

基于IWOA优化Res-BiGRU深度学习模型的海表温度预测方法

周国良^{1,2}, 周 锋^{2,3}

(1.浙江大学 海洋学院, 浙江 舟山 316021; 2.自然资源部第二海洋研究所 卫星海洋环境动力学国家重点实验室, 浙江 杭州 310012; 3.自然资源部 长三角海洋生态环境野外科学观测研究站, 浙江 舟山 316022)

摘 要:海表温度 (sea surface temperature, SST) 是研究海洋环境和气候变化的重点要素之一。现有基于深度学习的 SST 预测算法存在调参耗时和预测精度偏低的问题, 基于此本文提出了一种基于改进鲸鱼优化 (improved whale optimization algorithm, IWOA) 的残差双向门控循环神经网络 (residual bidirectional gated recurrent neural network, Res-BiGRU) 的 SST 预测算法, 并选取东海海域的二维月均 SST 空间分布和一维日均 SST 时间序列数据验证了算法的可行性和有效性。针对二维月均 SST 空间分布数据, 首先利用 Res-BiGRU 神经网络提取 SST 序列信号的空间和时间特征, 在解码阶段引入时空注意力机制对提取特征的权重分配进行调整。其次采用 IWOA 算法对训练参数进行寻优, 提高训练参数调整效率和 Res-BiGRU 模型精度。针对一维日均 SST 时间序列数据, 引入集合经验模态分解法 (ensemble empirical mode decomposition, EEMD) 对 SST 序列信号进行预处理, 然后利用 IWOA-Res-BiGRU 进行特征提取和 SST 预测。实验结果表明, 本研究提出的基于 IWOA 优化 Res-BiGRU 算法对二维月均 SST 和一维日均 SST 的预测均方误差分别可达到 0.35 °C 和 0.09 °C, 预测结果优于其他模型, 可成为 SST 预测的重要工具。

关键词:海表温度; 鲸鱼优化; 深度学习; 时序信号; 残差连接

中图分类号:P731.3

文献标识码:A

文章编号:1007-6336(2024)05-0806-11

Sea surface temperature prediction method based on an IWOA optimized Res-BiGRU deep learning model

ZHOU Guoliang^{1,2}, ZHOU Feng^{2,3}

(1.College of Ocean, Zhejiang University, Zhoushan 316021, China; 2.Second Institute of Oceanography, State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou 310012, China; 3.Observation and Research Station of Yangtze River Delta Marine Ecosystem, Ministry of Natural Resources, Zhoushan 316022, China)

Abstract: Sea surface temperature (SST) is one of the key elements in the study of marine environment and climate change. The existing SST prediction algorithm based on deep learning has the problems of time-consuming parameter adjustment and low prediction accuracy. Based on this, a SST prediction algorithm based on improved whale optimization (IWOA) residual bidirectional gated recurrent neural network (Res-BiGRU) is proposed. The two-dimensional monthly average SST spatial distribution and one-dimensional SST time series data in the East China Sea are selected to verify the feasibility and effectiveness of the algorithm. For the two-dimensional monthly average SST spatial distribution data, the Res-BiGRU neural network is used to extract

收稿日期:2023-05-31, 修订日期:2024-02-02

基金项目:国家自然科学基金联合基金重点支持项目 (U23A2033); 自然资源部中德海洋科技合作项目 (NECO)

作者简介:周国良 (1996—), 男, 河南洛阳人, 硕士, 主要研究方向为海洋人工智能、深度学习, E-mail: zhoughuoliang@zju.edu.cn

the spatial and temporal features of the SST sequence signal, and the time and space attention mechanism is introduced in the decoding stage to adjust the weight distribution of the extracted features. Secondly, the IWOA algorithm is used to optimize the training parameters to improve the efficiency of training parameter adjustment and the accuracy of Res-BiGRU model. Aiming at the one-dimensional daily average SST time series data, the ensemble empirical mode decomposition (EEMD) method is introduced to preprocess the SST sequence signal, and then the IWOA-Res-BiGRU is used for feature extraction and SST prediction. The experimental results show that the proposed Res-BiGRU algorithm based on IWOA optimization can predict the mean square error of two-dimensional monthly mean SST and one-dimensional daily mean SST by 0.35 °C and 0.09 °C, respectively. The prediction results are better than other models and can become one of the important tools for SST prediction.

Key words: sea surface temperature; whale optimization; deep learning; timing signal; residual connection

海表温度是研究海气交换和海水热量的重要变量,对气候变化、海洋生态系统和渔业发展等都具有重要影响。研究表明,厄尔尼诺、海平面上升等现象和海水温度变化存在密切关系,因此,研究 SST 时空特征的预测算法在气候预测、海洋环境安全保障、海洋生态预警等领域具有积极意义^[1-2]。

按照设计方式的差异,可将目前的 SST 预测算法分为数值模拟法和数据驱动法^[3-4]。前者是利用海洋的热力学和动力学方程,通过提供初值和边界资料进行预测,具有明确的物理意义,但建模较为困难且计算复杂度高;后者是基于统计学方法,从 SST 数据中挖掘时空特征,通过建立数据驱动模型来拟合输入输出的非线性关系,代表性方法有线性回归^[5]、支持向量机^[6]、LightGBM^[7]等机器学习模型。

近年来,人工智能技术在各领域都有广泛应用,如智能驾驶、故障诊断、灾害预测等,给其他学科提供了新的研究方向^[8]。传统机器学习模型存在处理大规模数据乏力,且无法准确提取 SST 数据中的时空分布特征等问题。当前,基于深度学习的 SST 预测算法得到了广泛研究。Zhang 等^[9]考虑到 SST 数据的时间关联性,提出了一种基于长短期记忆网络(long and short term memory network, LSTM)的 SST 预测方法。Yang 等^[10]在 LSTM 的基础上,为进一步融合 SST 数据的空间特征,将卷积神经网络(convolutional neural network, CNN)和 LSTM 相结合用于 SST 预测。张雪薇等^[11]采用卷积门控神经单元(convolution gated neural unit, ConvGRU)实现了

对西北太平洋部分海域 SST 的预测,平均预测精度达到 97.31%。Zheng 等^[12]提出了一种用于 SST 预测的深度学习模型,基于 CNN 创建了 4 个堆叠的复合层来接收不同分辨率的 SST 图像,并在模型输出中增加偏差修正来提高 SST 预测精准度。但上述算法并未考虑不同 SST 数据在时空特征上的不均衡性, SST 预测结果可能受某些时空特征的影响更大,例如冬季 SST 更关注离大陆较近的低温区域。因此,王丽娜等^[13]将注意力机制(attention mechanism, AM)引入模型的解码器阶段,调整提取到的 SST 时空特征的权重,但其采用的 AM 不能动态关注不同时刻、不同空间的局部特征。因此,从现有的 SST 预测的深度学习算法来看, SST 的时间与空间特征挖掘仍然不够充分,且没有对深度学习模型的训练参数进行优化,导致人工调参成本较高。

综上所述,为了有效提取 SST 的时空特征和提高 SST 模型的预测精度,本研究提出了一种基于深度学习的 SST 预测方法,该方法采用集合经验模态分解算法对 SST 时序信号进行预处理,降低 SST 序列的多尺度复杂度和提高 SST 序列的平稳性,然后在特征提取阶段,采用一维残差神经网络(one-dimensional residual neural network, 1D ResNet)和双向门控循环单元(bidirectional gated recurrent unit, BiGRU)相结合的网络结构(Res-BiGRU)能有效提取 SST 的空间和时间特征。在解码阶段,该方法利用时空注意力机制(time and space attention mechanism, TSAM)对提取的 SST 时空特征图进行调整,以关注时空特征图中的重要信息,最后经过全连接层输出预测

的 SST。同时,为获取最优的模型训练参数,将 IWOA 算法引入 Res-BiGRU 训练参数优化中,以提高调参的效率和模型精度。

1 材料与方法

1.1 数据来源与预处理

1.1.1 数据来源

实验数据来自最佳内插海表温度(optimum interpolation sea surface temperature, OISST)数据集,由美国国家海洋和大气管理局地球系统研究实验室物理科学部提供(<https://www.esrl.noaa.gov/psd/>)。本文选取了其中的 1982 年 1 月—2022 年 12 月的二维月均 SST 数据和 2017 年 1 月 1 日—2022 年 12 月 31 日的一维日均 SST 数据。

选择东海海域作为研究对象,该海域的 SST 具有显著的季节特征,同时具有较强的锋面结构,时空分布特征有较好的代表性。该海域的具体经纬度范围为 23°N—33°N、117°E—130°E,图像空间分辨率为 0.25°×0.25°,二维月均 SST 对应的空间格点为 41×53。一维日均 SST 数据为东海的部分区域平均后的时间序列,区域平均的经纬度范围为 28°N—29°N、124°E—126°E。

初步处理后,二维月均 SST 数据为 480 个,以 30 个月的历史月均 SST 观测值来预测未来 6 个月的月均 SST 值;一维日均 SST 数据长度为 2190 d,以 35 d 的历史日均 SST 观测值来预测未来 7 d 的日均 SST 值。

1.1.2 EEMD 分解

EEMD 分解算法首先在原始信号中加入高斯白噪声序列 $e_i(t)$, 然后利用经验模态分解(empirical mode decomposition, EMD)分解出不同特征尺度、尖峰和波动幅度降低的时序分量数据^[14]。本文对 SST 时序信号进行 EEMD 分解,降低了 SST 信号噪声对深度学习模型的影响。具体计算步骤如下:

(1) 向原始 SST 序列 $x(t)$ 中加入 $e_i(t)$ 后得到新序列 $x_i(t) = x(t) + e_i(t)$;

(2) 利用 EMD 算法对 $x_i(t)$ 进行多尺度分解,得到 k 个模态分量 $c_{i,j}(t)$ 和 1 个残余分量 $r_0(t)$;

(3) 向原始 SST 序列 $x(t)$ 中重新加入高斯白

噪声,重复步骤(2),并对 $c_{i,j}(t)$ 进行均值化处理,得到:

$$\begin{cases} x_i(t) = \sum_{j=1}^k c_{i,j}(t) + r_0(t) \\ c_j(t) = \frac{1}{N} \sum_{i=1}^N c_{i,j}(t) \end{cases} \quad (1)$$

(4) 得到最终分解后的序列:

$$x(t) = \sum_{j=1}^k c_j(t) + r(t) \quad (2)$$

式中: N 表示第 N 次加入高斯白噪声。

如图 1 所示,2022 年前 200 d 的一维日均 SST 数据经过 EEMD 可以分解为两个模态分量和一个残余分量。

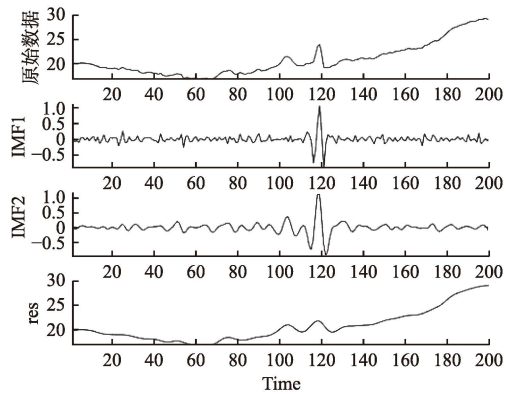


图 1 一维日均 SST 时序数据 EEMD 分解结果

Fig. 1 EEMD decomposition results of one-dimensional SST time series data

1.2 改进鲸鱼优化算法

常见的参数优化算法包括麻雀搜索算法(sparrow search algorithm, SSA)^[15]和鲸鱼优化算法(whale optimization algorithm, WOA)^[16]等。WOA 是 Mirjalili 等^[17]模拟鲸鱼捕食特点而设计的一种参数优化算法,具有算法原理简单、参数设置步骤少等优点。

以本文对神经元单元数和初始学习率两个参数进行优化为例,确定合适的损失函数 $loss$, 设 20 组初始参数值,即鲸鱼的种群数量。每一组初始值包含神经元单元数和初始学习率,代表一只鲸鱼的初始位置。优化过程共分为包围猎物、螺旋气泡网攻击和搜索目标猎物 3 个阶段。

第一阶段,包围猎物阶段。计算全部鲸鱼个

体与目标的距离,即每组参数的 $loss$ 。以 $loss$ 最小的鲸鱼为当前步的最优解,其他鲸鱼个体根据当前最优解更新自身位置。该过程如下所示^[18]:

$$\begin{cases} D = |CX^*(k) - X(k)| \\ X(k+1) = X^*(k) - AD \end{cases} \quad (3)$$

式中: D 表示某一鲸鱼个体与当前最优解的距离; k 表示迭代步数; $X^*(k)$ 和 $X(k)$ 分别表示当前最优解的位置(即当前最优的参数值)和某一鲸鱼个体的位置(即某一组参数值); A 和 C 为控制鲸鱼个体位置更新的参数向量,由下式定义:

$$\begin{cases} A = a(2r_1 - 1) \\ C = 2r_2 \\ a = 2(1 - k/k_{\max}) \end{cases} \quad (4)$$

式中: r_1 和 r_2 表示 $[0,1]$ 的随机值; $k_{\max}$ 表示最大迭代次数。

第二阶段,螺旋气泡网攻击阶段。鲸鱼个体靠近目标后(即 $loss$ 减小),改为通过螺旋式由外向内更新位置,数学模型如下:

$$X(k+1) = De^{bl} \cos(2\pi l) + X^*(k) \quad (5)$$

式中: l 为 $[0,1]$ 间的随机数; b 为常数。

第三阶段,搜索目标猎物阶段。通过判别控制参数 $|A|$ 的值来决定执行全局搜索($|A| \geq 1$)还是局部搜索($|A| < 1$)。执行全局搜索时,将会随机选择一只鲸鱼个体(即一组参数值)代替当前最优解作为其他鲸鱼个体位置更新的依据,以此避免局部最优的情况,计算式如下:

$$\begin{cases} D = |CX_{rand}(k) - X(k)| \\ X(k+1) = X_{rand}(k) - AD \end{cases} \quad (6)$$

式中: $X_{rand}(k)$ 表示第 k 步迭代随机选择的鲸鱼个体的位置。

由上述分析可知,参数 a 控制了 WOA 算法的寻优进程,但参数 a 在 WOA 算法中是线性递减的,导致其在面对复杂问题时寻优时间增加,且难以跳出局部最优解,因此引入正弦非线性控制对参数 a 的迭代过程进行优化。改进后 a 的迭代计算式如下:

$$a = 2 \left[1 - \sin \left(\frac{\pi k}{2k_{\max}} \right) \right] \quad (7)$$

如图 2 所示,改进后的 a 在迭代过程的猎物搜索阶段能够快速下降,提高全局搜索速度;而在螺旋气泡网攻击阶段能够降低下降速率,进而

提高局部搜索最优解的效果。

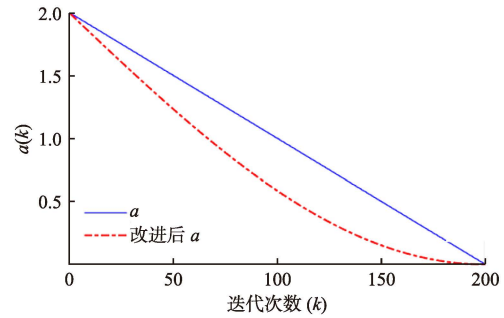


图 2 改进后 a 的迭代曲线

Fig. 2 The iterative curve of a after improvement

1.3 网络模型

1.3.1 一维残差卷积

CNN 中的卷积层通过卷积运算和非线性激活函数能够有效提取原始数据的深层特征,该过程可以表示为:

$$C_l = f(W_l * x_{l-1} + b_l) \quad (8)$$

式中: $f(\cdot)$ 为非线性的激活函数; W_l 和 b_l 分别为第 l 卷积层的权重和偏置。

理论上, CNN 隐藏层的层数越多,其非线性表达能力越强,但过多的网络层数会增加模型训练中梯度消失和梯度爆炸的概率。因此,为降低网络模型参数,本文采用一维残差卷积块对 SST 信号的空间特征进行提取^[19],其结构如图 3 所示,每层卷积层(1D Conv)进行批归一化(batch normalization, BN)处理,激活函数选择线性整流函数 ReLU。

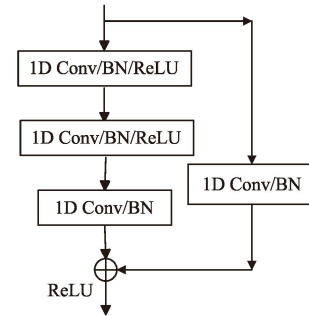


图 3 一维残差卷积块

Fig. 3 One-dimensional residual convolution block

可以发现,一维残差卷积块在正向计算过程中,通过跳跃连接将每个单元的输入和映射输出一起输入下一个单元,即假定一维残差卷积块的

输入为 x , 经残差单元计算后的输出为 $F(x)$, 则下一单元的输入 $H(x)$ 可表示为:

$$H(x) = F(x) + x \quad (9)$$

1.3.2 BiGRU

门控神经单元(gated recurrent unit, GRU)是一种能够捕捉时间序列中依赖关系的神经网络模型, 是 LSTM 的简化版^[20], 主要由重置门和更新门两部分组成, 该方法的结构如图 4 所示。

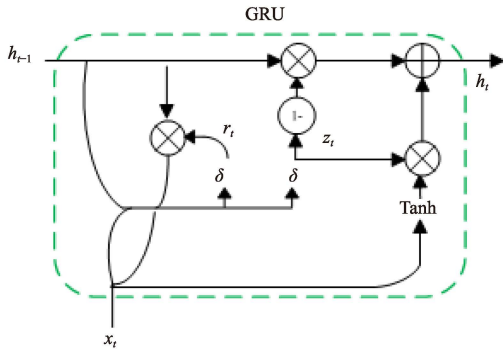


图 4 GRU 内部结构^[20]

Fig. 4 Internal structure of GRU

其中, 更新门的输出由式(10)定义:

$$z_t = \delta(x_t w_z + h_{t-1} w_z + b_z) \quad (10)$$

式中: h_t 和 h_{t-1} 分别表示当前时刻和前一时刻的隐藏状态; $\sigma(\cdot)$ 表示 Sigmoid 函数, 用于控制 z_t 处于 $[0, 1]$ 之间, 0 和 1 分别表示完全舍弃和保留 h_{t-1} ; w_z 和 b_z 分别表示更新门权重和偏置。

重置门 r_t 用于控制过去信息的遗忘程度, 由

下式定义:

$$r_t = \delta(x_t w_r + h_{t-1} w_r + b_r) \quad (11)$$

因此, GRU 的当前时刻输出可以表示为:

$$\begin{cases} \tilde{h}_t = \tanh(x_t w_h + r_t h_{t-1} w_h + b_h) \\ h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t \end{cases} \quad (12)$$

标准的 GRU 网络虽然能够学习到时序信号的部分历史信息, 却无法对后向信息进行利用, 因此 BiGRU 通过将 GRU 进行反向并联, 即将正向输出 $\vec{h}_t$ 和反向输出 $\overleftarrow{h}_t$ 直接相加作为 BiGRU 的隐藏状态输出, 能同时有效提取时序信号的前、后向特征^[21-22], 因此本文采用 BiGRU 对 SST 时序信号进行时间特征提取。

1.3.3 时空注意力机制

为加强对 SST 数据关键细节时空特征的提取, 本文提出了一种时空注意力机制 TSAM。其中, 空间注意力机制(spatial attention mechanism, SAM)原理如图 5 所示。输入特征首先经过平均池化和最大池化, 然后对拼接得到的特征图进行不同的卷积操作(卷积核分别为 3 和 5), 并经过 ReLU 函数进行非线性映射, 通过矩阵相加来实现特征融合, 从而实现多尺度特征信息的提取。SAM 的计算原理可表达为:

$$M_s(F) = \delta\{\text{ReLu}[f_{3 \times 3}(PL)] \oplus \text{ReLu}[f_{5 \times 5}(PL)]\} \quad (13)$$

式中: PL 表示平均池化和最大池化后得到的特征图; $f_{a \times a}$ 表示卷积核大小为 a 的卷积操作; $\oplus$ 表示矩阵加法计算。

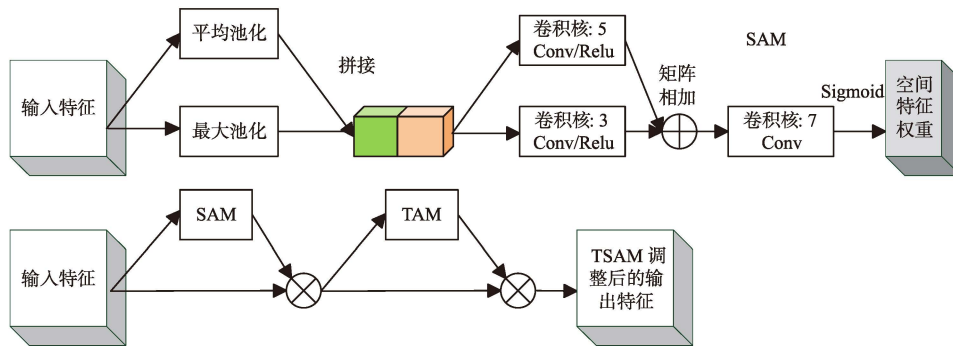


图 5 SAM 原理图

Fig. 5 SAM block diagram

时间注意力机制(time attention mechanism, TSM)与 SAM 类似, 输入特征 F 经过平均池化后, 沿时间维度压缩为一维矩阵 $T \in R^{1 \times C}$, 其中

F 沿时间维度的第 i 个元素的池化公式如下:

$$T_i = \frac{1}{C} \sum_{i=1}^C F_i(i) \quad (14)$$

然后, 对 T 进行卷积操作和 Sigmoid 函数归一化处理, TSM 的计算原理如下:

$$M_T(F) = \delta[f_{3 \times 3}(T)] \quad (15)$$

因此, TSAM(串行排列)的计算公式如下:

$$\begin{cases} F_1 = M_S(F) \odot F \\ F_2 = M_T(F_1) \odot F_1 \end{cases} \quad (16)$$

式中: F_1 和 F_2 分别表示经过 SAM 和 TSAM 优化后的特征图。

1.3.4 Res-BiGRU

本文提出了一种用于 SST 时序数据预测的深度学习网络模型(Res-BiGRU), 其网络结构如图 6 所示, 其中 None 表示动态批量大小。

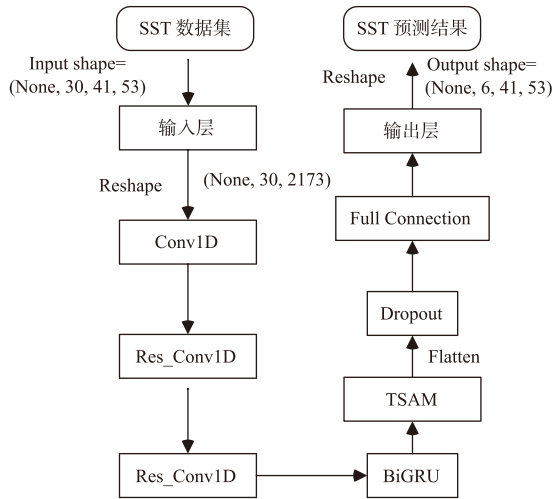


图 6 Res-BiGRU 的模型结构

Fig. 6 The model structure of Res-BiGRU

对于二维月均 SST 数据, 训练集样本经过输入层进入 Res-BiGRU, 输入尺寸(Input shape)为 (None, 30, 41, 53); 通过一维卷积(Cov1D)和两层残差卷积(Res_Conv1D)提取 SST 数据的空间特征; 然后进入 BiGRU 层, 提取 SST 数据的时间特征, 并经过多头注意力(Attention)层调整提取的特征图, 以突出 SST 图像的有效特征; 最后, 将矩阵展平(Flatten)后经随机失活(Dropout)处理通过全连接层(Full connection)输出预测的 SST 数据, 输出尺寸(Output shape)为 (None, 6, 41, 53)。

对于一维日均 SST 数据, 长度为 2190 的一维日均 SST 数据经过 EEMD 分解后转变为 7 个模态分量和 1 个残余分量, 利用 Res-BiGRU 对其进行建模, 输入层经矩阵变形(Reshape)后

变为 (None, 35, 8), 输出尺寸(Output shape)为 (None, 7, 1)。

1.3.5 算法流程

根据上述分析, 基于 IWOA 优化 Res-BiGRU 的海洋表面温度预测算法流程如图 7 所示。

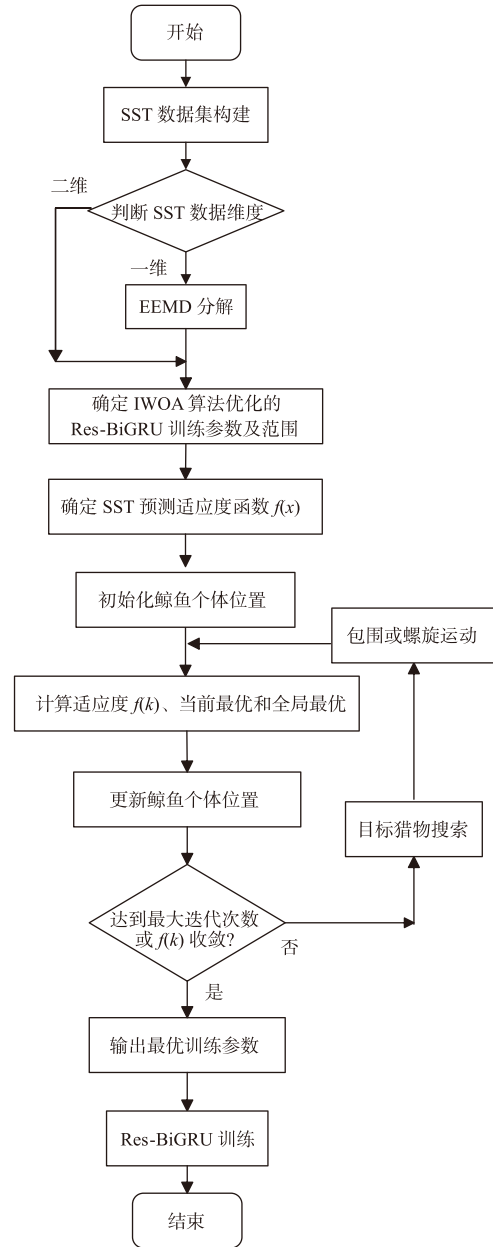


图 7 算法流程

Fig. 7 Algorithmic procedure

(1) 判断 SST 数据维度, 一维日均 SST 数据可以利用 EEMD 分解, 降低 SST 数据中的噪声分量, 并利用 Min-Max 对输入的 SST 数据进行归一化处理。

(2) 确定 IWOA 算法优化 Res-BiGRU 模型的参数, 包括 BiGRU 的单元数和初始学习率 Lr 。

(3) 确定适应度函数 $f(x)$ 为均方根误差 (root mean square error, $RMSE$), 计算式如下:

$$f(x): RMSE = \frac{1}{N \times M} \sqrt{\sum_{n=1}^N \sum_{m=1}^M (y_n^m - \widetilde{y}_n^m)^2} \quad (17)$$

式中: n 表示预测样本数; m 表示某样本第 m 个特征; $f(x)$ 收敛值越低, 表明优化的参数值越有效。

(4) 开始利用 IWOA 算法优化 Res-BiGRU 模型, 当达到最大迭代次数或适应度 $f(k)$ 已经收敛时, 则输出最优参数。

(5) 利用最优训练参数训练 Res-BiGRU, 并保存训练模型供测试调用。

2 结果与讨论

2.1 评价指标

为评价不同 SST 预测模型的性能, 选用一维平均绝对百分比误差均值 (one-dimensional averaged mean absolute percentage error, 1_AMAPE)、 2_AMAPE 和 $RMSE$ 作为评价指标, 计算表达式如下:

$$1_AMAPE = \frac{1}{N \times I} \sum_{n=1}^N \sum_{i=1}^I \frac{|y_i(i) - y_p(i)|}{|y_i(i)|} \quad (18)$$

$$2_AMAPE = \frac{1}{N} \sum_{n=1}^N \frac{1}{I \times J} \sum_{i=1}^I \sum_{j=1}^J \frac{|y_i(i, j) - y_p(i, j)|}{|y_i(i, j)|} \quad (19)$$

式中: i 和 j 分别表示空间格点的坐标; N 表示样本数; $y_i(i, j)$ 和 $y_p(i, j)$ 分别表示对应空间格点的观测值和预测值。

2.2 实验配置与训练参数优化

2.2.1 实验配置

实验计算机配置为 Windows 10, CPU 为 Intel Core i7-9700F, 深度学习框架为 TensorFlow2.10.0。训练集、验证集和测试集的划分比例为 8 : 1 : 1。

2.2.2 训练参数优化

Res-BiGRU 的超参数设置如下: Dropout 概率为 0.3, Batch size 为 64, Cov1D 和两层 Res_Conv1D 的卷积核大小分别为 5、3 和 3。损失函数 $loss$ 为 $RMSE$, 优化器为随机梯度下降 SGD,

训练轮数为 200, BiGRU 的单元数为 256, 初始学习率 Lr 为 0.002。

其中, BiGRU 的单元数和初始学习率 Lr 分别对比了 IWOA、WOA 和 SSA 对两个训练参数的寻优效果。Res-BiGRU 利用 3 种优化算法对一维日均 SST 数据进行训练, $loss$ 收敛曲线如图 8 所示。

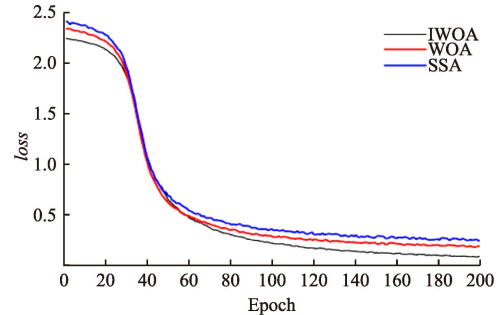


图 8 不同优化算法 $loss$ 收敛曲线

Fig. 8 Different optimization algorithms $loss$ convergence curves

由图 8 可知, Res-BiGRU 利用 IWOA 算法优化的训练参数进行训练, 其 $loss$ 收敛值更小, 相较于 SSA 和 WOA 分别降低了约 0.18 和 0.1, 进而降低了人工调参成本。

2.3 消融实验

为验证 Res-BiGRU 中各模块对模型精度的影响, 开展了消融实验, Res-BiGRU 对测试集二维月均 SST 数据消融实验结果如表 1 所示。

表 1 测试集消融实验结果

Tab.1 Ablation test results in test set

模型编号	替换模块	去除模块	$loss/^{\circ}\text{C}$
<i>a</i>	/	/	0.35
<i>b</i>	/	TSAM(串行)	0.40
<i>c</i>	TSAM(并行)	TSAM(串行)	0.36
<i>d</i>	AM ^[13]	TSAM(串行)	0.38
<i>e</i>	Cov1D	Res_Cov1D	0.46
<i>f</i>	GRU	BiGRU	0.43

其中, 模型 *a* 为本文算法, 其中 TSAM 采用图 5 所示的串行排列, 其 $loss$ 收敛值最低; 模型 *b* 为 Res-BiGRU 去除掉 TSAM 层, 其 $loss$ 收敛值上升了 0.05, 表明 TSAM 层(串行)的引入提高了模型精度; 模型 *c* 为 Res-BiGRU 将 TSAM 层

的排列方式由串行变为并行,其 $loss$ 上升了 0.01; 模型 d 为文献 [13] 采用的 AM 替换 TSAM 层 (串行), 其 $loss$ 上升了 0.03; 模型 e 将 Res_Conv1D 层替换为 Conv1D, 其 $loss$ 达到 0.46; 模型 f 为用 GRU 替换掉 BiGRU 层, 其 $loss$ 达到 0.42, 低于采用 BiGRU 层的训练结果。

表 1 模型 a 、 b 、 e 和 f 的实验结果验证了 TSAM、Res_Conv1D 和 BiGRU 层引入的有效性。从模型 a 和 c 实验结果可知, 串行 TSAM 排列的 SST 预测精度更高; 从模型 a 、 c 和 d 实验结果可知, 本研究提出的 TSAM 更能有效提高 Res-BiGRU 的性能。

2.4 实验结果分析

2.4.1 一维日均 SST 数据

为验证本文提出的算法在一维日均 SST 预测中的性能(EEMD+IWOA-Res-BiGRU), 将该算法与文献提出的 VMD-CNN-LSTM^[23] 和 VMD-LSTM-BLS 模型^[24] 进行比较。使用 2022 年 3 月 1 日—2022 年 4 月 4 日和 2022 年 7 月 1 日—2022 年 8 月 4 日的日均 SST 数据, 预测未来一个星期的 SST 数据。

同时, 为了验证 EEMD 分解算法应用于 SST 时序数据预处理的有效性, 对比了去除 EEMD 分解环节 (IWOA-Res-BiGRU) 对预测结果的影响。4 种算法模型均采用相同的训练集进行训练和相同的测试集进行测试, 实际观测值用 True 表示, 实验结果如图 9 和图 10 所示。表 2 为 4 种算法预测测试集一维日均 SST 数据的评价指标平均值。

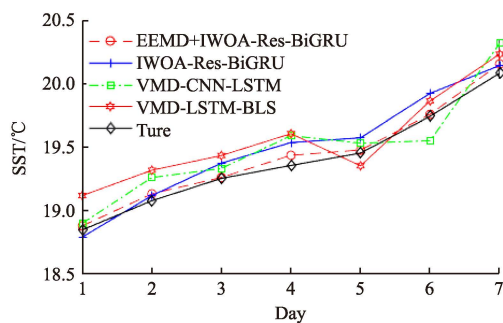


图 9 不同模型对 2022 年 4 月 5 日至 4 月 11 日 SST 的预测结果

Fig. 9 Prediction results of different models for SST in the April 5-11, 2022

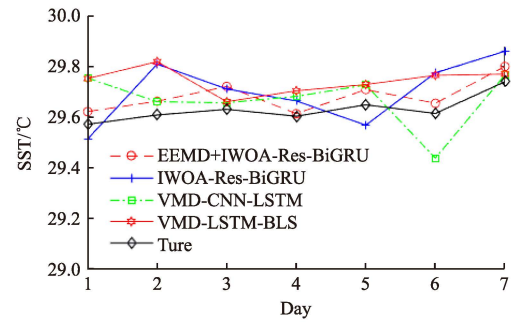


图 10 不同模型对 2022 年 8 月 5 日至 8 月 11 日 SST 的预测结果

Fig. 10 Prediction results of different models for SST in the August 5-11, 2022

表 2 不同算法评价指标

Tab.2 Different algorithm evaluation index

算法编号及类型	$RMSE/^\circ C$	$1_AMAPE/(%)$
1 IWOA-Res-BiGRU	0.13	0.46
2 EEMD+IWOA-Res-BiGRU	0.09	0.22
3 VMD-LSTM-BLS	0.19	0.53
4 VMD-CNN-LSTM	0.13	0.31

结合表 2、图 9 和图 10 可知, 针对一维日均 SST 时序数据, 本文提出的 EEMD+IWOA-Res-BiGRU 模型的预测精度最优, 其 $RMSE$ 和 1_AMAPE 分别为 $0.09^\circ C$ 和 0.22% , 明显优于其他模型。对比算法 1 和算法 2 的实验结果可知, EEMD 分解算法的引入有效降低了 SST 预测误差, $RMSE$ 和 1_AMAPE 分别降低了 $0.04^\circ C$ 和 0.24% , 这与 VMD-CNN-LSTM 和 VMD-LSTM-BLS 算法采用 VMD 分解的结论一致, 即引入 SST 预处理技术能够有效提高深度学习模型的性能。

对比算法 1、算法 3 和算法 4 的结果可知, 采用“卷积神经网络-循环神经网络”架构能够更有效地提取 SST 数据的时空特征, 同时残差卷积和双向门控神经元的特征提取效率更高。

2.4.2 二维月均 SST 数据

为验证本文提出的 IWOA-Res-BiGRU 模型在二维月均 SST 数据中的性能, 将其与文献提出的 CFCC-LSTM^[10] 和 ConvGRU 模型^[11] 进行对比。

使用 2019 年 7 月至 2021 年 12 月的月均 SST 数据预测未来半年的月均 SST 数据, 结果如

图 11 和图 12 所示。表 3 为 3 种算法预测测试集二维月均 SST 数据的评价指标平均值。

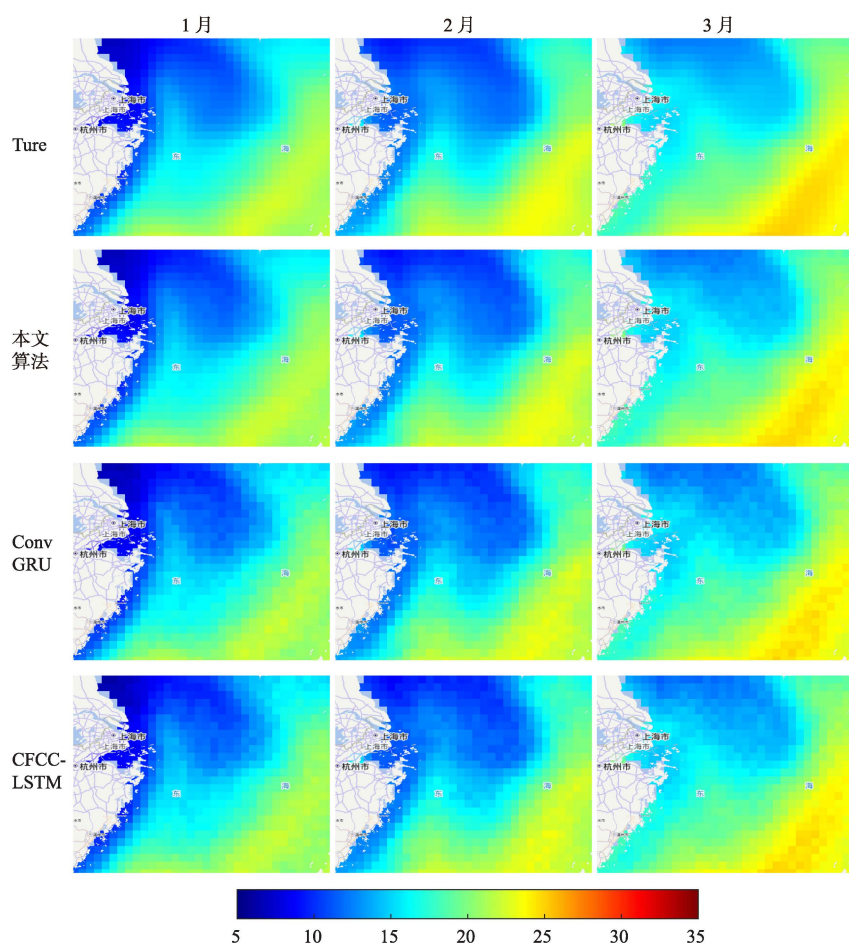


图 11 不同模型对 2022 年 1 月至 3 月 SST 的预测结果 [审图号: 国审字(2024)00691 号]

Fig. 11 Prediction results of different models for SST in the January to March 2022

结合表 3、图 11 和图 12 可知, 在 3 种模型中, Conv GRU 的预测效果较好; CFCC-LSTM 的预测误差最大, 其 $RMSE$ 和 2_AMAPE 分别达到 $0.67\text{ }^{\circ}\text{C}$ 和 3.0% , 例如, 2022 年 2 月的 SST 预测图像与 SST 真实图像存在较大偏差; IWOA-Res-BiGRU 的性能最好, 其 $RMSE$ 和 2_AMAPE 分别达到 $0.35\text{ }^{\circ}\text{C}$ 和 1.1% 。

相较于没有引入残差卷积和注意力机制的其他两种模型, IWOA-Res-BiGRU 预测的 SST 图像更接近真实值, 且能够关注到 SST 数据的重点区域。例如, 图 11 中 2022 年 1 月的 SST 预测结果, IWOA-Res-BiGRU 对低温区域的还原度明显更高; 图 12 中 2022 年 5 月的 SST 预测图像, IWOA-Res-BiGRU 对高温区域的还原更加清

楚。因此, 上述对比实验结果验证了本文所提算法的有效性。

3 结论

(1) 为提高 SST 预测精度和降低深度学习模型调参成本, 本文提出了一种 IWOA-Res-BiGRU 深度学习模型用于 SST 预测。对于一维日均 SST 时间序列, 可以利用 EEMD 算法来降低 SST 观测数据中的噪声分量, 提高深度学习模型的预测精度。

(2) 本文提出的 IWOA-Res-BiGRU 模型能够充分提取出 SST 的时空特征, 结合时空注意力机制能够进一步提升模型性能。

(3) 当然, SST 的预测算法仍然有较大提升

空间。在物理上,影响 SST 的因素包括海水盐度、气温、太阳辐射和风场等,未来的算法可能需要考虑上述因素之间的彼此影响,因此下一步可以开展基于多特征融合的 SST 预测算法研究。

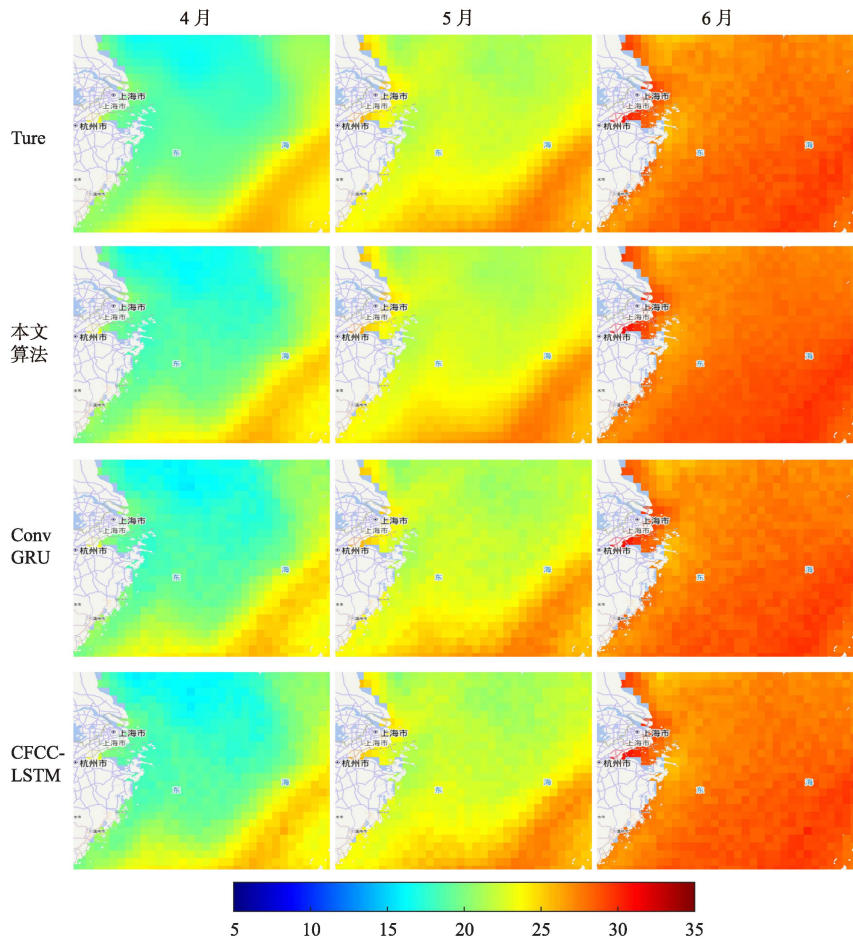


图 12 不同模型对 2022 年 4 月至 6 月 SST 的预测结果 [审图号:国审字(2024)00691 号]

Fig. 12 Prediction results of different models for SST in the April to June 2022

表 3 不同算法评价指标			
Tab.3 Different algorithm evaluation index			
算法编号及类型		RMSE/℃	2_AMAPE/(%)
1	IWOA-Res-BiGRU	0.35	1.1
5	CFCC-LSTM	0.67	3.0
6	ConvGRU	0.48	1.5

参考文献:

[1] ZHANG Z, PAN X L, JIANG T, et al. Monthly and quarterly sea surface temperature prediction based on gated recurrent unit neural network[J]. Journal of Marine Science and Engineering, 2020, 8(4): 249.

[2] 罗德杨, 郑 飞, 陈权亮. 基于深度学习的全球平均表面温度年际信号时间序列的预测[J]. 气候与环境研究, 2022,

27(1): 94-104.

[3] 冉 林, 熊建军, 赵 照, 等. 基于机器学习的电热防除冰表面温度变化趋势预测[J]. 装备环境工程, 2021, 18(8): 29-35.

[4] PATIL K, DEO M C. Basin-Scale prediction of sea surface temperature with artificial neural networks[J]. Journal of Atmospheric and Oceanic Technology, 2018, 35(7): 1441-1455.

[5] 刘 闯, 金仁瀚, 巩二磊, 等. 多元线性回归方法预测燃气轮机发电机组性能[J]. 中国电机工程学报, 2017, 37(16): 4731-4738.

[6] ZHOU S Y, CHU X, CAO S B, et al. Prediction of the ground temperature with ANN, LS-SVM and fuzzy LS-SVM for GSHP application[J]. Geothermics, 2020, 84: 101757.

[7] 汤贵艳, 朱善良, 周伟峰, 等. 基于高斯混合聚类和LightGBM算法的印度洋次表层温度反演研究[J]. 青岛科技大学学报(自然科学版), 2023, 44(2): 116-126.

[8] 彭豪杰, 周 杨, 胡校飞, 等. 基于深度学习与随机森林的

- PM_{2.5}浓度预测模型[J]. 遥感学报, 2023, 27(2): 430-440.
- [9] ZHANG Q, WANG H, DONG J Y, et al. Prediction of sea surface temperature using long short-term memory[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10): 1745-1749.
- [10] YANG Y T, DONG J Y, SUN X, et al. A CFCC-LSTM model for sea surface temperature prediction[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(2): 207-211.
- [11] 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.
- [12] ZHENG G, LI X F, ZHANG R H, et al. Purely satellite data-driven deep learning forecast of complicated tropical instability waves[J]. Science Advances, 2020, 6(29): eaba1482.
- [13] 王丽娜, 葛 鹏, 刘会滔, 等. 基于注意力机制的HDC-Bi-GRU海表温度预测模型[J]. 海洋环境科学, 2022, 41(6): 947-956.
- [14] 杜向楠, 杨 飞, 郭宁明, 等. 基于EMMD局域波分解的网络行波测距方法研究[J]. 电力系统保护与控制, 2018, 46(21): 175-181.
- [15] 刘 威, 郭直清, 王 东, 等. 改进鲸鱼算法及其在浅层神经网络搜索中的权值阈值优化[J]. 控制与决策, 2023, 38(4): 1144-1152.
- [16] 孙 林, 黄金旭, 徐久成, 等. 基于自适应鲸鱼优化算法和容错邻域粗糙集的特征选择算法[J]. 模式识别与人工智能, 2022, 35(2): 150-165.
- [17] PARAMBANCHARY D, RAO V M. WOA-NN: a decision algorithm for vertical handover in heterogeneous networks[J]. Wireless Networks, 2020, 26(1): 165-180.
- [18] 王维高, 魏云冰, 滕旭东. 基于VMD-SSA-LSSVM的短期风电预测[J]. 太阳能学报, 2023, 44(3): 204-211.
- [19] 赵敬娇, 赵志宏, 杨绍普. 基于残差连接和1D-CNN的滚动轴承故障诊断研究[J]. 振动与冲击, 2021, 40(10): 1-6.
- [20] LIU J J, JIN B G, WANG L, et al. Sea surface height prediction with deep learning based on attention mechanism[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1501605.
- [21] XIAO C J, CHEN N C, HU C L, et al. A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data[J]. Environmental Modelling & Software, 2019, 120: 104502.
- [22] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis: Association for Computational Linguistics, 2019: 4171-4186.
- [23] 韩 莹, 孙凯强, 张 栋, 等. 一种基于深度学习的海表温度混合预测方法[J]. 海洋环境科学, 2022, 41(5): 791-798.
- [24] 韩 莹, 孙凯强, 闫加宁, 等. 基于变分模态分解-长短时记忆网络-宽度学习系统混合模型的东海海温预测[J]. 激光与光电子学进展, 2023, 60(07): 59-67.

(本文编辑: 胡莹莹)