

DOI:10.11705/j.issn.1672-643X.2023.04.06

基于数据处理与若干群体算法优化的 GRU/LSTM 水质时间序列预测

杨坪宏¹, 胡奥², 崔东文³, 杨杰⁴

(1. 云南省水文水资源局, 云南 昆明 650106; 2. 临沧润汀水资源科技服务有限公司, 云南 临沧 677000; 3. 云南省文山州水务局, 云南 文山 663000; 4. 北京全路通信信号研究设计院集团有限公司昆明分公司, 云南 昆明 650041)

摘要: 为提高水质时间序列预测精度, 提出一种基于小波包变换(WPT)和变色龙优化算法(CSA)、猎豹优化(CO)算法和山瞪羚优化(MGO)算法的优化门限循环控制单元(GRU)、长短期记忆神经网络(LSTM)的预测模型。首先利用WPT对pH值、DO、COD_{Mn}、NH₃-N时间序列进行平稳化处理, 得到若干个规律性较强的子序列分量; 其次简要介绍了CSA、CO、MGO算法原理, 利用CSA、CO、MGO分别寻优GRU、LSTM超参数, 建立WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU、WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM模型; 最后利用所建立的模型对pH值及DO、COD_{Mn}、NH₃-N浓度各分量进行预测和重构, 并建立WPT-GRU、WPT-LSTM和WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM模型作对比分析模型, 以云南省昆明市观音山断面为例, 通过pH值及DO、COD_{Mn}、NH₃-N浓度预测对模型进行了验证。结果表明: WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU、WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM模型对实例pH值及DO、COD_{Mn}、NH₃-N浓度的预测精度优于其他对比模型, 具有较好的预测效果, 其中尤以WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU模型的预测精度最高; CSA、CO、MGO能有效调优GRU、LSTM超参数, 显著提高GRU、LSTM预测性能; 所构建的6种模型预测精度高且计算规模小, 是有效的水质时间序列预测模型, 可为相关水质预测研究提供参考。

关键词: 水质预测; 门限循环控制单元; 长短期记忆神经网络; 小波包变换; 变色龙优化算法; 猎豹优化算法; 山瞪羚优化算法

中图分类号: TV213.4

文献标识码: A

文章编号: 1672-643X(2023)04-0045-09

Prediction of GRU/LSTM water quality time series based on data processing and optimization of several swarm intelligence algorithms

YANG Pinghong¹, HU Ao², CUI Dongwen³, YANG Jie⁴

(1. Hydrology and Water Resources Bureau of Yunnan Province, Kunming 650106, China; 2. Lincang Runtong Water Resources Technology Service Co., Ltd., Lincang 677000, China; 3. Wenshan Prefecture Water Affairs Bureau of Yunnan Province, Wenshan 663000, China; 4. Kunming Branch of CRSC Research and Design Institute Group Co., Ltd., Kunming 650041, China)

Abstract: To improve the prediction accuracy of water quality time series, a type of prediction model based on wavelet packet transform (WPT), chameleon optimization algorithm (CSA), cheetah optimization (CO) algorithm and mountain gazelle optimization (MGO) algorithm is proposed to optimize the gated recurrent unit (GRU) and long short-term memory networks (LSTM). Firstly, the time series of pH, DO, COD_{Mn} and NH₃-N are stabilized by WPT, by which several regular subsequence components are obtained. Secondly, the CSA, CO and MGO algorithms are briefly introduced and then applied to the optimization of the hyper-parameters of GRU and LSTM, by which the WPT-CSA-GRU, WPT-CO-GRU, WPT-MGO-GRU, WPT-CSA-LSTM, WPT-CO-LSTM, WPT-MGO-LSTM models are estab-

收稿日期: 2022-11-27; 修回日期: 2023-02-27

基金项目: 国家重点研发计划项目(2021YFC3200903); 国家自然科学基金项目(51809288); 中国水利水电科学研究院基本科研业务费项目(WR0145B022021)

作者简介: 杨坪宏(1967—), 男, 云南兰坪人, 高级工程师, 主要从事水环境监测评价及水资源保护工作。

通讯作者: 崔东文(1978—), 男, 云南玉溪人, 学士, 正高级工程师, 主要从事水资源管理保护及智能算法在水文水资源系统中的应用研究等工作。

lished. Finally, the established models are used to predict and reconstruct the components of pH, DO, COD_{Mn}, NH₃-N, meanwhile, WPT-GRU, WPT-LSTM and WPT-CSA-SVM, WPT-CO-SVM, WPT-MGO-SVM models are established for comparative analysis. Then the models are verified by the prediction of pH, DO, COD_{Mn}, NH₃-N of Guanyinshan section in Kunming, Yunnan Province. The results show that the prediction accuracy of WPT-CSA-GRU, WPT-CO-GRU, WPT-MGO-GRU, WPT-CSA-LSTM, WPT-CO-LSTM, WPT-MGO-LSTM models for pH, DO, COD_{Mn}, NH₃-N of the case section are better than the comparison models with better prediction performance, of which WPT-CSA-GRU, WPT-CO-GRU, WPT-MGO-GRU models have the highest prediction accuracy; CSA, CO and MGO can effectively tune the hyper-parameters of GRU and LSTM, and significantly improve the prediction performance of GRU and LSTM; the six proposed models have high prediction accuracy and small calculation scale, which can be recognized as effective prediction models for water quality time series. This study can provide some reference for relevant water quality prediction research.

Key words: water quality prediction; gated recurrent unit (GRU); long short-term memory networks (LSTM); wavelet packet transform (WPT); chameleon swarm algorithm (CSA); cheetah optimization (CO) algorithm; mountain gazelle optimization (MGO) algorithm

1 研究背景

提高水质时间序列预测精度对于水环境治理和水生态保护具有重要意义^[1]。当前,除 MIKE 21 等水质模拟模型外,有关水质预测的方法可分为 3 类。第 1 类是以逐步回归为代表的传统预测方法;第 2 类是以机器学习为代表的预测方法,如 BP (back propagation) 神经网络^[2-3]、支持向量回归机^[4]、长短期记忆神经网络 (long short-term memory networks, LSTM)^[5-6]、非线性自回归神经网络^[7]等,此类方法提高了对非线性问题的处理能力,但权值阈值或关键参数的合理选取制约了其预测精度的提升,目前普遍采用群体智能算法进行寻优,并获得较好的预测效果。第 3 类是组合预测法,组合预测法分两种方式,一种是基于两个或两个以上预测器的组合预测,如 TimeGAN-CNN (convolutional neural networks)-LSTM 模型^[8]、GRNN (general regression neural network)-Markov 模型^[9]、STG (spatial-temporal graph)-LSTM 模型^[10]等,组合预测方法的特点是利用各模型的优势来克服单一预测模型的不足,但预测结果往往存在较大的不确定性;另一种是基于信号分解技术的预测方法,该方法利用分解技术将时序数据分解为若干更具规律的子序列分量,通过对子序列分量进行预测、重构后得到最终预测结果。由于基于信号分解技术的预测方法具有较好的预测精度,该方法已在水质时间序列预测研究中得到广泛应用,如孙铭等^[11]提出的小波分解-长短期记忆网络 (W-LSTM) 时间序列预测模型;金昌盛

等^[12]提出的奇异谱分析-遗传算法-BP 神经网络 (SSA-GA-BP) 预测模型;顾乾晖等^[13]提出的变分模态分解-粒子群优化算法-长短期记忆神经网络 (VMD-PSO-LSTM) 预测模型;李建文等^[14]提出的集合经验模态分解-支持向量回归机 (EEMD-SVR) 预测模型;白雯睿等^[15]提出的变分模态分解-卷积神经网络-长短期记忆神经网络 (VMD-CNN-LSTM) 组合水质预测模型。

近年来,随着大数据和人工智能的发展,LSTM、门控循环单元 (gated recurrent unit, GRU) 已被应用于水质预测^[15-16]的研究中。LSTM 是循环神经网络 (recurrent neural network, RNN) 的变种,其克服了 RNN 面临的梯度消失和梯度爆炸问题,在问题分类^[17]、预测研究^[18-21]等领域应用广泛。GRU 是 LSTM 的变种,与 LSTM 相比,GRU 拥有更低的复杂度以及更好的预测性能,被应用于各类预测研究^[22-24]之中。在实际应用中,GRU/LSTM 超参数 (包括层数、学习率、隐含层节点数、dropout、Learn-RateDropPeriod 等) 对 GRU/LSTM 预测性能具有较大影响,需要通过人工不断调整试错得到 GRU/LSTM 超参数,不仅要求操作人员具有丰富的调参经验,而且极易陷入局部极值,难以满足 GRU/LSTM 的预测需求。目前,海洋捕食者算法^[18]、粒子群优化算法等^[25-28]群体智能算法 (swarm intelligence algorithms, SIA) 已在寻优 GRU/LSTM 超参数中得到应用,并取得了较好的调优效果。

为提高 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时间序列的预测精度,改进 GRU/LSTM 的预测性能,拓展 SIA 在 GRU/LSTM 超参数调优中的应用范畴,本

文基于小波包变换(wavelet packet transform, WPT)分解技术和“分解算法+智能算法+预测模型+叠加重构”思想,研究提出采用变色龙优化算法(chameleon swarm algorithm, CSA)、猎豹优化(cheetah optimization, CO)算法、山瞪羚优化(mountain gazelle optimization, MGO)算法分别调优 GRU/LSTM 超参数的水质时间序列预测模型,并构建 WPT-GRU、WPT-LSTM 模型及基于 SVM 预测器的 WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM 模型进行对比分析,然后通过云南省昆明市观音山断面 pH 值、DO、COD_{Mn}、NH₃-N 预测实例对各模型进行检验。

2 数据来源与研究方法

2.1 数据来源

本文以云南省昆明市观音山断面 2004—2015 年逐周 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时间序列预测为例对各模型进行验证。原始 pH 值及 COD_{Mn}、DO、NH₃-N 浓度数据来源于原环境保护部数据中心的水质周报,共 624 组,对于个别缺失数据采用线性法进行插补,昆明市观音山断面水质时间序列 3D 图见图 1。从图 1 可以看出,该断面水质 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时序数据呈现出非线性、多尺度等特征,起伏变化激烈。

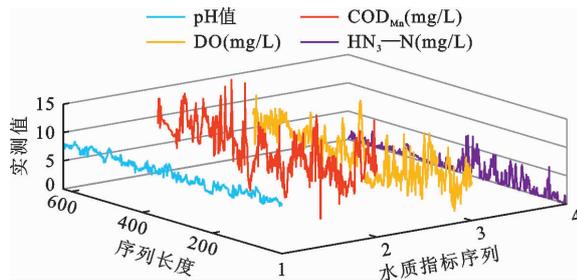


图1 昆明市观音山断面水质时间序列 3D 图

2.2 研究方法

2.2.1 小波包变换(WPT) WPT 能同时对信号低频部分和高频部分进行分解,更适用于水质时间序列分解。WPT 对 pH 值及 COD_{Mn}、DO、NH₃-N 浓度原始信号进行分解的公式为^[29-32]:

$$\begin{cases} d_i^{j,2n} = \sum_k h_{k-2l} d_k^{j-1,n} \\ d_i^{j,2n+1} = \sum_k g_{k-2l} d_k^{j-1,n} \end{cases} \quad (1)$$

重构算法为:

$$\bar{d}_i^{j,n} = \sum_k (\bar{h}_{l-2k} d_k^{j+1,2n} + \bar{g}_{l-2k} d_k^{j+1,2n+1}) \quad (2)$$

式中各参数意义见文献[32]。

2.2.2 变色龙优化算法(CSA) CSA 是 Malik 等于 2021 年提出的一种新型智能优化算法,该算法通过对变色龙搜索猎物、锁定猎物、攻击猎物 3 种行为进行数学建模来求解待优化问题^[33-34],CSA 数学描述具体见文献[34]。

2.2.3 猎豹优化(CO)算法 CO 算法是 Akbari 等^[35]于 2022 年受自然界猎豹狩猎启发而提出的一种新型智能优化算法。该算法通过模拟猎豹在狩猎过程中搜索、坐等和攻击 3 种策略来实现位置更新,即待优化问题的求解。CO 算法数学描述如下:

(1) 初始化。与其他群体智能优化算法类似,CO 算法也是从种群初始化开始。设在 d 维搜索空间中,猎豹初始化位置描述为:

$$X_{i,j} = LB_j + rand \cdot (UB_j - LB_j) \quad (3)$$

$(i = 1, 2, \dots, n; j = 1, 2, \dots, d)$

式中: $X_{i,j}$ 为第 i 头猎豹第 j 维位置; UB_j 、 LB_j 为第 j 维搜索空间上、下限值; $rand$ 为介于 0 和 1 之间的随机数; n 为猎豹种群规模; d 为问题维度。

(2) 搜索策略。猎豹在其领地(搜索空间)或周围区域进行全范围扫描或主动搜索,以找到猎物。该策略数学描述为:

$$X_{i,j}^{t+1} = X_{i,j}^t + \tilde{r}_{i,j}^{-1} \alpha_{i,j}^t \quad (4)$$

式中: $X_{i,j}^t$ 为第 i 头猎豹第 t 次迭代第 j 维位置; $\tilde{r}_{i,j}$ 为第 i 头猎豹第 j 维呈正态分布的随机数; $\alpha_{i,j}^t$ 为第 i 头猎豹第 t 次迭代第 j 维的搜索步长; T 为算法最大迭代次数。

(3) 坐等策略。在搜索模式下,猎物可能会暴露在猎豹视野中,在这种情况下,猎豹的每一个动作都可能会导致猎物逃跑。为避免该情况发生,猎豹采取坐等伏击策略(躺在地上或躲进灌木丛)以接近猎物。该策略数学描述为:

$$X_{i,j}^{t+1} = X_{i,j}^t \quad (5)$$

该策略不但可提高狩猎成功率(获得最优解),而且还避免了 CO 过早收敛。

(4) 攻击策略。在 CO 算法中,种群中每头猎豹都可以根据逃跑猎物、领头猎豹或附近猎豹的位置来调整自己的位置,以获得最佳攻击位置。该策略数学描述为:

$$X_{i,j}^{t+1} = X_{B,j}^t + \tilde{r}_{i,j} \cdot \beta_{i,j}^t \quad (6)$$

式中: $X_{B,j}^t$ 为第 t 次迭代第 j 维猎物位置,即当前最佳位置; $\tilde{r}_{i,j}$ 为第 i 头猎豹第 j 维转向因子; $\beta_{i,j}^t$ 为第 i 头猎豹第 t 次迭代第 j 维交互因子,反映猎豹之间或猎

豹与领头猎豹之间的互动。

2.2.4 山瞪羚优化 (MGO) 算法 MGO 是 Abodollahzadeh 等^[36] 于 2022 年受自然界中瞪羚生存能力启发而提出的一种新型群体智能优化算法。该算法通过对瞪羚觅食 (开发阶段, 在没有捕食者或在捕食者跟踪情形下)、逃离 (探索阶段, 发现捕食者) 行为进行数学建模求解待优化问题。MGO 算法数学描述如下:

(1) 初始化。与其他群体智能优化算法类似, MGO 也是从种群初始化开始。设在 d 维搜索空间中, 瞪羚种群为 N 的个体初始化位置描述为:

$$x_{i,j} = LB_j + rand \cdot (UB_j - LB_j) \quad (7)$$

$$(i = 1, 2, \dots, N; j = 1, 2, \dots, d)$$

式中: $x_{i,j}$ 为第 i 只瞪羚第 j 维位置; UB_j, LB_j 为第 j 维搜索空间上、下限值; $rand$ 为介于 0 和 1 之间的随机数; N 为瞪羚种群规模; d 为问题维度。

(2) 觅食 (开发) 阶段。MGO 中, 假设在没有捕食者或在捕食者跟踪情形下, 瞪羚悠闲地觅食, 并在觅食时以布朗运动进行移动, 瞪羚位置更新数学描述为:

$$\vec{X}_{t+1} = \vec{X}_t + s \cdot \vec{R} \cdot \vec{R}_b \cdot (\vec{X}_{Elite} - \vec{R}_b \cdot \vec{X}_t) \quad (8)$$

式中: \vec{X}_{t+1} 为瞪羚第 $t+1$ 次迭代位置; \vec{X}_t 为瞪羚当前位置; s 为瞪羚觅食移动速度, km/h; \vec{R}_b 为包含了布朗运动的随机向量; \vec{X}_{Elite} 为具有最佳适应值的瞪羚位置, 即精英瞪羚位置; \vec{R} 为 $[0, 1]$ 范围内均匀随机向量。

(3) 逃离 (探索) 阶段。MGO 中, 瞪羚一旦发现捕食者就迅速奔跑, 并通过快速奔跑和突然改变方向来逃离捕食者。假设瞪羚起步阶段采用布朗运动, 奔跑阶段改用 Lévy 飞行。瞪羚发现捕食者时, 其位置更新描述为:

$$\vec{X}_{t+1} = \vec{X}_t + S \cdot \mu \cdot \vec{R} \cdot \vec{R}_L \cdot (\vec{X}_{Elite} - \vec{R}_L \cdot \vec{X}_t) \quad (9)$$

式中: S 为瞪羚可以达到的最大速度, 本文取值为 88 km/h; \vec{R}_L 为基于 Lévy 分布的随机向量; μ 为方向突然系数; 其他参数意义同上。

捕食者追逐瞪羚时, 其位置更新描述为:

$$\vec{X}_{t+1} = \vec{X}_t + S \cdot \mu \cdot CF \cdot \vec{R}_b \cdot (\vec{X}_{Elite} - \vec{R}_L \cdot \vec{X}_t) \quad (10)$$

式中: CF 为捕食者移动的控制参数, 描述为 $CF = (1 - l/T)^{(2l/T)}$; l 为当前迭代次数; T 为最大迭代次数; 其他参数意义同上。

(4) 捕食成功率。研究表明, 虽然瞪羚不是濒危物种, 但其年存活率为 66%, 这意味着捕食者捕食成功率仅为 34%。MGO 通过捕食者成功率 PSR 来避免算法陷入局部最小值, 数学描述如下:

$$\vec{X}_{t+1} = \begin{cases} \vec{X}_t + CF[LB + \vec{R} \cdot (UB - LB)] \cdot \vec{U} & (r \leq PSR) \\ \vec{X}_t + [PSR(1 - r) + r] \cdot (\vec{X}_{r_1} - \vec{X}_{r_2}) & (\text{else}) \end{cases} \quad (11)$$

式中: r 为 $[0, 1]$ 范围内的随机数; \vec{U} 为随机生成的二进制向量, 若 $r < 0.2$, 则 $\vec{U} = 0$, 否则 $\vec{U} = 1$; r_1, r_2 为瞪羚矩阵随机指数; 其他参数意义同上。

2.2.5 长短期记忆神经网络 LSTM 在隐藏层单元中设计了遗忘门、输入门和输出门, 利用输出门来控制单元状态 c_t 的当前输出值 h_t , LSTM 结构如图 2 所示。

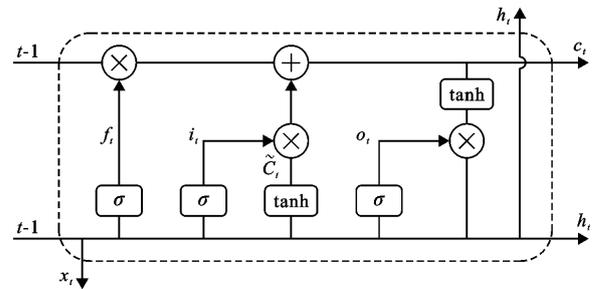


图 2 LSTM 结构示意图

设输入序列为 (x_1, x_2, \dots, x_T) , 隐含层状态为 (h_1, h_2, \dots, h_T) , 则在 t 时刻有^[18-22]:

$$i_t = \sigma(W_{hi}h_{t-1} + W_{xi}x_t + b_i) \quad (12)$$

$$f_t = \sigma(W_{hf}h_{t-1} + W_{xf}x_t + b_f) \quad (13)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot g(W_{hc}h_{t-1} + W_{xc}x_t + b_c) \quad (14)$$

$$o_t = \sigma(W_{ho}h_{t-1} + W_{xo}x_t + W_{co}c_t + b_o) \quad (15)$$

$$h_t = o_t \cdot g(c_t) \quad (16)$$

为使 LSTM 满足预测目的, 需加上一个线性回归层, 即:

$$y_t = W_{yo}h_t + b_y \quad (17)$$

上式中各参数意义见文献[18]、[19]。

2.2.6 门控循环单元 GRU 结构与 LSTM 类似, 是将 LSTM 的更新门和遗忘门整合为新的更新门, 相比于 LSTM, 其结构更加简单、训练参数更少, 因此收敛速度也就更快^[22-25], GRU 结构如图 3 所示。

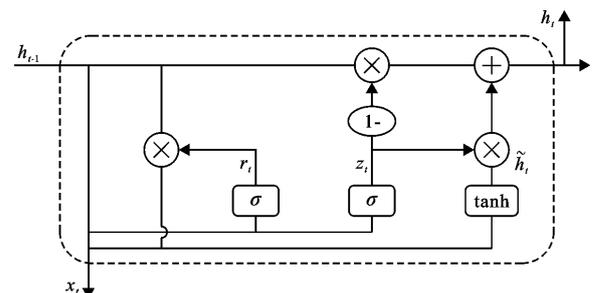


图 3 GRU 结构示意图

GRU 的迭代公式如下:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (18)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (19)$$

$$\tilde{h} = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (20)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h} \quad (21)$$

式中: z 为更新门; r 为重置门; W 为权重矩阵; σ 为 sigmoid 函数; h 为输出值; $*$ 为哈达玛积 (Hadamard product); 下标 t 表示 t 时刻。

2.3 预测步骤

WPT - CSA - GRU、WPT - CO - GRU、WPT -

MGO - GRU、WPT - CSA - LSTM、WPT - CO - LSTM、WPT - MGO - LSTM 模型的预测步骤归纳如下:

步骤 1: 利用 WPT 将 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时序数据进行 2 层分解, 分别得到 1 个趋势项分量 [2,4] 和 3 个周期项分量 [2,1] ~ [2,3], 其中趋势项分量可有效反映原始水质数据的趋势性, 周期项分量可反映原始水质数据的微小震荡变化, 如图 4 所示。

本文选取 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时序数据的 70% 作为训练样本, 剩余 30% 作为预测样本。

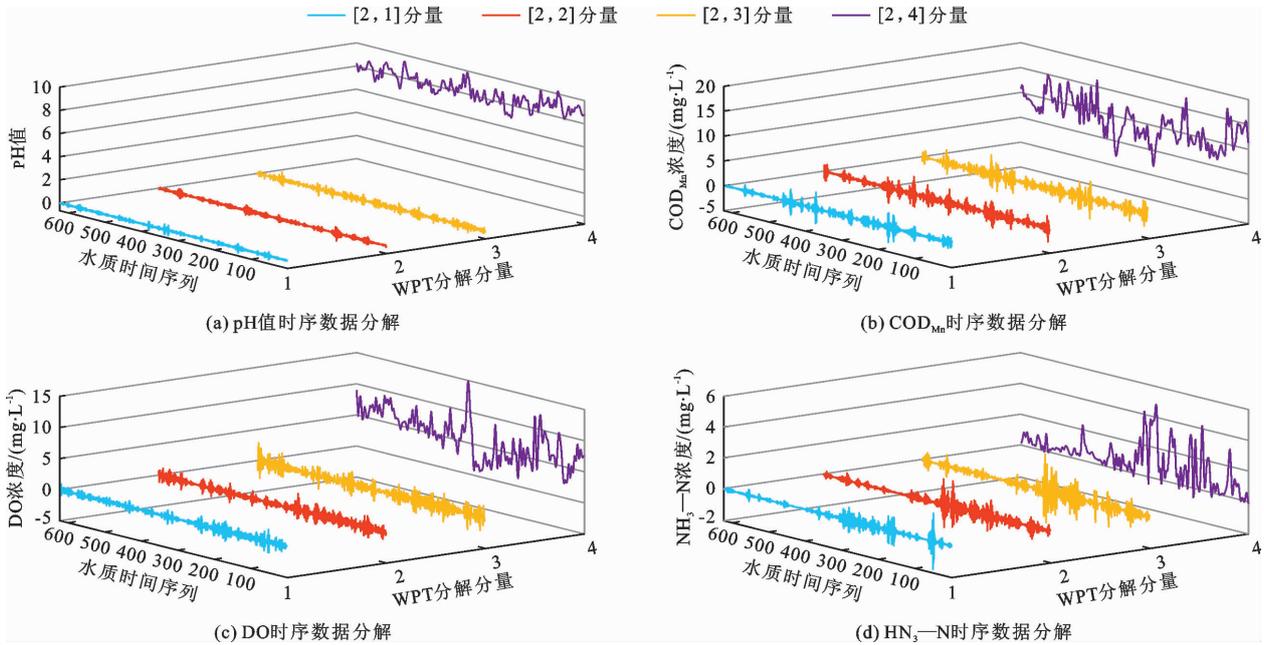


图 4 pH 值及 COD_{Mn}、DO、NH₃-N 浓度时间序列数据的 WPT 分解

步骤 2: 基于 SPSS 软件, 利用偏自相关函数 (partial autocorrelation function, PACF) 确定各分解分量的输入步长 k 。经计算, pH 值及 COD_{Mn}、DO、NH₃-N 浓度趋势项分量 [2,4] 的 k 值均为 1; 周期项分量 [2,1] ~ [2,3] 的 k 值均为 2, 即利用前 k 周的水质数据预测当前水质。各模型的输入、输出可表述为:

$$M_{\text{input}} = \begin{bmatrix} M_1 & M_2 & \cdots & M_k \\ M_2 & M_3 & \cdots & M_{k+1} \\ \vdots & \vdots & & \vdots \\ M_{l-k} & M_{l-k+1} & \cdots & M_{l-1} \end{bmatrix} \quad (22)$$

$$M_{\text{output}} = \begin{bmatrix} M_{k+1} \\ M_{k+2} \\ \vdots \\ M_l \end{bmatrix} \quad (23)$$

式中: l 为样本数量; k 为输入步长。

步骤 3: GRU/LSTM 预测性能的优劣取决于其超参数的选择。本文利用 2 层 GRU/LSTM 网络进行预测, 并利用各分量训练样本构建变色龙优化算法 (CSA)、猎豹优化 (CO) 算法、山瞪羚优化 (MGO) 算法调优 GRU/LSTM 超参数的适应度函数 f :

$$\begin{cases} \min f(m_1, m_2, dr_1, dr_2, c, L_1, L_2) = \frac{1}{l} \sum_{i=1}^l (\hat{M}_i - M_i)^2 \\ \text{s. t } m_1 \in [m_{1\min}, m_{1\max}], m_2 \in [m_{2\min}, m_{2\max}], dr_1 \in [dr_{1\min}, dr_{1\max}], dr_2 \in [dr_{2\min}, dr_{2\max}], \\ c \in [c_{\min}, c_{\max}], L_1 \in [L_{1\min}, L_{1\max}], L_2 \in [L_{2\min}, L_{2\max}] \end{cases} \quad (24)$$

式中: m_1 、 m_2 分别为第 1、第 2 隐藏层神经元数; dr_1 、 dr_2 分别为 dropout_1、dropout_2 参数; c 为学习

率; L_1 为 Learn Rate Drop Period 参数; L_2 为 Learn Rate Drop Factor 参数; 下标 min、max 分别为各调优参数搜索空间上、下限值; 其他参数意义同上。

步骤4: 设置变色龙优化算法(CSA)、猎豹优化(CO)算法、山瞪羚优化(MGO)算法种群规模为30, 最大迭代次数为100, 其他采用算法默认值初始化变色龙、猎豹、山瞪羚空间位置。

设置 GRU/LSTM 超参数 [$m_1, m_2, dr_1, dr_2, c, L_1, L_2$] 搜索空间下限为 [8, 8, 0.001, 0.001, 0.001, 0.0001, 0.1], 上限为 [128, 128, 0.5, 0.5, 0.5, 1, 1], 激活函数选用 tanh 函数; SVM 惩罚因子、核函数参数、损失系数搜索空间下限设置为 [0.01, 0.01, 0.001], 上限设置为 [100, 100, 0.1], 交叉验证折数设置为3, 原始数据采用 [-1, 1] 进行归一化处理。本文 WPT-LSTM、WPT-GRU 模型中的超参数采用试错法确定。

步骤5: 基于公式(24)计算变色龙优化算法(CSA)、猎豹优化(CO)算法、山瞪羚优化(MGO)算法计算个体的适应度值, 找到并保存最佳个体位置。令 $t = 1$ 。

步骤6: 分别利用上述变色龙优化算法(CSA)、猎豹优化(CO)算法、山瞪羚优化(MGO)算法位置更新算子更新个体新位置。

步骤7: 利用个体位置更新后的变色龙优化算法(CSA)、猎豹优化(CO)算法、山瞪羚优化(MGO)算法计算个体适应度值, 再比较并保存当前最佳个体位置。

步骤8: 令 $t = t + 1$, 若 $t = T$, 则结束算法; 否则返回步骤6。

步骤9: 输出最佳个体位置, 该位置即为最佳 GRU/LSTM 超参数向量。利用该向量建立 WPT-CSA-GRU 等6种模型对各分量进行预测和重构。

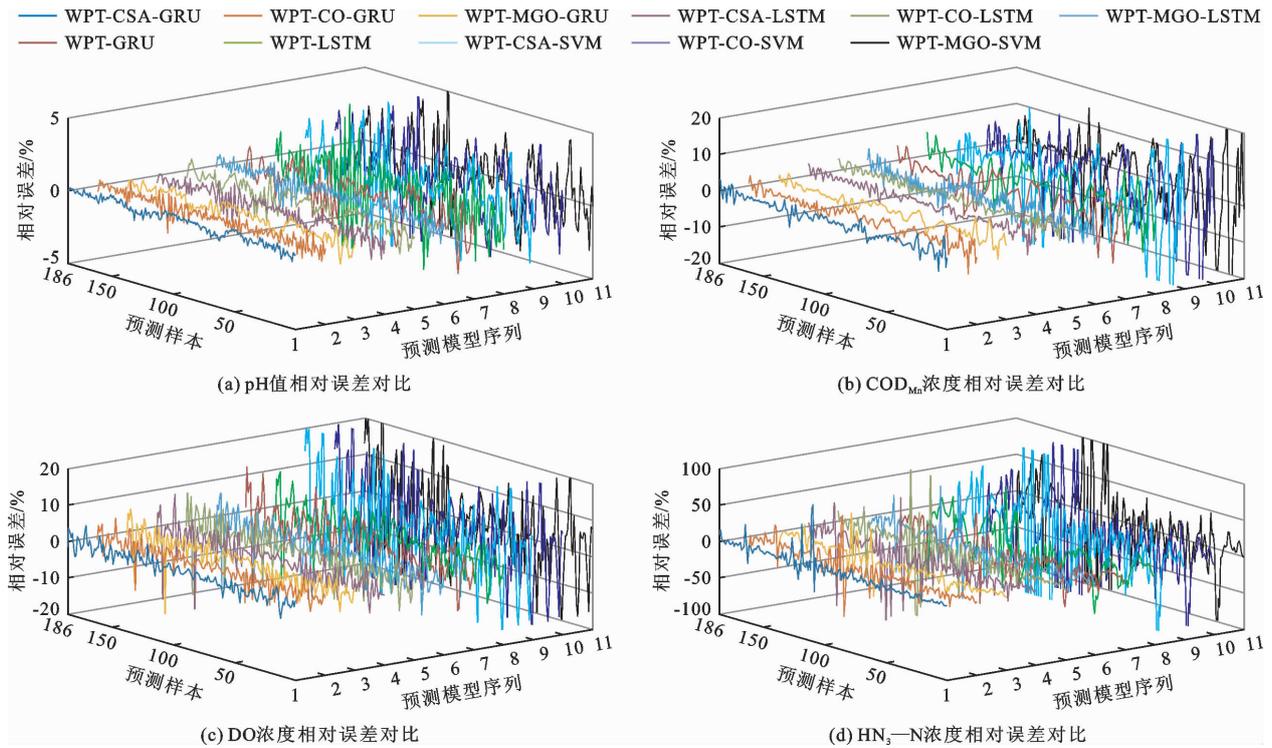
步骤10: 利用平均绝对百分比误差 MAPE、平均绝对误差 MAE、均方根误差 RMSE 和决定系数 DC 对模型进行评价。其中, MAPE、MAE 用于反映模型预测误差大小, RMSE 用于衡量观测值与真值之间的偏差, MAPE、MAE、RMSE 越小则说明模型性能越优、预测精度越高; DC 反映变量之间相关关系的密切程度, 其值越大则说明模型性能越好。

3 算法验证与结果分析

构建 WPT-CSA-GRU 等11种模型对实例断面(云南省昆明市观音山断面)pH值等4个水质指标进行训练及预测, 预测结果评价及对比见表1, 预测相对误差见图5。

表1 实例水质时间序列各模型预测结果评价及对比(pH值无量纲)

水质指标	模型	MAPE / %	MAE / (mg · L ⁻¹)	RMSE / (mg · L ⁻¹)	DC
pH 值	WPT-CSA-GRU	0.230	0.019	0.023	0.9982
	WPT-CO-GRU	0.304	0.024	0.032	0.9967
	WPT-MGO-GRU	0.310	0.025	0.032	0.9967
	WPT-CSA-LSTM	0.383	0.030	0.040	0.9948
	WPT-CO-LSTM	0.423	0.034	0.043	0.9938
	WPT-MGO-LSTM	0.459	0.037	0.046	0.9931
	WPT-GRU	0.836	0.066	0.086	0.9758
	WPT-LSTM	1.275	0.101	0.136	0.9397
	WPT-CSA-SVM	1.447	0.115	0.157	0.9193
	WPT-CO-SVM	1.445	0.115	0.156	0.9200
	WPT-MGO-SVM	1.443	0.115	0.156	0.9200
	WPT-CSA-GRU	0.991	0.096	0.129	0.9972
WPT-CO-GRU	1.018	0.103	0.138	0.9968	
WPT-MGO-GRU	1.093	0.110	0.148	0.9963	
WPT-CSA-LSTM	1.611	0.160	0.184	0.9942	
WPT-CO-LSTM	1.752	0.176	0.203	0.9930	
COD _{Mn}	WPT-MGO-LSTM	1.823	0.182	0.228	0.9911
	WPT-GRU	2.708	0.269	0.348	0.9793
	WPT-LSTM	2.879	0.289	0.373	0.9763
	WPT-CSA-SVM	6.542	0.626	0.851	0.8767
	WPT-CO-SVM	6.086	0.592	0.809	0.8886
	WPT-MGO-SVM	6.080	0.592	0.808	0.8890
DO	WPT-CSA-GRU	1.200	0.073	0.094	0.9965
	WPT-CO-GRU	1.451	0.084	0.106	0.9956
	WPT-MGO-GRU	1.891	0.113	0.143	0.9920
	WPT-CSA-LSTM	2.221	0.126	0.165	0.9893
	WPT-CO-LSTM	2.685	0.154	0.201	0.9842
	WPT-MGO-LSTM	2.417	0.135	0.183	0.9869
	WPT-GRU	3.193	0.196	0.254	0.9749
	WPT-LSTM	3.388	0.202	0.260	0.9737
	WPT-CSA-SVM	8.368	0.490	0.633	0.8436
	WPT-CO-SVM	8.363	0.488	0.629	0.8456
	WPT-MGO-SVM	8.324	0.485	0.630	0.8452
	WPT-CSA-GRU	5.391	0.014	0.019	0.9984
WPT-CO-GRU	7.030	0.019	0.026	0.9970	
WPT-MGO-GRU	5.403	0.015	0.020	0.9982	
WPT-CSA-LSTM	10.370	0.029	0.038	0.9933	
WPT-CO-LSTM	11.383	0.030	0.042	0.9921	
NH ₃ -N	WPT-MGO-LSTM	8.187	0.024	0.031	0.9956
	WPT-GRU	19.384	0.052	0.058	0.9847
	WPT-LSTM	20.193	0.060	0.072	0.9760
	WPT-CSA-SVM	40.242	0.107	0.153	0.8935
	WPT-CO-SVM	38.231	0.106	0.152	0.8944
	WPT-MGO-SVM	37.562	0.107	0.152	0.8941



注:预测模型序列1~11分别为WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU、WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM、WPT-GRU、WPT-LSTM、WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM模型

图5 各预测模型对实例断面4个水质指标预测结果的相对误差

依据表1及图5可以得出以下结论:

(1) WPT-CSA-GRU等6种模型对实例pH值预测的 $MAPE$ 为0.230%~0.459%、 MAE 为0.019~0.037 mg/L、 $RMSE$ 为0.023~0.046 mg/L,决定系数 $DC > 0.993$;对 COD_{Mn} 浓度预测的 $MAPE$ 为0.991%~1.823%、 MAE 为0.096~0.182 mg/L、 $RMSE$ 为0.129~0.228 mg/L,决定系数 $DC > 0.991$;对DO浓度预测的 $MAPE$ 为1.200%~2.685%、 MAE 为0.073~0.154 mg/L、 $RMSE$ 为0.094~0.201 mg/L,决定系数 $DC > 0.984$;对 NH_3-N 浓度预测的 $MAPE$ 为5.391%~11.383%、 MAE 为0.014~0.030 mg/L、 $RMSE$ 为0.019~0.042 mg/L,决定系数 $DC > 0.992$,预测精度优于WPT-GRU、WPT-LSTM,远优于WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM模型。以 $MAPE$ 评价指标为例,WPT-CSA-GRU等6种模型对pH值及 COD_{Mn} 、DO、 NH_3-N 浓度的预测精度较WPT-GRU、WPT-LSTM模型分别提高了45.1%、32.7%、15.9%、41.3%以上,较WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM模型分别提高了68.2%、70.0%、67.7%、69.7%以上。可见,本文提出的WPT-CSA-GRU等6种模

型具有较好的预测效果,应用于水质预测是可行的。其中,WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU模型的预测效果优于WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM模型,具有更高的预测精度和更好的泛化能力。

(2) WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU模型对pH值及 COD_{Mn} 、DO、 NH_3-N 浓度预测的 $MAPE$ 较WPT-GRU模型分别提高了62.9%、59.6%、40.8%、63.7%以上;WPT-CSA-LSTM、WPT-CO-LSTM、WPT-CSA-LSTM模型预测的 $MAPE$ 较WPT-LSTM模型分别提高了34.4%、36.7%、20.7%、43.6%以上,表明CSA、CO、MGO能有效优化GRU/LSTM超参数,不但提高了GRU/LSTM预测性能,而且还克服了人工调节超参数带来的高不确定性和效果不佳的缺点,其中CSA优化效果最佳。

(3)从模型预测效果对比来看,WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU模型的预测效果最佳,WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM模型预测效果次之,WPT-GRU、WPT-LSTM模型预测效果一般,WPT-CSA-SVM、WPT-CO-SVM、WPT-MGO-SVM

SVM 模型预测效果最差;从 4 个水质指标的时序数据变化及预测效果对比来看,pH 值时序数据起伏变化不大,预测效果最好;COD_{Mn}、DO 时序数据起伏变化较大,预测效果较好;NH₃-N 时序数据起伏变化最大,预测效果相对较差。

4 讨论

深度学习是由多个非线性映射隐含层组成的一种神经网络,能较好地挖掘数据之间的非线性。近年来,随着深度学习的快速发展,以 LSTM、GRU 为代表的深度学习已经成为当前研究的热点,并被广泛应用于水质预测^[37-39]领域。LSTM 是 RNN 的变种,克服了 RNN 面临的梯度消失和梯度爆炸问题,能够解决其他深度学习算法无法处理的时间长期依赖问题;GRU 是 LSTM 的变种,是为解决长期记忆和反向传播中的梯度等问题而提出来的,与 LSTM 相比,GRU 拥有更低的复杂度以及更好的预测性能。在实际应用中,LSTM、GRU 预测精度和收敛性在很大程度上依赖于超参数(如层数、学习率、隐含层节点数等),而超参数的设置普遍采用手工试错的方法,难以满足 GRU/LSTM 预测需求。CSA、CO、MGO 具有良好的全局搜索能力,应用 CSA、CO、MGO 调优 GRU/LSTM 超参数得到全局最优解,不但克服了 LSTM、GRU 易陷入局部最优从而影响预测精度的缺点,而且还使模型获得了较好的智能化水平。此外,水质时间序列具有较强的非线性和多尺度等特征,采用 WPT 进行分解处理,将水质时间序列分解为多个更具规律的模态分量,不但可降低模型预测器的预测难度,而且还大大减少了模型的预测误差。

GRU 与 LSTM 相比,其在结构上更具优势,从 WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU 模型与 WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM 模型的 pH 值及 COD_{Mn}、DO、NH₃-N 浓度预测结果对比来看,WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU 模型利用重置门控制当前信息和记忆信息的数据量,并生将成新的记忆信息继续向前传递,具有更好的适应性与预测能力。

当然,本文提出的 WPT-CSA-GRU 等 6 种预测模型存在训练时间长、系统资源消耗大等缺点和不足。同时本文仅利用了云南省昆明市观音山断面 pH 值及 DO、COD_{Mn}、NH₃-N 浓度的历史数据作为输入,未考虑滇池补水、水温以及降水等因素的影响,因此,

该模型及预测精度仍有进一步提升的空间。

5 结论

为提高水质时间序列的预测精度,本文提出了 WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU、WPT-CSA-LSTM、WPT-CO-LSTM、WPT-MGO-LSTM 预测模型,并构建了若干个对比模型,结合云南省昆明市观音山断面 2004—2015 年逐周 pH 值及 COD_{Mn}、DO、NH₃-N 浓度预测算例,得到如下结论:

(1) 本文提出的 WPT-CSA-GRU 等 6 种模型预测误差小于其他对比模型,具有较高的预测精度和较好的泛化能力,将其用于水质时间序列预测是可行的。其中 WPT-CSA-GRU、WPT-CO-GRU、WPT-MGO-GRU 模型对 pH 值及 COD_{Mn}、DO、NH₃-N 浓度预测的 MAPE 分别在 0.230%~0.310%、0.991%~1.093%、1.200%~1.891%、5.391%~7.030% 之间,DC 均在 0.99 以上,具有更好的预测效果。

(2) 针对随机性、波动性较强的水质时间序列,利用 WPT 对其进行平稳化处理,可得到更具规律的趋势项分量和周期项分量,大大减少了混合模型的预测误差。

(3) 针对 GRU/LSTM 因超参数选择不当而导致预测性能低下的问题,利用 CSA、CO、MGO 调优 GRU/LSTM 超参数,不但提高了 GRU/LSTM 的预测性能,而且克服了人工调节超参数带来的高不确定性和效果不佳的缺点。

(4) 本文提出的模型及方法可为相关时间序列数据处理、GRU/LSTM 超参数调优以及时间序列预测研究提供参考。

参考文献:

- [1] 崔东文,袁树堂.基于 WPD-AHA-ELM 模型的水质时间序列多步预测[J].三峡大学学报(自然科学版),2023,45(1):6-13.
- [2] 赵林,戴天骄,陈亮,等.基于 BP 神经网络的桃林口水库水质预测[J].安徽农业科学,2019,47(24):73-77+81.
- [3] 刘洁,祝榕婕,姜德迅,等.基于遗传-神经网络的实时水质预测模型[J].南水北调与水利科技(中英文),2020,18(6):93-100.
- [4] 薛同来,赵冬晖,韩菲.基于 GA 优化的 SVR 水质预测模型研究[J].环境工程,2020,38(3):123-127.
- [5] 李彦杰,贺鹏飞,冯巍巍,等.基于 LSTM 模型的海洋水质预测[J].计算机与数字工程,2020,48(2):437-441.

- [6] 孙开争,刘杨,姜腾龙,等.城市河流水质预测的 LSTM 方法适用性研究[J].河北环境工程学院学报,2022,32(6):78-82.
- [7] 唐亦舜,徐庆,刘振鸿,等.基于优化非线性自回归神经网络模型的水质预测[J].东华大学学报(自然科学版),2022,48(3):93-100.
- [8] 张丽娜,陈会娟,余昭旭.基于 TimeGAN-CNN-LSTM 模型的河流水质预测研究[J].自动化仪表,2022,43(8):11-15.
- [9] 徐睿,张晓斌,薛鹏松.基于改进的 GRNN-Markov 水质预测模型研究及应用[J].灌溉排水学报,2022,41(S1):104-110.
- [10] 宦娟,张浩,徐宪根,等.基于图卷积 STG-LSTM 的京杭运河水质时空预测研究[J].中国农村水利水电,2022(8):14-22.
- [11] 孙铭,魏守科,王莹洁,等.基于小波分解的 LSTM 水质预测模型[J].计算机系统应用,2020,29(12):55-63.
- [12] 金昌盛,邓仁健,刘俞希,等.基于奇异谱分析-遗传算法反向传播神经网络模型的湘江新港断面水质预测[J].环境污染与防治,2019,41(6):710-713+719.
- [13] 顾乾晖,曾斌,涂振宇.基于 VMD-PSO-LSTM 水质预测模型的应用研究[J].南昌工程学院学报,2021,40(4):23-29.
- [14] 李建文,孟连子,刘娜,等.基于 EEMD-SVR 的渔业养殖水质预测模型[J].天津理工大学学报,2018,34(5):14-19+24.
- [15] 白雯睿,杨毅强,郭辉,等.基于 VMD-CNN-LSTM 的珠江流域水质多步预测模型研究[J].四川轻化工大学学报(自然科学版),2022,35(4):66-74.
- [16] 邹可可,李中原,穆小玲,等.基于 LSTM-GRU 的污水水质预测模型研究[J].能源与环保,2021,43(12):59-63.
- [17] 余本功,许庆堂,张培行.基于 MAC-LSTM 的问题分类研究[J].计算机应用研究,2020,37(1):40-43.
- [18] 胡顺强,崔东文.基于海洋捕食者算法优化的长短期记忆神经网络径流预测[J].中国农村水利水电,2021(2):78-82+90.
- [19] 胡顺强,崔东文.基于 EMD-LSTM-ANFIS 模型的年径流预测研究[J].人民珠江,2021,42(3):46-52.
- [20] 史建楠,邹俊忠,张见,等.基于 DMD-LSTM 模型的股票价格时间序列预测研究[J].计算机应用研究,2020,37(3):662-666.
- [21] 宋刚,张云峰,包芳勋,等.基于粒子群优化 LSTM 的股票预测模型[J].北京航空航天大学学报,2019,45(12):2533-2542.
- [22] 徐冬梅,夏王萍,王文川.基于黏菌算法优化 VMD-CNN-GRU 模型的年径流预测[J].南水北调与水利科技(中英文),2022,20(3):429-439.
- [23] 黄永涛,刘晓青,季威,等. HCM-PSO-GRU 组合预测模型在大坝变形预测中的应用[J].水电能源科学,2021,39(12):120-123+61.
- [24] 孟安波,陈顺,王陈恩,等.基于混沌 CSO 优化时序注意力 GRU 模型的超短期风电功率预测[J].电网技术,2021,45(12):4692-4700.
- [25] 王秀杰,张帅,田福昌,等.基于 TVF-EMD-PSO-GRU 的月径流模型与应用研究[J].天津大学学报(自然科学与工程技术版),2022,55(8):802-810.
- [26] 段中兴,温倩,周孟,等.基于改进蝙蝠算法优化 LSTM 网络的短时客流预测[J].铁道科学与工程学报,2021,18(11):2833-2840.
- [27] 刘可真,阮俊泉,赵现平,等.基于麻雀搜索优化的 Attention-GRU 短期负荷预测方法[J].电力系统及其自动化学报,2022,34(4):99-106.
- [28] 吴铁洲,邹智,姜奔,等.基于 TLBGA-GRU 神经网络的短期负荷预测[J].计算机工程,2022,48(11):69-76.
- [29] 李新华,崔东文.基于 WPD-TSO-ELM 模型的月径流时间序列预测[J].水力发电,2022,48(9):9-15+44.
- [30] 杨琼波,崔东文.基于小波包分解的 AJS-GMDH 月径流时间序列预测研究[J].水力发电,2022,48(6):45-51.
- [31] 朱刘柱,王绪利,马静,等.基于小波包分解与循环神经网络的综合能源系统短期负荷预测[J].电力建设,2020,41(12):135-142.
- [32] 许建伟,崔东文. WPT-HPO-ELM 径流多步预报模型研究[J].水资源与水工程学报,2022,33(6):69-76.
- [33] BRAIK M S. Chameleon swarm algorithm: a bio-inspired optimizer for solving engineering design problems[J]. Expert Systems with Applications, 2021, 174: 114685.
- [34] 许建伟,崔东文.战争策略算法与变色龙算法优化极限学习机的输沙量时间序列预测[J].水力发电,2022,48(11):36-42.
- [35] AKBARI M A, ZAER M, AZIZIPANAH-ABARG-HOOEE R, et al. The cheetah optimizer: a nature-inspired metaheuristic algorithm for large-scale optimization problems[J]. Scientific Reports, 2022, 12: 10953.
- [36] ABDOLLAHZADEH B, GHAREHCHOPOGH F S, KHODADADI N, et al. Mountain gazelle optimizer: a new nature-inspired metaheuristic algorithm for global optimization problems[J]. Advances in Engineering Software, 2022, 174: 103282.
- [37] 郝玉莹,赵林,孙同,等.基于 RF-LSTM 的地表水体水质预测[J].水资源与水工程学报,2021,32(6):41-48.
- [38] 孙开争,刘杨,姜腾龙,等.城市河流水质预测的 LSTM 方法适用性研究[J].河北环境工程学院学报,2022,32(6):78-82.
- [39] 黎煜昭,刘启亮,邓敏,等.基于物理约束 GRU 神经网络的河流水质预测模型[J].地球信息科学学报,2023,25(1):102-114.