

基于差分分段 PCA 的多模态过程故障监测

谭帅¹ 王福利^{1,2} 常玉清^{1,2} 王姝¹ 周贺³

摘要 多模态的故障监测是一个复杂的问题,既需要考虑稳定模态下的故障监测,也需要考虑不同模态间的过渡故障监测。不同稳定模态下的数据具有不同的相关关系,对每个稳定模态需要建立不同的稳定模态模型。当稳定生产模态发生改变时,生产过程进入过渡模态,需要考虑过渡变量相关关系的变化。本文通过对过渡数据差分,得到变量相对变化信息。用 PCA 分段对差分变量的相关特性进行分析,提取相对变化的特征。最后以实际连续退火机组生产线为背景,用基于差分分段 PCA 的多模态方法对多模态过程进行故障监测,发现算法很好的反映了实际过渡过程机理,验证了算法的有效性。

关键词 差分矩阵,分段建模,多模态,故障监测,连续退火机组

DOI 10.3724/SP.J.1004.2008.xxxxx

Fault Detection of Multi-mode Process Using Segmented PCA Based on Differential Transform

TAN Shuai¹ WANG Fu-Li^{1,2} CHANG Yu-Qing^{1,2} WANG Shu¹ ZHOU He³

Abstract Fault detection for multi-model is a complicated problem, as the fault detection for both steady mode and transition mode should be taken into consideration. Different modes are need for different steady modes because different relations of variables are contained in each mode model. Transition mode is a dynamic process occurring when production changes operate mode. The dynamic characteristic reflects on not only the changing variables but also the changing relation of variables. The relative change information can be obtained by differential transform of transition data. Principle components can be extracted by analysing correlation of differential variables using PCA. At last, segmented modeling with PCA method is applied to monitor multi-mode process of continuous annealing line. The algorithm reflects transition process well and is proved to be efficient.

Key words difference matrix, segmented modeling, multi-mode, fault detection, continuous annealing line

统计过程监控 (statistical process monitoring, SPM) 是一种基于多元统计理论的过程控制技术。在实际应用中,由于大量测量变量的存在以及这些变量之间的高度相关性,需要从多变量的角度进行分析,从而形成了多元统计过程监控 (multivariate SPM, MSPM) 方法^[1]。二十世纪 80 年代末,以主成分分析 (Principal Component Analysis, PCA) 和偏最小二乘法 (Partial Least Squares, PLS) 等多变量投影技术为核心的多变量统计建模方法开始成为基于过程数据的统计过程监控的研究热点。因为 PCA 和 PLS 只需要正常工况下的历史数据来建立模型而不需要过程的机理知识,

同时 PCA 和 PLS 能够有效地剔除过程数据中的冗余信息、极大地降低数据维数,甚至可以将过程运行状态直接显示于二维的主成分监视图中,这类方法越来越受到研究人员和现场工程师的亲睐。目前,基于这类多变量统计模型的过程监测、故障诊断、控制器设计、质量控制等算法层出不穷,并且已成功地应用到各种连续生产过程中。

但是当前的 MSPM 技术往往假设过程本身仅存在一个标准操作范围,即假设变量间的相关关系具有相同的过程特征,只有这样,传统的 MSPM 才能基于历史数据准确的反映出过程特性。而实际工业中,生产过程往往不只有一个稳定工况,大部分生产面临的过程是多模态过程,即生产过程具有多个稳定工作点,并且不同稳定工作点之间变量的相关关系具有不同的特性。造成多模态工况的原因分很多方面:可能是原料性质、外界环境、过程负荷等条件的变化及设备磨损等因素,导致过程的操作条件发生变化,过程的正常工况会发生改变;或是过程本身生产方案变动,导致出现多个稳态操作点;还有可能是过程本身固有特性导致出现多个操作时段,比如发酵过程按细菌的生长周期可以划分为停滞期、指数生长期、静止期等几个阶段。以上每种情况均会导致生产过程出现多模态工况,每个模态下

收稿日期 2010-01-21; 录用日期 2010-05-13
Manuscript received January 21, 2010; accepted May 13, 2010
国家自然科学基金 (60774068); 973 计划子课题 (2009CB320601)
资助

Supported by National Natural Science Foundation of China (60774068); and the National Project for the Development of Key Fundamental Science in China (2009CB320601)

1. 东北大学信息科学与工程学院 沈阳 110004 2. 东北大学流程工业综合自动化教育部重点实验室 沈阳 110004 3. 辽河油田钻采工艺研究院中心试验所 盘锦 124000

1. School of Information Science and Engineering, Northeastern University, Shenyang 110004 2. Key laboratory of Process Industry Automation of MOE Northeastern University, Shenyang 110004 3. Drilling and Production Technology Research Institute of Liaohe Oilfield Company, Panjin 124000

的主导变量和过程特性并不相同,所对应的过程统计特征,如变量的均值、方差、相关性等往往有很大的差异。如果将现有的 MSPM 方法直接应用于这些过程往往导致很多的问题。因为过程是多模态过程,如果用一个线性模型对多个不同的模态进行描述,或者模型只能准确描述出某一模态的特性,导致在其余模态下出现大量的误报警;或者模型涵盖所有模态的操作范围,控制限很宽松,导致在某些模态下出现故障也不能及时报警,出现大量漏报警情况。针对以上出现的多种问题,本文对传统算法进行改进,提出一种全新的多模态生产过程过程监测及故障诊断方案。

针对复杂过程的多模态特性,一些学者基于传统的 PCA/PLS 方法提出了改进的多模态过程建模方法,如:韩国 POSTECH 大学的 Hwang DH 和 Han C.^[2],英国 Newcastle upon Tyne 大学的 Lane 等人^[3],基于整体建模的思想对多模态过程建立统一模型,但是这种思路所建立的模型无法准确的刻画所有的运行模态。Kosanovich 等人于 1994 指出,针对过程中两个明显具有不同特征的生产模态应该分别建立统计模型,可以更加准确、有效地监测并诊断过程中出现的异常工况^[4-5]。Dong 和 McAvoy 亦验证了 Kosanovich 等人的观点^[6]。在这种独立建模思想的基础上,新加坡国立大学的 Rajagopalan Srinivasan^[7]、Yew Seng Nga^[8]、Xuan-Tien Doan^[9],台湾中原大学的 Junhui Chen^[10],清华大学 Yongmao Xu^[11]、Shi Jian Zhao^[12] 等人,对每个稳定运行的模态独立建模,对稳定模态的监测取得了较好的监测效果。以上学者们是基于 PCA/PLS 方法对多模态监测问题的研究,解决的是单高斯模型的多模态监测。对于混合高斯模型的多模态过程监测问题,研究起步的比较晚,但是近年来也逐渐引起了不少学者的关注,如:韩国 Pohang 大学的 Sang Wook Choi^[13],韩国 Kyung Hee 大学的 Chang Kyoo Yoo^[14],荷兰 Radboud University Nijmegen 大学的 U. Thissen^[15],澳大利亚 Monash 大学的 RJ Hyndman^[16],美国 Southern California 大学的 Jie Yu^[17]、S. Joe Qin^[18] 等人均对基于混合高斯模型的多模态过程监测问题进行了深入的研究,S. Joe Qin 利用贝叶斯估计计算未知模态数据对应每种高斯分布的后验概率,在不需要已知在线运行模态的前提下,通过计算整体故障概率指标对混合高斯多模态模型进行过程监测。

当生产过程中存在多个运行模态时,从一个稳定运行模态到另外一个稳定运行模态,不可能发生突然的变化,单单考虑稳定模态是不全面的,还需要考虑到稳定运行模态之间的渐变过渡过程^[19]。过渡模态的过程特性与稳定模态的运行模态不同,过

渡模态会呈现出一种动态的渐变趋势,这种动态趋势不仅体现在过程变量的变化上,更体现在过程变量相关关系的变化上。深入研究后发现,相较于各个稳定模态的主要运行模态,模态间的过渡虽然并不代表主流的过程操作机理特性,但却是一种普遍现象并且是一种重要的过程行为。从总的趋势看来,在过渡开始初期,各时间点的过程特性与前一模态的运行特性比较接近;随着过渡的不断进行,在过渡后期,过程特性慢慢过渡到后一模态的运行模态^[20]。但同时应该指出,过渡中的过程潜在特性并不一定总是处于严格递变中,中间完全可能会出现反复与例外。此外,对于不同的运行模态,它们之间的过渡轨迹可能不同,过渡时期变量间的相关关系不同,历时也不相同。由于过渡运行模态的不稳定性,极易受到外界干扰而偏离正常的过程运行轨迹,进而影响生产安全和最终的产品质量,因此对过渡模态进行故障监测,确保模态间的平稳过渡具有重要的意义。鉴于过渡区域与稳定运行模态具有显著不同的运行模态和潜在特性,有必要将过渡区域分离出来,单独建模分析其过程特性的发展变化。

东北大学的 Ningyun L 和 Fuli Wang^[21] 等人通过“硬划分”确定稳定模态之间的过渡区域,利用与过渡模式相邻的两个子时段模型进行加权,来近似描述过渡模态的特性。在此基础上,东北大学的 Zhao C H 和 Fuli Wang^[21],香港科技大学的 Furong G, Yuan Yao^[22] 等人,进一步提出“软划分”的概念,定义 0-1 模糊隶属度作为与过渡模式相邻的两个子时段模型的权重系数,综合相邻两个子时段的过程特性,增强了过渡监测模型的鲁棒性。然而,这些方法或者无法准确判断故障与过渡,或者不能完整的描述过渡独有的特性,并没有真正根据过渡过程的数据反映出过渡区域自身的过程特性,因为过渡过程是一种模态向另一种模态渐变的动态过程,变量的相关关系具有自身的变化特性。本文从过渡模态本身的特性入手,分析过渡模态下变量的动态特性:过渡过程的变量是动态变化的,利用差分方法提取过渡过程变量的相对变化,因为相对变化更深刻的反映了过渡的动态趋势;过渡过程变量的相关关系也是动态变化的,本文把具有相同特性的过渡过程归为一个时段,用 PCA 分段方法建立过渡模型;正常的过渡过程具有一定的规律性,可以从大量的正常过渡数据建立统计模型。这样,基于相对变化对过渡模态分段建立 PCA 模型,通过分析相对变化的相关关系,可以更深入准确,更及时的反映出过渡的动态趋势。

在此,提出一套系统完整的解决多模态过程监测的方法,既考虑稳定模态过程监测,又考虑到过渡模态的监测。多模态过程中,稳定操作模态下的过

程变量相关关系并非随时间时刻变化,而不同操作模式下的变量相关关系由呈现出不同的规律性,基于此,对每个稳定模式建立不同的统计模型,使每个稳定模型可以涵盖该模式过程运行特性。对于稳定模式间的过渡,考虑到随时间推移过渡变量相关关系的变化,通过提取变量的相对变化信息,建立分段过渡子模型反映过渡时变量相关关系的动特性。

1 多模式过程数据模式识别

数据模式识别是对没有模式指示信息的数据进行分类划分,它是实现整个多模式过程监测的首要前提。无论是离线建模还是在线监测,都要先从大量的生产数据信息中识别出与其相对应的过程模式(稳定模式和过渡模式),然后在此基础上进行不同模式的离线建模或在线监测。离线建模数据的模式划分与识别就是将一组正常操作数据按照其生产模式不同进行分类划分,同时辨别出各类数据所对应的过程模式(某一稳定模式或某一过渡模式)。不同模式下的正常操作数据反映的是各自模式的过程特性,通常彼此之间具有较大的特征差异。因此,可以通过过程变量的特征提取(例如:变量的变异方向、相关性信息等等),找到能够反映过程特性变化的特征信息,并利用特征信息的相似度变化(例如:变量变异方向的夹角、变量相关系数的大小等等)对建模数据进行分类划分,实现对不同稳定模式数据与过程模式数据的聚类。然后,结合一定的过程知识(如:生产数据的时序信息、生产操作指示变量、专家经验等等),对各类建模数据所对应的稳定模式或过渡模式加以识别。多模式模型在线监测时,首先也需要判断出当前时刻数据的运行模式,然后才能选择与其对应的模式模型进行过程监测。如果在线运行的数据具有模式指示变量,则可以根据指示信息直接找到对应的模型;当无法在线获得模式指示变量时,需要通过匹配当前数据与各个模式运行特性的相似度来判断当前运行模式。根据以上分析,多模式过程监测可以分为如下几种情况,如图 1 所示:

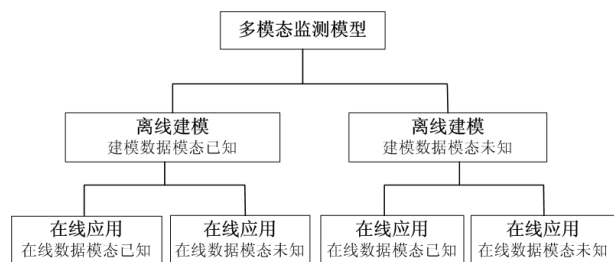


图 1 多模式过程的四种情况

Fig. 1 Four situations of multi-mode process

建模数据所对应的运行模式已知,在线监测时数据所对应的运行模式已知。这种情况主要是指过程在运行时有模式指示标签(例如:某个在线可测的过程变量),可以根据指示标签准确的判断出建模数据以及在线监测时当前工况所对应的运行模式。比如具有多时段特性的等长间歇过程,其固有的过程特性导致出现的多时段情况,离线数据和在线数据对应的运行模式均可以根据时间标签判断得到。

建模数据所对应的运行模式已知,在线监测时数据所对应的运行模式未知。在实际工业生产过程中,因为现场众多信号来源是多渠道的,很难完全做到在线同步,经常会出现某些指示标签信号延迟出现。如果模式指示标签信号不能及时得到,那么多模式模型在线运行时,就需要采取特殊的方法来判断当前时刻对应的运行模式,从而选取与其对应的过程监测及故障诊断模型。

建模数据所对应的运行模式未知,在线监测时数据所对应的运行模式已知。这种情况通常出现较少,但是却实际存在。工业生产中的数据库并不是所有历史数据都进行存储,如果模式指示信号可以做到在线同步,但是历史数据并没有对该信号进行存储,便会出现建模数据所对应的运行模式未知,而在线数据所对应的运行模式已知的情况。这时,需要对建模数据进行特殊的处理,分辨出不同运行模式下对应的正常历史数据,再分别建立过程监测及故障诊断模型。

建模数据所对应的运行模式未知,在线监测时数据所对应的运行模式未知。以上三种情况其实都可以看成最后这种情况的特例,所以对这种情况的研究最具有代表性和推广性。当建模数据和在线数据所对应的运行模式都未知时,既需要对建模数据的类型进行区分,同时还需要对在线的运行模式进行判断。

以上不管是哪一种多模式情况,基于多变量统计方法的多模式过程监测分为两部分:离线建模(如图 2 所示),以及在线监测(如图 3 所示)。离线建模包括三部分:①对建模数据所对应的运行模式进行区分;②识别聚类后数据所对应的运行模式;③对多个稳定模式和不同的过渡模式分别建立统计模型。进行过程在线监测时,也包括三部分:①根据当前数据对应的过程特性识别其所属的运行模式;②选择当前模式对应的离线模型;③调用对应的模型实现过程监测及故障诊断。如果多模式过程的建模数据所对应的运行模式已知,则可以省略图 2 中步骤 Step1 与 Step2;如果在线数据所对应的运行模式已知,则可以省略图 3 中步骤 Step1。本文从最简单的情况入手分析,假设现场有已知的模式指示变量:建模数据所对应的运行模式已知,在线监测时

数据所对应的运行模式已知，提出一套完整的解决多模态过程监测及故障诊断的方法。

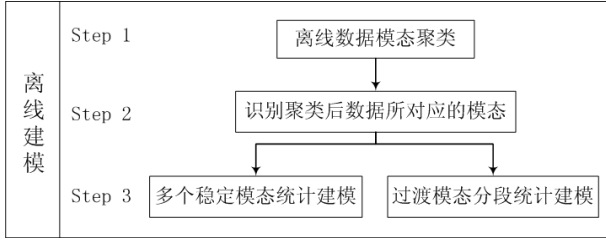


图 2 多模态过程离线统计建模步骤
Fig. 2 Offline steps of multi-mode process modeling

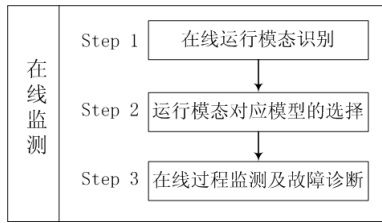


图 3 多模态过程在线监测及故障诊断步骤
Fig. 3 Online steps of multi-mode process monitoring

2 多模态生产过程故障监测模型的离线建模

2.1 基于多 PCA 模型的稳定模式建模

因为不同生产模式下变量的均值、方差、相关关系均有所不同，不可以用同一模型描述，所以对多个操作模式的过程进行监控，需要分别建立多个 PCA 模型。首先，根据已知条件，区分不同操作模式下的数据。假设生产过程一共包含 M 个操作模式，判断得到 m 模式的建模数据为 $X^{(m)}(N^{(m)} \times J)(m=1, \dots, M)$ ，每一行由某一采样时刻的所有过程变量测量值组成，每一列是某个过程变量在 m 模式下所有时刻的值。然后用 PCA 对建模数据分析，求取对应的得分矩阵 T 和负载矩阵 P 。在信息损失最少的前提下，对原始数据进行压缩，提取数据特征。经过主元分析，每个稳定模式下的数据 $X^{(m)}$ 被分解为如下表达，

$$X^{(m)} = T^{(m)} \cdot P^{(m)T} \quad (1)$$

其中， $T^{(m)}(N^{(m)} \times A^{(m)})$ 和 $P^{(m)}(A^{(m)} \times J)$ 分别是主元得分矩阵和负载矩阵； $A^{(m)}$ 是保留的主元个数。

利用 PCA 进行故障监测，传统的统计指标有两种：Hotelling- T^2 和 SPE。 T^2 统计量反映了每个主成分在变化趋势和幅值上偏离模型的程度，是对模型内部变化的一种度量，它可以用来对多个主元同时进行监测；SPE 统计量刻画了输入变量的测量

值对主元模型的偏离程度，是对模型外部变化的一种度量^[23]。第 m 个模式的统计量计算公式如下：

$$T^{2(m)} = \mathbf{t}^{(m)}(S^{(m)})^{-1}\mathbf{t}^{(m)T} \quad (2)$$

$$SPE^{(m)} = \left(X^{(m)} - \hat{X}^{(m)}\right) \left(X^{(m)} - \hat{X}^{(m)}\right)^T \quad (3)$$

其中， $\mathbf{t}^{(m)}$ 是 $T^{(m)}$ 的行向量； $S^{(m)}$ 是前 $A^{(m)}$ 个的特征值构成的对角矩阵。

由于两个统计量反映的是两个互补空间的数据特征，并且在实际监控中二者所占的比重并不完全一致。Hotelling- T^2 更多的是反映变量在投影尺度上蕴含的信息，SPE 侧重变量相关关系的信息。所以为了更好的平衡两个不同的统计量，在此，引用 Yue 和 Qin 提出的重构统计量作为过渡模式相对变化监测的指标^[24]。它主要由 Hotelling- T^2 和 SPE 加权组成，第 m 个模式的重构统计量计算如下：

$$\phi^{(m)} = \frac{T^{2(m)}}{C_{T^2}^{(m)}} + \frac{SPE^{(m)}}{C_{SPE}^{(m)}} = X^{(m)}\Phi^{(m)}X^{(m)T} \quad (4)$$

其中， $C_{T^2}^{(m)}, C_{SPE}^{(m)}$ 代表第 m 个模式 T^2 和 SPE 统计量的控制限（计算方法参见附录）； $\Phi^{(m)} = \frac{P^{(m)}(S^{(m)})^{-1}P^{(m)T}}{C_{T^2}^{(m)}} + \frac{(I-P^{(m)}P^{(m)T})}{C_{SPE}^{(m)}}$ 。确定第 m 个模式重构指标的控制限 $C_{\phi}^{(m)}$ ，方法分为两种：如果过程数据矩阵 $X^{(m)}$ 沿采样方向近似服从正态分布时，可以通过经验公式得到统计量的分布进而确定控制限 $\phi^{(m)} \sim g^{(m)}\chi^2(h^{(m)})$ ，其中 $g^{(m)} = \frac{\text{tr}(S^{(m)}\Phi^{(m)})^2}{\text{tr}(S^{(m)}\Phi^{(m)})}$ ； $h^{(m)} = \frac{[\text{tr}(S^{(m)}\Phi^{(m)})]^2}{\text{tr}(S^{(m)}\Phi^{(m)})^2}$ ； $S^{(m)} = \text{cov}(X^{(m)})$ ；如果过程数据不满足正态分布，可以通过核密度估计法确定统计量的分布进而确定其控制限。

2.2 基于差分分段 PCA 的过渡模式建模

2.2.1 过渡过程数据特点

不同稳定模式转换时，需要对模式间的过渡过程进行监测，监测过渡模式是实现整个过程监测的难点。通过深入研究发现，在过渡模式下，尽管过渡的运行特性会实时发生变化，但是正常过渡的演化服从某种固定轨迹。也就是说，在忽略随机干扰的前提下，变量在过渡时总体会沿着某一趋势有规律的渐变，变量在相同时间段内的相对变化有规律可循，这种规律揭示了过渡时变量时变的趋势与特性。用统计的方法学习相对变化的趋势，可以判断当前过渡是否与建模时的过渡规律一致，进而可以判断出当前时刻过渡模式是否运行在正常的轨迹下。

假设生产过程有已知条件，可以根据指示变量判断当前运行状态属于的模式。首先以生产过程有两个稳定运行模式为例分析：A 模式和 B 模式。不难理解，稳定模式间的过渡存在两种状态：一种过渡

是 A 稳定模态向 B 模态过渡,这类过渡区域定义为 AB 过渡模态;另一种过渡模态是 B 稳定模态向 A 模态过渡,这类过渡区域定义为 BA 过渡模态。不同模态间的过渡过程,数据变化必然会呈现出不同的相关特性,所以需要不同的过渡模态——AB 过渡和 BA 过渡两种模态,分别建立对应的过渡模型。

无论连续生产过程还是间歇生产过程,在一定前提假设下,可以认为相同过渡模态的过程数据具有重复性。“相同的过渡模态”是指运行状态从一个相同的稳定模态(起始模态)逐步过渡到下一个相同的稳定模态(终止模态)。以稳定模态 A 向稳定模态 B 的过渡为例,根据设备、操作工艺等条件可以设定一条生产中最佳的 AB 过渡轨迹,生产中每次正常的 AB 过渡模态遵循最佳过渡轨迹,所以每一次 A 模态向 B 模态的过渡可以认为具有一致的过程特性。相同的过渡模态数据具有重复性,所以累积多次(I 次)正常过渡的生产数据,可以得到某一过渡模态的建模数据矩阵 $X(I \times J \times K)$,数组三维分别表示过渡次数 I ,过程变量个数 J ,以及过渡过程的采样个数 K 。实际生产中, I 次过渡历经时间有长有短,数据会出现不等长现象,常见的处理手段是在建模之前,用一些方法将不等长数据变成等长的数据。其中最简单的处理方式,一是“最短长度法”,即找到最短的一次过渡操作数据,将其余的过渡数据截取使得它们都具有最短的数据长度。这种方法除了要求过程变量轨迹在公共部分保持一致之外,还要求过渡过程的主要操作及重要过程信息都被包含在公共时间段内。二是用一个合适的“指示”变量代替过程采样时间,采用插值的手段将变量轨迹统一成同样的长度。此外,近年来研究学者尝试采用动态时间扭曲(DTW)的思想通过适当地平移、拉伸或压缩两个不同信号轨迹的局部片断,从而达到同步两个不同长度甚至不同形状的信号轨迹的目的。但是,对原始过程轨迹的处理会在一定程度上掩盖异常工况的数据特征,降低故障的可检测性。分析实际数据特点,利用恰当的处理方法将不等长建模数据化成采样点个数均为 K ,得到某一过渡模态的建模数据阵 $X(I \times J \times K)$ 。

三维数据 $X(I \times J \times K)$ 沿时间采样 K 方向进行垂直切割,可以得到 K 个二维数据片。取步长为 h ,沿时间方向计算各个时间片的差分矩阵。以第 k 个时间的片 $X_k(I \times J)$ 为例,它对应的差分矩阵为 $X_{k-h}^* = X_k - X_{k-h+1}$,是通过第 k 个时刻的时间片,与之前间隔 h 个时刻的时间片作差获得的。如图 4 所示,最终可以得到三维数组对应的差分矩阵 $X^*(I \times J \times (K-h))$ 。

得到的差分矩阵强调的是过渡模态下,各时间

片沿时间方向上的相对变化,反映的是变量在过渡时,固定时间间隔下变化的趋势。差分步长的选择可以根据实际情况而定:过渡趋势变化比较缓慢,说明变量在过渡时,单位采样时间变化幅度比较小,可以选择差分步长比较大;过渡模态变化比较快速,说明变量在过渡时变化趋势比较大,差分步长可以选择比较小,以便很好的跟踪描述过渡细节变化。

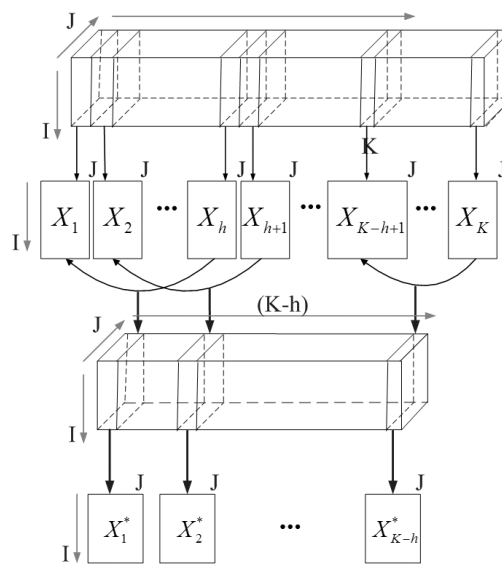


图 4 三维数据延时间轴求取差分矩阵

Fig. 4 Illustration of three-way data difference trajectory construction

2.2.2 基于差分分段 PCA 的过渡模态故障监测模型分段离线建模

$X^*(I \times J \times (K-h))$ 沿时间轴 $(K-h)$ 做时间切片,可以得到 $(K-h)$ 个二维时间切片 $X_k^*(I \times J)$,它是由第 k 个时间时刻与之前间隔 h 个时刻的时间片作差获得的。这 $(K-h)$ 个时间片描述了整个过渡的变量变化趋势,但是整个过渡过程变量的变化趋势并不是亘一不变。在过渡开始初期,变量的变化幅度会比较大,但是随着过渡的逐步完成,为了精准的达到下一个稳定模态,过渡后期变量的相对变化逐渐变小。变量的相对变化在整个过渡过程中会发生变化,变量的相关关系在过渡过程中也会发生变化。所以整个过渡过程每个时间片的差分矩阵不可以用统一的模型描述,需要根据不同时期,对过渡变量的不同变化特性进行归类,对整个过渡过程分阶段描述。这样建模的优势在于:差分矩阵强调的是变量相对变化,根据差分变量相关性分段,把变化模式相似的归为一类,比直接对过渡数据聚类减少了分段的数目;而且分段数目减少,对应每一子时段的时间片个数变多,相当于增加了建模采样的个数。实际生产中的过渡数据有限,基于差分矩

阵的分段算法很好的解决这个问题,增加了建模数据,使模型的稳定性得以提高。

对每一个二维时间片矩阵 X_k^* 利用 PCA 方法建立统计分析模型 $X_k^* = T_k^* P_k^{*T}$, 获得 $(K-h)$ 个负载矩阵 P_k^* , 它们表征了 $(K-h)$ 个差分数据阵变量相对变化之间的相关性信息。对这 $(K-h)$ 个负载矩阵, 用传统的 k-means 聚类算法进行聚类, 使变量相对变化具有相同特征关系的归为一个模型。这样, $(K-h)$ 个差分矩阵时间片被分成 C 个过渡子模型, 表征 C 种不同的过渡变化特性, 分别对 C 个子模型建模。

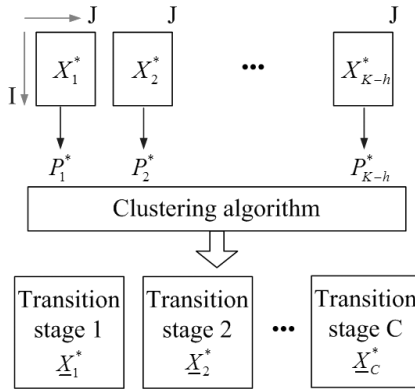


图 5 时间片聚类示意图

Fig. 5 Illustration of time slices partition algorithm

得到第 c 个子类的过渡数据差分三维数组 $X_c^*(I \times J \times n_c)$, 其中, I 代表过渡进行的次数; J 代表变量个数; n_c 为第 c 个过渡子模型包含的差分时间片的个数。将三维数组沿变量方向展开, 得到矩阵 $\tilde{X}_c^*(In_c \times J)$, 将过渡批次与时间方向上的数据揉和, 每一列代表在第 c 个过渡子模式下所有批次所有时刻上变量的波动。

用 PCA 对每个子模型下的差分数据阵 $\tilde{X}_c^*(In_c \times J)$ 建模, 提取每个子模型下过渡变量变化的特征。经过主元分析, 可以得到模型, $\tilde{X}_c^* = T_c P_c^T$ 。其中, $T_c(In_c \times A_c)$ 为 c 子模型差分矩阵的主元得分矩阵; A_c 是保留的主元个数; $P_c(J \times A_c)$ 为负载矩阵, 是主元的投影方向。

用公式 (4) 计算重构统计量, 对第 c 个过渡子模型进行故障监测, 如下:

$$\phi_c = \frac{T_c^2}{C_{T^2,c}} + \frac{SPE_c}{C_{SPE,c}} = \tilde{\mathbf{x}}_c^* \Phi_c \tilde{\mathbf{x}}_c^{*T} \quad (5)$$

其中, $T_c^2 = \tilde{\mathbf{x}}_c^* P_c S_c^{-1} P_c^T \tilde{\mathbf{x}}_c^{*T}$; $SPE_c = \tilde{\mathbf{x}}_c^* (I - P_c P_c^T) \tilde{\mathbf{x}}_c^{*T}$; S_c 是前 A_c 个的特征值构成的对角矩阵; $C_{T^2,c}$, $C_{SPE,c}$ 分别代表第 c 个过渡子模型 T^2 和 SPE 统计量的控制限; $\Phi_c = \frac{P_c S_c^{-1} P_c^T}{C_{T^2,c}} + \frac{I - P_c P_c^T}{C_{SPE,c}}$; $\tilde{\mathbf{x}}_c^*$ 是 \tilde{X}_c^* 的行向量。用第 2.1 节稳定模态控制限公式同

理求出第 c 个过渡子模型的重构指标控制限 $C_{\phi,c}$ 。

2.2.3 对 M 个模态间过渡过程监测的扩展

两个稳定运行模态 A 和 B 之间过渡模态有两种: 一种过渡是 A 稳定模态向 B 模态过渡; 另一种过渡模态是 B 稳定模态向 A 模态过渡。以此类推, M 个稳定模态之间的过渡模态有 $M \cdot (M-1)$ 种。

假设 i 模态向 j 模态过渡, 其中 $i \in [1, M], j \in [1, M], i \neq j$, 对应的过渡数据矩阵为 $X^{(ij)}(I \times J \times K)$ 。沿时间采样 K 方向进行切割得到 K 个二维数据片。取步长为 h , 沿时间方向计算各个时间片的差分矩阵, 得到三维数组对应的差分矩阵 $X^{(ij)*}(I \times J \times (K-h))$ 。按照节 2.2.2 所述方法, 对 $i-j$ 过渡过程进行分段, 分为 $C^{(ij)}$ 个过渡子模型, 则 $i-j$ 过渡过程的分段模型为: $\tilde{X}_c^{(ij)*} = T_c^{(ij)} P_c^{(ij)T}$ ($c = 1 \dots C^{(ij)}$), 同时可以得到每个子模型的统计量控制限。

3 多模态生产过程故障监测模型在线监测

对过渡模态进行在线故障监测, 首先根据已知条件判断当前生产过程对应模态。如果当前生产对应模态为稳定模态, 判断稳定操作模态种类, 调用对应的模型进行故障监测; 如果当前对应模态为过渡模态, 判断过渡过程前后模态, 调用对应的过渡模型进行故障监测。

假设当前运行模态为 A, 经过一段时间过渡模态, 最终过渡到新稳定模态 B。根据先验知识, 获取确定为 A 模态下的在线数据 $\mathbf{x}_{on}^{(A)}(1 \times J)$, J 表示过程变量个数。调用 A 模态监测模型, 计算在线的重构统计量 $\phi_{on}^{(A)}$:

$$\phi_{on}^{(A)} = \mathbf{x}_{on}^{(A)} \Phi_{on}^{(A)} \mathbf{x}_{on}^{(A)T} \quad (6)$$

其中, $\Phi_{on}^{(A)} = \frac{P^{(A)}(S^{(A)})^{-1}P^{(A)T}}{C_{T^2}^{(A)}} + \frac{I - P^{(A)}P^{(A)T}}{C_{SPE}^{(A)}}$ 。

当模型进入过渡模态, 运行 h 个采样时刻后, 计算第 h 时刻对应的在线差分向量 $\mathbf{x}_{on}^* = \mathbf{x}_{on}^h - \mathbf{x}_{on}^1$ 。下一步, 需要判断当前差分向量对应的是第几个过渡子时段, 根据差分后时间指针, 判断当前过渡时刻对应的过渡子时段, 调用对应的过渡子模型, 计算在线的重构统计量 $\phi_{on}^{(AB)*}$:

$$\phi_{on}^{(AB)*} = \mathbf{x}_{on}^* \Phi_{on}^{(AB)*} \mathbf{x}_{on}^{*T} \quad (7)$$

其中 $\Phi_{on}^{(AB)*} = \frac{P_{on}^{(AB)}(S_{on}^{(AB)})^{-1}P_{on}^{(AB)T}}{C_{T^2,on}^{(AB)}} + \frac{I - P_{on}^{(AB)}P_{on}^{(AB)T}}{C_{SPE,on}^{(AB)}}$;

$P_{on}^{(AB)}$, $S_{on}^{(AB)}$ 分别是当前时刻对应 AB 过渡子模型的负载矩阵和对角矩阵; $C_{T^2,on}^{(AB)}$, $C_{SPE,on}^{(AB)}$ 分别是当前时刻对应 AB 过渡子模型的统计量控制限。由于各种各样的原因, 实际工业过程不可能达到完全地

重复生产,每次过渡历经时间有长有短,因此在线过渡数据也会与建模数据不等长。为了满足在线实时的监测,本文用一个简单的算法处理。在线过渡时间短于建模过渡时间,不做处理;如果在线过渡时间长于建模时间,根据差分后时间标签调用对应子模态,长于建模时间的在线过渡时间段,调用最终子模态进行监测。

将实时获得的在线统计量 ϕ_{on}^* 与对应的控制限进行比较,判断当前时刻系统是否出现异常。当在线统计量超出对应控制限后,则认为当前过渡运行状况出现异常,系统发出报警信号,同时,可以调用对应的统计量贡献进行故障诊断。经过一段时间过渡模态,最终过渡到新稳定模态 B。获取确定为 B 模态下的在线数据 $\mathbf{x}_{on}^{(B)}(1 \times J)$ 。调用 B 模态监测模型,同理用公式 (6) 计算在线的重构统计量 $\phi_{on}^{(B)}$ 。

4 多模态过程故障监测离线建模、在线监测步骤

总结基于差分分段 PCA 的多模态故障监测离线建模与在线监测的步骤,如下:

离线建模

1). 分别建立 M 个稳定模态对应的模型 $X^{(m)} = T^{(m)} \cdot P^{(m)T}(m=1, \dots, M)$, 计算模态对应的重构统计量 $\phi^{(m)}$ 及其控制限 $C_{\phi}^{(m)}$;

2). 分别建立 $M \cdot (M-1)$ 种过渡模态对应的模型。得到 i 模态向 j 模态过渡时 I 次正常数据 $X^{(ij)}(I \times J \times K)$, 沿时间方向作步长为 h 的差分得到 $X^{(ij)*}(I \times J \times (K-h))$;

3). 对差分数据沿时间轴做切片,对每个时间片用 PCA 建模 $X_k^{(ij)*} = T_k^{(ij)*} P_k^{(ij)*T}$, 并根据负载矩阵 $P_k^{(ij)*}$ 聚类,将过渡模态分为 C 个子模态;

4). 每个子模态的数据 $X_c^{(ij)*}(I \times J \times n_c)$ 沿变量轴展成二维矩阵 $\tilde{X}_c^{(ij)*}(In_c \times J)$;

5). 建立每个子模态对应的 PCA 模型 $\tilde{X}_c^{(ij)*} = T_c^{(ij)} P_c^{(ij)T}$, 计算每个子模态对应的重构统计量 $\phi_c^{(ij)}$ 及其控制限 $C_{\phi}^{(ij)}$;

在线监测

1). 得到在线数据 $\mathbf{x}_{on}(1 \times J)$, 根据已知条件判断当前生产过程对应模态。如果当前生产对应模态为稳定模态 m , 调用对应的监测模型;

2). 计算在线的重构统计量 $\phi_{on}^{(m)}$, 当 $\phi_{on}^{(m)}$ 超出 $C_{\phi}^{(m)}$ 时, 认为当前稳定模态下出现故障, 利用贡献图进行故障诊断;

3). 如果当前对应生产为过渡模态, 判断过渡过程前后模态, 调用对应的过渡模型;

4). 在进入过渡模态运行 h 个时刻后, 计算在线差分向量 \mathbf{x}_{on}^* ;

5). 根据差分后时间指针, 判断当前过渡时刻对

应的过渡子时段, 调用对应的过渡子模型, 计算在线的重构统计量 ϕ_{on}^* ;

6). 当 ϕ_{on}^* 超出 $C_{\phi,c}$ 时, 认为当前过渡状态出现故障, 利用贡献图进行故障诊断;

5 算法验证及结果

在宝钢连续退火生产线中, 退火炉是进行退火热处理的重要设备。本方法用退火生产线的实际过程运行数据进行方法验证。退火炉内主要通过加热、均热、慢冷、快冷、过时效、终冷等过程, 如图 6 自右向左所示。其内部结构经历晶粒恢复、再结晶、晶粒长大、碳化物析出等几个阶段的组织变化过程, 使带钢的内在质量得到改善。

表 1 连续退火机组过程的建模变量

Table 1 Process variables for continuous annealing line

序号	变量名称	序号	变量名称
1	HF 加热炉带钢温度	20	1C 中央段速度
2	HF 中央段速度	21	1C 一冷炉带钢温度
3	HF 1 区炉温	22	1C 1 区炉温 (TOP)
4	HF 2 区炉温	23	1C 2 区炉温 (BOTTOM)
5	HF 3 区炉温	24	1C 3 区炉温 (TOP)
6	HF 4 区炉温	25	1OA-1 1 区炉温
7	HF 5 区炉温	26	1OA-1 2 区炉温
8	SF 中央段速度	27	1OA 中央段速度
9	SF 1 区炉温	28	1OA1 过时效炉带钢温度
10	SF 2 区炉温	29	1OA2 过时效炉带钢温度
11	SF 均热炉带钢温度	30	2OA-1 中央段速度
12	SCF 中央段速度	31	2OA-1 1 区炉温
13	SCF 缓冷炉带钢温度	32	2OA-1 2 区炉温
14	SCF 1 区炉温	33	2OA-2 中央段速度
15	SCF 2 区炉温	34	2OA-2 1 区炉温
16	SCF 3 区炉温 (BOTTOM)	35	2OA-2 2 区炉温
17	SCF 4 区炉温 (TOP)	36	2C 中央段速度
18	RH 中央段速度	37	2C 二冷炉带钢温度
19	RH 再加热炉带钢温度	38	2C 炉温

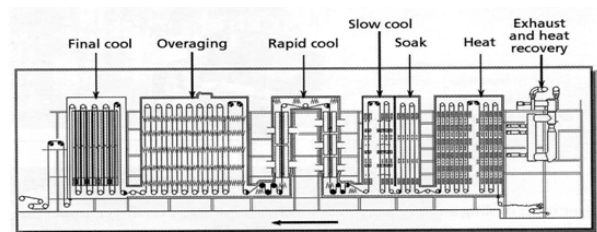


图 6 连续退火机组生产工艺示意图

Fig. 6 Flow diagram of the continuous annealing line

实际生产中, 退火带钢的调质度信息可以获得。本文以调质度为 T-3CA 和 T-4CA 的两种带钢为

例, 该生产线有四种生产模态: T-3CA 稳定运行模态、T-4CA 稳定运行模态、T-3CA 向 T-4CA 过渡模态、T-4CA 向 T-3CA 过渡模态。对该生产线的过渡模态进行故障监测, 首先需要离线分别建立 T-3CA 向 T-4CA 过渡监测模型、T-4CA 向 T-3CA 过渡监测模型。根据退火机组具体的生产工艺状况, 带钢在连续退火过程中历经的实际生产工艺阶段有 8 个, 分别为: 加热、均热、慢冷、重加热、1 冷、1 过时效、2 过时效、2 冷。带钢退火炉的生产质量是由退火过程各炉段共同决定, 选取 8 个工艺段中反映退火生产过程的 38 个关键变量作为模型输入, 现场变量的采集频率为 1 次/秒。

正常生产操作采集 T-3CA 稳定运行模态数据 $X_{T_3}(4000 \times 38)$, T-4CA 稳定运行模态数据 $X_{T_4}(4800 \times 38)$, 分别用第 2.1 节所提多模态方法对两种稳定模态建立 PCA 监测模型, 同时确定稳定模态统计量的控制限。因为每次过渡都是从相同的起始模态 (T-3CA) 过渡到相同的终止模态 (T-4CA), 所以正常的过渡数据基本服从一致的运行轨迹, 采集 40 次正常过渡数据构成 T-3CA 向 T-4CA 过渡的建模数据。但是实际生产中, 外界因素的干扰会造成每次过渡模态时间略有不同, 首先需要对不等长的建模数据进行预处理。结合工艺选择一次标准过渡数据 (标准数据的采样次数是 710), 用动态时间扭曲的方法对其余过渡数据进行等长处理, 分别获得 T-3CA 向 T-4CA 过渡模态的三维数组 $X_{T_3 \rightarrow T_4}(40 \times 38 \times 710)$, 以及 T-4CA 向 T-3CA 过渡模态的三维数组 $X_{T_4 \rightarrow T_3}(40 \times 38 \times 710)$ 。其中 38 个过程变量, 如表 1 所示。取步长为 10, 沿采样时刻方向取时间片差分, 得到差分矩阵为 $X(40 \times 38 \times 700)$ 。用 2.2 节所提差分分段 PCA 建模方法对两种过渡模态分别建立故障监测模型, 同时确定过渡模态统计量的控制限。

离线聚类 T-3CA 向 T-4CA 过渡过程, 分成了 3 个过渡子模态, 如图 7 所示: 1~65, 66~605, 606~700。究其原因, 不难发现, 在过渡刚开始时, 第一子模态内变量属于启动阶段, 与前一模态联系紧密; 当过渡到第二子模态时, 变量进入快速变化阶段, 各变量相对变化间相关关系联系紧密; 当过渡到达第三子模态时, 过渡变量变化进入相对缓慢阶段, 相关关系与后一模态较为紧密。同理, T-4CA 向 T-3CA 过渡过程分成了 3 个过渡子模态: 1~75, 76~638, 639~700, 如图 8 所示。

在线应用时, 获得在线数据 $\mathbf{x}_{new}(1 \times 38)$, 根据在线获得的调质度知识判断数据的模态。已知数据与第 221 秒采样时刻开始由调质度 T-3CA 向 T-4CA 过渡, 过渡开始 10 个采样时间后, 计算差分矩

阵 $\tilde{\mathbf{x}}_{new}(1 \times 38)$ 。根据时间标签判断数据处于过渡

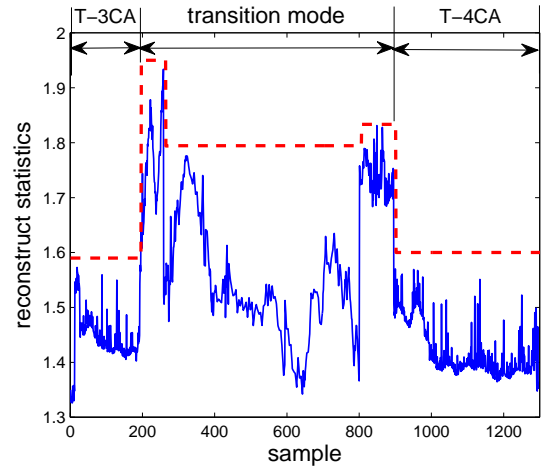


图 7 基于差分分段算法, T-3CA 向 T-4CA 过渡模态重构统计量故障监测图 (虚线代表重构统计量的控制限; 实线代表在线统计量)
Fig. 7 Monitoring plots for transition of T-3CA to T-4CA based on difference trajectory sub-phase model (Dashed line is control limit of statistic; solid line is online statistic.)

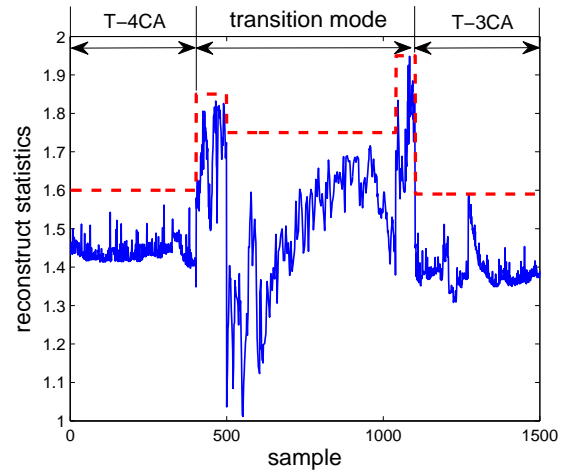


图 8 基于差分分段算法, T-4CA 向 T-3CA 过渡模态重构统计量故障监测图
Fig. 8 Monitoring plots for transition of T-4CA to T-3CA based on difference trajectory sub-phase model

子模态, 根据公式 (7) 计算在线统计量。T-3CA 向 T-4CA 过渡模态的监测结果如图 7 所示, (图 7 中, 虚线 “--” 代表重构统计量的控制限; 实线 “—” 代表在线统计量; 下文中图 8、图 9、图 10、图 11 仿真曲线含义与图 7 一致, 本文不再赘述) 过渡进行 710 秒以后结束, 进入稳定模态 T-4CA。T-4CA 向 T-3CA 过渡模态的检测如图 8 所示, 与第 402 秒开始由调质度 T-4CA 向 T-3CA 过渡, 过渡进行 710

秒以后结束,进入稳定模态 T-3CA。

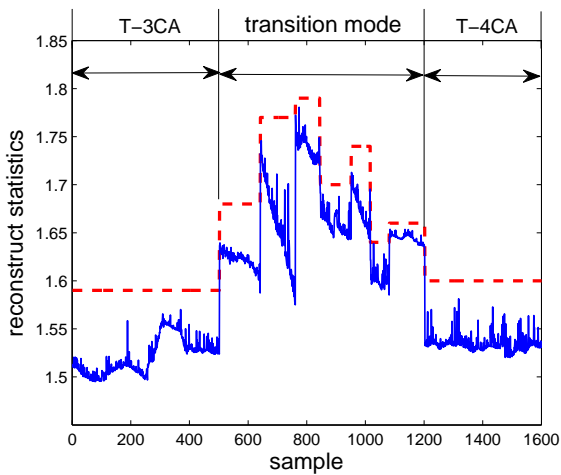


图9 基于传统子模态划分算法, T-3CA 向 T-4CA 过渡模态重构统计量故障监测图

Fig.9 Monitoring plots for transition of T-3CA to T-4CA based on tradition sub-phase model

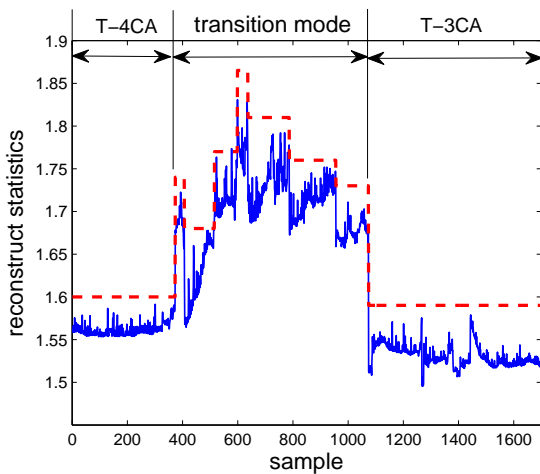


图10 基于传统子模态划分算法, T-4CA 向 T-3CA 过渡模态重构统计量故障监测图

Fig.10 Monitoring plots for transition of T-4CA to T-3CA based on tradition sub-phase model

为作比较,用 k-mean 聚类算法对过渡过程直接进行聚类,分析结果发现直接聚类的类别数明显比对差分矩阵聚类的数目多。如图 9 所示,对 T-3CA 向 T-4CA 过渡模态聚类得到 7 个子模态,如图 10 所示,对 T-4CA 向 T-3CA 过渡模态聚类也得到 7 个子模态。模态数目的增多不仅仅增加了过渡子模型的个数,而且会降低每个子模型的建模样本数目。对于采样数目有限的过渡过程而言,样本数目的减少会很大程度的影响模型稳定性。用 T-3CA 向 T-4CA 过渡在线数据 $\mathbf{x}_{new}(1 \times 38)$ 带入模型计算

在线监测结果如图 9 所示。同理, T-4CA 向 T-3CA 过渡在线监测结果如图 10 所示。

为了检测差分模型对过渡过程监测的有效性,以 T-4CA 向 T-3CA 过渡模型为例,获取在线非正常过渡数据。已知该次过渡由于机械故障导致在过渡第 253 秒采样时刻, HF 段 3 区、4 区炉温出现过渡异常。数据带入差分模型,统计量在过渡第 269 秒做出响应,忽略差分步长造成的滞后,即模型延后 6 秒准确指示异常情况的发生,如图 11(a) 所示。用 7 段过渡子模态对故障数据监测,如图 11(b) 所示,统计量在过渡第 344 秒超出控制限。比较 11(a)11(b) 发现基于差分的监测模型准确度和鲁棒性都优于 7 段子模态的过渡监控模型。

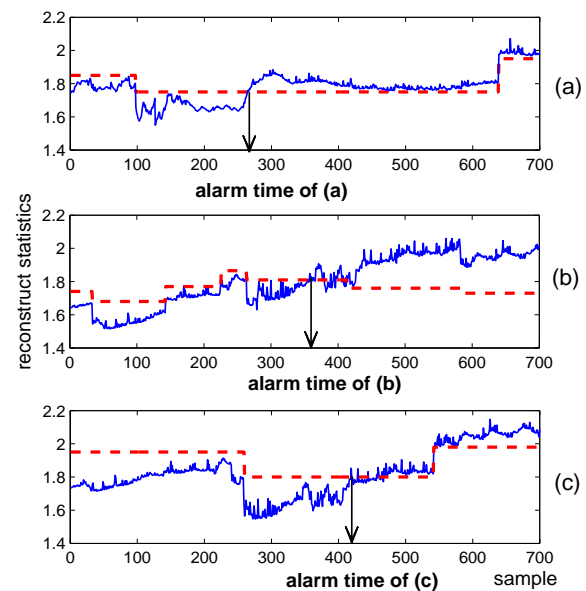


图11 T-4CA 向 T-3CA 过渡模态故障状态示意图 (a) 基于差分分段算法的监测图; (b) 基于 7 段子模型的监测图; (c) 基于 3 段子模型的监测图

Fig.11 Monitoring for transition of T-4CA to T-3CA (a)Monitoring plots of difference trajectory sub-phase model; (b) Monitoring plots of 7 sub-phases model; (c)Monitoring plots of 3 sub-phases model.

为了比较在相同过渡子模态数目下,基于差分数据监测的优势,离线把过渡段用聚类算法划分为 3 个子模态,如图 11(c) 所示,统计量在过渡第 416 秒超出控制限。比较 11(a)11(c) 发现划分相同数目的子模态建模,基于差分的监测模型准确度远远优于直接聚类的子模态监控模型。由以上对比得出结论,利用“相对变化”建立的过渡过程故障监测模型具备很有效的故障监测能力。

通过贡献图分析也可准确推断,发生故障的可能变量,如图 12 所示。图中所示贡献比较大的 1

号、5号、6号变量分别对应为：HF 加热炉带钢温度、HF3 区炉温、HF4 区炉温，与实际情况相符，证明了该算法的有效性。

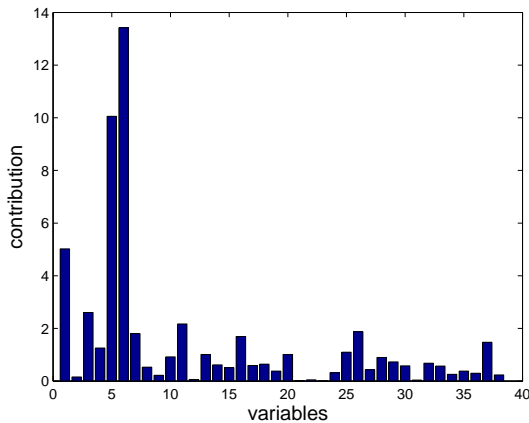


图 12 过渡第 279 秒采样时刻，对应变量贡献图

Fig. 12 Contribution plots in 279th transition sample

6 结论

多模态的过程监测是一个复杂的问题，既需要考虑稳定模态下的过程监测，也需要考虑不同模态间的过渡过程监测。不同操作模态的数据具有不同的相关关系，对每个模态，需要建立不同的稳定模态模型。稳定模态间的过渡过程更是一个复杂的过程，过渡模态最大的特点就是变量的动态特性。这种“动态”不仅体现在变量的变化上，也体现在变量相关关系的变化。本文通过对过渡数据差分，得到变量相对变化数据。根据差分变量相关关系分类，将变量“相对变化”相关关系相近的过渡段归为一类建模，分段描述过渡子模态。用 PCA 分段提取变量变化方向的主要信息，对每一段相对变化的特性进行分析，体现相关关系的动态变化。以连续退火机组生产线为背景，对不同调质度的生产过程分别建立不同的稳态模型，对于稳定模态间的过渡用差分分段 PCA 进行故障监测。用实际数据仿真发现离线的子模态分段特性能很好的反映实际过渡过程机理，并且对于过渡过程出现的过渡故障能够及时监测、准确诊断。因此，基于差分的分段 PCA 多模态故障监测方法对于多模态过程的故障监测，具有一定的实用价值。

附录

根据正文公式 (2) 和 (3) 分别求取变量对应的 Hotelling- T^2 和 SPE 统计量，在此给出统计量控制限的计算公式。分两种情况：

1) 如果过程数据矩阵 $X^{(m)}$ 沿采样方向近似服

从正态分布时，可以通过经验公式得到统计量的分布进而确定控制限，计算公式如下：

$$T^2(m) \sim \frac{A^{(m)}(N^{(m)} - 1)}{N^{(m)} - A^{(m)}} F_{\alpha}(A^{(m)}, N^{(m)} - A^{(m)}) \quad (8)$$

$$SPE(m) \sim g^{(m)} \chi_{h, \alpha}^{(m)2} \quad (9)$$

其中， $A^{(m)}$ 是第 m 个模态保留的主元个数； $N^{(m)}$ 是建模采样个数； $g^{(m)} = \frac{v^{(m)}}{2m^{(m)}}$ ； $h^{(m)} = \frac{2(m^{(m)})^2}{v^{(m)}}$ ； $m^{(m)}$ 、 $v^{(m)}$ 分别为第 m 个模态 SPE 统计量的均值和方差。

2) 如果过程数据不满足正态分布，二者的分布密度可以由核密度估计得出，其公式描述如下：

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K \left\{ \frac{x - x_i}{h} \right\} \quad (10)$$

其中， x 是要进行分布密度估计的分析对象，即 T^2 、SPE 统计量； x_i 为密度估计的样本点； h 为带宽参数； n 为可供估计的数据个数； $K(\cdot)$ 是核函数，在此选用高斯核函数：

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \quad (11)$$

得到统计量的密度分布后，计算统计量控制限为：

$$1 - \alpha = \int_0^L \hat{f}(x) dx \quad (12)$$

其中， α 为显著度， L 为控制限，即第 m 个模态的 Hotelling- T^2 和 SPE 的控制限 $C_{T^2}^{(m)}$ 、 $C_{SPE}^{(m)}$ 。

References

- 1 Yew Seng Ng, Rajagopalan Srinivasan. An adjoined multi-model approach for monitoring batch and transient operations. *Computers and Chemical Engineering*, 2009, **33**(4): 887–902.
- 2 Hwang D H, Han C. Real-time monitoring for a process with multiple operating modes. *Control Engineering Practice*, 1999, **7**(7): 891–902.
- 3 Lane, S.; Martin, E.B., Kooijmans, R., Morris, A.J. Performance monitoring of a multi-product semi-batch process. *Journal of Process Control*, 2001, **11**(1): 1–11.
- 4 Kosanovich K. A., Piovoso M. J. and Dahl K. S. Multiway PCA applied to an industrial batch process. *Process of American Control Conference*, 1994, (2): 1294–1298.
- 5 Kosanovich K.A., Dahl K.S. and Piovoso M.J. Improved process understanding using multiway principal component analysis. *Industrial & Engineering Chemistry Research*, 1996, **35**(1): 138–146.
- 6 Dong D., McAvoy T.J. Multi-stage Batch Process Monitoring. *Process of American Control Conference*, 1995, (3): 1857–1861.
- 7 Anshuman Bhagwat, Rajagopalan Srinivasan, P. R. Krishnaswamy. Multi-linear model-based fault detection during process transitions. *Chemical Engineering Science*, 2003, **58**(9): 1649–1670.

- 8 Yew Seng Nga, Rajagopalan Srinivasan. An adjoined multi-model approach for monitoring batch and transient operations. *Computers and Chemical Engineering*, 2009,**33**(4):887–902.
- 9 Xuan-Tien Doan, Rajagopalan Srinivasan. Online monitoring of multi-phase batch processes using phase-based multivariate statistical process control. *Computers and Chemical Engineering*, 2008,**32**(1–2):230–243.
- 10 Chen J, Liu J. Mixture principal component analysis models for process monitoring. *Industrial & Engineering Chemistry Research*, 1999,**38**(4):1478–1488.
- 11 Zhao S J, Zhang J, Xu Y M. Monitoring of processes with multiple operating modes through multiple principal component analysis models. *Industrial & Engineering Chemistry Research*, 2004,**43**(22):7025–7035.
- 12 Shi Jian Zhao, Jie Zhang, and Yong Mao Xu. Performance monitoring of processes with multiple operating modes through multiple PLS models. *Journal of Process Control*, 2006,**16**(7):763–772.
- 13 Choi SW, Park JH, Lee IB. Process monitoring using a Gaussian mixture model via principal component analysis and discriminant analysis. *Computers and Chemical Engineering*, 2004,**28**(8):1377–1387.
- 14 Yoo CK, Villez K, Lee IB, Rose 'n C, Vanrolleghem PA. Multimodel statistical process monitoring and diagnosis of a sequencing batch reactor. *Biotechnol Bioeng*, 2007,**96**(4):687–701.
- 15 Thissen U, Swierenga H, deWeijer AP, Wehrens R, Melssen WJ, Buydens LMC. Multivariate statistical process control using mixture modeling. *Journal of Chemometrics*, 2005,**19**(1):23–31.
- 16 Hyndman RJ. Computing and graphing highest density regions. *Am Stat*, 1996,**50**(2):120–126.
- 17 Jie Yu, S. Joe Qin. Multimode Process Monitoring with Bayesian Inference-Based Finite Gaussian Mixture Models. *AICHE*, 2008,**54**(7):1811–1829.
- 18 H. Henry Yue, S. Joe Qin. Reconstruction-Based Fault Identification Using a Combined Index. *Industrial & Engineering Chemistry Research*, 2001,**40**(20):4403–4414.
- 19 Lu, N.Y., Gao, F.R. and Wang F.L. A sub-PCA modeling and on-line monitoring strategy for batch processes. *AICHE*, 2004,**50**(1):255–259.
- 20 Anshuman Bhagwat, Rajagopalan Srinivasan, P. R. Krishnaswamy. Multi-linear model-based fault detection during process transitions. *Chemical Engineering Science*, 2003,**58**(9):1649–1670.
- 21 Zhao C H, Wang F L, Lu N Y, Jia M X. Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes. *Journal of Process Control*, 2007,**17**(9):728–741.
- 22 Yuan Yao, Furong Gao. Phase and transition based batch process modeling and online monitoring. *Journal of Process Control*, 2009,**19**(5):816–826.
- 23 Ningyun Lu, Fuli Wang, Furong Gao, Shu Wang. Statistical modeling and online monitoring for batch processes. *Acta Automatica Sinica*, 2006,**32**(3):400–410.
- (陆宁云, 王福利, 高福荣, 王妹. 间歇过程的统计建模与在线监测. *自动化学报*, 2006,**32**(3):400–410.)

- 24 Carlos F. Alcala, S. Joe Qin. Reconstruction-based contribution for process monitoring. *Automatica*, 2009,**45**(7):1593–1600.



谭帅 东北大学自动化研究所博士研究生. 主要研究方向为基于多元统计技术的过程监测、故障诊断及质量预测. 本文通信作者.

E-mail: tanshuai_neu@yahoo.cn

(TAN Shuai Ph.D. candidate at Northeastern University. Her research interest covers multivariate statistical process monitoring, fault diagnosis, and quality prediction. Corresponding author of this paper.)



王福利 东北大学教授. 主要研究方向为复杂工业过程建模与优化、过程监测.

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

(WANG Fu-Li Professor at Northeastern University. His research interest covers modeling and optimization of complex system, and process monitoring.)



常玉清 东北大学副教授. 主要研究方向为复杂工业过程建模, 过程监测及质量预测方法的研究与应用.

E-mail: changyuqing@mail.neu.edu.cn

(CHANG Yu-Qing Associate professor at Northeastern University. Her research interests covers process modeling, process monitoring and quality prediction and their applications in industry process.)



王妹 东北大学控制理论与控制工程博士生. 主要研究方向为统计过程建模、过程监测、故障诊断方法的研究与应用.

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

(WANG Shu Ph.D. candidate at Northeastern University. Her research interests covers multivariate statistical modeling, process monitoring, and fault diagnosis and their applications in industry process.)



周贺 辽河油田钻采工艺研究院中心试验所工程师, 从事采油工程专业研究.

E-mail: bluebaby-2@163.com

(ZHOU He Engineer in Drilling and Production Technology Research Institute of Liaohe Oilfield Company. Her research interest covers oil production engineering research.)