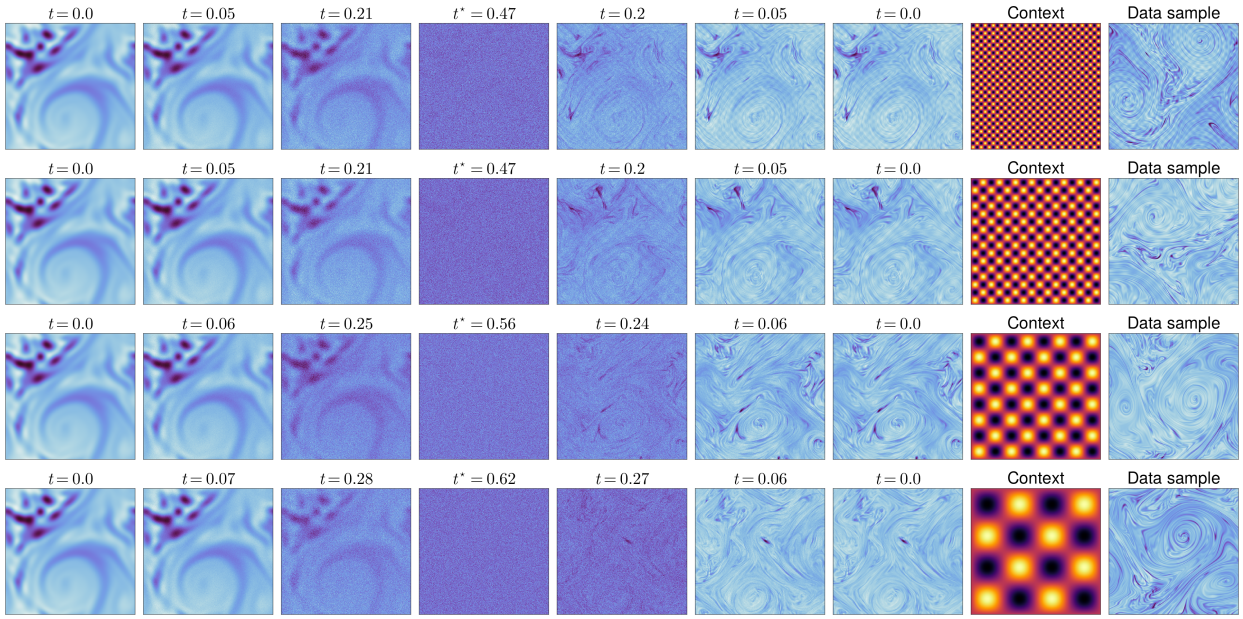

图 4：使用四种不同的上下文（按行排列），对同一个低分辨率**过饱和度示踪物**场进行降尺度。左侧是源图像，位于 $t = 0$。从左向右，相当于逐步推进降尺度过程：使用源域的前向模型对图像加噪，在 $t = t^\star$ 时切换到反向模型，积分回 $t = 0$。生成的高分辨率样本（位于 $t = 0$）在流体流动中具有不同的周期信号；这些信号由各自的上下文（从右数第二列）决定。最右列还展示了每种上下文随机选取的一个数据样本。所有加噪图像均已缩放至相同范围，这是必要的，因为所添加噪声的方差会随时间增大。对于过饱和场，正值（蓝色区域）对应模拟模型中凝结项处于活跃状态的区域（例如发生理想化降雨事件的区域）。过饱和场中的白色区域表示饱和亏缺区域（未发生理想化降雨事件）。

必须使用更大的 $t^\star$ 值，才能从上下文场中恢复信号。同时，该 $t^\star$值处于低分辨率信息正在丢失的范围内（如图3所示）。波数 $k_x = k_y = 4$的情况介于两者之间。源图像中的大尺度特征虽仍然存在，但已发生失真；与具有相同上下文的随机抽取数据样本相比，调制也不那么明显。第4节中，我们将展示低分辨率和高分辨率数据集的功率谱密度，以便更定量地讨论这一问题。

关于图4，请注意：(1) 即使每次降尺度模拟使用相同的上下文场，由于降尺度过程具有概率性质，生成的高分辨率图像也会有所不同；(2) 用于生成不同上下文样本的扩散模型只有一个。训练该模型时使用了所有上下文中的全部数据。

d). 空间均值偏差校正的捷径

正如引言中所讨论的，扩散模型可能难以生成空间均值正确的图像（即 RGB 图像中的“色偏”），同时又能生成逼真的空间变化（例如，功率谱看起来合理）。对此，推荐的解决方案是对模型参数采用记忆周期较长的指数移动平均（EMA）（Song and Ermon 2020）。某些得分网络架构中使用了注意力模块（例如 Ho等人（2020）），这也可能改善色偏，因为自注意力能够学习非局部特征（Wang等人 2018）。这些方法大多会增加训练期间的计算成本。虽然这未必会造成难以承受的负担，但仍有可能避免。

图像空间均值中的误差只能由得分函数空间均值中的误差导致。尽管用于得分建模的神经网络具备预测这一量的能力，但它们并不能高效地学会这样做。一种实用的解决方案是将网络的任务拆分为两个独立的任务：预测得分函数的空间均值，以及预测得分函数相对于其均值的空间变化。我们通过在网络的旁路层中预测得分函数的空间均值来实现这一点，该旁路层具有独立的参数。这样，我们便能保持神经网络架构简单，基本上只由一个基础 U-Net（Ronneberger 等，2015）构成，

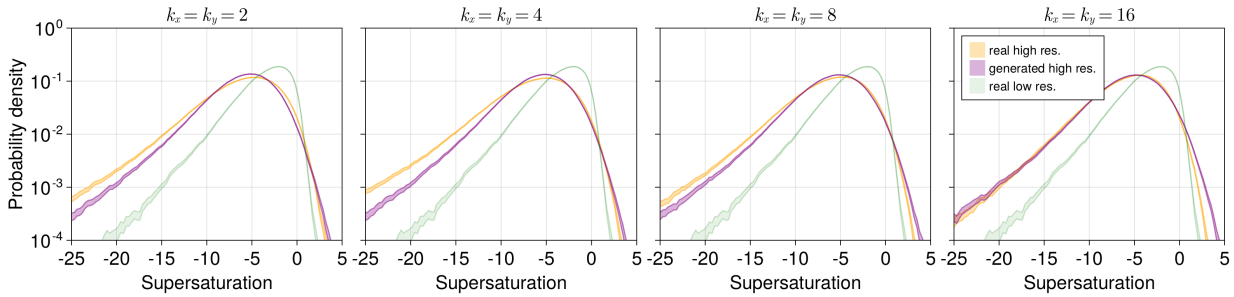

图 5：**过饱和度示踪物** 场取值的概率密度函数估计。各列对应不同的高分辨率数据子集，其中 k 表示饱和比湿的调制波数。橙色分布表示真实高分辨率样本的分布，紫色分布表示生成并降尺度后的高分辨率样本的分布，绿色分布表示真实低分辨率样本的分布。概率密度通过核密度估计进行估算。阴影区域根据 10000 个自助抽样样本在 99 % 置信区间下计算得出。

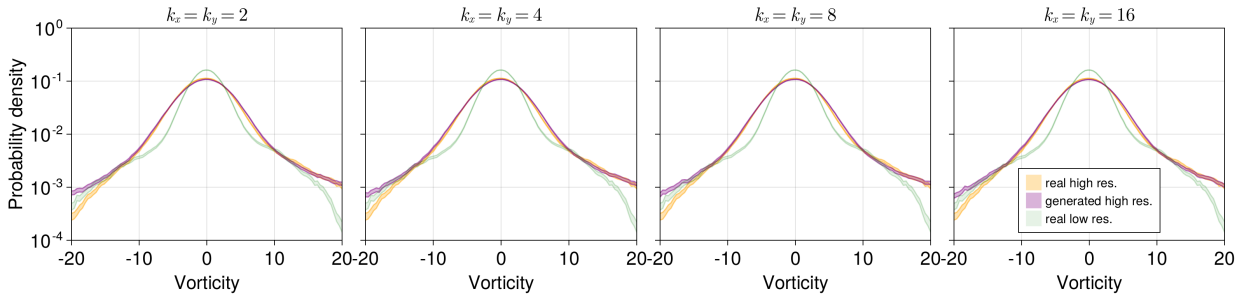

图 6：与图 5 相同，但此处绘制的是所有涡度**涡度**量。与过饱和度示踪场取值不同，涡度场取值的分布近似以零为中心对称，因为模拟中的涡度强迫也具有这一特性，而过饱和度示踪物则不然。

并且在训练过程中，无需像其他方法那样经历更长的训练过程。通过避免采用更复杂的解决方案，我们能够减少可训练参数的数量，从而使训练和采样过程尽可能保持较高的计算效率。因此，即使使用简单的欧拉–丸山采样器生成样本，我们生成的图像也不会出现明显的色偏。我们计划在另一项独立研究中进一步探讨这一架构选择的更多细节。

e). 降尺度扩散桥

为了实施我们的基于扩散的降尺度方法，我们针对高分辨率数据集训练了一个上下文扩散模型。由于前向加噪过程独立于得分函数，因此我们不需要针对低分辨率数据集训练模型（只有当我们希望生成低分辨率样本时才需要这样做）。换句话说，在层次自编码器设置中，加噪方向的作用类似于预训练的编码器（Luo 2022）。

关于扩散模型的构建、网络架构、损失函数、训练流程和采样方法的详细信息，见附录 B1、B2、C1、C2、C3 和 C4。这里需要指出的是，我们的网络架构无法保留流场的双周期特性。这是因为该特性是此数据集独有的，在大多数应用中并不存在，例如，对较大流体模拟中的局部区域进行降尺度。

4. 结果

本节将通过多项指标评估扩散桥生成样本的质量。在偏差校正方面，我们重点关注空间平均值偏差、中间尺度偏差，以及更极端尾部事件（例如分布尾部）中的偏差。此外，我们还量化大尺度信息

的保留情况，评估模型添加小尺度信息的效果，并探究模型对未见上下文场的泛化能力。

a). 过饱和度示踪物涡度值的分布

图 5 和图 6 展示了四种不同上下文场下过饱和度示踪物和涡度场的概率密度函数。 $k_x = k_y$的值表示上下文场中使用的调制波数，具体见公式 (1)。绿色概率密度函数显示真实低分辨率数据的数值分布，橙色概率密度函数显示真实高分辨率数据的数值分布。紫色概率密度函数显示使用上下文相关扩散桥方法对高分辨率样本进行降尺度（例如，生成）后所得数值的分布。

图 5 和图 6 显示，上下文相关扩散桥方法将从低分辨率数据集中计算出的场值分布移至接近从高分辨率数据集中计算出的数值分布的位置。可以看出，降尺度会调整分布的均值和方差，使生成数据的分布与高分辨率数据的分布比与低分辨率数据的分布更加接近。尽管如此，左尾部仍持续低估，误差约为 O(1) 倍。我们发现，随机生成的高分辨率图像（非 经降尺度的低分辨率图像）表现出相同的情况（未显示），这表明这些误差源自模型本身，而非降尺度过程。

b). 过饱和度示踪物和涡度的空间均值

图 7 展示了过饱和示踪场空间均值的概率密度函数估计结果。这些结果表明，即使在过饱和示踪场的空间均值上，低分辨率模拟与高分辨率模拟也存在差异。这说明，基于扩散桥的降尺度方法不仅能补充小尺度空间特征，还能校正大尺度特征中的偏差。尽管我们发现空间均值偏差得到了校正，但生成数据样本的空间平均值方差大于真实高分辨率数据。我们推测，改进均值旁路层或许有助于缓解这一差异。

图 8 展示了涡度场空间均值的概率密度函数估计结果。根据二维流体动力学模型的设计，涡度的空间均值围绕零值波动，并且几乎保持守恒。这一特性在生成的数据样本中得到了重现，这也是我们选择在神经网络设计中引入均值旁路层的结果，具体如第.....节所述。3d 和附录 C2。在我们的改进的 U-Net 架构中去除均值旁路层后，我们发现更难获得空间均值涡度离散程度最小的数据样本。

同样，我们确认随机生成的高分辨率图像（非降尺度后的低分辨率图像）表现出相同的行为（未展示），这表明这些误差源自模型本身，而非降尺度过程。

c). 功率谱密度与上下文信息的作用

我们的流体模拟通过过饱和度示踪方程依赖于上下文场，而涡度场不受此信息影响。然而，从像素值分布和空间均值（图 5 和图 7）无法看出上下文对高分辨率过饱和场的作用。本文通过流场的功率谱密度来探讨上下文信息的作用。该指标还能帮助我们了解降尺度方法在小尺度上补充信息、校正中间尺度偏差的效果。

图 9 和图 10 展示了上下文相关扩散桥降尺度算法在谱空间中的表现。图 9 展示了过饱和度示踪场的方位角平均功率谱密度均值，图 10 展示了涡度场的方位角平均功率谱密度均值。绿色谱线表示真实低分辨率数据的数值分布，橙色谱线表示真实高分辨率数据的数值。紫色谱线表示采用上下文相关扩散桥方法降尺度（即生成）的高分辨率样本的数值分布。可以看到，真实低分辨率数据的谱在空间波数相对较低时便迅速衰减。这是由于流体动力学模拟对小尺度的阻尼增强所致。上下文相关扩散桥方法不仅在比较低分辨率和高分辨率数据集时“补全”谱中缺失的部分，还能校正源自上下文信息的中间尺度偏差，即高分辨率模拟中背景饱和比湿场的调制。

对于所有波数，与真实高分辨率图像相比，生成图像在所有尺度上整体都存在功率不足的问题，尽管其整体形状是正确的。这表明模型已经学会了正确的空间模式，但生成图像的整体对比度略显不足。这一现象在训练过程中也有所观察到；我们推测，采用更精细的得分函数神经网络架构，并结合更多数据，可以缓解这种差异。同样，这种行为也出现在随机生成的图像中，因此其原因在于模型本身，而非降尺度过程。

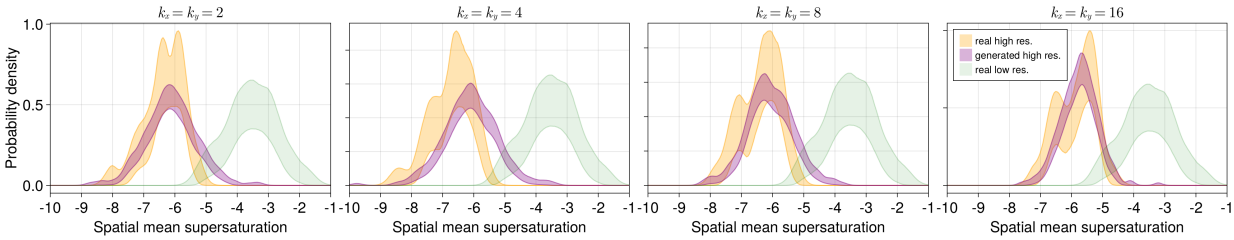

图 7：**过饱和度示踪物**场空间均值的概率密度函数估计。各列对应不同的高分辨率数据子集，其中 k 表示饱和比湿调制波数。橙色分布表示真实高分辨率样本的分布，紫色分布表示生成并降尺度后的高分辨率样本的分布，绿色分布表示真实低分辨率样本的分布。阴影区域根据 10000 个自助抽样样本计算，表示 99 % 置信区间。

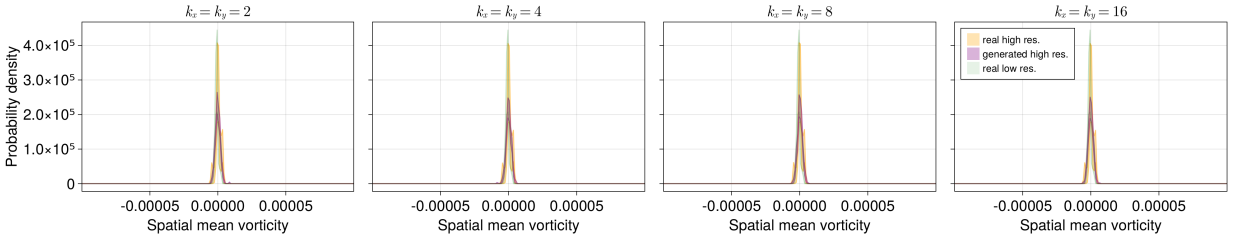

图 8：与图 7 相同，但此处绘制的是涡度的所有量。模拟中空间平均涡度并不守恒，而是在零值附近以较小振幅波动。

更重要的是，在波数$k_{x,\ y} = 4$和 $k_{x,\ y} = 2$处生成的图像，缺少上下文所规定的流场波数处的能量。这一点在图4中也有体现，并在第3c节中讨论过。这是因为我们的算法需要选择一个 $t^\star$ 值，以便在保留流场大尺度特征与添加中小尺度特征之间取得平衡。对于波数更大的上下文，这种平衡并不存在。也就是说，如果高分辨率和低分辨率流体流动在几乎所有尺度上都不同，我们的方法就无法奏效。这类转换是否仍可称为“降尺度”，尚不明确。我们验证了生成图像在上下文波数处能量不足并非由模型本身导致；纯生成式高分辨率模型（使用 $t^\star = 1$）在所有情况下都能正确添加上下文特征；而这样的 $t^\star$值会完全丢失我们试图在降尺度中保留的低分辨率特征。

这些图还展示了上下文在降尺度过程中的作用，因为它显然会影响生成图像的功率谱。我们初步测试了上下文扩散模型在未见过的上下文中的表现。为评估这一点，我们使用两个波数创建了一个新上下文，其中一个波数在训练期间未曾使用过，并执行了降尺度算法。生成的图像、上下文字段以及过饱和度示踪物的功率谱密度如图 11 所示。模型正确地施加了调制，但由于我们没有该上下文下真实流体的模拟结果，因此难以更定量地评估其性能。

d). 凝结速率分布

为了进一步评估我们降尺度方法的性能，我们计算了低分辨率和高分辨率数据集的凝结速率分布。为此，我们对数据中的正凝结速率进行了核密度估计。凝结速率的计算方法见附录 5a，在理想化模型中，可将其视为降雨形成速率。图 12 展示了降尺度算法在评估凝结速率分布时的表现。绿色分布表示低分辨率数据中的凝结速率，橙色分布表示真实高分辨率数据中的凝结速率。紫色分布表示采用依赖上下文的扩散桥方法降尺度（例如生成）的高分辨率样本中的凝结速率。可以看出，对于罕见事件，尤其是极其罕见的事件，低分辨率分布对高分辨率数据尾部的估计低了一到两个数量级。这是因为在流体流动的低分辨率数值模拟中，尖峰会被平滑掉。

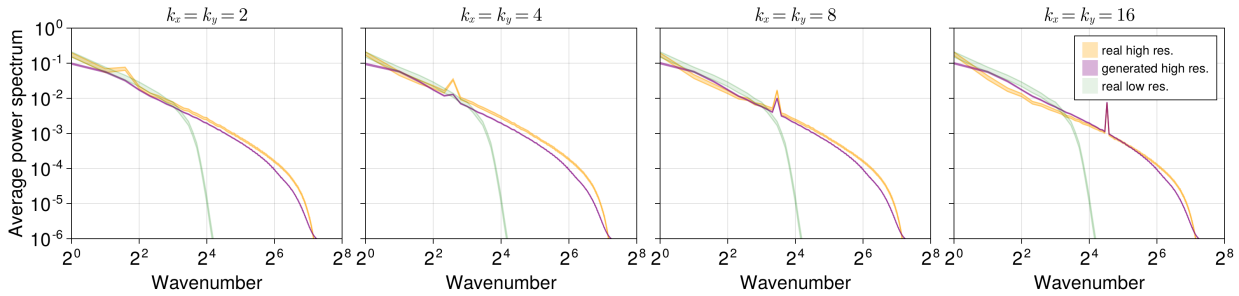

图 9：**过饱和度示踪物** 场值的方位角平均谱密度函数估计。各列对应不同的高分辨率数据子集，其中 k 表示饱和比湿调制波数。橙色谱表示真实高分辨率样本的谱，紫色谱表示生成并经降尺度的高分辨率样本的谱，绿色谱表示真实低分辨率样本的谱。阴影区域根据 10000 个自助抽样样本计算，置信区间为 99 %。

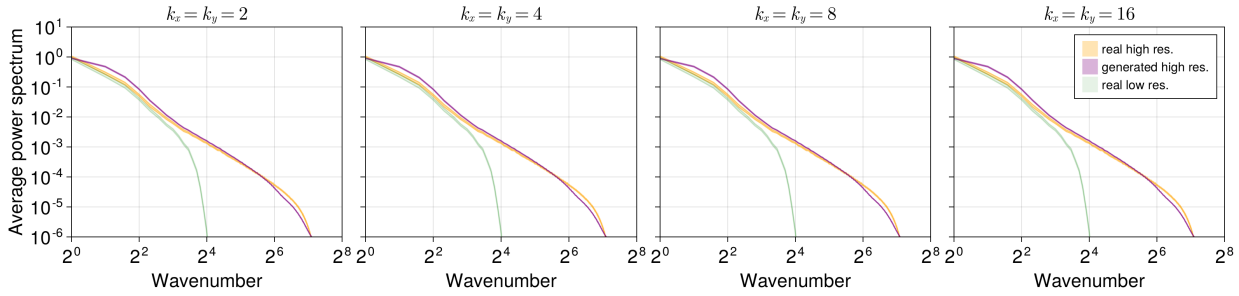

图 10：与图相同。图 9，但现在我们绘制的是所有涡度场值对应的物理量（例如涡量平方谱）。).

降尺度过程“抬升”了尾部，并减轻了凝结速率尾部事件中的偏差。

然而，我们发现，生成的样本高估了极其罕见事件（例如，1/1000 的事件）的发生频率。这种情况可能由多种原因造成。一般而言，由于缺乏来自数据域这一部分的训练数据，机器学习模型可能无法很好地处理分布尾部。然而，我们推测，除这一普遍存在的困难外，扩散模型还可能在生成的样本中留下少量肉眼无法察觉的残余噪声，但这种噪声会体现在尾部统计量中。这是由于许多扩散模型采用了特定的加噪调度，在样本生成过程中添加的最终噪声并不等于零。为了进一步改善这一问题，可能有必要寻找一种改进的加噪调度。由于当最终噪声振幅趋近于零时可能出现数值不稳定性，因此我们将这一技术挑战留待未来工作解决。我们还注意到，分布尾部的误差可能源于均值的误差。与真实图像相比，我们生成的图像具有更大的均值方差（图 7），这也可能导致分布尾部发生偏移。

e). 条件采样评估

根据第 3c 节中的符号表示法，我们期望一个合理的降尺度算法能够近似地从条件分布 $p(\mathbf{x}_\mathcal{T}|\mathbf{x}_\mathcal{S})$ 中生成样本，其中 T 是高分辨率数据域，S 是低分辨率数据域。更具体地说，我们期望$\mathbf{x}_\mathcal{T}$与 $\mathbf{x}_\mathcal{S}$之间的大尺度空间特征能够得到保留。为了检验我们的算法在多大程度上满足这一要求，我们使用逐像素 L^2-度量计算降尺度图像与其对应的低分辨率源图像之间的距离。此外，我们还对随机选取的两幅低分辨率图像计算相同的统计量，并比较这两种情况下所得距离的分布。

图 13 通过箱线图展示了所得的 L^2 度量值分布。这表明，下采样后的图像比两幅随机选取的低分辨率图像彼此之间更接近其低分辨率源图像，说明扩散桥算法保留了宽尺度的空间特征。需要注意的是，如图 7 、 9 和 10 所示，高分辨率数据集与低分辨率数据集之间存在偏差，因此我们在计算

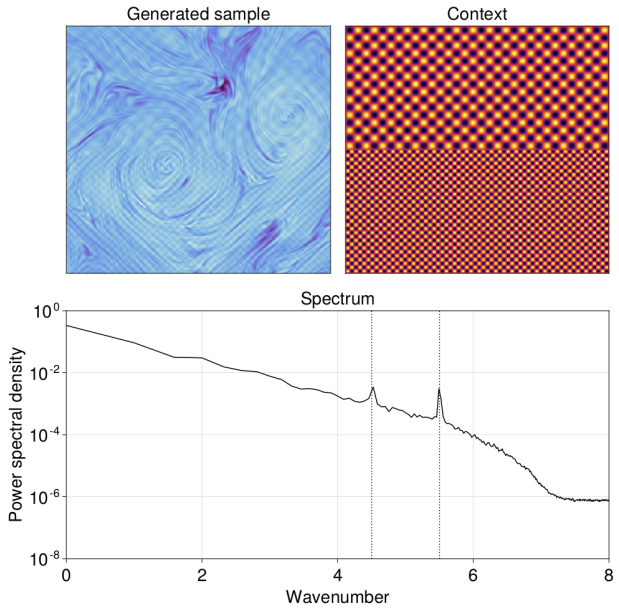

图 11：使用训练期间未见过的上下文场演示扩散桥算法。左侧显示降尺度的**过饱和度示踪物**场，右侧面板展示采样时使用的上下文，底部显示降尺度图像的功率谱密度。虚线表示两个上下文空间频率的波数。

距离度量之前，先进行了以下变换。我们对图像进行低通滤波，使其仅包含频率为 $k < k^{\star}$ 的空间频率。随后，我们使用相关数据域的平均像素值和像素值标准差对图像进行归一化。如果不进行这一步处理，下采样图像与真实源图像之间的 L^2 度量值可能会非常大，但这主要是由低分辨率数据的偏差造成的。

5. 结论

我们证明，采用依赖上下文的扩散桥的降尺度方法，能够校正空间均值偏差和中间尺度偏差，并提升理想化低分辨率流体动力学模拟的分辨率。此外，我们还证明，该方法有助于“拓展尾部”，改善低分辨率凝结率的尾部，使凝结率尾部的概率密度值至少得到一个数量级的校正。这表明，即使在更现实的场景中，且训练损失函数没有任何明确侧重的情况下，基于扩散的生成模型也可能能够校正极端事件发生率的偏差。我们还展示了扩散桥方法生成的降尺度图像，其方位角平均功率谱统计量以及过饱和度示踪物和涡度值的分布均与原始高分辨率数据相符。通过在扩散桥内用于建模反向扩散过程得分函数的神经网络中引入旁路连接，该方法缓解了空间均值偏差（例如色偏）问题，并保留了空间平均涡度值。这意味着，基于全局积分的守恒定律或许能由扩散模型自然满足，而无需在损失函数中额外明确强调。虽然这对于应用于现实气候模拟可能并不重要，因为这类应用每次只对流场的较小区域进行降尺度，但在其他上下文中或许会有用。

正如引言中所指出的，在流体流动的降尺度任务中，扩散模型相较于经典方法和其他生成式机器学习方法具有一定优势。我们发现，其优势可概括如下：

- 扩散模型灵活且可复用。本研究开发并应用的降尺度方法无需针对当前数据集进行任何特殊调优，而且在训练过程中完全不需要低分辨率数据。对于在其他模型中生成的数据或取自观测数据的域转换任务，只需为每个域训练一个扩散模型，从而降低训练所需的计算开销。

- 本研究用于训练扩散模型的损失函数具有通用性，且基本未作修改，因此并未特别侧重极端事件。训练我们的模型时，没有使用分位数损失或谱损失函数。

- 扩散桥能够近似地从高分辨率条件分布中生成样本。这在某些应用场景中很有用，例如需要计算复杂统计量，或训练完成后才知道需要计算哪些统计量的情况。

虽然这些优势中的一部分未必适用于每种建模场景，但我们发现，总体而言，基于扩散的模型具有很强的灵活性，因此在生成建模场景中颇具吸引力。

a). 替代方案与未来方向

本研究所使用的数据由二维受迫湍流流体和一种过饱和度示踪物组成。其中包含若干与更复杂的气候模拟相似的特征（过饱和度示踪物的非高斯分布统计特性，以及特定地点类似地形特征的影响）。然而，一个显而易见的下一步是

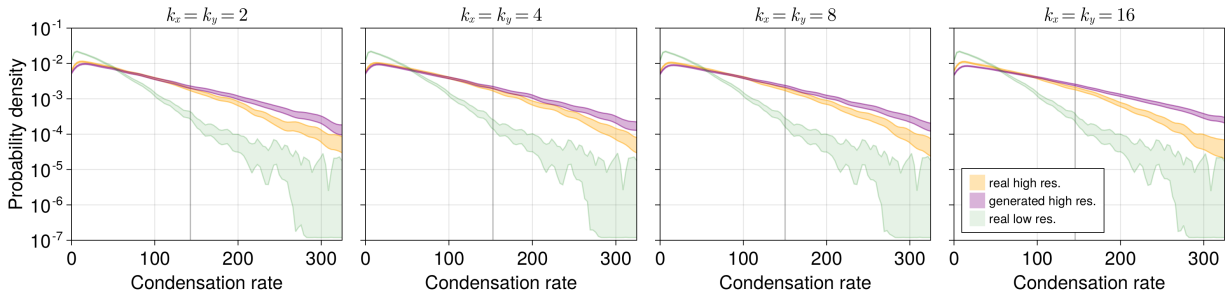

图 12：瞬时**凝结速率**的概率密度函数估计。各列对应不同的高分辨率数据子集，其中 k 表示饱和比湿调制波数。橙色分布表示真实高分辨率样本的分布，紫色分布表示生成并降尺度后的高分辨率样本的分布，绿色分布表示真实低分辨率样本的分布。阴影区域根据 10000 个自助抽样样本的 99 % 置信区间计算得出。竖线近似表示真实高分辨率凝结率的第 90 百分位数。

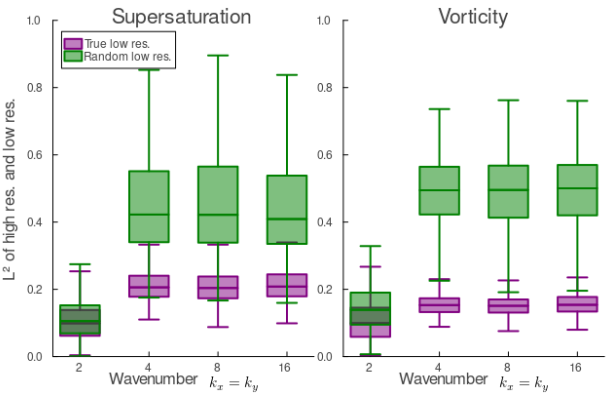

图 13：按通道和波数比较降尺度图像与其低分辨率源图像。经过滤波的降尺度图像与其经过滤波的源图像（低分辨率图像）之间逐像素 L^2-度量值的统计分布以紫色显示；两个随机低分辨率图像之间的相同统计分布以绿色显示。

是使用真实气候数据集检验本研究提出的方法，并将其性能与其他现有降尺度方法进行更直接的比较。此外，我们还确定了一些可能的研究方向和有待解决的问题：

- 无需借助扩散桥，也可以利用基于扩散的生成模型实现降尺度。另一种方法是在样本生成过程中采用投影步骤，这与 Song 等人（2022）所展示的方法类似。
- 采用扩散模型或许能够实现样本在时间上的连贯性，这些模型用于视频生成 Ho等人（2022a）。测试这些模型在物理系统上的表现会很有趣，但计算成本方面可能存在一些需要解决的缺点。
- Ho 和 Salimans（2022）提出的扩散模型引导采样技术，可能有助于生成具有其他理想特征的样本，例如某些气候指数具有较高值等。

不过，目前的工作仍有进一步拓展的空间。正如第4节所述，我们已经发现，与真实的高分辨率图像相比，降尺度图像存在一些细微差异，并确定这些差异主要是由模型本身造成的（而非降尺度过程）。改进网络架构和训练过程的研究可能会提升模型及其结果。此外，在训练中使用更加多样化的上下文，并真正证明模型能够泛化到样本外上下文，是下一步的重要工作。

总体而言，基于扩散的模型似乎是未来地球科学应用中颇具前景的候选方案。

致谢。 本研究得到了 Eric 和 Wendy Schmidt（经 Schmidt Futures 推荐）以及 Cisco 基金会的支持。作者感谢 Simone Silvestri、Ricardo Baptista、Nikola Kovachki、Honkai Zheng、Andrew Stuart 和 Tapio Schneider 就本研究展开的富有启发性的讨论。我们感谢 Andre Souza，他协助完善了本研究所用流体动力学模型的早期版本。Andre Souza 和 Tapio Schneider 对本文早期草稿提出了有益的意见。我们还感谢流体模拟代码 FourierFlows.jl 的作者，本研究的模拟基于该代码完成。本文所用的所有数值计算和模型训练，均在加州理工学院 Resnick 高性能计算中心的协助下完成。

数据可用性声明。我们用于扩散模型训练和分析的代码托管在 GitHub 仓库: https://github.com/CliMA/diffusion-bridge- downscaling。流体模拟的 GitHub 仓库位于: https://github.com/CliMA/SimpleMoisture.jl。

附录 A

数据生成模型

本研究所用的数据生成模型由一个动力系统构成，用于模拟地球热带外大气等熵面上的水汽平流和凝结。该模型虽经过理想化处理，但仍包含足够的复杂性，可用于检验本研究所述机器学习算法的性能。具体而言，该模型具有一些理想的动力学和统计特性。例如，过饱和场在空间上变化显著，其对应的理想化瞬时凝结速率服从近似具有指数尾部的分布（O'Gorman and Schneider 2006）。在评估生成式机器学习模型的谱精度和统计精度，尤其是其对极端事件的刻画能力时，这些特性十分有用。

该模型的研究动机及其行为已在 O'Gorman and Schneider（2006）中进行了详细阐述，因此，我们在此仅简要回顾该模型的主要组成部分。该模型的核心是在周期性区域上建立二维涡度方程，并通过随机强迫以及线性阻力和超扩散进行阻尼（谱滤波是另一种可选方法）。涡度场的控制方程如下：

$$\partial_t \zeta + \partial_y \Psi \partial_x \zeta - \partial_x \Psi \partial_y \zeta = f - a\zeta - \kappa \Delta^8 \zeta \tag{A1}$$

$$\Delta \Psi = \zeta, \tag{A2}$$

其中 ζ 表示涡度， Ψ 表示流函数， f 表示一种随机强迫，其波数谱具有各向同性，且能量集中在以 ζ 为中心、波数范围狭窄的环带内空间，带宽为 Δk， a 表示摩擦时间尺度， κ 则起超扩散率的作用。比湿场 q 的方程为

$$\partial_t q + \partial_y \Psi \partial_x q - \partial_x \Psi \partial_y q = e - c - \kappa \Delta^8 q \tag{A3}$$

$$c = \frac{1}{\tau}(q - q_s)\Theta(q - q_s), \tag{A4}$$

其中， Ψ 仍表示流函数， e 表示蒸发速率，假定其在时空中固定； c 表示瞬时凝结速率。这里， κ 与方程（A1）中的超扩散率相同。凝结速率 c 与比湿 q 和饱和比湿 q_s之差成正比，但只有在 $q > q_s$时才会发生凝结，这与地球大气中的情况一致。在我们的模拟中，考虑凝结时间尺度 τ 较小但有限的情形。换言之，过饱和（$q > q_s$）区域会在时间尺度τ内回调至饱和比湿 q_s。 τ 的有限性模拟了非平衡热力学过程，但并非本文结论成立的必要条件。正如正文所述，我们令 q_s 随空间变化，以模拟地球大气中沿等熵面 q_s 的衰减，并在不同长度尺度上施加空间非均匀性（方程（1））。当平均饱和亏缺较大（$q < q_s$）时，凝结事件很少发生，平均凝结速率趋近于零。当蒸发速率较大时，蒸发会压过湍流通过沿平均水汽梯度向上平流而生成未饱和流体团的能力（参见 O'Gorman and Schneider（2006））。

完整数据集由六个子集组成，包括一个低分辨率子集和五个高分辨率子集，共计 12,000 个数据点。用于生成完整数据集的参数总结见表 A1。

附录 B

B1. 扩散模型

如正文所述，我们的扩散模型的加噪过程在每个时间步向图像添加高斯噪声。我们采用了所谓的"方差爆炸"调度，其中

$$g(t) = \sigma_{\min}\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^t \sqrt{2\log\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)} \tag{B1a}$$

$$\sigma^2(t) = \sigma_{\min}^2\left[\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^{2t} - 1\right] \approx \sigma_{\min}^2\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^{2t}, \tag{B1b}$$

其中， $\sigma_{\min}$ 和 $\sigma_{\max}$ 是标量参数，用于决定方差随时间变化的形状。其他加噪过程也已得到应用，例如投影到一组基函数后在系数空间中进行加噪（Phillips 等人（2022）），以及通过模糊

表 A1：数据生成模型的参数值。本研究生成的完整数据集由六个子集组成：一个未经过饱和比湿调制的低分辨率数据集，以及五个调制波数各不相同的高分辨率子集。所有模拟均运行了 200,000 个时间步，并舍弃了前 100,000 个时间步作为自旋调整，以用于此处介绍的研究。报告的子集大小包含自旋调整部分。所有数值均为无量纲形式。

共享参数					
流体域尺寸（L）	时间步长（Δt）	时间步数（N_t）	阻力系数（a）	q_s梯度（γ）	蒸发速率（e）
2π	1e-3	100,000	1e-2	1.0	1.0
弛豫时间（τ）	强迫波数（k_f）	带宽（Δk）	能量输入率（ϵ）		
1e-2	3	2	0.1		
子集参数					
name	分辨率 (LxL)	振幅 (A)	波数 ($k_{x,\ y}$)	超扩散率 (κ)	子集大小 (N_d)
低分辨率	64x64	0	无调制	1e-8	2000
高分辨率-1	512x512	1	1	1e-16	2000
高分辨率-2	512x512	1	2	1e-16	2000
高分辨率-4	512x512	1	4	1e-16	2000
高分辨率-8	512×512	1	8	1e-16	2000
高分辨率-16	512x512	1	16	1e-16	2000

过程（Rissanen 等人（2023）；Hoogeboom 和 Salimans（2022）），但这些方法并未改变扩散模型的核心思想。

扩散建模通过参数化得分函数$\mathbf{s}_\theta(\mathbf{x},t) \approx s(\mathbf{x},t)$，并通过对适当选择的损失函数进行梯度下降来优化参数。在实践中，通常使用由以下内容定义的神经网络 f_θ 来表示得分函数 $\mathbf{s}_{\boldsymbol{\theta}}$。

$$\mathbf{s}_\theta(\mathbf{x},t) = \frac{f_\theta(\mathbf{x},t)}{\sigma(t)} \approx \mathbf{s}(\mathbf{x},t). \tag{B2}$$

这样做的好处是，神经网络的输出始终为 $O(1)$，这可以让神经网络更容易训练（而不必同时强迫它学习预先规定的 $\sigma(t)$ 依赖关系）。缺点是，理想情况下 $\sigma(t=0)$ 应为零，因此这会在 $\sigma(t=0)$ 处引入一个奇点。为避免这种情况，通常的做法（Song and Ermon 2020）是改为设定 $\sigma(t=0) = \sigma_{\min}$，如方程（B1）给出的 $\sigma^2(t)$ 表达式近似所示。

去噪得分匹配损失函数（Ho等人 2020；Song 等人 2021b）给出为

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)}\left[\lambda(t)^2\left(\frac{\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x},t)}{\sigma(t)} - \nabla_{\mathbf{x}}\log p(\mathbf{x}(t)|\mathbf{x}(0))\right)^2\right] \tag{B3a}$$

$$= \mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)}\left[\lambda(t)^2\left(\frac{\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x},t)}{\sigma(t)} - \frac{(\mathbf{x}(t)-\mathbf{x}(0))}{\sigma^2(t)}\right)^2\right] \tag{B3b}$$

$$= \mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)}\left[\frac{\lambda(t)^2}{\sigma(t)^2}\left(\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x},t) - \boldsymbol{\epsilon}\right)^2\right], \tag{B3c}$$

其中

$$\mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)} = \mathbb{E}_{t\sim U(0,1],\mathbf{x}(0)\sim p(\mathbf{x}(0)),\mathbf{x}(t)\sim p(\mathbf{x}(t)|\mathbf{x}(0))}, \tag{B4}$$

以及 $\mathbf{x}(t) = \mathbf{x}(0) + \sigma(t)\boldsymbol{\epsilon}$ 是时刻 t 的加噪图像，$\boldsymbol{\epsilon} \sim \mathbf{N}(\mathbf{0},\mathbf{1})$ 是一个高斯随机向量，而 $\lambda(t)$ 是一个加权因子，其取值设为 $\sigma(t)$（参见 Song 等人（2021a））。从方程（B3）的最后一步可以看出，得分匹配损失等价于使神经网络学习时刻 t 添加的噪声。请注意，尽管这涉及得分函数与条件分布对数梯度之间的 L_2 损失，但优化 $\mathrm{L}(\theta)$ 会得到对无条件分布真实得分函数的近似（Vincent（2011））。

我们对上述损失函数稍作修改，以监测空间均值及其周围变化所对应的具体损失值。由于方程（B3）中的 $\boldsymbol{\epsilon}$ 是随机高斯噪声，均值 $\bar{\boldsymbol{\epsilon}}$ 与

均值周围的变化 $\epsilon' = \epsilon - \bar{\epsilon}$ 相互独立。由于 $\mathbf{f}_\theta$ 旨在匹配 ϵ，我们预计网络训练充分后，它也会如此。在这种情况下，我们可以将损失改写为

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)}\left[\frac{\lambda(t)^2}{\sigma(t)^2}\left(\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x},t) - \boldsymbol{\epsilon}\right)^2\right] \tag{B5}$$

$$\approx \mathbb{E}_{t,\mathbf{x}(0),\mathbf{x}(t)}\left[\frac{\lambda(t)^2}{\sigma(t)^2}\left((\mathbf{f}'_{\boldsymbol{\theta}}(\mathbf{x},t) - \boldsymbol{\epsilon}')^2 + (\bar{\mathbf{f}}_{\boldsymbol{\theta}}(\mathbf{x},t) - \bar{\epsilon})^2\right)\right]. \tag{B6}$$

实际上，这不应影响训练过程；我们发现它主要有助于跟踪均值中的误差，而这些误差会导致色偏。

B2. 上下文扩散模型

在气候建模场景中，基于上下文信息（例如地形、水深地形或陆地表面属性）进行条件采样时，我们确实拥有成对的高分辨率和低分辨率图像数据点。将这些上下文图像记为 $\mathbf{x}_C$，我们可以从联合分布中获得样本

$$\mathbf{z}_\mathcal{S} = (\mathbf{x}_\mathcal{S}, \mathbf{x}_C) \sim p(\mathbf{z}_\mathcal{S}) \tag{B7a}$$

$$\mathbf{z}_\mathcal{T} = (\mathbf{x}_\mathcal{T}, \mathbf{x}_C) \sim p(\mathbf{z}_\mathcal{T}), \tag{B7b}$$

其中，与正文中一样，T 和 S 分别表示目标域和源域。随后，我们可以遵循 Song 等人（2021b）的方法进行条件采样。通过优化损失函数

$$\mathcal{L}(\boldsymbol{\theta}) = \mathbb{E}_{t,\mathbf{z}(0),\mathbf{z}(t)}\left[\lambda(t)^2\left(\frac{\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{z},t)}{\sigma(t)} - \nabla_\mathbf{x}\log p(\mathbf{z}(t)|\mathbf{z}(0))\right)^2\right], \tag{B8}$$

其中

$$\mathbb{E}_{t,\mathbf{z}(0),\mathbf{z}(t)} = \mathbb{E}_{t\sim U(0,1],\mathbf{z}(0)\sim p(\mathbf{z}(0)),\mathbf{z}(t)\sim p(\mathbf{z}(t)|\mathbf{z}(0))}, \tag{B9}$$

并且 $\mathbf{z}(t) = (\mathbf{x}(t), \mathbf{x}_C(t))$ 是一个元组，其中包含流体流动或气候模型的状态 $\mathbf{x}(t)$ 及其对应的上下文信息 $\mathbf{x}_C(t)$。我们选择不对上下文变量添加噪声，因此

$$p(\mathbf{z}(t)|\mathbf{z}(0)) = p(\mathbf{x}(t)|\mathbf{x}(0))\delta(\mathbf{x}_C(t) - \mathbf{x}_C(0)), \tag{B10}$$

因此，得分函数之间可以建立如下关系

$$\nabla_\mathbf{x}\log p(\mathbf{z}(t)|\mathbf{z}(0)) = \nabla_\mathbf{x}\log p(\mathbf{x}(t)|\mathbf{x}(0)). \tag{B11}$$

在这种上下文设置中，由此得到的得分函数是一个已知函数，就像无条件扩散模型的情况一样。正如 Batzolis 等人（2021）所示，优化这个损失函数等价于学习一个函数 $\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{z},t) = \nabla_\mathbf{x}\log p(\mathbf{x}(t)|\mathbf{x}_C)$，也就是说，该函数表示条件得分。在实现中，我们通过将 $\mathbf{x}_C$ 作为扩散模型输入的附加通道来实现这一点。关于架构的更多讨论见 C1 节。

附录 C

C1. 网络架构

我们的得分网络以 U-Net（Ron- neberger 等，2015）为基础，它将两个输入（$\mathbf{X}$，一个大小为 (N, N, C_{in}, B) 的张量，以及 t，一个大小为 $(B,)$ 的张量）映射为单个输出 $\mathbf{Y}$，一个大小为 (N, N, C_{out}, B) 的张量。也就是说，U-Net 返回

$$\mathbf{Y} = \mathcal{U}(\mathbf{X}, t; \boldsymbol{\theta}), \tag{C1}$$

其中，U 表示参数为 $\boldsymbol{\theta}$ 的 U-Net，具体说明见下文。

第一个输入 $\mathbf{X}$ 包含一批图像，第二个t 是一批时间； B 是批次大小。输入或输出的任一通道都是一幅大小为 (N, N) 的图像；输入通道数为 C_{in}，输出通道数为 C_{out}。对于我们的数据集，输入图像包含两个带噪通道：流体涡度和过饱和度示踪物浓度。计入上下文信息后，我们有 $C_{in} = 3, C_{out} = 2$。

在我们的默认配置中，U-Net 包含五个不同的部分。第一部分是初始提升层，也就是一个卷积层，它保留 $\mathbf{X}$ 的空间维度，但将通道数从 C_{in} 增加到 32。随后是三个下采样（卷积）层，它们将空间维度缩小为原来的 1/2，并将通道数增加 2 倍。这些变换后的数据经过八个残差块，这些残差块保留变换后数据的维度（He et al. 2016）。接着是三个上采样层，由最近邻上采样后接卷积组成；它们在减少通道数的同时增加空间维度，与下采样层相对应。最后，投影层将通道数减少到 C_{out}。我们使用 3x3 卷积核、组归一化（Wu and He 2018）以及 swish 函数作为非线性函数（Ramachandran et al. 2017）。

首先使用随机傅里叶投影对时间变量进行嵌入（Tancik 等人，2020）。随后，在每个上采样层和下采样层中，使用一个全连接网络对嵌入后的时间进行变换，然后将其加到上采样或下采样后的图像上。接着对求和结果进行组归一化，并按照 Song 等人（2021b）和 Ho等人（2020）的做法，使用 swish 函数处理，然后传递到下一层。

C2. 对 U-Net 的修改：均值旁路网络

我们对 C1 节中介绍的神经网络架构进行了修改，加入了旁路连接。输入批次被拆分为两部分：一部分包含空间变化（实际上是原始图像减去按通道和批次计算的空间均值），另一部分是每个通道和批次成员的空间均值。空间均值被输入旁路网络，而具有空间变化的部分则输入 U-Net。在最后一层，去除 U-Net 输出的空间均值后，将其输出与旁路网络的输出相加。这样，我们就有了一个完全独立的网络，分别处理空间均值和相对于均值的空间变化。结合我们对损失函数的选择，这还带来了一个额外优点：均值预测和空间变化预测成为完全独立的任务。

具体而言，我们按通道¯和批次成员计算输入$\mathbf{X}$的空间均值，记为$\bar{\mathbf{X}}$。相对于均值的空间¯变化记为$\mathbf{X}' = \mathbf{X} - \bar{\mathbf{X}}$，并按照 C1 节中的讨论由 U-Net 处理，以生成输出$\mathrm{U}(\mathbf{X}', t;\, \boldsymbol{\theta})$。然后，我们从该输出中减去空间均值（按通道和批次成员计算），即¯生成一个张量$\mathbf{Y}' = \mathrm{U}(\mathbf{X}', t;\, \boldsymbol{\theta}) - \bar{\mathrm{U}}(\mathbf{X}', t\boldsymbol{\theta})$，其¯空间均值为零。另一个函数 M $(\mathbf{X}, t;\, \phi)$使用¯可训练参数 ϕ 对$\mathbf{X}$ 和 t进行运算，并返回一个与$\mathbf{X}$大小相同的张量，记为$\mathbf{Y}$。在最后一层，我们将这些独立组件的输出组合起来，生成最终输出$\mathbf{Y}$，其形式为

$$\mathbf{Y} = \mathcal{U}(\mathbf{X}', t;\theta) - \bar{\mathcal{U}}(\mathbf{X}', t;\theta) + \mathcal{M}(\bar{\mathbf{X}}, t;\phi) = \mathbf{Y}' + \bar{\mathbf{Y}}. \quad \text{(C2)}$$

网络 M 由一个三层全连接前馈网络组成，包含两个线性变换，之后接归一化和非线性激活函数，最后再接一个线性变换，且不使用激活或归一化。嵌入的时间采用与 U 完全相同的处理方式：先经过线性变换，再与变换后的输入相加，然后进行归一化和激活。

需要注意的是，由于这一点，我们所实现的解决方案并未利用均值周围的空间变化与均值之间的相关性。如果输入的空间变化有助于预测得分函数的空间均值，或者反过来，那么我们的预测就无法利用这些信息。通过有限的测试，我们发现，在训练过程中允许 U 访问完整的输入 $\mathbf{X}$ 后，经过相同训练轮数的训练，其性能略差。为了将这些相关性纳入考虑，还需要进行进一步研究。

C3. 模型训练

我们在配置得分匹配去噪扩¯散模型的优化器时，遵循 Song 等人（2021b）和 Ho等人（2020）的建议。我们使用 Adam 优化器，学习率为 $\lambda_0 = 2\text{e-}4$，$\epsilon = 1\text{e-}8$，$\beta_1 = 0.9$，$\beta_2 = 0.999$。我们采用梯度范数裁剪，将其值设为 1.0。此外，我们还采用线性预热调度，使学习率在 5000 次梯度更新期间从 0 逐步增至 λ_0。所有运行的批大小均为 4。通常，我们训练 125 轮。在测试中，我们发现训练时间过长会导致某种过拟合，因此在残差层中使用了丢弃法，概率为 0.5，以帮助缓解这一问题。

关于原始图像的预处理，我们采用以下方法。首先，将每个数据样本分解为一个恒定均值图像和一个表示均值偏差的图像。对于数据的这两个组成部分，我们分别进行独立的最小-最大缩放，使得（所有预处理后数据中的）最小像素值为 -1，最大像素值为 1。然后，将这两个部分重新相加。由此得到的数据集不再具有恰好为 ± 1 的最小和最大像素值，但由于均值的最大值和最小值不一定与具有最大和最小空间偏差的图像相关联，因此像素值的分布仍主要集中在 $[-1, 1]$ 范围内（最坏情况下，则处于 $[-2, 2]$ 范围内）。

我们采用这一步预处理，主要是因为总涡量守恒，因此总涡量的分布是狄拉克δ函数。浮点误差会使其变成方差极小的高斯分布。接着，我们的预处理步骤会将其转化为宽得多的分布，从而更易于学习。不过，我们认为这通常都有益处，因为均值由一个独立的神经网络处理；这类似于按照惯例对该网络的输入进行预处理。

C4. 样本生成

为了生成此处展示的所有结果，我们使用具有固定时间步长的 Euler-Maruyama（EM）方法求解随机微分方程。对于随机微分方程

$$\mathrm{d}x = f(x,t)\mathrm{d}t + g(x,t)\mathrm{d}W, \quad \text{(C3)}$$

更新规则如下：

$$x(t+\Delta t) = x(t) + f(x,t)\Delta t + g(x,t)\eta\sqrt{\Delta t}, \quad \text{(C4)}$$

其中 $\eta \sim \mathrm{N}(0,1)$。所有模拟均使用 0.002 的固定时间步长，从$t = \epsilon = 1\text{e-}5$ 到 $t = 1$ 共对应 500 个时间步。图像生成并非本项目的主要计算成本，因此我们没有探究更改时间步进算法或时间步长的影响。测试其他时间步进方案是该领域一个活跃的研究方向。

附录 F

方位角平均功率谱密度

对于输入数据中的每个通道，我们都有一幅尺寸为 $N \times N$ 的二维图像。我们对该图像进行离散傅里叶变换，

$$\tilde{I}(k_x, k_y) = \sum_{x=-N/2}^{N/2-1} \sum_{y=-N/2}^{N/2-1} I(x,y) \exp\left[-i2\pi/N(k_x x + k_y y)\right]. \tag{F1}$$

波数（k_x，k_y）对应的功率谱为

$$\mathrm{PS}(k_x, k_y) = \frac{1}{N^4} \tilde{I}(k_x, k_y) \tilde{I}^*(k_x, k_y). \tag{F2}$$

这可以通过除以波数空间中的面积，将其转换为功率谱密度 $\mathrm{PSD}(k_x, k_y)$（Youngworth 等人，2005）。我们可以将其转换为极坐标 (k, ϕ)，其中 $k = \sqrt{k_x^2 + k_y^2}$。对于各向同性流，不同流场区域中 $|\tilde{I}(k, \phi)|$ 的期望值与 ϕ 无关。这意味着，对方位角 ϕ 进行积分不会造成信息损失（在期望意义上）。我们可以将方位角平均功率谱密度写为

$$\mathrm{PSD}(k) = \frac{1}{N^4} \frac{\int_0^{2\pi} \int_k^{k+1} \tilde{I}^* \tilde{I} k' dk' d\phi}{\int_0^{2\pi} \int_k^{k+1} k' dk' d\phi} \tag{F3}$$

$$\approx \frac{1}{N^4} \frac{\sum_{k_x} \sum_{k_y} \tilde{I}^* \tilde{I} \Theta(k^2 \le k_x^2 + k_y^2 < (k+1)^2)}{\sum_{k_x} \sum_{k_y} \Theta(k^2 \le k_x^2 + k_y^2 < (k+1)^2)}, \tag{F4}$$

其中，$\Theta(\text{condition})$ 是一个函数，当条件为真时返回 1，否则返回 0。对于具有优先方向或非均匀性的流场图像，该度量的参考价值会降低；在这种情况下，二维傅里叶变换后的图像本身可能更有用。

我们的算法 1 需要知道白噪声的功率谱。可以证明，如果 $I(x,y) \sim \mathrm{N}(0, \sigma^2)$，那么当 k_x 或 k_y 大于零时，$|\tilde{I}(k_x, k_y)^2| \sim \mathrm{Exp}[1/(\sigma^2 N^2)]$。其期望值为 $\sigma^2 N^2$。将其代入方程 (F3)，可知高斯白噪声的 $\mathrm{PSD}(k)$ 与波数 k 无关，其期望值为 σ^2/N^2（当 $k > 0$[2] 时）。

[2]当 $k_x = k_y = 0$，$|\tilde{I}(0,\ 0)^2|$ 并非来自指数分布。相反，$|\tilde{I}(0,\ 0)^2|/(\sigma^2 N^2) \sim \chi_1^2$ 服从卡方分布。我们在计算 PSD 之前减去均值，因此 $|\tilde{I}(0,\ 0)^2| \approx 0$。

参考文献

Abatzoglou, J. T. 和 T. J. Brown，2012：适用于野火应用的统计降尺度方法比较。国际气候学杂志，**32 (5)**，772–780。

Anderson, B. D.，1982：反向时间扩散方程模型。随机过程及其应用，**12 (3)**，313–326。Ballard, T. 和 G. Erinjippurath，2022：用于气候模型偏差校正和超分辨率的对比学习。arXiv:2211. 07555。

Batzolis, G., J. Stanczuk, C.-B. Schönlieb, and C. Etmann, 2021：基于得分的扩散模型条件图像生成。arXiv:2111.13606。

Brown, B. C. A., A. L. Caterini, B. L. Ross, J. C. Cresswell, and G. Loaiza-Ganem, 2023：验证图像数据的流形并集假设。arXiv:2207.02862。

Chetrite, R.，P. 穆拉托雷-Ginanneschi 和 K. 施维格，2021：E. 薛定谔1931年的论文《论自然定律的逆转》["论自然定律的逆转"，《普鲁士科学院会议报告》，物理数学类，8 N9 144–153]。《欧洲物理学杂志H辑》，**46**，1–29。

Choi, J.，J. Lee, C. Shin, S. Kim, H. Kim, and S. Yoon, 2022：Perception 优先训练扩散模型。*2022 年 IEEE/CVF 计算机视觉与模式识别会议（CVPR）*，IEEE 计算机学会，11 462–11471。

Choi, Y., M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, 2018：StarGAN：用于多域图像到图像转换的统一生成对抗网络。*IEEE 国际计算机视觉会议论文集*，8789–8797。

Dhariwal, P. 和 A. Nichol，2021：《扩散模型在图像合成方面胜过生成对抗网络》。神经信息处理系统进展，Curran Associates, Inc.，第 34 卷，8780 8794–。

Fowler, H. J., S. Blenkinsop 和 C. Tebaldi，2007：将气候变化建模与影响研究联系起来：水文建模降尺度技术的最新进展。国际气候学杂志，**27 (12)**，1547–1578。

Giannone, G.、D. Nielsen 和 O. Winther，2022：少样本扩散模型。arXiv:2205.15463。

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville 和 Y. Bengio，2020：生成对抗网络。*Commun. ACM*，**63 (11)**，139–144。

Groenke, B., L. Madaus, 和 C. Monteleoni, 2021：ClimAlign：通过归一化流场无监督统计降尺度气候变量。第十届气候信息学国际会议论文集，计算机协会，美国纽约州纽约市，60–66，CI2020。

Gutmann, E.，T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff 和 R. M. Rasmussen，2014：美国水资源评估中使用的统计降尺度方法比较。*Water Resour.*，**50 (9)**，7167–7186。

Hagemann, S., C. Chen, J. O. Haerter, J. Heinke, D. Gerten, 和 C. Piani, 2011：统计偏差订正对基于三个全球气候模式和两个水文模型获得的预估水文变化的影响。《水文气象学杂志》，**12 (4)**，556–578。

Harris, L., A. T. McRae, M. Chantry, P. D. Dueben, 和 T. N. Palmer, 2022：一种用于降水预报随机降尺度的生成式深度学习方法。《地球系统高级建模杂志》，**14 (10)**，e2022MS003120。

He, K., X. Zhang, S. Ren, and J. Sun, 2016：用于图像识别的深度残差学习。*IEEE*计算机视觉与模式识别会议论文集，770–778。

Ho, J., A. Jain, and P. Abbeel, 2020：去噪扩散概率模型。神经信息处理系统进展，Curran Associates, Inc.，第33卷，6840–6851。

Ho, J., and T. Salimans, 2022：无分类器扩散引导。arXiv:2207.12598。

Ho, J., T. Salimans, A. Gritsenko, W. Chan, M. Norouzi, and D. J. Fleet, 2022a：视频扩散模型。arXiv:2204.03458。

Ho, J.，及合著者，2022b：Imagen Video：使用扩散模型生成高清视频。arXiv:2210.02303。

Hoogeboom, E., and T. Salimans, 2022：模糊扩散模型。arXiv:2209.05557。

Hwang, S. 和 W. D. Graham，2014：评估将 GCM 日降水输出进行统计降尺度以模拟区域径流的替代方法。美国水资源协会期刊，**50 (4)**，1010–1032。

Hyvärinen, A.，以及 P. Dayan，2005：非……的估计 - 通过分数匹配对归一化统计模型进行建模。《机器学习研究杂志》，**6 (4)**。

Kingma, D. P., M. Welling, 和合著者，2019年：变分自编码器简介。《机器学习基础与趋势》，**12（4）**，307– 392。

Kong, Z., W. Ping, J. Huang, K. Zhao 和 B. Catanzaro，2021：DiffWave：一种用于音频合成的多功能扩散模型。arXiv: 2009.09761。

Luo, C.，2022：理解扩散模型：统一视角 . arXiv:2208.11970。

Maurer, E. P. 和 H. G. Hidalgo，2008年：每日与每月大尺度气候数据的效用：两种统计降尺度方法的比较。水文与地球系统科学，**12（2）**，551–563。

Maurer, E. P., 2014：偏差校正可以改变气候模型模拟的降水变化，而不会对集合平均产生不利影响。水文与地球系统科学，**18（3）**，915–925。

Meng, C., Y. He, Y. Song, J. Song, J. Wu, J.-Y. Zhu, and S. Ermon, 2022：SDEdit：基于随机微分方程的引导式图像合成与编辑。arXiv：2108.01073。

O'Gorman, P. A., and T. Schneider, 2006：均匀湍流中水分输运与凝结运动学的随机模型。大气科学杂志，**63 (11)**，2992–3005。

Pan, B., G. J. Anderson, A. Goncalves, D. D. Lucas, C. J. Bonfils, J. Lee, Y. Tian, and H.-Y. Ma, 2021：学习纠正气候预测偏差。《地球系统高级建模杂志》，**13（10）**，e2021MS002 509。

Panofsky, H. A., G. W. Brier, and W. H. Best, 1958：统计学在气象学中的一些应用。

Papamakarios, G., E. Nalisnick, D. J. Rezende, S. Mohamed 和 B. Lakshminarayanan，2021：用于概率建模和推断的归一化流。《机器学习研究杂志》，**22 (1)**，2617–2680。

Park, T.、A. A. Efros、R. Zhang 和 J.-Y. Zhu，2020：用于非配对图像到图像转换的对比学习。计算机视觉—*ECCV 2020*：第*16*届欧洲会议，英国格拉斯哥，*2020*年*8*月*23*日至*28*日，论文集，第*IX*卷 *16*，施普林格国际出版，319–345。Phillips, A.、T. Seror、M. Hutchinson、V. D. Bortoli、A. Doucet 和 E. Mathieu，2022：谱扩散过程。arXiv:2209.14125。Pierce, D. W.、D. R. Cayan 和 B. L. Thrasher，2014：使用局地构建类比（LOCA）进行统计降尺度。《水文气象学杂志》，**15 (6)**，2558–2585。

Pierce, D. W. 和合著者，2013：使用统计和动力降尺度对加利福尼亚州未来温度和降水变化进行概率估计。气候动力学，**40**，839–856。

Price, I., 和 S. Rasp，2022：《利用深度生成模型提高降水预报的准确性和分辨率》。第*25*届人工智能与统计国际会议论文集，PMLR，机器学习研究论文集，第151卷，10 555–10 571。

Ramachandran, P., B.Zoph, and Q.V. Le, 2017：搜索激活函数。arXiv:1710.05941.

Rissanen, S., M. Heinonen, 和 A. Solin，2023：具有逆热耗散的生成建模。arXiv:2206.13397.

Ronneberger, O., P. Fischer, and T. Brox, 2015：U-Net：用于生物医学图像分割的卷积网络。医学图像计算与计算机辅助介入——*MICCAI 2015*：第*18*届国际会议，慕尼黑，德国，*2015*年*10*月*5*日至*9*日，论文集，第*III*部分*18*，施普林格国际出版，234–241。

Saharia, C., W. Chan, H. Chang, C. Lee, J. Ho, T. Salimans, D. Fleet, and M. Norouzi, 2022：Palette：图像到图像扩散模型。*ACM SIGGRAPH 2022*会议论文集，计算机协会，美国纽约州纽约市，SIGGRAPH '22。

Saharia, C., J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, 2023：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**，4713–4726。

Salathé Jr, E. P., P. W. Mote, and M. W. Wiley, 2007：美国太平洋西北地区气候变化对水文学影响评估中的情景选择与降尺度方法综述。国际气候学杂志，**27 (12)**，1611–1621。

Sohl-迪克斯坦，J.，E. A. Weiss, N. Maheswaranathan, 和 S. Ganguli，2015：使用非平衡热力学进行深度无监督学习。JMLR.org，2256–2265，ICML'15。

Song, Y., C. Durkan, I. Murray, 和 S. Ermon，2021a：基于得分函数的扩散模型的最大似然训练。神经信息处理系统进展，CurranAssociates,Inc.，第34卷，1415–1428。

Song, Y.，以及 S. Ermon，2019：通过估计数据分布的梯度进行生成建模。神经信息处理系统进展，Curran Associates, Inc.，第 32 卷。

Song, Y.，以及 S. Ermon，2020：改进生成模型得分函数训练技术 - 。神经信息处理系统进展，Curran Associates, Inc.，第 33 卷，12438–12448。

Song, Y., J. Shen, 和 S. Ermon，2022Xing 利用基于评分的生成模型求解医学影像中的逆问题。arXiv2111.08005。

Song, Y., J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole，2021b：通过随机微分方程实现基于分数的生成建模。arXiv:2011.13456。Su, X., J. Song, C. Meng, and S. Ermon，2023：用于图像到图像转换的双重扩散隐式桥。arXiv:2203.08382。Tancik, M., and Coauthors，2020：傅里叶特征使网络能够学习低维域中的高频函数。神经信息处理系统进展，Curran Associates, Inc.，第33卷，7537–7547。Tashiro, Y., J. Song, Y. Song, and S. Ermon，2021：CSDI：用于概率时间序列插补的条件分数扩散模型。神经信息处理系统进展，Curran Associates, Inc.，第34卷，24 804–24 816。Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy，2012：利用分位数映射对气候模型模拟的日极端气温进行偏差校正。水文与地球系统科学，**16 (9)**，3309–3314。Vincent, P.，2011：分数匹配与去噪自编码器之间的联系。*Neural Comput.*，**23 (7)**，1661–1674。Wang, X., R. Girshick, A. Gupta, and K. He，2018：非局部神经网络。*Proc.IEEEInt.Conf.Comput.Vis.*，7794–7803。Wang, Z., J. Chen, and S. C. H. Hoi，2021：图像超分辨率深度学习综述。*IEEE*模式分析与机器智能汇刊，**43 (10)**，3365–3387。Wood, A. W., L. R. Leung, V. Sridhar, and D. Lettenmaier，2004：动力和统计方法对气候模型输出降尺度的水文影响。*Clim. change*，**62 (1-3)**，189–216。Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier，2002：美国东部长期试验性水文预报。*J. Geophys. Res. Atmos.*，**107 (D20)**，ACL–6。Wu, Y., and K. He，2018：组归一化。欧洲计算机视觉会议（*ECCV*）论文集，3–19。Youngworth, R. N., B. B. Gallagher, and B. L. Stamper，2005：功率谱密度（PSD）计算概述。*Opticalmanufacturing and testing VI*，**5869**，206–216。Zhu, J.-Y., T. Park, P. Isola, and A. A. Efros，2017：利用循环一致对抗网络进行非配对图像到图像转换。*Proc. IEEEInt.Conf.Comput.Vis.*，2242–2251。