

基于多模型平差接力的长时序风暴增水预测方法

杨彬¹, 艾波¹, 商杰^{2*}, 刘清容², 于梦超²

(1. 山东科技大学 测绘与空间信息学院, 山东 青岛 266590; 2. 自然资源部北海预报减灾中心, 山东 青岛 266061)

摘要: 利用神经网络模型预测未来风暴增水时, 预测精度会随预测时序的延伸不断降低。基于长短期记忆(Long Short-Term Memory)神经网络模型, 以风速、风向、气压和前时序的风暴增水数据作为模型输入, 利用多个模型接力预测风暴增水时间序列, 减小误差随模型的迭代累积, 建立基于多模型平差接力的长时序风暴增水预测方法。对比不同地点和不同台风下的风暴增水预测分析结果, 模型在渤海海北部区域每月的均方根误差为4~7 cm, 在黄海中部区域可控制在10 cm以内, 能够较准确预测未来24 h的风暴增水。

关键词: 风暴增水; 长短期记忆神经网络; 时间序列预测; 多模型平差接力

中图分类号: P731.34 **文献标识码:** A **文章编号:** 1003-0239(2023)02-0032-13

0 引言

风暴潮是一种由强烈的大气扰动(如热带气旋、温带气旋等)引起的海面异常升降的灾害性自然现象^[1-2]。若较大的风暴潮恰好与天文高潮相叠, 受其影响的海域会出现水位暴涨, 海水倒灌, 甚至泛滥成灾^[3-4]。据统计, 风暴潮已成为我国最严重的海洋灾害, 风暴增水也是最重要的致灾因子之一^[5-7]。2013年9月22日, 台风“天兔”在广东省汕尾市登陆并引发台风风暴潮, 造成全市225.25万人受灾, 直接经济损失达110.66亿元。2008—2017年, 我国因风暴潮造成的经济损失总额占海洋灾害经济损失总额的86%, 风暴潮也是造成我国经济损失最严重的海洋灾害^[8]。

目前风暴增水预测主要可分为经验预测方法、数值模式预测方法和机器学习方法3类^[9]。经验预测方法是根据预报员所掌握的专业知识和历史经验, 结合当前气压和风场实况进行主观判断和预报的方法。数值模式预测方法是在一定的初值和边

值条件下, 通过数值计算, 求解描写天气演变过程的流体力学和各类方程组, 预报未来网格化数据的方法^[10]。国内外学者建立了不同的增减水数值模式, 例如英国的“海模式”(Sea Model), 美国的SLOSH (Sea, Lake and Overland Surges from Hurricanes)模式, 荷兰的DSCM模式以及国内的反向传播模型等^[11]。机器学习是人工智能的核心, 可使计算机相对智能化, 可基于历史数据的学习并利用各种神经网络预测未来时间序列数据^[12]。朱佩京等^[6]基于随机森林模型对台风风暴潮进行预测, 分析了7个输入因子对台风风暴增水效果的影响, 为台风情景构建提供了重要依据。薛明等^[13]利用误差反向传播(Back-Propagation, BP)神经网络、小波神经网络、递归神经网络(Recurrent Neural Network, RNN)对2017年广东省阳江站风暴潮增水数据进行测试, 结果表明BP神经网络适用于多因子长时间预测, 小波神经网络更适合短时间预报, 递归神经网络在真实值与预报值的相关性方面表现更好。刘媛媛等^[12]利用2008—2017年深圳市赤湾潮位站

收稿日期: 2022-07-04; 修回日期: 2022-10-11。

基金项目: 国家自然科学基金面上项目(62071279); 山东科技大学科研创新团队支持计划资助(2019TDJH103)。

作者简介: 杨彬(1997-), 男, 硕士在读, 主要从事地理信息系统应用与开发研究工作。E-mail: yb18206550189@163.com

*通信作者: 商杰(1979-), 女, 高级工程师, 学士, 主要从事海洋监测、预报研究工作。E-mail: shangjie@ncs.mnr.gov.cn

的逐时潮位数据和赤湾站逐时气象数据,基于长短期记忆(Long Short-Term Memory, LSTM)神经网络,分析对比了不同因素的前时序数据对风暴潮的影响,结果表明前2 h的风速和风向过程、前7 h的潮位过程与未来1~3 h潮位的相关性最大,但预测的时间序列相对较短。直接利用神经网络模型进行预测,当预测时序只有几小时时精度相对较高,而当预测时序较长时精度则无法保证。本文重点研究如何调整LSTM神经网络模型,解决当预测时序较长时精度较低的问题。

本文通过直接预测未来24 h风暴增水时发现,预报的均方根误差会随着预测时间序列的延伸而不断增大,因此本文提出利用多个子模型进行平差接力预测的方法。从第一个子模型开始预测几小时后的风暴增水,随后在下一个子模型中将之前的输入要素加上前一子模型的预测数据作为输入要素,预测接下来几小时的风暴增水,期间不断进行误差纠正,最终完成未来24 h风暴增水预测,此方法称为基于多模型平差接力的长时序风暴增水预测。分析不同预测时序子模型之间的配合,选出最优的基于多模型的平差接力预测模型。经验证明,此方法减少了预测时长与预测精度之间的冲突,从而可以在预测时序较长的情况下,较大地提高风暴增水的预测精度。

1 数据与原理

1.1 数据提取

每年登陆我国的台风平均为7个左右,基本从

东南沿海登陆。风速和风向变化是引起风暴增水变化的主要原因,气压也是模型训练时不可忽略的影响因素。风速和风向是风暴增水产生的主导因素,向岸风会产生风暴增水,离岸风会产生风暴减水,风的大小决定增水大小^[14]。风暴增水也有一定的延迟性,会受前时序风暴增水过程的影响^[15]。本文研究区域为渤海和黄海中部,使用的部分站点数据包含风暴增水、风向、风速和表面气压。

风向、风速和表面气压数据提取于欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)的10 m纬向风分量、10 m经向风分量和海面气压的历史再分析格网数据,风暴增水数据使用了自然资源部北海预报减灾中心的风暴增水再分析数据。提取2011—2015年研究区域内包括成山头、芝罘岛、五号码头和岚山港的4个要素的逐时数据(见表1),其中2011—2014年数据作为训练集,2015年数据作为测试集。

1.2 模型原理

本文基于LSTM神经网络模型预测未来24 h风暴增水。LSTM神经网络是在RNN的基础上发展起来的^[16-17]。LSTM神经网络模型通过增加门限解决了RNN模型短期记忆的问题,使循环神经网络能够真正有效地利用长距离的时序信息^[18-19]。该模型拥有输入门、输出门、遗忘门3个逻辑控制单元,且各自连接到一个乘法元件上,通过设定神经网络的记忆单元与其他部分连接的边缘处的权值控制信息流的输入、输出及细胞单元的状态。该模型更擅长处理多个变量的问题,因此在实际研究中有广

表1 数据介绍

Tab.1 Description of data

| 数据名称 | 数据来源 | 来源类型 | 来源范围 | 数据分辨率 |
|------|--------------------|------|-----------------------------|-------|
| 风向 | ECMWF再分析10 m风场 | 格网数据 | 34°~41°N, 117.5°~125.5°E | 0.25° |
| 风速 | ECMWF再分析10 m风场 | 格网数据 | 34°~41°N, 117.5°~125.5°E | 0.25° |
| 表面气压 | ECMWF再分析表面压力 | 格网数据 | 34°~41°N, 117.5°~125.5°E | 0.25° |
| 风暴增水 | 自然资源部北海预报减灾中心再分析数据 | 站点数据 | 成山头、芝罘岛、 岚山港、五号码头 | — |

泛应用^[20-21]。LSTM神经网络模型内部结构首尾相接,同一层内前面单元的输出生作为后面单元的输

入,前一层的输出作为后一层的输入。模型结构如图1所示。

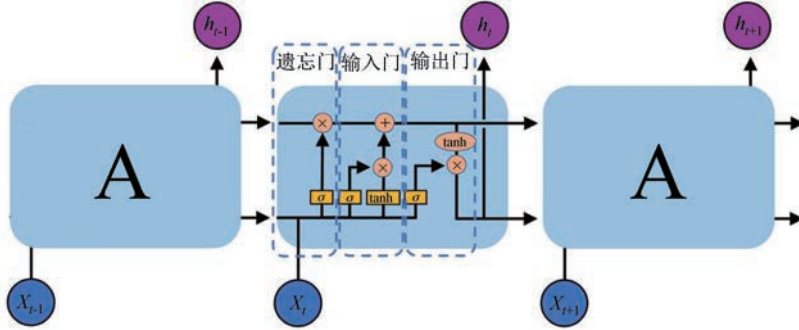


图1 LSTM结构图

Fig.1 Structure diagram of LSTM

训练模型属于LSTM神经网络模型中的多因素输入、单因素输出模型^[22]。每个子模型的输入时间步长为24 h,输出时间步长为24 h。神经网络的学习速率为0.1,隐藏层数、每层的神经元个数开始不做规定,需通过不断实验调试确定最优模型参数。

在进行数据预处理时,提取了2011—2015年威海成山头逐时风速、风向、气压和风暴增水影响因素数据,共53 824条。对样本进行归一化处理,这样可避免由于样本本身表达方式的原因导致的绝对数值中小数据被大数据“吃掉”的情况,可保证每个特征被平等对待^[13]。本实验将数据点映射到[0,1]区间。公式如下:

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

$$X_{\text{scaled}} = X_{\text{std}}(\max - \min) + \min \quad (2)$$

式中: X_{\min} 为一列中的最小值; X_{\max} 为一列中的最大值; \max 表示要映射到的区间最大值; \min 表示要映射到的区间最小值; X_{std} 为标准化结果; X_{scaled} 为归一化结果,数据在0~1之间^[23]。

分割数据集时,将数据样本按照0.8的比率划分成训练集和测试集。分割训练集和测试集的输入 X 和输出 Y ,并改造为LSTM神经网络模型的输出格式,即[samples, timesteps, features]^[24]。将数据放入定义好的模型中训练,迭代次数为200次,损失函数为平均绝对误差(Mean Absolute Error, MAE),优化器为自适应运动估计算法(Adaptive moments, Adam),可得出测试集的预测数据。

利用相关系数(Correlation Coefficient, CC)和均方根误差(Root Mean Square Error, RMSE)进行预测模型的评定。公式如下:

$$CC = \frac{\sum_{k=1}^n (y_k - \bar{y}_k)(\hat{y}_k - \bar{\hat{y}}_k)}{\sqrt{\sum_{k=1}^n (y_k - \bar{y}_k)^2 \sum_{k=1}^n (\hat{y}_k - \bar{\hat{y}}_k)^2}} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (4)$$

式中: y_k 表示网络预测输出序列; \hat{y}_k 表示再分析数据实际序列值; \bar{y}_k 表示预测输出序列的平均值; $\bar{\hat{y}}_k$ 表示实际序列的平均值; m 和 n 为预测风暴增水的数量; i 和 k 是变量表示序列中的位次; y_i 为实际预测数据; \hat{y}_i 为预测数据的平均值。CC的值越接近1,表示增水预测值与实际值序列之间的相关程度越大^[25]。

1.3 基于多模型平差接力的长时序风暴增水预测方法

直接预测24 h的风暴增水,预测结果的均方根误差会随时间序列的延后逐渐增大,造成后序预测精度持续降低,所以本文提出一种基于多模型的平差接力预测模型。通过训练单层LSTM神经网络模型预测24 h风暴增水,神经元个数从12开始依次递增4个,每小时的均方根误差会先变小再变大,找出其中误差最小的一个模型,在此模型基础上累加第二层LSTM神经网络,该层神经元个数同样从12开始依次递增4个,找出其中均方根误差最小的一个

双层模型。利用该方法可在预测未来 24 h、12 h 及 4 h 风暴增水的单双层 LSTM 神经网络模型中找出最优参数的模型作为子模型,每次只提取子模型预测的前 4 个时刻的数据,以此结合前序风暴增水时间序列数据作为新的子模型输入数据,进一步预测之后几个时刻的风暴增水,直至得出 24 h 的预测数据。在该方法中,风暴增水时间序列作为预测因素会不断地后移,但风速、风向和气压时间序列是不动的,所以风暴增水时间序列每前进一次,风暴增水和另外 3 个预测因素组成的预测数据集的时间序列排列是不同的组合,所以在预测过程中需要使用不同组合训练的子模型来输入数据。根据均方根误差在前几小时增长速度快的特点,将子模型的预报时间暂定为 4 h。

将第一个子模型命名为 pre4h_R4,将 $(t-24)$ — $(t-1)$ 时刻的风速、风向、气压和风暴增水数据输入该子模型中,预测 t — $(t+3)$ 时刻的风暴增水数据。选出前 4 h 的预测数据加到已有风暴增水数据中,这样风速、风向和气压 3 个影响因素时间序列和风暴增水时间序列相差 4 h,需重新训练子模型来继续预测 $(t+4)$ — $(t+7)$ 时刻的风暴增水数据。该子模型沿用了 pre4h_R4 模型各参数,被命名为 pre4h_R8。在 pre4h_R8 模型中输入数据为 $(t-24)$ — $(t-1)$ 时刻的风速、风向和气压 3 个影响因素序列和 $(t-20)$ — $(t+3)$ 时刻的风暴增水序列数据集,可得到 $(t+4)$ — $(t+7)$ 时刻的风暴增水。这样依次进行 4 h 的接力预测,经过 6 次预测最终可以得到 t — $(t+23)$ 时刻的风

暴增水数据。

由多模型接力预测得出的数据误差会通过每次模型迭代后累积,这也是导致预测误差逐渐增大的重要原因。本文在此基础上改进实验,结合实验误差特点,提出两种基于误差纠正的平差方法。方法一是将预测值统一加上子模型测试集所有样本真实值与预测值之差的均值,方法二是将预测值加上模型测试集所有样本对应时刻真实值与预测值之差的均值。两个方法的公式如下:

方法一:

$$p_e(t) = p_s(t) + \frac{1}{24T} \sum_{i=1}^{24} \sum_{t=1}^T (p_i(t) - \hat{p}_i(t)) \quad (5)$$

方法二:

$$p_e(t) = p_s(t) + \frac{1}{T} \sum_{i=1}^T (p_i(t) - \hat{p}_i(t)) \quad (6)$$

式中: $p_e(t)$ 表示误差纠正后在 t 时刻的预测值; $p_s(t)$ 表示误差纠正前的预测值; T 表示 t 时刻模型测试集中的预测值数量; i 表示序列中的位次; $p_i(t)$ 表示 t 时刻测试集中第 i 个预测值; $\hat{p}_i(t)$ 表示 t 时刻测试集中第 i 个真实值。

经误差平差后,第一个模型的预测数据误差变小,因此输入下一个模型后预测的数据误差本身会变小,经过平差后误差会更小,继续进入下一个模型,直到预测得到未来 24 h 的风暴增水。将此模型命名为 preRoll_4h,流程如图 2 所示。

由于基于多模型的平差接力预测需要多个子模型进行配合,虽然每个子模型在预测时序较短时误差较小,但是用到的子模型会较多;而当预测时

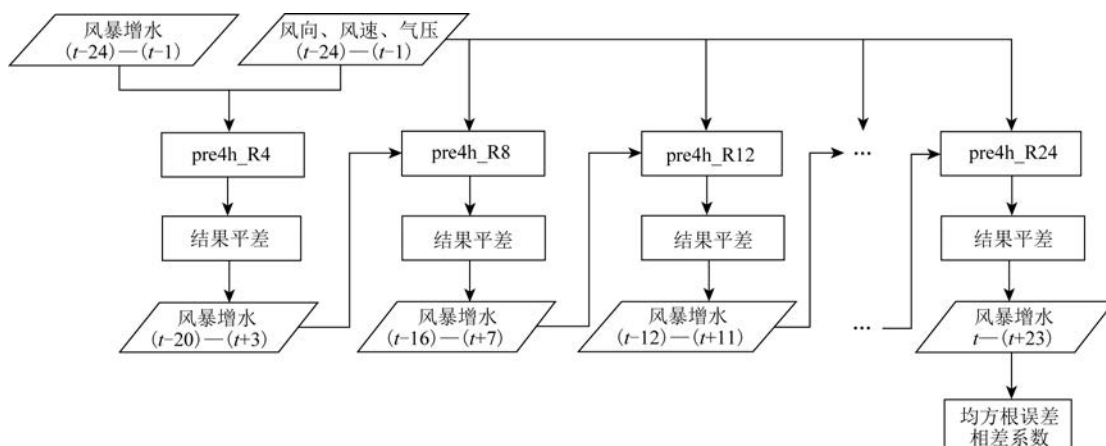


图2 preRoll_4h模型流程图

Fig.2 The flow chart of preRoll_4h model

序较长时虽然误差较大,但用到的子模型会较少,这就导致子模型预测时长和所需数量之间存在冲突。继续测试子模型在不同预测时长和不同数量之间配合预测的精度,基于多模型的平差接力模型预测未来 2 h、4 h、8 h 和 12 h 的风暴增水并进行比较,最终得出最优预测模型。模型训练及评价过程见图 3。

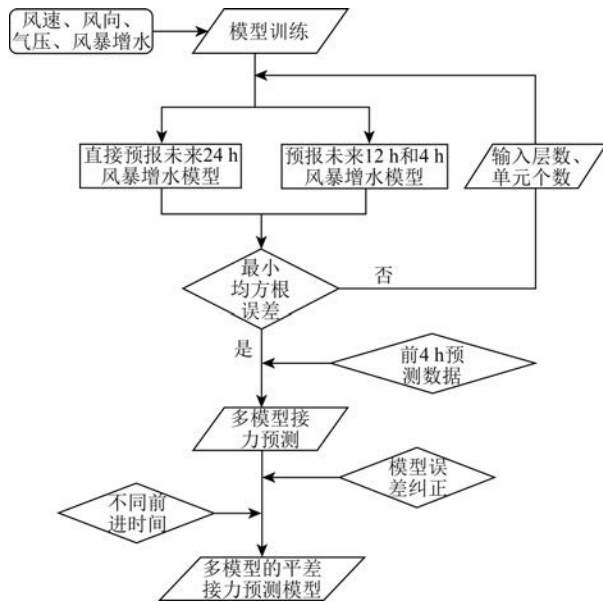


图3 实验过程

Fig.3 Experimental procedure

2 实验分析

2.1 神经网络 24 h 预测

提取 2011—2015 年山东威海成山头的风速、风向、表面气压及风暴增水数据。训练单 LSTM 层神经网络模型,输入 $(t-24) \sim (t-1)$ 时刻各因素数据,输出 $t \sim (t+23)$ 时刻的风暴增水数据。将模型 LSTM 层定义为一层,共训练 8 个单层 LSTM 神经网络模型。各模型命名见表 2。

pre24h_f40 模型的损失函数测试集虽然在初期下降较快,但在后期一直不稳定,该过程模型没有逐渐收敛^[26];其他模型的训练集损失函数下降速度趋于平缓,模型逐渐收敛^[27]。因此,计算除 pre24h_f40 模型以外的 7 个模型测试集所有样本 24 h 中每小时的均方根误差及预测数据和真实数据

表 2 单 LSTM 层模型名称

Tab.2 Model name of single LSTM layer

| 神经元个数 | 模型名称 |
|-------|------------|
| 12 | pre24h_f12 |
| 16 | pre24h_f16 |
| 20 | pre24h_f20 |
| 24 | pre24h_f24 |
| 28 | pre24h_f28 |
| 32 | pre24h_f32 |
| 36 | pre24h_f36 |
| 40 | pre24h_f40 |

的相关系数。

各模型预测精度存在差距,均方根误差越大则相关系数越小。综合选出均方根误差小且相关系数大的单层模型 pre24h_f32,在此基础上再加一层 LSTM 层预测未来 24 h 风暴增水。共训练 8 个双 LSTM 层神经网络模型,再继续得出两层 LSTM 层的最优模型。

单层和两层模型在前几个时刻的预测都是相对较准确的,均方根误差在 5 cm 以内,但是随着时间序列的不断延伸,均方根误差逐渐增大,最大可达到 10 cm 左右,相关系数逐渐减小,最小达到 0.8 左右,因此需对预测模型作进一步改进,使均方根误差和相关系数尽可能稳定。

2.2 基于多模型的平差接力预测

在直接预测 24 h 风暴增水时,每小时预测结果的均方根误差会逐渐增大,因此预测精度会随时间序列持续降低,所以本文提出一种基于多模型的平差接力预测模型。首先需要获得前 4 h 预测结果均方根误差最小的模型,我们分别从 24 h、12 h、4 h 风暴增水预测最优模型中选出前 4 h 预测结果均方根误差最小的 pre4h_f36_s28 模型。

为尽可能减少误差的累积,在完成子模型预测后会对数据进行纠正。将未平差和经两种方法平差后多模型接力预测的 2015 年 5—12 月风暴增水的月均方根误差和相关系数结果进行对比,结果见图 4。两者的对比结果表明,经方法二得到的预测结果的均方根误差最小,相关系数最大,未平差的预测结果的均方根误差最大,相关系数最小,所以

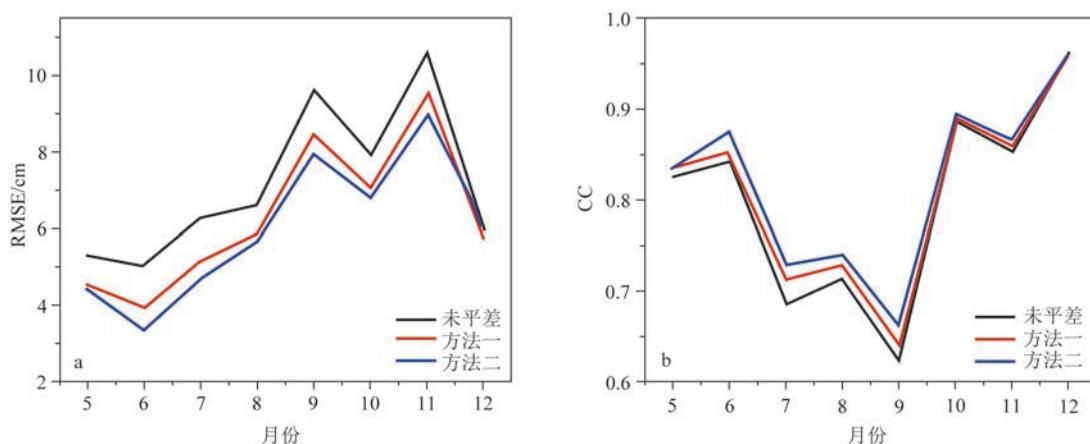


图4 平差前后均方根误差(a)和相关系数(b)对比

Fig.4 Comparison of root mean square error (a) and correlation coefficient (b) before and after adjustment

方法二的基于多模型的平差接力预测模型 preRoll_4h 能进一步提高预测精度。

应用 preRoll_4h 模型,利用增水变化幅度较大的 2015 年 12 月 1—30 日的 4 个预测因素数据集预测了 2015 年 12 月 2—31 日的风暴增水数据。和真实数据的对比结果表明,相较 pre24h_f32 模型,preRoll_4h 模型预测的风暴增水曲线和真实曲线的相似度更高。preRoll_4h 模型预测的 2015 年 5—12 月风暴增水的月均方根误差均值为 5.985 cm,pre24h_f32 模型为 6.298 cm,所以 preRoll_4h 模型预测精度更高。

2.3 不同前进时间分析

在基于多模型的平差接力预测模型中,子模型的预测时长和数量之间的不同配合可能会使得预测精度不同。继续实验预测 2 h、8 h 和 12 h 风暴增水的子模型,其中预测 8 h 和 12 h 的子模型需要用到预测未来 12 h 的模型 pre12h_f32 作为第一个子模型,利用其参数继续训练其他子模型。4 种模型如表 3 所示。

表 3 不同前进时间模型

Tab.3 Different advance time models

| 模型名称 | 前进时间/h | 基本参数模型 | 子模型数量/个 |
|-------------|--------|---------------|---------|
| preRoll_2h | 2 | pre4h_f36_s28 | 12 |
| preRoll_4h | 4 | pre4h_f36_s28 | 6 |
| preRoll_8h | 8 | pre12h_f32 | 3 |
| preRoll_12h | 12 | pre12h_f32 | 2 |

为比较 4 种模型的精度,我们分别预测了 2015 年 5—12 月各 30 d 的风暴增水,求出每月和每日均方根误差和相关系数的均值。4 种方案的对比结果表明,2015 年 5 月和 6 月 preRoll_8h 方案预测结果的月相关系数较小,月均方根误差最大,日均方根误差平均值最大。2015 年 7—12 月 preRoll_8h 方案预测结果的月相关系数和日相关系数平均值基本在较高水平,月均方根误差和日均方根误差的平均值基本为最小。通过实验可以得出,子模型预测时段的不同和最终预测的精度有一定关系,综合考虑后选择 preRoll_8h 模型预测未来 24 h 风暴增水,预测结果见图 5。

从图 6 可以看出,preRoll_8h 模型比初始 pre24h_f32 模型和 preRoll_4h 模型预测的 2015 年 5—12 月的月均方根误差以及每小时的均方根误差都要小,随着时序的后移,上述趋势更加明显。preRoll_8h 模型比另外两个模型的月相关系数和每小时的相关系数都要高,所以 preRoll_8h 模型的预测精度更高。

当前大部分研究得到的风暴增水预测时序只有几小时,且误差较大。为比较分析不同类型模型时序风暴增水预测结果,本文将 preRoll_8h 模型与 BP 神经网络和递归神经网络在学习率为 0.05、隐含层节点数为 9 的情况下进行对比。预测结果表明,preRoll_8h 模型对未来 24 h 风暴增水的预测效果有明显提升。利用 3 种模型对成山头 24 h 风暴增水进行预测,对比结果表明(见图 7),preRoll_8h 模型的

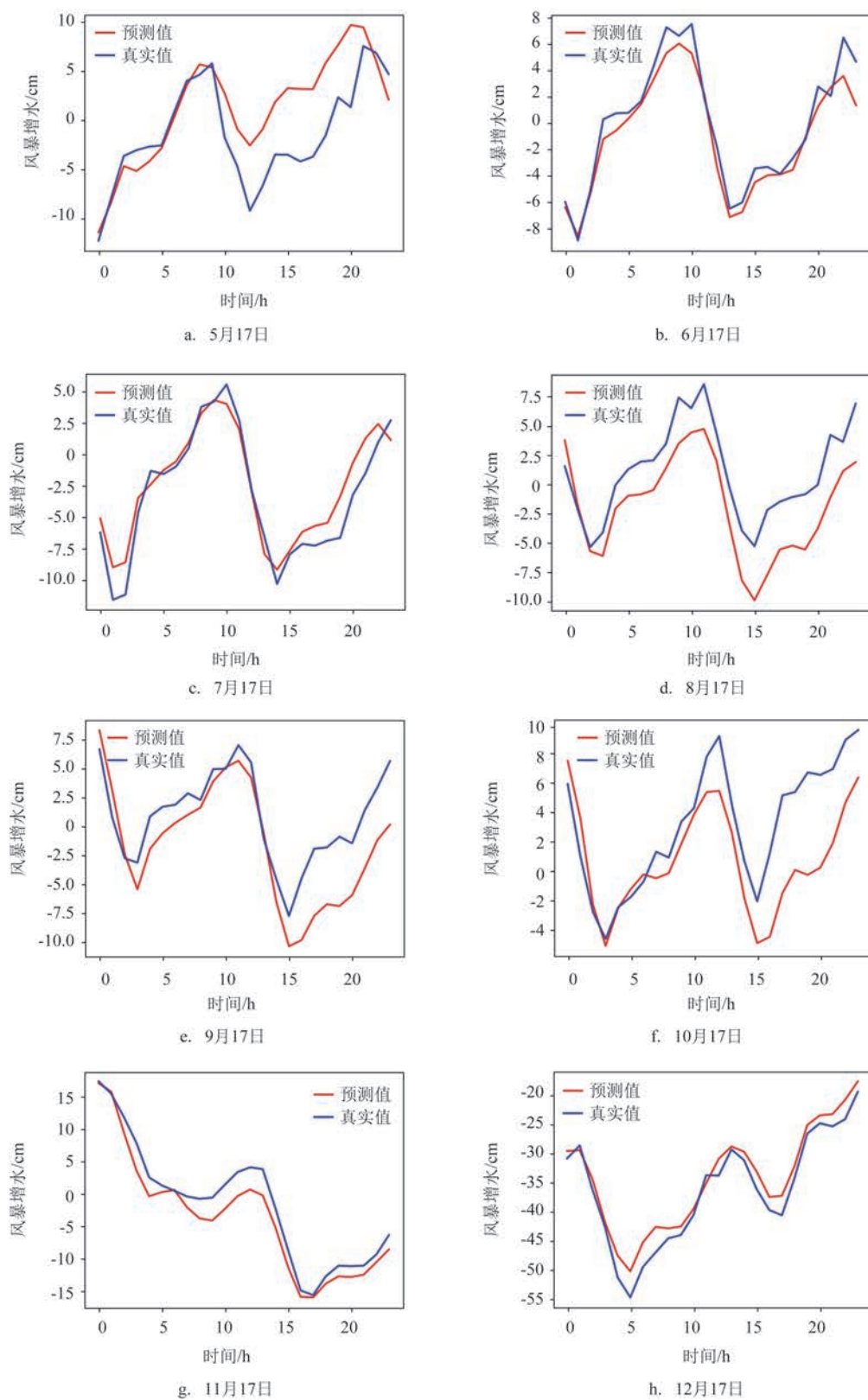


图5 preRoll_8h模型预测24 h风暴增水效果图

Fig.5 The effect of 24 h storm water increase predicted by the preRoll_8h model

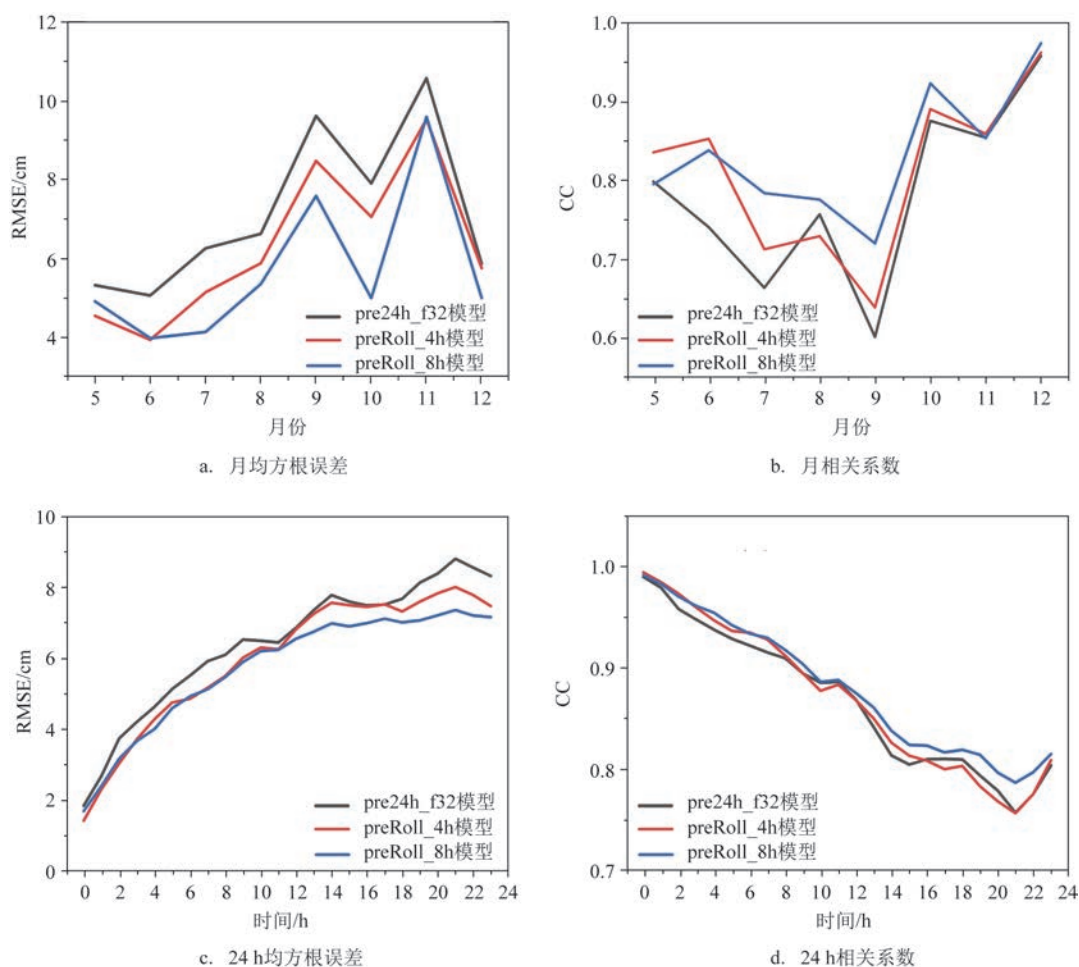


图6 3个预测实验模型误差对比

Fig.6 Error comparison of the three prediction experimental models

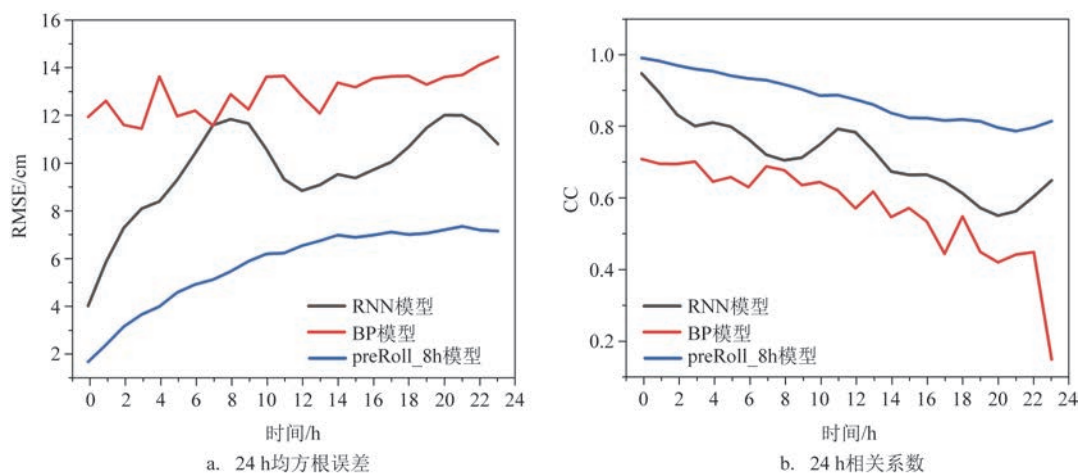


图7 模型预测精度对比图

Fig.7 The comparison of model prediction accuracy

均方根误差和相关系数指标均有明显优势,比其他两类模型预测精度更高。本文基于多个子模型进行接力预测,期间进行误差纠正,并选出最佳接力时段,有效提升了临近时序风暴增水的预测精度。

3 模型应用

3.1 不同地点模型应用

风暴增水的最终大小不仅由风速、风向和气压决定,同时也受岸线形状和海底地形影响^[28-29]。由于不同地点的岸线方向和海底地形不同,因此即使在相同的风速、风向和气压的作用下产生的风暴增水也有所不同^[30-31]。

选取不同地点的历史数据进行模型训练。为尽可能开展跨度大的测试,本研究提取2011—2015年山东半岛南北两侧岚山港、五号码头和芝罘岛的数据进行实验,其中2011—2014年数据作为训练集,2015年数据作为测试集。计算了3个站点利用初始模型和最终模型预测得到的2015年5—10月风暴增水的每小时均方根误差和每月均方根误差

结果并进行比较,结果见图8。

在3个站点中,preRoll_8h模型比pre24h_f32模型每小时的均方根误差小,可以较为有效地改进误差随时间序列的增大而增大的问题。preRoll_8h模型在除10月外的月均方根误差全部小于pre24h_f32模型(五号码头除外),五号码头和芝罘岛的月均方根误差全部在10 cm以下,岚山港、五号码头、芝罘岛的均方根误差依次从大到小。这主要由地理位置的南北差异导致的,越靠南受热带气旋等气候的影响越大,所以风暴增水整体起伏较大。均方根误差最大的岚山港整体月均方根误差在10 cm左右。通过3个站点每小时和每月的均方根误差对比可证明,preRoll_8h模型比pre24h_f32模型的预测精度有所提高,也可以较好地适应不同地点的风暴增水预测。

3.2 台风时期风暴增水预测

台风是加剧风暴增水的最主要原因。为了验证模型的通用性与准确性,将模型应用于台风发生时的风暴增水预测,并与真实数据做对比,对比结

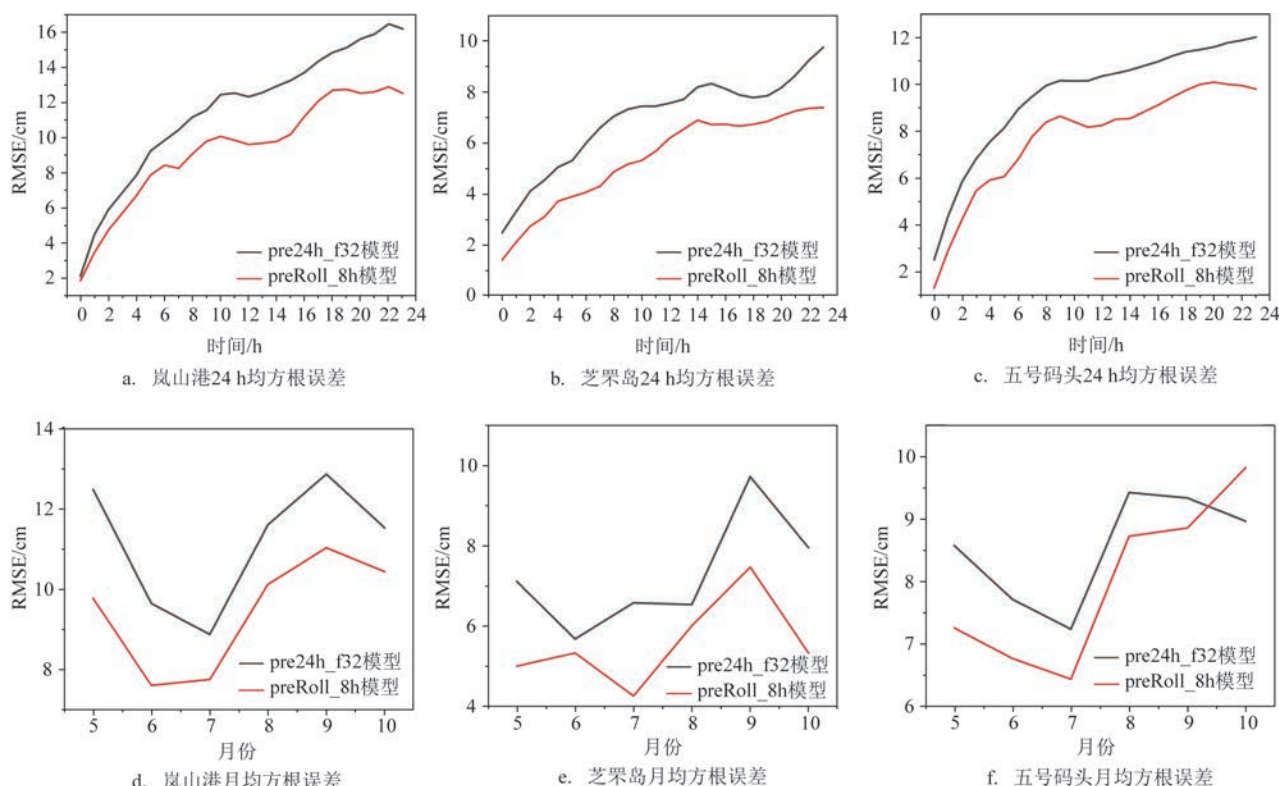


图8 不同地点对比图

Fig.8 Comparison of predictions at different locations

果是判断模型预测能力的重要指标^[32-33]。

研究区域内台风发生次数较少,我们通过中央气象台台风网提取台风数据,在2011年6月台风“米雷”经过研究区域,因此选取距离台风路线较近的成山头 and 芝罘岛站进行验证。提取2012—2015年两个站点的风速、风向、气压和风暴增水数据作为训练集输入preRoll_8h模型中,利用2011年6月的数据验证台风发生时preRoll_8h模型的预测能力。

从图9可以看出,2011年6月两个站点的均方根误差都比较稳定(26—27日台风“米雷”经过时除外),误差在6 cm以下。台风时预测结果的均方根误差较大,成山头达到了9 cm左右,芝罘岛达到了12 cm左右。成山头的风暴增水在6月27日07时(北京时,下同)左右达到最大值,预测值比真实值低10 cm左右,芝罘岛的风暴增水在6月27日05时左右达到最大值,预测值比真实值低3 cm左右。虽然台风时期的预报值和真实值的极值存在偏差,但

预测的整体趋势是相符合的,极值差也在合理范围内。

在台风经过时,preRoll_8h模型的整体预测趋势与真实值相符,但数值差距有所增大。主要原因是研究区域内台风等热带风暴发生次数较少,训练时偶尔的台风数据会被当成错误数据遗忘,所以模型没学习到由异常天气造成的增水剧烈变化的情况,但预测值与真实值在均方根误差和具体时刻数据对比中的差距都在合理范围之内。

4 结论

本文在选出4个预测因素后,分别进行了3个阶段实验。第一阶段为24 h风暴增水预测实验,分别训练LSTM层为单层和双层的模型,结果表明LSTM层为单层32个神经元的pre24h_f32模型预测24 h的风暴增水相对更精确,但随着时间序列的延

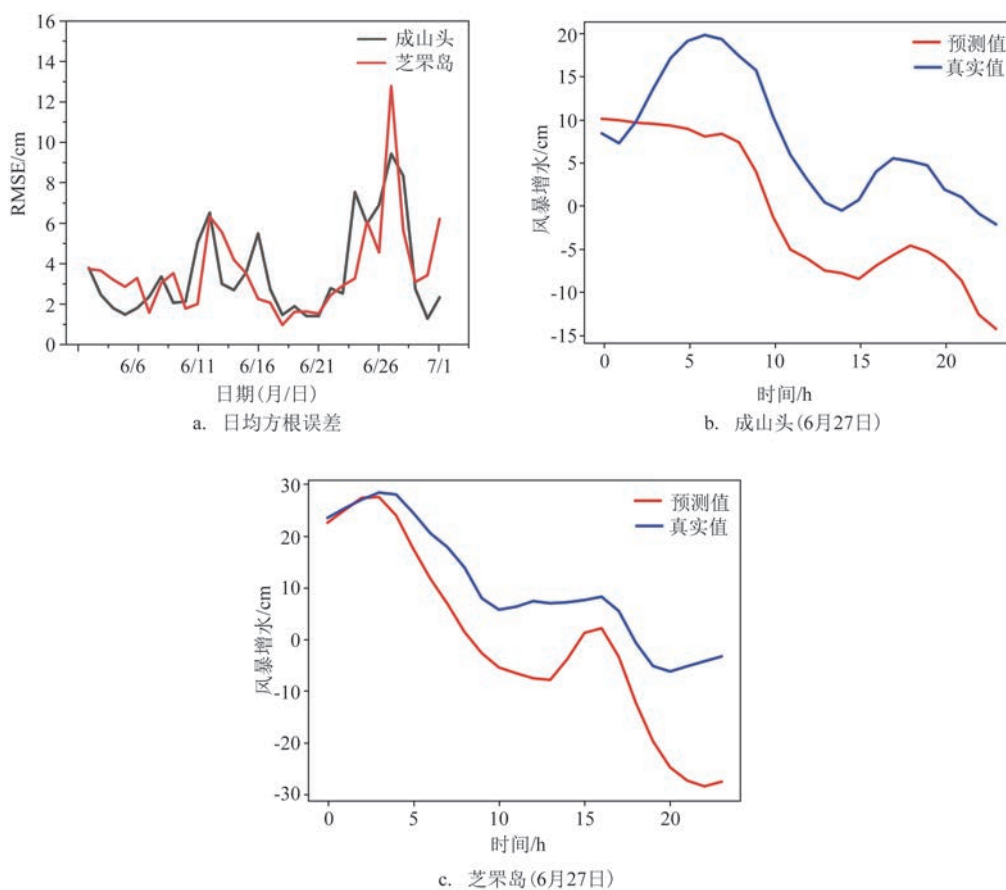


图9 台风时预测精度对比图

Fig.9 Comparison of prediction accuracy during typhoon

伸,越到后期预测误差越大。第二阶段为基于多模型的平差接力预测实验,从子模型预测数据中选出风暴增水时间序列之后4 h的预测数据作为下次输入数据,继续预测之后4 h的风暴增水数据,考虑在多模型接力预测中数据误差会通过每次模型预测迭代后累积,因此提出基于多模型的平差接力预测模型preRoll_4h,预测精度有了进一步提高。第三阶段为不同前进时间段预测实验,对每次前进2 h、4 h、8 h和12 h的风暴增水预测结果进行对比,包括预测的每月相关系数和均方根误差,最终得出基于多模型的平差接力模型preRoll_8h的预测精度最高。

利用preRoll_8h模型对2015年5—12月成山头站的风暴增水进行预测,每小时的均方根误差均值从pre24h_f32模型的6.4 cm降到了5.7 cm,每月均方根误差均值从7.17 cm减小到5.69 cm,相比其他神经网络模型,preRoll_8h模型的精度也是最高的。验证表明,在不同地点的风暴增水预测结果中,preRoll_8h模型比pre24h_f32模型每小时的均方根误差和每月均方根误差都小。在研究区域中部的芝罘岛,preRoll_8h模型的月均方根误差为4~7 cm,在南部的岚山港可控制在10 cm以内,因此相比于南部海域,北部海域的风暴增水较为稳定,更适合此方案。台风发生时的风暴增水变化较快且较大,preRoll_8h模型的预测值和真实值的极值时间较吻合,预测误差也在可接受的合理范围内,所以该模型在台风发生时也具有一定的适用性^[34]。本模型能够较准确地预测未来24 h的风暴增水,切实改进了风暴增水预测随时间序列延伸误差增大的问题,以此数据为基础开展防灾减灾工作部署,能够避免造成大规模的经济损失。如何采取有效措施继续提高台风风暴增水的预测准确性,是未来研究的重点工作。

参考文献:

- [1] JIN Y D, ZHAO X, MA L N. Storm surge characteristics and extreme parameters in the Chengshantou sea area[J]. IOP Conference Series: Earth and Environmental Science, 2017, 82: 012052.
- [2] 谭家辉,潘毅,于普兵,等.江苏沿海台风风暴潮过程中的波浪增水分布研究[J].海洋预报,2022,39(1): 48-55.
TAN J H, PAN Y, YU P B, et al. Study on the distribution of wave-induced setup along the coast of Jiangsu Province during typhoon processes[J]. Marine Forecasts, 2022, 39(1): 48-55.
- [3] WANG B, LIU S C, WANG B, et al. Multi-step ahead short-term predictions of storm surge level using CNN and LSTM network[J]. Acta Oceanologica Sinica, 2021, 40(11): 104-118.
- [4] 黄宝霞,胡金磊,郑淑贤,等.广东省徐闻沿海风暴潮数值模拟与淹没危险性分析[J].海洋预报,2022,39(1): 39-47.
HUANG B X, HU J L, ZHENG S X, et al. Numerical simulation and inundation risk assessment of storm surge along the coastal areas of Xuwen, Guangdong Province[J]. Marine Forecasts, 2022, 39(1): 39-47.
- [5] LIU Q R, RUAN C Q, ZHONG S, et al. Risk assessment of storm surge disaster based on numerical models and remote sensing[J]. International Journal of Applied Earth Observation and Geoinformation, 2018, 68: 20-30.
- [6] 朱佩京,罗年学,赵前胜.基于随机森林模型的台风风暴潮最大增水预测[J].测绘通报,2021(12): 71-74, 82.
ZHU P J, LUO N X, ZHAO Q S. Forecast of maximum water increase in typhoon storm surge based on random forest model[J]. Bulletin of Surveying and Mapping, 2021(12): 71-74, 82.
- [7] 高强,谷文凯,林亚琼,等.山东省风暴潮灾害经济损失预测——基于灰色一周期外延组合模型[J].海洋经济,2016,6(1): 46-51.
GAO Q, GU W K, LIN Y Q, et al. The financial loss prediction of storm surge in Shandong province based on the gray-periodic extensional combinatorial model[J]. Marine Economy, 2016, 6(1): 46-51.
- [8] 李程,王慧,李响,等.近10年福建省风暴潮灾害特征分析及社会经济影响[J].海洋经济,2019,9(1): 43-46.
LI C, WANG H, LI X, et al. Characteristics of storm surge disasters along Fujian coast in recent 10 Years[J]. Marine Economy, 2019, 9(1): 43-46.
- [9] 林林茂.基于ADCIRC模式的台风风暴潮数值模拟与危险性分析[D].哈尔滨:哈尔滨工业大学,2017.
LIN L M. Numerical simulation and hazard analysis of typhoon storm surge based on ADCIRC model[D]. Harbin: Harbin Institute of Technology, 2017.
- [10] HAN L, CHEN M X, CHEN K K, et al. A deep learning method for bias correction of ECMWF 24-240 h forecasts[J]. Advances in Atmospheric Sciences, 2021, 38(9): 1444-1459.
- [11] 郝婧,刘强.基于SSA-ELM模型的台风风暴潮灾害损失预评估[J].海洋科学,2022,46(2): 55-63.
HAO J, LIU Q. Pre-assessment of typhoon storm surge disaster loss based on the SSA-ELM model[J]. Marine Sciences, 2022, 46(2): 55-63.
- [12] 刘媛媛,张丽,李磊,等.基于多变量LSTM神经网络模型的风暴潮临近预报[J].海洋通报,2020,39(6): 689-694.
LIU Y Y, ZHANG L, LI L, et al. Storm surge nowcasting based on multivariable LSTM neural network model[J]. Marine Science Bulletin, 2020, 39(6): 689-694.

- [13] 薛明, 李醒飞, 成方林. 基于多种神经网络的风暴潮增水预测方法的比较分析[J]. 海洋通报, 2019, 38(3): 290-295.
- XUE M, LI X F, CHENG F L. Comparative analysis of storm surge water prediction methods based on multiple neural networks [J]. Marine Science Bulletin, 2019, 38(3): 290-295.
- [14] 范书鸣, 储廉, 蒋勤. 再分析风场修正及其在渤海湾典型温带风暴潮模拟中的应用[J]. 海洋预报, 2021, 38(4): 61-68.
- FAN S M, CHU A, JIANG Q. Optimization of reanalysis wind field and its application in the simulation of extratropical storm surge in the Bohai Bay[J]. Marine Forecasts, 2021, 38(4): 61-68.
- [15] CHEN K, KUANG C P, WANG L, et al. Storm surge prediction based on long short-term memory neural network in the East China Sea[J]. Applied Sciences, 2022, 12(1): 181.
- [16] CHEN S, GE L. Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction[J]. Quantitative Finance, 2019, 19(9): 1507-1515.
- [17] LAM M W Y, CHEN X, HU S K, et al. Gaussian process Lstm recurrent neural network language models for speech recognition [C]//2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton: IEEE, 2019: 7235-7239.
- [18] YU Y, SI X S, HU C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [19] TIGGELOVEN T, COUASNON A, VAN STRAATEN C, et al. Exploring deep learning capabilities for surge predictions in coastal areas[J]. Scientific Reports, 2021, 11(1): 17224.
- [20] ISHIDA K, TSUJIMOTO G, ERCAN A, et al. Hourly-scale coastal sea level modeling in a changing climate using long short-term memory neural network[J]. Science of the Total Environment, 2020, 720: 137613.
- [21] 吕忻, 丁骏. 基于深度学习的潮位预报订正技术研究[J]. 海洋预报, 2022, 39(2): 70-79.
- LYU X, DING J. Study on the correction technology of tide level forecast based on deep learning[J]. Marine Forecasts, 2022, 39(2): 70-79.
- [22] CHEN D W, ZHANG J H, JIANG S X. Forecasting the short-term metro ridership with seasonal and trend decomposition using loess and LSTM neural networks[J]. IEEE Access, 2020, 8: 91181-91187.
- [23] 余璇, 石绥祥, 徐凌宇, 等. 基于深度学习的赤潮发生预报方法研究[J]. 海洋通报, 2021, 40(5): 566-577.
- YU X, SHI S X, XU L Y, et al. Research on red tide occurrence forecast based on deep learning[J]. Marine Science Bulletin, 2021, 40(5): 566-577.
- [24] 董绍江, 裴雪武, 吴文亮, 等. 基于多层降噪技术及改进卷积神经网络的滚动轴承故障诊断方法[J]. 机械工程学报, 2021, 57(1): 148-156.
- DONG S J, PEI X W, WU W L, et al. Rolling bearing fault diagnosis method based on multilayer noise reduction technology and improved convolutional neural network[J]. Journal of Mechanical Engineering, 2021, 57(1): 148-156.
- [25] 李杰, 谭晓璇, 牛福新, 等. 天津沿海风暴潮特征及其与环境要素的关系[J]. 天津航海, 2021(1): 50-54.
- LI J, TAN X J, NIU F X, et al. The features of Tianjin coastal storm surge and its relations with the environmental elements[J]. Tianjin of Navigation, 2021(1): 50-54.
- [26] JIA F, YANG B L. Forecasting volatility of stock index: deep learning model with likelihood-based loss function[J]. Complexity, 2021, 2021: 5511802.
- [27] 代小红, 陈华江, 朱超平. 一种基于改进 Faster RCNN 的金属材料工件表面缺陷检测与实现研究[J]. 表面技术, 2020, 49(10): 362-371.
- DAI X H, CHEN H J, ZHU C P. Surface defect detection and realization of metal workpiece based on improved faster RCNN [J]. Surface Technology, 2020, 49(10): 362-371.
- [28] VAN ORMONT M, VAN DONGEREN A, ROELVINK D. A semi-empirical method for computing storm surges on open coasts during tropical cyclones[J]. Coastal Engineering, 2021, 165: 103839.
- [29] 杨万康, 杨青莹, 伊小飞, 等. 典型海湾风暴潮增水特征与机理研究[J]. 海洋通报, 2017, 36(5): 532-537.
- YANG W K, YANG Q Y, YI X F, et al. Characteristics and mechanism of storm surge in typical bay[J]. Marine Science Bulletin, 2017, 36(5): 532-537.
- [30] ENRÍQUEZ A R, WAHL T, MARCOS M, et al. Spatial footprints of storm surges along the global coastlines[J]. Journal of Geophysical Research: Oceans, 2020, 125(9): e2020JC016367.
- [31] JI T, LI G S, LIU Y X Y, et al. Spatiotemporal features of storm surge activity and its response to climate change in the Southeastern coastal area of China in the Past 60 years[J]. Journal of Geophysical Research: Atmospheres, 2021, 126(4): e2020JD033234.
- [32] HEO K Y, HA T, PARK K S. The effects of a typhoon-induced oceanic cold wake on typhoon intensity and typhoon-induced ocean waves[J]. Journal of Hydro-Environment Research, 2017, 14: 61-75.
- [33] 李健, 刘清容, 阮成卿, 等. 南海风暴潮对风场和外围海水敏感性研究[J]. 海洋科学, 2018, 42(2): 99-107.
- LI J, LIU Q R, RUAN C Q, et al. Study of storm surge sensitivity with typhoon and external seawater in the South China Sea[J]. Marine Sciences, 2018, 42(2): 99-107.
- [34] FENG X R, LI M J, YIN B S, et al. Study of storm surge trends in typhoon-prone coastal areas based on observations and surge-wave coupled simulations[J]. International Journal of Applied Earth Observation and Geoinformation, 2018, 68: 272-278.

Prediction method of long-term storm surge based on multi-model gradual adjustment

YANG Bin¹, AI Bo¹, SHANG Jie^{2*}, LIU Qingrong², YU Mengchao²

(1. College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China; 2. North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao 266061, China)

Abstract: When predicting storm surge using classical neural network models, the prediction accuracy will continue to decrease with the extension of the time series. To alleviate this problem, we proposed a method of long-term storm surge prediction based on multi-model gradual adjustment. The method is based on the Long Short-Term Memory neural network model, and integrates multiple models to predict the storm surge time series. Data from previous time periods were fed into the model, such as wind speed, wind direction, air pressure, and storm surge data, and the errors accumulated with model iterations were considered. According to the forecast analysis of storm surge in different locations and under different typhoons, the monthly root mean square error of the prediction model is between 4 and 7 cm in the northern part of the Bohai sea and Yellow Sea, and is less than 10 cm in the central part of the Yellow Sea, which indicate the model can predict storm surge in the next 24 hours accurately.

Key words: storm surge; Long Short-Term Memory neural network; time series prediction; multi-model gradual adjustment