FERMENTATION BATCH PROCESS MONITORING BY STEP-BY-STEP ADAPTIVE MPCA

Ning He, Lei Xie, Shu-qing Wang, Jian-ming Zhang

National Key Laboratory of Industrial Control Technology,
Zhejiang University, Hangzhou 310027, People's Republic of China

Abstract: Multi-way principal component analysis (MPCA) has been successfully applied to the monitoring of batch and semi-batch processes in most chemical industry. A new approach is presented to overcome the method MPCA's need for estimating or filling in the unknown part of the process variable trajectory deviations from the current time until the end. The approach is based on the Multi-block PCA method and processes the data in a sequential and adaptive manner. The adaptive rate is easily controlled by a parameter that represents the similarity between current and past data. The method is evaluated on industrial fermentation process data and is compared to the traditional MPCA. The method may have significant benefit when monitoring multi-stage batch process where the latent vector structure can change at several points during the batch.

Keyword: Statistical Process Control, Principal Component Analysis, Adaptive Algorithm

1. INTRODUCTION

Batch and semi-batch process play an important role in the chemical industry due to their low volume and high value products. In most of these processes product quality variables are only measured after the end of each batch, often hours later in a quality control laboratory. This makes it difficult to control the product quality or to monitor the progress of these batch processes. MacGregor (Kosanovich *et al*., 1994, Nomikos *et al* ,1994, Stefan Rännar, 1998), have presented very powerful process analysis, monitoring and diagnostic procedures which utilize these process variables trajectory data. These procedures, based on multi-way PCA (MPCA) method (Wold *et al*, 1982, 1994a), are now being widely adopted by the batch chemical industry.

The methods have proven to be very efficient for analysis historical data from past production and diagnosing operating problems. But when monitoring a new batch, the MPCA model, which assumes the complete history of the batch data, cannot be used directly. At any point during the batch, data on the deviations of the variables from its average trajectories for the remainder of the batch is not yet available. MacGregor and Nomikos (Nomikos *et al*, 1995) have proposed several approaches to handle this problem that have worked quite well in practice. However, it would be nice to have MPCA monitoring approaches that do not depend on having to fill in these missing data.

In this paper, we present a modification of the monitoring method that does not require estimates of the data for the uncompleted portion of the batch. The approach is based on a multi-block PCA algorithm that processes the data in a series manner. When monitoring future batches, one need only store the loading matrices for the local model at each point in time and the score vector from past. This step-by-step adaptive approach only requires fewer latent variable dimensions and appears to work as well as the existing methods. Application in monitoring an industrial fermentation process reveals that the proposed method gives more objective appraisal for new batch and may offer potential advantages in situation where the batch has several stages.

2. MULTI-WAY PRINCIPAL COMPONENT ANALYSIS

Multi-way principal component analysis (MPCA) (Nomikos *et al,* 1994) is an extension of conventional PCA to handle data in three-dimensional arrays. These data, which are

collected from batch processes, are organized into an array $\boldsymbol{X}$ of three-dimension ($I*J*K$), where, I is the number of batches, J is the number of variables, and K is the number of time samples over the duration of the batch. As illustrated in Fig.1, the array X can be unfolded in such a way as to put each of its vertical slices ($I*J$) contain the observed variables for all batches at a given time interval side by side to the right. The result, a wide and short two-dimensional matrix has dimensions ($I*JK$). MPCA is equivalent to performing ordinary PCA on unfolded $\boldsymbol{X}$ and it explains the variation of variables about their mean trajectories.

MPCA decomposes X as follow,

$$\boldsymbol{X} = \sum_{r=1}^{R} \boldsymbol{t}_r \otimes \boldsymbol{p}_r + \boldsymbol{E} \,, \tag{1}$$

where R is the number of principal components used in the analysis, $\boldsymbol{t}$ is score vectors and $\boldsymbol{p}$ is loading matrices. $\boldsymbol{E}$ is residual matrix. It accomplishes this decomposition in accordance with the principles of PCA and separates the data in an optimal way into two parts.

This decomposition represents a new coordination system obtained by rotating the original variables and projecting the data into the reduced space defined by the first few principal components, where the data are described adequately and in a simpler and more meaningful way (Wold *et al*, 1978). By doing this, MPCA utilizes not just the magnitude of the deviation of each variable from its mean trajectory but also the correlations among them. The appropriate number of principal components may be determined by cross- validation (Jackson.,1991).

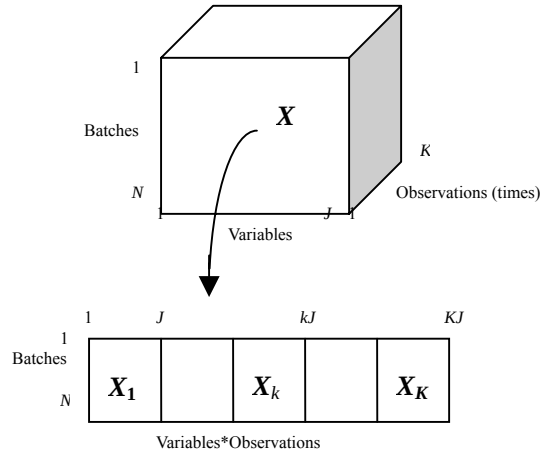

Fig.1. Overview of the MPCA method.

The MPCA algorithm derives directly from the nonlinear iterative partial least squares (NIPALS). NIPALS is a simple, fast and effective algorithm to extract the principal components in a sequential manner and is a variant of the power method for calculating eigenvectors of a matrix. The algorithm proceeds as follow. At first, scale $\boldsymbol{X}$ by subtracting from each column its mean and dividing by standard deviation, choose arbitrary a column of $\boldsymbol{X}$ as $\boldsymbol{t}$, arrange $\boldsymbol{E} = \boldsymbol{X}$, this vector is multiplied by score vector $\boldsymbol{t}$ to give the loading vector $\boldsymbol{P}$ and normalizes $\boldsymbol{P}$ to have length one.

$$\boldsymbol{P} = \boldsymbol{E}^{'}.\boldsymbol{t} \,, \tag{2}$$

where, $\boldsymbol{E}^{'}(\boldsymbol{j},\boldsymbol{i}.\boldsymbol{k}) = \boldsymbol{E}(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$,

$$\boldsymbol{P}(\boldsymbol{k},\boldsymbol{j}) = \sum_{i=1}^{I} \boldsymbol{E}^{'}(\boldsymbol{j},\boldsymbol{i},\boldsymbol{k})\boldsymbol{t}(\boldsymbol{i}) \,,$$

$$\boldsymbol{P} = \boldsymbol{P} \big/ \left\| \boldsymbol{P} \right\| \,, \tag{3}$$

$$\left\| \boldsymbol{P} \right\| = \sqrt{\sum_{k=1}^{K} \sum_{j=1}^{J} \boldsymbol{P}(\boldsymbol{k},\boldsymbol{j})^2} \,,$$

$$\boldsymbol{t} = \boldsymbol{E} \circ \boldsymbol{P} \,, \tag{4}$$

$$\boldsymbol{t}(\boldsymbol{i}) = \sum_{j=1}^{J} \sum_{k=1}^{K} \boldsymbol{E}(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})\boldsymbol{P}(\boldsymbol{k},\boldsymbol{j}) \,,$$

$$\boldsymbol{E} = \boldsymbol{E} - \boldsymbol{t} \otimes \boldsymbol{P} \,, \tag{5}$$

$$\boldsymbol{X} = \boldsymbol{t} \otimes \boldsymbol{P}, \boldsymbol{X}(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = \boldsymbol{t}(\boldsymbol{i})\boldsymbol{P}(\boldsymbol{j},\boldsymbol{k}) \,,$$

The new score vector t is calculated and the convergence of t is checked. If t has converged then to equation (5). Otherwise, one iterates equations (2)-(4).

Two approach for MPCA monitoring of process are used, one is the squared prediction error (SPE) (Jackson, 1991) of residual space as follows,

$$Q = \boldsymbol{E}^{'}\boldsymbol{E} \,, \tag{6}$$

The critical value for Q is,

$$Q_\alpha = \theta_1 \left[\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{1/h_0} \,, \tag{7}$$

where,

$$\theta_1 = \sum_{i=R+1}^{J} \lambda_i \,,$$

$$\theta_2 = \sum_{i=R+1}^{J} \lambda_i^2 \,,$$

$$\theta_3 = \sum_{i=R+1}^{J} \lambda_i^3 \,,$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \,,$$

where, λ_i is the eigenvalue of $\boldsymbol{X}$'s covariance.

Another index is the Hotelling T^2 test. T^2 value can be expressed as,

$$T_k^{\,2} = \boldsymbol{t_k}^{'} \boldsymbol{S}^{-1} \boldsymbol{t_k} \,, \tag{8}$$

The critical value for T^2 is,

$$T_{k,\alpha}^2 = \frac{R(I-1)}{I-R} F_\alpha (R, I-R) \,, \tag{9}$$

If an observation vector that produce a value of T_k^2 greater than $T_{k,\alpha}^2$, the process will be out of control.

However, one problem arises when MPCA monitors a new batch, only the data up to the present time is available and nothing is known about the remainder of the batch. In order to calculate the score vectors for the present batch, the missing data will have to be filled with some assumed approaches. MacGregor and Nomikos (Nomikos *et al*, 1995) suggests several

methods to solve this, usually, two methods are used. The first is put zeros in the vector for all remaining missing batch. It assumes that the process will continue as a normal batch from the present time until the end of the batch. The second approach fills in the future data with the current observed value that has been normalized. All these method have been seen to work in practice. However, it would be more satisfying if one would not have to make such assumptions when monitoring an on going batch. In the following sections, it is shown that by using a step-by-step adaptive MPCA, this can be accomplished.

3. STEP-BY-STEP ADAPTIVE MPCA

In the proposed Step-by-Step Adaptive MPCA algorithm the ***X***-block is divided into a number of blocks, which are the time slices from the three-dimensional batch data matrix and each block has *N* batches and *J* variables. The different between the step-by-step adaptive MPCA from ordinary PCA algorithm is that it only looks at one time slice each time rather than all of the blocks at once.

The step-by-step adaptive MPCA algorithm proceeds as follows. The initial step is to calculate one PCA model for the first time slice (k=1), giving the first block's score vector $\boldsymbol{T}_1$ and loading vector $\boldsymbol{P}_1$, then lead in a forgetting factor β_1 to decide l columns ahead of the score vector $\boldsymbol{T}_1$ should be retained to make a new score vector $\boldsymbol{T}_{1new}$ ($\boldsymbol{T}_{1new}$ captures greater than β_1 percentages of total variations among $\boldsymbol{T}_1$), the $\boldsymbol{T}_{1new}$ will join in the PCA model building of the next time slice. The value of l can be express as,

$$\sum_{i=1}^{l} \lambda_{i1} \geq \beta_1 \sum_{i=1}^{J} \lambda_{i1} , \qquad (10)$$

The algorithm begins with the second time slice and continues for the rest of the duration (k= 2, … K).

The first step in building the PCA model for time k is to combine the previous score $\boldsymbol{T}_{(k-1)new}$ vector that summarizes the recent prior history and the present $\boldsymbol{X}_k$ matrix together, then apply the PCA to $\boldsymbol{D}_k$, $\boldsymbol{T}_k$ is relative score vectors and $\boldsymbol{P}_k$ is loading vectors respectively.

$$\boldsymbol{D}_k = [\boldsymbol{T}_{(k-1)new}, \boldsymbol{X}_k], \qquad (11)$$

$$\boldsymbol{T}_k = \boldsymbol{D}_k \boldsymbol{P}_k , \qquad (12)$$

The second step is, through (13) decided the first l score vectors of $\boldsymbol{T}_k$ compose $\boldsymbol{T}_{knew}$, that will take part in the next time slice.

$$\sum_{i=1}^{l} \lambda_{ik} \geq \beta_k \sum_{i=1}^{J} \lambda_{ik}, 0 \leq \beta_k \leq 1, \qquad (13)$$

where λ_{ik} is the eigenvalue of the D_k 's covariance matrix.

The third step: define $k = k+1$. If k less than K then return the first step. Otherwise, ends.

During this algorithm, β_k is the forgetting factor. It controls the number of the columns of the score vector $\boldsymbol{T}_k$. A higher value of β_k will put more weight on the previous history and the model will adapt slowly, while a low value of β_k will make the model adapt fast. Setting β_k to zero will be equation to calculating separate PCA modes since it uses no memory of previous history of the batches. The value of β_k depends on the type of process to be monitored.

The criteria of choosing the forgetting factor β_k is described as follows, first lead-in the similarity (Manabu kano *et al*, 2002) between time slice k and k+1 of data sets.

When two data sets are similar to each other, the coefficient of similarity must be near one; the corresponding forgetting factor β_k is also high. However, the coefficient of similarity should be near zero when data sets are quite different from each other, the value of β_k is also low.

When the calculation as mentioned above is finished, K PCA models can be obtained. The Ith model relates to the Ith sample time of the batch process. In the monitoring phase two statistics with complementary information, the squared prediction error (SPE) and the Hotelling's statistics (T^2)(Jackson, 1991) can be used.

4. APPLICATION

In this section, an industrial typical multi-stage streptomycin fermentation process will be investigated to evaluate the performance of Step-by-Step Adaptive MPCA. The available on-line measurements for the employed process are shown in table1 and all these measurements are obtained at regular sample intervals.

Table 1 Online measurements obtained from the streptomycin fermentation process

1	Temperature	5	Nitrogen(N) Concentration
2	Air flow	6	Sugar Concentration
3	PH	7	Product Concentration
4	Viscosity		

With both good final product concentration (26000~30000) and normal variable trajectory, 20 good batches are used to train the model for process monitoring. Both variables are centred about their average trajectory and scaled to have unit variance prior to the analysis.

A special batch (the initial value of *PH* of this batch is abnormally high, due to the operator's operation,

the final product concentration is 24920) is employed for comparison between the proposed method and MacGregor's method (Nomikos *et al*, 1995). The Squared Predictive Error (SPE) (solid line) monitoring results are listed in following figures, where the dotted line represents the 95% SPE confidence level.

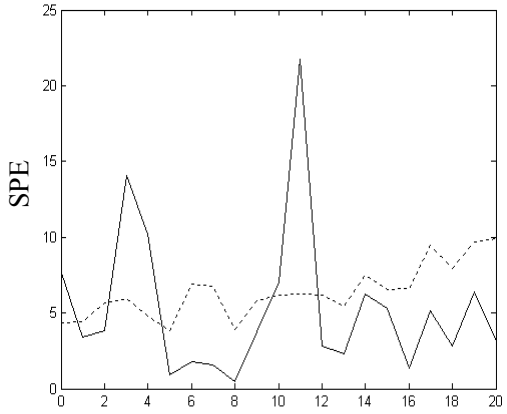

Fig 2 Monitoring of Special Batch Using Step-By-Step MPCA

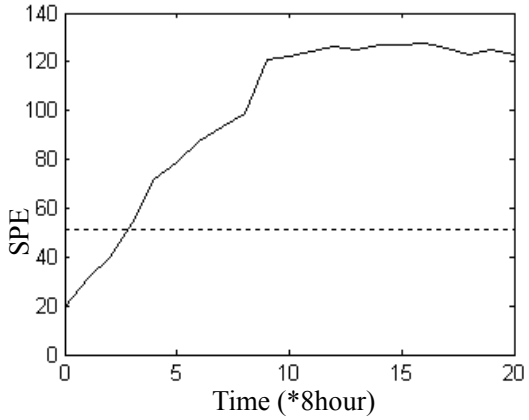

Fig 3 Monitoring of Special Batch Using MPCA (Method 1 to fill future data)

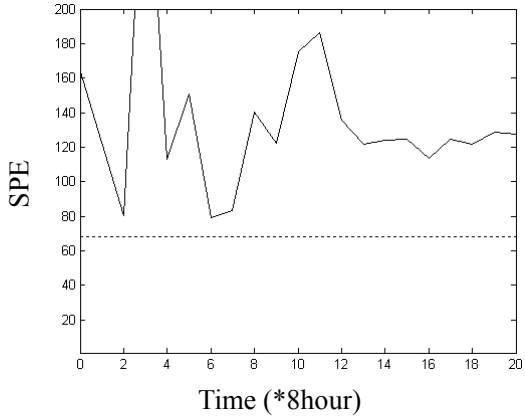

Fig 4 Monitoring of Special Batch Using MPCA (Method 2 to fill future data)

From figure 2, it's clear that Step-by-Step MPCA detects abnormal behaviours at the beginning of the bad batch. As we know, the initial condition of fermentation process has large effect on the final product quality, thus it's quite important to detect abnormal initial condition as soon as possible. Due to operators' efforts and compensations for the bad initial condition, this bad batch is drawn back to sub-normal operation trajectory. Compared with Step-by-Step MPCA, ignoring the multi-stage characteristics and taking the batch as whole, traditional MPCA method neglects the efforts of operators and evaluates this batch badly. Furthermore, using the first method to fill future data, MPCA is less sensitive to abnormal operation and the second method is too sensitive.

5. CONCLUSIONS

In this paper, a new Step-by-Step Adaptive MPCA algorithm has been presented for the purpose of monitoring batch processes. One advantage of this approach is that in the monitoring phase, there is no need to assume anything about the future deviations of the going batch from the normal trajectory. In traditional MPCA algorithm, one has to assume the operation of the future of the batch in order to full-fill the data vector from the current time period until the end of the batch. The step-by-step adaptive MPCA only works with the present and past data of the batch. Thus avoids the filling procedure. Another advantage is the model's ability of adapting to different stages of the batch process making it very suitable for monitoring multistage batch processes. In the industrial fermentation example involved in this paper, the step-by-step adaptive MPCA has represented its advantage. If one has a process with even more stages this method could be more advantageous since it will reduce the number of variables for monitoring and thereby simplifying the presentation of the monitoring results.

ACKNOWLEDGEMENTS

This project is supported by National High-Tech Program of China under grant No. 2001AA413110.

REFERENCES

Jackson J.E (1991), *A User's Guide to Principal Components*, Wiley Inter-science, New York,

Kosanovich K. A., Piovoso M. J (1994), Multi-way PCA applied to an industrial batch process, *In Am. Control conference,* **J.** 1294-1298

Nomikos P., MacGregor J. F (1994), Monitoring of batch processes using multi-way principal components analysis, *In Am. Inst. Chem. Eng.* **J. 40** 1361-1375

Nomikos P., MacGregor J. F (1995), Multivariate SPC charts for batch processes, *Techno-metrics* **37** 41-59

Stefan Rännar, MacGregor J. F (1998), Adaptive batch monitoring using hierarchical PCA, In *Chemo-metrics and Intelligent Laboratory System* **41** 73-81

Wold S. (1978), Cross-validation estimation of the number of components in factor and principal component analysis, *Techno-metrics* **20** 397-406

Wold S., Geladi P. (1982), Multi-way principal component analysis and LS analysis, *J. Chemo-metrics* 41-56

Wold S (1994), Exponentially weighted moving principal component analysis and projection to latent structures, *Chemo-metrics and Intelligent Laboratory System* **23** 149-161