

化学计量学在激光诱导击穿光谱分析中的研究进展

张天龙¹ 吴珊¹ 汤宏胜¹ 王康² 段忆翔³ 李华^{* 1,4}

¹(西北大学分析科学研究所, 西安 710069) ²(长安大学理学院, 西安 710064)

³(四川大学分析仪器研究中心, 成都 610064) ⁴(西安石油大学化学化工学院, 西安 710065)

摘要 激光诱导击穿光谱作为一种新兴的元素分析技术,具有实时在线、非接触、多元素同时检测等优点,是光谱分析领域内的一种前沿性分析手段。然而,如何从复杂大量的激光诱导击穿光谱数据提取有用信息,提高其定性、定量分析准确度是激光诱导击穿光谱技术目前面临的难题。化学计量学作为多学科交叉的化学分支学科,在数据处理、信号解析和模式识别等方面具有优势,能够解决传统化学研究方法难以解决的一些复杂问题。本文从光谱数据预处理、定性和定量分析三方面综述了近年来化学计量学方法在激光诱导击穿光谱中的研究进展。

关键词 激光诱导击穿光谱; 化学计量学; 数据预处理; 定性分析; 定量分析; 评述

1 引言

激光诱导击穿光谱(Laser induced breakdown spectroscopy, LIBS)是一种基于原子发射光谱且激光作为激发源的全新物质元素分析技术^[1-3]。激光器发射出低能量脉冲激光经平面镜反射和平凸透镜聚焦,在样品表面产生高温的等离子体,激光诱导等离子体在冷却过程中会辐射出不同频率的光,通过光纤进行收集并耦合至光谱仪,通过特征谱线的波长和强度进行定性与定量分析。该技术起源于红宝石晶体激光器^[4],激光因其具有高密度能量、单色性和相干性等特性,受到了众多科学领域的极大关注。1962年, Brech 和 Cross^[5]首次报道了 LIBS 技术,随着激光技术和光学检测技术的发展, LIBS 技术已经成为光谱学领域的研究热点。目前, LIBS 综述类的文章在逐年增多^[6],然而,有关化学计量学在 LIBS 技术应用方面的综述尚很少见。LIBS 因其具有快速、实时、原位、无损、遥测、多元素同步测试等优势^[7-13],已经被成功地应用于材料、冶金、燃烧、环境、考古、太空探测、医学、军事等诸多领域^[14-21]。

LIBS 能快速获得分析样品的组分和含量信息,已成为目前多元素快速分析方法之一。由于分析对象的多样性和化学成分的复杂性,使得其产生大量复杂的分析数据。然而, LIBS 技术目前面临的难题就是如何从原始数据中提取有用信息,减少主观解释。应用化学计量学技术处理 LIBS 分析所产生的大量数据,提高其应用的可靠性与稳定性,已成为 LIBS 技术发展的一个关键科学问题和主要研究方向之一^[22]。化学计量学是一门多学科交叉的化学分支学科,它应用数学和统计学的方法,设计和选择最优测量程序和实验方法,并最大程度地获取有用信息以解释化学数据。化学计量学已深入到分析化学的各个领域,成为现代分析化学工作者不可缺少的工具。化学计量学的主要任务包括最佳实验条件的选择和最大程度提取有用信息,它在 LIBS 领域的各项相关研究中逐渐发挥了其优势,但尚未见化学计量学在 LIBS 研究中的应用综述,因此,总结化学计量学在 LIBS 分析中的研究进展具有一定的参考价值。本文综述了化学计量学方法在 LIBS 光谱预处理、定性分析和定量分析等领域的应用研究进展,并对其在 LIBS 技术的应用前景做了展望。

2 光谱预处理

由于激光器激光能量的波动、光谱仪分辨率差异、外部环境的影响以及样品表面不均匀等因素的影响, LIBS 采集到的光谱数据包含大量的干扰信息与有效信息,而化学计量学方法可以消除光谱中受随

2014-12-30 收稿; 2015-04-14 接受

本文系国家重大科学仪器设备开发专项项目(No. 2011YQ030113) 和国家自然科学基金(Nos. 21175106, 21375105) 项目资助

* E-mail: huali@nwu.edu.cn

机因素影响产生的误差以及各种非目标因素对光谱的影响。因此,光谱数据预处理成为光谱分析过程中的重要环节。光谱数据预处理不仅需要区分有效信息和干扰信息,提取有效信息,更重要的是从有效信息中提取表征欲获取信息特性的特征信息,筛选数据变量,优化光谱范围,为建立光谱校正模型和预测未知样品组分浓度或性质提供真实准确的数据基础。LIBS 所涉及的预处理方法主要包括基线校正、噪声滤除、重叠峰分辨和数据压缩等。

2.1 基线校正

基线校正是 LIBS 技术中数据预处理的一个关键环节,它直接关系到 LIBS 分析质量的好坏,因而备受广大学者关注。LIBS 分析的是待测原子或者离子的发射光谱。由于黑体辐射、韧致辐射、负荷辐射^[23] 以及一些分子辐射的存在,使得其存在较大的连续背景干扰并导致基线漂移,进而对特征谱线的准确提取产生严重的影响,因此必须将背景光谱强度予以扣除实现基线校正。目前,研究者主要采用多项式拟合法实现 LIBS 的基线校正。Lazic 等^[24] 采用剔除谱线强度极弱或者背景噪声太强光谱数据的方法进行基线校正。李捷等^[25] 通过拟合特征谱线左右两端一个较窄波段内背景连续谱线的方法实现背景扣除。Gornushkin 等^[26] 将 LIBS 光谱分成 N 段,然后通过基于局部最小值的高阶多项式函数拟合连续背景。在此基础上,Sun 等^[27] 通过方差比设定阈值去扣除所有最小值,减少运算时间。若阈值过拟合,则产生的背景强度就会过拟合;反之亦然。Yaroshchuk 等^[28] 提出了基于 Model-free 算法实现连续背景的自动校正,该法无需输入变量且窗口大小不依赖光谱的复杂程度和类型,比较稳健而且易于实现自动化。由于 LIBS 的谱线展宽^[29] 主要由 Doppler 展宽(Gaussian 线型)、自然展宽与碰撞展宽(Lorentzian 线型)组成,其展宽轮廓为 Gaussian 与 Lorentzian 函数的卷积,以电子与离子碰撞导致的 Stark 展宽占主导作用。刘立拓等^[30] 利用 Levenberg-Marquardt(L-M) 算法对 Lorentzian 理论模型中的待定参数进行优化估值,扣除了光谱信号中的连续背景,校正和还原了峰位和峰强。

小波变换是一种新兴的信号处理方法,但它在 LIBS 信号背景扣除方面还没有开展系统的研究工作。Yuan 等^[31] 用小波系数代替重构后光谱信号,进行小波变换的 LIBS 背景扣除,提高了煤中高含量碳的定量分析精度。Zou 等^[32] 通过优化小波函数、分解层数和比例因子等实现 LIBS 信号的背景扣除。

2.2 滤噪

噪声是 LIBS 信号的主要组成部分,它的存在会对定量分析产生严重的影响。随机噪声可以通过平滑进行滤除,但对于一些特性接近有效信号的噪声,需要采用有效的方法进行滤除。常用的滤噪方法有 FFT、Savitzky-Golay、Boxcar 及现在被广泛关注的小波变换。与传统滤噪法相比,小波变换具有方法简单、计算时间短、占用内存少、失真度小、重现性好等优点,可以大大提高信号的信噪比且能够克服传统傅立叶变换无法描述信号局部频率特征的缺点,成为分析非平稳信号的有力工具。Schlenke 等^[33] 提出了基于自适应可变阈值的固定小波变换,实现了 LIBS 的噪声抑制和信号保留。Zhang 等^[34] 建立了小波滤噪的双阈值优化模型,并验证了该模型对合成和真实 LIBS 信号的滤噪效果。

2.3 重叠峰解析

重叠峰解析是目前分析化学尤其是光谱分析面临的难题之一,人们曾尝试通过提高光谱仪与探测器的分辨能力解决这一问题,但硬件设备代价较大且能达到的分辨效果也是有限的。通过数学解析的方法进行分解是解决光谱重叠分解的途径之一,常见的重叠峰解析方法有:曲线拟合法^[35]、小波变换法^[36]、Fourier 去卷积法^[37]、神经网络法^[38] 等。曲线拟合法通常用来获得定量的信息,但重叠峰的个数和相对强度都需要靠假设得来,因而可能引入严重的误差。Zhang 等^[39] 借助分数微分理论对曲线拟合的初始值(重叠峰的个数和相对强度)进行假设,将 Levenberg-Marquart 方法应用于曲线拟合,实现 LIBS 信号的重叠峰解析。化学计量学用于重叠峰解析就是借助某些数学或统计学方法把依赖化学方法和仪器未能完全分离的复合量测信号分解成几个单独组分的信号,实现从重叠谱中获取每个组分的相关信息。Body 等^[40] 利用去卷积的方法对 LIBS 多元素测量光谱中的光谱重叠干扰进行解析,取得了良好效果。但对于有较强噪声干扰的谱线分解,噪声的能量在去卷积过程中常被放大。LIBS 信号的重叠峰解析虽然取得较好的成果,但是还有一些问题亟需解决,比如重叠峰解析前后谱峰位置发生偏移,如何识别重叠峰等。

2.4 数据压缩

LIBS 技术一般使用较高频率的脉冲激光激发样品并产生大量的光谱数据。而 LIBS 光谱分析中,需要提取的只是光谱信号的突变部分,即特征峰的波长与强度信息等。因此在保证光谱基本特征不变前提下,尽可能多地剔除无关紧要的信息,大量压缩光谱数据,为数据的存储和后续处理带来方便。光谱数据压缩中常用的方法有加大取样间隔、主成分分析^[41]、傅里叶变换^[42]和小波变换等。小波变换由于其局部分辨能力高、压缩率大与失真度低,已被广泛应用于光谱谱图处理如色谱、红外和紫外可见等光谱分析^[43-45]。而 LIBS 分析以特征峰为对象,其频率特性相对稳定,变化频率介于系统随机误差变化频率与连续背景光谱频率之间,因此局部分辨能力强的小波变换也适用于 LIBS 光谱数据压缩处理研究。蒋梅城等^[46]采用小波变换对常用燃煤的 LIBS 光谱数据进行数据压缩,能很好的恢复原始光谱,特别是能够准确地反映特征峰的位置和强度,同时不影响特征峰的轮廓信息。同时考察了 db 小波阶数、分解层数和阈值化方法等参数对压缩效果的影响,以压缩分数、恢复分数、谱线强度相对误差为衡量准则,评价了压缩的效果。

3 定性分析

LIBS 采集到的是复杂发射光谱,其中蕴含着丰富的化学信息和相关信息,不同种类物质的 LIBS 光谱都存在差异,通过化学计量学方法可将差异信息提取出来并进行区分和判别物质所属种类,即化学模式识别。化学模式识别是一种可以揭示隐含在 LIBS 光谱数据内部规律的多元分析技术,其目的是找出不同种类物质光谱的差异信息,从而实现对未知样品的归属和分类。化学模式识别方法主要分为有监督模式识别和无监督模式识别。

3.1 有监督的模式识别

有监督的模式识别的基本思路是用已知类别的样本作为训练集,得到判别模型,对未知试样进行归属判别。常见的有监督模式识别方法有偏最小二乘判别分析(Partial least squares discriminate analysis, PLS-DA)^[47],簇类独立软模式识别(Soft independent modeling of class analogy, SIMCA)^[48]、K-最近邻法(K-Nearest neighbor, KNN)^[49],支持向量机(Support vector machines, SVM)^[50]、人工神经网络(Artificial neural network, ANN)^[51]、随机森林(Random forest, RF)^[52]等。

3.1.1 偏最小二乘判别分析 PLS-DA 方法是基于偏最小二乘法的一种判别分析方法,它利用训练样本的自变量矩阵 X 和分类变量 Y 建立训练模型,根据待分类样本的偏最小二乘预测值判断样本所属类别。它是一种与 LIBS 技术结合应用较广泛的化学计量学方法。生物医学方面,PLS-DA 用于异构肾结石的识别,有助于了解肾结石的形成以及肾结石疾病的质量与预防^[53];基于 PLS-DA 与判别因子法的多个不同独立变量模型用于活细菌标本的 LIBS 谱图灵敏和特定属类识别,通过向下选择独立变量法分别构建基于谱线总强度、单一强度比、多个强度比的 3 个模型,PLS-DA 能够更加有效地识别相似度小的细菌属^[54]。国家安全方面,PLS-DA 用于两个模拟神经毒剂(甲基膦酸二甲酯(DMMP)和二异丙基膦酸甲酯(DIMP))和 3 个常见推进剂(肼,一甲基肼(MMH)和不对称二甲基肼(UDMH))的识别,有助于化学战剂和推进剂的识别^[55]。太空探测方面,在仿真火星大气条件下,PLS-DA 与 SIMCA 用于 8 种常见盐、纯盐溶液以及盐冻溶液的识别,PCA 主要用于提取特征变量,PLS-DA 所建的模型优于 SIMCA,可以用于未来火星探测研究^[56]。Ollila 等^[57]将 PLS-DA 用于分析火星表面的化学物质,通过改进的 PLS-DA 新算法来建立流动预测模型,其正确分类率为 95%(传统算法的为 63%)。考古方面,LIBS 已成功地用于文化遗产研究,比如颜料与石头的识别以及考古金属的表征。PLS-DA 已应用于壁画颜料的识别,采用多变量方法能够提高壁画的识别率^[58]。军事安全方面,de Lucia 等^[59]采用 PLS-DA 方法对爆炸残留物进行判别分析,在氩气气氛下以铝为基质建立训练集,其正确分类率为 95% 以上。Martelli 等^[60]采用 PLS-DA 成功地判别出小麦组织(果皮、种皮、糊粉层等),并根据烧灼程度得出了组织凝聚力。

3.1.2 簇类独立软模式识别 SIMCA 方法是一种基于主成分分析建立的模式识别方法,其基本思路是先利用主成分分析的显示结果得到一个样本分类基本印象,然后对各类样本建立相应的类模型,进一步用这些类模型对未知进行判别分析分类。该方法在建立模型时没有考虑其它类,当类间差异远大于

类内差异时,能够很好的分类。其局限性就是类间差异和类内差异相差不大时会产生非优化鉴别模型。Silva 等^[61]采用 SIMCA 方法结合 LIBS 技术测定犯罪嫌疑人手中射击残留物的钡和铅。选用没有洗手和认真洗过手的射击者手背后的粘性磁带试样作为训练集建立模型,以验证集试样到模型中心的距离为评判标准,对预测集进行分类判别,识别率为 100%。Myakalwar 等^[62]利用 LIBS 技术探测药片的有效成分。采用 SIMCA 进行预测建模,将试样的残余方差与某一类整个方差相比较,其残余方差越小,则该试样与此类的相关性越大,以此进行分类,对生活中常见药片的识别率为 94%。Sirven 等^[63]利用 LIBS 技术结合化学计量学方法对黄饼的原产地进行分类。PCA 先对整个光谱数据进行分析,SIMCA 对每一类光谱构造模型,然后根据各类光谱方差进行预测辨识。

3.1.3 人工神经网络 神经网络是一种基于多输入多输出系统的黑箱建模工具,具有较强的适应、学习和容错能力等优势。它是通过对已知类别的训练集的拟合建模进行分类和预测。但是人工神经网络在网络训练速度、过拟合、全局寻优等问题都存在未知因素,值得进一步研究。在太空探测方面,Koujelev 等^[64]使用便携式 LIBS 仪器结合 ANN 方法对外界环境的大理石、土壤、岩石等材料进行判别,提出了粗略辨别、精细辨别和“根据需要学习”3 种模式,为火星探测奠定了理论基础。在生物医学方面,Manzoor 等^[65]将 ANN 与 LIBS 结合用于医院细菌感染同类菌株的识别。在地质方面,Vítková 等^[66]采用 ANN 结合遥测 LIBS 技术对生物矿物快速判别分析。为避免出现过拟合,先对数据进行 PCA 降维处理,最终得到了较好的预测结果。在工业方面,ANN 结合 LIBS 用于聚合物、聚氯乙烯(PVC)以及常见塑料的分类识别,有助于废塑料产品的回收和循环使用^[67]。在考古方面,Ramil 等^[68]使用 ANN 结合 LIBS 技术对赭色粘土陶器碎片自动进行判别分析,考察了两种不同的 ANN 输入结构(整个光谱和最大强度光谱)对识别结果影响,其正确识别率均高于 90%,但整个光谱作为输入变量具有更高识别率(100%)。

3.1.4 支持向量机 支持向量机是非常适合小样本学习的优秀算法。它以核函数为基础,以结构风险最小化为建模目标,采用优化算法训练得到的一个具有最大边界的模型,提高模型的泛化能力。Liang 等^[69]将 LIBS 结合 SVM 实现不同牌号钢材的快速识别,并在传统 SVM 基础上提出了一种改进的组合模型(多对多)算法,在充分发挥传统多分类方法优势的同时,避免了无用信息对识别过程的干扰,显著提高了预测准确率。Xu 等^[70]提出了一种基于模型集群分析的变量选择方法提高 SVM 的钢铁分类准确率,该法基于蒙特卡洛有放回的重采样技术最大限度的提取有效信息,从而显著提高了 SVM 的分类效率,具有更好的普遍性和稳定性。Dingari 等^[71]将最小二乘-支持向量机(LS-SVM)对于非处方药进行判别分析,并对比了 PLS-DA, SIMCA 等传统线性模式识别方法,LS-SVM 能够降低样品的消融和等离子发射过程中自吸收对分析的影响,使辨识准确率提高。当训练集不纳入预测集时,LS-SVM 的识别率对比 PLS-DA, SIMCA 有明显提高。Cisewski 等^[72]建立了已知元素含量炭疽杆菌孢子的 SVM 分类模型,小波变换用于 LIBS 光谱的阈值或标准化,然后通过已知元素组成孢子的线性模型方法进行降维,构建的 SVM 模型成功地用于可疑粉末中是否含有炭疽杆菌孢子的辨别。

3.1.5 K-最近邻法 K 最近邻法是一种适用于线性不可分情况的多类化合物模式识别方法。它计算待测样本到距其最近的 K 个已知样本的距离,根据距不同种类的已知样本的远近判断待测样本归属于哪一类样本。Godoi 等^[73]对比了 KNN, PLS-DA 和 SIMCA 等多种模式识别方法,并结合 LIBS 技术检测塑料玩具中的重金属是否超标。结果表明,KNN 方法是最好的校正模型,它逐一计算了待测样品与训练集各个样品的欧氏距离,并采用投票表决的方法进行判别,可以直接用于市场上玩具的检测。

3.1.6 随机森林 随机森林(RF)是一种统计学习理论,它是利用 Bootstrap 重抽样方法从原始样本中抽取多个样本,对每个 Bootstrap 样本进行决策树建模,然后组合多棵决策树的预测,通过投票得出最终预测结果。大量的理论和实证研究都证明了 RF 具有很高的预测准确率,对异常值和噪声具有很好的容忍度,且不容易出现过拟合。可以说,RF 是一种自然的非线性建模工具,是目前数据挖掘、生物信息学的热门的前沿研究领域之一。Remus 等^[74]通过对比 RF 与 PLS-DA 对比研究不同材料(4 个岩石样品和 1 个墨水样品)的分类和识别,RF 的分类效率(100%)高于 PLS。Sheng 等^[52]通过对比 SVM 和 RF 研究 10 种铁矿石样品的分类,结果表明,RF 具有更好的分类效果。

3.2 无监督模式识别

无监督的模式识别的基本思想是在多维空间中同类化合物靠近,而不同类化合物之间的距离大,从而达到分类的目的。常用的无监督模式识别方法是主成分分析(Principal component analysis, PCA)^[74]。PCA是通过降维用较少的主成分解释数据中的大部分信息,对数据进行主成分分析后,把前二或前三个主成分进行空间投影作图,即可在图中经不同类别的数据进行初步分类。Bousquet等^[75]采用PCA将220个光谱参量数据压缩到68个,取其中的主成分投影,实现判别分析。Unnikrishnan等^[76]采用PCA成功地识别了4种常见塑料(聚对苯二甲酸乙二酯(PET)、高密度聚乙烯(PE)、聚丙烯(PP)和聚苯乙烯(PS)),通过马氏距离和光谱残差等参数的优化提高PCA分类模型的准确度。

此外,K-平均法、线性判别分析、系统聚类分析等^[77~79]化学模式识别方法也被广泛应用于LIBS光谱数据处理和对未知样品判别分析。由于LIBS技术的无需样品预处理、分析样品速度快、远程分析和全元素分析等优势,在近几年发展迅速,LIBS与化学模式识别方法的结合也日益密切。目前,两者结合已经被广泛的应用于药品^[80]、地质勘探^[81]、爆炸残留物^[82]、文物保护^[83]等领域的分析。

4 定量分析

目前,LIBS的定量分析方法主要分为两类:基于定标曲线的定标方法和基于自由定标的非定标曲线方法^[84]。传统的定量分析方法一般采用一元回归分析方法建立元素浓度和特征谱线强度之间的关系式。对于复杂的基体样品,建立一元回归关系式往往得不到理想的结果。因为在现代光谱分析技术中,获取和处理的光谱信号由一系列数据组成,是一个矢量数据,其变量间存在相互依赖关系。这时需要考虑多个自变量的回归分析问题。在数据信息大幅度增加的条件下,获得更多对定量分析有用的信息。多变量分析方法可以充分利用LIBS光谱信息,提高定量分析的准确度。LIBS定量分析涉及的多变量校正方法有偏最小二乘(Partial least Squares, PLS)^[85]、多元线性回归(Multiple linear regression, MLR)^[86]、人工神经网络^[87]、支持向量机^[88]和随机森林^[89]等。

4.1 偏最小二乘法

偏最小二乘法是一种常用的多元统计分析方法,已成功地应用于分析化学的许多领域,如紫外光谱,气相色谱等。潜变量是PLS方法的重要参数之一,其个数一般少于原自变量的个数,因此PLS特别适用于自变量个数比试样个数多的情况。PLS本质上是一种基于特征变量的回归方法,同时进行测量矩阵X和响应矩阵Y的正交分解。PLS采用校准试样组确定参数,且在计算过程中无矩阵求逆运算,计算误差小、准确度高、运算速度快,具有较好的预测能力。

Clegg等^[90]利用PCA判别分析了18种性质各异的火成岩和高度变质的岩石,并用PLS建立定标模型,得到了良好的定量分析结果。张旭等^[91]采用五点平滑和中心化法对赣南脐橙的LIBS光谱进行了预处理,优化了赣南脐橙中Cd的PLS回归模型。Yaroshchuk等^[92]通过光谱归一化、异常值剔除和连续背景校正3种预处理方法提高铁矿石样本烧失量的PLSR定量分析性能。Wang等^[93]建立了一种基于黄铜合金样品中Pb和Zn不同谱线强度信息的多变量主成分非线性PLS模型,在PLS模型中引入非线性因素降低自吸收等非线性干扰对分析结果的影响,并与传统的线性PLS模型进行对比。结果表明,非线性PLS模型与传统的PLS模型具有相同的校正效果($R^2=0.999$),但非线性PLS模型分析结果优于传统的PLS模型。Andrade等^[94]分别研究了传统单变量校正和PLS方法定量分析黄铜样品中的Cu, Zn, Sn, Pb和Fe, PLS表现出较好的定量分析性能。

4.2 多元线性回归

多元回归方法是一种常规的定量分析方法,具有良好的统计特性,采用最小二乘法进行回归,同时可以充分利用光谱数据中的信息量提高其回归结果的准确度。Ayyalasomayajula等^[95]分别采用单一线性回归和多元线性回归方法建立的校正模型确定未知土壤样品的总碳量,多元线性回归建立的校正模型优于单一线性回归。陈添兵等^[96]建立了利用LIBS技术分析脐橙中Pb含量的多元线性回归模型,选取多个谱线强度作为自变量,建立了多元线性回归模型,并用方差分析和回归统计验证了该模型的可行性,并且得到了更好的定量分析结果。这说明多变量回归方法能比较充分地利用光谱中的信息,降低

了基体效应的影响,从而提高 LIBS 定量分析的准确度。多元线性回归的输入矩阵存在噪声,导致过度拟合,某种程度上降低了模型的预测能力。而支持向量机回归和神经网络回归等机器学习方法具有一定的容错能力,可以克服噪声的影响,减小预测结果的波动性。

4.3 人工神经网络法

人工神经网络利用神经网络自组织自学习的分析能力通过大量标准样本的训练学习之后构建一个合理可靠的网络模型,再利用这个模型对待测样品进行预测分析。神经网络以分析元素的特征谱线数据为网络的输入,通过运算输出分析元素含量或其种类信息等,Ferreira 等^[67]采用神经网络与 LIBS 技术结合对土壤中的 Cu 进行了分析,通过配置标准样品进行了模型的训练、验证与测试,得到 Cu 的检出限可达 2.3 mg/dm^3 。D'Andrea 等^[68]构建了多层向前神经网络模型定量分析铜合金的主量元素,向前特征选择用于减小输入变量(谱线强度)的个数。Anderson 等^[69]研究了 PLS、多层感知器 ANN 和级联相关 ANN 方法对 195 个岩板样品中主要成分的定量分析模型。人工神经网络存在一定的局限性,受网络结构复杂性和样本复杂性的影响较大,容易出现“过学习”或低泛化能力。由于实验中的 LIBS 数据属于小样本,易造成网络训练不充分,影响其预测精度。

4.4 支持向量机

支持向量回归是 SVM 的一种方法,也是一种基于统计学习理论的机器学习方法。该方法建模是将低维非线性输入映射到高维线性输出,并在这个空间进行线性回归,模型相对比较简单。支持向量机回归基于结构风险最小化,而不是经验风险最小化,可以根据有限样本信息在模型的复杂性和学习能力之间寻求最佳折中,在保证泛化能力的前提下,达到最优学习效果,使其泛化能力明显优于 BP 神经网络。支持向量机回归适用于小样品,可以融合先验知识,通过核函数和小波等先进算法相结合,已被广泛应用于各种领域。王春龙等^[70]建立水体重金属 Ni 的 MLR、ANN 和 SVR 3 种回归模型,支持向量机回归模型的定量分析效果优于其它两种方法。Zhang 等^[71]分别采用 PLS 和 SVM 方法构建了炉渣主要成分的定量分析回归模型,结果表明,构建的 SVM 定量分析模型性能优于 PLS。

4.5 随机森林

随机森林也是一种基于递归回归树的回归算法。它是以决策树为基础的集成学习模型,它包含多个有 Bagging 集成学习技术训练得到的决策树,当输入待回归的样本时,最终的回归结果由单个决策树的输出结果投票决定。它克服了决策树过拟合问题,对噪声和异常值有较好的容忍性,对高维数据分类问题具有良好的可扩展性和并行性。此外,随机森林是由数据驱动的一种非参数回归方法,只需通过对给定样本的学习训练分类规则,并不需要分类的先验知识。Zhang 等^[72]通过对比 PLS、SVM 和 RF 3 种建模方法定量分析钢铁样品中 5 种元素(Si, Mn, Cr, Ni 和 Cu)的同时分析,RF 定量分析模型明显优于 PLS 和 SVM。

5 展 望

以上综述了化学计量学方法在 LIBS 分析的数据预处理、定性和定量分析三方面的研究进展,化学计量学的引入能够有效进行光谱预处理和最大程度地提取有用信息(建模所需的输入变量),很大程度上提高了 LIBS 技术定性定量分析的准确度。尽管早期化学计量学只是用于解决 LIBS 分析的多变量同时分析问题,然而,近几年,LIBS 技术结合化学计量学方法在模式识别和分类分析方面逐步增多。化学计量学在 LIBS 中的应用处于起步阶段,然而,它在 LIBS 技术中复杂材料快速、工业分析领域已经体现出巨大潜力。LIBS 数据的化学计量学分析的未来进展无疑是继续提高 LIBS 技术在各个应用领域的性能。LIBS 技术目前仍存在一些理论问题有待解决,如光谱基体效应的扣除、重叠峰的分辨以及自吸收效应校正等问题,这些问题对化学计量学方法提出了严峻的挑战,需要科研人员的进一步探索和研究。

References

- 1 Kearnton B, Mattley Y. *Nat. Photonics.*, 2008, 2: 537-540
- 2 Cremers D A, Radziemski L J. *Handbook of Laser Induced Breakdown Spectroscopy*, Wilkey & Sons Ltd Press, 2006: 1-18

- 3 Singh J P, Thakur S N. *Application of Model Diagnostic Methods to Environmental Improvement*, 1st ed., Amsterdam: Elsevier Press, 2007: 42–43
- 4 Miziolek A W, Palleschi V, Schechter I. *Laser-Induced Breakdown Spectroscopy (LIBS): Fundamentals and Applications*, 1st ed., New York: Cambridge University Press, 2006: 1–39
- 5 Aragon C, Aguilera J A. *Spectrochim. Acta B*, 2008, 63: 893–916
- 6 Mainman T H. *Nature*, 1960, 187(4736): 493–494
- 7 Brech F, Cross L. *Appl. Spectrosc.*, 1962, 16: 59–62
- 8 Fortes F J, Moros J, Lucena P, Cabalín L M, Laserna J J. *Anal. Chem.*, 2012, 85(2): 640–669
- 9 Unnikrishnan V K, Nayak R, Aithal K, Kartha V B, Santhosh C, Gupta G P, Suri B M. *Anal. Methods*, 2013, 5(5): 1294–1300
- 10 Tognoni E, Palleschi V, Corsi M, Cristoforetti G. *Spectrochim. Acta B*, 2007, 57: 1115 –1130
- 11 Lanza N L, Clegg S M, Wiens R C, McInroy R E, Newsom H E, Deans M D. *Appl. Opt.*, 2012, 51: B74–B82
- 12 Cortez J, Pasquini C. *Anal. Chem.*, 2013, 85: 1547–1554
- 13 Cahoon E M, Almirall J R. *Anal. Chem.*, 2012, 84: 2239–2244
- 14 Hahn D W, Omenetto N. *Appl. Spectrosc.*, 2012, 66(4): 347–419
- 15 Haddad J E, Villot-Kadri M, Ismaël A, Gallou G, Michel K, Bruyère D, Laperche V, Canioni L, Bousquet B. *Spectrochim. Acta B*, 2013, 79–80: 51–57
- 16 Gurell J, Bengtson A, Falkenstrom M, Hansson B A M. *Spectrochim. Acta B*, 2012, 74–75: 46–50
- 17 Díaz D, Hahn D W, Molina A. *Appl. Spectrosc.*, 2012, 66: 99–106
- 18 Pathak A K, Kumar R, Singh V K, Agrawal R, Rai S, Rai A K. *Appl. Spectrosc. Rev.*, 2012, 47: 14–40
- 19 Fortes F J, Vadillo I, Stoll H, Jiménez-Sánchez M, Moreno A, Laserna J J. *J. Anal. At. Spectrom.*, 2012, 27: 868–873
- 20 Multari R A, Cremers D A, Bostian M L. *Appl. Opt.*, 2012, 51: B57–B64
- 21 Sunku S, Gundawar M K, Myakalwar A K, Kiran P P, Tewari S P, Rao S V. *Spectrochim. Acta B*, 2013, 79–80: 31–38
- 22 Brown S D. *Anal. Chem.*, 1990, 62: 84R–101R
- 23 XIN Ren-Xuan. *Spectroscopy Analysis of Plasma Emission*. Beijing: Chemistry and Engineering Press, 2004: 160–168
辛仁轩. 等离子体发射光谱分析. 北京: 化学工业出版社, 2004: 160–168
- 24 Lazic V, Colao F, Fantoni R, Spizzicchio V. *Spectrochim. Acta B*, 2005, 60: 1002–1013
- 25 LI Jie, LU Ji-Dong, XIE Cheng-Li, LIN Zhao-Xiang, ZHANG Wen-Yan, LI Yong. *Spectrosc. Spect. Anal.*, 2008, 28(4): 736–739
李捷, 陆继东, 谢承利. 光谱学与光谱分析, 2008, 28(4): 736–739
- 26 Gornushkin I B, Eagan P E, Novikov A B, Smith B W, Winefordner J D. *Appl. Spectrosc.*, 2003, 57(2): 197–207
- 27 Sun L, Yu H. *Spectrochim. Acta B*, 2009, 64(3): 278–287
- 28 Yaroshchuk P, Eberhardt J E. *Spectrochim. Acta B*, 2014, 99: 138–149
- 29 Cremers D A, Radiemski J. *Handbook of Laser Induced Breakdown Spectroscopy*, Wiley & Sons Ltd, Chichester, England, 2006: 1–18
- 30 LIU Li-Tuo, LIU Jian-Guo, ZHAO Nan-Jing, LU Cui-Ping, CHEN Dong, SHI Huan, WANG Chun-Long, ZHANG Yu-Jun, LIU Wen-Qing. *Spectrosc. Spectral. Anal.*, 2011, 31(12): 3285–3288
刘立拓, 刘建国, 赵南京, 鲁翠萍, 陈东, 石焕, 王春龙, 张玉钧, 刘文清. 光谱学与光谱分析, 2011, 31(12): 3285–3288
- 31 Yuan T B, Wang Z, Li Z, Ni W D, Liu J M. *Anal. Chim. Acta*, 2015, 807: 29–35
- 32 Zou X H, Guo L B, Shen M, Li X Y, Hao Z Q, Zeng Q D, Lu Y F, Wang Z M, Zeng X Y. *Opt. Express.*, 2014, 22(9): 10233–10238
- 33 Schlenke J, Hildebrand L, Moros J, Laserna J J. *Anal. Chim. Acta*, 2012, 754: 8–19
- 34 Zhang B, Sun L, Yu H, Xin Y, Cong Z, J. *Anal. At. Spectrom.*, 2013, 28(12): 1884–1893
- 35 Hu Y G, Li W, Hu J M. *Spectrochim. Acta A*, 2005, 62(1/3): 16–21
- 36 Zou X Y, Mo J Y. *Anal. Chim. Acta*, 1997, 340: 115–121
- 37 Kauppinen J K, Moffatt D J, Cameron D G. *Appl. Opt.*, 1981, 20(10): 1866–1879
- 38 Stuart R G, Steven P F, Steven M G. *Chemometr Intell. Lab.*, 1993 (18): 41–67

- 39 Zhang B, Yu H, Sun L, Xin Y, Cong Z. *Appl. Spectrosc.*, **2013**, 67(9): 1087–1097
- 40 Body D, Chadwick B L. *Spectrochim. Acta B*, **2001**, 56: 725–736
- 41 Harrington P B, Isenhour T L. *Appl. Spectrosc.*, **1987**, 41(3): 447–449
- 42 Owens P M, Isenhour T L. *Anal. Chem.*, **1983**, 55(9): 1548–1553
- 43 Dinc E, Buker E. *J. AOAC Int.*, **2012**, 95(3): 751–756
- 44 Hu T, Lu Y H, Cheng C G, Sun X C. *Spectroscopy*, **2011**, 26(3): 155–165
- 45 Wang G Q, Hou Z Y, Tang Y X, Zhao J B, Sun Y A, Dong C H, Fu D X. *Anal. Chim. Acta*, **2010**, 679(1–2): 43–48
- 46 JIANG Mei-Cheng, LU Ji-Dong, YAO Shun-Chun, PAN Sheng-Hua, CHEN Kai, DONG Mei-Rong. *Spectrosc Spect Anal.*, **2010**, 30(10): 2797–2801
- 蒋梅城, 陆继东, 姚顺春, 潘圣华, 陈凯, 董美蓉. *光谱学与光谱分析*, **2010**, 30(10): 2797–2801
- 47 De Lucia Jr. F C, Gottfried J L. *Spectrochim. Acta B*, **2011**, 66(2): 122–128
- 48 Myakalwar A K, Sreedhar S, Barman I, Dingari N C, Rao S V, Kiran P P, Tewari S P, Manoj K G. *Talanta*, **2011**, 87(10): 53–59
- 49 Godoi Q, Leme F O, Trevizan L C, Filho E R P, Rufini I A, Santos Jr D, Krug F J. *Spectrochim. Acta B*, **2011**, 26(11): 138–143
- 50 Dingari N C, Barman I, Myakalwar A K, Tewari S P, Gundawar M K. *Anal. Chem.*, **2012**, 84: 2686–2694
- 51 Vítková G, Novotný K, Prokeš L, Hrdlička A, Kaiser J, Novotný J, Malina R, Prochazka D. *Spectrochim. Acta B*, **2012**, 73: 1–6
- 52 Sheng L, Zhang T, Niu G, Wang K, Tang H, Duan Y, Li H. *J. Anal. At. Spectrom.*, **2015**, 30(2): 453–458
- 53 Oztoprak B G, Gonzalez J, Yoo J, Gulecen T, Mutlu N, Russo R E, Demir A. *Appl. Spectrosc.*, **2012**, 66(11): 1353–1361
- 54 Putnam R A, Mohaidat Q I, Daabous A, Rehse S J. *Spectrochim. Acta B*, **2013**, 87: 161–167
- 55 Stearns J A, McElman S E, Dodd J A. *Appl. Opt.*, **2010**, 49(13): C8–C15
- 56 Schröder S, Pavlov S G, Rauschenbach I, Jessberger E K, Hübers H W. *Icarus*, **2013**, 223(1): 61–73
- 57 Ollila A M, Lasue J, Newsom H E, Multari R A, Wiens R C, Clegg S M. *Appl. Opt.*, **2012**, 51(7): B130–B142
- 58 Duchêne S, Detalle V, Bruder R, Sirven J B. *Curr. Anal. Chem.*, **2010**, 6(1): 60–65
- 59 de Lucia F C Jr, Gottfried J L. *Spectrochim. Acta B*, **2011**, 66(2): 122–128
- 60 Martelli M R, Brygo F, Sadoudi A. *J. Agric. Food Chem.*, **2010**, 58(12): 7126–7134
- 61 Silva M J, Cortez J, Pasquini C, Honorato R S, Paim A P S, Pimentel M F. *J. Braz. Chem. Soc.*, **2009**, 20(10): 1887–1894
- 62 Myakalwar A K, Sreedhar S, Barman I, Dingari N C, Rao S V, Kiran P P, Tewari S P, Manoj K G. *Talanta*, **2011**, 87: 53–59
- 63 Sirven J B, Pailloux A, Baye Y M, Coulon N, Alpettaz T. *J. Anal. At. Spectrom.*, **2009**, 24: 451–459
- 64 Koujelev A, Sabsabi M, Motto-Ros V, Laville S, Lui S L. *Planet Space Sci.*, **2010**, 58(4): 682–690
- 65 Manzoor S, Moncayo S, Navarro-Villoslada F, Ayala J A, Izquierdo-Hornillos R, de Villena F J, Caceres J O. *Talanta*, **2014**, 121: 65–70
- 66 Boueri M, Motto-Ros V, Lei W Q, Zheng L J, Zeng H P. *Appl. Spectrosc.*, **2011**, 65(3): 307–314
- 67 Ramil A, Lopez A J, Yanez A. *Appl. Phys. A*, **2008**, 92: 197–202
- 68 Liang L, Zhang T, Wang K, Tang H, Yang X, Zhu X, Duan Y, Li H. *Appl. Opt.*, **2014**, 53(4): 544–552
- 69 Xu L, Liang L, Zhang T, Tang H, Wang K, Li H. *Anal. Methods*, **2014**, 6(20): 8374–8379
- 70 Dingari N C, Barman I, Myakalwar A K, Tewari S P, Gundawar M K. *Anal. Chem.*, **2012**, 84: 2686–2694
- 71 Cisewski J, Snyder E, Hannig J, Oudejans L. *J. Chemometr.*, **2012**, 26(5): 143–149
- 72 Godoi Q, Leme F O, Trevizan L C, Filho E R P, Rufini I A, Santos Jr D, Krug F J. *Spectrochim. Acta B*, **2011**, 66: 138–143
- 73 Remus J, Dunsin K S. *Appl. Opt.*, **2012**, 51(7): B49–B56

- 74 Cordeiro J R, Martinez M I V, Li R W C, Cardoso A P, Nunes L C, Krug F J, Paixão T R L C, Nomura C S, Gruber J. *Int. J. Electrochem. Sci.*, **2012**, 2012: 1–5
- 75 Bousquet B, Sirven J B, Canioni L. *Spectrochim. Acta B*, **2007**, 62: 1582–1589
- 76 Unnikrishnan V K, Choudhari K S, Kulkarni S D, Nayak R, Kartha V B, Santhosh C. *RSC Adv.*, **2013**, 3(48): 25872–25880
- 77 Anderso R B, Bell J F, Wiens R C, Morris R V, Clegg S M. *Spectrochim. Acta B*, **2012**, 70: 24–32
- 78 Pontes M J C, Cortez J, Galvão R K H, Pasquini C, Araújo M C U, Coelho R M, Chiba M K, Abreu M F de, Madari B E. *Anal. Chim. Acta*, **2009**, 642(1–2): 12–18
- 79 Anzano J, Bonilla B, Montull-Ibor B, Casas-Gonzalez J. *J. Appl. Polym. Sci.*, **2011**, 121(5): 2710–2716
- 80 Makoto Y, Shuntaro A. *Pharm. Tech. Jap.*, **2012**, 28(12): 2543–2551
- 81 Lui S L, Koujelev A, *J. Anal. Atom. Spectrom.*, **2011**, 26(12): 2419–2427
- 82 Gottfried J L. *Appl. Opt.*, **2013**, 52(4): B10–B19
- 83 Gaudiuso R, Dell'Aglio M, De Pascale O, Santagata A, De Giacomo A. *Spectrochim. Acta B*, **2012**, 74–75: 38–45
- 84 Ciucci A, Corsi M, Palleschi V, Rastelli S, Salvetti A, Tognoni E. *Appl. Spectrosc.*, **1999**, 53(8): 960–964
- 85 Andrade J M, Cristoforetti G, Legnaioli S, Lorenzetti G, Palleschi V, Shaltout A A. *Spectrochim. Acta B*, **2010**, 65(8): 658–663
- 86 Ayyalasomayajula K K, Dikshit V, Yueh F Y, Singh J P, Smith L T. *Anal. Bioanal. Chem.*, **2011**, 400(10): 3315–3322
- 87 Ferreira E C, Milori D M B P, Ferreira E J. *Spectrochim. Acta B*, **2008**, 63(10): 1216–1220
- 88 Zhang T, Wu S, Dong J, Wei J, Wang K, Tang H, Li H. *J. Anal. At. Spectrom.*, **2015**, 30(2): 368–374
- 89 Zhang T, Liang L, Wang K, Tang H, Yang X, Duan Y, Li H. *J. Anal. At. Spectrom.*, **2014**, 29(12): 2323–2329
- 90 Clegg S M, Sklute E, Dyar M D, Barefield J E, Wiens R C. *Spectrochim. Acta B*, **2009**, 64: 79–88
- 91 ZHANG Xu, YAO Ming-Yin, LIU Mu-Hua. *Acta Phys. Sin.*, **2013**, 62(4): 44211(1)–44211(6)
张旭, 姚明印, 刘木华. *物理学报*, **2013**, 62(4): 44211(1)–44211(6)
- 92 Yaroshchuk P, Death D L, Spencer S J. *Appl. Spectrosc.*, **2010**, 64(12): 1335–1341
- 93 Wang Z, Feng J, Li L Z, Ni W D, Li Z. *J. Anal. At. Spectrom.*, **2011**, 26: 2175–2182
- 94 Ayyalasomayajula K K, Yu-Yueh F, Singh J P, McIntyre D L, Jain J. *Appl. Opt.*, **2012**, 51(7): B149–B154
- 95 CHEN Tian-Bing, YAO Ming-Yin, LIU Mu-Hua, LIN Yong-Zeng, LI Wen-Bing, ZHENG Mei-Lan, ZHOU Hua-Mao, *Acta Phys. Sin.*, **2014**, 63(10): 104213(1)–104213(6)
陈添兵, 姚明印, 刘木华, 林永增, 黎文兵, 郑美兰, 周华茂. *物理学报*, **2014**, 63(10): 104213(1)–104213(6)
- 96 D'Andrea E, Pagnotta S, Grifoni E, Lorenzetti G, Legnaioli S, Palleschi V, Lazzerini B. *Spectrochim. Acta B*, **2014**, 99: 52–58
- 97 Anderson R B, Morris R V, Clegg S M, Bell III J F, Wiens R C, Humphries S D, Mertzman S A, Graff T G, McInroy R. *Icarus*, **2011**, 215(2): 608–627
- 98 WANG Chun-Long, LIU Jian-Guo, ZHAO Nan-Jing, MA Ming-Jun, WANG Yin, HU Li, ZHANG Da-Hai, YU Yang, MENG De-Shuo, ZHANG Wei, LIU Jing, ZHANG Yu-Jun, LIU Wen-Qing, *Acta Phys. Sin.*, **2013**, 62(12): 125201(1)–125201(6)
王春龙, 刘建国, 赵南京, 马明俊, 王寅, 胡丽, 张大海, 余洋, 孟德硕, 章炜, 刘晶, 张玉钧, 刘文清. *物理学报*, **2013**, 62(12): 125201(1)–125201(6)

Progress of Chemometrics in Laser-induced Breakdown Spectroscopic Analysis

ZHANG Tian-Long¹, WU Shan¹, TANG Hong-Sheng¹, WANG Kang², DUAN Yi-Xiang³, LI Hua^{* 1}

¹(*Institute of Analytical Science, Northwest University, Xi'an 710069, China*)

²(*College of Science, Chang'an University, Xi'an 710064, China*)

³(*Research Center of Analytical Instrumentation, Sichuan University, Chengdu 610064, China*)

⁴(*College of Chemistry and Chemical Engineering, Xi'an Shiyou University, Xi'an 710065, China*)

Abstract Laser induced breakdown spectroscopy (LIBS) as a new type of element analytical technology with the advantages such as real-time, online, non-contact and multiple elements simultaneous analysis is a frontier analysis technique in the field of spectral analysis. However, it is still the main problem for LIBS technology to improve the accuracy of qualitative and quantitative analysis by extracting the useful information from a large number of complex LIBS data. Chemometrics is a chemical sub-discipline of multi-interdisciplinary, which has the advantages in data processing, signal analysis and pattern recognition. It can solve some complicated problems that are difficult for traditional chemical methods. In the article, we reviewed the research progress of chemometrics methods in LIBS from the spectral data pre-processing, qualitative and quantitative analysis in recent years.

Keywords Laser-induced breakdown spectroscopy; Chemometrics; Data pre-processing; Qualitative analysis; Quantitative analysis; Review

(Received 30 December 2014; accepted 14 April 2015)

This work was supported by the National Major Scientific Instruments and Equipments Special Project (No. 2011YQ030113) and the National Natural Science Foundation of China (Nos. 21175106, 21375105)