

孔维恒, 曾令伟, 饶宇, 等. 基于预分类策略的激光诱导击穿光谱技术用于岩石样品定量分析[J]. 岩矿测试, 2023, 42(4): 760-770. doi: 10.15898/j.ykcs.202212190234.

KONG Weiheng, ZENG Lingwei, RAO Yu, et al. Laser-induced Breakdown Spectroscopy Based on Pre-classification Strategy for Quantitative Analysis of Rock Samples[J]. Rock and Mineral Analysis, 2023, 42(4): 760-770. doi: 10.15898/j.ykcs.202212190234.

基于预分类策略的激光诱导击穿光谱技术用于岩石样品定量分析

孔维恒¹, 曾令伟¹, 饶宇¹, 陈莎¹, 王旭¹, 杨燕婷², 段忆翔^{1*}, 樊庆文^{1*}

- (1. 四川大学机械工程学院, 四川 成都 610065;
2. 成都艾立本科技有限公司, 四川 成都 611930)

摘要: 岩石样品中复杂的基质效应严重影响激光诱导击穿光谱(LIBS)定量分析的准确性, 其原因是目标元素的发射特性会受到基质的影响, 导致其发射强度偏离理想的规律。为提高定量分析准确性, 本文提出一种基于岩性基质特性的预分类定量分析方法。该方法首先构建基于k近邻(kNN)与支持向量机(SVM)算法的多层分类模型识别样品的岩性进行分类, 通过kNN算法将样品分成碳酸盐和硅酸盐两大类, 再利用SVM算法将大类细分成6类, 而后针对不同岩性样品分别构建元素定量模型。通过采用预分类方法, 可以确保分析的样品具有相似的化学成分, 更好地确定分析时的基准线和校准曲线, 从而减少分析中的不确定度, 提高定量准确性。kNN算法通过交叉验证选取最优的k值, 同时使用网格寻优方法确定了SVM算法中关键惩罚参数C和RBF宽度参数 γ , 利用该分类模型对来自6类岩性的39个国标岩石样品和国标岩石混合样品中的Si、Ca、Mg和K元素进行分析, 岩性识别的准确率达100%, 保证了后续定量分析的准确性, 并针对不同岩性中的不同元素采用了合适的预处理方式提升光谱数据的稳定性。相比于传统标准曲线定量方法, 采用预分类方法可以减少不同岩性基质之间的相互影响, 从而减小样品基质非均匀性带来的误差。对比两种方法进行数据分析, 测试集样品的预测值与参考值相关性分析系数从0.231~0.664提高至0.994~0.999, 平均相对标准偏差从38.2%降低至8.6%。与传统定量分析方法相比较, 采用预分类定量分析方法所构建模型对上述4种元素定量分析结果准确性有着明显的提高, 为提高岩石元素定量分析准确性提供新的思路, 拓宽了LIBS技术的实际应用范围。

关键词: 激光诱导击穿光谱; 岩石检测; 定性定量分析; 分析模型; 预分类定量

要点:

- (1) 针对不同岩性中的不同元素采用合适的预处理方法可以提高光谱数据稳定性。
- (2) 通过预分类策略可以获得基质匹配的标准品以帮助构建与基质相关的标准曲线。
- (3) 将不同岩性的样品分类之后再定量分析可以有效地提升定量分析结果的准确性。

中图分类号: P597.3; O657.31

文献标识码: A

激光诱导击穿光谱(LIBS)是一种原子发射光谱技术, 其原理是以高能激光脉冲直接作用在样品上, 诱导样品产生等离子体, 利用光谱仪对等离子体发射光谱进行采集分析, 从而识别样品中元素组成和

收稿日期: 2022-12-19; 修回日期: 2023-03-31; 接受日期: 2023-05-16

基金项目: 四川省科技厅重点研发项目(2022YFG0235)

第一作者: 孔维恒, 硕士研究生, 主要研究方向为基于激光诱导击穿光谱的岩石定性定量分析方法。

E-mail: kongweiheng2021@163.com。

通信作者: 段忆翔, 博士, 教授, 主要研究方向为激光光谱、新型质谱分析方法及仪器的开发。E-mail: yduan@scu.edu.cn;

通信作者: 樊庆文, 博士, 研究员, 主要研究方向为机械设计、分析仪器设计、机电控制、图像处理技术及应用。

E-mail: fanqingwen@scu.edu.cn。

含量。LIBS具有系统结构简单、微损伤检测、多元素同时检测、可进行原位分析等诸多优点^[1-3],已经被应用于冶金分析^[4-5]、地球科学^[6-8]、生物科学^[9-11]等诸多领域。特别是近些年来,中国在新一轮找矿突破战略行动中提出重点聚焦1500米以浅矿产资源的勘查,LIBS技术因其原位分析和快速检测的优点,从而被广泛用于地质探测^[12-13],并且随着手持式LIBS仪器研发技术不断地成熟,LIBS检测也逐渐从实验室走向地质勘探现场,帮助实现地质找矿新突破。但是岩石样品中元素成分复杂多样^[14-15],在LIBS分析检测过程中由于基质效应和光谱波动会严重影响分析结果的准确性^[16-17],从而会对不同岩性的岩石元素定量分析形成一定的困难。迄今为止,LIBS研究人员已经进行大量的研究来解决这些问题,从2001年到2019年,每年发表的LIBS文章数量迅速增长,达到7336篇,LIBS的发展正向着更高灵敏度、更高分辨率、更低检出限不断进发。

国内外学者针对这些问题进行了诸多研究,其中包括LIBS定量算法^[18-19]、LIBS信号增强^[20-21]、时间分辨LIBS^[22]、LIBS图像辅助^[23-24]等领域,用于探究和减小基质效应带来的定量不准确的问题。Agresti等^[18]使用偏最小二乘回归(PLSR)和多层感知器(MLP)神经网络建立LIBS定量分析模型,将32个石灰石样品分成两组进行校准和验证,所获结果一致性证明了该模型在石灰石定量方面具有重大前景。Pagnotta等^[19]改进Franzini和Leoni算法(FL算法)用于纠正LIBS在地质材料成分测定中的基质效应,结果得出FL算法模型准确率高于单变量模型和PLS模型,且比人工神经网络的算法更简单和更可控。Hussain等^[20]利用磁约束和双脉冲-LIBS(DP-LIBS)两种增强技术对铝靶的LIBS进行对比研究,减小了基质效应对目标元素定量的影响,同时提高了LIBS灵敏度。Yang等^[21]构建了离焦共线双脉冲LIBS实验系统,并开展了预烧蚀和再加热等离子体光谱增强的相关实验,用于增强LIBS信号强度。Lu等^[22]研究了时间分辨率LIBS和卷积神经网络(CNN)的结合,以改善土壤中K元素的测定,通过提取和整合波长和时间维度的信息,提升了定量分析的准确性。Yan等^[23]提出一种图像定量分析(IQA)方法, R^2 由常规频谱定量分析(SQA)方法的0.9833提高到IQA方法的0.9996。Li等^[24]使用相应的图像强度对谱线强度进行标准化,用于水下LIBS定量分析,平均相对标准差(RSD)从10.71%

显著降低至5.76%,为解决由基质效应造成的定量检测困难提供了新的思路。上述学者所采用的方法,在一定程度上减小了基质效应的影响,提升定量分析的准确性,但忽视了不同岩性之间基质的差异,以及相同岩性之间基质的相关性,这样就会导致相同含量的元素在不同岩性的样品中具有不同的LIBS信号强度,从而影响分析结果的准确性。

本文提出了一种先预分类再定量的分析方法,根据不同类型岩石样品的基质,使用kNN和SVM双层分类模型对样品进行分类,先将所有样品数据使用kNN算法将其分为碳酸盐与硅酸盐两类,再对大类中的样品使用SVM算法将其细分为6类,获得基质相互匹配的样品^[25],然后对分类后的样品数据使用各自最合适的数据预处理方法进行预处理,建立校准曲线,对测试集的样品含量进行预测。该方法为地质探测等领域中岩石定量分析提供参考价值 and 借鉴意义。

1 实验部分

1.1 岩石压片样品制备

本实验使用由成都艾立本科技有限公司提供的来自6类岩性的共39个国标岩石和国标岩石混合粉末样品,编号为ALB1~ALB39,包括泥页岩(2个)、火成岩(4个)、白云岩(6个)、石灰岩(14个)、石膏岩(5个)和砂岩(8个),其中国标岩石混合样品采用相同岩性的两种国标岩石样品按照比例混合得到,这些岩石样品包含不同梯度的目标元素的含量,可以更好地对不同浓度下的元素进行定量分析研究。表1展示了不同岩石样品的编号、岩性以及4种目标元素的具体含量,其中元素含量由成都艾立本科技有限公司在第三方检测机构测得后提供。称取1.0g样品,使用压片机在8MPa压强下保压20s压制成直径20mm薄厚均匀的片状样品。

1.2 LIBS装置介绍及实验参数设置

本实验基于台式激光诱导击穿光谱仪(厂商:成都艾立本科技有限公司,型号:LT02A),该仪器采用1064nm的Nd:YAG激光器,重复频率1Hz,单次脉冲能量100mJ,仪器光谱探测范围185~960nm,光谱分辨率0.1nm。样品放置于三维自动样品台,激光束由聚焦透镜聚焦于样品表面诱导产生等离子体,等离子体的发射光由光纤探头接收传输至光谱仪,形成具有8192个光谱强度数值的光谱图。为减少连续背景对光谱质量的干扰,本实验将延时时间设置为1.5 μ s。

表1 39个国标岩石样品和国标混合岩石样品的种类及Si、Ca、Mg、K元素含量

Table 1 Types and elemental contents of Si, Ca, Mg and K for 39 national standard rock samples and national standard mixed rock samples.

企业编号	国标 GBW 编号	岩石类型	Si 元素含量(%)	Ca 元素含量(%)	Mg 元素含量(%)	K 元素含量(%)
ALB1	GBW03101a	泥页岩	23.32	0.09	0.28	0.66
ALB2*	GBWE070146	泥页岩	21.13	6.99	1.04	3.57
ALB3	GBW07109	火成岩	25.42	0.99	0.39	6.21
ALB4	GBW07110	火成岩	29.43	1.76	0.50	4.29
ALB5*	GBW07111	火成岩	27.85	3.37	1.69	2.90
ALB6*	GBW07121	火成岩	30.93	1.90	0.98	2.16
ALB7	GBW070157	白云岩	3.93	20.52	11.86	0.03
ALB8*	GBW070158	白云岩	0.87	21.59	12.51	0.01
ALB9	GBW070159	白云岩	1.01	21.54	12.55	0.02
ALB10*	GBW070160	白云岩	2.44	21.07	12.26	0.03
ALB11	GBW07114	白云岩	0.29	21.44	13.08	0.03
ALB12*	GBW07136	白云岩	3.85	23.62	10.80	0.01
ALB13	GBW03107a	石灰岩	1.89	35.78	1.07	0.35
ALB14*	GBW03108a	石灰岩	1.05	33.62	3.49	0.17
ALB15*	GBW07108	石灰岩	7.28	25.48	3.11	0.65
ALB16	GBW07120	石灰岩	3.10	36.50	0.43	0.12
ALB17	GBW07127	石灰岩	0.26	34.21	4.06	0.04
ALB18*	GBW07128	石灰岩	0.34	29.96	6.97	0.04
ALB19	GBW07130	石灰岩	0.50	38.63	0.85	0.04
ALB20	GBWE070149	石灰岩	1.41	36.15	1.73	0.14
ALB21*	GBWE070150	石灰岩	2.15	32.92	3.59	0.03
ALB22*	GBWE070151	石灰岩	1.55	36.59	1.46	0.08
ALB23	GBWE070152	石灰岩	0.62	38.42	0.70	0.05
ALB24*	GBWE070153	石灰岩	1.86	34.69	2.59	0.15
ALB25*	GBWE070154	石灰岩	1.10	37.11	0.74	0.14
ALB26	GBWE070155	石灰岩	0.45	38.71	0.49	0.07
ALB27*	GBW03109	石膏岩	0.78	28.03	1.04	0.08
ALB28*	GBW03109+GBW03111a (1+1)	石膏岩	0.54	25.55	1.26	0.05
ALB29*	GBW03109+GBW03111a (2+1)	石膏岩	0.62	26.38	1.19	0.06
ALB30	GBW03109+GBW03111a (3+1)	石膏岩	0.66	26.79	1.15	0.06
ALB31	GBW03111a	石膏岩	0.29	23.07	1.48	0.02
ALB32*	GBW03113	砂岩	44.68	0.12	0.06	0.56
ALB33	GBW03113+GBW03104 (4+1)	砂岩	42.24	0.13	0.13	1.07
ALB34	GBW03113+GBW03104 (9+1)	砂岩	43.46	0.13	0.09	0.81
ALB35	GBW03113+GBW03111a (4+1)	砂岩	35.80	4.71	0.34	0.45
ALB36	GBW03113+GBW03111a (9+1)	砂岩	40.24	2.42	0.20	0.50
ALB37	GBW03113+GBW070156 (4+1)	砂岩	35.82	7.78	0.23	0.45
ALB38	GBW03113+GBW07108 (4+1)	砂岩	37.20	5.19	0.67	0.57
ALB39*	GBW03114	砂岩	41.81	0.24	0.10	1.72

注：“*”代表测试集样品。国标混合样品编号后“()”中的数字代表混合比例。

1.3 光谱数据采集

针对每个岩石样品,随机地在样品表面 20 个位置采集光谱。为减少激光能量波动对光谱的影响,每个位置的光谱由 4 次激光产生的光谱平均而成,总共采集 780 张光谱。为验证模型的泛化能力,从每类岩石样品中选取部分样品数据作为测试集不参与模型的训练,共选择 18 个样品的光谱数据作为测

试集,其余光谱数据用于模型的训练。

1.4 光谱数据处理流程

岩石中复杂的基质效应会影响 LIBS 定量分析的准确性和稳定性^[26],不同岩性的样品基质不同,类似基质样品的光谱具有一定的共性。为减小基质效应的影响,提升定量分析的准确性,本研究采取先对不同基质的岩性样品进行分类,再根据不同的类

别分别对各元素进行定量分析的策略。整个数据处理的流程如图1所示。数据处理的大致流程如下:获取21个样品的LIBS光谱数据作为训练集,对这些数据进行相应的预处理(预处理方式有MinMax归一化^[27]、全谱归一化^[28]、分通道归一化^[29]),根据岩石含量的信息以及它们的光谱图呈现出的规律,使用两层分类模型,对6类样品进行分类,然后基于每个类别分类之后的数据建立单变量定标模型。完成建立预分类定量分析模型之后,将18个样品的测试集LIBS光谱数据经过对应的预处理方法,输入到分类模型中得出分类结果,再将分类之后的数据输入到对应分类类别结果的单变量定标模型中预测样品的含量。

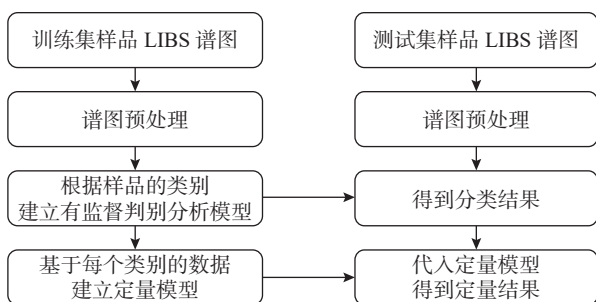


图1 光谱数据处理流程图

Fig. 1 Flow chart of spectral data processing.

1.5 多层分类模型的构建

根据1.4节中的数据划分,本研究基于训练集数据构建多层的定量分析模型,再使用测试集数据来评估模型的预测能力。整个模型的构建思路遵循先分类后定量的策略。

本研究使用k近邻算法和支持向量机算法来构建分类模型。其中k近邻(k-Nearest Neighbors, kNN)对测试集数据进行预测时,会在训练集数据中找到与该数据最相似的k个邻居,将k个近邻中占多数的类别认定为该数据的类别^[30]。支持向量机(Support-vector machines, SVM)会构造一个最优的超平面,使得线性分类形成的几何间隔最大化,将原始数据映射至内核函数构造的高维特征空间,在高维特征空间上构造线性分类^[31]。

根据样品中元素含量的差异,设计了一个双层的分类模型将LIBS光谱按照岩性类别分成6类:第一层采用kNN模型将所有光谱分为A、B两类,其中A类包括钙含量较高的白云岩、石灰岩和石膏岩, B类由硅含量较高的火成岩、泥页岩和砂岩所构成;

第二层使用径向基函数SVM模型将A类细分为白云岩、石灰岩和石膏岩,将B类细分为火成岩、泥页岩和砂岩。

本研究中,kNN模型的关键参数k值设定为5;SVM模型的关键参数惩罚参数(C)及径向基核函数(RBF)的宽度参数(γ)分别设定为1和0.1。这两个模型在测试集数据上的分类准确率均为100%,整个分类模型的准确率达到100%,表明该分类模型能够有效地识别不同岩性。

2 结果

2.1 不同岩性中不同元素的数据预处理结果对比

针对分类后的每个类别,使用标准曲线法分别建立各元素的定量模型。考虑到LIBS光谱中元素特征峰饱和与特征峰附近具有干扰峰等情况^[32-33],分别选择Si(I)250.69nm、Ca(I)585.745nm、Mg(I)516.732nm和K(I)766.489nm作为定量分析线。

在进行定量分析之前,为减小光谱之间噪声和干扰带来的差异,根据不同种类的岩石分别采用不同的预处理方法,包括全谱归一化、分通道归一化和MinMax归一化方法,分别见公式(1)、(2)和(3)。

$$I_{fn} = \frac{I_i}{\sum I_i} \quad (1)$$

式中: I_{fn} 代表全谱归一化后的峰强; I_i 代表第*i*个波长的光谱强度; $\sum I_i$ 表示所有波长的光谱强度之和。

$$I_{cn} = \frac{I_{ij}}{\sum I_{ij}} \quad (2)$$

式中: I_{cn} 代表分通道归一化后的峰强; I_{ij} 代表第*j*通道中第*i*个波长的光谱强度; $\sum I_{ij}$ 表示第*j*通道中所有波长的光谱强度之和。

$$I_{mn} = \frac{I_i - I_{\min}}{I_{\max} - I_{\min}} \quad (3)$$

式中: I_{mn} 代表MinMax归一化后的峰强; I_{\max} 表示整个光谱中最大光谱强度值; I_{\min} 表示整个光谱中最小光谱强度值; I_i 代表第*i*个波长的光谱强度。

由于基质效应的影响,单一的预处理方式并不适合所有待测元素的定量分析,因此为提高定量分析的准确度和稳定性,采用不同的方法对数据进行预处理。以所构建的单变量线性回归模型的决定系数(Coefficient of Determination, R^2)为评价指标^[34],选取了最佳预处理方法,结果如表2所示。

从表2中可以得出,白云岩和火成岩中的Si元素采用分通道归一化的数据预处理方式可以获得较高的 R^2 ,而对于石膏岩来说,无需对Si元素进行数

据预处理便可以获得较好的相关性。同样对于不同岩性中同一元素采用相同的数据预处理方式, 所求得 R^2 也大不相同, 泥页岩中的 Ca 元素采用全谱归一化的处理方法, R^2 可以达到 0.962, 远远大于砂岩中的 0.750。综合得出, 针对不同的预处理方式, Si、Ca、Mg、K 四个元素在 6 类岩石样品中的 R^2 值大部分大于 0.90, 有令人满意的结果。

2.2 未进行预分类和进行预分类两种方法定量结果对比

对相同的训练集和测试集数据采用该预分类定量模型和不分类直接进行定量分析的标准曲线法进行分析。其中对 21 个训练集样品中 Si、Ca、Mg 和 K 元素分别采用 MinMax 归一化、分通道归一化、全谱归一化和无数据预处理的预处理方式, 其传统标准曲线分析法 R^2 为 0.765、0.541、0.539、0.600, 预分

类定量方法 R^2 为 0.998、0.992、0.999、0.998, 对测试集预测效果有着较好的提升。使用两种策略进行定量分析的预测值与真值的相关分析结果如图 2 所示。

3 讨论

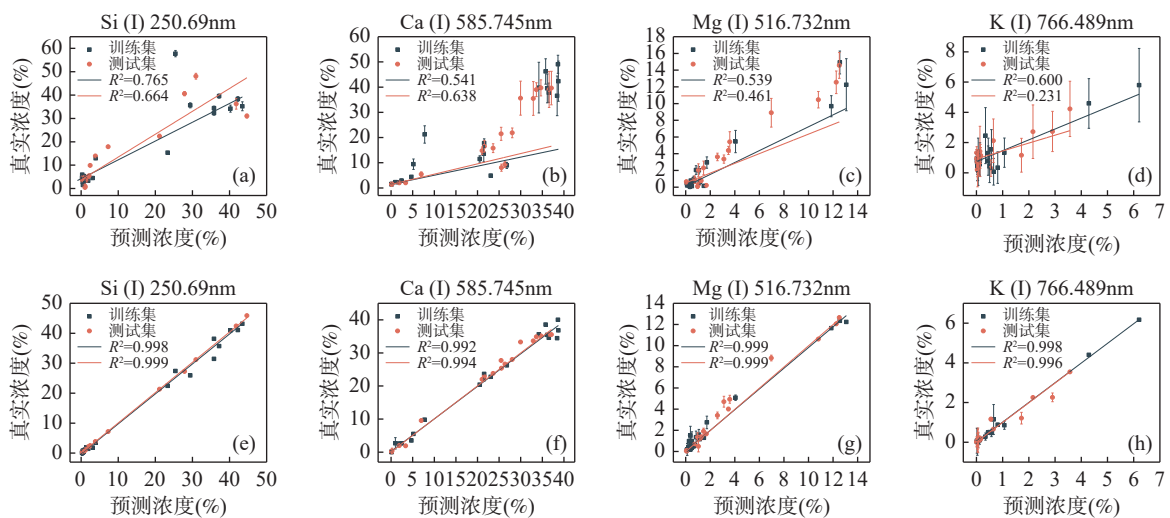
3.1 kNN 和 SVM 分类模型参数优化

关键参数的优化对于 kNN 和 SVM 双层分类模型的性能影响很大。对于 kNN 算法, 关键参数为 k 值, 代表的是算法所参考的近邻标签值的个数, k 取值较小时, 模型复杂度较高, 训练误差会减小, 泛化能力减弱; k 取值较大时, 模型复杂度低, 训练误差会增大, 泛化能力有一定的提高。在本文中采用交叉验证的方式选取最优的 k 值($k=5$)^[35], 用于岩石样品的预分类。对于 SVM 算法, 本文选择 RBF 函数作为核函数, 因为其具有良好的泛化能力, 可以避免

表 2 六种不同岩性岩石中不同元素的预处理方式结果对比

Table 2 Comparison of results of pretreatment methods for different elements in six different lithologies of rocks.

岩石岩性	Si (I) 250.69nm		Ca (I) 585.745nm		Mg (I) 516.732nm		K (I) 766.489nm	
	数据预处理方式	R^2	数据预处理方式	R^2	数据预处理方式	R^2	数据预处理方式	R^2
白云岩	分通道归一化	0.998	无数据预处理	0.929	全谱归一化	0.999	全谱归一化	0.970
火成岩	分通道归一化	0.985	无数据预处理	0.658	无数据预处理	0.813	无数据预处理	0.860
泥页岩	分通道归一化	0.778	全谱归一化	0.962	分通道归一化	0.998	MinMax 归一化	0.778
石膏岩	无数据预处理	0.981	MinMax 归一化	0.780	分通道归一化	0.998	无数据预处理	0.981
石灰岩	全谱归一化	0.977	无数据预处理	0.929	全谱归一化	0.984	全谱归一化	0.985
砂岩	全谱归一化	0.977	全谱归一化	0.750	全谱归一化	0.925	全谱归一化	0.934



a ~ d 为不分类的定量相关性曲线; e ~ h 为预分类定量相关性曲线。

图 2 两种方法在训练集和测试集上的定量分析结果对比

Fig. 2 The quantitative analysis results of the two methods are compared on the training set and test set. (a-d) is the correlation curve without classification, and (e-h) is the pre-classification quantitative correlation curve.

过拟合问题,提高模型的鲁棒性,同时与其他核函数相比,RBF核函数计算速度相对较快,可以提高模型分类的效率^[36],并且通过网格寻优方法确定了关键惩罚参数C和RBF宽度参数 γ 的值分别为1和0.1。

3.2 四种目标元素预处理结果讨论

采用不同预处理方法对4种目标元素进行预处理分析,可以提高光谱稳定性,增加定量分析准确性。但是由于火成岩中Na元素、砂岩中Si元素、石膏岩中Ca元素含量很高,在原始数据采集过程中,已经达到饱和^[37],所以在使用归一化方法对元素进行处理时,始终是以含量最高的元素作为最大值进行归一化,这样就会导致目标元素的光谱强度相对减弱,从而使 R^2 降低,表2中砂岩Ca元素的 R^2 较低,便是因为这个原因。本文对6类岩性中的4种元素均已采用合适的预处理方法。

3.3 四种目标元素多层模型定量结果讨论

从图2可以看出,无论是训练集还是测试集,以预分类定量策略所构建的多层模型对4种元素分析结果的相关性都有显著提高,均高于0.99。Si、Ca、Mg和K元素在测试集上定量分析结果的相关性从0.664、0.638、0.461和0.231分别提高到0.999、0.994、0.999和0.996。此外,从数据的分析处理中可以知道,传统定量分析模型的稳定性也比较差, Si、Ca、Mg和K等4种元素在测试集上的平均相对标准偏差(RSD)分别为3.4%、10.7%、48.2%、90.8%,而通过多层模型得到的4种元素RSD分别为1.5%、5.2%、10.3%、17.4%,其稳定性与传统定量分析模型相比有了很大提高。这是因为多层模型中的预分类定量策略,可以减小由基质效应所造成的定量结果的偏差。同时,相对单一的预处理方式,对不同岩性中不同元素采用不同的预处理方式能够有效地减小具有类似基质的光谱波动,从而降低定量分析的相对标准偏差^[38]。本研究还采用测试集中真实值和预测值偏差范围(公式(4))作为评价定量结果的指标。

$$\text{偏差范围} = \frac{\bar{x} - \mu_0}{\mu_0} \times 100\% \quad (4)$$

式中: \bar{x} 为元素预测含量, μ_0 为元素真实含量。

采用箱线图来表示未进行预分类和进行预分类的4种元素在测试集上偏差范围^[25],如图3所示。

如图3中的中心实线代表每个数据范围的中位数,方框显示数据范围中心的50%。该图展示了采用预分类策略比未采用预分类策略,4种元素的偏差范围都有明显地降低,尤其是对Si和K元素的定量效果,有着显著提升。基于这种结果,在进行定量分

析之前,根据岩石样品中基质的相似性进行预分类,可以减小不同基质之间的相互影响,增加定量准确性。

同时通过对测试集上各元素预测值与目标值之间的平均相对误差来更加直观地评估两种方法的差异,结果如表3所示。从表3可以看出,多层模型的对Si元素的预测效果最好,平均相对误差只有4.65%。虽然其他三个元素的平均相对误差超过10%,但是相比于传统的标准曲线模型,预测的准确度显著提升。结果表明,采用多层模型的方法,分别针对基质相似的同一样品构建定量分析模型,并且针对不同岩性中不同元素采用不同的预处理方式,能够减小基质效应造成的光谱差异,提升定量分析的准确度和稳定性。

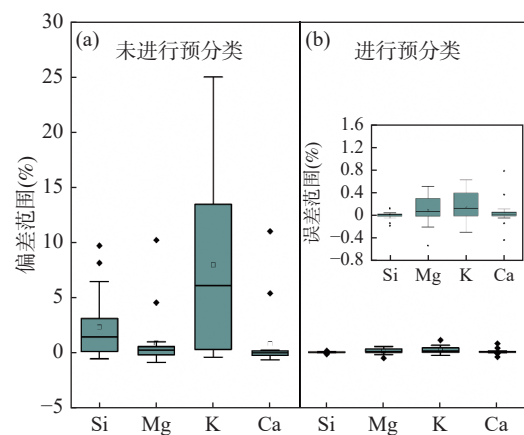


图3 (a)未进行预分类和(b)进行预分类两种方法中4种元素的偏差范围的箱线图对比

Fig. 3 Comparison of box line plots of deviation ranges of four elements in two methods: (a) Box line diagram of deviation ranges for Si, Mg, K and Ca without pre-classification method; (b) Box line diagram of deviation ranges of Si, Mg, K and Ca using pre-classification method.

表3 传统标准曲线模型和多层模型两种方法在测试集上的定量分析结果的平均相对误差对比

Table 3 Comparison of the mean value of relative errors of the quantitative analysis results on the test set between the two methods of traditional standard curve model and multi-layer model.

元素	传统标准曲线模型 (%)	多层模型 (%)
Si	367.41	4.65
Ca	212.66	12.40
Mg	66.60	19.90
K	953.85	26.95

4 结论

传统的标准曲线模型数据预处理方式单一,且定量的精度和稳定性较差,容易受到光谱波动和岩石中基质效应的影响。本文提出了一种先分类再定量的多层模型,采用kNN和SVM算法构建双层分类模型,采用交叉验证和网格寻优方法对kNN和SVM中关键参数进行优化。通过预分类模型对岩石样品进行预分类,获得基质相互匹配的岩石样品,对于不同种类的样品采用不同的预处理方式以减少光谱数据波动,再针对每种元素建立标准曲线模型,对6类

岩石样品中Si、Ca、Mg和K元素进行定量分析,得到结果与传统标准曲线模型进行对比,定量的准确度有了明显的提高。研究表明该模型不仅可以减小基质效应对定量分析产生的影响,还可以校正由于硬件、环境、样品等因素导致定量分析的不稳定性。

但是,目前的多层定量分析模型对于不同元素的预测效果还存在一定偏差,今后可以考虑采用结合多种不同算法建立预分类模型,达到更好的定量分析效果,为实际的野外岩矿分析作业提供有价值的帮助。

Laser-induced Breakdown Spectroscopy Based on Pre-classification Strategy for Quantitative Analysis of Rock Samples

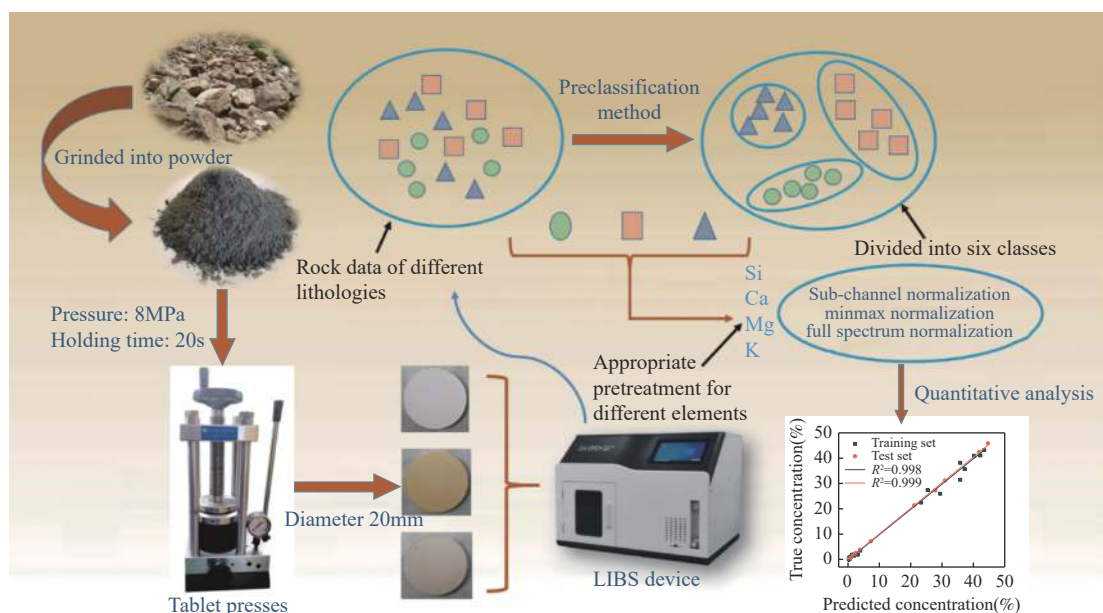
KONG Weiheng¹, ZENG Lingwei¹, RAO Yu¹, CHEN Sha¹, WANG Xu¹, YANG Yanting², DUAN Yixiang^{1*}, FAN Qingwen^{1*}

(1. School of Mechanical Engineering, Sichuan University, Chengdu 610065, China;

2. Chengdu Aliben Science & Technology Co., LTD., Chengdu 611930, China)

HIGHLIGHTS

- (1) Appropriate pre-processing methods for different elements in different lithologies can improve the stability of the spectral data.
- (2) Through pre-classification strategy, it is possible to obtain standard samples that are matched with the matrix to help construct a standard curve related to the matrix.
- (3) Sorting different rock types of samples before quantitative analysis can effectively improve the accuracy of the quantitative analysis results.



ABSTRACT

BACKGROUND: LIBS technology is a non-destructive, high sensitivity, high resolution spectroscopy technology that can be used to analyze the composition and structure of chemical substances and materials. It has extensive application in fields such as chemistry, materials science, life science, and geological exploration, and its emergence has provided new methods and technologies for the development of these fields. LIBS technology can be used to non-destructively analyze the chemical composition of underground rocks and minerals, helping geologists to better understand the composition and properties of underground resources, thus providing better guidance for geological exploration and development. In recent years, scholars at home and abroad have been exploring LIBS technology constantly, and through improving the detection system and optimizing laser pulse parameters, high sensitivity LIBS analysis at extremely low concentration has been achieved. By using finer spectral lines, higher sampling rate, and more precise laser pulse control, high resolution LIBS analysis at nanoscale has been achieved. The combination of LIBS technology with multi-spectral image processing technology can integrate information from multiple spectral channels to achieve a more comprehensive analysis of samples. However, the existence of matrix effects and spectral fluctuations always affects the accuracy of LIBS quantitative analysis, and poor reproducibility and high detection limits also need to be solved.

OBJECTIVES: To improve the accuracy of quantitative analysis of complex matrix samples.

METHODS: A multi-layer classification model based on k-nearest neighbors (kNN) and support vector machine (SVM) algorithms was constructed to identify the rock type of samples. The samples were divided into two major categories of felsic rocks and mafic rocks using the kNN algorithm, and then six categories were formed by the SVM algorithm. Different element quantitative models were constructed for each rock type. The kNN algorithm was selected using cross-validation to determine the optimal k value, and the key punishment parameter C and RBF width parameter γ of the SVM algorithm were determined using a grid search method. Then, appropriate pre-processing methods were adopted to improve the stability of spectral data for different elements in different rock types. Compared to the traditional standard curve quantitative method, using the pre-classification method can reduce the influence of different rock matrices on each other, thus reducing errors caused by the non-uniform matrix of samples.

RESULTS: Due to the influence of matrix effects, a single pre-processing method is not suitable for all elements in quantitative analysis. Therefore, in order to improve the accuracy and stability of quantitative analysis, different methods are used to pre-process the data. For different pre-processing methods, the R^2 values of four elements in six types of rock samples are mostly greater than 0.90, as shown in Table 2. After pre-processing, the correlation coefficients of the four elements are significantly improved, and they are all higher than 0.99. The correlation coefficients of Si, Ca, Mg, and K elements in the test set after quantitative analysis are increased from 0.664, 0.638, 0.461, and 0.231 to 0.999, 0.994, 0.999, and 0.996, respectively. In addition, it can be seen from the analysis of the data that the traditional quantitative analysis model has poor stability. The average relative standard deviation (RSD) of Si, Ca, Mg, and K elements in the test set are 3.4%, 10.7%, 48.2%, and 90.8%, respectively, while the RSD of four elements in the multi-layer model are 1.5%, 5.2%, 10.3%, and 17.4%, which shows a significant improvement in stability compared to traditional quantitative analysis models. At the same time, it can be more intuitively evaluated by comparing the average relative error between the predicted value and the target value of each element in the test set. As shown in Table 3, the prediction performance of Si element in the multi-layer model is the best, with an average relative error of only 4.65%. Although the average relative error of the other three elements is over

10%, it is significantly improved compared to the traditional standard curve model.

CONCLUSIONS: By utilizing a multi-layer classification model for preliminary categorization, standard rock samples that match the matrix are obtained. Subsequently, quantitative analysis models are developed for samples with similar matrices. Employing distinct preprocessing methods for different elemental compositions within various rock types helps mitigate spectral discrepancies caused by matrix effects, reduces spectral fluctuations and data noise, and enhances the accuracy and stability of quantitative analysis. Standard curve models are then established for each element, enabling quantitative analysis of Si, Ca, Mg, and K elements in six categories of rock samples. Results demonstrate a notable improvement in the accuracy of quantitative analysis compared to traditional standard curve models. This model not only diminishes the impact of matrix effects on quantitative analysis but also corrects instabilities arising from hardware, environmental conditions, and sample variations. Furthermore, it alleviates the workload of data analysis, simplifying the analytical process and thereby boosting efficiency. However, the current multi-layer quantitative analysis model still exhibits some deviations in regard to different elements. In the future, a potential avenue is to consider integrating various algorithms to establish preliminary classification models, aiming for even better quantitative analysis outcomes.

KEY WORDS: laser-induced breakdown spectroscopy (LIBS); rock detection; qualitative and quantitative analysis; analysis model; pre-categorized quantitation

参考文献

- [1] Guo L B, Zhang D, Sun L X, et al. Development in the application of laser-induced breakdown spectroscopy in recent years: A review[J]. *Frontiers of Physics*, 2021, 16(2): 1–25.
- [2] Harmon R S, Senesi G Sl. Laser-induced breakdown spectroscopy—A geochemical tool for the 21st century[J]. *Applied Geochemistry*, 2021, 128: 104929.
- [3] 杨春梅, 黄梓芸, 覃静雯, 等. 应用钻石观测仪-红外光谱仪-激光诱导击穿光谱仪鉴定无机材料充填翡翠[J]. *岩矿测试*, 2022, 41(2): 281–290.
Yang C M, Huang Z Y, Qin J W, et al. Identification of jadeite filled with inorganic materials using UV fluorescence, infrared spectroscopy and LIBS techniques[J]. *Rock and Mineral Analysis*, 2022, 41(2): 281–290.
- [4] Zhao S X, Zhang L, Hou J J, et al. Accurate quantitative CF-LIBS analysis of both major and minor elements in alloys *via* iterative correction of plasma temperature and spectral intensity[J]. *Plasma Science & Technology*, 2018, 20(3): 035502.
- [5] Che C, Lin X, Gao X, et al. Accuracy improvement for multimetal analysis by laser-induced breakdown spectroscopy with least squares support vector machine[J]. *Microwave and Optical Technology Letters*, 2021, 63(6): 1635–1641.
- [6] Ie A, Kra B, Jia C, et al. Measuring the concentration of gold in ore samples by laser-induced breakdown spectroscopy (LIBS) and comparison with the gravimetry/atomic absorption techniques[J]. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2021, 183: 106256.
- [7] Devangad P, Unnikrishnan V K, Nayak R, et al. Performance evaluation of laser induced breakdown spectroscopy (LIBS) for quantitative analysis of rare earth elements in phosphate glasses[J]. *Optical Materials*, 2016, 52: 32–37.
- [8] Alvarez-Llamas C, Purohit P, Moros J, et al. LIBS-acoustic mid-level fusion scheme for mineral differentiation under terrestrial and martian atmospheric conditions[J]. *Analytical Chemistry*, 2022, 94(3): 1840–1849.
- [9] Khan M N, Wang Q Q, Idrees B S, et al. A review on laser-induced breakdown spectroscopy in different cancers diagnosis and classification[J]. *Frontiers in Physics*, 2022, 10: 821057.
- [10] Wan X, Wang P. Analysis of heavy metals in organisms based on an optimized quantitative LIBS[J]. *Optik*, 2015, 126(19): 1930–1934.
- [11] Wang Q Q, Xiangli W T, Teng G, et al. A brief review of laser-induced breakdown spectroscopy for human and animal soft tissues: Pathological diagnosis and physiological detection[J]. *Applied Spectroscopy Reviews*, 2021, 56(3): 221–241.

- [12] Chen S, Pei H L, Pisonero J, et al. Simultaneous determination of lithology and major elements in rocks using laser-induced breakdown spectroscopy (LIBS) coupled with a deep convolutional neural network[J]. *Journal of Analytical Atomic Spectrometry*, 2022, 37(3): 508–516.
- [13] 秦燕,徐衍明,侯可军,等.铁同位素分析测试技术研究进展[J].*岩矿测试*,2020,39(2):151–161.
Qin Y, Xu Y M, Hou K J, et al. Progress of analytical techniques for stable iron isotopes[J]. *Rock and Mineral Analysis*, 2020, 39(2): 151–161.
- [14] 龚仓,丁洋,陆海川,等.五酸溶样-电感耦合等离子体质谱法同时测定地质样品中的稀土等28种金属元素[J].*岩矿测试*,2021,40(3):340–348.
Gong C, Ding Y, Lu H C, et al. Simultaneous determination of 28 elements including rare earth elements by ICP-MS with five-acid dissolution[J]. *Rock and Mineral Analysis*, 2021, 40(3): 340–348.
- [15] Wang X, Chen S, Wu M F, et al. Low-cost smartphone-based LIBS combined with deep learning image processing for accurate lithology recognition[J]. *Chemical Communications*, 2021, 57(58): 7156–7159.
- [16] Sun C, Xu W J, Tan Y Q, et al. From machine learning to transfer learning in laser-induced breakdown spectroscopy analysis of rocks for Mars exploration[J]. *Scientific Reports*, 2021, 11(1): 21379.
- [17] Zhang D, Chu Y W, Ma S X, et al. A plasma-image-assisted method for matrix effect correction in laser-induced breakdown spectroscopy[J]. *Analytica Chimica Acta*, 2020, 1107: 14–22.
- [18] Agresti J, Indelicato C, Perotti M, et al. Quantitative compositional analyses of calcareous rocks for lime industry using LIBS[J]. *Molecules*, 2022, 27(6): 1813.
- [19] Pagnotta S, Lezzerini M, Campanella B, et al. A new approach to non-linear multivariate calibration in laser-induced breakdown spectroscopy analysis of silicate rocks[J]. *Spectrochimica Acta Part B-Atomic Spectroscopy*, 2020, 166: 105804.
- [20] Hussain A, Xun G, Asghar H, et al. Enhancement of laser-induced breakdown spectroscopy (LIBS) signal subject to the magnetic confinement and dual pulses[J]. *Optics and Spectroscopy*, 2021, 129(4): 452–459.
- [21] Yang R Z, Bi L Q. Spectral enhancement mechanism and analysis of defocused collinear DP-LIBS technology[J]. *Optik*, 2021, 243: 167025.
- [22] Lu C X, Wang B, Jiang X P, et al. Detection of K in soil using time-resolved laser-induced breakdown spectroscopy based on convolutional neural networks[J]. *Plasma Science & Technology*, 2019, 21(3): 034014.
- [23] Yan J J, Hao Z Q, Zhou R, et al. A quantitative analysis method assisted by image features in laser-induced breakdown spectroscopy[J]. *Analytica Chimica Acta*, 2019, 1082: 30–36.
- [24] Li Q Y, Tian Y, Xue B Y, et al. Improvement in the analytical performance of underwater LIBS signals by exploiting the plasma image information[J]. *Journal of Analytical Atomic Spectrometry*, 2020, 35(2): 366–376.
- [25] Porizka P, Demidov A, Kaiser J, et al. Laser-induced breakdown spectroscopy for *in situ* qualitative and quantitative analysis of mineral ores[J]. *Spectrochimica Acta Part B-Atomic Spectroscopy*, 2014, 101: 155–163.
- [26] Sanchez-Esteva S, Knadel M, Labouriau R, et al. Total phosphorus determination in soils using laser-induced breakdown spectroscopy: Evaluating different sources of matrix effects[J]. *Applied Spectroscopy*, 2021, 75(1): 22–33.
- [27] Kosareva N, Krylovas A, Zavadskas E K. Statistical analysis of MCDM data normalization methods using Monte Carlo approach. The case of ternary estimates matrix[J]. *Economic Computation and Economic Cybernetics Studies and Research*, 2018, 52(4): 159–175.
- [28] Jia Y, Zhao N J, Fang L, et al. Online calibration of laser-induced breakdown spectroscopy for detection of heavy metals in water[J]. *Plasma Science & Technology*, 2018, 20(9): 095503.
- [29] Lu P, Zhuo Z, Zhang W H, et al. Determination of calorific value in coal by LIBS coupled with acoustic normalization[J]. *Applied Physics B-Lasers and Optics*, 2021, 127(6): 82.
- [30] 刘旭阳,张大成,冯中琦,等.基于远程激光诱导击穿光谱技术的航空合金鉴别[J].*光子学报*,2021,50(10):165–172.
Liu X Y, Zhang D C, Feng Z Q, et al. Discrimination of aviation alloy by remote laser-induced breakdown spectroscopy technology[J]. *Acta Photonica Sinica*, 2021, 50(10): 165–172.
- [31] Zeng Q D, Chen G H, Li W X, et al. Classification of steel based on laser-induced breakdown spectroscopy combined with restricted Boltzmann machine and support vector machine[J]. *Plasma Science & Technology*, 2022, 24(8): 084009.
- [32] Yi R X, Li J M, Yang X Y, et al. Spectral interference elimination in soil analysis using laser-induced breakdown spectroscopy assisted by laser-induced fluorescence[J]. *Analytical Chemistry*, 2017, 89(4): 2334–2337.

- [33] Liu X L, Rao Y M, Liu J G, et al. Quantitative analysis of gadolinium in thorium oxide matrix using laser-induced breakdown spectroscopy[J]. *Applied Physics B-Lasers and Optics*, 2022, 128(10): 192.
- [34] Martin M, Brice D, Martin S, et al. Inorganic characterization of switchgrass biomass using laser-induced breakdown spectroscopy[J]. *Spectrochimica Acta Part B-Atomic Spectroscopy*, 2021, 186: 106323.
- [35] Li J J, Chen F, Huang G Q, et al. Identification of Graves' ophthalmology by laser-induced breakdown spectroscopy combined with machine learning method[J]. *Frontiers of Optoelectronics*, 2021, 14(3): 321-328.
- [36] Liu X N, Che X Q, Li K Y, et al. Geographical authenticity evaluation of *Mentha haplocalyx* by LIBS coupled with multivariate analyzes[J]. *Plasma Science & Technology*, 2020, 22(7): 074006.
- [37] 侯佳佳, 张雷, 赵洋, 等. 基于共振与非共振双线的自吸收免疫LIBS技术研究[J]. *光谱学与光谱分析*, 2020, 40(1): 261-265.
- Hou J J, Zhang L, Zhao Y, et al. Investigation on resonance and non-resonance doublet based self-absorption-free LIBS technique[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(1): 261-265.
- [38] Guo Y M, Guo L B, Li J M, et al. Research progress in Asia on methods of processing laser-induced breakdown spectroscopy data[J]. *Frontiers of Physics*, 2016, 11(5): 114212.