

复现 DynaMo 方法应用于气候降尺度任务的研究

摘要

随着气候变化的影响日益加剧，气候降尺度作为气候研究中的一个关键任务，旨在将大尺度气候模型（GCMs）或区域气候模型（RCMs）的低分辨率输出转化为高分辨率的气候数据。传统的降尺度方法往往依赖大量标注数据，且难以充分利用气候数据的时空结构特征。本文基于 DynaMo 方法，实现了一种自监督学习框架，通过训练潜在表示并进行降尺度预测。DynaMo 利用时间序列动态建模，结合逆动态模型和前向动态模型，能够在没有外部标注的情况下学习气候系统的潜在动态变化，并提高气候降尺度任务的准确性。实验结果表明，DynaMo 在气候降尺度任务中，相较于传统方法，能够更高效地利用低分辨率数据，并提供更准确的高分辨率气候预测。

关键词：气候降尺度；DynaMo；自监督学习；时间序列建模

1 引言

在当今快速发展的科技时代，机器学习和人工智能技术在各个领域的应用日益广泛。特别是在机器人技术和自动化控制领域，模仿学习作为一种强大的工具，已经被证明能够有效地训练复杂的视觉-运动策略。然而，现有的模仿学习方法通常需要大量的专家演示来处理高维视觉观测，这在实际应用中往往面临着数据效率低下的问题。一个关键的原因是，当前的视觉表示方法要么是基于互联网等外部领域的预训练数据，要么是直接通过行为克隆目标进行训练，这两种方法在数据利用效率上都存在不足。

DynaMo [1] 作为一种新的自监督学习方法，旨在通过在特定领域内预训练视觉表示来提高视觉-运动控制的效率。这种方法的核心在于利用演示数据中的因果结构，联合学习编码器、逆动力学模型和前向动力学模型，从而在没有数据增强、对比采样或访问真实动作的情况下，预测下一帧的潜在空间表示。DynaMo 的一个显著特点是它不需要依赖于互联网数据集或跨身体数据集，而是专注于特定领域的的数据，这使得它在处理特定任务时能够更加高效地利用有限的专家演示数据。

ERA5 气候数据集作为全球气候数据的重要来源，提供了丰富的气象信息，包括温度、降水、风速等多种气象要素。这些数据不仅在气候研究和天气预报等领域有着广泛的应用，也为探索和验证新的机器学习方法提供了一个独特的实验平台。通过使用 ERA5 气候数据集复现 DynaMo 方法，我们不仅可以验证 DynaMo 方法在处理气候数据时的有效性，还可以探索其在气候数据分析和预测中的潜在应用，为未来的相关研究提供新的思路和方法。

本研究的目的是复现 DynaMo 方法，并将其训练过的编码器应用到 ERA5 气候数据集进行降尺度预测，以验证其在气象领域内的适用性和有效性。

2 相关工作

2.1 自监督学习

自监督学习是一种强大的学习范式，它通过利用数据本身的结构和模式来学习有用的特征表示，而不需要外部的监督信号。在视觉领域，自监督学习方法已经取得了显著的成果，例如通过预测图像的缺失部分、重建图像的原始版本或学习图像序列中的时间关系来学习视觉特征。这些方法在处理高维视觉数据时表现出色，尤其是在没有大量标注数据的情况下。

自监督学习在气象和气候领域的应用主要体现在通过学习历史天气数据中的潜在趋势和模式来预测未来的天气条件。例如，FourCastNet [2] 是一个基于视觉 Transformer 和自适应傅里叶神经网络算子 [3] (AFNO) 的气候预训练基础模型，它通过自监督预训练和自回归微调来实现高分辨率的预测和快速推理。PANGU-WEATHER [4] 则是一个基于 3D 地球特定 Transformer 的数据驱动模型，使用自监督学习来训练模型，使其能够从大量的气候数据中学习有用的特征表示，从而提高预测的准确性和可靠性。W-MAE [5] 是一个使用掩码自编码器 (Masked Auto-Encoder, MAE) 进行自监督训练的全球天气预测模型。它通过在训练过程中随机掩码部分数据来学习数据的潜在表示。CLIMAX [6] 是基于 Transformer 的气候预训练基础模型。它通过自监督预训练来学习气候数据中的潜在模式和关系，然后通过微调来适应特定的下游任务。这些方法展示了自监督学习在气象和气候数据理解中的潜力，尤其是在提高模型的泛化能力和预测精度方面。通过利用数据本身的结构和模式，这些方法能够有效地学习有用的特征表示，从而提高气候预测的准确性和可靠性。

2.2 气候降尺度

气候模型通常以全球或区域尺度进行模拟，其输出数据的空间分辨率相对较低。然而，许多实际应用需要更高分辨率的气候数据，以便更准确地理解和预测特定地区的气候条件。例如，水资源管理者需要了解特定流域的降水模式，农业规划者需要预测特定地区的温度和降水变化，城市规划者需要评估城市地区的热岛效应。因此，气候降尺度成为了一个关键的研究领域，旨在将低分辨率的气候模型输出转换为高分辨率的气候数据，以满足这些实际应用的需求。

气候降尺度是将全球气候模型 (GCMs) 或区域气候模型 (RCMs) 的低分辨率输出转换为更高分辨率的本地或区域尺度数据的过程。通常面临着诸多挑战，包括数据的不一致性、空间和时间分辨率的差异、以及模型的泛化能力等。全球气候模型和区域气候模型通常基于不同的物理过程和假设，这可能导致它们在不同地区的输出数据存在差异。此外，气候数据的空间和时间分辨率可能不一致，这使得降尺度模型需要能够处理不同分辨率的数据。降尺度模型同时需要能够泛化到不同的气候条件和区域，这要求模型具有足够的灵活性和鲁棒性。

为了解决降尺度这一关键气候中的问题，先前有许多研究工作也进行了探索。物理约束的深度学习方法通过将物理知识融入深度学习模型中，以提高降尺度模型的准确性和可靠性。MeshfreeFlowNet [7] 采用物理信息模型，将偏微分方程 (PDEs) 作为正则化项纳入损失函数，实现了时空降尺度。这种方法通过结合物理模型和机器学习技术，提高了降尺度模型的性能和可靠性。具体来说，他们使用了卷积神经网络 (CNN) 来学习气候数据的特征表示，并通过物理信息模型来约束模型的输出，从而提高了模型的泛化能力和预测精度。这种方法的优

点是它可以结合物理模型和机器学习技术，提高降尺度模型的性能和可靠性，但其缺点是需要大量的训练数据来训练模型。

不确定性建模的引入还允许在基于深度学习的模型中降低增益，显著提高效率和重建分辨率。ResDiff [8] 采用两步扩散模型方法，首先使用 U-Net 回归预测均值，然后使用扩散模型预测残差，从而实现千米尺度的大气降尺度。这种方法通过结合不同的机器学习技术，提高了降尺度模型的性能和可靠性。具体来说，他们使用了 U-Net 回归模型来预测气候数据的均值，并通过扩散模型来预测残差，从而提高了模型的泛化能力和预测精度。这种方法的优点是它可以结合不同的机器学习技术，提高降尺度模型的性能和可靠性，但其缺点是需要大量的计算资源来训练模型。

3 本文方法

3.1 Dynamo 概述

Dynamo 框架的关键之处在于通过对演示观察的动态建模来学习控制的良好视觉表示，而不需要增强、对比采样或访问地面真实动作。给定一系列原始视觉观测值 (o_1, \dots, o_T) ，联合训练编码器 $f_\theta : o_t \rightarrow s_t$ ，潜在的逆动力学模型 $q(z_{t:t+h-1} | s_{t:t+h})$ ，以及正向动力学模型 $p(\hat{s}_{t+1:t+h} | s_{t:t+h-1}, z_{t:t+h-1})$ 。模型将行为建模成未观测到的潜变量，并对所有模型进行端到端训练，在前向动力学预测上采用一致性损失。该架构如图1所示。

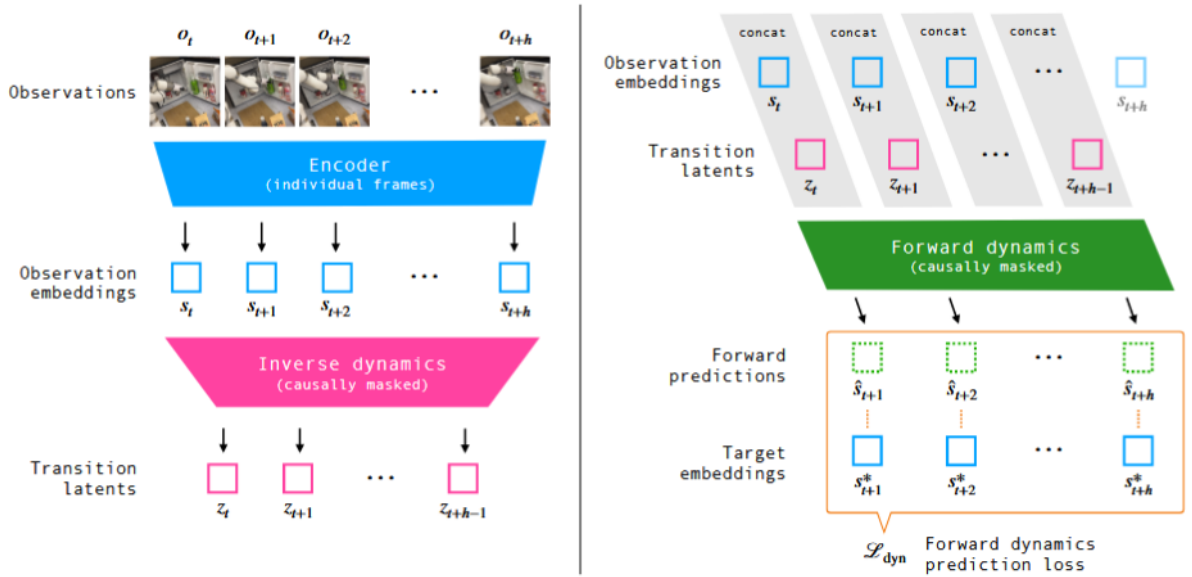


图 1. Dynamo 的架构。DynaMo 联合学习图像编码器、逆动力学模型和具有正向动力学预测损失的正向动力学模型。

3.2 视觉编码器

视觉编码器的作用是将高维的视觉图像映射到低维的潜在空间表示。从高维的原始视觉观测中提取低维嵌入 s_t ，将复杂的原始图像转化为紧凑的特征表示，为后续的动态建模模块提供输入。这种嵌入包含了图像中的重要视觉信息，并通过训练对任务关键特征敏感。这一过程通过残差网络 ResNet 实现，对于每个输入图像，编码器输出一个固定长度的向量，称为

潜在表示，该向量包含了图像中与任务相关的视觉信息，如物体的位置、大小和颜色等，为后续的动力学模型提供输入。

3.3 逆动力学模型 (Inverse Dynamics Model)

逆动力学模型的目标是从连续的视觉嵌入中推断出潜在的动态变化。具体来说，逆动态模型基于当前和下一时刻的嵌入 S_t 和 S_{t+1} ，推断出一个隐式的动态表示 Z_t ，表示从 S_t 到 S_{t+1} 之间的潜在变化。输入为编码器输出的当前状态表示和下一状态表示，输出为潜在的表示 Z_t 。这个潜在的动作表示不需要与真实的动作完全对应，而是用于捕捉动作的潜在结构和模式。

将这一部分应用到气候降尺度领域，可以从气候的时间序列数据中，基于状态转移推断出潜在的“动作”或“动态”。这种方式使得模型能够捕捉两个时刻之间的变化，而无需显式的动作标签。此外，气候系统的变化通常是高度非线性的。逆动态模型通过其非线性表示能力，有助于揭示气候变量（如温度、降水）在时间序列中的复杂变化。

3.4 前向动力学模型 (Forward Dynamics Model)

前向动态模型的核心任务是利用当前时刻的状态 S_t 和从逆动态模型获得的潜在动态变量 Z_t ，预测下一时刻的状态嵌入 S_{t+1} 。其目标是模拟时间序列的演化过程，即从当前状态出发，学习如何生成下一个时刻的状态。

这一组件在 DynaMo 中的作用非常重要，它决定了模型如何从当前时刻的状态和潜在动态变量中生成未来的状态。这一过程对于捕捉时间序列中的因果关系至关重要，通过学习时间序列中潜在的动态变化，前向动态模型可以提高高分辨率预测的准确性，特别是在低分辨率气候数据不足时。

4 复现细节

4.1 与已有开源代码对比

使用了https://github.com/jeffacce/dynamo_ssl的开源代码作为基础，修改了代码中关于分布式训练的部分，只使用了单卡进行训练，并且在 300epoch 迭代训练后，使用最后一轮保存的 Encoder 模型参数，编写代码进行下游的气候降尺度训练。修改后的 Dynamo 训练代码主要部分如图2所示。


```

def train(self):
    if self.cfg.use_lr_scheduling:
        lr = self.adjust_lr()
        self.log_append("metrics", 1, {"lr": lr})
    self.ssl.adjust_beta(self.epoch, self.cfg.num_epochs)
    pbar = tqdm.tqdm(
        self.train_loader,
        desc=f"Training epoch {self.epoch}",
        # disable=not self.accelerator.is_main_process,
        ncols=80,
    )
    for data in pbar:
        obs, _, _ = data
        obs=obs.to(self.device)
        (
            obs_enc,
            obs_proj,
            ssl_loss,
            ssl_loss_components,
        ) = self.ssl.forward(obs)
        self.log_append("ssl_train", len(obs), ssl_loss_components)
        ssl_loss.backward(ssl_loss)
        # self.accelerator.backward(ssl_loss, retain_graph=True)
        if self.cfg.clip_grad_norm:
            nn_utils.clip_grad_norm(self.encoder.parameters(), self.cfg.clip_grad_norm)
            nn_utils.clip_grad_norm(self.projector.parameters(), self.cfg.clip_grad_norm)
            nn_utils.clip_grad_norm(self.ssl.parameters(), self.cfg.clip_grad_norm)

        self.encoder_optim.step()
        self.projector_optim.step()
        self.ssl.step()
        self.encoder_optim.zero_grad(set_to_none=True)
        self.projector_optim.zero_grad(set_to_none=True)

```

图 2. 修改后的 Dynamo 训练代码

4.2 实验环境搭建

本实验在 Python3.8.17 的环境下进行，其他依赖包通过代码中的 conda_env.yml 文件安装。

4.3 使用说明

代码中的 dynamo_ssl-main 和 Data_generation 两个文件夹通过 Dynamo 训练 Encoder 并保存模型参数，实验中总共进行 300 轮训练迭代。downscaling 文件夹中代码读取 Dynamo 模型训练好的 Encoder 作为下游降尺度预测的编码器进行下游任务的微调训练。

4.4 创新点

DynaMo 原本是用于机器人模仿学习的视觉-运动控制的领域内动态预训练，通过对机器人操作任务的视觉表征进行预训练，下游策略学习能够使用跨模拟和真实世界机器人任务的有限域内数据。如图3，首先使用表征学习利用 Dynamo 框架自监督学习到域内数据潜在表征学习，然后在下游任务中微调以适应特定任务。

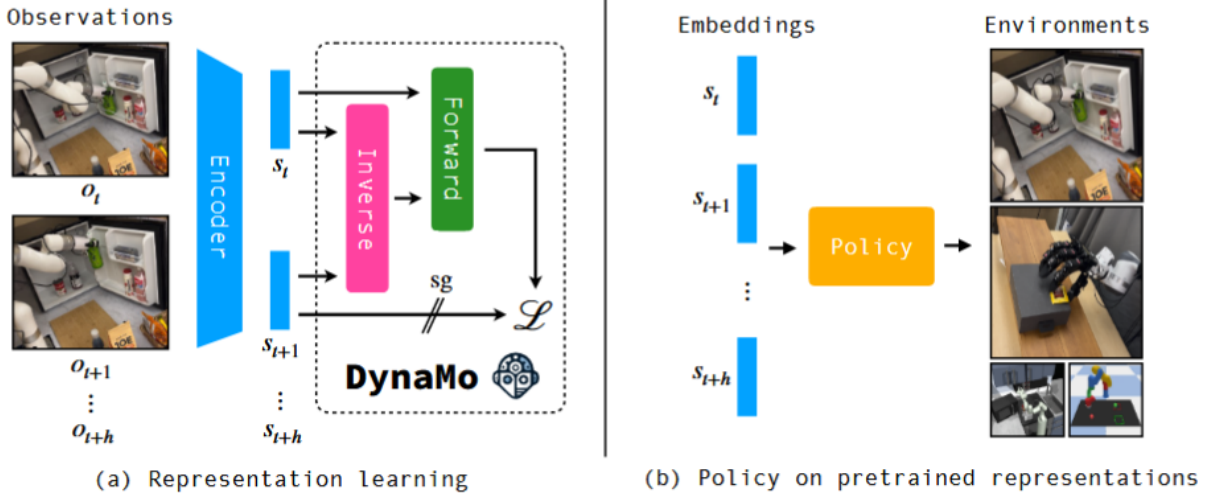


图 3. (a) 域内数据预训练；(b) 下游策略学习微调

将上述方法应用到 ERA5 气候数据集，由于气候数据也是时间序列，通过 DynaMo 的时间序列动态建模，隐式捕捉了气候系统中的物理连续性和因果性。将潜在动态表示作为中间层，将大尺度气候数据的复杂特征压缩到紧凑的嵌入 s_t 。这些嵌入并非直接用于某一个具体的气候任务，而是通过动态模型优化后，提供给不同的气候任务进行训练和预测。这种分离式的训练方法实现了跨任务学习，可以实现训练一个 Encoder 应用不同下游任务。

此外，高分辨率气候观测数据的稀缺性是降尺度任务中的关键瓶颈。DynaMo 利用自监督学习框架，通过大尺度数据时间序列的自监督优化，无需大量标注数据即可学习到强表达能力的潜在表示。自监督方法使得模型能够从已有的大尺度数据中挖掘时间序列间的因果和相关性，为高分辨率预测提供了强大的基础，突破了传统监督学习方法的限制，可以更高效地利用已有的大量低分辨率数据，缓解高分辨率观测数据稀缺的问题。

5 实验结果分析

将数据集换成 ERA5 数据，使用 Dynamo 框架预训练 Encoder，实验总共迭代 300 轮，可以在图4中观察到训练和评估损失的变化趋势。

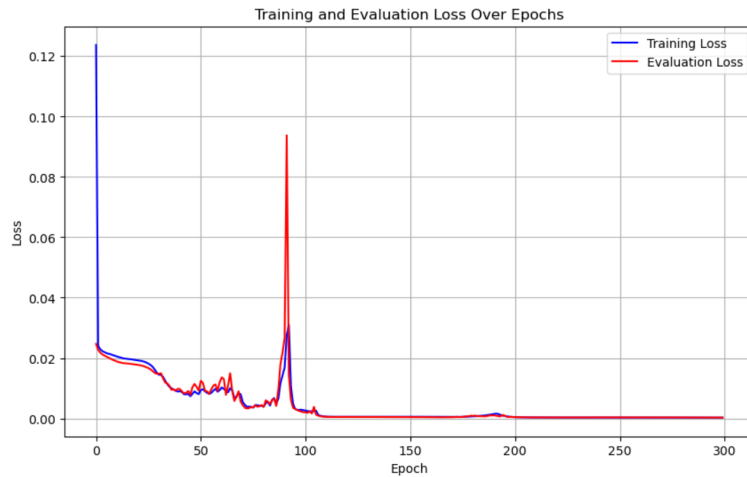


图 4. Dynamo 训练损失

可以发现在最初几个 epoch，训练损失和评估损失均处于较高水平，表明模型在初始阶段预测误差较大。然而，随着训练的进行，两条曲线均迅速下降，模型在早期阶段能够有效捕获数据的基础模式。在训练约 100 个 epoch 时，评估损失出现了一次显著的剧烈上升，同时训练损失也表现出轻微的波动。在异常峰值后，训练损失和评估损失均迅速回归到较低水平，并保持平稳。整体来看，损失曲线呈现出快速下降、稳定振荡以及最终收敛的典型特征。高峰波动阶段虽然显著，但最终模型能够很好地收敛，展现出较强的鲁棒性和泛化能力。

训练完成后使用 Encoder 模型参数对气候降尺度任务进行微调，训练结果与 DeepONet [9] 对比，结果如表1所示。

表 1. Dynamo 与 DeepONet 性能对比

Metric	Variable	DynaMo	DeepONet
MSE(↓)	Zonal wind speed at 10 meters (U10)	0.1499	0.1524
	Meridional wind speed at 10 meters (V10)	0.1921	0.1912
	Temperature at 2 meters (T2m)	0.0107	0.0108
	Temperature at 850 hPa (T850)	0.0073	0.0073
	Geopotential at 500 hPa (Z500)	0.0010	0.0010
	Total Variable Average	0.0722	0.0725
RMSE(↓)	Zonal wind speed at 10 meters (U10)	2.136	2.1535
	Meridional wind speed at 10 meters (V10)	2.0822	2.077
	Temperature at 2 meters (T2m)	2.2002	2.2087
	Temperature at 850 hPa (T850)	1.3397	1.3401
	Geopotential at 500 hPa (Z500)	108.1865	107.2406
	Total Variable Average	23.1889	23.0040
ACC(↑)	Zonal wind speed at 10 meters (U10)	0.8584	0.8622
	Meridional wind speed at 10 meters (V10)	0.8734	0.8720
	Temperature at 2 meters (T2m)	0.8908	0.8899
	Temperature at 850 hPa (T850)	0.9522	0.9542
	Geopotential at 500 hPa (Z500)	0.9943	0.9939
	Total Variable Average	0.9138	0.9145

从复现结果可以发现两种模型在不同气象变量的模拟上各有优势。总体来看，Dynamo 在均方误差方面表现更佳，尤其是在模拟 T2m 和 U10 等变量时，其误差平方平均值更低，显示出较好的预测精度。然而，DeepONet 在均方根误差和异常相关系数方面表现更优，在两者的平均总量中均表现更优，表明其在预测结果的稳定性和与真实数据的相关性方面更具优势，但同时 Dynamo 在个别变量如 T2m 上展现出其独特的优势。

6 总结与展望

本研究通过复现 DynaMo 方法，并将其应用于 ERA5 气候数据集，探索了其在气候降尺度任务中的潜力。与传统的降尺度方法不同，DynaMo 采用了自监督学习框架，通过训练潜在动态表示，无需大量标注数据，即可实现高效的降尺度预测。实验结果表明，DynaMo 在气候数据降尺度中具有显著优势，能够充分利用大尺度气候数据中的时空结构信息，提高降尺度结果的精度。

在应用过程中，DynaMo 模型成功捕捉了气候系统中的非线性动态变化，尤其是在温度、降水等气候变量的时序预测中表现出色。同时，逆动态模型和前向动态模型的联合训练，使得该方法能够有效地模拟气候数据的时间演化规律。通过潜在表示的学习，DynaMo 不仅提高了气候变量的降尺度精度，也在一定程度上突破了高分辨率气候观测数据稀缺的瓶颈。

然而，复现过程中也遇到了一些挑战，包括模型的收敛速度、训练数据的选择以及实验过程中对硬件资源的依赖等问题。未来的研究可以进一步优化模型架构，提升模型的计算效率，并尝试将该方法扩展到更复杂的气候模型和其他领域的应用。

参考文献

- [1] Zichen Jeff Cui, Hengkai Pan, Aadithya Iyer, Siddhant Haldar, and Lerrel Pinto. Dynamo: In-domain dynamics pretraining for visuo-motor control. *arXiv preprint arXiv:2409.12192*, 2024.
- [2] Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, et al. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*, 2022.
- [3] John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Anandkumar, and Bryan Catanzaro. Adaptive fourier neural operators: Efficient token mixers for transformers. *arXiv preprint arXiv:2111.13587*, 2021.
- [4] Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian. Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 619(7970):533–538, 2023.
- [5] Xin Man, Chenghong Zhang, Jin Feng, Changyu Li, and Jie Shao. W-mae: Pre-trained weather model with masked autoencoder for multi-variable weather forecasting. *arXiv preprint arXiv:2304.08754*, 2023.
- [6] Tung Nguyen, Johannes Brandstetter, Ashish Kapoor, Jayesh K Gupta, and Aditya Grover. Climax: A foundation model for weather and climate. *arXiv preprint arXiv:2301.10343*, 2023.
- [7] Soheil Esmailzadeh, Kamyar Azizzadenesheli, Karthik Kashinath, Mustafa Mustafa, Hamdi A Tchelepi, Philip Marcus, Mr Prabhat, Anima Anandkumar, et al. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–15. IEEE, 2020.
- [8] Morteza Mardani, Noah D Brenowitz, Yair Cohen, Jaideep Pathak, Chieh-Yu Chen, Cheng-Chin Liu, Arash Vahdat, Karthik Kashinath, Jan Kautz, and Mike Pritchard. Generative residual diffusion modeling for km-scale atmospheric downscaling. *CoRR*, 2023.
- [9] Lu Lu, Pengzhan Jin, and George Em Karniadakis. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193*, 2019.