

两种优化小波阈值算法对紫外可见光谱的水质COD数据去噪研究

黄慧波*, 李攀, 梁超, 刘连鲁诣

重庆理工大学, 重庆

收稿日期: 2022年2月24日; 录用日期: 2022年3月29日; 发布日期: 2022年4月29日

摘要

本文主要检测水质COD数据, 对数据进行去噪处理, 为建立预测水质COD模型提供数据预处理奠定基础。紫外可见光谱检测水质COD数据系统易产生光学仪器本身和外界散射光的噪声影响, 得到光谱数据存在不必要的噪声。为开展去噪效果评价, 实验稀释了COD浓度为20 mg/L的邻苯二甲酸氢钾溶液来得到了30组水质COD光谱数据, 将六种算法去噪效果相互对比, 得出灰狼优化小波阈值算法更适合用于去除水质COD数据噪声的结论。为了验证两种优化算法去噪的可行性, 采集某内陆河地表水质COD光谱数据进行实例去噪分析。比较两种优化算法的去噪参数评价, 灰狼优化小波阈值算法能更好抑制了噪声和提高了系统检测精度, 为紫外可见光谱法的水质COD数据的去噪处理提供了一种全新的解决办法, 证明了两种优化小波阈值算法适用于地表水质COD数据去噪分析的可行性。

关键词

紫外可见光谱, 化学需氧量, 小波阈值, 去噪, 优化算法

Study on Denoising of UV-VIS Spectrum Water Quality COD Data by Two Optimized Wavelet Threshold Algorithms

Huibo Huang*, Pan Li, Chao Liang, Lianluyi Liu

Chongqing University of Technology, Chongqing

Received: Feb. 24th, 2022; accepted: Mar. 29th, 2022; published: Apr. 29th, 2022

作者简介: 黄慧波, 重庆人, 出生于1996年8月, 硕士研究生, 研究方向: 光电检测技术与仪器。

*通讯作者 Email: 1084072982@qq.com

文章引用: 黄慧波, 李攀, 梁超, 刘连鲁诣. 两种优化小波阈值算法对紫外可见光谱的水质COD数据去噪研究[J]. 水资源研究, 2022, 11(2): 187-193. DOI: 10.12677/jwrr.2022.112020

Abstract

This paper mainly detects the COD data of water quality and denoises the data, which lays a foundation for data preprocessing to establish the COD model for predicting water quality. The UV-VIS spectrum detection water quality COD data system is easy to produce the noise influence of the optical instrument itself and the external scattered light, and there is unnecessary noise in the spectral data. In order to evaluate the denoising effect, 30 groups of water quality COD spectral data were obtained by diluting the potassium hydrogen phthalate solution with COD concentration of 20 mg/L. The denoising effects of the six algorithms were compared with each other, and it was concluded that the gray wolf optimized wavelet threshold algorithm was more suitable for removing the noise of water quality COD data. In order to verify the feasibility of the two optimization algorithms, COD spectral data of surface water quality of an inland river were collected for example denoising analysis. Comparing the denoising parameter evaluation of the two optimization algorithms, the gray wolf optimized wavelet threshold algorithm can better suppress the noise and improve the detection accuracy of the system. It provides a new solution for the denoising of water quality COD data by UV-VIS spectroscopy, and proves that the two optimized wavelet threshold algorithms are suitable for the denoising analysis of surface water quality COD data.

Keywords

Ultraviolet Visible Spectrum, Chemical Oxygen Demand, Wavelet Threshold, Denoising, Optimization Algorithm

Copyright © 2022 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

水源安全关乎人类生存安全, 发展经济的同时带来了严重的水体环境污染, 未经处理废水和污水排入地表水会加重水体净化负担。检测水质环境是治理所有水体污染繁杂工作的关键, 测定水质环境要检测出水质环境中各指标, 水质中指标[1]包含色度、浊度 FTU、酸碱度 PH、生化需氧量 BOD、化学需氧量 COD、电导率和总有机碳 TOC 等。本文主要开展检测水质中 COD 含量, 化学需氧量是指待测水质中需要被氧化的还原性物质质量, 测定 COD 含量便知道水质受到不同轻重程度的污染。本文以《地表水环境质量标准》(GB 3838-2002)为标准, 分析检测地表水 COD 数据的去噪效果, 为提高地表水污染治理奠定基础。

紫外可见光谱法比以传统化学方法为代表的检测水质 COD 方法更具有检测速度快、无二次污染和操作便捷等优点, 被广泛用在农产品[2]、微生物数量[3]、矿物石[4]和药材[5]鉴定等检测领域。在建立紫外可见光谱检测水质 COD 数据系统中, 存在自然环境干扰、电路接口信号衰减、光电仪器和光学系统等因素产生光谱数据噪声[6][7], 影响水质光谱数据的正确分析。对建立水质 COD 预测模型前, 先要进行光谱数据去噪处理来提高分析数据的精度。如何通过算法搜寻全局最优阈值便是去噪研究的核心环节, 仅凭单一方法难以完成这项工作, 故本文采用蚁群觅食行为和灰狼狩猎方式的算法来优化小波阈值。

各种优化算法在预测模型方面应用较多, 对紫外可见光谱水质 COD 数据的去噪研究应用较少。基于此, 本文提出两种优化算法搜寻最优小波阈值的去噪方法, 先进行光谱数据去噪, 并与中值滤波、均值滤波、滑动平均、小波阈值作对比分析。实验结果表明, 自适应全局搜寻最优阈值能体现出针对光谱数据卓越的去噪效果。

2. 两种算法优化小波阈值去噪原理

2.1. 蚁群优化算法原理

蚁群优化算法(Ant Colony Optimization, ACO)是一种蚂蚁群获取食物而找寻最短路径的智能启发式搜索算法,成千上万的蚂蚁组成庞大团队留下独特的信息传递标记——信息素。唐标[8]等用优化型蚁群算法对智能巡检机器人最优行走路径进行电力设备故障图片的识别。Wang Qingyang [9]等采用连续蚁群算法多次迭代选择影像最优镶嵌接缝线网络,节省了生成数据集时间。

ACO 算法原理[10]是计算蚂蚁选择下个城市的概率和走完每个城市边缘信息素增量两个关键参数,其表达式为:

$$P_{i,j}^n(m) = \begin{cases} \frac{T_{i,j}^x(m) \cdot \lambda_{i,j}^y(m)}{\sum_{s \in allowed_n} T_{i,s}^n(m) \cdot \lambda_{i,s}^y(m)}, & j \in allowed_n \\ 0, & otherwise \end{cases} \quad (1)$$

式中: $T_{i,j}^x(m)$ 是 i 、 j 城市间 $e_{i,j}$ 边信息素因数, $\lambda_{i,j}^y(m)$ 为 i 、 j 城市间能见度因数。信息素启发因子 x 和期望因子 y 表示信息重要程度和城市间距离程度。

$$\begin{cases} T_{ij}(t,t+1) = \rho T_{ij}(t) + \Delta T_{ij}(t,t+1) \\ \Delta T_{ij}(t,t+1) = \sum_{n=1}^m \Delta T_{ij}^n(t,t+1) \end{cases} \quad (2)$$

其中: $\Delta T_{ij}^k(t,t+1)$ 是蚂蚁 n 对边 e_{ij} 上信息素贡献增量, $\Delta T_{ij}(t,t+1)$ 表示所有蚂蚁通过边 e_{ij} 对边 e_{ij} 累积的信息素贡献量。

2.2. 灰狼优化算法原理

灰狼[11]狩猎的过程主要有三个方面:追踪和接近猎物、包围猎物、攻击猎物。灰狼优化算法(Grey Wolf Optimization, GWO)原理[12]如下:现有灰狼 a 、 b 、 c 、 d , a 狼位置代表最佳位置,也 GWO 算法求全局最优解问题相对应, b 狼位置表示 GWO 算法中的次级优解, c 狼位置在 GWO 算法中表现为第三优解, d 狼位置根据前面三狼所占位置后更替自己位置。若灰狼种群数为 N , 种群集合为 V 维, 则包围猎物行为按如下两式定义:

$$\begin{cases} V = |X_e^w(m) - W \cdot X_f^w(m)| \\ X_e^w(m+1) = X_e^w(m) - UV \end{cases} \quad (3)$$

式(3)表示计算灰狼捕猎距离和计算灰狼捕猎位置, 为迭代次数, 猎物位置为, 为灰狼当前位置, 表示灰狼接近猎物或远离猎物的收敛因子, 代表灰狼捕食猎物过程中受到外界干扰。灰狼改变位置计算式为:

$$\begin{cases} X_{e,a}^w(m+1) = X_a^w(m+1) - U_1 \cdot |X_e^w(m) - W_1 \cdot X_a^w(m)| \\ X_{e,b}^w(m+1) = X_b^w(m+1) - U_2 \cdot |X_e^w(m) - W_2 \cdot X_b^w(m)| \\ X_{e,c}^w(m+1) = X_c^w(m+1) - U_3 \cdot |X_e^w(m) - W_3 \cdot X_c^w(m)| \\ X_e^w(m+1) = \frac{X_{e,a}^w(m+1) + X_{e,b}^w(m+1) + X_{e,c}^w(m+1)}{3} \end{cases} \quad (4)$$

式(4)表示替换前四头灰狼的当前位置, 用灰狼优化算法对紫外可见光谱水质数据进行全局寻优得到最优小波阈值。

3. 采集数据实验

为证明蚁群优化小波阈值算法(ACOWT)和灰狼优化小波阈值算法(GWOWT)在紫外可见光谱数据方面的去噪可行性,实验采用 Sarspec LS-DWHP 大功率氙卤灯光源和 Maya-2000PRO 光谱仪组成采集系统来测量 COD 浓度为 20 mg/L 的邻苯二甲酸氢钾标液,总量为 3 ml 去离子水按等量均分为 30 组,每次均匀加入 10 μ l 去稀释邻苯二甲酸氢钾标液。在收集邻苯标液光谱数据中基础上加入高斯白噪声,实现去噪效果对比分析。图 1 中四幅图片分别为 30 组邻苯标液光谱图、某内陆河水水质 COD 光谱图、ACOWT 算法对 30 组邻苯标液去噪图、GWOWT 算法对 30 组邻苯标液去噪图。

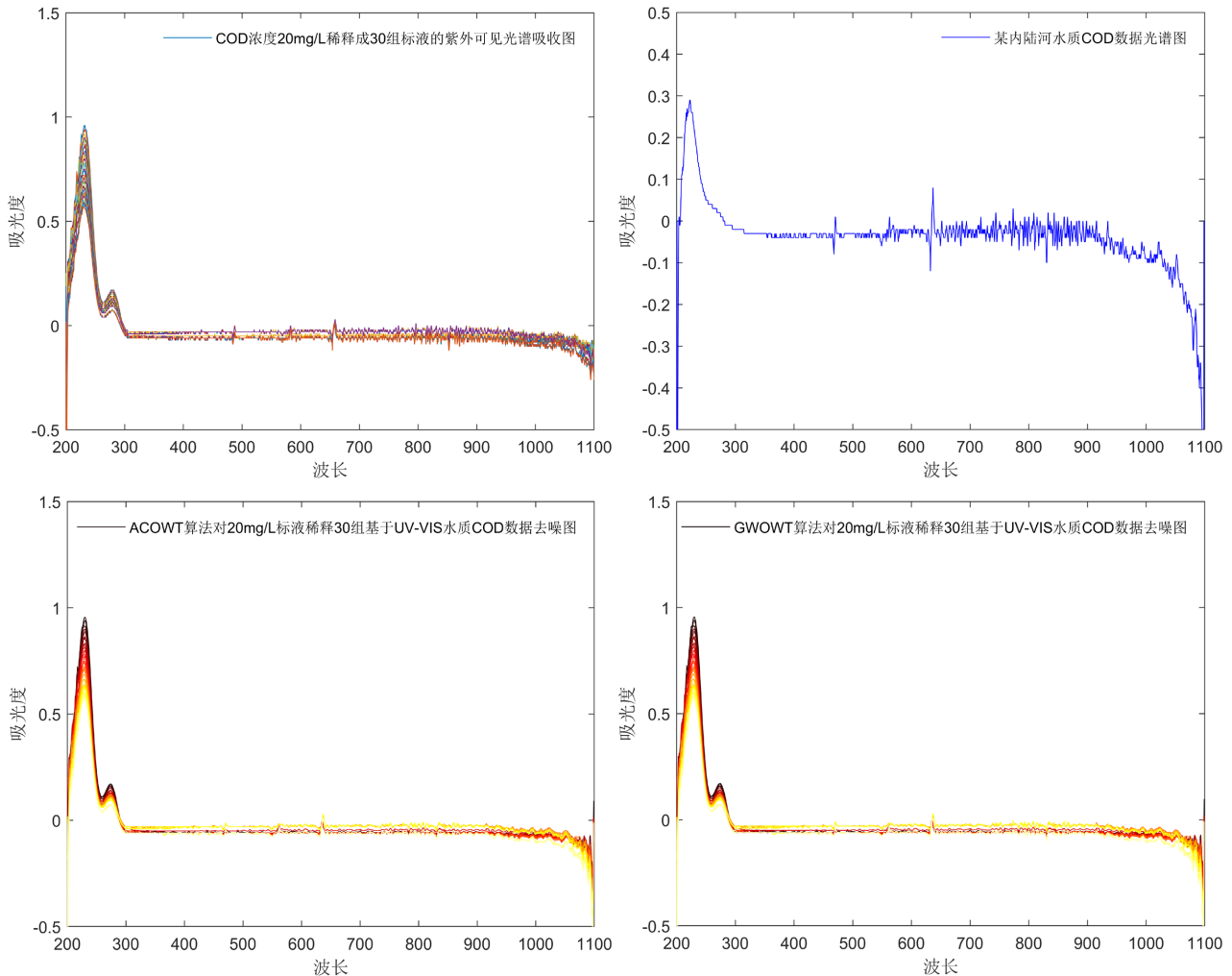


Figure 1. Experimental data acquisition diagram and denoising diagram

图 1. 实验数据采集图与去噪图

4. 结果与分析

与传统算法进行去噪对比后, GWOWT 算法和 ACOWT 算法不仅能有效去除噪声,还能保存光谱数据细节信息。为更好评价各种算法引入了信噪比 SNR、峰值信噪比 PSNR、标准偏差 STD 和均方差 MSE, 信噪比和峰值信噪比越大, 就说明了该算法的去噪质量好, 标准偏差和均方差越小, 就证明了该算法的去噪能力强。其表达式为:

$$\left\{ \begin{array}{l} \text{PSNR} = 10 \lg \frac{M \cdot X_{\max}^2(i)}{\sum_{i=1}^N [X_{sf}(i) - \overline{X}(i)]^2} \\ \text{STD} = \sqrt{\frac{\sum_{i=1}^N |X_{sf}(i) - \overline{X}(i)|}{N-1}} \end{array} \right. \left\{ \begin{array}{l} \text{SNR} = 10 \lg \frac{\sum_{i=1}^N X_{sf}^2(i)}{\sum_{i=1}^N [X_{sf}(i) - \overline{X}(i)]^2} \\ \text{MSE} = \sqrt{\frac{\sum_{i=1}^N [X_{sf}(i) - \overline{X}(i)]^2}{N}} \end{array} \right. \quad (5)$$

式中： $X_{sf}(i)$ 、 $X_{\max}(i)$ 、 $\overline{X}(i)$ 分别为去噪后的数据、样本去噪前的最大数据和每个样本数据的平均值， M 、 N 分别为为一组数据的长度和样本总量。表 1 对每种算法进行去噪效果对比。

Table 1. Analysis of denoising effect of 30 groups of water quality COD data of o-benzene standard solution

表 1. 邻苯标液 30 组水质 COD 数据去噪效果分析

去噪算法	SNR (dB)	STD	PSNR (dB)	MSE
中值滤波	18.235419	0.018215	41.440734	0.000355
均值滤波	17.979178	0.018478	42.184493	0.000365
滑动平均	13.376979	0.030340	40.582294	0.000959
小波阈值	16.405054	0.021793	43.610369	0.000504
ACOWT	28.416273	0.005435	45.621588	0.000032
GWOWT	29.677177	0.004728	46.88345	0.000024

如表 1 所示，GWOWT 算法对于紫外可见光谱数据体现出卓越的去噪效果，与中值滤波算法、均值滤波算法、滑动平均算法、小波阈值算法、ACOWT 算法比较，GWOWT 算法的信噪比提高了 11.441758、11.697999、16.300198、13.272123、1.260904 dB，标准偏差减少了 0.013487、0.013750、0.025612、0.017065、0.000707，峰值信噪比提高了 5.442717、4.698958、6.301157、3.273082、1.261863 dB，均方差降低了 0.000331、0.000341、0.000935、0.000480、0.000008。GWOWT 算法抑制了较多噪声也保留了光谱特征信息，误差进一步得到压减。不难看出，基于灰狼优化算法的小波阈值去噪方法更适用于紫外可见光谱水质 COD 数据的去噪分析。

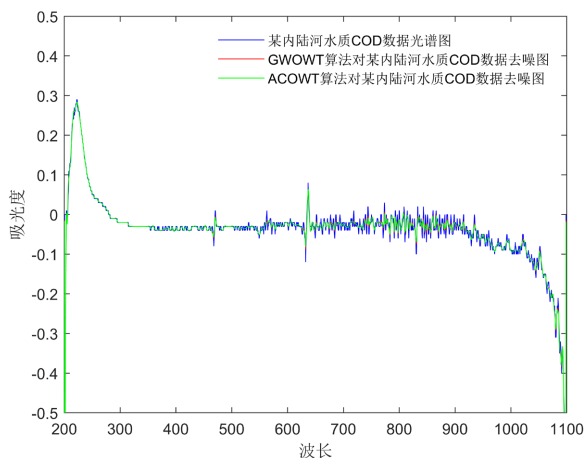


Figure 2. Comparison of denoising effects of two optimization algorithms for an inland river

图 2. 某内陆河采用两种优化算法去噪效果对比

图 2 是两种优化算法对某内陆河水质的光谱 COD 数据去噪效果图，某内陆河水在光谱波段 500~1000 nm 之

间存在大量跳变不稳定噪声, ACOWT 算法虽有效去除噪声但也丢失了部分光谱数据的细节特征, GWOWT 算法有效抑制了噪声和未产生光谱数据失真现象, 保留了细节光谱数据特征。表 2 中 GWOWT 算法的信噪比和峰值信噪比更高, 标准偏差和均方差更小, 证明了两种优化算法用于去噪的实用性, 也说明了 GWOWT 算法更适用于地表水质数据的噪声预处理。为建立水质 COD 预测模型提供良好的模型输入数据, 能进一步分析水体 COD 含量和治理水污染提供检测支撑。

Table 2. Denoising analysis of COD data of surface water quality by two optimization algorithms

表 2. 两种优化算法对地表水质 COD 数据的去噪分析

去噪算法	SNR (dB)	STD	PSNR (dB)	MSE
ACOWT	19.402277	0.009934	40.057108	0.000099
GWOWT	20756661	0.008500	41.411493	0.000072

5. 结论

紫外可见光谱采集系统的水质光谱 COD 数据中含有噪声, 针对小波阈值去噪不符合预期效果的问题, 提出两种优化小波阈值算法, 主要结论为:

1) 对邻苯标液标液在加入高斯白噪声基础上后进行去噪。采用了信噪比、峰值信噪比、标准偏差和均方差的四个量化指标来评价每个算法的去噪效果, 结果表明了灰狼优化小波阈值算法对 30 组邻苯溶液的去噪效果好于其他五种去噪算法, 也证实了灰狼优化小波阈值算法可用于地表水质的 COD 数据去噪的可行性。

2) 两种优化小波阈值算法在光谱特征信息得到保留的前提下, 有效抑制了噪声。对采集的某内陆河水质进行去噪测试结果证实了灰狼优化小波阈值算法的良好去噪效果, 为降低水质检测系统成本和提高检测速度提供新的解决办法。

参考文献

- [1] 楼显盛, 陈安瑶, 张研, 林强, 罗安程, 梁志伟. 浙江省农村生活污水电导率与水质指标的响应关系分析[J]. 环境监测管理与技术, 2022, 34(1): 64-67.
LOU Xiansheng, CHEN Anyao, ZHANG Yan, LIN Qiang, LUO Ancheng and LIANG Zhiwei. Analysis on the response relationship between conductivity and water quality index of rural domestic sewage in Zhejiang Province. Environmental Monitoring Management and Technology, 2022, 34 (1): 64-67. (in Chinese)
- [2] 孟庆龙, 尚静, 黄人帅, 撒雄星, 张艳. 不同成熟度猕猴桃糖度紫外可见光谱检测[J]. 食品研究与开发, 2020, 41(21): 167-172.
MENG Qinglong, SHANG Jing, HUANG Renshuai, SA Xiongxing and ZHANG Yan. Determination of sugar content of kiwi-fruit with different maturity by UV-VIS spectroscopy. Food Research and Development, 2020, 41(21): 167-172. (in Chinese)
- [3] 陈鑫, 周真, 杨旭, 郭抗抗, 李迪星. 紫外-可见光谱法定量检测牛奶菌群总数的研究[J]. 分析测试学报, 2017, 36(9): 1109-1113.
CHEN Xin, ZHOU Zhen, YANG Xu, GUO Kangkang and LI Dixing. Study on quantitative detection of total number of milk flora by UV-VIS spectroscopy. Journal of Analysis and Testing, 2017, 36(9): 1109-1113. (in Chinese)
- [4] 张金秋, 沈锡田. 热处理绿黄色磷灰石的紫外-可见光谱[J]. 宝石和宝石学杂志, 2018, 20(S1): 33-34.
ZHANG Jinqiu, SHEN Xitian. UV-VIS spectra of heat-treated green yellow apatite. Journal of Gemology and Gemology, 2018, 20(S1): 33-34. (in Chinese)
- [5] 周曙光, 甄汉深, 甄丹丹, 韦锦梅, 潘洪平. 玉米苞叶药材提取液紫外-可见光谱鉴别的初步研究[J]. 中国临床新医学, 2018, 11(9): 899-902.
ZHOU Shuguang, ZHEN Hanshen, ZHEN Dandan, WEI Jinmei and PAN Hongping. Preliminary study on identification of corn bract extract by UV-VIS spectroscopy. China New Clinical Medicine, 2018, 11(9): 899-902. (in Chinese)
- [6] 汤斌, 魏彪, 毛本将, 赵敬晓, 冯鹏. 紫外-可见吸收光谱法水质检测系统的噪声分析与处理研究[J]. 激光与光电子学进展, 2014, 51(4): 201-207.
TANG Bin, WEI Biao, MAO benjiang, ZHAO Jingxiao and FENG Peng. Study on noise analysis and treatment of water quality detection system by UV-VIS absorption spectroscopy. Progress in Laser and Optoelectronics, 2014, 51(4): 201-207. (in Chinese)

nese)

- [7] 孙然然, 林玉娥, 梁兴柱. 一种新的阈值和阈值函数的小波去噪算法[J]. 佳木斯大学学报(自然科学版), 2021, 171(2): 63-66.
SUN Ranran, LIN Yu'e and LIANG Xingzhu. A new wavelet denoising algorithm based on threshold and threshold function. Journal of Jiamusi University (Natural Science Edition), 2021, 171(2): 63-66. (in Chinese)
- [8] 唐标, 沈映泉, 黄绪勇, 肖登宇, 张粥. 基于改进型蚁群算法和图像识别的变电站机器人路径规划和设备缺陷识别研究[J]. 制造业自动化, 2022, 44(2): 46-52.
TANG Biao, SHEN Yingquan, HUANG Xuyong, XIAO Dengyu and ZHANG Zhou. Research on substation robot path planning and equipment defect recognition based on improved ant colony algorithm and image recognition. Manufacturing Automation, 2022, 44(2): 46-52. (in Chinese)
- [9] WANG, Q. Y., ZHOU, G. Q., SONG, R. H., XIE, Y. F., LUO, M. Y. and YUE, T. Continuous space ant colony algorithm for automatic selection of orthophoto mosaic seam-line network. Journal of Photogrammetry and Remote Sensing, 2022, 186(10): 201-217.
- [10] 吕琳, 尉永清, 任敏, 潘晓. 基于蚁群优化算法的凝聚型层次聚类[J]. 计算机应用研究, 2017, 34(1): 114-117.
LV Lin, WEI Yongqing, REN Min and PAN Xiao. Agglomerative hierarchical clustering based on ant colony optimization algorithm. Computer Application Research, 2017, 34(1): 114-117. (in Chinese)
- [11] 张英芝, 朱继微, 刘津彤, 翟粉莉, 牟黎明. 改进灰狼算法优化灰色预测模型在数控机床中的应用[J]. 制造技术与机床, 2022(3): 127-131.
ZHANG Yingzhi, ZHU Jiwei, LIU Jintong, ZHAI Fenli and MOU Liming. Application of improved grey wolf algorithm to optimize grey prediction model in NC machine tool. Manufacturing Technology and Machine Tools, 2022(3): 127-131. (in Chinese)
- [12] MOHAMMED, A. E., ABOBAKR, A. S. and NEHAL, E. Hybrid grey wolf optimization-based Gaussian process regression model for simulating deterioration behavior of highway tunnel components. Processes, 2021, 10(1): 36.
<https://doi.org/10.3390/pr10010036>