doi: 10.20236/j.CJIAC.2026.11.003

陈丁熙，柯川，何亚雄，等. 基于激光诱导击穿光谱(LIBS)与机器学习的水体低浓度铬污染等级分类[J]. 中国无机分析化学，2026，16(11)：1990-2001.

Chen Dingxi, Ke Chuan, He Yaxiong, et al. Classification of Low-concentration Chromium Pollution Grade in Water Based on Laser-induced Breakdown Spectroscopy (LIBS) Combined with Machine Learning[J]. Chinese Journal of Inorganic Analytical Chemistry, 2026, 16(11): 1990-2001.

基于激光诱导击穿光谱(LIBS)与机器学习的水体低浓度铬污染等级分类

陈丁熙[1] 柯 川[1,2,3*] 何亚雄[2,3] 温起帆[3] 钱自然[1] 赵栋烨[4]

(1.西南交通大学 物理科学与技术学院，成都 610031；
2.四川文理学院 智能制造学院，四川 达州 635000；
3.西南交通大学 电气工程学院，成都 610031；
4.内蒙古工业大学 新能源学院，呼和浩特 010051)

摘 要 水体重金属污染是环境监测领域的重要问题，其中铬(Cr)因在电镀、冶金、皮革加工及化工等工业中的广泛应用，易通过工业废水排入自然水体，对生态系统和人体健康构成潜在威胁。为实现水体Cr污染的快速、可靠识别，本文提出了一种基于激光诱导击穿光谱(Laser-induced Breakdown Spectroscopy，LIBS)与机器学习相结合的水体Cr污染等级分类方法。首先为降低样品分布不均及等离子体波动带来的影响，本研究提出在咖啡环沉积区域进行空间扫描(Mapping)点阵取样的策略，通过10×10矩阵系统采集每个样品的100条光谱，克服样品分布不均问题。采集的光谱数据经过最大-最小归一化(Min-Max Normalization)、Savitzky-Golay(S-G)滤波及非对称最小二乘(Asymmetric Least Squares，ALS)基线校正，显著增强了特征峰强度与Cr浓度之间的线性相关性，使线性拟合决定系数R^2由0.78提升至0.98。在机器学习建模环节，本研究将1 000条光谱数据随机划分为训练集、验证集和测试集，比例为8∶1∶1，分别用于模型训练、超参数调优以及最终性能评估。并选取偏最小二乘判别分析(PLS-DA)、随机森林(Random Forest，RF)、多层感知机(Multilayer Perceptron，MLP)以及支持向量机(Support Vector Machine，SVM)四种常用分类算法进行对比分析，以全面挖掘光谱特征与污染等级之间的判别关系，并评估不同算法在处理低浓度样本时的稳健性。结果显示，SVM模型在验证集和测试集上均表现出最优的分类性能，测试集整体准确率达到0.99，各污染等级样本尤其是低浓度类别样本均表现出高度稳定的判别能力。通过混淆矩阵与接收者操作特征(Receiver Operating Characteristic，ROC)曲线分析，进一步验证了SVM在类别边界判

收稿日期：2026-02-10 **修回日期**：2026-03-17 **录用日期**：2026-03-17 **在线出版日期**：2026-07-07

Received：2026-02-10 **Revised**：2026-03-17 **Accepted**：2026-03-17 **Online published**：2026-07-07

基金项目：国家重点研发计划项目(2022YFE03020002，2022YFE03200200)；四川省无人机产业创新基金资助项目(02-2025-02-0045)

Fund：Supported by the National Key Research and Development Program of China (2022YFE03020002, 2022YFE03200200); Sichuan UAV Industry Innovation Center Innovation Fund (02-2025-02-0045)

作者简介：陈丁熙，女，硕士研究生，主要从事激光诱导击穿光谱和拉曼光谱应用研究。E-mail：c121341679@163.com

Author：Chen Dingxi, female, master's student, mainly engaged in research on applications of Laser-Induced Breakdown Spectroscopy (LIBS) and Raman spectroscopy. E-mail：c121341679@163.com

*__通信作者__：柯川，男，副教授，主要从事激光诱导击穿光谱与机器学习结合应用研究及激光诱导击穿光谱装置及软件开发等研究。E-mail：cke@swjtu.edu.cn

*__Corresponding author__：Ke Chuan, male, associate professor, mainly engaged in research on the integration of laser-induced breakdown spectroscopy (LIBS)with machine learning, as well as development of LIBS instruments and software. E-mail：cke@swjtu.edu.cn

©Chinese Journal of Inorganic Analytical Chemistry Editorial Department 2026. This is an open access article under the CC-BY-NC-ND 4.0 License.

别、样本泛化能力以及多类别分类精度方面的优势。该方法不仅为工业废水中Cr污染的实时分级监测提供了可靠的技术方案，同时为LIBS技术从实验室分析向现场智能决策的转化提供了实践依据。未来可进一步结合多模态光谱数据融合，如拉曼光谱与LIBS数据联合分析，提升痕量重金属污染等级识别的精度与适用范围，从而为水环境监测和污染控制提供更高效的工具。

关键词　水体Cr污染；激光诱导击穿光谱(LIBS)；机器学习；污染等级分类

要点

1）建立了水体中低浓度Cr的污染等级分类模型，实现对低浓度Cr污染的快速、可靠识别。

2）结合光谱Mapping与机器学习方法，有效降低样品非均匀性和等离子体波动对检测结果的影响，提高模型判别稳定性与精度。

3）所提出的分析策略为水体重金属污染的快速监测与风险评估提供了新的技术思路。

中图分类号：O657.38　　**文献标志码：**A　　**文章编号：**2095-1035(2026)11-1990-12

水体重金属污染是当前环境监测领域面临的重要问题之一，其中Cr因其广泛应用于电镀、冶金、皮革加工及化工等工业过程，易通过工业废水排放进入自然水体，对生态系统和人体健康构成潜在威胁[1-4]。长期暴露于含Cr污染水体中，可能引发皮肤损伤、呼吸系统疾病甚至致癌风险，因此对水体Cr污染的快速、可靠识别对于环境风险评估与污染预警具有重要意义。当前，国家相关水质标准已对水体中Cr的限值及其污染等级作出明确规定，这使得检测技术不仅需要具备较高的灵敏度，同时还需要能够有效区分不同的污染水平。现有水体Cr污染的检测方法主要包括原子吸收光谱(Atomic Absorption Spectrometry，AAS)[5-6]、电感耦合等离子体质谱(Inductively Coupled Plasma-Mass Spectrometry，ICP-MS)[7-8]、电感耦合等离子体发射光谱(Inductively Coupled Plasma-Optical Emission Spectrometer，ICP-OES)[9-10]、拉曼光谱(Raman Spectra)[11-12]等。尽管这类实验室方法在痕量元素分析中具备极低的检出限，但其高昂的设备成本、复杂的样品前处理、较长的分析周期(通常需要数小时)，再或者是对实验室环境的依赖，使它们难以完全满足现场快速监测和大规模环境监测的需求[13]。此外，样品在运输和预处理过程中还可能引入额外误差，影响检测结果的时效性和真实性。

LIBS作为一种新兴的原子发射光谱技术，具有无需复杂前处理、响应速度快、多元素同时检测等优势[14-15]，近年来逐渐被应用于食品[16]、材料[17]及环境污染[18]分析等领域。然而，当激光直接作用于液体样品时，液体表面易产生飞溅，同时液体等离子体寿命较短、信号波动较大，从而限制检测灵敏度与重复性。因此，实际研究中常通过将液体样品转化为固体基底进行分析，以提高光谱信号的稳定性。例如，Hu等[19]将水样沉积在石墨片表面以富集重金属元素，再利用LIBS进行检测，结果表明样品在石墨片表面形成较均匀沉积层，在累积20次激光脉冲后获得稳定且重复性良好的光谱信号。Yu等[20]则以普通滤纸为吸附基底，将其浸入含$Cr(NO_3)_3$的溶液中富集Cr离子，并对干燥后的样品进行检测，检出限达到0.026 mg/L，相比直接液体检测灵敏度提高约2～3个数量级。尽管液转固方法能够改善信号稳定性，但由于水体中Cr通常以痕量形式存在，传统LIBS在灵敏度方面仍存在一定限制。因此，近年来研究者提出多种信号增强策略以进一步提升检测性能。一类方法是通过样品富集技术提高目标元素在检测区域的浓度，例如Xu等[21]将离子富集芯片与LIBS结合，实现Cr(Ⅵ)检出限约4 μg/L、总Cr约10 μg/L的检测能力。另一类方法是利用纳米材料或功能化基底实现离子吸附富集，例如Dai等[22]采用金纳米团簇修饰滤纸作为富集基底，实现Cr(Ⅲ)检测限约7.5 μg/L，并成功应用于实际水样分析；Luo等[23]利用金属三元纳米颗粒修饰滤纸进行富集后开展LIBS检测，使总Cr的检出限降低至0.21 μg/L。此外，实验条件优化同样是提升检测性能的重要途径。例如Chen等[24-25]利用飞秒激光LIBS在流动水溶液中实现了Cr等重金属元素的检测，并通过优化光谱积累次数等参数，使Cr检测限达到10^{-2} μg/mL量级。但上述多数研究仍聚焦于降低检出限，且实验过程中仍基于单点或少量采样策略，该策略易受等离子体波动影响，在低浓度样本判别稳定性方面存在不足，并主要依赖线性判别或经验阈值进行污染判断[26-27]。因此在样品非均匀性、等离子体波动以及实验重复性不足等因素影响下，污染等级识别的稳定性和可靠性仍有待进一步提升。

基于上述问题，本文针对水体Cr污染检测中样品空间分布不均以及污染等级判别稳定性不足等关键挑战，创新性地将统计采样思想引入LIBS检测：通过Mapping策略获取沉积区域内的空间光谱信息，从而表征Cr在样品表面的分布特征，并结合机器学习算法建立非线性判别模型，以减小单点采样带来的随机波动，提高污染等级识别的稳定性和可靠性，实现水体Cr污染等级的快速分类[28-29]。在此基础上，通过验证集对模型关键参数进行系统优化，并对不同算法的分类性能进行对比分析，最终在测试集上获得了较高的分类准确率和稳定的判别性能，为LIBS技术在水体Cr污染等级识别中的工程化应用提供了方法支持。

1 实验部分

1.1 LIBS设备与参数

本研究所采用的LIBS实验系统原理示意如图1所示。系统以Nd：YAG脉冲激光器(Nasor-600)作为激发光源，工作波长为1 064 nm，重复频率为10 Hz，单脉冲能量约为100 mJ。激光束经反射镜调节后，通过焦距为10 cm的聚焦透镜聚焦于样品表面，在局部区域产生高温等离子体后通过旁轴光学收集系统进行采集，并传输至光谱仪进行光谱分辨与记录。实验中采用Avantes AvaSpec-Mini4096CL光谱仪进行光谱采集，其有效工作波长为250～415 nm，分辨率约为0.1 nm，覆盖Cr的主要特征发射谱线区域。

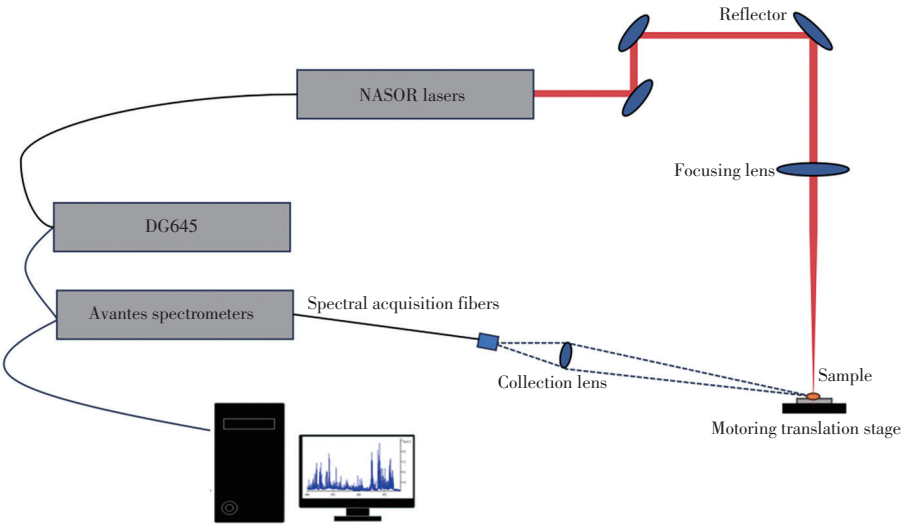


图1 LIBS装置示意图

Figure 1 Schematic diagram of the LIBS setup.

为获得稳定且具有较高信噪比的等离子体发射信号，实验中将光谱采集的积分时间设置为10 μs。系统时序由延迟发生器(DG645)进行统一控制，用于精确调节激光触发与光谱采集之间的延迟时间，实验中延迟时间设定为1 μs，以有效抑制等离子体早期连续谱背景对特征谱线分析的影响。

1.2 样品制备

依据《地表水环境质量标准》(GB 3838—2002)[30]及《污水综合排放标准》(GB 8978—1996)[31]中对六价铬Cr(Ⅵ)和总Cr浓度限值的相关规定，样品的具体制备流程如图2所示，以实验室去离子水作为溶剂，以重铬酸钾($K_2Cr_2O_7$)为Cr源，配制10种不同浓度的含Cr水样。样品浓度设置覆盖上述标准所涉及的不同污染水平区间，以反映水体中Cr污染程度的差异，并为后续污染风险等级的划分与分类建模提供数据基础，各样品的具体浓度分布及其对应的风险等级如表1所示。在光谱采集前，将配制好的含Cr溶液以定量方式滴加于洁净的铝板基底表面，并置于加热台上进行快速烘干处理，使液滴在基底表面形成稳定固体沉积区域。铝板基底具有良好的热稳定性和光谱背景一致性，有利于提高实验重复性。

1.3 LIBS元素面成像

本实验采用基于光谱Mapping的多点采集方式获取LIBS信号。实验过程中发现，溶液液滴在铝板基底上干燥后易形成明显的咖啡环结构，且咖啡环内外区域中Cr的空间分布存在不均匀性，如图3a所示。若仅在局部位置或随机选取少量测量点进行光

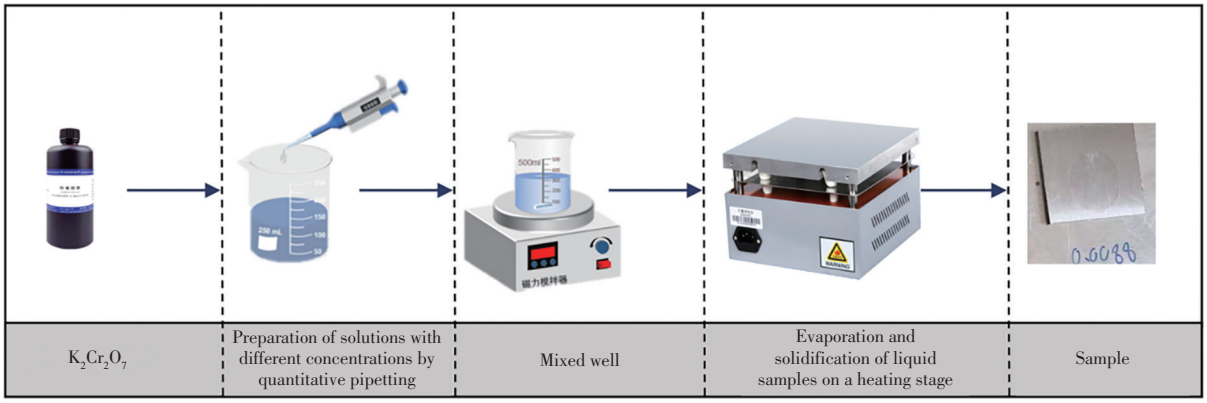


图2　样品制备流程

Figure 2　Sample preparation process.

表1　样品Cr含量及风险等级划分

Table 1　Cr concentrations in samples and classification of risk levels

Risk level	Water quality condition	Sample ID	Cr(Ⅵ) concentration/(mg · L^{-1})
Class Ⅰ	Not suitable for direct discharge	Sample 1	8. 843
		Sample 2	4. 422
Class Ⅱ	Meets total Cr/Cr(Ⅵ) discharge requirements	Sample 3	1. 105
		Sample 4	0. 884
		Sample 5	0. 442
Class Ⅲ	Surface water Class Ⅱ — Ⅴ	Sample 6	0. 088
		Sample 7	0. 044
		Sample 8	0. 011
Class Ⅳ	Surface water Class Ⅰ (very low pollution)	Sample 9	0. 008 8
		Sample 10	0. 005 5

谱采集，容易引入较大的随机误差，影响分析结果的稳定性。因此，如图3b所示，本研究以咖啡环内部区域为主要测量对象，在其区域内进行规则化Mapping采集。具体而言，每一个样品的咖啡环直径约为24 mm，三维移动平台步长设置为1. 5 mm，单点烧蚀直径约为1 mm，在约14. 5 mm×14. 5 mm的矩形区域内按照10×10矩阵布设测量点，共采集100个LIBS信号。通过Mapping方式对咖啡环内部区域进行系统采集，可较为全面地表征样品中Cr的空间分布特征。且所得各测量点光谱均作为独立样本参与后续数据分析与建模处理，从而降低由样品非均匀性及等离子体波动带来的随机误差，提高训练数据的代表性和模型分析的可靠性。

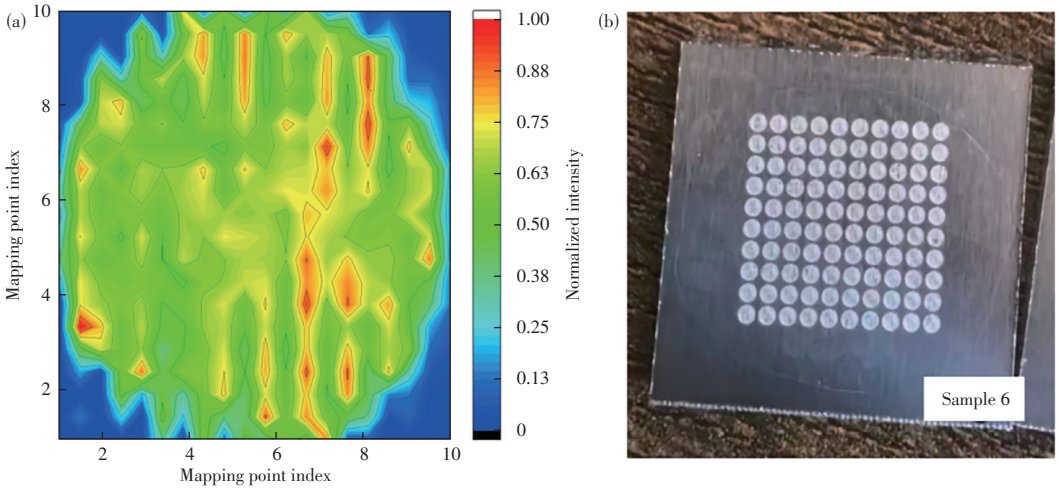


图3　(a) Mapping采集得到的特征峰归一化强度分布图，表明咖啡环内部Cr分布存在明显不均匀性；(b) 样品6咖啡环区域的10×10 mapping点阵实物图

Figure 3　(a) Mapping-acquired normalized intensity distribution of the characteristic peak, indicating a pronounced inhomogeneous distribution of Cr within the coffee-ring region; (b) Photograph of the 10×10 mapping grid in the coffee-ring area of Sample 6.

2 结果与讨论

2.1 谱线选择及预处理

为了确保定量模型输入的可靠性，本研究首先通过实验获取样品的LIBS光谱，并对光谱中可能存在的Cr特征峰进行初步识别。在观察光谱的基础上，将候选峰与美国国家标准与技术研究院(NIST)原子光谱数据库中的Cr谱线进行比对，以确认峰的元素归属，如图4所示为样品1在Cr(Ⅰ)357.860 nm处的峰，该峰较为稳定且强度显著，成为分析Cr含量的重要指标。

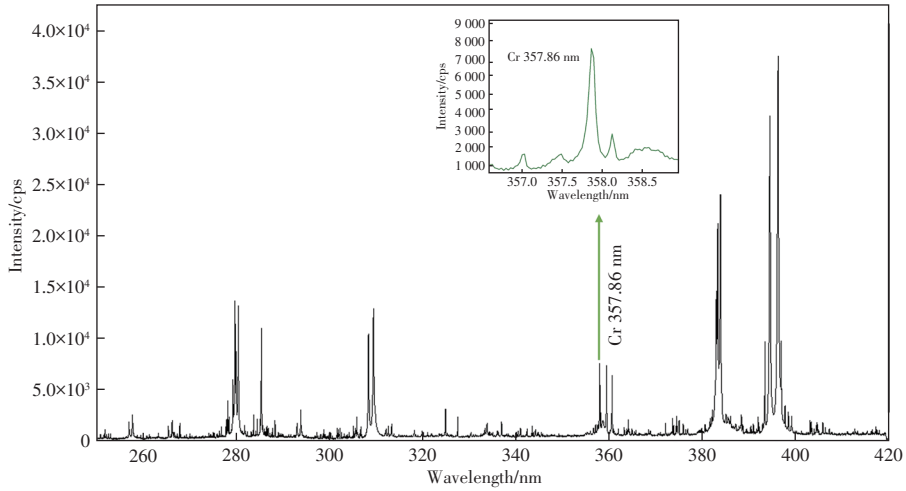


图4 样品1的LIBS全谱及Cr(Ⅰ)357.860 nm特征峰放大图

Figure 4 Full LIBS spectrum of Sample 1 and a zoomed-in view of the Cr(Ⅰ)357.860 nm characteristic peak.

考虑到基底材料对Cr特征谱线识别的潜在影响，本研究进一步采集了空白铝基板的LIBS光谱作为对照。如图5所示，空白铝基板的光谱在394和396 nm附近出现明显的铝元素特征峰，而在357.860 nm附近未观察到Cr相关发射峰。这表明铝基板本身不含Cr，其发射谱线也不会与Cr(Ⅰ)357.860 nm发生重叠干扰。因此，即使在激光烧蚀过程中存在少量基板材料参与等离子体形成，其对Cr特征谱线的影响可以忽略，从而保证了所选特征谱线用于Cr分析的可靠性。

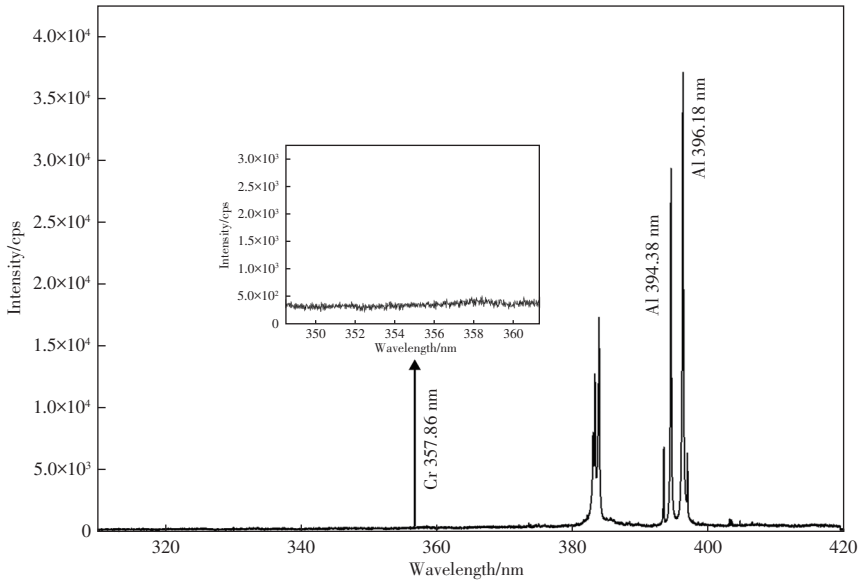


图5 空白铝基板的LIBS光谱，显示铝的特征峰(394、396 nm)，未见Cr在357.860 nm处的特征峰

Figure 5 LIBS spectrum of a blank aluminum substrate, showing aluminum peaks at 394 and 396 nm, with no characteristic Cr peak at 357.860 nm.

尽管原始LIBS光谱中已能够识别出Cr的特征谱线，但光谱数据仍不可避免地受到等离子体连续谱背景、随机噪声及基底效应等因素的影响，导致谱线强度波动较大，尤其在低浓度样品中更为明显。上述不确定性会削弱不同污染等级样品之间的光谱差异，不利于分类模型对类别边界的有效学习。因此，在锁定Cr特征谱线后，对原始光谱数据依次进行了归一化、平滑去噪及基线校正等预处理操作。具体而言，首先，对原始光谱进行Min-Max normalization。在LIBS实验中，激光能量波动及等离子体状态变化会引起光谱整体强度波动。该方法可将不同样本的光谱强度映射至统一范围，从而降低强度尺度差异对后续特征分析的影响。在此基础上，采用S-G滤波器对光谱进行平滑处理。该方法基于局部多项式拟合，在有效抑制高频随机噪声的同时能够较好地保持谱线峰形与峰位信息。随后采用ALS算法进行基线校正，以消除等离子体连续背景辐射及基线漂移，从而提高特征谱线的可辨识度。图6给出了Cr特征谱线周围在预处理前后的对比结果，其中不同颜色光谱表示不同Cr(Ⅵ)浓度样品，颜色由紫色逐渐过渡至黄色，对应浓度由低到高。且三维图中黑色折线为不同浓度样品在357.86 nm特征峰处的数据点连接线，用于直观展示各浓度下特征峰的强度关系。可以看出，经过预处理后，Cr(Ⅰ)357.860 nm处的特征峰更加突出，谱线形态更加平滑稳定。尤其是在低浓度样品中，预处理有效抑制了背景干扰和异常波动，避免了原始光谱中出现的低浓度样品特征峰强度反而高于高浓度样品的现象，使特征峰强度随浓度变化的趋势更加合理，从而增强了不同污染等级样品在特征空间中的可分性。同时，为定量评估预处理对特征有效性的影响，基于该特征峰强度与Cr浓度之间的关系进行了线性拟合，其决定系数R^2由原本的0.782 4提升至0.985 9(图7)。该结果表明，预处理显著增强了光谱特征对Cr浓度变化的响应一致性，使特征峰强度与浓度之间呈现出更加稳定的线性关系。为后续基于机器学习的污染等级分类模型提供了更加稳定、可靠的输入特征基础，有助于提升模型整体的判别性能与泛化能力。

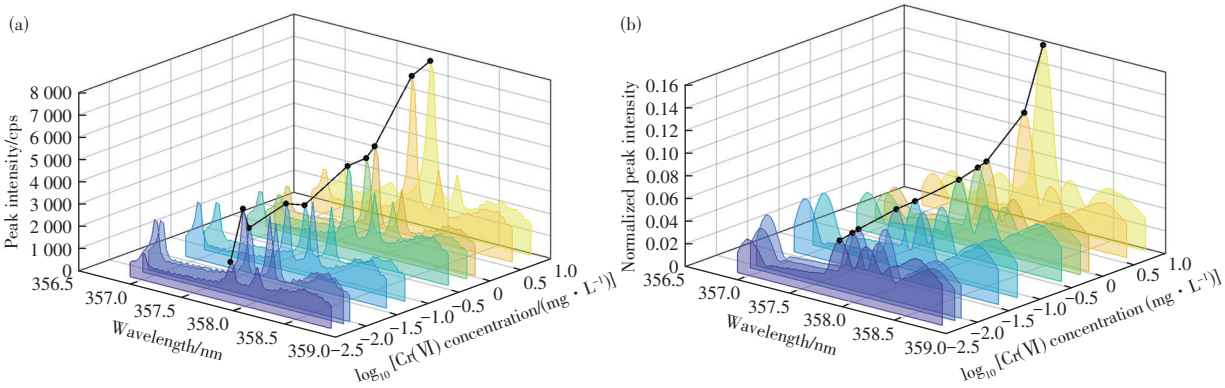


图6　Cr特征谱线周围区域预处理前后的三维分布对比图，(a)和(b)分别对应预处理前后的光谱数据

Figure 6　Three-dimensional distribution comparison around the Cr characteristic spectral line before and after preprocessing, with (a) and (b) corresponding to the spectra prior to and following preprocessing, respectively.

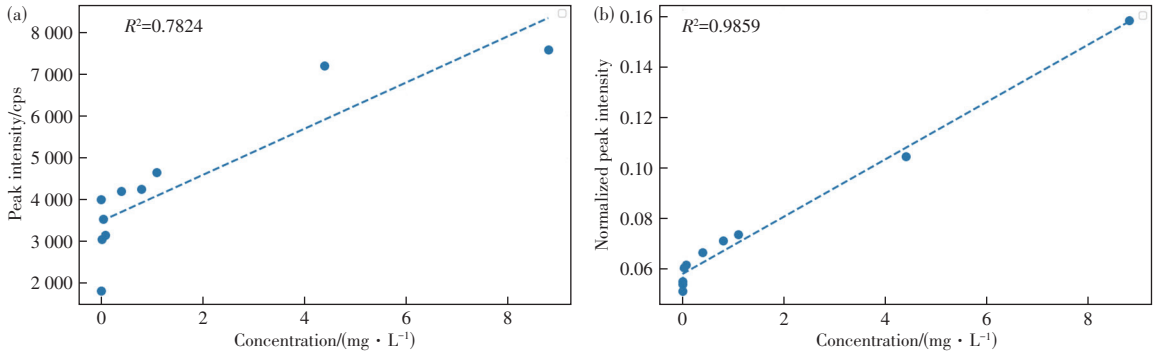


图7　Cr特征峰强度与浓度的线性拟合结果对比，(a)和(b)分别对应预处理前后的光谱数据

Figure 7　Comparison of linear fitting results between Cr characteristic peak intensity and concentration, with (a) and (b) corresponding to the spectra before and after preprocessing, respectively.

2.2 水体Cr污染等级分类

在完成样品污染等级标注以及数据预处理后，开始基于LIBS光谱数据构建水体Cr污染等级分类模型。具体而言，将表1中所列出的10个样品，每个通过Mapping方式采集100条光谱，共计1 000条光谱。为保证模型训练、参数优化与性能评估过程的独立性与可靠性，数据集按照8∶1∶1的比例随机划分为训练集、验证集和测试集。其中，训练集用于模型参数学习，验证集用于模型参数调优与模型选择，测试集仅用于最终分类性能评估。

为全面挖掘LIBS光谱特征与Cr污染等级之间的潜在判别关系，本文选取了四种具有不同建模思想和结构特点的经典分类算法进行对比研究，包括PLS-DA[32]、RF[33]、MLP[34]以及SVM[35-36]。其中，PLS-DA适用于高维光谱数据的线性判别建模，能够在降维的同时保留与类别相关的判别信息；RF通过集成多棵决策树，可有效刻画特征之间的非线性关系，并具有较强的抗过拟合能力；MLP作为典型的人工神经网络模型，能够学习复杂的非线性映射关系；SVM则通过最大化类别间隔，在小样本条件下具有良好的泛化性能。多模型联合建模有助于从不同角度刻画光谱特征与污染等级之间的判别机制。

在模型参数优化过程中，以验证集分类准确率(Accuracy)作为统一评价指标，对各模型的关键参数进行调节。图8分别给出了四种模型在不同参数设置下验证集准确率的变化情况。首先图8a展示了对于PLS-DA模型，验证集准确率随潜变量数(LV)的增加呈现出先快速提升、后趋于平稳的变化趋势。当LV较小时，模型难以充分提取光谱与类别之间的判别信息；当LV增加至约20以上时，分类准确率已接近饱和，进一步增加LV并未带来显著性能提升。综合判别性能与模型复杂度，选取位于性能平台区间内的LV作为最优参数。

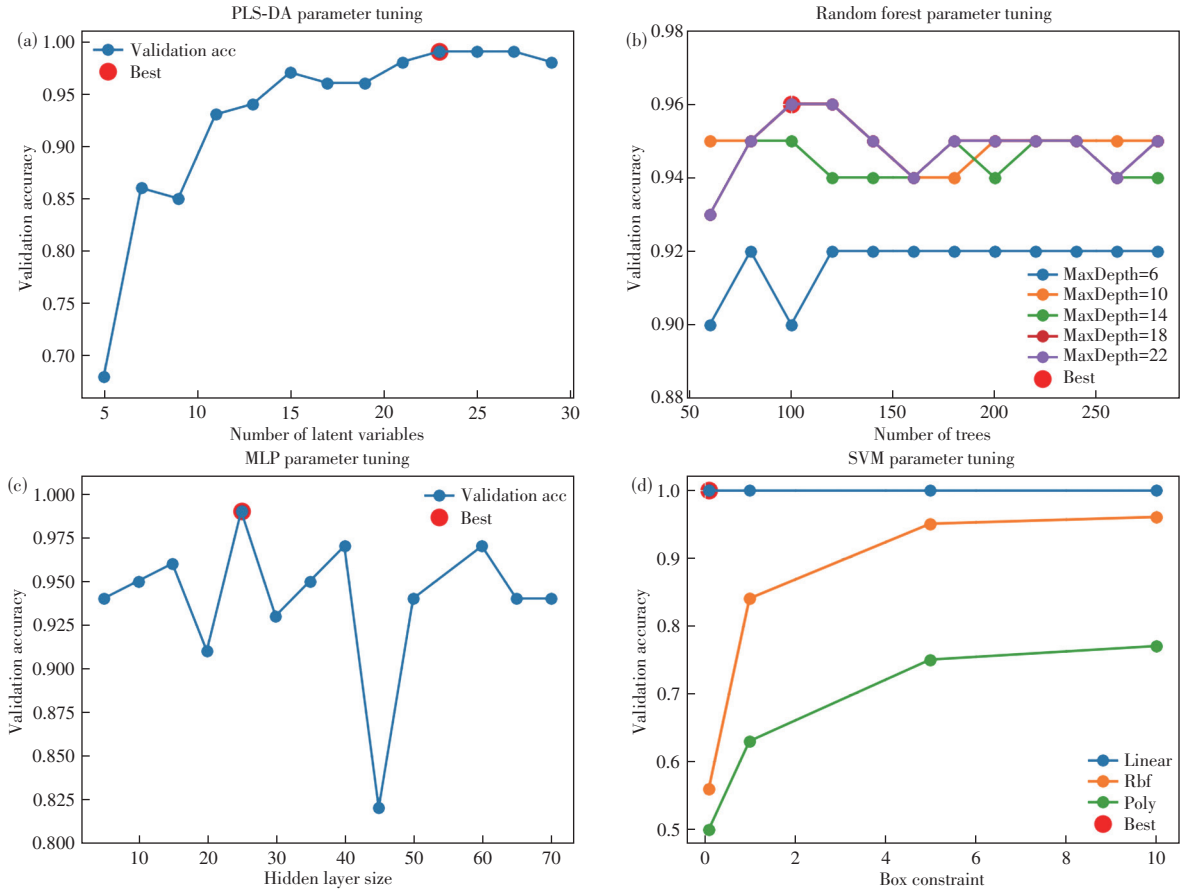


图8 四种分类模型在不同参数设置下的验证集准确率变化

Figure 8 Validation accuracy of the four classification models under different parameter settings.

图8b展示了RF模型在不同决策树数量和最大深度(MD)组合下的验证集准确率变化情况。结果表明，在决策树数量增加至一定规模后，模型性能趋于稳定，而最大深度对分类性能的影响更为显著。当MD

取中等偏大的数值时，模型在验证集上获得较高且稳定的准确率；过小或过大的MD均会导致性能下降或波动。因此，在保证较高分类准确率的同时，选取验证集性能稳定的参数组合作为RF模型的最终配置。

如图8c所示，MLP模型的验证集准确率随隐藏层神经元数量变化呈现一定波动特征。当隐藏层规模较小时，模型表达能力受限；而当节点数过多时，模型稳定性下降，验证集准确率出现明显波动。综合考虑分类性能与模型稳定性，选取验证集准确率较高且波动较小的隐藏层规模作为最优网络结构。

对于SVM模型，图8d比较了不同核函数在不同惩罚参数C下的分类性能。可以看出，线性核在较宽的C取值范围内均能获得接近1.00的验证集准确率，表现出最优且稳定的分类性能；相比之下，RBF核和多项式核对C的变化更为敏感，但整体性能相对较低。因此，最终选取线性核，并在对应的最优C值下构建SVM分类模型。

在此基础上，利用验证集和测试集的准确率对各模型的最终分类性能进行了评估，并将结果以柱状图形式呈现(图9)，以便直观比较不同模型在不同

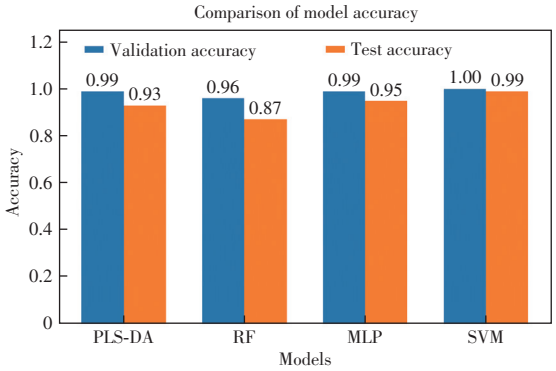


图9　四种分类模型在验证集和测试集上的最终分类准确率比较

Figure 9　Comparison of the final classification accuracy of the four models on the validation and test sets.

数据集上的分类精度。结果表明，SVM模型在测试集上仍保持较高的分类准确率，可达到0.99，因此该模型具有极强的泛化能力。

为进一步分析各模型在不同污染等级上的分类表现，将四种模型的混淆矩阵绘制如图10所示。结合图9的整体准确率可以看出，SVM在测试集上不仅总体准确率最高，也在各等级样本上识别均接近理想状态，误判最少；MLP在各污染等级样本上仍存

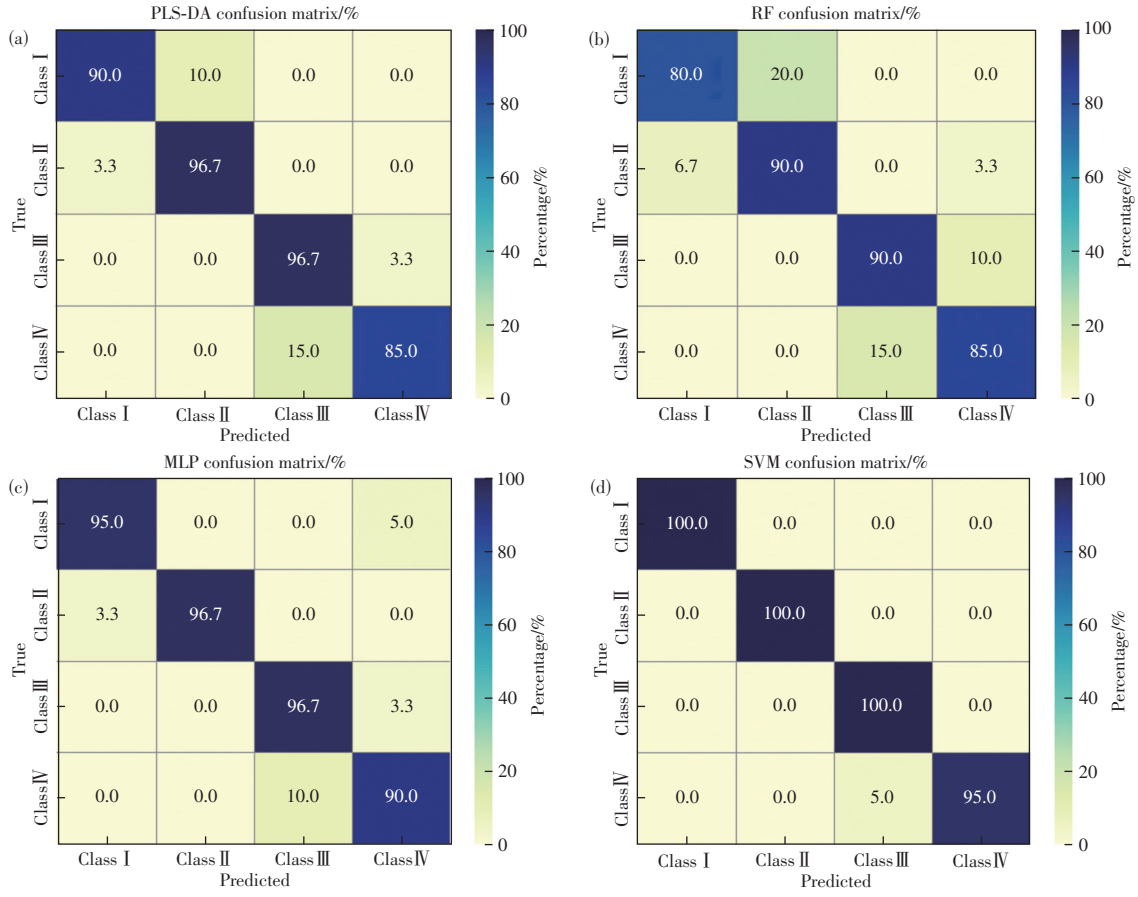


图10　四种分类模型的混淆矩阵

Figure 10　Confusion matrices of the four classification models.

在少量误判，其中低浓度样本的判误情况较为明显，这可能与低浓度样本光谱信号较弱有关；而RF模型和PLS-DA模型基本在每个等级上判误情况都较为明显，这可能与模型对局部特征的敏感性及信号非线性波动有关。总体来看，图9的准确率趋势与图10的混淆矩阵模式相一致，各模型的误判特点与其建模特性密切相关。

如图11所示，绘制了各模型的多分类ROC曲线及对应的AUC值。在多分类情形下，采用“一对其余”策略(one-*vs*.-rest，OvR)，分别计算每个等级的ROC曲线，并通过调整决策阈值观察模型对不同类别的敏感性。结果显示，SVM在所有等级上均具有最高的AUC值，判别能力和稳定性最佳；MLP次之，而PLS-DA和RF在部分类别的AUC略低，尤其是在低浓度或边界等级样本上。综合柱状图、混淆矩阵和ROC分析可知，SVM在整体精度、各等级判别能力及泛化性能方面均表现最佳，因而成为水体Cr污染等级快速、可靠识别的优选模型。

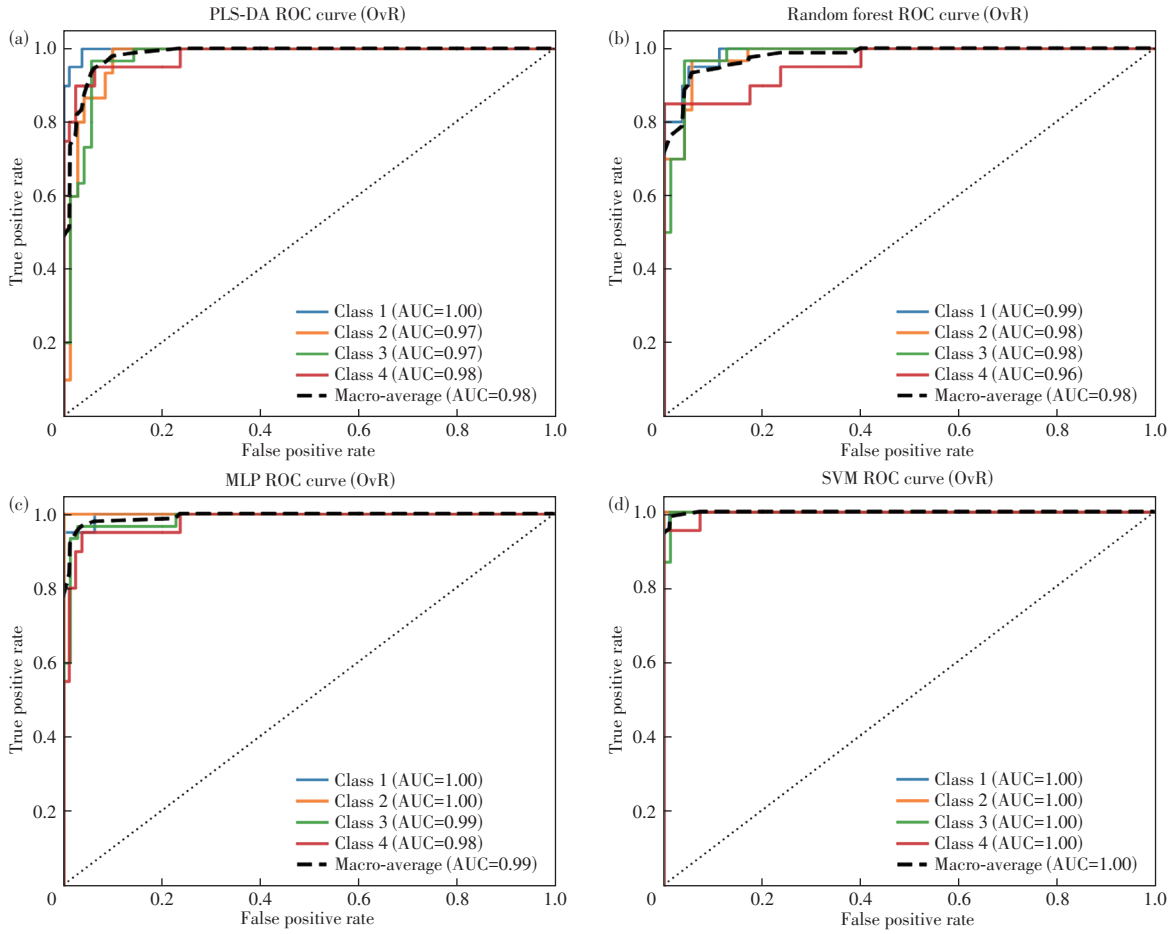


图11　四种分类模型的多分类ROC曲线及对应AUC值

Figure 11　Multiclass ROC curves and corresponding AUC values of the four classification models.

3　结论

本研究成功建立了一种结合LIBS-mapping采样与机器学习分类的分析方法，有效解决了水体低浓度Cr污染等级判别中的稳定性难题。采用咖啡环区域系统化采样策略，通过10×10点阵布点与统计平均，显著降低了样品空间分布不均带来的误差，为可靠分类奠定了坚实基础。在数据处理方面，光谱预处理流程的优化使特征信号稳定性显著提升。经过归一化、平滑去噪及基线校正处理后，Cr特征峰强度与浓度的相关性决定系数R^2从0.78提高至0.98，为后续机器学习建模提供了高质量输入。通过对四种分类模型的对比研究，发现SVM在污染等级分类任务中表现最优。该模型在测试集上的整体准确率达到0.99，且在各污染等级特别是低浓度样本上均展现出稳定的判别能力。混淆矩阵与多分类ROC曲线分析进一步证实，SVM模型在处理类别边界样本时具有明显优势，其非线性分类能力显著优于传统

线性判别方法。本研究的创新价值在于将空间统计采样思想与机器学习算法相融合，突破了传统LIBS技术依赖单点测量的局限，实现了从“定性检测”到“智能分级”的方法突破。这一策略不仅提升了低浓度Cr污染的判别可靠性，也为LIBS技术在环境现场监测中的实际应用提供了新的技术路径。

展望未来，可在多模态数据融合方向进一步探索，通过结合LIBS与Raman等光谱技术，实现Cr价态区分与定量分析的双重突破。同时，开发轻量化分类模型以适应便携式设备的应用需求，将推动该技术在水质实时监测领域的广泛应用。

参考文献

[1] Markiewicz B, Komorowicz I, Sajnóg A, et al. Chromium and its speciation in water samples by HPLC/ICP-MS-technique establishing metrological traceability: a review since 2000 [J]. Talanta, 2015, 132: 814-828.

[2] Zhang Peng, Yang Mingjie, Lan Jingjing, et al. Water quality degradation due to heavy metal contamination: health impacts and eco-friendly approaches for heavy metal remediation[J]. Toxics, 2023, 11(10): 828. DOI: 10. 3390/toxics11100828.

[3] Hadi Najafabadi H, Irani M, Roshanfekr Rad L, et al. Removal of Cu^{2+}, Pb^{2+} and Cr^{6+} from aqueous solutions using a chitosan/graphene oxide composite nanofibrous adsorbent[J]. RSC Advances, 2015, 5(21): 16532-16539.

[4] Erdem E, Güngörmüş H, Kılınçarslan R. The investigation of some properties of cement and removal of water soluble toxic chromium(Ⅵ) ion in cement by means of different reducing agents[J]. Construction and Building Materials, 2016, 124: 626-630.

[5] Rubio R, Sahuquillo A, Rauret G, et al. Determination of chromium in environmental and biological samples by atomic absorption spectroscopy: a review[J]. International Journal of Environmental Analytical Chemistry, 1992, 47(2): 99-128.

[6] Rossi E, Errea M I, de Cortalezzi M M F, et al. Selective determination of Cr(Ⅵ) by on-line solid phase extraction FI-SPE-FAAS using an ion exchanger resin as sorbent: an improvement treatment of the analytical signal[J]. Microchemical Journal, 2017, 130: 88-92.

[7] Zhang Nan, Suleiman J S, He Ming, et al. Chromium(Ⅲ)-imprinted silica gel for speciation analysis of chromium in environmental water samples with ICP-MS detection[J]. Talanta, 2008, 75(2): 536-543.

[8] Jia Xinyan, Gong Derong, Xu Bin, et al. Development of a novel, fast, sensitive method for chromium speciation in wastewater based on an organic polymer as solid phase extraction material combined with HPLC-ICP-MS[J]. Talanta, 2016, 147: 155-161.

[9] Frois S R, Tadeu Grassi M, de Campos M S, et al. Determination of Cr(vi) in water samples by ICP-OES after separation of Cr(iii) by montmorillonite[J]. Analytical Methods, 2012, 4(12): 4389. DOI: 10. 1039/c2ay26125a.

[10] Vu H C, Dwivedi A D, Le T T, et al. Magnetite graphene oxide encapsulated in alginate beads for enhanced adsorption of Cr(Ⅵ) and As(Ⅴ) from aqueous solutions: role of crosslinking metal cations in pH control[J]. Chemical Engineering Journal, 2017, 307: 220-229.

[11] 杨增霞，宋露露，杨林. 拉曼光谱法测定废水中六价铬的含量[J]. 理化检验-化学分册，2022，58(4)：485-486.
Yang Zengxia, Song Lulu, Yang Lin. Determination of hexavalent chromium in waste water by Raman spectrometry[J]. Physical Testing and Chemical Analysis (Part B: Chemical Analysis), 2022, 58(4): 485-486.

[12] Dvoynenko O, Lo S L, Chen Yijie, et al. Speciation analysis of Cr(Ⅵ) and Cr(Ⅲ) in water with surface-enhanced Raman spectroscopy[J]. ACS Omega, 2021, 6(3): 2052-2059.

[13] Vinay B K, Bagchi S, Suranjan T R. Chromium detection in water using optical methods: a study of reagent and reagentless approaches[J]. Critical Reviews in Analytical Chemistry, 2025: 1-38.

[14] 何亚雄，王一钦，韩晶阳，等. 飞秒-纳秒双脉冲激光诱导击穿光谱(LIBS)技术对合金的定量分析[J]. 中国无机分析化学，2024，14(2)：214-222.
He Yaxiong, Wang Yiqin, Han Jingyang, et al. Quantitative analysis of alloys by femtosecond-nanosecond dual-pulse laser-induced breakdown spectroscopy (LIBS)[J]. Chinese Journal of Inorganic Analytical Chemistry, 2024, 14(2): 214-222.

[15] Pathak A K, Kumar R, Singh V K, et al. Assessment of LIBS for spectrochemical analysis: a review[J]. Applied Spectroscopy Reviews, 2012, 47(1): 14-40.

[16] Casanova L, Beldjilali S A, Bilge G, et al. Evaluation of limits of detection in laser-induced breakdown spectroscopy: demonstration for food[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2023, 207: 106760. DOI: 10. 1016/j. sab. 2023. 106760.

[17] He Yuxin, Wen Qingfu, Ke Chao, et al. Depth-resolved and time-resolved investigations of Er_2O_3 ceramics after static lithium corrosion by using LIBS at different pressures[J]. Ceramics International, 2023, 49(22):

35802-35811.

[18] Stipe C B, Miller A L, Brown J, et al. Evaluation of laser-induced breakdown spectroscopy (LIBS) for measurement of silica on filter samples of coal dust[J]. Applied Spectroscopy, 2012, 66(11): 1286-1293.

[19] Hu Lei, Zhao Nanjing, Liu Wenqing, et al. Quantitative analysis of heavy metals in water based on LIBS with an automatic device for sample preparation[J]. Plasma Science and Technology, 2015, 17(8): 699-703.

[20] Yu Yongli, Zhou Weidong, Qian Honggang, et al. Simultaneous determination of trace lead and chromium in water using laser-induced breakdown spectroscopy and paper substrate[J]. Plasma Science and Technology, 2014, 16(7): 683-687.

[21] Xu Fenghua, Ma Shuxing, Tian Hongwei, et al. Detection of chromium in different valence states in water and soil using laser-induced breakdown spectroscopy combined with an ion enrichment chip[J]. Journal of Analytical Atomic Spectrometry, 2023, 38(7): 1442-1449.

[22] Dai Xin, Song Chengbiao, Ma Shuxing, et al. Rapid determination of Cr^{3+} and Mn^{2+} in water using laser-induced breakdown spectroscopy combined with filter paper modified with gold nanoclusters[J]. Biosensors, 2024, 14(6): 267. DOI: 10.3390/bios14060267.

[23] Luo Guocheng, Zhao Changjiang, Ma Shuxing, et al. Rapid detection of total chromium in water bodies using AuAgPd trimetallic nanoparticles-modified filter paper combined with laser-induced breakdown spectroscopy[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2025, 234: 107336. DOI: 10.1016/j.sab.2025.107336.

[24] Chen Yitao, Guo Shujie, Jiang Yufeng, et al. Direct analysis of heavy metal elements in liquid water using femtosecond laser-induced breakdown spectroscopy for high-sensitivity detection[J]. Talanta, 2025, 286: 127512. DOI: 10.1016/j.talanta.2024.127512.

[25] Chen Yitao, Jiang Yufeng, Han Junhao, et al. Sensitivity, precision, and accuracy of fs-LIBS for heavy metal detection in flowing aqueous solutions[J]. Optics Letters, 2024, 49(11): 3106-3109.

[26] Rai N K, Rai A K. LIBS: an efficient approach for the determination of Cr in industrial wastewater[J]. Journal of Hazardous Materials, 2008, 150(3): 835-838.

[27] Chen Yitao, Si Junrong, Wan Xingtao, et al. Sub-ppb detection limit of Cr, Pb, and Cu in water by surface-enhanced LIBS with femtosecond laser[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2024, 213: 106880. DOI: 10.1016/j.sab.2024.106880.

[28] Sun Chao, Tian Yang, Gao Lei, et al. Machine learning allows calibration models to predict trace element concentration in soils with generalized LIBS spectra[J]. Scientific Reports, 2019, 9: 11363. DOI: 10.1038/s41598-019-47751-y.

[29] Zhang Xianglin, Zhang Feng, Kung H T, et al. Estimation of the Fe and Cu contents of the surface water in the ebinur lake basin based on LIBS and a machine learning algorithm[J]. International Journal of Environmental Research and Public Health, 2018, 15(11): 2390. DOI: 10.3390/ijerph15112390.

[30] GB 3838—2002 地表水环境质量标准[S].
GB 3838—2002 Environmental quality standards for surface water[S].

[31] GB 8978—1996 污水综合排放标准[S].
GB 8978—1996 Integrated wastewater discharge standard[S].

[32] Liu Wenjie, Sun Zongtao, Chen Jingyu, et al. Raman spectroscopy in colorectal cancer diagnostics: comparison of PCA-LDA and PLS-DA models[J]. Journal of Spectroscopy, 2016, 2016: 1603609. DOI: 10.1155/2016/1603609.

[33] Belgiu M, Drăguţ L. Random forest in remote sensing: a review of applications and future directions[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 114: 24-31.

[34] Almeida L B. Multilayer perceptrons[M]//Handbook of Neural Computation. Bristol: IOP Publishing Ltd., 2004.

[35] Pisner D A, Schnyer D M. Support vector machine[M]// Machine Learning. Amsterdam: Elsevier, 2020: 101-121.

[36] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.

Classification of Low-concentration Chromium Pollution Grade in Water Based on Laser-induced Breakdown Spectroscopy (LIBS) Combined with Machine Learning

Chen Dingxi[1], Ke Chuan[1,2,3*], He Yaxiong[2,3], Wen Qifan[3], Qian Ziran[1], Zhao Dongye[4]

(*1. School of Physical Science and Technology, Southwest Jiaotong University, Chengdu, Sichuan 610031, China;*
2. School of Intelligent Manufacturing, Sichuan University of Arts and Science, Dazhou, Sichuan 635000, China;
3. School of Electrical Engineering, Southwest Jiaotong University, Chengdu, Sichuan 610031, China;
4. School of New Energy, Inner Mongolia University of Technology, Hohhot, Inner Mongolia 010051, China)

Abstract Heavy metal pollution in water bodies is a critical concern in environmental monitoring. Among these metals, chromium (Cr), widely used in industries such as electroplating, metallurgy, leather processing, and chemical manufacturing, can enter natural water systems through industrial wastewater, posing significant risks to ecosystems and human health. To achieve rapid and reliable identification of Cr contamination in water, this study proposed a classification method for water Cr pollution levels based on laser-induced breakdown spectroscopy (LIBS) combined with machine learning. To minimize the effects of sample inhomogeneity and plasma fluctuations, a mapping point matrix sampling strategy was employed in the coffee-ring deposition area. For each sample, 100 spectra were systematically collected using a 10×10 matrix, effectively addressing sample heterogeneity. The acquired spectra were processed using min-max normalization, Savitzky-Golay (S-G) smoothing, and asymmetric least squares baseline correction. These preprocessing steps substantially enhanced the linear correlation between characteristic peak intensities and Cr concentrations, improving the coefficient of determination (R^2) from 0.78 to 0.98. For machine learning modeling, 1 000 spectral data points were randomly divided into training, validation, and test sets at a ratio of 8∶1∶1, used for model training, hyperparameter optimization, and performance evaluation, respectively. Four widely used classification algorithms-partial least squares discriminant analysis (PLS-DA), random forest (RF), multilayer perceptron (MLP), and support vector machine (SVM) were compared to fully explore the relationship between spectral features and pollution levels and to assess the robustness of each algorithm, particularly for low-concentration samples. The results showed that the SVM model achieved the highest classification performance on both the validation and test sets, with an overall test accuracy of 0.99. Samples from all pollution levels, especially those in the low-concentration category, demonstrated consistently stable discrimination. Analysis using the confusion matrix and receiver operating characteristic (ROC) curves further confirmed the superiority of SVM in boundary discrimination, generalization capability, and multi-class classification accuracy. This approach provided a reliable technical solution for real-time grading of Cr pollution in industrial wastewater and supports the translation of LIBS technology from laboratory analysis to on-site intelligent decision-making. Future work may integrate multi-modal spectral data, such as combining Raman spectroscopy with LIBS, to further improve the accuracy and applicability of trace heavy metal pollution level identification, thereby offering more efficient tools for water quality monitoring and pollution control.

Keywords water Cr pollution; laser-induced breakdown spectroscopy (LIBS); machine learning; pollution level classification

HIGHLIGHTS

1) A classification model for low-concentration Cr pollution in water was established, enabling rapid and reliable identification of pollution levels.
2) The integration of spectral mapping and machine learning effectively reduced the influence of sample heterogeneity and plasma fluctuations, improving model stability and classification accuracy.
3) The proposed analytical strategy provides a new technical approach for rapid monitoring and risk assessment of heavy metal pollution in water.