

基于气象数值预报和卷积神经网络近岸定点海温预报研究

郭婷婷, 翁少佳*, 蔡锦海, 李惠, 罗荣真

(自然资源部汕头海洋中心, 广东 汕尾 516600)

摘 要: 基于遮浪海洋站实测海温数据和气象数值预报数据,对传统的卷积神经网络(CNN)进行优化,采用优化CNN方法搭建近岸定点短期海温逐时预报模型,并开展建模研究。研究显示:将观测数据与气象数值预报数据相结合用于训练神经网络模型,能够有效提升模型海温预测效果;在CNN参数调节中选择合适的卷积核数量,模型可发挥出最优的海温预报性能;相比传统CNN、长短期记忆网络(LSTM)和持续性预报方法,优化CNN方法新增归一化层,模型在训练过程中能够更快更好地收敛,在测试数据集中具备更优的海温预测效果,在0~24 h、24~48 h、48~72 h预报时效内,该模型预报结果的平均绝对误差分别为0.276 °C、0.412 °C、0.487 °C,均方根误差分别为0.459 °C、0.632 °C、0.714 °C,相关系数分别为0.994 7、0.990 0、0.987 2。

关键词: 海温预报;气象数值预报;卷积神经网络;归一化层

中图分类号: P731.31 **文献标识码:** A **文章编号:** 1003-0239(2026)04-0043-10

0 引言

随着海洋经济的不断发展,人们对精细化的海洋预报需求日益增强^[1-2]。近岸海域海温精确预报一直是海温预报的难点。目前海温预报方法主要为数值预报、统计预报和经验预报3种^[3]。虽然基于海洋动力学的数值预报模式作为海温预报的主要方法已经实现业务化运行,但是,由于受到陆地和人类活动影响,近岸海域海洋环境状况复杂,数值预报模式在近岸和浅水区域预报精度会大大降低,达不到预报业务化运行需求^[4]。目前,沿海城市在渔业养殖、海水浴场、滨海旅游度假区开展的海温预报主要依靠经验预报和传统统计预报方法,即预报员根据近岸海域海洋观测数据进行统计分析,结合未来天气系统变化,对预报区域开展日最高、日最低、日平均海温预报^[5-6]。传统的统计预报方法由于数据处理方法简单,难以充分挖掘海温潜在变化规律,预报准确性和可靠性较低;经验预报准确度

取决于预报员的工作经验,预报结果不稳定,存在一定误差。上述两种方法都难以满足精细化海温预报业务化需求,因此,探索新的海温预报技术,提高近岸海域海温预报准确性迫在眉睫。

近年来将人工神经网络算法应用于海温预报已成为研究热点,该算法以强大的数据处理和分析能力在近岸海域海温预报中表现十分出色^[7-10]。匡晓迪等^[11]以数值预报和观测数据为研究对象,利用反向传播(Back Propagation, BP)人工神经网络方法搭建舟山站海温数值预报释用模型,有效提高了定点海温预报精度。Zhang等^[7]首次使用长短期记忆网络(Long Short-Term Memory, LSTM)方法预测了渤海近岸表层海温(Sea Surface Temperature, SST),证明LSTM在近岸海温预报中具备可行性。林小刚等^[12]基于区域天气研究预报模式(Weather Research and Forecasting, WRF)的气象预报数据和海洋站点海温观测数据,利用LSTM方法搭建粤东近岸日均海温预报模型,预报效果较好,平均绝对误差为

收稿日期: 2025-07-07。

基金项目: 广东省平台基地及科技基础条件建设项目(2021B1212050025)。

作者简介: 郭婷婷(1983-),女,工程师,硕士,主要从事海洋观测预报研究。E-mail: 53811841@qq.com

*通信作者: 翁少佳(1990-),男,工程师,硕士,主要从事海洋观测预报研究。E-mail: w_shaojia@163.com

0.32 °C。李竞时等^[13]基于浮标海温观测数据和 WRF 模式气象预报数据,利用主成分分析方法(Principal Components Analysis, PCA)和 LSTM 方法搭建荣成、海阳两个站未来 5 日的 SST 预报模型,预报效果优于数值预报和经验预报方法。何恩业等^[14]基于 SST 和气象数值预报数据,利用 PCA 和 BP 神经网络相结合的方法搭建荣成近岸海域海温预报模型,可在业务化运行中提高预报速度,降低预报误差。赖晓倩等^[15]搭建了一种集成差分回归可迁移长短期记忆网络(Differential Regression Transferable Long Short-Term Memory, DR-TLSTM)模型,用于三沙湾渔业水域的水温业务化预报,有效解决了该区域缺乏历史水温监测数据和水温突变预测困难的问题,具有良好的推广应用潜力。Jiao 等^[16]利用 LSTM 方法搭建了小麦岛海洋站未来 3 日 SST 预报模型,并将该方法应用于渤海和黄海沿岸其他海洋站点 SST 预报中,与数值模式相比,该方法能够降低 SST 预报误差,提高预报精确度。

与数值预报模式不同,人工神经网络方法不涉及复杂的数学和物理过程,能够直接从大量数据中学习隐含的变化规律,并对预测对象作出快速准确的预测。神经网络算法可以视为数据驱动方法,数值预报模式可视为理论驱动方法^[17]。虽然两者属于不同领域,但可以将它们融合,即利用人工神经网络对数值模式输出的预报数据进行学习,从预报数据中提取有用的预报信息,提高人工神经网络预报的准确性。目前应用于近岸海域海温预报研究的神经网络算法主要集中于 BP 神经网络和 LSTM,后者已成为许多海温预报研究的首选神经网络方法。这是因为近岸海域海温预报研究主要基于近岸站点或者浮标观测得到的时序数据,LSTM 具有更好的记忆性能,能够留存前文输入信息,从而充分捕捉前后序列之间的关联性,在处理时序数据方面具有明显优势;但其缺点是无法对数据进行并行化处理,网络训练速度慢。

卷积神经网络(Convolutional Neural Network, CNN)作为图像和视频处理领域公认的主力军,过去被认为不适合于时间序列分析建模,主要原因是其卷积核在处理信息时没有记忆功能。但是近年来越来越多的研究显示,将 CNN 用于时序预测也能取得很好的效果^[18-20]。在近岸海域海温预报方面,

CNN 已经被应用于定点的表层海温预报,且预报效果能与 LSTM 相媲美^[21]。该方法基于历史观测数据可实现近岸单点的日平均、日最高、日最低表层海温预报,但缺点是仅依靠历史观测数据难以全面学习海温变化规律,大大制约了海温预报的准确度。

为了进一步提高 CNN 在近岸海温预报中的准确度,本文以遮浪海洋站 SST 为研究对象,尝试将 CNN 与气象数值预报模式相结合,构建预报时效为 72 h 的近岸站点海温预报模型并开展研究。首先对传统的 CNN 结构进行优化,使其更适用于时序数据学习;其次分析测数据和气象预报数据对建模效果的影响,建立预报因子最优训练方案;然后研究不同卷积核数量下的模型预测性能变化,确定卷积核数量最优值;最后基于最优配置方案和测试数据,将优化后的 CNN 算法与传统的 CNN、LSTM 算法开展对比试验,验证优化后的 CNN 模型的有效性。

1 模型与数据

1.1 CNN 模型结构

在近岸定点海温预报研究中,输入数据为按时间顺序排列的数值,一般会采用一维 CNN 来搭建模型。传统的一维 CNN 结构一般由输入层、卷积层、池化层、扁平(flatten)层、输出层组成。本文对传统的一维 CNN 结构进行优化,以归一化层、激活层代替池化层,即优化后的一维 CNN 结构主要由输入层、卷积层、归一化层、激活层、扁平(flatten)层、输出层组成。两种 CNN 模型结构见图 1。

卷积层采用适合处理时间序列数据的一维卷积运算,即卷积核只沿着输入序列的时间维度进行滑动计算。卷积操作可以有效学习时序数据中的局部特征,通过设置多个卷积核组合,可以提取更加复杂的特征。

在传统的一维 CNN 结构中,池化层位于卷积层之后,卷积层对输入数据进行特征提取后传输给池化层,池化层则对输入特征进行压缩,降低特征维度,减少参数量,提取更高级的特征。该过程不可避免导致一部分时序特征信息(如细节信息和时间维度上的位置信息)丢失。这是因为在处理时序数据时,池化操作主要采用最大值池化或平均值池化,即选取数据区域中某个特定值来代表该区域,

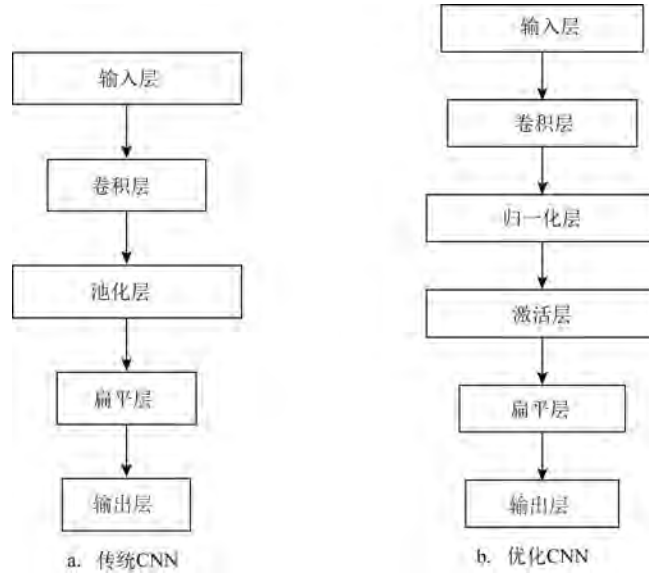


图1 CNN模型结构图

Fig.1 Framework of CNN model

这种简化过程容易忽略区域中其他潜在的重要信息,导致部分细节信息丢失。

为了避免上述问题,本文在优化CNN结构中,使用归一化层、激活层代替池化层。归一化层采用了应用最广泛最经典的批量归一化(Batch Normalization, BN)操作。批量归一化由Ioffe等^[22]于2015年提出,是近年来神经网络架构领域一项创新技术。数据集在输入模型前会划分为若干个批量(batch),批量归一化操作是对上一层输出的每个batch求均值和标准差,即进行标准化操作,使每个batch中样本服从标准正态分布,然后在该操作基础上对每个batch进行拉伸和偏移。具体公式为:

$$\hat{x} = \frac{x - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (1)$$

$$\text{BN}_{\gamma, \beta}(x) = \gamma \times \hat{x} + \beta \quad (2)$$

式中: μ_B 、 σ_B^2 分别表示一个batch中所有样本的均值、方差。式(2)引入可学习的参数 γ 、 β ,并在反向传播过程中进行更新,可通过移动和缩放归一化值帮助网络获取最佳预测结果。

未引入归一化操作时,神经网络在训练过程中,每一层的参数和数据分布会持续更新变化,再传输给下一层,使得下一层在提取数据特征时需要不断适配新的数据分布,最终导致神经网络训练过

程不稳定。使用归一化层代替池化层后,批量归一化会对上一层输出的每个batch数据进行标准化处理,使其数据分布不会发生较大变化,保证每个batch内特征维度分布维持一定的稳定性,有助于网络在训练过程中降低梯度消失或梯度爆炸的可能性,使模型更快更好地收敛,提升模型训练稳定性。此外,批量归一化中引入可学习的参数 γ 和 β ,通过前向传播和反向传播完成参数调整,使得网络能够更好地适配不同的数据分布,更易学习数据内在特征,进一步增强模型的适应性和学习能力,提高模型预测性能。

激活层主要通过引入激活函数,对前一层的输出结果进行非线性变换,赋予输出特征非线性属性,帮助网络学习和拟合复杂数据,有效提升神经网络模型的非线性拟合能力。本文的激活函数采用常用的ReLU函数。扁平层和输出层负责将多维特征向量转换成一维向量并完成输出。

1.2 数据

1.2.1 数据来源

本文使用的海温数据来源于汕尾遮浪半岛遮浪海洋站2020—2023年逐时观测SST数据。该数据通过自动化方式采集获取,能够较好地反映遮浪半岛近岸海温环境变化。

气象数值预报数据来自自然资源部南海预报减灾中心业务化运行的中尺度天气预报 WRF 模式输出结果。该模型空间范围为 $96.1^{\circ}\sim 135.0^{\circ}\text{E}$, $-5^{\circ}\sim 31.5^{\circ}\text{N}$, 模型采用规则网格, 预报时长 78 h, 输出要素时间间隔 1 h。本文选取该模型 2020—2023 年海面 2 m 气温 (T_2)、海平面气压 (Sea Level Pressure, SLP)、2 m 相对湿度 (Q_2)、10 m 风速 (U_{10})、风向 (V_{10}) 5 个要素开展研究, 并利用双线性内插法将气象要素数据插值到台站对应位置, 构建逐时气象预报数据集。

1.2.2 数据集构造与处理

将海温观测数据和气象预报数据划分为训练数据和测试数据, 分别用于模型训练和预报效果评估。其中 2020 年 1 月 1 日 00 时 (北京时, 下同)—2022 年 12 月 31 日 24 时逐时数据为训练数据, 2023 年 1 月 1 日 00 时—12 月 31 日 24 时逐时数据为测试数据。所有数据输入模型前需要进行统一预处理, 将各要素标准化。标准化处理所用的均值和标准偏差, 均基于剔除异常值和缺失值后计算求得。采用滑动窗口方法将数据集分割成相同时间长度的数据样本及对应预报实测值, 并剔除包含异常值和缺测值的数据样本, 其中训练数据集按小时滑动分割, 测试数据集按日滑动分割, 经筛选共得到 363 个测试样本。

2 实验与分析

本实验选择平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Square Error, RMSE) 和相关系数 (r) 作为预报效果评估指标。将未来 3 d 逐时 SST 预报划分为 0~24 h、24~48 h、48~72 h 三个预报时效, 根据各时效训练对应预报模型并计算评估指标。

实验选择传统 CNN 模型、LSTM 模型与本文优化 CNN 模型进行比较。模型基于 Python 和模块化神经网络库 keras 搭建。模型训练最大迭代次数为 15 000 次, 学习率设置为 0.000 1, 损失函数选择 MAE, 输出层使用 ReLU 激活函数。CNN 模型卷积核大小为 3, 池化层采用平均值池化, 池化窗口大小为 2; LSTM 模型隐藏层数量设置为 4。

由于神经网络训练过程存在随机因素, 模型在

数据集上的预测结果会有一定的浮动。本文将每组实验模型重复运行 6 次, 对各项评估指标取平均值作为最终预测结果。

2.1 不同训练方案的对比分析

为了研究海温观测数据集和气象预报数据集各要素对海温预报结果的影响, 本文对输入因子进行筛选组合, 设计 3 组不同训练方案 (见表 1)。方案 1 为历史 SST 训练方案, 即选用海洋站过去 24 h 观测 SST 作为模型输入预报因子, 未来 24 h 实测 SST 作为模型输出期望值。方案 2 为气象数值预报要素训练方案, 即选用气象预报模型未来 24 h 预报的 T_2 、SLP、 Q_2 、 U_{10} 和 V_{10} 作为模型输入预报因子, 未来 24 h 观测 SST 作为模型输出期望值。方案 3 为 SST 观测+气象数值预报要素联合训练方案, 即选用海洋站过去 24 h 观测 SST, 以及气象预报模型未来 24 h 预报的 T_2 、SLP、 Q_2 、 U_{10} 、 V_{10} 作为模型输入预报因子, 未来 24 h 观测 SST 作为模型输出期望值。

表 1 模型训练方案

Tab.1 Training cases for CNN

训练方案	输入因子						输出期望值
	SST	T_2	SLP	Q_2	U_{10}	V_{10}	SST
方案 1	√	—	—	—	—	—	√
方案 2	—	√	√	√	√	√	√
方案 3	√	√	√	√	√	√	√

注: “√”表示参与训练, “—”表示不参与训练。

将 3 种方案的预报因子分别输入优化 CNN 模型进行训练, 使用训练完成的模型对 2023 年 SST 进行预报, 3 种方案预报效果见表 2。3 种方案中, 方案 2 模型预报误差较大, MAE 和 RMSE 均超过 1.2°C , 相关系数较低, 整体预报效果较差。可见仅依靠气象数值预报要素, 神经网络难以充分学习海温变化内在规律。方案 1 模型 0~24 h 预报误差远低于方案 2, MAE 和 RMSE 分别为 0.313°C 和 0.520°C , 误差较方案 2 降低近 1°C 。说明海温时序变化与历史海温变化密切相关, 神经网络可以从历史海温数据挖掘变化规律, 并对未来海温进行较好预报。相较于方案 1, 方案 3 在实测 SST 基础上引入气象预报数

表2 不同训练方案的预报结果对比

Tab.2 Comparison of training cases

训练方案	0~24 h		
	MAE/℃	RMSE/℃	<i>r</i>
方案1	0.313	0.520	0.9932
方案2	1.268	1.580	0.9420
方案3	0.276	0.459	0.9947

注:黑色加粗为最优值。

据,训练所得模型能够有效改善海温预报精度、降低预报误差、提高相关系数,其中MAE降低0.037℃,RMSE降低0.061℃。

绘制2023年4月12—15日方案1和方案3的逐时海温预报曲线(见图2)。在海温剧烈变化时段,方案1预报曲线相对实测曲线存在明显滞后,预报误差较大。这是因为方案1模型高度依赖历史海温数据,神经网络无法依托历史信息捕捉海温突变规律,海温发生突变时模型易产生惯性预报。方案3引入气象数值预报数据,模型能够从气象信息中提前捕捉海温未来变化趋势并进行准确预测。综合误差和相关系数两项指标,方案3构建的预测模型效果更优。

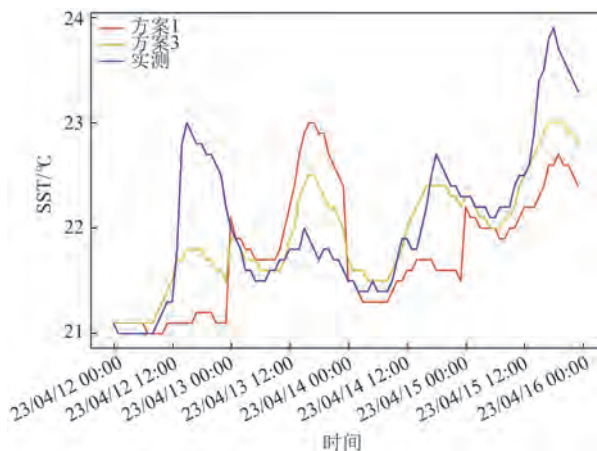


图2 4日预报结果与实测海温比较

Fig.2 Comparison of predicted and observed values in the prediction of 4 days

2.2 气象预报数据时长对建模的影响分析

为进一步研究输入气象预报数据长度对建模效果的影响,本文在方案3的基础上新增两组训练

方案,比较不同气象预报数据长度下,模型在24~48 h、48~72 h时效的海温预报效果。方案3-1:输入因子为过去24 h观测 SST 和未来24~48 h/48~72 h气象预报数据,输出期望因子为对应时段观测 SST。方案3-2:输入因子为过去24 h观测 SST 和未来0~48 h/0~72 h气象预报数据,输出期望因子为未来24~48 h/0~72 h观测 SST。将两组方案分别输入优化 CNN 模型训练并预报2023年 SST,预报效果见表3。

表3 两组训练方案的 SST 预报结果对比

Tab.3 Comparison of two training cases under different input numerical weather prediction data lengths

训练方案	气象预报 数据时长/h	0~24 h		
		MAE/℃	RMSE/℃	<i>r</i>
方案3-1	24~48	0.439	0.678	0.988 4
方案3-2	0~48	0.412	0.632	0.990 0

训练方案	气象预报 数据时长/h	24~48 h		
		MAE/℃	RMSE/℃	<i>r</i>
方案3-1	48~72	0.555	0.794	0.984 0
方案3-2	0~72	0.487	0.714	0.987 2

注:黑色加粗为最优值。

相较于方案3-1,方案3-2能够有效提升海温预报模型、降低预报误差;其中48~72 h时效的预报误差降低最明显,MAE降低0.068℃,RMSE降低0.08℃,*r*提高0.003 2。天气变化是影响海温变化的重要因素之一,气象要素对海温的影响具有滞后性,前期天气变化会影响后续海温状态。这意味着预报时段的海温不仅与该时段气象条件有关,还受前置时刻气象要素共同影响;同时海温随时间的变化存在一定连续性和关联性,前一时刻的海温预报值会影响下一时刻的预报结果。方案3-1仅输入预报时段对应的气象数据,模型学习到的海温变化规律有所欠缺,预报起始时刻误差较大导致该时段整体误差偏高;方案3-2增加了预报时段之前的气象数据,模型能够完整捕捉海温变化规律,对海温的连续性变化进行很好预测,降低预报误差。综合误差和相关系数来看,方案3-2在24~48 h、48~72 h时效中预报效果更好。

2.3 卷积核数量对建模的影响分析

卷积核是CNN中重要的超参数,卷积核数量的设置会影响模型的性能。本文基于方案3和方案3-2,对比3种CNN算法在不同卷积核数量下未来3 d逐时海温预报结果(见表4)。其中CNN-AP为带平均池化层的传统CNN;CNN-BN为本文提出的引入批量归一化操作的优化CNN;CNN为移除归一化层的CNN-BN基础结构。从表中看出,随着卷积核数量增加,3种CNN模型0~72 h预报效果得到改善,误差MAE、RMSE有所下降, r 整体上升。其中CNN-AP的MAE平均下降0.021 °C, RMSE平均下降0.024 °C, r 平均提高0.000 8; CNN的MAE平均下降0.019 °C, RMSE平均下降0.025 °C, r 平均提高0.000 9; CNN-BN的MAE平均下降0.019 °C, RMSE平均下降0.020 °C, r 平均提高0.000 6。卷积核数量由8增加到10时,3种模型的预报效果提升不明显,少部分的预报效果甚至变差。CNN中增加卷积核数量能够让网络更全面理解输入数据,学习多样性特征,提高模型性能;然而,卷积核数量并不是越多

越好。实验结果表明模型性能随卷积核数量增加先上升后下降,主要原因是过多卷积核会导致模型过拟合,降低模型泛化能力。由此CNN模型需选择合适的卷积核数量,既充分提取数据特征,又避免过拟合问题,使模型发挥最优性能。本实验中,根据模型的性能变化,3种CNN模型的卷积核数量最优值为8~10。

2.4 不同算法对比分析

为进一步检验本文提出的卷积神经网络结构在建模过程中的效果,我们比较3种CNN方法未来3 d逐时海温预报效果,并与LSTM方法、持续性预报方法进行比较。持续性预报是将过去24 h观测得到的SST作为未来0~24 h、24~48 h、48~72 h三个时段的预报SST。

从表中可以看出,CNN-AP和CNN各项预报指标比较接近。CNN-BN较CNN-AP、CNN、LSTM和持续性预报方法可以获得较优的预报效果。相较CNN-AP,CNN-BN在0~72 h预报中MAE平均下降0.031 °C, RMSE平均下降0.045 °C, r 平均提高

表4 3种CNN模型的SST预报效果

Tab.4 Results of three CNN forecast methods

模型	卷积核数量	MAE/°C			RMSE/°C			r		
		0~24 h	24~48 h	48~72 h	0~24 h	24~48 h	48~72 h	0~24 h	24~48 h	48~72 h
CNN-AP	4	0.321	0.473	0.536	0.518	0.719	0.776	0.993 3	0.987 0	0.984 9
	6	0.299	0.457	0.531	0.492	0.709	0.762	0.993 9	0.987 4	0.985 4
	8	0.296	0.454	0.538	0.489	0.693	0.776	0.994 0	0.988 0	0.984 9
	10	0.303	0.440	0.538	0.492	0.689	0.765	0.993 9	0.988 1	0.985 4
CNN	4	0.308	0.464	0.536	0.504	0.708	0.776	0.993 7	0.987 4	0.984 8
	6	0.301	0.442	0.519	0.487	0.683	0.753	0.994 0	0.988 3	0.985 8
	8	0.300	0.433	0.523	0.491	0.674	0.750	0.994 0	0.988 5	0.985 9
	10	0.300	0.445	0.519	0.490	0.687	0.749	0.994 0	0.988 1	0.986 0
CNN-BN	4	0.283	0.437	0.507	0.474	0.658	0.728	0.994 4	0.989 1	0.986 9
	6	0.281	0.425	0.511	0.469	0.656	0.734	0.994 5	0.989 1	0.986 5
	8	0.274	0.411	0.496	0.463	0.637	0.725	0.994 6	0.989 8	0.986 9
	10	0.276	0.412	0.487	0.459	0.632	0.714	0.994 7	0.990 0	0.987 2

注: 黑色加粗为最优值。

表5 5种方法的SST预报效果
Tab.5 Results of five forecast methods

模型	MAE/℃			RMSE/℃			<i>r</i>		
	0~24 h	24~48 h	48~72 h	0~24 h	24~48 h	48~72 h	0~24 h	24~48 h	48~72 h
CNN-AP	0.296	0.440	0.531	0.489	0.689	0.762	0.994 0	0.988 1	0.985 4
CNN	0.300	0.443	0.519	0.490	0.674	0.749	0.994 0	0.988 5	0.986 0
CNN-BN	0.276	0.412	0.487	0.459	0.632	0.714	0.994 7	0.990 0	0.987 2
LSTM	0.310	0.480	0.566	0.502	0.709	0.771	0.993 7	0.987 5	0.985 6
持续性预报	0.426	0.636	0.787	0.669	0.922	1.075	0.988 8	0.978 9	0.971 4

0.001 5; 相较 CNN, CNN-BN 在 0~72 h 预报中 MAE 平均下降 0.026 ℃, RMSE 平均下降 0.036 ℃, *r* 平均提高 0.001 1; 相较 LSTM, CNN-BN 在 0~72 h 预报中 MAE 平均下降 0.06 ℃, RMSE 平均下降 0.059 ℃, *r* 平均提高 0.001 7。相较持续性预报, CNN-BN 在 0~72 h 预报中 MAE 平均下降 0.225 ℃, RMSE 平均下降 0.287 ℃, *r* 平均提高 0.010 9。

我们对 3 种时效逐时海温误差进行统计分析, 即对同一时效下所有天数同一时刻的海温预报误差求取平均值, 绘制 5 种方法的预报误差随预报时效增加的变化曲线图, 结果见图 3。从表 5 结果和图 3 变化曲线可以看出, 5 种方法的预报误差随着预报时效延长呈明显递增趋势。其中, 0~72 h 预报中持续性预报方法预报误差曲线递增最快、误差最大; CNN-AP 和 CNN 方法预报误差曲线差距不大; CNN-BN 方法预报误差曲线递增最慢, 总体误差小于其他 4 种方法。

我们分别统计了 363 个样本在 0~72 h 时效内的平均预报误差, 绘制 CNN-BN 与其他 3 种神经网络方法误差比较直方图。在 363 个样本中, CNN-BN 方法预报误差总体低于其他 3 种神经网络方法, 其中 CNN-BN 方法预报误差低于 CNN-AP 方法预报误差的样本占比为 64.7%, 低于 CNN 方法的样本占比为 63.1%, 低于 LSTM 方法的占比为 72.2%。在 363 个样本中, CNN-AP 方法的最大 MAE 为 2.84 ℃, CNN 方法为 2.64 ℃, CNN-BN 方法为 2.41 ℃, LSTM 方法为 2.49 ℃。CNN-BN 方法的最大 MAE 均小于其他 3 种神经网络方法。

从海温预报整体效果来看, 持续性预报方法的预报误差最大, 预报效果最差。这是因为在近岸海

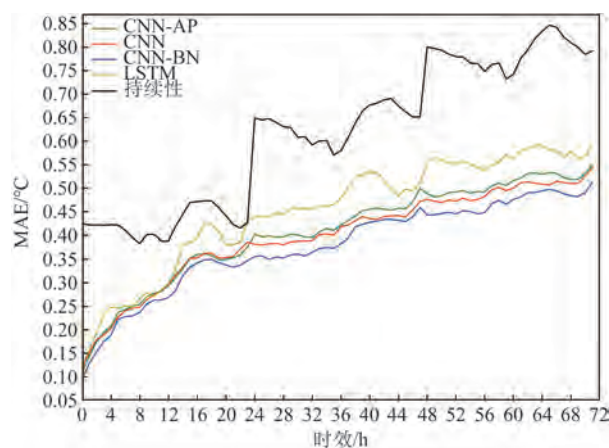


图3 5种方法的预报误差随预报时效增加的变化曲线图

Fig.3 MAE variation of five forecast methods

域中, 海温日变化规律存在一定差异, 并且随着天数间隔增大, 海温日变化差异也逐渐增大。仅依靠持续性预报方法开展近岸站点海温预报存在很大缺陷。CNN方法略优于CNN-AP方法, CNN-BN方法优于其他3种神经网络方法, 主要原因在于它们的网络结构设计存在区别。相较于CNN、CNN-AP结构, CNN-BN利用归一化层代替了池化层, 不仅能够避免池化操作在特征提取过程中造成特征信息丢失, 还能够使模型在训练过程中保持稳定, 更快从时序中学习数据特征, 提升模型海温预报效果。LSTM在处理时序数据时, 随着序列长度增加会出现梯度消失、特征信息捕捉不足等问题, 导致模型在训练过程中收敛缓慢, 预测准确度不高。CNN-BN采用卷积运算提取数据特征, 能够高效挖掘时序数据中的局部依赖关系, 更充分地从数据中捕捉预报因子与海温变化的关系。

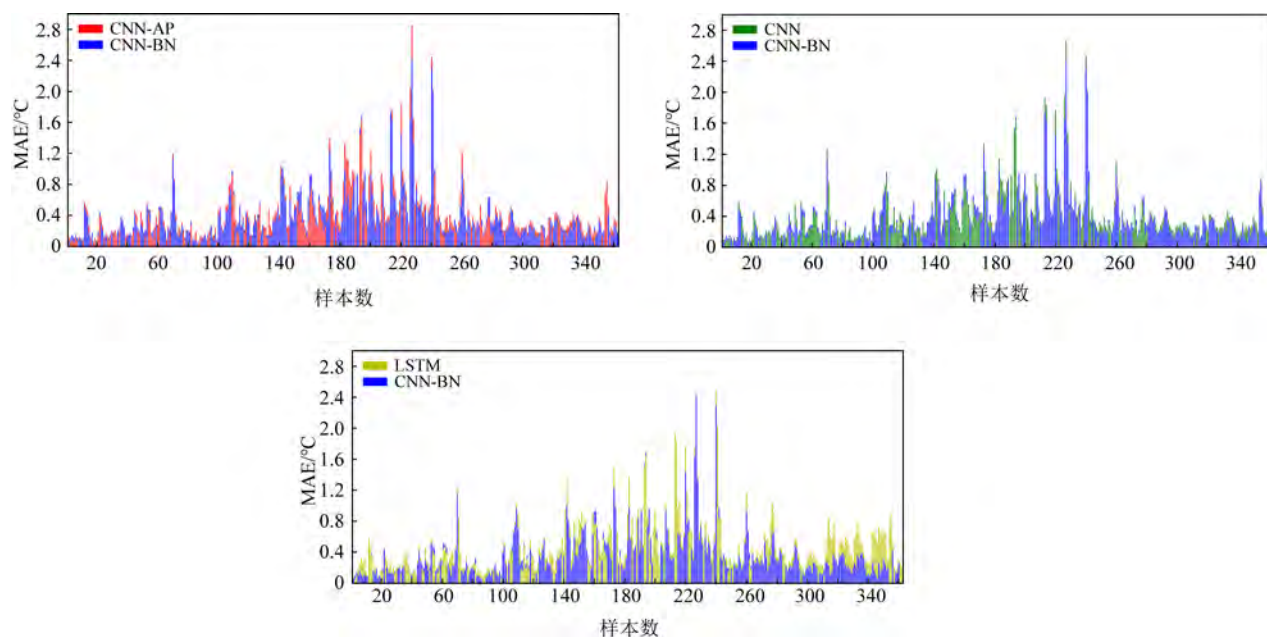


图4 4种神经网络方法在0~72 h时效中平均预报误差直方图

Fig.4 The 0~72 h MAE values of four neural network methods in the test set

3 结论与讨论

本文基于遮浪海洋站 SST 观测数据和 WRF 模式输出逐时气象预报数据,对传统的 CNN 方法进行优化,引入归一化层,利用优化后 CNN 方法搭建预报时效为 72 h 的逐时海温预报模型,研究输入数据、卷积核数量、归一化层对建模效果的影响,并将其与传统 CNN、LSTM 和持续性预报方法进行对比分析。结论如下:

①相比单一采用海温观测数据或气象数值预报数据训练模型,将两种数据相结合训练得到的模型,在海温预报中能够很好地捕捉海温未来变化趋势并完成准确预测,有效提升模型预报性能、降低预报误差、提高预测值和实测值相关性。

②在 24~48 h、48~72 h 预报时效中,在输入数据中增加预报时段之前的气象预报数据用于模型训练,可以增强模型对海温预测的连续性,降低海温预报误差。

③在 CNN 建模过程中,选择合适的卷积核数量,能够帮助模型提取多样性特征,使模型发挥最优性能。基于 CNN 开展海温预报建模中,卷积核数

量为 8~10 时,模型海温预报性能达到最佳。

④本文提出的引入归一化层的优化 CNN 模型,相较于传统 CNN、LSTM 模型和持续性预报方法,在近岸定点海温预报中表现出最优效果,全年海温平均预报误差最小,预测值和实测值相关性最好,预报误差增幅最小。

从本文的研究可以看出,输入数据质量、网络结构是影响神经网络算法海温预报建模效果的重要因素。筛选并加入影响海温变化的相关因子,是提升输入数据质量的方式之一。本文引入气象数值预报,将历史海温数据和气象预报数据相结合,利用模型建立预报目标与历史观测数据、气象预报数据之间的非线性关系,实验证实引入气象数值预报可以提升模型的海温预测性能。但模型的预测效果及稳定性会受气象数值预报结果的影响。只有气象数值预报结果准确反映天气变化实况,模型才能从数据中捕捉到预报数据与海温变化的关系。从预报误差直方图可以看出,4 种模型的较大误差主要集中在夏季。这是因为夏季遮浪站的海温除受天气变化影响外,还会受到低温上升流影响,海温变化不稳定^[23]。这意味仅通过引入气象数值预报来提升遮浪站海温预报模型性能存在局限性。

在网络结构方面,本文仅对传统CNN结构进行优化,并没有构造层次更深、结构更复杂的神经网络模型。构建深层复杂神经网络会大幅增加模型参数量,占用更多计算资源,提升训练难度,同时对输入数据的数量和质量提出更高要求。搭建海温预测模型时,盲目堆积当下主流网络结构、一味追求网络深度,并不能提升模型预测性能。过深、过复杂的网络会导致模型在训练过程中出现过拟合,降低模型性能。神经网络模型的搭建需要根据任务需求以及数据集特点进行选择和优化。后续研究将全面深入梳理近岸站点海温变化的相关因子,提升输入数据质量,并根据海温变化特性尝试引用其他神经网络算法搭建预报模型,持续提升预报精度。

参考文献:

- [1] 刘娜,王辉,凌铁军,等. 全球业务化海洋预报进展与展望[J]. 地球科学进展, 2018, 33(2): 131-140.
Liu Na, Wang Hui, Ling Tiejun, et al. Review and prospect of global operational ocean forecasting[J]. *Advances in Earth Science*, 2018, 33(2): 131-140.
- [2] 张永垂,陈诗尧,王宁,等. 全球业务化海洋预报系统进展[J]. 地球科学进展, 2022, 37(4): 344-357.
Zhang Yongchui, Chen Shiyao, Wang Ning, et al. Progress of global operational ocean forecasting systems[J]. *Advances in Earth Science*, 2022, 37(4): 344-357.
- [3] 张建华. 海温预报知识讲座 第一讲 海水温度预报概况[J]. 海洋预报, 2003, 20(4): 81-85.
Zhang Jianhua. Lecture on sea water temperature prediction: Lecture 1 overview of sea water temperature prediction[J]. *Marine Forecasts*, 2003, 20(4): 81-85.
- [4] 王兆毅,李云,王旭. 中国近岸海域基础预报单元海温预报指导产品研制[J]. 海洋预报, 2020, 37(4): 59-65.
Wang Zhaoyi, Li Yun, Wang Xu. Development of forecast guidance product for sea temperature of basic forecast units in the Chinese coastal waters[J]. *Marine Forecasts*, 2020, 37(4): 59-65.
- [5] 张建华. 海温预报知识讲座 第二讲 数理统计方法在海温预报中的应用[J]. 海洋预报, 2004, 21(1): 85-90.
Zhang Jianhua. Lecture on sea water temperature prediction: Lecture 2 application of mathematical statistical methods on sea water temperature prediction[J]. *Marine Forecasts*, 2004, 21(1): 85-90.
- [6] 赵强,王擎宇,舒志光. 基于SARIMA模型的近岸海表温度短期预报研究[J]. 海洋预报, 2024, 41(1): 42-49.
Zhao Qiang, Wang Qingyu, Shu Zhiguang. A study of short-term forecast of nearshore SST based on SARIMA model[J]. *Marine Forecasts*, 2024, 41(1): 42-49.
- [7] Zhang Qin, Wang Hui, Dong Junyu, et al. Prediction of sea surface temperature using long short-term memory[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1745-1749.
- [8] Zhang Xiaoyu, Li Yongqing, Frery A C, et al. Sea surface temperature prediction with memory graph convolutional networks [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 8017105.
- [9] 贺琪,查铖,宋巍,等. 基于STL的海表面温度预测算法[J]. 海洋环境科学, 2020, 39(6): 918-925.
He Qi, Zha Cheng, Song Wei, et al. Sea surface temperature prediction algorithm based on STL model[J]. *Marine Environmental Science*, 2020, 39(6): 918-925.
- [10] 杨乐晴,王丽娜,张红春,等. 基于STL的南海海表温度组合预测模型[J]. 海洋环境科学, 2024, 43(1): 109-118.
Yang Leqing, Wang Lina, Zhang Hongchun, et al. Combining forecasting model for sea surface temperature in the South China Sea based on STL[J]. *Marine Environmental Science*, 2024, 43(1): 109-118.
- [11] 匡晓迪,王兆毅,张苗茵,等. 基于BP神经网络方法的近岸数值海温预报释用技术[J]. 海洋与湖沼, 2016, 47(6): 1107-1115.
Kuang Xiaodi, Wang Zhaoyi, Zhang Miaoyin, et al. An interpretation scheme of numerical near-shore sea-water temperature forecast based on BPNN[J]. *Oceanologia et Limnologia Sinica*, 2016, 47(6): 1107-1115.
- [12] 林小刚,王兆毅,李竞时,等. 基于LSTM神经网络方法的粤东近岸海温预报[J]. 海洋预报, 2022, 39(5): 27-36.
Lin Xiaogang, Wang Zhaoyi, Li Jingshi, et al. Sea temperature forecasting based on LSTM neural network along the coast of eastern Guangdong[J]. *Marine Forecasts*, 2022, 39(5): 27-36.
- [13] 李竞时,匡晓迪,李琼,等. 基于主成分分析和LSTM神经网络的海温预报模型[J]. 海洋预报, 2023, 40(2): 1-10.
Li Jingshi, Kuang Xiaodi, Li Qiong, et al. SST forecasting model based on principal component analysis and LSTM neural network [J]. *Marine Forecasts*, 2023, 40(2): 1-10.
- [14] 何恩业,李琼,张聿柏,等. 基于PCA-BP特征工程的近海单点海温预报技术及应用[J]. 海洋预报, 2023, 40(3): 35-44.
He Enye, Li Qiong, Zhang Yubai, et al. Technology and application of offshore SST prediction based on PCA-BP feature engineering[J]. *Marine Forecasts*, 2023, 40(3): 35-44.
- [15] 赖晓倩,余镒琦,梁中耀,等. 基于差分回归模型和可迁移长短期记忆网络集成的三沙湾水温预测[J]. 海洋学报, 2023, 45(4): 165-178.
Lai Xiaoqian, Yu Yiqi, Liang Zhongyao, et al. Water temperature prediction in the Sansha Bay based on the integration of differential regression model and transportable long short-term memory network[J]. *Haiyang Xuebao*, 2023, 45(4): 165-178.
- [16] Jiao Yan, Li Ge, Zhao Peng, et al. Construction of sea surface temperature forecasting model for Bohai Sea and Yellow Sea coastal stations based on long short-time memory neural network

- [J]. *Water*, 2024, 16(16): 2307.
- [17] 李双林, 张仲石, 王惠. 数值天气预报的未来是人工智能与数学物理模型的融合?[J]. *地球科学*, 2022, 47(10): 3919-3921.
- Li Shuanglin, Zhang Zhongshi, Wang Hui. Will the future of numerical weather prediction be a fusion of artificial intelligence and mathematical and physical modeling? [J]. *Earth Science*, 2022, 47(10): 3919-3921.
- [18] Van Den Oord A, Dieleman S, Zen Heiga, et al. WaveNet: A generative model for raw audio[C]//*Proceedings of the 9th ISCA Workshop on Speech Synthesis (SSW 9)*. Sunnyvale: ISCA, 2016.
- [19] Borovykh A, Bohte S, Oosterlee C W. Conditional time series forecasting with convolutional neural networks[PP/OL]. V2. arXiv (2017-05-16)[2024-09-29]. <https://arxiv.org/abs/1703.04691v2>.
- [20] Bai Shaojie, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[PP/OL]. V2. arXiv (2018-4-19) [2024-09-29]. <https://doi.org/10.48550/arXiv.1803.01271>.
- [21] 翁少佳, 蔡锦海, 庞运禧, 等. 卷积神经网络在近岸表层海温预报中的应用[J]. *热带海洋学报*, 2024, 43(1): 40-47.
- Weng Shaojia, Cai Jinhai, Pang Yunxi, et al. Application of convolutional neural network to sea surface temperature prediction in the coastal waters[J]. *Journal of Tropical Oceanography*, 2024, 43(1): 40-47.
- [22] 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*. Online: PMLR, 2015: 448-456.
- [23] 张云翼, 江毓武. 汕尾外侧冷水跨架输送的形成机制[J]. *厦门大学学报(自然科学版)*, 2012, 51(4): 746-752.
- Zhang Yunyi, Jiang Yuwu. The mechanism of cold water cross-shelf transport in the continental shelf off Shanwei[J]. *Journal of Xiamen University (Natural Science)*, 2012, 51(4): 746-752.

A study of sea surface temperature prediction in the nearshore stations based on Numerical Weather Prediction and Convolutional Neural Network

GUO Tingting, WENG Shaojia*, CAI Jinhai, LI Hui, LUO Rongzhen

(Shantou Marine Center, Ministry of Natural Resources, Shanwei 516600, China)

Abstract: Based on the sea surface temperature (SST) observational data of Zhelang Station and the Numerical Weather Prediction (NWP) results, this study optimizes the traditional Convolutional Neural Networks (CNN) to construct a short-term forecasting model for nearshore hourly SST forecasts. The experimental results show that using the SST observational data and the NWP results to train the neural network model improves the SST prediction performance. The model performs best when the number of convolution kernel is tuned. Compared with the traditional CNN, LSTM model and persistence forecasting method, the optimized CNN model has better SST prediction in the testing set and faster learning speed through involving the normalized layer. The mean absolute errors of 0~24 h, 24~48 h, and 48~72 h forecasts are 0.276 °C, 0.412 °C and 0.487 °C, the corresponding root mean square errors are 0.217 °C, 0.396 °C and 0.567 °C, and the corresponding correlation coefficients are 0.994 7, 0.990 0, and 0.987 2, respectively.

Key words: sea temperature forecasting; numerical weather prediction; convolutional neural networks; normalized layer