

基于粒子群优化随机森林算法的辉钼矿浮选精矿品位预测模型

宛鹤¹, 杨亮亮¹, 屈娟萍², 江鹏³

- 西安建筑科技大学 资源工程学院, 陕西 西安 710055;
- 奥卢大学 奥卢矿业学院, 芬兰 奥卢 FI-90014;
- 中煤科工集团南京设计研究院有限公司, 江苏 南京 210031

中图分类号: TD954; TD923 文献标识码: A 文章编号: 1001-0076(2025)03-0076-11
DOI: 10.13779/j.cnki.issn1001-0076.2025.08.026

摘要 矿物浮选精矿品位是评估浮选效果的关键技术指标之一, 其精确测定备受业内关注。然而, 现有的辉钼矿浮选精矿品位在线检测方法检测精度不高且实施成本昂贵。为此, 提出了一种基于粒子群优化随机森林算法(PSO-RF)的辉钼矿浮选精矿品位预测模型。该模型通过结合粒子群优化(PSO)算法与随机森林(RF)算法, 有效提升了模型预测精度和泛化能力。首先通过辉钼矿浮选实验, 分析了不同浮选药剂用量和磨矿细度对精矿品位的影响, 并以此构建了多输入输出的随机森林预测模型; 然后应用粒子群优化算法对随机森林模型进行参数优化, 显著提升了模型的预测精度, 在验证集上, $RMSE$ 值为 0.036 9, MAE 值为 0.024 5, R^2 值为 0.980 2。相比随机森林模型(RF), PSO-RF 模型 R^2 提升了 1.83%, $RMSE$ 值和 MAE 值分别降低了 19.43% 和 16.95%。最后通过实验对 PSO-RF 模型进行了验证, 预测值与实际值的最大相对误差在 4.28% 以内, 表现出较高的预测精度和良好的泛化能力。同时, 该模型还具备实时检测能力以及较低的实施成本, 在工业生产中具有广阔的应用潜力。

关键词 辉钼矿; 品位预测; 粒子群算法; 随机森林模型

引言

矿物浮选是一种基于矿物表面疏水性差异的选择性分离工艺。浮选精矿品位是评价浮选效果好坏的重要依据^[1]。然而, 浮选过程中矿石的物理化学性质、矿浆中矿物颗粒的粒度分布以及浮选药剂的作用等都具有高度的差异性, 且不同的操作条件(如气泡尺寸、搅拌速度、气体流量等)会交互作用, 形成复杂的非线性关系, 使得浮选性能指标难以通过具体的关系模型进行准确预测^[2]。早期矿物浮选过程中品位的检测主要是通过人工取样进行化学分析检测, 所需时间周期长且由于取样过程和分析技术的局限性, 存在一定的精度误差, 满足不了现代化生产的需要。20 世纪 50 年代, X 射线荧光光谱(XRF)技术开始被用于矿物浮选过程中浮选

品位的检测, 这种基于 X 射线管和晶体光谱仪的在线分析方法检测周期短且精度高^[3]。然而, 这种光谱仪过于复杂且昂贵, 通常在工厂的中心实验室中进行。传统的浮选精矿品位在线检测方法, 所需设备造价昂贵, 稳定性较差, 维护工作也相对困难, 检测时间周期也无法满足实时性要求^[4]。

随着人工智能技术的发展, 针对传统在线检测方法的不足, 一些学者开始研究浮选过程中浮选泡沫特征(尺寸、颜色、数量等)与浮选精矿品位间的相关性。提出了基于机器视觉技术的浮选精矿品位预测方法, 通过建立泡沫图像特征与浮选精矿品位之间的关系模型, 利用提取的泡沫图像特征对浮选精矿品位进行预测。例如, Bartolacci G 等人通过多元图像分析(MIA)法及灰度共生矩阵(GLCM)和小波变换分析(WTA)来提取锌浮选泡沫图像的

收稿日期: 2025-01-13

基金项目: 国家自然科学基金面上项目(52274271)

作者简介: 宛鹤(1982—), 男, 辽宁锦州人, 博士, 教授, 博士生导师, 主要从事资源综合利用、浮选药剂、选矿智能化等方面研究, E-mail: wanhe@xauat.edu.cn。

颜色、形态、纹理等特征来预测锌浮选品位,该方法用于锌浮选过程中取得了良好的经济效益^[5]。Ren C 等人提出了一种基于显微图像颜色特征和最小二乘支持向量回归模型(LS-SVR)的铜精矿品位间接测量的方法,以色相为输入参数的 LS-SVR 模型可以估算铜精矿品位,且具有满意的精度^[6]。Cao B 等人通过分析铝土矿不同浮选阶段的图像数据,利用多核最小二乘支持向量机和小波极值学习机建立精矿品位预测模型和残差补偿模型,最后通过集成预测精矿品位,验证结果表明集成预测模型具有良好的泛化能力^[7]。Jahedsaravani A 等人分析了硫化铜矿浮选过程中泡沫特征(泡沫大小、泡沫速度、泡沫颜色、泡沫稳定性)与浮选品位之间的关系,通过神经网络建立铜精矿品位预测模型,实验表明该模型具有较高的准确性^[8]。

基于泡沫特征的机器视觉预测方法具有实时监测的优势,能够在浮选过程中即时获取泡沫的尺寸、颜色和数量等信息,从而实现对浮选品位的动态预测。然而,该方法也存在一些显著的缺点。首先,泡沫特征与品位和回收率之间的关系复杂且非线性,不同矿石类型和工艺条件下模型的泛化能力较差。其次,光照变化、矿浆颜色波动以及泡沫的动态性会干扰图像特征的稳定性,影响预测精度。此外,高分辨率工业摄像头和高性能计算设备的高昂成本和维护难度也限制了该方法的广泛应用。

针对基于机器视觉预测方法的不足,一些学者通过研究药剂用量、矿浆 pH、矿浆浓度、原矿品位、充气量、磨矿细度等关键参数与浮选精矿品位之间的关系,提出了基于关键浮选参数的机器学习精矿品位预测模型。基于关键浮选参数的预测模型,易于实现且实施成本低,不受现场环境(如光线变化、泡沫流速等)的影响,具有良好的泛化能力和适应性。例如,Jing Z 等人以铁矿浮选中的矿浆流量、矿石品位、磨矿细度、矿浆浓度、矿浆温度和浮选流量为输入变量,浮选精矿品位为输出变量,提出了一种将混沌蚁群算法与 BP 神经网络相结合的浮选性能指标预测模型,在仿真和实验结果中表现出良好的预测性能^[9]。Wang J S 等人以给料品位、矿浆流量、给料浓度、给料粒度、药剂流量为输入变量,精矿品位为输出变量,提出了一种基于 PSO-GSA 算法的浮选过程前馈神经网络(FNN)软测量模型,仿真结果验证了该软测量模型的有效性^[10]。Gomez-Flores A 等人通过研究浮选过程中物理化学变量

(接触角、颗粒直径、气泡直径、颗粒电荷)和操作变量(捕收剂浓度、浮选时间和矿物种类及其含量与浮选品位之间的关系,并使用不同机器学习算法(MLR、KNN、DT、RF)建立预测模型,实验表明随机森林(RF)模型预测性能最优^[11]。

基于关键浮选参数的精矿品位预测方法,在一定程度上克服了基于泡沫特征方法的局限性,但在模型选择和优化方面仍存在不足。现有研究主要采用多种机器学习模型进行浮选品位的预测,常见的包括支持向量机(SVM)、神经网络和随机森林(RF)等。SVM 模型在处理小样本数据时表现出色,能够有效处理高维数据,但在面对复杂的非线性关系时,其表现略显不足。神经网络具有强大的非线性建模能力,适用于复杂的浮选过程,但其对大量数据和长时间训练的需求使得在实际应用中存在一定挑战。RF 模型具有处理非线性关系和高维数据的优势,且具备良好的泛化能力,在浮选精矿品位预测中表现出较好的预测效果。尽管随机森林模型在浮选精矿品位预测中表现优异,但其在参数优化方面存在一定的局限性。其预测性能在很大程度上依赖于决策树的数量、最大深度、最小样本分割数等参数的合理设置。传统的参数调优方法(如网格搜索、随机搜索)往往计算成本高且效率低。

粒子群优化(PSO)算法作为一种高效的全局优化算法,能够通过模拟群体智能行为,快速搜索参数空间,找到全局最优解。将粒子群优化算法应用于随机森林模型的参数优化,能够在较短时间内找到最优的决策树数量和其他关键参数设置,从而克服随机森林模型在参数选择上的不足,可以显著提升模型的预测精度和泛化能力。

在矿物浮选中,浮选药剂用量影响矿物表面化学性质和泡沫稳定性,过量或不足会降低浮选效率,且导致精矿品位不稳定。磨矿细度决定矿石颗粒的比表面积和反应活性,过细增加能耗和设备磨损,过粗则降低矿物可浮性,影响回收率。因此,本研究通过辉钼矿浮选过程中不同浮选药剂用量和磨矿细度对浮选精矿品位的影响,分析不同浮选药剂及磨矿细度间的交互作用,利用随机森林算法建立了辉钼矿浮选精矿品位预测模型。针对建立的预测模型预测精度不高等问题,进一步利用粒子群优化算法(PSO)来优化随机森林模型,以提升预测性能,优化后的模型在预测精度上有了明显提高。

1 实验与方法

1.1 辉钼矿浮选

浮选流程如图 1 所示。研究不同浮选药剂用量及磨矿细度对辉钼矿浮选效果的影响,浮选实验在实验室条件下进行,研究因素包括捕收剂用量、起泡剂用量、分散剂用量和磨矿细度。参考现场辉钼矿实际生产工况,其中浮选药剂设定 4 个水平,磨矿细度设定为 2 水平,共设计 128 组实验,每组实验辉钼矿用量为 350 g。浮选药剂用量捕收剂(煤油)为 80、100、120、140 g/t,起泡剂(一元醇)为 30、40、50、60 g/t,分散剂(硅酸钠)为 0、500、1 500、2 000 g/t;磨矿细度为-0.074 mm 含量 60%、65%。浮选设备为 XFD 实验室单槽浮选机,规格为 1 L,主轴转速为 1 992 r/min,球磨机为 PK/ZQM(BM) 系列智能球磨机,选矿用水为自来水,矿浆质量浓度为 41%,浮选时间为 4 min。同时考虑到不同矿石成分对浮选精矿品位的影响较大,因此本研究所有实验中的矿样均来源于某浮选厂同一批次矿石。为了减少其他因素干扰,确保实验结果准确可靠,每组实验进行三个时间段的取样,分别为前 30 s、30~60 s 以及 60~240 s,每组实验最终结果由三次取样的金属量累计值与三次取样产率累计值之比计算得出。每组数据记录了不同浮选药剂用量和磨矿细度对应的辉钼矿浮选精矿品位, X_1 表示捕收剂用量、 X_2 表示起泡剂用量、 X_3 表示分散剂用量、 X_4 表示磨矿细度(-0.074 mm 含量)。数据集的具体格式如表 1 所示。

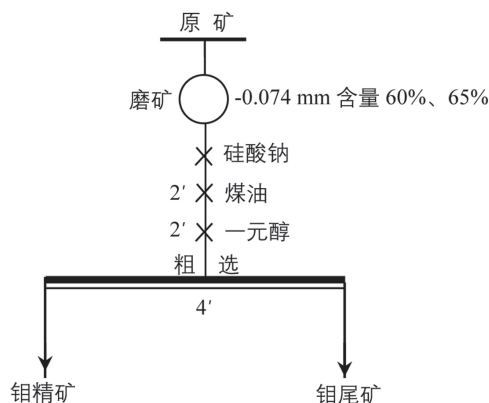


图1 浮选实验流程

Fig. 1 Flowchart of the flotation experiment

1.2 方法

1.2.1 随机森林 (RF) 算法

随机森林算法^[12]是一种集成学习方法,它通过

表 1 实验数据集

Table 1 Dataset for the experiment

实验编号	$X_1/(g \cdot t^{-1})$	$X_2/(g \cdot t^{-1})$	$X_3/(g \cdot t^{-1})$	$X_4/\%$	平均品位/%
1	80	30	0	65	1.18
2	100	30	0	65	1.22
3	120	30	0	65	1.28
4	140	30	0	65	1.66
.....					
125	80	60	2000	60	1.49
126	100	60	2000	60	1.68
127	120	60	2000	60	1.79
128	140	60	2000	60	1.52

构建多个决策树来进行预测。在训练过程中,每棵决策树都会从训练数据中随机抽取一个子集,并且在每个节点的划分时,只考虑从特征空间中随机选择的特征子集。最终的预测结果是通过对所有决策树的预测结果进行平均(对于回归任务)或投票(对于分类任务)得到的。随机森林模型的预测性能主要受决策树的数量、树的深度以及随机种子等超参数的影响。算法模型如图 2 所示。

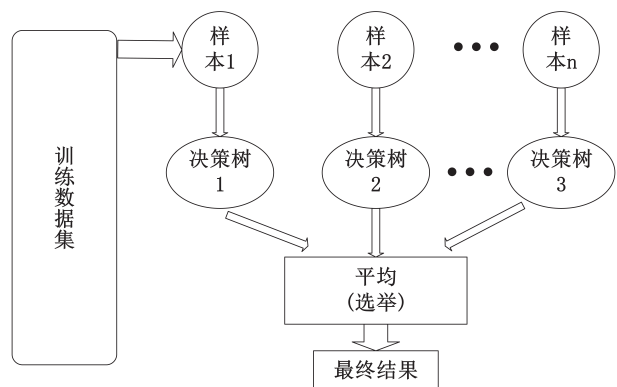


图2 随机森林算法

Fig. 2 Random forest algorithm

1.2.2 粒子群优化 (PSO) 算法

粒子群算法^[13](Particle Swarm Optimization, PSO)是一种基于群体智能的优化算法,其原理主要模拟了鸟群捕食的行为,鸟群在森林中沿着不同方向寻找食物,并且进行信息共享,经过一段时间的搜索后就可以找到森林中哪个位置的食物量最多。在粒子群算法中,每个解被称为一个粒子,每个粒子都代表搜索空间中的一个候选解,并具有位置、速度和适应度值三个属性。假设有 N 个粒子在 D 维空间中搜寻最优解,则速度更新公式可以表示为:

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_d^k - x_{id}^k) \quad (1)$$

位置更新公式表示为:

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (2)$$

式中: i —粒子序号, $i=1,2,3\cdots N$; d —粒子维度, $d=1,2,3\cdots D$; k —迭代次数; ω —惯性权重; c_1 —个体学习因子; c_2 —群体学习因子; r_1, r_2 —区间 $[0,1]$ 内的随机数, 增加搜索的随机性; v_{id}^k —粒子 i 在第 k 次迭代中第 d 维的速度向量; x_{id}^k —粒子 i 在第 k 次迭代中第 d 维的位置向量; p_{id}^k —粒子 i 在第 k 次迭代中第 d 维的历史最优位置, 即在第 k 次迭代后, 第 i 个粒子(个体)搜索得到的最优解; p_d^k —粒子 i 在第 k 次迭代中第 d 维的历史最优位置, 即在第 k 次迭代后, 整个粒子群体中的最优解。

1.2.3 PSO-RF 模型框架流程

使用粒子群优化算法来搜索最佳的随机森林模型参数($n_estimators$, max_depth , $random_state$)组合, 每个粒子代表一个随机森林参数组合。在迭代过程中, 把每次随机森林模型在训练集上的预测值与实际值之间均方误差视为近似适应函数, 均方误差值作为粒子的适应度值, 用于评估每组参数的

好坏。当达到预设的迭代次数时, 优化过程结束, 并输出找到的最佳参数组合, 使用最佳参数组合重新训练模型, 提高模型的性能和泛化能力。在 PSO 算法中, 惯性权重和学习因子通过控制粒子的速度和方向, 在平衡探索能力与收敛速度中起关键作用。在本研究中, 通过反复实验和性能评估, 选定了合理的惯性权重和学习因子, 从而增强了粒子群优化算法的全局搜索能力和收敛速度。图3展示了粒子群优化随机森林模型的算法过程。

2 实验结果与讨论

2.1 浮选参数与浮选指标间的相关性分析

通过辉钼矿浮选实验数据, 建立不同浮选药剂用量和磨矿细度与浮选品位之间的响应面模型。模型拟合的决定系数为 $R^2=0.973$, 这表明所建立的模型可以较好地解释自变量和因变量之间的关系。模型的方差分析结果见表2。 F 值越大, P 值越小, 表明建立的模型拟合效果显著, 且 P 值为 0.3628, 大于 0.05, 表明不存在不匹配因素, 因此可以用于预测 4 个选定变量与辉钼矿浮选品位之间的关系。

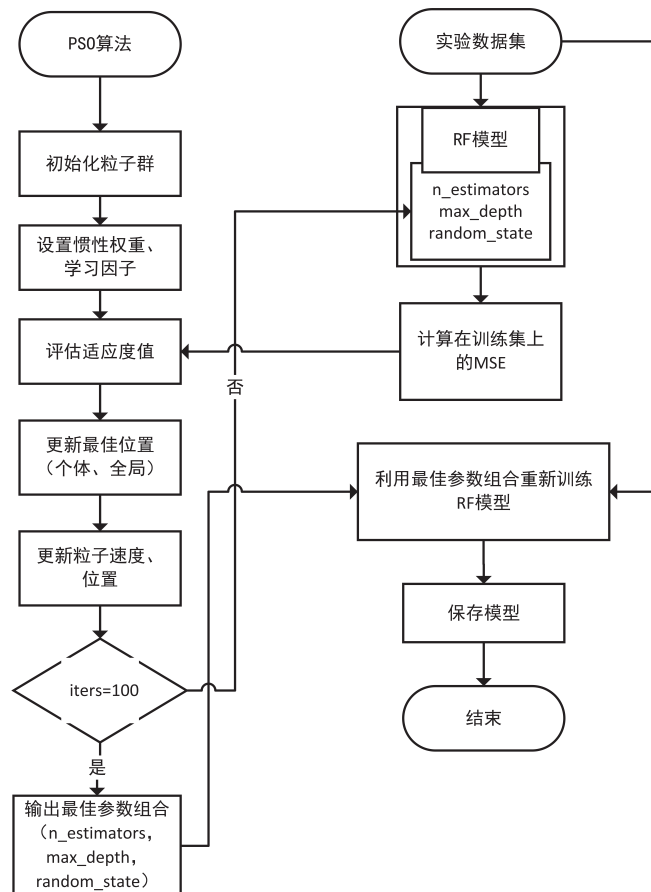


图3 粒子群优化随机森林算法流程

Fig. 3 Random forest algorithm based on particle swarm optimization

模型拟合的多元回归方程见公式 3。通过响应面模型拟合的多元回归方程,能够有效地捕捉输入变量之间的相互作用及其对响应变量的影响。在本研究中,为了扩充各机器学习模型训练数据集,通过拟合回归方程,在原始实验数据基础上生成新的数据样本,丰富训练数据集,有助于提高各机器学习模型的预测精度和泛化能力^[14]。从表中分析可知,磨矿细度(X_4)、起泡剂(X_2)、分散剂(X_3)对模型影响最为显著。各因素对响应值的影响 $X_4 > X_3 > X_2 > X_1$,交互项的影响顺序为 $X_2X_3 > X_3X_4 > X_1X_2 > X_2X_4 > X_1X_4 > X_1X_3$ 。

$$y_{\text{grade}} = 1.6854 - 0.0005 \cdot X_1 + 0.1363 \cdot X_2 - 0.0020 \cdot X_3 - 2.1170 \cdot X_4 + 0.0002 \cdot X_1 \cdot X_2 - 2.8544 \times 10^{-7} \cdot X_1 \cdot X_3 + 0.0220 \cdot X_1 \cdot X_4 + 7.5039 \times 10^{-6} \cdot X_2 \cdot X_3 - 0.1816 \cdot X_2 \cdot X_4 + 0.0031 \cdot X_3 \cdot X_4 - 0.0001 \cdot X_1^2 - 0.0005 \cdot X_2^2 - 1.0162 \times 10^{-7} \cdot X_3^2 \quad (3)$$

图 4 展示了捕收剂、起泡剂、分散剂、磨矿细度与浮选精矿品位之间的响应面图。响应面越陡,影响越显著,从图 4 中可以看出,当捕收剂(X_1)和起泡剂(X_4)、分散剂(X_3)逐渐增大时,浮选精矿品位先增大后减小;从图 4c、4e、4f 中可以看出,磨矿细度逐渐增大,精矿品位逐渐减小。

通过诊断分析能进一步检验模型的准确性,从图 5a 中可以看出大部分数据点紧密地分布在参考直线附近,这表明模型的残差服从正态分布,说明模型的拟合是合理的;图 5b、5c 中可以看出残差的

偏差或变化趋势较小,随机分布在水平带内,表明模型拟合良好,适应性强;从图 5d 中可以看出所有数据点都分布在对角线附近,且没有显著的偏离,表明模型能够很好地预测响应变量,实际值与预测值之间的差异较小,表明模型的预测准确性较高。因此,建立的模型能够有效地预测新的数据。

2.2 模型评价指标概述

为了量化各模型的准确性,在回归问题中均方误差(RMSE)、平均绝对误差(MAE)和决定系数(R^2)是用于评价回归模型性能的主要指标,如公式(4)、(5)、(6)所示。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

其中, y_i 表示真实值, $\hat{y}_i$ 表示预测值, $\bar{y}_i$ 表示各真实值的平均值, n 表示样本数量。

RMSE 计算了真实值和预测值之间的平均误差平方,反映了预测值与真实值之间的平均偏离程度,并且对异常值敏感。MAE 计算了真实值和预测

表 2 辉钼矿浮选精矿品位的方差分析

Table 2 Analysis of variance for the grade of molybdenite flotation concentrates

变异来源	平方和	自由度	均方	F值	P值	备注
模型	0.723 8	13	0.055 7	26.11	<0.000 1	显著
X_1	0.009 8	1	0.009 8	4.58	0.058 0	
X_2	0.044 4	1	0.044 4	20.82	0.001 0	
X_3	0.073 6	1	0.073 6	34.51	0.000 2	
X_4	0.259 1	1	0.259 1	121.47	<0.000 1	
X_1X_2	0.060 3	1	0.060 3	28.27	0.000 3	
X_1X_3	0.000 6	1	0.000 6	0.278 7	0.609 1	
X_1X_4	0.003 0	1	0.003 0	1.43	0.259 4	
X_2X_3	0.094 7	1	0.094 7	44.40	<0.000 1	
X_2X_4	0.054 2	1	0.054 2	25.43	0.000 5	
X_3X_4	0.075 1	1	0.075 1	35.22	0.000 1	
X_1^2	0.027 4	1	0.027 4	12.83	0.005 0	
X_2^2	0.053 3	1	0.053 3	25.00	0.000 5	
X_3^2	0.024 6	1	0.024 6	11.53	0.006 8	
失拟项	0.012 4	5	0.002 5	1.39	0.362 8	不显著
残差	0.008 9	5	0.001 8			

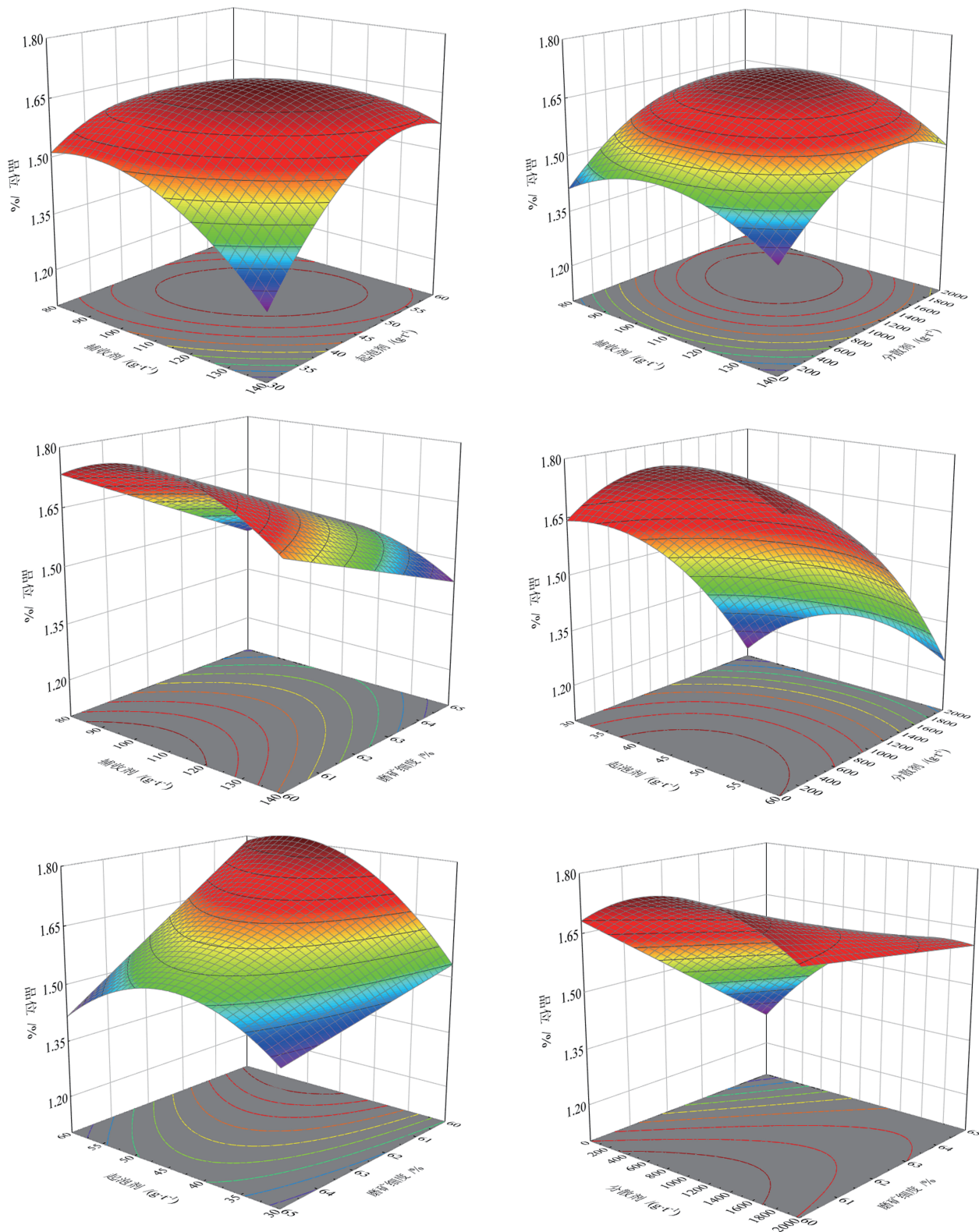


图4 辉钼矿浮选精矿品位响应面
Fig. 4 Response surface of molybdenite flotation concentrate grade

值之间的平均偏差,反映了预测值与真实值之间的平均绝对差距。 R^2 表示模型能够解释数据方差的比例,衡量模型的拟合效果,范围在 0~1 之间。 R^2 值越接近 1 和 $RMSE$ 、 MAE 值越小,表明模型的准

确性越高。

2.3 各机器学习模型预测性能分析

通过 python 的 sklearn、numpy 以及 torch 等库

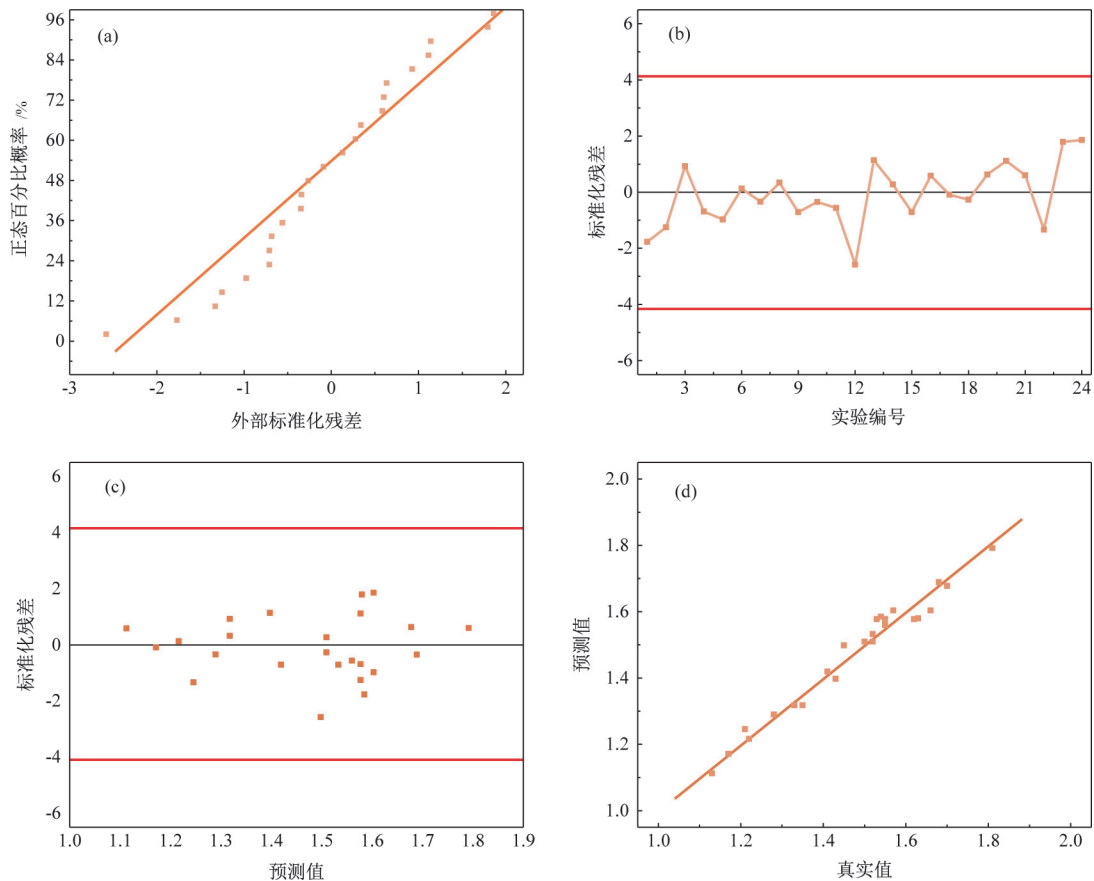


图5 模型诊断结果

Fig. 5 Model diagnostic plot

建立各机器学习模型。数据集共 256 组数据,由 128 组实际实验数据和 128 组生成数据组成,其中训练集与验证集划分比为 8 : 2,利用训练集和验证集对各模型进行训练和验证,各模型在验证集上的结果统计见表 3。以 R^2 为主要评估指标对预测结果进行排序,支持向量机(SVM)、线性回归(LR)、K 近邻对数据的拟合效果都较差,其中支持向量机(SVM)的 R^2 仅为 0.529 8。随机森林算法(RF)对数据的预测效果最好, R^2 达到了 0.969 6,同时 $RMSE$ 和 MAE 值相比于其他模型最小,预测效果表现见图 6。

从图 4 可以看出,捕收剂、起泡剂、分散剂和磨矿细度之间存在复杂的交互作用,不同比例组合

表 3 各模型在验证集上的预测效果

Table 3 Predictive performance of each model on the validation set

Model	$RMSE$	MAE	R^2
SVM	0.1785	0.1251	0.5298
LR	0.1763	0.1342	0.5793
KNN	0.1595	0.1187	0.6507
RF	0.0458	0.0295	0.9696

的捕收剂、起泡剂、抑制剂和磨矿细度浮选效果不同。如当起泡剂用量增加时,适量添加分散剂能够显著提升浮选精矿品位;反之,当起泡剂用量过高或分散剂用量不足时,可能导致矿物颗粒分散不均,进而降低浮选精矿品位,起泡剂和分散剂之间交互作用显著。同时在实验过程中,在人为、机械等因素的影响下,实验结果可能存在异常值。而随机森林模型通过构建多棵决策树,每棵树在构建过程中都会随机选择特征进行划分,相比其他算法能够更好地捕捉到这些特征之间的非线性复杂关系。随机森林模型通过集成多个决策树的预测结果取平均值作为最终预测结果,对异常值具有较强的鲁棒性,能够大幅减少异常值对整体预测的影响。虽然随机森林算法在各模型中的预测效果最好,但从图 6 中可以看出,部分样本的预测值与实际值存在明显偏差,且预测值与真实值的误差曲线波动较大,部分单点误差较高。可能是由于随机森林模型在训练过程中树的数量、树的深度、随机种子等参数设置不当。可以通过优化随机森林模型参数($n_estimators$, max_depth , $random_state$)组合,有效地提升模型的预测效果,减少单点误差。

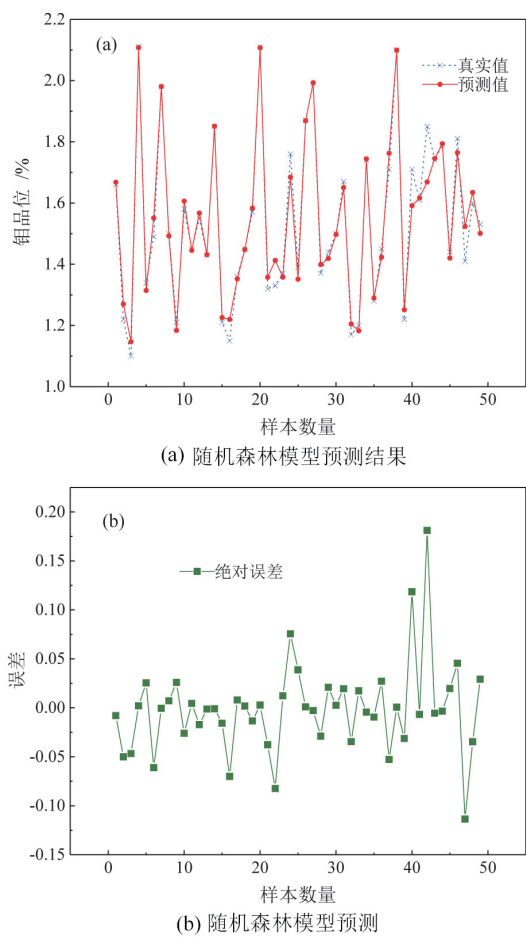


图6 随机森林模型预测效果
Fig. 6 Predictive performance of the random forest model

2.4 PSO-RF 预测模型性能分析

使用粒子群算法优化随机森林模型(PSO-RF)参数 $n_estimators$ 、 max_depth 、 $random_state$, 模型参数值设定范围分别为 $[1,200]$ 、 $[1,50]$ 、 $[1,50]$ 。粒子群算法中粒子数量设定为 100, 参数维度设定为 3, 学习因子 C_1 、 C_2 值设定为 0.7, 惯性权重 ω 值设定为 0.9。迭代次数设定为 100 次。当迭代次数达到 100 次时, 停止迭代优化, 输出最佳参数组合, 使用最佳参数组合重新训练模型。以均方误差(RMSE)、平均绝对误差(MAE)和决定系数(R^2)为评价指标, 验证重新训练后的 PSO-RF 模型在验证集上的性能, 其中 RMSE 值为 0.036 9, MAE 值为 0.024 5, R^2 值为 0.980 2。预测效果表现见图 7。

随机森林模型、PSO-RF 模型在验证集上的预测性能指标统计结果如表 4 所示。N、M、R 分别表示随机森林模型中参数 $n_estimators$ 、 max_depth 、 $random_state$ 的数量。PSO-RF 模型性能指标相比 RF 模型更优, RMSE 值和 MAE 值相对于 RF 模型分别降低了 19.43% 和 16.95%, 同时 R^2 值提升了

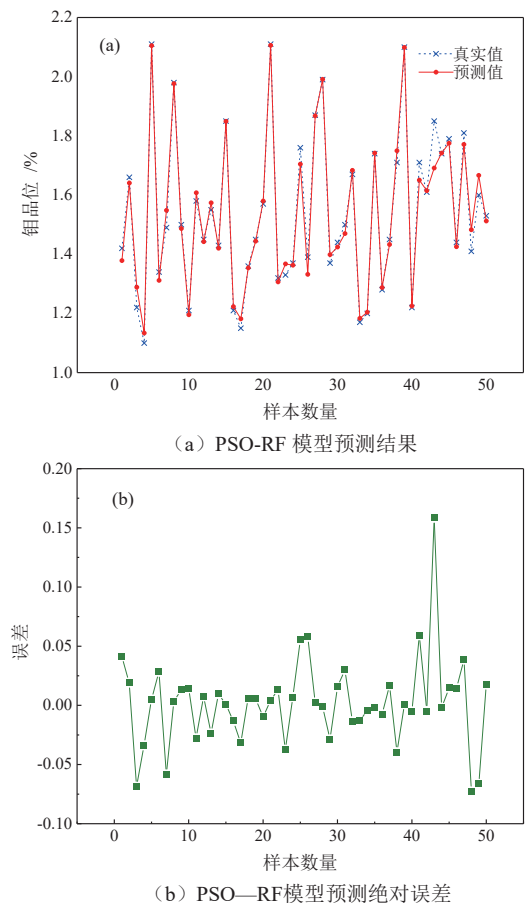


图7 PSO-RF 模型预测效果
Fig. 7 Predictive performance of the PSO-RF model

表 4 随机森林模型、PSO-RF 模型在验证集上的预测性能指标统计
Table 4 Statistical performance metrics of the random forest model and PSO-RF model on the validation set

模型	RMSE	MAE	R^2	优化时/s	最佳参数组合 NMR		
RF	0.045 8	0.029 5	0.962 6	—	—	—	—
PSO-RF	0.036 9	0.024 5	0.980 2	21.34	58	38	32

1.83%。此外,从图 6、图 7 中可以看出, PSO-RF 模型的预测值也更接近真实值。分析结果表明, PSO-RF 模型相比文章中的其他预测模型, 模型性能更优, 预测结果更准确。

2.5 PSO-RF 预测模型实验验证

为了验证 PSO-RF 模型在实际应用中预测的准确性, 从某辉钼矿浮选厂采集矿样进行浮选, 利用保存性能最优的模型预测了在浮选过程中 15 组不同比例组合的捕收剂、起泡剂、分散剂和磨矿细度对应的浮选精矿品位, 并对 15 组浮选实验进行实际检测, 比较预测值与实际值之间的误差, 检验模型在实际应用中的准确性。PSO-RF 模型对 15

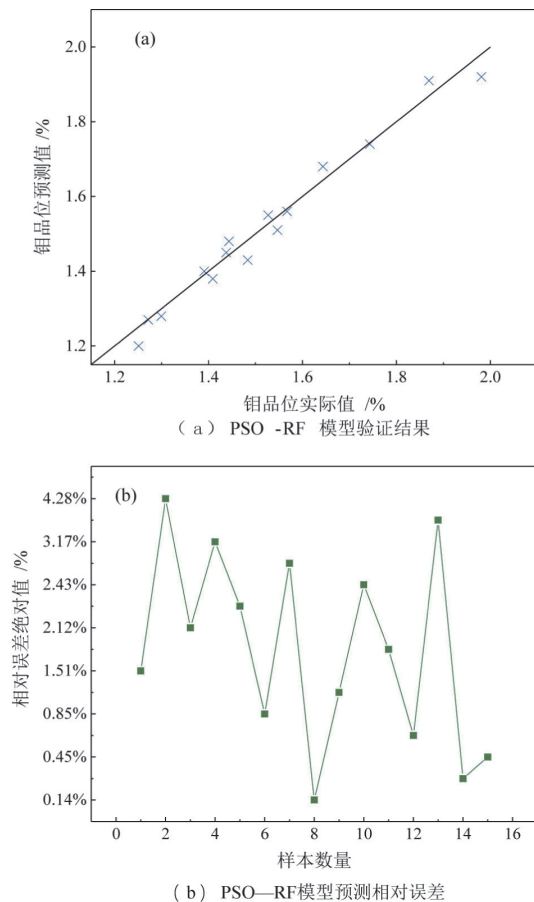


图8 PSO-RF 模型验证效果
Fig. 8 Validation effect of the PSO-RF model

组浮选实验的预测结果见图8。从实验结果中可以看出, PSO-RF 模型对品位的预测值和实际值之间的相对误差绝对值范围在 0.14%~4.28%, 表明模型对辉钼矿浮选品位有良好的预测效果。因此, PSO-RF 模型能够通过不同浮选药剂用量和磨矿细度准确地预测浮选品位, 在实际应用中表现出良好的性能。

3 结论

目前辉钼矿浮选精矿品位检测手段检测周期长, 在线检测设备成本昂贵且检测精度相对不高, 中小型浮选厂主要依靠人工取样化验的方法进行品位检测。为此, 本文提出了一种基于粒子群算法结合随机森林模型的方法, 用于预测辉钼矿浮选过程中的精矿品位。其中 $RMSE$ 值为 0.036 9, MAE 值 0.024 5, R^2 值为 0.980 2, 表明该模型具有较高的预测精度。PSO-RF 预测模型通过降低对在线检测设备的依赖, 显著减少了设备成本且具备实时预测能力, 提升了品位检测效率, 有望为中小型浮选厂浮选品位在线检测提供一种新方法。

在模型验证阶段, 选取某辉钼矿浮选厂的矿样开展浮选实验, 以评估所构建模型的预测性能。实验结果表明, 该模型的预测精度较高, 能够满足工业生产需求, 并对浮选工艺的优化及过程控制具有重要的指导价值。本研究为解决辉钼矿浮选过程中的品位在线检测问题提供了一种高效且低成本的技术方案, 在提升浮选工艺智能化水平方面具有重要意义。此外, 所提出的预测模型不仅适用于辉钼矿浮选过程中精矿品位的预测, 其研究方法亦具备一定的适用潜力, 可为构建通用的矿物浮选精矿品位预测模型提供理论参考。该研究的成果为矿物加工行业向智能化、高效化方向发展提供了新的技术路径。未来研究应进一步探讨更多浮选参数对精矿品位的影响, 以扩充数据集并优化 PSO-RF 模型的预测性能和泛化能力。同时, 还应收集不同类型矿石的浮选数据, 评估 PSO-RF 模型在不同矿石浮选过程中的精矿品位预测效果, 从而进一步验证其在多种矿物浮选工艺中的适用性和鲁棒性。

参考文献:

- [1] MCKEE D J. Automatic flotation control—a review of 20 years of effort[J]. Minerals Engineering, 1991, 4(7/8/9/10/11): 653–666.
- [2] JO K, JE J, LEE D, et al. Prediction of multi-stage froth flotation efficiency of complex lead–zinc sulfide ore using an integrated ensemble neural network–random forest model[J]. Minerals Engineering, 2024, 210: 108669.
- [3] KOSKINEN J, LAPPALAINEN S, LESKINEN T, et al. Performance of on-stream analyzers at Outokumpu concentrators, Finland[J]. CAN Min Metall Bull, 1973, 66(730): 37–47.
- [4] BENDAOUIA A, QASSIMI S, BOUSSETTA A, et al. Hybrid features extraction for the online mineral grades determination in the flotation froth using deep learning[J]. Engineering Applications of Artificial Intelligence, 2024, 129: 107680.
- [5] BARTOLACCI G, PELLETIER JR P, TESSIER JR J, et al. Application of numerical image analysis to process diagnosis and physical parameter measurement in mineral processes—part I: Flotation control based on froth textural characteristics[J]. Minerals Engineering, 2006, 19(6/7/8): 734–747.
- [6] REN C, YANG J, LIANG C. Estimation of copper concentrate grade based on color features and least-squares support vector regression[J]. Physicochemical Problems of Mineral Processing, 2015, 51(1): 163–172.
- [7] CAO B, XIE Y, GUI W, et al. Integrated prediction model

- of bauxite concentrate grade based on distributed machine vision[J]. [Minerals Engineering](#), 2013, 53: 31–38.
- [8] JAHEDSARAVANI A, MARHABAN M H, MASSINAEI M. Prediction of the metallurgical performances of a batch flotation system by image analysis and neural networks[J]. [Minerals Engineering](#), 2014, 69: 137–145.
- [9] JING Z, JINLING N, YONG Z. Prediction of concentrate grade and recovery rate of tailings in the process of production based on chaotic ant colony algorithm[C]// 2018 Chinese Control and Decision Conference (CCDC). IEEE, 2018: 5308–5313.
- [10] WANG J S, HAN S. Feed - Forward Neural Network Soft - Sensor modeling of flotation process based on particle swarm optimization and gravitational search algorithm[J]. [Computational Intelligence and Neuroscience](#), 2015, 2015(1): 147843.
- [11] GOMEZ-FLORES A, HEYES G W, ILYAS S, et al. Prediction of grade and recovery in flotation from physicochemical and operational aspects using machine learning models[J]. [Minerals Engineering](#), 2022, 183: 107627.
- [12] RIGATTI S J. Random forest[J]. [Journal of Insurance Medicine](#), 2017, 47(1): 31–39.
- [13] WANG D, TAN D, LIU L. Particle swarm optimization algorithm: An overview[J]. [Soft computing](#), 2018, 22(2): 387–408.
- [14] MEHRABANI J V, NOAPARAST M, MOUSAVI S M, et al. Process optimization and modelling of sphalerite flotation from a low-grade Zn–Pb ore using response surface methodology[J]. [Separation and Purification Technology](#), 2010, 72(3): 242–249.

A PSO–RF–Based Prediction Model for Molybdenite Flotation Concentrate Grade

WAN He¹, YANG Liangliang¹, QU Juanping², JIANG Peng³

1. School of Resources Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;

2. Oulu Mining School, University of Oulu, Oulu FI-90014, Finland;

3. China Coal Science and Industry Group Nanjing Design and Research Institute Co., Ltd, Wuhan 210031, China

Abstract: Mineral flotation concentrate grade is one of the crucial technical indicators for evaluating flotation efficiency, and its measurement accuracy plays a significant guiding role in optimizing production processes, adjusting operations, and controlling product quality. For many years, the accurate real-time measurement of flotation concentrate grade has been a key research focus within the mining industry. However, current online detection methods for molybdenite flotation concentrate grades generally exhibit notable shortcomings, such as inadequate measurement accuracy, high equipment costs, and complex maintenance procedures, severely limiting their widespread implementation in industrial production. Therefore, developing an online detection method characterized by high accuracy, low implementation cost, and ease of industrial deployment has significant practical importance. To overcome these limitations, this study proposed a molybdenite flotation concentrate grade prediction model based on a Particle Swarm Optimization–Random Forest (PSO–RF) algorithm. The objective of this study was to address the accuracy and cost issues inherent to existing detection approaches and to provide a more economical and efficient predictive tool suitable for industrial practice. The proposed PSO–RF model effectively combined the global optimization capability of the particle swarm optimization (PSO) algorithm and the superior ability of random forest (RF) models in handling complex data, significantly enhancing the model's predictive accuracy and generalization performance. Initially, flotation experiments involving molybdenite were conducted to systematically investigate the influence of critical flotation parameters, including reagent dosage and grinding fineness, on the concentrate grade. Based on the experimental data obtained from these trials, a multi-input and multi-output random forest prediction model was constructed. Subsequently, considering that the prediction performance of random forest models is sensitive to hyperparameter selection (specifically, *n_estimators*, *max_depth*, and *random_state*), the particle swarm optimization algorithm was employed to globally optimize these critical hyperparameters. This optimization approach effectively overcame the limitations of traditional manual hyperparameter tuning, substantially improving model prediction performance and stability. In the model validation stage, the optimized PSO–RF model currently exhibits excellent predictive performance. Specifically, on the validation dataset, the PSO–RF model achieves a root mean square error (RMSE) of 0.036 9, mean absolute error (MAE) of 0.024 5, and a coefficient of determination (R^2) of 0.980 2. Compared to the unoptimized random forest model, the proposed model improves the R^2 value by 1.83%, reduces the RMSE by 19.43%, and decreases the MAE by 16.95%. Furthermore, additional experimental validation confirms that the maximum relative error between predicted and actual measured concentrate grades is consistently below 4.28%, demonstrating high prediction accuracy and strong generalization capability. In conclusion, the PSO–RF prediction model proposed in this study effectively addresses the existing deficiencies of conventional online detection methods. This model not only realizes accurate real-time prediction of flotation concentrate grades but also features lower implementation costs and ease of industrial application. Consequently, the proposed PSO–RF model demonstrates substantial potential for practical application and economic value in industrial molybdenite flotation processes.

Keywords: molybdenite; grade prediction; particle swarm optimization; random forest model

引用格式:宛鹤, 杨亮亮, 屈娟萍, 江鹏. 基于粒子群优化随机森林算法的辉钼矿浮选精矿品位预测模型[J]. 矿产保护与利用, 2025, 45(3): 76–86.

WAN He, YANG Liangliang, QU Juanping, JIANG Peng. A PSO–RF–based prediction model for molybdenite flotation concentrate grade[J]. Conservation and Utilization of Mineral Resources, 2025, 45(3): 76–86.

投稿网址: <http://kcbhyly.xml-journal.net>

E-mail: kcbh@chinajournal.net.cn