

冶金生产过程 质量监控理论与方法

徐金梧 等著



冶金工业出版社
Metallurgical Industry Press

冶金生产过程 质量监控理论与方法

徐金梧 等著

北 京
冶金工业出版社
2015

内 容 提 要

本书结合冶金生产过程的特点,系统地介绍适用于钢铁企业的产品质量监控理论与方法,具体包括:数据样本和数理统计的基础知识、单变量统计过程控制、多变量统计过程控制、历史数据集的构造、生产过程的优化控制、非线性预测与诊断、全流程质量监控系统的框架等,并在此基础上结合生产实际数据给出应用实例。

本书可为钢铁企业从事产品质量分析的工程技术人员和其他科研人员提供指导和参考,同时也适合高等院校及相关科研单位的研究人员使用。

图书在版编目(CIP)数据

冶金生产过程质量监控理论与方法/徐金梧等著. —北京:
冶金工业出版社, 2015. 5

ISBN 978-7-5024-6901-6

I. ①冶… II. ①徐… III. ①冶金—生产工艺—质量
控制 IV. ①TF1

中国版本图书馆 CIP 数据核字(2015)第 096825 号

出 版 人 谭学余

地 址 北京市东城区嵩祝院北巷 39 号 邮编 100009 电话 (010)64027926

网 址 www.cnmp.com.cn 电子信箱 yjcb@cnmp.com.cn

责任编辑 戈 兰 唐晶晶 美术编辑 彭子赫 版式设计 孙跃红

责任校对 石 静 责任印制 李玉山

ISBN 978-7-5024-6901-6

冶金工业出版社出版发行;各地新华书店经销;三河市双峰印刷装订有限公司印刷

2015 年 5 月第 1 版, 2015 年 5 月第 1 次印刷

787mm×1092mm 1/16; 19.25 印张; 465 千字; 297 页

78.00 元

冶金工业出版社 投稿电话 (010)64027932 投稿信箱 tougao@cnmp.com.cn

冶金工业出版社营销中心 电话 (010)64044283 传真 (010)64027893

冶金书店 地址 北京市东四西大街 46 号(100010) 电话 (010)65289081(兼传真)

冶金工业出版社天猫旗舰店 yjgybs.tmall.com

(本书如有印装质量问题,本社营销中心负责退换)

前 言

在我国钢铁工业从钢铁大国迈向钢铁强国的进程中，钢铁企业对高端产品的自主研发和高端产品质量的稳定性提出了迫切需求。高端产品的质量已成为钢铁企业市场竞争力的关键要素，尤其是钢铁行业在拓展产业服务链过程中，产品质量的稳定性成为企业核心竞争力。在市场竞争的五大要素——品种、质量、价格、服务和交货期中，决定竞争胜负的要素是质量。正如美国质量管理专家朱兰（J. M. Juran）于1994年在美国质量管理学会会上所说的，20世纪以“生产力的世纪”载入史册，21世纪将是“质量的世纪”。

冶金、化工是典型的大型复杂流程工业，由功能不同但又相互关联、相互支撑、相互制约的各个工序和各种装置，通过工序间串联、并联方式集成后构成完整的复杂生产制备系统。针对冶金工业的特点，如何寻求工艺参数与产品质量间的统计规律，建立更加精准、适用性更强的质量控制模型；如何建立产品质量在线监控系统，实时检测产品质量状况，避免造成批量的质量判废；如何通过优化工艺流程和工艺参数，确保产品的质量、减少企业的经济损失，已成为我国钢铁工业重大的研究课题。

我国大型钢铁企业从20世纪末开始贯彻 6σ 质量管理标准，要求将传统的事后检验提升为对生产全过程的监控，采用基于统计过程控制的质量管理方法来提高产品质量。统计过程控制（Statistical Process Control, SPC）是一种借助数理统计方法的过程控制工具。统计过程控制实质上是建立工艺参数与质量指标间的统计规律，通过监测原料数据、各工序的工艺参数来推断目前的产品质量状况，诊断产品质量异常的原因，优化产品质量的控制。但是，由于受测量技术以及数据存储和分析技术的限制，传统的统计过程控制主要采用单变量统计过程控制方法，只对生产过程中的少数几个重要指标单独地实施统计过程控制，比如单变量休哈特控制图。

直到20世纪80年代，一些新的多元分析方法，如主成分分析、偏最小二乘法，尤其是基于核（kernel）的非线性分析方法取得了重大突破，将多变量统计分析方法融入传统的统计过程控制，形成了多变量统计过程控制的基本框

架。主成分分析和偏最小二乘法采用多元投影方法,将工艺参数和质量数据从高维数据空间投影到低维特征空间,保留原始数据的特征信息,大大降低了多元统计分析中的复杂性,提高了分析过程的准确性,是一种高维数据分析处理的有效工具。最近几年,基于核函数映射的机器学习方法为解决多变量的非线性统计过程控制提供了新的途径。这些新的理论和方法为解决高维、强耦合、非线性的连续过程的多变量统计过程控制提供了有效的手段。

随着计算机系统、数据库技术的普及与应用,我国钢铁企业信息化系统中拥有了丰富的生产数据资源,从而也提出了采用各种数值分析方法对大量的工艺过程数据和产品质量数据进行全流程产品质量监控的迫切需求。目的是通过大量生产数据分析来揭示、总结生产过程的内在规律,为提高产品质量提供各种信息,从而把数据资源转化为企业的经济效益和产品质量优势,提高产品的市场竞争力。在大数据时代,如何从海量数据中提取出有价值的信息,从而完善产品质量管理体系,实现全流程产品质量在线监控,已成为全行业提高管理水平和经济效益的必然趋势。因此,建立适用于钢铁企业的质量监控理论与方法,研发全流程冶金产品质量分析与过程监控系统,对加速新钢种研发的进程、减少研发成本以及控制产品质量稳定性、提高产品的竞争力具有十分重要的现实意义。

本书结合冶金生产过程的特点,介绍了一些适用于钢铁企业的质量监控理论与方法,并结合生产实际数据给出应用实例。另外,还讨论了冶金全流程质量监控系统的基本框架,为钢铁企业从事产品质量分析的工程技术人员和其他科研人员提供一些指导和参考。本书第1章~第8章由徐金梧、黎敏撰写,第9章由何飞撰写,第10章由徐钢撰写,最后由徐金梧负责统稿。本书的研究内容和写作过程,包含了历届研究生和很多同事的研究成果,在此要特别感谢姚林、王建国、赵晨熙等博士和郑杰等硕士,以及吕志民教授、阳建宏副教授张文兴讲师等同事。书中的部分实例来源于钢铁企业的实际生产数据,在本书出版之际,作者愿向这些企业的科技工作者致以诚挚的谢意。

本书的出版得到了北京科技大学“十二五”教材建设基金的资助,本书的顺利出版还得到了冶金工业出版社的大力支持和帮助,感谢他们为本书所付出的心血和辛勤工作。

由于作者的学术水平有限,写作过程中难免存在不足之处,敬请读者批评指正。

作 者

2015年1月于北京科技大学

目 录

符号表	1
1 绪论	3
1.1 冶金生产过程的特点	3
1.2 质量监控技术的现状与发展趋势	5
1.3 本书各章节内容	7
2 数据样本的基础知识	9
2.1 向量	9
2.1.1 向量的定义	9
2.1.2 向量的长度	9
2.1.3 向量的夹角	9
2.1.4 向量的投影	10
2.2 矩阵	11
2.2.1 矩阵的定义	11
2.2.2 矩阵的运算	12
2.2.3 行列式	12
2.2.4 逆矩阵	13
2.2.5 特征值与特征向量	14
2.3 样本空间	14
2.3.1 基本定义	14
2.3.2 样本点间的欧氏距离	16
2.3.3 样本点间的统计距离	16
2.3.4 样本点间的马氏距离	18
2.4 变量空间	19
2.4.1 变量的均值	19
2.4.2 变量的方差	20
2.4.3 变量间的协方差	20
2.4.4 变量间的相关系数	20
2.5 数据的预处理	21
2.5.1 中心化处理	21
2.5.2 无量纲化处理	21
2.5.3 标准化处理	22

2.6 应用举例	22
2.6.1 均值	22
2.6.2 方差	23
2.6.3 协方差	23
2.6.4 相关系数	23
2.6.5 欧氏距离	24
2.6.6 统计距离	24
2.6.7 马氏距离	25
2.6.8 无量纲化	25
2.7 小结	26
3 数理统计的基础知识	27
3.1 基本概念	27
3.1.1 总体与个体	27
3.1.2 概率密度函数	28
3.1.3 数学期望和方差	29
3.2 几个重要分布	30
3.2.1 正态分布	30
3.2.2 χ^2 分布	32
3.2.3 F 分布	33
3.2.4 β 分布	35
3.2.5 t 分布	36
3.3 参数估计	38
3.3.1 总体均值的点估计	38
3.3.2 总体方差的点估计	38
3.3.3 总体均值的区间估计	39
3.3.4 总体方差的区间估计	40
3.4 假设检验	42
3.4.1 零假设与研究假设	42
3.4.2 显著性水平	43
3.4.3 假设检验与置信区间的关系	43
3.5 应用举例	45
3.5.1 回归模型的线性关系检验—— F 检验	45
3.5.2 总体回归参数的检验—— t 检验	46
3.6 小结	47
4 单变量统计过程控制	49
4.1 统计过程控制简介	49
4.1.1 统计过程控制的基本概念	49
4.1.2 工业过程中的各种变化	50

4.2 统计控制图	53
4.2.1 控制图的定义	53
4.2.2 控制图的基本原理	53
4.2.3 常用控制图	55
4.2.4 控制图的风险	62
4.2.5 判稳准则和判异准则	63
4.2.6 应用统计控制图时需要注意的问题	67
4.3 过程能力指数	68
4.3.1 过程能力	68
4.3.2 过程能力指数的计算	69
4.3.3 过程能力指数与不合格品率之间的关系	72
4.3.4 给定置信度下的过程能力指数	74
4.3.5 过程能力分析的功能与步骤	75
4.4 应用举例	76
4.4.1 实例1	76
4.4.2 实例2	78
4.5 小结	81
5 多变量统计过程控制	82
5.1 多变量统计过程的意义和研究现状	82
5.2 霍特林 T^2 控制图	84
5.2.1 霍特林 T^2 控制图的提出	84
5.2.2 T^2 统计量的分布特性	87
5.2.3 T^2 的正交分解	89
5.3 基于主元模型的多变量统计控制图	94
5.3.1 主成分分析方法 (PCA)	94
5.3.2 偏最小二乘法 (PLS)	96
5.3.3 多变量统计过程的监控	101
5.3.4 多变量统计过程的诊断	105
5.4 多变量统计过程应用实例	106
5.4.1 镀锌板表面粗糙度统计过程控制	106
5.4.2 某化工过程统计过程控制	109
5.5 三种统计量之间的关系	113
5.5.1 几何解释	113
5.5.2 数学推导	114
5.6 小结	115
6 历史数据集的建立	116
6.1 建立历史数据集的过程及数据预处理	116
6.2 数据采集	117
6.2.1 冶金生产数据的基本特征	117

6.2.2 数据的采集方式	118
6.2.3 数据的预处理	119
6.3 变量间的多重相关性	119
6.3.1 多重相关性产生的原因	120
6.3.2 多重相关性的影响	120
6.3.3 多重相关性的检验方法	121
6.3.4 多重相关性的解决方法	122
6.4 变量的自相关性	123
6.4.1 自相关性的数学描述	123
6.4.2 自相关性产生的原因	124
6.4.3 自相关性的影响	126
6.4.4 自相关性的检验方法	128
6.4.5 自相关性的解决方法	130
6.5 删除异常点	134
6.5.1 异常点产生的原因	134
6.5.2 异常点的检验方法	135
6.5.3 异常点的剔除方法	137
6.5.4 应用实例	142
6.6 小结	143
7 生产过程的优化控制	145
7.1 优化控制的基本概念	145
7.1.1 优化控制的几个环节	145
7.1.2 优化控制的基本框架	146
7.2 基于规则的工艺参数设定与优化	147
7.2.1 常用的工艺参数设定方法	147
7.2.2 基于神经网络的规则提取流程	148
7.2.3 规则的提取	152
7.2.4 工艺参数的优化	157
7.3 基于数据驱动的工艺参数动态调整	161
7.3.1 数据驱动的基本概念	162
7.3.2 数据驱动控制器的设计	163
7.3.3 数据驱动控制器的应用	167
7.4 工艺流程的优化	169
7.4.1 工艺装备的能力	170
7.4.2 工艺装备的能力分析	170
7.4.3 工艺装备的能力优化	171

7.5 小结	171
8 非线性预测与诊断	172
8.1 核函数的基本原理	172
8.1.1 核函数	172
8.1.2 核矩阵	175
8.1.3 常用的核函数及其特点	175
8.1.4 核函数的性质和特征空间的计算	177
8.1.5 核函数方法实施步骤	178
8.2 核主成分分析的过程监控与诊断	179
8.2.1 核矩阵与协方差矩阵	179
8.2.2 核主成分	182
8.2.3 核空间的数据重构	183
8.2.4 基于核方法的监控模型	184
8.2.5 应用实例	185
8.3 核偏最小二乘的质量预测方法	188
8.3.1 核偏最小二乘法的基本原理	189
8.3.2 核偏最小二乘的预测模型	190
8.3.3 应用实例	192
8.4 支持向量机的质量预测方法	194
8.4.1 支持向量机的基本原理	194
8.4.2 支持向量回归	198
8.4.3 应用实例	200
8.5 数据缺失情况下质量预测方法	201
8.5.1 流形半监督学习	201
8.5.2 核岭回归的基本原理	203
8.5.3 基于流形学习的半监督核岭回归预测模型	204
8.5.4 应用实例	206
8.6 小结	207
9 案例分析	208
9.1 汽车用钢案例分析	208
9.1.1 质量建模方法与流程	208
9.1.2 IF 钢力学性能分析与预测	210
9.1.3 质量监控与诊断	219
9.2 热轧带钢头部拉窄案例分析	223
9.2.1 变量选择和数据采集	223
9.2.2 变量的匹配	223
9.2.3 核主成分分析监控与诊断	224

9.3 小结 226

10 全流程质量监控系统框架 227

10.1 系统基本架构 227

10.1.1 数据采集与重整平台 229

10.1.2 数据集成平台 229

10.1.3 实时质量监控与预警平台 229

10.1.4 离线数据分析平台 229

10.2 数据采集、预处理和数据利用 231

10.2.1 数据采集与重整 231

10.2.2 数据集成平台 231

10.3 主要功能模块 235

10.3.1 质量监控与在线判定 235

10.3.2 质量分析与诊断 237

10.3.3 过程质量在线优化 238

10.3.4 人工判定与综合判定 240

10.3.5 产品质量卡建档 241

10.3.6 质量报表生成 242

10.3.7 质量追溯与诊断 244

10.3.8 系统仿真与质量优化 245

10.4 小结 247

附录 A 概率分布表 248

附录 B Matlab 核心代码..... 273

参考文献 291

符号表

符 号	物 理 含 义
X	一元随机变量
$E(X)$	随机变量 X 的期望
$D(X)$	随机变量 X 的方差
$\boldsymbol{X} = (X_1, X_2, \cdots, X_p)$	p 元随机变量
$x_{(i)}$	第 i 个样本点 ($i=1, 2, \cdots, n$)
x_j	第 j 个变量 ($j=1, 2, \cdots, p$)
$x_i^{(l)}$	第 l 个子集 (容量为 n) 的 i 个样本点 ($l=1, 2, \cdots, m$)
$\bar{x}_j$	第 j 个变量的样本均值
$\bar{\boldsymbol{x}}^{(l)}$	第 l 个子集 (容量为 n) 的样本均值 ($l=1, 2, \cdots, m$)
$R^{(l)}$	第 l 个子集 (容量为 n) 的极差 ($l=1, 2, \cdots, m$)
s_j	第 j 个变量的样本标准差
s_{jk}	第 j 个变量和第 k 个变量的协方差
$\boldsymbol{S}$	样本的协方差矩阵
r_{jk}	第 j 个变量和第 k 个变量的相关系数
μ	总体的均值
σ	总体的标准差
$\boldsymbol{\Sigma}$	总体的协方差矩阵
ρ_{jk}	在总体中, 第 j 个变量和第 k 个变量的相关系数
$N(\mu, \sigma^2)$	均值为 μ 、方差为 σ^2 的一元正态分布
$N_p(\mu, \boldsymbol{\Sigma})$	均值为 μ 、协方差矩阵为 $\boldsymbol{\Sigma}$ 的 p 元正态分布
$\chi^2(n)$	自由度为 n 的卡方分布
$W_p(n, \boldsymbol{\Sigma})$	自由度为 n 、协方差矩阵为 $\boldsymbol{\Sigma}$ 的 p 维中心 Wishart 分布
$F(n, m)$	自由度为 n 和 m 的 F 分布
$\Lambda_p(n, m)$	自由度为 n 和 m 的 p 维 Wilks 分布
$\beta(n, m)$	自由度为 n 和 m 的 β 分布
$t(n)$	自由度为 n 的 t 分布

续表

符 号	物 理 含 义
$T_p^2(n)$	自由度为 n 的 p 维 T^2 分布
α	显著性水平
C_p	过程能力指数
C_{pU}	上单侧过程能力指数
C_{pL}	下单侧过程能力指数
C_{pK}	偏移度为 K 时的修正过程能力指数
$T_U、T_L$	规范上下限
T_i^2	第 i 个样本点的 T^2 值
T_{ij}^2	第 i 个样本点的第 j 个变量 X_j 的 T^2 值
$T_{\bar{y},k}^2$	第 i 个样本点在已知变量 X_k 时, 变量 X_j 的 T^2 值
l_j	第 j 个主方向向量 ($j=1, 2, \cdots, p$)
L	主方向矩阵
t_j	第 j 个主成分 ($j=1, 2, \cdots, p$)
T	主成分矩阵
E_h	提取 h 个主成分后的残差矩阵
CPV_h	前 h 个主成分的累计贡献率
$T^2_{\text{Hotelling}}$	霍特林 T^2 统计量
T^2_{PCA}	基于 PCA 的 T^2 统计量
SPE	平方预测误差统计量
$Contr_{ij}^T$	第 i 个样本点的第 j 个变量对 h 个主成分的总贡献值
$Contr_{ij}^{SPE}$	第 i 个样本点的第 j 个变量对 SPE 统计量的总贡献值
VIF_j	变量 X_j 的方差膨胀因子
$k(x, z)$	核函数
$\phi: \mathbf{x} \rightarrow \phi(\mathbf{x})$	从原始 $\mathbf{X}$ 空间到高维特征空间的映射
K	核矩阵
ξ_i, ξ_i^*	松弛变量
W	权值矩阵

1 绪 论

随着我国制造业向高端化转型，用户对钢材的质量要求越来越高，指标越来越严格，钢铁产品的质量已成为钢铁企业市场竞争力的关键要素。尤其是钢铁行业在拓展产业服务链过程中，产品质量的稳定性成为企业核心竞争力。此外，我国钢铁企业由于产品质量不合格所造成的损失也十分惊人。据粗略估计，我国钢铁产品由于质量异议索赔和质量改判所造成的经济损失约为每吨钢 10 元，且高端产品因质量不合格所造成的经济损失将更加惨重。因此，有的学者认为：21 世纪是“质量的世纪”。

激烈的市场竞争促使企业更加关注质量问题，各个钢铁企业都在不断地努力提高产品的质量，越来越多的学者也开始研究提高产品质量的方法：寻求工艺参数与产品质量的关系，建立更加精准、适用性更强的质量预测与控制模型；建立产品质量在线监控系统，实时检测产品质量状况，一旦出现质量问题，及时找出引起异常的原因，避免造成批量的质量判废，减少企业的经济损失。目前钢铁企业在全流程产品质量监控中存在的主要问题有：

(1) 生产过程中，存在产品质量不稳定、批次之间质量差异大、批次内质量波动大等一系列问题，尤其是高端产品对质量规范要求更加严格，如何确保产品质量的稳定性是钢铁企业需要解决的问题。

(2) 在新产品开发过程中，如何解决各工序中质量设计、质量控制、质量检验、判定放行等一系列问题，成为企业实际生产中亟待解决的难题。

因此，建立适用于钢铁企业的质量监控理论与方法，研发冶金全流程产品质量分析与过程监控的实时系统，对加速新钢种研发的进程、减少研发成本以及控制产品质量稳定性、提高产品的竞争力具有十分重要的意义。

1.1 冶金生产过程的特点

冶金和化工是典型的大型复杂流程工业，由功能不同但又相互关联、相互支撑、相互制约的各个工序和各种装置及相关设备，通过工序间串联、并联方式集成后构成完整的复杂生产制备系统。针对这类大型流程工业的特点，解析多流程、多尺度、多装置间的相互关系，建立有效的质量预测与控制模型，寻求最优的质量控制策略，监控产品的生产过程状态，确保产品的质量，已成为我国钢铁工业重大的研究课题。

研究产品质量与生产过程参数的关系，建立产品质量监控模型对生产过程本质特性的研究和实际生产中的过程控制、质量预测、质量诊断和工艺优化有着重要的现实意义。在实际生产中，产品质量受到操作水平、设备能力、原材料、工艺参数、生产环境等多方面因素的综合影响，产品质量的波动是不可避免的。而波动又分正常波动和异常波动两种，过程监控的目的就是要监控实际生产过程的质量波动状况，消除、避免异常波动，使

生产过程处于正常波动范围之内。图 1-1 给出了影响产品质量的主要因素。

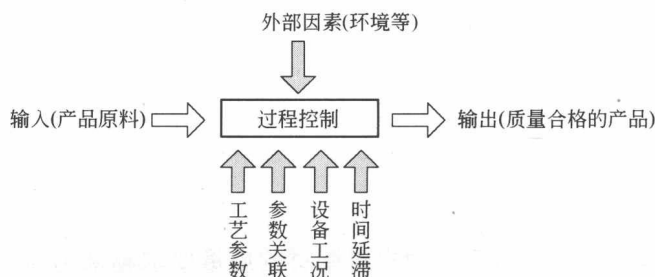


图 1-1 影响产品质量的主要因素

- (1) 输入环节：主要是原料准备，如铁水、钢水成分，板坯质量等；
- (2) 外部因素：如环境温度、湿度、尘埃等，尤其对高端产品需考虑环境影响；
- (3) 工艺参数：设置合理的过程控制的工艺参数是确保产品质量的关键因素；
- (4) 参数关联：在非线性、强耦合情况下，需掌握工艺参数间及工艺参数与质量间的关系；

(5) 设备工况：设备在服役过程中工况会发生变化，使产品质量出现偏移；

(6) 时间延滞：在慢过程中，控制变量的瞬间变化存在时间滞后，需考虑延滞性。

在钢铁生产过程中，各工序将根据不同产品的质量要求制定多项质量规范要求，而影响这些质量规范的工艺参数也是多方面的，主要包括原料的各种参数、操作过程工艺参数等。如果将某一工序的生产过程看成是一个系统，则所有的工艺参数（包括原料参数）可作为系统的输入，产品质量指标作为系统的输出，而工艺装备的过程能力和工况、外部环境因素、操作人员水平等可视为系统的特征。产品质量的监控模型就是寻求在确定的系统特征下，建立生产过程中各种工艺参数与产品的各种质量指标之间的关系，即根据生产过程的输入输出数据建立质量监控系统的数学模型。但是，由于冶金生产过程的复杂性，常常难以建立系统的机理模型，因此基于实际生产数据建立统计过程质量控制模型成为必然的选择。

钢铁生产过程具有如下几个特点：

(1) 多变量。生产过程中涉及的工艺控制变量和产品质量指标多达成千上万个，如轧制温度、速度、轧制力等工艺参数，力学性能、合金成分、组织结构、尺度精度、表面质量等质量指标。对于不同品种、规格、工序等需确定不同的工艺参数与质量指标。

(2) 强耦合。由于变量间的耦合关系，一个变量发生变化时将会引起其他多个相关变量发生相应变化。另一种情况是，当一个工艺参数调整时，其他参数也需调整才能确保产品质量。例如，温度的变化引起轧制力的变化，而轧制力又影响了塑性变形率，反过来又引起温度的变化。

(3) 非线性。生产过程中工艺参数间、质量指标间、工艺参数与质量指标之间往往存在着多重非线性关系，采用常规的线性分析方法会造成质量预测和质量诊断过程的偏差，因此，需要采用非线性分析方法。

(4) 遗传性。上个工序出现的质量问题将会影响下个工序的产品质量，甚至本工序中前阶段的过程控制对下阶段的质量控制都会有影响。例如，铸坯的内部夹杂物和偏析对热

轧板带的质量和组织性能带来影响,进而影响冷轧的产品质量。

(5) 高速性。生产过程中带钢的轧制速度最高可达到 20 ~ 30m/s,这对工艺参数控制的时效性和准确性提出了更高要求,也对质量监控系统的实时性提出了更高的要求。

(6) 时变性。当装备工况、环境条件发生改变时,工艺参数和质量指标也会随之产生变化。尤其像转炉、连铸机、轧辊等装备随着服役时间的增加,其工况会产生细微的变化。

(7) 多态性。不同的工艺流程、不同的装备,甚至同一型号,但不同的工艺装备,由于流程和装备自身的不同特性,具有不同的特征形态。因此在设定工艺参数时,必须考虑产品、工艺、装备本身的特点,才能制造出合格的产品。

实现精准、高效的产品质量监控需要具备一些必要的基本条件:

(1) 数据采集(基础)。完整、准确、可靠、快捷地采集和整合工艺参数与质量数据,包括各种实时和离线工艺参数与质量数据,才能保证质量监控系统发挥其作用,成为企业质量在线监控、质量在线判定、质量分析与工艺参数优化等业务协同平台的数据支撑。

(2) 分析方法(工具)。正确运用各种统计模型、机理模型和智能模型等分析方法和数学工具,包括各种数据预处理、质量预测、过程监控和质量诊断、分类与聚类、参数优化等算法。特别要注意,不同的分析方法适合于不同的对象和场合,这需要大量的实践经验和理论指导。

(3) 专业知识(依据)。发挥不同工序冶金专家的知识 and 经验是必不可少的,尤其是在质量设计、规则的建立、质量标准的制定、质量判定等过程中,领域专家的参与是非常重要的。同时,冶金领域专家也应当借助各种新的数学分析方法,从大量数据分析中不断完善和丰富自己的专业知识。

随着计算机系统、数据库技术的普及与应用,钢铁企业信息化系统中拥有了丰富的生产数据资源,从而也提出了采用各种数值分析方法对大量的工艺过程数据和产品质量数据进行全流程产品质量监控的迫切需求。目的是通过大量生产数据分析来揭示、总结生产过程的内在规律,为提高产品质量提供各种信息,从而把数据资源转化为企业的经济效益和产品质量优势,提高产品的市场竞争力。在大数据时代,如何从每个钢铁企业每年所产生的几十 TB 海量数据中提取出有价值的信息,从而完善产品质量管理体系,实现全流程产品质量在线监控,已成为全行业提高管理水平的必然趋势。

1.2 质量监控技术的现状与发展趋势

世界经济正向全球一体化的方向发展,世界市场的竞争日趋激烈。在市场竞争的五大要素——品种、质量、价格、服务和交货期中,决定竞争胜负的要素是质量。正如美国质量管理专家朱兰(J. M. Juran)于 1994 年在美国质量管理学会上所说的,20 世纪以“生产力的世纪”载入史册,21 世纪将是“质量的世纪”。产品质量已成为工业生产中最为核心的竞争力之一。在实际的工业生产过程监测和产品质量控制中,很多情况下对工业过程机理的精确建模非常困难,而基于过程数据分析的各种统计方法和人工智能方法已被广泛应用于工业生产实践。

基于 6 σ 的管理理念追求严格的质量管理,要求将传统的事后检验提升为对生产全过

程的监控,并采用基于统计过程控制的质量管理方法。统计过程控制(Statistical Process Control, SPC)是一种借助数理统计方法的过程控制工具。它对生产过程进行分析评价,根据反馈信息及时判断出现异常因素的征兆,并采取措施消除其影响,使生产过程保持在受控状态,以稳定产品的质量。当生产过程仅受到偶然因素的影响时,过程处于受控状态;当过程中存在系统因素的影响时,过程可能处于失控状态。由于过程变量的波动具有统计规律性,当过程受控时,过程特性一般服从稳定的随机分布;失控时,过程变量的分布将发生改变。SPC正是利用过程波动的统计规律性对生产过程进行分析和控制。

传统的统计过程控制采用单变量统计过程控制方法。在统计过程控制早期应用中,由于受测量技术以及数据存储和分析技术的限制,只对生产过程中的少数几个重要指标单独地实施统计过程控制,比如为这些指标建立单变量休哈特控制图。单变量统计过程监控主要有:休哈特控制图(测量变量对时间作图)、移动平均控制图(Moving Average, MA, 测量变量的滑动平均值对时间作图)、指数加权滑动平均控制图(Exponentially Weighted Moving Average, EWMA, 变量的指数加权滑动平均值对时间作图)、累积和控制图(Cumulative Sum, CUMSUM, 测量值与目标值偏差的累加对时间作图)等。Thompson 和 Bissell 等对单变量统计过程监控方法做了详细的描述。但是,随着测量技术的发展,人们已经能够对越来越多的产品性能指标进行测量,同时用户对产品质量要求也越来越严格,这就要求对更多的产品质量指标和过程变量进行监控。若需要监控的多个过程变量之间存在相关性,仅仅采用多个单变量统计过程监控,其结果往往是不可靠的。原因是,变量之间耦合关系将改变基于变量间互相独立的休哈特图的统计分布规律。对于钢铁生产过程而言,由于变量之间必须满足能量流、物质流等各种内在的关系,相互之间往往存在着多重相关性,即当一个变量改变时,相关的其他变量也应做相应的调整才能确保质量指标达到规范要求。

将多变量统计分析方法融入传统的统计过程控制,形成了多变量统计过程控制的基本框架。多变量统计过程控制 MSPC (Multivariate Statistical Process Control) 综合考虑各变量相关关系,实现多变量生产过程的质量监控。对多变量统计控制图的研究最早可以追溯到 20 世纪 40 年代中期。霍特林(Hotelling)于 1947 年针对多变量过程控制问题,首次提出了多变量 T^2 控制图,开创了多变量控制图研究与应用的先河。霍特林的多变量 T^2 控制图使用 T^2 统计量,在显著性水平 α 下,同时监控多个变量。 T^2 控制图的基本原理是:如果多变量过程控制中没有异常点存在,则过程中各个样本点到均值的统计距离应该保持受控状态。随后,Healy 等人提出了适用于监控过程微小偏移的多变量累积和控制图(MCUSUM)及多变量指数移动平均控制图(MEWMA),推动了多变量统计控制图的进一步发展。

随着统计数据降维技术的发展,多变量统计控制图的研究出现了新的方向,过程控制的对象从基于距离的统计量,转向了一些基于统计降维技术构建的综合变量的统计量。利用统计降维的原理, Jackson 等人提出了基于主成分分析(Principal Component Analysis, PCA)方法的多变量统计控制图,随后又有人提出了基于偏最小二乘法(Partial Least Squares, PLS)的多变量统计控制图。主成分分析和偏最小二乘法采用多元投影方法,将工艺参数数据和质量数据从高维数据空间投影到低维特征空间,所得到的特征变量保留了原始数据的特征信息,摒弃了冗余信息,是一种高维数据分析处理的有效工具。对于高维