

基于 SPC 的高炉炉况异常检测研究

肖维民, 臧俊, 袁志祥, 任诗流

安徽工业大学 计算机科学与技术学院, 安徽 马鞍山 243032

摘要: 针对高炉冶炼过程复杂多变, 影响高炉炉况的因素众多且运行过程复杂, 为保证高炉炉况稳定顺行, 开发了一种基于主元分析(PCA)和统计过程控制(SPC)的高炉炉况异常检测模型。由于高炉运行参数具有高耦合和非高斯的特点, 该模型首先采用主元分析算法对高炉实际生产的历史离线数据进行聚类分析, 然后应用基于 T^2 统计量的多元控制图和单值控制图对聚类后的新变量和相关参数进行监测, 从而达到监测高炉出现异常炉况的目的。该模型可以实时监测高炉炉况的异常波动, 其中 PCA 算法将高炉本身的高维数据降到低维, 大幅去除数据中的噪声和不重要特征, 在实际生产和应用中, 节省了大量的成本和时间。选取马钢某高炉炼铁过程为应用场景, 结合数据特点调整和改进算法, 通过历史数据模拟和实时在线运行验证模型的可靠性和算法的有效性, 同时也对指导高炉实际操作技术做出了一定的贡献。

关键词: 高炉; 统计过程控制; 主元分析; 过程监测

中图分类号: TP391 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2023.0001.010

Research on Abnormal Detection of Blast Furnace Condition Based on SPC

XIAO Weimin, ZANG Jun, YUAN Zhixiang, REN Shiliu

School of Computer Science and Technology, Anhui University of Technology, Anhui Ma' anshan 243032, China

Abstract: The smelting process of blast furnace is complex and changeable, many factors affect the conditions of blast furnace, and the operation process is complex. In order to ensure the stable and forward flow of the blast furnace condition, an abnormal detection model of blast furnace conditions based on principal component analysis (PCA) and statistical process control (SPC) has been developed. Due to the high coupling and non-Gaussian characteristics of blast furnace operating parameters, the model first used a principal component analysis algorithm to perform cluster analysis on the historical offline data of the actual production of the blast furnace. Then, the multivariate control chart based on T^2 statistics and single value control chart were used to monitor the new variables and related parameters after clustering, so as to monitor abnormal furnace conditions in the blast furnace. The model can monitor the abnormal fluctuations of blast furnace conditions in real time. The PCA algorithm reduced the high-dimensional data of the blast furnace to low-dimensional, greatly removing noise and unimportant features in the data, saving a lot of cost and time in actual production and application. The ironmaking process of a certain blast furnace in Ma' anshan Iron and Steel was selected as the application scenario, and the algorithm was adjusted and improved in combination with the data characteristics, and the reliability of the model and the effectiveness of the algorithm were verified through historical data simulation and real-time online operation. The achievements in this paper make certain contributions to guiding the actual operation technology of the blast furnace.

Keywords: blast furnace; statistical process control; principal component analysis; process monitoring

收稿日期: 2021-03-05 **修回日期:** 2021-05-18 **文章编号:** 1672-058X(2023)01-0059-05

基金项目: 国家自然科学基金青年项目(61806005).

作者简介: 肖维民(1962—), 男, 安徽无为, 副教授, 本科, 从事计算机网络应用研究.

通讯作者: 臧俊(1997—), 男, 安徽马鞍山, 硕士研究生, 从事计算机应用研究. Email: 1570651154@qq.com.

引用格式: 肖维民, 臧俊, 袁志祥, 等. 基于 SPC 的高炉炉况异常检测研究[J]. 重庆工商大学学报(自然科学版), 2023, 40(1): 59—63.

XIAO Weimin, ZANG Jun, YUAN Zhixiang, et al. Research on abnormal detection of blast furnace condition based on SPC[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2023, 40(1): 59—63.

1 引言

钢铁工业是我国国民经济的基础产业,也是我国发展现代化建设的重要角色之一。高炉炼铁是钢铁工业中必不可少的关键部分,它不仅在我国炼铁工艺中占据绝对主导地位,同样也处于世界炼铁工艺的主体地位^[1]。在高炉冶炼过程中,原燃料条件、送风制度、装料制度、造渣制度等因素中任何一个的变化都会导致炉况出现或大或小的波动。若高炉长时间处于异常状态,轻则造成铁水质量异常,产能损失,重则严重浪费人力、物力,同时安全风险高,易出现安全事故。因此,对高炉炼铁过程的炉况进行实时监控和预警,降低高炉出现异常炉况的概率从而预防高炉事故的发生,对钢铁行业来说有着重要的意义^[2]。

不同的高炉类型、高炉的不同时期、不同高炉冶炼特点的不同,以及其高炉冶炼具有滞后性,所以高炉冶炼是一个复杂多变的物理化学过程。高炉可以看作是一个“黑匣子”,操作者只能通过监控和分析高炉各项参数的变化对高炉的生产过程进行评估,而操作者自身的经验水平、主观认识都不相同,这使得对高炉炉况的认识缺乏科学的方法。因此,科学、客观地判断炉况是炼铁工作者急需解决的课题。而早在 20 世纪 80 年代,日本首先开发出了高炉专家系统,日本川崎的 GO-STOP 系统通过 8 个高炉指数采取信息,构建了专家经验规则,对高炉冶炼过程进行评判^[3]。现代大多数高炉专家系统都采用了 GO-STOP 系统的理论基础。黄波^[4]等在武钢应用专家系统的前提下,于 2005 年根据多变量技术对高炉生产过程中的过程参数进行处理和数据分析。该方法有效地预判了炉温的变化趋势,也能提前 2 h 对铁水硅含量进行预报。

此外,随着技术日新月异的发展,基于机器学习的高炉炉况检测方法也有了很大的发展,取得了较多的研究成果。2010 年, Tian^[5] 等将引导聚集算法和 SVM 算法集成,建立了一种新型高炉炉况诊断系统。该方法证明了 SVM 集成方法的性能要优于单一 SVM 方法。2011 年, Liu^[6] 等改进 DPSP 算法加以优化 LS-SVM 算法的性能,阐述了一种基于成本意识的 LS-SVM 多类分类方法。该方法加快了处理速度,优化了分类精度,具有良好的适应能力。2020 年,陈伟^[7] 等以高炉冶炼原理和专家经验为基础,通过线性规划和 BP 神经网络等技术方法,开发了炉料结构智能优化系统。在保证烧结矿质量和铁水产能前提下,降低了炼铁成本。

但以上方法都存在一定的不足之处。专家系统是针对性地去解决部分问题;机器学习方法需要大量的高炉故障数据作为训练集,才能保证其准确率。然而高炉实际生产中发生异常的次数不多,且历史故障数据保存不完整,因此很难提供大量的故障数据作为训练集。相比之下,基于统计分析的方法优点在于不

需要了解很多的机理知识和过程的前因后果关系,且对异常数据没有过多的要求。因此,本文结合高炉数据的特点,调整和改进主元分析算法(Principal Component Analysis, PCA)和统计过程控制方法(Statistical Process Control, SPC),构建高炉炉况异常检测模型,通过离线数据模拟和在线运行进行实验。结果表明,本方法可快速监测高炉炉况,减少高炉工长自身主观意识,对保持高炉顺行提供了重要的保障。

2 PCA 算法

PCA 是一种常用于高维数据降维的数据分析方法。PCA 保留了绝大部分主要信息,减少噪声干扰和去除一些不重要的特征,从而达到加快数据处理速度的目的。

对训练集进行 PCA 前,首先要对训练集进行中心化处理,均值必须从训练集计算而来。假设中心化后的数据为矩阵 $\mathbf{X}(M \times N)$,其中 M 为变量数目, N 为训练集的样本数目,如下:

$$\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{pmatrix} \quad (1)$$

计算 M 个变量的相关系数矩阵 $\mathbf{R} = (r_{ij})_{(MN)}$, r_{ij} 是原第 i 个变量和第 j 个变量的相关系数。计算公式如下:

$$r_{ij} = \frac{\sum_{k=1}^n (X_{ik} - X_i)(X_{jk} - X_j)}{\sqrt{\sum_{k=1}^n (X_{ik} - X_i)^2 \sum_{k=1}^n (X_{jk} - X_j)^2}} \quad (2)$$

系数矩阵 \mathbf{R} 的特征值通过特征方程 $|\lambda_i - \mathbf{R}| = 0$ 得到,并对其降序排列, $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$;继而得到相应的特征向量 \mathbf{e} ,最后,主元可以表示如下:

$$\mathbf{Z}_i = \sum_{j=1}^n \alpha_{ij} X_j \quad (3)$$

计算各个主元的贡献率和累计贡献率,第 P 个主元贡献率计算公式如下:

$$g_p = \frac{\lambda_p}{\sum_{k=1}^m \lambda_k} \quad (4)$$

则前 P 个主元的累积贡献率计算公式如下:

$$S_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^m \lambda_k} \quad (5)$$

累计贡献率越大,说明新的向量空间保留了越多的数据信息,一般当累计贡献率达到或大于 85% 时,可取这些成分代替原始向量空间^[8-11]。

PCA 模型一般使用 T^2 统计量或 SPE 统计量对过程进行故障判断。其中 T^2 统计量和 SPE 统计量的计算公式如下:

$$T^2 = \mathbf{z}^T \mathbf{S}^{-1} \mathbf{z} \quad (6)$$

$$SPE = \|(I - PP^T)X\|^2 \quad (7)$$

其中, z 为主元向量, $S = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$, S^{-1} 表示对矩阵中的非零元素取-1 指数。 X 是新样本归一化后的数据。

若样本近似高斯分布, T^2 统计量可作为服从 F 分布, T^2 统计量和 SPE 统计量的控制限公式如下:

$$T^2_{UCL} = \frac{p(N-1)}{N-p} F_{\alpha}(p, N-p) \quad (8)$$

$$SPE_{UCL} = \theta_1 \left[\frac{c_{\alpha} h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (9)$$

其中, $F_{\alpha}(p, N-p)$ 是服从第一自由度为 P , 第二自由度为 $N-P$ 的 F 分布。 $\theta_r = \sum_{j=k+1}^m \lambda_j^r (r = 1, 2, 3)$, $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$, α 是标准正态分布的置信极限, 一般取值 0.01 左右。可通过上述统计量是否超过置信极限下的阈值来判断过程有无发生异常^[9]。

3 SPC 基本原理

SPC 最早是由美国的 Shewhart W A 博士提出, 并率先使用的控制图理论。但 SPC 真正得到发展和应用是在 20 世纪 70 年代, 战后经济极度萎缩的日本采用 SPC 快速恢复了经济, 并跃居全球质量和生产率的领先地位。自此, SPC 开始大规模应用到制造业中, 并逐渐成为企业中不可缺少的工具和质量保证手段^[12]。美国著名的质量管理大师 Joseph h Juran 说: “21 世纪是质量的世纪^[13]。”

SPC 借助数理统计的方法对过程中的信息和数据进行监测, 而过程中的数据一定会出现变差, SPC 关注的是变差。将变差数据用图表的形式表现出来可直观看出一定的规律, 从而有一定的预警作用。SPC 主要采用控制图理论预防生产过程中错误的发生, 拉依达准则是控制图理论的判定方法, 由上控制限 (Upper Control Limit, UCL)、下控制限 (Lower Control Limit, LCL) 和中心线 (Center Line, CL), 并有按时间排序的样本数值的序列构成^[14-16]。

控制图根据不同的数据类型可分为计数型和计量型, 不同的情况使用不同类型的控制图。针对高炉生产过程中质量特性数据多、获取费时且花费成本也很高的问题本文采用单值-移动极差图中的单值控制图对高炉炉况的测量值进行监测。单值控制图控制上限和控制下限计算公式如下:

$$UCL_x = X + E_2 \bar{R} \quad (10)$$

$$LCL_x = X - E_2 \bar{R} \quad (11)$$

单值移动极差图将数据按时间序列标记在图上, 便于迅速发现过程中的异常波动, 并对其采取相应的措施, 这很适合高炉过程的实际应用情况。而另一方

面将单值控制图和基于 PCA 的 T^2 统计量控制图相结合, 二者共同决策出什么因素更为主要地导致了高炉炉况的波动。

4 基于现场数据的离线测试

4.1 数据预处理

本文的实验数据来源于某钢铁厂的一座 2 500 m³ 的 2 号高炉。在高炉炼铁过程中每 5 s 对其过程数据进行采集, 含有 15 个高炉炉况主要参数, 一天共 17 280 条数据。具体参数如表 1 所示。本文选用了从 2019 年 1 月 10 日到 2 月 10 日, 共 32 天的数据。

表 1 2 500 m³ 高炉生产参数
Table 1 Process parameters of 2 500 m³ BF

序号	简称	参数名/单位符号
1	BF2_FW	风温/°C
2	BF2_TQXZS	K 值(透气性指数)
3	BF2_RFYL	热风压力/kPa
4	BF2_LDYL	炉顶压力/kPa
5	BF2_QYC	全压差/MPa
6	BF2_SD	湿度/g/m ³
7	BF2_LDWD	炉顶温度/°C
8	BF2_MQLYL	煤气利用率/%
9	BF2_Z	Z 值(中心流指数)
10	BF2_W	W 值(边缘流指数)
11	BF2_FY	富氧/m ³ /h
12	BF2_CCT	CCT 值(十字测温中心温度)/°C
13	BF2_TC	探尺差/mm
14	BF2_FJFL	风机风量/m ³ /min
15	BF2_LFLL	冷风流量/m ³ /min

这些高炉生产参数从自身来说存在着高耦合、非稳态、非高斯的复杂多样性数据, 影响高炉炉况的因素众多且关系复杂。利用本文第 2 节中的 PCA 算法步骤, 基于以上 15 个变量参数, 对其进行降维处理。在保留多数信息的情况下用少量的几个综合变量取代原变量参数进行分析, 其结果如表 2 所示。

表 2 降维后的主元贡献率
Table 2 Contribution rate of principal components after dimension reduction

主元	λ_i	g_p	S_p
Z_1	5.103 0	0.340	0.340
Z_2	2.677 2	0.178	0.519
Z_3	1.965 0	0.131	0.650
Z_4	1.384 0	0.092	0.742
Z_5	1.138 5	0.076	0.818
Z_6	0.985 0	0.066	0.884
Z_7	0.673 5	0.045	0.928

由表 2 可看出, 前 6 个主元已经包括了原数据的 88% 的信息, 根据主元分析理论, 已经可以通过前 6 个主元来代替原始数据集研究高炉炉况的原因。同时也可得到 15 个变量和 6 个主元之间的关系, 具体如表 3 所示。

表 3 主元系数矩阵

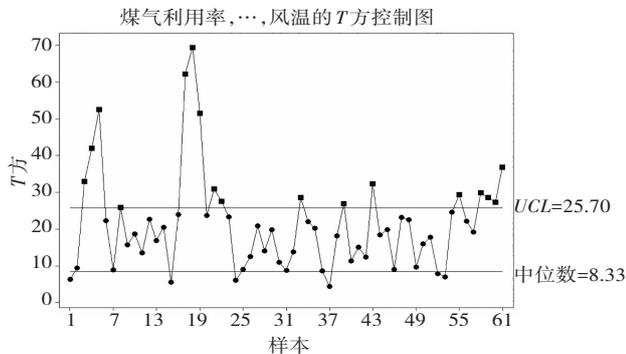
Table 3 Matrix of principal component coefficients

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
X_1	0.423	0.023	0.002	-0.087	0.002	-0.060
X_2	-0.170	-0.215	0.524	0.035	0.319	-0.146
X_3	0.038	0.071	-0.247	-0.577	0.219	-0.420
X_4	0.025	0.059	0.061	-0.299	0.627	0.554
X_5	-0.352	0.078	0.231	-0.143	0.145	-0.347
X_6	0.357	-0.078	0.018	0.237	0.328	0.081
X_7	-0.227	0.319	-0.273	0.329	0.182	0.084
X_8	0.007	0.050	-0.368	0.322	0.433	-0.464
X_9	0.335	-0.313	0.143	-0.055	0.203	-0.033
X_{10}	0.373	-0.047	-0.345	-0.051	-0.119	0.101
X_{11}	-0.253	-0.260	-0.400	0.102	0.206	0.141
X_{12}	0.254	0.460	0.177	0.066	0.069	-0.117
X_{13}	-0.250	-0.470	-0.158	-0.081	-0.058	0.113
X_{14}	-0.189	0.427	0.061	0.053	0.026	0.282
X_{15}	0.092	-0.214	0.196	0.501	0.039	-0.067

在实际生产中,由于每个主元都有 15 个影响因素,很难直接监测到主元,且当某个主元出现异常时,很难对其调整,因此,将每个主元中影响系数最高的变量作为该主元的关键因素。由表 3 可知,风温、风机风量、 K 值、冷风流量、煤气利用率和炉顶压力分别为对应前 6 个主元的关键控制参数。

4.2 构建控制图

根据上一小节得到的前 6 个主元,通过式(6)、式(8)计算出 T^2 统计量的控制上限,并构造出控制图,如图 1 所示。

图 1 T^2 控制图Fig. 1 T^2 control chart

由图 1 可看出,共有 17 个样本点不在控制界限内。但是这仅仅只能得出哪些时刻样本是否处于控制界限内,但无法具体看出是哪些变量影响了样本点,导致不受控制。因此,根据式(10)、式(11)再对以上 6 个影响主元的关键因素构建单值控制图。

由图 1—图 6 分析,图 1 中的第 3、4、5 样本点出现了失控情况,而在图 2 中,炉顶压力单值控制图也出现了失控情况,说明炉顶压力在样本点 3、4、5 的失控状况产生了很大的影响,而在样本点 3 中,图 6 中 K 值也出现失控情况,说明 K 值也对该点失控也有一定的影响。通过分析, T^2 统计量是各变量影响的最终结果,因此,可通过观察该值来判断高炉炉况是否处于失控状

态,若发现某个时刻的 T^2 统计量超出控制限,则查看各变量的单值控制图,从而达到确定某个因素导致的高炉炉况不在控制范围内。

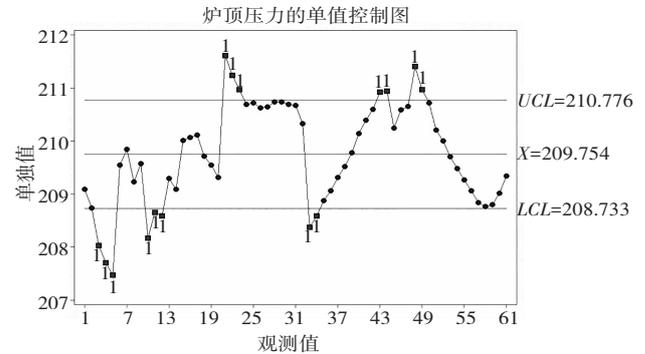


图 2 炉顶压力单值控制图

Fig. 2 Single value control chart of TP

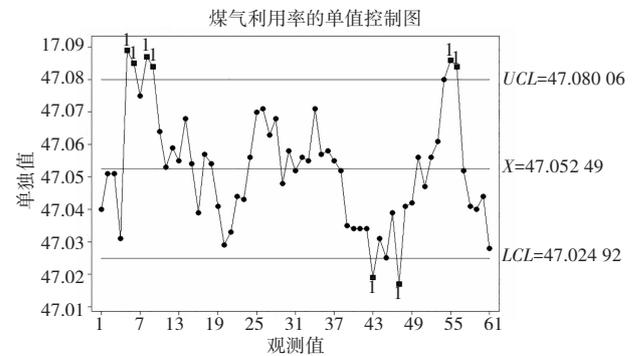


图 3 煤气利用率单值控制图

Fig. 3 Single value control chart of utilization rate of gas

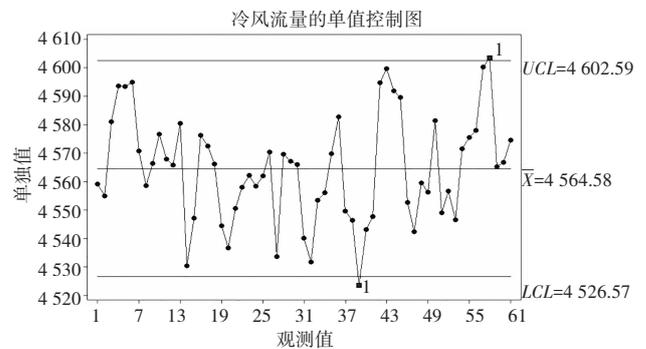


图 4 冷风流量单值控制图

Fig. 4 Single value control chart of cold-wind volume

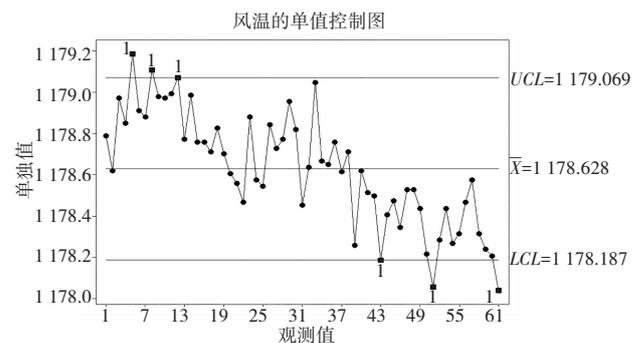


图 5 风温单值控制图

Fig. 5 Single value control chart of BT

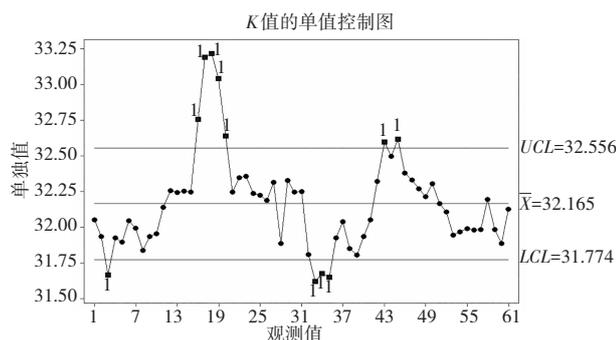


图 6 K 值单值控制图
Fig. 6 Single value control chart of K

5 结束语

通过某钢铁厂历年生产数据,采用 SPC 中的 T^2 控制图和单值控制图建立了高炉炉况异常监测模型, T^2 控制图主要监测高炉炉况的变化情况,单值控制图主要提供高炉工长分析什么因素导致的炉况波动,二者相互起到了互补作用,可以对高炉过程控制参数进行监控并建立起相关标准值,有效地改善了某钢铁厂高炉体检制度的全新体检标准。同时将该模型应用到了基于 Spark 的大数据平台中。根据实时运行参数与建立起来的标准值进行对比,更为直观地表现出高炉顺行状态。最后形成的参数监控模式更为有效地监控高炉炉况,保证了最后的产品质量的稳定性。本文证实了 SPC 技术在高炉冶炼过程在线实时监测的可能性,为钢铁企业把握铁水质量提供了一种新思路。

参考文献 (References):

- [1] 王新东,郝良元. 现代炼铁工艺及低碳发展方向分析[J]. 中国冶金, 2021, 31(5): 1-10.
WANG Xin-dong, HAO Liang-yuan. Analysis of modern ironmaking technology and lowcarbon development direction [J]. China Metallurgy, 2021, 31(5): 1-10.
- [2] 赵宏博,霍守锋,郝经伟,等. 高炉炉缸的安全预警机制[J]. 钢铁, 2013, 48(4): 24-31.
ZHAO Hong-bo, HUO Shou-feng, HAO Jing-wei, et al. Discussion on security early warning mechanism of BF hearth [J]. Iron and Steel, 2013, 48(4): 24-31.
- [3] YAMAMOTO T, SAWADA T, SHINOHARA K, et al. Blast furnace operational system with the application of advanced go/stop system at Mizushima works [C]//IISC. The Sixth International Iron and Steel Congress, 1990, 2: 364-371.
- [4] 黄波,汪卫. 基于高炉专家系统的多变量炉温智能控制系统[J]. 钢铁, 2005, 40(4): 21-27.
HUANG Bo, WANG Wei. Multi-variable furnace temperature intelligent control system based on BF expert system[J]. Iron and Steel, 2005, 40(4): 21-27.
- [5] TIAN H, WANG A. A novel fault diagnosis system for blast furnace based on support vector machine ensemble [J]. ISIJ International, 2010, 50(5): 738-742.
- [6] LIU L M, WANG A N, SHA M, et al. Multi-class classification methods of cost-conscious LS-SVM for fault diagnosis of blast furnace [J]. Journal of Iron and Steel Research (International), 2011, (10): 20-26, 36.
- [7] 陈伟,武鹏飞,王宝祥,等. 烧结-高炉炼铁全过程炉料结构智能优化系统[J]. 烧结球团, 2020, 45(5): 8-13, 24.
CHEN Wei, WU Peng-fei, WANG Bao-xiang, et al. Intelligent optimization system of burden structure in the whole process of sintering-blast furnace iron-making [J]. Sintering and Pelletizing, 2020, 45(5): 8-13, 24.
- [8] ZHANG T S, YE H, WANG W, et al. Fault diagnosis for blast furnace iron-making process based on two-stage principal component analysis [J]. ISIJ International, 2014, 54(10): 2334-2341.
- [9] 张杰,阳宪惠. 多变量统计过程控制[M]. 北京: 化学工业出版社, 2000: 24, 25, 60-62.
ZHANG Jie, YANG Xian-hui. Multivariate statistical process control [M]. Beijing: Chemical Industry Press, 2000: 24, 25, 60-62.
- [10] 徐海鹏. 基于数据驱动的复杂工业过程故障诊断方法研究[D]. 北京: 清华大学自动化系, 2012.
XU Hai-peng. Research on fault diagnosis method of complex industrial process based on data-driven [D]. Beijing: Department of Automation, Tsinghua University, 2012.
- [11] 刘慧,李沛然,包哲静,等. 高炉生产过程的智能预测[J]. 中南大学学报, 2012, 43(5): 1787-1794.
LIU Hui, LI Pei-ran, BAO Zhe-jing, et al. Intelligent predictive modeling of blast furnace system [J]. Journal of Central South University, 2012, 43(5): 1787-1794.
- [12] 钱夕元,荆建芬,侯旭暹. 统计过程控制 (SPC) 及其应用研究[J]. 计算机工程, 2004, 30(19): 144-145, 154.
QIAN Xi-yuan, JING Jian-fen, HOU Xu-xian. Research on statistical process control (SPC) and its application [J]. Computer Engineering, 2004, 30(19): 144-145, 154.
- [13] JURAN J M, B GODFREY A. 朱兰质量手册 (第五版) [M]. 焦书斌,等,译. 北京: 中国人民大学出版社, 2014.
JURAN J M, BLANTON G A. Juran's quality handbook (fifth edition) [M]. JIAO Shu-bing, et al. Translate. Beijing: China Renmin University Press, 2014.
- [14] LIANG J, QIAN J X. Multivariate statistical process monitoring and control: recent developments and applications to chemical industry [J]. Chinese Journal of Chemical Engineering, 2003, 11(2): 191-203.
- [15] 黄忠山,田凌,向东,等. 基于 PCA 和 SPC-动态神经网络的风电机组齿轮箱油温趋势预测[J]. 清华大学学报(自然科学版), 2018, 58(6): 539-546.
HUANG Zhong-shan, TIAN Ling, XIANG Dong, et al. Prediction of oil temperature variations in a wind turbine gearbox based on PCA and an SPC-dynamic neural network hybrid [J]. Journal of Tsinghua University (Science and Technology), 2018, 58(6): 539-546.
- [16] 田学森. SPC 统计过程控制在铸造过程质量管理中的应用[J]. 铸造设备与工艺, 2020(3): 53-57, 64.
TIAN Xue-sen. Application of SPC statistical process control in casting process quality management [J]. Foundry Equipment & Technology, 2020(3): 53-57, 64.