

DOI:10.16867/j.issn.1673-9264.2021313

刘媛媛,刘业森,张丽,等.基于神经网络模型对风暴潮特征分析及预测——以深圳赤湾站和南澳站为例[J].中国防汛抗旱,2022,32(7):66-71.LIU Yuanyuan, LIU Yesen, ZHANG Li, et al.Characteristics analysis and tide level prediction of storm surge based on neural network model——Taking Chiwan and Nan'ao in Shenzhen City as an example[J].China Flood & Drought Management, 2022, 32(7):66-71.(in Chinese)

基于神经网络模型对风暴潮特征分析及预测

——以深圳赤湾站和南澳站为例

刘媛媛^{1,2} 刘业森^{1,2} 张丽³ 力梅³ 穆杰^{1,2}

(1. 中国水利水电科学研究院,北京100038;2. 水利部防洪抗旱减灾工程技术研究中心,北京100038;

3. 深圳市国家气候观象台,深圳519082)

摘要:我国沿海城市经常遭受台风的影响,台风带来的强降雨如果遭遇风暴潮高潮位,会造成严重的洪涝灾害。在台风到来之前,对风暴潮位进行科学准确的分析和预测,及时合理地调度工程,对降低沿海城市的洪涝灾害风险具有重要的意义。在分析了风暴潮增水影响因素的基础上,在赤湾站和南澳站分别建立了长短期记忆人工神经网络预测模型。结果表明,处于不同位置的潮位站,风暴潮位影响因素及前时序 n 各不相同,利用前3h潮位和风速预测南澳站的风暴潮位,误差最小,而赤湾站则是利用前6h的潮位和风速预测,误差最小。

关键词:风暴潮;LSTM神经网络;智能预测;深圳

中图法分类号:P731.23

文献标识码:A

文章编号:1673-9264(2022)07-66-06

1 研究背景

台风是我国沿海城市经常出现的一种灾害天气,对城市的影响主要包括强风、暴雨和风暴潮^[1]。暴雨造成的河道洪水和内涝积水受高潮位顶托,排泄不畅,造成严重的城市洪涝灾害。随着沿海经济的飞速发展,洪涝灾害带来的经济损失也随之逐年增加,已成为严重影响经济发展的主要灾害之一^[2]。因此,对风暴潮位进行科学准确的预测预警,合理地调度工程,及时安排相关人员及物资服务救灾,对降低城市洪涝灾害风险具有重要的意义。

目前,对风暴潮预测的研究主要包括数值模型和人工神经网络模型。从20世纪80年代,我国就开展了风暴潮数值预报技术研究^[3-4],风暴潮的数值研究和应用发

展迅速,并取得了较好的成果^[5-6]。但是由于数值预报模型对数据的要求较高,而海底地形、风场等数据的不确定性,在一定程度上影响了数值模拟的精度。随着人工智能技术的发展,神经网络模型对风暴潮进行预测,逐渐成为重要的研究方向^[7]。Kim等^[8]用神经网络模型模拟了日本Sanin海岸水位变化,薛彦广等^[9]、王云如等^[10]都利用了人工神经网络的方法预测了风暴潮。类似的还有小波神经网络^[11]、递归神经网络^[12]等预测方法,这些神经网络模型基本都是基于历史潮位数据进行分析预测的。

风暴潮潮位的高低,不仅跟潮位数据有关,还和气象条件、海岸形态、沿岸地形等因素有关。这些预测模型虽然都有一定的效果,但预测误差会随着预测时间的增加或气象条件的变化逐步增大,且预测结果与实测值的相位差会

收稿日期:2021-11-02

第一作者信息:刘媛媛,女,正高级工程师,博士,E-mail:50237479@qq.com。

通信作者信息:刘业森,男,高级工程师,博士,E-mail:liuys@iwhr.com。

基金项目:国家自然科学基金项目(52009147)。

越来越大。由于不同岸段的海岸形态相差较大,以某一点的风暴潮样本训练出的模型,对较长岸段的风暴潮进行预测,误差也会比较明显。

近几年深度学习技术日新月异,在语音识别、音乐生成、文本生成等方面取得了成功的应用。这些成功的应用都离不开循环神经网络RNN(Recurrent Neural Network)模型。RNN作为一种新型的网络结构,非常擅长处理序列数据问题。RNN模型有很多变种模型,长短期记忆模型LSTM(Long-Short Term Memory)就是其中一种,可以有效解决RNN模型梯度消失或者爆炸的问题,擅长处理多因素的时间序列预测问题,而风暴潮的预测,就是多因素的时间序列预测问题。

本文首先以深圳赤湾站为例,分析了1964-2019年赤湾站附近沿岸风暴潮资料,探讨了风暴潮风速、气压等相关关系。在此基础上,以赤湾站和南澳站为例,分岸段建立了LSTM神经网络风暴潮预测模型。由于台风降水具有历时短、强度大、雨量集中的特点,未来1~3h的潮位,对深圳市排涝的影响最大。因此,本文对赤湾和南澳站未来1~3h的潮位进行了预测,并对不同岸段、不同影响因素以及前时序n的敏感性进行了分析。结果表明,风暴潮位不仅和前时序潮位相关,和气象因素以及海岸形态、沿岸地形等均有很强的相关性,需要考虑不同影响因素,对不同岸段分别建立神经网络模型。

2 LSTM神经网络预测模型

目前神经网络技术使用得比较成熟的,主要包括卷积神经网络CNN(Convolutional Neural Networks)和循环神经网络RNN。其中CNN主要用于图像处理,RNN主要用于序列数据处理。但是传统的RNN模型在处理时间序列上距离较远的节点之间的参数时,往往出现梯度消失或者梯度爆炸的问题。Hochreiter等^[13]对RNN模型进行了改进,提出了LSTM预测模型。LSTM在隐藏层中增加了一个隐藏状态 C_t ,增加了3个门:遗忘门、输入门和输出门,解决了RNN模型中梯度消失或者梯度爆炸的问题,模型结构图如图1所示。

遗忘门的输出为 f_t ,如式(1)所示:

$$f_t = \sigma(W_f S_{t-1} + U_f x_t + b_f) \quad (1)$$

由式(1)可知,遗忘门的输出 f_t 由本时序的样本 x_t 和上一时序隐藏层的输出 S_{t-1} 决定,其中 σ 为sigmoid函数。

输入门的输出如式(2)、式(3)所示,其中式(2)中的激活函数为sigmoid函数,式(3)中的激活函数为tanh函数。

$$i_t = \sigma(W_i S_{t-1} + U_i x_t + b_i) \quad (2)$$

$$a_t = \tanh(W_a S_{t-1} + U_a x_t + b_a) \quad (3)$$

由这两部分输出更新 C_t 的状态:

$$C_t = C_{t-1} \odot f_t + i_t \odot a_t \quad (4)$$

其中, \odot 为哈达玛积。

输出门的输出如式(5)所示,其中激活函数为sigmoid函数:

$$o_t = \sigma(W_o S_{t-1} + U_o x_t + b_o) \quad (5)$$

则隐含层的输出 S_t 由输出门的输出 o_t 和 C_t 乘积得到:

$$S_t = o_t \odot \tanh(C_t) \quad (6)$$

预测输出如式(7)所示,激活函数为sigmoid函数:

$$\hat{y}_t = \sigma(V S_t + b_t) \quad (7)$$

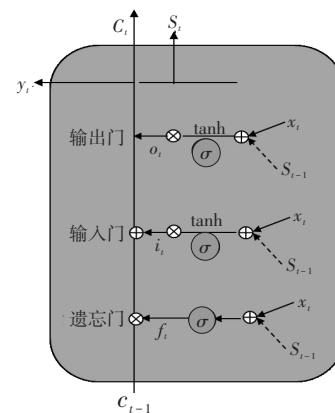


图1 LSTM模型隐藏层结构图

3 基于LSTM的风暴潮位预报模型构建

风暴潮位主要是由大气剧烈扰动所致,受台风强度、风速、风向等气象因素影响^[14],还和前时序天文潮位有关。如果台风登陆遇到涨潮期,即使台风强度不大,也可能造成严重的影响;相反,即使台风强度较大,但是若遇到天文低潮,也不会造成很严重的影响。因此,前时序潮位和气象因素对风暴潮位都有影响,在建立神经网络预测模型之前,需要对这些影响因素进行分析,以选择最适合的影响因素。

3.1 影响风暴潮位的相关因子分析

赤湾站位于珠江口东侧,南邻香港、澳门、珠海,1964年建立,有超50a的风暴潮资料,可以较好地代表赤湾沿岸风暴潮的变化特点。本文以1964-2019年赤湾站沿岸风暴潮增

水超过 50 cm,造成较大危害的风暴潮资料为样本,对影响风暴潮增水的相关因子进行分析。资料包括热带气旋登陆时的中心气压、最大风速以及当时赤湾站沿岸的最大增水情况。数据主要来自《台风暴雨洪水风暴潮手册》第三册风暴潮部分和2013-2019年的《广东省海洋灾害公报》。

赤湾站沿岸 1964-2019年台风登陆引起增水超过 50 cm的风暴潮共有 82次,都发生在 5-11月,其中 7-9月发生的频率最高,占总次数的 72%,具体如图 2 所示。

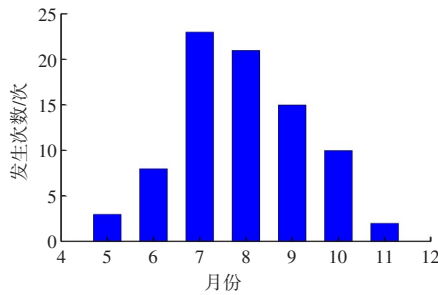


图2 赤湾站风暴潮月次变化

图 3、图 4 分别是风暴潮最大增水和登陆时最大风速和气压的对比图。如图 3 所示,赤湾站风暴潮最大增水和登陆时的风速变化趋势一致,具有很大的相关性,登陆时的风速越大,引起的风暴潮增水也更高。图 4 中,风暴潮增水和登陆时热带气旋中心气压变化趋势相反,中心气压较低时,风暴潮增水较高,而中心气压较高时,风暴潮增水较低。由此可知,登陆时气旋的最大风速和登陆沿岸的风暴潮增水相关性很强,而中心气压和风暴潮增水变化趋势相反。

为了进一步分析风暴潮增水和登陆风速以及气旋中心气压的相关性,分别计算风暴潮增水和气旋登陆时的风速以及气旋中心气压的相关系数,相关系数的计算如式 8 所示:

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (8)$$

式中: r 为相关系数; x 为最大风暴潮增水, cm; y 分别为气旋登陆时的最大风速和气旋中心最低气压, m/s、hpa。经计算,风暴潮最大增水和登陆时的风速的相关系数达到 0.62,和登陆时热带气旋中心最低气压的相关系数为 -0.58。也就是说,风暴潮最大增水和登陆时的风速强正相关,而和气旋中心气压强负相关。

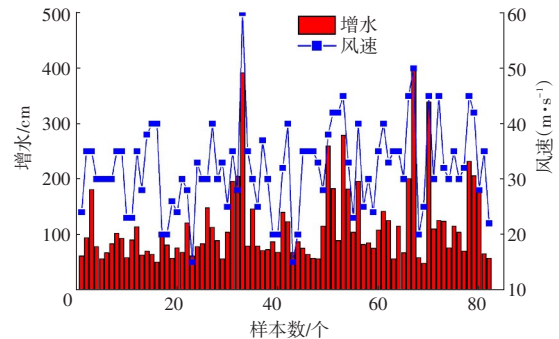


图3 风暴潮增水和登陆时风速对比图

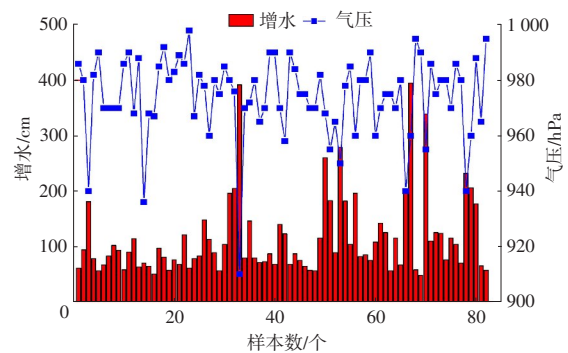


图4 风暴潮增水和登陆时中心最低气压对比图

3.2 LSTM 预报模型

不同岸段的地理形态、潮汐规律、风暴潮增水和风、浪的情况分布各不相同,而且影响风暴潮位的影响因素也各有不同,需要分别建模型分析。本文选择数据时间序列较长、质量较好,分别位于深圳河湾岸段的赤湾站和大鹏湾岸段的南澳站建立神经网络预测模型。

风暴潮增水是台风活动时期的实测潮位和天文潮位的差,而受台风影响海区的潮位站只有实时水位观测数据,根据实时水位观测数据和天文潮位的差得出的风暴潮增水误差较大^[15],因此本文选择实测潮位值(天文潮叠加风暴潮增水)作为样本,进行模型的训练,并对未来 1~3 h 的总潮位(天文潮+风暴潮增水)进行预测。

如本文第 3.1 节分析,风暴潮增水和前时序潮位以及登陆时最大风速相关,因此建立风暴潮预测模型时,选择赤湾和南澳潮位站 2008-2019 年逐小时的风速以及潮位数据为训练样本,对未来 1~3 h 潮位进行预测。

将赤湾站和南澳站 2008-2019 年共 96 433 组样本中的 90% 作为训练样本,余下的 10% 作为测试样本,对模型进行测试。为更好地说明模型的预测效果,引起较大风暴潮的

201713号台风“天鸽”和201822号台风“山竹”的样本数据不参与模型训练,作为预测样本,利用训练好的模型对这两场台风引起的总潮位进行预测,以检验模型的预测能力。

模型中, t 为预测的时刻, $X_{t-i} = \{x_{t-i}^1, x_{t-i}^2\}$ ($i=1,$

$2, \dots, n$) 为 t 之前 i 个时刻的 2 个影响因素,包括前时序潮位、风速。模型的输入层的神经元 cell 个数为 n 个,隐藏层神经元设为 50 个。取 t 时刻的输出 Y_t 作为预测结果。训练时,为了避免各类数据数值区间的差异性,在训练学习之前,对原始数据做归一化处理,公式如下:

$$x^{*i} = \frac{x^i - x_{\min}^i}{x_{\max}^i - x_{\min}^i} \quad (9)$$

式中: x^{*i} 为某一类 factors 归一化之后的值, x^i 为原始数值, x_{\min}^i 为该类 factor 中的最小值, x_{\max}^i 为其中最大值。

训练时,模型迭代次数 Epochs 为 100,损失函数为平均绝对误差 (MAE),优化器为 Adam 算法。最后通过计算预测数据与真实数据的均方根误差 (RMSE) 来评估模型的训练结果。经计算,赤湾站和南澳站的预测模型,训练次

数在 25 次之内,误差都迅速下降,在训练次数 40 次左右,误差曲线均达到了拐点,误差下降速度逐渐趋于平缓,模型逐渐收敛。说明本文选择的 LSTM 神经网络算法合理,输入的影响因素相关性强,训练模型的数据样本质量可靠。

4 结果分析

为了选择最优的输入因子及前时序值,以训练出更智能的预测模型,本文选取前 2~13 h (前时序 $n=2 \sim 13$) 的不同影响因素,作为模型输入条件,对模型进行训练,并利用训练好的预测模型,以台风“山竹”和“天鸽”为例,对赤湾站和南澳站未来 1~3 h 的总潮位进行预测,预测结果的均方误差如图 5 所示。

从图 5 可以看出,南澳站和赤湾站预测模型,单独考虑前时序 $n=2 \sim 13$ h 的天文潮位、单独考虑风速以及同时考虑前时序天文潮位和风速的因素,对未来 1~3 h 的总潮位都可以很好地预测。其中,对未来 1 h 总潮位预测的平均误差在 0.18 m 左右,而对未来 3 h 总潮位预测的平均误差在 0.3 m 左右,基本满足日常应急工作的要求。

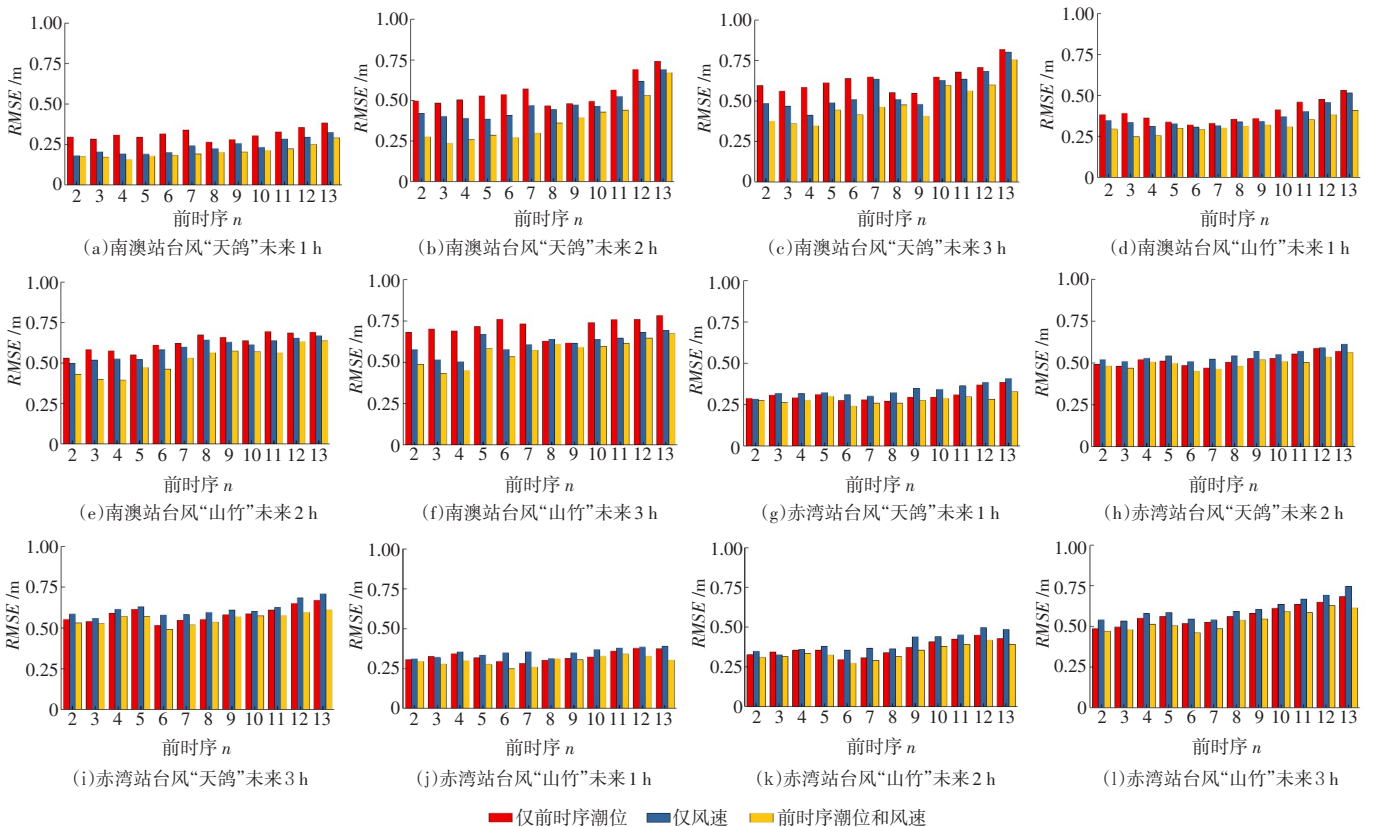


图5 南澳站和赤湾站潮位预测误差图

图5中红色的柱状图是单独考虑前时序潮位因素时,用前2~13 h的潮位预测未来1~3 h总潮位的均方误差。从图中可以看出,对于南澳站,如果仅考虑前时序潮位,当 $n=8\sim 9$ 时,误差最小。而赤湾站是 $n=6\sim 7$ 时,预测效果较好。即南澳站利用前8~9 h的潮位预测未来1~3 h总潮位,误差最小,而赤湾站利用前6~7 h的潮位预测未来1~3 h的总潮位,误差最小。

这主要是因为南澳站和赤湾站在沿岸的位置不同,潮位特征不同。南澳站位于深圳东部岸段,平均落潮历时5~6 h,平均涨潮历时7~8 h。赤湾站位于深圳中部岸段,涨落潮历时,均在6~7 h。因此,南澳站在8 h之内会出现一次完整的涨落潮过程,而赤湾站则在7 h之内。对于神经网络来说,训练样本包含的特征越多,训练出来的模型越智能,预测结果越好。因此,单独考虑前时序潮位因素时,南澳站前时序 $n=8\sim 9$ 时效果最好,而赤湾站 $n=6\sim 7$ 的效果最好。

图5中蓝色的柱状图是单独考虑风的因素时,用前2~13 h的风速预测未来1~3 h总潮位的均方误差。从图中可以看出,当 $n=3\sim 4$ 时,南澳站总潮位预测的误差较小,赤湾站当 $n=2\sim 3$ 时误差较小。即南澳站利用前3~4 h的风速预测未来1~3 h总潮位,误差最小,而赤湾站利用前2~3 h的风速预测未来1~3 h的总潮位,误差最小。

这主要是因为气旋在深圳东部沿岸登陆的情况较多,位于东部沿岸的南澳站比位于中部沿岸的赤湾站对

风速更敏感。南澳站在台风登陆前3~4 h就会受到风的影响,产生较大的涌浪。因此,单独考虑风速时,南澳站 $n=3\sim 4$,和风暴潮的相关性更强,模型的预测误差更小。

图5中黄色的柱状图是同时考虑了前时序潮位和风速预测模型的均方误差,从图5中可以看出,对于南澳站和赤湾站来说,同时考虑前时序潮位和风速影响的预测模型,在这3种模型中,预测误差最小。该模型中,南澳站 $n=3$ 时,预测误差最小,赤湾站 $n=6$ 时,预测误差最小。即南澳站选择前3 h的潮位和风速、赤湾站选择前6 h的潮位和风速作为输入因素,预测结果误差最小。

由此可见,由于各潮位站所处的地理位置不同,影响因素及前时序 n 各不相同。而且并不是考虑的影响因素越多、前时序取得越长,神经网络模型预测效果就越好。需要对各潮位站分别分析,选择相关性更强、误差最小的因素,作为各潮位站预测模型的输入因素。

本文利用该模型对台风“天鸽”和“山竹”未来1~3 h的总潮位进行了预测,结果如图6所示。从图6中可以看出,该模型对南澳站和赤湾站未来1~3 h总潮位的预测结果,和实测值吻合得都很好,其中,未来1 h最高潮位时间相位误差在0.16 h之内,未来3 h最高潮位时间相位误差在0.3 h之内。而且从图中可以看出,该模型不仅可以准确预测天文潮位,对于台风过程中的潮位的极值,包括高潮极值和低潮极值,也可以准确地预测。

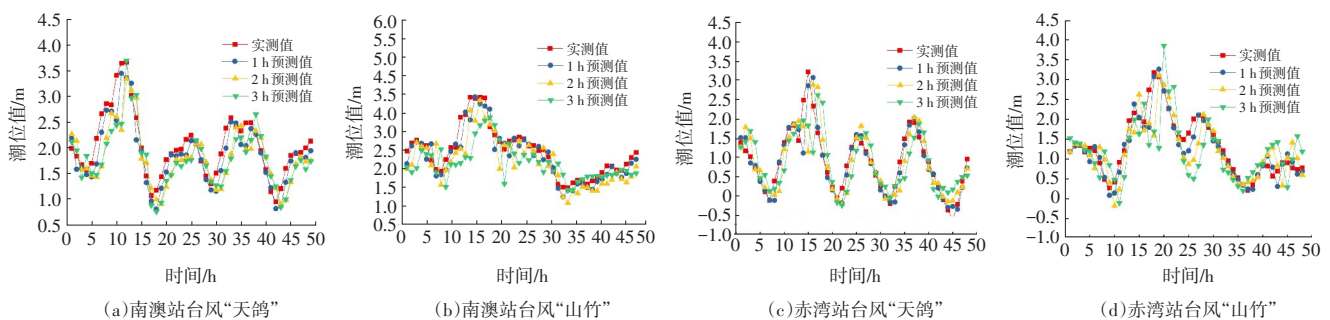


图6 实测值和未来1~3 h潮位预测值对比图

5 结 论

影响风暴潮位的因素很多,不同因素的前时序 n ,对风暴潮的影响也不同。本文首先分析了影响风暴潮的主要因素,在此基础上建立了赤湾站和南澳站风暴潮位LSTM人工神经网络预测模型,并对不同潮位站风暴潮影响因素

及其前时序 n 进行了分析,结论如下:

(1)基于LSTM模型的风暴潮智能预报模型,可以对深圳沿岸南澳站和赤湾站未来1~3 h的潮位进行准确、合理的预测。对未来1 h总潮位预测的平均误差在0.18 m左右,而对未来3 h总潮位预测的平均误差在0.3 m左右,基本满足沿海城市排涝应急工作的需要。

(2)仅考虑前时序潮位或者仅考虑前时序风速因素的人工神经网络模型,对潮位的预测误差较大,而综合考虑前时序潮位和风速因素的模型,对潮位的预测误差最小。

(3)由于不同的潮位站所处的地理位置不同,其气象条件和潮汐特征也各不相同,需要分别针对潮位站各自的特点,确定不同影响因素。经分析,热带气旋登陆时的最大风速对南澳站潮位的影响比对赤湾站的大,而前时序潮位对赤湾站潮位影响较大。南澳站利用前3 h的潮位和风速预测未来1~3 h总潮位,误差最小,而赤湾站利用前6 h的潮位和风速预测未来1~3 h的总潮位,误差最小。

参考文献

- [1] 张月霞,王辉.台风风暴潮灾害风险评估研究综述[J].海洋预报,2016,33(2):81-88.
- [2] 王秀杰,胡冰,苑希民,等.洪水与风暴潮共同作用下的溃堤洪水一维、二维耦合模型及应用[J].南水北调与水利科技,2017,15(5):43-49.
- [3] 侯一筠,尹宝树,管长龙,等.我国海洋动力灾害研究进展与展望[J].海洋与湖沼,2020,51(4):759-767.
- [4] 黄金池.中国风暴潮灾害研究综述[J].水利发展研究,2002,2(12):63-65.
- [5] 易笑园,余文韬,闫智超,等.几种台风风暴潮预报方法在实际预报中的运用及比较[J].海洋预报,2006,23(4):82-87.
- [6] 宗先国,刘志雨,陈翠英,等.渤海湾东南海岸风暴潮成因分析及预报[J].水文,2007,27(1):40-43.
- [7] 张娟,周水华,黄宝霞,等.神经网络在台风风暴潮模拟中的解释应用[J].海洋预报,2016,33(2):60-65.
- [8] Kim S Y, Matsumi Y, Shiozaki S, et al. A study of a real-time storm surge forecast system using a neural network at the Sanin Coast [C]//OCEANS,2012:2204-2210.
- [9] 薛彦广,沙文钰,徐海斌,等.神经网络在风暴潮增水预测中的应用[J].海洋预报,2005,22(2):33-37.
- [10] 王如云,雷磊,占飞.基于高低潮的风暴增水人工神经网络预报模型[J].海洋预报,2014,31(6):23-27.
- [11] 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.
- [12] 雷森,史振威,石天阳,等.基于递归神经网络的风暴潮增水预测[J].智能系统学报,2017,12(5):640-644.
- [13] Hochreiter S, Schmidhuber J. Long Short-Term Memory [J]. Neural Computation, 1997(9):1735-1780.
- [14] 庞古乾,黄玉明,何健,等.珠海市台风及其次生风暴潮特征分析与评估[J].热带气象学报,2020,36(1):42-50.
- [15] 朱元胜.上海防洪(潮)安全风险分析和管理的[J].水利学报,2002,33(8):21-28.

Characteristics analysis and tide level prediction of storm surge based on neural network model——Taking Chiwan and Nan'ao in Shenzhen City as an example

LIU Yuanyuan^{1,2}, LIU Yesen^{1,2}, ZHANG Li³, LI Mei³, MU Jie^{1,2}

(1. China Institute of Water Resources and Hydropower Research, Beijing 100038; 2. Research Center on Flood and Drought Disaster Reduction of the Ministry of Water Resources, Beijing 100038; 3. Shenzhen National Climate Observatory, Shenzhen 518040)

Abstract: Coastal cities of China are often affected by typhoons. Extreme rainfall brought by typhoons will cause serious flood and waterlogging if it encounters the high tide of storm surge. Before the arrival of typhoon, scientific and accurate analysis and prediction of storm tide level and timely and reasonable scheduling of projects are of great significance to reduce the risk of waterlogging disasters in coastal cities. Based on the analysis of the influencing factors of storm surges, the LSTM (Long-Short Term Memory) artificial neural network prediction models of Chiwan station and Nan'ao station are established. The results show that the influencing factors and pre time sequence n of storm tide level are different at different tide level stations. The error is the smallest by using the tide level and wind speed in the first 3 hours to predict the storm tide level at Nan'ao station, and predicted with the minimum error by using the tide level and wind speed in the first 6 hours at Chiwan station.

Keywords: Storm surge; Long Short-Term Memory; Intelligent prediction; Shenzhen

责任编辑 杜晓鹤