

doi: 10.3969/j.issn.1003-2029.2024.02.001

基于 ConvGRU 和自注意力的海表温度 偏差订正研究

高丽媛, 黄丙湖, 何亚文, 费童涵

(中国石油大学(华东)海洋与空间信息学院, 山东 青岛 266580)

摘要: 海表温度 (Sea Surface Temperature, SST) 是一个重要的海洋物理量, 其准确预报对于水产养殖和预测海洋环境信息等海洋相关领域的研究至关重要。数值预报是目前预报海表温度的一种常用方法, 但数值预报模型所产生的预报结果往往与实际观测数据有偏差, 因此有必要对数值预报产品的偏差进行订正。本文提出了一种结合 ConvGRU 神经网络与自注意力 (Self-Attention, SA) 的新型时空混合海表温度订正模型 (ConvGRU-SA), 对南海海表温度预报数据进行偏差订正, 该模型适用于利用卫星遥感数据对海表温度数值预报产品进行订正。经与 ConvLSTM、ConvGRU 等网络模型对比, 证明了 ConvGRU-SA 模型的优越性, 设置不同的超参数进行实验, 提高模型订正准确率。订正后该区域的海表温度预报均方根误差从 0.52 °C 降低至 0.32 °C, 准确率提高了 38.4%, 优于现有模型。

关键词: SST; 偏差订正; ConvGRU; 注意力机制

中图分类号: P732 **文献标识码:** A **文章编号:** 1003-2029 (2024) 02-0001-09

海洋覆盖地球上大约 70% 的地区, 是陆地的主要水汽来源。海洋为人类提供了丰富的化学资源、矿产资源、动力资源和生物资源, 与人类的生活息息相关。海表温度是指海洋表面附近的水温。在全球气候研究、海洋生态系统研究等海洋相关应用中, 海表温度是一个重要的物理量^[1], 是海洋热力状态的一个重要表征参数, 在海洋环流、大气环流、海气相互作用等方面起着关键作用。海表温度预报精度对海洋灾害预防、航海运输、海洋渔业、国防军事等领域至关重要^[2]。目前, 海表温度预报方法主要是利用海洋数值模式进行预报, 而由于气压场、地形等初始场的不确定, 以及数值模式本身存在的缺陷^[3], 有时预报会产生较大误差。因此, 对数值预报产品进行误差订正是十分有必要的^[4]。

目前, 数值产品订正方法主要分为 3 类: 统计学

订正法、传统机器学习法和深度学习订正法。统计学订正方法是通过建立线性回归方程, 构建海表温度要素订正模型来提高数值预报结果精度。VAN-NITSEM S^[5], TIAN D 等^[6]分别利用模式输出统计 (Mode Output Statistics, MOS) 方法根据回归方程对预报数据进行订正得到订正温度, 效果理想。金巍等^[7]采用最优滑动周期订正方法来对气温模式预报进行订正, 并分析了季节误差对提高准确率的作用。王辉赞等^[8]利用卡尔曼滤波方法对副热带高压数值预报产品进行了订正, 结果表明该方法在研究区域内有较好效果。统计学线性回归方法虽在一定程度上提高了精度, 但存在无法准确描述海洋中非线性物理过程的问题。

传统机器学习是一种基于数据的模型, 能够捕捉要素和目标结果之间的映射关系。孙全德等^[9]利

收稿日期: 2023-10-09

基金项目: 国家自然科学基金资助项目 (41976184)

作者简介: 高丽媛 (1998—), 女, 硕士研究生, 主要从事海表温度预报订正研究。E-mail: 727186309@qq.com

通讯作者: 黄丙湖 (1977—), 男, 博士, 副教授, 主要从事海洋信息可视化研究。E-mail: huangbhu@163.com

用最小绝对收缩和选择算法 (Least Absolute Shrinkage and Selection Operator, LASSO) 及随机森林方法订正近地面 10 m 风速, 订正效果均好于 MOS 方法。曾静^[10]利用支持向量机 (Support Vector Machine, SVM) 基于欧洲中期天气预报中心温度集合控制预报历史资料及实况站点温度资料对模式预报温度进行订正, 有效提高了预报精度。王在文等^[11]利用 SVM 对昌平和奥林匹克公园的 MM5V3 模式预报地面气温和风速进行建模处理, 预报精度有显著提升。门晓磊等^[12]使用岭回归、随机森林等方法对华北气温进行分析, 效果显著。以上研究表明机器学习在数值预报订正领域存在巨大潜力, 但在模型预报精度方面仍有待提升。

近年来, 随着人工智能算法的迅猛发展, 深度学习在许多领域得到广泛应用, 利用深度学习算法进行预报订正成为新的研究趋势^[13]。卷积神经网络 (Convolutional Neural Networks, CNN) 是一种含有卷积运算的前馈神经网络, 适用于空间特征提取问题。循环神经网络 (Recurrent Neural Network, RNN) 的本质是全连接网络, 每个时刻都会对之前信息进行选择并按一定权重保留到当前输出, 1982 年, 美国加州理工学院物理学家 John Hopfield 最早提出了一种用于解决组合优化问题的单层反馈神经网络, 这是 RNN 的雏形。WANG T 等^[14]提出了一种残差单隐层前馈神经网络, 用于基于残差学习的海事数值模型海洋预报订正, 具有明显订正效果。RNN 在处理序列数据上有一定优势, 但存在梯度消失、梯度爆炸等问题。为解决梯度消失、梯度爆炸等问题, RNN 逐步演变为长短期记忆网络 (Long-Short Term Memory, LSTM)^[15-16]。LSTM 最早由 Hochreiter & Schmidhuber 提出, 是一种特殊类型的 RNN, 与 RNN 使用不同函数计算隐状态。LSTM 引入门结构来去除或增加信息到细胞状态, 达到了控制信息选择的效果, 解决了长期依赖问题。滕志伟^[17]根据 LSTM 对雷达回波外推问题的实践, 提出了一种基于 LSTM 的 RET-RNN 模型, 并对 RET-RNN 模型的结构和超参数进行优化, 在外推时效较长时获得较理想效果。2014 年, CHO K 等^[18]提出门控循环单元 (Gated Recurrent Unit, GRU), 并根据 ConvLSTM 模型逐渐修改形成 ConvGRU。张雪薇等^[19]使用 ConvGRU 模型对

西北太平洋部分海域的 Argo 温度数据进行预测分析, 研究结果表明 ConvGRU 模型对 Argo 温度数据的变化趋势具有较好的模拟能力, 预测精度较高且可以很好地解决传统模型因训练数据大无法训练的问题。近年来, 注意力机制广泛应用到深度学习预测中, 例如常见的卷积块注意力模块 (Convolutional Block Attention Module, CBAM)、极化自注意力 (Polarized Self-Attention, PSA) 等, 曾静^[10]使用 LSTM 结合注意力机制建模对多气象因子模式预报温度进行订正, 结果证明该方法订正后温度准确率有明显提升。上述数值预报订正方法虽然都取得了一定成果, 但没有考虑到海表温度要素作为时空相关变量在时间空间上的变化与联系, 同时在模型算法的精度等方面也有待提高。

基于以上问题, 本文在 ConvGRU 中引入了自注意机制, 将 SA 嵌入到一个标准的 ConvGRU 中, 构建 ConvGRU-SA, 提出了一种新的自我注意记忆来记忆在空间和时间域上具有长期依赖性的特征。基于自注意机制, SA 通过聚合输入本身和记忆特征的所有位置的特征来产生特征。该方法不仅考虑了数据集的空间分布影响, 同时考虑了时间信息的重要性, 可以更为准确地校正海表温度数值预报产品。利用模式预报数据和再分析数据进行实验^[20], 证明该方法的可行性。

1 数据获取与预处理

1.1 数据获取

本文所采用数据主要由模式预报数据和遥感卫星数据制作的再分析海表温度数据两部分组成^[21]。模式预报数据为美国国家海洋和大气管理局 (National Oceanic and Atmospheric Administration, NOAA) 发布的全球 0.08° 混合坐标海洋模式 HYCOM 预报数据, 实验中使用的 HYCOM 模式预报产品为未来 24 小时预报产品, 包含海洋温度、盐度和海流等海洋要素, 采用其作为待订正预报数据, 空间分辨率为 0.08°, 时间分辨率为 3 h。再分析海表温度数据为 NOAA 下属国家气候数据中心 (National Climatic Data Center, NCDC) 发布的 Daily OI SST Analysis 数

数据集, 采用其作为真值评估订正精度。

NCDC 提供了两种基于最优插值法 (Optimum Interpolation, OI) 的高分辨率的海表温度数据集, 包括版本 2 (OI.v2) 和版本 2.1 (OI.v2.1)。实验中使用版本 2 (OI.v2) 的海表温度数据集作为真值来评估订正精度^[22]。其空间分辨率为 0.25° , 时间分辨率为 24 h。实验选用时间范围 2021 年 1 月至 2021 年 12 月数据, 空间范围为 12°N — 16°N , 110°E — 114°E 。图 1 为实验数据所在区域。

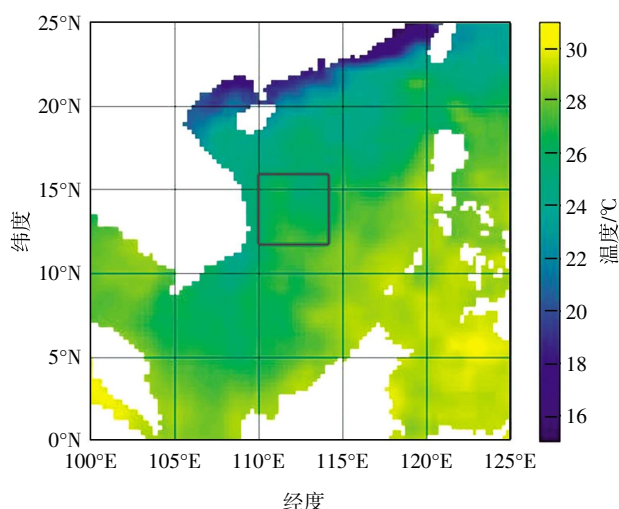


图 1 实验区域

1.2 数据预处理

为统一模式预报数据和遥感卫星数据的时空分辨率, 选取合适的插值方法对数据进行时空分辨率匹配。由于预报数据和标签数据都是格网数据, 选用双线性插值法^[23]对数据进行时空分辨率匹配。双线性插值法的核心思想是在两个方向上分别进行线性插值。双线性插值可以分别在不同方向进行两次单线性插值, 分别求出插值点, 在两个方向插值点的基础上可得出所求点值。时间上, 对 HYCOM 预报数据进行日平均, 将时间分辨率统一为 24 h。空间上, 将 OI SST 数据插值到日平均后的 HYCOM 预报数据格点上, 将空间分辨率统一为 $0.08^\circ \times 0.08^\circ$ 。

为消除不同量级数据间数值的差异, 提升模型训练速度, 对统一时空分辨率后的数据进行 Min-Max 标准化处理。此方法是对输入的数据采用线性变换, 将数据输出结果值映射到 $[0, 1]$, 计算公式如下。

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中, X' 是标准化后的输出数据; X 是输入数据; $X_{\min}$ 是输入数据中的最小值; $X_{\max}$ 是输入数据中的最大值。

2 研究方法

2.1 模型整体结构

本文提出一种新的基于深度学习的混合海表温度数值预报偏差订正方法, 该方法基于 ConvGRU 模型, 并在此基础上添加了自注意力机制, 充分利用输入的时空信息和特征信息, 主要分为空间特征提取、时间特征提取、自注意力机制和输出结果 4 个阶段。其主要思想是利用 ConvGRU 学习海表温度变化过程中的时空关系, 利用卷积操作提取融合多个变量的空间特征, 利用自注意力机制提高三维卷积网络的空间特征利用率, 展现不同环境变量因素对结果的重要性, 同时调整变量中不同历史时刻信息的重要性。自注意力机制的加入可以增强特征提取能力, 提高预报精度。本模型不仅考虑了海表温度场数据的空间相关性, 还考虑了不同时间海表温度场数据之间的时间依赖性与海洋环境变量之间的交互作用。模型框架如图 2 所示。

2.2 ConvGRU 模型

GRU 作为 LSTM 的一种变体, 将遗忘门和输入门合成为更新门, 输出门变为重置门。由更新门控制决定丢弃信息量, 接收来自前一个输入和当前输入的信息, 输出经过 sigmoid 激活函数输出 $0 \sim 1$ 之间的数值, 分别表示信息的遗忘和记忆; 将获取到的门状态送入重置门, 由重置门控制候选状态的计算是否依赖上一时刻状态, 值接近 0 说明不依赖于上一时刻的信息, 接近于 1 则说明上一时刻状态将被保留^[24]。GRU 结构图如图 3 所示。

相对于 LSTM 而言, GRU 参数更少, 收敛速度更快, 大大加快了训练速度。由于 GRU 无法进行空间特征提取, 为了解决海表温度数值预报校正问题, 本文选择由 RNN 扩展算法和卷积神经网络结合的 ConvGRU 模型, 该模型以 GRU 为基础, 将全连接结构改为卷积运算, 从而得到 ConvGRU。其核

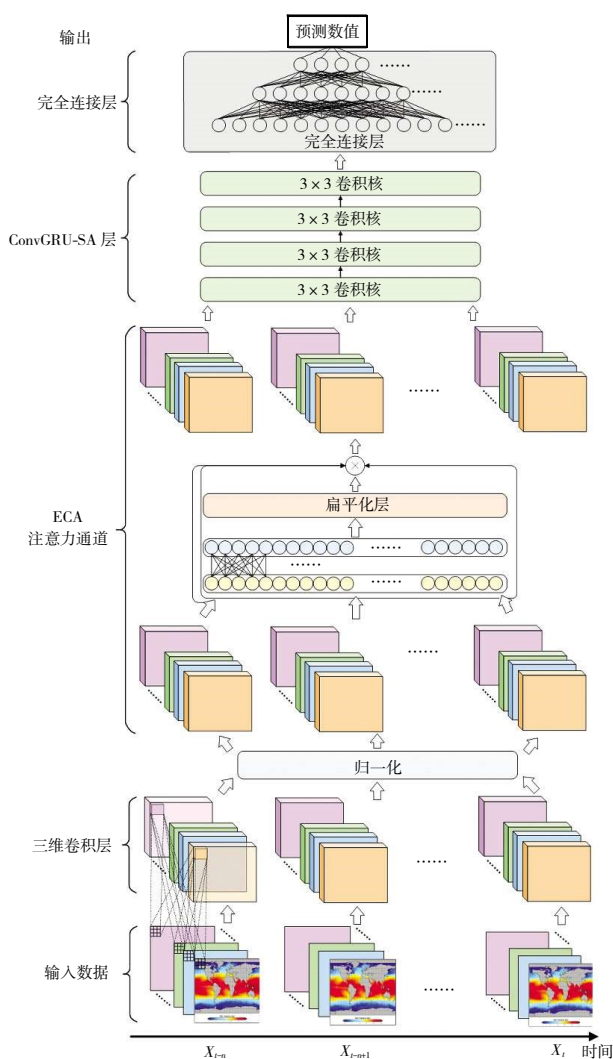


图2 模型框架图

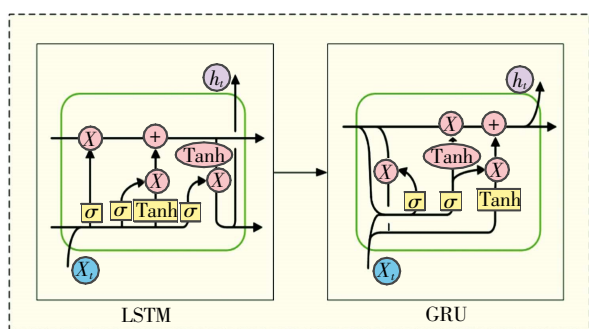


图3 GRU 结构图

思想是将矩阵运算加上卷积操作,既能利用GRU得到时序特征,又能利用卷积计算提取空间特征。ConvGRU模型采用门结构控制信息流动,权重的一部分放到了卷积核内,另一部分放入循环层的循环核。ConvGRU模型计算公式如下。

$$Z_t = \sigma(W_{xz} x_t + W_{hz} H_{t-1}) \quad (2)$$

$$R_t = \sigma(W_{xr} x_t + W_{hr} H_{t-1}) \quad (3)$$

$$H'_t = f(W_{xh} x_t + W_{hz} + R_t \circ (W_{hh} H_{t-1})) \quad (4)$$

$$H'_t = (1 - Z_t) \circ H'_t + Z_t \circ H_{t-1} \quad (5)$$

式中, Z_t 为更新门; R_t 为重置门; H'_t 为候选门(记忆状态); $\circ$ 为卷积操作; $\odot$ 为哈达玛积; x_t 为输入; f 为非线性激活函数; σ 为 sigmoid 激活函数; H 和 W 为输入张量的高和宽。 H_t 为第 t 个时间步的输入向量,即输入序列的第 t 个分量,它会经过一个线性变换(与权重矩阵相乘)。 H_{t-1} 保存的是前一个时间步 $t-1$ 的信息。更新门将这两部分信息相加并投入到 sigmoid 激活函数中,因此将激活结果压缩到 0 到 1 之间。重置门决定了如何将新的输入信息与前面的记忆相结合,更新门帮助模型决定到底要将多少过去的信息传递到未来,或到底前一时间步和当前时间步的信息有多少是需要继续传递的。

2.3 自注意力机制

注意力机制模型模拟了人类快速提取信息的过程,它的核心目标就是从众多信息中选择出对当前任务目标更关键的信息。传统的注意力机制^[29]发生在目标元素和来源中所有元素之间,权重的计算需要目标参与,在解码编码的过程中,注意力机制权值的计算同时需要解码编码两部分的隐状态。而自注意力机制是输入元素或者输出元素之间的注意力机制,在计算权重参数时,只需要在输入部分进行矩阵操作,不需要目标元素中的信息。自注意力机制首次提出是在机器翻译领域,用于绘制输入的全局依赖关系。在计算机视觉任务领域,自注意力机制能够通过计算二元关系函数中特征映射的不同位置之间的成对关系来捕获随机的时空依赖性。通过捕捉到的依赖性计算出参与的特征。

自注意力机制可以解决全连接神经网络对于多个相关输入无法建立相关性的问题,神经网络的输入包含大小不一的向量,向量之间存在千丝万缕的联系,训练过程中经常存在输入向量之间关系无法建立而导致模型训练效果差的问题,引入自注意力机制可以让机器在学习过程中注意到输入中不同部分之间存在的相关性。自注意力机制结构图如图4所示。

自注意力机制计算公式如下。

$$Att(Q, K, V) = \omega(QK^T)V \quad (6)$$

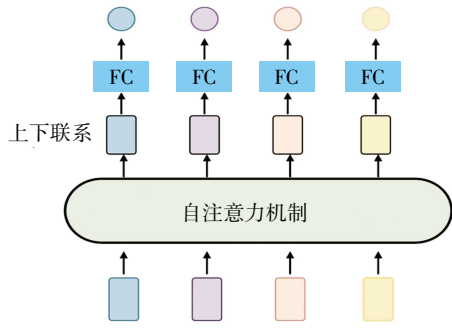


图4 自注意力机制结构图

式中, $Att(Q, K, V)$ 是得到注意力的值, Q 、 K 、 V 分别是查询向量矩阵、键向量矩阵和值向量矩阵, 这 3 个矩阵中每一行分别代表一个对应的向量。 Q 、 K 、 V 通过将输入序列 X 分别乘矩阵 W^q 、 W^k 、 W^v 得到。

2.4 ConvGRU-SA 模型

ConvGRU-SA 模型是自注意力机制和 ConvGRU 的组合, 基础模型是由两部分直接联级建立的, 模型表示如下。

$$\hat{X}_t = Y(x_t), \hat{H}_{t-1} = Y(H_{t-1}) \quad (7)$$

$$Z_t = \sigma(W_{xz}x_t + W_{hz}H_{t-1}) \quad (8)$$

$$R_t = \sigma(W_{xr}x_t + W_{hr}H_{t-1}) \quad (9)$$

$$H'_t = f[W_{xh}x_t + R_t \circ (W_{hh}H_{t-1})] \quad (10)$$

$$\hat{H}'_t = (1 - Z_t) \circ H'_t + Z_t \circ H_{t-1} \quad (11)$$

式中, Y 为自注意力模块; $\hat{X}_t$ 和 $\hat{H}_{t-1}$ 是通过自注意力模块聚合的当前输入信息和上一状态信息。在每一个时间步长内, 注意力模块会选择性地将每个位置的输入特征与所有位置信息加权求和, 使全局依赖性在卷积层所有传播过程中得以捕捉。将自注意力模块嵌入到 ConvGRU 中得到 ConvGRU-SA, 模型结构图如图 5 所示。如将 Y 模块删减, ConvGRU-SA 将退化为标准的 ConvGRU, 同时 Y 模块可灵活地嵌入其他模型中。

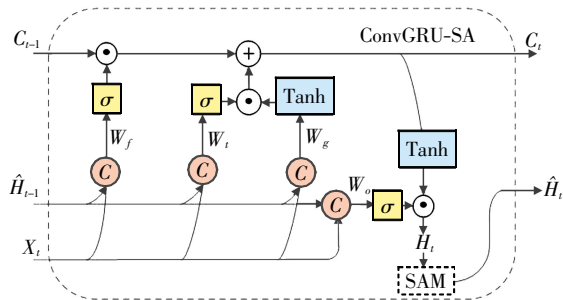


图5 ConvGRU-SA 模型图

3 实验与结果分析

实验运行环境为 Windows10 系统下, 处理器为 Intel Core i5 11/2.4 GHz/16G RAM, 算法使用 Python 语言实现, 采用 PyTorch 框架^[26]搭建神经网络预测模型。根据实验区域的经纬度范围读取相应的预报和标签数据, 选取前 235 组数据作为训练集, 训练新的订正模型参数, 其余 59 组数据作为测试集, 验证模型的学习效果, 测试集所占比例为 20%。

3.1 实验流程

实验训练订正模型数据选用 HYCOM 预报模式数据集和 Daily OI SST Analysis 数据集。通过 PyTorch 实现提取空间数据特征, 使用 ConvGRU 和自注意力机制处理海表温度数据的时空序列关系, 构建混合订正模型并对模型有效性进行评价。构建训练数据集, 对获取到的模式预报数据和再分析数据进行预处理, 根据确定的时间步长和选定的实验区域构建训练模型所需的数据集和标签序列。训练构建好的订正模型, 将构建好的数据集按照 20% 测试集、80% 训练集的比例划分, 将训练集中的 X_{t-n} ($n = 1, 2, 3 \dots$) 作为输入数据, Y_t 作为输出数据, 选取合适的损失函数和评价指标。在实验中, 模型的卷积部分包括一个 3D-CNN 层和一个批量归一化层 (Batch Normalization, BN), BN 层的主要作用是加快模型训练和收敛速度, 缓解梯度爆炸和梯度消失问题, 防止过拟合。网络设置模型 ConvGRU-SA 部分中卷积核的大小为 $3 \times 3 \times 3$, 采用 Adam 优化器进行自适应优化。训练过程中调试学习率, 时间步长等超参数及神经元个数、网络层数等参数, 选择模型训练结果最好的最优组合。记录每次训练后的各项评价指标, 选择效果最好的模型作为最终订正模型。订正模型完成训练后, 将测试数据输入到模型中进行测试, 测试模型结果输出后需经反归一化处理, 最终得到模型的偏差订正结果。通过比较订正前后的评价指标来检验模型的订正效果。

3.2 评价指标

为验证模型的有效性, 本文采用均方误差 (Mean Square Error, MSE)、均方根误差 (Root Mean Square Error, RMSE)、平均绝对误差 (Mean Absolute

Error, MAE) 和平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 4 个指标对该模型进行了评价^[27]。MSE、RMSE、MAE 和 MAPE 可反映订正值和真实值之间的相对误差、计算订正前后的误差离散程度、订正值和真实值之间的平均误差大小, 以及订正值和真实值之间的相对误差百分比, 计算如下。

$$M(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (12)$$

$$R(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

$$A(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (14)$$

$$P(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (15)$$

式中, M 为均方误差; R 为均方根误差; A 为平均绝对误差; P 为平均绝对百分比误差; $\hat{y}$ 为实际观测的平均值; $\hat{y}_i$ 为修正值。

3.3 实验结果

为验证本文所提出的海表温度数值预报偏差订正模型的有效性, 设置对比实验, 将订正结果与 LightGBM、Random Forest 等 4 种常规方法的海表温度订正结果进行比较。以下模型参数均为经过多次试验调整得到的最佳参数, 对于 LightGBM、Random Forest 模型使用机器学习包 sklearn 进行订正; ConvGRU 模型, 学习率 learning rate 设置为 0.009, batch size 设置为 4, 迭代次数 epoches 设置为 49 000, 损失函数选择 MSE 损失函数; 3DCNN-CBAM-CONVLSTM-AT 模型 learning rate 设置为 0.01, batch size 设置为 4, epoches 设置为 100, 损失函数选择 MSE 损失函数。对于提出的 ConvGRU-SA 网络模型, learning rate 设置为 0.001, batch size 设置为 4, epoches 设置为 49 000, 时间步长设置为 4 天, 损

失函数使用 MSE 损失函数。

实验结果对比如表 1 所示。在使用传统机器学习订正方法进行订正时, Random Forest 准确率较 LightGBM 高, RMSE 值为 0.44 °C, MAE 值为 0.34, MAPE 值为 1.23。除本文提出的方法外, ConvGRU 模型结果最好, 其 RMSE 值为 0.43 °C, MAE 值为 0.33, MAPE 值为 1.17。张培军等^[28]在 2020 年提出的海表温度偏差订正模型也利用卫星遥感数据对南海海表温度数值预报模式进行订正。订正后的海表温度预报结果均方根误差由 0.80 °C 降至 0.50 °C, 降低了 37.5%, 而经本文提出的 ConvGRU-SA 订正后, 预报数据集与再分析数据集之间的平均绝对误差为 0.25, 均方误差为 0.10, 均方根误差为 0.32 °C, 相较于未订正之前, 均方根误差降低了 38.4%。与 LightGBM、Random Forest 等方法相比, ConvGRU-SA 模型在订正海表温度数值预报产品时均具备良好性能, 验证了该方法的有效性。为进一步证明本文提出的海表温度数值预报产品偏差订正模型的有效性。将各方法订正后与标签数据的差值进行了可视化, 如图 6 所示。从整体来看, 图 6(d) 所示的订正结果差值与图 6(a) 所示的标签数据差值具有高度相似性, 结合图 6 和表 1, 可以直观地看出, 本文所提出的海表温度数值预报产品偏差订正模型的订正结果最接近标签值。

3.4 参数分析

学习率 (Learning Rate, LR) 是一个重要的超参数, 为使梯度下降法有较好的性能, 学习率的值需要设定在合适的范围内。学习率控制着参数更新速度, 决定目标函数收敛到局部最小值的时间。学习率越大, 参数更新越快, 但也可能错过最优解; 反之学习率越小, 不容易错过最优解, 但更新效率低, 会导致算法长时间无法收敛。基于以上设置了学习率分析实验, 实验结果如表 2 所示。

表 1 海表温度订正实验结果对比

方法名称	P	M	$R/^\circ\text{C}$	A	R 降低/%
订正前	1.43	0.27	0.52	0.40	—
LightGBM	1.17	0.17	0.41	0.32	21.2
Random Forest	1.23	0.19	0.44	0.34	15.4
ConvGRU	1.17	0.18	0.43	0.33	17.3
3DCNN-CBAM-CONVLSTM-AT	1.17	0.19	0.43	0.33	17.3
ConvGRU-SA	0.90	0.10	0.32	0.25	38.4

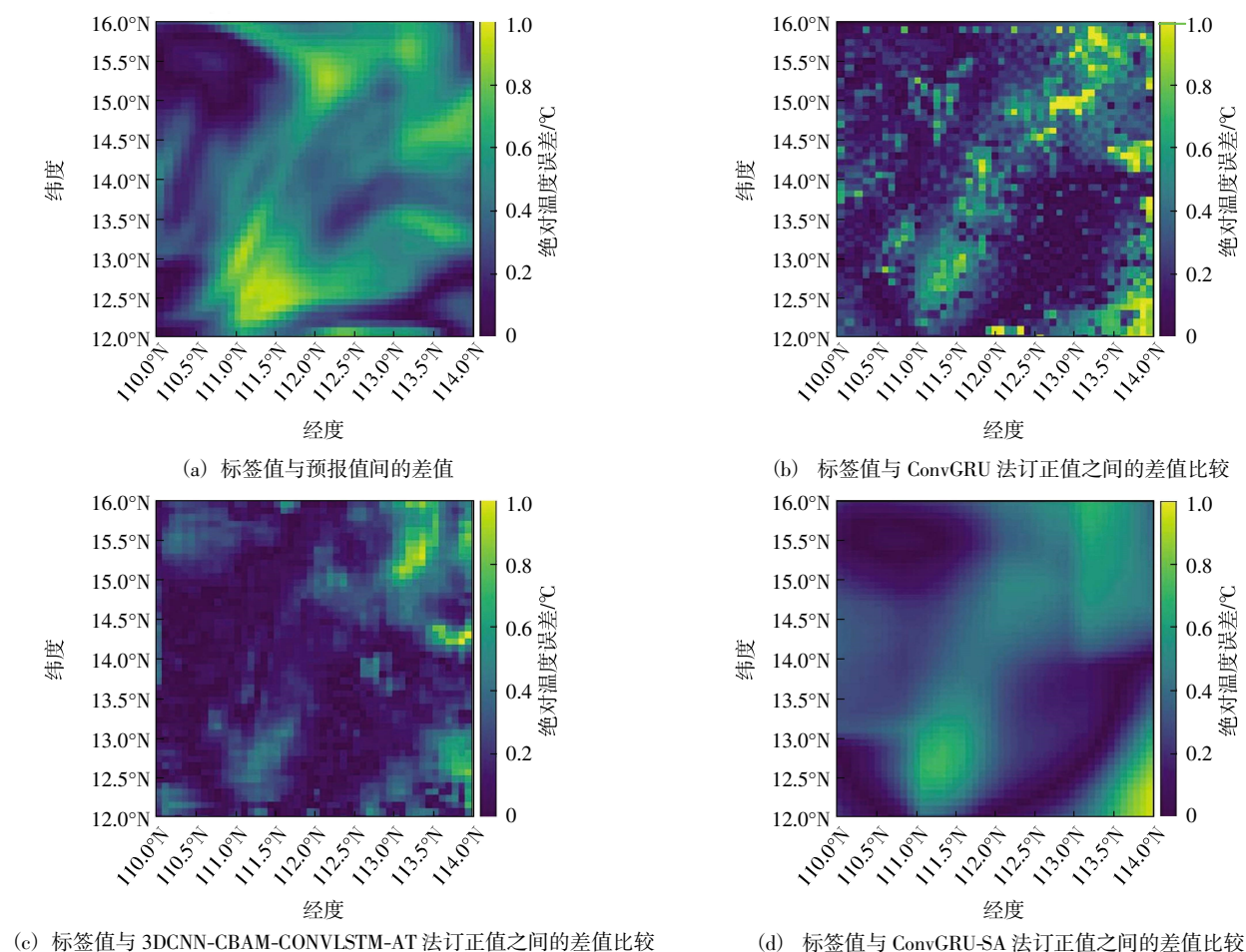


图 6 标签值与订正值之间的差值比较

表 2 不同学习率下海表温度订正实验结果对比

学习率	P	M	$R/^\circ\text{C}$	A
0.012	1.06	0.21	0.45	0.29
0.002	0.93	0.12	0.35	0.26
0.001	0.90	0.10	0.32	0.25
0.000 8	1.19	0.21	0.46	0.33
0.000 5	1.26	0.21	0.45	0.35

从表 2 可以看出, RMSE、MAPE 等指标随 LR 变化而变化, 在设置的对比实验中, 学习率设置为 0.001~0.002 量级时, 均方误差、平均绝对误差、均方根误差、平均绝对百分比误差与 0.000 8 数量级相差较大, 说明本模型在使用的数据集中, 学习率在 0.001~0.002 数量级下效果最好。其中, 当学习率设置为 0.001 时, 各项指标最好, 因此, 使用 0.001 作为本次实验的学习率进行实验。不同学习率下实验结果变化趋势如图 7 所示。

迭代次数 (epoch) 会影响模型的训练效果, epoch 太小会欠拟合, 反之 epoch 太大会出现过拟

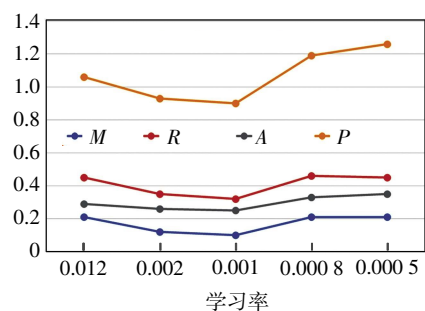


图 7 订正模型在不同学习率的实验结果

合现象, 导致模型泛化能力不高, 测试效果不理想。对不同的数据集和网络模型来说, 最优的迭代次数也是不同的。为确定本数据集的最佳迭代次数, 设置了不同迭代次数进行实验, 实验结果如表 3 所示。

如表 3 所示, 模型迭代次数设置为 49 000~50 000 时, 均方误差、平均绝对误差、均方根误差、平均绝对百分比误差均处于下降趋势, 当迭代次数为 49 000 时效果最好。因此, 使用 49 000 作为本次实验的迭代次数进行实验。不同迭代次数下实验结

果变化趋势如图8所示。

表3 不同迭代次数下海表温度订正实验结果对比

迭代次数	P	M	$R/^\circ\text{C}$	A
50 000	0.98	0.13	0.36	0.27
49 000	0.90	0.10	0.32	0.25
48 000	1.13	0.20	0.44	0.31
47 000	0.96	0.13	0.37	0.27
46 000	0.96	0.13	0.36	0.27

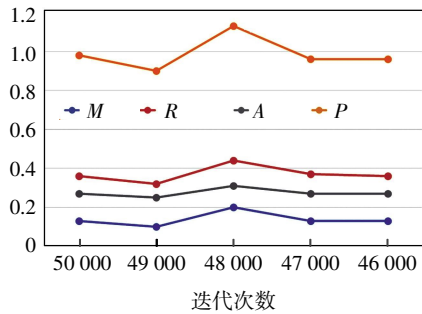


图8 订正模型在不同迭代次数的实验结果

4 结 论

本文采用新的混合海表温度校正模型对海表温

度数值预报产品进行了订正,并对其效果进行了评价。该模型结合了时空信息和海洋环境变量信息,有效订正了海表温度数值预报产品,提高了海表温度预报的准确性。将海表温度预报问题定义为时空序列预测问题,采用卷积门循环单元网络结合自注意力机制(ConvGRU-SA)对时空序列的三维数据进行时间和空间的特征挖掘,避免了传统时间回归模型忽略空间影响带来的预报结果可能在空间上存在的不连续问题。实验证明本文提出的模型较之前的同类模型具有更好的订正效果,订正后该区域的海表温度预报 RMSE 从 0.52°C 降低到 0.32°C ,订正后 RMSE 值降低了 38.4%。

未来将对模型进行进一步的改进,目前模型只订正了海表温度,但海洋内部的温度同样重要。在后续的研究中将把三维空间的数据纳入模型中,实现对海洋三维温度的订正。同时,本文只对部分海域的海表温度预报订正结果进行了订正,为验证方法的普适性,后续研究中将扩大订正区域,纳入更多的海洋要素,实现更大区域更多海洋要素的预报订正。

参考文献:

- [1] FUNK C, HOELL A, et al. The leading mode of observed and CMIP5 ENSO residual sea surface temperatures and associated changes in indo-pacific climate[J]. Journal of Climate, 2015, 28(11): 4309–4329.
- [2] 孙艺迪, 王德, 高志强, 等. 1982—2021 年黄、渤海表层温度变化的线性和非线性特征分析[J]. 海洋科学, 2023, 47(5): 41–53.
- [3] LIU X, WANG Q, MU M, et al. Optimal initial error growth in the prediction of the kuroshio large meander based on a high-resolution regional ocean model[J]. Advances in Atmospheric Sciences, 2018, 35(11): 1362–1371.
- [4] 李奕雯. 对接“智慧海洋”工程发展大湾区海洋预报业务[J]. 海洋与渔业, 2019, 6: 80–82.
- [5] VANNITSEM S. Dynamical properties of MOS forecasts: Analysis of the ECMWF operational forecasting system[J]. Weather and Forecasting, 2008, 23(5): 1032–1043.
- [6] TIAN D, MARTINEZ C J, GRAHAM W D, et al. Statistical downscaling multimodel forecasts for seasonal precipitation and surface temperature over the southeastern united states[J]. Journal of Climate, 2014, 27(22): 8384–8411.
- [7] 金巍, 刘卫华, 高凌峰, 等. 辽宁地区 ECMWF 模式气温预报检验及误差订正研究[J]. 气象与环境学报, 2020, 36(6): 50–57.
- [8] 王辉赞, 张韧, 王彦磊, 等. 基于 Kalman 滤波的副热带高压数值预报误差修正[J]. 热带气象学报, 2006, 22(6): 661–666.
- [9] 孙全德, 焦瑞莉, 夏江江, 等. 基于机器学习的数值天气预报风速订正研究[J]. 气象, 2019, 45(3): 132–142.
- [10] 曾静. 基于机器学习的多气象因子模式预报温度订正模型[D]. 金华: 浙江师范大学, 2020.
- [11] 王在文, 陈敏, MONACHE L D, 等. 相似集合预报方法在北京区域地面气温和风速预报中的应用[J]. 气象学报, 2019, 77(5): 869–884.
- [12] 门晓磊, 焦瑞莉, 王鼎, 等. 基于机器学习的华北气温多模式集合预报的订正方法[J]. 气候与环境研究, 2019, 24(1): 116–124.

- [13] 毛远宏, 孙琛琛, 徐鲁豫, 等. 基于深度学习的时间序列预测方法综述[J/OL]. 微电子学与计算机, 2023, 4: 8–17[2023-08-15]. <https://doi.org/10.19304/J.ISSN1000-7180.2022.0725>.
- [14] WANG T, GAO S, XU J, et al. Correcting predictions from oceanic maritime numerical models via residual learning[C]//OCEANS OCEANS-MTS/IEEE. Kobe: Kobe Techno-Oceans, 2018: 5–8.
- [15] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[J]. Eprint Arxiv, 2014, 8(3): 3–6.
- [16] HOCHBERG S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735–1780.
- [17] 滕志伟. 基于深度学习的多普勒气象雷达回波外推算法研究[D]. 长沙: 湖南师范大学, 2017.
- [18] CHO K, MERRIENBOER B V, BAHDANAU D, et al. On the properties of neural machine translation: En-coder-decoder approaches[EB/OL]. (2014-07-24)[2023-08-15]. <https://arxiv.org/pdf/1409.1259.pdf>.
- [19] 张雪薇, 韩震. Argo 温度数据的 ConvGRU 模型预测分析[J]. 海洋环境科学, 2022, 41(4): 628–635.
- [20] KUPILIK M, WITMER F, MACLEOD E A, et al. Gaussian process regression for arctic coastal erosion forecasting[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(6): 1256–1264.
- [21] BREKKE C, SOLBERG A. Oil spill detection by satellite remote sensing[J]. Remote Sensing of Environment, 2005, 95(1): 1–13.
- [22] REYNOLDS R W, SMITH T M, LIU C, et al. Daily high-resolution-blended analyses for sea surface temperature[J]. Journal of Climate, 2007, 20(22): 5473–5496.
- [23] 辛凯凯. 基于旅行时双线性插值三维地震波场数值模拟[D]. 西安: 长安大学, 2022.
- [24] 张雪薇, 韩震. 基于 ConvGRU 深度学习网络模型的海表面温度预测[J]. 大连海洋大学学报, 2022, 37(3): 531–538.
- [25] ITTI L, KOCH C. Computational modelling of visual attention[J]. Nature Reviews Neuroscience, 2001, 2(3): 194–203.
- [26] 李逸. 基于时空序列网络的气象预报格点订正算法研究[D]. 长沙: 湖南师范大学, 2019.
- [27] FEI T H, HUANG B H, WANG X, et al. A hybrid deep learning model for the bias correction of SST numerical forecast products using satellite data[J]. Remote Sensing, 2022, 14(3): 12–16.
- [28] 张培军, 周水华, 梁昌霞. 基于卫星遥感海表温度数据的南海 SST 预报误差订正[J]. 热带海洋学报, 2020, 39(6): 59–67.

Study of SST Bias Revision Based on ConvGRU and Self-Attention

GAO Liyuan, HUANG Binghu, HE Yawen, FEI Tonghan

(College of Oceanography and Space Informatics, China University of Petroleum, Qingdao 266580, China)

Abstract: Sea surface temperature (SST) is an important physical quantity of the ocean, and accurate forecasting of SST is crucial for research in marine-related fields such as aquaculture and predicting information about the marine environment, and numerical prediction is now a common method for predicting SST. However, the prediction results produced by numerical prediction models often deviate from the actual observations, so it is necessary to revise the deviation of numerical prediction products. In this paper, a new spatio-temporal hybrid SST revision model (ConvGRU-SA) is proposed to be constructed by combining ConvGRU neural network and the attention mechanism to revise the deviation of SST forecast data in the South China Sea, which is suitable for revising the numerical SST prediction products using satellite remote sensing data. Comparison with network models such as ConvLSTM and ConvGRU proves the superiority of the ConvGRU-SA model, and different hyper-parameters are set to conduct experiments to improve the model revision accuracy. The root-mean-square error (RMSE) of the region is reduced from 0.52 °C to 0.32 °C after the revision, and the accuracy is improved by 38.4%.

Key words: SST; bias revision; ConvGRU; attention mechanism