DOI: 10.19666/j.rlfd.202509008

基于 TrAdaBoost 迁移学习的跨仪器 LIBS 煤质定量分析模型研究

邹祥波[1,2]，饶睦敏[1]，陈公达[1]，谭淑雯[3]，秦士伟[1]，
匡 草[1]，叶 骥[1]，姚顺春[3]，覃淮青[3]
（1. 广东能源集团科学技术研究院有限公司，广东 广州 510630；
2. 广东省能源集团有限公司，广东 广州 510730；
3. 华南理工大学电力学院，广东 广州 510641）

［摘 要］【目的】激光诱导击穿光谱（LIBS）技术由于具有无需复杂样品预处理、多指标同步检测、快速分析等优点，在煤质快速分析领域极具应用潜力。然而，由于不同仪器对光谱信号响应差异，基于主仪器光谱数据训练的定量分析模型应用于其他从仪器时精确度严重下降。因此，本研究基于 TrAdaBoost 迁移学习，结合不同机器学习算法构建跨仪器 LIBS 的煤质定量分析模型。【方法】将 2 台 LIBS 煤质快速检测仪分别视为主仪器和从仪器，在主从仪器上采集不同数量煤样的 LIBS 光谱。分别利用随机森林（RF）、支持向量机（SVR）以及 TrAdaBoost 结合 RF（TrA-RF）和 TrAdaBoost 结合 SVR（TrA-SVR）构建煤质定量分析模型，并在从仪器上分析未知煤样煤质对比模型性能。【结果】结果表明，与 RF 和 SVR 模型相比，TrA-RF 和 TrA-SVR 模型预测煤质准确性均显著提高。其中，TrA-RF 模型预测准确性最好，与 RF 模型对比，发热量、灰分和含碳量预测平均绝对误差由 1.390 MJ/kg、4.774%和 3.826%减小到 0.654 MJ/kg、2.338%和 1.927%；TrA-SVR 模型预测挥发分的准确性最高，与 SVR 模型对比，预测平均绝对误差由 2.722%减小到了 2.524%。【结论】研究结果表明，迁移学习结合合适的基础学习器能够提高煤质定量分析模型对不同仪器的适应性。

［关 键 词］激光诱导击穿光谱；迁移学习；煤质；定量分析；机器学习

［引用本文格式］邹祥波，饶睦敏，陈公达，等. 基于 TrAdaBoost 迁移学习的跨仪器 LIBS 煤质定量分析模型研究[J]. 热力发电，2026，55(5)：147-156. ZOU Xiangbo, RAO Mumin, CHEN Gongda, et al. Research on cross-instrument LIBS quantitative analysis model for coal properties based on TrAdaBoost transfer learning[J]. Thermal Power Generation, 2026, 55(5): 147-156.

Research on cross-instrument LIBS quantitative analysis model for coal properties based on TrAdaBoost transfer learning

ZOU Xiangbo[1,2], RAO Mumin[1], CHEN Gongda[1], TAN Shuwen[3], QIN Shiwei[1], KUANG Cao[1], YE Ji[1], YAO Shunchun[3], QIN Huaiqing[3]
(1. Guangdong Energy Group Science and Technology Research Institute Co., Ltd., Guangzhou 510630, China;
2. Guangdong Energy Group Co., Ltd., Guangzhou 510730, China;
3. School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: [Objective] Laser-induced breakdown spectroscopy (LIBS) holds significant potential for application in the field of coal property analysis, due to its advantages of eliminating complex sample pretreatment, enabling multi-parameter synchronous detection, and offering rapid analysis. However, discrepancies in spectral responses

收 稿 日 期：2025-09-03 修 回 日 期：2025-10-03 接 受 日 期：2025-10-15
基 金 项 目：国家重点研发计划项目（2021YFF0601001、2024YFC3909002）；广东省新型电力系统技术创新项目（1688950422168）；国家自然科学基金青年科学基金项目（22403032）；国家自然科学基金联合基金重点项目（U22B20119）
Supported by: National Key Research and Development Program (2021YFF0601001, 2024YFC3909002); Technological Innovation Project for New Power Systems in Guangdong Province (1688950422168); Youth Science Fund Project of National Natural Science Foundation of China (22403032); Joint Funds of the National Natural Science Foundation of China (U22B20119)
第一作者简介：邹祥波（1988），男，高级工程师，主要研究方向为碳减排与资源化利用技术，zouxiangbo@geg.com.cn。
通信作者简介：覃淮青（1991），男，助理研究员，主要研究方向为能源低碳清洁转化过程智能感知，202390441212@scut.edu.cn。

exist among different instruments. These discrepancies cause severe accuracy degradation when a quantitative model trained on spectra acquired by a master instrument is applied to slave instruments. Therefore, this study constructed cross-instrument LIBS quantitative analysis models of coal property by integrating TrAdaBoost transfer learning with various machine-learning algorithms. [Methods] Two LIBS-based coal analyzers were designated as the master and slave instruments respectively, and LIBS spectra were collected from different numbers of coal samples on both devices. Random forest (RF), support-vector regression (SVR), and their TrAdaBoost-enhanced counterparts (TrA-RF and TrA-SVR) were employed to build quantitative analysis models. Model performance was evaluated by predicting the coal properties of unknown coal samples on the slave instrument. [Results] The results indicated that both TrA-RF and TrA-SVR models significantly outperformed their non-transfer counterparts. The TrA-RF model achieved the highest accuracy for calorific value, ash content, and carbon content. Compared with RF model, the mean absolute errors decreased from 1.390 MJ/kg, 4.774 %, and 3.826 % to 0.654 MJ/kg, 2.338%, and 1.927%, respectively. TrA-SVR model yielded the highest accuracy for volatile matter prediction. Compared with the SVR model, the mean absolute error decreased from 2.722% (SVR) to 2.524%. [Conclusion] These findings demonstrate that coupling transfer learning with an appropriate base learner markedly enhances the adaptability of LIBS-based coal property models across different instruments.

Key words: laser-induced breakdown spectroscopy; transfer learning; coal property; quantitative analysis; machine learning

在“双碳”战略目标的推动下，构建以清洁能源为主导的新型电力系统已成为明确方向，这也要求现役煤电机组加快功能转型，承担起电力兜底保障与灵活调峰的双重角色[1-2]。煤质特性是实现煤电低碳灵活运行的重要决策依据，其直接影响配煤掺烧、制粉系统运行及锅炉燃烧效率，进而关系到机组的负荷调节能力与运行经济性。

常规煤质检测方法需经历取样、制样和化验等多个环节，耗时长达数小时，难以满足运行优化对数据实时性的要求。目前，主流的在线煤质快速检测系统主要基于瞬发γ中子活化分析（prompt gamma neutron activation analysis，PGNAA）[3]和 X 射线荧光光谱（X-ray fluorescence spectroscopy，XRF）[4]技术构建。虽然这些技术已相对成熟，但其普遍存在设备结构复杂、日常维护成本高昂，以及因使用放射源而带来的潜在安全风险等固有局限。激光诱导击穿光谱（laser-induced breakdown spectroscopy，LIBS）作为近年来新兴物质分析技术，因其无需复杂的样品预处理、设备系统简洁，且能够同步快速检测多种元素和指标的特点，在煤质快速检测领域展现出广阔的应用前景[5]。然而，LIBS 技术在实际工业应用中也面临挑战。强烈的基体效应以及激光与样品相互作用的不稳定性，使采集到的光谱信号与煤质关键指标（如灰分、挥发分、热值）之间的映射关系变得高度复杂和非线性。为建立可靠的预测模型，通常需要依赖大量已知化学成分的煤样及其对应的光谱数据来训练定量分析模型[6]。然而，由于测量环境波动、仪器自身状态差异以及系统噪声特性不同等因素，不同 LIBS 仪器即便对同一种煤样进行检测，其获得的光谱信号强度、背景、特征峰形等也存在显著差异。这种差异导致基于主仪器训练的定量分析模型直接用于另一台从仪器时，其预测精度出现大幅下降。为解决这一问题，传统做法是为每一台仪器单独采集大量有标签样本的光谱数据，并训练独立的定量模型。但这种做法需要投入较高的时间、人力和经济成本。在某些苛刻的工业现场条件下，可能根本无法为每台新仪器获取足够数量且标注精确的样本数据来重新训练一个高精度、高可靠性的模型[7]。

有必要探索一种方便快捷且对数据要求较低的方法，解决工业现场从仪器繁琐的“重建模”问题。迁移学习是机器学习的一个重要研究方向，其根本思路在于将一个领域（源域）中学习到的知识或模型结构，有效地应用于另一个具有相关性但数据分布或任务定义存在差异的新领域（目标域），从而显著增强目标模型的学习效能与预测能力。这种方法显著降低了目标领域对大规模标注数据的依赖，体现出良好的知识复用特性[8]。迁移学习算法已被引入 LIBS 检测技术中，旨在应对因样品物理状态差异、空间环境波动及仪器响应特性不一致等因素导致的分析精度下降问题。Shabbir 等人[9]提出了一种基于特征表示的迁移学习方法，通过将表面形态不同的 2 种样品特征结合成一组无显著差异共同特征，结合多元回归算法将预测的相对误差从单变量模型的 112%降低至 16.3%。Cui 等人[10]提出了一种基于迁移学习的多任务正则化卷积神经网络方法，在源模型训练中引入多个相关任务（如煤中的碳、灰分、挥发分和硫含量）作为正则化，约束

模型学习更通用的先验知识，解决了 LIBS 定量分析中的小样本问题。陈翊翔等[11]将迁移学习算法 TrAdaBoost 应用于 LIBS 定量分析中，在建模时将从煤粉压片样品上获取的高稳定性光谱数据中学习到的知识应用于煤粉颗粒流光谱数据中，有效校正因颗粒流状态不可控所导致的光谱信号随机波动，提高 LIBS 检测煤粉颗粒流工业分析指标的准确性。Yu 等人[12]将 TrAdaBoost 与极限学习机算法相结合，通过加权 ELM 子模型，迭代调整源域和目标域样本的权重，提升了定量分析模型对不同仪器的适用性。

综上所述，迁移学习算法能够学习不同仪器数据的共同特征，减少从仪器对大量数据的依赖，提高模型对不同仪器的适应性。本研究基于 TrAdaBoost 迁移学习的跨仪器 LIBS 煤质定量分析模型，探究了不同基础学习器选择对模型预测性能的影响，实现主仪器上训练的定量分析模型在从仪器上准确检测煤质，提高了煤质定量分析模型对不同仪器适应性。

1 实　验

1.1 实验装置

本文的煤炭 LIBS 光谱数据采集在自研的 LIBS 煤质快速检测一体机上进行[13]，仪器的构成如图 1 所示。本实验系统的核心组成部分主要包括振动给粉机、激光器、光谱仪以及一系列光学元件。为产生稳定且连续的下落煤粉颗粒流，将给粉机的振动频率设定为 252 Hz。波长为 1 064 nm 的调 Q Nd:YAG 脉冲激光器（Beamtech Optronics Dawa-200）以 10 Hz 的频率运行，其脉冲宽度为 5 ns，输出单脉冲能量为 65 mJ 的激光。该激光束垂直入射到下落的煤粉颗粒流上，通过烧蚀作用激发产生等离子体。采用与激光入射同轴收光方式将等离子体辐射耦合进光纤并传输至四通道光纤光谱仪（ULS2048-CL-4-EVO-RM，Avantes AvaSpec）进行分光和光电转换，光谱仪探测波长范围为 200~851 nm，分辨率为 0.08~0.28 nm，积分时间设置为 1.05 ms。光谱仪触发信号由激光器发出，激光发射与光谱仪采集信号之间的延迟时间设置为 2 μs，主从仪器在每个样品上均采集 200 次脉冲的光谱数据，后续将结合有效光谱筛选方法，剔除无效光谱后在每个煤样所有的脉冲光谱数据中保留有效激发的光谱数据。分别在 2 台仪器采集煤样的光谱数据，将其中一台仪器作为主仪器，采集满足模型训练需求的足够煤样光谱数据，所组成的数据集称为源域；另一台仪器为从仪器，只采集部分煤样光谱数据，无法满足模型训练需求，得到的数据集称为目标域。

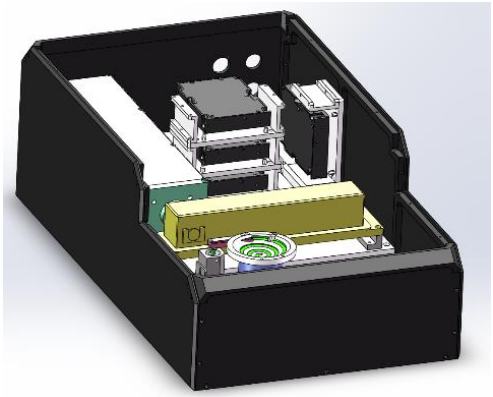


图 1 LIBS 煤质快速检测一体机示意
Fig.1 Schematic diagram of the LIBS coal property rapid detection integrated machine

1.2 实验样品

实验选取了从不同燃煤电厂收集的 95 个煤样，根据《煤样的制备方法》[14]进行制样，得到粒径小于 0.2 mm 的空气干燥基煤样。然后，依据国家标准《煤的工业分析方法》[15]、《煤的发热量测定方法》[16]和《煤的元素分析》[17]规定的实验方法，对煤样的发热量、灰分、挥发分和含碳量进行测定。所得结果作为本研究的参考基准值，具体见表 1。同时，为确保形成稳定的颗粒流并防止落粉管堵塞，将煤粉与石英砂以 1:4 的质量比进行均匀混合。其中，在主仪器上采集 70 个煤样的光谱（记为 X_s）用于训练定量分析模型，在从仪器上采集 25 个煤样光谱，其中 15 个煤样作为定量分析模型训练集（记为 X_a），剩余 10 个煤样作为检验定量分析模型的测试集（记为 X_t）。

表 1 煤样煤质指标分布情况
Tab.1 Distribution of coal property indexes of the coal samples

项目	最小值	最大值	平均值	SD 值
训练集（X_s+X_a）				
发热量/(MJ·kg^{-1})	15.79	34.37	25.18	3.61
灰分/%	0.80	44.66	20.95	9.84
挥发分/%	7.18	54.59	28.07	10.39
含碳量/%	40.46	84.89	62.98	8.92
测试集（X_t）				
发热量/(MJ·kg^{-1})	23.25	28.48	26.37	1.59
灰分/%	9.12	27.54	16.28	5.34
挥发分/%	12.37	46.82	32.06	9.08
含碳量/%	58.56	72.22	65.27	3.78

2 方法原理

2.1 光谱预处理

在利用 LIBS 检测煤粉颗粒流时，观测系统的视角是保持固定的。但颗粒流存在分散特性，这致使等离子体的激发位置随机波动。这种位置的不确定性易导致所采集的光谱中包含不能有效反映样品组分信息的信号，即无效光谱。为提升光谱数据的质量与可靠性，本文引入了一种基于特征谱线强度标准偏差（standard deviation，SD）的数据筛选方法[18]，用以识别并剔除这些无效光谱。其中，SD 值的计算公式为：

$$S_{\mathrm{D}}=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(I_i-\bar{I})^2} \tag{1}$$

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

在光谱数据筛选中，SD 值是判别光谱有效性的关键指标。若该值低于预设阈值，则判定其对应光谱为无效数据，予以剔除。SD 阈值的具体设定需综合考虑实际检测条件，主要取决于所用光谱仪的分辨率及系统固有的背景噪声水平等。本文选用特征峰 C I 247.86 nm 计算 SD 值，并设定阈值为 500，特征峰 SD 值小于 500 的光谱予以剔除。由于样品内部化学成分分布不均、入射激光能量存在起伏以及等离子体形态的动态演化等因素，采集到的光谱常出现基线漂移和信号强度无规律波动，这些干扰会显著降低后续定量分析模型的精度与可靠性。为抑制此类干扰，提升数据质量，本研究采用通道归一化方法[19]对经过有效性筛选后的光谱数据进行处理，随后将每个样品对应的多组光谱进行平均。

2.2 TrAdaBoost 迁移学习算法

TrAdaBoost 是一种基于 AdaBoost 框架改进的迁移学习算法。该算法利用少量有标注的目标域数据和大量来自源域的数据进行模型训练。通过动态调整样本权重的策略，增强对目标域回归任务有益样本的影响力，同时抑制分布不一致样本的负面影响。因此，该算法特别适用于目标域标记样本稀缺的应用场景。其算法流程主要包括权重初始化、迭代训练、误差计算与权重更新等[20]。

1）权重初始化　在模型训练的初始阶段，将源域训练数据（X_{s}）和目标域训练数据（X_{a}）合并，构成联合数据集 X_{sa}，随后，对该联合数据集中的所有样本赋予初始权重，形成初始权重向量值 W_i^1。

$$W_i^1=\begin{cases}\dfrac{1}{N},\ i=1,2,...,n\\[2mm]\dfrac{1}{m},\ i=n+1,...,n+m\end{cases} \tag{2}$$

式中：n 和 m 分别为源域 X_{s} 和目标域 X_{a} 的样本数量，且 $n>m$。

2）误差计算　在迭代中建立回归模型，并计算每个个体回归器的预测误差 ε_t。

$$\varepsilon_t=\sum_{i=n+1}^{n+m}\frac{w_1^t\left|f_t(x_i)-h(x_i)\right|}{\max\left|f_t(x_i)-h(x_i)\right|} \tag{3}$$

式中：$f_t(x_i)$为个体回归器的预测值；$h(x_i)$为样品对应指标的真实值；w_1^t为样品权重；$\max|f_t(x_i)-h(x_i)|$为每次迭代中预测值与真实值的最大误差；同时设定当误差率 $\varepsilon_t<10^{-5}$ 时停止迭代。

3）权重更新　在每一次迭代过程中，预测误差较低的样本权重会相应增加，从而使后续回归器更加关注这些难以精确拟合的数据点；反之，预测误差较大的样本权重则会降低，以减少其对当前模型训练过程的负面影响。调整后的样品权重 W_i^{i+1}，个体回归器权重 α_t 按下式计算：

$$\beta_t=\frac{\varepsilon_t}{1-\varepsilon_t} \tag{4}$$

$$W_i^{t+1}=\begin{cases}W_i^t\beta_t^{|f_t(x_i)-h(x_i)|},\ i=1,2,...,n\\W_i^t\beta_t^{-|f_t(x_i)-h(x_i)|},\ i=n+1,...,n+m\end{cases} \tag{5}$$

$$\alpha_t=\frac{1}{2}\ln\frac{1}{\beta_t} \tag{6}$$

式中：t 为总迭代次数，即所集成的个体回归器总数；β_t 为迭代参数。

4）加权集成　将训练得到的个体弱回归器 f_t 进行加权集成，构建出最终的强回归模型 $F(x)$。为评估不同基础回归器的效果，本文对比了随机森林（random forest，RF）和支持向量机（support vector regression，SVR）作为个体回归器时的预测情况，对应的迭代过程示意如图 2 所示。

2.3 模型评价指标

以训练集决定系数 R_t^2、预测均方根误差（root mean square error of prediction，RMSEP）δ_{RMSEP}、预测平均绝对误差（mean absolute error of prediction，MAEP）δ_{MAEP} 作为模型的评价指标。计算公式为：

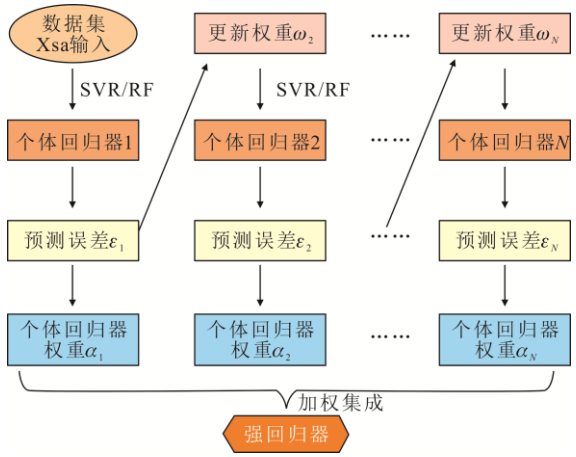


图 2　基于 TrAdaBoost 建模示意

Fig.2 Schematic diagram of modeling based on TrAdaBoost

$$R_t^2 = 1 - \frac{\sum_{j=1}^{n_1}(c_j - y_j)^2}{\sum_{j=1}^{n_1}(c_j - \bar{c})^2} \tag{7}$$

$$\delta_{\mathrm{RMSEP}} = \sqrt{\frac{\sum_{j=1}^{n_2}(c_j - y_j)^2}{n_2}} \tag{8}$$

$$\delta_{\mathrm{MAEP}} = \frac{\sum_{j=1}^{n_2}|c_j - y_j|}{n_2} \tag{9}$$

式中：c_j 为第 j 个样品的煤质指标参考值；y_j 为第 j 个样品的煤质指标预测值；$\bar{c}$ 为所有样品煤质指标参考值的平均值；n_1 为训练集样品数量；n_2 为测试集样品数量。

3 结果与讨论

3.1 不同仪器光谱对比

采用 SD 值有效光谱筛选方法剔除主从仪器采集的无效光谱，经过有效性筛选后其中一个煤样在主从仪器上采集的原始光谱如图 3 所示。可以看到主从仪器的光谱强度存在显著差异，这是造成主从仪器光谱数据分布差异的主要因素之一，也导致了主仪器上稳定的定量模型无法适用于从仪器光谱。此外，主仪器光谱相对于从仪器光谱在部分通道存在着基线抬升现象。为了减少光谱波动，采用通道归一化方法对主从仪器有效光谱进行预处理。预处理后的光谱如图 4 所示，虽然经过预处理后主仪器光谱相对从仪器光谱基线漂移得到改善，但是主从仪器光谱强度还是存在较大差异，导致基于主仪器光谱数据训练的定量分析模型难以适应从仪器。

3.2 不同基础学习器的定量分析结果

由于基础学习器的选择和性能对整体模型的迁移效果和预测精度有着重要的影响，本文分别以随机森林（random forest，RF）和支持向量机回归（support vector regression，SVR）为基础学习器结合 TrAdaBoost 构建煤质定量分析模型，同时直接采用 RF 和 SVR 构建煤质定量分析模型，对比不同模型预测的发热量、灰分、挥发分和含碳量。为分析 4 个煤质指标，RF 模型的决策树数量和叶节点最小样品数分别设置为 260 和 3、150 和 1、200 和 2、200 和 3；SVR 模型采用多项式核函数，利用粒子群寻优算法（particle swarm optimization，PSO）对多项式阶数 d 和核参数 g 进行自动寻优，PSO 算法进化代数设置为 200、种群数目为 30、d 寻优范围为[1,6]、g 寻优范围为[1×10^{-5},10]。TrAdaBoost 算法的迭代次数根据损失函数的收敛情况而定，设置为 10~20。图 5 为 RF 和 TrA-RF 模型预测煤质指标的结果，图 6 为 SVR 和 TrA-SVR 模型预测煤质指标的结果。

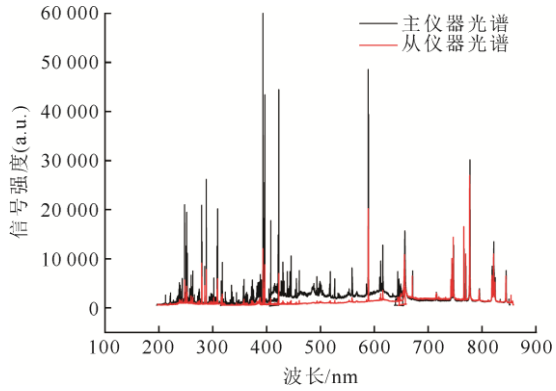


图 3　主从仪器原始光谱对比

Fig.3 Comparison of raw spectra between master and slave instruments

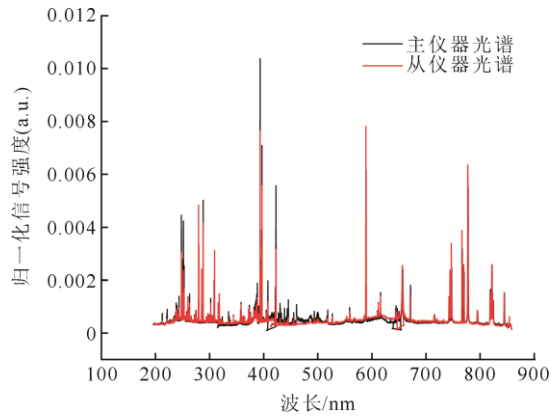


图 4　主从仪器预处理后光谱对比

Fig.4 Comparison of spectra after pretreatment between master and slave instruments

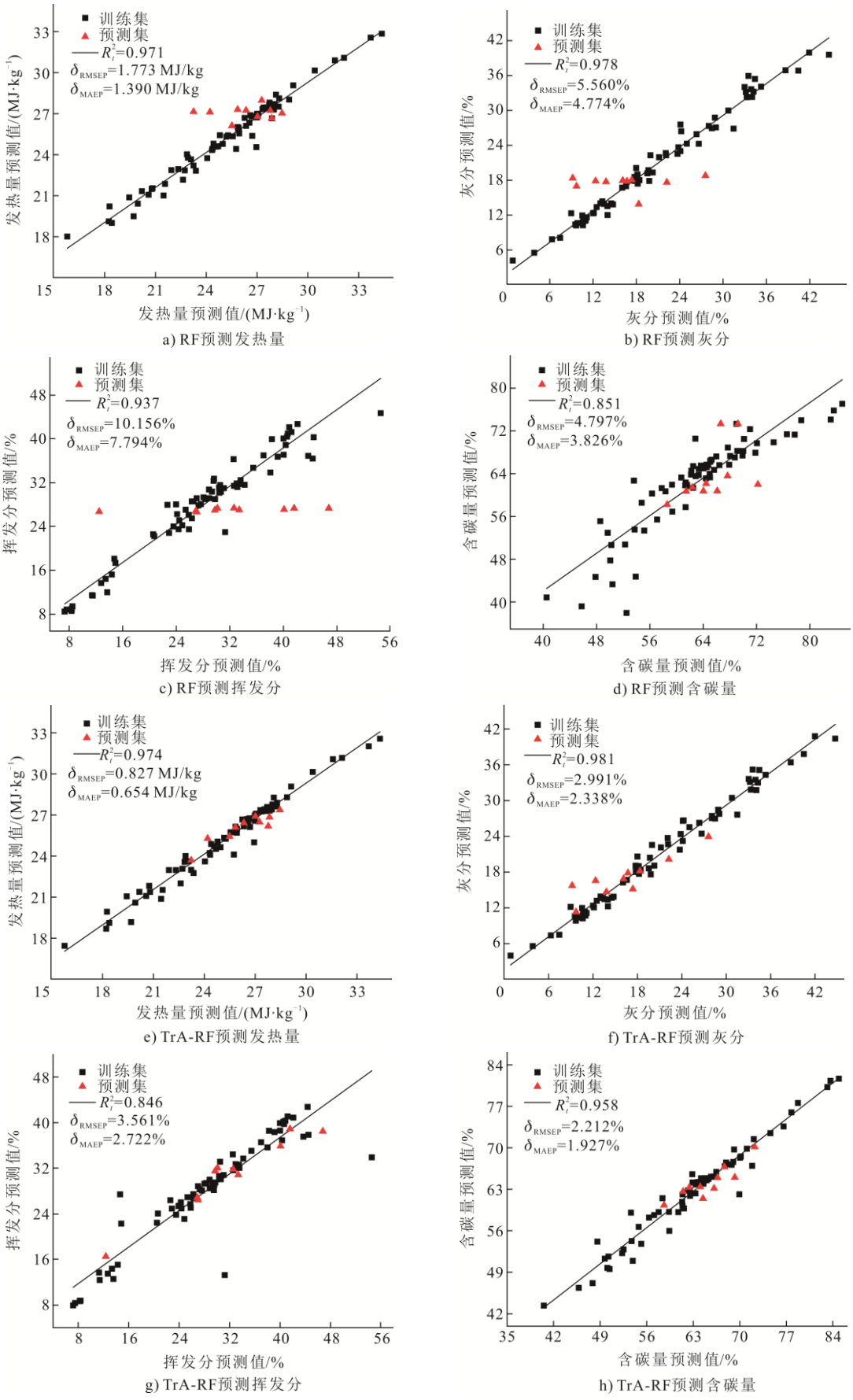


图 5 RF 模型和 TrA-RF 模型煤质分析结果对比

Fig.5 Coal property analysis results using RF model and TrA-RF model

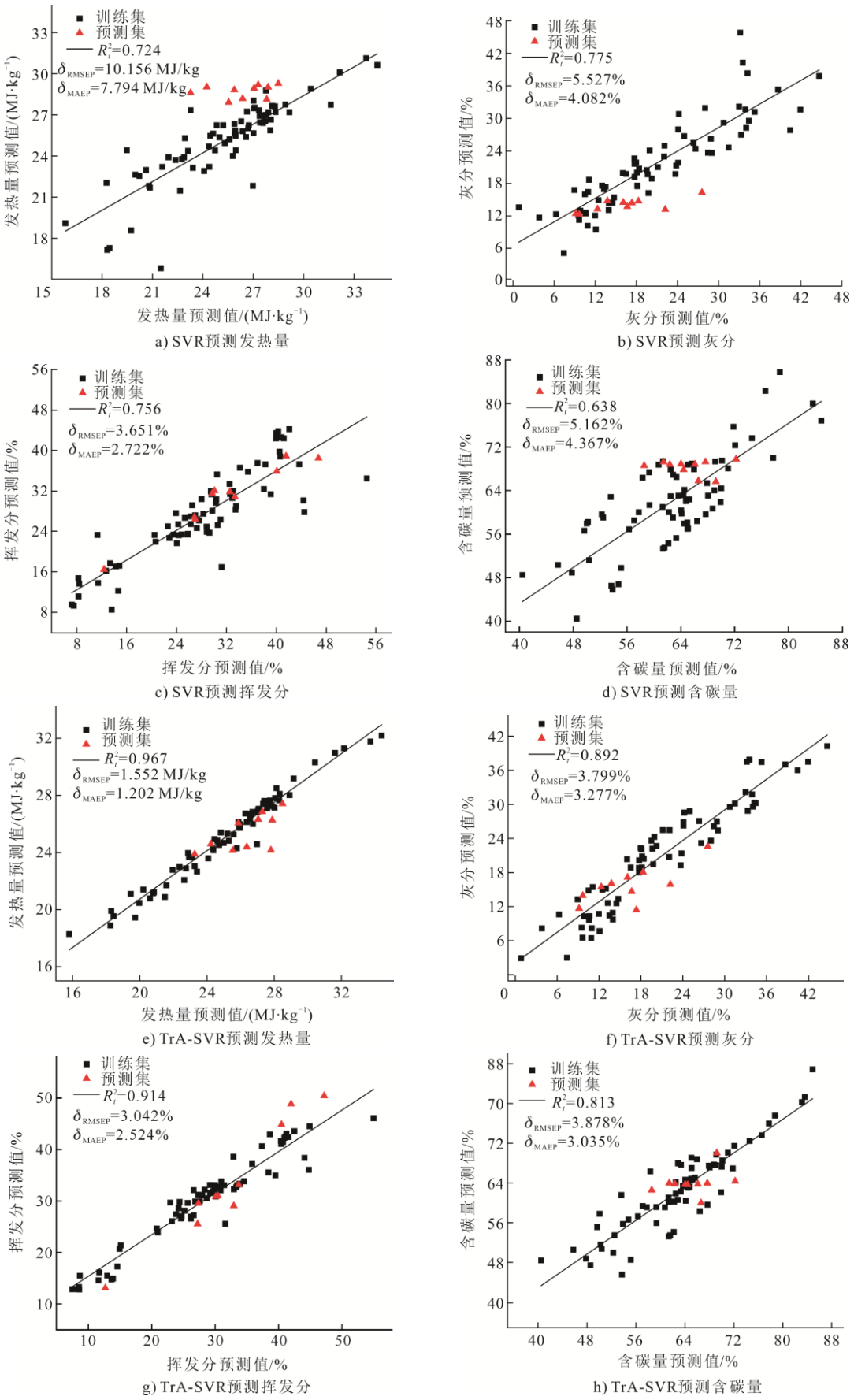


图 6 SVR 模型和 TrA-SVR 模型煤质分析结果对比

Fig.6 Coal property analysis results using SVR model and TrA-SVR model

从图 5 和图 6 可以看出，引入 TrAdaBoost 迁移学习算法提高了定量分析模型在从仪器的预测性能，但不同基础学习器对模型的性能有显著影响。虽然 RF 模型的 R_t^2 较高，但是 RMSEP 和 MAEP 都较大，表明模型过拟合，而 SVR 模型的 R_t^2、RMSEP 和 MAEP 都较差。这是由于主从仪器光谱数据存在显著差异，直接使用主仪器光谱数据训练的定量分析模型无法有效捕捉到从仪器光谱数据特征，从而导致模型预测性能大幅下降。通过采用 TrAdaBoost 迁移学习框架对主从仪器间的特征知识进行迁移与适配，最终构建的预测模型在各项性能指标上均展现出实质性的改进与优化。其中，TrA-RF 模型分析发热量、灰分和含碳量准确性最高。相对于 RF 模型，R_t^2 均得到显著提高，发热量的 RMSEP 和 MAEP 由 1.773 MJ/kg 和 1.390 MJ/kg 减小到 0.827 MJ/kg 和 0.654 MJ/kg，灰分的 RMSEP 和 MAEP 由 5.561%和 4.774%减小到 2.991%和 2.338%，含碳量的 RMSEP 和 MAEP 由 4.797%和 3.826%减小到 2.212%和 1.927%。然而，分析挥发分准确性最高的是 TrA-SVR 模型。相对于 SVR 模型，挥发分的 R_t^2 由 0.756 提升到 0.914，RMSEP 和 MAEP 由 3.561%和 2.722%减小到 3.042%和 2.524%。上述结果表明，TrAdaBoost 迁移学习算法能够有效提升煤质定量分析模型在从仪器上的预测性能，但基于不同基础学习器的模型分析不同煤质指标的表现存在差异，这与基础学习器的特性相关。RF 是一种集成学习方法，能够通过多棵决策树的投票机制捕捉复杂的非线性关系，所以适用于处理光谱数据与煤质指标之间复杂非线性关系。SVR 借助核函数方法将原始输入数据转换到更高维的特征空间，并在该空间内构造一个线性的决策函数以执行回归分析。该方法在处理变量间呈显著线性关联的问题时表现出色。然而，当数据间存在复杂的非线性相互作用时，其拟合能力会受到一定限制，难以充分刻画特征间的内在规律。因此，SVR 模型通常更适合表征光谱数据与煤质指标之间的线性关系。

3.3 权重可视化及相关性分析

在 TrAdaBoost 迁移学习框架下，通过动态调整样本的权重分布，促使模型在迭代过程中持续关注并学习从仪器的数据特征，从而有效实现知识从主仪器向从仪器的迁移。这一权重调整策略直接决定了跨域特征对齐的效果和最终模型的泛化性能。通过迭代调整主仪器和从仪器样品的权重，TrAdaBoost 能够逐步降低源域中分布不一致的样本对模型的影响，同时增强目标域样本的贡献，从而实现模型从主仪器到从仪器的有效迁移。为了进一步揭示其样品权重调整对模型的影响，以发热量为研究对象，对主仪器样品在构建定量分析模型中经过 TrA-RF 迭代后的权重进行可视化处理，并计算每个主仪器样品与从仪器测试集样品的皮尔逊系数平均值，用于表征主仪器样品与从仪器样品之间的相关程度。较高的皮尔逊系数表明该主仪器样品的光谱特征与从仪器样品的光谱特征具有较高的相似性，在迁移过程中能够为目标域提供更有价值的信息，结果如图 7 所示。

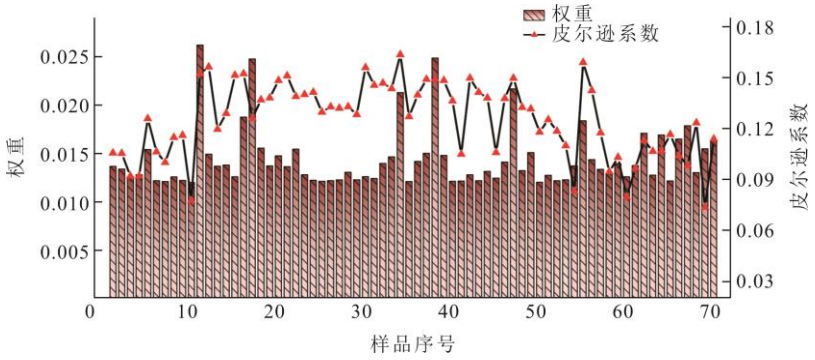


图 7 主仪器样品权重及其与从仪器测试集样品相关性

Fig.7 Weights of master instrument samples and their correlation with test set samples of the slave instrument

由上图可知，主仪器样品迭代后的权重与皮尔逊系数基本呈现正相关趋势。具体而言，那些与从仪器光谱表现出较高线性相关性的主仪器样品，在迁移学习算法中被分配了相对更大的权重；反之，线性相关性较弱的主仪器样品，其所获得的权重也相应较低。这是因为 TrAdaBoost 算法提取主从仪器

分布一致的知识，同时抑制分布不一致的部分，在主从仪器光谱数据迁移过程中，通过主仪器样品的权重调整识别那些能够更好地适应从仪器光谱特征的样本，从而增强模型的泛化能力。这种迁移机制不仅提高了模型在从仪器上的分析性能，还减少了对从仪器大量标注数据的依赖，为小样本条件下训练定量分析模型提供了有效的解决方案。

4 结　论

1）针对定量分析模型对不同 LIBS 仪器适应性差的问题，本文研究了基于 TrAdaBoost 迁移学习算法，构建跨仪器 LIBS 煤质定量分析模型，并研究了基础学习器对定量分析模型性能的影响。选择 RF 和 SVR 作为基础学习器，分别基于 RF、SVR 以及 RF 结合 TrAdaBoost（TrA-RF）和 SVR 结合 TrAdaBoost（TrA-SVR）构建定量分析模型，对比了不同模型在从仪器上预测煤质的表现。

2）引入 TrAdaBoost 迁移学习算法能够显著提高定量分析模型在从仪器光谱数据上的预测准确性。其中，TrA-RF 模型在发热量、灰分和含碳量的预测中表现最优异，MAEP 分别从 RF 模型的 1.390 MJ/kg、4.774%、3.826%下降至 0.654 MJ/kg、2.338%、1.927%。TrA-SVR 模型则在挥发分的预测中表现最优，MAEP 从 SVR 模型的 2.722%下降至 2.524%。

3）进一步研究发现，主从仪器样品光谱特征相关性高的样品被赋予了更高的迁移权重，从而强化模型学习主从仪器共同数据特征，提高了定量分析模型对仪器的适应性。本研究提出的方法能够减少模型训练过程对从仪器数据的依赖，提高了定量分析模型对不同 LIBS 仪器的适应性，有助于 LIBS 产业化发展。

4）模型的部署应用需考虑其计算效率与硬件需求。本研究中的 TrA-RF 与 TrA-SVR 模型在训练阶段因涉及多轮迭代与权重更新，需消耗较多的计算资源与时间，但其预测阶段效率较高，可满足在线快速分析的需求。TrA-RF 模型在集成多棵树时虽会占用较多内存，但支持并行计算，适于在配备多核 CPU 的工业工控机上部署。而 TrA-SVR 模型在核函数计算上需一定开销，但模型体积相对较小。未来工作将聚焦于模型压缩与量化技术研究，在保持精度的同时进一步降低计算复杂度与内存占用，使其更适配于嵌入式系统或边缘计算设备，推动 LIBS 煤质检测仪器的智能化与小型化发展。

5）本研究主要关注模型预测精度的提升，但仍存在几点局限性：首先，实验仅在 2 台同类型 LIBS 仪器间进行迁移验证，未来需在更多型号及新旧程度不同的仪器平台上进行广泛验证，以进一步检验所提方法的普适性与鲁棒性；其次，本研究仅采用了 RF 和 SVR 2 种基础学习器，未来可探索更多先进的机器学习模型（如轻量化神经网络、集成学习模型）作为 TrAdaBoost 的基础学习器，以期获得更优的迁移性能和泛化能力。

［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 on National Economic and Social Development of the People's Republic of China in 2023[EB/OL]. (2024-02-29)[2024-06-28]. http://www1.xinhuanet.com/politics/2024 0301/2c4a1c0966a647abae6a864aecc68ef9/c.html.

［2］ 刘云. 煤电“三改联动”技术综述及讨论[J]. 洁净煤技术, 2024, 30(增刊 1): 82-90.
LIU Yun. Overview and discussion of coal power “Three Reforms and Linkage” technology[J]. Clean Coal Technology, 2024, 30(Suppl.1): 82-90.

［3］ ZHAO L, XU X, LU J B, et al. Optimization of PGNAA device and algorithm for testing basicity index of sinter mixture[J]. Nuclear Science and Techniques, 2021, 32(6): 1-12.

［4］ 章连香, 符斌. X-射线荧光光谱分析技术的发展[J]. 中国无机分析化学, 2013, 3(3): 1-7.
ZHANG Lianxiang, FU Bin. Advances in X-ray fluorescence spectrometry[J]. Chinese Journal of Inorganic Analytical Chemistry, 2013, 3(3): 1-7.

［5］ GUO N, XU L, GAO W, et al. Progress in the application of laser-induced breakdown spectroscopy in coal quality analysis[J]. Energies, 2024, 17(14): 3559-3595.

［6］ 侯宗余, 宋惟然, 宋玉洲, 等. 激光诱导击穿光谱煤质在线分析技术现状与展望[J]. 煤质技术, 2023, 38(1): 1-12.
HOU Zongyu, SONG Weiran, SONG Yuzhou, et al. Current situation and prospect of online coal analysis by laser-induced breakdown spectroscopy[J]. Coal Quality Technology, 2023, 38(1): 1-12.

［7］ RAO Y, ZENG L W, WU M F, et al. Transfer learning based on dynamic time warping algorithms to improve qualitative analysis and quantitative prediction of rocks over multiple LIBS instruments[J]. Journal of Analytical Atomic Spectrometry, 2023, 38(3): 693-703.

［8］ 王惠. 迁移学习研究综述[J]. 电脑知识与技术, 2017, 13(32): 203-205.
WANG Hui. A review of transfer learning research[J]. Computer Knowledge and Technology, 2017, 13(32): 203-205.

［9］ SHABBIR S, ZHANG Y Q, SUN C, et al. Transfer learning improves the prediction performance of a LIBS model for metals with an irregular surface by effectively correcting the physical matrix effect[J]. Journal of Analytical Atomic Spectrometry, 2021, 36(7): 1441-1454.

[10] CUI J C, SONG W R, HOU Z Y, et al. A transferred multitask regularization convolutional neural network (TrMR-CNN) for laser-induced breakdown spectroscopy quantitative analysis[J]. The Royal Society of Chemistry, 2022, 37(10): 2059-2068.
[11] 陈翊翔, 董美蓉, 蔡俊斌, 等. 激光诱导击穿光谱技术结合 TrAdaBoost 算法的煤粉颗粒流定量分析[J]. 中国电机工程学报, 2023, 43(24): 9638-9646.
CHEN Yixiang, DONG Meirong, CAI Junbin, et al. Quantitative analysis of coal particle flow by laser induced breakdown spectroscopy based on TrAdaBoost algorithm[J]. Proceedings of the CSEE, 2023, 43(24): 9638-9646.
[12] YU Y, YAO M B, HUANG J P, et al. When process analysis technology meets transfer learning: a model transfer strategy between different spectrometers for quantitative analysis[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-19.
[13] LU W Y, YANG Q, YU Z Y, et al. Developing a compact coal quality rapid detection integrated machine based on laser-induced breakdown spectroscopy[J]. Microwave and Optical Technology Letters, 2024, 66(1): e33991.
[14] 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.
Method for preparation 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.
[15] 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.
Proximate 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.
[16] 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.
Determination of calorific value 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.
[17] 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.
Ultimate 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.
[18] 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.
[19] YAO S C, XU J L, BAI K J, et al. Improved measurement performance of inorganic elements in coal by laser-induced breakdown spectroscopy coupled with internal standardization[J]. Plasma Science and Technology, 2015, 17(11): 938-943.
[20] 刘翠玲, 徐金阳, 孙晓荣, 等. 基于 TrAdaBoost 算法的近红外光谱模型传递研究[J]. 农业机械学报, 2022, 53(2): 239-245.
LIU Cuiling, XU Jinyang, SUN Xiaorong, et al. Near infrared spectroscopy calibration transfer based on TrAdaBoost algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 239-245.

（责任编辑　邓玲惠）