



基于紫外可见吸收光谱的水质检测算法研究

林春伟, 郭永洪, 何金龙

(中国计量大学, 浙江 杭州 310018)

摘要: 为实时有效地检测地表水中硝酸根离子和亚硝酸根离子的变化过程, 提出一种基于紫外可见吸收光谱的水质检测算法。针对水质光谱数据受到干扰易出现波动误差的问题, 采用小波变换对其进行分解以滤除高频噪声, 并通过主成分分析对数据特征进行降维以防止模型复杂度较高导致过拟合。水质光谱数据经预处理后采用支持向量机对其进行建模, 通过非线性自适应调整变异收缩因子对差分进化算法进行改进, 并采用改进差分进化算法对水质预测模型进行参数优化。通过与采用其他常用算法所建模型进行对比分析, 实验结果表明: 基于该算法所建的硝酸根离子和亚硝酸根离子模型具有更高的预测精度, 且其能够以更快的收敛速度使模型达到全局最优。

关键词: 水质检测; 光谱分析; 支持向量机; 改进差分进化算法; 小波变换; 主成分分析

中图分类号: TP391

文献标志码: A

文章编号: 1674-5124(2019)05-0079-06

Research on water quality detection algorithm based on UV-vis absorption spectral

LIN Chunwei, GUO Yonghong, HE Jinlong

(China Jiliang University, Hangzhou 310018, China)

Abstract: In order to detect the change process of nitrate ions and nitrite ions in surface water effectively and in real time, a water quality detection algorithm based on UV-vis absorption spectra was proposed. Aiming at the problem that water quality spectral data are disturbed and easy to appear wave error, wavelet transform is used to decompose the water quality spectral data to filter out the high frequency noise, and the principal component analysis is used to reduce the dimension of data features to prevent the over-fitting caused by high complexity of the model. The water quality spectral data are modeled by support vector machine after pretreatment, and the differential evolution algorithm is improved by nonlinear adaptive adjustment of variation shrinkage factor, and the parameters of the water quality prediction model are optimized by the improved differential evolution algorithm. By comparing with the models built by other common algorithms, the experimental results show that this algorithm can make the nitrate ion and nitrite ion models have higher prediction accuracy, and it can make the model achieve global optimization with faster convergence speed.

Keywords: water quality detection; spectral analysis; support vector machine; improved differential evolution algorithm; wavelet transform; principal component analysis

收稿日期: 2018-09-06; 收到修改稿日期: 2018-11-09

基金项目: 浙江省自然科学基金(Y14F010075)

作者简介: 林春伟(1994-), 男, 河南鹤壁市人, 硕士研究生, 专业方向为检测技术。

0 引言

传统水质检测常采用化学法通过人工采集水样进行实验分析,随着物联网技术的快速发展,能够实时有效检测整个水质变化过程的水质在线监测系统已成为该领域的重要发展趋势^[1]。化学法虽然分析精度较高,但由于其操作复杂且测量周期长而难以实现水质在线监测。紫外可见吸收光谱法则通过对标准水样的吸收光谱进行分析建模,无需添加化学试剂即可直接测定水样中各组分浓度,具有检测速度快和无二次污染等优点,故该方法广泛应用于水质在线监测系统研究中^[2]。

基于紫外可见吸收光谱进行水质检测的关键在于光谱数据的建模与优化。当水样中各组分浓度较低时,根据朗伯比尔定律其浓度与吸光度呈线性关系,此时通常采用多元线性回归^[3]或偏最小二乘法^[4]对光谱数据进行建模;但当水样中各组分浓度较高时,各吸光粒子距离变小导致其电荷分布相互影响,故水样中各组分浓度与吸光度呈非线性关系,此时通常采用神经网络^[5]或支持向量机(SVM)^[6]对光谱数据进行建模。在地表水质检测的应用研究中,由于水样中各组分浓度变化范围较大,故通常采用非线性建模算法对光谱数据进行处理。其中,神经网络虽然具有强大的非线性拟合能力,但其存在学习速度慢、易陷入局部最优和调参复杂等问题;而支持向量机基于统计学习理论具有更强的泛化能力,且其需要优化的参数也相对较少,故支持向量机在光谱分析领域具有极为广泛的应用^[7]。在模型参数优化方面,Huo等^[8]和Wang等^[9]分别采用遗传算法和粒子群算法优化支持向量机建立水质预测模型,通过实验分析均得出了其相对于BP神经网络具有更高预测精度的结论。而差分进化算法虽然在水质预测模型优化方面应用较少,但Civicioglu等^[10]通过标准测试函数对上述3种优化算法进行了对比分析,实验结果表明在收敛速度方面差分进化算法表现更优。

为使水质预测模型具有较高的预测精度和收敛速度,本文采用差分进化算法优化支持向量机对光谱数据进行建模。为防止模型陷入局部最优并进一步提高其收敛速度,通过非线性与自适应调整变异缩放因子相结合的策略对差分进化算法进行改进。由于光谱数据大多带有扰动噪声,且高维特征会增加模型复杂度从而导致过拟合,故本文在对光谱数

据进行建模优化之前,分别采用小波变换和主成分分析对其进行去噪和降维以进一步提高模型的预测精度。

1 数据采集与预处理

1.1 数据采集

虽然物质对光的吸收具有选择性,但其吸收光谱是带状光谱而非单一的特征谱线,因此水样中各组分吸收光谱会存在相互重叠和干扰的问题。本文根据《地表水环境质量标准》所述,选取硝酸根离子、亚硝酸根离子和氯离子作为实验对象采集数据,以上3种离子在3类及以上地表水中的浓度检测限值分别为10 mg/L、1 mg/L、250 mg/L。其中,硝酸根离子和亚硝酸根离子是引起水体富营养化的主要离子,而氯离子是地表水中含量最多的离子之一,且上述3种离子的吸光波段均有重叠,故将硝酸根离子和亚硝酸根离子作为检测的目标离子,而氯离子作为检测的干扰离子。

由于钠离子吸收光谱与上述3种离子均不重叠,故以硝酸钠(分析纯)、亚硝酸钠(分析纯)、氯化钠(分析纯)和去离子水作为实验材料。硝酸根离子、亚硝酸根离子和氯离子的浓度梯度分别设置为1 mg/L、0.1 mg/L、25 mg/L,根据此浓度梯度配制以上3种离子的单组分标准液与多组分混合液,使用紫外可见分光光度计(UV2550)对其进行光谱扫描。其中,单组分标准液的吸收光谱主要用于研究各离子吸光特性;根据正交试验设计配制的100种多组分混合液的吸收光谱作为模型训练样本,另配制与训练样本水样中各离子浓度配比不同的30种多组分混合液的吸收光谱作为模型测试样本。当水样中硝酸根离子、亚硝酸根离子和氯离子浓度分别为6 mg/L、1 mg/L、100 mg/L时,以上3种离子单组分标准液和多组分混合液的吸收光谱如图1所示。

1.2 数据预处理

1.2.1 光谱去噪

在使用紫外可见分光光度计进行光谱扫描的过程中会由于机械抖动使数据产生一定误差,且由图1可知,水样中各组分吸光度越大,由于仪器检测范围和精度等原因造成其数据波动程度越剧烈。小波变换是一种时间窗和频率窗均可改变的信号分析方法,其通过小波函数的伸缩和平移变换对信号进行多尺度细化分析,可以聚焦到信号的任意局部特征^[11]。为减小实验误差对水质预测模型的影响,

本文采用小波变换对光谱数据进行分解以滤除高频噪声。光谱数据经小波变换处理前后对比如图 2 所示。

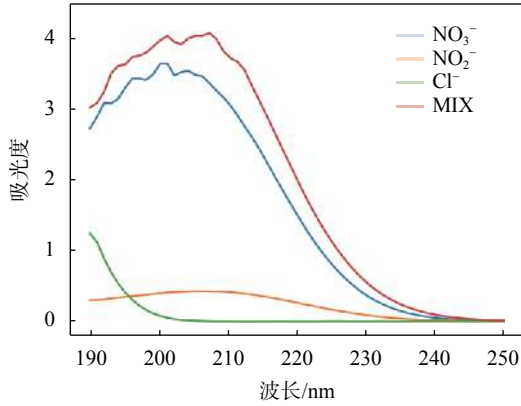


图 1 单组分吸收光谱和多组分吸收光谱

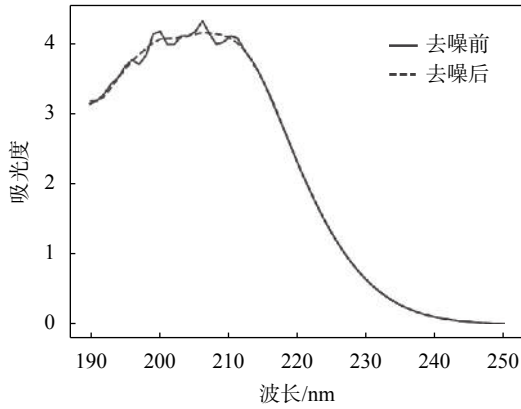


图 2 光谱去噪前后对比

1.2.2 特征降维

水样中各组分吸收光谱相互重叠和干扰的情况造成有时无法使用吸收峰波进行建模,且单一谱线无法全面反映吸光物质的特征信息,故通常对全光谱数据提取特征波段进行建模。较高的特征维数会增加模型复杂度导致其过拟合,而主成分分析将数据进行正交变换,通过少数几个主成分尽可能多的保留原始变量信息,不但可以降低数据特征的维数,还可以减少冗余信息对模型的影响^[12]。硝酸根离子和亚硝酸根离子的吸光波段均集中在 190~250 nm,而氯离子的吸光波段则集中在 190~210 nm。对于硝酸根离子,为减小氯离子对其吸光特性的影响,选取 216~235 nm 作为特征波段,对以上 20 维数据运用主成分分析进行降维,选取前 3 个主成分作为模型输入变量,其累计解释方差总和为 99.9%。对于亚硝酸根离子,由于其在浓度检测范围内吸光

度较小,从而导致其在多组分吸收光谱中的特征信息被硝酸根离子完全覆盖,为减小硝酸根离子对其吸光特性的影响,直接选取吸收峰波 210 nm 作为特征波长,并选取 220 nm 和 230 nm 作为补偿波长,以上 3 个维度即为模型输入变量。

2 算法理论

2.1 支持向量机

支持向量机是一种基于统计学习理论和结构风险最小化原理的监督学习算法,其基本模型是定义在特征空间上间隔最大的线性分类器,当引入核函数将数据非线性映射到高维空间之后,支持向量机也可用于非线性分类^[13]。除此之外,支持向量机也可用于解决回归问题,其基本思想都是通过间隔最大化学习得到最优超平面: $f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$ 。

支持向量机解决回归问题根据模型输出 $f(\mathbf{x}_i)$ 与真实值 y_i 之间的差别来计算损失,其允许 $f(\mathbf{x}_i)$ 与 y_i 之间最多有 ε 的偏差,仅当 $f(\mathbf{x}_i) - y_i > \varepsilon$ 时才计算损失,当引入松弛变量 ξ_i^+, ξ_i^- 后,其优化问题为:

$$\begin{aligned} \min & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N (\xi_i^+ + \xi_i^-) \\ \text{s.t.} & f(\mathbf{x}_i) - y_i \leq \varepsilon + \xi_i^+, \\ & f(\mathbf{x}_i) - y_i \geq \varepsilon - \xi_i^-, \\ & \xi_i^+ \geq 0, \xi_i^- \geq 0, i=1, 2, \dots, N \end{aligned} \quad (1)$$

当引入拉格朗日乘子 α^+, α^- 后,根据拉格朗日对偶性,原始问题转化为求解 α^+, α^- 的凸二次规划问题,其最终优化问题为:

$$\begin{aligned} \min & \sum_{i=1}^N [y_i(\alpha_i^- - \alpha_i^+) - \varepsilon(\alpha_i^- + \alpha_i^+)] - \\ & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i^- - \alpha_i^+)(\alpha_j^- - \alpha_j^+) \mathbf{x}_i^T \mathbf{x}_j \\ \text{s.t.} & \sum_{i=1}^N (\alpha_i^- - \alpha_i^+) = 0, 0 \leq \alpha_i^+, \alpha_i^- \leq C \end{aligned} \quad (2)$$

若求解以上问题得到最终解为 $\alpha^+ = (\alpha_1^+, \alpha_2^+, \dots, \alpha_N^+)$, $\alpha^- = (\alpha_1^-, \alpha_2^-, \dots, \alpha_N^-)$, 则有:

$$f(\mathbf{x}) = \sum_{i=1}^N (\alpha_i^- - \alpha_i^+) \mathbf{x}_i^T \mathbf{x} + b \quad (3)$$

若通过引入核函数解决非线性问题,则上式可以表示为:

$$f(\mathbf{x}) = \sum_{i=1}^N (\alpha_i^- - \alpha_i^+) K(\mathbf{x}_i, \mathbf{x}) + b \quad (4)$$

2.2 差分进化算法及改进

差分进化算法(DE)是一种基于群体差异的并行优化算法,该算法与遗传算法(GA)类似,都包括变异、交叉和选择等操作,而粒子群算法(PSO)是通过模仿生物个体间竞争与合作的启发式群体智能来进行优化搜索^[14]。差分进化算法通过差分策略实现个体变异,这也是区别于遗传算法的重要标志,该算法具体实现流程如下:

1) 种群初始化,在解空间中随机均匀地产生 m 个个体作为第 0 代种群,其中每个个体包含 n 个变量,标记为:

$$\mathbf{X}_i(0) = [x_{i,1}(0), x_{i,2}(0), \dots, x_{i,n}(0)], i = 1, 2, \dots, m \quad (5)$$

2) 变异,在第 k 次迭代中,从种群中随机选择 3 个个体 $\mathbf{X}_{j_1}(k), \mathbf{X}_{j_2}(k), \mathbf{X}_{j_3}(k)$, 且 $j_1 \neq j_2 \neq j_3 \neq i$, 则其变异规则为:

$$\mathbf{H}_i(k) = \mathbf{X}_{j_1}(k) + F \cdot (\mathbf{X}_{j_2}(k) - \mathbf{X}_{j_3}(k)), i = 1, 2, \dots, m \quad (6)$$

其中, $\mathbf{X}_{j_2}(k) - \mathbf{X}_{j_3}(k)$ 为差分向量, F 为变异缩放因子,用于控制差分向量的影响力。

3) 交叉,在 $[0, 1]$ 区间内产生 m 个随机数,确定交叉概率因子 CR,其交叉方法如下:

$$\mathbf{V}_i(k) = \begin{cases} \mathbf{H}_i(k), \text{rand}(0, 1) \leq \text{CR} \\ \mathbf{X}_i(k), \text{else} \end{cases}, i = 1, 2, \dots, m \quad (7)$$

其中, $\text{CR} \in [0, 1]$, $\text{rand}(0, 1)$ 是 $[0, 1]$ 区间上服从均匀分布的随机数。

4) 选择,根据适应度函数选择 $\mathbf{V}_i(k)$ 或 $\mathbf{X}_i(k)$ 作为 $\mathbf{X}_i(k+1)$,其选择方法如下:

$$\mathbf{X}_i(k+1) = \begin{cases} \mathbf{V}_i(k), f(\mathbf{V}_i(k)) \leq f(\mathbf{X}_i(k)) \\ \mathbf{X}_i(k), \text{else} \end{cases}, i = 1, 2, \dots, m \quad (8)$$

5) 判断算法是否达到终止条件。若是,将最佳个体作为最优解输出;若否,则转至步骤 2) 进行第 $k+1$ 次迭代。

差分进化算法通过差分策略实现个体变异,变异收缩因子 F 主要用于控制差分向量对变异个体的影响力, F 值较大时可以通过增加个体间差异提升种群多样性使算法达到全局最优; F 值较小时可以增强个体的局部开发能力并加快其收敛速度,故针对差分进化算法的改进一般采取非线性递减变异收缩因子的策略。但该改进策略还存在一定问题,比如当算法前期没有搜索到较优的解,后期减小 F 值会使算法陷入局部最优;而对于适应度值较高的个体,算法前期较大的 F 值也会影响其收敛速度。

针对以上问题,本文提出一种基于非线性递减与自适应调整变异收缩因子相结合的改进差分进化算法(IDE)。改进后的差分进化算法在每一次迭代后计算种群本代个体适应度值的平均值,当个体适应度值小于种群平均适应度值时,基础 F 值通过加上非线性变化因子以增强算法的全局搜索能力;当个体适应度值大于种群平均适应度值时,基础 F 值通过减去非线性变化因子以增强算法的局部开发能力。改进差分进化算法中 F 值的具体调整方法如下:

$$F = \begin{cases} \frac{1}{2}(F_{\max} + F_{\min}) + \frac{1}{2}(F_{\max} - F_{\min})e^{-\frac{t}{T}}, f_i \leq \bar{f} \\ \frac{1}{2}(F_{\max} + F_{\min}) - \frac{1}{2}(F_{\max} - F_{\min})e^{-\frac{t}{T}}, f_i > \bar{f} \end{cases} \quad (9)$$

其中, T 为最大迭代次数, t 为当前迭代次数, f_i 为种群中第 i 个个体的适应度值, \bar{f} 为种群平均适应度值, F_{\max} 为最大变异收缩因子, F_{\min} 为最小变异收缩因子。

改进后的差分进化算法可以根据个体适应度值动态调整变异收缩因子,对于种群中较优个体可以加快其收敛速度,而对于种群中较差个体可以通过扩大搜索范围防止其陷入局部最优。

3 实验与分析

基于紫外可见吸收光谱构建水质预测模型主要使用紫外可见分光光度计采集标准水样的光谱数据,采用小波变换对光谱数据进行分解去噪以减小数据误差的影响,并对水样中各组分进行特征降维以确定模型输入变量,通过以上数据预处理操作之后,构建基于支持向量机的水质预测模型,并使用改进差分进化算法对模型进行参数优化以使其达到最优。水质预测模型架构如图 3 所示。

若水样中各组分吸光波段相互重叠,当水样中各组分浓度较高时,多组分吸收光谱中各波长的吸光度即为各组分在该波长的吸光度的非线性叠加,故本文选取高斯核函数作为支持向量机的核函数。由于经过特征降维后硝酸根离子和亚硝酸根离子的输入变量不同,故采用支持向量机对两种目标离子分别进行建模,并采用改进差分进化算法对模型进行参数优化,支持向量机中待优化的参数分别为惩罚系数 C 与核参数 γ 。为防止模型过拟合造成预测性能下降,限定惩罚系数 C 与核参数 γ 的搜索范围为 $(0, 1000)$ 。设置改进差分进化算法中种群个体数为 20,交叉概率 CR 为 0.3,最大变异收缩因子 F_{\max} 为 0.9,最小变异收缩因子 F_{\min} 为 0.1。选取拟

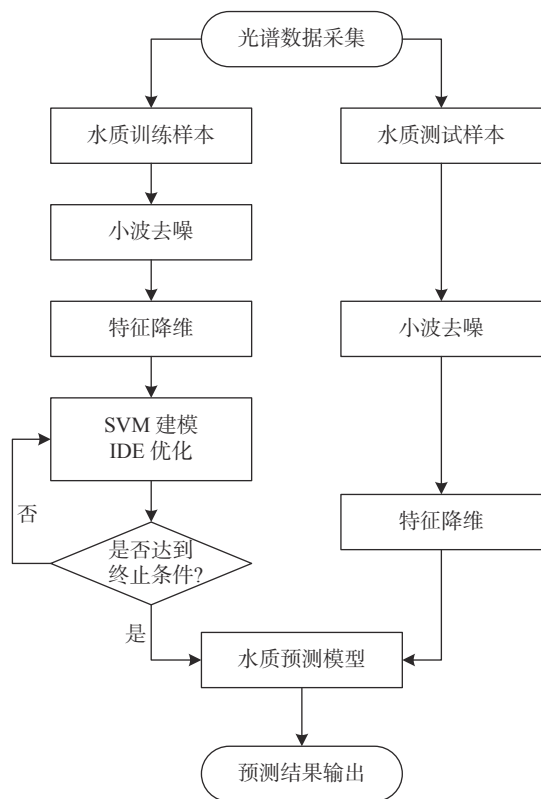


图 3 水质预测模型架构

合优度 R^2 作为评估模型好坏的适应度函数。

为测试本文基于 IDE-SVM 算法所建模型的预测精度与学习速度, 设定模型训练的最大迭代次数为 100, 选取模型训练时间 t 、模型分别在训练样本和测试样本上的拟合优度 R^2 与均方根误差 RMSE 等作为性能指标, 并分别与基于 DE-SVM、PSO-SVM 和 GA-SVM 算法所建模型在相同种群个体数和迭代次数下的性能指标作对比分析。其中, 差分进化算法中交叉概率 CR 为 0.3, 变异收缩因子 F 为 0.5; 粒子群算法中学习因子 w 为 2, 惯性权重 c 为 1; 遗传算法中交叉概率 pc 为 0.8, 变异概率 pm 为 0.2。本文测试环境为: Intel i5-6300HQ CPU 2.30 GHz, RAM 8 GB, Pycharm5.0.3。硝酸根离子和亚硝酸根离子分别基于 4 种算法所建模型的最优参数和性能指标对比如表 1 和表 2 所示。

硝酸根离子由于在浓度检测范围内吸光度较大, 故其在多组分吸收光谱中的特征信息比较明显; 而亚硝酸根离子由于其在多组分吸收光谱中的特征信息被硝酸根离子覆盖, 故针对其所建模型的预测精度会受到一定影响。由表 1 和表 2 可知, 针对以上两种目标离子分别基于 4 种算法所建模型在训练样本上均能取得较高的预测精度, 但相比之下

表 1 4 种模型最优参数与性能指标对比 (硝酸根离子)

| 参数指标 | IDE-SVM | DE-SVM | PSO-SVM | GA-SVM |
|---------------------|---------|---------|---------|---------|
| C | 18.61 | 18.39 | 6.37 | 86.97 |
| γ | 9.02 | 9.16 | 89.47 | 11.24 |
| t/s | 10.72 | 10.81 | 15.24 | 19.68 |
| $R^2(\text{train})$ | 0.997 9 | 0.997 8 | 0.997 5 | 0.996 3 |
| $R^2(\text{test})$ | 0.996 9 | 0.996 2 | 0.995 8 | 0.995 5 |
| RMSE(train) | 0.132 1 | 0.135 4 | 0.142 5 | 0.175 1 |
| RMSE(test) | 0.159 6 | 0.176 2 | 0.185 4 | 0.191 5 |

表 2 4 种模型最优参数与性能指标对比 (亚硝酸根离子)

| 参数指标 | IDE-SVM | DE-SVM | PSO-SVM | GA-SVM |
|---------------------|---------|---------|---------|---------|
| C | 1.17 | 1.29 | 1.59 | 3.36 |
| γ | 81.17 | 77.62 | 28.57 | 41.27 |
| t/s | 10.37 | 10.19 | 14.65 | 18.92 |
| $R^2(\text{train})$ | 0.970 2 | 0.963 8 | 0.956 3 | 0.958 4 |
| $R^2(\text{test})$ | 0.959 1 | 0.953 2 | 0.947 6 | 0.940 4 |
| RMSE(train) | 0.049 6 | 0.054 7 | 0.060 1 | 0.058 6 |
| RMSE(test) | 0.058 1 | 0.062 2 | 0.065 8 | 0.070 1 |

基于 IDE-SVM 算法所建模型的预测精度最高, 且其在测试样本上 R^2 值最大、RMSE 值最小, 故该模型相对于其他 3 种模型具有更强的泛化能力; 而在相同的学习迭代次数下, 由于改进差分进化算法相对于原算法只是在变异收缩因子的选取方式上有所不同, 故分别基于以上两种算法所建模型的训练运行时间基本一致, 但其相对于粒子群算法和遗传算法的模型训练运行时间明显较短。

模型训练运行时间较短只能说明本文所提的改进差分进化算法在每次迭代中的解更新速度较快, 而判定优化算法优劣的标准主要是其搜索最优解的收敛速度以及所得解是否是全局最优。为测试本文所采用优化算法的收敛速度, 硝酸根离子和亚硝酸根离子分别基于 4 种算法训练模型的收敛性能对比如图 4 和图 5 所示。其中, 为方便观察各算法的收敛变化情况, 将图中纵轴设置为对数坐标轴, 并选取 $1-R^2$ 作为新的适应度值。由图 4 和图 5 可知, 基于 IDE-SVM 算法所建模型搜索到最优解所需的迭代次数最少, 且其所得解相对于其他 3 种算法最优, 故本文所提改进差分进化算法具有更快的收敛速度且能防止算法陷入局部最优。

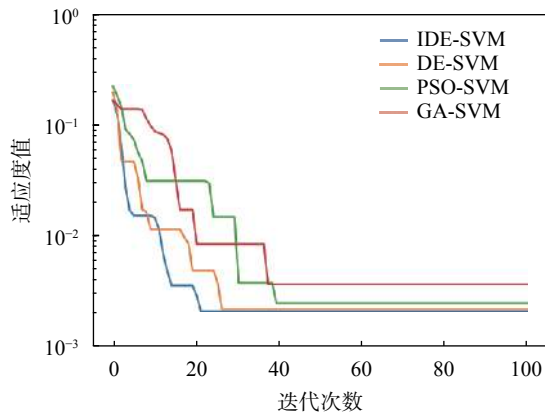


图4 收敛性能对比(硝酸根离子)

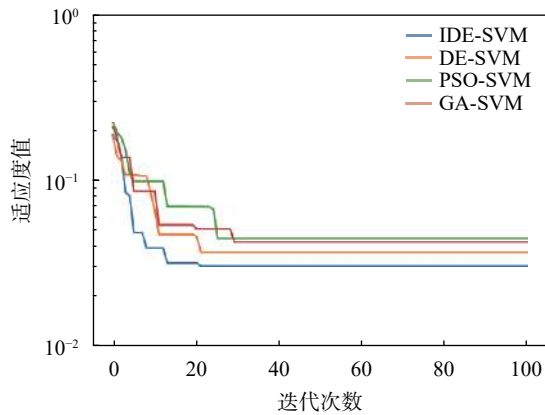


图5 收敛性能对比(亚硝酸根离子)

4 结束语

本文针对水质在线监测系统能够对水质进行实时有效连续检测的性能要求,提出一种基于紫外可见吸收光谱的水质检测算法。该算法通过小波变换和主成分分析对水质光谱数据进行去噪和降维,采用支持向量机对预处理后的数据进行建模,并采用改进差分进化算法对模型进行参数优化。通过实验对比分析可以得出以下结论:本文基于 IDE-SVM 算法所建模型具有更高的预测精度,且通过非线性自适应调整变异收缩因子对差分进化算法进行改进能够显著加快其收敛速度并防止算法陷入局部最优。由于本文所建模型在测试样本上的预测精度较高,可能存在训练样本和测试样本的相似度较高造成数据冗余的问题,故解决数据的冗余性问题将成为本文后续工作的重点。

参考文献

[1] ADU-MANU K S, TAPPARELLO C, HEINZELMAN W, et al. Water quality monitoring using wireless sensor networks:

current trends and future research directions[J]. *ACM Transactions on Sensor Networks*, 2017, 13(1): 1-41.

[2] STOREY M V, VAN DER GAAG B, BURNS B P. Advances in on-line drinking water quality monitoring and early warning systems[J]. *Water Research*, 2011, 45(2): 741-747.

[3] JIAO L, DONG D, ZHENG W. Determination of thiophanate-methyl using UV absorption spectra based on multiple linear regression[J]. *Optik*, 2014, 125(1): 183-185.

[4] SKOU P B, BERG T A, AUNSBJERG S D, et al. Monitoring process water quality using near infrared spectroscopy and partial least squares regression with prediction uncertainty estimation[J]. *Applied Spectroscopy*, 2017, 71(3): 410-421.

[5] OZBALCI B, BOYACI IH, TOPCU A, et al. Rapid analysis of sugars in honey by processing Raman spectrum using chemometric methods and artificial neural networks[J]. *Food Chemistry*, 2013, 136(3-4): 1444-1452.

[6] DEVOS O, RUCKEBUSCH C, DURAND A, et al. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation[J]. *Chemometrics and Intelligent Laboratory Systems*, 2009, 96(1): 27-33.

[7] 曾甜玲, 温志渝, 温中泉, 等. 基于紫外光谱分析的水质监测技术研究进展 [J]. *光谱学与光谱分析*, 2013, 33(4): 1098-1103.

[8] HUO A D, ZHANG J, QIAO C L, et al. Multispectral remote sensing inversion for city landscape water eutrophication based on genetic algorithm-support vector machine[J]. *Water Quality Research Journal of Canada*, 2014, 49(3): 285-293.

[9] WANG X, LÜ J K, XIE D T. A hybrid approach of support vector machine with particle swarm optimization for water quality prediction[J]. *International Conference on Computer Science and Education*, 2010, 31(7): 1158-1163.

[10] CIVICIOGLU P, BESDOK E. A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms[J]. *Artificial Intelligence Review*, 2013, 39(4): 315-346.

[11] 汤斌, 魏彪, 毛本将, 等. 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): 1-7.

[12] 陈扬, 张太宁, 郭澎, 等. 基于主成分分析的复杂光谱定量分析方法的研究 [J]. *光学学报*, 2009, 29(5): 1285-1291.

[13] CHANG C C, LIN C J. LIBSVM: A library for support vector machines[J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(3): 1-27.

[14] 曹飞凤, 许月萍. 基于改进差分进化算法的水文模型参数多目标优选研究 [J]. *系统工程理论与实践*, 2014, 34(12): 3268-3273.

(编辑:刘杨)