

基于CNN-STLSTM-CNN模型的有效波高预测

王丽娜^{1,2}, 齐致远³, 张红春¹, 董昌明^{2,4}

[1.南京信息工程大学 人工智能学院(未来技术学院), 江苏 南京 210044; 2.南方海洋科学与工程广东省实验室(珠海), 广东 珠海 519080; 3.南京信息工程大学 长望学院, 江苏 南京 210044; 4.南京信息工程大学 海洋科学学院, 江苏 南京 210044]

摘要:有效波高 (significant wave height, SWH) 是海洋的重要参数之一, 对其的精确预测对渔业发展、海上交通和海洋生态系统具有重要意义。为了提高有效波高的预测精度, 本文提出了一种基于卷积神经网络-时空长短时记忆神经网络-卷积神经网络 (convolutional neural network-spatiotemporal long short-term memory-convolutional neural network, CNN-STLSTM-CNN) 的有效波高预测模型。该模型由编码器 (Encoder)、解释器 (Translator) 和解码器 (Decoder) 构成。Encoder 通过卷积神经网络提取 SWH 数据的空间特征, Translator 通过时空长短时记忆神经网络 (spatiotemporal long short-term memory, STLSTM) 提取 SWH 数据的空间特征在时间上的变化特性, Decoder 通过卷积神经网络的转置卷积模块重建预测结果。对东海和南海海域的二维有效波高数据进行建模, 实验结果表明 CNN-STLSTM-CNN 模型的均方根误差 (root mean squared error, RMSE)、平均绝对误差 (mean absolute error, MAE)、均方根误差均值 (mean of root mean squared error, M_RMSE) 和平均绝对误差均值 (mean of mean absolute error, M_MAE) 等指标值均低于已有的方法, 验证了 CNN-STLSTM-CNN 模型的有效性。

关键词:有效波高; 卷积神经网络; 时空长短时记忆神经网络; 转置卷积

中图分类号: P731.3 文献标识码: A 文章编号: 1007-6336(2024)03-0417-13

Prediction of significant wave height based on CNN-STLSTM-CNN model

WANG Lina^{1,2}, QI Zhiyuan³, ZHANG Hongchun¹, DONG Changming^{2,4}

[1.School of Artificial Intelligence (School of Future Technology), Nanjing University of Information Science and Technology, Nanjing 210044, China; 2.Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519080, China; 3.Changwang School of Honors, Nanjing University of Information Science and Technology, Nanjing 210044, China; 4.School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China]

Abstract: Significant wave height (SWH) is one of the important parameters of the ocean, and its accurate prediction is of great significance to fishery development, maritime traffic and the marine ecosystem. In order to improve the prediction accuracy of significant wave height, this paper proposes a significant wave height prediction model based on convolutional neural network-spatiotemporal long short-term memory-convolutional neural network (CNN-STLSTM-CNN). The model consists of an Encoder, a Translator and a Decoder. The Encoder extracts the spatial features of the SWH data through the convolutional neural network, the Translator extracts the temporal variation characteristics of the spatial features of the SWH data through the spatiotemporal long short-term memory neural network (STLSTM), and the Decoder reconstructs the prediction results through

收稿日期: 2023-05-31, 修订日期: 2023-12-18

基金项目: 南方海洋科学与工程广东省实验室(珠海)资助项目(SML2020SP007); 国家自然科学基金项目(42192562)

作者简介: 王丽娜(1979-), 女, 山西晋中人, 副教授, 博士, 主要研究方向为海洋人工智能、机器学习, E-mail: wangln@nuist.edu.cn

the transposed convolution module of convolutional neural network. Modeling the two-dimensional significant wave height data in the East China Sea and South China Sea, the experimental results show that *RMSE*, *MAE*, *M_RMSE* and *M_MAE* values of the CNN-STLSTM-CNN model are lower than the existing methods, which verifies the effectiveness of the CNN-STLSTM-CNN model on SWH prediction.

Key words: significant wave height; convolutional neural network; spatiotemporal long short-term memory neural network; transposed convolution

海洋活动不仅对人类的生产、生活有着深刻的影响,对沿海生态系统和沿海地形也同样具有重要意义^[1]。海洋活动的一个重要特征是海浪,而有效波高(significant wave height, SWH)是海浪的重要要素,对其的精确预测对渔业发展、海上交通和海洋生态系统具有重要意义。

SWH 预测主要有两类模型:数值海浪模型和数据驱动模型。目前第三代数值海浪模型 WAVE WATCH III (WW3)^[2] 被广泛应用于业务化的海浪预报系统。该数值海浪模型提高了预测精度,但是需要设置大量参数,且计算过程复杂、运行时间长,难以实现浪高的快速精确预测。而数据驱动模型仅对历史数据建模,可以实现快速精确预测。深度学习方法能够学习数据的内在规律和表示层次,拟合复杂的非线性过程,可为浪高预报这类物理机制复杂的非线性问题提供有力支撑。随着深度学习技术的快速发展,已有一些研究将深度学习应用到 SWH 预测。Fan 等^[3] 使用长短期记忆神经网络(long short-term memory, LSTM)预测多个时刻的 SWH,结果证明了 LSTM 对 SWH 数据具有较强的建模能力。Pirhooshyaran 等^[4] 将 LSTM 与编码器-解码器结构结合构建了一个序列到序列网络,并将其用于 SWH 的预测和重建,结果表明,该序列到序列模型在 SWH 预测任务上有良好的性能。Pokhrel 等^[5] 将 Transformer 深度学习模型与第三代数值海浪模型结合,使用 Transformer 预测 SWH 观测值与 WW3 模式所得 SWH 的一阶差分,取得了较优的预测性能。上述研究均为基于单点观测数据的 SWH 预测。而二维 SWH 具有较强的空间相关性,Zhou 等^[6] 提出基于多层 ConvLSTM 对东海-南海海域的 SWH 时空数据建模,并取得较为准确的预报结果。

为了进一步探究相关特征对 SWH 预测的影响,一些研究将海洋相关特征和大气相关特征引

入 SWH 预测。Pokhrel 等^[5] 结合 SWH、平均波周期(mean wave period, MWP)、主导波周期(dominant wave period, DWP)和波方向(wave direction)等特征进行 SWH 预测。Han 等^[7] 认为 SWH 主要受风向(wind direction)、风速(wind speed)和气压(atmospheric pressure)等因素的影响,其中风速和风向是影响浪高的最主要因素。Gao 等^[8] 将 SWH、平均波长(mean wave length, MWL)、平均波方向(mean wave direction, MWD)、方向谱宽度(directional spectrum width, DSW)和 10 m 风的 *U*、*V* 分量作为模型输入,进行 SWH、平均波长、平均波方向和方向谱宽度等的预测。

为了充分提取 SWH 数据的时空特征,提升 SWH 预测精度,本文提出了由编码器(Encoder)、解释器(Translator)和解码器(Decoder)构成的 CNN-STLSTM-CNN 预报方法。在 Encoder 中,卷积神经网络(convolutional neural network, CNN)可以较为充分地提取数据的空间特征,但难以获取时序数据之间的关联信息。在 Translator 中,采用时空长短时记忆神经网络(spatiotemporal long short-term memory, STLSTM)充分提取空间时序数据的空间特征在时间上的变化特性。在 Decoder 中,采用 CNN 的转置卷积还原编码器中卷积层下采样导致的数据空间尺寸的减小,从而重建时序数据。本文将该方法应用于东海和南海海域(3.5°N—35°N, 99.5°E—131°E)二维 SWH 数据,分别建立输入序列为 3 h、预测 3 h 和输入序列为 6 h、预测 6 h 的 CNN-STLSTM-CNN 模型。实验结果表明,其预测性能均优于已有的方法,验证了方法的有效性。

1 材料与方法

1.1 数据来源和研究区域

SWH 数据为美国国家海洋和大气管理

局 (National Oceanic Atmospheric Administration, NOAA) 提供的 WAVE WATCH III 数值模式的数据集^[9], 其时间分辨率为 1 h, 空间分辨率为 $0.5^{\circ} \times 0.5^{\circ}$ 。

影响浪高的特征有很多, 包括 10 m 风的 U 分量与 V 分量、MWD、MWP、平均海平面压力 (mean sea level pressure, MSLP) 数据等。这些特征来自第五代欧洲中期天气预报中心再分析数据集 ERA5 中的 10 m 风的 U 分量和 V 分量、MWD、MWP、MSLP 数据^[10]。ERA5 提供大气、海浪和地表参数的每小时估算值。10 m 风的 U 、 V 分量和 MSLP 的空间分辨率均为 $0.25^{\circ} \times 0.25^{\circ}$, MWD 和 MWP 的空间分辨率均为 $0.5^{\circ} \times 0.5^{\circ}$ 。所有特征的时间分辨率都为 1 h。

选取东海和南海海域 ($3.5^{\circ}N - 35^{\circ}N, 99.5^{\circ}E -$

$131^{\circ}E$) 作为研究对象。研究时间为 2011 年 1 月 1 日 00:00 至 2021 年 12 月 31 日 24:00。在特征选择方面, SWH 为必选特征, 10 m 风的 U 分量与 V 分量、MWD、MWP、MSLP 为可选特征。

1.2 模型构建

1.2.1 模型结构

本文的 CNN-STLSTM-CNN 预测模型包含 Encoder、Translator 和 Decoder 3 个部分, 如图 1 所示。Encoder 包含多组由 Conv3d、BatchNorm 和 LeakyReLU 层构成的模块; Translator 包含多个 STLSTM 层; Decoder 包含多组由 Conv3dTranspose、BatchNorm 和 LeakyReLU 层构成的模块。其中, Conv3d 为 3D 卷积, BatchNorm 为批归一化, LeakyReLU 为激活函数, STLSTM 为时空长短时记忆神经网络, Conv3dTranspose 为转置卷积。

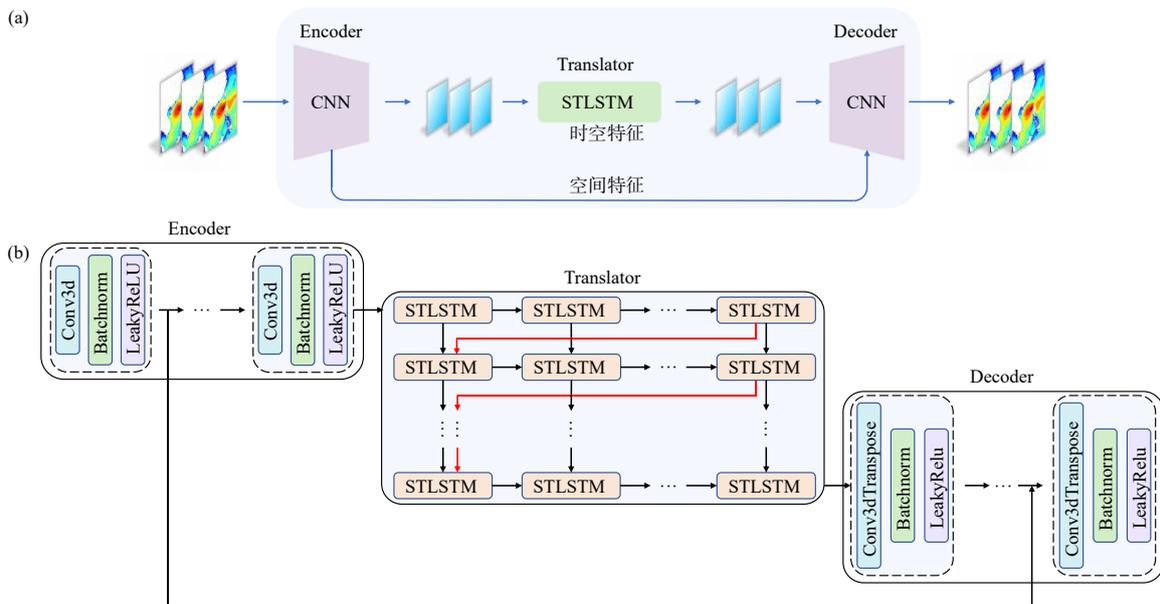


图 1 CNN-STLSTM-CNN 模型结构

Fig. 1 The structure of CNN-STLSTM-CNN model

Encoder 在接受输入的二维 SWH 数据后, 通过 CNN 提取数据的空间特征。

Translator 堆叠了多个 STLSTM 层, 提取 SWH 数据的空间特征在时间上的变化特性。首先, Translator 接受 Encoder 的连续多个时刻输出特征, 其次, 多个 STLSTM 神经元构建的 STLSTM 网络能够在垂直和水平方向传递存储器状态, 即时空记忆。该存储器状态在图 1b 的 Translator 模块中以红色线方向流动, 构成 Z 字形。即空间

信息沿着不同层间传递, 随后随着时间向前移动。这样空间信息能够有效流动, 实现了标准时间记忆和时空记忆的同时流动。

Decoder 接受 Translator 的输出, 经 3D 转置卷积重建 Translator 输出的空间特征在时间上的变化特性, 并通过批归一化和 LeakyReLU 激活函数输出预测结果。

另外, 本文将 Encoder 中由 Conv3d、BatchNorm 和 LeakyReLU 层构成的第一个模块输出的特征

输入 Decoder 中最后一个由 Conv3dTranspose、BatchNorm 和 LeakyReLU 构成的模块,以防特征丢失。

1.2.2 Encoder

Encoder 由多个卷积模块构成。CNN 是由 LeCun 等提出的一种深度神经网络^[11]。本文所使用的卷积模块包括卷积层、批归一化层^[12]和 LeakyReLU 激活函数^[13]。在卷积层中,卷积核以滑动窗口的方式对输入数据进行卷积,提取输入数据的特征。批归一化层对输入的特征进行均值为 0,方差为 1 的归一化。卷积的计算如公式(1)所示。

$$O = w \times x + b \quad (1)$$

式中: O 表示输出特征图; x 表示卷积层的输入; w 表示卷积的权重矩阵; \times 代表点积; b 表示偏差向量。

批归一化的计算如公式(2)所示。

$$m = \frac{O - E[O]}{\sqrt{\text{Var}[O]}} * \gamma + \beta \quad (2)$$

式中: m 表示批归一化的输出; 输出特征图 O 为批归一化的输入; γ 是尺度因子; β 是偏移因子, γ 和 β 是可学习的参数; 在训练过程中, $E[O]$ 表示批的均值, $\text{Var}[O]$ 表示批的方差, 而在

测试过程中, $E[O]$ 表示训练集所有数据的均值, $\text{Var}[O]$ 表示训练集所有数据的方差。LeakyReLU 函数如公式(3)所示。

$$\text{LeakyReLU}(m_{ij}) = \begin{cases} m_{ij}, & m_{ij} \geq 0 \\ \alpha m_{ij}, & m_{ij} < 0 \end{cases} \quad (3)$$

式中: 批归一化输出 m_{ij} 为激活函数的输入; α 表示权重因子, 取值范围是 (0,1)。

1.2.3 Translator

Translator 由多个 STLSTM 层连接构成。STLSTM 是长短时记忆神经网络(long short-term memory, LSTM)的一种改进^[14]。由于 LSTM 在处理时序数据时忽略了空间相关性, Shi 等^[15]提出了 ConvLSTM 模型, 关注空间和时间特征。为了解决信息流动的单向传递问题, STLSTM 引入时空记忆, 使得通过不同层获取的空间信息随着时间向前传递。

STLSTM 神经元如图 2a 所示, STLSTM 网络如图 2b 所示。在 STLSTM 网络中, L_1 和 L_2 均为 STLSTM 神经元, 红色箭头表示时空记忆 M 的传播方向。首先 M 按层间传递, 然后随时间向前传递。细胞状态 C 和隐藏状态 H 的传递方式与 LSTM 相同, 在同层之间按照时间向前传递。

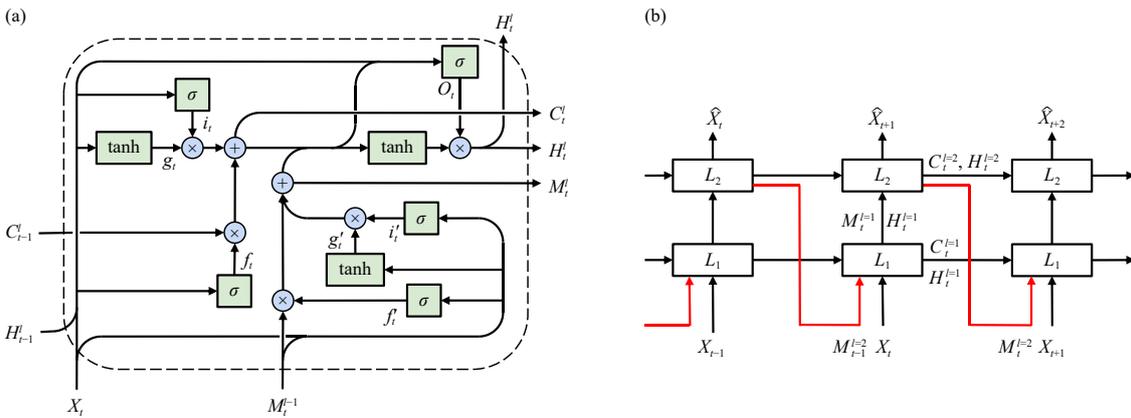


图2 STLSTM 网络

Fig. 2 STLSTM network

STLSTM 神经元的计算如公式(4)–(13)所示。

$$g_t = \tanh(W_{xg} * X_t + W_{hg} * H_{t-1}^l + b_g) \quad (4)$$

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1}^l + b_i) \quad (5)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1}^l + b_f) \quad (6)$$

$$C_t^l = f_t \odot C_{t-1}^l + i_t \odot g_t \quad (7)$$

$$g_t' = \tanh(W'_{xg} * X_t + W_{mg} * M_t^{l-1} + b'_g) \quad (8)$$

$$i_t' = \sigma(W'_{xi} * X_t + W_{mi} * M_t^{l-1} + b'_i) \quad (9)$$

$$f_t' = \sigma(W'_{xf} * X_t + W_{mf} * M_t^{l-1} + b'_f) \quad (10)$$

$$M_t^l = f_t' \odot M_t^{l-1} + i_t' \odot g_t' \quad (11)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1}^l + W_{co} * C_t^l + W_{mo} * M_t^l + b_o) \quad (12)$$

$$H_t^l = o_t \odot \tanh(W_{1 \times 1} * [C_t^l, M_t^l]) \quad (13)$$

式中: $\sigma()$ 和 $\tanh()$ 分别表示 Sigmoid 和 tanh 激活函数; W_{xg} 、 W_{hg} 、 W_{xi} 、 W_{hi} 、 W_{xf} 、 W_{hf} 、 W'_{xg} 、 W_{mg} 、 W'_{xi} 、 W_{mi} 、 W'_{xf} 、 W_{mf} 、 W_{xo} 、 W_{ho} 、 W_{co} 、 W_{mo} 、 $W_{1 \times 1}$ 表示卷积的权重; X_{t-1} 表示 $(t-1)$ 时刻的输入; H_t^l 表示 t 时刻第 l 层的隐藏状态; \widehat{X}_t 为第 t 时刻的输出; C_{t-1}^l 表示 $(t-1)$ 时刻第 l 层的细胞状态; C_t^l 表示 t 时刻第 l 层的细胞状态; M_t^{l-1} 表示 t 时刻 $(l-1)$ 层的时空记忆; M_t^l 表示 t 时刻第 l 层的时空记忆; b_g 、 b_i 、 b_f 、 b'_g 、 b'_i 、 b'_f 、 b_o 表示偏差。

1.2.4 Decoder

Decoder 由多个 Conv3dTranspose、BatchNorm 和 LeakyReLU 层构成的卷积模块组成。首先经 3D 转置卷积把 STLSTM 提取的空间特征在时间上的变化特性重建, 随后通过 BatchNorm 和 LeakyReLU 激活函数获得预测结果。转置卷积用于还原卷积层下采样导致图像尺寸减小的图片^[16]。

1.2.5 评价指标

为了验证模型的有效性, 选择均方根误差 (root mean squared error, *RMSE*)、平均绝对误差 (mean absolute error, *MAE*)、均方根误差均值 (mean of root mean squared error, *M_RMSE*)、平均绝对误差均值 (mean of mean absolute error, *M_MAE*) 和绝对误差 (*absolute error*) 5 个评价指标进行评价。其中 *M_RMSE* 和 *M_MAE* 分别给出测试样本空间上的均值来评估预测值和观测值之间的误差和偏离程度。各指标的计算如公式(14)–(18)所示。

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (h_p(i, j) - h_o(i, j))^2} \quad (14)$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |h_p(i, j) - h_o(i, j)| \quad (15)$$

$$M_RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T \sqrt{\frac{1}{I \times J} \sum_{i=1}^I \sum_{j=1}^J (h_p(i, j) - h_o(i, j))^2}} \quad (16)$$

$$M_MAE = \frac{1}{T} \sum_{t=1}^T \frac{1}{I \times J} \sum_{i=1}^I \sum_{j=1}^J |h_p(i, j) - h_o(i, j)| \quad (17)$$

$$absolute\ error = |h_p(i, j) - h_o(i, j)| \quad (18)$$

式中: i, j 表示空间格点坐标; T 表示测试样例总数; I 表示经向总格点数; J 表示纬向总格点数; $h_p(i, j)$ 表示对应空间格点模型预报的 SWH 值; $h_o(i, j)$ 表示对应空间格点的 SWH 观测值。

2 结果与讨论

2.1 实验设置

本文实验的 CPU 为 Intel(R) Xeon(R) Platinum 8255C, GPU 为 RTX 3090, Ubuntu 版本为 20.04, Python 版本为 3.8.10, TensorFlow 版本为 2.9.0, CUDA 版本为 11.2。

训练集选取 2011 年 1 月 1 日 0 时至 2018 年 12 月 31 日 23 时的 SWH, 验证集选取 2019 年 1 月 1 日 0 时至 2019 年 12 月 31 日 23 时的 SWH, 测试集选取 2020 年 1 月 1 日 0 时至 2021 年 12 月 31 日 23 时的 SWH。SWH 数据的空间分辨率均为 $0.5^\circ \times 0.5^\circ$, 时间分辨率均为 1 h。实验选取 Adam 优化器^[17], 其中 $\beta_1 = 0.9$, $\beta_2 = 0.999$ 。

SWH 数据应用最大最小归一化方法进行预处理, 如公式(19)所示。

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (19)$$

式中: x 为观测样本; $\min(x)$ 为样本最小值; $\max(x)$ 为样本最大值; x_{scaled} 为归一化后的样本。

由于 10 m 风的 U 分量与 V 分量、平均海平面压力的空间分辨率较高, 在预处理时, 对其进行采样, 使其空间分辨率由 $0.25^\circ \times 0.25^\circ$ 降为 $0.5^\circ \times 0.5^\circ$ 。这里, 要求 SWH 与其他特征要素维度一致。

2.2 最优模型性能分析

本文最优模型结构为: Encoder 是两组包含由 Conv3d、BatchNorm 和 LeakyReLU 层构成的卷积块, Translator 是两层 STLSTM, Decoder 是两组包含 Conv3dTranspose、BatchNorm 和 LeakyReLU 层构成的卷积块。

经过测试, 给出最优模型参数设置。Encoder 和 Decoder 中卷积核大小为 $1 \times 2 \times 2$, 过滤器个数为 64, Translator 的卷积核大小为 2×2 , 过滤器个数为 64。本文给出了输入序列为 3 h 和 6 h 的模型预测结果(表 1 和表 2)。结果表明, 输入序列

表1 输入序列为3 h的模型在测试集上的误差统计

Tab.1 Error statistics of the model with an input sequence of three hours in the test set

预测时刻	M_RMSE/m	M_MAE/m
1	0.0284	0.0166
2	0.0386	0.0198
3	0.0512	0.0252

为3 h和6 h的模型在测试集上的 M_RMSE 和 M_MAE 值均较小。

表2 输入序列为6 h的模型在测试集上的误差统计

Tab.2 Error statistics of the model with an input sequence of six hours in the test set

预测时刻	M_RMSE/m	M_MAE/m
1	0.0370	0.0234
2	0.0376	0.0243
3	0.0411	0.0253
4	0.0493	0.0298
5	0.0604	0.0367
6	0.0701	0.0415

为了进一步展示本文 CNN-STLSTM-CNN 模型的预测性能,以输入序列为3 h的模型为例,在东海海域和南海海域分别选择一个观测点进行分析。东海海域选择 27°N、126°E 的观测点,记为观测点1;南海海域选择 9°N、113°E 的观测点,记为观测点2。表3和表4分别展示了两个观测点测试集上的 $RMSE$ 和 MAE 指标值。

表3 两个观测点测试集上的 $RMSE$ 值Tab.3 $RMSE$ values of 2 observation points in the test set

位置	$RMSE/m$		
	第1 h	第2 h	第3 h
观测点1	0.0440	0.0759	0.1104
观测点2	0.0257	0.0359	0.0468

表4 两个观测点测试集上的 MAE 值Tab.4 MAE values of 2 observation points in the test set

位置	MAE/m		
	第1 h	第2 h	第3 h
观测点1	0.0229	0.0368	0.0536
观测点2	0.0146	0.0190	0.0243

由表中数据可知,在相同时刻,观测点2的 $RMSE$ 和 MAE 值都低于观测点1,说明本文模型在南海观测点的预测性能优于东海观测点。图3展示了两个观测点在测试集上2020年1月1日0时至12月31日24时的预测结果。由图3可以看出,在东海和南海海域选择的观测点上,本文模型得到的SWH预测值和观测值相近,仅在浪高极值位置上,不同时刻的预测值和观测值差异较大,如图3(a)所示2020年9月附近的极值点。放大2020年7月15日至7月20日这个时间段的预测值,发现位于东海的观测点1上,未来第1h的SWH预测值接近观测值,未来第2h的SWH预测值偏离观测值比未来第1h要远,未来第3h的SWH预测值偏离观测值最远,预测性能最差。位于南海的观测点2上,不同时刻的预测值与观测值基本一致,预测精度高。综上所述,无论是南海海域还是东海海域上,本文模型均取得较高的预测精度。

2.3 不同模型预测性能比较

2.3.1 输入序列为3 h的模型性能

为了分析所提出的 CNN-STLSTM-CNN 模型的预测性能,选取了 ConvLSTM^[6]、PredRNN^[14]、SimVP^[18] 3个模型与本文模型比较。选择输入序列为3 h、输出序列为3 h的SWH数据建模分析。不同模型在 M_RMSE 和 M_MAE 上的误差指标值如表5和表6所示。

表5和表6展示了 ConvLSTM、PredRNN、SimVP 和 CNN-STLSTM-CNN 4个模型在测试集上未来第1 h、第2 h和第3 h的预测误差。CNN-STLSTM-CNN 模型在测试集上各时刻的 M_RMSE 分别为 0.0284 m、0.0386 m 和 0.0512 m,在所有模型中本文提出的模型的预测误差最低。同理,本文提出的模型在各时刻的 M_MAE 分别为 0.0166 m、0.0198 m 和 0.0252 m,在所有模型中 M_MAE 值最小。表明在输入序列3 h、输出序列3 h的模型中,本文提出的模型的预测性能最优。

2.3.2 输入序列为6 h的模型性能

选择输入序列为6 h、输出序列为6 h的SWH数据建模分析。图4展示了 ConvLSTM、PredRNN、SimVP 和 CNN-STLSTM-CNN 4个模

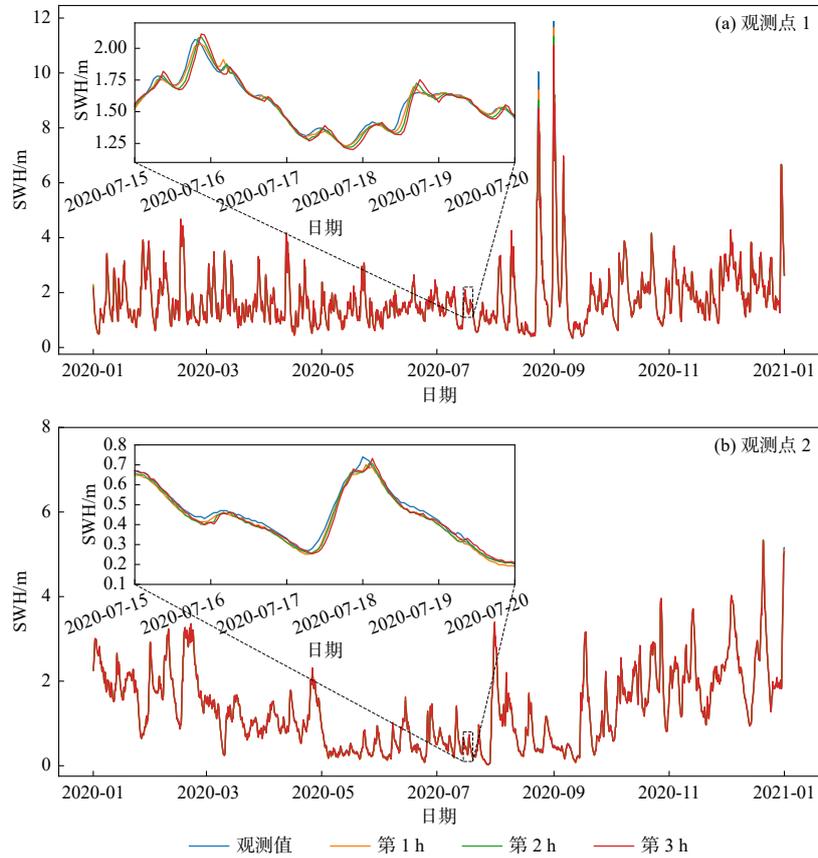


图 3 不同观测点在测试集上 2020 年 1 月 1 日 0 时至 12 月 31 日 24 时的预测值

Fig. 3 Predicted values of different observation points in the test set from 00:00 on January 1, 2020 to 24:00 on December 31, 2020

表 5 不同模型在测试集上的 M_{RMSE} 值

Tab.5 Comparisons of M_{RMSE} values among different algorithms in the test

模型	M_{RMSE}/m		
	第1 h	第2 h	第3 h
ConvLSTM ^[6]	0.0666	0.0900	0.1323
PredRNN ^[14]	0.0427	0.0515	0.0645
SimVP ^[18]	0.0375	0.0445	0.0530
CNN-STLSTM-CNN	0.0284	0.0386	0.0512

型在测试集上未来第 1 h 至第 6 h 的 M_{RMSE} 和 M_{MAE} 上的误差指标值。

CNN-STLSTM-CNN 模型在测试集上各时刻的 M_{RMSE} 分别为 0.0370 m、0.0376 m、0.0411 m、0.0493 m、0.0604 m 和 0.0701 m, 在所有模型中 M_{RMSE} 值最低。同理, 本文提出的模型在各时刻的 M_{MAE} 分别为 0.0234 m、0.0243 m、0.0253 m、0.0298 m、0.0367 m 和

表 6 不同模型在测试集上的 M_{MAE} 值

Tab.6 Comparisons of M_{MAE} values among different algorithms in the test

模型	M_{MAE}/m		
	第1 h	第2 h	第3 h
ConvLSTM ^[6]	0.0400	0.0600	0.0925
PredRNN ^[14]	0.0247	0.0295	0.0368
SimVP ^[18]	0.0212	0.0242	0.0285
CNN-STLSTM-CNN	0.0166	0.0198	0.0252

0.0415 m, 在所有模型中该误差指标值最低。表明在所有输入序列 6 h、输出序列 6 h 的模型中, 本文提出的模型的预测性能最优。

通过 4 个模型未来 3 h 和 6 h 的预测实验对比分析可知, ConvLSTM 模型关注时间和空间特征, 其具有处理时空序列数据的能力, 但其空间信息仅限于按层间传递。而 PredRNN 模型通过 STLSTM 引入时空记忆, 空间信息首先按层间传

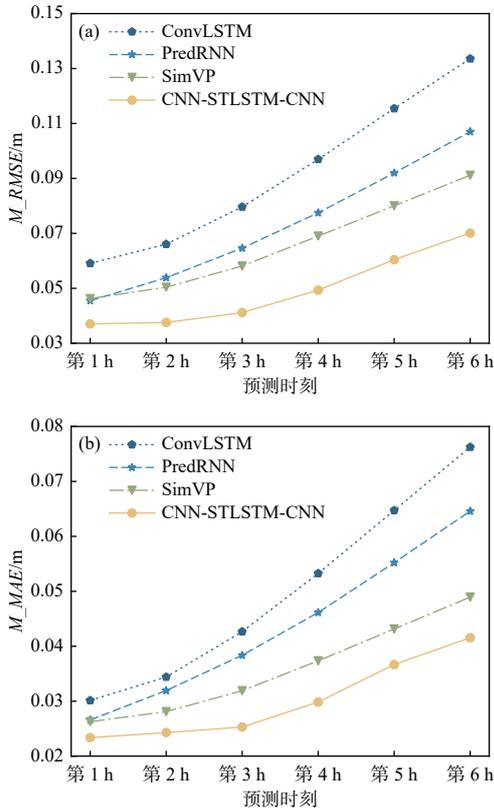


图 4 在东海和南海研究区域不同模型未来 6 h 的指标值对比

Fig. 4 Comparisons of index values obtained from different models in the selected areas of ECS and SCS

递, 然后随时间向前传递, 解决了信息流动的局限性问题。这里, PredRNN 引入 STLSTM, 模型参数显著增加。整体上, PredRNN 模型预测性能优于 ConvLSTM 模型。SimVP 模型由 Encoder、Translator 和 Decoder 3 个部分组成。其中 Encoder 通过多组 CNN 模块提取空间特征, Translator 通过堆叠 CNN 模块学习时间演化, Decoder 通过多组 CNN 的转置卷积模块集成时空信息预测 SWH 值。SimVP 模型减少了模型参数, 预测性能有了进一步提升。本文提出的 CNN-STLSTM-CNN 模型由 Encoder、Translator 和 Decoder 3 个部分组成。首先, 在 Encoder 中通过多组 CNN 提取空间特征; 其次, 在 Translator 中通过堆叠 STLSTM 层, 引入时空记忆, 实现了标准时间记忆和时空记忆的同时流动; 最终, 在 Decoder 中通过 CNN 的多组转置卷积重建预测数据。实验结果表明本文提出的模型的 SWH 预测精度明显高于其他模型, 验证了方法的有效性。

2.4 模型复杂度分析

应用访存量 (Memory) 和浮点运算次数 (floating points of operations, FLOPs) 两个指标来分别评价本文模型的空间和时间复杂度。Memory 是指模型的参数数量, FLOPs 表示模型的计算复杂度, 其主要包括乘法和加法运算次数。给出输入尺寸为 $1 \times 64 \times 64$ 的时序图片数据, 输入序列分别为 3 h 和 6 h 的不同模型的 Memory 和 FLOPs 结果列于表 7 和表 8。由表 7 和表 8 的数据分析可知: 当处理输入尺寸为 $1 \times 64 \times 64$ 的时序图片数据时, CNN-STLSTM-CNN 模型占用存储空间为 2.3167 MB, 比 PredRNN 模型占用空间少, 比 SimVP 占用空间多; 同时本文提出的模型的计算量比 PredRNN 模型少, 但远多于 SimVP 模型。因此, 本文提出的模型需要计算量较大, 复杂度较高。

表 7 输入序列为 3 h 的模型的 Memory 和 FLOPs 统计

Tab.7 Memory and FLOPs statistics of the model with an input sequence of three hours

模型	Memory/MB	FLOPs/G
ConvLSTM	2.8264	20.0588
PredRNN	8.1563	43.7885
SimVP	0.8379	1.5514
CNN-STLSTM-CNN	2.3167	21.1834

表 8 输入序列为 6 h 的模型的 Memory 和 FLOPs 统计

Tab.8 Memory and FLOPs statistics of the model with an input sequence of six hours

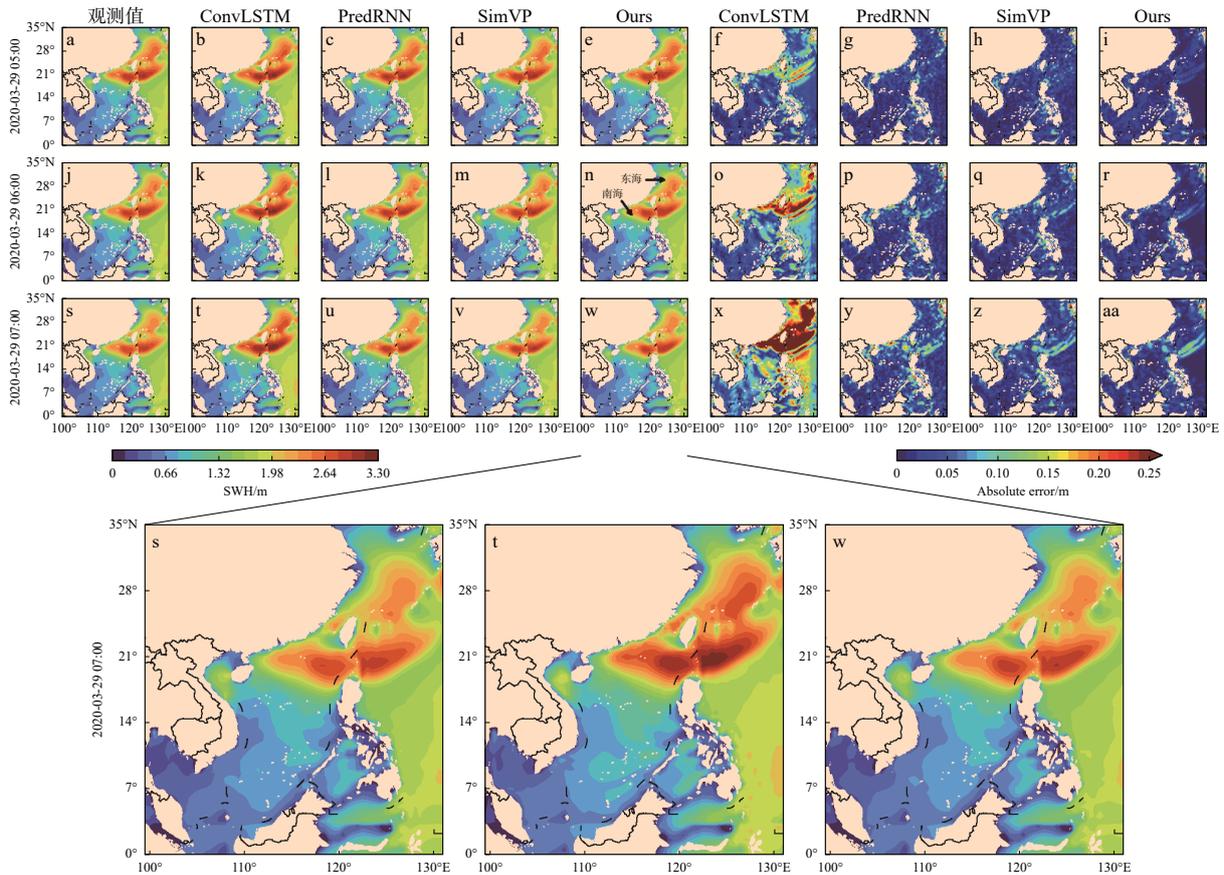
模型	Memory/MB	FLOPs/G
ConvLSTM	2.8264	40.1255
PredRNN	8.1563	96.3348
SimVP	0.9697	3.0660
CNN-STLSTM-CNN	2.3167	46.5253

2.5 可视化分析

为了展示 CNN-STLSTM-CNN 模型的优势, 选择输入序列为 3 h 的本文模型、ConvLSTM、PredRNN 和 SimVP 4 个模型进行可视化分析。在测试集上随机挑选 3 h 的 SWH 值, 图 5 是不同模型在东海和南海研究区域 2020 年 3 月 29 日 5 时至 7 时的预测值与绝对误差的空间分

布。不同空间格点预测值的颜色不同。在预测值图上, 蓝色越深表示 SWH 值越接近于 0 m, 红色越深表示 SWH 值越大; 在绝对误差值图上, 蓝色越深表示绝对误差越小, 红色越深表示绝对误差越大。图 5a、图 5j、图 5s 为 5 时至 7 时的观测值, 图 5b、图 5k、图 5t 和图 5f、图 5o、图 5x 分别为 ConvLSTM 模型 3 个时刻的预测值和绝

对误差的空间分布, 图 5c、图 5l、图 5u 和图 5g、图 5p、图 5y 分别为 PredRNN 模型 3 个时刻的预测值和绝对误差的空间分布, 图 5d、图 5m、图 5v 和图 5h、图 5q、图 5z 分别为 SimVP 模型 3 个时刻的预测值和绝对误差的空间分布, 图 5e、图 5n、图 5w 和图 5i、图 5r、图 5aa 分别为本文模型 3 个时刻的预测值和绝对误差的空间分布。



注: 本文研究海域范围为 3.5°N — 35°N 和 99.5°E — 131°E

图 5 不同模型在 2020 年 3 月 29 日 5 时至 7 时的预测值、预测值与观测值的绝对误差对比 [审图号:GS(2023) 4355 号]

Fig. 5 Comparisons of the predictive values among different models and the absolute errors between predictive values and observed values from 5:00 to 7:00 on March 29, 2020

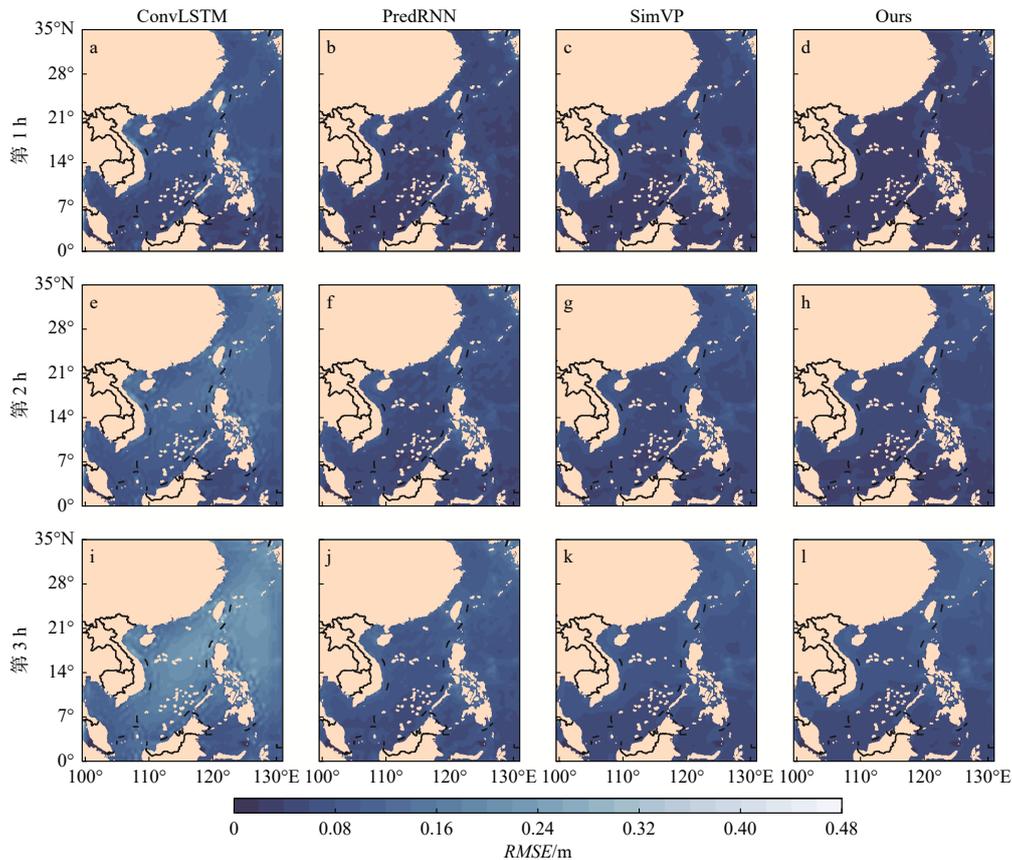
由图 5 观测到: 随着预测时间的增加, 所有模型的绝对误差在增大。ConvLSTM 模型在未来第 3 h 的预测性能均最差, 绝对误差值最大达到 0.5516 m。PredRNN 模型每小时的预测性能比 ConvLSTM 模型较优, 绝对误差空间分布图中每小时的蓝色区域较多, 尤其是第 3 h 的绝对误差图没有出现绝对误差大的深红色区域, 说明

预测误差更小。在 SimVP 模型每小时的绝对误差空间分布中, 蓝色区域比 PredRNN 模型更多, 尤其是第 3 h 蓝色区域明显增多, 说明 SimVP 模型的预测性能得到进一步提升。放大第 3 h 的观测值(图 5s)、ConvLSTM 模型图(图 5t)和本文模型图(图 5w), 发现无论东海海域还是南海海域, 本文模型预测的 SWH 值与观测值最接近,

ConvLSTM 模型预测效果最差。CNN-STLSTM-CNN 模型每小时的绝对误差深蓝色的区域最多,说明绝对误差最小,表明在所有模型中预测性能最优。

图 6 为 4 个模型在测试集上 $RMSE$ 的空间分布的可视化结果。蓝色越深,说明误差越小,

红色越深,说明误差越大。由图 6 观察到:所有模型的 $RMSE$ 值显示几乎都是蓝色,仅 ConvLSTM 模型在第 3 h 的 $RMSE$ 值有个别浅红色点。与其他模型相比,CNN-STLSTM-CNN 模型每小时的 $RMSE$ 空间分布中的深色区域最多,说明其预测误差最小,表明本文模型的预测性能最优。



注: 本文研究海域范围为 3.5°N—35°N 和 99.5°E—131°E

图 6 不同模型在测试集上 $RMSE$ 的空间分布 [审图号:GS(2023)4355 号]

Fig. 6 The spatial distribution of the $RMSE$ among different models in the test set

2.6 不同特征对模型预测性能的影响

为了分析不同特征对本文模型预测性能的影响,将 SWH 分别与 10 m 风的 U 分量与 V 分量、MWD、MWP、MSLP 4 个特征要素组合建模,输出未来 3 h 的 SWH 预测值。

图 7 展示了加入不同特征的模型的 M_RMSE 和 M_MAE 值。在加入 10 m 风的 U 、 V 分量特征后,每小时的 M_RMSE 和 M_MAE 值均低于加入 MWD、MWP 和 MSLP 特征的模型,10 m 风的 U 、 V 分量能显著地提升 SWH 的预测性能。MWD 和 MWP 特征对预测性能有一定的提升,

两者的预测误差指标值相近。MWP 特征在第 1 h 的预测优于 MWD 特征,在第 2 h 和第 3 h 的预测比 MWD 特征差。加入 MSLP 特征与未加特征的模型每小时的 M_RMSE 值相近,而在第 2 h 和第 3 h 的 M_MAE 值比未加特征模型更大,说明 MSLP 特征未有效提升预测性能。选择加入全部特征,所得模型的预测性能和加入 10 m 风的 U 、 V 分量的模型相近,未提升预测精度。

图 8 是加入不同特征的本文模型在 2020 年 3 月 29 日 5 时至 7 时的预测值与绝对误差的空

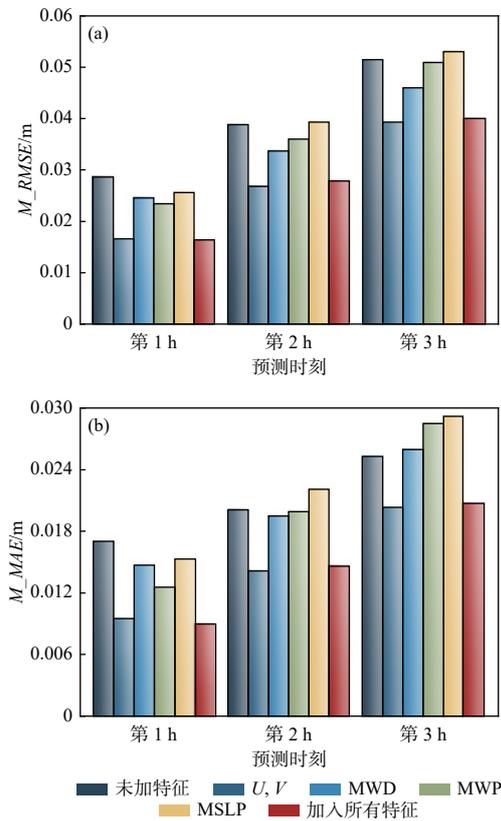


图 7 加入不同特征的模型在测试集上的误差统计

Fig. 7 Error statistics of the model adding different features on the test set

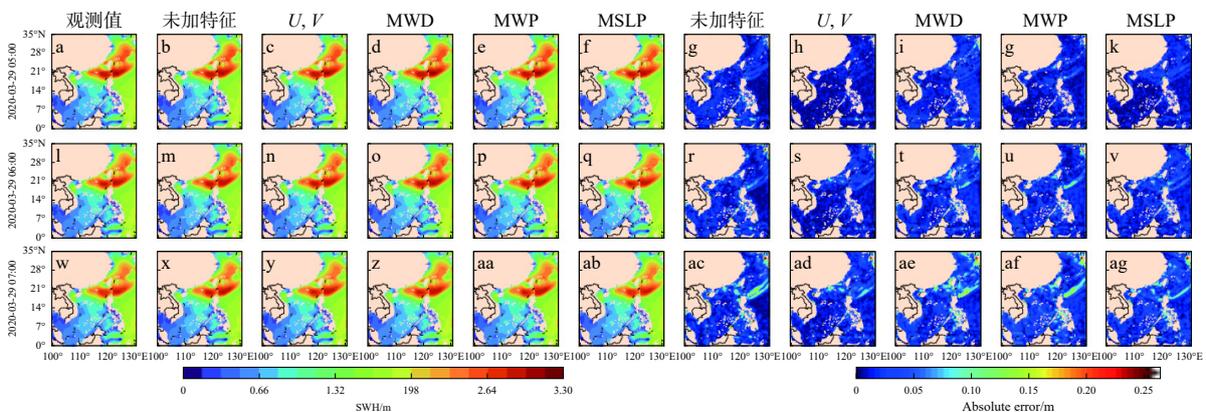
间分布。

图 8a、图 8l、图 8w 为观测值, 图 8b、图 8m、图 8x 和图 8g、图 8r、图 8ac 分别为未加特征的

预测值与绝对误差的空间分布, 图 8c、图 8n、图 8y 和图 8h、图 8s、图 8ad 分别为加入 10 m 风 U 、 V 分量的预测值与绝对误差的空间分布, 图 8d、图 8o、图 8z 和图 8i、图 8t、图 8ae 分别为加入 MWD 的预测值与绝对误差的空间分布, 图 8e、图 8p、图 8aa 和图 8j、图 8u、图 8af 分别为加入 MWP 的预测值与绝对误差的空间分布, 图 8f、图 8q、图 8ab 和图 8k、图 8v、图 8ag 分别为加入 MSLP 的预测值与绝对误差的空间分布。

由图 8 观察到: 在绝对误差图中, 蓝色越深, 误差越小; 在由蓝色渐变到红色的过程中, 误差逐渐增大; 红色越深, 误差越大。在加入 10 m 风 U 、 V 分量的模型每小时绝对误差空间分布中, 深蓝色的区域比其他特征的模型都多, 说明预测误差最小, 表明每小时的预测值最接近观测值。而在加入 MWD 的模型每小时绝对误差图中, 浅色区域比不加特征的模型多, 说明预测误差比未加特征的模型大。在加入 MWP 的模型绝对误差图中, 第 1 h 深蓝色的区域多于加 MWD 特征的模型, 略少于加 10 m 风 U 、 V 分量的模型, 说明预测误差小, 第 2 h 和第 3 h 的模型与加 MWD 特征的模型性能相近。加入 MSLP 特征的模型的绝对误差与加入 MWP 特征的模型相近。

综上所述, 加入 10 m 风 U 、 V 分量特征能显



注: 本文研究海域范围为 3.5°N—35°N 和 99.5°E—131°E

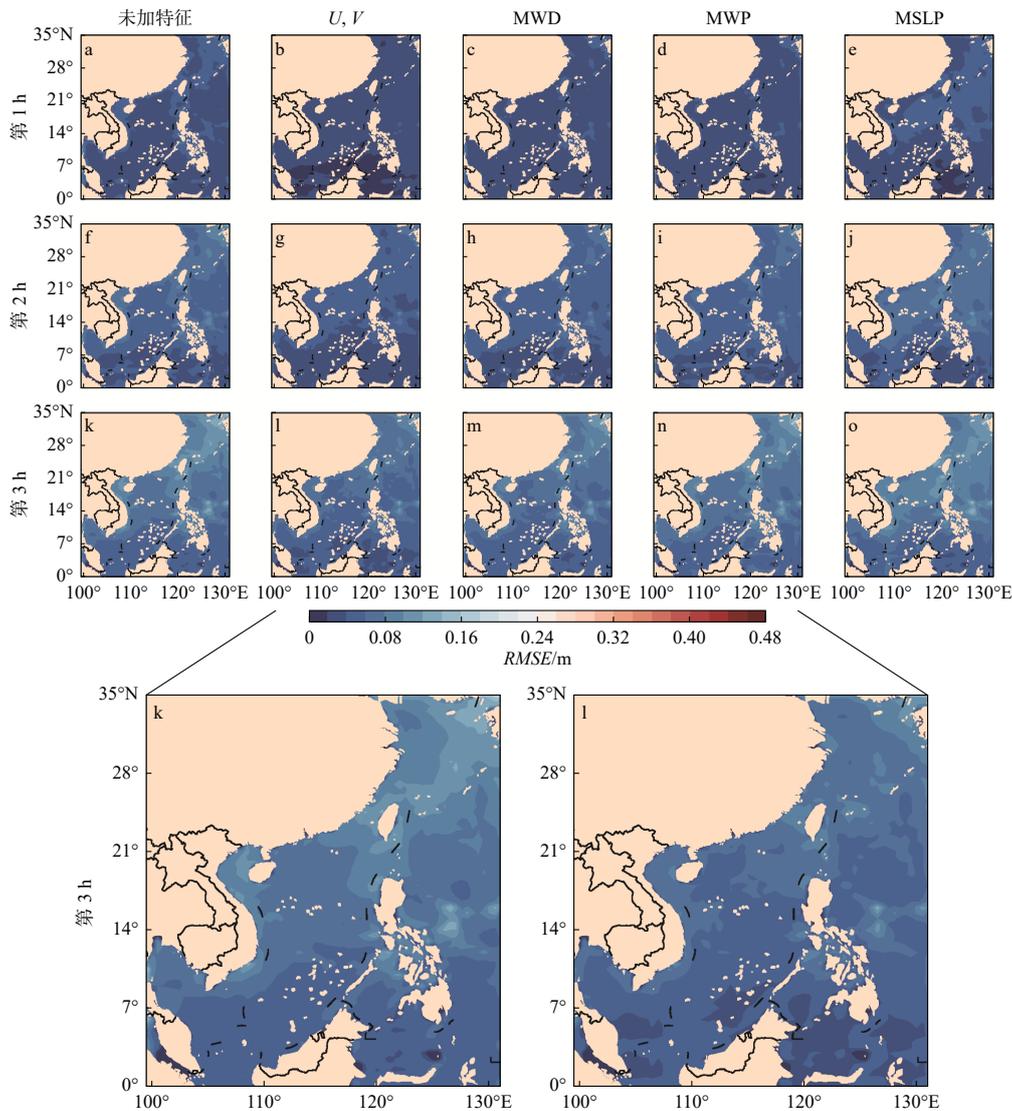
图 8 加入不同特征的模型在 2020 年 3 月 29 日 5 时至 7 时的预测值、预测值与观测值的绝对误差对比 [审图号: GS(2023)4355 号]

Fig. 8 Comparisons of the predictive values among the models with different features and the absolute errors between predictive values and observed values from 5:00 to 7:00 on March 29, 2020

著提升 SWH 的预测性能。

图 9 展示了加入不同特征的模型在测试集上每小时 $RMSE$ 的空间分布。蓝色越深,说明误差越小,红色越深,说明误差越大。在加入 10 m 风 U 、 V 分量的模型每小时分布图上深蓝色的区域最多,说明预测误差小于其他模型。放大第 3 h 未加特征和加入 10 m 风 U 、 V 分量特征的两

个 $RMSE$ 空间分布图,发现在加入 10 m 风 U 、 V 分量后,在南海海域的低纬度上,深蓝色的区域比未加特征的多,说明预测误差更小。而在东海海域,在加入 10 m 风 U 、 V 分量后,相应蓝色区域颜色深于未加特征的模型,说明预测误差小。上述结果表明本文模型选择加入 10 m 风 U 、 V 分量建模,可以进一步提升模型预测性能。



注: 本文研究海域范围为 3.5°N—35°N 和 99.5°E—131°E

图 9 加入不同特征的模型在测试集上 $RMSE$ 的空间分布 [审图号:GS(2023)4355 号]

Fig. 9 The spatial distribution of the $RMSE$ of the model with different features in the test set

3 结论

(1) 本文利用 $RMSE$ 、 MAE 、 M_RMSE 和 M_MAE 评价指标,给出预测值、预测值与观测

值的绝对误差、测试集 $RMSE$ 的可视化分析,评估本文模型的预测性能,并将其与其他深度学习模型进行比较。结果表明,本文提出的 CNN-STLSTM-CNN 模型预测性能优于其他模型,验

证了方法的有效性。

(2) 将 SWH 分别与 10 m 风 U 分量、 V 分量、平均波方向、平均波周期和平均海平面压力等要素结合, 评估加入不同特征的模型预测性能, 结果表明 10 m 风 U 、 V 分量特征能显著提升模型预测性能。

(3) 在所提及的模型中, CNN-STLSTM-CNN 模型的预测性能最优, 但其计算量较大, 训练时间较长。未来应考虑更加高效的 SWH 预测模型。

参考文献:

- [1] CHEN F D, LIU L L, GUO F. Sea surface height estimation with multi-GNSS and wavelet de-noising[J]. *Scientific reports*, 2019, 9(1): 15181.
- [2] TOLMAN HL. Distributed-memory concepts in the wave model WAVEWATCH III[J]. *Parallel Computing*, 2002, 28(1): 35-52.
- [3] FAN S T, XIAO N H, DONG S. A novel model to predict significant wave height based on long short-term memory network[J]. *Ocean Engineering*, 2020, 205: 107298.
- [4] PIRHOOSHYARAN M, SNYDER L V. Forecasting, hindcasting and feature selection of ocean waves via recurrent and sequence-to-sequence networks[J]. *Ocean Engineering*, 2020, 207: 107424.
- [5] POKHREL P, IOUP E, SIMEONOV J, et al. A transformer-based regression scheme for forecasting significant wave heights in Oceans[J]. *IEEE Journal of Oceanic Engineering*, 2022, 47(4): 1010-1023.
- [6] ZHOU S Y, XIE W H, LU Y X, et al. ConvLSTM-based wave Forecasts in the South and East China Seas[J]. *Frontiers in Marine Science*, 2021, 8: 680079.
- [7] HAN L, JI Q Y, JIA X Y, et al. Significant wave height prediction in the South China Sea based on the ConvLSTM algorithm[J]. *Journal of Marine Science and Engineering*, 2022, 10(11): 1683.
- [8] GAO Z Y, LIU X, YV F J, et al. Learning wave fields evolution in North West Pacific with deep neural networks[J]. *Applied Ocean Research*, 2023, 130: 103393.
- [9] CHEUNG K F. Wave Watch III (WW3) global wave model distributed by the Pacific Islands Ocean Observing System (PacIOOS)[EB/OL]. (2010-06-20). http://pacioos.org/meta-data/ww3_global.html.
- [10] HERBACH H, BELL B, BERRISFORD P, et al. The ERA5 global reanalysis[J]. *Quarterly Journal of the Royal Meteorological Society*, 2020, 146(730): 1999-2049.
- [11] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. *计算机学报*, 2017, 40(6): 1229-1251.
- [12] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille: JMLR. org, 2015: 448-456.
- [13] MAAS A L, HANNUN A Y, NG A Y, et al. Rectifier nonlinearities improve neural network acoustic models[C]//Proceedings of the 30th International Conference on Machine Learning. Atlanta: JMLR, 2013: 3.
- [14] WANG Y B, LONG M S, WANG J M, et al. PredRNN: Recurrent Neural Networks for Predictive Learning using Spatiotemporal LSTMs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 879-888.
- [15] SHI X J, CHEN Z R, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2015: 802-810.
- [16] DUMOULIN V, VISIN F. A guide to convolution arithmetic for deep learning [J]. arXiv preprint arXiv: 1603.07285, 2016.
- [17] KINGMA D P, BA J. Adam: A method for stochastic optimization[C]//Proceedings of the 3rd International Conference on Learning Representation. San Diego: ICLR, 2015: 1-15.
- [18] GAO Z Y, TAN C, WU L R, et al. SimVP: Simpler Yet Better Video Prediction[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 3160-3170.

(本文编辑: 胡莹莹)