

基于LSTM-ResNet模型的定点有效波高预测

李自立, 蒙素素

(广西师范大学电子工程学院, 广西 桂林 541004)

摘要: 基于北部湾单站位浮标采集数据, 提出一种基于长短期记忆网络(LSTM)和残差网络(ResNet)相融合的网络模型, 将研究结果运用到短时波高预测中, 并将模型的数值预测结果与LSTM网络、反向传播(BP)网络和ResNet网络在短时波高预测中的数值计算结果进行对比分析。结果表明: 该模型在短时波高预测中, 预测结果偏差较小且实用性较高, 能够在一定条件下提高有效波高短期预测数值的有效性。

关键词: 北部湾; 波高预测; LSTM-ResNet网络; LSTM网络

中图分类号: P731.22 **文献标识码:** A **文章编号:** 1003-0239(2022)02-0080-06

1 引言

海洋波浪高度的短期预测对人类海洋活动的开展有着极为重要的意义^[1-2], 例如, 对于海上航行、海上施工、渔业捕捞和海上军事等活动, 波高短期预测精度越高, 海上活动相关状况的应对就会越充分。在海洋数值预测研究中, 有效波高的短期预测一直都是热点和难点问题^[3]。海浪有效波高预测方法主要分为经验预测法和数值模型预测法^[4]。经验预测法是通过历史观测得到的海浪数据按照时间先后进行外推, 进而得到相关海浪状态的预期数值^[5], 这种预测方法主要依靠历史数据和传统分析相结合, 在实用性和准确度方面都存在较大局限性。随着数学科学的发展, 海浪波高相关研究在数值建模新领域获得了一些具有较高研究价值的数值模型, 如SWAN (Simulating WAve Nearshore)、WAVEWATCH-III以及WAM (Wave Model)等。数值模型的结果是基于大量数据计算得到的, 对计算量和计算速度有着较高的要求, 在数据不是很充分的情况下, 其计算精度存在较大的局限。随着近年神经网络和深度学习技术的兴起, 人们尝试搭建

各种神经网络或混合模型来预测海浪高度。比如Makarynsky等^[6-7]以波高数据为输入搭建了神经网络; Tsai等^[8]将反向传播(Back Propagation, BP)神经网络技术应用于短期海浪预测和数据补充, 取得了较好的效果; Fan等^[9]将长短期记忆网络(Long Short-Term Memory, LSTM)应用于波高预测, 提高了有效波高的预报精度; Mandal等^[10]利用时域单变量建模和递归神经网络预测波高; Zhang等^[11]构造了受限波尔兹曼机-深度置信网络模型, 具有较好的波高短期预测能力。本文在前人研究的基础上, 把深度残差网络(Residual Network, ResNet)理论引入到长短期记忆网络技术中, 对北部湾特定海域的波浪数据进行建模研究, 尝试获得一个置信度较高的波高数值模型, 并依据此网络对有效波高进行短期预测分析, 通过多网络算法对比验证模型的有效性与预测的准确度。

2 数据准备

本实验研究数据由国家海洋局南海信息中心提供, 为2016—2018年北部湾海域定点浮标实测数

收稿日期: 2021-01-06; 修回日期: 2021-03-10。

基金项目: 国家自然科学基金项目(61661009)。

作者简介: 李自立(1979-), 男, 副教授, 博士, 主要从事海洋雷达信号的研究以及教学工作。E-mail: zlienishi@mailbox.gxnu.edu.cn

通讯作者: 蒙素素(1994-), 女, 硕士, 主要从事遥感信息处理研究。E-mail: 2840580225@qq.com

据,有效波高的最小采样间隔为1 h,数据的样本点总数为26 277个。为了能够准确地分析波高的数学特征和分布状况,我们对实测有效波高进行了特征分析,结果见表1。

表1 定点浮标有效波高数据值分析结果

Tab.1 The analysis results of the effective wave height data of the fixed-point buoy

数据清洗 前后	缺失值 个数	均值 /m	标准差 /m	最小值 /m	最大值 /m
原始波高值	242	0.98	0.76	0.00	19.2
清洗后波高值	0	0.98	0.75	0.00	6.00

表1中原始数据的缺失值为242个,缺失数据只占总值的9.21%,对整体数据分析的误差影响极小。为了批量数据分析的便捷,将缺失值用前向填充法对其进行插值处理。经过数据清洗后有效波高的数值统计特征如下:均值为0.98 m,标准差为0.75 m,最小值为0 m。当有效波高数值达到6 m及以上时,可定义为灾难性海浪^[12-13],此海况出现的几率小,有很强的偶然性,对长时的模型预测研究意义不大。为了方便研究有效波高的长时规律,本文将超过6 m的数值全部限幅为6 m。本文设定模型输入为当前时刻6 h前的波高数据,网络标签数据设置为预测当前时刻未来6 h的波高数据。实验时,选取总数据的80%为训练集,用于模型建模训练;剩余的20%为模型的测试集,用于模型预测效果测试分析。

3 基于LSTM-ResNet的模型原理

随着层数的增加,单一堆叠的LSTM网络模型会发生梯度消失和网络性能的退化问题,因此,本文引入ResNet模块接入LSTM层生成LSTM-ResNet模型。该模型由ResNet模块和LSTM模块组成。模型总体构图如图1所示。

图1中上方LSTM层中32表示隐藏层数量,整流线性单元(Rectified Linear Unit, ReLU)为激活函数,第一个箭头左边的数字代表当前输入6个时间刻度和1个变量,这里省略了Batch_Size维度。

3.1 长短期记忆网络

循环神经网络(Recurrent Neural Network,

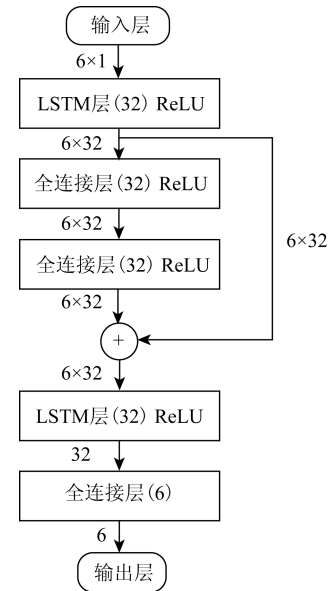


图1 LSTM-ResNet模型总体结构图

Fig.1 Overall structure of the LSTM-ResNet model

RNN)是一种节点定向连接成环的人工神经网络。与前馈神经网络(Feedforward Neural Network, FNN)不同的是,RNN可以利用它内部的记忆来处理任意时序的输入序列,这让其可以更容易处理如不分段的手写识别或语音识别等。然而,当连续时间数据的输入序列加长时,就会造成网络训练时梯度爆炸和梯度消失的问题,从而使RNN网络丧失长期记忆的功能。Hochreiter等^[14]在1997年提出LSTM模型,通过设计门结构来避免梯度消失等问题,门结构使得LSTM单元可以保存和获取长时间周期的上下文信息。LSTM单元结构图如图2所示^[9]。

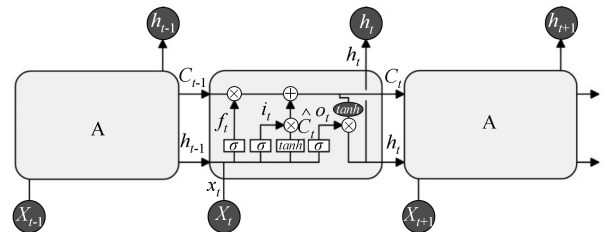


图2 LSTM单元的原理图

Fig.2 Schematic diagram of the LSTM unit

图2中各参数的计算公式如下:

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

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_i = \sigma(W_o \cdot [h_{t-1}, x_i] + b_o) \quad (3)$$

$$\hat{C}_i = \tanh(W_c \cdot [h_{t-1}, x_i] + b_c) \quad (4)$$

$$C_i = f_i C_{i-1} + i_i \cdot \hat{C}_i \quad (5)$$

$$h_i = o_i \cdot \tanh(C_i) \quad (6)$$

式中, i_i 、 f_i 和 o_i 分别为输入门、遗忘门和输出门; x_i 为当前时刻的输入特征; W_i 、 W_f 、 W_o 和 W_c 为待训练参数矩阵; b_i 、 b_f 、 b_o 和 b_c 为训练偏置项; σ 表示激活函数; h_{t-1} 为隐藏层前一时刻的输出。

式(1)~(3)经过 Sigmoid 激活函数使得 3 个门的大小在 0~1 之间; 式(4)是当前信息和前一步信息归纳形成的新记忆 \hat{C}_i , 从而实现信息的更新; 式(5)通过遗忘门去忘记一些旧信息再同时加入一些新信息; 式(6)主要是把过滤好的信息输出, 当有多层循环网络时, 当前层网络的输入 x_i 是前一层网络提取出来的有用信息 h_i 。

3.2 残差网络

残差网络作为 2015 年 ImageNet 竞赛 (Image Net Large-Scale Visual Recognition Challenge) 的冠军, 其 top5 错误率 (预测的前 5 个类别中不包含正确类别的比例) 为 3.57%。VGGNet (Visual Geometry Group Net) 和 GoogLeNet (Google Inception Net) 的成功, 说明网络越深, 模型表现越良好。然而, 单纯堆叠神经网络层数会使网络模型退化, 导致后面的特征丢失了前边特征的原本模样, 产生梯度消失或者梯度爆炸等问题, 造成训练和测试效果变差。为了解决模型“退化”的问题, 崔文植^[15]提出了残差结构 (见图 3)。

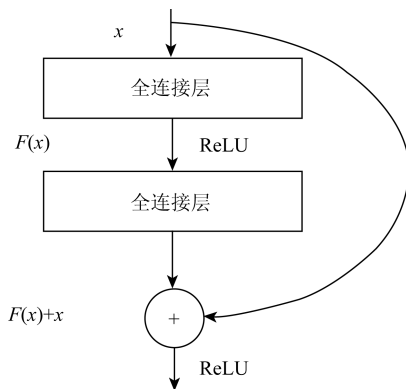


图 3 残差网络的核心结构

Fig.3 The core structure of the residual network

图 3 是 ResNet 的核心结构图。如图所示, ResNet 加入了一根跳连线后, 其输入可以直接短连接到非线性层的输出上, 输出结果为直接连接过来的恒等映射 x 和非线性输出 $F(x)$ 两路值元素的对应相加。这个步骤能有效缓解神经网络模型堆叠导致的梯度消失问题。

4 仿真和讨论

4.1 评估指标

为了评价不同神经网络的预测效果, 本文采用均方根误差 (Root Mean Square Error, RMSE)、平均绝对误差 (Mean Absolute Error, MAE)、均方误差 (Mean Squared Error, MSE) 和相关系数 R^2 来反映预测效果。公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - u_i)^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |o_i - u_i| \quad (8)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n |o_i - u_i|^2 \quad (9)$$

$$R^2 = \frac{\frac{1}{n} \sum_{i=1}^n (o_i - \bar{o})(u_i - \bar{u})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - \bar{o})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (u_i - \bar{u})^2}} \quad (10)$$

式中, o_i 为预测值; u_i 为观测值; \bar{o} 为预测值的平均值; \bar{u} 为观测值的平均值; n 为样本数量。

根据 MAE、MSE、RMSE 和 R^2 的结果来判断模型的好坏, 结果的取值范围均为 [0, 1]。前 3 个评价指标的值越接近 0, 说明模型拟合效果越好; R^2 越接近 1, 表示模型拟合效果越好。

4.2 有效波高的短期时刻预测结果

为了验证模型的预测性能, 采用表 1 清洗过的有效波高序列为研究对象, 将其输入 LSTM-ResNet 模型、LSTM 模型和 ResNet 模型以及目前应用最广泛的 BP 模型进行对比试验。4 种模型的隐藏层数量均设置为 32 个, 且都采用 ReLU 激活函数。

4.3 结果分析

表2列出了4种算法的有效波高短期预测结果,最优结果以粗体显示。数值结果表明LSTM-ResNet网络的短期预测效果最优,在1 h的预测中MAE低至0.08 m, R^2 高至0.96。ResNet网络、LSTM网络和BP网络的总体预测效果与实测数据基本一致,但是预测精度没有LSTM-ResNet网络高。随着预测时间的逐渐加长,数据的有效性不断降低,4种算法的MAE数值逐渐增大, R^2 数值不断减小,LSTM-ResNet网络的性能衰减程度相对较慢,尽管MAE最高达到了0.23 m, R^2 最多下降至0.70,但是预测结果还是处于可以接受的精度范围内。

图4为浮标实测值和4种算法的有效波高1 h预测数值对比图,图中的曲线变化对比表明LSTM-ResNet网络在预测波高的变化趋势和数值稳定性方面都表现出优于其他算法的效果。基于LSTM网络的周期性结构和门机制,本文设计的网络在信息选取上具有优势,相较于其他算法能够获得最优信息;在加入ResNet网络并对网络进行层数加深设置后,网络计算的数值精度得到较大的提高,综合以

上算法优点取得了波高预测与浮标数据拟合度较好的结果。与之对比的LSTM网络、BP网络和ResNet网络,预测结果与浮标数值虽然也具有一定的一致性,但是在某些时刻或者时间段上出现了预测值相对滞后的情况。

图5为2018年5月6日3:00—9月8日9:00基于LSTM-ResNet网络的有效波高间隔1 h和6 h的数值预测对比图。相对于6 h的预测结果,1 h的预测结果在数值上与浮标测量值更贴近。虽然6 h的预测结果偏差略大,但总体预测的曲线变化趋势与浮标测量值基本一致,进一步验证了LSTM-ResNet网络在短间隔的有效波高预测方面的有效性。

5 结论

本文运用神经网络的相关模型方法,对北部湾海洋有效波高的变化规律进行了数值建模,并进行了数值预测研究与结果对比分析。将4种网络模型方法的波高数据建模预测数值结果进行对比,结果表明本文采用的LSTM-ResNet网络模型在短时有效波高预测中能够取得比较好的结果。在浮标实测数值的结果对比分析中,采用4个统计指标对预

表2 短期预测的性能结果比较

Tab.2 Comparison of performance results for short-term forecasting

算法	时间跨度	平均绝对误差/m	均方根误差/m	均方误差/m	相关系数 R^2
LSTM+ResNet	1 h	0.081 8	0.120 3	0.014 4	0.960 4
LSTM		0.142 1	0.198 4	0.039 4	0.892 4
BP		0.165 4	0.219 4	0.048 2	0.868 4
ResNet		0.152 5	0.216 0	0.046 6	0.872 4
LSTM+ResNet	2 h	0.118 9	0.171 0	0.029 2	0.920 0
LSTM		0.174 3	0.240 3	0.057 8	0.842 1
BP		0.191 7	0.266 1	0.070 8	0.806 4
ResNet		0.243 6	0.308 9	0.095 4	0.739 2
LSTM+ResNet	4 h	0.189 7	0.262 2	0.069 8	0.809 9
LSTM		0.227 9	0.308 3	0.095 0	0.740 6
BP		0.221 1	0.300 5	0.090 3	0.753 4
ResNet		0.283 6	0.389 2	0.151 5	0.586 4
LSTM+ResNet	6 h	0.238 9	0.325 5	0.105 9	0.707 5
LSTM		0.269 4	0.362 3	0.131 3	0.641 0
BP		0.322 9	0.407 6	0.166 2	0.546 7
ResNet		0.333 8	0.451 1	0.203 5	0.444 7

注:粗体表示最优结果

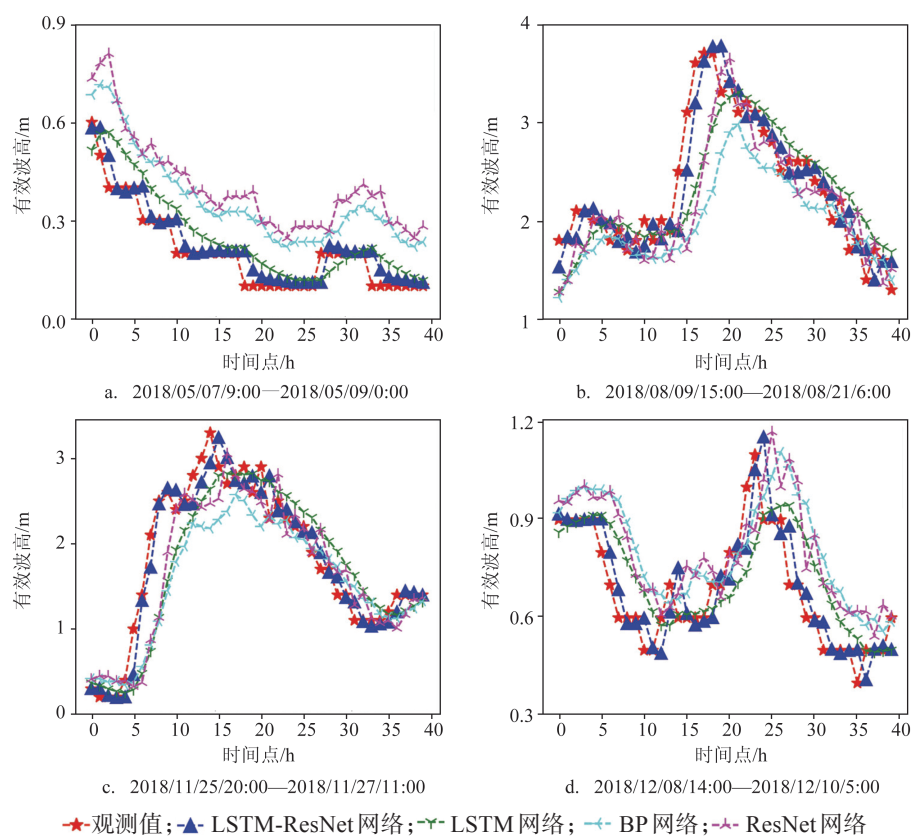


图4 1 h时间范围内观测结果与算法结果的比较图

Fig.4 Comparison of observation results and algorithm results in the time range of 1 hour

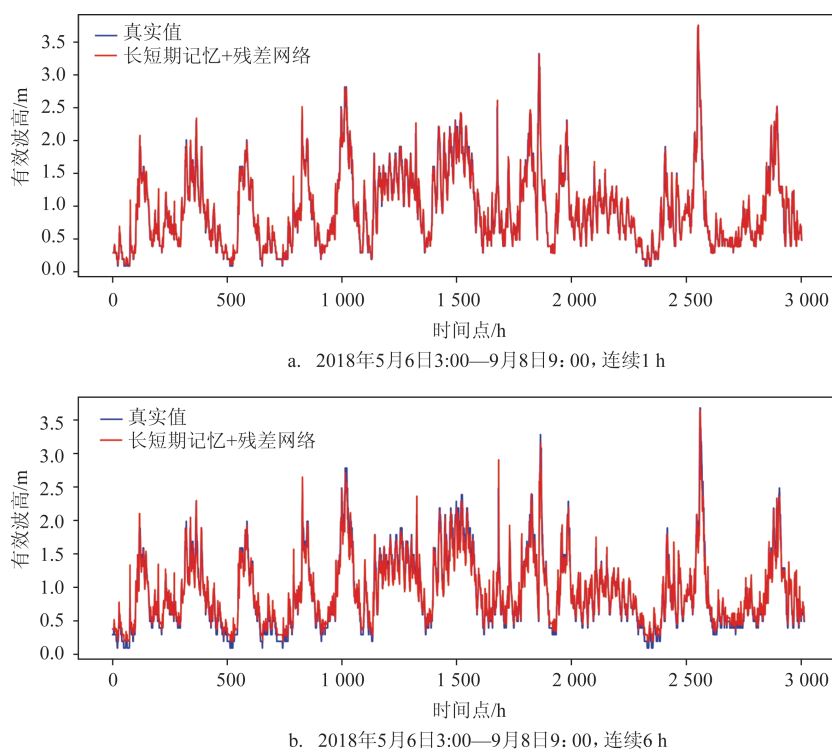


图5 预测值与观测值对比图

Fig.5 Comparison of predicted and observed values

测结果进行综合评价,本文采用的LSTM-ResNet网络预测结果与实测数值的一致性最好,预测的时间容许范围最大,相较于其余对比方法,在短时有效波高预测上具有一定的实用性,因此在海洋数值预测中具有一定的应用潜力。

参考文献:

- [1] Duan W Y, Han Y, Huang L M, et al. A hybrid EMD-SVR model for the short-term prediction of significant wave height[J]. Ocean Engineering, 2016, 124: 54-73.
- [2] Altunkaynak A, Özger M. Temporal significant wave height estimation from wind speed by perceptron Kalman filtering[J]. Ocean Engineering, 2004, 31(10): 1245-1255.
- [3] 陈希, 沙文钰, 李妍, 等. 人工神经网络技术在海浪预报中的应用[J]. 海洋通报, 2002, 21(2): 11-15.
- Chen X, Sha W Y, Li Y, et al. Application of the artificial neural network in the sea wave forecast[J]. Marine Science Bulletin, 2002, 21(2): 11-15.
- [4] Meng L, He Y J, Chen J N, et al. Neural network retrieval of ocean surface parameters from SSM/I data[J]. Monthly Weather Review, 2007, 135(2): 586-597.
- [5] 孟雷, 闻斌, 姜洪峰, 等. 神经网络方法对海浪有效波高数值模拟的改进[J]. 海洋预报, 2010, 27(2): 8-14.
- Meng L, Wen B, Jiang H F, et al. Neural network method to numerical simulation of significant wave height improvements[J]. Marine Forecasts, 2010, 27(2): 8-14.
- [6] Makarynskyy O. Improving wave predictions with artificial neural networks[J]. Ocean Engineering, 2004, 31(5-6): 709-724.
- [7] Londhe S N, Panchang V. One-day wave forecasts based on artificial neural networks[J]. Journal of Atmospheric and Oceanic Technology, 2006, 23(11): 1593-1603.
- [8] Tsai C P, Lin C, Shen J N, et al. Neural network for wave forecasting among multi-stations[J]. Ocean Engineering, 2002, 29(13): 1683-1695.
- [9] 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.
- [10] Mandal S, Prabakaran N. Ocean wave forecasting using recurrent neural networks[J]. Ocean Engineering, 2006, 33(10): 1401-1410.
- [11] Zhang X N, Dai H. Significant wave height prediction with the CRBM-DBN model[J]. Journal of Atmospheric and Oceanic Technology, 2019, 36(3): 333-351.
- [12] 许小峰, 顾建峰, 李永平. 海洋气象灾害[M]. 北京: 气象出版社, 2009: 79-82.
- Xu X F, Gu J F, Li Y P. Marine meteorological disaster[M]. Beijing: China Meteorological Press, 2009: 79-82.
- [13] 周媛媛, 周林, 关皓, 等. 基于浮标资料的中国东部海域最大波高特征分析[J]. 海洋预报, 2019, 36(2): 21-29.
- Zhou Y Y, Zhou L, Guan H, et al. Characteristic analysis of maximum wave height in the eastern China Sea based on buoy data[J]. Marine Forecasts, 2019, 36(2): 21-29.
- [14] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [15] 崔文植. 基于长短期记忆与残差网络的航班延误预测[D]. 杭州: 杭州师范大学, 2019.
- Cui W Z. Flight delay prediction based on long short-term memory and residual network[D]. Hangzhou: Hangzhou Normal University, 2019.

Prediction of significant wave height of a fixed point based on LSTM-ResNet model

LI Zili, MENG Susu

(College of Electronic Engineering, Guangxi Normal University, Guilin 541004 China)

Abstract: Based on the data collected by single-station buoys in the Beibu Gulf, this paper proposes a network model using the fusion of Long Short-Term Memory (LSTM) and Residual Network (ResNet), and applies the research results to short-term wave height forecasting; Thereafter, the numerical prediction results of the model are compared with the numerical calculation results of LSTM network, Back Propagation (BP) network and ResNet in the prediction of short-term wave height. Finally, the research results show that the LSTM-ResNet model have the characteristics of small deviation and high practicability in predicting the short-term wave height, and could improve the effectiveness of the short-term prediction value of the significant wave height under certain conditions.

Key words: Beibu Gulf; wave height prediction; LSTM-ResNet; LSTM network