研究简报

因子分析及其在过程监控中的应用

赵忠盖，刘 飞
（江南大学自动化研究所，江苏 无锡 214122）

关键词：因子分析；监控指标；主元分析；TE 过程
中图分类号：TP 277 **文献标识码**：A **文章编号**：0438－1157（2007）04－0970－05

Factor analysis and its application to process monitoring

ZHAO Zhonggai, LIU Fei
(*Institute of Automation, Southern Yangtze University, Wuxi* 214122, *Jiangsu, China*)

Abstract: Principal component analysis (PCA) has already been widely applied to process monitoring. However, PCA model is only a special case of probabilistic principal component analysis (PPCA) model and the latter itself is a special case of factor analysis (FA) model. Compared with PCA and PPCA models, FA model has less restriction and can do better to reveal essential features of the data. A FA model was built by the expectation maximum (EM) algorithm, and was introduced into industrial process monitoring. Monitoring indices based on FA were proposed to monitor the process factors space and residual space, respectively. A method was presented to select the number of factors by means of the property that the explanation ratio for the process information was convergent with the increasing number of factors. A contrastive study with PCA and PPCA was carried out in the Tennessee Eastman (TE) process, which showed the FA-based method's superiority either in missed detection rate or in the sensitivity for fault.

Key words: factor analysis; monitoring indices; principal component analysis; TE process

引 言

随着各种传感器和计算机技术的发展，现代工业过程中收集了大量过程变量的在线和离线数据，如何从这些数据中提取出反应过程运行状况的有用信息是工业界和学术界的一个研究热点。基于此，各种基于数据驱动的监控方法应运而生[1-2]。这类方法将高维的过程数据压缩到低维，并将获得的过程有用信息以一些统计数字量提供给现场监控人员，从而大大改善了过程的监控系统。主元分析（PCA）是一种最常用的数据驱动监控方法，其在保证数据信息丢失最小的情况下将高维空间的数据投影到低维主元空间，并通过对主元空间和噪声空间中统计量的分析，实现对过程的监控，在工业过程中得到了广泛的应用[3-4]。但是在应用 PCA 对数据进行处理时，需要假定噪声向量各向同性，且噪声方差最小，因此 PCA 模型具有很大的局限性。近年来，概率主元分析（PPCA）被提出，克服了

2006－06－07 收到初稿，2006－08－22 收到修改稿。
联系人：刘飞。**第一作者**：赵忠盖（1976—），男，博士研究生。
基金项目：新世纪优秀人才支持计划项目（NCET-05-0485）。

Received date: 2006－06－07.
Corresponding author: Prof. LIU Fei. **E－mail**: fliu@thmz.com
Foundation item: supported by the Program for New Century Excellent Talents in University (NCET-05-0485).

PCA 的部分缺点[5-6]。PPCA 定义了数据的生成模型，认为经过归一化处理的过程数据 $\boldsymbol{x}$ 是由满足标准正态分布的不相关主元变量 $\boldsymbol{t}$ 的线性组合加上满足一定正态分布的噪声 $\boldsymbol{e}$ 得到，即 $\boldsymbol{x}=\boldsymbol{P}\cdot\boldsymbol{t}+\boldsymbol{e}$，式中 $\boldsymbol{P}\in\boldsymbol{R}^{m\times k}$，为负荷向量；$k<m$ 为主元个数；$\boldsymbol{t}\sim N(0,\boldsymbol{I})$；$\boldsymbol{e}\sim N(0,\lambda\boldsymbol{I})$，其中 λ 为噪声方差值；使用期望最大化（EM）算法估计出参数 $\boldsymbol{P}$ 和 λ 后，PPCA 模型即通过生成模型表示出来。这里 λ 为常参数，$\lambda\boldsymbol{I}$ 为对角线元素相同的对角阵，因此 PPCA 需要噪声向量各向同性的假定，但是克服了要求噪声方差最小的限制，当限制噪声方差最小即 $\lambda\approx 0$ 时，PPCA 就转化成 PCA，所以说 PCA 为 PPCA 的一种特殊形式。当噪声方差矩阵（在 PPCA 中为 $\lambda\boldsymbol{I}$）为任意的对角阵时，通过生成模型建立的就是因子分析（FA）模型，也就是说 PPCA 是 FA 的一个特例。从 PPCA、PCA 和 FA 的模型建立过程可以看出，FA 没有 PPCA 和 PCA 的约束条件，因此更能反应数据中的本质特征。目前 PPCA已被引入过程监控[7]，但是更具广泛意义的 FA 的应用却还只停留在信号处理和经济分析等方面[8-9]，关于 FA 在过程监控中的应用尚没有报道。

本文提出基于 FA 的监控方法，利用 EM 算法建立过程的 FA 模型[10]，在此基础上，本文的主要贡献在于：(1) 提出了两个基于 FA 模型的监控指标，分别监控过程的受控状态以及模型关系的变化，并结合两个监控指标提出一个综合指标，简化了监控程序，减少了监控量；(2) 根据因子对过程信息的解释率收敛的特性，提出了因子个数的选取方法；(3) 将 FA 应用到 TE 过程中，对比了和 PCA、PPCA 监控方法的应用效果。

1　FA 模型的建立

设经过归一化预处理后的过程观测数据 $\boldsymbol{x}\in R^{m\times n}$，其中 m 为过程变量个数，n 为采样数。由生成模型有 $\boldsymbol{P}\in R^{m\times k}$ 为负荷向量，$k<m$ 为因子个数。令 $\boldsymbol{e}\sim N(0,\boldsymbol{\Psi})$，其中 $\boldsymbol{\Psi}$ 为任意对角矩阵，则所有因子的联合概率密度函数为

$$p(\boldsymbol{t})=(2\pi)^{-k/2}\exp\left\{-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\right\}$$

观测数据关于因子的条件分布可以表示为

$$p(\boldsymbol{x}\mid\boldsymbol{t})=(2\pi)^{-m/2}\mid\boldsymbol{\Psi}\mid^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{P}\boldsymbol{t})^{\mathrm{T}}\boldsymbol{\Psi}^{-1}(\boldsymbol{x}-\boldsymbol{P}\boldsymbol{t})\right\}$$

观测数据的分布为

$$p(\boldsymbol{x})=\int p(\boldsymbol{x}\mid\boldsymbol{t})p(\boldsymbol{t})\mathrm{d}\boldsymbol{x}=$$

$$(2\pi)^{-m/2}\mid\boldsymbol{C}\mid^{-1/2}\exp\left\{-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{x}\right\}\tag{1}$$

其中，$\boldsymbol{C}=\boldsymbol{\Psi}+\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}$ 为观测数据的方差。由贝叶斯定理得因子的后验分布

$$p(\boldsymbol{t}\mid\boldsymbol{x})=(2\pi)^{-\frac{k}{2}}\mid\boldsymbol{M}\mid^{1/2}\exp\left\{-\frac{1}{2}[\boldsymbol{t}-\beta\boldsymbol{x}]^{\mathrm{T}}(\boldsymbol{M})[\boldsymbol{t}-\beta\boldsymbol{x}]\right\}\tag{2}$$

式中

$$\boldsymbol{M}^{-1}=\boldsymbol{I}-\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{\Psi}+\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{P},\beta=\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{\Psi}+\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}})^{-1}$$

根据生成模型，$\boldsymbol{x}\sim N(0,\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}+\boldsymbol{\Psi})$，如果能够确定出 $\boldsymbol{P}$ 和 $\boldsymbol{\Psi}$，则变量的分布函数就能确定。迭代 EM 算法即是一种有效的算法，一般分为 E-（求期望）步和 M-（最大化）步。EM 算法假定因子向量为遗失数据，完整的数据为 $(\boldsymbol{x},\boldsymbol{t})$，则所有完整数据的 lg 概率函数为

$$\boldsymbol{L}_{\mathrm{C}}=\sum_{i=1}^{n}\ln\{p(\boldsymbol{x}_i,\boldsymbol{t}_i)\}$$

其中

$$p(\boldsymbol{x}_i,\boldsymbol{t}_i)=p(\boldsymbol{x}_i\mid\boldsymbol{t}_i)p(\boldsymbol{t}_i)=(2\pi)^{-(m+k)/2}\mid\boldsymbol{\Psi}\mid^{-1/2}\times$$
$$\exp\left\{-\frac{1}{2}(\boldsymbol{x}_i-\boldsymbol{P}\boldsymbol{t}_i)^{\mathrm{T}}\boldsymbol{\Psi}^{-1}(\boldsymbol{x}_i-\boldsymbol{P}\boldsymbol{t}_i)\right\}\exp\left\{-\frac{1}{2}\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_i\right\}\tag{3}$$

在 E-步，根据 $p(\boldsymbol{t}_i\mid\boldsymbol{x}_i,\boldsymbol{P},\boldsymbol{\Psi})$ 的期望求出 $\boldsymbol{L}_{\mathrm{C}}$ 的期望值 $\langle\boldsymbol{L}_{\mathrm{C}}\rangle$ 如下

$$\langle\boldsymbol{L}_{\mathrm{C}}\rangle=c-(n/2)\lg\boldsymbol{\Psi}-\sum_{i=1}^{n}\{(1/2)\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{x}_i-$$
$$\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{P}E[\boldsymbol{t}_i\mid\boldsymbol{x}_i]+(1/2)tr(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{P}E[\boldsymbol{t}_i\boldsymbol{t}_i^{\mathrm{T}}\mid\boldsymbol{x}_i]\}\tag{4}$$

式中

$$E[\boldsymbol{t}_i\mid\boldsymbol{x}_i]=\beta\boldsymbol{x}_i$$
$$E[\boldsymbol{t}_i\boldsymbol{t}_i^{\mathrm{T}}\mid\boldsymbol{x}_i]=\boldsymbol{I}-\beta\boldsymbol{P}+\beta\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}\beta^{\mathrm{T}}$$

在 M-步，求 $\langle\boldsymbol{L}_{\mathrm{C}}\rangle$ 关于 $\boldsymbol{P}$ 和 $\boldsymbol{\Psi}$ 的偏导，并令其等于 0，求得使 $\langle\boldsymbol{L}_{\mathrm{C}}\rangle$ 达到最大的新参数值 $\tilde{\boldsymbol{P}}$ 和 $\tilde{\boldsymbol{\Psi}}$

$$\tilde{\boldsymbol{P}}=(\sum_{i=1}^{n}\boldsymbol{x}_iE[\boldsymbol{t}_i\mid\boldsymbol{x}_i]^{\mathrm{T}})(\sum_{i=1}^{n}E[\boldsymbol{t}_i\boldsymbol{t}_i^{\mathrm{T}}\mid\boldsymbol{x}_i])^{-1}\tag{5}$$

$$\tilde{\boldsymbol{\Psi}}=(1/n)\mathrm{diag}(\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}-\tilde{\boldsymbol{P}}E[\boldsymbol{t}_i\mid\boldsymbol{x}_i]\boldsymbol{x}_i)\tag{6}$$

反复迭代式（5）、式（6）直到收敛，得到 $\boldsymbol{P}$ 和 $\boldsymbol{\Psi}$，至此 FA 模型建立完成。

2　基于 FA 模型的过程监控

2.1　因子个数的选择

因子是反应数据信息的主要成分，当数据中大部分的信息已经被解释后，因子数目的增加不会引起解释的数据信息太大的变化。利用这个特征，本文提出一种因子个数选择方法。数据中包含的所有变量的总信息是 sum［diag（$\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}+\boldsymbol{\Psi}$）］，为常数，

其中 sum（）为求和算子，diag（）为取对角线数值算子。假设因子个数为 l，通过 EM 算法得出参数 $\boldsymbol{P}$，则当前因子中包含的信息为 sum［diag（$\boldsymbol{PP}^{\mathrm{T}}$）］，因此在因子个数为 l 时因子对原始数据信息的解释率如下

$$\boldsymbol{Q}(l)=\mathrm{sum}[\mathrm{diag}(\boldsymbol{PP}^{\mathrm{T}})]/\mathrm{sum}[\mathrm{diag}(\boldsymbol{PP}^{\mathrm{T}}+\boldsymbol{\Psi})] \quad (7)$$

逐个增大因子个数，并重新由 EM 算法估计出参数 $\boldsymbol{P}$，求出取不同因子个数下因子对过程数据的解释率，当 $\boldsymbol{Q}$（l+1）$-\boldsymbol{Q}$（l）$<\varepsilon$ 时，其中 ε 为阈值，因子对过程数据的解释率收敛，则选择此时的 l 即为因子个数 k。

2.2 监控指标

2.2.1 对因子空间的监控（T^2 的扩展指标 GT^2）

因子和主元都反应了影响过程变化的主要因素，对因子空间的监控与 PCA 中对主元空间的监控一样，都是对过程数据的主要成分进行监控，从而评估过程的受控状态，如同 PCA 监控中的 T^2 监控指标。当因子或主元变化不大时，过程处于受控状态；否则，过程的主要因素变化，过程出现异常。按照 T^2 定义一个 T^2 的扩展监控指标 GT^2。通过生成模型只能得出因子满足一定的概率分布，其在任意采样时刻（不失一般性，设为第 i 时刻）的确切值没法确定，仿效文献［4］，采用 $\boldsymbol{t}_i$ 的估计值 $\hat{\boldsymbol{t}}_i=\boldsymbol{E}[\boldsymbol{t}_i\mid\boldsymbol{x}_i]=\beta\boldsymbol{x}_i$ 代替 $\boldsymbol{t}_i$，则 GT^2 定义为

$$GT^2=\|\hat{\boldsymbol{t}}_i\|^2=\boldsymbol{x}_i^{\mathrm{T}}\beta^{\mathrm{T}}\beta\boldsymbol{x}_i\leqslant\chi^2_{(\alpha,k)} \quad (8)$$

式中 $\chi^2_{(\alpha,k)}$ 为置信度为 α，自由度为 k 下的 χ^2 分布的值。$\|\hat{\boldsymbol{t}}_i\|^2$ 满足 χ^2 分布，因此 $\chi^2_{(\alpha,k)}$ 为 $\|\hat{\boldsymbol{t}}_i\|^2$ 在置信度为 α 下的控制限。以下各监控指标中不等式后面的一项即为该监控指标的控制限。建立模型时，所有过程变量都预先经过归一化处理，因此其均值向量为 0 向量，而因子是过程变量的线性组合，因此因子的均值向量也为 0 向量，则由式（8）可得，GT^2 即为因子对中心点（均值向量 0）的欧几里德距离，但是因子满足标准正态分布，因此 GT^2 也是对因子与其中心点的马氏距离的检验。GT^2 用来检验过程的工作条件是否发生改变。

2.2.2 对噪声空间的监控（SPE 的扩展指标 GSPE） 噪声反应了过程变量与模型的拟合程度，当噪声变量小时，表示当前过程变量严格符合 FA 模型的关系，反之则认为当前过程变量不符合 FA 模型，过程变量之间的关系改变，过程出现了故障。PCA 监控中的 SPE 监控指标是对噪声强度的量度，但是与 SPE 采用欧几里德距离作为量度不同的是，本文采取对噪声与其中心（0 向量）的马氏距离作为噪声空间的监控指标（定义为 SPE 的扩展指标 GSPE）。通过模型只能得到噪声向量的正态分布密度函数，但是其在每个采样时刻的值没法确定，因此采用 $\boldsymbol{e}_i$ 基于测量值 $\boldsymbol{x}_i$ 的估计值 $\hat{\boldsymbol{e}}_i=\boldsymbol{E}[\boldsymbol{e}_i\mid\boldsymbol{x}_i]=[\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}+\boldsymbol{\Psi})^{-1}\boldsymbol{P}^{\mathrm{T}}]\boldsymbol{x}_i$ 代替 $\boldsymbol{e}_i$，则 GSPE 定义如下

$$GSPE=\|\boldsymbol{\Psi}^{-1/2}\hat{\boldsymbol{e}}_i\|^2\leqslant\chi^2_{(\alpha,m)} \quad (9)$$

在基于 PPCA 的监控中，假定噪声各向同性，即 $\boldsymbol{\Psi}=\lambda\boldsymbol{I}$，因此对噪声的监控虽然采用的是马氏距离，但是其实质还是欧几里德距离。而在式（9）中，噪声没有了各向同性的限制，估计出的 $\boldsymbol{\Psi}$ 反应了噪声的方差，假设 $\boldsymbol{\Psi}$ 对角线上的元素分别为 λ_1，λ_2，…，λ_m，则有 $[\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}]_{ii}=\lambda_i$，且 $[\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}]_{ij}=0$，$i\neq j$，因此 GSPE 充分考虑了噪声向量与建模噪声结合的紧密程度，而这一点对满足正态分布的过程数据而言，恰恰是能将其与正常数据分开的关键，也就是能够更好的对数据是否满足过程模型进行检测。

2.2.3 综合监控指标（ST） 对因子和噪声的监控指标采用的量度都是马氏距离，可以联合起来对过程运行状况进行判断。根据生成模型，$\boldsymbol{x}_i$ 是由因子和噪声线性混合组成的，直接对测量值进行检测能反映对因子和噪声的检测结果。$\boldsymbol{x}_i\sim N(0,\boldsymbol{PP}^{\mathrm{T}}+\boldsymbol{\Psi})$，将 $\boldsymbol{x}_i$ 对其中心（0 向量）的马氏距离作为综合监控指标 ST，如下

$$ST=\|(\boldsymbol{PP}^{\mathrm{T}}+\boldsymbol{\Psi})^{-1/2}\boldsymbol{x}_i\|^2\leqslant\chi^2_{(\alpha,m)} \quad (10)$$

式（10）综合了式（8）、式（9），为减少监控量，并且为了对比某些采样时刻过程运行状况的好坏，在过程监控中，可以单独使用 ST 进行监控。在 ST 值较大的采样时刻，过程的运行状况往往要比在 ST 值小的采样时刻糟糕。

3 仿真实例

TE 过程是一个公认的比较各种控制和监控方案的平台，共设有 52 个过程变量，21 种故障模式。运行一次的时间为 48 h，采样间隔时间为 3 min，因此运行一次共能采集 960 组数据。过程的详细描述，工艺流程图以及其故障形式的具体介绍和使用 PCA、动态 PCA 等对各故障的监控效果参见文献［11］。

采用 960 组正常数据建立 FA 的模型。首先对正常数据进行归一化处理，然后通过 EM 算法分别估计出因子个数从 1 到 30 的 FA 模型参数 $\boldsymbol{P}$ 和 $\boldsymbol{\Psi}$，计算出取不同因子个数下因子对建模数据的信息解释率，如图 1 所示。

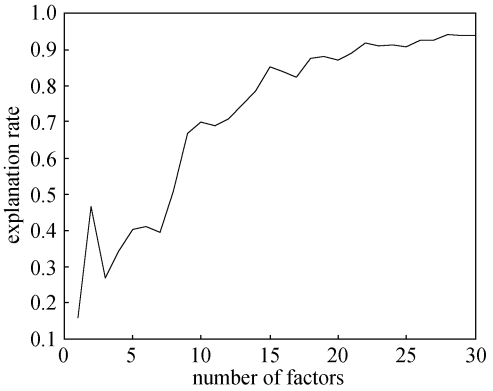

图 1　不同因子个数下因子对过程信息的解释率

Fig. 1　Explanation rate for process information under different number of factors

由图 1 可以看出，在因子个数大于 15 时，因子对建模数据的解释率收敛，因此取因子个数为 15。需要说明的是，在取不同因子个数时，FA 模型都需要通过 EM 算法得到，但是 EM 算法为局部优化算法，所以解释率随着因子个数的增加会有些波动，但是解释率的变化趋势是上升的。

以故障 5 为例，对比 FA、PCA 和 PPCA 的监控效果。冷凝器冷却水入口温度在 8 h，也就是采样时刻 160 的阶跃变化引起故障 5，故障引起冷凝器冷却水流量（过程变量 52）也产生阶跃变化。故障发生后，冷凝器出口流量增加，导致汽液分离器的温度增加，且分离器冷却水出口温度（过程变量 22）也有增加。取因子个数为 15，建立过程的 FA 模型，将经过归一化处理的故障数据代入 FA 模型，通过式（8）和式（9）得到的 GSPE 和 GT^2 的监控图见图 2。

由图 2 可知，故障 5 的 GSPE 值和 GT^2 值分别在采样时刻 162 和 161 超出各自的控制限，表明从采样时刻 162 开始过程变量不再符合模型关系，在第 161 个采样时刻处于不受控状态，与故障 5 引入的时间和状况相符。另外，出现故障时，过程的 GSPE 值和 GT^2 值远远大于正常情况的值，表明故障被明显区分开来，便于监控人员做出正确的判断。将故障数据代入式（10）得到 ST 监控指标图见图 3。

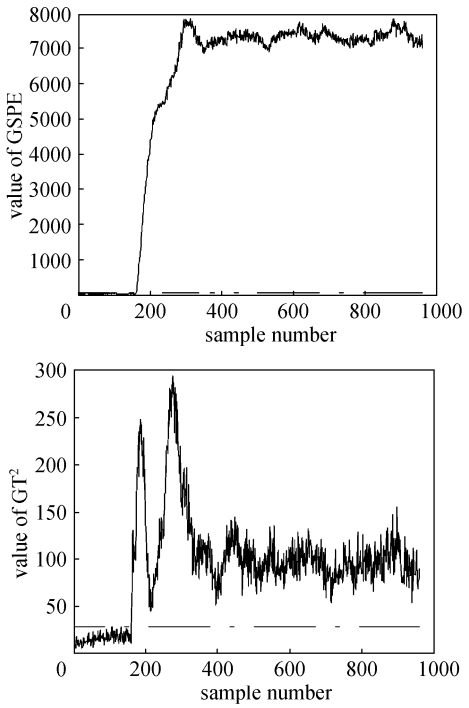

图 2　基于 FA 对故障 5 的 GSPE 和 GT^2 监控图

Fig. 2　GSPE and GT^2 monitoring charts for fault 5 using FA

由图 3，ST 值在 161 时刻超出控制限，可以判断出过程在第 161 采样时刻出现故障，与通过图 2 两张监控图的检测结果一致，并且因为 ST 综合了 GSPE 和 GT^2 的监控效果，因此故障状况下的 ST 值很明显大于正常状况下的值。在实际故障检测中，现场操作人员只需要监控图 3 即可评价过程的运行状况，减少了一半的监控工作量，这一点在多个模型或者分段监控的情况下显得尤为重要。

基于 PPCA 监控中，对噪声和主元的监控分别相当于 PCA 中的 SPE 和 T^2 监控指标，而对过程变量白化值的监控相当于基于 FA 监控中的 ST

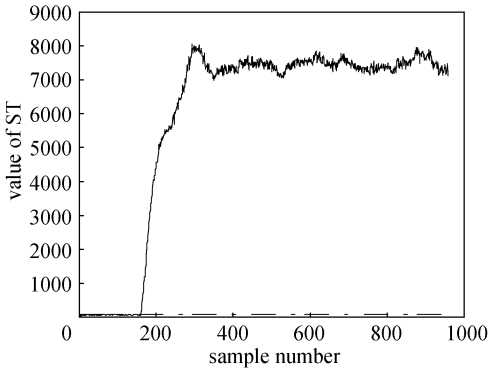

图 3　基于 FA 对故障 5 的 ST 监控图

Fig. 3　ST monitoring chart for fault 5 based on FA

指标。选择主元个数为 15，按文献［7］建立模型并进行监控。由 PPCA 得到的监控效果图见图 4。

由图 4 可以看出，由 PPCA 监控方法，过程在 163 时刻不再符合模型，在 172 时刻处于不受控状态，而过程变量白化值在 162 时刻超出控制限。对比图 2～图 4，基于 FA 的方法不仅对故障更为敏感，更能将故障与正常情况区分开，能更及时反应出故障状况，而且没有漏检率。另外，在 FA 监控中，故障 5 在监控图上以故障形式持续到过程结束，也就是故障持续时间很长，这对现场监控人员及时找出故障原因，实现故障诊断很关键。基于 PCA 对 TE 过程故障 5 的监控效果图及说明参见文献［11］，效果明显没有 FA 好，本文不再赘述。

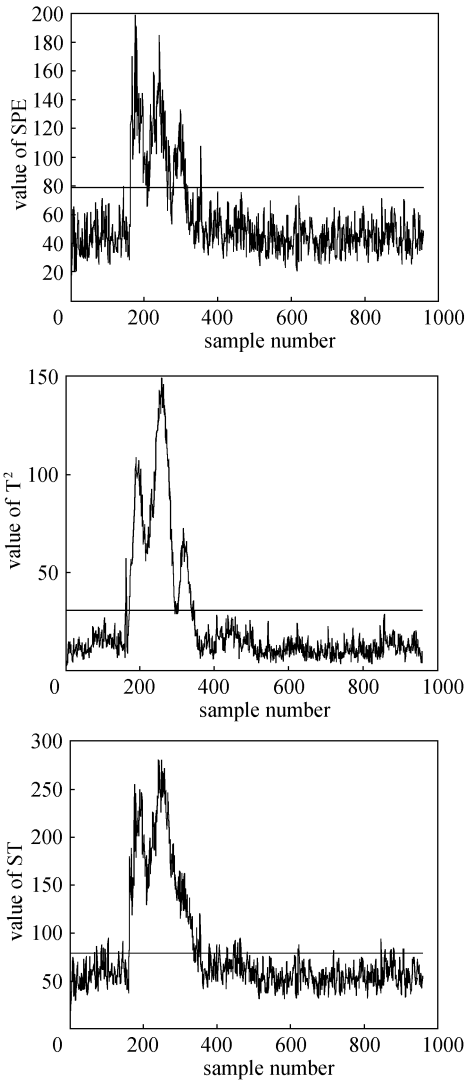

图 4 基于 PPCA 对故障 5 的监控图

Fig. 4 Monitoring charts for fault 5 based on PPCA

4 结 论

提出了一种因子个数的选择方法。在 FA 模型的基础上，提出了过程受控状态，过程变量是否符合模型的监控指标，并通过马氏距离将这两个监控指标合成一个对测量值监控的综合指标。最后将 FA 模型引入了对 TE 过程的监控中，从与 PPCA、PCA 的监控效果对比情况可以看出，在故障检测的及时性、漏检率、故障的持续时间等方面都比后两种方法好，其原因在于 FA 的约束条件比 PPCA 和 PCA 少，因此更能反映数据的内部特征，基于 FA 的监控方法能有效地评价过程的运行状况。

References

［1］ Piovoso M J，Kosanovich K A. Applications of multivariate statistical methods to process monitoring and controller design. *Int. J. Control*，1994，**59**：743-765

［2］ Kourti Thedora. Application of latent variable methods to process control and multivariate statistical process control in industry. *Int. J. Adapt. Control Signal Process*，2005，**19**：213-246

［3］ Wang Haiqing（王海清），Song Zhihuan（宋执环），Li Ping（李平）. Improved PCA with application to process monitoring and fault diagnosis. *Journal of Chemical Industry and Engineering*（*China*）（化工学报），2001，**52**：471-475

［4］ Kramer M A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.*，1991，**37**：233-243

［5］ Michael E T. Probabilistic principle component analysis. *Journal of the Royal Statistical Society*，1999，**3**（1）：71-86

［6］ Michael E T，Christopher M B. Mixtures of principal component analyzers. *Neural Computation*，1999，**11**（2）：443-482

［7］ Kim Dongsoon，Lee In-Beum. Process monitoring based on probabilistic PCA. *Chemometrics and Intelligent Laboratory Systems*，2003，**67**：109-123

［8］ Wang Hsiao-Fan，Kuo Ching-Yi. Factor analysis in data mining. *Computers and Mathematics with Applications*，2004，**40**（10/11）：1765-1778

［9］ Jean Boivin，Serena Ng. Are more data always better for factor analysis? *Journal of Econometrics*，2006，**132**（1）：169-194

［10］ Ghahramani Zoubin，Geoffrey E H. The EM algorithm for mixtures of factor analyzers：Technical Report CRG-TR-96-1［R］. University of Toronto，1997

［11］ Chiang L H，Russell E L，Braatz R. D. Fault Detection and Diagnosis in Industrial Systems. London：Springer-Verlag London Limited，2001：103-120