

1-2-2019

Improved CPLS Algorithm and Its Application in Process Monitoring

Qinghua Li

Key Laboratory of Advanced Process Control for Light Industry Ministry of Education, Jiangnan University, Wuxi 214122, China;

Pan Feng

Key Laboratory of Advanced Process Control for Light Industry Ministry of Education, Jiangnan University, Wuxi 214122, China;

Zhonggai Zhao

Key Laboratory of Advanced Process Control for Light Industry Ministry of Education, Jiangnan University, Wuxi 214122, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Improved CPLS Algorithm and Its Application in Process Monitoring

Abstract

Abstract: Concurrent PLS (CPLS) further extracts information from the residuals of input variables and quality variables drawn by PLS, thus the raw data are projected into five subspaces. The process monitoring based CPLS provides a whole framework for the monitoring of input variables and quality variables. The model for residuals is developed by a deterministic manner while the residuals are inherently stochastic; therefore a probabilistic model is more proper for describing their features. *This paper introduces factor analysis (FA) into CPLS, in which FA instead of PCA is used to analyze the residuals to develop the improved CPLS model, and the monitoring indices for checking the validity of variables satisfying Gaussian distribution are built to improve the consistence between the modeling objective and the monitoring indices.*

Keywords

CPLS, factor analysis, expectation maximization (EM) algorithm, process monitoring

Recommended Citation

Li Qinghua, Pan Feng, Zhao Zhonggai. Improved CPLS Algorithm and Its Application in Process Monitoring[J]. Journal of System Simulation, 2018, 30(2): 622-628.

一种改进 CPLS 算法及其在过程监控中的应用

李庆华, 潘丰, 赵忠盖

(江南大学轻工过程先进控制教育部重点实验室, 江苏 无锡 214122)

摘要: CPLS(Concurrent PLS)对 PLS 分解的过程变量和质量变量的残差和主元进行进一步的提取, 从而将变量投影到五个子空间, 并由此构建了对过程变量和质量变量信息的完整监控框架。但是, 在 CPLS 中, 假设残差为可以求解的确定值, 而残差本质上为随机分布量。因此, 采用随机模型及其基于随机模型的监控更能反应残差的特性。在基于 CPLS 的过程监控中, 采用因子分析(FA)算法对 PLS 中的残差进行分析, 建立了基于 FA 的改进 CPLS 模型, 并构建了符合正态分布假设条件的监控指标, 提高了监控指标与建模指标的一致性。

关键词: CPLS; 因子分析; 期望最大化(EM)算法; 过程监控

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2018) 02-0622-07

DOI: 10.16182/j.issn1004731x.joss.201802031

Improved CPLS Algorithm and Its Application in Process Monitoring

Li Qinghua, Pan Feng, Zhao Zhonggai

(Key Laboratory of Advanced Process Control for Light Industry Ministry of Education, Jiangnan University, Wuxi 214122, China)

Abstract: Concurrent PLS (CPLS) further extracts information from the residuals of input variables and quality variables drawn by PLS, thus the raw data are projected into five subspaces. The process monitoring based CPLS provides a whole framework for the monitoring of input variables and quality variables. The model for residuals is developed by a deterministic manner while the residuals are inherently stochastic; therefore a probabilistic model is more proper for describing their features. This paper introduces factor analysis (FA) into CPLS, in which FA instead of PCA is used to analyze the residuals to develop the improved CPLS model, and the monitoring indices for checking the validity of variables satisfying Gaussian distribution are built to improve the consistence between the modeling objective and the monitoring indices.

Keywords: CPLS; factor analysis; expectation maximization (EM) algorithm; process monitoring

引言

在工业过程中, 为保证安全稳定的生产, 过程监控技术受到了广泛的重视。基于数据驱动的过程监控方法对过程的机理知识依赖少, 仅通过提取蕴

藏在过程日常运行数据中的潜在信息, 即可实现对过程的监控, 可实现性强, 近年来得到了极大的发展^[1]。其中, 偏最小二乘算法(PLS)已逐步成为数据信息提取进而实现过程监控的常用算法之一。PLS 通过提取过程变量与质量变量相关性最大的特征信息^[2], 将过程变量和质量变量划分为主元子空间和残差子空间, 实现了对数据的压缩与提取。基于 PLS 的过程监控方法将实时数据投影到过程变量的主元子空间和残差子空间, 并构建监控指标



收稿日期: 2015-12-03 修回日期: 2016-03-05;
基金项目: 国家自然科学基金(61273131);
作者简介: 李庆华(1976-), 女, 湖北, 博士生, 研究方向为工业过程的建模与监控; 潘丰(1963-), 男, 江苏, 博士, 教授, 博导, 研究方向为工业过程建模与软测量等。

<http://www.china-simulation.com>

T^2 和 SPE 分别度量过程变量的主元和残差的信息变化^[3], 从而评估过程的运行状态。但是, 该方法忽略了质量变量的信息变化, 当质量变量的残差信息发生变化时, 会产生漏报现象。实际上, 相较于过程变量, 对质量变量信息变化的监控更加重要。另一方面, 建立 PLS 模型所涉及的优化目标是使过程变量和质量变量的主元相关性最大, 而对残差没有任何约束, 因此无法保证过程变量和质量变量的残差方差最小, 导致残差中可能包含了大量的过程变量和质量变量的信息, 采用 SPE 指标无法满足对该残差的监控。TPLS(Total, PLS)对采用 PLS 算法得到的过程变量的主元空间进行进一步划分^[4], 分为与质量变量有关的和与质量变量无关的子空间, 然后再对由 PLS 得到的过程变量和质量变量的残差信息进行进一步划分, 分为主要信息和残差信息。基于 TPLS 的监控方法能够有效监控过程变量与质量变量相关的信息, 以及不相关信息中的主要信息和残差信息。但是该方法无法对质量变量的信息进行监控, 并且将过程变量划分为四个子空间, 涉及重复划分^[5]。CPLS(Concurrent PLS)算法在 PLS 提取的基础上, 对与质量变量不相关的过程变量信息、无法预测的质量变量信息分别进行主元分析, 得到 5 个子空间^[5]: 过程变量与质量变量相关信息的子空间(相关主元子空间)、过程变量主元空间、过程变量残差空间、质量变量主元空间、质量变量残差子空间。基于 CPLS 的过程监控方法设计了 SPE 和 T^2 指标分别监控这 5 个子空间的信息, 取得了较好的效果。

CPLS 中, 对 PLS 的过程变量和质量变量残差信息的提取采用了 PCA 算法。PCA 算法将残差进行投影, 将残差分离成主元和真正的残差, 所有变量为确定性(deterministic)变量。而残差变量本质上是一种随机量, 因此采用随机变量更能反应残差的特性。另一方面, 在 PCA 监控中, 采用残差的欧几里德范数作为监控指标 SPE(或 Q), 也不符合残差的统计特性, 且与统计监控的假设条件不一致^[6]。相较于 PCA, 因子分析(FA)模型是一

种随机模型^[7], 其算法基于变量的统计分布假设, 与基于 FA 的监控方法的指标度量完全一致, 因此更具广泛意义。本文采用 FA 代替 PCA, 提出一种改进 CPLS 算法。该算法对过程变量和质量变量的残差信息进行提取, 分离出 PLS 模型残差信息的因子空间和残差子空间, 然后构建 SPE 和 T^2 指标分别进行度量, 保证了对 PLS 模型残差信息提取和监控的一致性, 提高了监控的准确性。最后, 在 TE 过程平台上进行了应用说明。

1 CPLS 算法及其缺点

1.1 CPLS 算法

考虑过程数据为 $X = \{x_1, x_2, \dots, x_N\} \in \mathfrak{R}^{m \times N}$, 质量数据为 $Y = \{y_1, y_2, \dots, y_N\} \in \mathfrak{R}^{k \times N}$, 其中 N, m, k 分别为样本个数, 过程变量个数和质量变量个数, x_n 和 y_n 为第 n 个采样值。为简略记, x_n 和 y_n 简记为 x 和 y 。PLS 模型如下:

$$\begin{cases} x = Pt + e \\ y = Qt + f \end{cases} \quad (1)$$

PLS 没有限制残差 e 和 f , 因此 e 和 f 中可能分别包含了过程变量和质量变量的大量信息。此外, t 同时可能有与 y 无关的信息^[5]。PLS 的具体算法请参见参考文献[8]。

CPLS 基于模型(1), 对主元 t 的分解提取, 并采用 PCA 对 e 和 f 分别进行再次提取, 分离 e 和 f 中的主要成分和残差。CPLS 模型如下:

$$\begin{cases} X = U_c R_c^+ + T_x P_x^T + \tilde{X} \\ Y = U_c Q_c^T + T_y P_y^T + \tilde{Y} \end{cases} \quad (2)$$

式中: U_c 为 X 与 Y 的相关主元; T_x 表示与 Y 无关的过程变量的主元; P_x 表示与 Y 无关的过程变量的投影向量; T_y 表示无法由 X 预测的 Y 的主元; P_y 表示无法预测的 Y 的投影向量; \tilde{X} 和 \tilde{Y} 分别表示 X 和 Y 的残差; R_c 和 Q_c 分别为 PLS 模型投影向量。基于 PLS 模型, 针对质量变量的预测矩阵 $\hat{Y} = QT$ 进行奇异值分解, 如下^[5,9]:

$$\hat{Y} = U_c D_c V_c^T \equiv U_c Q_c \quad (3)$$

式中： U_c 为相关主元矩阵； $Q_c = D_c V_c^T$ 。其中， Q_c 包含了所有不为零特征值所对应的特征向量，去除了与 Y 不相关的信息。因此，可得主元如下：

$$U_c = \hat{Y} V_c D_c^{-1} = X P Q^T V_c D_c^{-1} \equiv X R_c \quad (4)$$

式中： $R_c = P Q^T V_c D_c^{-1}$ 。由式(4)可得

$$\hat{X} = U_c R_c^+ \quad (5)$$

式中： $R_c^+ = (R_c^T R_c)^{-1} R_c^T$ 为相关矩阵的投影矩阵。得到了与 Y 的相关主元后，进一步，采用PCA算法，对过程变量残差 $E = X - \hat{X}$ 和质量变量残差 $F = Y - \hat{Y}$ 进行分解得到投影矩阵 P_x 和 P_y 。

1.2 CPLS 算法的缺点

CPLS在对残差信息进行进一步提取中，采用的优化指标分别是使 $\|E - T_x P_x^T\|^2$ 和 $\|F - T_y P_y^T\|^2$ 最小^[10]。而实际残差包含了随机噪声，本质为一种随机量。因此，采用随机模型更能反应残差的特性。另一方面，针对残差的监控采用基于PCA的监控指标，如下：

$$\begin{cases} Q_x = \tilde{x}_k \tilde{x}_k^T \\ Q_y = \tilde{y}_k \tilde{y}_k^T \end{cases} \quad (6)$$

显然，对采样时刻 k 的残差信息 \tilde{x}_k 和 \tilde{y}_k 的监控指标采用的是欧几里德距离度量，隐藏的假设条件是残差满足均匀分布。而基于PCA的统计过程监控方法的应用前提是在过程正常运行状况下，残差满足正态分布，因此，对残差的监控存在量度的不一致性^[6,11]。

2 改进 CPLS 算法及其应用

2.1 改进 CPLS 算法

以对 E 的进一步提取为例，说明改进CPLS算法。FA假设 E (向量表示为 e)由如下生成方程产生：

$$e = W_x t_x + \varepsilon \quad (7)$$

其中 W_x 为负荷向量矩阵， t_x 为因子向量，服从标准正态分布， ε 为残差，服从 $\varepsilon \sim N(0, \Psi_x)$ ，其中 $\Psi_x = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ ， $e \sim N(0, \Psi_x + W_x W_x^T)$ 。模型(7)中 e 由CPLS求出，需要求取参数 Ψ_x 和 W_x ，FA算法采用 e 的极大似然指标求取参数。但是 Ψ_x

和 W_x 不与 e 直接相关，而是通过因子向量 t_x 决定 e ，并且 t_x 未知，给FA的极大似然求解带来了困难。假设完整数据为 $[t_x, e]$ ，将 t_x 作为缺失数据，则完整数据的log似然函数为 $L_c = \sum_{k=1}^N \ln(t_{x,k}, e_k)$ ，式中 $t_{x,k}$ 和 e_k 分别表示采样时刻 k 的因子和PLS残差。采用期望最大化(EM)算法求解 e 的极大似然函数，得到参数 Ψ_x 和 W_x 。EM算法由对 t 的期望和似然最大化两步组成，如下：

期望：假设参数 Ψ_x 和 W_x 给定，完整数据的log似然函数为

$$Q = -\frac{N}{2} \ln |\Psi_x| - \sum_{k=1}^N \left(\frac{1}{2} e_k^T \Psi_x^{-1} e_k - e_k^T \Psi_x^{-1} W E[t_{x,k} | e_k] + \frac{1}{2} \text{tr}(t_{x,k}^T \Psi_x^{-1} E[t_{x,k} t_{x,k}^T | e_k]) \right) \quad (8)$$

式(8)忽略了似然函数的一些常数项。其中 tr 为迹操作算子。因此，完整数据的log似然函数取决于以下两个期望：

$$E[t_{x,k} | e_k] = W_x^T C^{-1} e_k \quad (9)$$

$$E[t_{x,k} t_{x,k}^T | e_k] = I - W_x^T C^{-1} W_x + W_x^T C^{-1} e_k e_k^T C^{-1} W_x \quad (10)$$

式中： $C = W_x W_x^T + \Psi_x$ 。

最大化：对完整数据的log似然函数进行最大化处理：关于参数 Ψ_x 和 W_x 对期望求导，并令为0，得到新的参数如下

$$\hat{W}_x = \left(\sum_{k=1}^N e_k E(t_{x,k} | e_k) \right) \left(\sum_{k=1}^N E(t_{x,k} t_{x,k}^T | e_k) \right)^{-1} \quad (11)$$

$$\hat{\Psi}_x = \frac{1}{N} \text{diag} \left(\sum_{k=1}^N (e_k e_k^T - \hat{W} E(t_{x,k} | e_k) e_k^T) \right) \quad (12)$$

式(9)、(10)、(11)、(12)构成了EM算法。反复迭代期望步和最大化步直到收敛，即可求得参数。关于因子个数的选取参见文献[11]。很明显，在针对残差的FA建模过程中， e 、 t_x 和 ε 满足正态分布是基本假设条件，并且优化指标是针对正态分布的似然函数，因此模型的求解与统计过程监控的假设条件一致。

与 E 的建模相同，也可采用FA对 F 进行建模。改进CPLS算法如下：

1) 对归一化后的 X 和 Y 进行 PLS 提取, 得到 T, P, Q 。

2) 对 Y 的预测值 $\hat{Y} = TQ^T = XPQ^T$ 进行奇异值分解得到 $\hat{Y} = U_c D_c V_c^T \equiv U_c Q_c^T$, 其中 $U_c = \hat{Y} V_c D_c^{-1} = XPQ^T V_c D_c^{-1} \equiv XR_c$ 。

3) 求出 X 中不可预测的部分为 $\tilde{X}_c = X - U_c R_c^+$, 并采用 FA 算法对 \tilde{X}_c 进行因子提取, 得到 \tilde{X}_c 的因子模型 $\tilde{X}_c = W_x T_x + \tilde{X}$ 。

4) 与步骤 3 一样, 求出 Y 中不可预测的部分为 $\tilde{Y}_c = Y - \hat{Y}$, 并采用 FA 算法对 \tilde{Y}_c 进行因子提取, 得到 \tilde{Y}_c 的因子模型 $\tilde{Y}_c = W_y T_y + \tilde{Y}$ 。

2.2 改进 CPLS 算法在过程监控中的应用

经过改进 CPLS 算法的分析后, 过程变量和质量变量被划分为 5 个部分: 两者的相关主元子空间, 过程变量因子子空间, 过程变量残差子空间, 质量变量因子子空间, 质量变量残差子空间。这五个子空间代表了实时数据的所有信息, 通过监控这 5 个子空间的信息即可实现对过程的监控。

由 2.1 节可知, 在 CPLS 模型中, 相关主元、因子、残差均满足正态分布, 反应了过程正常运行状况下的状态^[6]。基于 CPLS 模型的过程监控即判断实时数据是否符合模型, 也就是将当前的相关主元、因子、残差是否满足正常的正态分布。为监测实时数据的相关主元、因子和残差是否属于正常分布, 则必须定义指标进行定量地衡量。由于正态分布的信息量包含在均值和方差中, 为在监控中综合应用均值和方差的信息, 采用各参数的 Mahalanobis 距离对其进行衡量, 如下:

对相关主元的监控指标:

$$T_c^2 = u_c^T \Lambda_c^{-1} u_c \quad (13)$$

式中 Λ_c 为 u_c 的协方差矩阵。该监控指标与 CPLS 中关于相关主元的监控相同, 主要用于检测过程变量和质量变量相关信息的变化。

对过程变量因子子空间的监控指标为:

$$T_x^2 = t_x^T t_x \quad (14)$$

t_x 为满足正态分布的随机量, 根据基于 FA 的监控

方法, 通常采用期望 $\bar{t}_x = E[t_x | x]$ 代替 t_x ^[6,11]。该指标用于检测过程变量与质量变量无关的主要信息变化。

对过程变量残差的监控指标

$$T_{\tilde{x}}^2 = \tilde{x}^T \Psi_x^{-1} \tilde{x} \quad (15)$$

在 CPLS 中, 过程变量的残差是过程变量中剔除与质量相关信息以及无关主要信息后的残留信息, 该部分信息的变化反映了实时数据是否符合模型定义的相关关系, 如果不符合, 则残差中包含的信息量大, 表现为监控指标偏大。

对质量变量因子子空间的监控指标为

$$T_y^2 = t_y^T t_y \quad (16)$$

同样地, t_y 为满足正态分布的随机量, 通常采用期望 $\bar{t}_y = E[t_y | y]$ 代替 t_y 。该指标用于检测与过程变量无关的质量变量的主要信息变化。

对过程变量残差的监控指标

$$T_{\tilde{y}}^2 = \tilde{y}^T \Psi_y^{-1} \tilde{y} \quad (17)$$

式中: Ψ_y 为 \tilde{y} 的协方差矩阵。与 $T_{\tilde{x}}^2$ 一样, 体现了实

时变量相关关系与模型定义的符合程度。

在过程正常运行工况下, u_c, t_x, \tilde{x}, t_y 和 \tilde{y} 满足期望为 0, 方差分别为 $\Lambda_c, I, \Psi_x, I, \Psi_y$ 的正态分布。过程监控的实质为判断实时采样变量的因子和残差是否符合相应的分布。上述五个监控指标均采用 Mahalanobis 范数, 度量采样时刻 k 下 $u_{c,k}, t_{x,k}, \tilde{x}_k, t_{y,k}$ 和 \tilde{y}_k 与正常正态分布的距离, 如果五个距离都在各自阈值(控制限)内, 则标明当前时刻过程正常运行。相较于欧几里德范数, Mahalanobis 范数不仅考虑了当前样本点与正常分布期望向量之间的距离, 而且也考虑了正常分布的方差。为评估采样变量是否满足正态分布, 只有综合考虑期望和方差才能准确评估在线变量的状态。因此, 欧几里德范数适合度量当前样本与均匀分布的距离, 而 Mahalanobis 范数适合度量当前样本与正态分布的距离, Mahalanobis 范数更符合统计过程监控方法正态分布的假设条件。

3 仿真研究

TE 过程是对实际化工过程的仿真模拟, 目前已逐渐成为一个评估过程控制、建模、监控与优化等技术的公认仿真平台。该过程包括反应器、产品冷凝器、气液分离器、循环压缩机和汽提塔五个操作单元, 涉及四种气态组分反应, 最终形成两种产品和一种副产品。气态组分在反应器反应后, 经过冷凝器冷凝后, 流入气液分离器被分成两部分, 被压缩的组分到产品提馏段进行精馏提纯, 而没有被压缩的组分则通过压缩机被回收作为组分的原料。整个过程涉及到能量和物料的变化。具体的描述参见文献[12]。

过程采样周期为 3 分钟, 运行周期为 48h, 过程设置了 21 种故障模式。过程包含 52 个投运变量, 其中 11 个操作变量, 41 个测量变量。采用 700 组正常样本建立模型, 200 组数据进行校验, 然后再利用故障数据进行监控算法的验证。在 CPLS 建模中, 根据文献[5], 47 个变量(XMEAS(1-36)和XMV(1-11))作为 X 变量, 5 个变量(XMEAS(37-41))作为 Y 变量。通过交叉检验, X 与 Y 相关的主元个数为 5 个, 而 X 与 Y 无关的主元个数为 32 个, 无法由 X 预测的 Y 的主元个数为 5 个, T_y^2 为空。针对正常运行数据, 采用 PLS 进行分解后, 得到的流 1 中 A 进料(XMEAS(1))和质量变量产品 E(XMEAS(38))的残差见图 1。

可见, 残差中包含了大量的有用信息, 需要进行进一步的提取。采用 CPLS 得到的残差如图 2。其中, 由于无法由 X 预测的 Y 的主元个数为 5, 因此经过 CPLS 提取后, 产品 E 的残差信息为 0。显然, CPLS 充分提取了 PLS 的残差中的信息。

TE 过程中, 故障 8 表示流 4 中的进料成分 A, B, C 有一个随机变化, 该故障在采样点 300 处引入, 并保持到过程运行结束。采用 CPLS 的监控效果见图 3。

分析故障 8 可知, 流 4 中的进料成分 A、B、C 发生变化, 对模型相关主元、过程变量和质量变量的因子以及残差都会产生影响。从监控图中可以

看出, 监控指标 T_c^2 , T_y^2 , T_x^2 都能够较准确及时的检测出过程故障, 但是关于残差的监控指标 Q_x 的监控效果较差有大量的漏报, 检测也存在延迟。

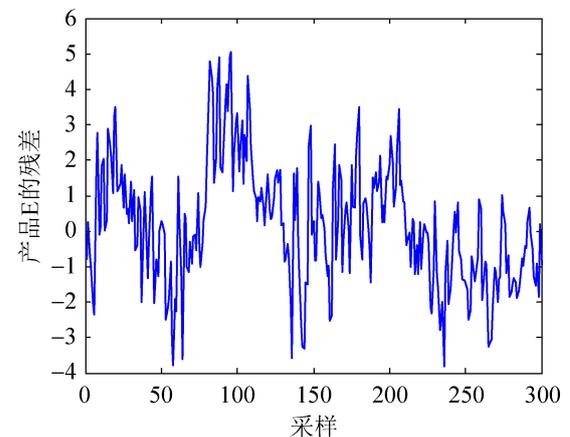
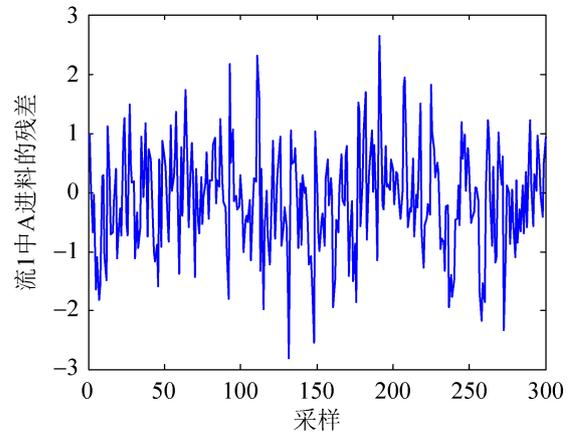


图1 经过 PLS 提取后过程变量和质量变量中的残差
Fig. 1 Residual of process variables and quality variables based on the PLS-based extraction

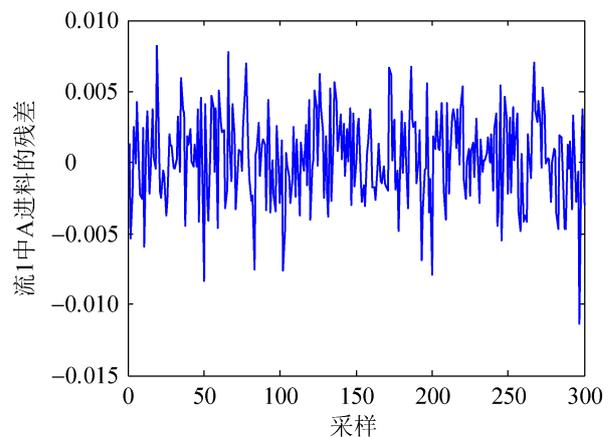


图2 经过 CPLS 提取后过程变量和质量变量中的残差
Fig. 2 Residual of process variables and quality variables based on the CPLS-based extraction

将 FA 引入到 CPLS, 参见文献[8], 得到 X 与 Y 无关的因子个数为 30 个, 无法由 X 预测的 Y 的因子个数为 5 个, 监控效果见图 4。对比图 3 和图 4 发现, 将 FA 算法引入到针对 PLS 残差的建模中, 其监控指标 T_x^2 和 T_y^2 的监控效果与基于 CPLS 的监控效果相近, 而对残差的监控指标 T_x^2 则要大大优于基于 CPLS 的监控指标 Q_x 。究其原因: 无法预

测的质量变量的因子个数为 5, 导致 T_y^2 包含了无法预测的质量变量的所有信息, 因此两者的 T_y^2 监控效果基本一样; 相比 CPLS, 改进 CPLS 算法对与质量无关的过程变量进行了新的提取(前者是 PCA 提取, 后者是 FA 提取), 使最终的残差更能反映故障数据状况下, 过程变量之间相关性的变化, 因此 T_x^2 要优于 CPLS 的 Q_x 。

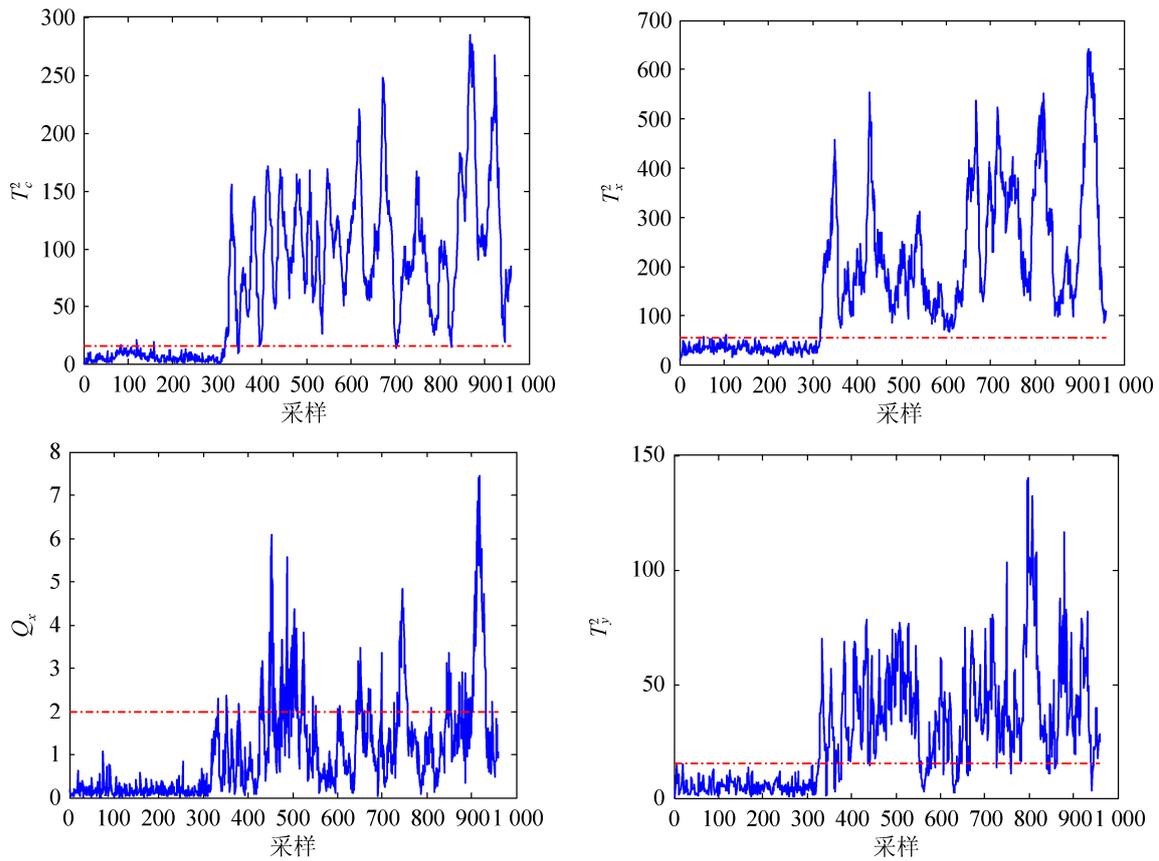


图 3 采用 CPLS 的监控效果

Fig. 3 Monitoring performance based on CPLS

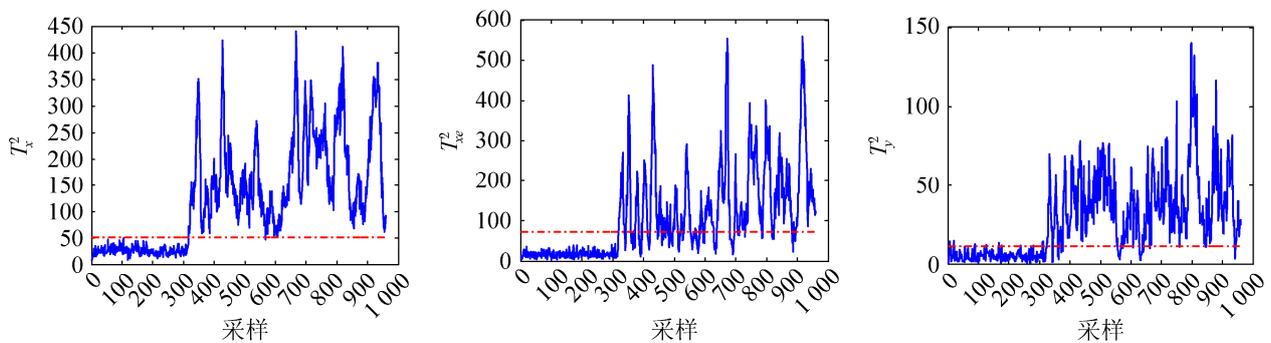
图 4 采用改进 CPLS 的监控效果(为表示的方便, 图中 T_x^2 用 T_{xe}^2 表示)

Fig. 4 Monitoring performance based on the improved CPLS

4 结论

针对 CPLS 在对 PLS 残差的建模以及监控上存在量度不一致、不能提高符合残差特性的随机模型的问题, 将 FA 引入到 CPLS 中, 建立了过程变量和质量变量两者不相关信息的 FA 模型, 并采用基于 FA 的监控方法对该部分信息进行了监控, 联合常规 CPLS 监控方法中关于过程变量和质量变量相关信息的监控, 形成了对过程状况的完整监控。

参考文献:

- [1] MacGregor J F, Kourti T. Statistical process control of multivariate processes[J]. *Control Engineering Practice* (S0967-0661), 1995, 3(3): 403-414.
- [2] Wold S, Sjostrom M, Eriksson L. PLS-regression: a basic tool of chemometrics[J]. *Chemometrics and Intelligent Laboratory Systems* (S0169-7439), 2001, 58(2): 109-130.
- [3] Godoy J L, Vega J R, Marchetti J L. A fault detection and diagnosis technique for multivariate processes using a PLS-decomposition of the measurement space[J]. *Chemometrics and Intelligent Laboratory Systems* (S0169-7439), 2013, 128: 25-36.
- [4] Zhou D, Li G, Qin S J. Total projection to latent structures for process monitoring[J]. *AIChE Journal* (S1547-5905), 2010, 56(1): 168-178.
- [5] Qin S J, Zheng Y. Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures[J]. *AIChE Journal* (S1547-5905), 2013, 59(2): 496-504.
- [6] Kim D, Lee I B. Process monitoring based on probabilistic PCA[J]. *Chemometrics and Intelligent Laboratory Systems* (S0169-7439), 2003, 67(2): 109-123.
- [7] Ghahramani Z, Geoffrey E H. The EM algorithm for mixtures of factor analyzers[R]. Technical Report CRG-TR-96-1, University of Toronto, 1997: 1-8.
- [8] Hoskuldsson A. PLS regression methods[J]. *Journal of Chemometrics* (S0886-9383), 1988, 2(3): 211-228.
- [9] Zhao Z, Li Q, Huang M, et al. Concurrent PLS-based process monitoring with incomplete input and quality measurements[J]. *Computers and Chemical Engineering* (S0098-1354), 2014, 67: 69-82.
- [10] Zhou L, Chen J, Song Z, et al. Probabilistic latent variables regression model for process-quality monitoring[J]. *Chemical Engineering Science* (S0009-2509), 2014, 116: 296-305.
- [11] 赵忠盖, 刘飞. 因子分析及其在过程监控中的应用[J]. *化工学报*, 2007, 58(4): 970-974.
Zhao Z G, Liu F. Factor analysis and its application to process monitoring[J]. *Journal of Chemical Industry and Engineering (China)* (S0438-1157), 2007, 58(4): 970-974.
- [12] Chiang L H, Russell E L, Braatz R D. *Fault detection and diagnosis in industrial systems*[M]. Springer: London. 2001: 103-112.