

研究论文

# 基于 PICA 的过程监控方法

葛志强, 宋执环

(浙江大学工业控制技术国家重点实验室, 工业控制研究所, 浙江 杭州 310027)

**摘要:** 工业过程中普遍存在噪声污染, 本文在概率主元分析方法 (PPCA) 的基础上, 把该方法推广到非高斯过程, 提出一种新的基于概率独立成分分析 (PICA) 的过程监控方法。针对过程的非高斯和噪声信息, 分别建立其对应的统计量  $I^2$  和  $MR$ 。通过对 Tennessee Eastman (TE) 过程的仿真研究, 验证了该方法的可行性和有效性, 较好地改善了过程的监控效果, 从而更好地保证过程运行的安全、稳定性。

**关键词:** 概率独立成分分析; 非高斯; 噪声污染

**中图分类号:** TP 277

**文献标识码:** A

**文章编号:** 0438-1157 (2008) 07-1665-06

## PICA based process monitoring method

GE Zhiqiang, SONG Zhihuan

(State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control,  
Zhejiang University, Hangzhou 310027, Zhejiang, China)

**Abstract:** Noise corruption always exists in the industrial process. Based on the probabilistic principal component analysis (PPCA) method, a new process monitoring method based on probabilistic independent component analysis (PICA) was proposed, which extends PPCA to the non-Gaussian process. Two statistical quantities ( $I^2$  and  $MR$ ) were constructed for monitoring non-Gaussian and noise information of the process. A case study of the Tennessee Eastman (TE) process showed that the proposed method was feasible and efficient. The process monitoring performance was evidently improved, thus enhancing the reliability and stability of the TE process.

**Key words:** probabilistic independent component analysis; non-Gaussian; noise corruption

### 引 言

随着集散控制系统 (DCS) 在工业中的广泛应用, 海量的工业过程数据被收集和存储起来。如何从这些数据中提取出有用的信息来对过程进行监控已经成为目前的研究热点之一。在基于模型、基于知识和基于数据驱动这三大类方法中, 基于数据驱动的过程监控方法是目前应用最为广泛的一种监控方法, 其中以主元分析/偏最小二乘 (PCA/PLS)

为代表的多变量统计过程控制 (MSPC) 自 20 世纪 90 年代以来受到了学术界和工业界的广泛重视, 并在很多工业过程中得到了成功的应用<sup>[1]</sup>。但是, 在应用 PCA 等传统方法对过程进行监控的时候, 需要假设噪声向量各向同性且方差最小, 这就极大地限制了该方法的应用范围。针对 PCA 需要噪声方差最小的假设限制, 有学者提出了概率 PCA (PPCA) 模型<sup>[2]</sup>。在此基础上, 赵忠盖等<sup>[3]</sup>提出了基于因子分析的过程监控方法, 克服了 PCA 需

2007-12-02 收到初稿, 2008-01-06 收到修改稿。

**联系人:** 宋执环。第一作者: 葛志强 (1982-), 男, 博士研究生。

**基金项目:** 国家自然科学基金项目 (60774067)。

**Receive date:** 2007-12-02.

**Corresponding author:** Prof. SONG Zhihuan. E-mail: zhsong@iipc.zju.edu.cn

**Foundation item:** supported by the National Natural Science Foundation of China (60774067).

要噪声各向同性的限制。但是，这些方法都是假设过程数据服从高斯分布，这在实际过程中是很难得到满足的。

独立成分分析 (independent component analysis, ICA) 作为一种新的非高斯统计信号处理方法，在生物医学信号处理、混合语音分离、盲源信号分离、图像处理以及人脸识别领域已获得了成功的应用<sup>[4]</sup>，在化工过程性能监控方面也有一定的应用前景。Kano 等<sup>[5]</sup>把 ICA 方法引入化工过程监控领域，Xie 等<sup>[6-8]</sup>也在这一领域做了很多工作。虽然带噪声的 ICA 模型在其他领域（特别是语音信号处理领域）已经得到了一定的研究成果，在过程监控领域，研究在噪声环境下的 ICA 监控模型目前尚没有报道。

针对传统 MSPC 及其改进方法的不足，本文提出一种概率 ICA (PICA) 监控方法。该方法首先利用期望极大方法 (EM) 获得模型参数的估计，然后分别构造  $I^2$  和 MR 统计量对过程进行监控。在确定  $I^2$  统计量的控制限时，本文采用了核密度估计的方法。为了验证该方法的有效性，把 PICA 在 Tennessee Eastman (TE) 过程的仿真结果和其他方法进行了比较。

## 1 基本方法

### 1.1 独立成分分析 (ICA)

独立成分分析假定  $k$  时刻  $m$  维观测向量  $\mathbf{Z}(k) = [z_1(k), \dots, z_m(k)]^T$  可由未知的独立分量  $\mathbf{S}(k) = [S_1(k), \dots, S_r(k)]^T$  线性表示

$$\mathbf{Z}(k) = \mathbf{A}\mathbf{S}(k) \quad (1)$$

式中  $\mathbf{A} \in \mathbf{R}^{m \times r}$  ( $m \geq r$ ) 为混迭矩阵。它的目的就是从小混合信号中找到未知的混迭矩阵  $\mathbf{A}$ ，实现源信号的盲源分离。在  $\mathbf{S}$  统计独立的假设下，ICA 算法寻求满足式 (2) 的线性变换矩阵

$$\hat{\mathbf{S}}(k) = \mathbf{W}\mathbf{Z}(k) \quad (2)$$

ICA 算法是在  $\mathbf{S}$  和  $\mathbf{A}$  未知的情况下，寻找分离矩阵  $\mathbf{W}$ ，并从  $\mathbf{Z}$  中分离源信号使得分离后的信号最大程度独立。本文采用负熵的方法评价这种信号分离的非高斯性。限于篇幅，ICA 算法的详细描述可以见文献 [4]，另外，在提取独立成分的方法上，一种快而有效的方法是 Hyvarinen<sup>[9]</sup> 提出的定点快速 ICA 算法。但是，该方法需要事先确定独立成分的数目，而且提取出来的独立成分并不是按照非高斯性进行排列。最近提出的基于微粒群优化

算法的 PSO-ICA 方法<sup>[6]</sup>能较好地解决上述问题，因此本文采用 PSO-ICA 算法提取过程的独立成分。

### 1.2 核密度估计

由于本文在确定统计量  $I^2$  的控制限时，只需对一维的核密度进行估计，因此，在此只对一维核密度估计方法进行简要介绍。假设  $y_1, y_2, \dots, y_N$  为一维随机变量的采样样本，通过下列形式可估计  $y$  的概率密度函数

$$\hat{f}(y, \mathbf{H}) = \frac{1}{N} \sum_{i=1}^N K[\mathbf{H}^{-1/2}(y - y_i)] \quad (3)$$

其中， $y_i \in \mathbf{R}$  为第  $i$  个数据向量， $\mathbf{H}$  为相应核函数的带宽 (bandwidth) 参数矩阵，可以通过优化相关的目标函数进行确定， $K$  为核函数。这里选取核函数为高斯函数  $K(y) = (2\pi)^{-1/2} \exp(-y^T y / 2)$ ，带宽参数矩阵选为  $\mathbf{H} = \text{diag}(h_1^2, h_2^2, \dots, h_p^2)$ 。关于多元核密度估计以及详细的介绍见文献 [10]。

## 2 概率 ICA (PICA) 模型的建立

标准的独立成分分析只考虑理想的情况，即过程数据不存在噪声 (或者噪声可以忽略)。但是，实际工业应用中采集到的数据往往受到噪声的污染。因此，更一般的情况是同时考虑过程数据的噪声项，具有噪声的 ICA 模型可以表示为

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{e} \quad (4)$$

其中， $\mathbf{x} \in \mathbf{R}^m$  表示  $m$  维观测向量的一个样本； $\mathbf{s} \in \mathbf{R}^r$  表示独立成分，服从非高斯分布； $\mathbf{e}$  为过程数据的噪声项。一般情况下，假设噪声与各个独立成分之间是相互独立的，而且为高斯噪声，均值为零，方差为  $\mathbf{\Sigma}$ 。

由于概率 PCA 和 FA 是在高斯分布的框架下进行建模，其对应主元和因子的后验概率分布很容易获得，因此，用期望极大 (EM) 算法对其进行参数估计也就很容易实现。但是，对于独立成分，源信号的概率分布事先是未知的，由于不是严格服从高斯分布，它的后验概率分布很难用一种解析的形式给出。文献 [11] 提出了用拉普拉斯近似方法将源信号的后验分布表示为高斯分布的形式，然后求取其一阶和二阶统计量，进而构造 EM 算法估计参数  $\mathbf{A}$  和  $\mathbf{\Sigma}$ 。该算法是在极大似然意义下完成的，也就是最大化下列似然函数

$$Q(\mathbf{A}, \mathbf{\Sigma}, \mathbf{s}) = \sum_{i=1}^N \lg p(x_i | \mathbf{A}, \mathbf{\Sigma}) \quad (5)$$

$$p(\mathbf{x}_i | \mathbf{A}, \boldsymbol{\Sigma}) = \int p(\mathbf{x}_i | \mathbf{s}_i, \mathbf{A}, \boldsymbol{\Sigma}) p(\mathbf{s}_i) d\mathbf{s}_i \quad (6)$$

这个 EM 算法分两步完成：期望步（E 步）和最大步（M 步）。E 步的目的是求源信号后验分布的充分统计量（一阶和二阶），M 步的目的是求参数  $\mathbf{A}$  和  $\boldsymbol{\Sigma}$  的估计。

在 E 步，根据拉普拉斯近似，可以把源信号的后验分布近似表示为高斯分布的形式

$$p(\mathbf{s} | \mathbf{x}, \mathbf{A}, \boldsymbol{\Sigma}) \approx (2\pi)^{-\frac{M}{2}} |H(\hat{\mathbf{s}})|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{s} - \hat{\mathbf{s}})^T H(\hat{\mathbf{s}})(\mathbf{s} - \hat{\mathbf{s}})\right\} \quad (7)$$

其中， $\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} \{\lg p(\mathbf{s} | \mathbf{x}, \mathbf{A}, \boldsymbol{\Sigma})\}$ 。因此，独立成分的求取可以采用极大后验估计来实现，即求  $\hat{\mathbf{s}}$  使得下列似然函数最大

$$L(\mathbf{s}) = \sum_{i=1}^N \left\{ -\frac{1}{2}(\mathbf{x} - \mathbf{A}\mathbf{s})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \mathbf{A}\mathbf{s}) + \phi(\mathbf{s}) \right\} + C \quad (8)$$

其中， $\phi(\mathbf{s})$  是与先验分布有关的某个函数， $C$  是与先验分布无关的某个常数，可以忽略。综合上面各式，后验分布的条件统计量可以表示为

$$E\{\mathbf{s} | \mathbf{x}\} = \hat{\mathbf{s}}$$

$$E\{\mathbf{s}\mathbf{s}^T | \mathbf{x}\} = H(\hat{\mathbf{s}})^{-1} + \hat{\mathbf{s}}\hat{\mathbf{s}}^T \quad (9)$$

其中

$$H(\hat{\mathbf{s}}) = -\nabla_{\mathbf{s}} \nabla_{\mathbf{s}} L(\hat{\mathbf{s}}) = \mathbf{A}^T \boldsymbol{\Sigma} \mathbf{A} - \nabla_{\mathbf{s}} \nabla_{\mathbf{s}} \phi(\hat{\mathbf{s}}) \quad (10)$$

式中  $\nabla_{\mathbf{s}}$  表示对于  $\mathbf{s}$  的梯度算子。在求取独立成分的数学期望的时候，最简单的方法是采用梯度算法，即求似然函数关于  $\mathbf{s}$  的梯度

$$\nabla_{\mathbf{s}} L(\mathbf{s}) = \mathbf{A}^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \mathbf{A}\mathbf{s}) + \nabla_{\mathbf{s}} \phi(\mathbf{s}) \quad (11)$$

其中， $\nabla_{\mathbf{s}} \phi(\mathbf{s})$  是与先验分布有关的某个非线性函数，类似于普通的 ICA 算法，对于不同的过程可以选取不同的非线性函数。一般地，有以下几种非线性函数可以选取<sup>[11]</sup>：

对于超高斯过程

$$\Phi(\mathbf{s}) = \nabla_{\mathbf{s}} \phi(\mathbf{s}) = \begin{cases} -\tanh \mathbf{s} \\ -\mathbf{s} - \tanh \mathbf{s} \\ -\text{sighs} \end{cases} \quad (12)$$

对于亚高斯过程

$$\Phi(\mathbf{s}) = \nabla_{\mathbf{s}} \phi(\mathbf{s}) = \begin{cases} 2\tanh \mathbf{s} - 2\tanh(\mathbf{s} + 2) - 2\tanh(\mathbf{s} - 2) \\ -\mathbf{s} + \tanh \mathbf{s} \\ -\mathbf{s}^3 \end{cases} \quad (13)$$

相对于语音识别，信号处理领域，工业过程数据大多呈现亚高斯的情况。具体选取的时候应视情况而定，必要时也可以把两类非线性函数综合在一

起，自动实现超高斯和亚高斯信号的识别。在求得似然函数关于  $\mathbf{s}$  的梯度之后，可以得到关于  $\mathbf{s}$  的迭代公式

$$\mathbf{s}^j = \mathbf{s}^{j-1} + \beta \nabla_{\mathbf{s}} L(\mathbf{s}^{j-1}) \quad (14)$$

其中， $\beta$  为表示学习率， $\mathbf{s}^{j-1}$  表示上一步迭代得到的估计值。一旦得到了后验分布的统计量，就可以实施 EM 算法对模型的参数进行估计了。求似然函数  $Q(\mathbf{A}, \boldsymbol{\Sigma}, \mathbf{s})$  关于参数  $\mathbf{A}$  和  $\boldsymbol{\Sigma}$  的导数并令其为零，可以得到参数  $\mathbf{A}$  和  $\boldsymbol{\Sigma}$  的迭代估计值分别如下

$$\mathbf{A}^j = \left\{ \sum_{i=1}^N \mathbf{x}_i E\{\mathbf{s} | \mathbf{x}_i\}^T \right\} \left\{ \sum_{i=1}^N E\{\mathbf{s}\mathbf{s}^T | \mathbf{x}_i\} \right\}^{-1} \quad (15)$$

$$\boldsymbol{\Sigma}^j = \frac{1}{N} \sum_{i=1}^N \{ \mathbf{x}_i \mathbf{x}_i^T - \mathbf{A}^j E\{\mathbf{s} | \mathbf{x}_i\} \mathbf{x}_i^T \} \quad (16)$$

由 EM 算法的收敛性质可知，只要给定合适的初值，每一步迭代都能使似然函数值增大。因此，只要交替进行 E 步和 M 步，即反复对式 (14) ~ 式 (16) 进行迭代直至收敛，就可以获得模型参数  $\mathbf{A}$  和  $\boldsymbol{\Sigma}$  以及独立成分  $\mathbf{s}$  的估计。

### 3 基于 PICA 的过程监控方法

利用 ICA 算法提取过程得非高斯独立成分，如何选取非高斯独立成分的个数是一个重要且难以解决的问题，在以往的方法中，有利用 PCA 事先选取独立成分个数的，也有根据分离矩阵  $\mathbf{W}$  的范数大小进行排序的<sup>[11]</sup>，这些方法都没有考虑到独立成分的非高斯量度。本文利用 PSO-ICA 算法提取过程的非高斯独立成分<sup>[6]</sup>，该方法可以实现独立成分按照非高斯性度量自动排序，这样就可以选取最重要的（即非高斯性最强的）几个独立成分对过程进行监控。在得到独立成分  $\mathbf{s}$  的估计后，本文参考文献 [12] 构造  $I^2$  统计量如下

$$I^2 = \mathbf{s}^T \mathbf{s} \quad (17)$$

在得到正常工况下各个采样点的  $I^2$  值后，利用核密度估计方法<sup>[6]</sup> 获得其概率密度分布函数，进而确定其统计置信限，即  $I^2$  统计量的控制限  $I_{lim}^2$ 。由于 PICA 方法对过程噪声进行了建模，对过程残差部分的监控可以构造马氏距离统计量

$$MR = \|\boldsymbol{\Sigma}^{-1/2} \mathbf{e}\|^2 = \mathbf{e}^T \boldsymbol{\Sigma}^{-1} \mathbf{e} \quad (18)$$

$$\mathbf{e} = \mathbf{x} - \mathbf{A}\mathbf{s}$$

相比之下，传统的 SPE 统计量采用的是欧氏距离统计量，这是由于建立模型的时候没有考虑过程的噪声项。当过程的建模样本足够大时，统计量  $MR$  服从  $\chi^2$  分布，因此，其控制限可由式 (19) 确定

$$MR = \|\Sigma^{-1/2} e\|^2 \leq \chi_{\alpha, m}^2 = MR_{lim} \quad (19)$$

其中,  $\alpha$  为概率置信度,  $m$  为自由度 (等于变量个数)。

建立好监控模型之后, 就可以对过程进行在线监控了。假设当前采样数据为  $x_{new}$ , 首先提取独立成分

$$s_{new} = Wx_{new} \quad (20)$$

其中,  $W$  为分离矩阵。计算当前样本的  $I^2$  统计量

$$I_{new}^2 = s_{new}^T s_{new} \quad (21)$$

然后计算残差统计量  $MR$  的值如下

$$MR_{new} = \|\Sigma^{-1/2} e_{new}\|^2 = e_{new}^T \Sigma^{-1} e_{new} \quad (22)$$

$$e_{new} = x_{new} - As_{new}$$

当  $I_{new}^2 > I_{lim}^2$ ,  $MR_{new} < MR_{lim}$  时, 过程有可能发生某种扰动, 但也有可能是由于过程工作点变化所引起, 导致该统计量偏离建模控制限。当  $MR_{new} > MR_{lim}$ ,  $I_{new}^2 < I_{lim}^2$  时, 说明过程变量之间的模型关系已经被打破, 可以认为过程已经发生了某种故障。当  $MR_{new} > MR_{lim}$ ,  $I_{new}^2 > I_{lim}^2$  时, 说明过程已经严重偏离正常的运行工况, 过程变量之间的关系已经不服从模型关系, 这种情况发生故障的可能性最大。而当两个统计量都在控制限之内, 即  $MR_{new} < MR_{lim}$ ,  $I_{new}^2 < I_{lim}^2$  时, 可以认为过程处于统计可控状态, 此时无须对过程做出任何调整。在对过程进行监控的时候, 上述 4 种情况均有可能发生, 当有统计量发生超限时, 过程操作工程师需引起注意, 并结合具体过程知识对各种情况进行分析。如果确定过程发生故障, 需要对发生的故障进行识别、分离和修复, 使得过程回到统计可控即正常状态。

## 4 TE 过程仿真研究

Tennessee Eastman (TE) 平台是一个开放的非常具有挑战性的化工模型仿真平台, 由连续搅拌式反应器、冷凝器、气液分离器、解析塔、再沸器和离心式压缩机等多个操作单元组成, 近年来已经成为一个标准 (benchmark) 测试平台<sup>[1]</sup>。TE 过程的流程示意图如图 1 所示。

TE 过程共有 41 个测量变量 (其中包括 22 个连续过程测量变量和 19 个组分变量) 和 12 个操控变量, 仿真时人为地设定了 21 种故障工况, 详细资料见文献 [1]。这里, 选取 16 个连续过程变量作为监测变量, 如表 1 所示。采集过程平稳运行时的 960 个样本作为训练样本, 模拟故障工况时, 过

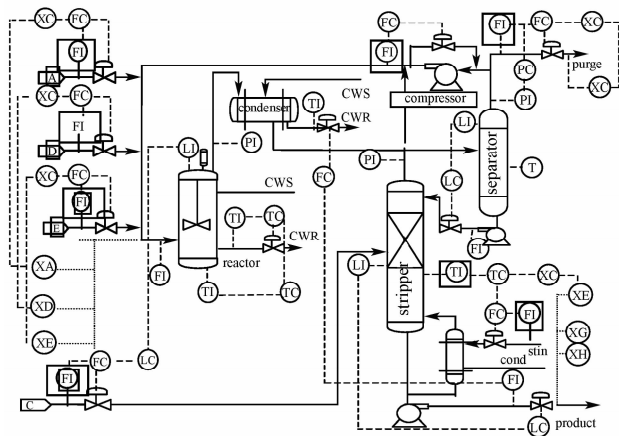


图 1 TE 过程流程图

Fig. 1 Flowsheet of TE process

程开始运行在无故障状态, 平稳运行 8 个仿真小时之后加入故障 (即前 160 个样本点为无故障样本), 样本的采样间隔均设定为 3 min, 整个采集过程中系统共运行 48 h, 采集 960 个样本点。

表 1 监测变量

Table 1 Monitoring variables

No.	Measured variables	No.	Measured variables
1	A feed	10	product separator pressure
2	D feed	11	product separator underflow
3	E feed	12	stripper pressure
4	A and C feed	13	stripper temperature
5	recycle flow	14	stripper steam flow
6	reactor feed rate	15	reactor cooling water outlet temperature
7	reactor temperature	16	separator cooling water outlet temperature
8	purge rate		
9	product separator temperature		

首先采用正常工况下的 960 个样本建立 PICA 模型, 选取的独立成分个数为 3。为了和其他方法的效果作比较, 同时建立了传统 ICA、FA、PPCA 和 PCA 模型。所有方法均选取 99% 的统计控制限, 对 TE 过程监控的结果如表 2 所示。表中给出了所有 21 种故障的漏报率, 在所有统计量中, 给出的最小漏报率已经用黑体标出, 结果表明 PICA 方法获得了最好的监控效果。其中对于故障 3~5、7、9 和 21, PICA 的  $I^2$  统计量获得了最好的监控效果, 说明这些故障引起了过程操作条件的变化, 或者使得过程运行状况出现了扰动, 但是过程变量之间的关系还是符合物料、能量等平衡关系。相比之下, 其余故障 (除了故障 12) 引起了  $MR$  统计量的超限, 说明这些故障引起了过程变量之间关系的变化, 之前建立的 PICA 模型关系已经

表 2 TE 过程监控结果(漏报率)

Table 2 Monitoring results of TE process (leak detection rate)

Mode/statistic	PICA_ $I^2$	PICA_ $MR$	ICA_ $I^2$	ICA_ $SPE$	FA_ $T^2$	FA_ $SPE$	PPCA_ $T^2$	PPCA_ $SPE$	PCA_ $T^2$	PCA_ $SPE$
1	0.266	<b>0</b>	0.460	0.008	0.006	0.003	0.010	0.006	0.345	0.005
2	0.024	<b>0.011</b>	0.018	0.019	0.040	0.016	0.065	0.016	0.078	0.016
3	<b>0.853</b>	0.956	0.965	0.989	0.989	0.998	0.999	0.989	0.983	0.976
4	<b>0.921</b>	0.975	0.993	0.989	0.988	0.999	1.000	0.994	0.991	0.983
5	<b>0.699</b>	0.750	0.785	0.781	0.761	0.819	0.798	0.781	0.794	0.763
6	0.025	<b>0</b>	0.013	0	0.009	0	0.009	0	0.034	0
7	<b>0.525</b>	0.673	0.665	0.656	0.610	0.745	0.680	0.700	0.626	0.668
8	0.093	<b>0.021</b>	0.069	0.026	0.053	0.026	0.126	0.033	0.161	0.031
9	<b>0.894</b>	0.943	0.978	0.975	0.976	0.998	1.000	0.981	0.981	0.971
10	0.508	<b>0.193</b>	0.331	0.456	0.628	0.265	0.864	0.728	0.714	0.664
11	0.741	<b>0.534</b>	0.961	0.633	0.969	0.654	0.960	0.631	0.859	0.628
12	0.080	0.018	0.065	<b>0.016</b>	0.024	0.093	0.055	0.025	0.043	0.041
13	0.058	<b>0.043</b>	0.056	0.059	0.055	0.058	0.084	0.054	0.074	0.054
14	0.108	<b>0</b>	0.171	0.001	0.905	0.000	0.180	0	0.059	0
15	<b>0.828</b>	0.946	0.968	0.971	0.983	0.996	0.995	0.986	0.964	0.969
16	<b>0.543</b>	0.645	0.721	0.836	0.733	0.869	0.939	0.858	0.794	0.828
17	0.161	<b>0.028</b>	0.286	0.056	0.395	0.045	0.246	0.055	0.174	0.056
18	0.114	<b>0.096</b>	0.104	0.103	0.109	0.101	0.111	0.101	0.115	0.101
19	0.861	<b>0.634</b>	0.915	0.831	0.975	0.845	1.000	0.756	0.991	0.746
20	0.715	<b>0.359</b>	0.521	0.553	0.431	0.489	0.918	0.695	0.831	0.664
21	<b>0.408</b>	0.550	0.415	0.791	0.479	0.700	0.819	0.584	0.685	0.591

被打破。从表 2 中可以看出，PICA 方法对 TE 过程监控的改进效果是明显的，特别对于故障 10、11、19~21，漏报率比其他方法有明显的减少。其中，对于故障 10 的监控结果如图 2 所示，从图中可以看出，特别是  $MR$  统计量，监控效果明显好于其他统计量的监控效果。

### 5 结 论

本文提出了一种基于 PICA 的过程监控方法，针对受到噪声污染的过程，在原有 PPCA 和 FA 方法的基础上，考虑了过程变量的非高斯性。相比已有方法，PICA 模型更确切地描述了过程的信息，

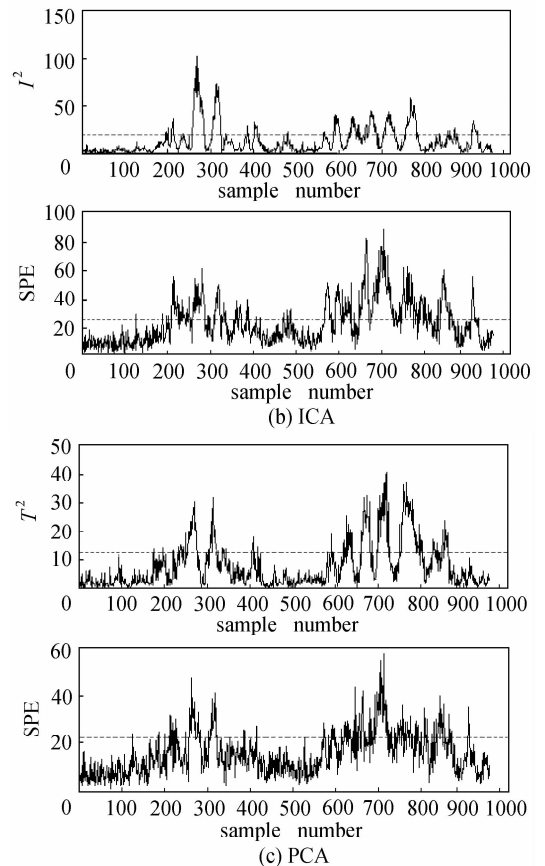
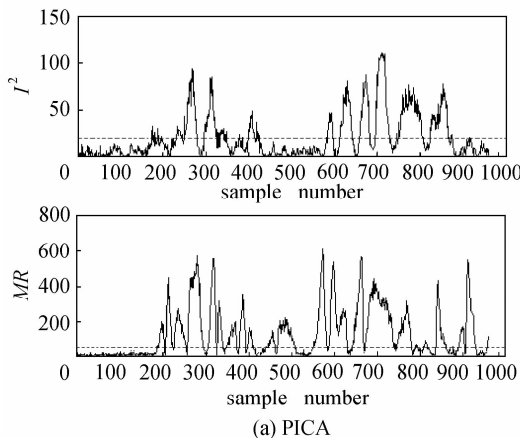


图 2 故障 10 监控结果

Fig. 2 Monitoring results of fault 10

在 TE 过程的仿真研究也说明了 PICA 方法的有效性, 过程监控的效果得到了较大的改善。但是, 本文提出的方法只适用于线性非高斯过程, 而实际工业过程往往是非线性的, 如何把该方法扩展为非线性的形式, 对下一步的工作提出了挑战。

## References

- [1] Chiang L H, Russell E L, Braatz R D. Fault Detection and Diagnosis in Industrial Systems. London: Springer-Verlag, 2001
- [2] Kim D, Lee I B. Process monitoring based on probabilistic PCA. *Chemometrics and Intelligent Laboratory Systems*, 2003, **67**: 109-123
- [3] Zhao Zhonggai (赵忠盖), Liu Fei (刘飞). Factor analysis and its application to process monitoring. *Journal of Chemical Industry and Engineering (China)* (化工学报), 2007, **58** (4): 970-974
- [4] Hyvarinen A, Oja E. Independent component analysis: algorithms and applications. *Neural Network*, 2000, **13**: 411-430
- [5] Kano M, Tanaka S, Hasebe S, Hashimoto I, Ohno H. Monitoring independent components for fault detection. *AIChE Journal*, 2003, **49** (4): 969-976
- [6] Xie L, Wu J. Gobar optimal ICA and its application in MEG data analysis. *Neurocomputing*, 2006, **69**: 2438-2442
- [7] Chen Guojin (陈国金), Liang Jun (梁军), Qian Jixin (钱积新). ICA and its application to chemical process monitoring and fault diagnosis. *Journal of Chemical Industry and Engineering (China)* (化工学报), 2003, **54** (10): 1474-1477
- [8] He Ning (何宁). Research on performance monitoring and fault diagnosis for process industry based on ICA-PCA method [D]. Hangzhou: Zhejiang University, 2004: 31-54
- [9] Hyvarinen A. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 1997, **9**: 1483-1492
- [10] Chen Q, Kruger U, Andrew T Y Leung. Regularised kernel density estimation for clustered process data. *Control Engineering Practice*, 2004, **12**: 267-274
- [11] Zhong Mingjun (钟明军). Some algorithms for independent component analysis and their application to fMRI data analysis [D]. Dalian: Dalian University of Technology, 2004: 31-42
- [12] Lee J M, Yoo C K, Lee I B. Statistical process monitoring with independent component analysis. *Journal of Process Control*, 2004, **14**: 467-485