海洋学报 Haiyang Xuebao
ISSN 0253-4193
CN 11-2055/P

基于Swin–Transformer和时空融合注意力机制的ENSO预测

张霄智，方巍，王淏西

ENSO prediction based on Swin-Transformer and spatio-temporal fusion attention mechanism

Zhang Xiaozhi, Fang Wei, Wang Haoxi

在线阅读 View online: https://doi.org/10.12284/hyxb2024127

张霄智, 方巍, 王淏西. 基于 Swin-Transformer 和时空融合注意力机制的 ENSO 预测[J]. 海洋学报, 2024, 46(12): 111–121, doi:10.12284/hyxb2024127

Zhang Xiaozhi, Fang Wei, Wang Haoxi. ENSO prediction based on Swin-Transformer and spatio-temporal fusion attention mechanism[J]. Haiyang Xuebao, 2024, 46(12): 111–121, doi:10.12284/hyxb2024127

基于 Swin-Transformer 和时空融合注意力机制的 ENSO 预测

张霄智[1]，方巍[1, 2, 3*]，王淏西[1]

(1. 南京信息工程大学 计算机学院/软件学院, 江苏 南京 210044; 2. 中国气象局武汉暴雨研究所 中国气象局流域强降水重点开放实验室/暴雨监测预警湖北省重点实验室, 湖北 武汉 430205; 3. 中国气象科学研究院 灾害天气国家重点实验室, 北京 100081)

摘要: 厄尔尼诺–南方涛动预测是气候变化研究的热点问题之一。本文将 Swin-Transformer 模型与时空融合注意力机制相结合, 采用 1850–2014 年 CMIP6 多模式模拟历史数据、1871–1979 年 SODA 同化数据和 1980–2023 年 GODAS 同化数据, 构建厄尔尼诺–南方涛动预测模型, 即 ENSO-STformer。该模型通过在 CMIP6 和 SODA 数据集上进行充分的训练, 并在 GODAS 数据上进行评估, 结果表明: 本文模型在提前 11 个月的 Niño3.4 指数相关技巧的平均值上分别比 CanCM4、CCSM3、GFDLaer04 动力预报系统高出 5.1%、21.6% 和 12.4%, 同时, 在中长期的 Niño3.4 指数相关技巧上显著优于其他深度学习模型, 并可以进行长达 24 个月的有效 ENSO 预测, 此外, 在对 2015–2016 年厄尔尼诺事件模拟中表现出较强的应对春季预报障碍能力。

关键词: 深度学习; ENSO 预测; 时空融合注意力机制; 卷积神经网络; Niño3.4 指数

中图分类号: P467　　**文献标志码:** A　　**文章编号:** 0253–4193(2024)12–0111–11

1 引言

厄尔尼诺–南方涛动(El Niño-Southern Oscillation, ENSO)是发生在热带中东太平洋地区的海表温度异常(Sea Surface Temperature Anomaly, SSTA)现象, 并且主要表现为厄尔尼诺(El Niño)和拉尼娜(La Niña)两个相位。当发生 El Niño 时太平洋赤道中东地区海水温度升高, 而与此相反, 当 La Niña 发生时, 相应地区海水温度降低。除此之外, ENSO 的发生往往是不规律的, 但是有着明显的周期性, 大约平均 2～7 a 发生一次[1–4]。

目前, ENSO 往往由海洋尼诺指数(Oceanic Niño Index, ONI)来表征其发生情况, 其中 ONI 由 Niño3.4 区域(5°N～5°S, 120°～170°W)SSTA 的 3 个月滑动平均值计算而来。如果 ONI 持续不少于 5 个月大于 0.5℃, 则被认定为 El Niño; 反之, 如果 ONI 持续不少于 5 个月小于 0.5℃, 则被认定为 La Niña[5]。

尽管 ENSO 事件发生于热带太平洋地区, 但是它对全球气候变化有着巨大的影响[6]。研究表明, ENSO 通过海–气的遥相关对太平洋地区及全球其他地区的降水、气候和生态系统造成了巨大影响, 并可在全球范围内诱发极端天气事件[7–8]。因此, ENSO 的准确预

收稿日期: 2024–07–24; **修订日期:** 2024–10–30。

基金项目: 国家自然科学基金面上项目 (No.42475149); 中国气象局流域强降水重点开放实验室开放研究基金 (No.2023BHR–Y14); 南京气象科技创新研究院北极阁开放研究基金 (BJG202306); 灾害天气国家重点实验室开放课题 (2024LASW-B19); 江苏省研究生科研与实践创新计划项目(NO.KYCX24_1533)。

作者简介: 张霄智(1998—), 男, 江苏省南通市人, 研究方向为深度学习、ENSO 预测。E-mail: 2201083714@qq.com

*** 通信作者:** 方巍, 教授, 博士生导师, 研究方向为人工智能、大数据分析、机器学习和计算机视觉。E-mail: fangwei@nuist.edu.cn

测对减少气候变化带来的风险，保护社会、经济和生态系统具有重要意义。其中，ENSO 的研究经历十几年的发展，其方法大致分为 3 类：基于统计学模型的方法、基于动力模型的方法和基于深度学习模型的方法。实践表明，基于统计和动力模型的 ENSO 预测都有一定的准确率，两者都能反映大气运动的部分规律[9]，但由于 ENSO 时空演化的变化性和多样性，传统的 ENSO 预测方法很难实现提前 1 a 的有效预测。但是，随着地球科学技术与观测设备的快速发展，人们通过雷达、浮标和遥感卫星等途径得到了大量的气候数据，并使用人工智能技术进行分析研究。其中，以深度学习为代表的数据驱动方法为气候预测提供了一种新的方法[10–14]，这也包括 ENSO 预测。然而，现有的基于深度学习的 ENSO 预测方法依旧存在着两大困难和挑战。首先，ENSO 的气候因子数据存在着大量的时空信息，以往基于深度学习的 ENSO 预测模型尚未对时空信息进行充分的特征提取，以至于模型无法有效地捕获气候数据背后隐藏的时空关系。其次，以往的模型仅仅是对数据进行多层卷积或者是自注意力计算，前者只能捕获气候因子中的局部信息而忽视了全局信息，使得模型无法模拟气候变化中的遥相关[15]，后者虽然可以得到气候因子在整个空间的权重系数，但是这无疑会忽略重点区域的潜在演变，进而使得模型的在长期预测上的相关技巧较低。

为了解决上述问题，本文主要工作如下：

（1）首次以 Swin-Transformer[16] 为 ENSO 预测模型架构并结合时空融合注意力机制（spatio-temporal fusion attention mechanism, STFAM），其命名为 ENSO-STformer。该模型通过层次化结构和移动窗口操作，以捕获多尺度特征，并通过时空融合注意力机制建立时间和空间关系，在保留时空信息与减少计算量的同时，建立起时间和空间维度之间的联系，以更有效地提取气候数据的时空信息，从而加强模型对 ENSO 变化中复杂规律的学习。

（2）ENSO-STformer 模型将原有的多层感知器（Multilayer Perceptron, MLP）进行改进，设计出 Niño3.4 指数预测器（Niño3.4 Index Predictor, NIP），通过对多层提取后的特征进行非线性映射，以增强时空特征之间的表达能力，使得模型可以更好地适用 ENSO 事件的预测，并最终得到未来 Niño3.4 指数。

2 相关工作

2.1 基于统计学模型的 ENSO 预测

在传统的预测方法中，基于统计学模型的 ENSO 方法往往是对历史观测数据进行分析，通过将 ENSO 信号与 SSTA、次表层的温度异常和海表风应力异常等其他气候因子之间建立联系，进而对 ENSO 进行预测。其中，常见的统计方法有典型相关分析（Canonical Correlation Analysis, CCA）[17]、奇异谱分析（Singular Spectrum Analysis, SSA）[18]、主振荡分析（Principal Oscillation Pattern Analysis, POP）[19]、线性反演模式（Linear Inverse Model, LIM）[20] 等。虽然上述的统计方法可以在一定程度上模拟 ENSO 变化，但是，这些方法统尚未有效利用 ENSO 变化中隐含着的物理规律，进而使得模型预测的时效性与精确度较低。

2.2 基于动力学模型的 ENSO 预测

另一种传统的预测方法是基于物理动力学模型构建的复杂气候模型，该方法通过模拟全球大气环流、海洋循环和生态系统变化等过程，使得效果要优于基于统计模型，并已达到 6～12 个月的可靠预测，但是该方法需要耗费大量资源，性能提升不易，与此同时，数值模式预测技巧很大程度上依赖于初始场，这使得预测系统研发极其复杂[21]。除此之外，由于 ENSO 时空演变的多变性和复杂的非线性，基于物理动力学模型在预测 ENSO 仍存在很大的不确定性[22]，尤其在长期预测上表现的尤为明显。

2.3 基于深度学习的 ENSO 预测

近几年来，人工智能的飞速发展为 ENSO 预测研究提供了新视角。与基于动力学模型相比，基于深度学习的 ENSO 预测方法不需要考虑复杂的物理关系，进而极大地提高 ENSO 建模的非线性与准确性，使得 ENSO 的预测研究得到进一步发展。

Ham 等[23] 首次将深度学习技术应用到 ENSO 预测之中，通过使用卷积神经网络（Convolutional Neural Network, CNN）对全球 SSTA 与海洋热含量异常（Ocean Heat Content Anomaly, HCA）进行 ENSO 建模，通过 CNN 的平移不变性使得模型学习到 ENSO 的前兆信息，并且其预测效果超过了世界上最先进的动力预报系统，但是该模型较为简单，无法挖掘气候因子中更为复杂的非线性特征。周佩等[24] 采用长短期记忆网络（Long Short-Term Memory, LSTM）对 Niño3.4 指数进行时序预测，使得春季预报障碍（Spring Predictability Barrier, SPB）得到了明显缓解，然而该方法仅仅考虑到时序间的依赖关系，忽略了空间信息。Geng 和 Wang[25] 将 ENSO 预测看作时空预测问题，并设计了基于密集卷积长短时记忆（Dense Convolution Long ShortTerm Memory, DC-LSTM）模型，通过充分的特征提取，得到充足的深层特征信息，并在验证数据上显

示出更高的 Niño3.4 指数全季节相关技巧，尽管如此，该方法中忽略了时间与空间之间的关联，造成模型获得的时空信息不够充分。Ye 等[26]首次将 Transfomer 引入到 ENSO 预测之中，通过将 CNN 和 Transfomer 相结合，即 ENSOTR，有效提取了气候数据中的浅层与高级语义特征，实现了对 Niño3.4 指数提前一年半的高相关性预测，但是，该方法仅仅在分辨率更小的数据集上获取的效果，其模型无法学习到的丰富的空间信息并缺乏对时空信息的学习。Rui 等[27]针对以往基于深度学习的 ENSO 预测模型没有充分考虑月份和季节属性的问题，提出了一个两阶段时空（Two-Stage SpatioTem-poral，TSST）自回归模型，通过将气候因子预测与 ENSO 指标预测相结合，以保留隐藏在气象序列中的模式，该方法考虑到了不同时间尺度上的相互关系，并在一定程度上考虑到了时空信息，但是由于采用自回归的方式，未能对历史序列进行有效学习，造成模型无法充分提取历史序列中隐藏的非线性时空信息。

本文提出一种时空融合注意力机制的 ENSO 预测算法。与以往的 ENSO 预测模型相比，本文模型首次以 Swin-Transformer 为 ENSO 的预测框架，通过 Swin-Transformer 模型独特的层次化结构与移动窗口注意力机制，既可以捕获多尺度下气候数据的特征信息，又可以兼顾全局与区域信息权重进而提高模型提取空间特征的能力；与此同时，通过引入时空融合注意力机制，在保留时空信息与减少计算量的同时，建立起通道和空间维度之间的联系，以加强模型对气候因子中隐藏着的复杂规律的学习；最后，通过设计 Niño3.4 指数预测器模块，将时空数据压缩为一维数据并进行多层的非线性映射，提升模型泛化能力。

3　数据和方法

3.1　数据集

本文使用经度范围为 90°E～30°W，纬度范围为 30°S～30°N 的 SSTA 与海表下 150 m 温度异常（Sea Surface Temperature Anomaly at 150 m Depth，SSTA150）作为预测因子，其分辨率为 1°×1°。

数据来自第六次国际耦合模式比较计划（Coupled Model Intercomparison Project Phase 6，CMIP6）（1850–2014 年）[28]中 16 个模式的逐月历史模拟数据、简单海洋数据同化版 2.2.4（Simple Ocean Data Assimilation，SODA）[29]（1871–1979 年）的逐月再分析数据和全球海洋数据同化系统（Global Ocean Data Assimilation System，GODAS）[30]（1980–2023 年）的逐月再分析数据。尽管 SODA、GODAS 和 CMIP6 之间的时间有重叠，但这 3 个数据集之间不存在依赖关系，其中历史模拟数据和再分析数据信息如表 1 和表 2 所示。

表 1　CMIP6 数据信息

Table 1　CMIP6 data information

序号	模式名称	所属国家	研发机构
1	ACCESS-CM2	澳大利亚	CSIRO-ARCCSS
2	ACCESS-ESM1-5	澳大利亚	CSIRO
3	CAMS-CSM1-O	中国	CAMS
4	CanESM5-CanOE	加拿大	CCCma
5	E3SM-1-O	美国	E3SM-Project LLNL UCI
6	BCC-CSM2-HR	中国	BCC
7	FGOALS-f3-L	中国	CAS
8	FGOALS-g3	中国	CAS
9	FIO-ESM-2-0	中国	FIO-QLNM
10	AWI-CM-1-1-MR	德国	AWI
11	IPSL-CM6A-LR	法国	IPSL
12	MIROC6	日本	MIROC
13	MIROC-ES2L	日本	MIROC
14	MRI-ESM2-0	日本	MRI
15	NESM3	中国	NUIST
16	NorESM2-MM	挪威	NCC

表 2　SODA 和 GODAS 数据信息

Table 2　Data information of SODA and GODAS

序号	再分析数据	所属国家	研发机构
1	SODA	美国	UMD
2	GODAS	美国	NCEP

本实验将数据中所有缺失值与异常值都设为 0，并且通过使用滑动窗口的方法增加样本数量，其中窗口大小为 29 步长，间隔为 1，即前 3 个步长作为模型输入，后 26 个步长用于模型输出验证，时间间隔都为 1 个月。因此，CMIP6、SODA 和 GODAS 中样本数分别为 31 652、1 628 和 452 个样本数据。鉴于 CMIP6 中不同的模式数据存在偏差，并与再分析数据存在较大差异，因此，本文通过迁移学习方法对 ENSO-STformer 进行训练。首先，将 ENSO-STformer 在 CMIP6 数据上进行预训练，然后，将预训练模型在 SODA 数据上进行再训练，最终在 GODAS 数据上进行评估，其中 CMIP6 与 SODA 的数据划分如表 3 所示。

表 3　CMIP6 和 SODA 数据划分

Table 3　Data division between CMIP6 and SODA

数据集	训练样本数	验证样本数
CMIP6	25 321	6 331
SODA	1 302	325

3.2　ENSO 预测问题定义

ENSO 预测是利用历史气候因子数据以进行未来 Niño3.4 指数预测。本文利用深度学习方法，通过将历史 3 个月的 SSTA 与 SSTA150 两个气候因子数据在通道维度进行拼接，并对气候数据的时空挖掘，以预测 26 个月的 Niño3.4 指数。假设给定一个气候因子的历史步数为T的时空数据序列为$X_{1,T}=\{X_1, X_2,\cdots,X_T\}$，其中$X_1$和$X_T$分别表示过去第 1 时间点和过去第$T$时间点的气候数据。为了简化公式表示，ENSO 预测模型可以表示为$f(\cdot)$，它把历史 T步的气候数据$X_{1,T}$作为输入，然后提取历史气候数据的特征并进行未来N步的 Niño3.4 指数预测 $Y_{T+1,T+N}=\{Y_{T+1}, Y_{T+2},\cdots,Y_{T+N}\}$。上述过程可被公式 1 所描述：

$$Y_{T+1,T+N}=f(X_{1,T}), \tag{1}$$

式中，T为输入模型的气候数据的历史步数，N为模型输出 Niño3.4 指数的未来步数，$f(\cdot)$表示 ENSO 深度学习模型。

3.3　ENSO-STformer 模型网络架构

本文模型采用的是 Swin-Transformer 模型结构，包含一个下采样模块用于提取历史气候因子浅层特征以及降低图像分辨率、多个 ENSO-STformer 骨干网络，即 Swin STFA Block，用于提取包含时空信息在内的多维度深层特征，以及一个 Niño3.4 指数预测器用于将多维度融合信息进行充分的非线性映射，并得到未来的 Niño3.4 指数。

首先，使用卷积核为3×3的 CNN 进行下采样操作，以达到减小图片分辨率和最大程度保留信息；其次，为了避免过深的网络结构带来的模型过拟合，这里采用 2 个骨干网络进行时空数据挖掘；最后，通过 Niño3.4 指数预测器进行未来 26 个月的 Niño3.4 指数预测。

本文模型网络如图 1 所示，在下采样操作上，由于 SSTA 与 SSTA150 为海洋数据，其分布之中有着大量数值为 0 的陆地数据，因此原始的 Swin-Transformer 中将图片进行简单卷积以降低图片分辨率的做法，不能够充分地提取气候因子信息，进而降低注意力机制的计算效率。

因此，ENSO-STformer 中下采样主要有一组卷积模块构成，通过卷积操作以及最大池化（MaxPool）的

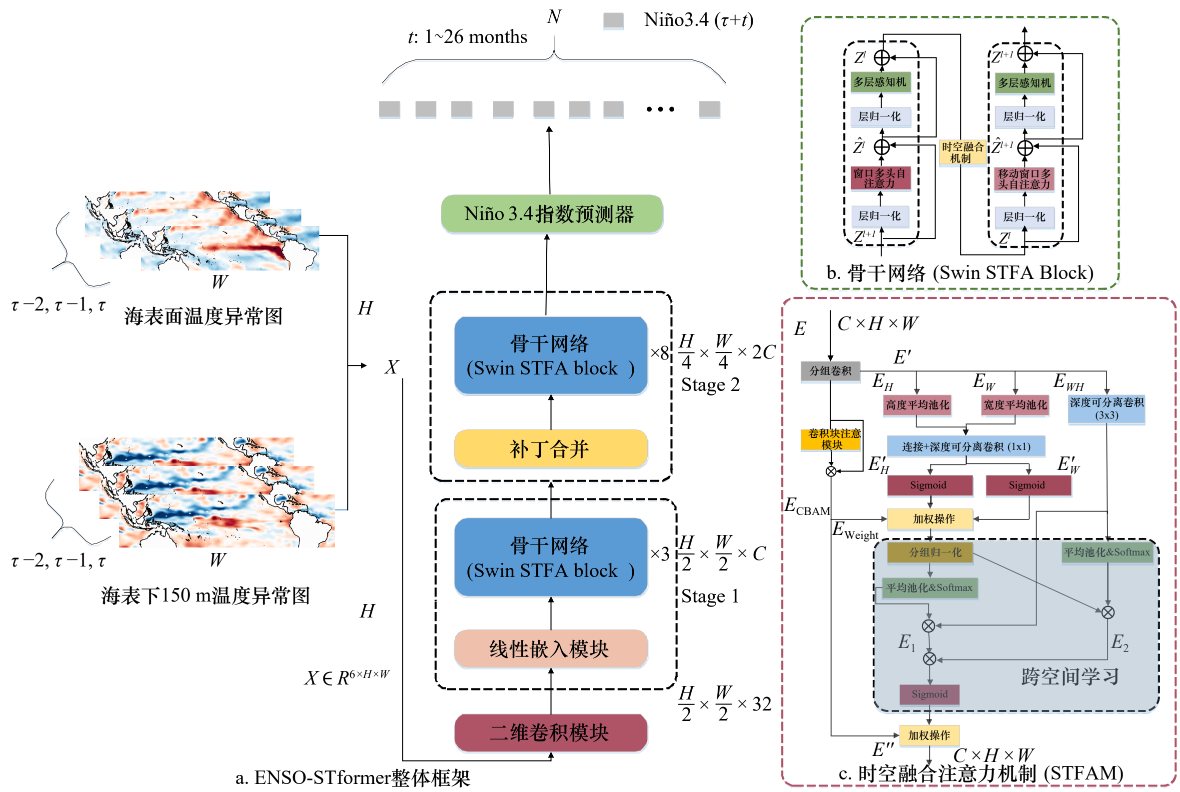


图 1　本文模型结构

Fig. 1　Model structure of this paper

方法达到最大程度过滤陆地数据和降低图片分辨率的目的，其过程如图 2 所示。其中图 2a 表示了整体过程，图 2b 表示下采样后的空间映射。

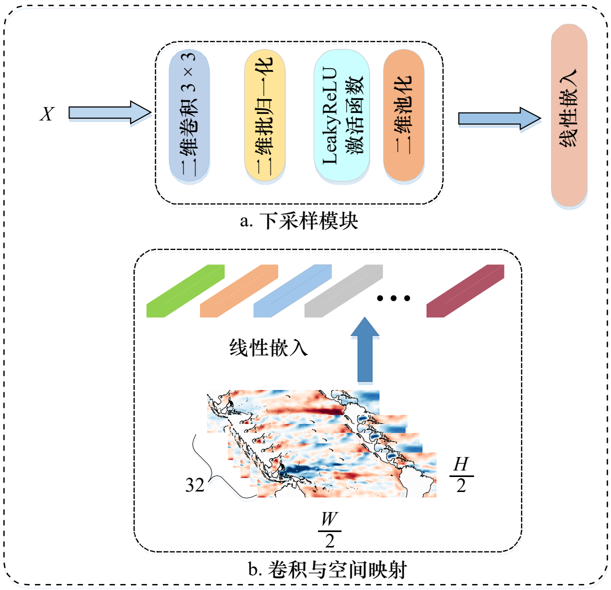


图 2　气候数据下采样与映射

Fig. 2　Climate data downsampling and mapping

3.3.1　窗口注意力机制

窗口注意力机制是窗口多头自注意力(Windows Multi-head Self-Attention, W-MSA)和移动窗口多头自注意力(Shifted Windows Multi-Head Self-Attention, SW-MSA)。

如图 1b 所示，窗口多头自注意力能够将注意力计算限制在单个不重叠的区域之中，充分获取区域信息，但是该方法造成不同窗口之间信息无法相互影响，并导致模型对全局特征的提取不够充分，进而使得模型无法模拟气候变化中的遥相关。因此，在移动窗口多头自注意力中使用了移动窗口策略，使得信息可以在相邻窗口之间相互传递信息，进而增强模型对全局信息的获取能力，其计算过程为

$$z^l = \mathrm{MSA}^W(z^{l-1}), \tag{2}$$

$$z^{l+1} = \mathrm{MSA}^{SW}(\mathrm{STFAM}(z^l)), \tag{3}$$

式中，MSA_W 表示窗口多头自注意力，MSA^W 表示移动窗口多头自注意力；STFAM 表示时空融合机制。

3.3.2　时空融合注意力机制

由于，以往通道或空间注意力机制往往通过降低通道维度来建立跨通道之间的关系，这可能给深度模型中的特征表示带来了消极影响[31]。除此之外，在对气候数据进行空间信息提取的同时，需要对时空信息进行充分的关联，以提高模型对时空数据的表达能力。如图 1c 所示时空融合注意力机制，使用了并行子网络以聚合多尺度空间信息，并通过跨空间学习以增强时空信息融合。

首先，通过分组卷积进行初次的特征提取，时空融合注意力机制的子网络主要由四条子网络构成，其中一条平行引入了卷积块注意模块(Convolutional Block Attention Module, CBAM)，并通过残差连接的方式，得到浅层时空融合信息，其中两条平行路线将两个编码特征沿图像高度和宽度方向进行全局平均池化(AvgPool)，并使用 1×1 卷积核与 Sigmoid 函数将空间信息嵌入到含有时间学习的通道之中，进而建立起时间与空间信息的关联，与此同时，为了提取深层的特征与提高计算效率，通过使用 3×3 深度可分离卷积(Depthwise Separable Convolution, DSC)，提升模型的计算效率，该模块的计算过程为

$$E' = \mathrm{GroupConv2d}^{1\times1}(E), \tag{4}$$

$$E_H = \mathrm{Avg}^H(E'), E_W = \mathrm{Avg}^W(E'), \tag{5}$$

$$E_{WH} = \mathrm{DSConv2d}^{3\times3}(E')), \tag{6}$$

$$E'_W, E'_H = \mathrm{Sigmoid}(\mathrm{DSConv2d}^{1\times1}(\mathrm{cat}(E_W, E_H))), \tag{7}$$

$$E_{\mathrm{CBAM}} = \mathrm{CBAM}(E') + E', \tag{8}$$

$$E_{\mathrm{Weight}} = E_{\mathrm{CBAM}} \odot (E'_W \times E'_H), \tag{9}$$

式中，$\mathrm{GroupConv2d}^{k\times k}$ 表示使用不同的卷积核的分组卷积；$\mathrm{DSConv2d}^{k\times k}$ 表示使用不同的卷积核的深度可分离卷积；CBAM 表示卷积块注意力；Avg 表示全局平均池化；Sigmoid 表示非线性激活函数。

时空融合学习主要是将并行子网络中的两个分支的输出向量进行不同空间维度方向的跨空间学习聚合。首先，使用分组归一化将输出向量转化为对应维度形状，并通过平均池化(AvgPool)对全局空间信息进行编码，最后在 Sigmoid 的激活下与 3×3 分支下的特征向量进行点积运算，以此获得第一个空间注意力图。与此同时，通过相似的方法对 3×3 分支的输出向量进行相同操作，获得第二个空间注意力图，并在联合激活前进行相加，从而实现时空信息融，该模块的计算过程为

$$E_1 = \mathrm{GroupNorm}(\mathrm{Softmax}(\mathrm{Avg}(E_{\mathrm{Weight}}))) \odot E_{HW}, \tag{10}$$

$$E_2 = \mathrm{GroupNorm}(E_{\mathrm{Weight}}) \odot \mathrm{Softmax}(\mathrm{Avg}(E_{HW})), \tag{11}$$

$$E'' = E_{\mathrm{CBAM}} \odot \mathrm{Sigmoid}(E_1 + E_2), \tag{12}$$

式中，GroupNorm 为分组归一化；$\odot$ 为哈达玛积运算，Avg 表示全局平均池化；Sigmoid 和 Softmax 表示激活函数。

3.3.3　Niño3.4 指数预测器

由于众多的补丁(Patches)中包含了时空融合信息，

因此之前 Swin-Transformer 中将多个补丁进行 AvgPool 的方式会造成模型在预测时忽略重要的时空信息。

针对该问题，本文设计了 Niño3.4 指数预测器模块，其结构如图 3 所示，通过对补丁的特征向量进行 Maxpool 得到每一个补丁中数值更大的特征信息，与此同时，通过多层的非线性映射将包含时空信息的特征向量压缩为一维向量，增强时空特征之间的联合表达，进而提高模型的预测能力与鲁棒性，除此之外，在 Niño3.4 指数预测器的输出层添加正则化操作(Dropout)，通过随机失活神经元以模拟大气−海洋之中不确定性并避免模型过拟合，该模块的计算过程为

$$P_{\max} = \mathrm{MaxPool(Patches)}, \tag{13}$$

$$P_{\mathrm{dropout}} = \mathrm{Dropout(LN(MLP}(P_{\max}))), \tag{14}$$

$$\mathrm{Output} = \mathrm{FC}(\mathrm{P}_{\mathrm{dropout}}), \tag{15}$$

式中，MaxPool 为最大池化；LN 为层归一化；Dropout 表示随机失活神经元；FC 表示全连接。

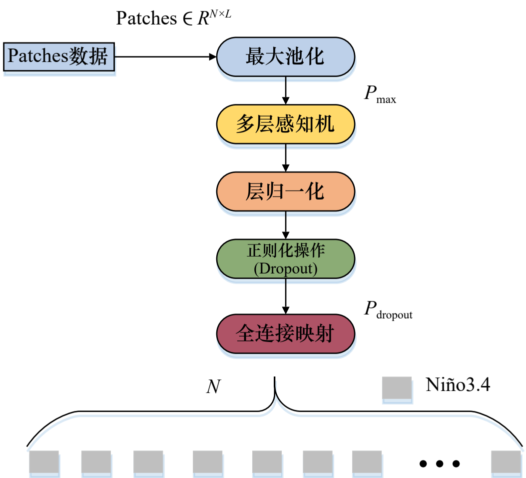


图 3　Niño3.4 指数预测器模块

Fig. 3　Niño3.4 Index Predictor module

4　验证与结果分析

4.1　实验环境与设置

本文实验在 Ubuntu22.04 系统下进行，模型使用 PyTorch 开源神经网络框架进行实现，在 NVIDIA GeForce RTX 3060 设备上进行训练。

实验中，每个输入包含 3 个时间步长的 2 个气候因子，并预测未来 26 个时间步长的 Niño3.4 指数。在模型训练过程中，初始学习率(Learning Rate)设置为 0.001 6，时期(Epoch)的最大值设置为 100。

在模型超参数设置方面，考虑输入数据大小，因此，本实验设置 CNN 中卷积核大小为 3，隐藏维度为 32。除此之外，借鉴原 Swin-Transformer[16] 中参数设置，将每一个补丁的映射维度为 96，其次，设置 2 个阶段中层数分别为 3 和 8 的骨干网络，其中每一个骨干网络中的滑动窗口大小分别为 3 和 5，移动窗口多头自注意力中多头大小设置为 8，时空融合注意力机制中的分组卷积的分组数设置为 4。批大小(Batch-size)设置为 3，并选择 Adam Optimizer 梯度下降优化算法作为模型的优化器，其权重衰减系数为 0.001。

4.2　实验性能指标

为充分评估模型性能，本文采用平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)以及皮尔逊相关系数(Pearson Correlation Coefficient, PCC) 3 个指标来评估模型性能，MAE、RMSE 与 PCC 3 个指标的具体计算公式如下所示：

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{16}$$

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{17}$$

$$\mathrm{PCC} = \frac{\sum_{i=1}^{n}(X_i - \hat{X})(Y_i - \hat{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \hat{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}}, \tag{18}$$

式中，n表示样本数量，y_i表示第i个真实值，$\hat{y}_i$表示第i个预测值，X和Y分别表示两个变量的取值，$\hat{X}$和$\hat{Y}$表示两个变量的平均值；MAE 和 RMSE 值越小表示模型预测效果越好，PCC 值是计算真实与预测的 Niño3.4 指数的相关系数，它的范围为 0 到 1，数值越高则表示模型预测 ENSO 的相关技巧越高。

4.3　多模型的 ENSO 相关技巧对比

为了比较 ENSO-STformer 的 ENSO 预测能力，本文将 ENSO-STformer 与基于深度学习的 ENSO 预测模型和北美多模式集合(North American Multi-Model Ensemble, NMME)[32] 项目中的多个动力预报系统进行对比，并通过计算各个模型与预报系统在 1984−2017 年期间 ONI 的相关技巧来评估其预测能力，其中基于深度学习的各模型均在 1984−2017 年期间的 GODAS 数据进行评估。用于比较的深度学习模型分别是 CNN[23]、ENSOTR[26] 和 GCNTR[33]。用于比较的动力预报系统 CanCM4、CCSM3 与 GFDL-aer04 来自于 NMME。

不同类型模型的 ENSO 相关技巧如图 4 所示，横坐标为不同提前期，纵坐标为 ONI 预测值与真实值之间的相关系数，其数值越高说明模型预测能力越强。ENSO-STformer 中的 ONI 的相关技巧在提前

11 个月的情况下超过了 NMME 中 3 个动力预报系统，并且与其他的深度学习模型相比，随着提前月数的增加，ONI 的相关技巧的下降幅度最小，除此之外，最为明显的是，在提前 14 个月后，ENSO-STformer 的相关技巧远超过以往基于深度学习的 ENSO 预测模型，并可以进行长达两年的有效预测。这表明 ENSO-STformer 在 ENSO 预测上有着明显的优势，尤其在中长期的 ENSO 预测上。

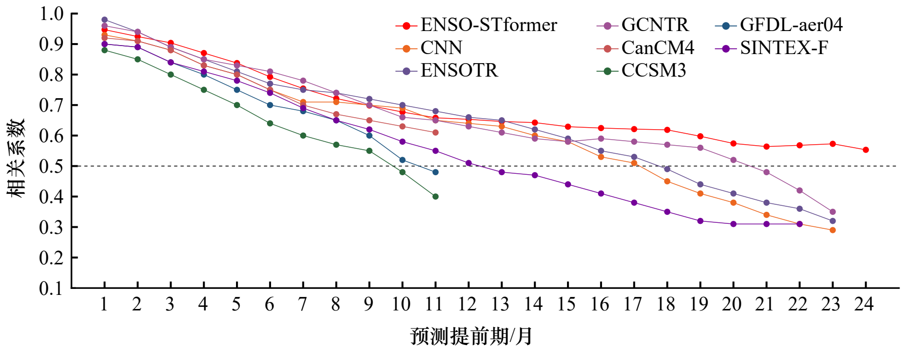


图 4　ENSO-STformer、CNN[23]、ENSOTR[26]、GCNTR[33] 和动力预报系统 CanCM4[32]、CCSM3[32]、GFDL-aer04[32] 与 SINTEX-F[33] 的 ONI 相关技巧情况

Fig. 4　ONI related skills of ENSO-STformer, CNN[23], ENSOTR[26], GCNTR[33], CanCM4[32], CCSM3[32], GFDL-aer04[32] and SINTEX-F[33]

4.4　ENSO-STformer 不同提前期的相关技巧

图 5a 至图 5f 是分别 ENSO-STformer 对 1985−2017 年的 Niño3.4 指数提前 1、3、6、12、15 和 18 个月的预测结果。

从图 5 中可以看出，ENSO-STformer 在不同提前期的预测能力，随着提前期的增加，其相关技巧随之下降，但是依旧可以在提前 18 个月上达到 0.6 的高技巧。与此同时，模型在提前 12 个月的预测之中，能够较好地拟合 ENSO 处于 El Niño 时的各种峰值，并且预测曲线没有表现出明显的后滞情况，这说明模型对 El Niño 事件的捕获较为可靠。但是，从预测结果来看，模型对 ENSO 的冷事件捕获能力较弱，并最早于提前期 6 个月中便无法预测 La Niña 的最低值。这可能归咎于模型中多次的 MaxPool 操作，导致模型更多地捕获到暖事件的信号，从而降低了对冷事件的学习能力。

4.5　模拟 2015−2016 年 El Niño 事件

为了评估 ENSO-STformer 应对具体 ENSO 事件的能力，本文将模型模拟 2015−2016 年 El Niño 事件，同时为了检验模型应对 SPB 的能力，本文分别从 2015 年 2 月、2015 年 3 月、2015 年 4 月以及 2015 年 5 月开始对 2015−2016 年的 El Niño 进行预测，并通过 ONI 来比较模型预测的情况，其中观测值是 GODAS 中相应时间的 ONI。除此之外，本文将相同时刻起报的动力预报系统 CMC1-CanCM3 中的预报结果与 ENSO-STformer 的预测结果进行比较，CMC1-CanCM3 预报结果为该系统中 10 个模式预报的 SSTA 的距平值的平均值并以此计算出相应的 ONI，其结果如图 6 所示。

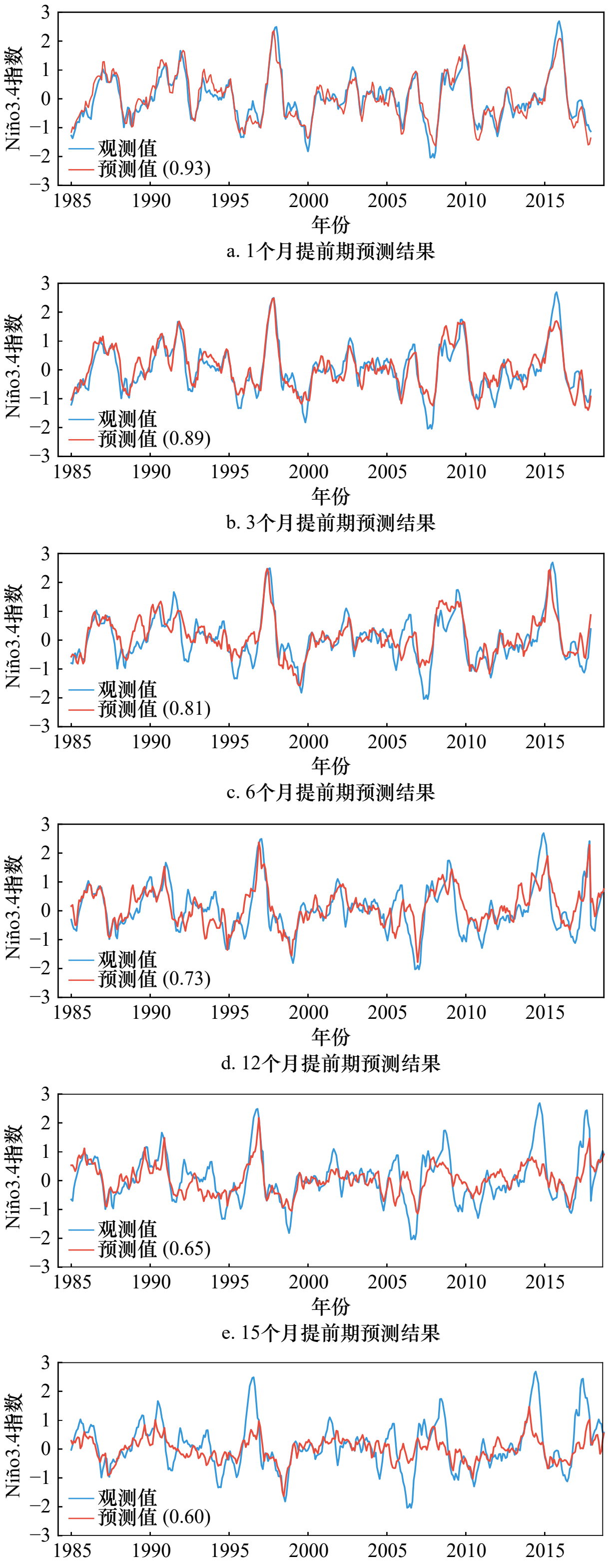
从图 6 中可以看出，从 2015 年 2 月至 2015 年 3 月起报的结果可以看出，ENSO-STformer 对 El Niño 预测强度偏弱，这情况在 2015 年 8 月至 2015 年 11 月之间最为明显，但是，这两次的起报依旧可以大致预测出本次 El Niño 的变化趋势，并且与观测值相比，没有出现明显的滞后性。与此同时，2015 年 4 月与 2015 年 5 月的起报结果可以很好地拟合观测值的变化趋势，除此之外，ENSO-STformer 与 CMC1-CanCM3 在各时间点的起报情况相比，ENSO-STformer 的预测效果要远远好于 CMC1-CanCM3 的预报效果，这表明 ENSO-STformer 相比于动力预报系统而言，可以较好地应对 SPB 带来的消极影响。

4.6　消融实验

为了验证本文对原始 Swin-Transformer 改进的有效性，本文针对所有改进模块进行多组消融实验，并在 GODAS 数据集上进行验证。

表 4 中，第一组是不做任何修改的 Swin-Transformer 模型；“本文（w/时空融合注意力机制）”表示只保留时空融合注意力机制所得到的实验结果；“本文（w/Niño3.4 指数预测器）”表示只保留 Niño3.4 指数预测器所得到的实验结果；“Swin-Transformer（w/下采样）”表示替换原始 Swin-Transformer 中图片下采样操作后所得到结果。

实验结果如表 4 所示，修改 Swin-Transformer 中下采样方式，确实可以为模型带来了不小的提升，这也说明了，针对数据处理上的优化依旧是一项重要的



a. 1个月提前期预测结果

b. 3个月提前期预测结果

c. 6个月提前期预测结果

d. 12个月提前期预测结果

e. 15个月提前期预测结果

f. 18个月提前期预测结果

图 5 ENSO-STformer 模型在 GODAS 数据集上不同提前期的预测效果以及对应的相关系数

Fig. 5 Prediction effect of ENSO-STformer model in different lead times and corresponding correlation coefficient on GODAS data set

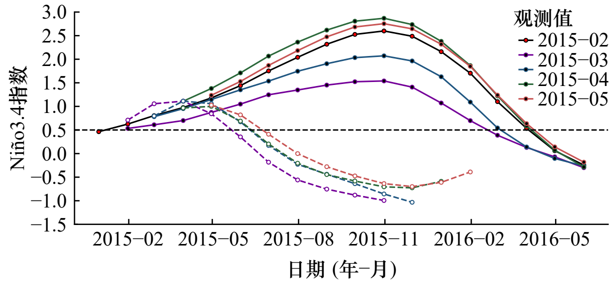


图 6 ENSO-STformer 和 CMC1-CanCM3 在不同时期下 ONI 的预测值与真实值比较

Fig. 6 ONI predicted values of ENSO-STformer and CMC1-CanCM3 in different periods are compared with the real values

实线和虚线分别表示本文模型和 CMC1-CanCM3 的预测结果，不同颜色表示不同提前时期的情况

Solid and dashed lines represent the predicted results of the model in this paper and CMC1-CanCM3 respectively, and different colors represent the situations in different advance periods

表 4 各模型评价指标对比结果

Table 4 Comparison results of evaluation indicators of each model

模型	RMSE ↓	MAE↓	PCC↑
Swin-Transformer	1.11	0.95	0.52
Swin(w/下采样)	0.91	0.77	0.64
ENSO-STformer	**0.76**	**0.63**	**0.73**
本文(w/时空融合注意力机制)	0.83	0.68	0.69
本文(w/Niño3.4指数预测器)	0.88	0.72	0.66

注：表中RMSE、MAE和PCC数据为验证集中所有样本在相应模型预测下的平均值，向下的箭头表示指标值越低模型性能越好，向上的箭头表示指标值越高模型性能越好，所加粗的数据是该指标下最优的情况。

工作。除此之外，时空融合注意力机制的引入比 Niño3.4 指数预测器的引入性能提升更高，而将所有改进模块融合的本文模型获取了最佳的性能提升，这有力的说明了：将时空维度的信息建立关联，从而获取气候因子在时空演化中隐藏的复杂信息对于模型预测性能的提高至关重要，与此同时，通过添加大气−海洋变化中的不确定性，从而为模型进行一定的物理约束，使得模型能够较为合理地模拟真实世界中复杂的非线性变化。

5 结语

针对以往基于深度学习的 ENSO 预测模型中时空信息未能充分挖掘，受到通道与空间注意力机制的启发，本文提出一种时空融合注意力机制的预测模型，在 Swin-Transformer 的基础上引入时空融合注意

力机制并提出了 ENSO-STformer 模型。通过对 SSTA 与 SSTA150 两个气候因子进行 ENSO 建模，并利用 CMIP6、SODA 和 GODAS 数据对 ENSO-STformer 进行训练与验证。

实验结果表明：在提前 11 月之中，其相关技巧超过于动力预报系统，并相比于 CMC1-CanCM3 动力预报模型而言，本模型对现有的 SPB 问题有较好的应对能力；与此同时，ENSO-STformer 在提前 14 个月之后的预测技巧上超越以往基于深度学习的 ENSO 预测模型，并能够在长达 24 个月 ENSO 预测的相关技巧上可以达到 0.5 以上；除此之外，通过消融实验，可以得出本文在 Swin-Transform 基础上的诸多改动的有效性，例如引入时空融合注意力机制（Spatio-Temporal Fusion Attention Mechanism，STFAM）与 Niño3.4 指数预测器（Niño3.4 Index Predictor，Niño3.4 指数预测器）。

尽管 ENSO-STformer 在 ENSO 预测上显示出优秀的性能，但是依旧存在以下两点明显问题：

（1）本文仅仅使用了两个气候因子，即 SSTA 与 SSTA150，但是 ENSO 的演变是有众多气候因子共同影响的，因此，在之后的研究之中应当尝试更多的气候因子，并定量评估相关的气候因子对 ENSO 预测的贡献程度；

（2）本文通过数据驱动的方式对 ENSO 事件进行建模，其过程尚未融入物理机制，从而导致模型的可解释性较差，这十分不利于深入探讨 ENSO 变化的潜在机制，因此，如何将更多的物理定律作为先验知识融入深度学习模型？这些问题有待于进一步研究。

参考文献：

[1] 石世玮, 智海, 林鹏飞, 等. 热带太平洋盐度年际变化对海表温度异常作用比较: 1997/1998、2014/2015 和 2015/2016 年 El Niño 事件 [J]. 大气科学, 2020, 44(5): 1057−1075.
Shi Shiwei, Zhi Hai, Lin Pengfei, et al. Contrasting salinity interannual variations in the tropical pacific and their effects on recent El Niño events: 1997/1998, 2014/2015, and 2015/2016[J]. Chinese Journal of Atmospheric Sciences, 2020, 44(5): 1057−1075.

[2] 祁莉, 毛欣, 张文君. 北太平洋海温 Victoria 模态与 ENSO 年际关系的非对称特征 [J]. 大气科学学报, 2022, 45(2): 280−291.
Qi Li, Mao Xin, Zhang Wenjun. Asymmetric characteristics of interannual relationship between Victoria mode of North Pacific SST and ENSO[J]. Transactions of Atmospheric Sciences, 2022, 45(2): 280−291.

[3] 万云霞, 晏红明, 金燕, 等. 两次不同 ENSO 背景下云南冬季极端冷事件的成因分析 [J]. 大气科学学报, 2023, 46(4): 575−586.
Wan Yunxia, Yan Hongming, Jin Yan, et al. Causes of two extremely cold events in Yunnan during different ENSO events[J]. Transactions of Atmospheric Sciences, 2023, 46(4): 575−586.

[4] Zhang Ronghua, Zhou Lu, Gao Chuan, et al. A transformer-based coupled ocean-atmosphere model for ENSO studies[J]. Science Bulletin, 2024, 69(15): 2323−2327.

[5] Wang Gaige, Cheng Honglei, Zhang Yiming, et al. ENSO analysis and prediction using deep learning: a review[J]. Neurocomputing, 2023, 520: 216−229.

[6] Cai Wenju, Santoso A, Gollins M, et al. Changing El Niño-Southern Oscillation in a warming climate[J]. Nature Reviews Earth & Environment, 2021, 2(9): 628−644.

[7] Anttila-Hughes J K, Jina A S, Mccord G C. ENSO impacts child undernutrition in the global tropics[J]. Nature Communications, 2021, 12(1): 5785.

[8] Cai Wenju, Ng B, Geng Tao, et al. Anthropogenic impacts on twentieth-century ENSO variability changes[J]. Nature Reviews Earth & Environment, 2023, 4(6): 407−418.

[9] 任宏利, 郑飞, 罗京佳, 等. 中国热带海−气相互作用与 ENSO 动力学及预测研究进展 [J]. 气象学报, 2020, 78(3): 351−369.
Ren Hongli, Zheng Fei, Luo Jingjia, et al. A review of research on tropical air-sea interaction, ENSO dynamics, and ENSO prediction in China[J]. Acta Meteorologica Sinica, 2020, 78(3): 351−369.

[10] Dong Changming, Xu Guangjun, Han Guoqing, et al. Recent developments in artificial intelligence in oceanography[J]. Ocean-Land-Atmosphere Research, 2022, 2022: 9870950.

[11] Zhu Yuchao, Zhang Ronghua, Moum J N, et al. Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations[J]. National Science Review, 2022, 9(8): nwac044.

[12] Zheng Gang, Li Xiaofeng, Zhang Ronghua, et al. Purely satellite data-driven deep learning forecast of complicated tropical instability waves[J]. Science Advances, 2020, 6(29): eaba1482.

[13] Ling Fenghua, Luo Jingjia, Li Yue, et al. Multi-task machine learning improves multi-seasonal prediction of the Indian Ocean Dipole[J]. Nature Communications, 2022, 13(1): 7681.

[14] Zhao Sen, Jin Feifei, Stuecker M F, et al. Explainable El Niño predictability from climate mode interactions[J]. Nature, 2024, 630(8018): 891−898.

[15] 黄刚, 胡开明, 唐颢苏, 等. 从能量学角度理解气候背景场对 ENSO 热带和热带外遥相关的影响 [J]. 大气科学, 2024, 48(1): 218−227
Huang Gang, Hu Kaiming, Tang Haosu, et al. Understanding the influence of background mean-state field on ENSO tropical and ex-

tratropical teleconnection from an energetic perspective[J]. Chinese Journal of Atmospheric Sciences, 2024, 48(1): 218−227
[16] Liu Ze, Lin Yutong, Cao Yue, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 9992−10002.
[17] Fekadu K. Ethiopian seasonal rainfall variability and prediction using canonical correlation analysis (CCA)[J]. Earth Sciences, 2015, 4(3): 112−119.
[18] Biabanaki M, Eslamian S S, Koupai J A, et al. A principal components/singular spectrum analysis approach to ENSO and PDO influences on rainfall in western Iran[J]. Hydrology Research, 2014, 45(2): 250−262.
[19] Zhou Lu, Zhang Ronghua. A hybrid neural network model for ENSO prediction in combination with principal oscillation pattern analyses[J]. Advances in Atmospheric Sciences, 2022, 39(6): 889−902.
[20] Kido S, Richter I, Tozuka T, et al. Understanding the interplay between ENSO and related tropical SST variability using linear inverse models[J]. Climate Dynamics, 2023, 61(3/4): 1029−1048.
[21] 杨淑贤, 零丰华, 应武杉, 等. 人工智能技术气候预测应用简介 [J]. 大气科学学报, 2022, 45(5): 641−659.
Yang Shuxian, Ling Fenghua, Ying Wushan, et al. A brief overview of the application of artificial intelligence to climate pre-diction[J]. Transactions of Atmospheric Sciences, 2022, 45(5): 641−659.
[22] Fang Wei, Sha Yu, Sheng V S. Survey on the application of artificial intelligence in ENSO forecasting[J]. Mathematics, 2022, 10(20): 3793.
[23] Ham Y G, Kim J H, Luo Jingjia. Deep learning for multi-year ENSO forecasts[J]. Nature, 2019, 573(7775): 568−572.
[24] 周佩, 黄颖婕, 胡冰逸, 等. 基于 LSTM 深度学习的 ENSO 预测及其春季预报障碍研究 [J]. 北京大学学报 (自然科学版), 2021, 57(6): 1071−1078.
Zhou Pei, Huang Yingjie, Hu Bingyi, et al. Spring predictability barrier phenomenon in ENSO prediction model based on LSTM Deep learning algorithm[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2021, 57(6): 1071−1078.
[25] Geng Huantong, Wang Tianlei. Spatiotemporal model based on deep learning for ENSO forecasts[J]. Atmosphere, 2021, 12(7): 810.
[26] Ye Min, Nie Jie, Liu Anan, et al. Multi-year ENSO forecasts using parallel convolutional neural networks with heterogeneous architecture[J]. Frontiers in Marine Science, 2021, 8: 717184.
[27] Rui Chuang, Sun Zhengya, Zhang Wensheng, et al. Enhancing ENSO predictions with self-attention ConvLSTM and temporal embeddings[J]. Frontiers in Marine Science, 2024, 11: 1334210.
[28] Lu Yunlong, FENG Junqiao, HU Dunxin. CMIP6 models simulation of the connection between North/South Pacific Meridional Mode and ENSO[J]. Journal of Oceanology and Limnology, 2024, 42(2): 439−453.
[29] Lyu Pumeng, Tang Tao, Ling Fenghua, et al. ResoNet: robust and explainable ENSO forecasts with hybrid convolution and transformer networks[J]. Advances in Atmospheric Sciences, 2024, 41(7): 1289−1298.
[30] Behringer D, Xue Yan. Evaluation of the global ocean data assimilation system at NCEP: the Pacific Ocean[C]//Eighth Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and LandSurface, AMS 84th Annual Meeting. Seattle, 2004.
[31] Ouyang Daliang, He Su, Zhang Guozhong, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes Island: IEEE, 2023: 1−5.
[32] Kirtman B P, Min D, Infanti J M, et al. The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction[J]. Bulletin of the American Meteorological Society, 2014, 95(4): 585−601.
[33] 杜先君, 李河. 基于集成 GCN-Transformer 网络的 ENSO 预测模型 [J]. 海洋学报, 2023, 45(12): 156−165.
Du Xianjun, Li He. ENSO prediction model based on integrated GCN-Transformer network[J]. Haiyang Xuebao, 2023, 45(12): 156−165.

ENSO prediction based on Swin-Transformer and spatio-temporal fusion attention mechanism

Zhang Xiaozhi[1], Fang Wei[1,2,3], Wang Haoxi[1]

(1. *School of Computer Science* & *School of Software* , *Nanjing University of Information Science and Technology*, *Nanjing* 210044, *China*; 2. *China Meteorological Administration*, *China Meteorological Administration Basin Heavy Rainfall Key Laboratory/Hubei Key Laboratory for Heavy Rain Monitoring and Warning Research*, *Institute of Heavy Rain*, *Wuhan* 430205, *China*; 3. *State Key Laboratory of Severe Weather*, *Chinese Academy of Meteorological Sciences*, *Beijing* 100081, *China*)

Abstract: The prediction of El Niño-Southern Oscillation is one of the hot issues in climate change research. This paper combines swin-transformer model with spatio-temporal fusion attention mechanism, and uses CMIP6 multi-model simulation historical data from 1850 to 2014, SODA assimilated data from 1871 to 1979 and GODAS assimilated data from 1980 to 2023 to construct El Niño-Southern Oscillation prediction model—ENSO-STformer. The model was fully trained on CMIP6 and SODA datasets and evaluated on GODAS data. The results show that the average skill of this model in predicting the Niño3.4 index at 11-month lead times exceeds those of CanCM4, CCSM3, and GFDLaer04 by 5.1%, 21.6%, and 12.4% respectively. Meanwhile, the Niño3.4 index related skills of the proposed model are significantly better than other deep learning models in the medium and long term. Effective ENSO forecasts can be made for up to 24 months, and the 2015−2016 El Niño event simulation shows strong ability to cope with spring forecast obstacles.

Key words: deep learning; ENSO predicting; spatio-temporal fusion attention mechanism; convolutional neural network; Niño3.4 index