

基于TSA-BP模型的温州站台风风暴潮增水预测

周寅杰, 刘 强, 张晓琪

(中国海洋大学 工程学院, 山东 青岛 266100)

摘 要: 台风风暴潮灾害通常会对沿海地区造成巨大损失, 因此, 准确预测台风风暴潮增水对沿海地区的防灾、减灾工作具有现实意义。本文根据现有风暴潮增水预测研究的成果, 建立了基于被囊群算法 (tunicate swarm algorithm) 优化的 BP 神经网络模型, 将该模型应用于台风风暴潮增水预测研究中。本文选取影响温州验潮站的 3 个台风作为研究对象, 收集并建立了 3 个台风影响验潮站过程的 129 个逐时数据样本。利用新模型对温州站进行风暴潮增水预测, 结果表明, 该模型与 BP 神经网络相比克服了陷入局部最优解的缺陷, 与粒子群优化的 BP 神经网络模型相比, 提升了模型收敛速度, 具有更好的预测精度及稳定性。

关键词: 风暴潮增水; 预测; BP 神经网络; 被囊群算法 (TSA)

中图分类号: X43

文献标识码: A

文章编号: 1007-6336(2022)05-0807-06

Prediction of typhoon storm surge at Wenzhou station based on TSA-BP model

ZHOU Yin-jie, LIU Qiang, ZHANG Xiao-qi

(College of Engineering, Ocean University of China, Qingdao 266100, China)

Abstract: Typhoon storm surge disasters usually cause huge losses to coastal areas. Therefore, accurate prediction of typhoon storm surge is of practical significance for disaster prevention and mitigation in coastal areas. In this paper, a BP neural network model based on the optimization of the tunicate swarm algorithm is established based on the results of existing storm surge prediction studies, and the model is applied to the study of typhoon storm surge prediction. In this paper, 129 time-by-time data samples of typhoons affecting the Wenzhou tide station were collected and established by selecting three typhoons affecting the Wenzhou tide station as the research object. The results show that the new model overcomes the defect of BP neural network falling into local optimal solutions, and improves the convergence speed of BP neural network model based on the optimization of particle swarm optimization algorithm. The TSA-BP model performs better prediction accuracy and stability.

Key words: storm surge; prediction; BP neural network; tunicate swarm algorithm

台风风暴潮是海平面受台风剧烈的大气扰动而出现的潮位异常增水现象^[1]。若风暴潮与天文潮耦合相互作用, 将造成海水水位暴涨, 引发特大潮灾。合理、快速地预测风暴潮增水, 对

沿海地区的防灾、减灾工作具有实用价值。

台风风暴潮增水预测方法可归结为以下两类: 经验预报法与数值预报法。经验预报法包括预报人员的主观经验预报与经验统计预报, 前者

收稿日期: 2021-12-23, 修订日期: 2022-03-03

基金项目: 国家自然科学基金项目 (41072176, 41371496)

作者简介: 周寅杰 (1998—), 男, 浙江宁波人, 硕士研究生, 主要从事海洋灾害风险管理, E-mail: zhouyinjiegreg@foxmail.com

通讯作者: 刘 强 (1961—), 男, 山东青岛人, 教授, 主要从事海洋灾害风险管理, E-mail: liuqiang@ouc.edu.cn

多利用预报人员的工作经验进行增水预报,后者主要运用数理统计方法,建立预测区域气象因子(如风、气压等)与风暴潮的经验函数关系进行增水预报^[2]。曾德美^[3]研究了青岛港多年的实测统计数据,提出了该港风暴潮极值的增、减水预报公式;高清清等^[4]对南通沿海台风风暴潮进行归类,分析实测统计数据后,提出该区域各类别台风风暴潮的增水预报公式。

数值预报法作为当前主流的台风风暴潮增水预测方法,主要采取数值模拟的形式。周才扬等^[5]采用 ERA-Interim 背景风场与 FVCOM 模式建立了长江口三维风暴潮数学模型,探究该区域台风风速、强度等影响因素与风暴潮增水的关系;梁连松等^[6]使用 ADCIRC 模式对温州鳌江站风暴潮增水进行数值模拟,探究台风“1808 玛莉亚”移速变化对风暴潮增水的影响。

人工智能的发展将机器学习方法引入风暴潮增水预测领域中,以台风相关因子作为机器学习的输入指标进行增水预测,为风暴潮增水预测提供了新的辅助手段。Lee 等^[7]选取台风风速、风向、气压及天文潮作为 BP 神经网络的输入,验证了神经网络应用于风暴潮增水预测的可行性。You 等^[8]将聚类神经网络应用于韩国多个沿海区域的风暴潮增水预测;Rajasekaran 等^[9]利用支持向量机进行预报,可以有效地预测风暴潮的偏差。刘媛媛等^[10]将气象因素数据与前时序潮位数据作为网络输入,建立了以 LSTM 神经网络为基础的风暴潮临近预报模型;雷森等^[11]使用递归神经网络挖掘风暴潮数据间的时序关联特性,减小了神经网络应用于风暴潮增水预测的误差。薛明等^[12]对比 BP 神经网络、小波神经网络及递归神经网络于广东阳江站风暴潮增水的预测应用,提出三种神经网络的预测范围。

上述学者的研究成果为台风风暴潮增水预测提供了丰富的理论基础,但采用的神经网络多属于基础算法。为进一步提高神经网络在台风风暴潮增水预测中的准确性,本文以机器学习理论为基础,提出了基于被囊群算法改进的 BP 神经网络模型。训练结果表明,相较于粒子群优化算法,该模型对风暴潮增水预测的准确性有较大提升,同时在稳定性方面表现优异。

1 材料与方法

1.1 数据来源与变量选取

由于不同的地理位置受台风的影响也有明显的差异,因此,为控制验潮站地理位置这一重要因素,本文选取温州验潮站作为研究对象,并对其在台风风暴潮影响期间的增水进行预测。基于对近年台风路径的分析与判断,选取位于温州验潮站(28.03°N, 120.65°E)东北侧,路径为西北走向的3个台风“1909 利奇马”“1918 米娜”及“2004 黑格比”作为研究对象。台风数据及温州验潮站的水位数据主要源于国家海洋科学数据中心、中央气象台台风网、中国海洋信息网。风暴潮增水受多种因素影响,各因素的影响效果也不尽相同,学者们^[7-8]多采用台风路径经纬度、台风中心气压、台风移速、海平面压力等因素作为相关模型的输入,以进行风暴潮增水的预测。也有学者^[6]提出,风暴潮的增水与最大风圈半径、台风路径与验潮站的夹角等因素的密切相关。

基于数据的可获取性与可利用性,本文选取台风中心的经度、纬度、最低气压、最大风速、移动速度、最大风圈半径、验潮站与台风中心的直线距离、验潮站与台风的夹角、该时刻的增水值作为输入变量,将下一时刻的增水值作为输出变量。

1.2 数据处理

针对台风数据的时间密度与增水数据的时间密度不一致的问题,鉴于三次样条插值法对增水数据有良好的拟合性与数据精细化的作用^[13],因此,本文采用该方法对数据进行加密,以获得时序间距为1h的增水数据。由于目前没有标准方法将风暴潮增水从潮位数据中提取出来,本文选取变分模态分解方法^[14](自动准正交)进行提取,相较于经验模态分解方法^[15]的提取形式,该方法可以改善经验模态分解方法在潮汐数据分析与提取方面的模态混叠问题^[16]。验潮站与台风的夹角、验潮站与台风中心的直线距离通过验潮站经纬度和台风中心经纬度进行测算,验潮站与台风中心的距离采用 haversin 函数进行测算,其计算公式如下:

$$\text{haversin}\left(\frac{d}{R}\right) = \text{haversin}(\text{lat}2 - \text{lat}1) + \cos(\text{lat}1) \cos(\text{lat}2) \text{haversin}(\text{lng}2 - \text{lng}1) \quad (1)$$

式中: R 表示地球半径; $\text{lat}1$ 、 $\text{lat}2$ 分别表示验潮站与台风中心的纬度; $\text{lng}1$ 、 $\text{lng}2$ 分别表示验潮站与台风中心的经度。

鉴于所采用的数据计量单位差异较大,为保证不同类别数据间的可比性,在输入前将各项变量归一化处理至 $[0,1]$, 公式如下:

$$y_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (2)$$

式中: y_i 表示输出的归一化数据; x_i 表示初始数据; $\min(x_i)$ 、 $\max(x_i)$ 分别表示初始数据的最小值与最大值。

1.3 BP 神经网络原理

BP 神经网络(back propagation neural network, BPNN)由 Rumelhart 等^[17]学者于 1986 年提出,是具有误差反向传播特性的前向型神经网络学习算法,其发明灵感来源于人脑神经的拓扑结构。该神经网络以输入层、隐含层、输出层为基本结构,由信息前向输入与误差反向传播两个过程实现最优结果的目标,图 1 为一个 3 层的 BP 神经网络结构示意图。

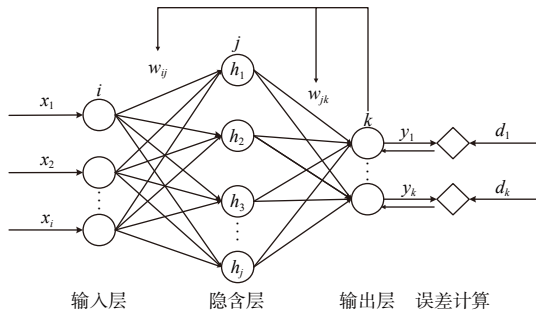


图 1 BP 神经网络结构示意图

Fig. 1 BP neural network structure diagram

信息前向输入指信息 $(x_1, \dots, x_i)$ 由输入层接收后通过隐含层神经元多次加权与函数激活于输出层获取输出值的过程。由于前向过程中的权值 (w) 与阈值均为随机初始化,输出值 $(y_1, \dots, y_k)$ 通常与真实值间误差 $(d_1, \dots, d_k)$ 较大,为保证参数将信息利用最大化, BP 神经网络采用误差反向传播以减小误差。误差反向传播指神经网络以减小误差为目标,将误差由输出层反向传导通过隐含层后反馈到输出层,将隐含层中的

权值与阈值基于梯度最小原则进行调整。上述过程是一个循环往复的过程,当其使神经网络的输出误差低于试验设定标准时,停止调整。

1.4 TSA 原理

被囊群算法(tunicate swarm algorithm, TSA)是 Kaur 等^[18]于 2020 年根据深海被囊生物群体觅食规律提出的一种元启发式算法,其生物原理为:被囊生物群体在一定空间范围内呈分散状分布,其个体间具有位置依赖性。个体在觅食过程中通过搜寻离食物源位置最近的同伴并以喷气推进的形式向该同伴运动,经过个体位置的不断更迭,最后所有被囊生物会共同群聚在食物源周围。TSA 算法与粒子群(particle swarm optimization, PSO)算法等智能优化算法类似,摆脱了特定函数限制且采用无梯度的优化技术,加速寻优速度。同时,作为一种以群体智能算法,其具有良好的空间搜索能力和全局最优解获取能力。

TSA 算法主要从被囊生物群体的个体喷气推进行为与生物群体聚集行为进行设计,建模步骤如下。

(1) 个体喷气推进行为

①为防止喷气过程中与其他个体相互碰撞,被囊生物会对其他个体位置进行判断,算法中使用这一位置向量进行表示:

$$\vec{A} = \frac{\vec{G}}{\vec{M}} \quad (3)$$

$$\vec{G} = c_2 + c_3 - \vec{F} \quad (4)$$

$$\vec{F} = 2 \cdot c_1 \quad (5)$$

$$\vec{M} = [P_{\min} + c_1 \cdot (P_{\max} - P_{\min})] \quad (6)$$

式中: $\vec{G}$ 表示被囊生物所受重力; $\vec{F}$ 表示深海中的水平流力; c_1 、 c_2 、 c_3 均为 $[0,1]$ 内的随机数组; $\vec{M}$ 表示被囊生物个体间相互作用力; $P_{\min}$ 与 $P_{\max}$ 分别表示被囊生物个体相互作用的初始速度与从属速度。

②保证个体间避免碰撞后,被囊生物将规划向位置最优邻居进行喷气推进移动,此过程根据个体与食物源之间的距离 $\vec{PD}$ 表进行判断:

$$\vec{PD} = |\vec{FS} - r_{\text{and}} \cdot \vec{P_p}(x)| \quad (7)$$

式中: $\vec{FS}$ 表示食物源所处位置; $\vec{P_p}(x)$ 表示被

囊个体当前所处位置; r_{and} 为 $[0,1]$ 内的随机数。

③待个体确定位置最优邻居后, 将开始向相关位置喷气推进, 最终到达该个体此次迭代的最优位置, $\overrightarrow{P_p(x)}$ 也会在此时进行更新。

$$\overrightarrow{P_p(x)} = \begin{cases} \overrightarrow{FS} + \overrightarrow{A} \cdot \overrightarrow{PD}, & \text{if } r_{and} \geq 0.5 \\ \overrightarrow{FS} - \overrightarrow{A} \cdot \overrightarrow{PD}, & \text{if } r_{and} < 0.5 \end{cases} \quad (8)$$

(2) 生物群体行为

待个体喷气推进行为满足避免碰撞、最优移动方向确定及移动至个体最优邻居 3 个条件后, 所有个体将以群聚形式共同向食物源进发, 其最终更新位置将依据个体预更新位置 $\overrightarrow{P_p(x+1)}$ 及前一次迭代 $\overrightarrow{P_p(x)}$ 的更新位置进行确定:

$$\overrightarrow{P_p(x+1)} = \frac{\overrightarrow{P_p(x)} + \overrightarrow{P_p(x+1)}}{2 + c_1} \quad (9)$$

1.5 TSA-BP 神经网络模型建立

BP 神经网络对初始权值与阈值具有较强的敏感性, 但其初始权值与阈值以随机形式进行确定易引发神经网络陷入局部最优解的困境。而 TSA 算法利用生物群体寻优的原理, 经过多次迭代将被囊生物群体确定至最优位置, 可发挥 TSA 算法深入探索搜索空间与返回全局最优解的能力。因此, 本文将 TSA 算法引入 BP 神经网络中, 寻求最优初始权值与阈值, 减少 BP 神经网络陷入局部最优解的可能性, 增强 BP 神经网络的稳定性。模型建立步骤如下。

(1) 确立 BP 神经网络的基本结构, 根据数据建立以 m 个输入层神经元, h 个隐含层神经元及输出层神经元节点数 n 的三层拓扑网络结构, 故可确定搜索空间维度 $k = m * h + h * n + h + n$ 。

(2) 根据搜索空间维度 k 初始化被囊种群位置 $\overrightarrow{P_p}$, P 选择 TSA 初始参数及最大迭代次数。

(3) 进行 TSA 算法训练, 执行被囊生物喷气行为, 根据式(3)与(7)保证个体在喷气过程中的避免碰撞及最优邻居确定, 并参照式(8)对个体位置 $\overrightarrow{P_p(x)}$ 进行迭代更新。

(4) 待被囊生物喷气行为执行完毕, 开始执行被囊生物群体行为, 根据式(9)计算出每次群体行为结束后的被囊个体的位置。

(5) 计算适应度函数值, 若满足算法结束条件则输出 TSA 算法优化的权值和阈值, 若不满足算法结束条件则重新从第(3)步进行。

(6) 将得到的群体最优解赋值于 BP 神经网络的初始权值与阈值。

2 结果与讨论

本文以“1909 利奇马”“1918 米娜”及“2004 黑格比”3 个台风的 129 组风暴潮增水样本作为研究总样本, 训练集与测试集比例为 2:1。本文将总样本中每 3 h 的前 2 h 样本作为训练集, 后 1 h 样本作为测试集, 此设置不仅体现风暴潮的时序特征, 还避免了数据来源及趋势对试验结果产生影响。本试验在 MATLAB2019b 平台进行。

试验中, 根据输入变量与输出变量的选取数量, 建立了一个 9 节点输入, 单输出的 TSA-BP 神经网络模型。该模型隐含层为单隐含层, 根据经验公式 $h = \sqrt{m+n} + a$ (其中 $a=1,2,\dots,10$) 确定隐含层神经元数量 h 可取范围为 $[5,14]$, 通过比较该范围内 h 取整时的模型最优状况, 确定隐含层神经元的个数为 14, 故建立的 TSA-BP 神经网络为 9-14-1 基本结构形式, 其空间搜索维度为 155。鉴于 TSA 算法的最优参数选取方法现阶段尚未被提出, 本文通过控制变量法确定被囊个体相互影响初始速度 $P_{\min}=1$, 被囊个体相互影响从属速度 $P_{\max}=6$, 种群数量 $n=70$ 及最大迭代次数 $t=40$ 。

为有效分析试验结果, 本文选用均方误差 (MSE) 与相关系数 (R^2) 两个评价指标对 TSA-BP 模型进行评价, 其公式如下:

$$MSE = \frac{1}{m} \sum_{i=1}^m (x_i - \hat{x}_i)^2 \quad (10)$$

$$R^2 = \frac{\left(n \sum_{i=1}^n \hat{x}_i x_i - \sum_{i=1}^n \hat{x}_i \sum_{i=1}^n x_i \right)^2}{\left[n \sum_{i=1}^n \hat{x}_i^2 - \left(\sum_{i=1}^n \hat{x}_i \right)^2 \right] \left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right]} \quad (11)$$

式中: $\hat{x}_i$ 表示样本 i 的风暴潮增水预测值; x_i 表示样本 i 的风暴潮增水真实值; 样本总数为 n 。MSE 越小, 表明预测值和真实值的差距越小, 预测模型表现越好; R^2 范围在 $[0,1]$, 数值越趋近于 1, 表明模型性能越好。

为对 TSA-BP 神经网络预测模型在风暴潮增水预测的表现进行合理评判, 本文将该模型

与 PSO-BP 神经网络预测模型及 BP 神经网络预测模型进行预测准确性与稳定性对比, 对比的神经网络模型均采用模型运行最优参数。

图 2 为 TSA-BP 神经网络模型、PSO-BP 神经网络模型与 BP 神经网络模型测试集的拟合结果, 由图 2 可知, BP 模型的整体拟合效果较差, PSO-BP 拟合效果良好, TSA-BP 模型拟合效果最佳。在部分极端点位置 TSA-BP 模型表现优于其他模型, 如样本 16, 23, 32, 尤其在极端样本点 32 优势明显。这表明, 经优化的神经网络模型具有更好的拟合表现, 且 TSA-BP 神经网络模型在风暴潮增水预测的各个样本中都表现出优势。

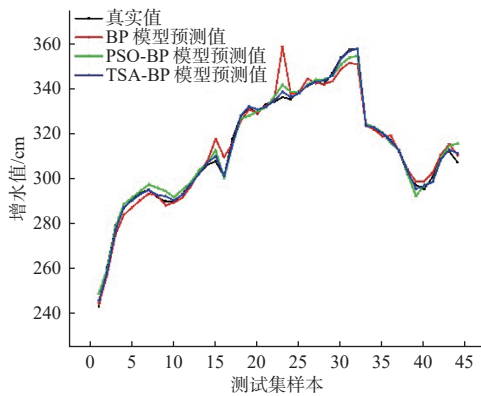


图 2 三种预测模型试验预测结果对比

Fig. 2 Comparison of prediction results of three models

图 3 表示 3 种模型在 100 次预测过程的均

方误差图, 均方误差图反映了不同模型的预测结果稳定性。由图 3 可知, BP 神经网络模型在 100 次预测中均方误差 MSE 多处于 $15 \sim 50 \text{ cm}^2$, 跨度区间大且波动性强; PSO-BP 神经网络在 100 次预测过程中均方误差 MSE 多处于 $8 \sim 20 \text{ cm}^2$, MSE 总体较 BP 减少但波动性仍较强; TSA-BP 神经网络在 100 次预测过程中均方误差 MSE 多处于 $2 \sim 10 \text{ cm}^2$, 预测结果明显优于前二者, 且波动性最小。以上表明, TSA-BP 神经网络模型在风暴潮增水预测中具有良好的稳定性。

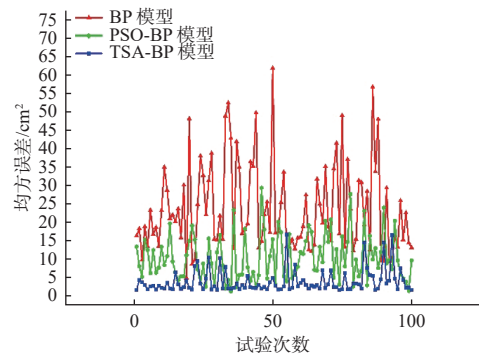


图 3 三种模型 100 次试验预测均方误差对比

Fig. 3 comparison diagram of prediction MSE of 100 tests of three models

本文从均方误差 (MSE)、相关系数 (R^2) 及单次模型运算时间 3 项评判指标对模型的预测效果进行描述, 结果见表 1。

表 1 不同模型效果对比

Tab.1 Comparison of different model effects

算法名称	MSE / cm^2		R^2		单次模型运算时长/s
	训练集	测试集	训练集	测试集	
TSA-BP神经网络	3.69	3.94	0.9957	0.9917	82.75
PSO-BP神经网络	7.66	10.43	0.9923	0.9795	106.26
BP神经网络	14.90	23.67	0.9836	0.9520	2.30

对表 1 从训练集和测试集的 MSE 、 R^2 以及单次模型运行时长来对预测模型的表现进行分析。从 MSE 角度分析, 误差大小排序为 TSA-BP 模型 < PSO-BP 模型 < BP 模型, 相对于其他两个模型, TSA-BP 模型的训练集和测试集表现最优, 测试集为 3.94 cm^2 , 而 BP 模型的测试集误差高

达 23.67 cm^2 ; R^2 值表现与 MSE 表现一致, TSA-BP 模型训练集和测试集分别 0.9957 和 0.9917, 明显优于其他两种模型; 在单次模型运算时长上 TSA-BP 模型较 PSO-BP 算法快 23.51 s, 但由于 TSA-BP 模型优化过程需要寻找 BP 最优的权值与阈值, 故其较 BP 模型运算时间长。因此, 从

3种模型的训练精度、预测精度以及收敛速度上分析, TSA-BP模型均体现出良好的拟合性能和运算速度, 反映出 TSA-BP神经网络模型在风暴潮增水预测方面具有良好的适用性。

3 结论

本文从影响温州验潮站的台风中选取3个路径相似的台风为研究对象, 提取129个逐时样本数据进行风暴潮增水预测, 提出了基于被囊群搜索算法优化的BP神经网络风暴潮增水预测模型。试验证明, 经TSA算法优化后的BP神经网络, 可有效防止BP神经网络陷入局部最优解, 在训练精度、预测精度及预测结果的稳定性方面提升明显, 且较PSO优化的BP神经网络收敛速度快。TSA-BP神经网络模型在风暴潮增水预测方面具有一定适用性, 但由于影响台风风暴潮增水是一个持续的动态过程, 如何利用动态数据结合机器学习以提高模型的预测精度尚需进一步研究。

参考文献:

- [1] 冯士筭. 风暴潮导论[M]. 北京: 科学出版社, 1982.
- [2] 王喜年. 风暴潮预报知识讲座 第四讲 风暴潮预报技术(1)[J]. 海洋预报, 2001, 18(4): 63-69.
- [3] 曾德美. 青岛港风暴潮经验统计预报[J]. 海洋预报, 1992, 9(3): 66-73.
- [4] 高清清, 曹 兵, 高鑫鑫, 等. 南通沿海台风风暴潮分析及其经验预报初探[J]. 海洋预报, 2014, 31(1): 29-35.
- [5] 周才扬, 殷成团, 章卫胜, 等. 长江口登陆台风增水数值模拟——以“安比”为例[J]. 水利水运工程学报, 2021 (1): 70-77.
- [6] 梁连松, 张 钊, 顾冬明, 等. 典型路径下台风移速调整对鳌江站增水的数值分析[J]. 海洋预报, 2020, 37(5): 59-66.
- [7] LEE T L. Predictions of typhoon storm surge in Taiwan using artificial neural networks[J]. Advances in Engineering Software, 2009, 40(11): 1200-1206.
- [8] YOU S H, SEO J W. Storm surge prediction using an artificial neural network model and cluster analysis[J]. Natural Hazards, 2009, 51(1): 97-114.
- [9] RAJASEKARAN S, GAYATHRI S, LEE T L. Support vector regression methodology for storm surge predictions[J]. Ocean Engineering, 2008, 35(16): 1578-1587.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 曹艳艳, 姚日升. 样条插值在台风业务中的应用[C]//第26届中国气象学会年会热带气旋科学研讨会分会场论文集. 杭州: 中国气象学会, 2009.
- [14] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [15] 杨金湘, 袁方超, 李郅明, 等. EMD方法在中国沿岸风暴潮增水分析中的应用[J]. 海洋通报, 2016, 35(2): 157-162, 169.
- [16] 胡亮亮, 王 进, 刘 欢, 等. GNSS浮标海潮高滤波光滑算法[J]. 科学技术与工程, 2021, 21(27): 11463-11471.
- [17] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [18] KAUR S, AWASTHI L K, SANGAL A L, et al. Tunicate Swarm Algorithm: a new bio-inspired based metaheuristic paradigm for global optimization[J]. Engineering Applications of Artificial Intelligence, 2020, 90: 103541.