arXiv:2302.03173v1 [physics.ao-ph] 7 Feb 2023

利用深度神经算子学习气候模型的偏差校正

Aniruddha Bora[*1]、**Khemraj Shukla**[*1]、**Shixuan Zhang**[2]、**Bryce Harrop**[2]、**Ruby Leung**[2] 和 **George Em Karniadakis**[1]

[1] 布朗大学应用数学系，霍普街170号，普罗维登斯，RI-02906 george karniadakis@brown.edu[2]太平洋西北国家实验室，巴特尔大道902号，里奇兰，WA 99354

摘要

气候建模中需要解析所有重要尺度的数值模拟，计算成本极高。因此，为了避免这一问题，通常会进行低分辨率模拟，随后利用再分析数据（ERA5）进行偏差校正，这一过程称为松弛校正。现有的松弛校正方法采用基于松弛的方法，处理低分辨率数据与 ERA5 数据之间的代数差异。本研究以基于深度算子网络（DeepONet）的代理模型取代偏差校正过程。DeepONet（深度算子网络）学习从松弛前的状态（一个泛函）到松弛趋势（另一个泛函）的映射。松弛趋势是维度极高的数据，尽管其中包含许多低能量模态。为此，我们将 DeepONet 与基于卷积的自动编码器—解码器（AED）架构结合，以便在较低维的潜在空间中高效地学习松弛趋势。我们使用 E3SMv2（能源百亿亿次地球系统模型）获得的松弛趋势对 DeepONet 模型的准确性进行测试，结果显示二者吻合良好。本研究的总体目标是在线部署 DeepONet 模型，以取代 E3SM循环中的松弛模块，从而提升效率和准确性。

引言

大气环流具有复杂的多尺度物理过程，因此给数值模拟带来了诸多挑战。这些模拟过程可能与真实状态不一致，且存在显著偏差，这些偏差被称为“偏差”。与观测到的大气状态相比，气候模型的空间分辨率有限，物理参数化也不准确，因此会产生这些偏差。低分辨率模拟对计算资源的要求较低，但其偏差限制了模拟的实用性。这种权衡会导致气候预测的不确定性增大。通常，这类偏差会通过后处理步骤消除，例如用于天气预报的模式输出统计（Glahn and Lowry，1972年）、用于季节预测的集合偏差校正（Arribas等，2011年；Stockdale等，1998年）等。在现有的在线偏差校正方法中，偏差校正基于数据驱动方法，其泛化能力较弱（Watt-Meyer等，2021年）。在本研究中，我们提出一种基于神经算子的代理模型，以学习偏差校正，并提升其泛化能力。具体而言，我们采用深度算子网络（DeepONet）（Lu等，2021年），并通过引入自编码器（Oommen等，2022年）扩展现有的神经网络架构。

方法

本节将讨论数据生成、训练 DeepONet 的方法，以及所提出代理模型的架构细节。

数据生成

能源百亿亿次地球系统模型 E3SM（E3SM 项目 2018）中物理量 X_m 随时间的演化由下式给出（Sun 等，2019）：

$$\frac{\partial X_m}{\partial t} = F(X_m), \tag{1}$$

其中， X_m 是状态变量。状态变量的松弛趋势为

$$\left(\frac{\partial X_m}{\partial t}\right)_{ndg} = \begin{cases} \frac{X_p(t_p) - X_m(t)}{\tau}, & t = t_i, \\ 0, & t \neq t_i, \end{cases} \tag{2}$$

其中， t_i 是 i^{th} 小时、 $ERA5$ 数据可用的时间。用于训练和测试 DeepONet 的数据通过执行 EAMv2 模拟生成，这些模拟会向参考数据集（即每 3 小时 ERA5 再分析数据（Hersbach 等，2020））进行松弛。该随时间变化的松弛趋势由模型 X_m 、参考状态 X_p 以及松弛时间尺度τ共同确定，如公式（2）所示。根据公式（2），每 30 分钟计算一次松弛趋势。在公式（2）中，通过将每 3 小时 ERA5 再分析数据线性插值到模型时间点，获得 X_p 。在本研究中，EAMv2 模拟仅针对 U（纬向风）和 V（经向风）进行松弛，我们使用了三组数据集，分别为：（i）松弛前数据；（ii）松弛趋势；以及（iii）ERA5 再分析数据集。这些数据集之间的关系表示为：

$$\left(\frac{\partial X_m}{\partial t}\right)_{\mathbf{cor}} = \left(\frac{\partial X_m}{\partial t}\right)_{\mathbf{uncor}} + \left(\frac{\partial X_m}{\partial t}\right)_{ndg}, \tag{3}$$

[*]这些作者贡献相同。版权所有 © 2022年，人工智能促进协会 (www.aaai.org)。保留所有权利。

其中，$(\partial X_m/\partial t)_{uncor}$ 和$(\partial X_m/\partial t)^{ndg}$ 分别表示推搡前状态和松弛趋势。$(\partial X_m/\partial t)_{cor}$ 表示状态变量 X_m的校正状态，该状态应接近 ERA5。按照构造方式，公式（3）旨在使模型的预测逐渐趋近于真实观测。

深度算子网络的架构

要学习用于预测松弛趋势的代理模型，需要学习一个从空间 $\mathcal{G}$ **U**（松弛前数据）到**V**（松弛趋势）的函数，该函数以时空坐标（x，y，t）为变量。DeepONet 是一种基于通用逼近定理（Chen 和 Chen（1995））的神经算子回归技术。DeepONet 的目标是学习从输入空间 **U**(η) 到输出空间 **V**(ζ) 的映射。传统的 DeepONet 架构首先由两个神经网络组成：分支网络（分支网络）和主干网络（主干网络）。分支网络和主干网络分别将样本 $u(\eta) \in$ **U**(η) 和 ζ 作为输入。DeepONet 的输出可表示为算子 $\mathcal{G}$：**U** $\rightarrow$ **V**，使得 $v(\zeta) = \mathcal{G}(u(\zeta))$，其中 $u(\zeta) \in$ **U** 以及 $v(\zeta) \in$**V**。松弛趋势包含许多低能量模态、较少的相干结构，并且具有非常高的维度，因此很难使用神经网络进行逼近。这是由于特征与标签数据之间的相关性非常低。Oommen等（Oommen 等（2022））已证明，可以利用编码器发现的潜在空间成功学习这类问题。在本研究中，我们使用两个自编码器来降低**U** 和**V** 的空间维度，从而在平滑的潜流形中表示松弛趋势，使其更易于学习。训练 DeepONet 所需的步骤如下

i. 第一步是训练两个自编码器：一个用于 DeepONet 输入，即松弛前变量（$\psi_{bf}(z,\ \bar{x},\ \bar{y},\ t)$），另一个用于 DeepONet 输出，即松弛趋势（$\psi_{ndg}(z,\ \bar{x},\ \bar{y},\ t)$）变量。两个自编码器均使用均方误差损失进行训练，如下所示：

$$\mathcal{L}_{AE1} = ||\psi_{bf}(z,\bar{x},\bar{y},t) - \psi_{bf}(z,\bar{x},\bar{y},t;\theta_{AE1})||_2,$$
$$\mathcal{L}_{AE2} = ||\psi_{ndg}(z,\bar{x},\bar{y},t) - \psi_{ndg}(z,\bar{x},\bar{y},t;\theta_{AE2})||_2,$$

其中，AE1 和 AE2 分别表示松弛前变量和松弛趋势变量对应的自编码器。此外，θ_{AE1} 和 θ_{AE2} 分别表示相应自编码器的可训练参数。

ii. 在完成编码器和解码器的训练与验证后，我们在潜在空间中学习 DeepONet，其特征和标签分别定义为 $\psi_{bf}(z,\ \bar{x},\ \bar{y},\ t)$ 和 $\psi_{ndg}(z \times ld,\ t)$。DeepONet 的输出 $\psi_{ndg\ don}(z \times ld,\ t)$ 通过最小化如下所示的均方误差函数，针对 $\psi_{ndg}(z \times ld,\ t)$ 进行训练。

$$\mathcal{L} = ||\hat{\psi}_{ndg_don}(z \times ld, t) - \hat{\psi}_{ndg}(z \times ld, t)||_2. \quad (4)$$

iii. 为了恢复 DeepONet 在原始空间中的输出，我们将其与解码器组合。

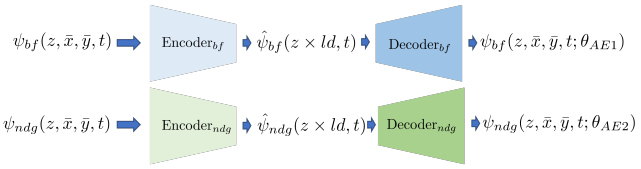


图1：自编码器整体结构的框图。这里，$\psi_{bf}(z,\bar{x},\bar{y},t)$ 和 $\psi_{ndg}(z,\bar{x},\bar{y},t)$ 分别是编码器 $Encoder_{bf}$（nudging前数据的编码器）和 $Encoder_{ndg}$（nudging趋势的编码器）的输入。$\psi_{bf}(z \times ld,t)$、$\psi_{ndg}(z \times ld,t)$ 分别是对应输入的潜在维度表示；$\psi_{bf}(z,\bar{x},\bar{y},t;\ \theta_{AE1})$、$\psi_{ndg}(z,\bar{x},\bar{y},t;\ \theta_{AE2})$ 分别是通过 $Decoder_{bf}$（nudging前数据的解码器）和 $Decoder_{ndg}$（nudging趋势的解码器）重建的变量。

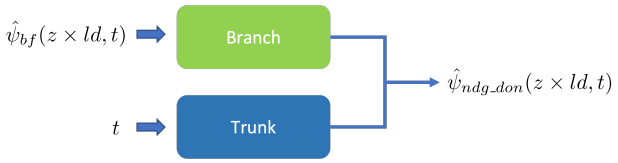


图2：ˆDeepONet的高层视图框图。$\psi_{bf}(z \times ld,\ t)$（松弛前状态的潜在空间表示）是分支网络的输入，$t$ 是主干网络的输入时间，即我们希望预测松弛趋势ˆ $\psi_{ndg\ don}(z \times ld,\ t)$的时间。

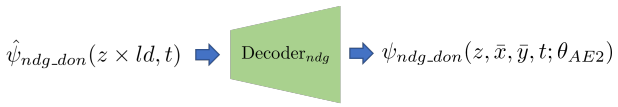


图3：DeepONet完成预测后的解码过程高层视图框图。在潜在空间中获得松弛趋势的预测后，将其传入 $Decoder_{ndg}$，以获得原始空间表示 $\psi_{ndg\ don}(z,\ \bar{x},\ \bar{y},\ t;\ \theta_{AE2})$。

算法1：基于自编码器的DeepONet算法

输入：ψ_{bf}，ψ_{ndg}，t**参数**：θ_{AE1}，θ_{AE2}，$\theta_{DeepONet}$**输出**：$\psi_{ndg\ don}$

ˆ1: ψ_{bf} $\rightarrow AE1 \rightarrow \psi_{bf}$ˆ

ˆ2: $\psi_{bf}, t \rightarrow DeepONet \rightarrow \psi_{ndg\ don}$

3: $\psi_{ndg\ don} \rightarrow AE2 \rightarrow \psi_{ndg\ don}$4: **返回** $\psi_{ndg\ don}$

结果

为测试我们的模型，我们针对地球的一个特定子区域，使用状态变量 U（纬向风）和 V（经向风）训练 DeepONet。所选子区域的纬度和经度范围分别为北纬 10° 至 80° 和西经 120° 至 50°。选择该子区域是因为

层	特征	卷积核	激活函数
二维卷积	16	(3，3)	$ReLU$
最大池化		（2，2）	
二维卷积	8	(3，3)	$ReLU$
最大池化		（2，2）	
二维卷积	4	(3，3)	$ReLU$
最大池化		(2,	
展平		2)	
全连接层			

表1：编码器架构详情表。

层	特征	卷积核	激活函数
全连接层			$ReLU$
重塑			
转置二维卷积	4	(3，3)	$ReLU$
转置二维卷积	8	(3，3)	$ReLU$
转置二维卷积	16	(3，3)	$ReLU$
二维卷积	4	(3, 3)	
重塑			
全连接层			

表2：表示解码器架构细节的表格。

分支网络		
层		激活函数
二维卷积（特征=32，K=(3,3)）		$ReLu$
二维卷积（特征=16，K=(3,3)）		$ReLu$
展平		
线性层(ld)		tanh
线性层(32)		tanh
线性层(32)		tanh
线性层(32)		tanh
线性层($M \times ld$)		
主干网络		
线性层(32)		tanh
线性层(32)		tanh
线性层(32)		tanh
线性层(32)		tanh
线性层($M \times ld$)		

表 3：DeepONet 架构示意表。

该区域的大气状态 U 和 V 具有更大的变异性，并且受历史气旋和风暴的显著影响。DeepONet 模型的训练时间跨度为十月，其中包含了 2012 年著名的飓风桑迪。基于 DeepONet 的模型能够准确预测 nudging 趋势项，相关系数高于 0.7，置信度为 95%。对桑迪风暴路径的正确预测展现了 DeepONet 模型的稳健性。

图5中左上/左下和右上/右下的子图分别表示 E3SM 和 DeepONet 预测的 nudging U（纬向风）和 V（经向风）的趋势，分别为。

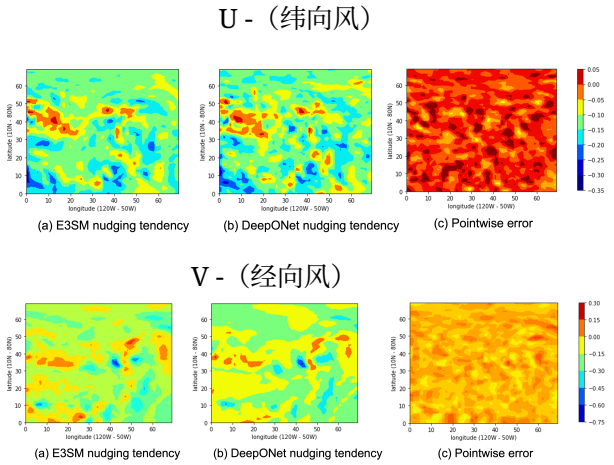


图4：展示 DeepONet 对 U（纬向风）和 V（经向风）预测的松弛趋势，并与测试域中相应的 E3SM 松弛趋势进行比较的图。列（a）显示 E3SM 松弛趋势，（b）显示 DeepONet 预测的松弛趋势，（d）显示 U 和 V 的逐点误差，分别对应。

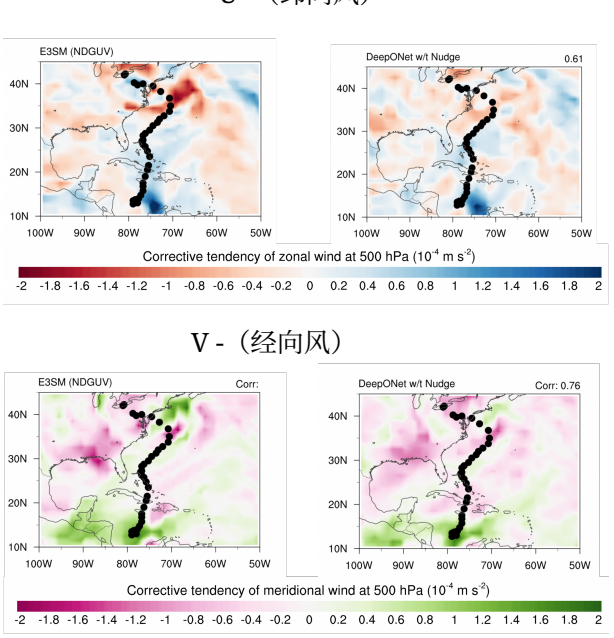


图5：展示2012年飓风桑迪路径的图。左列子图来自E3SM松弛UV模拟，右列子图分别为使用松弛趋势训练的DeepONet模拟，对应U（纬向风）和 V（经向风）。

摘要

在这项工作中，我们开发了一种基于 DeepONet 的代理模型，该模型成功地学习了气候模型中的偏差校正。我们首先训练了两个自编码器，分别为 U 和 V 提供潜在表示。DeepONet 学习了 U 和 V 的潜在表示的演变。特征和标签的潜在表示有助于 DeepONet 快速高效地学习系统动力学。通过测试 DeepONet 模型对飓风桑迪路径的预测能力，我们验证了其准确性。目前的工作包括将训练好的 DeepONet 模型集成到 E3SM 中，以便在线校正偏差，从而获得更优质的模拟。

致谢

作者感谢 DARPA 通过“人工智能辅助气候临界点建模（ACTM）”项目资助本研究，协议编号为 HR00112290029。K.S. 和 A.B. 还要感谢 Vivek Oommen（布朗大学）。K.S. 和 A.B. 还要感谢使用布朗大学 CCV 的 Oscar 高性能计算集群。

参考文献

Arribas, A.; Glover, M.; Maidens, A.; Peterson, K.; Gordon, M.; MacLachlan, C.; Graham, R.; Fereday, D.; Camp, J.; Scaife, A.; 等. 2011年. 用于季节预测的 GloSea4 集合预报系统。《月刊天气评论》, 139(6): 1891–1910. Chen, T.; and Chen, H. 1995. 神经网络利用任意激活函数逼近非线性算子及其在动力系统中的应用。《*IEEE* 神经网络汇刊》, 6(4): 911–917. E3SM Project, D. 2018. 能源百亿亿次地球系统模型 v1.0. [计算机软件] https://doi.org/10.11578/ E3SM/dc.20180418.36. Glahn, H. R.; and Lowry, D. A. 1972年. 模式输出统计（MOS）在客观天气预报中的应用。《应用气象与气候学杂志》, 11(8): 1203–1211. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Hor´anyi, A.; Mu˜noz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; 等. 2020. ERA5 全球再分析。《英国皇家气象学会季刊》, 146(730): 1999–2049. Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; and Karniadakis, G. E. 2021年. 基于算子通用逼近定理，通过 DeepONet 学习非线性算子。《自然·机器智能》, 3(3): 218–229. Oommen, V.; Shukla, K.; Goswami, S.; Dingreville, R.; and Karniadakis, G. E. 2022年. 利用神经算子和自编码器架构学习两相微观结构演化。《*npj* 计算材料》, 8(1): 190. Stockdale, T. N.; Anderson, D. L.; Alves, J. O. S.; and Balmaseda, M. A. 1998年. 使用海洋–大气耦合模型进行全球季节降雨预报。《自然》, 392(6674): 370–373. Sun, J.; Zhang, K.; Wan, H.; Ma, P.-L.; Tang, Q.; and Zhang, S. 2019年. 松弛策略对受约束 EAMv1 模拟的气候代表性和回报技巧的影响。地球系统建模进展期刊, 11(12): 3911–3933. Watt-Meyer, O.; Brenowitz, N. D.; Clark, S. K.; Henn, B.; Kwa, A.; McGibbon, J.; Perkins, W. A.; and Bretherton, C. S. 2021年. 通过机器学习松弛历史模拟来校正天气和气候模型。《地球物理研究快报》, 48(15): e2021GL092555.