

基于CNN-LSTM的珠江河口台风过程实时滚动修正预报

邓志弘¹, 刘丙军^{1,2*}, 张卡¹, 胡仕焜¹, 曾慧³, 张明珠³, 李丹³

(1. 中山大学土木工程学院, 广东 珠海 519085; 2. 中山大学水资源与环境研究中心, 广东 广州 510275; 3. 广州市水务科学研究所, 广东 广州 510220)

摘 要: 为改善台风预报精度, 基于实时滚动修正预报思路, 利用卷积神经网络嵌套长短期记忆神经网络(CNN-LSTM)和误差校正(EC)技术, 搭建了珠江河口台风实时预报模型。研究结果表明:“滚动预报”比单次预报有更好的路径和强度预报效果, 随着模型滚动时间的延长, 预报整体精度有逐渐改善的趋势。路径预报结果的均方根误差比单次预报减小了25.67%, 强度预报结果的平均绝对误差比单次预报减小了65.04%; 考虑误差校正的CNN-LSTM-EC的路径、强度“滚动预报”效果均优于CNN-LSTM, 前者的路径预报误差较后者减小了22.57%, 强度预报误差减小2.5%。

关键词: 实时滚动预报; 台风; 珠江河口; 深度学习; 误差校正

中图分类号: P457.8 **文献标识码:** A **文章编号:** 1003-0239(2024)01-0094-10

0 引言

台风是河口海岸地区最频繁的自然灾害之一。在全球气候变暖的背景下, 台风持续时间和强度明显增加^[1], 且强台风和超强台风所占比例逐渐增大, 这将引发更加严重的风暴潮灾害^[2-3]。珠江河口地区作为我国人均GDP最高、经济实力最强的地区之一, 位于亚热带地区大陆东岸, 邻近热带海洋, 多年来常受台风侵袭^[4-5]。准确预测珠江河口地区的台风轨迹和强度, 对台风登陆风险评估及其后续安全防范具有重要作用。

机器学习算法能在建模中较好地处理变量间的非线性关系, 已有不少研究者将其应用于台风预报领域^[6]。曹祥村等^[7]和ALI等^[8]分别基于多层前馈神经网络(Back Propagation Neural Network, BPNN)和多层感知机(Multi-Layer Perceptron, MLP)搭建台风路径预报模型, 证明了机器学习在台风预报中的可行性; ALEMANY等^[9]、KORD-MAHALLEH等^[10]和PAN等^[11]将台风运动视为时间

序列, 将循环神经网络(Recurrent Neural Network, RNN)融进预报模型中, 在提高预报精度的同时延长了预见期。为减少模型预报随着预见期延长可能出现的累积误差问题, 误差校正(Error Correction, EC)技术被广泛应用于台风过程实时预报, 如GUO等^[12]使用EC技术对2009年和2010年的台风轨迹进行实时校正预测, 结果表明经过EC校正的短时间预测效果优于未校正模型。

当前, 基于机器学习的台风预报模型已通过改善学习算法、EC技术等手段获得了更高的预报精度, 但台风过程变化的影响因素很多, 如海气相互作用和大尺度环流因素等, 这些难以解释的因素会给模型预测带来较大的不确定性^[13-14]。实时滚动修正预报(以下简称“滚动预报”)能随着时间推移, 不断利用实时数据作为模型输入来更新预报对象的未来趋势, 可有效防止长期预测造成的误差累积, 从而在满足时效性的同时提升预报精度^[15]。LIU等^[16]基于长短期记忆神经(Long Short Term Memory, LSTM)网络建立了一种河流短期水位滚动预测方

收稿日期: 2022-11-15。

基金项目: 广州市水务科技项目(GZSWKJ-2020-2); 国家自然科学基金资助项目(52179029, 51879289)。

作者简介: 邓志弘(1999-), 男, 硕士, 主要从事台风模拟与预报工作。E-mail: dengzh36@mail2.sysu.edu.cn

*通信作者: 刘丙军(1976-), 男, 教授, 博士, 主要从事水资源系统分析研究。E-mail: liubj@mail.sysu.edu.cn

法,证实了LSTM在水位实时“滚动预报”中的可行性;YANG等^[17]基于卷积神经网络(Convolutional Neural Networks, CNN)和LSTM建立“滚动预报”模型,改善了短期降水预报效果;ALASALI等^[18]基于机器学习算法建立了水位实时“滚动预报”模型,该模型准确率比固定预测模型高近93.5%。

本文以珠江河口地区作为研究区域,针对近年来影响珠江河口地区的热带气旋,选用CNN-LSTM结合EC技术建立台风路径、强度预报模型,并利用实时“滚动预报”思路建立“滚动预报”方法,对台风过程进行“滚动预报”,以期为该河口区台风预报提供一定的理论方法与技术参考。

1 研究区域及数据来源

珠江河口地区为亚热带气候,暖和湿润,热带气旋一般发生在7—9月,年平均为4~5次^[19]。据统计,近年来珠江河口地区台风灾害频发,强台风以上级别的发生频数增加趋势明显,给当地发展造成了巨大的影响^[20]。典型个例包括1713号台风“天鸽”和1822号台风“山竹”,两场台风的最大增水分别达到340 cm和279 cm,导致珠江河口多数站点的潮位超过200年一遇,给当地造成超百亿元的直接经济损失。

本文所采用数据为2004—2020年温州台风网的真实台风路径信息,其风场数据来源于美国国家环境预报中心,包含台风中心位置、台风等级、近中心最大风速、中心最低气压、移动速度、移动方向等

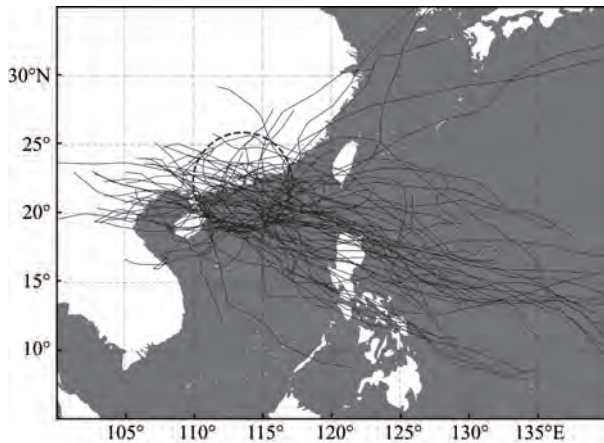


图1 2004—2020年影响珠江河口地区的历史台风路径

Fig.1 Historical typhoons' tracks affecting the Pearl River estuary from 2004 to 2020

信息,时间间隔为1~3 h。经过Pearson系数、Spearman系数、Kendall系数相关性分析以及多次测试优选,确定前12 h的台风中心位置、近中心最大风速及中心最低气压为建模的预报因子,同时将影响珠江河口地区的热带气旋定义为其中心进入以珠海站(113.58°E, 22.28°N)为圆心,影响半径为400 km范围内的热带气旋^[20]。对数据集进行处理分析,挑选出2004—2020年影响珠江河口地区的61场台风数据用于后续预报模型的训练与验证。

2 研究方法

2.1 CNN-LSTM模型

CNN被广泛认为是图像识别、自然语言处理和文本分类等领域的成熟技术^[21-22],在数据空间和时序上都具有良好的特征提取能力。CNN单个输出神经元的输出信息计算公式为:

$$y_{\text{CNN}} = g\left(\sum_{i,j,k} W_{\text{CNN}}^{i,j,k} x_{\text{CNN}}^{i,j,k} + b_{\text{CNN}}\right) \quad (1)$$

式中: y_{CNN} 是输出层中单个神经元的结果; g 为非线性激活函数; $W_{\text{CNN}}^{i,j,k}$ 为滤波器的权重; $x_{\text{CNN}}^{i,j,k}$ 为输入信息; b_{CNN} 为该神经元的偏置参数; i 、 j 、 k 分别代表输入信息的宽度、高度和图像通道数。

LSTM由HOCHREITER等^[23]于1997年提出,是RNN的变形,具有从时间序列中学习的能力,已被广泛应用于时间序列预测的研究中。LSTM的网络结构由输入门、遗忘门和输出门组成,控制信息的传递。

输入门负责控制当前时刻需要保存到神经元状态中的信息,公式为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (4)$$

遗忘门负责控制神经元状态中信息的取舍,公式为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

输出门负责控制神经元状态中信息的输出,公式为:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \times \tanh(C_t) \quad (7)$$

式中: i_t 、 f_t 、 o_t 为输入门、遗忘门和输出门; C_t 为当

前时刻的神经元状态, \tilde{C}_i 为神经元状态新的输入信息; h_i 和 h_{i-1} 为当前时刻和前一时刻的输出信息; x_i 为当前时刻的输入向量; W_i 、 W_c 、 W_f 和 W_o 为输入门、当前神经元状态、遗忘门和输出门中对应的权重矩阵; b_i 、 b_c 、 b_f 和 b_o 为对应门中的偏置向量; σ 和 \tanh 为激活函数。

本文通过 CNN 挖掘数据之间的关系, 并从中消除噪声及不稳定因素后输入至 LSTM 网络中, 搭建起台风路径预报、强度预报和 EC 模型。按照时间顺序将 2004—2020 年影响珠江河口地区的 80% 的台风数据作为训练集训练台风预报模型, 剩余 20% 的台风数据作为验证集验证模型性能。

2.2 实时滚动修正预报模型框架

误差自回归修正法基于模型预报误差前后相依规律, 利用历史误差系列的自相关性对未来误差进行预报, 并以此修正模型原始预报结果, 在 EC 技术中较为常见^[24]。考虑到台风预报误差序列可能存在较大的不确定性与非线性特征, 本文假定台风路径预报、强度预报的误差序列存在一定自相关性, 基于 CNN-LSTM 并根据前几个时段的已知预报误差建立自回归模型, 外延出预报时段内的模型误差, 将其与模型原预报值相加, 所得即为修正后的预报结果。

为解决传统机器学习模型单次预报方法中预报精度低、不确定性较大等问题, 同时还需要满足台风跟踪预报的时效性要求等, 本文利用实时“滚动预报”思想建立了实时滚动修正预报框架, 将 CNN-LSTM 与 EC 技术进行耦合, 随着时间推移不断利用实时数据进行更新并作为模型新的输入端进行预报, 以期逐时改善台风登陆前的预报效果, 更好地满足实际应用要求。该框架的基本原理与构建方法见图 2, 假定当前时刻为 T , 需要对台风路径与强度过程进行“滚动预报”, 设置每个滚动时刻的预见期为 12 h。具体操作流程为:

①基于历史台风数据, 构建基于 CNN-LSTM 的台风预报模型对历史数据进行预报, 与历史实际值相减获得历史序列预测误差, 然后根据自相关性分析确定误差修正模型输入端的自回归项, 从而建立误差修正模型。

②基于已有数据提取、计算出预报模型以及误差自回归修正模型所需要的输入端数据, 以此为基

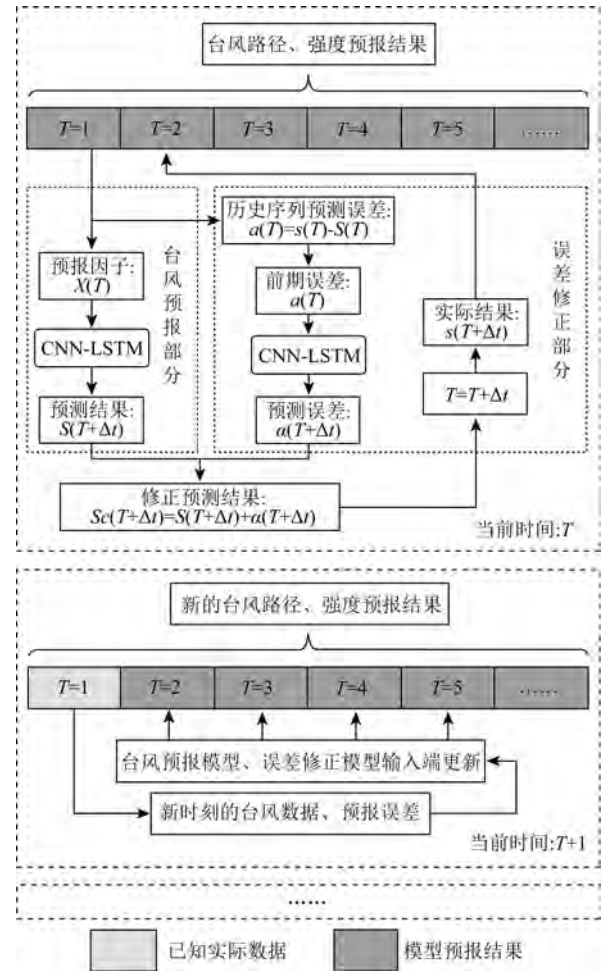


图 2 实时滚动修正预报流程图

Fig.2 Flow chart of real-time rolling corrected forecasts

础预报出下一时刻的路径、强度和误差结果, 再通过下一时刻的结果对下一时刻进行预报, 以此类推, 最终获得未来 12 h 内的台风预报结果以及误差预报结果。

③将上述结果进行合并获得未来 12 h 内的修正台风过程。

④将时间推移至 $T = T + 1$, 利用当前时刻最新的台风实际数据替换上一步该时刻的模型预报值, 同时计算出实际预报误差并替换上一步该时刻的误差预报值, 然后重复步骤②与③的预报过程, 最终获得对应当前滚动时刻新的未来 12 h 内的台风过程。

⑤以此类推, 随着时间逐时推移, 重复上述步骤, 不断对台风过程预报结果进行更新。

2.3 评价方法

本节选用4种传统指标评价模型的台风预报精度,即均方根误差(Root Mean Squared Error, RMSE)、平均绝对误差(Mean Relative Error, MAE)和平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)。计算公式为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (8)$$

$$MAE = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \quad (9)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{O_i - P_i}{O_i} \right| \times 100\% \quad (10)$$

式中: P_i 与 O_i 分别为第 i 时刻的台风预报值与实际值; \bar{O} 为实际值的平均值; n 是样本总数。RMSE与MAE均用于衡量预报值与实际值的差异性,随着实际数据与预报结果差异的增大,两者将从0逐渐增加到 $+\infty$,通常RMSE略大于MAE;MAPE是一种无量纲指标,取值范围为 $[0, +\infty)$,其中MAPE=0表明模型预报值与实际值完全吻合,当实际数据中存在0值时(即分母为0),无法使用该评价因子。

3 结果分析

3.1 台风路径预报结果分析

3.1.1 验证集预报结果

基于CNN-LSTM-EC建立实时滚动修正预报

模型,设置滚动间隔为1 h,每个滚动时刻的预见期为12 h,对验证集中的每场历史台风分别进行路径“滚动预报”,并对不同滚动次数下各场台风的路径预报误差进行均值计算。为了定量评估“滚动预报”在台风预报中的可靠性,此处计算了单次预报和不同滚动次数下验证集历史台风路径过程的RMSE和MAE均值(见图3),结果表明,“滚动预报”比单次预报有更好的路径预报效果。整体而言,随着模型滚动次数的增加,“滚动预报”的预报误差逐渐降低,各滚动次数下的RMSE和MAE基本都小于单次预报。例如,第1次“滚动预报”的RMSE为146.78 km,比单次预报误差减小3.79%,第24次滚动时,RMSE降低为127.74 km,减小幅度增大到16.27%,滚动期间RMSE的减小幅度最大达到25.67%,平均减小幅度为12.80%。此外,随着滚动次数的增加,预报误差变化趋势平稳,整体下降,这说明基于CNN-LSTM-EC搭建的路径“滚动预报”模型具有良好的稳定性。

计算可得在CNN-LSTM-EC方法下验证集整体预报的RMSE与MAE分别为68.25 km、46.61 km, CNN-LSTM方法下的误差分别为93.82 km、74.69 km,由此可见误差校正能提升路径预报模型的整体精度。表1展示了不同滚动次数下CNN-LSTM-EC与CNN-LSTM路径“滚动预报”的误差结果对比,前者相较于后者预报误差的减小幅度均在10%以上,最大可达22.57%,第5~19次滚动的误差减小幅度基本维持在15%~20%。

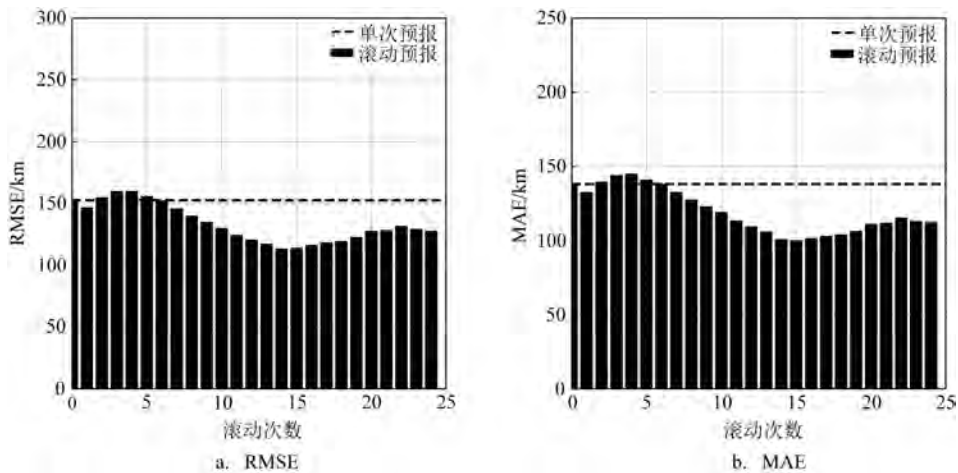


图3 验证集台风路径预报平均误差结果

Fig.3 Mean errors of typhoons' track forecasts in the validation set

表1 CNN-LSTM与CNN-LSTM-EC路径预报结果的RMSE(单位:km)

Tab.1 RMSE of typhoon's track forecasts between CNN-LSTM and CNN-LSTM-EC (unit:km)

预报方法	第1次	第2次	第3次	第4次	第5次	第6次	第7次	第8次
CNN-LSTM	186.403	182.708	183.974	184.854	185.356	182.074	178.319	173.832
CNN-LSTM-EC	146.782	154.910	159.966	160.016	155.828	152.380	146.081	139.997
减小幅度	21.26%	15.21%	13.05%	13.44%	15.93%	16.31%	18.08%	19.46%
预报方法	第9次	第10次	第11次	第12次	第13次	第14次	第15次	第16次
CNN-LSTM	168.335	161.648	156.445	152.983	150.186	146.452	143.854	143.322
CNN-LSTM-EC	134.927	130.158	124.520	120.832	117.423	113.399	113.971	116.298
减小幅度	19.85%	19.48%	20.41%	21.02%	21.81%	22.57%	20.77%	18.86%
预报方法	第17次	第18次	第19次	第20次	第21次	第22次	第23次	第24次
CNN-LSTM	143.561	143.561	144.92	146.784	148.149	150.074	151.328	152.871
CNN-LSTM-EC	118.383	119.450	122.755	127.661	128.370	131.802	129.277	127.740
减小幅度	17.54%	16.79%	15.29%	13.03%	13.35%	12.18%	14.57%	16.44%

注:加粗数据表示两种预报方法中RMSE较小者

3.1.2 台风天鸽预报结果

同理,以2017年8月22日07时(北京时,下同)为起报时间,对1713号台风“天鸽”进行路径“滚动预报”,此时台风“天鸽”已强化至强热带风暴等级。从图4可以看出,在个例预报中“滚动预报”同样比单次预报有更好的路径预报效果,随着模型滚动时间的延长,台风路径预报误差逐渐降低,预报结果更接近实际情况,整体精度有不断升高的趋势。在“滚动预报”中,前几次滚动的预报结果与单次预报结果相差不大,沿台风前进方向的误差均有向北偏移的情况,随着滚动时长增加,预报前进方向逐渐趋近实际,在第12次滚动之后,预报前进方向已与实际情况十分接近。具体来看,第24次滚动时,两者预报误差的RMSE降低为25.03 km,减小幅度增大到69.64%,滚动期间减小幅度最大达到85.74%,平均减小幅度为45.40%(见图5)。

3.2 台风强度预报结果分析

3.2.1 验证集预报结果

与路径预报同理,对验证集中的每场历史台风分别进行强度“滚动预报”,并对不同滚动次数下各场台风的强度预报误差进行均值计算。结果表明(见图6),“滚动预报”比单次预报有更好的强度预报效果,随着模型滚动次数的增加,台风强度预报

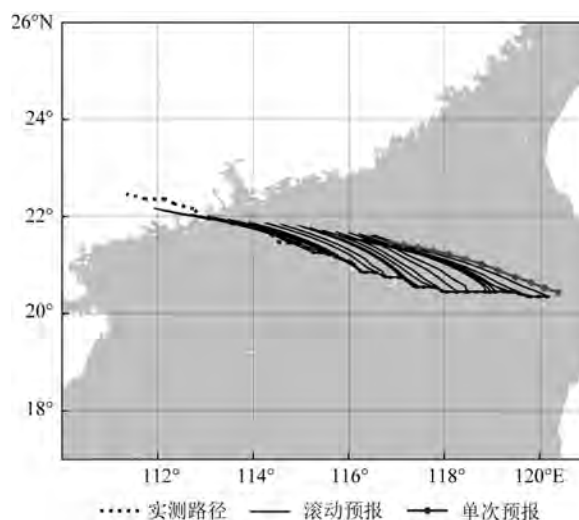


图4 台风天鸽路径预报结果

Fig.4 Track forecasts of Typhoon "Hato"

的整体精度有逐渐改善的趋势,不同滚动次数下台风强度预报结果的RMSE、MAE和MAPE均小于单次预报。例如,第1次滚动预报结果的RMSE、MAE和MAPE分别为3.35 m/s、2.83 m/s和8.60%,相对于单次预报误差的减小幅度仅为6.26%、8.13%和8.15%;第24次滚动时,RMSE、MAE和MAPE分别降低为2.84 m/s、1.39 m/s和4.12%,相对于单次预报误差的减小幅度提升了20.39%、54.98%和55.97%,平

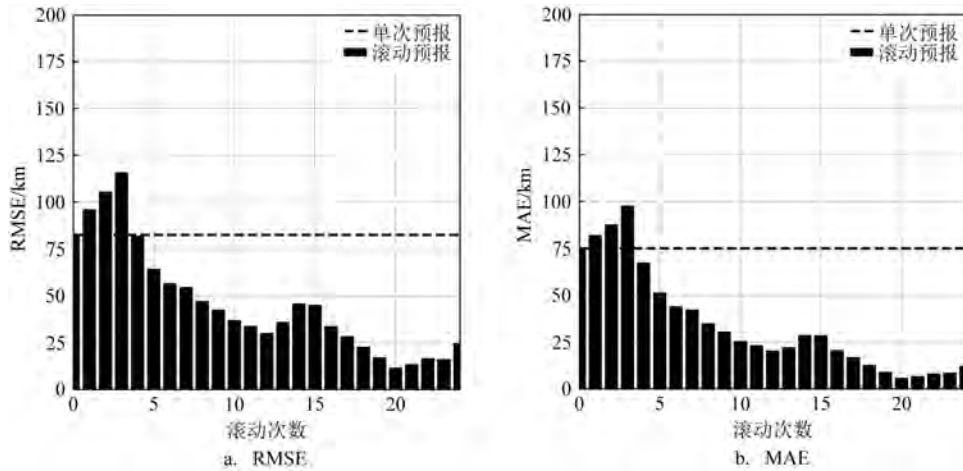


图5 台风天鸽路径预报误差结果

Fig.5 Errors of track forecasts of Typhoon "Hato"

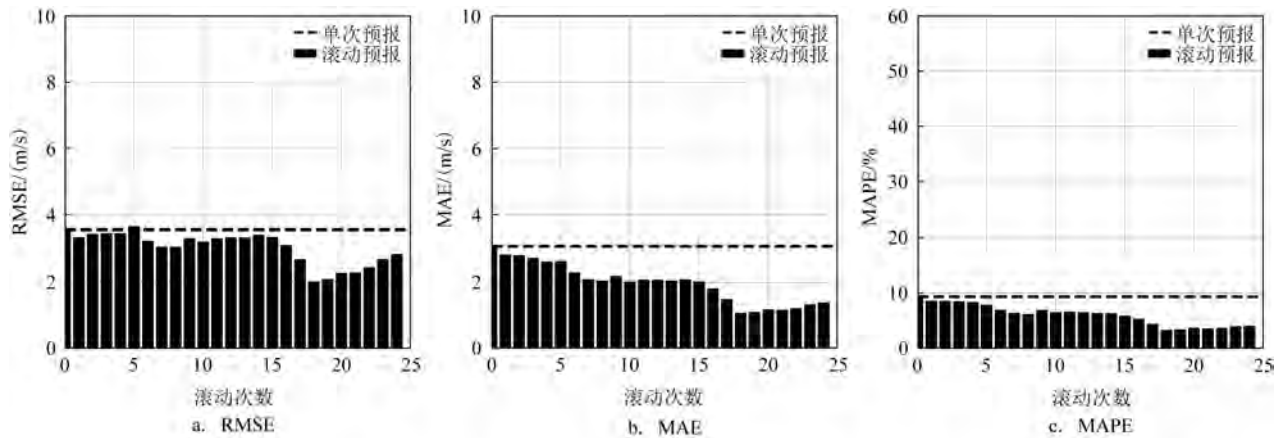


图6 验证集台风强度预报平均误差结果

Fig.6 Mean errors of typhoons' intensity forecasts in the validation set

均减小幅度分别为15.39%、37.38%、36.55%。

计算可得在CNN-LSTM-EC方法下验证集整体预报误差的RMSE、MAE和MAPE分别为1.45 m/s、0.94 m/s、3.78%，CNN-LSTM方法下的整体预报误差分别为1.67 m/s、1.19 m/s、4.56%，可见误差校正同样能提升强度预报模型的整体精度。表2展示了不同滚动次数下CNN-LSTM-EC与CNN-LSTM强度“滚动预报”的误差结果对比，可见EC技术同样能对强度“滚动预报”模型起到改善作用，但减小幅度基本保持在2.5%以内，改善效果不如路径“滚动预报”模型显著。

3.2.2 台风天鸽预报结果

对1713号台风“天鸽”进行强度“滚动预报”，并

对比“滚动预报”与传统单次预报的预报效果。从图7可以看出，在个例预报中“滚动预报”同样比单次预报有更好的强度预报效果。随着滚动次数的增加，预报整体精度有着逐渐改善的趋势。在前6次滚动中，“滚动预报”与单次预报均未能预测出台风的快速增强过程，推测原因在于本模型未考虑海表面温度等海洋条件以及其他大气环流因素对台风强度变化过程的影响，导致模型在输入端强度普遍偏低的情况下未能捕捉到台风的快速增强过程。但是在第7次滚动的时候，随着预报时效缩短、实际数据补充，预测的强度变化趋势出现明显改变，且与实际情况相近，大致预测出了台风的增强、削弱趋势。具体而言，模型在第7次滚动的时候，“滚动

表2 CNN-LSTM与CNN-LSTM-EC强度预报结果的RMSE(单位:m/s)

Tab.2 RMSE of typhoon's intensity forecasts between CNN-LSTM and CNN-LSTM-EC (unit:m/s)

预报方法	第1次	第2次	第3次	第4次	第5次	第6次	第7次	第8次
CNN-LSTM	3.324	3.482	3.490	3.449	3.651	3.286	3.135	3.118
CNN-LSTM-EC	3.346	3.446	3.471	3.470	3.655	3.240	3.059	3.060
减小幅度	-0.69%	1.03%	0.55%	-0.61%	-0.09%	1.39%	2.41%	1.86%
预报方法	第9次	第10次	第11次	第12次	第13次	第14次	第15次	第16次
CNN-LSTM	3.350	3.244	3.328	3.337	3.329	3.433	3.392	3.156
CNN-LSTM-EC	3.318	3.213	3.320	3.342	3.335	3.419	3.356	3.112
减小幅度	0.94%	0.95%	0.23%	-0.14%	-0.18%	0.40%	1.05%	1.40%
预报方法	第17次	第18次	第19次	第20次	第21次	第22次	第23次	第24次
CNN-LSTM	2.738	2.067	2.126	2.296	2.313	2.466	2.698	2.843
CNN-LSTM-EC	2.687	2.014	2.089	2.271	2.289	2.444	2.691	2.842
减小幅度	1.86%	2.53%	1.71%	1.12%	1.04%	0.90%	0.27%	0.02%

注:加粗数据表示两种预报方法中RMSE较小者

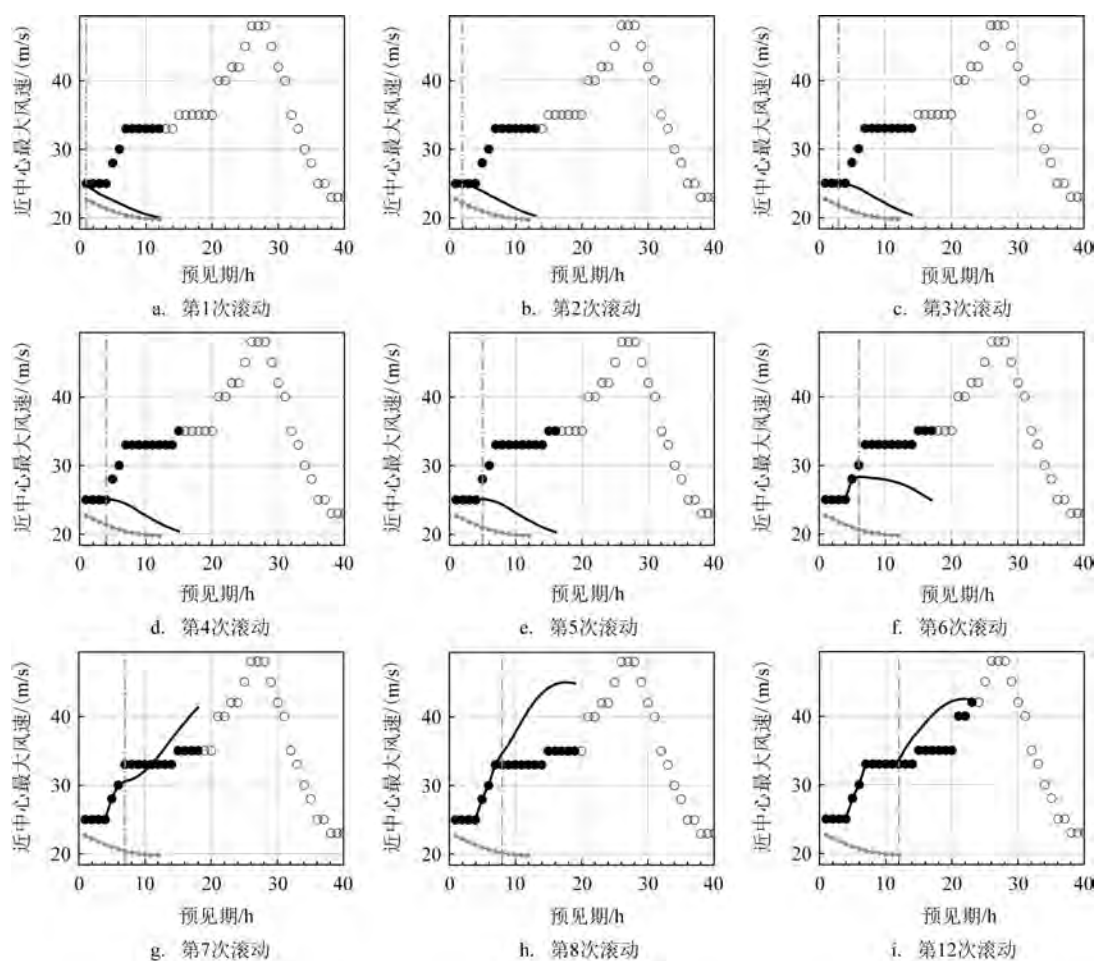


图7 台风“天鸽”强度预报结果

Fig.7 Intensity forecasts of Typhoon "Hato"

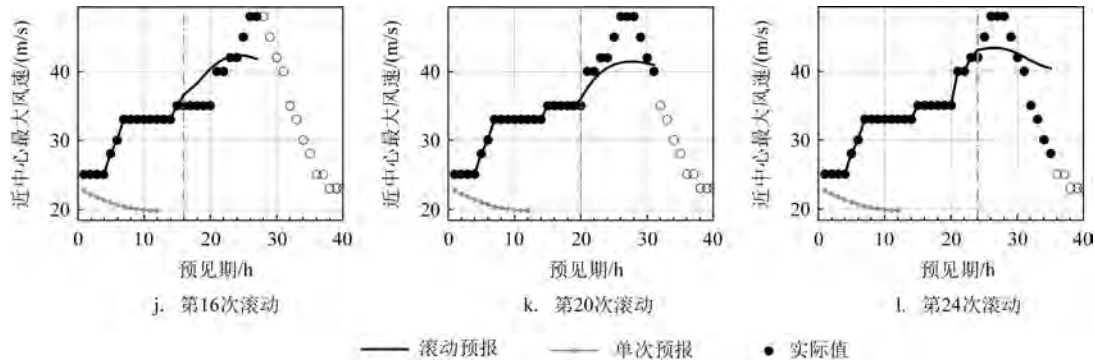


图7 (续)

Fig.7 (Continued)

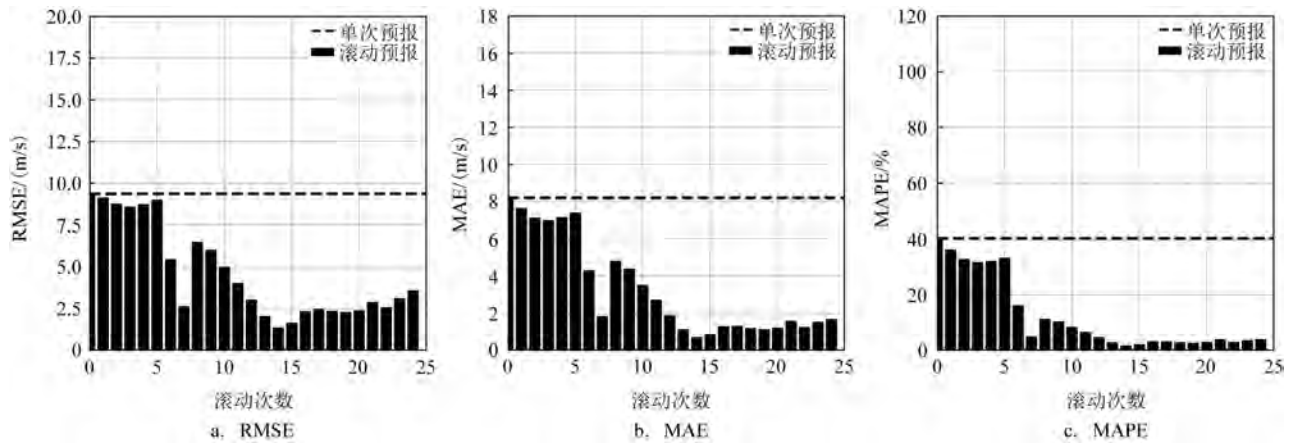


图8 台风“天鸽”强度预报误差结果

Fig.8 Errors of intensity forecasts of Typhoon "Hato"

预报”与实际结果的RMSE、MAE和MAPE分别降低为2.66 m/s、1.83 m/s和5.00%，相对于单次预报误差的减小幅度分别从第6次滚动的41.56%、47.38%和60.00%变为71.61%、77.61%和87.60%（见图8），与路径“滚动预报”结果不同，这一现象主要源于台风强度变化趋势预测结果的大幅改变，这也进一步印证了“滚动预报”在实际应用中的必要性。

4 结论

为解决传统机器学习模型单次预报方法中预报精度低、不确定性较大等问题，本文利用实时“滚动预报”的思想，基于CNN-LSTM和EC技术搭建了误差校正预报模型，通过验证集历史台风以及1713

号台风“天鸽”为实例进行验证分析，并对比了误差校正在实时“滚动预报”模型中的预报效果。主要结论如下：

①“滚动预报”比单次预报有更好的路径预报效果，随着模型滚动时间的延长，预报结果更接近于实际情况，整体精度有着不断升高的趋势。在验证集预报结果中，路径“滚动预报”结果的RMSE从第1次滚动的146.78 km降低至第24次滚动的127.74 km，相对于单次预报误差减小幅度达到16.27%，滚动期间的减小幅度最大达到25.67%，平均减小幅度达12.80%。

②“滚动预报”比单次预报有更好的强度预报效果，特别是随着预报时效的缩短，预报结果明显改善。在验证集预报结果中，强度“滚动预报”结果的MAE从第1次滚动的2.83 m/s降低至第24次滚

动的 1.39 m/s, 相对于单次预报误差的减小幅度达到 54.98%, 滚动期间的减小幅度最大达到 65.04%, 平均减小幅度达 37.38%。

③CNN-LSTM-EC 的路径、强度“滚动预报”效果均优于 CNN-LSTM。在不同滚动次数下, CNN-LSTM-EC 相较于 CNN-LSTM 的路径“滚动预报”误差减小幅度均在 10% 以上, 最大可达 22.57%, 强度预报误差减小幅度基本保持在 2.5% 以内。

参考文献:

- [1] EMANUEL K. Increasing destructiveness of tropical cyclones over the past 30 years[J]. *Nature*, 2005, 436(7051): 686-688.
- [2] HOYOS C D, AGUDELO P A, WEBSTER P J, et al. Deconvolution of the factors contributing to the increase in global hurricane intensity[J]. *Science*, 2006, 312(5770): 94-97.
- [3] WEBSTER P J, HOLLAND G J, CURRY J A, et al. Changes in tropical cyclone number, duration, and intensity in a warming environment[J]. *Science*, 2005, 309(5742): 1844-1846.
- [4] 王洁, 杨奕杰, 王杰, 等. 基于近 20a 历史数据的中国沿海城市台风灾害风险评估[J]. *海洋预报*, 2021, 38(5): 24-30.
WANG J, YANG Y J, WANG J, et al. Typhoon disaster risk assessment of coastal cities in China based on historical data over the past 20 years[J]. *Marine Forecasts*, 2021, 38(5): 24-30.
- [5] 陈煜, 杨剑, 段忠东, 等. 粤港澳大湾区台风危险性分析[J]. *自然灾害学报*, 2022, 31(2): 26-38.
CHEN Y, YANG J, DUAN Z D, et al. Typhoon hazard analysis of the Guangdong-Hong Kong-Macao Greater Bay Area[J]. *Journal of Natural Disasters*, 2022, 31(2): 26-38.
- [6] CHEN R, ZHANG W M, WANG X. Machine learning in tropical cyclone forecast modeling: a review[J]. *Atmosphere*, 2020, 11(7): 676.
- [7] 曹祥村, 邵利民. 一种利用 BP 网络预报台风路径的新方法[J]. *海洋预报*, 2007, 24(3): 75-82.
CAO X C, SHAO L M. A new method of forecasting typhoon paths using BP Network[J]. *Marine Forecasts*, 2007, 24(3): 75-82.
- [8] ALI M M, KISHTAWAL C M, JAIN S. Predicting cyclone tracks in the north Indian Ocean: an artificial neural network approach[J]. *Geophysical Research Letters*, 2007, 34(4): L04603.
- [9] ALEMANY S, BELTRAN J, PEREZ A, et al. Predicting hurricane trajectories using a recurrent neural network[C]//Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. Honolulu: AAAI, 2019: 58.
- [10] KORDMAHALLEH M M, SEFIDMAZGI M G, HOMA-IFAR A. A sparse recurrent neural network for trajectory prediction of Atlantic hurricanes[C]//Proceedings of Genetic and Evolutionary Computation Conference 2016. Denver: ACM, 2016: 957-964.
- [11] PAN B, XU X, SHI Z W. Tropical cyclone intensity prediction based on recurrent neural networks[J]. *Electronics Letters*, 2019, 55(7): 413-415.
- [12] GUO R, QI L B, GE Q Q, et al. A study on the ensemble forecast real-time correction method[J]. *Journal of Tropical Meteorology*, 2018, 24(1): 42-48.
- [13] LIN I I, CHEN C H, PUN I F, et al. Warm ocean anomaly, air sea fluxes, and the rapid intensification of tropical cyclone Nargis (2008)[J]. *Geophysical Research Letters*, 2009, 36(3): L03817.
- [14] SANDERY P A, BRASSINGTON G B, CRAIG A, et al. Impacts of ocean-atmosphere coupling on tropical cyclone intensity change and ocean prediction in the Australian region[J]. *Monthly Weather Review*, 2010, 138(6): 2074-2091.
- [15] LIU Y, WANG H, LEI X H, et al. Real-time forecasting of river water level in urban based on radar rainfall: A case study in Fuzhou City[J]. *Journal of Hydrology*, 2021, 603: 126820.
- [16] LIU Y, WANG H, FENG W W, et al. Short term real-time rolling forecast of urban river water levels based on LSTM: a case study in Fuzhou city, China[J]. *International Journal of Environmental Research and Public Health*, 2021, 18(17): 9287.
- [17] YANG R Y, MU J L, WANG S D, et al. Hourly rolling correction of precipitation forecast via convolutional and long short-term memory networks[J]. *Atmospheric Science Letters*, 2022, 23(10): e1100.
- [18] ALASALI F, TAWALBEH R, GHANEM Z, et al. A sustainable early warning system using rolling forecasts based on ANN and golden ratio optimization methods to accurately predict real-time water levels and flash flood[J]. *Sensors*, 2021, 21(13): 4598.
- [19] 刘天绍, 刘孙俊, 杨玺, 等. 1951—2015 影响广东沿海台风的统计分析[J]. *海洋预报*, 2018, 35(4): 68-74.
LIU T S, LIU S J, YANG X, et al. Statistical analysis of the typhoon influencing Guangdong province during 1951-2015[J]. *Marine Forecasts*, 2018, 35(4): 68-74.
- [20] 叶荣辉, 戈军, 张文明, 等. 影响粤港澳大湾区的热带气旋统计分析[J]. *水利水电技术*, 2020, 51(S1): 37-43.
YE R H, GE J, ZHANG W M, et al. Statistical analysis on impact from tropical cyclone on Guangdong-HongKong-Macao Greater Bay Area[J]. *Water Resources and Hydropower Engineering*, 2020, 51(S1): 37-43.
- [21] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. *Neural Computation*, 1989, 1(4): 541-551.
- [22] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [23] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [24] 周梦, 陈华, 郭富强, 等. 洪水预报实时校正技术比较及应用研

究[J]. 中国农村水利水电, 2018(7): 90-95.

ZHOU M, CHEN H, GUO F Q, et al. The application of real-time

correction techniques for flood forecasting[J]. China Rural Water and Hydropower, 2018(7): 90-95.

Real-time rolling correction forecasting of typhoon process in the Pearl River estuary based on CNN-LSTM

DENG Zhihong¹, LIU Bingjun^{1,2*}, ZHANG Ka¹, HU Shikun¹, ZENG Hui³, ZHANG Mingzhu³, LI Dan³

(1. School of Civil Engineering, Sun Yat-sen University, Zhuhai 519085, China; 2. Water Resources and Environment Research Center of Sun Yat-sen University, Guangzhou 510275, China; 3. Guangzhou Hydraulic Research Institute, Guangzhou 510220, China)

Abstract: In order to improve the accuracy of typhoon forecasting, this paper introduces a real-time rolling corrected typhoon forecasting model in the Pearl River Estuary utilizing Convolutional Neural Network Long Short-Term Memory (CNN-LSTM) neural network and Error Correction (EC) method. The results show that the rolling forecasts have better performances on typhoon's track and intensity than the single-time forecasts. The overall accuracy of the rolling forecasts increases gradually along with the prolong of the rolling time of the model. In comparison with the single-time forecasts, the root mean squared error of typhoon's track rolling forecasts decreases by 25.67% and the mean absolute error of typhoon's intensity rolling forecasts decreases by 65.04%. The real-time rolling corrected forecasts of typhoon's track and intensity based on CNN-LSTM-EC are better than those based on CNN-LSTM. Compared with the latter, the forecasting error of the former decreases by 22.57% on the typhoon's track and by 2.5% on the typhoon's intensity.

Key words: real-time rolling forecast; typhoon; Pearl River estuary; deep learning; error correction