烧结矿低温还原粉化率的影响因素及预测模型

郑兆颖[1]，邢相栋[1]，王荪璇[2]，周小辉[3]
(1. 西安建筑科技大学 冶金工程学院，陕西 西安 710055；2. 陕西省食品药品监督检验研究院，陕西 西安 710016；
3. 山东钢铁股份有限公司 莱芜分公司，山东 济南 250000)

摘要： 针对烧结矿低温还原粉化率传统检测方法时间滞后、过程繁杂、耗时较长等问题，本文以某钢铁厂 400 m^2 烧结机烧结矿的性能指标为研究对象，选用 60 组训练集进行因子分析，建立以 Al_2O_3、SiO_2、MgO、TiO_2、FeO 质量分数以及 R、$w(MgO)/w(Al_2O_3)$ 和 $w(CaO)/w(TFe)$ 为自变量，低温还原粉化率为因变量的烧结矿低温还原粉化率预测模型。结果表明：影响烧结矿低温还原粉化率的主要因素是 $w(Al_2O_3)$、$w(SiO_2)$、$w(FeO)$、R 和 $w(MgO)/w(Al_2O_3)$；优化后预测模型为 $Y_{RDI}=1.137w(FeO)+5.56w(SiO_2)+26.44R-20.19w(MgO)/w(Al_2O_3)-6.07w(Al_2O_3)$；模型的 R^2 为 99.90%，具有较高的精确度，能够快速预测烧结矿的低温还原粉化率，并能在生产中改善烧结矿质量。本文建立的预测模型可为现场烧结生产提供理论指导和技术支持。

关键词： 烧结矿；低温还原粉化率；影响因数；主成分分析法；回归模型；预测模型
中图分类号： TF046.4；TP202$^+$.2 **文献标识码：** A **文章编号：** 1000－8764(2023)02－0025－08
doi： 10.13403/j.sjqt.2023.02.019

Influencing factors and prediction model of low temperature reduction disintegration index of sinter

ZHENG Zhaoying[1], XING Xiangdong[1], WANG Sunxuan[2], ZHOU Xiaohui[3]
(1. School of Metallurgical Engineering, Xi´an University of Architecture and Technology, Xi´an 710055, Shaanxi, China; 2. Shaanxi Institute for Food and Drug Control, Xi´an 710016, Shaanxi, China; 3. Laiwu Branch of Shandong Iron and Steel Co., Ltd., Jinan 250000, Shandong, China)

Abstract: Aiming at such problems as time lag, complicated process and high time consumption of traditional detection methods for low-temperature reduction disintegration index of sinter, the performance index of sinter of 400 m^2 sintering machine of a steel plant is taken as the research object, the 60 training sets are selected for factor analysis, and a prediction model of low-temperature reduction disintegration index of sinter is established with Al_2O_3, SiO_2, MgO, TiO_2, FeO mass fractions, R, $w(MgO)/w(Al_2O_3)$ and $w(CaO)/w(TFe)$ as independent variables and low-temperature reduction disintegration index as dependent variables. The results show that the main factors affecting the low-temperature reduction disintegration index of sinter are $w(Al_2O_3)$, $w(SiO_2)$, $w(FeO)$, R and $w(MgO)/w(Al_2O_3)$. The optimized prediction model is $Y_{RDI}=1.137w(FeO)+5.56w(SiO_2)+26.44R-20.19w(MgO)/w(Al_2O_3)-6.07w(Al_2O_3)$. The R^2 of the model is 99.90%, and the model has high accuracy, which can quickly predict the low-temperature reduction disintegration index of sinter, and improve the quality of sinter in production. The prediction model can provide theoretical guidance and technical support for on-site sintering production.

Key words: sinter; low temperature reduction disintegration index; impact factor; principal component analysis; regression model; prediction model

收稿日期： 2021－11－18；**修回日期：** 2021－12－21
基金项目： 国家自然科学基金资助项目(52174325)；陕西省重点研发计划资助项目(2019TSLGY05－09)
作者简介： 郑兆颖(1998—)，女，硕士研究生，从事冶金工程数学模型及大数据技术方面的研究。
通信作者： 邢相栋(1983—)，男，从事复杂难处理矿产资源综合利用等方面的研究。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

据统计，高炉炼铁产量占世界生铁总产量的95%以上，而烧结矿作为我国高炉炼铁的主要含铁炉料，其低温还原粉化率是评价矿石冶金性能的质量标准之一，对高炉操作和固体燃料消耗均有着重要的影响[1]。然而，传统烧结矿低温还原粉化率检验方法时间滞后、过程繁杂，且耗费时间较长[2-3]。因此，快速预测烧结矿的低温还原粉化率，实时调控改善生产操作的工艺参数，对于获取高品质烧结矿具有重要的指导作用。

主成分分析法是从较大数据集中标识较小数量的不相关变量，使用最少数量的主分量来解释最大量的方差。杨浩等[4]采用该方法对张家港市13条重要河道中的水质指标进行分析，发现氨氮和磷的质量分数、高锰酸盐指数与电导率呈现显著的正相关关系，该方法能有效识别主导水体变化的环境因子。因子分析法作为主成分分析法的推广和发展，其过程简单、耗时较短，近年来在工业生产数据分析和性能指标预测等领域得到越来越多的应用[5-6]。吕航等[7]采用该方法对水化学环境的影响因素进行分析，将各变量归纳为5个主因子，快速表征了石油烃生物降解作用对水化学环境的影响。WANG等[8]结合因子分析和聚类分析方法，对内蒙古某地区采集的土壤样品的每个集群进行因子分析，用于识别新的地球化学模式。

目前，铁矿烧结领域的指标预测研究和应用多集中在固体燃料消耗、成品率和转鼓指数等方面，而对烧结矿冶金性能较少。本文通过分析60组烧结矿的低温还原粉化率及化学成分，采用因子分析法得到影响烧结矿低温还原粉化率的主成分，并进行最佳子集回归分析；同时，利用测试集对预测模型进行误差检验，得到最佳子集回归模型，建立预测回归方程，采用基于观测值的杠杆率、Cook距离和DFITS值来优化预测模型，为改善烧结矿低温还原粉化率提供参考。

1 模型因子分析

本文通过分析60组烧结矿的低温还原粉化率和化学成分进行因子分析，以转鼓强度、$w(CaO)$、$w(TFe)$、$w(SiO_2)$、$w(FeO)$、$w(MgO)$、$w(Al_2O_3)$、$w(S)$、$w(TiO_2)$、$w(MgO)/w(Al_2O_3)$、$w(SiO_2)/w(Al_2O_3)$、$w(CaO)/w(TFe)$、R共13种因子作为自变量，采用主成分分析法来计算因子载荷，确定每个因子与载荷较高自变量的联系；结合变量-因子方差最大法，保证每个因子中具有最高载荷的变量数及解释变量所需的因子数最少。

为明确自变量对低温还原粉化率的影响大小，利用降维的思想从上述13种因子中提炼出有限个彼此之间相关性不明显的自变量，尽量规避线性回归中自变量之间的共线性问题。本文基于方差贡献率和因子载荷分析，并筛选掉相关性高的因子，建立旋转后的载荷和公因子方差，结果如表1所示。公因子的累积方差贡献率达到80%即能满足模型预测要求。由表1可以看出：前5个公因子的方差贡献率分别为29.1%、17.9%、15.4%、14.5%和9.3%，能够解释原始指标总信息的86.2%，符合模型预测要求。在公因子1中，$w(TFe)$、$w(CaO)$和$w(CaO)/w(TFe)$三者的相关性高，$w(CaO)/w(TFe)$的载荷绝对值最大，其与低温还原粉化率的相关性最高；同样地，公因子2中，$w(SiO_2)$和$w(S)$的相关性较强，$w(SiO_2)$的载荷绝对值最大；公因子4中，$w(Al_2O_3)$和$w(SiO_2)/w(Al_2O_3)$的相关性较强，$w(Al_2O_3)$的载荷绝对值最大；公因子5中，转鼓强度和$w(FeO)$的载荷相关性较强，$w(FeO)$的载荷绝对值最大。因此，筛选后得到$w(Al_2O_3)$、$w(SiO_2)$、$w(MgO)$、$w(TiO_2)$、R、$w(FeO)$、$w(MgO)/w(Al_2O_3)$、$w(CaO)/w(TFe)$共8种因子。

为验证因子分析结果，本文绘制13种因子的特征值变化，结果如图1所示。由图1可以看出，当因子数为8时，折线开始趋于平缓，后面的特征值变化不大，所以本文选择因子数为5。

2 模型建立与检验

2.1 最佳子集回归模型建立

最佳子集回归算法根据所有预测变量集生成最高R^2值的子集模型，能够有效识别最优模型。本文将前述因子分析结果作为自变量，烧结矿低温还原粉化率作为因变量，建立预测回归模型。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1　旋转后的载荷和公因子方差

Table 1　Rotated load and common factor variance

变量名称	公因子 1	公因子 2	公因子 3	公因子 4	公因子 5
转鼓强度	0.203 0	0.124 0	0.155 0	0.292 0	−0.751 0
w(TFe)	−0.811 0	0.350 0	0.219 0	0.122 0	−0.055 0
w(CaO)	0.896 0	−0.324 0	−0.164 0	−0.120 0	0.018 0
w(SiO_2)	0.198 0	−0.907 0	0.297 0	0.016 0	−0.025 0
w(Al_2O_3)	0.180 0	−0.198 0	0.085 0	−0.930 0	0.022 0
w(S)	0.330 0	−0.670 0	−0.312 0	0.082 0	−0.060 0
w(MgO)	0.664 0	−0.271 0	−0.534 0	−0.075 0	0.156 0
w(TiO_2)	−0.020 0	0.170 0	−0.888 0	−0.078 0	−0.015 0
R	0.661 0	0.525 0	−0.453 0	−0.118 0	0.030 0
w(FeO)	−0.215 0	−0.181 0	−0.113 0	−0.212 0	−0.767 0
w(MgO)/w(Al_2O_3)	0.607 0	−0.202 0	−0.599 0	0.273 0	0.152 0
w(SiO_2)/w(Al_2O_3)	−0.003 0	−0.458 0	0.153 0	0.855 0	−0.030 0
w(CaO/w(TFe)	0.896 0	−0.338 0	−0.185 0	−0.126 0	0.029 0
方差	3.778 8	2.329 3	2.000 6	1.879 0	1.210 5
方差贡献率	0.291 0	0.179 0	0.154 0	0.145 0	0.093 0

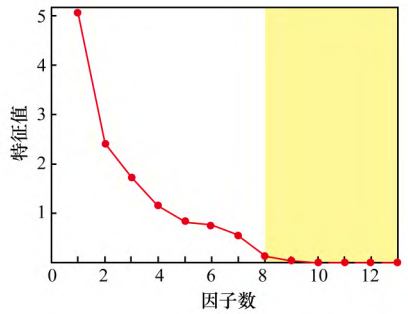


图 1　13 种因子的特征值变化

Fig. 1　The changes of eigenvalue of 13 kinds of factors

本文定义 3 个参数来评价模型，C_p 为评估线性回归模型优良性的标准，R^2 为确定系数，$R^2_{预测}$ 为预测确定系数。设置如下筛选条件：① C_p 接近且小于预测变量个数；② R^2 值达到最大；③ $R^2_{预测}$ 值达到最大。计算方法分别如式(1)～(3)所示。

$$C_p = \frac{\sigma_{SSE}}{\sigma_{MSE}} - n + 2p = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{i,p})^2}{[1/(n-K)]\sigma_{SSE}} - n + 2p = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{i,p})^2}{[1/(n-K)]\sum_{i=1}^{n}(y_i - \hat{y}_{i,K})^2} - n + 2p \quad (1)$$

$$R^2 = 1 - \frac{S_{误差}}{S_{合计}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (2)$$

$$R^2_{预测} = 1 - \frac{\sigma_{PRESS}}{S_{合计}} = 1 - \frac{\sum_1^n [e_i/(1-h_i)]^2}{\sum_1^n (y_i - \bar{y})^2} \quad (3)$$

式中：σ_{SSE} 为残差平方和；σ_{MSE} 为平均误差平方；σ_{PRESS} 为预测平方和；$S_{误差}$ 为误差距离平方和；$S_{合计}$ 为总误差平方和；n 为观测值个数，个；p 为模型中的项数，个；y_i 为第 i 个观测响应值；$\hat{y}_{i,K}$ 为在含 K 个自变量的线性回归模型中第 i 个 y 样本数据的拟合值；$\bar{y}$ 为平均响应值；$\hat{y}_i$ 为第 i 个拟合响应值；e_i 为第 i 个残差值；h_i 为 $X(X^TX)^{-1}X^T$（X 为设计矩阵，X^T 为设计矩阵的转置）中第 i 个对角线元素。

以前述因子分析中 8 种因子为自变量，烧结矿的低温还原粉化率为因变量，建立最佳子集回归模型如表 2 所示。由表 2 可以看出：三变量模型{w(MgO)、R、w(CaO)/w(TFe)}的 C_p 值为 2.9，与模型项数 $p=3$ 很接近，其在估计真实回归系数和预测未来响应时较为精确且无偏倚；四变量模型{w(SiO_2)、w(MgO)、w(FeO)、w(CaO)/w(TFe)}的 C_p 值为 3.2，与模型项数 $p=4$ 较接近，其 R^2 和 $R^2_{预测}$ 值较大，表明模型拟合优度高，预测能力出色；五变量模型{w(SiO_2)、w(Al_2O_3)、R、w(FeO)、w(MgO)/w(Al_2O_3)}的

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表2 烧结矿低温还原粉化率的最佳子集回归模型

Table 2 The best subset regression model of low temperature reduction disintegration index of sinter

变量个数	R^2	σ_{PRESS}	$R^2_{预测}$	C_p	$w(SiO_2)$	$w(Al_2O_3)$	$w(MgO)$	$w(TiO_2)$	R	$w(FeO)$	$w(MgO)/w(Al_2O_3)$	$w(CaO)/w(TFe)$
1	4.5	688.3	0.0	12.7						√		
1	4.2	686.8	0.0	12.9					√			
1	3.7	690.6	0.0	13.3			√					
1	3.5	690.3	0.0	13.4							√	
1	1.3	709.5	0.0	15.0								√
2	18.4	599.4	11.3	4.7			√					√
2	15.3	630.1	6.8	7.0			√		√			
2	15.2	623.1	7.8	7.0					√		√	
2	13.0	643.2	4.8	8.6							√	√
2	9.8	665.0	1.6	10.9					√	√		
3	25.0	570.2	15.6	2.0	√		√		√			
3	23.8	576.6	14.7	2.9			√		√			√
3	22.9	583.6	13.6	3.5	√		√					√
3	21.0	599.8	11.2	4.9			√			√		√
3	20.3	603.9	10.6	5.4	√				√		√	
4	27.9	564.8	16.4	1.9	√		√		√	√		
4	26.9	570.5	15.6	2.6			√		√	√		√
4	26.0	577.2	14.6	3.2	√		√			√		√
4	25.6	582.0	13.9	3.5	√		√		√			√
4	25.2	585.3	13.4	3.8	√		√	√	√			
5	28.4	576.1	14.8	3.5	√		√		√	√		√
5	28.3	579.9	14.2	3.6	√		√	√	√	√		
5	27.9	582.7	13.8	3.9	√	√	√		√	√		
5	27.9	582.1	13.9	3.9	√		√		√	√	√	
5	27.5	587.2	13.1	4.2	√	√			√	√	√	
6	28.9	590.6	12.6	5.2	√		√	√	√	√		√
6	28.4	595.7	11.9	5.5	√	√	√		√	√		√
6	28.4	595.2	11.9	5.5	√		√		√	√	√	√
6	28.3	598.4	11.4	5.6	√	√	√	√	√	√		
6	28.3	598.1	11.5	5.6	√		√	√	√	√	√	

注: 表中“√”代表该因数在模型中被选用,空白表示未被选用。

C_p 值为4.2，与模型项数 $p=5$ 较接近，其与其余五变量模型相比，在估计真实回归系数和预测未来响应时较为精确且无偏倚。

2.2 模型检验

2.2.1 误差检验

本文采用损失函数来定量比较模型{$w(MgO)$、R、$w(CaO)/w(TFe)$}、{$w(SiO_2)$、$w(MgO)$、$w(FeO)$、$w(CaO)/w(TFe)$}和{$w(SiO_2)$、$w(Al_2O_3)$、R、$w(FeO)$、$w(MgO)/w(Al_2O_3)$}的预测精度。各损失函数的具体定义如式(4)~(7)所示。

$$\sigma_{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{4}$$

$$\sigma_{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{5}$$

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$$\sigma_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{6}$$

$$\sigma_{\mathrm{MAPE}}=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i}{y_i}-1\right| \tag{7}$$

式中：σ_{MSE}为平均误差平方(mean squared error，MSE)；σ_{RMSE}为均方根误差(root mean square error，RMSE)；σ_{MAE}为平均绝对误差(mean absolute error，MAE)；σ_{MAPE}为平均相对百分比误差(mean absolute percentage error，MAPE)。

使用 10 个测试集检验上述 3 个模型的准确性，对比各损失函数值，结果如表 3 所示。由表 3 可以看出，平均误差平方最大为 17.17，最小为 1.54，五变量模型{$w(\mathrm{SiO_2})$、$w(\mathrm{Al_2O_3})$、R、$w(\mathrm{FeO})$、$w(\mathrm{MgO})/w(\mathrm{Al_2O_3})$}的预测性能最好。同样地，其均方根误差为 1.24，与三、四变量模型相比也最小，训练精度最高。五变量模型的平均绝对误差仅为 1.02，误差最低；其平均相对百分比误差最小为 0.014 65，与最大值相差 0.036 67。

表 3　烧结矿低温还原粉化率预测回归模型的误差评判

Table 3　Error evaluation of forecast regression model for low temperature reduction disintegration index of sinter

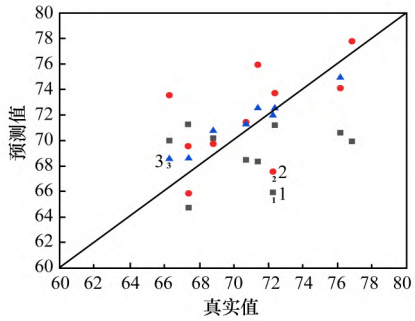
模型	σ_{MSE}	σ_{RMSE}	σ_{MAE}	σ_{MAPE}
{$w(\mathrm{MgO})$、R、$w(\mathrm{CaO})/w(\mathrm{TFe})$}	17.17	4.14	3.68	0.051 32
{$w(\mathrm{SiO_2})$、$w(\mathrm{MgO})$、$w(\mathrm{FeO})$、$w(\mathrm{CaO})/w(\mathrm{TFe})$}	11.02	3.32	2.61	0.037 41
{$w(\mathrm{SiO_2})$、$w(\mathrm{Al_2O_3})$、R、$w(\mathrm{FeO})$、$w(\mathrm{MgO})/w(\mathrm{Al_2O_3})$}	1.54	1.24	1.02	0.014 65

3 个模型的预测值和实际值对比如图 2 所示。由图 2 可以看出，五变量模型的预测值与图中直线更接近，预测更精准。由此可知，五变量模型的预测效果最好。

2.2.2　异常值检验

模型输入数据通常会受到主、客观因素的影响，产生部分异常值点，对模型预测结果产生影响[14-16]。为了检测出这些异常点，需要对模型的杠杆率、Cook 距离、DFITS 值和残差值等统计量进行诊断。

定义第 i 个观测值的杠杆率为帽子矩阵 $\boldsymbol{H}$ 的第 i 个对角线元素，$\boldsymbol{H}$ 定义如式(8)所示。



1—三变量模型；2—四变量模型；3—五变量模型。

图 2　3 个模型的真实值与预测值对比

Fig. 2　Comparison of true value and predicted value of the 3 kinds of models

$$\boldsymbol{H}=\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} \tag{8}$$

Cook 距离的定义如式(9)所示。

$$D_i=\frac{e_i^2}{ps^2}\left[\frac{h_i}{(1-h_i)^2}\right]=\frac{(\boldsymbol{b}-\boldsymbol{b}_{(i)})^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{b}-\boldsymbol{b}_{(i)})}{ps^2} \tag{9}$$

式中：D_i 为保留和删除第 i 个观测值计算的系数之间的距离；s^2 为均方误差；$\boldsymbol{b}$ 为系数向量；$\boldsymbol{b}_{(i)}$ 为删除第 i 个观测值之后计算的系数向量。

DFITS 值的定义如式(10)所示。

$$\sigma_{\mathrm{DFITS}}=\frac{\hat{y}_i-\hat{y}_{i(i)}}{\sqrt{h_i}\sqrt{\sigma_{MSE_{(i)}}}}=e_i\sqrt{\frac{n-p-1}{\sigma_{\mathrm{SSE}}(1-h_i)-e_i^2}}\left[\sqrt{\frac{h_i}{1-h_i}}\right] \tag{10}$$

式中：σ_{DFITS}为每个观测值对回归和方差分析模型中拟合值的影响程度；$\hat{y}_{i(i)}$ 代表删除第 i 个观测值计算的拟合值；$\sigma_{\mathrm{MSE}(i)}$ 为删除第 i 个观测值计算的平均误差平方。

误差检验得到的最优预测回归模型如式(11)所示。

$$Y_{\mathrm{RDI}}=5.71w(\mathrm{SiO_2})-6.26w(\mathrm{Al_2O_3})+27.81R+1.111w(\mathrm{FeO})-24.33w(\mathrm{MgO})/w(\mathrm{Al_2O_3}) \tag{11}$$

根据杠杆率、Cook 距离和 DFITS 值等诊断统计量，发现观测值 2、36、44、46 为异常观测值，如表 4 所示。

对比删除异常值前后的回归结果，分析回归方程的统计量，结果如表 5 所示。由表 5 可得：删除异常值前，$w(\mathrm{FeO})$ 的 P 值为 0.102(大于

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 4 异常观测值的拟合和诊断

Table 4 Fitting and diagnosis of abnormal observations

观测值	低温还原粉化率/%	拟合值/%	残差	标准化残差	杠杆率	Cook 距离	DFITS 值
2	67.920	73.783	-5.863	-2.01	0.043 39	0.036 52	-0.439 80
36	71.310	71.574	-0.264	-0.10	0.082 73	0.044 40	-0.477 65
44	80.740	74.364	6.376	2.18	0.039 27	0.038 76	0.456 31
46	83.540	74.285	9.255	3.15	0.032 17	0.065 90	0.628 20

表 5 删除异常值前后的回归结果对比

Table 5 Comparison of regression results before and after removing outliers

异常值状态	S 值	R^2/%	$R^2_{调整}$/%	$R^2_{预测}$/%	模型方程系数值				P 值		
					$w(FeO)$	$w(SiO_2)$	R	$w(MgO)/w(Al_2O_3)$	$w(Al_2O_3)$	$w(FeO)$	$w(Al_2O_3)$
删除前	2.987 82	99.84	99.83	99.82	1.111	5.71	27.81	-24.33	-6.26	0.102	0.072
删除后	2.440 79	99.90	99.89	99.88	1.137	5.65	26.44	-20.19	-6.07	0.047	0.038

注: P 值为显著性概率; S 值为模型的响应程度。

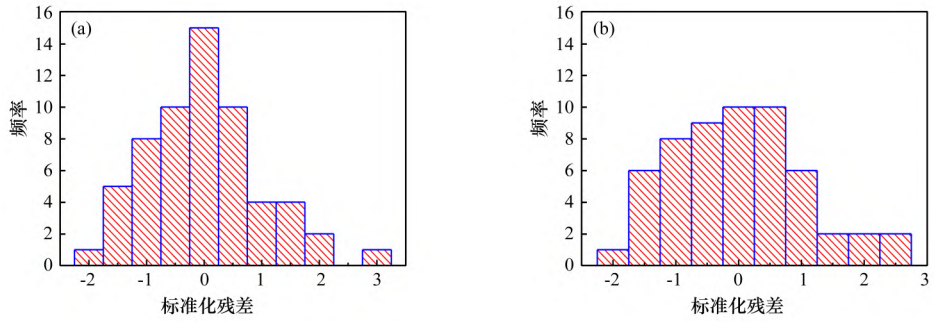
0.05 的显著性水平)，$w(Al_2O_3)$ 的 P 值为 0.072，效应均不显著; 删除异常值后，重新建立回归模型，$w(FeO)$ 的 P 值为 0.047，$w(Al_2O_3)$ 的 P 值为 0.038，均变成显著性因子。新模型的 R^2、$R^2_{调整}$和 $R^2_{预测}$均提高了 0.06%，S 值降低了 0.547 03，预测更为精准。删除异常值后的预测回归模型如式(12)所示。

$$Y_{RDI}=1.137\,w(FeO)+5.56w(SiO_2)+26.44R-20.19w(MgO)/w(Al_2O_3)-6.07w(Al_2O_3) \quad (12)$$

删除异常数据前后的标准化残差直方图、正态概率分布如图 3、4 所示。由图 3 可知: 删除异常数据前，标准化残差为 0 的频率为 15，与标准化残差为 -2 和 3 的频率相差较远较远; 而删除异常数据后，频率最大为 10，最小为 1，不存在某个标准化残差的频率与其他相距较远，删除异常值后效果明显变好，所以新模型的拟合优度更好。图 4(a)中右上角最后一点偏离直线过远，而图 4(b)中未存在某个点偏离直线过远，且各点更加接近直线。可知删除异常值后，样本的正态分布可以接受。删除异常数据前、后的拟合值与残差关系如图 5 所示。图 5(a)中存在两个点在 $y=2$ 的直线上方，且距离均匀分散区域较远。图 5(b)中所有点均匀地分散在 $y=\pm2$ 的直线内，未出现曲线，拟合效果较好。

3 结 论

(1) 采用主成分分析法得到与烧结矿低温还原粉化率相关性大的 8 个影响因素: $w(Al_2O_3)$、$w(SiO_2)$、$w(MgO)$、$w(TiO_2)$、R、$w(FeO)$、$w(MgO)/w(Al_2O_3)$ 和 $w(CaO)/w(TFe)$。

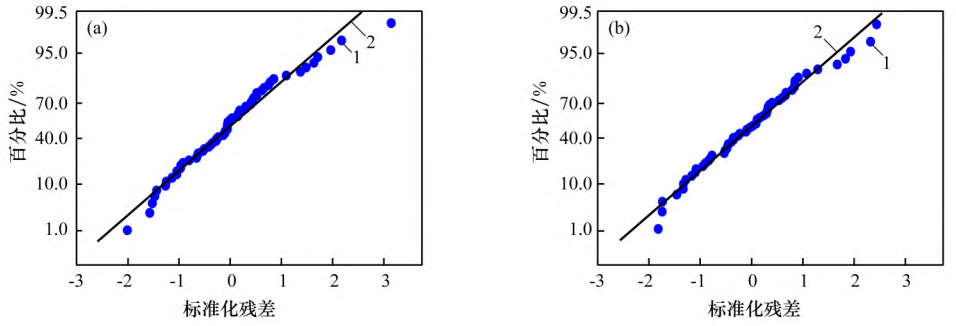


(a) 删除前; (b) 删除后

图 3 删除异常数据前、后的标准化残差直方图

Fig. 3 Standardized residual histogram under complete data and after deleting abnormal data

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

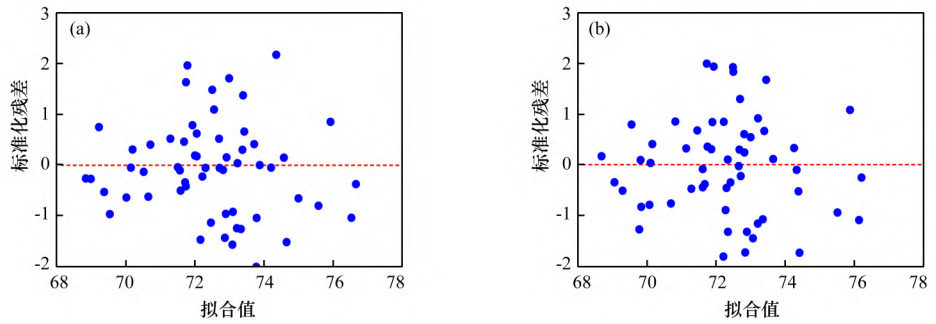


（a）删除前；（b）删除后

1—百分比数值；2—参照线。

图 4　删除异常数据前后的标准化残差正态概率图

Fig. 4　Standardized residual normal probability plot under complete data and after deleting abnormal data



（a）删除前；（b）删除后

图 5　删除异常数据前后的拟合值与残差关系

Fig. 5　Relationship between fitted values and residuals under complete data and after deleting abnormal data

（2）采用最佳子集回归得到 3 种的烧结矿低温还原粉化率预测模型，经过误差函数检验，得到最优模型为{$w(SiO_2)$、$w(Al_2O_3)$、R、$w(FeO)$、$w(MgO)/w(Al_2O_3)$}。

（3）根据杠杆率、Cook 距离和 DFITS 值等诊断统计量对模型进行异常值处理，最终烧结矿低温还原粉化率预测模型 $Y_{RDI}=1.13w(FeO)+5.56w(SiO_2)+26.44R-20.19w(MgO)/w(Al_2O_3)-6.07w(Al_2O_3)$ 的 P 值均小于 0.05，R^2 达到 99.90%，$R^2_{预测}$达到 99.88%，残差效果好。优化后模型效果显著提高，更为精准，能够快速有效地预测和改善烧结矿低温还原粉化率。

参考文献:

[1] 刘然，王龙浩，严照照，等. 化学成分对烧结矿低温还原粉化影响的研究现状[J]. 烧结球团，2018，43(1)：1－5.
LIU Ran，WANG Longhao，YAN Zhaozhao，et al. Present situation of research on impact of chemical composition on low temperature reduction degradation of sinter[J]. Sintering and Pelletizing，2018，43(1)：1－5.

[2] 李杰，信自成，刘卫星，等. 钒钛烧结矿低温还原粉化性能预测[J]. 烧结球团，2017，42(5)：1－5.
LI Jie，XIN Zicheng，LIU Weixing，et al. Prediction of vanadium titanium sinter in low temperature reduction disintegration properties[J]. Sintering and Pelletizing，2017，42(5)：1－5.

[3] 杨广洲，褚利晴. 影响烧结矿低温还原粉化指数因素的定量分析[J]. 烧结球团，2015，40(1)：17－19.
YANG Guangzhou，CHU Liqing. Quantitative analysis of factors influencing low temperature reduction degradation index of sinter[J]. Sintering and Pelletizing，2015，40(1)：17－19.

[4] 杨浩，李一平，蒲亚帅，等. 张家港市河道水质时空分布特征研究分析[J]. 环境科学学报，2021，41(10)：4064－4073.
YANG Hao，LI Yiping，PU Yashuai，et al. Spatio-temporal distribution characteristics and the river water quality of Zhangjiagang city[J]. Acta Scientiae Circumstantiae，

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

2013,44(5):2150 – 2156.

[5] 陈蓉,王明果,黄林培,等. 滇池浮游植物碳氮同位素时空分布特征及其影响因子分析[J]. 中国环境科学,2021(2):843 – 853.
CHEN Rong, WANG Mingguo, HUANG Linpei, et al. Spatio-temporal distribution and influencing factors of stable carbon and nitrogen isotopes of phytoplankton in Dianchi lake [J]. China Environmental Science, 2021(2):843 – 853.

[6] 蒋大军. 应用因子分析探索影响烧结返矿率的主要因素及控制对策[J]. 烧结球团,2013,38(4):17 – 22.
JIANG Dajun. Research on main factors influencing sinter return fines rate by applying factor analysis and counter measures [J]. Sintering and Pelletizing, 2013, 38(4):17 – 22.

[7] 李金珊,杨敏芳,朱维耀,等. 川南筠连沐爱地区煤层含气量预测及控制因素分析[J]. 东北大学学报(自然科学版),2015,36(5):724 – 727;732.
LI Jinshan, YANG Minfang, ZHU Weiyao, et al. Coalbed gas content prediction and controlling factors analysis of coalbed in Junlian Muai area at south of Sichuan [J]. Journal of Northeastern University (Natural Science), 2015,36(5):724 – 727;732.

[8] YANG Xuemei, HAN Henry. Factors analysis of protein O-glycosylation site prediction [J]. Computational Biology And Chemistry, 2017, 71: 258 – 263.

[9] 吕航,刘明遥,苏小四,等. 主因子分析方法在确定地下水石油烃生物降解过程中的应用[J]. 中南大学学报(自然科学版),2013,44(8):3552 – 3560.
LÜ Hang, LIU Mingyao, SU Xiaosi, et al. Determination of biogeochemical processes in oil-contaminated aquifer using principal component analysis [J]. Journal of Central South University (Science and Technology), 2013, 44(8):3552 – 3560.

[10] WANG Jian, ZUO Renguang, CAERS J. Discovering geochemical patterns by factor-based cluster analysis [J]. Journal of Geochemical Exploration, 2017, 181: 106 – 115.

[11] ZHAO Yonghong, WENG Yang, PENG Ningqi, et al. Prediction of mechanical properties of hot rolled strip by using semi-parametric single-index model [J]. Journal of Iron and Steel Research (International), 2013, 20(7): 9 – 15.

[12] NCIBI M C. Applicability of some statistical tools to predict optimum adsorption isotherm after linear and non-linear regression analysis [J]. Journal of Hazardous Materials, 2008, 153(1 – 2): 207 – 212.

[13] SHI Zhenqin, DENG Wei, ZHANG Shaoyao. Spatio-temporal pattern changes of land space in Hengduan mountains during 1990—2015 [J]. Journal of Geographical Sciences, 2018, 28(4): 529 – 542.

[14] 高放,包燕平,王敏,等. 基于FA-ELM的转炉终点磷含量预测模型[J]. 钢铁,2020,55(12):24 – 30.
GAO Fang, BAO Yanping, WANG Min, et al. Prediction model of end-point phosphorus content of converter based on FA-ELM [J]. Iron and Steel, 2020, 55(12): 24 – 30.

[15] IVANESCU A E, LI P, GEORGE B, et al. The importance of prediction model validation and assessment in obesity and nutrition research [J]. International Journal of Obesity, 2016, 40(6): 887 – 894.

[16] MUDHAFAR A, WATHEQ J. Bayesian and LASSO regressions for comparative permeability modeling of sandstone reservoirs [J]. Natural Resources Research, 2019, 28(1): 47 – 62.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net