

谢晓啸, 白洋, 张九零, 等. 鞍山式磁铁矿含水量高光谱预测模型的建立[J]. 岩矿测试, 2024, 43(6): 901-913. DOI: [10.15898/j.ykcs.202409070183](https://doi.org/10.15898/j.ykcs.202409070183).

XIE Xiaoxiao, BAI Yang, ZHANG Jiuling, et al. Establishment of Hyperspectral Prediction Model of Water Content in Anshan-Type Magnetite[J]. Rock and Mineral Analysis, 2024, 43(6): 901-913. DOI: [10.15898/j.ykcs.202409070183](https://doi.org/10.15898/j.ykcs.202409070183).

鞍山式磁铁矿含水量高光谱预测模型的建立

谢晓啸¹, 白洋^{2*}, 张九零¹, 贾玉娜²

(1. 华北理工大学应急管理与安全工程学院, 河北唐山 063210;

2. 华北理工大学矿业工程学院, 河北唐山 063210)

摘要: 铁矿的高含水量会降低其可加工性, 不利于选矿、烧结、冶炼及尾矿处理等环节的顺利进行。因此, 合理控制铁矿的含水量对于提高矿业生产效率、降低能源消耗和减少原材料浪费至关重要。然而, 由于铁矿成分和性质的复杂性, 传统检测技术(如干燥失重法和电阻法)的灵敏度和准确度存在不足。基于此, 本文选取三种颗粒度的河北唐山某地区的鞍山式磁铁矿, 在不同含水量(0~40.0%)条件下测试其高光谱数据, 应用S-G平滑滤波(S-G)、多元散射校正(MSC)、标准正态变换(SNV)、二阶导数(SD)、倒数对数(LR)和包络线去除(CR)预处理数据, 深入分析了光谱特征及其与含水量的相关性。为进一步提高模型的预测能力, 采用竞争性自适应重加权法(CARS)筛选特征波段, 结合随机森林回归(RFR)、最小二乘支持向量回归(LSSVR)、粒子群算法优化最小二乘法支持向量回归(PSO-LSSVR)建立预测模型, 比较了不同磁铁矿含水量模型的预测效果, 最终筛选出最佳模型以提升选矿和冶炼过程中含水量检测的精度。结果表明: ①不同颗粒度的鞍山式磁铁矿样品在含水量变化时, 其光谱曲线变化趋势总体一致, 反射率与含水量呈负相关, 并在990nm、1440nm、1920nm附近表现出明显的吸收特征; 经过MSC和SNV预处理后的光谱数据, 与含水量的皮尔逊相关系数(r)最高可分别达到-0.950(412nm)和-0.964(421nm); ②三种模型中, PSO-LSSVR模型最稳定, 优选出颗粒度0.3~0.5mm的SNV-CARS-LSSVR模型和颗粒度0.5~2mm的MSC-CARS-PSO-LSSVR模型, 模型的预测集决定系数(R^2)分别为0.778、0.789, 预测集均方根误差(RMSEP)分别为5.45%、5.41%。与以往研究相比, 本文通过结合数据预处理、CARS特征筛选和非线性回归算法, 构建了更加稳定的鞍山式磁铁矿含水量预测模型, 为矿业生产过程中的含水量检测提供更高的精度支持。

关键词: 高光谱数据; 磁铁矿; 含水量; 回归模型; 预测精度

要点:

- 磁铁矿的光谱在990nm、1440nm和1920nm附近存在吸收特征, 可作为区分不同矿物成分的标志。
- 利用光谱变换突出光谱特征, 采用CARS减少冗余信息, 提高了模型的预测性能和稳定性。
- 建立磁铁矿含水量识别模型, 分析模型精度和误差, 优选最佳模型, 解决了传统方法在检测灵敏度和准确度方面的不足, 提升了模型在复杂矿石监测中的精确度和可靠性。

中图分类号: TB16; TB2

文献标识码: A

中国铁矿石资源丰富, 磁铁矿作为钢铁生产的关键原材料, 其含水量直接影响到矿石的加工、运输以及冶炼效率^[1-2]。高含水量不仅增加了干燥成本和能耗, 还可能降低冶炼效率和产品质量, 同时加剧

收稿日期: 2024-09-07; 修回日期: 2024-10-16; 接受日期: 2024-10-18; 网络出版日期: 2024-10-31

基金项目: 河北省教育厅基金项目(QN2020149); 唐山市科技局基金项目(21130204C)

第一作者: 谢晓啸, 硕士研究生, 研究方向为矿山安全与智能矿山。E-mail: 1556952492@qq.com。

通信作者: 白洋, 讲师, 研究方向为遥感与光谱解混。E-mail: 58610355@qq.com。

设备腐蚀并增加环境负担。因此,在磁铁矿的开采、选矿及冶炼等环节中,需要充分考虑含水量对生产过程的影响,并采取相应措施进行严格管理和控制,以确保生产效率和产品质量。然而,干燥失重法、电阻法等传统检测技术存在速度慢、准确度低等问题,导致目前难以高效地识别磁铁矿的含水量^[3]。因此,亟需加强对现代识别技术的研究,以满足矿业生产的需求。

光谱技术已被广泛应用于植物、土壤和岩石等不同领域的含水量分析,高光谱技术因具备无损、快速且精确的识别能力,近年来备受瞩目^[4]。已有学者研究发现农作物的光谱在 900nm、1300nm 波段附近存在吸收特征^[5-6];岩石的光谱吸收峰强度在 1400nm 和 1900nm 处与含水量呈正相关^[7-9]。当水分子与物质的相互作用时,通过氢键效应引起光谱中水吸收带的变化,因而在 1300~1600nm 和 1900nm 波段形成了—OH 键的吸收带特征^[10-12]。基于水分子的光谱特性,学者们将其应用于矿物含水量的识别领域,发现含水量与矿物光谱存在一定的线性关系^[13-14]。Maurais 等^[15]对尾矿的含水量演变过程进行研究,发现蒸发残留物的光谱特征反映了其表面干燥速率的变化;虞茉莉等^[8]则发现尾矿砂的含水量与 350~1200nm 波段的光谱反射率呈显著的负相关,并建立了含水量预测模型,决定系数(R^2)为 0.798,均方根误差(RMSE)为 0.077;梁业恒等^[16-17]利用水体重金属遥感模型对 350~950nm 的反射率与实测值进行对比, R^2 为 0.964;曹粤等^[18]建立的铁尾矿含水量模型验证精度的 R^2 为 0.92,实地验证 R^2 为 0.79。上述研究表明,通过分析水在特定波段的吸收光谱特性,可以有效地预测物质的含水量,然而,目前针对铁矿含水量的研究较少,且现有模型在精度方面仍有提升空间。

鞍山式铁矿的储量大、矿物组成均匀等特点使其成为冶炼高质量钢铁产品的理想原料,然而高含水量的磁铁矿在冶炼时会产生大量的水蒸气,影响冶炼效率和安全性。基于此,本文以品位为 38% 的鞍山式磁铁矿为研究对象,设计 0~40.0% 的 20 个含水量等级,测试不同含水量下的磁铁矿高光谱通过预处理突出光谱特征,采用竞争性自适应重加权法(CARS)筛选特征波段,结合随机森林回归(RFR)、最小二乘支持向量回归(LSSVR)、粒子群算法优化最小二乘法支持向量回归(PSO-LSSVR)建立含水量预测模型,探讨模型预测精度和误差,从而优选出能快速、精准识别磁铁矿含水量的最佳模型,以实现

冶炼中对磁铁矿含水量的实时监测,保障矿产行业的安全性。

1 实验与方法

1.1 样品采集与处理

采集河北唐山某地区的鞍山式磁铁矿作为样品,原始样品剔除树叶等杂物,混合均匀后,品位约为 38%,其主要成分为 Fe_3O_4 。通过球磨机对样品进行破碎,考虑实验室筛网目数和实际生产中适用性普遍的颗粒度^[19],本文利用标准筛将磁铁矿分为三个不同颗粒度等级,分别为等级 1 (0.15~0.3mm)、等级 2 (0.3~0.5mm)、等级 3 (0.5~2mm)。将烘干后的磁铁矿粉储存在黑色密封袋中。在实验前,缓慢向磁铁矿粉中加入水,确保其表面自由水完全消失,之后进行饱和度测定,计算出磁铁矿粉的饱和度趋于 35.0%。根据测定结果,设计了 0~40.0% 之间的 20 个不同的含水量等级,模拟不同湿度条件下的磁铁矿样品,以研究其光谱特征与含水量的关系。样品含水量等级列于表 1。

表 1 样品含水量的等级

Table 1 Moisture content grade of the samples.

含水量等级	含水量 (%)	含水量等级	含水量 (%)
等级 1	0	等级 11	21.0
等级 2	3.0	等级 12	24.0
等级 3	5.0	等级 13	25.0
等级 4	6.0	等级 14	28.0
等级 5	9.0	等级 15	30.0
等级 6	10.0	等级 16	31.0
等级 7	12.0	等级 17	33.0
等级 8	15.0	等级 18	35.0
等级 9	18.0	等级 19	38.0
等级 10	20.0	等级 20	40.0

1.2 磁铁矿样品含水量的光谱测试

磁铁矿样品的含水量光谱测试采用美国 Analytical Spectral Devices 公司生产的便携式地物光谱仪(FieldSpec4 型),其光谱范围为 350~2500nm。实验在黑暗避光的环境下进行,以减少外界光线对光谱测试的干扰。将配置好的矿粉样品充分摇匀后,置于黑色避光盒中。采用蒸发法每间隔一段时间后测定样品的含水量,以获取不同含水量等级的样品。在光谱测定过程中,光谱仪的镜头垂直 90°照射样品,重复采集 10 条光谱曲线,取平均值作为原始数据,实验过程如图 1 所示。由于在 350~399nm 和

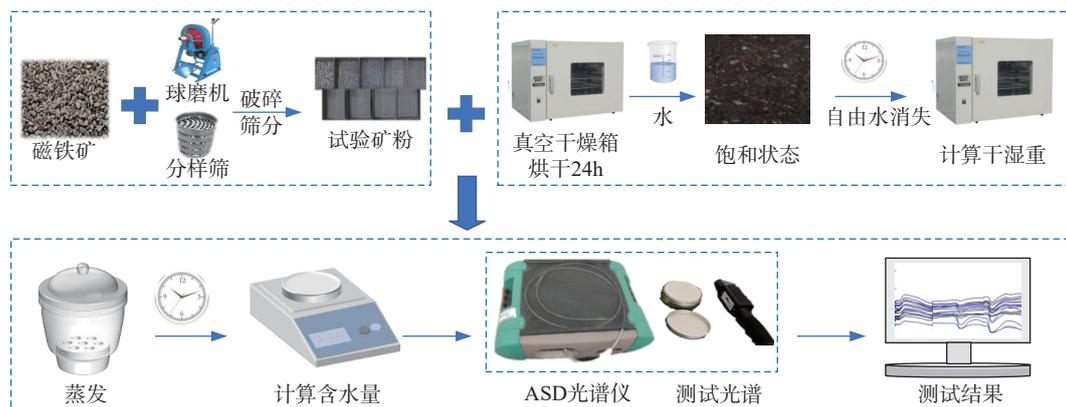


图1 磁铁矿的含水量光谱测试实验流程

Fig. 1 Experimental flow chart of spectrum measurement of water content in magnetite.

2401 ~ 2500nm 波段的数据存在较大噪声且稳定性低^[20], 因此, 仅选择 400 ~ 2400nm 波段的光谱数据进行数据分析和建模。

1.3 光谱数据预处理与皮尔逊相关性分析

在光谱采集过程中, 仪器的状态、环境温度等因素可能引入噪声, 从而影响模型的稳定性和准确性。因此, 在模型构建之前通常需要对光谱数据进行预处理^[21-22]。本文采用 6 种不同的方法对原始光谱数据进行了处理, 包括 S-G 滑滤波 (S-G)、多元散射校正 (MSC)、标准正态变换 (SNV)、二阶导数 (SD)、倒数对数 (LR)、包络线去除 (CR), 以去除噪声并改善数据质量。为评估不同预处理方法的效果, 本文使用皮尔逊 (Pearson) 相关系数分析磁铁矿含水量与预处理后的光谱反射率之间的相关性, 并在 $p=0.01$ 显著性水平上进行显著性检验, 从而量化不同光谱变换对含水量预测的影响, 为后续建模提供支持^[23-25]。相关系数 (r) 计算公式如式 (1)^[26-27] 所示。

$$r = \frac{\sum_{i=1}^n (x - \bar{x}) \times (y - \bar{y})}{\sqrt{\sum_{i=1}^n (x - \bar{x})^2 \times \sum_{i=1}^n (y - \bar{y})^2}} \quad (1)$$

式中: n 为样本数量; x 为含水率 (%); y 为波段反射率; r 为 x 和 y 的相关系数。

1.4 竞争性自适应重加权算法 (CARS) 优选特征波段

光谱数据具有高维度、结构复杂且包含大量噪声等特点, 这不仅增加了冗余信息, 还提升了计算复杂度, 从而影响模型的精度表现^[28]。因此, 通常需要对光谱数据进行特征波段筛选。本文采用竞争性自适应重加权算法 (CARS) 进行特征波段优选。

CARS 算法是基于蒙特卡洛采样法和偏最小二乘回归模型 (PLSR) 选择特征波长, 利用指数衰减函数结合自适应加权采样, 计算回归系数绝对值的权重, 保留权重较大且共线性较小的波长变量, 构建新的变量子集, 选取交叉验证均方根误差 (RMSE) 最小的 PLSR 模型对应的波段, 作为最终特征波段^[29]。

1.5 含水量预测模型构建与精度评价

目前常用的建模方法可分为线性回归和非线性回归两大类。线性回归包括多元线性回归 (MLR)、主成分回归 (PCR) 和偏最小二乘回归 (PLSR); 非线性回归方法则涵盖了人工神经网络 (ANN)、随机森林回归 (RFR) 和最小二乘支持向量回归 (LSSVR) 等^[30]。线性回归能够同时考虑物理化学值和光谱数据之间的关系, 但其基于严格的假设, 如线性关系和独立同分布的误差项, 在实际应用中这些假设通常难以完全满足, 限制了线性回归模型的适用性。而非线性回归能够处理复杂变量间的关系, 适合拟合线性回归无法捕捉的非线性模式, 具有更强的灵活性和拟合能力^[31-33]。RFR 通过构建多个回归树并综合这些树的预测结果, 能够有效地处理高维数据, 并捕捉自变量与因变量间的复杂非线性关系^[34-35]。LSSVR 作为支持向量机的改进算法, 将二次优化问题简化为线性方程组的求解, 理论框架完善且计算效率高^[36-37]。利用粒子群算法 (PSO) 优化 LSSVR 的超参数选择, PSO-LSSVR 有效地避免了模型陷入局部最优, 从而提高模型的性能^[38-39]。基于鞍山式磁铁矿晶体形态多样、嵌布粒度细等特点, 非线性回归算法能够更好地应对其复杂的矿石特性, 提高模型的预测准确性。因此, 本文采用非线性回归中的 RFR、LSSVR 和 PSO-LSSVR 建立鞍山式磁铁矿含水量的预测模型。

本文每次将200个样本按3:1的比例随机划分为校正集和预测集,训练集样本占总体样本的75%(150个),测试集为剩余的50个样本。根据校正集和预测集的决定系数(R_c^2 、 R_p^2)及均方根误差(RMSEC、RMSEP)评估模型性能,通常 R^2 越接近1, RMSE越小,模型的稳定性和预测能力越强^[40-41]。

2 结果与讨论

2.1 鞍山式磁铁矿不同含水量的光谱曲线特征

图2为鞍山式磁铁矿反射光谱在不同含水量条件下的变化趋势。通过对比图2中a、b、c可以发现,不同颗粒度等级的磁铁矿样品在光谱曲线上的变化趋势相似,反射率范围大致在0.05~0.025之间。总体上,含水量与光谱反射率整体呈负相关,当含水量为0%时,反射率较高;随着含水量的增加,反射率下降。然而,含水量的变化对光谱曲线的整体形态和趋势影响较小。在990nm、1440nm和1920nm附近出现了明显的光谱吸收特征。其中,990nm处的吸收带主要与磁铁矿中铁离子有关,而1440nm和1920nm处的吸收峰则与水分子中的—OH键振动相关^[42-43]。特别是在1440nm和1920nm处,吸收峰强度随含水量增加而增强。由此可见,含水量的变化显著影响了磁铁矿的反射率和吸收峰深度,含水量越高,反射率越低,吸收峰深度越深,这表明光谱吸收特征可用于表征磁铁矿样品的含水量。

2.2 光谱数据预处理与皮尔逊相关性分析结果

不同颗粒度等级的磁铁矿的光谱数据的变换结果如图3所示。图3中a1、b1、c1为S-G平滑后的光谱曲线,变化趋势与原始光谱基本一致,特别是在400~550nm波段,噪声的滤除效果显著。图3中a2、b2、c2为MSC变换结果,增强了1440nm和1920nm处的吸收特征。图3中a3、b3、c3和图3中a6、b6、c6分别为SNV和CR变换后的曲线,在440、990、1440和1920nm附近的光谱特征更加清晰。图3中a4、b4、c4的SD变换在440nm、990nm和1920nm的特征最为突出,但在噪声滤除效果上稍显不足。图3中a5、b5、c5的LR变换在1920nm附近显示出显著的差异,不同含水量的反射率高低区分更加明显。通过这些变换,光谱数据的特征得到了有效增强,噪声得以抑制,从而为后续建模提供了更具代表性和可靠性的光谱信息。

含水量与经过不同光谱变换后的光谱反射率的相关系数如图4所示。整体上,6种变换的相关性趋

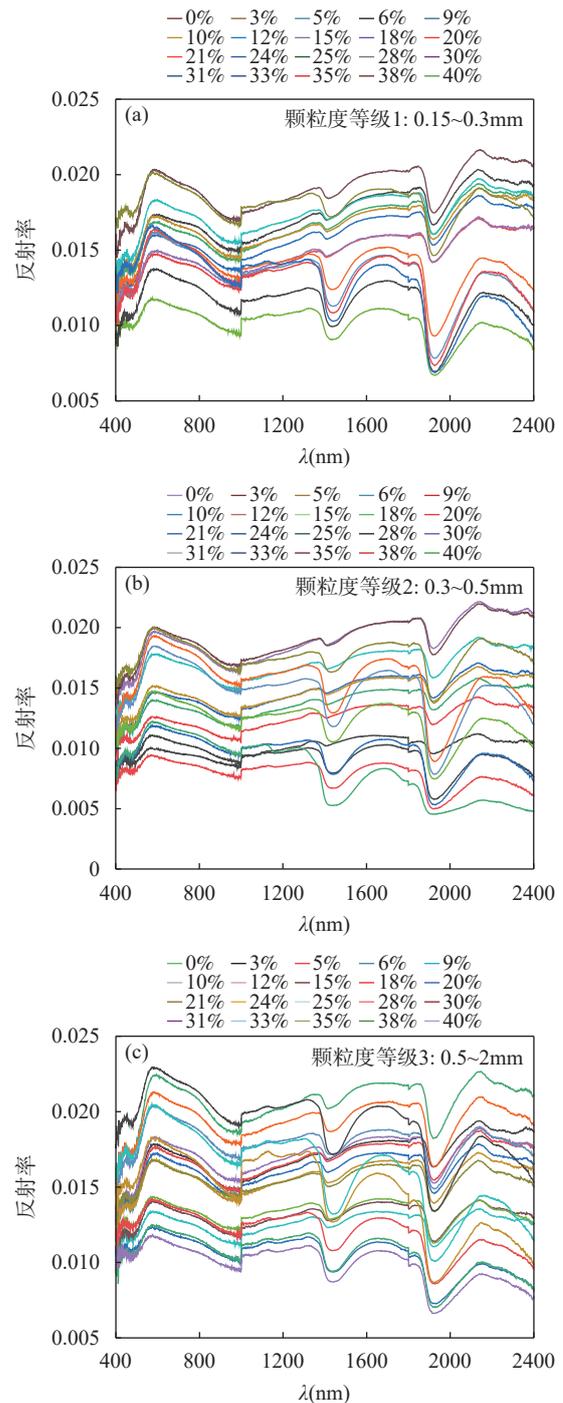


图2 各颗粒度不同含水量的鞍山式磁铁矿高光谱曲线:
(a) 颗粒度等级1; (b) 颗粒度等级2; (c) 颗粒度等级3

Fig. 2 Hyperspectral curves of Anshan magnetite with different water contents in different granularity grades:
(a) Granularity grade 1; (b) Granularity grade 2;
(c) Granularity grade 3.

势大致相同,但LR变换显示部分反向趋势。经过MSC、SNV、LR、CR变换后,大部分波段相关性得到显著提升, MSC和SNV变换相关系数(r)绝对值

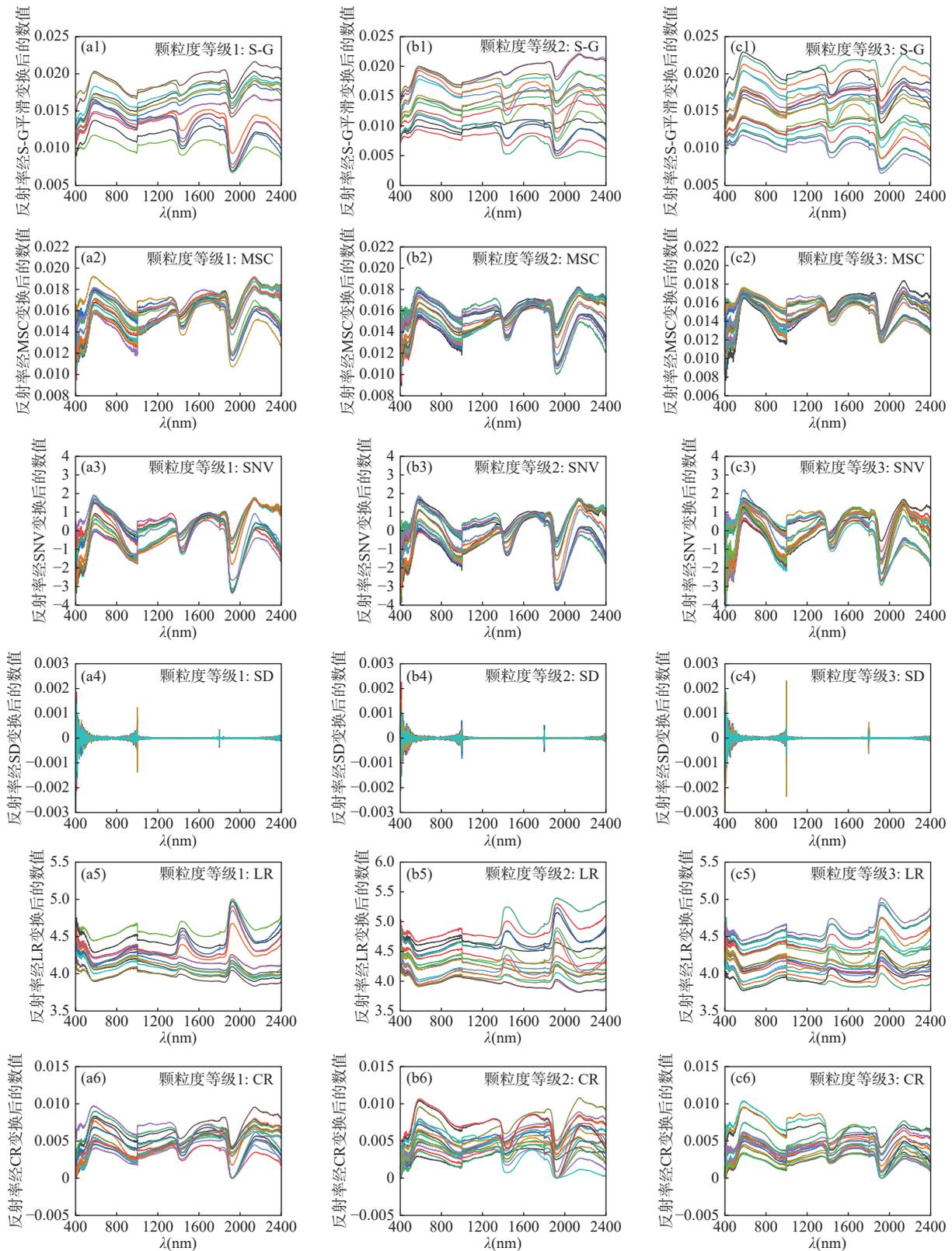


图3 不同颗粒度磁铁矿的光谱变换结果: (a1、b1、c1) S-G 变换; (a2、b2、c2) MSC 变换; (a3、b3、c3) SNV 变换; (a4、b4、c4) SD 变换; (a5、b5、c5) LR 变换; (a6、b6、c6) CR 变换

Fig. 3 Spectral transformation results of magnetite with different particle sizes: (a1, b1, c1) S-G transformation; (a2, b2, c2) MSC transformation; (a3, b3, c3) SNV transformation; (a4, b4, c4) SD transformation; (a5, b5, c5) LR transformation; (a6, b6, c6) CR transformation.

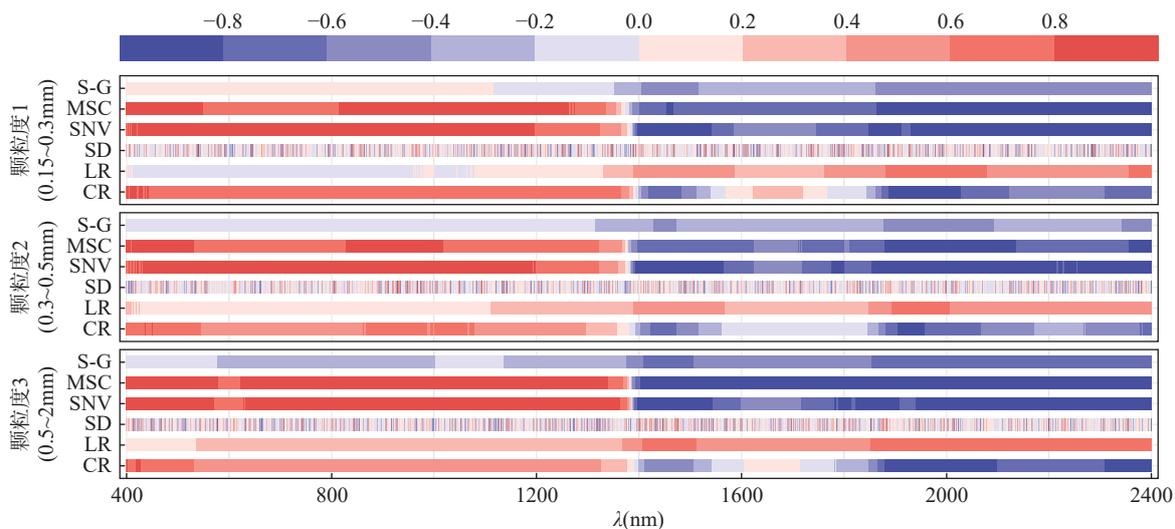


图4 含水量与不同光谱变换的相关性热图

Fig. 4 Correlation thermogram between water content and different spectral transformations.

达到 0.8 左右, S-G、LR 和 CR 变换相关系数在 0.6 左右, SD 变换相关系数在 0.5 左右。颗粒度 1 (0.15 ~ 0.3mm)、颗粒度 2(0.3 ~ 0.5mm)、颗粒度 3 (0.5 ~ 2mm) 在 MSC 变换中最大相关系数分别为 -0.950(412nm)、-0.903(435nm) 和 -0.946(438nm), 在 SNV 变换中最大相关系数分别为-0.892(457nm)、-0.902(492nm) 和-0.964(421nm); 颗粒度 3 在 CR 变换中最大相关系数为-0.901(1506nm)。

2.3 CARS 算法筛选特征波段提升数据质量

CARS 算法设置的蒙特卡洛采样次数为 50, 随着迭代次数的增加, 样本被选中的波段数量、交叉验证的均方根误差 (RMSECV) 以及各波段回归系数路径均发生变化。如图 5 所示, 在迭代初期, 由于无关变量的逐步剔除, 模型的精度逐渐提升, 表现为 RMSECV 降低。随着迭代次数增加, 剩余变量减少, 模型精度逐渐下降, RMSECV 开始增加。在经过多次迭代后, 选取了一个较优结果。如图 5c 所示, 在第 36 次回归系数路径显示该次迭代为最优迭代次数。

经过 CARS 算法筛选后 6 种光谱变换所得的特征波段位置如图 6 所示。S-G、MSC 和 LR 变换所选的波段主要集中在 440nm 和 900nm 附近, SD 变换的特征波段则集中在 440 ~ 500nm 范围, SNV 和 CR 变换筛选出的波段则主要分布在 400nm 和 1350nm 左右。通过筛选剔除冗余信息, 不仅提升了数据质量, 也为进一步分析提供了更为可靠的数据基础。

2.4 含水量预测模型建立与优选的结果

基于 CARS 算法筛选出的 6 种光谱变换特征波段, 结合 RFR、LSSVR 和 PSO-LSSVR 三种模型, 估

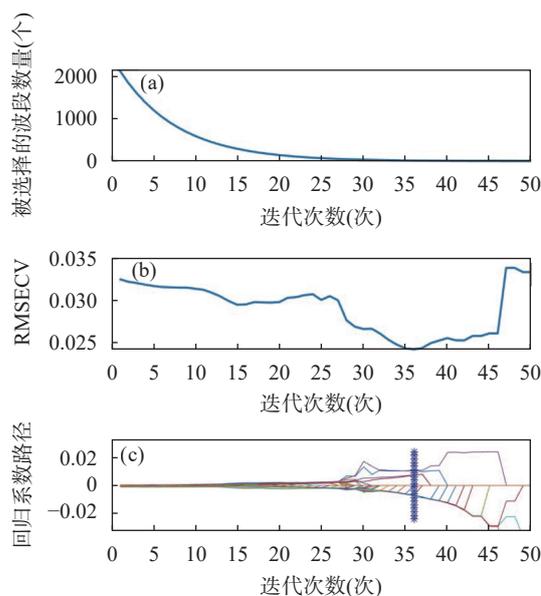


图5 CARS 算法筛选过程: (a) 被选择的波段数量变化; (b) RMSECV 变化; (c) 回归系数路径

Fig. 5 CARS algorithm screening process: (a) Changes of selected band number; (b) Changes of RMSECV; (c) Regression coefficient path.

测了三种不同颗粒度等级磁铁矿的含水量, 共产生 54 种不同的结果。图 7 为三种颗粒度等级下, 较优光谱变换模型的实测值与预测值。由图 7 可知, RFR 和 PSO-LSSVR 模型的数据点较为密集, 更接近 1 : 1 线; 而 LSSVR 模型的数据点较为离散, 出现了偏离现象, 易导致模型不稳定。从模型表现来看, PSO-LSSVR 的模型效果最佳, 训练集 R^2 达到 0.980

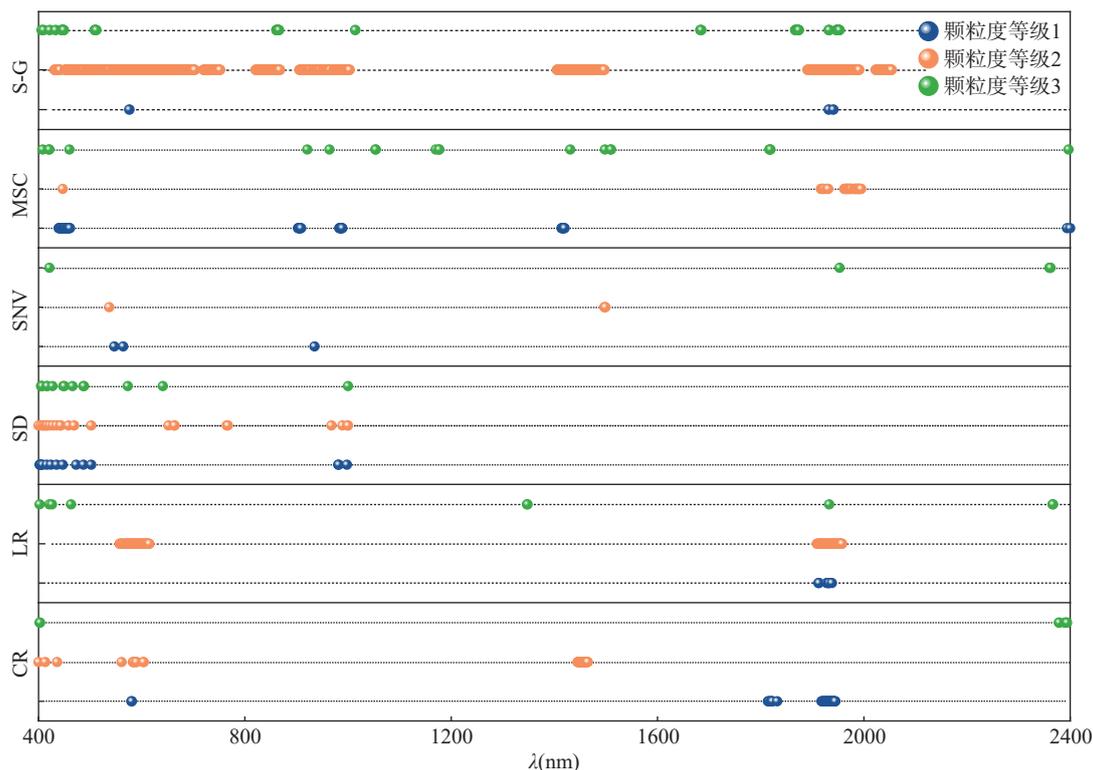


图6 CARS 算法处理三种颗粒度预处理后的光谱特征波段筛选结果

Fig. 6 CARS algorithm screening results of the spectral characteristic band after pretreatment with three kinds of granularity.

以上, RMSE 在 1.22% 以下; RFR 模型和 LSSVR 模型的效果次之。对 PSO-LSSVR 模型优选结果 (图 7 中 b, c, f, i) 进行分析, 训练集的 R^2 分别为 0.994、0.999、0.983、0.992, RMSEC 分别为 0.85%、0.03%、1.22%、0.35%; 预测集的 R^2 分别为 0.648、0.659、0.685、0.789, RMSEP 分别为 6.75%、6.53%、6.81%、5.41%。经过变换后, 模型精度得到了不同程度地提升, 其中 MSC、SNV 模型的效果佳。在 LSSVR 模型中 (图 7 中 d, e, h), 训练集 R^2 分别为 0.674、0.797、0.656, RMSEC 分别为 5.56%、6.40%、7.54%; 预测集 R^2 分别为 0.668、0.778、0.682, RMSEP 分别为 4.91%、5.45%、8.28%。在 RFR 模型中 (图 7 中 a, g), 训练集 R^2 分别为 0.959、0.910, RMSEC 分别为 2.39%、3.18%; 预测集 R^2 分别为 0.653、0.554, RMSEP 分别为 4.23%、5.46%。综合比较三种模型, 随机验证样本对鞍山式磁铁矿含水量的预测精度顺序为: PSO-LSSVR > LSSVR > RFR。

虞茉莉等^[8]建立的尾砂含水量预测模型 R^2 为 0.798, 均方根误差(RMSE)为 0.077; 曹粤等^[18]建立的铁尾矿含水量模型实地验证 R^2 为 0.79。通过上述模型效果的分析对比, 最终优选出颗粒度 2(0.3 ~ 0.5mm) 的 SNV-CARS-LSSVR 模型和颗粒

度 3(0.5 ~ 2mm) 的 MSC-CARS-PSO-LSSVR 模型, 预测集的 R^2 分别为 0.778、0.789, RMSEP 分别为 5.45%、5.41%。

3 结论

通过高光谱实验, 研究了三种颗粒度 (含水量在 0% ~ 40.0% 范围) 的鞍山式磁铁矿样品, 结果表明磁铁矿的光谱反射率与含水量总体上呈负相关, 并在 990nm、1440nm、1920nm 附近表现出明显的吸收特征, 这些吸收主要归因于 Fe^{3+} 离子和—OH 键的影响。通过 6 种预处理方法 (S-G、MSC、SNV、SD、LR、CR) 和相关性分析, 有效地突出了 990nm 和 1920nm 的吸收特性, 而 CARS 算法则进一步筛选了特征波段, 减少了冗余信息的干扰。三种模型 (RFR、LSSVR、PSO-LSSVR) 均可有效地反演磁铁矿含水量, 其中 PSO-LSSVR 模型表现最稳定, 分别优选出颗粒度等级 2(0.3 ~ 0.5mm) 和颗粒度等级 3(0.5 ~ 2mm) 的 SNV-CARS-LSSVR 模型和 MSC-CARS-PSO-LSSVR 模型, 预测集 R^2 分别为 0.778、0.789, RMSEP 分别为 5.45%、5.41%, 充分发挥了 SNV、MSC、CARS、LSSVR 和 PSO-LSSVR 模型的优势。

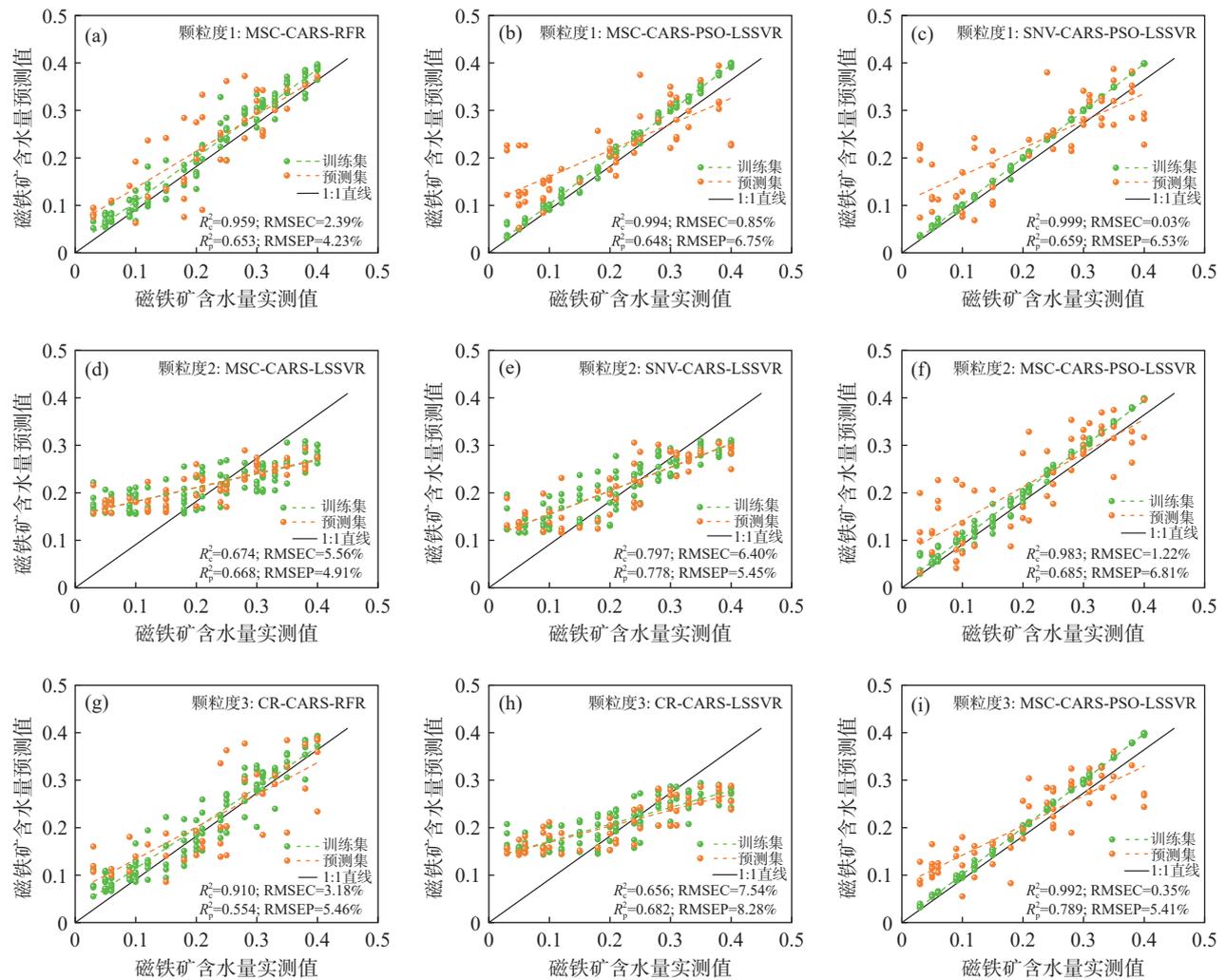


图7 三种模型预测的含水量结果：(a) 颗粒度 1 的 MSC-CARS-RFR 模型；(b) 颗粒度 1 的 MSC-CARS-PSO-LSSVR 模型；(c) 颗粒度 1 的 SNV-CARS-PSO-LSSVR 模型；(d) 颗粒度 2 的 MSC-CARS-LSSVR 模型；(e) 颗粒度 2 的 SNV-CARS-LSSVR 模型；(f) 颗粒度 2 的 MSC-CARS-PSO-LSSVR 模型；(g) 颗粒度 3 的 CR-CARS-RFR 模型；(h) 颗粒度 3 的 CR-CARS-LSSVR 模型；(i) 颗粒度 3 的 MSC-CARS-PSO-LSSVR 模型

Fig. 7 Predicting results of water content by three models: (a) MSC-CARS-RFR model with granularity 1; (b) MSC-CARS-PSO-LSSVR model with granularity 1; (c) SNV-CARS-PSO-LSSVR model with granularity 1; (d) MSC-CARS-LSSVR model with granularity 2; (e) SNV-CARS-LSSVR model with granularity 2; (f) MSC-CARS-PSO-LSSVR model with granularity 2; (g) CR-CARS-RFR model with granularity 3; (h) CR-CARS-LSSVR model with granularity 3; (i) MSC-CARS-PSO-LSSVR model with granularity 3.

本文优选的高光谱预测模型在鞍山式磁铁矿含水量预测方面表现出较高的精度, 为矿产行业在复杂环境下精确识别磁铁矿含水量提供了理论支撑。

然而, 本研究所使用的样品类型较为单一, 今后可以扩展至不同类型的样品, 以进一步提高模型的适用性和推广性。

Establishment of Hyperspectral Prediction Model of Water Content in Anshan-Type Magnetite

XIE Xiaoxiao¹, BAI Yang^{2*}, ZHANG Jiuling¹, JIA Yuna²

(1. College of Emergency Management and Safety Engineering, North China University of Science and Technology, Tangshan 063210, China;

2. College of Mining Engineering, North China University of Science and Technology, Tangshan 063210, China)

HIGHLIGHTS

- (1) The spectrum of magnetite has absorption characteristics around 990nm, 1440nm and 1920nm, which can be used to distinguish different mineral components.
- (2) Using spectral transformation to highlight spectral characteristics and CARS to reduce redundant information, the prediction performance and stability of the model are improved.
- (3) Establishing the identification model of magnetite water content, analyzing the precision and error of the model, and optimizing the best model, solves the shortcomings of traditional methods in detection sensitivity and accuracy, and improves the accuracy and reliability of the model in complex ore monitoring.

ABSTRACT: The high water content of iron ore will reduce its machinability, which is not conducive to the smooth progress of mineral processing, sintering, smelting and tailings treatment. Therefore, it is very important to control the water content of iron ore reasonably for improving mining production efficiency, reducing energy consumption and reducing waste of raw materials. However, due to the complexity of iron ore composition and properties, traditional detection techniques (such as loss on drying method and resistance method) have shortcomings in sensitivity and accuracy. Three kinds of Anshan-type magnetite from a certain area in Tangshan, Hebei Province, were selected to test hyperspectral data under different water contents (0–40.0%). Using S-G smoothing filtering (S-G), multivariate scattering correction (MSC), standard normal transformation (SNV), second derivative (SD), reciprocal logarithm (LR) and continuum removal (CR) to preprocess the data, the spectral characteristics and their correlation with water content were analyzed. In order to further improve the prediction ability of the model, the competitive adaptive reweighting method (CARS) was used to optimize the characteristic band, and a prediction model was established by combining random forest regression (RFR), least squares support vector regression (LSSVR) and particle swarm optimization least squares support vector regression (PSO-LSSVR). The prediction effects of different magnetite water content models were compared, and finally the best model was selected to improve the accuracy of water content detection in mineral processing and smelting. The results show that: (1) when the water content of Anshan-type magnetite samples with different particle sizes changes, the change trend of their spectral curves is generally consistent, and the reflectivity is negatively correlated with the water content, it shows obvious absorption characteristics around 990nm, 1440nm and 1920nm; the Pearson correlation coefficient (r) of spectral data pretreated by MSC and SNV can reach $-0.950(412\text{nm})$ and $-0.964(421\text{nm})$, respectively. (2) Among the three models, the PSO-LSSVR model is the most stable, and the SNV-CARS-LSSVR model with granularity of 0.3–0.5mm and the MSC-CARS-PSO-LSSVR model with granularity of 0.5–2mm are preferred. The prediction set determination coefficients (R^2) of the models are 0.778 and 0.789, and the root mean square error (RMSE) were 5.45% and 5.41%, respectively. Compared with previous studies, a more stable water

content prediction model of Anshan magnetite was constructed by combining data preprocessing, CARS feature screening and nonlinear regression algorithm, which provides higher precision support for water content detection in mining production.

KEY WORDS: hyperspectrum; magnetite; water content; regression model; precision of prediction

参考文献

- [1] 鲁银鹏, 孟郁苗, 黄小文, 等. 宁芜盆地玢岩型铁矿尾矿元素与矿物组成特征[J]. *岩矿测试*, 2024, 43(2): 259–269.
- Lu Y P, Meng Y M, Huang X W, et al. Element and mineral characteristics of tailings in the porphyry-type iron deposit from Ningwu Basin[J]. *Rock and Mineral Analysis*, 2024, 43(2): 259–269.
- [2] 王伟, 李勇, 樊金虎, 等. 辽河群富铁表壳岩系磁铁矿微量元素组成对古元古代铁矿成因的制约——以周家地区为例[J]. *岩石学报*, 2024, 40(10): 3103–3113.
- Wang W, Li Y, Fan J H, et al. Trace element geochemistry of the magnetite from the iron-rich supracrustal rocks of the Liaohe Group: Constraints on the genesis of oaleoproterozoic iron ores, a case study from the Zhoujia area[J]. *Acta Petrologica Sinica*, 2024, 40(10): 3103–3113.
- [3] 刘善军, 王东, 毛亚纯, 等. 智能矿山中的岩矿光谱智能感知技术与研究进展[J]. *金属矿山*, 2021(7): 1–15.
- Liu S J, Wang D, Mao Y C, et al. Intelligent spectrum sensing technology and research progress of rock and ore in intelligent mine[J]. *Metal Mine*, 2021(7): 1–15.
- [4] 黄华雨, 丁启东, 张俊华, 等. 基于地面高光谱的宁夏银北地区农田不同土层盐碱化信息反演[J/OL]. *应用生态学报* (2024-09-30). <https://doi.org/10.13287/j.1001-9332.202411.017>.
- Huang H Y, Ding Q D, Zhang J H, et al. Ground-based hyperspectral inversion of salinization and alkalization information of different soil layers in farmland in Yinbei area, Ningxia, China[J/OL]. *Chinese Journal of Applied Ecology* (2024-09-30). <https://doi.org/10.13287/j.1001-9332.202411.017>.
- [5] 王延仓, 朱玉晨, 齐焱鑫, 等. 离散小波去噪后冬小麦叶片含水量高光谱估算[J]. *光谱学与光谱分析*, 2024, 44(9): 2559–2567.
- Wang Y C, Zhu Y C, Qi Y X, et al. Study on quantitative inversion of leaf water content of winter wheat based on discrete wavelet technique[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(9): 2559–2567.
- [6] LaCour R A, Heindel J P, Head-Gordon T. Predicting the Raman spectra of liquid water with a monomer-field model[J]. *The Journal of Physical Chemistry Letters*, 2023, 14(51): 11742–11749.
- [7] 蒋航, 郭娜, 张柯凡, 等. 花岗伟晶岩型稀有金属矿床蚀变系统与矿物光谱-地球化学特征耦合性研究——以川西打枪沟矿区为例[J]. *岩石学报*, 2024, 40(1): 197–214.
- Jiang H, Guo N, Zhang K F, et al. Coupled study of alteration system and spectral-geochemical characteristics of granite-pegmatitic type rare metal deposits, associated with Daqianggou mining area in Western Sichuan Province[J]. *Acta Petrologica Sinica*, 2024, 40(1): 197–214.
- [8] 虞茉莉, 刘善军, 宋亮, 等. 不同含水量尾砂的光谱特征与遥感模型[J]. *光谱学与光谱分析*, 2019, 39(10): 3096–3101.
- Yu M L, Liu S J, Song L, et al. Spectral characteristics and remote sensing model of tailings with different water contents[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(10): 3096–3101.
- [9] 王东升, 王海龙, 张芳, 等. 砂岩的近红外光谱特征及其含水量反演[J]. *光谱学与光谱分析*, 2022, 42(11): 3368–3372.
- Wang D S, Wang H L, Zhang F, et al. Near-infrared spectral characteristics of sandstone and inversion of water content[J]. *Spectroscopy and Spectral Analysis*, 2022, 42(11): 3368–3372.
- [10] 唐振豪, 张智勇, 沙鑫, 等. 近红外水光谱组学技术及其应用近十年研究进展[J]. *分析测试学报*, 2024, 43(7): 1086–1096.
- Tang Z H, Zhang Z Y, Sha X, et al. Research progress in

- near infrared aquaphotomics technology and its applications in the past decade[J]. *Journal of Instrumental Analysis*, 2024, 43(7): 1086–1096.
- [11] 陈定芳, 吴月峰, 桂卉, 等. 水光谱组学研究现状及对中药归经理论的特殊影响[J]. *湖南中医药大学学报*, 2021, 41(12): 1986–1992.
- Chen D F, Wu Y F, Gui H, et al. Current status of hydrospectromics research and its special impact on meridian tropism theory for the Chinese materia medica[J]. *Journal of Hunan University of Chinese Medicine*, 2021, 41(12): 1986–1992.
- [12] 赵汇珍, 陈勇, 涂聪, 等. 柳江盆地髻髻山组凝灰岩地球化学与熔体包裹体水含量特征[J/OL]. *岩矿测试* (2024-09-28). <https://doi.org/10.15898/j.ykcs.202404030074>.
- Zhao H Z, Chen Y, Tu C, et al. Geochemical characteristics and water content of melt inclusions in the tuff of the Tiaojishan Formation, Liujiang Basin[J/OL]. *Rock and Mineral Analysis* (2024-09-28). <https://doi.org/10.15898/j.ykcs.202404030074>.
- [13] 孟俊贞, 杨小权, 李志萍. 基于遥感技术的地下水埋深和储变量监测评估研究进展[J/OL]. *华北水利水电大学学报(自然科学版)* (2024-09-30). <http://kns.cnki.net/kcms/detail/41.1432.TV.20240929.1327.002.html>
- Meng J Z, Yang X Q, Li Z P. Overview of the research progress of groundwater depth and storage volume assessment based on remote sensing[J/OL]. *Journal of North China University of Water Resources and Electric Power (Natural Science Edition)* (2024-09-30). <http://kns.cnki.net/kcms/detail/41.1432.TV.20240929.1327.002.html>
- [14] 高源, 孙兰香, 李翔宇, 等. 基于 LIBS 在线分析烧结矿混合料成分及校正水分影响[J]. *中国激光*, 2023, 50(19): 198–206.
- Gao Y, Sun L X, Li X Y, et al. On-line analysis of sinter mixture composition and correction of moisture influence based on laser-induced breakdown spectroscopy[J]. *Chinese Journal of Lasers*, 2023, 50(19): 198–206.
- [15] Maurais J, Orban F, Dauphinais E, et al. Monitoring moisture content and evaporation kinetics from mine slurries through albedo measurements to help predict and prevent dust emissions[J]. *Royal Society Open Science*, 2021, 8(7): 210414.
- [16] 梁业恒, 邓孺孺, 梁钰婕, 等. 重金属污染水体背景下的底质反射率光谱特征及其对离水反射率贡献影响分析[J]. *光谱学与光谱分析*, 2024, 44(1): 111–117.
- Liang Y H, Deng R R, Liang Y J, et al. Spectral characteristics of sediment reflectance under the background of heavy metal polluted water and analysis of its contribution to water-leaving reflectance[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(1): 111–117.
- [17] 梁业恒, 邓孺孺, 黄靖岚, 等. 典型重金属污染水体光谱特征分析——以广东省大宝山尾矿水为例[J]. *光谱学与光谱分析*, 2019, 39(10): 3237–3244.
- Liang Y H, Deng R R, Huang J L, et al. The spectral characteristic analysis of typical heavy metal polluted water—A case study of mine drainage in Dabao Mountain, Guangdong Province, China[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(10): 3237–3244.
- [18] 曹粤, 包妮沙, 周斌, 等. 基于实测光谱和国产高分五号高光谱卫星的铁尾矿表层含水率遥感反演方法研究[J]. *光谱学与光谱分析*, 2023, 43(4): 1225–1233.
- Cao Y, Bao N S, Zhou B, et al. Research on remote sensing inversion method of surface moisture content of iron tailings based on measured spectra and domestic Gaofen-5 hyperspectral high-resolution satellites[J]. *Spectroscopy and Spectral Analysis*, 2023, 43(4): 1225–1233.
- [19] 刘海琪, 刘善军, 丁瑞波. 颗粒度对高品位赤铁矿可见光-近红外光谱的影响研究[J]. *金属矿山*, 2022(4): 158–162.
- Liu H Q, Liu S J, Ding R B. Effect of particle size on visible-near infrared spectral of high grade hematite[J]. *Metal Mine*, 2022(4): 158–162.
- [20] 李想, 张永彬, 刘明月, 等. 滨海湿地土壤质地高光谱估测模型对比分析[J]. *光谱学与光谱分析*, 2024, 44(9): 2568–2576.
- Li X, Zhang Y B, Liu M Y, et al. Comparative analysis of hyperspectral estimation models for soil texture in coastal wetlands[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(9): 2568–2576.

- [21] 杨新艳, 李东东, 叶文清, 等. 基于激光诱导击穿光谱的标准加入法研究进展[J/OL]. 激光技术 (2024-09-30). <http://kns.cnki.net/kcms/detail/51.1125.tn.20240830.0959.002.html>.
Yang X Y, Li D D, Ye W Q, et al. Research progress of standard addition method based on laser-induced breakdown spectroscopy[J/OL]. Laser Technology (2024-09-30). <http://kns.cnki.net/kcms/detail/51.1125.tn.20240830.0959.002.html>.
- [22] Engel J, Gerretzen J, Szymańska E, et al. Breaking with trends in pre-processing[J]. *TrAC Trends in Analytical Chemistry*, 2013, 50: 96–106.
- [23] Upadhyay R, Gupta A, Mishra H N, et al. At-line quality assurance of deep-fried instant noodles using pilot scale visible-NIR spectroscopy combined with deep-learning algorithms[J]. *Food Control*, 2022, 133: 108580.
- [24] 陈蓓, 郑恩让, 郭拓. 多种光谱变量筛选算法在红参提取近红外建模中的应用[J]. *光谱学与光谱分析*, 2021, 41(8): 2443–2449.
Chen B, Zheng E R, Guo T. Application of various algorithms for spectral variable selection in NIRS modeling of Red Ginseng extraction[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(8): 2443–2449.
- [25] Li H D, Liang Y Z, Xu Q S, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. *Analytica Chimica Acta*, 2009, 648(1): 77–84.
- [26] Chapman J, Tomasello B, Carr S. Bifurcation in correlation length of the Ising model on a Toblerone lattice[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2024(9): 93214–93214.
- [27] 庞江, 张焯毓, 黄毅, 等. 铁含量对白云石拉曼光谱特征的影响[J]. *岩矿测试*, 2023, 42(4): 852–862.
Pang J, Zhang Y Y, Huang Y, et al. Effect of Fe content on Raman spectral characteristics of dolomite[J]. *Rock and Mineral Analysis*, 2023, 42(4): 852–862.
- [28] Yang D, Hu J. A detection method of oil content for maize kernels based on CARS feature selection and deep sparse autoencoder feature extraction[J]. *Industrial Crops and Products*, 2024, 222: 119464.
- [29] Bocklitz T W, Salah F S, Vogler N, et al. Pseudo-HE images derived from CARS/TPEF/SHG multimodal imaging in combination with Raman-spectroscopy as a pathological screening tool[J]. *BMC Cancer*, 2016, 16: 1–11.
- [30] 黄华丹, 邓健豪, 曹庸, 等. 基于近红外光谱技术快速检测广式酱油发酵过程中主要理化指标的含量[J/OL]. *现代食品科技* (2024-09-30). <https://doi.org/10.13982/j.mfst.1673-9078.2024.12.0716>.
Huang H D, Deng J H, Cao Y, et al. Rapid determination of main physicochemical indexes in the fermentation process of Cantonese soy sauce based on near infrared spectroscopy[J/OL]. *Modern Food Science and Technology*(2024-09-30). <https://doi.org/10.13982/j.mfst.1673-9078.2024.12.0716>.
- [31] 洗丽铎, 朱薪蓉, 卢德浩, 等. 联合运用多光谱和激光雷达技术构建的林分生物量估算模型[J]. *东北林业大学学报*, 2024, 52(8): 85–94.
Xian L H, Zhu X R, Lu D H, et al. Estimation models of forest stand biomass using combined multi-spectral and LiDAR technologies[J]. *Journal of Northeast Forestry University*, 2024, 52(8): 85–94.
- [32] 郭军, 曲亮, 邵丹, 等. 基于机器学习的地方鸡产蛋曲线拟合探索[J]. *中国畜牧兽医*, 2024, 51(8): 3428–3437.
Guo J, Qu L, Shao D, et al. Exploration of egg production curve fitting of local chickens based on machine learning[J]. *China Animal Husbandry & Veterinary Medicine*, 2024, 51(8): 3428–3437.
- [33] 李璇, 甘淑, 袁希平, 等. 洱海东岸海滨三种典型湿地植被光谱特征分析与识别建模[J]. *光谱学与光谱分析*, 2024, 44(9): 2439–2444.
Li X, Gan S, Yuan X P, et al. Spectral characteristic and identification modelling of three typical wetland vegetation along the seashore of the of the east coast of the Erhai Lake[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(9): 2439–2444.
- [34] 王巧玲, 李双成. 云南省碳排放时空演变特征及影响因素分析[J/OL]. *中国环境科学* (2024-09-30). <https://doi.org/10.19674/j.cnki.issn1000-6923.20240928.001>.
Wang Q L, Li S C. Dynamics of carbon emissions in

- Yunnan Province: Spatiotemporal characteristics and influencing factors[J/OL]. *China Environmental Science* (2024-09-30). <https://doi.org/10.19674/j.cnki.issn1000-6923.20240928.001>.
- [35] 王蕾,李斌,吴飞,等.基于改进极限学习机的电力市场实时电价预测方法[J].*电子设计工程*,2024,32(20):21–25,30.
Wang L, Li B, Wu F, et al. Real time electricity price prediction method in the electricity market based on improved extreme learning machine[J]. *Electronic Design Engineering*, 2024, 32(20): 21–25, 30.
- [36] 崔峰,何仕凤,来兴平,等.基于LSSVR与灰色理论的急倾斜巨厚煤层群开采冒落高度与时滞特征研究[J].*岩石力学与工程学报*,2024,43(4):822–837.
Cui F, He S F, Lai X P, et al. Study on collapse height and time delayed characteristics in the mining of steeply inclined extra-thick coal seam group based on LSSVR and grey theory[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2024, 43(4): 822–837.
- [37] 王小辉,李圣普.基于GBO和LSSVR的丝网印刷产品墨量转移预测模型研究[J].*数字印刷*,2022(6):79–84.
Wang X H, Li S P. Prediction model of ink transfer rate for screen prints based on garden balsam optimization and least squares support vector regression[J]. *Printing and Digital Media Technology Study*, 2022(6): 79–84.
- [38] 蔡鸣,朱光,李论,等.基于PSO-LSSVR的机器人磨抛材料去除模型[J].*组合机床与自动化加工技术*,2024(1):174–177,182.
Cai M, Zhu G, Li L, et al. Robot polishing material removal model based on PSO-LSSVR[J]. *Modular Machine Tool & Automatic Manufacturing Technique*, 2024(1): 174–177, 182.
- [39] Chen J, Hu Y, Zhu Q, et al. A novel battery health indicator and PSO-LSSVR for LiFePO₄ battery SOH estimation during constant current charging[J]. *Energy*, 2023, 282: 128782.
- [40] 李志尚,赵龙,宗洪祥,等.机器学习型分子力场在金属材料相变与变形领域的研究进展[J].*金属学报*,2024,60(10):1388–1404.
Li Z S, Zhao L, Zong H X, et al. Machine-learning force fields for metallic materials: Phase transformations and deformations[J]. *Acta Metallurgica Sinica*, 2024, 60(10): 1388–1404.
- [41] 常爱莲,黄乐,张郡铄,等.页岩气储层中甲烷运移的变分数阶导数模型研究[J].*力学季刊*,2024,45(3):825–833.
Chang A L, Huang L, Zhang J S, et al. Study of variable-order fractional derivative model for CH₄ transport in shale gas reservoirs[J]. *Chinese Quarterly of Mechanics*, 2024, 45(3): 825–833.
- [42] 李孟倩,李鸣铎,汪金花,等.铁矿粉铁品位高光谱精确估测方法研究[J].*金属矿山*,2023(3):206–213.
Li M Q, Li M D, Wang J H, et al. Study on hyperspectral accurate estimation method of iron grade for iron ore powder[J]. *Metal Mine*, 2023(3): 206–213.
- [43] 毛亚纯,文杰,曹旺,等.基于鞍山式铁矿成像光谱的融合算法研究[J].*光谱学与光谱分析*,2024,44(9):2620–2625.
Mao Y C, Wen J, Cao W, et al. Fusion algorithm research based on imaging spectrum of Anshan iron ore[J]. *Spectroscopy and Spectral Analysis*, 2024, 44(9): 2620–2625.