

# 数据驱动的最优运行状态鲁棒评价方法及应用

褚菲<sup>1,2</sup> 赵旭<sup>1,2</sup> 代伟<sup>2</sup> 马小平<sup>2</sup> 王福利<sup>3,4</sup>

**摘要** 在现代复杂工业生产过程中, 细致而稳健的运行状态评价及非优因素识别对指导工业生产具有十分重要的实际意义. 考虑到复杂工业过程难以建立准确的数学模型和实际工业过程数据噪声及离群点污染比较严重的问题, 本文提出一种全潜鲁棒偏 M 估计的复杂工业过程最优状态的鲁棒评价方法. 在建立离线评价模型时, 通过对过程数据主元和残差子空间的进一步分解, 提取出能够反映与原材料、生产消耗和产品质量等因素相关的经济指标的变化信息, 同时采用样本数据加权的方法消除离群点对评价模型的不利影响, 提高算法的鲁棒性; 在线评价时, 针对生产过程中存在不确定性因素, 引入在线数据窗口及相似度分析进行在线评价, 并给出在线评价的准则和流程, 提高评价结果的可靠性, 当评价结果非优时, 通过计算相应变量的贡献率识别非优因素. 最后, 通过重介质选煤过程验证了所提方法的有效性.

**关键词** 复杂工业过程, 数据驱动, 运行状态评价, 全潜鲁棒偏 M 估计, 非优因素

**引用格式** 褚菲, 赵旭, 代伟, 马小平, 王福利. 数据驱动的最优运行状态鲁棒评价方法及应用. 自动化学报, 2020, 46(3): 439–450

DOI 10.16383/j.aas.c180018



开放科学(资源服务)标识码(OSID):

## Data-driven Robust Evaluation Method for Optimal Operating Status and Its Application

CHU Fei<sup>1,2</sup> ZHAO Xu<sup>1,2</sup> DAI Wei<sup>2</sup> MA Xiao-Ping<sup>2</sup> WANG Fu-Li<sup>3,4</sup>

**Abstract** In the process of modern complex industrial production, a detailed and robust evaluation method of operation state is of great significance for guiding the production. Considering the difficulty to establish an accurate principle model and the process data which are easily polluted by noise and outliers, this paper proposes a robust optimal evaluation method for complex industrial processes based on total partial robust M-regression. In the off-line modeling stage, by further decomposing the principal and residual subspaces of the process data, the process variation information related to the economic indexes reflecting the factors such as raw materials, production consumption and product quality is extracted, and the adverse effects of the outliers are eliminated by sample data weighting to improve the robustness of the algorithm. In the stage of online evaluation, the online data window and similarity analysis are introduced for the uncertain factors of the production process, and the framework and procedure of online evaluation are given to improve the reliability of the evaluation results. If the evaluation results are not optimal, then the non-optimal factors are identified by calculating the contribution rates of the corresponding variables. Finally, the effectiveness of the proposed method is illustrated by a process of dense medium coal preparation.

**Key words** Complex industrial process, data-driven, operational status evaluation, total partial robust M-regression, non-optimal factor

**Citation** Chu Fei, Zhao Xu, Dai Wei, Ma Xiao-Ping, Wang Fu-Li. Data-driven robust evaluation method for optimal operating status and its application. *Acta Automatica Sinica*, 2020, 46(3): 439–450

收稿日期 2018-02-27 录用日期 2018-06-10  
Manuscript received February 27, 2018; accepted June 10, 2018  
国家自然科学基金(61973304, 61503384, 61873049), 江苏省六大大人才高峰项目(DZXX-045), 江苏省科技计划项目(BK20191339), 徐州市科技创新计划项目(KC19055), 矿冶过程自动控制技术国家重点实验室开放课题, 前沿课题专项项目(2019XKQYMS64) 资助  
Supported by National Natural Science Foundation of China (61973304, 61503384, 61873049), Selection and Training Project of High-level Talents in the Sixteenth “Six Talent Peaks” of Jiangsu Province (DZXX-045), Science and Technology Plan Project of Jiangsu Province (BK20191339), Science and Technology Innovation Plan Project of Xuzhou (KC19055), Open Foundation of State Key Laboratory of Process Automation in Mining and Metallurgy, and Fundamental Research Funds for

the Central Universities (2019XKQYMS64)

本文责任编辑 王卓

Recommended by Associate Editor WANG Zhuo

1. 中国矿业大学地下空间智能控制教育部工程研究中心 徐州 221116
  2. 中国矿业大学信息与控制工程学院 徐州 221116
  3. 东北大学信息科学与工程学院 沈阳 110819
  4. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819
1. Research Center of Underground Space Intelligent Control Engineering of the Ministry of Education, China University of Mining and Technology, Xuzhou 221116
  2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116
  3. College of Information Science and Engineering, Northeastern University, Shenyang 110819
  4. State Key Laboratory of Integrated Automation for Process Industries, Northeastern University, Shenyang 110819

随着现代工业的高度复杂化,依据物理化学机理模型建立准确数学模型的传统控制方法已经不能满足实际生产需求<sup>[1-2]</sup>.因此,数据驱动的思想逐渐流行于现代控制领域.数据驱动的主要思想是利用受控系统大量的离线、在线数据,实现对系统的预报、评价、诊断以及优化等控制功能<sup>[2]</sup>.

在市场竞争日益激烈的今天,良好的工业过程运行状态是有效保证企业产品质量及综合经济效益的重要前提.然而,在复杂工业生产过程中,运行状态常常受到扰动等各种不确定因素的影响,导致其偏离最优工况点.因此,为了保证产品质量和提高企业综合经济效益,引入过程监控技术对生产过程进行实时监控是很有必要的<sup>[3-4]</sup>.然而,传统的过程监控研究<sup>[5-8]</sup>只关注如何区分“正常”和“故障”两种运行状态,这一粗犷的评价方法明显不能满足现在生产绿色高效的需求.因此研究细致而稳健的运行状态评价及非优因素识别方法对指导复杂工业过程的高效运行具有十分重要的实际意义.

目前,在故障预测与健康管理(Prognostic and health management, PHM)领域中机械设备的健康监测方面涌现出很多运行状态评价方法的研究成果<sup>[9-10]</sup>.例如:文献[11]提出了一种基于随机过程模型一般表达式的RUL(Remaining useful life)预测的方法,该方法通过退化过程模拟产生一系列退化轨迹,然后将降解轨迹的状态输入到解析解中来估计RUL的概率密度函数,从而进行RUL预测.文献[12]从数据采集到健康指标构建、健康等级划分、剩余使用寿命预测等方面系统地介绍了机械设备健康预测的各种方法及目前的研究现状与挑战等.对于复杂工业过程而言,针对过程的关键装备或核心控制器的运行性能等<sup>[13-15]</sup>,近年来也出现了部分运行状态评价方法的研究成果,采用的方法涉及模糊理论<sup>[16]</sup>、主成分分析(Principal component analysis, PCA)技术和Fisher判别分析法<sup>[17]</sup>等.基于模糊理论,文献[18]提出了一种模糊综合评判(Fuzzy comprehensive evaluation, FCE)的方法,并将其应用于变压器的状态评估中.该方法依据模糊数学隶属度理论将定性评价转变为定量评价,利用不同指标对待评定对象的隶属度情况进行综合评价,一方面能够考虑到评价对象的差异性,使得评判指标、影响因子的模糊性得以体现,另一方面在评价过程中又结合人们的先验知识,使得评价结果更加客观.但是,该评判方法过度依赖于领域知识,且容易造成信息冗余,在一定程度上影响了评价结果的准确性和可靠性.采用灰色关联分析法(Grey relational analysis, GRA)<sup>[19-20]</sup>进行状态评价,该方法是基于灰色系统理论,通过对评价指标相关的

原始数据进行无量纲化处理,得出关联系数和关联度值,并根据关联度值对待评价指标做出排序<sup>[21]</sup>.此方法最明显的优势是对数据量的要求不高,在很大程度上弥补了因信息不对称而造成的损失.但是,使用该方法需要先确定各评价指标的最优值,使得人为评价的主观性过强,此外,在应用过程中也同样存在信息冗余的问题.文献[22]等提出,将PCA应用于运行状态评价中,通过PCA的数据降维作用,将多个评价指标转换为少数几个综合指标,即主成分,其中每个主成分都可以反映原始数据大部分信息,且所含信息互不重复,从而消除各评价指标之间的相关性,从而有效地避免了信息冗余的情况发生.

如何对复杂工业过程的整体运行状态进行评价,从而更加深入的了解复杂过程的运行机制和关键影响因素,对于提高企业的经济效益至关重要<sup>[8, 23-30]</sup>.文献[26]提出一种基于高斯混合模型的复杂工业过程的在线操作安全性和最优性评估的概率框架.该方法虽然可以提供有效的在线运行状况评估,但不能通过有效的方法识别非最优运行状态的原因.文献[27]提出了一种针对多模态复杂工业过程的运行状态评估的方法,其中最优化评估指标被定义并用于评估过程操作的最优化.然而,评估指标是根据优化的目标函数的价值构建的,这些在实际生产中通常无法在线获得,特别是耗时和大规模的复杂工业过程;文献[24]提出了一种基于Fisher判别分析的复杂工业过程运行状态在线评价方法.通过求取样本之间的相似度将未知类别的样本进行相应的归属划分,该方法既能有效地减小相似数据间的差异,又能扩大不同类别数据间的差异,从而可以识别过程整体运行状态的优劣,提高在线识别的效率.然而,该方法未能将过程运行状态与产品质量或综合经济性能指标相关联,无法识别导致过程非最优的因素.因此,文献[25]以湿法冶金过程为背景,通过计算在线数据与离线评价模型之间的欧氏距离,实现复杂过程经济指标最优的在线评价.在建立离线评价模型部分,对每一个运行状态等级分别利用数据驱动的方法建立评价模型;在线评价时,则通过比较在线窗口内数据与离线模型之间的数据相似度识别状态等级<sup>[25]</sup>.文献[29]提出了一种针对非高斯过程的运行最优性评估及非最优因素识别策略,该方法首次对生产过程过渡阶段进行在线评估,并结合各操作量的贡献率对非最优因素进行识别.文献[28]提出了一种新的两级多块操作性能优化评估方法.通过适当定义全局和子块级别的性能等级,而不需要确定难以获得的子块级别评估指标.针对非最优性能等级,在两级多块结构的框架下提出了基于匹配度函数的原因识别技术.负责的子块和变量可以迅速

定位, 并且可以大大提高可解释性.

然而, 在实际生产过程中由于生产原料的波动、外部生产环境的改变及生产过程中存在各种形式的噪声干扰和不确定性因素等, 导致来自工业现场的过程数据质量不高, 存在噪声大、数据缺失和离群点等问题<sup>[25]</sup>. 如果继续采用上述已有评价方法, 无法建立准确可靠的运行状态评价模型, 进而导致评价模型的泛化能力变差, 精度降低, 在线评价的准确度不理想. 因此, 本文提出一种基于全潜鲁棒偏 M 估计 (Total partial robust M-regression, Total-PRMR) 的复杂过程最优状态的鲁棒评价方法. 在离线部分, 提出了一种改进的全潜鲁棒偏 M 估计算法, 该方法通过对鲁棒偏 M 估计 (Partial robust M-regression, PRMR) 主元子空间的进一步分解, 从历史数据提取出能够反映与原材料、生产消耗和产品质量等因素相关的综合经济指标的过程变化信息, 同时通过在回归函数中对离群点的选择性惩罚措施, 提高模型的鲁棒性和精度. 本文结合产品质量指标或经济指标对历史数据进行划分, 离线建立各状态等级的鲁棒评价模型. 在线评价时, 根据实际需求合理制定在线窗口的宽度, 利用各等级鲁棒评价模型对在线窗口内的数据进行综合鲁棒加权, 并计算加权后窗口数据与各评价等级之间的相似度的大小进行在线评价, 当在线评价结果非优时, 通过计算各操作量的贡献度, 实现非优 (Non-optimal) 因素的准确判别. 最后, 以原煤重介分选过程为实际应用背景, 通过仿真实验, 验证该方法的有效性.

## 1 全潜鲁棒偏 M 估计算法

### 1.1 单输出 Total-PLS 算法

全潜结构投影法 (即 Total-PLS 算法) 是对偏最小二乘 (Partial least squares, PLS) 算法的发展<sup>[31]</sup>. Total-PLS 算法结合输入矩阵与输出矩阵的相关性, 对 PLS 的主元空间和残差空间进行分解, 从输入矩阵得分中进一步分离出与输出相关和正交的不同分量. 同时从输入矩阵的偏差中, 分离出携带较多过程信息的部分和过程噪声.

设过程输入数据矩阵为  $X \in \mathbf{R}^{N \times J}$ ,  $N$  为样本数,  $J$  为过程变量数. 输出数据矩阵为  $\mathbf{y} \in \mathbf{R}^{N \times 1}$ , 代表综合经济效益指标, 或能够反映经济效益的产品的重要属性, 为了方便起见, 本文讨论的是  $\mathbf{y}$  为单变量的情况, 对于多个经济指标的情况, 可以通过加权等方式转换成单变量形式.

通过非线性迭代的偏最小二乘法 (Nonlinear iterative partial least squares, NIPALS), 将标准化后的  $(X, \mathbf{y})$  投影到一个低维空间中, 则 PLS 模型可

描述为<sup>[31]</sup>

$$\begin{cases} X = TP^T + E \\ \mathbf{y} = T\mathbf{q}^T + \mathbf{f} \end{cases} \quad (1)$$

其中,  $T$ ,  $P$  和  $E$  分别是矩阵  $X$  的得分矩阵、负载矩阵和残差矩阵;  $\mathbf{q}$  和  $\mathbf{f}$  分别是  $\mathbf{y}$  的负载矩阵和残差矩阵.

在上述 PLS 模型的基础上, 单输出 Total-PLS 模型可表示为<sup>[31]</sup>

$$\begin{cases} X = \mathbf{t}_y \mathbf{p}_y^T + T_o P_o^T + T_r P_r^T + E_r \\ \mathbf{y} = \mathbf{t}_y + \mathbf{f} \end{cases} \quad (2)$$

其中,  $\mathbf{t}_y$  为在  $T$  中和输出  $\mathbf{y}$  直接相关的部分,  $T_o$  为得分矩阵  $T$  中与输出即经济指标无关的部分,  $T_r$  为 PLS 模型的残差中残留的过程变化信息的得分矩阵,  $E_r$  为白噪声或无法建模部分.

### 1.2 鲁棒偏 M 估计算法

鲁棒偏 M 估计 (PRMR) 是一种单输出变量的加权迭代偏最小二乘算法的鲁棒形式<sup>[32]</sup>, 通过迭代计算, 赋予样本不同的样本权值 (离群点分配接近于 0 的权值, 正常数据分配接近于 1 的权值), 用来克服疑似离群点数据对模型的影响, 并能保留该部分数据的有用信息. 在鲁棒偏 M 估计算法中, 离群点分为远离输入空间中心 (即输出正常, 输入异常) 的高杠杆 (High leverage) 点和测量值与预测值具有较大偏差 (即输入正常, 输出异常) 的高残差 (High residual) 点, 鲁棒偏 M 估计算法采用不同的方法对以上两种离群点赋予惩罚项 (即权值)<sup>[32]</sup>. 设  $w_i^x$  为第  $i$  个样本数据的杠杆权值, 可表示为

$$w_i^x = f \left( \frac{\|\mathbf{t}_i - \text{med}_{L_1}(T)\|}{\text{med}_i \|\mathbf{t}_i - \text{med}_{L_1}(T)\|}, c \right) \quad (3)$$

且

$$f(z, c) = \frac{1}{(1 + |\frac{z}{c}|)^2} \quad (4)$$

其中,  $T$  是样本数据的得分矩阵, 其第  $i$  行即为  $\mathbf{t}_i$  (得分向量),  $\text{med}_{L_1}$  是  $T$  的空间中位值,  $\|\cdot\|$  是欧几里得度量 (即欧氏距离),  $c$  为自然常数, 取值方法见文献 [33–34]. 类似  $w_i^x$  的定义, 设残差权值  $w_i^r$ , 计算式为

$$w_i^r = f \left( \frac{\mathbf{r}_i}{\tilde{r}}, c \right) \quad (5)$$

其中,  $\mathbf{r}_i$  为模型预测误差,  $\tilde{r}$  为模型预测误差的估计值或中值<sup>[24]</sup>

$$\tilde{r} = \text{med}_i |\mathbf{r}_i - \text{med}_j(\mathbf{r}_j)|, \quad i, j = 1, 2, \dots, n \quad (6)$$

通过计算上述两种权值可以得出第  $i$  个样本的综合权值  $w_i$ , 具体为

$$w_i = \sqrt{w_i^x w_i^r} \quad (7)$$

PRMR 算法通过反复迭代更新样本数据的权值直至收敛<sup>[32]</sup>.

### 1.3 全潜鲁棒偏 M 估计算法

本文针对过程数据中的离群点问题, 提出全潜鲁棒偏 M 估计 (Total-PRMR) 算法.

设过程数据矩阵为  $X \in \mathbf{R}^{N \times J}$ ,  $N$  为样本数,  $J$  为过程变量数, 输出数据矩阵为  $\mathbf{y} \in \mathbf{R}^{N \times 1}$ , 由于过程数据受离群点影响严重, 采用 PRMR 算法对  $X$ ,  $\mathbf{y}$  进行加权处理, 对加权后的数据进行 PLS 回归分析

$$\begin{cases} X_W = TP^T + E \\ \mathbf{y}_W = T\mathbf{q}^T + \mathbf{f} \end{cases} \quad (8)$$

在 PLS 模型中, 得分矩阵  $T \in \mathbf{R}^{N \times A}$  ( $A$  为主元个数, 可由交叉验证的方法确定<sup>[34]</sup>) 关于  $X$  的线性关系可以表示为  $T = XW(P^TW)^{-1}$ , 其中  $W$  为原权重系数矩阵, 令  $R = W(P^TW)^{-1}$ , 得  $T = XR$ . 由式 (3), (5), (7), 可计算出  $w_i^x$ ,  $w_i^r$ ,  $w_i$ , 其中残差为  $r_i = y_{wi} - \mathbf{t}_i \mathbf{q}^T$ . 通过反复迭代运算更新  $T$ ,  $\mathbf{r}$ , 从而更新权值, 直至满足收敛条件为止. 对加权后的数据  $X_W$  进一步分解, 提取与输出相关的过程变量

$$\begin{cases} X_W = \mathbf{t}_{wy} \mathbf{p}_{wy}^T + T_{wo} P_{wo}^T + T_{wr} P_{wr}^T + E_{wr} \\ \mathbf{y}_W = \mathbf{t}_{wy} + \mathbf{f}_w \end{cases} \quad (9)$$

其中,  $\mathbf{t}_{wy}$  是经 PRMR 算法处理后的回归模型得分矩阵, 是与经济指标  $\mathbf{y}_W$  直接相关的过程变化信息,  $T_{wo}$  为回归模型得分矩阵  $T$  中与经济指标  $\mathbf{y}_W$  无关的内容.  $T_{wr}$  是残差矩阵  $E$  中残留的过程信息.  $E_{wr}$  是白噪声部分.  $\mathbf{p}_{wy}$ ,  $P_{wo}$ ,  $P_{wr}$  为相应的负载矩阵. Total-PRMR 将输入数据矩阵  $X_W$  分解为 4 个子空间, 分别为  $\mathbf{p}_{wy}$ ,  $P_{wo}$ ,  $P_{wr}$  和  $(I - P_{wr} P_{wr}^T)(I - PR^T)$ , 对应的维度分别为 1,  $A - 1$ ,  $A_r$ ,  $J - A - A_r$ ,  $A_r$ . 同样可由交叉验证的方法确定. 对于一个新的样本数据向量  $\mathbf{x}_w$

$$\mathbf{x}_w = \mathbf{x}_{wy} + \mathbf{x}_{wo} + \mathbf{x}_{wr} + \tilde{\mathbf{x}}_{wr} \quad (10)$$

将分解为以下 4 个部分:

$$\mathbf{x}_{wy} = \mathbf{p}_{wy} \mathbf{q} R^T \mathbf{x}_w \quad (11)$$

$$\mathbf{x}_{wo} = (P - \mathbf{p}_{wy} \mathbf{q}) R^T \mathbf{x}_w \quad (12)$$

$$\mathbf{x}_{wr} = P_{wr} P_{wr}^T (I - PR^T) \mathbf{x}_w \quad (13)$$

$$\tilde{\mathbf{x}}_{wr} = (I - P_{wr} P_{wr}^T) (I - PR^T) \mathbf{x}_w \quad (14)$$

其中,  $\mathbf{x}_{wy}$  为样本数据中与经济指标  $\mathbf{y}_W$  直接相关的数据部分,  $\mathbf{x}_{wo}$  为与  $\mathbf{y}_W$  无关的数据部分,  $\mathbf{x}_{wr}$  为偏差数据中残留的过程数据部分,  $\tilde{\mathbf{x}}_{wr}$  为白噪声数据部分. 样本数据向量  $\mathbf{x}_w$  的得分向量计算如下:

$$\mathbf{t}_{wy} = \mathbf{q}^T R^T \mathbf{x}_w \quad (15)$$

$$\mathbf{t}_{wo} = P_{wo}^T (P - \mathbf{p}_{wy} \mathbf{q}^T) R^T \mathbf{x}_w \quad (16)$$

$$\mathbf{t}_{wr} = P_{wr}^T (I - PR^T) \mathbf{x}_w \quad (17)$$

Total-PRMR 算法能进一步分解 PRMR 的高维主元子空间和残差子空间. 因此, Total-PRMR 能够准确地提取出与输出相关的过程变化信息. Total-PRMR 算法具备了 PRMR 算法的鲁棒性和 Total-PLS 算法对过程变化信息的精准提取能力.

Total-PRMR 算法的具体步骤如下:

**步骤 1.** 标准化输入数据矩阵  $X$  和输出数据矩阵  $\mathbf{y}$ .

**步骤 2.** 对标准化的输入数据矩阵  $X$  和输出数据矩阵  $\mathbf{y}$  实施 PRMR 算法.

1) 初始化权值  $W_i$ , 利用式 (5) 计算残差权值  $w_i^r$ , 其中  $r_i = y_i - \text{med}_j |y_j|$ , 利用式 (3) 计算杠杆权值  $w_i^x$ , 其中得分向量  $\mathbf{t}_i$  由  $\mathbf{x}_i$  代替 ( $0 < i < n$ ), 最后由式 (7) 计算综合权值  $W_i$ ;

2) 分别对输入数据矩阵  $X$  和输出数据矩阵  $\mathbf{y}$  进行加权计算, 得到矩阵  $X_W$  和输出矩阵  $\mathbf{y}_W$

$$\begin{cases} X_W = W_i X \\ \mathbf{y}_W = W_i \mathbf{y} \end{cases}$$

3) 利用非线性迭代 NIPALS 对  $X_W$  和  $\mathbf{y}_W$  进行回归分析, 并计算残差

$$\begin{cases} r_i = y_i - \mathbf{t}_i \mathbf{q}^T \\ \tilde{r} = \text{med}_i |r_i - \text{med}_j (r_j)| \end{cases}$$

利用式 (3), (5), (7) 更新综合权值  $W_i$ , 当连续两个  $\mathbf{q}$  的差值小于一个指定的阈值 (即满足收敛条件) 时, 转入步骤 3, 否则返回 2), 循环计算直至满足阈值条件为止.

**步骤 3.** 从  $\mathbf{y}_W$  中提取收敛的  $\mathbf{u}_i$ , 令  $i = 1$ ,  $X_{W_i} = X_W$ ,  $\mathbf{y}_{W_i} = \mathbf{y}_W$ , 并计算  $X_{W_i}$  的得分向量  $\mathbf{t}_i = X_{W_i} \mathbf{u}_i$ , 以及  $\mathbf{y}_{W_i}$  的得分向量  $\mathbf{u}_i = \mathbf{y}_{W_i} \mathbf{q}_i$ .

**步骤 4.** 计算  $X_{W_i}$  的负载矩阵:  $\mathbf{p}_i = X_{W_i}^T \mathbf{t}_i$ , 同时提取出全部主元, 计算  $T$ ,  $P$ ,  $U$ ,  $\mathbf{q}$ .

**步骤 5.** 得到以下结果:

$$\begin{cases} X_W = TP^T + E \\ \mathbf{y}_W = U\mathbf{q}^T + \mathbf{f} \end{cases}$$

**步骤 6.** 对主元  $TP^T$  运行 PCA 算法, 得

$$TP^T = \mathbf{t}_{wy} \mathbf{p}_{wy}^T + T_{wo} P_{wo}^T$$

**步骤 7.** 对步骤 5 中的  $E$  进一步运行主成分分解, 得

$$E = T_{wr} P_{wr}^T + E_{wr}$$

Total-PRMR 模型如式 (9) 所示. 对于新的数据  $\mathbf{x}_{new}$ , 其模型参数  $\mathbf{t}_{wy,new}$  为

$$\mathbf{t}_{wy,new} = \mathbf{q}^T R^T \mathbf{x}_{new} = \mathbf{g}_y \mathbf{x}_{new} \in \mathbf{R}^1 \quad (18)$$

由于得分向量  $\mathbf{t}_{wy,new}$  中包含了与经济指标指标直接相关的过程变化信息, 因此利用 Total-PRMR 算法求得的  $\mathbf{t}_{wy,new}$ , 可以方便地获得过程运行状态离线评价模型.

## 2 基于 Total-PRMR 的最优状态的鲁棒评价方法

### 2.1 离线建模

离线建模时, 先根据选取的实际生产过程的经济指标, 将建模数据集根据运行状态的等级分为几个子集, 其中每一个子集代表相应的运行状态等级. 对于复杂工业过程, 运行状态等级的设置需要考虑实际运行管理的需求, 过多或过少的评价等级都不合适, 过多的评价等级将会使问题复杂化, 评价流程过于复杂, 而较少的等级数, 不利于对过程运行状态的细致描述<sup>[25]</sup>. 因此, 建议一般复杂工业过程的状态等级可根据过程特性和运行管理的实际需求分为 3~6 个等级 (但不限于). 本文根据重介选煤过程的特性和产品煤灰分的分布情况, 考虑将其分为 4 个状态等级, 分别为优 (Optimal)、良 (Fine)、中 (Medium)、差 (Poor), 分别用序号 1, 2, 3, 4 表示. 不失一般性, 下面以 4 个状态等级为例介绍本文所提算法. 对应每一个运行状态等级的建模数据子集表示为  $(X_c, \mathbf{y}_c)$ ,  $c = 1, 2, 3, 4$ . 其中,  $X_c = [\mathbf{x}_c(1), \mathbf{x}_c(2), \dots, \mathbf{x}_c(N_c)]^T \in \mathbf{R}^{N_c \times J}$ ,  $\mathbf{y}_c \in \mathbf{R}^{N_c}$ ,  $N_c$  为样本数据的个数. 基于各状态等级的建模数据子集, 利用本文提出的 Total-PRMR 进行回归分析和信息提取. 由式 (9) 和式 (18) 可知, 每个状态等级的运行评价模型可由  $\mathbf{g}_y^c$  表示,  $c = 1, 2, 3, 4$ , 同时各运行状态等级离线建模数据的得分向量可由式 (18) 求得,  $\mathbf{t}_{wy}^c(i) = \mathbf{g}_y^c \mathbf{x}_c(i)$ ,  $i = 1, 2, \dots, N_c$ ,  $c$  为运行状态等级的数目.

### 2.2 在线评价策略

对复杂工业过程进行在线评价时, 考虑到离群点和过程不确定性等因素的影响, 引入一个在线滑动数据窗口用于在线评价, 滑动窗口的宽度  $H$  取

值要适中, 过小不易保证在线评价的可靠性, 过大容易造成算法灵敏度降低, 过渡状态识别不及时和评价结果滞后等,  $H$  值需要在复杂过程的实际运行特性在状态评价的准确性和及时性之间取合理的折中方案. 为了保证在线评价的实时性, 可以根据窗口内数据距离当前时刻的距离 (时间) 赋予遗忘权值, 距离当前时刻越近的数据权值越大, 反之较小; 同时为了增加在线评价的鲁棒性, 可以利用各等级的离线评价模型对在线窗口内数据进行鲁棒加权, 减小离群点和强噪声对在线评价结果的不利影响. 最后, 设置一个用于区分 Optimal, Fine, Medium, Poor 四个稳定状态等级及介于两稳定状态等级之间的过渡过程的相似度控制限, 记为  $\varepsilon$ , 通常取  $0.5 < \varepsilon < 1$ <sup>[25]</sup>.

在线评价流程可以总结为:

1) 构建当前时刻 ( $k$ ) 在线滑动数据窗口矩阵  $X_{w,k} = [\mathbf{x}_{w,k-H+1}, \dots, \mathbf{x}_{w,k}]^T$  作为当前时刻状态评价的基本分析单元, 求取标准化之后的在线窗口数据矩阵,  $\tilde{X}_{w,k}^c$ ,  $c = 1, 2, 3, 4$ . 根据窗口内数据距离当前时刻的距离 (时间) 赋予遗忘权值  $w_{hk1}$ ,  $h = 1, 2, \dots, H$ , 减少离群点对在线评价结果的影响, 利用离线训练好的各等级评价模型对滑动窗口的数据进行鲁棒估计并根据估计值赋予各数据不同的鲁棒权值  $\tilde{w}_{hk2}^c$ ,  $c = 1, 2, 3, 4$  (估计式参考式 (3)~(5)), 对窗口内的数据进行加权仍记为  $\tilde{X}_{w,k}^c$ ,  $c = 1, 2, 3, 4$ .

2) 由式 (18) 计算在线滑动数据窗口矩阵  $\tilde{X}_{w,k}^c$  的得分  $\mathbf{t}_{wy,k}^c = \mathbf{g}_y^c \tilde{\mathbf{x}}_{w,k}^c$ , 并计算该得分与各运行状态等级离线评价模型的距离<sup>[17]</sup>,  $d_{w,k}^c = |\mathbf{t}_{wy,k}^c - \bar{\mathbf{t}}_{wy}^c|^2$ , 其中,  $\bar{\mathbf{t}}_{wy}^c$  是  $\mathbf{t}_{wy}^c$  的均值化向量, 而  $\bar{\mathbf{t}}_{wy}^c = \mathbf{0}$ , 则  $d_{w,k}^c = |\mathbf{t}_{wy,k}^c|^2$ .

3) 利用  $d_{w,k}^c$  可求取数据窗口与评价等级  $c$  之间的相似度: 假设  $d_{w,k}^c \neq 0$ , 则  $\gamma_{w,k}^c = \frac{1/d_{w,k}^c}{\sum_{p=1}^4 1/d_{w,k}^p}$ , 其中,  $\sum_{c=1}^4 \gamma_{w,k}^c = 1$ ,  $0 \leq \gamma_{w,k}^c \leq 1$ , 并利用评价准则进行在线状态评价.

4) 对于非最佳运行状态进行非优因素的识别.

运行状态在线评价可能会受噪声及过程不确定性因素的影响而出现个别错误的评价结果. 为了提高在线评价结果的可靠性, 除了对滑动窗口数据进行鲁棒加权外, 可以根据制定的三条线评价准则进行综合评价<sup>[25]</sup>, 具体如下:

**准则 1.** 假设  $\gamma_{w,k}^p = \max_{1 \leq c \leq 4} (\gamma_{w,k}^c) > \varepsilon$ , 则可以判定该在线数据窗口是属于某一个稳定的运行状态等级  $p$ .

**准则 2.** 假设不满足准则 1, 而满足

$$p = \arg \max \{ \gamma_{w,k}^p | \gamma_{w,k-l+1}^p < \dots < \gamma_{w,k}^p \}$$

$$1 \leq p \leq 4\}$$

其中,  $l$  为正整数, 大小根据实际生产需求确定, 则确定当前时刻开始进入相邻等级之间的过渡。

**准则 3.** 假如不满足以上两个条件, 则认为当前时刻滑动数据窗口受过程不确定性等未知因素影响, 无法给出合乎逻辑的评价结果, 考虑到过程运行的连续性和不同等级之间不能跳变, 沿用前一刻的在线评价结果。

### 2.3 Non-optimal 因素追溯

当生产过程的在线运行状态评价结果为非优 (Non-optimal) 时, 即 Fine、Medium 或 Poor 时, 就需要识别导致生产过程处于非优状态的原因, 从而为生产过程的优化控制提供依据<sup>[23]</sup>。

本文采用基于变量贡献率的方法<sup>[23]</sup>进行非最优因素的识别。在线窗口数据与最优运行状态等级之间的欧氏距离可由下式得出<sup>[35-36]</sup>:

$$\begin{aligned} d_{w,k}^* &= (\bar{\mathbf{t}}_{wy,k}^*)^2 = \left( \frac{1}{H} \sum_{h=k-H+1}^k \mathbf{t}_{wy,k}^*(h) \right)^2 = \\ & \left( \frac{1}{H} \sum_{h=k-H+1}^k \tilde{\mathbf{g}}_y^* \mathbf{x}_{w,k}^*(h) \right)^2 = \\ & (\bar{\mathbf{g}}_y^* \bar{\mathbf{x}}_{w,k}^*)^2 = \left( \sum_{j=1}^J \tilde{g}_{y,j}^* \bar{x}_{w,k,j}^* \right)^2 \end{aligned} \quad (19)$$

式中, 上标 \* 代表 Optimal 离线模型数据,  $\mathbf{x}_{w,k}^*(h)$  是  $k$  时刻的在线数据,  $\mathbf{t}_{wy,k}^*(h)$  是  $\mathbf{x}_{w,k}^*(h)$  的得分向量;  $\tilde{g}_{y,j}^*$  是  $\tilde{\mathbf{g}}_y^*$  的第  $j$  个元素;  $\bar{x}_{w,k,j}^*$  是  $\bar{\mathbf{x}}_{w,k}^*$  的第  $j$  个变量, 且  $\bar{\mathbf{x}}_{w,k}^*$  是  $\mathbf{x}_{w,k}^*(h)$  的平均向量,  $h = k - H + 1, \dots, k$ 。对于  $d_{w,k}^*$ , 第  $j$  个变量的贡献率由下式求出<sup>[36]</sup>:

$$\text{Contr}_j = (\tilde{g}_{y,j}^* \bar{x}_{w,k,j}^*)^2, \quad j = 1, 2, \dots, J \quad (20)$$

根据式 (19) 和式 (20) 可以看出, 相应变量的  $\text{Contr}_j$  值和  $d_{w,k}^*$  值随着  $\bar{x}_{w,k,j}^*$  值的增大而增大, 同时  $\gamma_{w,k}^*$  值会随着  $\bar{x}_{w,k,j}^*$  值的增大而减小, 从而导致生产过程处于 Non-optimal 运行状态。因此, 通过比较操作变量贡献度的大小, 可以进行 Non-optimal 因素的识别。基于全潜鲁棒偏 M 估计的在线运行状态评价以及非优因素识别流程如图 1 所示。

### 3 仿真实验分析

原煤重介质分选技术是当前最先进的一种选煤技术, 其工艺流程基于阿基米德原理。由于重介质选煤过程是一个极其复杂的过程, 实际运行中需要对重介质密度等众多过程参数进行实时控制以保证产

品煤的质量<sup>[27]</sup>。然而, 由于生产环境比较恶劣, 干扰众多, 导致原选煤过程可能脱离事先设计好的最优运行状态。因此, 研究细致且稳健的运行状态鲁棒评价以及过程非优因素的识别的方法对保障重介质选煤过程的绿色高效运行生产, 提高工业企业的综合经济效益, 具有十分重要的实际应用价值。本文将提出的基于全潜鲁棒偏 M 估计的复杂工业过程运行状态评价方法应用于重介质选煤过程, 并进行仿真实验分析。

#### 3.1 两产品重介质原煤分选过程

以某选煤厂的重介质选煤过程为研究对象, 其工艺流程如图 2 所示<sup>[37-38]</sup>, 原煤与介质同时送入混料桶, 通过两产品的重介质旋流器进行分选, 获得精煤和矸石。

采用两产品旋流器, 合介桶中的重介质悬浮液密度是决定分离密度的关键因素, 也是重介分选工艺的主控对象, 主要通过改变注水量、介质 (例如磁铁矿粉) 添加量和分流的办法进行调节, 介质密度的设定值主要通过灰分的离线检测结果进行给定。这种控制方式往往使得原煤分选过程的控制效果不理想, 产品品质不稳定<sup>[27]</sup>。

#### 3.2 建立 Total-PRMR 评价模型

本文结合实际生产情况对原煤重介分选过程的参数检测条件进行细致地分析, 最终选取介质桶液位、主系统介质密度、主旋流器入口压力、煤泥的含量和磁性物的浓度这 5 个过程变量作为重介质选煤过程运行状态评价的在线测量变量, 选取离线化验的灰分作为经济指标, 验证所提方法的有效性。

这里需要说明的是, 除了灰分, 影响重介质选煤过程的综合经济指标的还有硫分度以及设备耗能指标等。但考虑到目前仿真实验的复杂度和指标获取的难易问题, 本文仅以原煤灰分度代表重介质选煤过程的综合经济指标, 以验证所提算法思想的有效性和可行性, 在后续的研究中将进一步讨论影响重介质选煤过程经济效益的其他因素。

考虑选煤厂的生产需求, 根据灰分等级将重介质选煤过程分为 4 个状态等级, 分别是 Optimal、Fine、Medium、Poor, 以及不同状态等级之间的过渡状态。过程状态划分标准如表 1 所示<sup>[38]</sup>, 共采集 8 000 余组数据用于离线建立评价模型。

建模数据集  $(X_1, \mathbf{y}_1)$ ,  $(X_2, \mathbf{y}_2)$ ,  $(X_3, \mathbf{y}_3)$ ,  $(X_4, \mathbf{y}_4)$  分别对应 Optimal、Fine、Medium 和 Poor 这 4 个运行状态等级, 对 4 个状态等级分别建立 Total-PRMR 离线评价的模型  $\mathbf{g}_y^c$ ,  $c = 1, 2, 3, 4$ , 并获得建模数据集的得分  $\mathbf{t}_y$ 。

实际过程数据由于噪声、干扰和人为因素等各

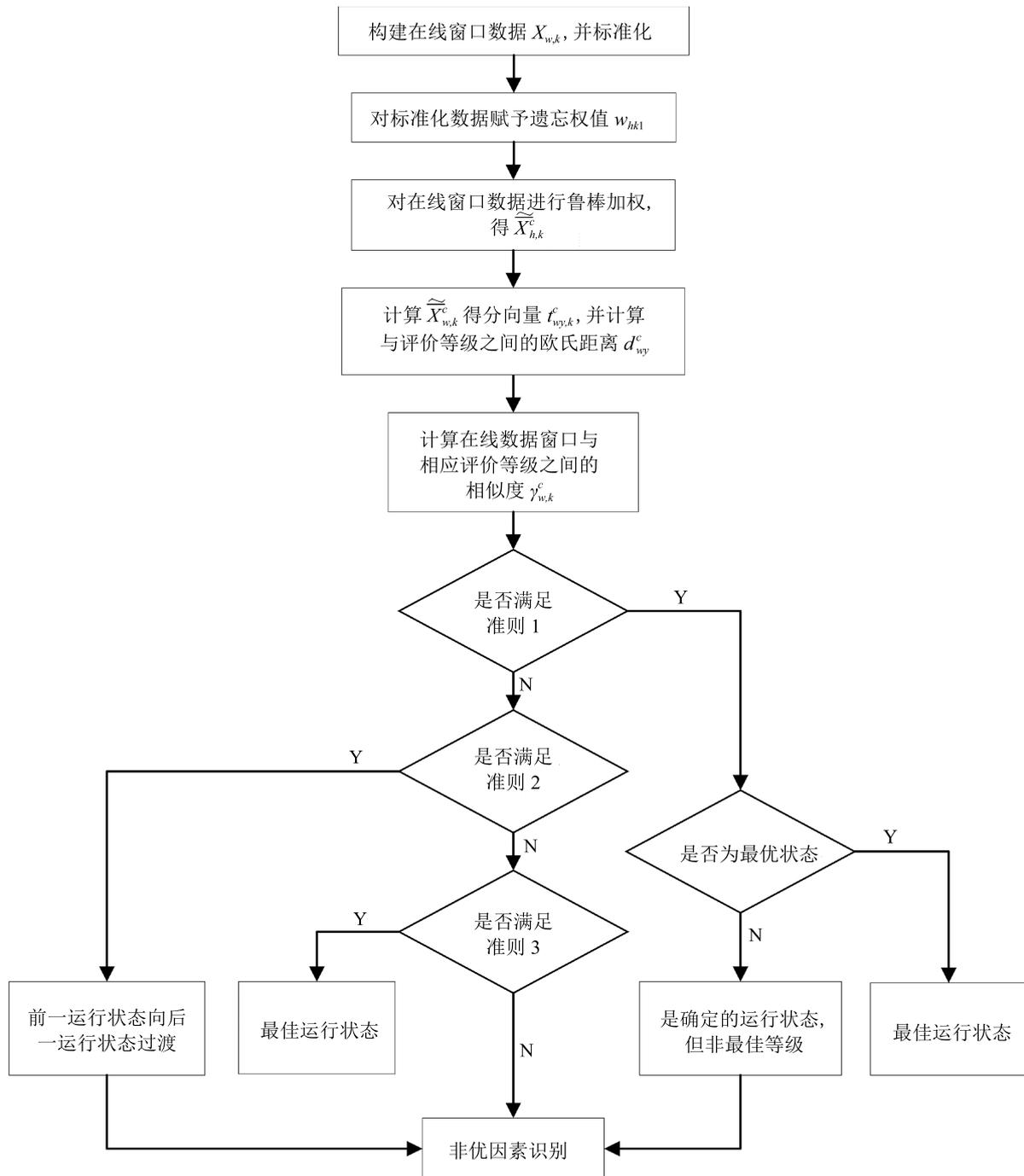


图1 基于全潜鲁棒偏 M 估计的复杂工业过程在线运行状态评价流程  
 Fig. 1 Online operation state evaluation process of complex industrial process based on total partial robust M-regression

种原因而常常含有离群点<sup>[24]</sup>. 文献 [21] 中也提到了“离群点”这个概念, 并指出: 为了确保模型的稳定性, 在建立离线评价模型之前将这些离群点从建模数据中删除, 但是在实际噪声污染严重的工业生产现场中, 离群点数量比较大, 精准识别和剔除都比较困难, 并且这些离群点本身还可能包含一定的生产

过程信息, 若直接去除, 很有可能造成相关生产信息的损失.

这里需要说明的是, 本文采用的实验数据来自选煤厂的实际运行数据, 由于现场生产环境比较恶劣, 建模数据当中已经含有比较强的噪声干扰. 此外, 为了模拟粗大误差等不确定性因素引入的离群

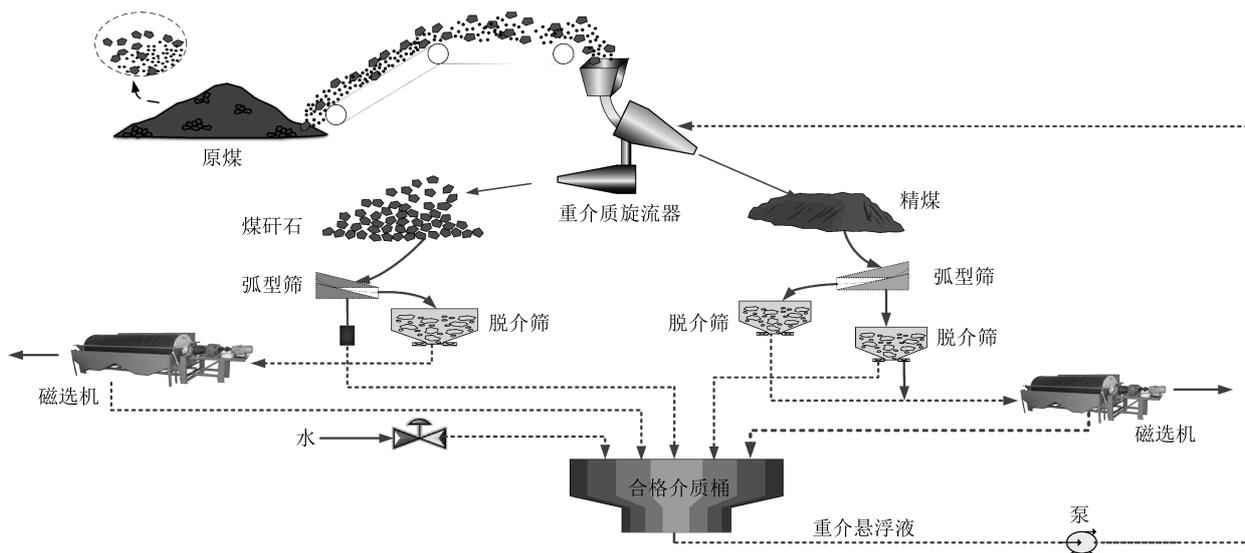


图2 重介质选煤工艺流程图

Fig.2 Process flow diagram of dense medium coal preparation

表1 原煤灰分与状态等级

Table 1 Raw coal ash and state level

原煤灰分化验值 (%)	状态等级及过渡过程
6.0~6.5	Optimal
6.5~6.7	Optimal 到 Fine 过渡
6.7~7.2	Fine
7.2~7.5	Fine 到 Medium 过渡
7.5~8.0	Medium
8.0~8.2	Medium 到 Poor 过渡
8.2~9.0	Poor

点, 这里选择在建模数据中随机选出小部分数据人工加入扰动的方式来模拟. 在生产数据的实际采集过程中, 状态等级为 Optimal 和 Fine 的建模数据要远少于 Medium 和 Poor, 因此, 在状态等级为 Optimal 和 Fine 的建模数据中随机挑选出 5% 的样本数据加入扰动<sup>[17]</sup>. 增加扰动的方法是将选定的 5% 的样本数据分成等数量的两部分:  $A$  子集和  $B$  子集, 对  $A$  子集的过程变量和  $B$  子集的灰分离线化验值分别加入幅度为 30% 左右的扰动. 同时, 对状态等级为 Medium 和 Poor 的建模数据中随机挑选出 10% 的样本数据, 并按照上述方法, 增加幅度为 30% 左右的扰动, 然后, 对加入离群点的建模数据重新划分数据集, 并用于离线评价模型的建立.

### 3.3 基于 Total-PRMR 的重介选煤过程运行状态在线评价

针对重介质选煤过程, 另外采集 605 组测试

数据用于算法验证. 这些测试数据中仍然包含表 1 中所有运行状态, 其中包括: Optimal、Fine、Medium、Poor 这 4 个稳定状态等级, 分别对应 98, 132, 208 和 102 组测试数据; Fine 到 Optimal 过渡、Medium 到 Fine 过渡、Poor 到 Medium 过渡 3 种过渡运行状态分别对应 21, 22 和 22 组测试数据. 在线测试的数据也按同样方式随即挑选 5% 的样本加入扰动, 模拟离群点.

为了验证本文所提方法的可行性, 分别用本文方法 (记为 Total-PRMR 方法) 与文献 [21] 中所提方法 (即基于 Total-PLS 的运行状态评价方法, 记为 Total-PLS 方法) 对上述数据进行建模和运行状态在线评价. 在线评价时窗口宽度均设置为  $H = 3$ , 评价结果如图 3 和图 4 所示. 对比图 3 和图 4 可以看出, 当数据中存在离群点时, 已有方法 (Total-PLS) 失去了原有的泛化能力. 而本文所提方法虽然出现了少许的误识别, 但基本与无离群点情况下的评价结果保持一致. 可以看出本文提出的基于全潜鲁棒偏 M 估计的最优状态的鲁棒评价方法可以有效地去除离群点的不利影响, 提高算法在线状态评价的准确率.

表 2 和表 3 对两种方法的运行状态评价识别准确率进行了对比. 这里, 在满足相似度阈值的情况下, 通过计算测试样本个数除以总的测试样本数得到识别准确率. 从表中可以看出, 在阈值范围为 0.6~0.8 时, Total-PRMR 算法的评价识别准确率要明显优于现有方法 (Total-PLS). 需要注意的是, 相似度阈值的选取, 会影响评价识别准确率. 因此, 需要结合实际的工业过程, 设定合理的相似度阈值.

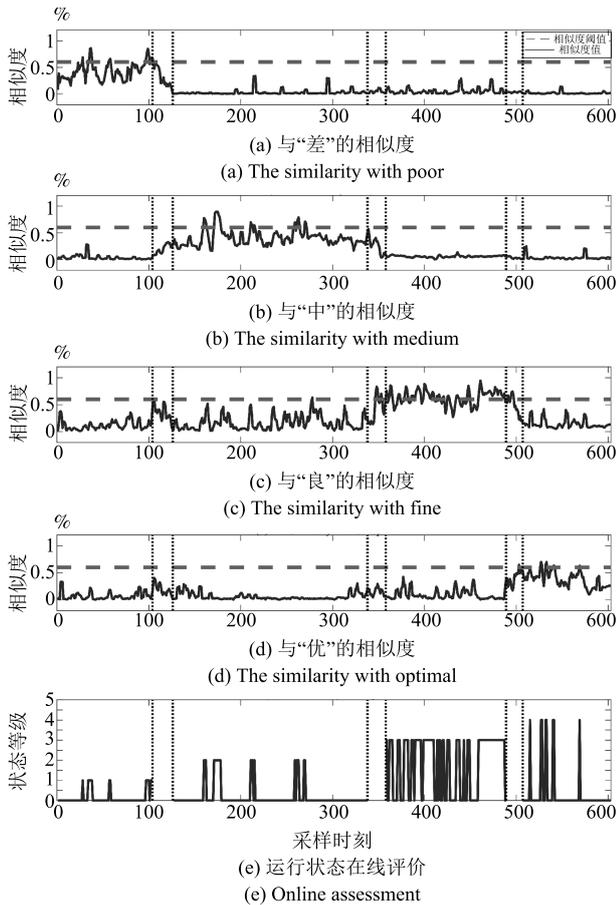


图3 Total-PLS方法的运行状态评价结果

Fig.3 Evaluation results of the operating state of Total-PLS method

表2 现有方法 (Total-PLS) 评价识别准确率 (%)

Table 2 The assessment identification accuracy rate of the existing method (Total-PLS based) (%)

相似度阈值	Poor (差)	Medium (中)	Fine (良)	Optimal (优)
$\epsilon \geq 0.6$	11	9	55	8
$\epsilon \geq 0.7$	4	5	33	1
$\epsilon \geq 0.8$	1	3	7	0

表3 本文方法 (Total-PRMR) 评价识别准确率 (%)

Table 3 The assessment identification accuracy rate of the proposed method (Total-PRMR based) (%)

相似度阈值	Poor (差)	Medium (中)	Fine (良)	Optimal (优)
$\epsilon \geq 0.6$	100	100	100	100
$\epsilon \geq 0.7$	93	88	81	94
$\epsilon \geq 0.8$	90	78	72	90

### 3.4 重介选煤过程的 Non-optimal 因素识别

Non-optimal 状态是指处于 Optimal 状态以外

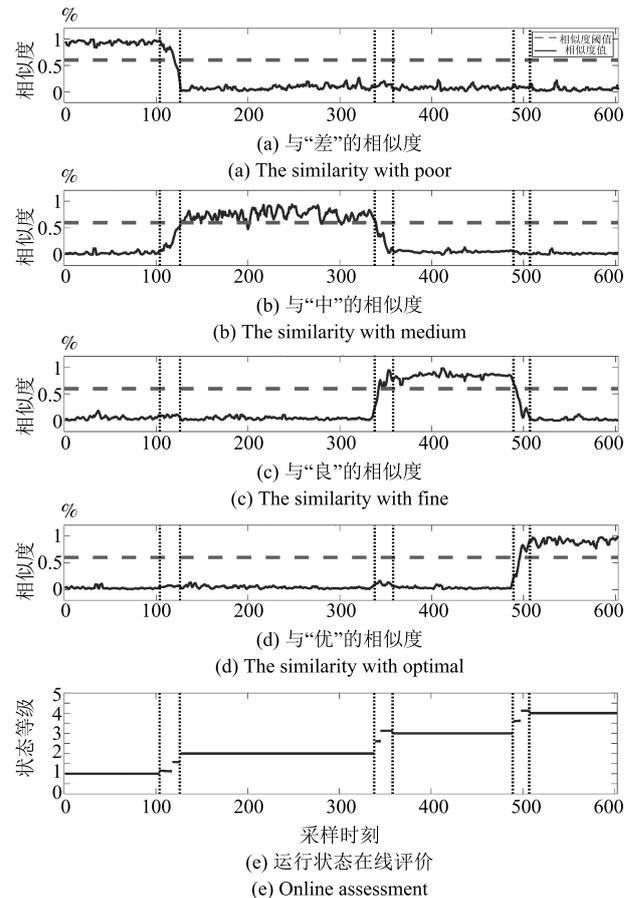


图4 Total-PRMR方法的运行状态评价结果

Fig.4 Evaluation results of the operating state of Total-PRMR method

的一种运行状态,可能是 Fine、Medium 或 Poor 稳定运行状态,也可能是过渡运行状态. 当实际选煤过程处于 Non-optimal 状态时,需要对导致该状态的非优因素进行识别.

本文选取了 50, 110, 300, 320, 400, 492 这几个非优状态等级的样本点进行非优因素的识别,仿真结果如图 5 所示. 其中, Z 轴为过程变量的贡献率,主系统介质密度、介质桶液位、主旋流器入口压力、磁性物的浓度和煤泥的含量这 5 个过程变量分别对应 X 轴上的 1, 2, 3, 4, 5 坐标, Y 轴是本文所选的分别取自不同非优状态等级的 6 个样本点. 以其中一个 492 样本点为例,从图中可以看出,在第 492 个样本点处,过程变量 3 的贡献率明显高于其他过程变量,所以认为过程变量 3 是导致运行非优的主要因素;进一步结合过程机理知识可以分析得出导致过程运行状态非优的实际原因与图中所提算法的识别结果一致,即:对于第 492 个样本点,在实际生产过程中由于主旋流器入口压力的突然增加,导致主旋流器入口处进料量的突然增加,这一扰动导致选煤过程出煤品质出现波动,此外进料量的突然改变

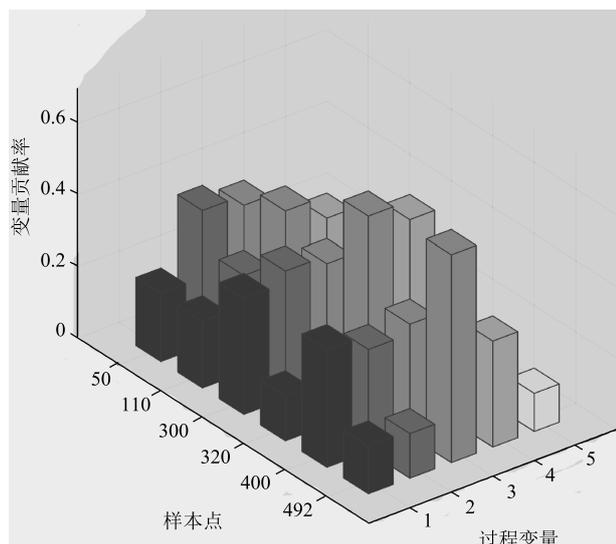


图5 重介质选煤过程非优因素识别

Fig. 5 Identification of non-optimal factors in the process of dense medium coal preparation

还会导致混料桶中磁性物溶度的变化,从图中可以看出除了变量3,变量4的贡献率也明显高于其他3个变量.类似的,可以分析验证其他5个样本点的非优因素识别结果也与实际运行情况的结果一致.因此,利用本文所提方法可以通过对非优因素的在线识别与追溯及时有效的发现导致重介质选煤过程运行非优的主要原因,进而给出合理的控制策略调控方案,保证选煤过程的高效运行.

综上所述,利用本文提出的运行状态评价方法能够及时对过程的运行状态做出评价,并识别导致过程非优的变量因素,这为实际生产过程中及时发现、调整策略并提高生产效率提供了理论依据,保证原煤分选过程的出煤品质和提高企业的综合经济效益.

## 4 结束语

复杂工业过程运行状态的在线准确且稳健的评价对指导生产过程的运行优化具有十分重要的实际意义.本文提出了一种基于数据驱动的复杂过程最优状态的鲁棒评价方法,并将所提方法应用于重介质选煤过程中.该方法能够有效克服离群点对评价模型精度的影响.通过从过程数据中挖掘出与产品综合经济指标相关的过程信息,利用这些信息建立评价模型,并在线结合三条评价准则进行复杂工业过程的在线运行状态评价,同时针对非最优运行状态,进行非优因素的识别,为现场操作人员实时调整生产策略提供依据.最后通过仿真实验,验证了所提方法的有效性.

## References

- Wen Chen-Lin, Lv Fei-Ya, Bao Zhe-Jing, Liu Mei-Qin. A review of data driven-based incipient fault diagnosis. *Acta Automatica Sinica*, 2016, **42**(9): 1285–1299 (文成林, 吕菲亚, 包哲静, 刘妹琴. 基于数据驱动的微小故障诊断方法综述. *自动化学报*, 2016, **42**(9): 1285–1299)
- Yin S, Gao H J, Qiu J B, Kaynak O. Fault detection for nonlinear process with deterministic disturbances: a just-in-time learning based data driven method. *IEEE Transactions on Cybernetics*, 2017, **47**(11): 3649–3657
- del Rio-Chanona E A, Zhang D D, Vassiliadis V S. Model-based real-time optimisation of a fed-batch cyanobacterial hydrogen production process using economic model predictive control strategy. *Chemical Engineering Science*, 2016, **142**: 289–298
- Rossi F, Copelli S, Colombo A, Pirola C, Manenti F. Online model-based optimization and control for the combined optimal operation and runaway prediction and prevention in (fed-)batch systems. *Chemical Engineering Science*, 2015, **138**: 760–771
- Ji H Q, He X, Shang J, Zhou D H. Incipient fault detection with smoothing techniques in statistical process monitoring. *Control Engineering Practice*, 2017, **62**: 11–21
- Frangos M. Uncertainty quantification for cuttings transport process monitoring while drilling by ensemble Kalman filtering. *Journal of Process Control*, 2017, **53**: 46–56
- Liu Yang, Zhang Guo-Shan. Chemical process monitoring and fault diagnosis based on sensitive sparse principal component analysis. *Control and Decision*, 2016, **31**(7): 1213–1218 (刘洋, 张国山. 基于敏感稀疏主元分析的化工过程监测与故障诊断. *控制与决策*, 2016, **31**(7): 1213–1218)
- Zou X Y, Wang F L, Chang Y Q, Zhang B. Process operating performance optimality assessment and non-optimal cause identification under uncertainties. *Chemical Engineering Research and Design*, 2017, **120**: 348–359
- Mosallam A, Medjaher K, Zerhouni N. Component based data-driven prognostics for complex systems: methodology and applications. In: *Proceedings of the 1st International Conference on Reliability Systems Engineering (ICRSE)*. Beijing, China: IEEE, 2015. 1–7
- Jia X D, Zhao M, Di Y, Yang Q B, Lee J. Assessment of data suitability for machine prognosis using maximum mean discrepancy. *IEEE Transactions on Industrial Electronics*, 2018, **65**(7): 5872–5881
- Li N P, Lei Y G, Guo L, Yan T, Lin J. Remaining useful life prediction based on a general expression of stochastic process models. *IEEE Transactions on Industrial Electronics*, 2017, **64**(7): 5709–5718
- Lei Y G, Li N P, Guo L, Li N B, Yan T, Lin J. Machinery health prognostics: a systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 2018, **104**: 799–834
- Qi Liang-Cai, Li Hong-Guang. A data-driven approach to multi-attribute assessment of controller performance. *Journal of East China University of Science and Technology (Natural Science Edition)*, 2014, **40**(2): 244–249 (齐良才, 李宏光. 基于数据驱动的控制性能多属性评价方法. *华东理工大学学报(自然科学版)*, 2014, **40**(2): 244–249)

- 14 Yan Xiang-Wu, Li Yan-Yan, Zhang He-Chuan, Wang Li-Na. Comprehensive evaluation of EV charging equipment performance base on variable weights analysis method. *Journal of Hunan University (Natural Sciences)*, 2014, **41**(5): 86–93  
(颜湘武, 李艳艳, 张合川, 王丽娜. 基于变权分析方法的电动汽车充电设备性能综合评价. 湖南大学学报(自然科学版), 2014, **41**(5): 86–93)
- 15 Melhem M, Ananou B, Ouladsine M, Pinaton J. Regression methods for predicting the product's quality in the semiconductor manufacturing process. *IFAC-PapersOnline*, 2016, **49**(12): 83–88
- 16 de Souza M F Z, Reis Y, Almeida A B, Lima I, de Souza A C Z. Load margin assessment of systems with distributed generation with the help of a neuro-fuzzy method. *IET Renewable Power Generation*, 2015, **9**(4): 331–339
- 17 Shirali G A, Mohammadfam I, Ebrahimipour V. A new method for quantitative assessment of resilience engineering by PCA and NT approach: a case study in a process industry. *Reliability Engineering and System Safety*, 2013, **119**: 88–94
- 18 Liao Rui-Jin, Wang Qian, Luo Si-Jia, Liao Yu-Xiang, Sun Cai-Xin. Condition assessment model for power transformer in service based on fuzzy synthetic evaluation. *Automation of Electric Power Systems*, 2008, **32**(3): 70–75  
(廖瑞金, 王谦, 骆思佳, 廖玉祥, 孙才新. 基于模糊综合评判的电力变压器运行状态评估模型. 电力系统自动化, 2008, **32**(3): 70–75)
- 19 Zhang Kun, Shen Hai-Bo, Zhang Hong, Jiang Li-Ming, Zhong Yi. Synthesis evaluation method for node importance in complex networks based on grey relational analysis. *Journal of Njing University of Science and Technology (Natural Science)*, 2012, **36**(4): 579–586  
(张琨, 沈海波, 张宏, 蒋黎明, 衷宜. 基于灰色关联分析的复杂网络节点重要性综合评价方法. 南京理工大学学报(自然科学版), 2012, **36**(4): 579–586)
- 20 Luo Yi, Zhou Chuang-Li, Liu Xiang-Jie. Application of the multi-level grey relational analysis method in operation assessment of thermal power units. *Proceedings of the CSEE*, 2012, **32**(17): 97–103  
(罗毅, 周创立, 刘向杰. 多层次灰色关联分析法在火电机组运行评价中的应用. 中国电机工程学报, 2012, **32**(17): 97–103)
- 21 Wu Qi. Research on Heavy Medium Coal Preparation Process Operating Performance Assessment Method Based on Process Data [Master thesis], China University of Mining and Technology, China, 2016.  
(吴奇. 基于过程数据的重介质选煤过程运行状态评价方法研究 [硕士学位论文], 中国矿业大学, 中国, 2016.)
- 22 Vo H X, Durlafsky L J. Data assimilation and uncertainty assessment for complex geological models using a new PCA-based parameterization. *Computational Geosciences*, 2015, **19**(4): 747–767
- 23 Liu Y, Wang F L, Chang Y Q. Operating optimality assessment and cause identification for nonlinear industrial processes. *Chemical Engineering Research and Design*, 2017, **117**: 472–487
- 24 Liu Yan, Chang Yu-Qing, Wang Fu-Li. Online assessment of process operating performance based on Fisher discriminant analysis. *Control and Decision*, 2014, **29**(9): 1655–1660  
(刘炎, 常玉清, 王福利. 基于 Fisher 判别分析的过程运行状态在线评价. 控制与决策, 2014, **29**(9): 1655–1660)
- 25 Liu Y, Chang Y Q, Wang F L. Online process operating performance assessment and nonoptimal cause identification for industrial processes. *Journal of Process Control*, 2014, **24**(10): 1548–1555
- 26 Paalanen P, Kamarainen J K, Ilonen J, Kälviäinen H. Feature representation and discrimination based on Gaussian mixture model probability densities-practices and algorithms. *Pattern Recognition*, 2006, **39**(7): 1346–1358
- 27 Ye L B, Liu Y M, Fei Z S, Liang J. Online probabilistic assessment of operating performance based on safety and optimality indices for multimode industrial processes. *Industrial and Engineering Chemistry Research*, 2009, **48**(24): 10912–10923
- 28 Zou X Y, Wang F L, Chang Y Q, Zhao L P, Zheng W. Two-level multi-block operating performance optimality assessment for plant-wide processes. *The Canadian Journal of Chemical Engineering*, DOI: 10.1002/cjce.23159, 2018.
- 29 Liu Y, Wang F L, Chang Y Q, Ma R C. Operating optimality assessment and nonoptimal cause identification for non-Gaussian multimode processes with transitions. *Chemical Engineering Science*, 2015, **137**: 106–118
- 30 Liu Y, Wang F L, Chang Y Q. Operating optimality assessment based on optimality related variations and nonoptimal cause identification for industrial processes. *Journal of Process Control*, 2016, **39**: 11–20
- 31 Zhou D H, Li G, Qin S J. Total projection to latent structures for process monitoring. *AIChE Journal*, 2010, **56**(1): 168–178
- 32 Jia Run-Da, Mao Zhi-Zhong, Wang Fu-Li. Identification of hybrid models for batch reaction processes based on kernel partial robust M-regression. *Journal of Chemical Engineering of Chinese Universities*, 2014, **28**(1): 115–122  
(贾润达, 毛志忠, 王福利. 基于核偏鲁棒 M-回归的间歇反应过程混合模型辨识. 高校化学工程学报, 2014, **28**(1): 115–122)
- 33 Chu F, Ma X P, Wang F L, Jia R D. Novel robust approach for constructing Mamdani-type fuzzy system based on PRM and subtractive clustering algorithm. *Journal of Central South University*, 2015, **22**(7): 2620–2628
- 34 Jia R D, Mao Z Z, Chang Y Q, Zhang S N. Kernel partial robust M-regression as a flexible robust nonlinear modeling technique. *Chemometrics and Intelligent Laboratory Systems*, 2010, **100**(2): 91–98
- 35 Zhao Chun-Xiang, Ye Gui-Sen. Study of heavy medium coal preparation process control model and control algorithm. *Journal of China Coal Society*, 2000, **25**(z1): 196–200  
(赵春祥, 叶桂森. 重介质选煤过程控制模型及控制算法的研究. 煤炭学报, 2000, **25**(z1): 196–200)
- 36 Peng K X, Zhang K, Li G, Zhou D H. Contribution rate plot for nonlinear quality-related fault diagnosis with application to the hot strip mill process. *Control Engineering Practice*, 2013, **21**(4): 360–369
- 37 Cao Zhen-Guan. Study of Prediction Control on Heavy Medium Density in the Process of Coal Preparation [Ph. D. dissertation], China University of Mining and Technology, China, 2014.  
(曹珍贵. 重介选煤过程中重介质的密度预测控制研究 [博士学位论文], 中国矿业大学, 中国, 2014.)

- 38 Chu F, Wu Q, Dai W, Ma X P, Wang F L. Latent variable techniques based operating performance assessment for the dense medium coal preparation process. In: Proceedings of the 35th Chinese Control Conference (CCC). Chengdu, China: IEEE, 2016. 9821–9826



**褚 菲** 中国矿业大学信息与控制工程学院副教授. 2014 年获中国东北大学控制理论与控制工程博士学位. 主要研究方向为复杂工业过程的建模, 控制与优化, 统计过程监测及运行状态评价. 本文通信作者.

E-mail: chufeizhufei@sina.com

(**CHU Fei** Associate professor at the College of Information and Control Engineering, China University of Mining and Technology. He received his Ph.D. degree from Northeastern University in 2014. His research interest covers modeling, control and optimization of complex industrial process, statistical process monitoring, and operating performance assessment. Corresponding author of this paper.)



**赵 旭** 中国矿业大学信息与控制工程学院硕士研究生. 2017 年获三江学院机械与电气工程学院学士学位. 主要研究方向为复杂工业过程运行优化及最优状态评价. E-mail: zhao\_xu1994@126.com

(**ZHAO Xu** Master student at the College of Information and Control Engineering, China University of Mining

and Technology. He received his bachelor degree from San Jiang College in 2017. His research interest covers optimization of complex industrial process and optimal performance assessment.)



**代 伟** 中国矿业大学信息与控制工程学院副教授. 2015 年获中国东北大学控制理论与控制工程博士学位. 主要研究方向为复杂工业过程的运行优化控制.

E-mail: weidai@cumt.edu.cn

(**DAI Wei** Associate professor at the College of Information and Control Engineering, China University of Mining and Technology. He received his Ph.D. degree from Northeastern University in 2015. His research interest covers operation optimization control of complex industrial process.)



**马小平** 中国矿业大学信息与控制工程学院教授. 2001 年获中国矿业大学信息与电气工程博士学位. 主要研究方向为过程控制, 网络化控制系统及故障检测.

E-mail: xpma@cumt.edu.cn

(**MA Xiao-Ping** Professor at the College of Information and Control Engineering, China University of Mining and Technology. He received his Ph.D. degree from China University of Mining and Technology in 2001. His research interest covers process control, networked control system, and fault detection.)



**王福利** 东北大学教授. 1988 年获东北大学自动化系博士学位. 主要研究方向为复杂工业系统的建模、控制与优化, 过程监测和故障诊断.

E-mail: wangfuli@ise.neu.edu.cn

(**WANG Fu-Li** Professor at Northeastern University. He received his Ph.D. degree from Northeastern University in 1988. His research interest covers modeling, control and optimization of complex industrial process, process monitoring, and fault diagnosis.)