

基于循环变量筛选非线性偏最小二乘的 LIBS 铁矿浆定量分析

尚栋^{1,2,3,4}, 孙兰香^{1,2,3*}, 齐立峰^{1,2,3}, 谢远明^{1,2,3,5}, 陈彤^{1,2,3,4}

¹中国科学院沈阳自动化研究所机器人学国家重点实验室, 辽宁 沈阳 110016;

²中国科学院网络化控制系统重点实验室, 辽宁 沈阳 110016;

³中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169;

⁴中国科学院大学, 北京 100049;

⁵沈阳化工大学, 辽宁 沈阳 110142

摘要 激光诱导击穿光谱(LIBS)技术因其在线、原位、多元素同时测量等优点,在物质成分检测上得到广泛应用。但是,LIBS技术常受到自吸收及基体效应的干扰,分析的准确度较低,同时,随着光谱仪分辨率的不断提高,数据维度越来越高,其中包括大量对成分分析无用的冗余信息,这就增加了建模的复杂度。为了降低建模的复杂度,减少光谱数据维度以提取最有力的光谱信息,同时减少自吸收及基体效应的非线性干扰对定量分析精度的影响,在传统偏最小二乘(PLS)方法的基础上,提出了利用循环筛选特征变量来校正自吸收及基体效应影响的非线性 PLS 模型。以铁精矿矿浆样本为分析对象,结果表明,与传统 PLS 方法相比,所提出的基于循环变量筛选的非线性 PLS 模型的定量分析精度显著提高,测试样品的均方根误差(RMSE)从 1.15% 降到 0.70%,决定系数 R^2 从 0.51 提高到 0.86。

关键词 光谱学; 激光诱导击穿光谱; 非线性偏最小二乘模型; 变量筛选; 自吸收效应; 基体效应

中图分类号 O657.31

文献标志码 A

doi: 10.3788/CJL202148.2111001

1 引言

在矿物加工过程中,准确获取矿物质含量信息对于充分利用原材料和能源、精准控制生产过程至关重要。以铁矿石为例,铁矿石原矿到最后变成钢铁要经过一系列复杂的选矿流程,其中矿物浮选是根据目标矿物与杂质理化性质的不同,将目标矿物与杂质分离,进而从原矿浆中提取出来的选矿方法^[1]。铁矿浆品位(成分含量)不同直接影响浮选的效果及最后产品的质量和产出效益,因此对铁矿浆成分进行准确的品位分析至关重要^[2]。

目前对矿浆品位进行分析的常用方法一般有化学分析法和 XRF 荧光分析法。化学分析法是以物质的化学反应为基础的经典分析方法,分析精度高,

但是该方法的操作复杂且分析周期长,分析效率低,无法做到实时在线检测^[3]。XRF 荧光分析法能够实现在线检测,但是在线检测无法检测原子序数 20 之前的元素,且 X 射线具有放射性,具有潜在的危害性。而激光诱导击穿光谱(LIBS)技术作为一种新兴检测方法,因具有可分析元素范围广,对待测物质的形态和样品制备要求低,可实时在线检测、多元素同时分析等优点受到越来越多研究人员的关注,在诸多领域得到广泛应用^[4-10]。LIBS 技术将高能激光聚焦于待测物质表面,激发待测物质产生等离子体,然后对等离子体的发射光谱进行收集,通过对发射光谱的处理和分析,从而得出待分析物质的元素信息。

一些研究通过对 LIBS 矿浆设备开发和实验参

收稿日期: 2021-03-15; 修回日期: 2021-04-12; 录用日期: 2021-04-19

基金项目: 国家重点研发计划(2016YFF0102502)、中国科学院前沿科学重点研究计划(QYZDJ-SSW-JSC037)、中国科学院青年创新促进会和辽宁省“兴辽英才计划”资助项目(XLYC1807110)

通信作者: *sunlanxiang@sia.cn

数优化,取得了较为满意的效果,证明了 LIBS 应用于铁矿浆分析的可行性^[11-15]。Khajehzadeh 等^[16]利用 LIBS 技术对铁矿尾矿浆品位进行分析,利用偏最小二乘(PLS)方法对光谱数据建模,分析了铁矿尾矿浆的多种成分,其中 Fe 元素的决定系数 R^2 达到 0.826,但均方根误差(RMSE)相对较高,仅为 2.86%。

在利用 LIBS 技术进行定量分析时,由于光谱数据包含大量干扰噪声,同时受自吸收及基体效应的影响,单变量定量分析精度并不理想。而作为多变量线性回归方法,PLS 方法以其简单快速、定量精度较高等优点,成为目前应用广泛的定量分析方法,PLS 尤其适合处理样本量较少、变量维度高且变量之间存在多重共线性的问题^[17]。丛智博等^[18]利用 PLS 方法对合金钢组分进行分析,证明了 PLS 在进行受复杂基体影响的元素分析时具有优势;但是其将全谱数据进行建模并加入大量与待分析元素无关的信息,容易造成过拟合。特征提取是一种常用的降维方法,能够降低光谱数据的维度,降低建模的复杂度^[19-21]。但是,作为一种线性处理方法,传统的 PLS 模型并不能解决自吸收及基体效应对光谱数据造成的非线性影响,从而限制了该方法在定量分析精度上的进一步提高。Wang 等^[22]提出一种多元非线性 PLS 模型,以标准铜样为分析对象,通过将 Cu 和 Pb 的多条特征谱线强度的多项式加入 PLS 建模中,再结合全谱数据进行修正,降低了自吸收和元素间的干扰,与传统 PLS 相比,该方法的

总体 RMSE 从 2.81% 降到 1.05%。

在利用 PLS 进行 LIBS 的定量分析时,现有研究分别采用谱线特征选择方法来降维或单独采用非线性修正方法进行改进。为同时降低数据维度,减少冗余信息的干扰并校正数据本身的非线性问题,本文结合谱线特征选择和非线性修正方法,提出一种基于循环变量筛选的非线性 PLS 模型以提高铁矿浆品位分析精度。

2 装置及样本

2.1 仪器及参数

LIBS 铁矿浆实验装置采用中国科学院沈阳自动化研究所研制的 SIA-LIBSlurry 分析仪,原理图和实物照片如图 1 所示。等离子体激发光源为 Nd:YAG 双脉冲激光器(DPS-1064-BS-D,长春新产业),每个激光器的输出波长均为 1064 nm,每个激光器的最大输出能量均为 100 mJ,激光脉冲经过焦距均为 750 mm 的会聚透镜在样品表面产生等离子体。光谱探测装置采用 Avantes 公司生产的 AvaSpec-2048 光谱仪,波长范围为 227 ~ 784 nm,分辨率为 0.1 ~ 0.2 nm,通过本课题组自行研制的多通道时序控制器来调节激光器激发时间与光谱仪收集光谱时间之间的延迟,激光器、光谱仪和时序控制器均被封装在控制柜中。采用上海高晋流体科技有限公司生产的 QBY3-15S 气动隔膜泵作为矿浆循环系统动力源,矿浆搅拌装置采用台湾世汇公司生产的气动搅拌机,以保证矿浆成分均匀,矿浆喷嘴直径为 7 mm。

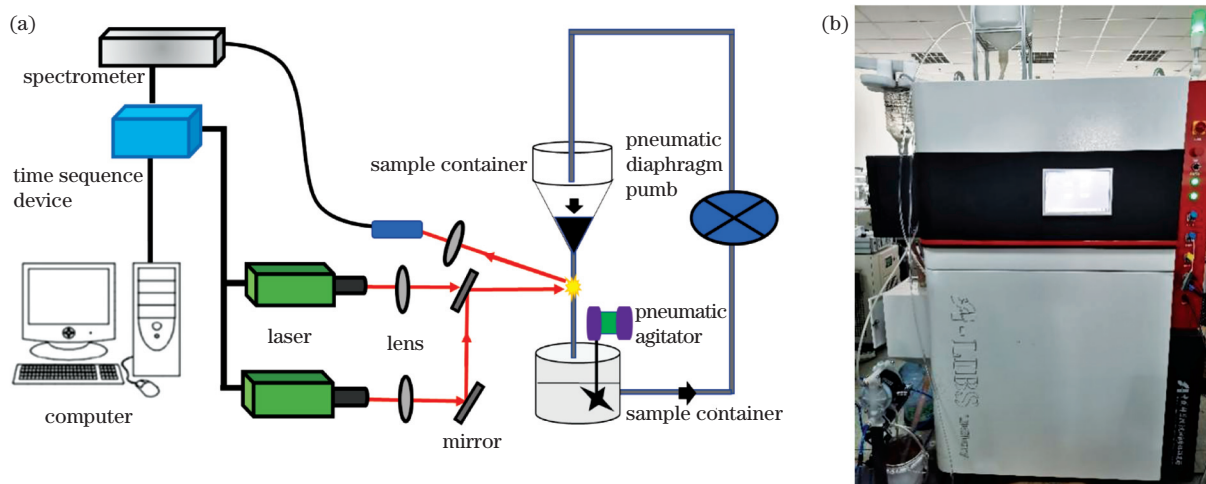


图 1 LIBS 铁矿浆系统。(a)原理图;(b)实验装置实物照片

Fig. 1 LIBS iron ore slurry system. (a) Schematic of the system; (b) photo of the experimental device

2.2 样品制备

所用铁矿浆样本采自某铁矿选矿厂,共 95 个铁精矿矿浆样品,每个样本大约 5 L,如图 2 所示。Fe

元素的品位范围为 61.3% ~ 68.6%。实验过程中,首先选择不同的激光脉冲能量、采集延迟时间、两个激光器间的触发时间间隔进行实验,将最优的实验

参数作为最终铁精矿矿浆的实验条件,每个激光器的能量均为 60 mJ,即双激光脉冲总能量为 120 mJ,光谱采集延时 0.8 μ s,激光器间延时 0 μ s;同时气动隔膜泵的进气量会影响矿浆液柱的流速和稳定性,经过反复试验最终确定当矿浆在漏斗容器的液位为 1/3 时实验结果最好。



图 2 铁精矿矿浆样品

Fig. 2 Iron ore concentrate slurry samples

在 10 Hz 频率下对每个样本采集 1100 张谱,每张谱共 6116 维,去掉前 100 张谱,对剩下的 1000 张谱取平均,95 个样本最终得到 95 张谱,矿浆样品光谱示例如图 3 所示。

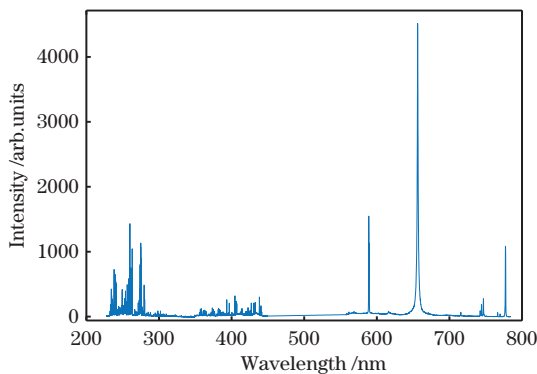


图 3 铁精矿矿浆样品光谱示例

Fig. 3 Plasma spectrum example of iron ore concentrate slurry sample

3 建模方法

3.1 PLS 方法

PLS 作为一种多元统计分析方法,其在主成分分析的基础上,结合了多元线性回归和典型相关分析,不仅可以对变量进行主成分分析,还可以根据解释变量和被解释变量之间的关系进行线性建模。对于传统 PLS 模型,浓度和谱线强度的关系式可表示为

$$C = \sum_{i=1}^n \beta_i I_i + E, \quad (1)$$

式中: C 为元素浓度; β_i 为回归系数; I_i 为特征谱线

强度值; n 表示选择的特征谱线数; E 为残差。

3.2 非线性 PLS 方法

为降低自吸收对待分析元素定量分析的影响,选取 m 条待分析 Fe 元素的特征谱线强度;同时为了降低其他元素的干扰,而铁矿浆中的主要干扰元素为 Si,因此选择 h 条 Si 元素的特征谱线强度。将 Fe 和 Si 元素特征谱线强度的三次多项式加入 PLS 的建模中,以近似校正自吸收和基体效应带来的非线性影响,模型表达式为

$$C = \sum_{i=1}^e \sum_{j=0}^m \alpha_{i,j} I_{Fe,j}^i + \sum_{i=1}^e \sum_{j=0}^h \beta_{i,j} I_{Si,j}^i + \sum_{i=1}^k \lambda_i I_i + \lambda_0, \quad (2)$$

式中: e 为最高次数; $\alpha_{i,j}$ 为 Fe 的线回归系数; I_{Fe} 为 Fe 线强度; $\beta_{i,j}$ 为 Si 的线回归系数; I_{Si} 为 Si 线强度; k 为全谱变量数; λ 为全谱谱线的回归系数; I_i 为全谱谱线强度。

3.3 基于循环变量筛选的非线性 PLS 方法

在非线性 PLS 模型的基础上,为减少变量冗余信息的干扰,降低模型复杂度,通过循环变量筛选的方式确定最优变量,模型表达式为

$$C = \sum_{i=1}^e \sum_{j=0}^m \alpha_{i,j} I_{Fe,j}^i + \sum_{i=1}^e \sum_{j=0}^h \beta_{i,j} I_{Si,j}^i + \sum_{k=1}^u \gamma_k I_k + \gamma_0, \quad (3)$$

式中: u 为从全谱中筛选的谱线数; γ 为筛选后谱线的回归系数; I_k 为筛选后谱线的强度。

基于循环变量筛选的非线性 PLS 方法步骤如下:

- 1) 数据预处理,即全谱和归一化,小波降噪;
- 2) 对训练样本选取分析线(m 条待分析元素 Fe 的特征谱线和 h 条主要干扰元素 Si 的特征谱线)强度的三次多项式以及全谱数据进行 PLS 建模,主成分数通过校验集的最小 RMSE 确定;
- 3) 将变量按回归系数的绝对值从大到小排列;
- 4) 去除回归系数绝对值最小的 50 个变量,将剩下的变量数据进行 PLS 建模,记录校验集的 RMSE;
- 5) 重复步骤 4) 直到建模变量数小于设定的最少变量数 100;

6) 根据校验集的 RMSE 确定建立非线性 PLS 模型的最终变量,建立循环筛选变量后的非线性 PLS 模型,主成分数由校验集最小 RMSE 确定。

基于循环变量筛选的非线性 PLS 算法流程如图 4 所示。

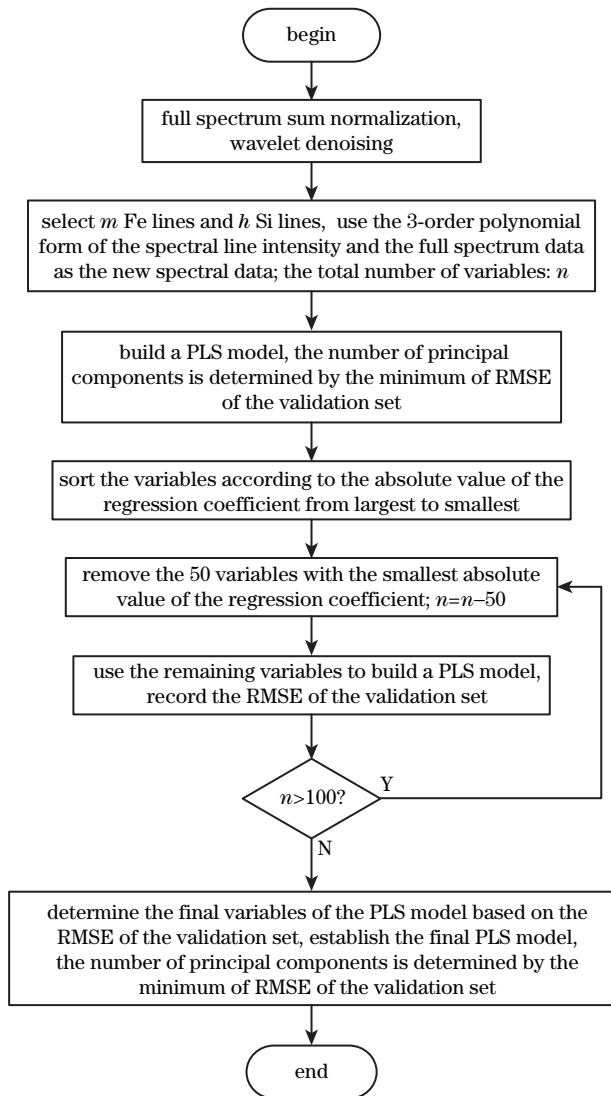


图 4 基于循环变量筛选的非线性 PLS 算法流程图
Fig. 4 Flow chart of the nonlinear PLS algorithm based on cyclic variable filtering

4 结果与讨论

实验对 95 个铁精矿矿浆样本进行分析,选取

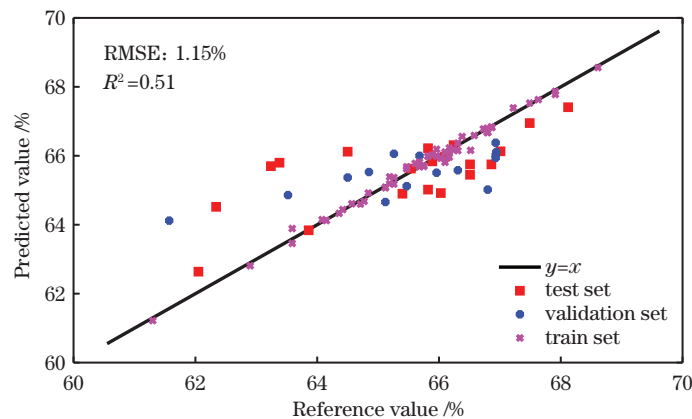


图 6 传统 PLS 模型的 Fe 元素质量分数分析结果
Fig. 6 Analysis results of Fe mass fraction in the traditional PLS model

15 个样本作为校验样本,利用该样本来优化模型的参数,20 个样本作为测试样本,用于评价最终模型的预测精度。在划分校验样本和测试样本时,使其均匀分布在总样本的浓度范围之内,以达到最充分地评估模型性能的效果。将其余 60 个样本作为训练样本进行建模,分别采用传统 PLS 模型、加入特征谱线三次多项式的非线性 PLS 模型以及基于循环变量筛选的非线性 PLS 模型进行分析比较。相关算法在 Matlab R2017a 环境下编写。

4.1 传统 PLS 模型

将选定的 60 个样本作为定标样本建立传统 PLS 模型。将获得的大小为 60×6116 的光谱数据经过全谱和归一化与两层小波降噪等预处理,作为传统 PLS 模型的输入,对应样本 Fe 元素的质量分数作为输出,通过校验集的最小 RMSE 确定模型的最优主成分。校验集 RMSE 随主成分数 N 的变化如图 5 所示。

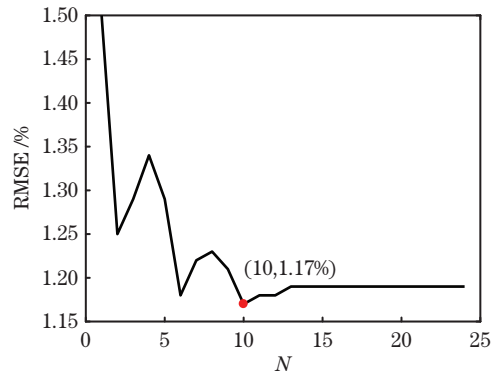


图 5 传统 PLS 模型的校验集 RMSE 与主成分数的关系
Fig. 5 RMSE of validation set as a function of number of components in the traditional PLS model

如图 5 所示,当主成分数为 10 时,校验集的 RMSE 最小,因此,本实验选择 10 个主成分来建立传统 PLS 定标模型,校验集和测试集的真实值与预测值如图 6 所示。

从图 6 可以看到,传统的 PLS 建模方法对铁精矿矿浆测试集的决定系数 R^2 仅为 0.51,其 RMSE 为 1.15%,而实际应用中 RMSE 需要低于 1%,因此测量准确性需要进一步提高。因铁精矿矿浆基体成分复杂,受到严重的谱线干扰和自吸收影响,光谱数据包含大量冗余信息,传统的 PLS 建模方法不能满足工业现场的精度要求。

4.2 非线性 PLS 模型

为降低自吸收和基体效应对 Fe 元素定量分析的影响,选取 Fe 元素的 10 条特征谱线和主要干扰元素 Si 的 5 条特征谱线,将其三次多项式加入 PLS 的建模中,以近似校正自吸收和基体效应带来的非线性影响,选取的 Fe 谱线和 Si 谱线如表 1 所示,其

表 1 选择的分析线

Table 1 Selected analytical lines

Serial No.	Emission lines /nm	E_i /eV	E_j /eV
1	Fe I 252.2849	0	4.913304
2	Fe I 271.9027	0	4.558830
3	Fe I 322.7796	2.425597	6.265889
4	Fe II 259.9396	0	4.768628
5	Fe II 261.1874	0.047711	4.793558
6	Fe II 261.3825	0.106958	4.849263
7	Fe II 273.9548	0.986398	5.511082
8	Fe II 274.6484	1.076312	5.589570
9	Fe II 274.9321	1.040538	5.549138
10	Fe II 275.5737	0.986398	5.484502
11	Si I 250.6897	0.009562	4.954129
12	Si I 251.9202	0.009562	4.929980
13	Si I 252.4108	0.009562	4.920417
14	Si I 252.8509	0.027670	4.929980
15	Si I 288.1577	0.781011	5.082689

中 E_i 和 E_j 为粒子所在的激发能级。

将 60 个训练样本作为定标样本建立非线性 PLS 模型。将获得的 60×6116 维光谱数据和挑选的 15 条特征谱线的三次多项式共 60×6161 维数据进行非线性 PLS 建模,经过全谱和归一化和两层小波降噪等预处理后作为非线性 PLS 模型的输入,对应样本 Fe 元素的质量分数作为输出,最优主成分数由校验集的最小 RMSE 确定。校验集 RMSE 随主成分数 N 的变化如图 7 所示。

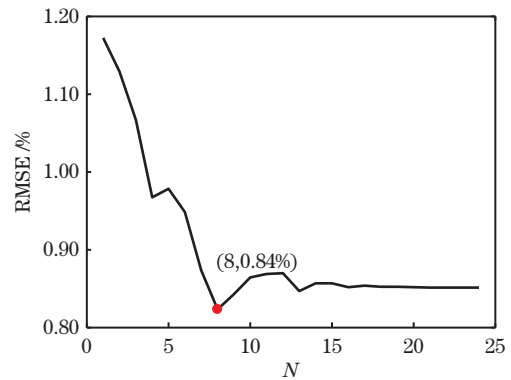


图 7 非线性 PLS 模型的校验集 RMSE 与主成分数的关系
Fig. 7 RMSE of validation set as function of number of components in the nonlinearity PLS model

由图 7 确定的最优主成分数为 8,建立的非线性 PLS 模型如图 8 所示。

比较传统 PLS 和非线性 PLS 两种方法的预测结果,测试样本的 RMSE 从 1.15% 降到 0.85%,同时其 R^2 从 0.51 升高到 0.73,说明加入特征谱线三次多项式的非线性 PLS 模型在校正自吸收和基体效应带来的非线性干扰方面有明显的效果。

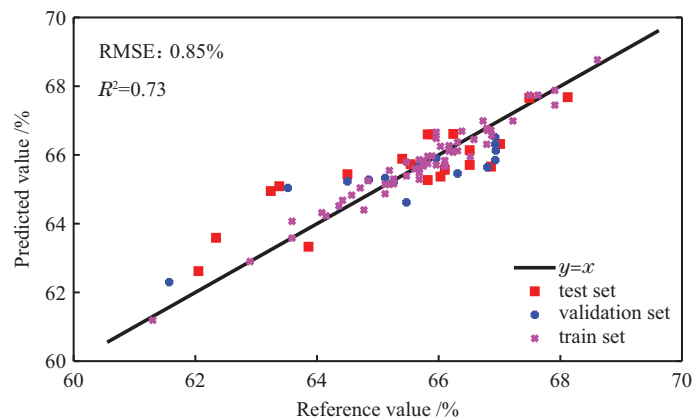


图 8 非线性 PLS 模型的 Fe 元素质量分数分析结果

Fig. 8 Analysis results of Fe mass fraction in the nonlinear PLS model

4.3 基于循环变量筛选的非线性 PLS 模型

对训练集的 60×6116 维光谱数据和挑选的

15 条特征谱线的三次多项式共 60×6161 维数据进行非线性 PLS 建模,同样经过全谱和归一化和两层

小波降噪预处理后作为非线性 PLS 模型的输入, 对应样本 Fe 元素的质量分数作为输出, 最优主成分数仍然根据图 7 所示的校验集最小 RMSE 确定为 8。

将光谱数据按回归系数绝对值进行排序, 每次变量筛选去除绝对值最小的 50 个光谱数据后, 再对剩下的数据进行 PLS 建模, 记录校验集的 RMSE。校验集 RMSE 和变量筛选次数 T 的关系如图 9 所示。

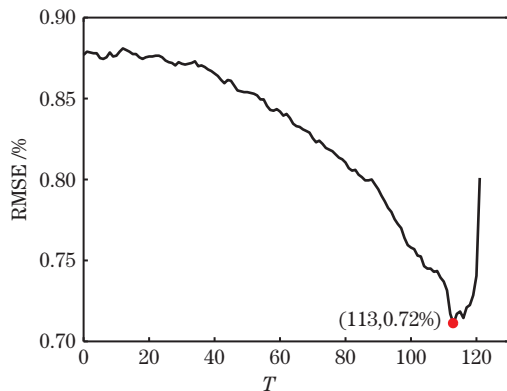


图 9 基于循环变量筛选的非线性 PLS 模型校验集 RMSE 与变量筛选次数的关系

Fig. 9 RMSE of validation set as function of variable selection times in the nonlinear PLS model based on cyclic variable filtering

从图 9 得出当筛选次数为 113 次, 即变量个数为 511 时, 校验样本的 RMSE 达到最小值 0.72%。

确定好最优变量数后, 在新变量构成的数据集下, 再根据校验集的 RMSE 确定最佳主成分数, RMSE 随主成分数 N 的变化如图 10 所示。

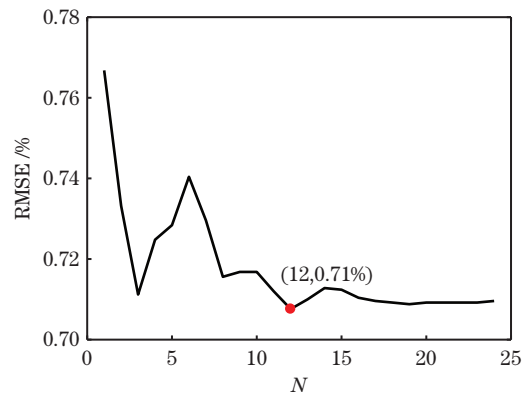


图 10 筛选变量后校验集的 RMSE 与主成分数的关系
Fig. 10 RMSE of validation set after filtering variables as a function of number of components

从图 10 可以看到, 当主成分数等于 12 时, 校验集的 RMSE 达到最小值, 因此确定主成分数 12 为筛选变量后非线性 PLS 模型的最优主成分数。

在筛选后的新变量和主成分数 12 下, 用训练集建立非线性 PLS 模型, 校验集和测试集的分析结果如图 11 所示。

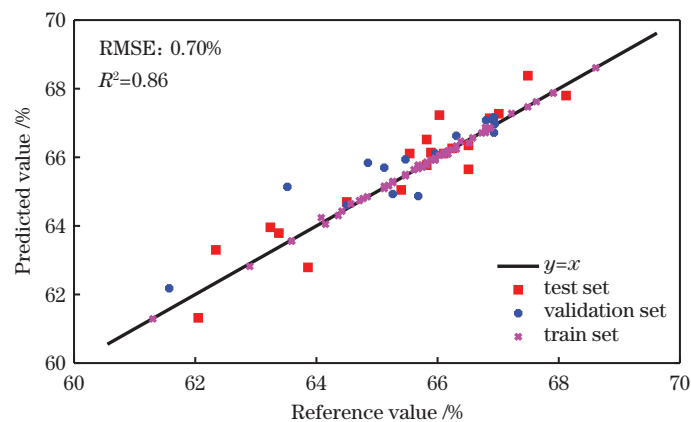


图 11 基于循环变量筛选的非线性 PLS 模型的 Fe 元素质量分数分析结果

Fig. 11 Analysis results of Fe mass fraction in the nonlinear PLS model based on cyclic variable filtering

从图 11 可以看到, 经过循环变量筛选的非线性 PLS 模型对铁精矿矿浆中 Fe 元素定量分析的准确度显著提高, 测试集的 RMSE 降低到 0.70%, 其决定系数 R^2 达到 0.86。3 种算法在测试集上的结果对比如表 2 所示。传统 PLS 方法的建模时间为 1.08 s, 所提出的基于循环变量筛选的非线性 PLS 方法的建模时间为 1.32 s, 所提方法的建模时间相较于传统 PLS 方法略有增长。

表 2 三种算法的结果对比

Table 2 Comparison of the results of the three algorithms

Algorithm	RMSE / %	R^2
PLS	1.15	0.51
Nonlinear PLS	0.85	0.73
Nonlinear PLS based on cyclic variable filtering	0.70	0.86

5 结 论

利用 LIBS 技术获得的光谱数据常受到自吸收及基体效应的非线性干扰,从而影响定量分析的精度,为了解决 LIBS 技术在利用 PLS 方法进行回归分析时存在的信息冗余和无法处理非线性数据的问题,提出一种基于循环变量筛选的非线性 PLS 模型,以铁精矿矿浆中的 Fe 元素为分析对象,并传统 PLS 建模方法进行比较。结果表明,基于变量筛选的非线性 PLS 模型虽在建模时间上略有增加,但对铁精矿矿浆中 Fe 元素的分析精度显著提高,说明该方法可用于受基体效应和自吸收影响较大的元素的定量分析,其对在线建模应用具有参考意义。

参 考 文 献

- [1] Wills B A, Hopkins D W, et al. Wills' mineral processing technology [M]. 7th ed. Amsterdam: Elsevier, 2006:267-270.
- [2] Khajehzadeh N, Haavisto O, Koresaar L. On-stream mineral identification of tailing slurries of an iron ore concentrator using data fusion of LIBS, reflectance spectroscopy and XRF measurement techniques [J]. Minerals Engineering, 2017, 113: 83-94.
- [3] Geng W R. The development status of automatic measuring instruments in concentrators [J]. Modern Mining, 2014, 30(7): 172-174.
耿文瑞. 选矿厂自动化测量仪表的发展现状 [J]. 现代矿业, 2014, 30(7): 172-174.
- [4] Chen J Z, Wang J, Song G J, et al. Recent development and application of laser induced breakdown spectroscopy [J]. Chinese Science Bulletin, 2016, 61(10): 1086-1098.
陈金忠, 王敬, 宋广聚, 等. 激光诱导击穿光谱的近期发展与应用 [J]. 科学通报, 2016, 61(10): 1086-1098.
- [5] Zheng P C, Liu H D, Wang J M, et al. Online mercury determination by laser-induced breakdown spectroscopy with the assistance of solution cathode glow discharge [J]. Journal of Analytical Atomic Spectrometry, 2015, 30(4): 867-874.
- [6] Lin Z X, Liu L M, Liu L W. Validation of the solidifying soil process using laser-induced breakdown spectroscopy [J]. Optics & Laser Technology, 2016, 83: 13-15.
- [7] Pan L J, Chen W F, Cui R F, et al. Adaptive selection method for analytical lines in laser-induced breakdown spectra [J]. Chinese Journal of Lasers, 2020, 47(8): 0811001.
- [8] Zhong Z M, Lu J D, Zheng J P, et al. Study on properties of plasma spectroscopy of laser-induced carbon materials with different chemical forms [J]. Acta Optica Sinica, 2011, 31(12): 1230001.
钟子铭, 陆继东, 郑建平, 等. 激光诱导不同化学形态碳等离子体光谱特性研究 [J]. 光学学报, 2011, 31(12): 1230001.
- [9] Liu S M, Xiu J S, Liu Y Y. Rapid quantitative analysis of element content ratios in Cu(In, Ga)Se₂ thin films using laser-induced breakdown spectroscopy [J]. Chinese Journal of Lasers, 2019, 46(9): 0911001.
刘世明, 修俊山, 刘云燕. 基于激光诱导击穿光谱技术的铜铟镓硒薄膜中元素含量比的快速定量分析方法 [J]. 中国激光, 2019, 46(9): 0911001.
- [10] Yang Y L, Wang L, Ma C H. Quantitative analysis of Mn in molten steel with multi-factor calibration input using laser-induced breakdown spectroscopy [J]. Laser & Optoelectronics Progress, 2019, 56(22): 223003.
杨友良, 王禄, 马翠红. 多元校正输入的钢液 Mn 元素 LIBS 定量分析 [J]. 激光与光电子学进展, 2019, 56(22): 223003.
- [11] Guo L B, Cheng X, Tang Y, et al. improvement of spectral intensity and resolution with fiber laser for on-stream slurry analysis in laser-induced breakdown spectroscopy [J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2019, 152: 38-43.
- [12] Michaud D, Leclerc R, Proulx É. Influence of particle size and mineral phase in the analysis of iron ore slurries by laser-induced breakdown spectroscopy [J]. Spectrochimica Acta Part B, 2007, 62(12): 1575-1581.
- [13] Cheng X, Yang X Y, Zhu Z H, et al. On-stream analysis of iron ore slurry using laser-induced breakdown spectroscopy [J]. Applied Optics, 2017, 56(33): 9144-9149.
- [14] Ayyalasmayajula K K, Dikshit V, Yueh F Y, et al. Quantitative analysis of slurry sample by laser-induced breakdown spectroscopy [J]. Analytical and Bioanalytical Chemistry, 2011, 400(10): 3315-3322.
- [15] Oh S Y, Yueh F Y, Singh J P, et al. Preliminary evaluation of laser induced breakdown spectroscopy for slurry samples [J]. Spectrochimica Acta Part B, 2009, 64(1): 113-118.
- [16] Khajehzadeh N, Haavisto O, Koresaar L. On-stream and quantitative mineral identification of tailing slurries using LIBS technique [J]. Minerals Engineering, 2016, 98: 101-109.

- [17] Luo D Z. Research on LIBS spectrum recognition technology based on PLS algorithm [D]. Chengdu: University of Electronic Science and Technology of China, 2018: 54-59.
罗德志. 基于偏最小二乘算法的激光诱导击穿光谱识别技术研究[D]. 成都: 电子科技大学, 2018: 54-59.
- [18] Cong Z B, Sun L X, Xin Y, et al. Quantitative analysis of alloy steel based on laser induced breakdown spectroscopy with partial least squares method [J]. Spectroscopy and Spectral Analysis, 2014, 34(2): 542-547.
丛智博, 孙兰香, 辛勇, 等. 基于激光诱导击穿光谱的合金钢组分偏最小二乘法定量分析[J]. 光谱学与光谱分析, 2014, 34(2): 542-547.
- [19] Kong H Y, Sun L X, Hu J T, et al. Automatic method for selecting characteristic lines based on genetic algorithm to quantify laser-induced breakdown spectroscopy[J]. Spectroscopy and Spectral Analysis, 2016, 36(5): 1451-1457.
- [20] Amador-Hernández J, García-Ayuso L E, Fernández-Romero J M, et al. Partial least squares regression for problem solving in precious metal analysis by laser induced breakdown spectrometry [J]. Journal of Analytical Atomic Spectrometry, 2000, 15(6): 587-593.
- [21] Wang G D. Quantitative analysis and grade identification of alloy materials for portable LIBS system[D]. Beijing: University of Chinese Academy of Sciences, 2020: 24-38.
王国栋. 面向便携 LIBS 系统的合金材料定量分析与牌号鉴别[D]. 北京: 中国科学院大学, 2020: 24-38.
- [22] Wang Z, Feng J, Li L Z, et al. A non-linearized PLS model based on multivariate dominant factor for laser-induced breakdown spectroscopy measurements [J]. Journal of Analytical Atomic Spectrometry, 2011, 26(11): 2175-2182.

Quantitative Analysis of Laser-Induced Breakdown Spectroscopy Iron Ore Slurry Based on Cyclic Variable Filtering and Nonlinear Partial Least Squares

Shang Dong^{1,2,3,4}, Sun Lanxiang^{1,2,3*}, Qi Lifeng^{1,2,3}, Xie Yuanming^{1,2,3,5}, Chen Tong^{1,2,3,4}

¹State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

²Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

³Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, Liaoning 110169, China;

⁴University of Chinese Academy of Sciences, Beijing 100049, China;

⁵Shenyang University of Chemical Technology, Shenyang, Liaoning 110142, China

Abstract

Objective From iron ore to the final steel processing, accurate mineral content data is essential to maximize raw materials and energy accurately control the manufacturing. Mineral flotation is a beneficiation method in which target minerals and impurities are separated based on the physical and chemical properties of target minerals and impurities and then extracted from the original ore slurry. Content of iron ore slurry directly affects the flotation effect and quality and output benefit of the final product. Therefore, conducting an accurate quantitative analysis of the iron ore slurry composition is essential. Laser-induced breakdown spectroscopy (LIBS) has been widely used to detect material composition owing to its advantages such as online, *in situ*, and simultaneous measurement of multiple elements. However, self-absorption and matrix effects in LIBS affect the accuracy of the analysis. Simultaneously, with the continuous improvement of the spectrometer's resolution, the data dimension is increasing, including a large amount of redundant information that is unnecessary for component analysis. When using PLS and LIBS for quantitative analysis, the existing research uses spectral line feature selection to reduce dimensionality and nonlinear correction to make improvements separately. To simultaneously reduce the data dimension and correct the nonlinear problem of the data itself, we build a nonlinear PLS model to reduce the influence of self-absorption and matrix effects on the accuracy of quantitative analysis. In addition, the characteristic variables are cyclically filtered to reduce the modeling complexity.

Methods PLS is widely used in the quantitative analysis of material components, but as a linear processing method, it cannot resolve the nonlinear effects of self-absorption and matrix effects on the spectrum, reducing the accuracy of quantitative analysis. The characteristic spectrum line n -order polynomial form was proposed to be added to the PLS model. Thus, we can reduce the dimensionality of the data to extract the most useful information and reduce the complexity of the model by filtering feature variables. Taking the iron (Fe) element in the iron ore concentrate slurry as the analysis object, to reduce the influence of self-absorption on the quantitative analysis of the element to be analyzed, 10 characteristic spectral lines of Fe were selected. Simultaneously, to reduce the interference of other elements, 5 characteristic spectral lines of silicon (Si) were selected, and their three-order polynomial form was added to the modeling of PLS to correct the nonlinear influence caused by self-absorption and matrix effect. The regression coefficients of the variables were sorted according to the absolute value, and the optimal variables were determined by cyclically filtering the variables to reduce the interference of redundant information of the variables and reduce the model's complexity.

Results and Discussions Using the training set to build the model and determining the optimal variables and the number of principal components according to the root mean square error (RMSE) of the validation set, we made predictions on the prediction set and compared the traditional PLS model, the nonlinear PLS model with the characteristic spectrum lines three-order polynomial form, and the model proposed in this paper. The RMSE of the traditional PLS model is 1.15%, and the coefficient of determination R^2 is only 0.51 (Fig.6). However, the RMSE of the nonlinear PLS model is reduced by 0.85%, and the coefficient of determination R^2 is 0.73 (Fig. 8). Furthermore, the RMSE of the cyclic filtering variable nonlinear PLS model proposed in this paper is reduced to 0.70%, and the coefficient of determination R^2 is increased to 0.86 (Fig.11).

Conclusions We propose a nonlinear PLS model based on cyclic variable filtering to address the problem that LIBS is used for composition analysis, which is often affected by self-absorption and matrix effects and data redundancy caused by excessively high spectral data dimensions. The analysis object is the Fe element in the iron ore concentrate slurry, compared with the traditional PLS modeling method (Table 2). As a result, the RMSE of validation set is reduced from 1.15% to 0.70%, and the coefficient of determination R^2 increased from 0.51 to 0.86. The result shows that the nonlinear PLS model based on cyclic variable filtering can significantly improve the analysis accuracy of Fe in iron concentrate slurry, indicating that this method has evident effects on the quantitative analysis of elements that are greatly affected by the matrix effect and self-absorption.

Key words spectroscopy; laser-induced breakdown spectroscopy; non-linear partial least squares model; variable filtering; self-absorption effect; matrix effect;

OCIS codes 300.6365; 140.3440; 140.3510