

一种基于 DIOCVA 的过程监控方法

曹玉苹¹ 黄琳哲² 田学民¹

摘要 传统基于典型变量分析的过程监控方法无法判断故障是否影响产品质量. 为此, 本文提出一种基于动态输入输出典型变量分析 (Dynamic input-output canonical variate analysis, DIOCVA) 的过程监控方法. 该方法利用典型变量分析提取数据之间的相关性, 并进一步考虑方差信息和时序相关性, 将过程数据和质量数据映射到 5 个子空间: 输入输出相关子空间, 不相关输入主元子空间, 不相关输入残差子空间, 不相关输出主元子空间和不相关输出残差子空间. 所提方法能够精细区分影响质量的过程故障和不影响质量的过程故障. 以 Tennessee Eastman 过程为例对所提方法的有效性进行了验证.

关键词 过程监控, 质量监控, 故障检测, 典型变量分析, 方差

引用格式 曹玉苹, 黄琳哲, 田学民. 一种基于 DIOCVA 的过程监控方法. 自动化学报, 2015, 41(12): 2072–2080

DOI 10.16383/j.aas.2015.c150058

A Process Monitoring Method Using Dynamic Input-output Canonical Variate Analysis

CAO Yu-Ping¹ HUANG Lin-Zhe² TIAN Xue-Min¹

Abstract Traditional process monitoring methods based on canonical variate analysis cannot judge whether faults affect product quality or not. Therefore, a process monitoring method using dynamic input-output canonical variate analysis is proposed. Canonical variate analysis is used to analyze correlations between data. Furthermore, variance information and serial correlations in data are taken into account. Process data and quality data are projected to five subspaces: input-output correlated subspace, uncorrelated input principal subspace, uncorrelated input residual subspace, uncorrelated output principal subspace and uncorrelated output residual subspace. The proposed method can distinguish between the faults that affect product quality and those that do not. The effectiveness of the proposed method is demonstrated through a case study of the Tennessee Eastman process.

Key words Process monitoring, quality monitoring, fault detection, canonical variate analysis, variance

Citation Cao Yu-Ping, Huang Lin-Zhe, Tian Xue-Min. A process monitoring method using dynamic input-output canonical variate analysis. *Acta Automatica Sinica*, 2015, 41(12): 2072–2080

工业过程日益复杂, 研究过程监控方法及时有效地发现过程中存在的异常状况具有重要意义^[1–5]. 基于数据驱动的过程监控方法^[6–10]主要包括主元分析 (Principal component analysis, PCA)^[11]、独立元分析 (Independent component analysis, ICA) 和典型变量分析 (Canonical vari-

ate analysis, CVA)^[12–14]等. 在传统方法的基础上, 研究者围绕非线性^[15–18]、时序相关性^[19–20]、离群点^[21–22]等问题提出了改进方法. 过程测量数据在线采集, 数据量较大; 而质量测量数据往往离线采集, 采集频率较慢且存在时间滞后. 上述数据驱动方法在进行过程监控时, 仅能判断过程测量变量是否异常, 无法判断产品质量变量是否异常^[23]. 如果产品质量没有发生异常, 报警常常被视为误报, 这在一定程度上降低了故障检测的可信度.

实际生产过程中故障的存在形式可以分为三种, 第一种是过程测量变量发生异常, 并进一步导致与之关联的质量测量变量发生异常; 第二种是只有过程测量变量发生异常, 由于控制器的抑制作用, 质量测量变量没有受到影响; 第三种是只有质量测量变量发生异常, 过程测量变量正常. 因此, 过程测量数据空间和质量测量数据空间可以依据它们之间的相关性, 各自分解为相关子空间与不相关子空间. 如何对这些子空间进行精细监控值得深入研究. 基于偏最小二乘 (Partial least squares or projection to

收稿日期 2015-01-30 录用日期 2015-10-19
Manuscript received January 30, 2015; accepted October 19, 2015

国家自然科学基金 (61273160, 61403418), 山东省自然科学基金 (ZR2014FL016), 中央高校基本科研业务费专项资金 (14CX02174A) 资助

Supported by National Natural Science Foundation of China (61273160, 61403418), Natural Science Foundation of Shandong Province (ZR2014FL016) and the Fundamental Research Funds for the Central Universities (14CX02174A)

本文责任编辑 钟麦英

Recommended by Associate Editor ZHONG Mai-Ying

1. 中国石油大学 (华东) 信息与控制工程学院 青岛 266580 2. 中国石油工程建设公司华东设计分公司 青岛 266071

1. College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580 2. East-China Design Branch, China Petroleum Engineering and Construction Corporation, Qingdao 266071

latent structures, PLS) 的方法利用质量测量数据引导过程测量数据空间的分解^[24], 其改进方法 Total-PLS (T-PLS) 方法^[25] 和 Concurrent PLS (CPLS) 方法^[23, 26-29] 改进了空间划分和统计量构造方式, 能同时监控过程测量数据和质量测量数据.

传统基于典型变量分析的故障检测方法最大化了过去数据与未来数据的相关性, 但是无法判断故障是否影响产品质量. 传统 CVA 故障检测方法使用 T_s^2 统计量、平方预测误差 (Squared prediction error, SPE) 统计量、 T_r^2 统计量进行监控. 过程数据残差空间可能包含变化较大的项, 这会影响 SPE 统计量的监控效果. T_r^2 统计量对协方差矩阵求逆计算比较敏感^[12]. 针对上述问题, 本文提出了基于动态输入输出 CVA 的故障精细检测方法. 该方法利用 CVA 算法对动态扩展后的输入输出数据进行相关分析, 将原始输入输出空间映射到输入输出相关子空间、不相关输入主元子空间、不相关输入残差子空间、不相关输出主元子空间和不相关输出残差子空间 5 个子空间, 对这 5 个子空间进行精细监控. 以 Tennessee Eastman 过程为对象对所提方法的有效性进行了验证.

1 传统基于 CVA 的故障检测方法

1.1 CVA 原理

典型变量分析又称典型相关分析 (Canonical correlation analysis, CCA), 是一种线性降维技术, 它将两个变量集的相关性最大化. 给定随机向量 $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$ 与 $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$, 寻找一对投影方向 $\mathbf{J}_1 = [J_{11}, J_{21}, \dots, J_{m1}]^T$ 与 $\mathbf{L}_1 = [L_{11}, L_{21}, \dots, L_{n1}]^T$, 使得投影 $c_1 = \mathbf{J}_1^T \mathbf{x}$ 与 $d_1 = \mathbf{L}_1^T \mathbf{y}$ 具有最大的相关性, 称这种相关为典型相关, c_1, d_1 为第一对典型变量. 通过求解优化问题

$$\begin{aligned} \max \rho(c_1, d_1) &= \frac{E(\mathbf{J}_1^T \mathbf{x} \mathbf{y}^T \mathbf{L}_1)}{\sqrt{E(\mathbf{J}_1^T \mathbf{x} \mathbf{x}^T \mathbf{J}_1) \cdot E(\mathbf{L}_1^T \mathbf{y} \mathbf{y}^T \mathbf{L}_1)}} \\ \text{s.t. } \text{var}(c_1) &= \text{var}(\mathbf{J}_1^T \mathbf{x}) = 1 \\ \text{var}(d_1) &= \text{var}(\mathbf{L}_1^T \mathbf{y}) = 1 \end{aligned} \quad (1)$$

实现典型变量提取. 继而可以由 \mathbf{x}, \mathbf{y} 出发, 寻找第二对典型变量 c_2, d_2 , 使其与第一对典型变量 c_1, d_1 不相关, 且 c_2 与 d_2 之间具有最大相关性, 如此下去, 使 \mathbf{x} 与 \mathbf{y} 的相关性特征提取完毕为止. 上述优化问题可使用广义奇异值分解实现^[12], 设 \mathbf{x}, \mathbf{y} 的协方差矩阵分别为 Σ_{xx}, Σ_{yy} , 互协方差矩阵为 Σ_{xy} , 则存在

矩阵 $\mathbf{J} \in \mathbf{R}^{m \times m}$ 和 $\mathbf{L} \in \mathbf{R}^{n \times n}$ 满足^[13]:

$$\begin{cases} \mathbf{J} \Sigma_{xx} \mathbf{J}^T = \mathbf{I}_{\bar{m}}, \mathbf{L} \Sigma_{yy} \mathbf{L}^T = \mathbf{I}_{\bar{n}} \\ \mathbf{J} \Sigma_{xy} \mathbf{L}^T = \mathbf{D} = \text{diag}\{\gamma_1, \dots, \gamma_r, 0, \dots, 0\} \end{cases} \quad (2)$$

其中 $\gamma_1 \geq \dots \geq \gamma_r$, $\bar{m} = \text{rank}(\Sigma_{xx})$, $\bar{n} = \text{rank}(\Sigma_{yy})$, \mathbf{D} 包含典型相关系数 γ_i , $\mathbf{I}_{\bar{m}} \in \mathbf{R}^{m \times m}$ 是前 \bar{m} 个对角元素是 1 而其余对角元素均为 0 的对角阵, $\mathbf{I}_{\bar{n}} \in \mathbf{R}^{n \times n}$ 是前 \bar{n} 个对角元素是 1 而其余对角元素均为零的对角阵. 典型向量 $\mathbf{c} = \mathbf{J} \mathbf{x}$ 包含一系列不相关的随机变量, 且有协方差矩阵: $\Sigma_{cc} = \mathbf{J} \Sigma_{xx} \mathbf{J}^T = \mathbf{I}_{\bar{m}}$; 典型向量 $\mathbf{d} = \mathbf{L} \mathbf{y}$ 包含一系列不相关的随机变量, 且有协方差矩阵: $\Sigma_{dd} = \mathbf{L} \Sigma_{yy} \mathbf{L}^T = \mathbf{I}_{\bar{n}}$; \mathbf{c} 和 \mathbf{d} 之间的互协方差矩阵是一个对角阵: $\Sigma_{cd} = \mathbf{J} \Sigma_{xy} \mathbf{L}^T = \mathbf{D}$. 通过广义奇异值分解:

$$\Sigma_{xx}^{-\frac{1}{2}} \Sigma_{xy} \Sigma_{yy}^{-\frac{1}{2}} = \mathbf{U} \Sigma \mathbf{V}^T \quad (3)$$

得出

$$\mathbf{J} = \mathbf{U}^T \Sigma_{xx}^{-\frac{1}{2}}, \mathbf{L} = \mathbf{V}^T \Sigma_{yy}^{-\frac{1}{2}}, \mathbf{D} = \Sigma \quad (4)$$

1.2 故障检测方法

传统基于 CVA 的故障检测方法通过最大化过去数据空间与未来数据空间的相关性, 提取出具有最佳预测能力的低维典型变量, 对低维典型变量进行统计监控. 假设过程输入输出向量为 $\mathbf{u}_t \in \mathbf{R}^{m_u}$, $\mathbf{y}_t \in \mathbf{R}^{m_y}$, 在某时刻 t , 构造过去信息向量 $\mathbf{p}_t = [\mathbf{y}_{t-1}^T, \mathbf{y}_{t-2}^T, \dots, \mathbf{y}_{t-h}^T, \mathbf{u}_{t-1}^T, \mathbf{u}_{t-2}^T, \dots, \mathbf{u}_{t-h}^T]^T$, 以及包含现在和未来输出信息的向量 $\mathbf{f}_t = [\mathbf{y}_t^T, \mathbf{y}_{t+1}^T, \dots, \mathbf{y}_{t+l}^T]^T$, 其中 h 和 l 表示滞后阶次. 通过典型变量分析可得过程状态的 CVA 估计

$$\mathbf{x}_t^c = \mathbf{J}_k \mathbf{p}_t = \mathbf{U}_k^T \Sigma_{pp}^{-\frac{1}{2}} \mathbf{p}_t \quad (5)$$

式中 k 为模型阶次, \mathbf{J}_k 为矩阵 \mathbf{J} 的前 k 行, \mathbf{U}_k 为矩阵 \mathbf{U} 的前 k 列, Σ_{pp} 为向量 \mathbf{p}_t 的协方差矩阵.

利用统计量 T_s^2, T_r^2 和 SPE

$$\begin{cases} T_s^2 = \mathbf{p}_t^T \mathbf{J}_k^T \mathbf{J}_k \mathbf{p}_t \\ T_r^2 = \mathbf{p}_t^T \mathbf{J}_q^T \mathbf{J}_q \mathbf{p}_t \\ SPE = \mathbf{p}_t^T (\mathbf{I} - \mathbf{J}_k^T \mathbf{J}_k)^T (\mathbf{I} - \mathbf{J}_k^T \mathbf{J}_k) \mathbf{p}_t \end{cases} \quad (6)$$

监控过程的变化趋势. 式 (6) 中, \mathbf{J}_q 为矩阵 \mathbf{J} 的后 $q = h(m_u + m_y) - k$ 行. T_s^2 统计量对 CVA 状态空间内的变化进行度量. T_r^2 统计量对 CVA 状态空间外的变化进行度量, 其检测效果最好. 但是当协方差矩阵 Σ_{pp} 含有较小值时, T_r^2 统计量受协方差矩阵求逆计算影响较大^[12]. SPE 统计量为平方预测误差, 可以用来监控残差空间的变化. 对于 T_r^2 和 SPE

统计量监控的两个数据空间, 虽然变量对未来输出向量的预测能力较小, 但是变量变化可能较大, 利用 T_r^2 和 SPE 统计量是不恰当的.

2 基于 DIOCVA 的过程监控

在实际生产过程中, 有的过程测量变量异常影响与之关联的质量测量变量发生异常; 有的过程测量变量异常被控制器抑制, 质量测量变量没有受到影响; 有的故障只有质量测量变量发生异常, 过程测量变量正常. 按照过程测量变量和质量测量变量的关系对过程故障进行区分, 能够更加精细地进行故障检测, 以满足多种过程故障检测需求. 本文利用典型变量分析将过程测量数据空间划分为与质量测量变量相关的子空间和与质量测量变量不相关的子空间, 将质量测量数据空间划分为与过程测量变量相关的子空间和与过程测量变量不相关的子空间, 对这些子空间进行精细监控, 并进一步考虑测量变量的时序相关性, 提出基于动态输入输出 CVA (Dynamic input-output CVA, DIOCVA) 的过程监控方法.

2.1 输入输出 CVA 方法

定义过程测量向量 $\mathbf{x}' = [x'_1, x'_2, \dots, x'_{m'}]^T$ 为输入向量, 质量测量向量 $\mathbf{y}' = [y'_1, y'_2, \dots, y'_{n'}]^T$ 为输出向量, 对 s 次测量数据进行标准化处理使之均值为 0 方差为 1, 得到过程测量数据矩阵 $X' \in \mathbf{R}^{s \times m'}$ 和质量测量数据矩阵 $Y' \in \mathbf{R}^{s \times n'}$. 对输入输出数据进行典型变量分析可得相关性最大的前 k 对典型向量 \mathbf{c}' 和 \mathbf{d}' ,

$$\mathbf{c}' = [c'_1, c'_2, \dots, c'_k]^T = U_k^T \Sigma_{x'x'}^{-\frac{1}{2}} \mathbf{x}' \quad (7)$$

$$\mathbf{d}' = [d'_1, d'_2, \dots, d'_k]^T = V_k^T \Sigma_{y'y'}^{-\frac{1}{2}} \mathbf{y}' \quad (8)$$

由 \mathbf{c}' 和 \mathbf{d}' 张成的子空间称为典型变量子空间. \mathbf{c}' 返回到 m' 维测量空间的投影为 $\hat{\mathbf{x}}'$, \mathbf{d}' 返回到 n' 维测量空间的投影为 $\hat{\mathbf{y}}'$,

$$\hat{\mathbf{x}}' = \Sigma_{x'x'}^{-\frac{1}{2}} U_k U_k^T \Sigma_{x'x'}^{-\frac{1}{2}} \mathbf{x}' \quad (9)$$

$$\hat{\mathbf{y}}' = \Sigma_{y'y'}^{-\frac{1}{2}} V_k V_k^T \Sigma_{y'y'}^{-\frac{1}{2}} \mathbf{y}' \quad (10)$$

则估计残差 $\mathbf{e}_{x'} = \mathbf{x}' - \hat{\mathbf{x}}'$, $\mathbf{e}_{y'} = \mathbf{y}' - \hat{\mathbf{y}}'$, 对应的数据矩阵形式为

$$E_{x'} = X' - \hat{X}' \quad (11)$$

$$E_{y'} = Y' - \hat{Y}' \quad (12)$$

其中 $E_{x'} \in \mathbf{R}^{s \times m'}$, $E_{y'} \in \mathbf{R}^{s \times n'}$ 表示残差矩阵, $\hat{X}' \in \mathbf{R}^{s \times m'}$, $\hat{Y}' \in \mathbf{R}^{s \times n'}$ 分别表示典型向量 \mathbf{c}' 和 \mathbf{d}' 返回相应测量空间的投影矩阵. 由 $\mathbf{e}_{x'}$ 和 $\mathbf{e}_{y'}$ 张

成的子空间称为典型变量残差子空间. 为了以示区分, 称 \mathbf{c}' 和 \mathbf{d}' 张成的子空间为输入输出相关子空间 $\Phi_{x'y'}$, $\mathbf{e}_{x'}$ 张成的子空间为不相关输入子空间 $\Phi_{e_{x'}}$, $\mathbf{e}_{y'}$ 张成的子空间为不相关输出子空间 $\Phi_{e_{y'}}$.

输入输出相关子空间 $\Phi_{x'y'}$ 描述了过程变量和质量变量相关的信息, 对该子空间构造 T^2 统计量:

$$T_s^2 = \mathbf{c}'^T \mathbf{c}' \quad (13)$$

假设正态性成立, 则 T_s^2 服从如下分布

$$T_{s,\alpha}^2 = \frac{k(s^2 - 1)}{s(s - k)} F_{\alpha}(k, s - k) \quad (14)$$

选定置信水平 $100(1 - \alpha)\%$, 利用式 (14) 计算 T_s^2 统计的阈值. T_s^2 统计量对输入输出相关子空间中的变化进行度量, 统计量超出阈值表明与产品质量向量相关的过程测量向量的变化是异常的 (即过程存在故障), 进而可判断产品质量的变化出现异常, 即过程故障影响产品质量发生变化. 过程测量变量的采集速率一般比产品质量变量的采集速率快, 由式 (7) 和 (13) 可以看出 T_s^2 统计量只利用了过程测量变量, 因此其在线监控的实时性优于利用产品质量的统计量.

对于不相关输入子空间 $\Phi_{e_{x'}}$, 传统 CVA 监控方法利用式 (6) 中的 T_r^2 统计量和 SPE 统计量进行监控. 当协方差矩阵 $\Sigma_{x'x'}$ 含有较小值时, T_r^2 统计量对 $\Sigma_{x'x'}$ 的求逆计算过度敏感. 在利用典型变量分析方法进行数据分析时, 并没有考虑变量方差的大小. 通过算法研究发现, 在子空间 $\Phi_{e_{x'}}$ 中可能存在方差较大的项, 这会影响 SPE 统计量的监控效果. 针对上述问题, 本文利用主元分析方法对不相关输入子空间 $\Phi_{e_{x'}}$ 进行方差分析, 从而形成不相关输入主元子空间 $\Phi_{e_{x'}}^P$ 和不相关输入残差子空间 $\Phi_{e_{x'}}^R$, 对子空间 $\Phi_{e_{x'}}^P$ 构建 T^2 统计量, 对子空间 $\Phi_{e_{x'}}^R$ 构建 SPE 统计量进行监控.

对不相关输入子空间 $\Phi_{e_{x'}}$ 的数据矩阵 $E_{x'}$ 进行主元分析^[1, 12]

$$E_{x'} = \hat{E}_{x'} + E_{x'}^R = T_{x'} P_{x'}^T + E_{x'}^R \quad (15)$$

假设主元个数为 $k_{x'}$, 则 $P_{x'} \in \mathbf{R}^{m' \times k_{x'}}$ 表示载荷矩阵, $T_{x'} = E_{x'} P_{x'}$ 表示主元得分矩阵, $\hat{E}_{x'}$ 表示 $T_{x'}$ 返回不相关输入子空间 $\Phi_{e_{x'}}$ 的投影, $E_{x'}^R$ 表示主元残差矩阵. $\hat{E}_{x'}$ 张成的子空间为不相关输入主元子空间 $\Phi_{e_{x'}}^P$, 得分向量 $\mathbf{t}_{x'} = P_{x'}^T \mathbf{e}_{x'}$, 构造统计量:

$$T_{x'}^2 = \mathbf{t}_{x'}^T \Lambda_{x'}^{-1} \mathbf{t}_{x'} \quad (16)$$

进行监控, 其中 $\Lambda_{x'}$ 表示 $k_{x'}$ 个主元的特征值组成的对角矩阵. $E_{x'}^R$ 张成的子空间为不相关输入残差子

空间 $\Phi_{e_{x'}}^R$, 残差向量 $\mathbf{e}_{x'}^R = \mathbf{e}_{x'} - \hat{\mathbf{e}}_{x'} = \mathbf{e}_{x'} - P_{x'}\mathbf{t}_{x'}$, 构造统计量

$$SPE_{x'} = (\mathbf{e}_{x'} - \hat{\mathbf{e}}_{x'})^T (\mathbf{e}_{x'} - \hat{\mathbf{e}}_{x'}) \quad (17)$$

进行监控. 统计量 $T_{x'}^2$ 度量了与输出不相关的输入子空间中的系统变化, 统计量 $SPE_{x'}$ 度量了同一子空间中的随机变化. 统计量 $T_{x'}^2$ 或 $SPE_{x'}$ 超出阈值表明与质量向量不相关的过程测量向量的变化是异常的 (即过程存在故障), 该故障不会影响产品质量, 如故障被控制器抑制. 产品质量是否异常需根据其其他统计量判断.

不相关输出子空间 $\Phi_{e_{y'}}$, 在进行监控时, 存在与不相关输入子空间相同的问题. 因此, 本文同样采用主元分析方法来分析不相关输出子空间 $\Phi_{e_{y'}}$ 的方差信息, 从而形成不相关输出主元子空间 $\Phi_{e_{y'}}^P$ 和不相关输出残差子空间 $\Phi_{e_{y'}}^R$, 对子空间 $\Phi_{e_{y'}}^P$ 构建 T^2 统计量, 对子空间 $\Phi_{e_{y'}}^R$ 构建 SPE 统计量进行监控.

对不相关输出子空间 $\Phi_{e_{y'}}$ 的数据矩阵 $E_{y'}$ 进行主元分析

$$E_{y'} = \hat{E}_{y'} + E_{y'}^R = T_{y'} P_{y'}^T + E_{y'}^R \quad (18)$$

假设主元个数为 $k_{y'}$, 则 $P_{y'} \in \mathbf{R}^{n' \times k_{y'}}$ 表示载荷矩阵, $T_{y'} = E_{y'} P_{y'}$ 表示主元得分矩阵, $\hat{E}_{y'}$ 表示 $T_{y'}$ 返回不相关输出子空间 $\Phi_{e_{y'}}$ 的投影, $E_{y'}^R$ 表示主元残差矩阵. $\hat{E}_{y'}$ 张成的子空间为不相关输出主元子空间 $\Phi_{e_{y'}}^P$, 得分向量 $\mathbf{t}_{y'} = P_{y'}^T \mathbf{e}_{y'}$, 构造统计量:

$$T_{y'}^2 = \mathbf{t}_{y'}^T \Lambda_{y'}^{-1} \mathbf{t}_{y'} \quad (19)$$

进行监控, 其中 $\Lambda_{y'}$ 表示 $k_{y'}$ 个主元的特征值组成的对角矩阵. $E_{y'}^R$ 张成的子空间为不相关输出残差子空间 $\Phi_{e_{y'}}^R$, 残差向量 $\mathbf{e}_{y'}^R = \mathbf{e}_{y'} - \hat{\mathbf{e}}_{y'} = \mathbf{e}_{y'} - P_{y'}\mathbf{t}_{y'}$, 构造统计量:

$$SPE_{y'} = (\mathbf{e}_{y'} - \hat{\mathbf{e}}_{y'})^T (\mathbf{e}_{y'} - \hat{\mathbf{e}}_{y'}) \quad (20)$$

进行监控. 统计量 $T_{y'}^2$ 度量了与输入不相关的输出子空间中的系统变化, $SPE_{y'}$ 统计量对同一子空间中的随机变化进行度量. 统计量 $T_{y'}^2$ 或 $SPE_{y'}$ 超出阈值表明与过程测量向量不相关的质量测量数据子空间的变化是异常的, 该质量异常与过程测量数据无关, 如引起该质量异常的关键变量未测量.

综上, 数据空间 X' 和 Y' 可精细分解为

$$\begin{cases} X' = \hat{X}' + E_{x'} = \hat{X}' + \hat{E}_{x'} + E_{x'}^R \\ Y' = \hat{Y}' + E_{y'} = \hat{Y}' + \hat{E}_{y'} + E_{y'}^R \end{cases} \quad (21)$$

各子空间及相应监控统计量见图 1, 统计量的控制限可由核密度估计方法^[30] 求取. 过程变量的采集速率一般较快, 每获得一个时刻的过程数据, 利用过程变量监控输入输出相关子空间 $\Phi_{x'y'}$ 、不相关输入子空间 $\Phi_{e_{x'}}$ 的变化. 如果统计量 T_s^2 超出控制限, 表明过程故障对产品质量的影响较大, 应予以重视. 如果统计量 T_s^2 正常, 不相关输入子空间 $\Phi_{e_{x'}}$ 的统计量 $T_{x'}^2$ 或 $SPE_{x'}$ 超出控制限, 表明过程发生故障, 但该故障对产品质量没有影响或者影响很小. 如果质量数据可得, 再利用质量数据监控不相关输出主元子空间 $\Phi_{e_{y'}}^P$ 和不相关输出残差子空间 $\Phi_{e_{y'}}^R$ 的变化, 验证实时过程监控的判断结果是否正确.

2.2 动态输入输出 CVA 方法

实际过程往往存在动态特性, 过程测量数据在线采集, 采样时间间隔较短, 相邻采样时刻的数据往往具有相关性. 输入输出 CVA 方法在分析相关性时仅考虑了变量间的相关性, 并没有考虑变量的时序相关性. 因此, 对 IOCVA 方法进一步改进, 将过程测量数据按照式 (22) 进行动态增广, 便可以得到动态输入输出 CVA 方法. 由于质量测量数据的采样时间间隔一般较长, 相邻采样时刻的数据相关性很小, 可以忽略. 因此, 不对质量测量数据进行增广.

$$X'(h') = [X'(t') \ X'(t' - 1) \ \dots \ X'(t' - h')] =$$

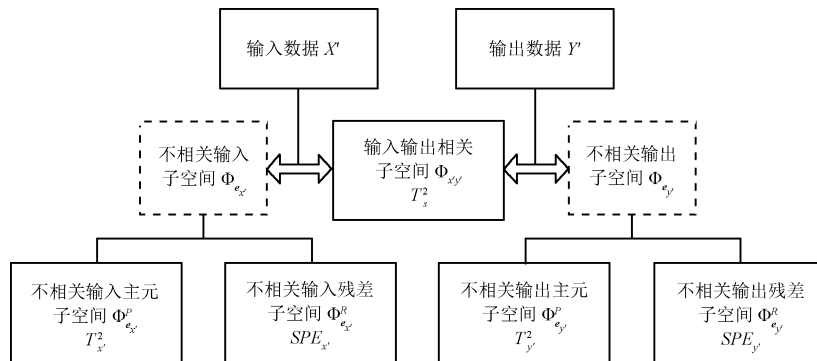


图 1 IOCVA 方法的各子空间及监控统计量

Fig. 1 Subspaces and monitoring statistics of the IOCVA method

$$\begin{bmatrix} \mathbf{x}'_{t'}^T & \mathbf{x}'_{t'-1}^T & \cdots & \mathbf{x}'_{t'-h'}^T \\ \mathbf{x}'_{t'-1}^T & \mathbf{x}'_{t'-2}^T & \cdots & \mathbf{x}'_{t'-h'-1}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}'_{t'+h'-s}^T & \mathbf{x}'_{t'+h'-s-1}^T & \cdots & \mathbf{x}'_{t'-s}^T \end{bmatrix} \quad (22)$$

其中, t' 为当前时刻, h' 为滞后阶次, h' 可利用动态 PCA 滞后阶次的求解方法^[19] 确定. 然后即可按照 IOCVA 方法的步骤进行故障检测.

3 仿真研究

田纳西-伊斯曼 (Tennessee-Eastman, TE) 过程^[1,31] 由伊斯曼化学品公司创建, 广泛用于测试和比较过程控制、故障检测方法的性能. 过程包括五个主要单元: 反应器、冷凝器、压缩机、分离器和汽提塔; 包含八种成分: A、B、C、D、E、F、G 和 H. 反应物 A、C、D 和 E 以及惰性组分 B 一起进入反应器形成产品 G 和 H 以及副产品 F. 过程包括 41 个测量变量和 12 个控制变量, 所有过程测量值均包含高斯噪声. 过程仿真包括 21 个预设定的故障, 故障从第 161 个采样时刻加入. 仿真数据可从 <http://brahms.scs.uiuc.edu> 下载. 分别利用 IOCVA 方法、DIOCVA 方法、CVA 方法、PCA 方法和动态 PCA (Dynamic PCA, DPCA) 方法进行故障检测. 对于 IOCVA 方法和 DIOCVA 方法, 选择 TE 过程测量变量 XMEAS(1~36) 和控制变量 XMV(1~11) 作为输入向量 \mathbf{x}' , 测量变量 XMEAS(37~41) 作为输出向量 \mathbf{y}' . 对于 CVA 方法, 选择测量变量 XMEAS(1~36) 和控制变量 XMV(1~11) 作为输入向量 \mathbf{u} , 测量变量 XMEAS(37~41) 作为输出向量 \mathbf{y} . 考虑到最终产品质量变量测量实时性差, 为了实现过程监控的实时性, PCA 方法和 DPCA 方法利用测量变量 XMEAS(1~36) 和控制变量 XMV(1~11) 进行故障检测^[23].

故障 4 是反应器冷却水入口温度阶跃变化, 5 种方法的故障检测效果见图 2~图 8. 故障发生时, 反应器温度突然升高, 由于闭环控制进行补偿, 冷却水流速阶跃变化, 导致输出变量保持稳定. 图 2~图 4 显示 PCA、DPCA 和 CVA 可以检测到故障, 但无法进一步区分故障是否影响产品质量. 从图 5~图 8 的 IOCVA 和 DIOCVA 检测图可以看出, 输入输出相关子空间统计量 T_s^2 没有检测出故障, 不相关输入子空间的统计量 $T_{x'}^2$ 和 $SPE_{x'}$ 检测出故障, 表明故障仅存在于与输出不相关的输入子空间中, 故障受到反馈抑制, 没有影响最终产品的质量.

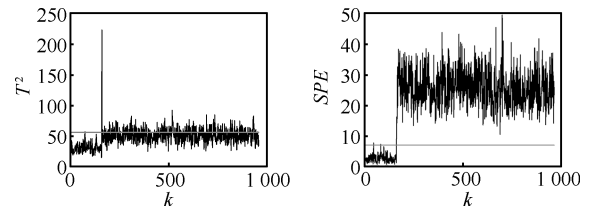


图 2 故障 4 的 PCA 监控图

Fig. 2 PCA-based monitoring results for Fault 4

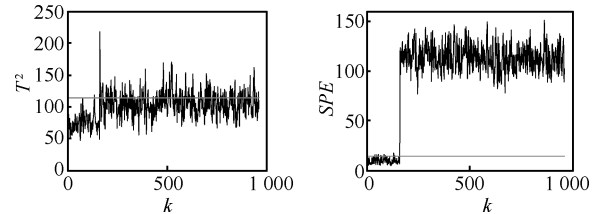


图 3 故障 4 的 DPCA 监控图

Fig. 3 DPCA-based monitoring results for Fault 4

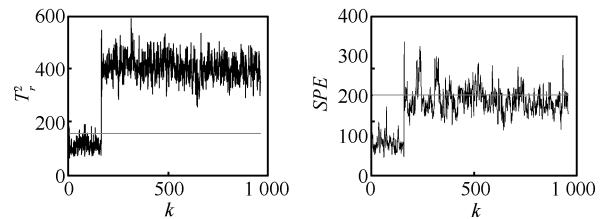
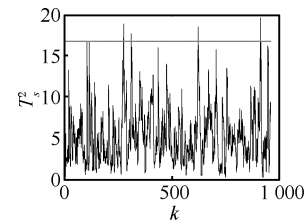


图 4 故障 4 的 CVA 监控图

Fig. 4 CVA-based monitoring results for Fault 4

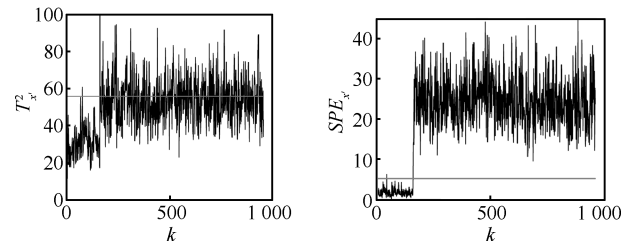
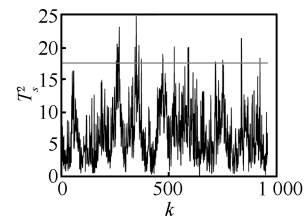


图 5 故障 4 利用新过程数据的 IOCVA 监控图

Fig. 5 IOCVA-based monitoring results for Fault 4 using new process data

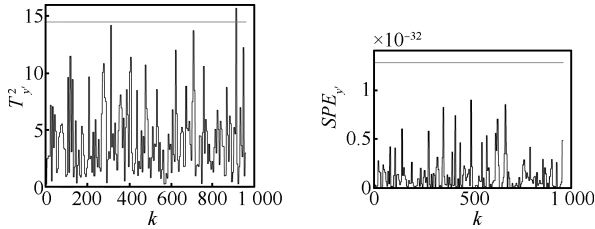


图 6 故障 4 利用新质量数据的 IOCVA 监控图

Fig.6 IOCVA-based monitoring results for Fault 4 using new quality data

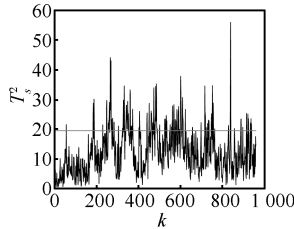


图 7 故障 4 利用新过程数据的 DIOCVA 监控图

Fig.7 DIOCVA-based monitoring results for Fault 4 using new process data

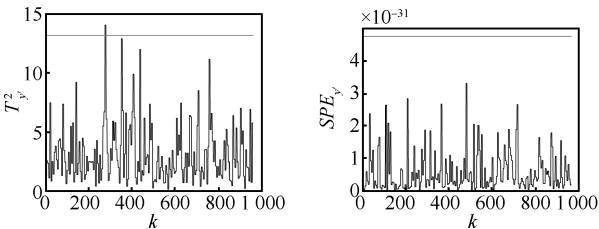


图 8 故障 4 利用新质量数据的 DIOCVA 监控图

Fig.8 DIOCVA-based monitoring results for Fault 4 using new quality data

故障 7 为成分 C 可用性的阶跃变化. 5 种方法对故障 7 的监控结果见图 9~图 15. 可以看出, 5 种方法都能检测出故障. 在 PCA、DPCA 和 CVA 方法的监控结果图 9~图 11 中, 可以看出故障发生一段时间之后统计量超限程度明显减小, 可是由于检测的不够精细, 无法确定产品质量是否受到影响. 进一步分析 IOCVA/DIOCVA 的监控结果图 12~图 15 可以发现, 故障在第 500 个时刻以后仅存在于不相关输入子空间, 说明此故障经过控制系统一段时间的调节已经被补偿, 输出变量回归正常水平.

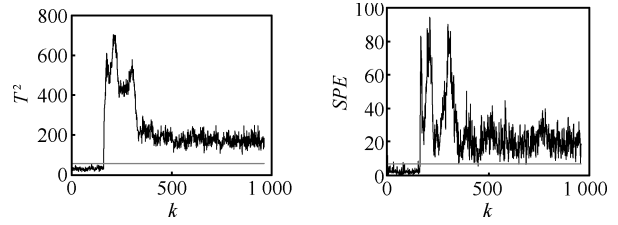


图 9 故障 7 的 PCA 监控图

Fig.9 PCA-based monitoring results for Fault 7

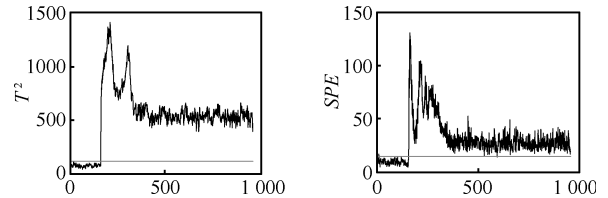


图 10 故障 7 的 DPCA 监控图

Fig.10 DPCA-based monitoring results for Fault 7

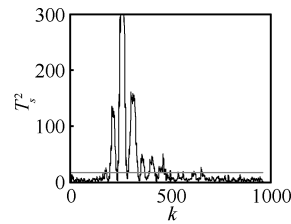


图 11 故障 7 的 CVA 监控图

Fig.11 CVA-based monitoring results for Fault 7

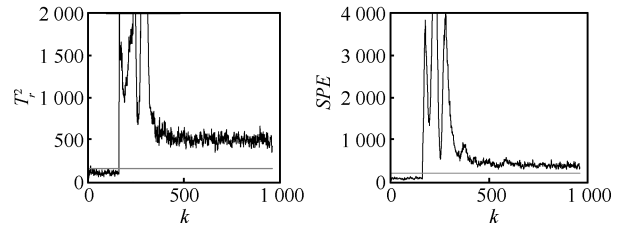
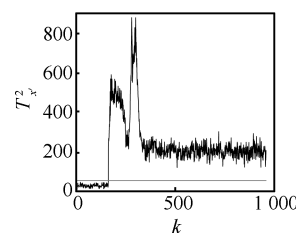
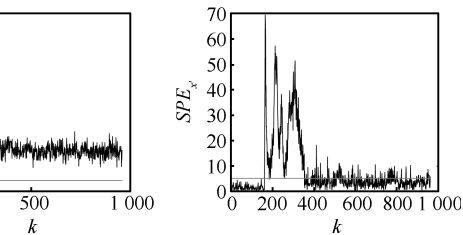
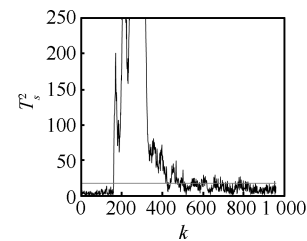


图 12 故障 7 利用新过程数据的 IOCVA 监控图

Fig.12 IOCVA-based monitoring results for Fault 7 using new process data



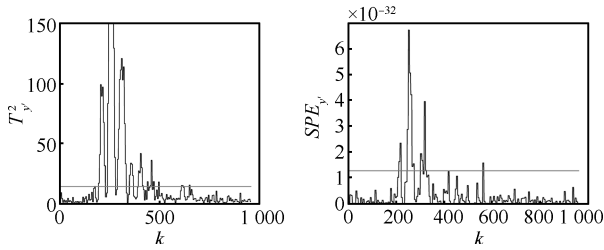


图 13 故障 7 利用新质量数据的 IOCVA 监控图

Fig. 13 IOCVA-based monitoring results for Fault 7

using new quality data

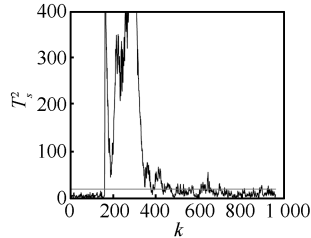


图 14 故障 7 利用新过程数据的 DIOCVA 监控图

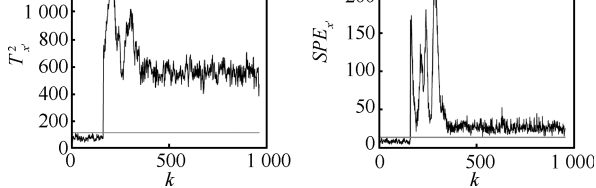


图 14 DIOCVA-based monitoring results for Fault 7

using new process data

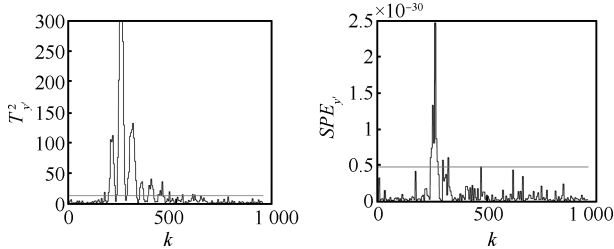


图 15 故障 7 利用新质量数据的 DIOCVA 监控图

Fig. 15 DIOCVA-based monitoring results for Fault 7

using new quality data

故障 13 由反应动力学常数慢偏移引发, 5 种方法的故障检测效果如图 16~图 22 所示, 可以看到 5 种方法都能够检测出故障. PCA 方法的故障检出率为 95.25%, 检出时刻为 201, 误报率为 0.625%; DPCA 的故障检出率为 95.75%, 检出时刻为 197, 误报率为 1.25%; CVA 的故障检出率为 95.88%, 检出时刻为 199, 误报率为 1.25%; IOCVA 的故障检出率为 95.50%, 检出时刻为 198, 误报率为 0; DIOCVA 的故障检出率为 95.88%, 检出时刻为 195, 误报率为 0.625%. 从检出率、检出时间和误报率的对比可以看出, IOCVA/DIOCVA 方法的检测效果优于 PCA、DPCA 和 CVA 方法. 通过

IOCVA/DIOCVA 监控结果可以看出输入输出相关子空间存在故障, 即存在导致质量异常的过程故障.

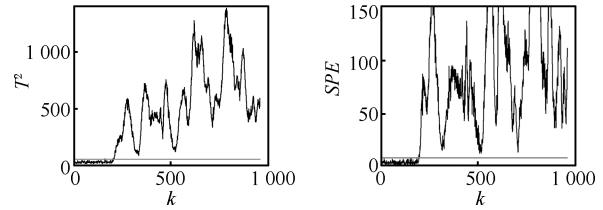


图 16 故障 13 的 PCA 监控图

Fig. 16 PCA-based monitoring results for Fault 13

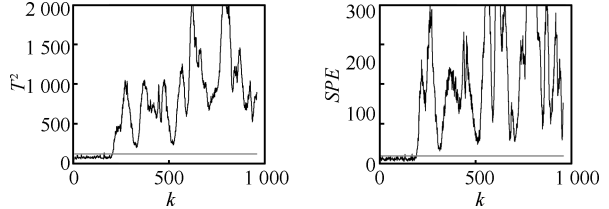


图 17 故障 13 的 DPCA 监控图

Fig. 17 DPCA-based monitoring results for Fault 13

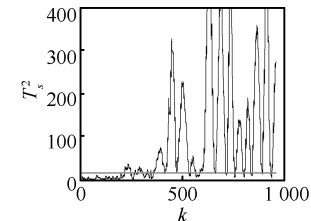


图 18 故障 13 的 CVA 监控图

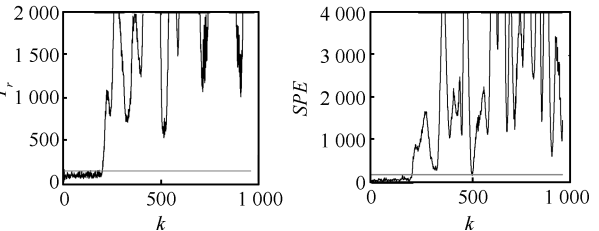


图 18 CVA-based monitoring results for Fault 13

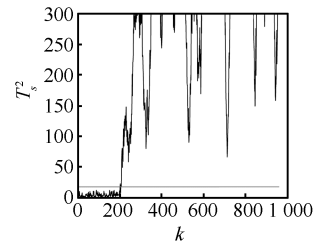


图 19 故障 13 利用新过程数据的 IOCVA 监控图

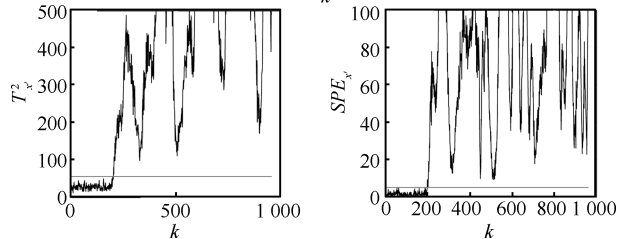


图 19 IOCVA-based monitoring results for Fault 13 using new process data

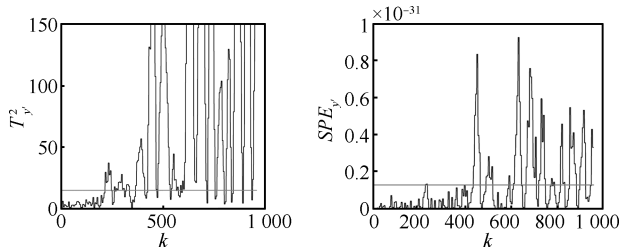


图 20 故障 13 利用新质量数据的 IOCVA 监控图

Fig. 20 IOCVA-based monitoring results for Fault 13 using new quality data

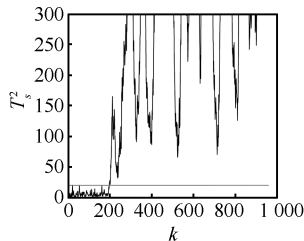


图 21 故障 13 利用新过程数据的 DIOCVA 监控图

Fig. 21 DIOCVA-based monitoring results for Fault 13 using new process data

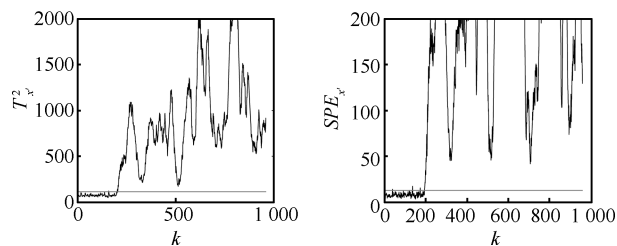


图 22 故障 13 利用新质量数据的 DIOCVA 监控图

Fig. 22 DIOCVA-based monitoring results for Fault 13 using new quality data

4 结论

本文提出的基于 DIOCVA 的过程监控方法利用 CVA 方法对过程测量数据和测量数据之间的相关性进行分析, 利用 PCA 方法进一步分析了数据的方差特性, 最终将原始数据空间映射到 5 个子空间, 对 5 个子空间分别构造统计量进行精细监控. 从 TE 过程的仿真结果可以看出, 与 PCA、DPCA 和 CVA 方法相比, DIOCVA 方法提供了更加精细的过程监控手段. 通过对输入输出数据空间的精细

化分解, 使监控系统在检测故障的同时, 能够有效地分析过程故障与产品质量的关系, 对实际工业应用具有重要价值.

References

- Chiang L H, Russell E L, Braatz R D. *Fault Detection and Diagnosis in Industrial Systems*. London: Springer-Verlag, 2001. 35–98
- Zhou Dong-Hua, Liu Yang, He Xiao. Review on fault diagnosis techniques for closed-loop systems. *Acta Automatica Sinica*, 2013, **39**(11): 1933–1943 (周东华, 刘洋, 何潇. 闭环系统故障诊断技术综述. *自动化学报*, 2013, **39**(11): 1933–1943)
- Zhou Dong-Hua, Shi Jian-Tao, He Xiao. Review of intermittent fault diagnosis techniques for dynamic systems. *Acta Automatica Sinica*, 2014, **40**(2): 161–171 (周东华, 史建涛, 何潇. 动态系统间歇故障诊断技术综述. *自动化学报*, 2014, **40**(2): 161–171)
- Li Yue-Yang, Zhong Mai-Ying. Fault detection filter design for linear discrete time-varying systems with multiple packet dropouts. *Acta Automatica Sinica*, 2015, **41**(9): 1638–1648 (李岳扬, 钟麦英. 具有多测量数据包丢失的线性离散时变系统故障检测滤波器设计. *自动化学报*, 2015, **41**(9): 1638–1648)
- Wang Jing, Hu Yi, Shi Hong-Bo. Fault detection for batch processes based on Gaussian mixture model. *Acta Automatica Sinica*, 2015, **41**(5): 899–905 (王静, 胡益, 侍洪波. 基于 GMM 的间歇过程故障检测. *自动化学报*, 2015, **41**(5): 899–905)
- Li Han, Xiao De-Yun. Survey on data driven fault diagnosis methods. *Control and Decision*, 2011, **26**(1): 1–9 (李晗, 箫德云. 基于数据驱动的故障诊断方法综述. *控制与决策*, 2011, **26**(1): 1–9)
- Qin S J. Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 2012, **36**(2): 220–234
- Ge Z Q, Song Z H, Gao F R. Review of recent research on data-based process monitoring. *Industrial and Engineering Chemistry Research*, 2013, **52**(10): 3543–3562
- Yin S, Ding S X, Xie X C, Luo H. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 2014, **61**(11): 6418–6428
- Ding S X. Data-driven design of monitoring and diagnosis systems for dynamic processes: a review of subspace technique based schemes and some recent results. *Journal of Process Control*, 2014, **24**(2): 431–449
- Garcia-Alvarez D, Fuente M J, Sainz G I. Fault detection and isolation in transient states using principal component analysis. *Journal of Process Control*, 2012, **22**(3): 551–563
- Russell E L, Chiang L H, Braatz R D. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2000, **51**(1): 81–93
- Juricek B C, Seborg D E, Larimore W E. Fault detection using canonical variate analysis. *Industrial and Engineering Chemistry Research*, 2004, **43**(2): 458–474
- Stubbs S, Zhang J, Morris J. Fault detection in dynamic processes using a simplified monitoring-specific CVA state space modelling approach. *Computers and Chemical Engineering*, 2012, **41**: 77–87
- Lee J M, Yoo C K, Choi S W, Vanrolleghem P A, Lee I B. Nonlinear process monitoring using kernel principal component analysis. *Chemical Engineering Science*, 2004, **59**(1): 223–234

- 16 Deng X G, Tian X M. Multivariate statistical process monitoring using multi-scale kernel principal component analysis. In: Proceedings of the 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes. Beijing, China: IFAC, 2006. 108–113
- 17 Tian X M, Zhang X L, Deng X G, Chen S. Multiway kernel independent component analysis based on feature samples for batch process monitoring. *Neurocomputing*, 2009, **72**(7–9): 1584–1596
- 18 Deng Xiao-Gang, Tian Xue-Min. Nonlinear process fault diagnosis based on kernel canonical variate analysis. *Control and Decision*, 2006, **21**(10): 1109–1113
(邓晓刚, 田学民. 基于核规范变量分析的非线性故障诊断方法. 控制与决策, 2006, **21**(10): 1109–1113)
- 19 Ku W F, Storer R H, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 1995, **30**(1): 179–196
- 20 Stefatos G, Ben H A. Dynamic independent component analysis approach for fault detection and diagnosis. *Expert Systems with Applications*, 2010, **37**(12): 8606–8617
- 21 Wang D, Romagnoli J A. Robust multi-scale principal components analysis with applications to process monitoring. *Journal of Process Control*, 2005, **15**(8): 869–882
- 22 Cai L F, Tian X M, Chen S. A process monitoring method based on noisy independent component analysis. *Neurocomputing*, 2014, **127**: 231–246
- 23 Qin S J, Zheng Y Y. Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures. *AIChE Journal*, 2013, **59**(2): 496–504
- 24 Gunther J C, Conner J S, Seborg D E. Process monitoring and quality variable prediction utilizing PLS in industrial fed-batch cell culture. *Journal of Process Control*, 2009, **19**(5): 914–921
- 25 Zhou D H, Li G, Qin S J. Total projection to latent structures for process monitoring. *AIChE Journal*, 2010, **56**(1): 168–178
- 26 Sun R R, Fan Y P, Zhang Y W. Fault monitoring of nonlinear process based on kernel concurrent projection to latent structures. In: Proceedings of the 33rd Chinese Control Conference. Nanjing, China: IEEE, 2014. 5184–5189
- 27 Liu Q, Qin S J, Chai T Y. Quality-relevant monitoring and diagnosis with dynamic concurrent projection to latent structures. In: Proceedings of the 19th International Federation of Automatic Control World Congress. Cape Town, South Africa: IFAC, 2014. 2740–2745
- 28 Liu Q, Qin S J, Chai T Y. Multiblock concurrent PLS for decentralized monitoring of continuous annealing processes. *IEEE Transactions on Industrial Electronics*, 2014, **61**(11): 6429–6437
- 29 Zhao Z G, Li Q H, Huang M, Liu F. Concurrent PLS-based process monitoring with incomplete input and quality measurements. *Computers and Chemical Engineering*, 2014, **67**: 69–82
- 30 Odiowei P P, Cao Y. State-space independent component analysis for nonlinear dynamic process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 2010, **103**(1): 59–65
- 31 Lyman P R, Georgakis C. Plant-wide control of the Tennessee Eastman problem. *Computers and Chemical Engineering*, 1995, **19**(3): 321–331

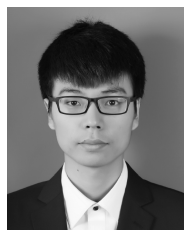


曹玉苹 中国石油大学(华东)信息与控制工程学院讲师。主要研究方向为过程故障诊断与预测。本文通信作者。

E-mail: caoy@upc.edu.cn

(CAO Yu-Ping Lecturer at the College of Information and Control Engineering, China University of Petroleum (East China). Her research interest

covers process fault diagnosis and prediction. Corresponding author of this paper.)



黄琳哲 中国石油工程建设公司华东设计分公司助理工程师。主要研究方向为过程故障诊断。

E-mail: joeyseraph@gmail.com

(HUANG Lin-Zhe Assistant engineer at the East-China Design Branch, China Petroleum Engineering Construction Corporation. His research interest

is process fault diagnosis.)



田学民 中国石油大学(华东)信息与控制工程学院教授。主要研究方向为过程建模、先进控制和故障诊断。

E-mail: tianxm@upc.edu.cn

(TIAN Xue-Min Professor at the College of Information and Control Engineering, China University of Petroleum (East China). His research

interest covers process modelling, advance control and fault diagnosis.)