

控制与决策

Control and Decision

复杂工业过程运行状态评价方法回顾与展望

褚菲, 郝莉莉, 王福利

引用本文:

褚菲, 郝莉莉, 王福利. 复杂工业过程运行状态评价方法回顾与展望[J]. *控制与决策*, 2024, 39(3): 705–718.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0359>

您可能感兴趣的其他文章

Articles you may be interested in

[工业信息物理系统安全风险动态表现分析量化评估模型](#)

Quantitative evaluation model for dynamic performance analysis of security risk in industrial cyber physics systems

控制与决策. 2021, 36(8): 1939–1946 <https://doi.org/10.13195/j.kzyjc.2019.1479>

[基于马尔可夫过程的多部件系统劣化状态空间划分模型](#)

Multi-component system state space partition model based on Markov process

控制与决策. 2021, 36(2): 418–428 <https://doi.org/10.13195/j.kzyjc.2019.0480>

[数据驱动的综合能源系统运行优化方法研究综述](#)

Review of research of data-driven methods on operational optimization of integrated energy systems

控制与决策. 2021, 36(2): 283–294 <https://doi.org/10.13195/j.kzyjc.2020.0725>

[基于多块信息提取的AUV资源勘查系统故障检测](#)

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790–800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

[基于随机系数回归模型的退化过程及维修策略](#)

Degradation process and maintenance planning based on random coefficient regression model

控制与决策. 2021, 36(3): 754–760 <https://doi.org/10.13195/j.kzyjc.2019.1087>

复杂工业过程运行状态评价方法回顾与展望

褚菲^{1,2,3†}, 郝莉莉^{1,2}, 王福利⁴

1. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116;
2. 中国矿业大学 地下空间智能控制教育部工程研究中心, 江苏 徐州 221116;
3. 中国矿业大学 人工智能研究院, 江苏 徐州 221116;
4. 东北大学 信息科学与工程学院, 沈阳 110819)

摘要: 准确感知和认知复杂工业过程的运行状态对于实现过程智能控制和优化决策至关重要, 是当前实现工业人工智能需要解决的关键问题之一. 传统过程监测理论系统已不能满足现代工业生产过程对过程运行状态认知的精细化及准确化的需求, 因此, 复杂工业过程运行状态评价技术应运而生, 近几年受到学术界和工业界广泛关注并快速发展. 对此, 首先从复杂工业过程的主要特性以及数据提取过程中面临的问题出发, 回顾基于数据驱动的相关工业过程运行状态评价方法; 然后根据最优性评价结果总结导致状态非“优”的原因, 并进一步给出相关非优因素追溯方法; 最后对现有研究内容和这一领域中值得进一步研究的发展方向做出总结和展望.

关键词: 复杂工业过程; 运行状态评价; 数据驱动; 非优因素

中图分类号: TP273 文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0359

引用格式: 褚菲, 郝莉莉, 王福利. 复杂工业过程运行状态评价方法回顾与展望[J]. 控制与决策, 2024, 39(3): 705-718.

Review and prospect of operation performance assessment methods for complex industrial processes

CHU Fei^{1,2,3†}, HAO Li-li^{1,2}, WANG Fu-li⁴

1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China;
2. Research Center of Underground Space Intelligent Control Engineering of the Ministry of Education, China University of Mining and Technology, Xuzhou 221116, China;
3. Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou 221116, China;
4. College of Information Science and Engineering, Northeastern University, Shenyang 110819, China)

Abstract: It is of great significance to accurate perception and cognition of the operating performance of complex industrial processes for the realization of process intelligent control and optimization decision. This is also one of the key issues to be solved for the current realization of industrial artificial intelligence. The traditional process monitoring theory system cannot meet the demands of the modern industrial production process for the refinement and accuracy of the process operating state cognition. Hence, the operation performance assessment technology of complex industrial processes comes into being. In recent years, it has attracted the attention of academia and industry and developed rapidly. Based on the main characteristics of complex industrial processes and the problems faced in the process of data extraction, this paper reviews the operation performance assessment methods of related industrial processes based on data-driven. Then, according to the results of non-optimal evaluation, the reasons leading to non-optimal performance are summarized, and the traceability method of relevant non-optimal factors is further given. Finally, a summary and outlook of existing research and directions of development in this area that merit further research is provided.

Keywords: complex industrial process; operation performance assessment; data-driven; non-optimal factors

0 引言

流程工业^[1]广泛存在于矿物加工、冶金、石化等各个领域. 随着工业化与网络化信息的全面交融, 在

工业 4.0 背景下, 过程工业正朝着绿色、高效、集成化方向发展, 并具有向智慧企业发展的趋势^[2]. 过程工业实现自动化运行有助于企业具备更强的国际竞争

收稿日期: 2023-03-27; 录用日期: 2023-06-28.

基金项目: 国家自然科学基金项目(61973304, 61873049, 62073060); 江苏省六大人才高峰项目(DZXX-045).

责任编辑: 牛玉刚.

†通讯作者. E-mail: chufei@cumt.edu.cn.

力,并且可以加快国家实现节能减排降耗的步伐.但是,针对复杂工业过程往往具有生产原料成分频繁变化,生产设备状态不稳定,生产环境复杂多变等一系列的不确定因素导致生产过程无法采用精准的数据模型来解释,进一步导致生产参数和产品质量等无法得到保障.因此,实现复杂工业过程的平稳高效运行成为进一步实现其智能控制和优化决策的关键.

运行状态稳定是提升产品质量和综合经济效益的必要条件.然而,对于运行在恶劣开放环境中的过程而言,例如矿加、冶金、选煤等过程,不确定性因素及扰动频繁发生,导致其无法长期稳定运行在最优工作点附近,常常需要人的不断干预才能保证其正常运行,生产效率和综合经济效益难以保证.实际过程中的各种不确定干扰,包括原材料的波动和环境变化,导致操作点偏离最佳条件,使流程工业面临着高资源消耗、低附加值产品、高环境污染等问题.另外,工业系统变得越来越昂贵,对运行性能、生产率、可用性、可靠性和安全性也提出了更高的要求.因此,在工业过程运行正常的基础上,进一步评价和精细识别当前运行状态具有重要意义,这也与倡导过程工业智能化、生产高质量化的目标一致.同时,工业过程的运行状态的感知和认知技术也被视为工业人工智能3个关键技术之一^[3-5].目前,过程监控技术已成功地应用于矿加、冶金、石化等重要工业领域,相关学者已做了大量的研究工作^[6-7].

从实际应用的角度出发,随着企业对生产要求的日益严格以及计算机设备的强大功能性和大数据运维的发展趋势,大部分企业对工业生产要求不但要保证产品质量和生产效率,而且也对工业生产过程能安全平稳地运行提出了越来越高的要求.现代工业过程不仅在产品工艺、过程、技术以及设施等都逐渐迈向自主化、复杂化生产,而且容易受工业生产过程本身结构复杂、工艺流程长、产品工艺和设备可调节性、外部环境变化等多个因素干扰.为了减少整个生产过程因为异常变化而造成严重的生产事故,避免产品质量问题,建立精确可靠的状态监控模型,开发工业过程运行状态监测方法成为研究人员的热点话题.传统的工业过程状态监测是指分析、了解和观察运行设备的生产过程状态,其中包括对生产设备的检测、观察、运行数据记录和及时做出运行环境的评估.结合整个运行工况的现状以及历史数据,考虑外界环境的原因,对整个运行过程状态进行评估,判断设备和运行环境是否处于良好和安全的运行水平^[8].运行状态评价方法主要是从传统过程监测方法

的基础上发展而来的,是指对处于正常的工业过程状态进一步细分为“优”“良”“中”“差”等不同的等级.之所以说过程运行状态评价是从过程监测的基础上发展起来的,主要是因为在实际生产过程中运行状态评价等级对应的运行状态的变化幅度远远小于过程监测中的“正常”与“故障”之间的变化幅度.因此,这对传统的过程监测方法提出了更严苛的要求,从而促进更多学者提出具有更高灵敏性和辨别细微差别的能力的分析方法,进而可以更精准提取过程状态不同等级之间的差异度信息,能够更加灵活地适应实际生产工业过程运行状态评价问题的需求.进一步,以综合经济效益作为评价依据,就需要提取与综合经济指标密切相关的过程变化信息,这也是过程运行状态评价方法与传统过程监测方法最本质的区别所在.目前,专门针对工业过程运行状态评价的研究热度正呈现逐年上升的趋势.

早在20世纪中期,人们就将专家经验作为一种评价准则并运用到评价问题中,专家评分法是基于定量和定性分析打分的形式给出一些相对规则化的评价.之后的20年间科学技术飞速发展,越来越多的评价方法应运而生并应用到生产生活的各个领域.到了20世纪末期,人们通过纵向挖掘,使得更加专业化的理论分析方法(通过结合其他领域专业知识的新的评价方法)也被逐渐开发出来.例如,早期通过建立拓扑结构将决策相关的指标进行分解并以层次的形式表示的层次分析法(analytic hierarchy process, AHP)^[9],现在通过结合概率统计分析的多元统计分析法以及神经网络拓展的人工智能神经网络分析法等.生产过程运行状态评价方法有很多,但每类方法的研究思路大体是一致的,首先了解评价对象,其次确定评价指标,然后设置每个评价指标的权重并建立评价机制,最后观察评价结果等.其中,确定评价指标、设置权重参数、建立评价机制是生产过程运行状态体系中的重要部分.

随着计算机技术的发展,状态评价应用已经发展到社会生活的各个领域,渗透到日常生活的方方面面,例如人员的综合素质评价^[10]和国家综合经济水平评价^[11]等.由此可见,状态评价与人们的日常生活紧密相关.过程运行状态评价方法是通过了解实际生产过程运行状况,判断将来发展结果,比较实际目标得出差异,并进一步分析引起生产过程外部和内部条件变化的根本因素.随着分布式控制系统的普及,大量的流程工业数据被采集下来,在这种历史条件下,服务于工业过程的控制、调度与优化.因此,为了

满足各类实际生产过程变化产生的原因分析,很多运行状态评价方法应运而生. 这些方法主要分为两大类:基于多元统计的运行状态评价和基于神经网络的运行状态评价,如表1所示. 其中基于多元统计的运行状态评价又结合工业过程运行特点分为基于过

程变量间具有相关关系的评价、基于多模态的评价、基于定向定量信息共存的评价、基于动态感知的评价等. 接下来的讨论主要围绕复杂工业过程的运行状态评价方法展开,首先简单介绍一下复杂工业过程的基本特性.

表1 运行状态评价方法的优缺点对比

方法	特征	优点	缺点
多元统计分析法	通过利用各种统计工具(如方差、协方差、回归分析等)对运行状态数据进行分析	容易理解,解释性更强,易于实现,具有一定的鲁棒性	对数据的规律和特征要求较高,易受样本容量影响,缺乏预测性,处理复杂数据类型时结果不理想
神经网络分析法	使用神经网络对运行状态数据进行学习和预测,根据预测结果评价运行状态	当系统非线性、数据高维时,预测能力强;能够学习到数据特征和规律,模型的鲁棒性强	数据需求较高,需要大量有标记的训练数据进行训练;解释性能力较弱;模型结构设计复杂

1 复杂工业过程描述

关于复杂工业过程,目前还没有具体的定义给出. 一般根据复杂性可以分为两类:客观复杂性和认识复杂性. 客观复杂性主要是指工业过程在运行中某种运动或模态的过渡变化,即跨流程分层次后整合的、不可还原的新的模态和相互关系;认识复杂性可以理解为是工业过程中复杂性的有效理解和表达^[12-13]. 因此,传统的矿加、冶金、化工是依据具体的实验验证和专家经验科学,通过采用实验室反复实验累计的数据进一步提高生产效率,或者结合数学理论将过程变量的可解释性提高,即采用微分方程、数学分析等对生产过程产生的复杂现象进行合理化解,进一步指导过程工艺. 由于复杂工业过程存在全流程包裹着众多子过程的特点,而子过程之间关系错综复杂,跨流程分层次的结构也复杂并且子过程之间的耦合程度很高,很难采用一般的线性或者简单的非线性关系来描述. 复杂工业过程一般存在以下几种特性:数据维数高、非线性、非高斯和生产过程多模态等.

2 工业过程运行状态评价技术回顾

本节面向复杂工业过程运行状态评价理论和方法展开概述. 首先介绍基本的复杂工业过程数据处理技术;在此基础上,从复杂工业过程特性的角度出发,分别基于不同过程特点回顾数据驱动的运行状态评价方法;进而分析非优原因追溯查找非优变量,给技术人员提供进一步的操作依据.

2.1 运行状态优性评价的基础理论

针对实际生产数据复杂维度高,变量间存在强非线性和强耦合关系,数据具有动态特性等特点,本节着重介绍几种典型的数据分析方法.

2.1.1 主成分分析(principal component analysis, PCA)

PCA算法^[14]是一种常见的降维方法,PCA的主要思想是将 N 维特征映射到 K 维上并且两两正交,这 K 维正交的特征被称为主成分. PCA是通过线性变化获取主成分变量,实现从高维空间映射到低维空间的操作并期望尽可能多地保留原始数据的信息量,以使用较少的数据维度,同时保留较多的原数据的特性.

PCA由一个二维数据矩阵 $X \in R^{N \times J}$ 构成,表示由过程变量测量值构成的数据集. 通过从 X 中提取携带大量信息的成分向量 t ,获取主成分的方差信息 $\text{Var}(t) = t^T t$ 来求取 X 中的主成分向量 t_1, t_2, \dots, t_K ,其中 K 是保留主成分的个数. 通常情况下主成分呈现为 $\|t_1\| \geq \|t_2\| \geq \dots \geq \|t_K\|$,即对应的方差信息的特征值依次减小. 主成分分解的公式可以表示为

$$X = TP^T + E = \sum_{k=1}^A t_k p_k^T + E. \quad (1)$$

其中: $t_k \in R^{N \times 1}$ 代表得分向量, $p_k \in R^{J \times 1}$ 代表负载向量, $T \in R^{N \times K}$ 和 $P \in R^{J \times K}$ 分别代表得分矩阵和负载矩阵, $E \in R^{N \times J}$ 代表 X 的残差矩阵.

2.1.2 偏最小二乘(partial least square, PLS)

PLS算法^[15]用来处理数据共线性、回归建模和两组变量间的相关分析等问题. 假设存在一组自变量 $X = \{x_1, \dots, x_p\}$ 和一组因变量 $Y = \{y_1, \dots, y_q\}$, p 和 q 分别是自变量和因变量的个数. PLS算法分别从 X 和 Y 中提取相应的得分向量 t 和 u ,在保证 t 和 u 尽可能携带各自数据信息的前提下,使 t 和 u 的协方差最大化,即

$$\begin{aligned} \max_{w_1, c_1} \langle Xw, Yc \rangle &= \max\{\text{cov}(t, u)\}; \\ \text{s.t.} \quad &\begin{cases} w^T w = 1, \\ c^T c = 1. \end{cases} \end{aligned} \quad (2)$$

其中 w_1 和 c_1 分别是 X 和 Y 的第1个投影方向, 并且满足 $\|w_1\| = 1$ 和 $\|c_1\| = 1$. 然后建立 X 和 Y 对 t_1 和 u_1 的回归方程, 如下式所示:

$$X = t_1 p_1^T + E_1, \quad (3)$$

$$Y = u_1 q_1^T + F_1. \quad (4)$$

其中: t_1 和 u_1 为 X 和 Y 的第1得分向量, 向量 p_1 和 q_1 分别是与之对应的负载向量.

对 E_1 和 F_1 再次提取得分向量, 直到 E_a 和 F_a 不含主要的信息, 如下式所示:

$$\begin{cases} X = TP^T + E_a = \sum_{i=1}^a t_i p_i^T + E_a, \\ Y = UQ^T + F_a = \sum_{i=1}^a u_i q_i^T + F_a. \end{cases} \quad (5)$$

其中: T 和 U 是对 X 和 Y 提取的协方差最大的特征, P 和 Q 表示 X 和 Y 在此方向上相关性最大的向量. 由此得到 PLS 模型为

$$Y = TQ^T + F. \quad (6)$$

2.1.3 慢特征分析(slow feature analysis, SFA)

SFA^[16] 是一种对时序数据分析的模型, 其目的是从时序数据中识别缓慢变化的成分, 挑选慢度最小的特征为最主要的特征, 其优化问题可描述如下:

$$\min_{s_j} \langle \dot{s}_j^2(t) \rangle \text{ or } \langle \Delta s_j^2(t) \rangle; \quad (7)$$

其目标函数的约束为

$$\langle s_j^2(t) \rangle = 1, \quad (8)$$

$$\langle s_j(t) \rangle = 0, \quad (9)$$

$$\forall i \neq j, \langle s_i(t) s_j(t) \rangle = 0. \quad (10)$$

其中: $\dot{s}_j(t) = s_j(t) - s_j(t-1)$ 表示慢度, $\langle \cdot \rangle_t$ 是对时间的期望. 约束(8)和(9)使提取的特征具有单位协方差和零均值, 约束(10)确保每次提取的特征之间都线性无关. 对 SFA 进行 SVD 分解, 有

$$R = UAU^T. \quad (11)$$

其中: U 是特征向量所组成的矩阵, Λ 是相对应的特征值为主对角线的矩阵. 基于式(11), 可以得到白化矩阵 $Q = \Lambda^{-1/2}U^T$, 同时白化过程可表示为

$$Z = \Lambda^{-1/2}U^T X = QX. \quad (12)$$

可知式(12)满足 $\langle zz^T \rangle_t = Q \langle xx^T \rangle Q^T = I$ 以及 $\langle z \rangle_t = 0$. 因此, 优化问题变成求解矩阵 W , 使其满足

$$s = WX = WQ^{-1}Z. \quad (13)$$

结合约束(8)可得

$$\langle ss^T \rangle_t = I = P \langle zz^T \rangle P^T = PP^T = I. \quad (14)$$

因此, SFA 的优化问题是求解正交矩阵 P 使得 $\langle s_t^2 \rangle_t$ 最小化. 再次进行 SVD 分解来解决优化问题, 得到

$$\langle \dot{z}z^T \rangle_t = P^T \Omega P. \quad (15)$$

经上述分析, 基于 SFA 方法建立的模型可以表示为

$$W = PQ = P\Lambda^{-1/2}U^T. \quad (16)$$

2.2 基于多元统计分析的运行状态评价

随着现代智能工业过程技术的高速发展, 大多数工业生产过程的传感测量设备都已经完善, 可以在线获得大量完整的过程变量数据, 例如浓度、压力、流量等. 这些过程变量所隐含的数据信息有一些可能严重影响整个生产过程平稳运行, 所以即时对整个生产过程运行状态进行评价是重要且有意义的. 然而在实际生产过程中存在一些特殊的工业流程, 同时也会存在由于外部因素的干扰而影响内部变量的变化无法实时感知的问题. 在这种情况下, 可能需要借助离线建模分析、专家经验、软测量等多个方法的结合来获取可利用的数据信息. 但是, 无论利用何种方法获取数据信息, 在复杂工业过程运行评价的分析中, 一些主要的问题也会显现出来. 图1给出了复杂工业过程运行状态评价的基本流程.

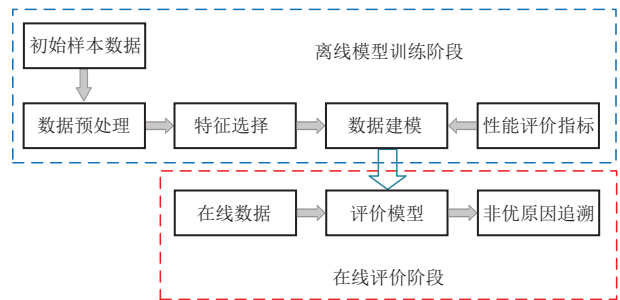


图1 复杂工业过程运行状态评价流程

2.2.1 基于过程变量间相关关系的运行状态评价

近10年, 针对过程生产的运行状态评价问题的研究逐渐受到相关学者的广泛关注. 在工业过程中大量的变量之间有着复杂的相关性, 这里主要讨论过程变量间线性相关和非线性相关. 首先, 为了追求最优的综合经济效益 (comprehensive economic benefit, CEB), 工艺运行绩效的在线优化评价成为关键问题. Liu等^[17]提出了一种基于状态相似度的在线经营性能评估和基于可变贡献的非最优原因识别方法,

通过应用全投影潜结构(total projection to latent structures, T-PLS)于过程变量线性相关的性能评估,利用T-PLS对删除异常值之后的数据进行重建,即通过T-PLS模型提取与CEB指标密切相关的变量进行过程的运行状态评价.在线评估结果不仅包含了确定性的性能等级,还包含了现有评估方法未涉及的性能等级转换;当评估结果非最优时,开发了一种基于变量贡献的非最优原因自动识别策略,对指导进一步的生产调整具有重要意义.但是,在考虑建立T-PLS模型之前需要对建模过程数据与CEB指标进行时序上的对应调整,对于复杂工业过程数据而言,这不仅是一件耗时耗力的工作,而且还大大降低了工作效率.因此,Liu等^[18]进一步提出了一种基于最优性相关变化和非最优原因识别方法的工业过程操作最优性在线评估方法.通过分析各稳定性能等级之间常见和唯一的变化,提取各稳定性能等级的最优相关变量,避免了耗时的数据比对,提高了与经营绩效相关的过程变量和排除不相关过程变量信息的能力,有效提高了该算法对劣性无关信息改变的稳健性和对劣性相关信息改变的敏感性.并且基于在线数据的最优性变化与各稳定性能等级之间的相似性,可将工艺操作性能评价为稳定性能等级或性能等级之间的转换过程,这为深入理解工艺操作提供了更多的信息.

实际工业过程中由于过程扰动和一些不确定性因素的存在,导致过程数据非线性特性显著,因此,基于线性假设的特征提取算法的评价结果并不理想.为了有效提取非线性数据中的信息,通过引入核技术,并将上述所提到的两种经典的线性相关评价方法延伸到非线性领域,Liu等^[19]针对复杂工业过程,通过将核投影潜结构(kernel projection to latent structure, KPLS)算法^[20]与T-PLS算法相结合,提出了一种基于核全投影潜结构(kernel total projection to latent structure, KT-PLS)的评价方法,并建立了各性能等级的评估模型,使其能够对非线性过程进行状态评价,通过改进T-KPLS模型结构不仅可以快速提取过程数据的非线性,而且还能精准提取与CEB相关的过程变量.由于基于KT-PLS的评价过程中也存在数据对齐问题,Liu等^[21]提出了一种基于非线性最优性信息(nonlinear optimality related variation information, NORVI)的运行状态评价方法,该方法利用多空间主成分分析(multi-space principal component analysis, MsPCA)模型,将过程数据分解为特有信息和共有信息并建立了状态评价模型,在不

借助具有综合指标数据的前提下,提高算法对无关信息的鲁棒性以及对相关信息的敏感性,从而依据各性能等级的实验数据与建模数据的非线性最优性能等级的相似性建立评价指标,并在湿法冶金过程工艺中进行了验证.除此之外,基于核主元分析(kernel principal component analysis, KPCA)^[22]、支持向量机(support vector machine, SVM)^[23]、核独立主元分析(kernel independent component analysis, KICA)^[24-25]等非线性特征提取方法也得到了相应的改进.最近,Chu等^[26]针对在解决工业过程中的非线性通常会忽略各个状态等级的局部邻域信息的问题,提出了一种核局部线性嵌入偏最小二乘(kernel locally linear embedding partial least squares, KLLEPLS)方法,通过在离线建模部分将核局部线性嵌入提取的局部信息嵌入到偏最小二乘中,提出KLLEPLS算法.该算法具有PLS算法最大化过程变量与综合经济指标(comprehensive economic indicators, CEI)相关性的能力,同时具有KLLE算法保持局部非线性结构的能力,从而实现了工业过程中具有强非线性关系的运行状态评价.尽管通过非线性特征提取算法能够提取反映过程状态的主要内容,但在实际工业现场应用时评价结果并不理想.其中一个因素是受到各种现场复杂环境的干扰,使得过程数据中包含大量噪声甚至是离群点,从而造成原先建立状态评价模型框架的泛化性被大大降低,从而削弱了模型评价结果的准确性.

虽然工业过程通常在设计条件下表现完美,但由于参数漂移、环境干扰等原因,它们可能会偏离最佳操作点.为此,褚菲等^[27]针对实际工业过程中由于外部环境变化而引起过程数据出现离群点、数据不准确以及噪声大等问题导致无法建立可靠的评价模型,提出了一种全潜鲁棒偏 M 估计(total partial robust M -regression, Total-PRMR)方法,自适应地为不同的样本数据分配不同的权重,从而克服异常值对模型的影响.进一步,Chu等^[28]基于性能相似度特性同时考虑具有非线性和异常值的复杂工业过程操作最优性评估问题,提出了一种改进的潜在结构全鲁棒核投影算法(robust total kernel partial M -regression, T-KPRM)的工业过程运行状态在线评价方法.该方法有效减少了运行过程状态评价由于环境噪声和离群点造成的不利影响,并利用滑动数据窗口与最优评价水平之间的欧氏距离来衡量变量的贡献率.该方法在重介质选煤过程运行状态等级划分的应用上也体现出很好的优越性.

2.2.2 多模态工业过程的运行状态评价

以上方法都是针对单一模态过程进行研究的,然而,在复杂的工业过程中,由于生产流程长、规模宏大、变量因素众多等问题,以及负荷变化、原料变化和参数漂移等,许多工业过程,包括连续和批量过程,都具有多模特性.即整个运行过程不止存在一种单一的运行模态,因为产品设定值的变化、原料的波动以及根据实际生产需求的调整,都会产生过程运行于不同的工作点的现象.在实际工业过程中,操作人员为了得到不同的产品会相应地调整操作条件,从而导致不同的稳定模式.从一种稳定模式到另一种稳定模式的调整过程可视为过渡过程.稳定模式是生产效率高、占用生产时间最多的主要过程,稳定模式良好的运行性能是CEI提高的保证.然而,由于生产条件不稳定,过渡过程通常会产生不合格的产品,这可能会影响到CEI.如果能及时准确地评估非最优过渡模式,则管理者和经营者会采取适当的措施来调整和改过渡过程,从而缩短过渡时间,减少废品率和经济损失.总之,稳定模式和过渡模式都是多模式工艺不可缺少的组成部分,对多模式工艺进行在线运行最优性评估,对于进一步调整生产运行和提高工艺性能具有重要意义.

近年来, Ye等^[29]提出了一个连续多模式工业过程操作安全性和最优性评估的概率框架.通过在离线训练阶段采集稳定状态下的历史数据,利用EM算法构建了表征多种运行模式的高斯混合模型(Gaussian mixture model, GMM).在此基础上, Liu等^[30]针对具有过渡的多模态过程,提出了一种新的运行最优性评估和非最优原因识别方法.通过针对稳定模式和过渡模式不同的工艺特性,分别建立了运行最优性评估模型.然而,具有过渡模态的工艺过

程通常假设每一个稳定模态都遵循多元高斯分布,并且使用有限GMM来表征多个稳定模型.这无疑限制了整个工业过程某一个或多个模态下的数据服从非高斯分布情况的方法的应用,并且进一步考虑稳定模态下的非高斯运行评价,过渡模态下的运行状态评价也不应该被忽略.于是,针对多模态下的非高斯问题, Xu等^[31]提出了一种动态贝叶斯独立主成分分析(dynamic Bayesian independent component analysis DBICA)方法来弥补传统独立主成分分析只能处理单一运行模态和忽略过程数据动态性特征的问题.在DBICA方法中,通过动态增广矩阵提取原始数据的动态信息,然后利用贝叶斯推理和独立主成分分析来建立概率混合框架,进而可以处理多模态非高斯数据的工业过程运行监测问题.为此,本文归纳了基于高斯混合模型和非高斯模型的运行状态评价方法,如表2所示.为了进一步研究此类问题, Liu等^[32]提出了一种针对数据分布为非高斯的过程运行性能评价策略,不仅针对稳定模态,而且包括了过渡模态,进一步实现了带有过渡模态的非高斯工业过程运行状态最优性评价.为了分析现代工业过程中需要同时考虑综合经济效益和多种模式的过程监测问题, Liu等^[33]提出了一种基于预测CEI的多模态过程运行状态评价方法,首先针对稳定模态和过渡模态进行识别,然后结合CEI给出预测,最后对运行过程进行在线评价. Fan等^[34]提出了多模态概率可预测特征分析的复杂工业过程监控方法,通过结合深度自编码器(DAE)和GMM的策略,利用DAE提取低维特征以及识别概率模型的运行模式,并根据得到的特征将原始训练数据样本进行合理分类为相应的高斯分布,并在三相流装置中进行实验,得到了优越的异常情况监测性能.

表2 多模态工业过程运行状态评价方法

模型方法	划分依据	优点	缺点
高斯混合模型 ^[30-31]	根据各个模态下的数据特性均服从高斯分布	可以同时建立多种运行模式下的稳定模态和过渡模态的运行性能评估框架	训练数据特性受限,模型不稳定,容易陷入局部最优
非高斯模型 ^[32-33]	由于过程动态性质导致在过渡模态下呈现出非高斯特性	可以拟合非线性和非高斯分布的数据,更具有灵活性	模型复杂,要求训练数据量庞大,推理困难

以上从流程工业的多个方面阐述了多模态下所面临的问题以及对应的解决方法,同样在间歇工业过程生产中也同样面临类似的问题. Liu等^[35-36]分别针对批次过程的运行状态评价问题给出了相应的解决方法.首先,分多个阶段考虑整个批次过程的运行性能的联合作用,从阶段内和阶段间进行评估模型,分

析各个阶段的过程特征以及阶段间的相关性和相互作用对整个批次过程的运行性能的影响;然后,针对实际生产过程中存在动态特性以及长度不等的批次问题,提出了一种高斯过程回归以及贝叶斯推理的多阶段批次过程生产绩效评价方法,该方法能有效地处理长度不等的批次问题,并通过建立多个局部高斯回

归评价模型来表征所识别到的多个阶段的不同动态关系;最后,利用变量贡献图的方法对非优原因进行深入回溯。

2.2.3 定性与定量信息共存的运行状态评价

虽然上述方法已涵盖了工业过程中的大部分生产类型,但其存在一个共同的特点:以定量数据信息作为过程运行状态分析的基础。实际上,复杂工业过程存在一部分可在线测量的变量以及一部分不可在线测量的变量,在线测量的变量存在定量测量值,不可在线测量的变量以定性状态的形式表示。因此,定性与定量变量共存现象普遍存在。定量信息是指采用数值信息定量地描述变量信息,传统定量技术的评价方法主要通过建立变量间的相关性和准确的数学模型来对工业过程进行评价,其评价方法精度高、预测性能较好。定性信息是指利用客观描述性的语义对变量信息进行定性分析,其优点在于解释性强、容易建立模型并且可以处理一些不精确的信息。

随着计算机技术、智能设备技术和工业生产技术的高速发展,许多复杂工业过程可以在线得到大量的过程数据,例如压力、入料、流量等测量值。一些复杂的实际工业过程往往受到各种外界的影响,例如生产环境恶劣、测量值不准确等因素,从而造成一些常见的关键过程变量无法测量。传统单一的评价方法难以直接应用,于是针对当影响过程运行状态的关键变量无法在线获得或短期无法获得时,依赖过程数据的评价方法无法直接应用于实际现场等问题,衍生出一系列基于数据与知识相融合的运行状态评价方法,即定量与定性信息共存的评价方法,包括贝叶斯推断等基于定性描述的方法。首先,Zou等^[37]针对实际复杂工业过程中存在既有定性信息又有定量信息的问题,为了减少信息损失,避免计算复杂度的扩大,又可以最大程度地挖掘信息量,提出从两个角度出发,通过在线测量或者估计等数据驱动的形式提取信息,以及另一种通过建立变量之间的关联时间来推理相应的结果,因此,通过引入一种改进的动态因果图来尽可能地确保定量信息和定性信息变量的利用,从而提高运行性能的评价精度。然后,Zou等^[38]针对确定性和不确定性并存的复杂化工过程在线评估问题,开发了一种区分稳定性能等级和等级切换过渡的DCD方案,将多源信息的使用赋予了可解释性,并提高了改进DCD的准确性和鲁棒性。进而,Zou等^[39]针对在定量和定性共存、包含因果关系的过程,提出了一种基于改进的过程运行状态评价方法——将PCA与DCD相结合的评价方法。由于流程工业过程往往存在生

产周期长、规模庞大、变量众多、变量相关性复杂等特点,仅通过使用分块的多元统计和概率论的方法来解决流程工业的性能评价问题已不能满足实际流程工业生产的需求。邹筱瑜等^[40]针对流程工业进行性能评价过程中常见的问题,如未考虑定量、定性变量共存的问题,全流程的评价问题难以解决子块的评价问题,子块的等级难以定义等问题,提出了一种基于两层分块的混合模型,即利用GMM定量地获取子块内各个运行等级数据分布的概率密度函数,再通过概率粗糙集(probabilistic rough set)定性地得到子块内的运行状态等级的推理规则,从而建立了一个两层分块GMM-PRS(Gaussian mixture model-probabilistic rough set)模型,可以根据每个子块的特征选择合适的评价模型,提高有效信息的利用率,保证模型的评价准确度。上述方法均在国内大型实际工业过程中得到了应用,并验证了方法的有效性。

在实际的工业过程中,不仅存在一个状态下的信息共存现象,而且在每一个状态下都起着或多或少的的作用,所以无论是单一的稳定模态下,还是过渡等多模态下,都存在着定量和定性信息共存的现象,因此,针对此类问题研究评价方法也已经出现。为了解决含有多数定性变量和少数定量变量的单一模态的过程评价问题,充分利用定量和定性信息,Zou等^[41]提出了一种改进粗糙集方法,对于复杂工业过程中定性与定量信息共存的情况,分别针对单一模态和多模态给出了完整的运行状态评价方案,不仅提高了评价精度,也降低了判断结果对参数选取的依赖程度。针对定性和定量信息共存的多模态复杂工业过程,Chang等^[42]提出了一种先将生产过程的模式大致划分为包含初始生产条件且定性变量较多的模式指标变量,再利用包含大量定量信息的过程变量进行模式的细划分,进而使得多模态过程的运行状态可以被识别出相应的过程模式。

2.2.4 基于动态感知信息的运行状态评价

在实际生产过程中,由于操作条件的波动、控制器的调节、操纵人员的调整等,复杂工业过程一般表现出典型的动态特性。过程的时间动态表明了时变特性,反映了时间相关性。然而,时间动力学在经营绩效评估中很少被探索。动态信息可以反映工况的变化、控制行为引起的时间序列的自相关性等,对揭示过程的运行状态和性能水平具有重要意义。基于多元统计过程控制动态信息感知的分析方法在过程监测和故障诊断领域已得到广泛的应用,例如动态PCA^[43-44]、动态PLS等^[45-48]。然而,经典的动态潜变

量模型中,数据稳态值分布和变化速率分布的信息混杂在一起,无法清晰区分过程动态信息与稳态信息,因而对动态异常不敏感。

21世纪前期,人们在人类大脑的自组织机制的研究中提出了SFA算法^[46],SFA主要是通过利用一阶差分信息的方差来描述过程信号变化的快慢,并从时序数据中提取变化最缓慢的成分作为本质特征,因此称其为“慢特征(slow feature, SF)”。近几年,越来越多的研究人员利用SFA来处理各种复杂过程中的动态问题。2015年,Shang等^[49]提出了一种基于SFA的过程监测策略,可以同时用来监测操作点偏差和过程动力学异常两种状态,并且潜变量的缓慢特征可以描述缓慢变化的动态,进而区分运行条件的变化是正常的还是真实的故障发生,此方法有效降低了误报警率,大大提高了监测精度。Shang等^[50]基于SFA对特征 s 的稳态分布与 s 导数的动态分布分别进行了统计建模,利用稳态分布描述过程变量的静态特性即数据分布的中心,利用动态分布描述过程动态特性即数据分布的离散程度,赋予监控统计量不同的物理含义,提供更全面的监控信息。此外,SFA与现代控制理论存在着深刻联系,因而近年来在过程数据解析中得到了广泛应用^[51-52]。

当过程表现出典型的动态变化时,传统的性能评估方法可能无法描述过程的时变特性,因此不再适用于精细和敏感的过程性能建模和评估。为了评估运行性能,进行具有动态特征的运行性能评价是很有必要的。最近,Zou等^[53]提出了一种工业过程操作性能的静态和动态并行评估方法,结合静态和动态性能分析,深入掌握工艺特性,从而提供细致的操作性能评估和灵敏的性能水平识别。进一步,结合典型变量分析(canonical variate analysis, CVA)和慢特征分析,对过程的静态和动态特性进行了深入分析。因为同时使用了静态和动态信息,所以能够识别控制器的调节动作、性能切换、运行状态的变化等。进一步,褚菲等^[54]针对复杂工业过程中难以直接精确地获取过程动态信息,从而导致漏报、误报等问题,提出了一种基于关键性能指标驱动的SFA算法,对整个工艺过程进行运行状态评价。该方法能有效利用深层次特征信息的相似度和一阶差分信息的相似度,实现对工业过程稳定模态和过渡模态的评价。为了能更广泛地处理复杂工业过程面临的动态以及非线性问题,Wang等^[55]结合高斯过程回归模型(Gaussian process regression, GPR)提出了(slow feature analysis and Gaussian process regression, SFA-GPR)运行状态

评价方法,该方法不仅能提取工业过程潜在驱动力特征,也可以捕获非线性和动态特征信息,由于GPR具有显著的易操作、自适应超参数获取以及概率输出等优势,使得工业过程运行状态评价的精度得到有效提高。尽管SFA在过程监控领域展示了其良好的性能和有意义的数据模型,但运行状态评价更希望有效地获取状态的潜在变化,以便更准确地表达运行状态。

事实上,在实际的工业过程中,采集到的数据往往具有一定的动态特性,即各个变量之间除了互相关性之外,还表现或多或少的自相关性。针对过程数据的动态特性,传统的多元统计过程监测方法虽然考虑了当前时刻与之前若干个时刻对应数据的自相关性,但容易造成计算复杂度的急剧上升和不能有效地描述过程的动态特性。因此,基于子空间辨识的过程动态信息感知技术应运而生。子空间辨识技术通过利用数据的低维结构,将高维数据转化为小型子空间的形式,从而避免了大维空间的计算难题,同时还可以通过对数据的子空间分析,挖掘数据背后隐藏的知识和规律。为此,何雨辰等^[56]提出了一种基于多个子空间划分的慢特征分析模型的动态大规模过程集成监控策略,通过变换矩阵揭示过程数据与慢特征之间的关系,并构造每个子空间中的 S^2 统计量,实现对过程集成监控。Guo等^[57]将慢特征分析扩展到概率图模型,以线性高斯状态空间的形式建立概率慢特征分析(probabilistic slow feature analysis, PSFA),其中慢特征作为状态量随自回归模型演化,因而PSFA提供了状态变化潜在驱动力的明确表示,极大减少了过程数据中不利因素的影响。

随着运行状态技术的不断提高,基于时间序列的动态信息感知方法也逐渐被开发,由于运行状态性能的变化,可以通过挖掘系统的动态变化和时间序列的自相关等信息来揭示。全流程工业由于反馈控制系统作用,使得过程观测值具有显著的时序相关性,除表现出典型的多维时间序列特点外,过程运行本身也存在明显的动态变化^[57]。然而,典型的静态分析模型,如PCA、T-PLS等,都是基于工业过程在稳定状态下运行的假设,只是简单对数据的稳态值分布进行分析,而对过程的动态变化信息不敏感。为此,王佩^[58]提出了一种基于孪生网络的运行状态评价方法,该方法将孪生网络的子网络结构特定为长短期记忆(long short-term memory, LSTM)网络,以增强模型对工业过程动态变化的感知能力。

2.3 基于神经网络的运行状态评价

建模任务一般可分为无监督学习和有监督学习两类. 在无监督学习中, 建立描述模型来描述输入数据中的底层结构, 主要用于监控过程数据的分布. 在监督学习中, 包括回归和分类, 建立了输入与输出之间的函数映射, 特别关注输出的预测精度. 近年来, 越来越多的研究集中在表征学习或特征学习上^[59], 其中在建立模型时结合领域特定知识尤为重要. 因此, 模型的可解释性得到显著增强, 这进一步提高了模型的性能. 表征学习是通过将原始数据转化为机器学习所能处理的一种数据类型, 其主要目标是将复杂的数据类型进行简化, 然后提取更加精准的数据表达, 使得后期的数据模型的建立更有助于提高模型性能^[60]. 对于复杂工业过程而言, 数据驱动的好处就是能够快速建立过程模型, 无需对过程工艺进行充分了解.

随着深度学习在机器学习领域的卓越表现, 让人们看到了深度学习强大的学习能力、覆盖范围广、可迁移性强等优势, 使得神经网络日趋成为解决工业过程中非线性问题的研究热点. 自编码器(autoencoder, AE)的应用也逐渐扩展到各个领域, 其中, 褚菲等^[61]针对工业过程中与过程变量相关的CEI难以提取的问题, 通过将稀疏学习和降噪处理引入AE模型中, 提出了一种具有稳健可靠的工业运行状态评价模型, 同时考虑到变量过程之间的强非线性以及综合经济指标问题, 引入基于综合经济指标驱动的稀疏学习AE模型, 将特征提取与评价模型相结合, 使得所建立的网络模型结构不仅可以挖掘原始数据的特征, 而且还能反向调节运行状态评价模型. Bu等^[62]采用一种基于堆叠式稀疏降噪自编码器(stacked sparse denoising autoencoder, SSDA)的运行状态评价方法, 通过利用降噪AE模型增强鲁棒性、堆叠自编码器(stacked autoencoder, SAE)模型学习过程数据的稀疏表达, 并结合softmax分类器实现了对复杂工艺过程运行状态的评价. 进一步, 在面向多源信息的工业过程进行过程评价时, 单一的特征提取工具已不适应于所有的来源数据信息. 对此, Bu等^[63]提出了一种多源异构特征信息融合的运行状态评价方法, 针对来自不同数据源的信息进行处理, 使得各个信息源可以转换成同一种张量的信息. 首先, 同时对图像采用卷积神经网络进行信息的特征提取; 然后, 利用SAE提取电流的特征信息并采用双向长短期记忆(bi long short-term memory, Bi-LSTM)神经网络对具有时间序列的声音信息进行提取; 最后, 采用融合

机制对多源信息进行特征融合, 进而通过评价模型进行状态评价. 基于神经网络建模的运行状态评价可以避免使用过多的先验知识和过程数据等, 其评价结果精度高、可解释性强. 因此, 有效地开发基于神经网络知识, 例如, 基于图卷积神经网络、长短期记忆网络^[64-65]等深度神经网络的复杂工业过程运行状态评价方法, 可以实现大规模数据以及数据空间的拓扑结构等深层次信息挖掘.

2.4 全流程分层分级运行状态评价

复杂工业过程又称为流程工业, 在实际生产过程中往往面临着一个很现实的问题: 生产流程长, 生产规模庞大以及过程变量复杂. 复杂工业生产过程, 通常根据变量间的耦合关系将整个流程工业分割成若干个生产单元. 因此, 传统的过程工业评价方法往往不能准确地处理此类问题. 进而, 一类针对流程工业特性的评价方法, 例如多层级、分子块和跨流程等, 通过物理特性将过程划分为不同层次和子块. 一个流程工业过程的层次和子块的划分^[66-67]如图2所示.

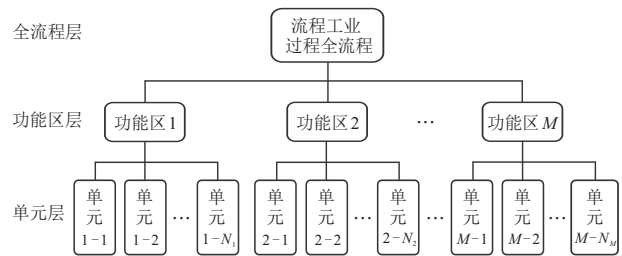


图2 全流程分层分级流程图

Liu等^[68]提出了一种新的基于潜在结构分层多块全投影的全流程运行状态评价方法, 通过将过程变量划分为几个工艺块, 并将块信息整合到全局范围内过程运行状态评价中, 大大减少了在线评价工作量, 提高了评价模型的可解释性, 并能快速定位非优的工艺块以缩小搜索范围, 提高非优原因识别结果的准确性. Macgregor等^[69]和Jiang等^[70]分别提出了分块的多元统计方法和分块的概率论方法, 用来处理流程工业过程性能评价问题. 但是, 流程工业往往面临着存在不确定性信息的风险, 对此, 邹筱瑜等^[40,71]提出了一种分层分块结构的过程运行状态评价方法, 通过引入粗糙集对每一个子块进行建模, 然后根据不同子块数据特征运用不同的模型. 这种方法不仅考虑了工业过程中存在不确定性的信息, 而且还从某一方面简化了模型规模, 提高了在非优状态下的非优原因定位的精确度和模型的解释性. 在不同的子块内, 由于存在规则不一致等问题, 导致无法直接应用统一的模型进行信息处理. Chang等^[72-73]和Zou等^[74]进一步提

出了概率粗糙集的GMM运行状态评价模型,通过利用多分块级联,分层分阶段地获取局部和全局信息的方式来构建整个长流程工艺环境,并且针对多模态问题、定性定量信息共存问题、不确定性等问题,分别结合高斯模型、概率粗糙集对全流程工业过程建立评价指标,构建混合评价模型,使得模型可以通过综合考虑不同工艺特性减少信息损失,提高模型精度,进而对整个运行工艺进行有效评估.近几年,这种分层分块的工业过程方法也逐渐被应用到其他领域中,例如过程工业的过程监测问题^[75-76].

2.5 非优因素回溯

在复杂工业过程中,非优因素回溯对于运行状态的评价具有重要意义.首先,对于复杂工业过程而言,过程运行状态的优劣是根据评价指标的大小来确定的,因此,非优原因回溯过程可以定位那些导致评价指标低于阈值的原因变量,而这些原因可能涉及到多种输入因素、控制参数和输出变量,有许多原因可能导致异常状态的出现,如设备故障、操作不当、原材料质量等.在这种情况下,如果只是解决表面上的问题,则问题可能没有从根本上解决,且问题可能会反复出现.非优因素回溯对于深入了解事件的根本原因非常有帮助,可以帮助操作人员找到真正的缺陷原因而不是解决表面问题.其次,非优因素回溯可以为复杂工业过程的持续改进提供指导.如果一个问题反复发生,则通过回溯以往的历史记录可以发现问题产生的根源,从而采取相应的改进措施.这些措施可能包括修改特定的操作步骤、改变运行参数、更新设备和维护等,从而有效地控制此类问题的发生.最后,通过运行优性评估对过程非优因素进行追溯,现场操作人员和管理者可以实时掌握过程运行绩效,对实际生产过程的生产进行调整以及对性能改进,从而更好地做好预防工作,使得整个过程更加高效地运行.

很早就有学者通过贡献图的方法对统计过程监测进行非优原因回溯,如图3所示^[77-79].目前,这种非优原因追溯方法也被大量应用于运行状态评价.基于贡献图的非优原因追溯方法是根据过程运行状态非优时,针对在线数据对于状态等级“优”的评价指标,构造每个过程变量对其的贡献,并将具有较大贡献的过程变量确定为导致过程运行状态非优的原因变量.由于复杂工业过程中的变量数量庞大、变量关系复杂和强大的耦合关系,使用变量贡献图的方式往往会因局部异常覆盖全局而导致整个工艺过程出现严重的涂抹效应,从而使得非关键变量产生不可忽视的贡献,进而对非优原因追溯工作产生扰乱.

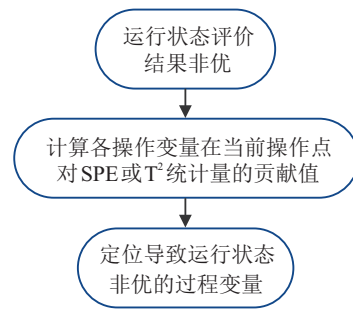


图3 基于贡献图的非优因素回溯流程

近年来,一些基于重构贡献图方法开始得到应用. Liu等^[80]为了确定故障区块位置变量,首次提出一种新颖的基于重构的区块贡献,并且提出了区块内的变量贡献图算法来确定故障变量.重构贡献图算法主要是沿着某一个可变方向的重构贡献率,从某一程度上减少变量贡献图的涂抹效应.但是当特征维数只有一维时,这种方法就失效了.为了解决上述问题,基于稀疏学习的方法进入了人们的视野,其原理是过程数据稀疏性往往存在于过程状态恶化的方向.与传统的贡献图方法相比,基于稀疏学习的方法可以有效抑制噪声,减轻无关变量的影响,具有清晰、明确、可解释性强的特点. Shang等^[81]通过结合正则化技术将稀疏正则化的方法应用于故障诊断问题,从而改进了传统的稀疏学习只能针对单一变量进行评估的弊端.但是,运行状态评价过程中往往采用滑动窗口来记录当前和过去数据,导致Lasso技术不能直接应用于运行评价.褚菲等^[54]将组套索Group Lasso技术应用于复杂工业过程的运行状态评价的非优原因追溯问题,以避免使用变量贡献图进行非优原因追溯所带来的涂抹效应.

3 总结与展望

本文对复杂工业过程运行状态评价的背景意义和研究现状进行了较为全面的综述和梳理,将现有的研究成果划分为基于多元统计分析的运行状态评价和基于神经网络的运行状态评价以及非优原因追溯.进一步总结了数据驱动运行状态评价的相关问题,例如:变量之间存在线性或非线性相关关系、多模态问题、定量与定性信息共存问题、动态信息感知的评价问题、全流程分层分级运行状态评价以及非优因素回溯等问题.

近年来,随着计算机的高速发展以及相关理论的不不断改进,图神经网络因其强大的对复杂网络的处理能力吸引了大量的学者来研究,并在工业过程控制中得到了广泛的应用.由于数据之间错综复杂的关系网络,图神经网络结构已变得无处不在.从这些数据

(称为图的节点)中提取相关信息,对于有效处理机器学习任务至关重要,如链接预测、节点聚类等.节点之间的特征关系可以用节点或网络连接边表示,节点间关系的强弱可以使用权值或二进制的方式表示.它们通过迭代传播邻居信息学习目标节点的表示,直到达到一个稳定的不动点.这个过程的成本很高,最近已经有越来越多的研究在克服这些挑战,特别是图卷积神经网络已受到了广泛的关注,但是图卷积神经网络性能需要与先验知识紧密连接.特别地,由于图神经网络的无监督学习的优势,有力地推进了图卷积神经网络的发展与应用.基于编码-解码的设计方案,有效地将原始数据映射到低维特征空间并保证原始数据空间结构与特征数据空间结构的相似性最大.图神经网络结构的应用也将在工业过程中发挥积极作用.面向实际工业过程数据非线性特性显著和图结构的运行状态评价模型的解释性好等方向,提出下一步的研究方向和目标.

1) 针对复杂工业过程的运行数据维度高、非线性特性明显等问题,为了能更好地提取过程非线性特性,在本文提到的特征提取方法中,往往采用引入核函数或局部嵌套等方法.但是,无论对数据进行升维还是局部线性化等方法,都势必破坏数据原本的动态特性,因此,如何有效提取耦合非线性数据中的动态特性仍是一个需要研究的方向.

2) 随着工业发展,对过程运行状态的认知会越来越精准,对最优状态的要求也会越来越高,各个状态等价之间差异越来越小,此时对于单个状态很难获得足够的用于离线建模,进而无法建立高效的评价模型.事实上,目前实际工业现场的运行状态评价很大程度上依赖于专家的经验知识,因此,如何将专家经验知识有效融入到数据驱动的运行状态评价模型中,探索人机交互式的运行状态评价方法,是一个具有理论与实际应用价值的研究方向.

参考文献(References)

- [1] 柴天佑. 制造流程智能化对人工智能的挑战[J]. 中国科学基金, 2018, 32(3): 251-256.
(Chai T Y. Artificial intelligence research challenges in intelligent manufacturing processes[J]. Bulletin of National Natural Science Foundation of China, 2018, 32(3): 251-256.)
- [2] 柴天佑. 工业过程控制系统研究现状与发展方向[J]. 中国科学: 信息科学, 2016, 46(8): 1003-1015.
(Chai T Y. Research status and development direction of industrial process control system[J]. Scientia Sinica: Informationis, 2016, 46(8): 1003-1015.)
- [3] 陈龙, 刘全利, 王霖青, 等. 基于数据的流程工业生产
过程指标预测方法综述[J]. 自动化学报, 2017, 43(6): 944-954.
(Chen L, Liu Q L, Wang L Q, et al. Data-driven prediction on performance indicators in process industry: A survey[J]. Acta Automatica Sinica, 2017, 43(6): 944-954.)
- [4] 桂卫华, 岳伟超, 谢永芳, 等. 铝电解生产智能优化制造研究综述[J]. 自动化学报, 2018, 44(11): 1957-1970.
(Gui W H, Yue W C, Xie Y F, et al. A review of intelligent optimal manufacturing for aluminum reduction production[J]. Acta Automatica Sinica, 2018, 44(11): 1957-1970.)
- [5] 钱锋, 桂卫华. 人工智能助力制造业优化升级[J]. 中国科学基金, 2018, 32(3): 257-261.
(Qian F, Gui W H. Boosting optimization and upgrade for manufacturing industry by artificial intelligence[J]. Bulletin of National Natural Science Foundation of China, 2018, 32(3): 257-261.)
- [6] 郭方洪, 林凯, 窦云飞, 等. 基于多维特征评价的风机齿轮箱早期故障诊断[J]. 控制与决策, DOI: 10.13195/j.kzyjc.2022.1388.
(Guo F H, Lin K, Dou Y F, et al. Incipient fault diagnosis for wind turbine gearbox based on multidimensional feature evaluation[J]. Control and Decision, DOI: 10.13195/j.kzyjc.2022.1388.)
- [7] Shang C, Chen W H, Stroock A D, et al. Robust model predictive control of irrigation systems with active uncertainty learning and data analytics[J]. IEEE Transactions on Control Systems Technology, 2020, 28(4): 1493-1504.
- [8] 赵春晖, 余万科, 高福荣. 非平稳间歇过程数据解析与状态监控——回顾与展望[J]. 自动化学报, 2020, 46(10): 2072-2091.
(Zhao C H, Yu W K, Gao F R. Data analytics and condition monitoring methods for nonstationary batch processes—Current status and future[J]. Acta Automatica Sinica, 2020, 46(10): 2072-2091.)
- [9] Xu S B. Analytic hierarchy process principle[M]. Tianjin: Tianjin University Press, 1988.
- [10] 叶立新. 模糊综合评价在会计人员素质评价中的应用[J]. 辽宁工程技术大学学报: 社会科学版, 2006, 8(5): 499-501.
(Ye L X. Application of comprehensive fuzzy evaluation in the quality of accounting personnel[J]. Journal of Liaoning Technical University: Social Science Edition, 2006, 8(5): 499-501.)
- [11] 张莉. 综合国力评价初探[J]. 统计与决策, 2002(5): 9-10.
(Zhang L. Preliminary study on comprehensive national strength evaluation[J]. Statistics and Decision, 2002(5): 9-10.)
- [12] 王福利, 常玉清, 刘炎, 等. 复杂工业过程运行状态优性评价[M]. 北京: 科学出版社, 2019.
(Wang F L, Chang Y Q, Liu Y, et al. Operation

- performance optimality assessment for complex industrial process[M]. Beijing: Science Press, 2019.)
- [13] 刘炎. 工业过程运行状态最优性评价及非优原因追溯方法的研究[D]. 沈阳: 东北大学, 2016.
(Liu Y. Study on the optimal evaluation of industrial process operation state and the tracing method of non-optimal reasons[D]. Shenyang: Northeastern University, 2016.)
- [14] Dunteman G H. Principal component analysis[M]. London: SAGE Publication, 1989.
- [15] Wold S. Nonlinear partial least squares modelling. II. Spline inner relation[J]. Chemometrics and Intelligent Laboratory Systems, 1992, 14(1): 71-84.
- [16] Wiskott L, Sejnowski T J. Slow feature analysis: Unsupervised learning of invariances[J]. Neural Computation, 2002, 14(4): 715-770.
- [17] Liu Y, Chang Y Q, Wang F L. Online process operating performance assessment and nonoptimal cause identification for industrial processes[J]. Journal of Process Control, 2014, 24(10): 1548-1555.
- [18] Liu Y, Wang F L, Chang Y Q. Operating optimality assessment based on optimality related variations and nonoptimal cause identification for industrial processes[J]. Journal of Process Control, 2016, 39: 11-20.
- [19] Liu Y, Chang Y Q, Wang F L, et al. Complex process operating optimality assessment and nonoptimal cause identification using modified total kernel PLS[C]. The 26th Chinese Control and Decision Conference. Changsha, 2014: 1221-1227.
- [20] Zhang Y W, Teng Y D, Zhang Y. Complex process quality prediction using modified kernel partial least squares[J]. Chemical Engineering Science, 2010, 65(6): 2153-2158.
- [21] Liu Y, Wang F L, Chang Y Q. Industrial process operating optimality assessment based on Gaussian mixture model[C]. The 29th Chinese Control and Decision Conference. Chongqing, 2017: 713-718.
- [22] Cho J H, Lee J M, Choi S W, et al. Fault identification for process monitoring using kernel principal component analysis[J]. Chemical Engineering Science, 2005, 60(1): 279-288.
- [23] Zhang Y W. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM[J]. Chemical Engineering Science, 2009, 64(5): 801-811.
- [24] 孔祥玉, 杨治艳, 刘佑民, 等. 基于独立成分分析及其扩展模型的工业过程监测方法综述[J]. 控制与决策, 2022, 37(4): 799-814.
(Kong X Y, Yang Z Y, Liu Y M, et al. Overview of industrial process monitoring methods based on independent component analysis and its extended model[J]. Control and Decision, 2022, 37(4): 799-814.)
- [25] Lee J M, Qin S J, Lee I B. Fault detection and diagnosis based on modified independent component analysis[J]. AIChE Journal, 2006, 52(10): 3501-3514.
- [26] Chu F, Mo S S, Hao L L, et al. Operating performance assessment of complex nonlinear industrial process based on kernel locally linear embedding PLS[J]. IEEE Transactions on Automation Science and Engineering, 2022(99): 1-13.
- [27] 褚菲, 赵旭, 代伟, 等. 数据驱动的最优运行状态鲁棒评价方法及应用[J]. 自动化学报, 2020, 46(3): 439-450.
(Chu F, Zhao X, Dai W, et al. Data-driven robust evaluation method for optimal operating status and its application[J]. Acta Automatica Sinica, 2020, 46(3): 439-450.)
- [28] Chu F, Dai W, Shen J, et al. Online complex nonlinear industrial process operating optimality assessment using modified robust total kernel partial M -regression[J]. Chinese Journal of Chemical Engineering, 2018, 26(4): 775-785.
- [29] Ye L B, Liu Y M, Fei Z S, et al. Online probabilistic assessment of operating performance based on safety and optimality indices for multimode industrial processes[J]. Industrial & Engineering Chemistry Research, 2009, 48(24): 10912-10923.
- [30] Liu Y, Wang F L, Chang Y Q. Operating optimality assessment and nonoptimal cause identification for multimode industrial process with transitions[J]. The Canadian Journal of Chemical Engineering, 2016, 94(7): 1342-1353.
- [31] Xu Y, Deng X G. Fault detection of multimode non-Gaussian dynamic process using dynamic Bayesian independent component analysis[J]. Neurocomputing, 2016, 200: 70-79.
- [32] Liu Y, Wang F L, Chang Y Q, et al. Operating optimality assessment and nonoptimal cause identification for non-Gaussian multimode processes with transitions[J]. Chemical Engineering Science, 2015, 137: 106-118.
- [33] Liu Y, Wang F L, Chang Y Q, et al. Comprehensive economic index prediction based operating optimality assessment and nonoptimal cause identification for multimode processes[J]. Chemical Engineering Research and Design, 2015, 97: 77-90.
- [34] Fan W, Ren S J, Yu C, et al. A mixture of probabilistic predictable feature analysis for multi-mode dynamic process monitoring[J]. Journal of the Taiwan Institute of Chemical Engineers, 2023, 143: 104635.
- [35] Liu Y, Ma R C, Wang F L, et al. Inner-phase and inter-phase analysis based operating performance assessment and nonoptimal cause identification for multiphase batch processes[J]. Chemical Engineering Research and Design, 2018, 134: 292-308.
- [36] Liu Y, Wang X J, Wang F L, et al. Gaussian process regression and Bayesian inference based operating performance assessment for multiphase batch processes[J]. Industrial & Engineering Chemistry

- Research, 2018, 57(21): 7232-7244.
- [37] Zou X Y, Chang Y Q, Wang F L, et al. Process operating performance optimality assessment with coexistence of quantitative and qualitative information[J]. The Canadian Journal of Chemical Engineering, 2018, 96(1): 179-188.
- [38] Zou X Y, Wang F L, Chang Y Q, et al. Process operating performance optimality assessment and non-optimal cause identification under uncertainties[J]. Chemical Engineering Research and Design, 2017, 120: 348-359.
- [39] Zou X Y, Wang F L, Chang Y Q, et al. Process operation performance optimality assessment and cause identification based on PCA-DCD[C]. Proceeding of the 28th Chinese Control and Decision Conference. Kunming, 2016: 43-48.
- [40] 邹筱瑜, 王福利, 常玉清, 等. 基于分层分块结构的流程工业过程运行状态评价及非优原因追溯[J]. 自动化学报, 2019, 45(2): 315-324.
(Zou X Y, Wang F L, Chang Y Q, et al. Plant-wide process operating performance assessment and non-optimal cause identification based on hierarchical multi-block structure[J]. Acta Automatica Sinica, 2019, 45(2): 315-324.)
- [41] Zou X Y, Wang F L, Chang Y Q, et al. Two-level multi-block operating performance optimality assessment for plant-wide processes[J]. The Canadian Journal of Chemical Engineering, 2018, 96(11): 2395-2407.
- [42] Chang Y Q, Ma R X, Wang F L, et al. Multimode process mode identification with coexistence of quantitative information and qualitative information[J]. IEEE Transactions on Automation Science and Engineering, 2020, 17(3): 1516-1527.
- [43] Dong Y N, Qin S J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring[J]. Journal of Process Control, 2018, 67: 1-11.
- [44] Dong Y N, Qin S J. Dynamic latent variable analytics for process operations and control[J]. Computers & Chemical Engineering, 2018, 114: 69-80.
- [45] Chiplunkar R, Huang B. Output relevant slow feature extraction using partial least squares[J]. Chemometrics and Intelligent Laboratory Systems, 2019, 191: 148-157.
- [46] Liu Q, Qin S J, Chai T Y. Unevenly sampled dynamic data modeling and monitoring with an industrial application[J]. IEEE Transactions on Industrial Informatics, 2017, 13(5): 2203-2213.
- [47] Zhao C H, Sun H. Dynamic distributed monitoring strategy for large-scale nonstationary processes subject to frequently varying conditions under closed-loop control[J]. IEEE Transactions on Industrial Electronics, 2019, 66(6): 4749-4758.
- [48] Zhao S Y, Ma Y J, Huang B. Robust FIR state estimation of dynamic processes corrupted by outliers[J]. IEEE Transactions on Industrial Informatics, 2019, 15(1): 139-147.
- [49] Shang C, Yang F, Gao X, et al. Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis[J]. AIChE Journal, 2015, 61(11): 3666-3682.
- [50] Shang C, Huang B, Yang F, et al. Slow feature analysis for monitoring and diagnosis of control performance[J]. Journal of Process Control, 2016, 39: 21-34.
- [51] Zhang S M, Zhao C H. Slow-feature-analysis-based batch process monitoring with comprehensive interpretation of operation condition deviation and dynamic anomaly[J]. IEEE Transactions on Industrial Electronics, 2019, 66(5): 3773-3783.
- [52] Ma Y J, Huang B. Bayesian learning for dynamic feature extraction with application in soft sensing[J]. IEEE Transactions on Industrial Electronics, 2017, 64(9): 7171-7180.
- [53] Zou X Y, Zhao C H. Concurrent assessment of process operating performance with joint static and dynamic analysis[J]. IEEE Transactions on Industrial Informatics, 2020, 16(4): 2776-2786.
- [54] 褚菲, 许杨, 尚超, 等. 基于静-动态特性协同感知的复杂工业过程运行状态评价[J]. 自动化学报, 2023, 49(8): 1621-1634.
(Chu F, Xu Y, Shang C, et al. Evaluation of complex industrial process operating state based on static-dynamic cooperative perception[J]. Acta Automatica Sinica, 2023, 49(8): 1621-1634.)
- [55] Wang H D, Wang X, Wang Z L. Performance assessment method of dynamic process based on SFA-GPR[J]. Journal of Process Control, 2022, 111: 27-34.
- [56] 何雨辰, 王云, 刘涛, 等. 基于公共特有子空间提取的工业设备多模式运行过程故障检测方法[J]. 控制与决策, 2022, 37(6): 1469-1478.
(He Y C, Wang Y, Liu T, et al. Multimode process monitoring for industrial equipments utilizing common-specific information extraction strategy[J]. Control and Decision, 2022, 37(6): 1469-1478.)
- [57] Guo F H, Shang C, Huang B, et al. Monitoring of operating point and process dynamics via probabilistic slow feature analysis[J]. Chemometrics and Intelligent Laboratory Systems, 2016, 151: 115-125.
- [58] 王佩. 煤泥浮选过程模型仿真及其运行状态评价[D]. 徐州: 中国矿业大学, 2022.
(Wang P. Model simulation of slime flotation process and its operation state evaluation[D]. Xuzhou: China University of Mining and Technology, 2022.)
- [59] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [60] Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]. Proceedings of the 27th International Conference on Machine Learning (ICML-10). Haifa, 2010: 807-814.

- [61] 褚菲, 傅逸灵, 赵旭, 等. 基于ISDAE模型的复杂工业过程运行状态评价方法及应用[J]. 自动化学报, 2021, 47(4): 849-863.
(Chu F, Fu Y L, Zhao X, et al. Operating performance assessment method and application for complex industrial process based on ISDAE model[J]. Acta Automatica Sinica, 2021, 47(4): 849-863.)
- [62] Bu K Q, Wang F L, Liu Y. Operating performance assessment based on stacked sparse denoising auto-encoder for gold hydrometallurgy processes[C]. Proceeding of the 32th Chinese Control and Decision Conference. Hefei, 2020: 2904-2909.
- [63] Bu K Q, Liu Y, Wang F L. Operating performance assessment based on multi-source heterogeneous information with deep learning for smelting process of electro-fused magnesium furnace[J]. ISA Transactions, 2022, 128: 357-371.
- [64] Yuan X F, Li L, Wang Y L, et al. Deep learning for quality prediction of nonlinear dynamic processes with variable attention-based long short-term memory network[J]. The Canadian Journal of Chemical Engineering, 2020, 98(6): 1377-1389.
- [65] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [66] Liu Q, Qin S J, Chai T Y. Multiblock concurrent PLS for decentralized monitoring of continuous annealing processes[J]. IEEE Transactions on Industrial Electronics, 2014, 61(11): 6429-6437.
- [67] Deng X G, Wang L. Multimode process fault detection method using local neighborhood standardization based multi-block principal component analysis[C]. The 29th Chinese Control and Decision Conference. Chongqing, 2017: 5615-5621.
- [68] Liu Y, Wang F L, Gao F R, et al. Hierarchical multiblock T-PLS based operating performance assessment for plant-wide processes[J]. Industrial & Engineering Chemistry Research, 2018, 57(43): 14617-14627.
- [69] MacGregor J F, Jaeckle C, Kiparissides C, et al. Process monitoring and diagnosis by multiblock PLS methods[J]. AIChE Journal, 1994, 40(5): 826-838.
- [70] Jiang Q C, Yan X F. Monitoring multi-mode plant-wide processes by using mutual information-based multi-block PCA, joint probability, and Bayesian inference[J]. Chemometrics and Intelligent Laboratory Systems, 2014, 136: 121-137.
- [71] 邹筱瑜, 王福利, 常玉清, 等. 基于两层分块GMM-PRS的流程工业过程运行状态评价[J]. 自动化学报, 2019, 45(11): 2071-2081.
(Zou X Y, Wang F L, Chang Y Q, et al. Plant-wide process operating performance assessment based on two-level multi-block GMM-PRS[J]. Acta Automatica Sinica, 2019, 45(11): 2071-2081.)
- [72] Chang Y, Ma R, Zhao L, et al. Online operating performance evaluation for the plant-wide industrial process based on a three-level and multi-block method[J]. The Canadian Journal of Chemical Engineering, 2019, 97: 1371-1385.
- [73] Chang Y Q, Zou X Y, Wang F L, et al. Multi-mode plant-wide process operating performance assessment based on a novel two-level multi-block hybrid model[J]. Chemical Engineering Research and Design, 2018, 136: 721-733.
- [74] Zou X Y, Wang F L, Chang Y Q, et al. Two-level multi-block operating performance optimality assessment for plant-wide processes[J]. The Canadian Journal of Chemical Engineering, 2018, 96(11): 2395-2407.
- [75] Ren H, Chen Z W, Jiang Z H, et al. An industrial multilevel knowledge graph-based local-global monitoring for plant-wide processes[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-15.
- [76] Tian Y, Yao H, Li Z Q. Plant-wide process monitoring by using weighted copula-correlation based multiblock principal component analysis approach and online-horizon Bayesian method[J]. ISA Transactions, 2020, 96: 24-36.
- [77] Kourti T, MacGregor J F. Process analysis, monitoring and diagnosis, using multivariate projection methods[J]. Chemometrics and Intelligent Laboratory Systems, 1995, 28(1): 3-21.
- [78] Alcalá C F, Qin S J. Reconstruction-based contribution for process monitoring[J]. Automatica, 2009, 45(7): 1593-1600.
- [79] Ji H Q, He X, Zhou D H. On the use of reconstruction-based contribution for fault diagnosis[J]. Journal of Process Control, 2016, 40: 24-34.
- [80] Liu Q, Chai T Y, Qin S J. Fault diagnosis of continuous annealing processes using a reconstruction-based method[J]. Control Engineering Practice, 2012, 20(5): 511-518.
- [81] Shang C, Ji H Q, Huang X L, et al. Generalized grouped contributions for hierarchical fault diagnosis with group Lasso[J]. Control Engineering Practice, 2019, 93: 104193.

作者简介

褚菲(1984—), 男, 教授, 博士生导师, 从事复杂工业过程建模、控制与优化、机器学习及运行状态评价等研究, E-mail: chufei@cumt.edu.cn;

郝莉莉(1994—), 女, 博士生, 从事复杂网络协同控制和复杂工业过程运行状态评估的研究, E-mail: haolihaoli@cumt.edu.cn;

王福利(1957—), 男, 教授, 博士生导师, 从事复杂工业系统的建模、控制与优化、过程监测和故障诊断等研究, E-mail: wangfuli@ise.neu.edu.cn.