

DOI: 10.19666/j.rlfed.202408177

激光诱导击穿光谱结合机器学习的 煤质多指标同步快速检测

熊凯^{1,2}, 邹祥波^{1,2}, 陈创庭², 杨先茂³, 陈公达², 张爽¹,
卢伟业⁴, 陈小玄^{3,4}, 卢志民³, 姚顺春³

(1.广东省能源集团有限公司, 广东 广州 510730;

2.广东能源集团科学技术研究院有限公司, 广东 广州 510630;

3.华南理工大学电力学院, 广东 广州 510641;

4.广东省特种设备检测研究院顺德检测院, 广东 佛山 528300)

[摘要] 煤质快速全面检测对锅炉燃烧优化和燃煤电厂数智化转型具有重要意义, 激光诱导击穿光谱 (laser-induced breakdown spectroscopy, LIBS) 在煤质快速检测中具有巨大应用潜力。为满足煤炭现场快速检测的应用目标, 利用煤粉颗粒流式 LIBS 检测实验装置采集了不同电厂的 46 组煤样光谱数据, 系统开展了 LIBS 与机器学习相结合的煤质多指标同步快速检测研究。针对颗粒流状态下光谱波动较大的特点, 对单脉冲光谱采集个数进行了优化, 并建立无效光谱筛选、光谱平均、光谱归一化数据预处理方法, 进一步对比了 PLSR、SVR、PSO-SVR 和 LSTM 4 种机器学习算法以及全光谱、特征波段、强度积分和 PCA 提取 4 种光谱特征输入对模型预测煤质多指标性能的影响。结果表明: 单次检测累计采集 200 个单脉冲光谱进行光谱平均, 光谱信号不确定性可控制在 5% 以内; 煤质多指标定量分析中 PSO-SVR 算法的预测性能最佳, 利用 PCA 算法对光谱数据进行降维, 在减少模型计算量的同时提高了模型预测性能, 两者结合所建立的模型对煤发热量预测均方根误差为 0.289 MJ/kg, 平均绝对误差为 0.231 MJ/kg; 对煤碳质量分数、灰分和挥发分预测结果同样较理想, 均方根误差分别为 0.987%、1.310% 和 1.612%, 平均绝对误差分别为 0.839%、1.014%、1.033%。研究结果表明结合合适的机器学习算法, LIBS 技术可以实现煤质多指标同步精准快速检测, 在煤炭高效清洁利用场景中具有广阔的应用前景。

[关键词] 煤质; 激光诱导击穿光谱; 机器学习; 高效清洁; 颗粒流

[引用本文格式] 熊凯, 邹祥波, 陈创庭, 等. 激光诱导击穿光谱结合机器学习的煤质多指标同步快速检测[J]. 热力发电, 2025, 54(4): 129-139. XIONG Kai, ZOU Xiangbo, CHEN Chuangting, et al. Simultaneous rapid determination of multiple coal quality indicators using laser-induced breakdown spectroscopy combined with machine learning[J]. Thermal Power Generation, 2025, 54(4): 129-139.

Simultaneous rapid determination of multiple coal quality indicators using laser-induced breakdown spectroscopy combined with machine learning

XIONG Kai¹, ZOU Xiangbo^{1,2}, CHEN Chuangting², YANG Xianmao³, CHEN Gongda²,
ZHANG Shuang¹, LU Weiye⁴, CHEN Xiaoxuan^{3,4}, LU Zhimin³, YAO Shunchun³

(1.Guangdong Energy Group Co., Ltd., Guangzhou 510730, China;

收稿日期: 2024-08-07

基金项目: 国家重点研发计划项目 (2021YFF0601001); 广东省新型电力系统技术创新项目 (1688950422168); 广东省自然科学基金杰出青年项目 (2021B1515020071); 中央高校自然科学类项目 (2023ZYGXZR090); 广东省自然科学基金项目 (2022A1515010741)

Supported by: National Key Research and Development Program (2021YFF0601001); Guangdong S&T Program (1688950422168); Guangdong Basic and Applied Basic Research Foundation (2021B1515020071); Fundamental Research Funds for the Central Universities (2023ZYGXZR090); Guangdong Natural Science Foundation (2022A1515010741)

第一作者简介: 熊凯 (1973), 男, 硕士, 高级工程师, 主要研究方向为电力行业绿色低碳技术研究与应用, xionгкаi@geg.com.cn.

通信作者简介: 姚顺春 (1983), 男, 博士, 教授, 博士生导师, 主要研究方向为能源清洁转化过程的智能感知和调控, epscyao@scut.edu.cn.

- 2.Guangdong Energy Group Science and Technology Research Institute Co., Ltd., Guangzhou 510630,China;
3.School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China;
4.Guangdong Institute of Special Equipment Inspection and Research Shunde Branch, Foshan 528300, China)

Abstract: The rapid and comprehensive determination of coal quality is of great significance for the optimization of boiler combustion and the digital transformation of coal-fired power plants. Laser-induced breakdown spectroscopy (LIBS) has the potential to be applied effectively in the rapid determination of coal quality. In order to meet the application goal of rapid coal inspection, 46 sets of spectral data of coal samples from different power plants were collected by the experimental device of coal particle flow LIBS, and the research of simultaneous rapid inspection of multiple indicators of coal quality by combining LIBS with machine learning was carried out systematically. In view of the considerable spectral fluctuations observed in the particle flow state, the number of single-pulse acquisitions was optimized. In addition, invalid spectral screening, spectral averaging and spectral normalization data preprocessing methods were established. Furthermore, four machine learning algorithms (PLSR, SVR, PSO-SVR, and LSTM) and four spectral feature inputs (full spectra, eigenbands, intensity integration, and PCA extraction) were compared in terms of their performance in predicting multiple indicators of coal quality. The results demonstrate that the uncertainty of the spectral signals can be maintained at a maximum of 5% when 200 single-pulse spectra are collected for spectral averaging in a single test. The PSO-SVR algorithm exhibits the most optimal prediction performance in the quantitative analysis of coal quality indicators, and the PCA algorithm reduces the dimensionality of the spectral data, which reduces the amount of model computation and at the same time improves the prediction performance of the model, and the model established by combining both of them has the best performance, the root mean square error (RMSEP) of the coal heat content is 0.289 MJ/kg, and the mean absolute error (MAE) is 0.231 MJ/kg. The coal carbon mass fraction, ash content and volatile matter content are also predicted satisfactorily, with the RMSEP of 0.987%, 1.310% and 1.612%, and the MAE of 0.839%, 1.014%, and 1.033%, respectively. The results show that, combined with appropriate machine learning algorithms, the LIBS technique can achieve simultaneous accurate and rapid determination of multiple indicators of coal quality, which has a broad application prospect in the scenario of efficient and clean coal utilization.

Key words: coal quality; laser-induced breakdown spectroscopy; machine learning; efficient cleaning; particle flow

2023 年我国能源消费总量达到 57.2 亿吨标准煤, 其中煤炭消费量占比为 55.3%^[1]。立足我国能源资源禀赋和能源安全战略, 短期内我国对煤炭资源仍将保持较高依赖, 快速全面煤质检测对推进煤炭清洁高效利用与煤炭工业数智化转型具有重要意义。燃煤电厂通过实时煤质信息可以及时调整燃烧过程中的煤耗、氧气供应和炉膛温度等参数, 实现最佳的燃烧效果和热能利用效率, 从而降低燃料消耗和污染物排放, 同时可根据煤的发热量、挥发分含量等参数, 结合负荷需求, 调节煤炭入炉量和燃烧工况, 提高机组的深度灵活调节能力, 保持发电系统的稳定运行^[2]。此外, 快速准确获取煤质信息可以缩短煤炭企业碳排放核算周期, 为制定科学合理的碳减排策略提供依据。

现有煤质快速检测技术主要基于光谱学方法, 如瞬发 γ 中子活化分析 (prompt gamma neutron activation analysis, PGNA) ^[3]、X 射线荧光光谱 (X-ray fluorescence spectroscopy, XRF) 技术^[4]、拉曼光谱 (Raman) 技术^[5]等。但上述技术在实际应用中仍存在检测指标不全面、安全性不足、现场适应性差等问题, 尚不能完全满足工业现场煤质快速全面

检测需求。近年来, 激光诱导击穿光谱 (laser-induced breakdown spectroscopy, LIBS) 技术凭借全元素同步分析、方便快捷、无辐射危害等优势, 在煤质在线检测领域快速发展^[6-8]。目前, 已工程化应用的 LIBS 煤质检测模式主要包括煤块、煤粉压片和煤粉颗粒流 3 类^[9-10], 其中, 煤块和煤粉压片基本实现了 LIBS 在电厂输送带侧的煤质快速检测, 但仍普遍存在取样代表性不足、集成设备较为庞大的问题。而煤粉颗粒流式 LIBS 检测无需复杂的制样流程, 可以更好发挥 LIBS 快捷优势, 测量代表性好, 且已有研究发现, 相比于压片, 煤粉颗粒流式 LIBS 检测可以有效降低基体效应影响^[11]。

由于煤炭结构组成复杂, 颗粒流形态下, 煤粉粒径及空间分布不均匀导致等离子体形态波动较大, 影响 LIBS 煤质定量分析性能。为解决该问题, 研究人员一方面通过改进实验设备, 如采用火花放电增强^[12]、靶体辅助增强^[13]、空间限制^[14]等方法, 实现煤粉颗粒流等离子体的稳定调控, 但这些方法一定程度上会增加实验装置的复杂性; 另一方面, 结合机器学习的化学计量学方法在 LIBS 煤质检测中越来越多被采用, 如偏最小二乘回归 (partial least

squares regression, PLSR)^[15]、支持向量回归(support vector regression, SVR)^[16]、核极限学习机(kernel based extreme learning machine, KELM)^[17]、人工神经网络(artificial neural network, ANN)^[18]等,这些算法可以从大样本数据中获取光谱与煤质指标映射关系,实现煤质指标的回归预测。目前,针对颗粒流式 LIBS 煤质分析研究中,相关定量分析研究主要集中在单一机器学习模型对单一指标的预测上,缺乏多模型对多指标预测的比较研究,且较少讨论不同光谱特征输入对模型预测影响。

基于上述背景,本文通过颗粒流式 LIBS 实验系统采集不同燃煤电厂的 46 个煤样激光诱导击穿光谱数据,并建立光谱数据预处理和 4 种机器学习预测模型,对煤样的发热量、碳质量分数、灰分质量分数和挥发分质量分数 4 个指标进行同步定量分析,通过多种评价指标对不同模型的预测效果进行对比,同时探讨不同光谱特征输入对模型预测误差和预测耗时的影响。

1 实验方法

1.1 实验系统

本实验所使用的颗粒流式 LIBS 实验系统原理如图 1 所示。

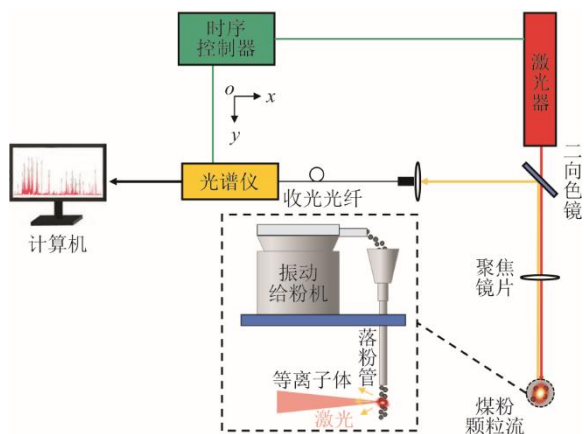


图 1 颗粒流 LIBS 实验系统原理
Fig.1 Schematic diagram of the particle flow LIBS experimental setup

实验系统主要组成器件有: 镭宝光电纳秒 Nd:YAG 脉冲激光器 Dara-200, 输出波长为 1 064 nm, 重复频率 1~20 Hz, 最大脉冲能量 100 MJ; Avantes 四通道型光纤光谱仪 AvaSpec-UL2048-4, 光谱探测范围为 200~850 nm, 分辨率为 0.05~0.10 nm, 最小积分时间为 30 μ s; Stanford Research System 时序信号发生器 DG535, 以及颗粒流样品单元、光路系统

和计算机。其中颗粒流样品单元由 SANKI 压电式振动给粉机 PEF-90A、调速控制器和落粉管组成。实验时激光透过光路系统中的长波通二向色镜和一个焦距为 100 mm 的聚焦透镜, 聚焦至颗粒流区域。图 1 中脉冲激光的入射光轴与颗粒流中心轴垂直相交, 煤粉颗粒在下落过程中被激光烧蚀击穿, 从而形成等离子体。采用了同轴收光方式采集光谱信号, 等离子体辐射光先通过透镜变成平行光, 再经过二向色镜以 90°反射, 被另一个焦距为 100 mm 的聚焦透镜汇集至信号收集器, 然后通过光纤传输到光谱仪的多个通道中进行分析。

1.2 实验样品

实验选取了不同燃煤电厂收集的 46 个煤样, 根据《煤样的制备方法》(GB/T 474—2008)^[19]进行制样, 得到粒径小于 0.2 mm 的空气干燥基煤样, 再根据《煤的工业分析方法》(GB/T 212—2008)^[20]、《煤的发热量测定方法》(GB/T 213—2008)^[21]和《煤的元素分析》(GB/T 31391—2015)^[22]对煤样的发热量、碳质量分数、灰分质量分数和挥发分质量分数进行测定。46 个煤样的发热量、碳质量分数、灰分质量分数和挥发分质量分数的参考值见表 1, 发热量为 19.23~29.14 MJ/kg, 碳质量分数为 48.50%~72.08%, 灰分质量分数为 8.91%~38.65%, 挥发分质量分数为 12.67%~46.82%, 其中随机选取 39 个煤样(C1—C39)作为训练集样本, 剩下的 7 个煤样(V1—V7)作为预测集样本。

1.3 煤样光谱数据采集

利用颗粒流式 LIBS 实验系统对 46 个煤样进行等离子体激发与光谱采集, 实验中激光器能量设置为 65 MJ, 脉宽为 10 ns, 重复频率为 5 Hz。光谱仪收光延迟设置为 2.0 μ s, 积分时间为 1.05 ms。图 2 为煤样 C1 的颗粒流 LIBS 单脉冲光谱。

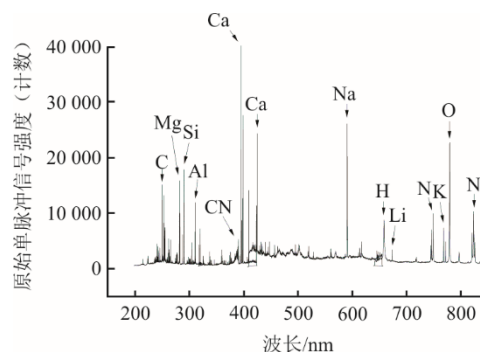


图 2 煤样原始单脉冲光谱
Fig.2 Raw single pulse spectroscopy of coal sample

经过与美国国家标准与技术研究院原子光谱数据库^[23]比对, 确认光谱中包含了煤主要元素 C、H、O、N 以及金属元素 Mg、Si、Al、Ca、Fe、Na、K 等 50 余条对应特征谱线, 说明颗粒流式激光诱导击穿光谱实验系统可以有效激发出煤粉等离子体光谱。

但由于颗粒流为非连续、不均匀的物质形态, 颗粒流 LIBS 检测中可能会出现无效激发的情况, 此外仪器自身产生的随机噪声、基线漂移等也会对 LIBS 光谱造成干扰, 导致其不能充分代表煤质特性, 因此在进行煤质定量分析之前, 煤样 LIBS 光谱的数据预处理必不可少^[24]。

表 1 煤样 4 个指标的参考值

Tab.1 Reference values for calorific value, carbon content, ash and volatile matter of coal samples

煤样编号	发热量/ (MJ·kg ⁻¹)	碳质量 分数/%	灰分 质量分数/%	挥发分质量 分数/%	煤样编号	发热量/ (MJ·kg ⁻¹)	碳质量 分数/%	灰分 质量分数/%	挥发分质量 分数/%
C1	26.27	65.08	14.71	38.27	C24	25.86	64.76	18.14	29.75
C2	27.03	64.03	12.29	40.11	C25	25.52	64.43	9.12	46.82
C3	27.65	63.19	9.13	40.33	C26	19.23	50.60	35.88	21.97
C4	27.72	64.05	9.61	40.88	C27	25.29	64.53	17.70	33.49
C5	28.00	64.40	8.91	40.66	C28	26.28	66.66	17.11	29.48
C6	28.04	64.13	10.46	38.05	C29	27.00	68.99	12.05	32.53
C7	27.65	62.26	9.44	41.20	C30	23.10	59.31	23.78	28.84
C8	27.22	62.47	10.28	39.60	C31	24.71	62.49	22.08	28.87
C9	27.00	62.17	12.96	37.03	C32	24.22	61.43	22.19	29.73
C10	23.39	60.97	26.96	33.90	C33	24.72	62.53	22.47	28.45
C11	19.35	48.50	38.03	23.99	C34	26.61	70.12	14.46	44.54
C12	25.52	64.26	18.75	29.58	C35	23.25	61.36	23.67	31.26
C13	21.57	55.12	35.25	12.67	C36	29.14	72.08	10.92	33.10
C14	22.92	58.32	31.44	13.34	C37	20.20	50.19	34.01	24.06
C15	26.36	66.43	16.51	30.36	C38	25.62	63.35	21.11	28.32
C16	27.18	68.36	14.13	30.57	C39	27.34	67.62	18.16	26.99
C17	25.20	63.66	20.22	27.62	V1	27.70	64.51	9.78	40.01
C18	22.91	58.11	28.60	23.86	V2	27.74	63.35	7.38	42.13
C19	24.34	61.30	26.30	20.63	V3	27.04	68.74	13.93	30.52
C20	20.78	52.46	38.65	14.81	V4	24.98	62.86	22.03	25.84
C21	21.48	53.85	33.14	23.90	V5	26.35	66.63	17.32	30.11
C22	19.71	50.36	33.51	24.83	V6	24.37	61.58	24.96	26.99
C23	21.94	56.30	28.43	26.26	V7	27.28	67.66	18.28	26.89

1.4 光谱数据预处理

针对煤颗粒流 LIBS 光谱特点, 本实验通过无效光谱筛选、光谱平均、光谱归一化 3 个步骤进行煤样 LIBS 光谱数据预处理。由于基于特征谱线强度的标准偏差 (standard deviation, SD) 值法剔除无效光谱具有较好的适用性, 本文采用该方法剔除无效光谱, SD 值法计算公式为^[25]:

$$\delta_{SD} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (I_i - \bar{I})^2} \quad (1)$$

式中: N 为特征峰内像素点的个数; I_i 为特征峰内第 i 个像素点的光谱强度; \bar{I} 为特征峰内 N 个波长点光谱强度的平均值。

SD 值法认为当特征峰强度的 SD 值低于设定阈值时, 则对应的光谱为无效光谱, SD 阈值的选取

在实际测量中与光谱仪的分辨率和背景噪声有关, 具有一定经验性。本文选用煤主要组成元素碳元素特征谱线 C I 247.86 nm 计算 SD 值, 将特征峰 SD 值小于 400 的光谱定义为无效光谱, 予以剔除, 随后将保留下来的有效光谱进行光谱平均处理, 单个煤样光谱平均个数将在下一节进行具体优化分析。

此外, 在 LIBS 分析中, 为克服等离子体时空演化的不稳定及样品之间的非化学差异和仪器响应的影响, 通常采用归一化方法对数据进行预处理。本文采用最大最小归一化方法, 将原始数据映射到 [0,1] 内, 最大最小归一化公式为:

$$X' = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (2)$$

式中: X' 为归一化后的数据; X 为原始数据; X_{\min} 为

原始数据中的最小值； X_{\max} 为原始数据中的最大值。

图3为煤样C1光谱数据经上述3个步骤预处理后的谱线图。经对比，预处理后元素特征谱线的相对强度不同与原始单脉冲的谱线，原因在于预处理步骤有效降低了光谱数据的随机性和偶然误差、并剔除了无效光谱信息，同时归一化使得光谱数据更具可比性。这些预处理方法共同提升了光谱数据的准确性和可靠性，为后续煤质多指标定量分析提供了更加准确和可靠的光谱数据信息。

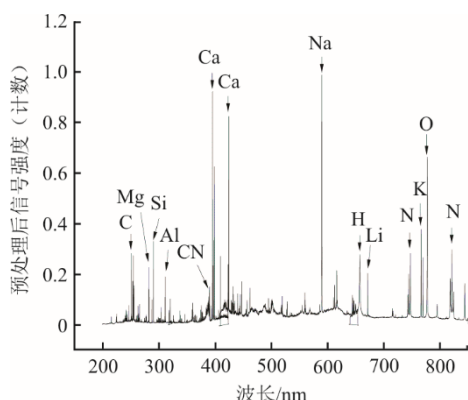


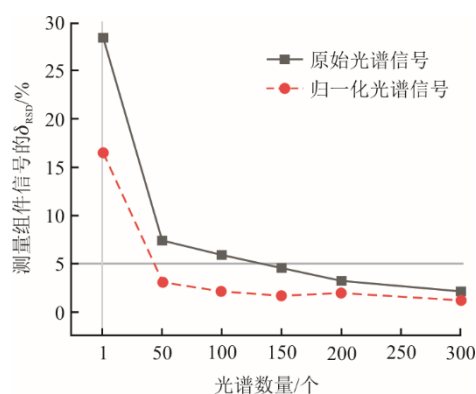
图3 单个煤样预处理后 LIBS 光谱
Fig.3 LIBS spectra of individual coal samples after pre-processing

1.5 单个煤样光谱采集个数优化

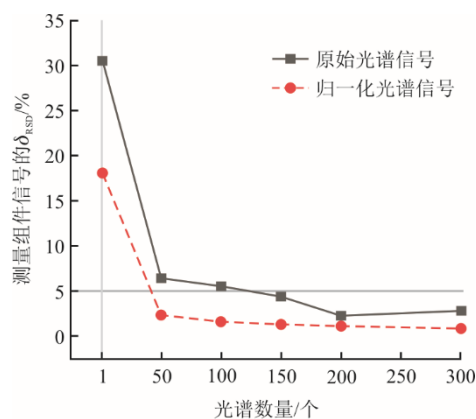
在 LIBS 测量中，单个脉冲光谱信号往往带有一定随机性，不能很好地反映样品原本的元素组成信息，平均多个脉冲光谱信号可以降低光谱信号随机波动的影响，提升定量分析稳定性^[8]，但平均的光谱数量直接影响煤质分析速度。因此，为尽可能降低煤粉颗粒流光谱信号的不确定性，同时保证煤质定量分析速度，对单次测量采集的光谱数量进行了优化。在固定测量参数条件下，对煤样 C1 累计采集 900 个单脉冲光谱，并进行多次分组，每次分组分别得到多组含 50、100、150、200、300 个单脉冲光谱的小组。将每组有效光谱进行平均，作为代表单个煤样的平均光谱。

光谱信号不确定性是指在同一样品上多次重复测量信号的波动，一般通过计算测量与测量平均（每次测量平均了多个脉冲）光谱信号的相对标准偏差（relative standard deviation, RSD） δ_{RSD} 来评价。图4呈现了 C I 247.86 nm、Si I 288.16 nm、Al I 309.22 nm 和 CN 388.29 nm 多条特征谱线峰值信号强度测量组之间的 δ_{RSD} 随平均光谱数量的变化规律，包括原始和经过归一化处理后的峰值信号强

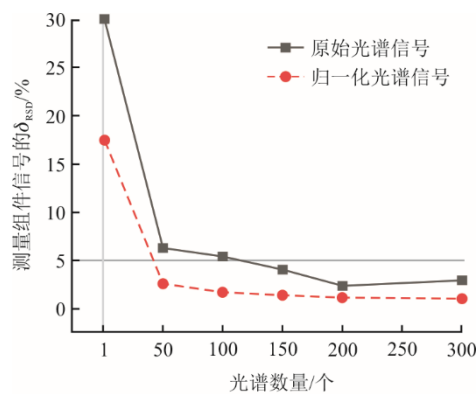
度。由图4分析可知，光谱归一化可以有效降低各谱线测量组间的信号 δ_{RSD} ，进而减少光谱信号波动性的影响；在未进行多光谱平均（平均光谱数量为1）时谱线测量组间 δ_{RSD} 超过了15%，信号存在较大波动，随着平均光谱数量的增加测量组间 δ_{RSD} 减小，并逐步趋于稳定；当单次测量的平均光谱数量达到150个以上时，特征谱线信号强度的测量组间 δ_{RSD} 均下降到5%以内，平均光谱数量达到200个左右时， δ_{RSD} 最低。根据该结果，对每个煤样分别采集200个 LIBS 光谱数据进行后续分析，在激光脉冲频率为5 Hz 的条件下，单个煤样的光谱采集耗时可以控制在1 min 以内。



a) 谱线C I 247.86 nm



b) 谱线Si I 288.16 nm



c) 谱线Al I 309.22 nm

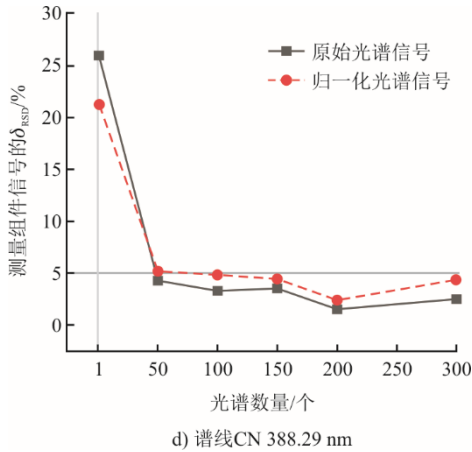


图 4 平均光谱数量对光谱信号波动性的影响
Fig.4 Influence of the average spectral number on the volatility of the spectral signal

1.6 定量分析模型构建

图 5 为煤质多指标定量分析模型的构建流程。原始光谱数据经预处理后，再利用特征工程方法对输入机器学习模型数据特征进行选取，接着对数据划分训练集与预测集，建立不同机器学习回归模型对煤发热量、碳质量分数、灰分质量分数和挥发分质量分数 4 个指标的预测分析结果。

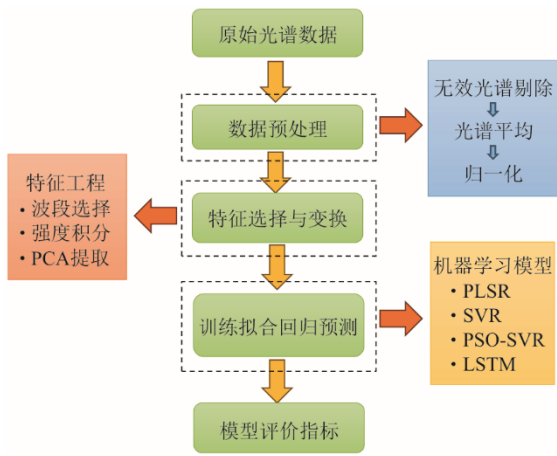


图 5 模型构建流程
Fig.5 Model building process

1.6.1 机器学习模型

为提高 LIBS 测量的准确度，研究人员已将不同的化学计量方法用于 LIBS 的定量分析中，通过机器学习算法对光谱数据的训练拟合，LIBS 技术不仅可以实现传统的元素分析，还可以对煤炭灰分、挥发分等工业指标进行预测测量。本文结合现有研究并综合考虑 LIBS 光谱数据特点选择建立偏最小二乘回归 (PLSR)、支持向量回归 (SVR)、粒子群优化支持向量回归 (particle swarm optimization-support

vector regression, PSO-SVR)、长短时记忆神经网络 (long short-term memory, LSTM) 4 种机器学习算法对煤质多指标进行回归预测，并比较在相同特征输入下不同模型分析性能。其中，PLSR 是融合了主成分分析法和多元线性回归的算法，可以有效消除多元线性回归中各变量的共线性，剔除了冗余信息，可提高计算效率^[26]；SVR 是一种非线性回归算法，将样本集从原始特征空间映射到高维特征空间，在高维空间中对样本集进行回归分析，在处理高维特征空间上优势较好^[27]；PSO-SVR 是一种改进的支持向量回归算法，基于 PSO 群体智能算法实现 SVR 的超参数自动寻优选择，可避免人工调参的盲目性，减少调参工作量^[28]；LSTM 是一种递归神经网络 (recursive neural network, RNN) 的变体，通过多个 LSTM 单元的堆叠来构建深层网络，能够拟合复杂的非线性关系^[29]。

根据训练集均方根误差，对定量模型的超参数进行系统调优。调优后 PLSR 模型中主成分数量选择为 10；SVR 模型关键参数惩罚因子 c 、径向基核函数参数 g 分别选取为 100 和 0.001；LSTM 模型中 LSTM 层选择为 6 层，最大迭代次数选取为 2 500。PSO-SVR 模型利用 PSO 算法可实现 c 、 g 参数的自动寻优，但仍需对基础参数进行设置：PSO 算法最大进化代数设置为 100，种群数目为 20，惩罚因子 c 寻优范围为 [0.1, 1 000]，径向基核函数参数 g 寻优范围为 [0, 100]。

1.6.2 光谱特征选取方法

煤炭等离子体光谱数据具有高维、海量、非线性的特点^[30]，本实验选择的四通道光谱仪，单个样品单张光谱有 8 192 个光谱特征变量，直接选择全光谱数据输入，模型计算复杂度大，计算耗时长。而采取合适的特征选取方法不仅可以实现数据降维，还可以校正由不相关变量引起的波动干扰，有助于建立更加稳健、更加准确的回归预测模型^[31]。基于全光谱，本文采取了特征选择和特征变换的不同特征工程方法，特征选择为选择包含煤炭主要元素特征谱线的光谱波段作为特征变量输入，本实验对单个样品选取了包含 C、H、O、Si、Al、Ca、Fe、Mg 等元素的 55 个特征波段共 805 个特征输入；特征变换为对选择的光谱波段进行强度-波长积分运算，即利用主要元素特征谱线的光谱面积作为特征输入；以及对光谱经主成分分析 (principal component analysis, PCA) 降维之后的特征作为输

入。PCA 是一种无监督降维方法，可实现将庞大的数据集简化为低维主成分^[32]，由于不同煤质指标对应的最佳主成分个数不同，本实验对单个煤样数据先降维至 30 个主成分，再根据后续模型预测结果选择输入模型主成分个数，表 2 为光谱特征选取方法汇总。

表 2 光谱特征选取方法
Tab.2 Spectral feature selection methods

输入特征	特征选取方法	特征数量
全光谱	无	8 192
特征波段	主要元素特征谱线的光谱波段	805
强度积分	光谱波段积分	55
PCA 提取	PCA 算法降维	<30 (根据各指标预测误差选取)

1.6.3 模型评价指标

以训练集决定系数 R^2 、预测集决定系数 R_p^2 、预测均方根误差 (root mean square error of prediction, RMSEP) δ_{RMSEP} 、平均绝对误差 (mean absolute error, MAE) δ_{MAE} 来评价模型对煤炭多指标预测准确度性能。其中决定系数越接近 1、 δ_{RMSEP} 和 δ_{MAE} 越接近 0 意味着模型准确度越高。 R^2 、 δ_{RMSEP} 和 δ_{MAE} 的计算公式为：

$$R_t^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3)$$

$$R_p^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2} \quad (4)$$

$$\delta_{RMSEP} = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}} \quad (5)$$

$$\delta_{MAE} = \frac{\sum_{i=1}^m |y_i - \hat{y}_i|}{m} \quad (6)$$

式中： n 为训练集样品数量； m 为预测集的样品数量； y_i 为样品指标的参考值； \bar{y}_i 为样品指标参考值的平均值； \hat{y}_i 为样品指标的预测值。

2 结果与讨论

2.1 4 种机器学习模型预测结果对比

表 3 为全光谱特征输入下 4 种机器学习模型的对煤质发热量、碳质量分数、灰分质量分数和挥发分质量分数 4 项指标预测结果。结果表明，PLSR 模型除了对煤炭碳质量分数有较小预测误差外，对发热量、灰分质量分数、挥发分质量分数指标均出现过拟合现象， δ_{RMSEP} 和 δ_{MAE} 均较大。SVR 模型在

预测 4 个煤质指标时， R_t^2 均接近 1，有较好的训练拟合度，但预测误差均较大，同样出现较为明显的过拟合现象；针对复杂的煤质光谱数据以及多指标同步分析需求，LSTM 模型因模型复杂度过高而出现过拟合，且训练时间较长。相比之下，PSO-SVR 模型结合了粒子群寻优算法 (PSO) 和支持向量回归 (SVR) 的优势，PSO 算法通过自动优化 SVR 模型的超参数，避免了传统方法中参数选择的盲目性与繁琐性，显著提高了模型的泛化能力和预测精度，在保证较好训练拟合度的同时，对预测集中 4 个煤质指标的预测误差也均最小，其中发热量、碳质量分数、灰分质量分数和挥发分质量分数的 δ_{RMSEP} 分别达到 0.328 MJ/kg、1.235%、1.568%、2.000%， δ_{MAE} 分别达到 0.266 MJ/kg、1.043%、1.340%、1.330%。

表 3 4 种机器学习模型对煤样指标预测结果
Tab.3 Prediction results of four machine learning models on coal sample indicators

机器学习模型	模型评价指标	发热量/ (MJ·kg ⁻¹)	碳质量分数/%	灰分质量分数/%	挥发分质量分数/%
PLSR	R_t^2	0.991	0.985	0.994	0.983
	R_p^2	0.763	0.775	0.822	0.758
	δ_{RMSEP}	0.603	1.175	3.008	3.008
	δ_{MAE}	0.488	1.071	2.193	2.421
SVR	R_t^2	0.998	0.998	0.998	0.998
	R_p^2	0.656	0.855	0.859	0.696
	δ_{RMSEP}	4.104	1.407	3.002	3.586
	δ_{MAE}	2.957	1.156	2.475	2.42
PSO-SVR	R_t^2	0.915	0.865	0.932	0.925
	R_p^2	0.931	0.750	0.928	0.892
	δ_{RMSEP}	0.328	1.235	1.568	2.000
	δ_{MAE}	0.266	1.043	1.340	1.330
LSTM	R_t^2	0.987	0.839	0.999	0.980
	R_p^2	0.556	0.414	0.865	0.739
	δ_{RMSEP}	0.826	1.896	2.156	3.119
	δ_{MAE}	0.768	1.609	1.618	2.297

2.2 不同光谱特征输入预测结果对比

根据不同机器学习模型预测对比结果，PSO-SVR 模型表现最佳，本小节进一步对比不同光谱特征 (全光谱、特征波段、强度积分和 PCA 提取) 输入 PSO-SVR 模型对煤样发热量、碳质量分数、灰分质量分数和挥发分质量分数的预测结果。由于 PCA 在降维过程中能够去除部分噪声和冗余信息，但不同煤质指标对噪声和冗余信息的敏感程度可

能不同,不同指标对应的主要信息分布在不同的主成分上。因此,在降维过程中保留的主成分个数需根据具体指标进行调整以优化预测结果。根据预测集 δ_{MAE} 进行调优选取各指标保留的主成分个数,具体如图6所示。由图6可知,发热量、碳质量分数、灰分质量分数、挥发分质量分数4个指标分别在主成分个数为11、30、7、28时预测 δ_{MAE} 最小。

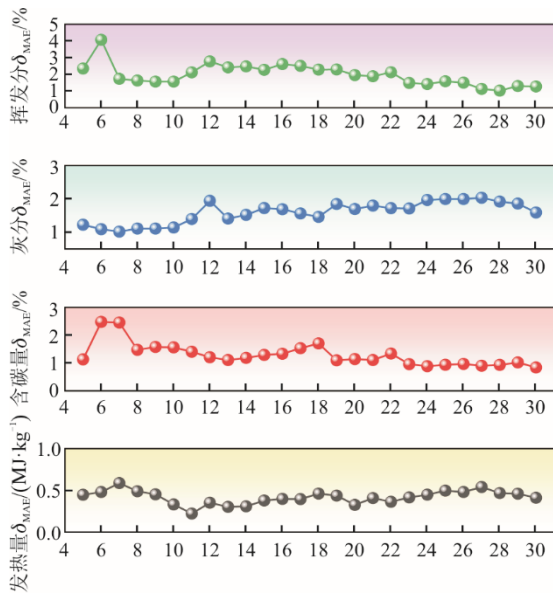


图6 主成分个数与各指标预测 δ_{MAE} 关系

Fig.6 Relationship between the number of principal components and the predicted δ_{MAE} value of each indicator

表4为不同特征输入下的预测结果对比。结果表明,对比全谱输入,特征波段及其强度积分输入均会导致模型预测误差增加,预测效果反而变差。而利用PCA提取特征输入,发热量、碳质量分数、

灰分质量分数和挥发分质量分数的 R_p^2 由全光谱下的 0.931、0.750、0.928 和 0.892 提高至 0.946、0.841、0.950 和 0.930, δ_{RMSEP} 分别由全光谱下的 0.328 MJ/kg、1.235%、1.568% 和 2.000% 降低到 0.231 MJ/kg、0.839%、1.014%、1.033%, δ_{MAE} 分别由全光谱下的 0.266 MJ/kg、1.043%、1.340%、1.330% 降低至了 0.231 MJ/kg、0.839%、1.014%、1.033%。煤质测量指标准确度与国家标准对比见表5。由表5可见,上述检测准确度优于国家标准《中子活化型煤炭在线分析仪》(GB/T 29161—2012) [33],且与其他文献报道的煤粉压片 LIBS 检测精度相近 [9-10]。

表4 不同特征输入 PSO-SVR 模型对煤样指标预测结果
Tab.4 Prediction results of coal sample indexes by different spectral features input into PSO-SVR model

输入特征	模型评价指标	发热量/(MJ·kg ⁻¹)	碳质量分数/%	灰分质量分数/%	挥发分质量分数/%
全光谱	R_p^2	0.931	0.750	0.928	0.892
	δ_{RMSEP}	0.328	1.235	1.568	2.000
	δ_{MAE}	0.266	1.043	1.340	1.330
特征波段	R_p^2	0.868	0.582	0.943	0.883
	δ_{RMSEP}	0.452	1.598	1.402	2.089
	δ_{MAE}	0.409	1.114	1.022	1.443
强度积分	R_p^2	0.681	0.644	0.900	0.628
	δ_{RMSEP}	0.702	2.608	1.855	3.272
	δ_{MAE}	0.535	1.897	1.639	2.501
PCA提取	R_p^2	0.946	0.841	0.950	0.930
	δ_{RMSEP}	0.289	0.987	1.310	1.612
	δ_{MAE}	0.231	0.839	1.014	1.033

表5 煤质测量指标准确度与国家标准对比

Tab.5 Comparison of the accuracy of coal quality measurement indicators with national standards

测量指标	《中子活化型煤炭在线分析仪》(GB/T 29161—2012)		本文 (PCA 提取输入 PSO-SVR 模型)	指标差距
	静态准确度	动态准确度		
发热量/(MJ·kg ⁻¹)	0.65	0.8	0.231	优于标准
碳质量分数/%	40~85: 1.5		0.839	优于标准
	45~65: 1.0			
灰分质量分数/%	<15: 1.2	<15: 1.8	1.014	优于标准
	15~30: 1.6	15~30: 2.3		
	>30: 2.0	>30: 2.8		
挥发分质量分数/%			1.033	标准未规定

2.3 不同光谱特征输入模型运行耗时对比

在煤炭工业应用现场,对煤质分析时效性有较高要求。本文还对比了不同光谱特征输入 PSO-SVR 模型对煤质指标预测耗时。全光谱输入耗时 27.43 s,特征波段输入模型从数据输入到获取最终预测结

果,4个指标的平均运行时长降低至 4.73 s,积分强度与 PCA 提取特征输入,进一步降低为 2.59 s 和 2.75 s,相比全光谱特征输入模型计算耗时减少近 10 倍。由此可见,通过特征筛选与特征变换降低特征维度,可以显著降低模型训练预测耗时。

综合以上分析,PCA 由于通过线性变换将高维光谱数据映射到低维空间,保留数据中最重要的信息,使得特征空间更为简洁和有效,将 PCA 降维后的数据输入 PSO-SVR 模型不仅预测精度高而且计算耗时少,有利于在实际热力发电应用场景中,实现快速精准响应。图 7 为采用 PCA 方法对煤样光谱数据进行降维后输入 PSO-SVR 进行煤质发热量、碳质量分数、灰分质量分数、挥发分质量分数预测结果。结果表明该模型在分析速度和准确性上都可以满足大多数煤炭工业应用煤质快速检测需求。

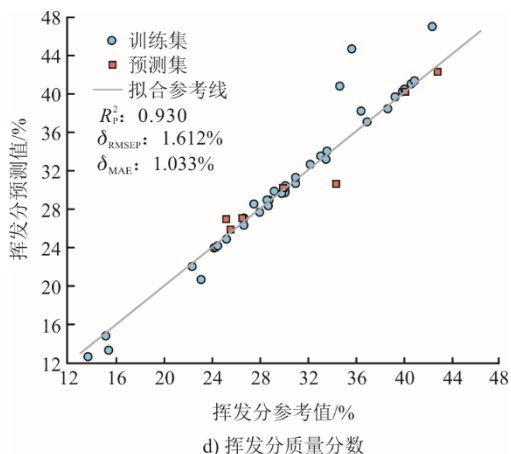
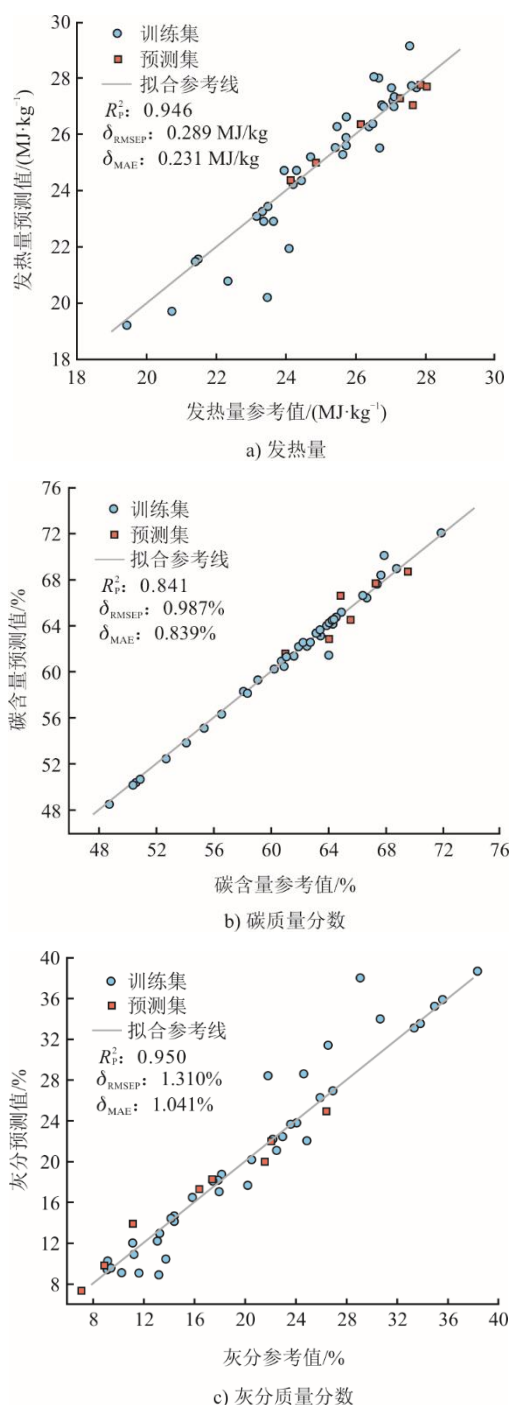


图 7 PCA 提取输入 PSO-SVR 模型煤质指标预测结果
Fig.7 Prediction results of coal quality indexes for the PCA feature extraction input PSO-SVR model

3 结 论

本文基于颗粒流式 LIBS 检测实验系统,较为系统地开展了煤颗粒流等离子体光谱采集与数据预处理工作,并结合特征工程与机器学习算法,建立了偏最小二乘回归(PLSR)、支持向量回归(SVR)、粒子群优化支持向量回归(PSO-SVR)和长短时记忆网络(LSTM)模型对煤炭发热量、碳质量分数、灰分质量分数和挥发分质量分数 4 个指标进行了定量分析,分析结果如下。

1) 数据预处理对于煤质定量分析至关重要,光谱平均、光谱归一化可以有效降低光谱信号的不确定性。单次煤样检测采集 200 个单脉冲光谱,特征谱线信号强度的测量组间 δ_{RSD} 在 5% 以内。定量分析算法中,相比其他 3 类模型,PSO-SVR 可以利用粒子群优化算法自动寻优选出模型的最佳关键参数,在避免煤质多指标预测下的人工调参复杂工作的同时,模型对煤炭,4 个指标预测误差均最小。

2) 通过特征选择与提取降低光谱数据特征维度,可以有效减少模型计算耗时,相比于全光谱特征输入,通过 PCA 进行特征提取后输入 PSO-SVR 模型,单个煤质指标的预测时间降低至 2.75 s。同时,PCA 特征提取可以实现数据降维的同时保留数据中最重要的信息,使得特征空间更为简洁和有效,有效提高了 PSO-SVR 模型预测性能,煤炭发热量、碳质量分数、灰分质量分数和挥发分质量分数 4 个指标预测 δ_{MAE} 分别由全光谱下的 0.266 MJ/kg、1.043%、1.340%、1.330% 降低至 0.231 MJ/kg、0.839%、1.014%、1.033%。

3) 颗粒流式 LIBS 结合适当的特征工程与机器

学习算法可以实现煤炭碳质量分数、发热量、灰分质量分数、挥发分质量分数的多指标同步快速检测。其中,实验光谱采集端对煤样等离子体多光谱采集后进行平均,耗时仅需1 min以内;数据处理端仅需15 s以内可以计算得到煤质多指标检测结果,且多指标检测准确性能够满足煤质日常快速检测需求。该技术可作为一种实时、快速的煤炭检测方法,应用于煤炭生产与运输、煤炭分选、煤炭加工利用等多种场合。

参考文献

- [1] 国家统计局. 中华人民共和国2023年国民经济和社会发展统计公报[EB/OL]. (2024-02-29)[2024-06-28]. <http://www1.xinhuanet.com/politics/20240301/2c4a1c0966a647abae6a864aecc68ef9/c.html>. National Bureau of Statistics. Statistical Bulletin of the People's Republic of China on National Economic and Social Development in 2023[EB/OL]. (2024-02-29)[2024-06-28]. <http://www1.xinhuanet.com/politics/20240301/2c4a1c0966a647abae6a864aecc68ef9/c.html>.
- [2] 王洪磊, 郭鑫, 张亦凡, 等. 煤质煤量全面在线检测技术发展现状及应用进展[J]. 煤炭科学技术, 2024, 52(2): 219-237. WANG Honglei, GUO Xin, ZHANG Yifan, et al. Research progress and application of online coal quality and coal quantity analyses[J]. Coal Science and Technology, 2024, 52(2): 219-237.
- [3] 周海渊, 郭世明, 宋青锋, 等. 中子活化煤质分析仪在沙曲选煤厂的应用[J]. 山西焦煤科技, 2021, 45(6): 8-10. ZHOU Haiyuan, GUO Shiming, SONG Qingfeng, et al. Application of neutron activated coal quality analyzer in Sha Qu coal preparation plant[J]. Shanxi Coking Coal Science & Technology, 2021, 45(6): 8-10.
- [4] 林春强, 于波, 刘志刚, 等. 四棵煤矿中的应用[J]. 数字技术与应用, 2023, 41(3): 58-61. LIN Chunqiang, YU Bo, LIU Zhigang, et al. Application of X-ray ash meter in Sikeshe Coal Mine[J]. Digital Technology & Application, 2023, 41(3): 58-61.
- [5] 于鹏峰, 苏攀, 刘佳薇, 等. 基于PL-Raman光谱分析的煤质快速检测方法[J]. 动力工程学报, 2022, 42(3): 215-220. YU Pengfeng, SU Pan, LIU Jiawei, et al. Rapid evaluation method of coal property using PL-Raman spectroscopy[J]. Journal of Chinese Society of Power Engineering, 2022, 42(3): 215-220.
- [6] 徐水秀, 喻子彧, 覃淮青, 等. 基于激光诱导击穿光谱的煤质快速分析研究及应用[J]. 量子电子学报, 2021, 38(6): 727-750. XU Shuixiu, YU Ziyu, QIN Huaiqing, et al. Research and application of rapid analysis of coal quality by laser-induced breakdown spectroscopy[J]. Chinese Journal of Quantum Electronics, 2021, 38(6): 727-750.
- [7] 陈小玄. 基于激光诱导击穿光谱的生物质燃料特性分析研究[D]. 广州: 华南理工大学, 2021: 1. CHEN Xiaoxuan. Study on the biomass fuel properties based on laser-induced breakdown spectroscopy[D]. Guangzhou: South China University of Technology, 2021: 1.
- [8] 荏方, 王庆松, 李承峻, 等. 基于LIBS和NIRS信号同步采集和融合的入炉煤炭发热量测量研究[J]. 热力发电, 2023, 52(7): 92-98. CHI Fang, WANG Qingsong, LI Chengjun, et al. Study on the measurement of coal as fired calorific value based on synchronous collection and fusion of LIBS and NIRS signals[J]. Thermal Power Generation, 2023, 52(7): 92-98.
- [9] ZHANG L, GONG Y, LI Y F, et al. Development of a coal quality analyzer for application to power plants based on laser-induced breakdown spectroscopy[J]. Spectrochimica Acta Part B, 2015, 113: 167-173.
- [10] SONG W, HOU Z, GU W, et al. Industrial at-line analysis of coal properties using laser-induced breakdown spectroscopy combined with machine learning[J]. Fuel, 2021, 306: 121667.
- [11] YU Z, YAO S, JIANG Y, et al. Comparison of the matrix effect in laser induced breakdown spectroscopy analysis of coal particle flow and coal pellets[J]. Journal of Analytical Atomic Spectrometry, 2021, 36(11): 2473-2479.
- [12] YAO S, XU J, ZHANG X, et al. Real-time measurement of constituents in solid materials using particle flow spark induced breakdown spectroscopy[J]. Journal of Analytical Atomic Spectrometry, 2018, 33(6): 986-991.
- [13] YU Z, YAO S, ZHANG L, et al. Surface-enhanced laser-induced breakdown spectroscopy utilizing metallic target for direct analysis of particle flow[J]. Journal of Analytical Atomic Spectrometry, 2019, 34(1): 172-179.
- [14] CHEN H, DONG M, CAI J, et al. An optimization method based on spatial confinement for direct detection of laser-induced particle flow[J]. Journal of Analytical Atomic Spectrometry, 2023, 38(6): 1224-1231.
- [15] 刘谕霖. 基于机器学习及特征工程的激光诱导击穿光谱技术用于煤质的研究[D]. 济南: 山东大学, 2023: 1. LIU Yulin. Investigation of coal quality by using laser-induced breakdown spectroscopy assisted with machine learning and feature engineering[D]. Jinan: Shandong University, 2023: 1.
- [16] ZHANG L, GONG Y, LI Y, et al. Development of a coal quality analyzer for application to power plants based on laser-induced breakdown spectroscopy[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2015, 113: 167-173.
- [17] YAN C, LIANG J, ZHAO M, et al. A novel hybrid feature selection strategy in quantitative analysis of laser-induced breakdown spectroscopy[J]. Analytica Chimica Acta, 2019, 1080: 35-42.
- [18] YAO S, MO J, ZHAO J, et al. Development of a rapid coal analyzer using laser-induced breakdown spectroscopy (LIBS)[J]. Applied Spectroscopy, 2018, 72(8): 1225-1233.
- [19] 煤样的制备方法: GB/T 474—2008[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 2008: 1. Preparation method of coal sample: GB/T 474—2008[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, 2008: 1.
- [20] 煤的工业分析方法: GB/T 212—2008[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 2008: 1. Methods of industrial analysis of coal: GB/T 212—2008[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's

- Republic of China, 2008: 1.
- [21] 煤的发热量测定方法: GB/T 213—2008[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 2008: 1.
Method of determining the heat content of coal: GB/T 213—2008[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, 2008: 1.
- [22] 煤的元素分析: GB/T 31391—2015[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 2015: 1.
Elemental analysis of coal: GB/T 31391—2015[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, 2015: 1.
- [23] KRAMIDA A, RALCHENKO Y, READER J, et al. NIST atomic spectra database (version 5.11)[EB/OL] (2023-05-11)[2024-07-13]. <https://physics.nist.gov/asd>.
- [24] 覃淮青. 基于煤炭特性的激光诱导击穿光谱检测煤质优化方法研究[D]. 广州: 华南理工大学, 2023: 1.
QIN Huaiqing. Study on optimization method of coal property analysis by laser-induced breakdown spectroscopy based on coal characteristics[D]. Guangzhou: South China University of Technology, 2023: 1.
- [25] YAO S C, ZHANG L F, YIN K J, et al. Identifying laser induced plasma emission spectra of particles in a gas-solid flow based on the standard deviation of intensity across an emission line[J]. *Journal of Analytical Atomic Spectrometry*, 2018, 33(10): 1676-1682.
- [26] LU Z, CHEN X, YAO S, et al. Feasibility study of gross calorific value, carbon content, volatile matter content and ash content of solid biomass fuel using laser-induced breakdown spectroscopy[J]. *Fuel*, 2019, 258: 116150.
- [27] ZHANG Z, HONG W C, DONG Y. Multi-hyperplane twin support vector regression guided with fuzzy clustering[J]. *Information Sciences*, 2024, 666: 120435.
- [28] WU W, CHEN K, TSOTSAS E. Prediction of particle mixing in rotary drums by a DEM data-driven PSO-SVR model[J]. *Powder Technology*, 2024, 434: 119365.
- [29] 杨奕, 刘金源, 陈天民, 等. 基于 PSO-SVR 模型的短期天然气负荷预测[J]. *科学技术与工程*, 2023, 23(35): 15210-15216.
YANG Yi, LIU Jinyuan, CHEN Tianmin, et al. Short-term natural gas load forecasting based on PSO-SVR model[J]. *Science Technology and Engineering*, 2023, 23(35): 15210-15216.
- [30] 李艳坤, 董汝南, 张进, 等. 光谱数据解析中的变量筛选方法[J]. *光谱学与光谱分析*, 2021, 41(11): 3331-3338.
LI Yankun, DONG Runan, ZHANG Jin, et al. Variable selection methods in spectral data analysis[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(11): 3331-3338.
- [31] 郭广盟. 紧凑型激光诱导击穿光谱 (LIBS) 仪器研制与溶液、土壤分析方法研究[D]. 长春: 吉林大学, 2019: 1.
GUO Guangmeng. Development of compact laser-induced breakdown spectroscopy (LIBS) instrument and analysis methods of solutions and soils[D]. Changchun: Jilin University, 2019: 1.
- [32] 张杰, 曲洪权, 柳长安, 等. 基于双子空间 PCA 降维的脑力负荷分类[J]. *科学技术与工程*, 2024, 24(11): 4433-4438.
ZHANG Jie, QU Hongquan, LIU Chang'an, et al. Classification of mental workload based on dimension reduction of PCA in two subspaces[J]. *Science Technology and Engineering*, 2024, 24(11): 4433-4438.
- [33] 中子活化型煤炭在线分析仪: GB/T 29161—2012[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 2012: 1.
Specifications of on-line analyzer for coal based on neutron activation analysis: GB/T 29161—2012[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, 2012: 1.

(责任编辑 邓玲惠)