Batch Process Monitoring Using Two-Dimensional Hidden Semi-Markov Models

J. Chen, *Member, IAENG* and Y.-C. Jiang

***Abstract*—In this paper, a novel monitoring method for the repetitive batch operation with two-dimensional dynamic behavior is proposed. It combines dynamic multiway principal component analysis (DMPCA) and hidden segmental semi-Markov models (HSMM) to resolve the problem caused by the two-dimensional behavior of batch processes. DMPCA utilizes the batch-to-batch dynamic characteristics and eliminates the batch correlation among process variables. HSMM is used to construct the temporal behavior among process variables during each batch run. The proposed method has the temporal property of HSMM and the batch-to-batch dynamic characteristics of DMPCA. To demonstrate the performance of the proposed method, data from the monitoring practice in a fed-batch penicillin cultivation process are conducted.**

***Index Terms*—Batch monitoring; Hidden segmental semi-Markov model; Multiway principal component analysis**

I. INTRODUCTION

In order to increase capabilities of meeting the need of constantly changing market situations, batch and semi-batch processes play an important role in most industries. This especially rings true in the processes mainly involved in the production and processing of high-quality and value-added specialty chemicals (e.g. polymers, pharmaceuticals, biochemicals, food, semiconductors and agricultural chemicals). Batch processes are characterized by the precise sequencing and automation of all stages in the sequence. They convert raw materials into products within the finite duration. Therefore, in order to have safe operation and consistent good quality products, on-line process monitoring is imperative. In recent years, several statistical techniques for mining process information of the operating measurement profiles have been developed. Nomikos and MacGregor (1994) first used multiway principal component analysis (MPCA) to analyze three-way batch data [1]. Since then, batch process monitoring based on the multivariable statistical control process had been developed increasingly because of easy implementation of associated theoretical concepts [2,3]. In fact, the batch system is considered as a two-dimensional model that represents the functions of the operation number in the trial domain and of the operation time in the time domain. This assumption of the entire stage bounded as a single one does not necessarily hold true for the behavior in a two-dimensional manner.

Due to high dimensionality in time space and variable space, multiphase characteristics of each batch, and batch-to-batch dynamic variation, MPCA is not able to deal with the two-dimensional batch operation because it assumes that the batch data are independent and identically distributed. Despite the characteristics of the two-dimensional batch process, less effort has been devoted to improving the monitoring scheme of batch processes in the past. The moving window MPCA approach was a strategy that generated a model by incorporating batch-to-batch information [4]. Another two-dimensional PCA (2DPCA) approach representing the batch process in a two-dimensional space was presented [5]. However, the above methods assumed that the statistical indices followed Gaussian distribution for easy construction of the control limits. The assumption is not realistic in the practical situation.

Hidden Markov models (HMM) are recognized as being appropriate for time sequence data [6]. However, the modeling of HMM is limited by the Markov property in processing realistic duration of different stages. It cannot accurately provide representation of the structure of multiphase operations due to the limitation of state duration of HMM to one with exponential nature. An extended model of HMM, called the hidden semi-Markov model (HSMM), has been developed [7]. Several problems related to HSMM were further investigated [8]. To our best knowledge, the full potential of HSMM models has not yet been recognized in the development of the batch monitoring.

With the integration of the HSMM framework, a statistical model of two-dimensional dynamic process variables for the batch process is developed using HSMM and dynamic MPCA (DMPCA) in this paper. The dynamic data array is constructed by incorporating both static and dynamic process characteristics using the prior and the current batches. The strongest relations of the scores are extracted by DMPCA. The HSMM model is trained with the extracted scores obtained from MPCA method to enhance the capability of statistical process monitoring. Thus, the control charts for a two-dimensional batch are developed in this research. The charts can not only analyze the two-dimensional spatial-temporal measurements but also capture the statistical characteristics of the practical data.

II. TWO-DIMENSIONAL DMPCA-HSMM MODELS

In many batch processes, the same run is carried out

This work was sponsored in part by National Science Council, R.O.C.

J. Chen is with Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taiwan 320, Republic of China (corresponding author to provide fax: 886-3-265-4199; e-mail: jason@wavenet.cycu.edu.tw).

Y.-C. Jiang is with Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taiwan 320, Republic of China.

repeatedly. The same time varying trajectories are used batch after batch. Batch process operations are by nature much more dynamic and they involve several transitions that cover large operating envelopes of phase durations. This means that the multiplicity of the operation phases is an inherent nature of many batch processes and each phase exhibits significantly different underlying behaviors.

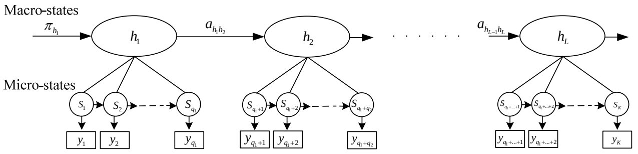


Fig. 1. The structure of hidden semi-Markov models

A. *Hidden Semi-Markov Model* (*HSMM*)

HSMM is modified from HMM, allowing arbitrary state-durations to be used here [9]. The basic structure of HSMM is illustrated in Fig. 1. At the bottom of Fig. 1, it consists of a pair of discrete-time stochastic processes $\{s_k\}$ and $\{y_k\}$, $k \in \{1 \quad \cdots \quad K\}$. For simple explanation and convenience of notation reasons, the observation process $\{y_k\}$ is assumed to be a single variable in the above expression. The extension to the multivariable case will be treated later. The observed process $\{y_k\}$ is linked to the hidden, i.e. non-observed state process $\{s_k\}$ by the conditional distribution $b_{s_k}(y_k)$ depending on the state process,

$$b_{s_k}(y_k) = P(y_k | s_k) \text{ and } \int_{y_k} b_{s_k}(y_k) dy_k = 1 \qquad (1)$$

which implies the fact that the output process at time k depends only on the state of the underlying semi-Markov chain at time k . Due to the sojourn in each of the state, the entire sequence of states $\{s_1, s_2, \cdots, s_K\}$ can be reduced into the sequences of state $\{h_1, h_2, \cdots, h_L\}$ which have been visited

$$\begin{aligned} & h_1 \equiv \{s_1, s_2, \cdots, s_{q_1}\} \\ & h_2 \equiv \{s_{q_1+1}, s_{q_1+2}, \cdots, s_{q_1+q_2}\} \\ & \vdots \\ & h_L \equiv \{s_{q_1+\cdots+q_{L-1}+1}, s_{q_1+\cdots+q_{L-1}+2}, \cdots, s_{K-1}\} \end{aligned} \qquad (2)$$

where the state h_l , $l = 1, 2, \cdots, L$ is macro-states shown in Fig. 1. Each macro-state consists of several single states, which are called micro-states ($\{s_{q_{l-1}+1}, s_{q_{l-1}+2}, \cdots, s_{q_{l-1}+q_l}\}$). Thus, HSMM is like an HMM except that each state can emit a sequence of observations. Note the sojourn time q_l is a discrete non-negative random variable with a occupancy distribution $p_{h_l}(q_l)$. The sojourn of the unobserved process of the time length $q_l \in \min(\tau, k)$ in the state is denoted by a distribution

$$\begin{aligned} p_{h_l}(q_l) = P(& s_{k+q_l+1} \neq h_l, s_{k+q_l-p} = h_l, p = 0, 1, \cdots, q_l - 2 | \\ & s_{k+1} = h_l, s_k \neq h_l) \end{aligned} \qquad (3)$$

where τ is the maximum duration with the state h_l . For a fixed macro-state sequence, the observation probability of a batch run given the model λ can be written

$$P(y_1^K, h_1^L | \lambda) = \pi_{h_1} p_{h_1}(q_1) \prod_{l=2}^{L} a_{h_{l-1}h_l} p_{h_l}(q_l) \prod_{k=1}^{K} b_{s_k}(y_k) \qquad (4)$$

where $y_1^K = \{y_1 \quad y_1 \quad \cdots \quad y_K\}$ and $h_1^L = \{h_1 \quad h_2 \quad \cdots \quad h_L\}$ are the observed as well as the state sequences of a batch duration. The parameters of the HSMM model λ include the initial state $\pi = \{\pi_{h_1}\}$, the state transition probability $A = \{a_{h_{l-1}h_l}\}$, the observation probability $B = \{b_{s_k}\}$ and the state duration probability $D = \{p_{h_l}(q_l)\}$. In this study, the state output is a mixture of Gaussian distribution characterized by the mean $\mu_{h_l,v}$, the variance $\Sigma_{h_l,v}$ and the mixing weight $c_{h_l,v}$ for each node (v), and the duration distribution is a single Gaussian distribution characterized by the mean m_{h_l} and the variance σ_{h_l}

$$\begin{aligned} & b_{s_k}(y_k) = \sum_{v=1}^{V} c_{h_l,v} N\left(y_k; \mu_{h_l,v}, \Sigma_{h_l,v}\right) \\ & p_{h_l}(q_l) = N\left(q_l; m_{h_l}, \sigma_{h_l}\right) \end{aligned} \qquad (5)$$

To estimate the model λ , the expectation-maximization (EM) algorithm for HSMM can be formulated as

$$\lambda = \underset{\lambda}{\arg\max}\, P(y_1^K | \lambda) \qquad (6)$$

This probability ($P(y_1^K | \lambda)$) is the summation of the probability over all the possible state sequences ($s_1^K = \{s_1 \quad s_2 \quad \cdots \quad s_K\}$).

$$P(y_1^K | \lambda) = \sum_{all\ s_1^K} P(y_1^K, s_1^K | \lambda) \qquad (7)$$

The join probability of y_1^K and s_1^K , i.e. the probability that y_1^K and s_1^K will occur simultaneously, is the product of $P(y_1^K | s_1^K, \lambda)$ and $P(s_1^K | \lambda)$. The training consists of two steps: finding the HSMM parameters λ that maximize the probability and computing the joint probability. It is a recursive algorithm. Each iteration of the EM algorithm increases $P(y_1^K | \lambda)$ and generally, the sequence of the re-estimated parameters λ converges to a local maximum of $P(y_1^K | \lambda)$.

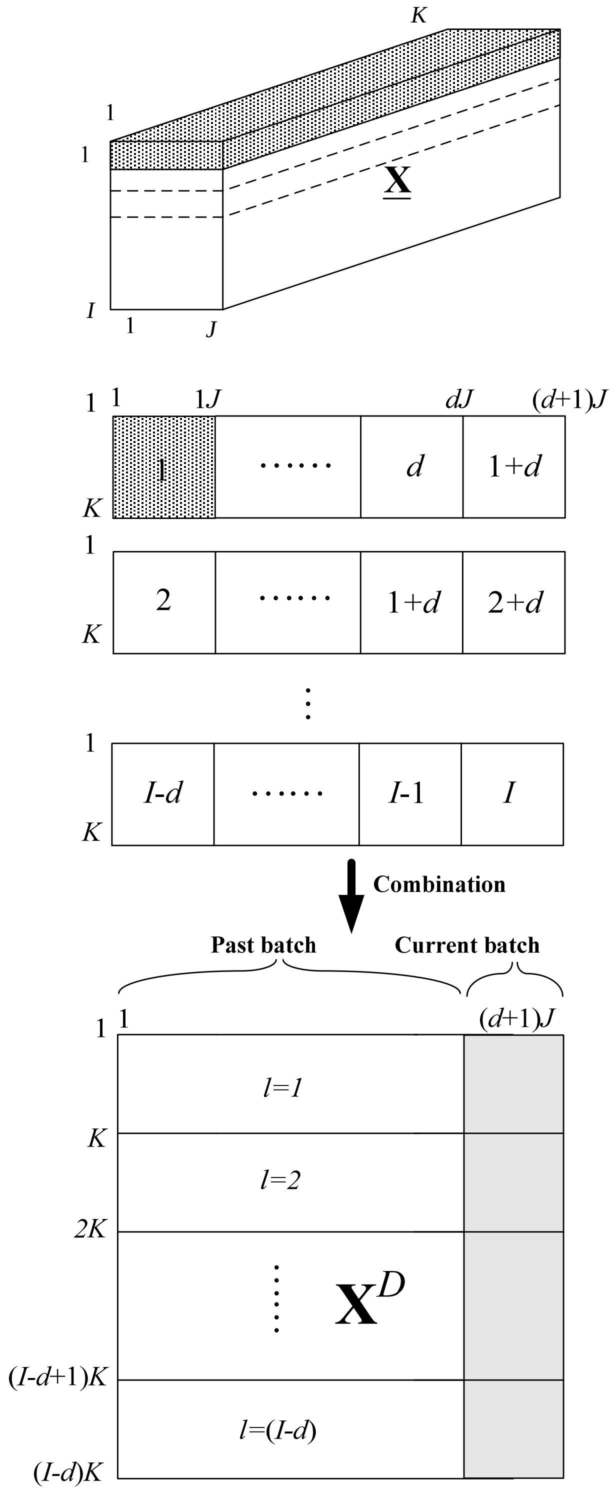


Fig. 2. Time-wise unfolding used for constructing the batch lagged window

B. *Dynamic Batch HSMM Model*

To capture information in the previous batches, a dynamic batch model should be constructed. It is not only concerned with the correlation in the sequence batch but also with the autocorrelation of the process variables. The serial correlation is taken into account when all the variables and their time histories of the previous batches are augmented into the data matrix for the current batch. Fig. 2 shows an approach to an incorporation of batch-to-batch information. Each horizontal slice $K \times J$ is put side by side to the right, starting with the slice corresponding to the first batch run until $d+1$ batch window lengths. The resulting two-dimensional matrix has dimensions of $K \times (d+1)J$. Likewise, the second two-dimensional matrix is constructed starting with the second batch run. Keep doing the same procedures until $I-d$ two-dimensional matrices are completed. Concatenate these two-dimensional matrices vertically, padding the matrices automatically to make a large two-dimensional array,

$$\mathbf{X}^D = \begin{bmatrix} \mathbf{X}_1^D \\ \vdots \\ \mathbf{X}_l^D \\ \vdots \\ \mathbf{X}_{I-d}^D \end{bmatrix} \tag{8}$$

and

$$\mathbf{X}_l^D = \begin{bmatrix} \mathbf{X}_l & \mathbf{X}_{l+1} & \cdots & \mathbf{X}_{l+d} \end{bmatrix},\ l = 1,2,\ldots,I-d \tag{9}$$

$$\mathbf{X}_l = \begin{bmatrix} x_{1,1}^l & x_{1,2}^l & \cdots & x_{1,J}^l \\ x_{2,1}^l & x_{2,2}^l & \cdots & x_{2,J}^l \\ & & \ddots & \\ x_{K,1}^l & x_{K,2}^l & \cdots & x_{K,J}^l \end{bmatrix} \tag{10}$$

where the observation of variable j at the time point k in batch run l is represented as $x_{j,k}^i$. Here all the batches have equal duration and are synchronized. To reduce the variables, the data array is decomposed. the PCA model is now considered as

$$\mathbf{X}^D = \underbrace{\mathbf{T}^D \mathbf{P}^T}_{explained} + \underbrace{\mathbf{E}}_{unexplained} \tag{11}$$

where $\mathbf{P}$ is the loading matrix which retains the first R principal component loadings. $\mathbf{E}$ is a residual vector which cannot be explained by the PCA model. The score matrix is

$$\mathbf{T}^D = \begin{bmatrix} \mathbf{T}_1^D \\ \mathbf{T}_2^D \\ \vdots \\ \mathbf{T}_{I-d}^D \end{bmatrix} \tag{12}$$

and

$$\mathbf{T}_l^D = \begin{bmatrix} \mathbf{t}_{l,1}^D \\ \mathbf{t}_{l,2}^D \\ \vdots \\ \mathbf{t}_{l,K}^D \end{bmatrix}, \quad l = 1,2,\cdots,I-d \tag{13}$$

Next, each $\mathbf{T}_l^D$ in the batch matrix $\mathbf{T}^D$ is extracted to form a observation sequence $\left[\mathbf{T}_l^D\right]_1^K = \left\{\mathbf{t}_{l,1}^D \quad \mathbf{t}_{l,2}^D \quad \cdots \quad \mathbf{t}_{l,K}^D\right\}$, $l = 1,2,\cdots,I-d$. Now the new variables $\left[\mathbf{T}_l^D\right]_1^K$ that are linear combinations of the original ones are uncorrelated with each other. The above procedure enables us to easily model HSMM because it is only necessary to perform the diagonal covariance matrices ($\mathbf{\Sigma}_{h_l,v}$) in Eq. (5), in contrast to the standard ones. Furthermore, instead of building HSMM upon the training data of one batch set only, data tying across all simultaneous batch sets are applied to training the HSMM model in the reduced space. Tying increases the amount of training data to avoid over-fitting the HSMM model and achieve the robust training result.

III. MONITORING CONTROL CHART BASED ON MPCA-HSMM

Once the HSMM model is trained, the most likely state sequence for an observed sequence could be found. Each observed sequence infers a corresponding hidden state path. However, there are potentially many state paths that can generate the same sequence. For the description of the process operation, the best state path of HSMM for each batch should be found in the trained model. The best state path should also be able to maximize the path probability for a given observation. The optimal state sequence is used instead of all the possible state sequences for more efficient classification. For a given model and an observation score sequence $\left[\mathbf{T}^D\right]_1^K = \left\{\mathbf{t}_1^D \quad \mathbf{t}_2^D \quad \cdots \quad \mathbf{t}_K^D\right\}$, the most likely underlying state sequence ($\left(h_1^L\right)^* = \left\{h_1^* \quad h_2^* \quad \cdots \quad h_L^*\right\}$) can be found.

$$\left(h_1^L\right)^* = \left\{h_1^* \quad h_2^* \quad \cdots \quad h_L^*\right\} = \arg\max_{h_1^L} P(\left[\mathbf{T}^D\right]_1^K, h_1^L \middle| \lambda) \qquad (14)$$

A recursive Viterbi algorithm with dynamic programming can be used to find out the most likely state sequence [10]. Repeat the same procedure for all the collected sets to compute the most likely state sequences for each set. Not all the sets follow the same best path. Different selected states for different sets account for the fact that the observed values of the process variables do not perfectly conform to the deterministic model. This variability is resulted from the uncertain variations and disturbances among the hidden variables that affect the system. Fig. 3 shows a pictorial display of the output score distribution ($P(\mathbf{t}_{l,k}^D\left(s_1^K\right)\middle| s_1^K, \lambda)$) based on the selected optimal state path at each time point for all the batches. The distribution can help assess what the status of the current operating batch is and if the operation batch is in control.

With the developed probability distribution that reflects the normal operation, the control limit for the selected state at each time point is required to detect any departure of the process from its standard behavior

$$\int_{P(\mathbf{t}_c^D(k)|\left(h_1^L\right)^*, \lambda) > P^{th}(k)} P(\mathbf{t}_c^D(k) \,|\, \left(h_1^L\right)^*, \lambda) d\mathbf{t}_k \le 0.95 \qquad (15)$$

$$k = 1, 2, \cdots, K$$

The score matrix by projecting the data ($\mathbf{x}_c^D(k)$) mentioned above onto the loading matrix ($\mathbf{P}$) can be computed:

$$\mathbf{t}_c^D(k) = \mathbf{x}_c^D(k)\mathbf{P} \qquad (16)$$

$$\mathbf{x}_c^D(k) = \left[\mathbf{x}_{c-d}(k) \quad \mathbf{x}_{c-d+1}(k) \quad \cdots \quad \mathbf{x}_c(k)\right] \qquad (17)$$

$$\mathbf{x}_c(k) = \left[x_{k,1}^c \quad x_{k,2}^c \quad \cdots \quad x_{k,J}^c\right], \; c = 1, 2, \ldots, I-d \qquad (18)$$

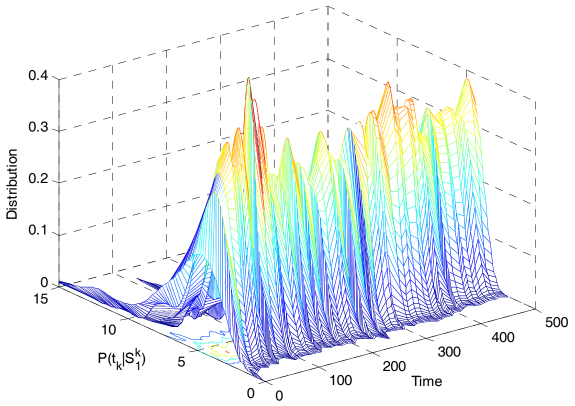


Fig. 3. The output distribution based on the selected optimal state path at each time point

The data corresponding to 95% of the confidence limit is taken to get a likelihood threshold ($P^{th}(k)$) at time point k .

The corresponding unexplained part can be gotten by applying DMPCA

$$\mathbf{e}_c^D(k) = \left[\mathbf{e}_{c-d}(k) \quad \mathbf{e}_{c-d+1}(k) \quad \cdots \quad \mathbf{e}_c(k)\right] = \mathbf{x}_c^D(k) - \mathbf{t}_{c,k}^D\mathbf{P}^T \qquad (19)$$

The *SPE* statistics for the deviations from the model behavior (residuals) of the process variation at sample k is defined

$$SPE(k) = \mathbf{e}_c(k)\mathbf{e}_c^T(k) \qquad (20)$$

Note that the row vector $\mathbf{e}_c(k)$ represents the residuals at time point k of the current batch. The confidence limits of $SPE(k)$ cannot be determined directly from a particular distribution. The kernel density estimation is used to determine the confidence limits [11].

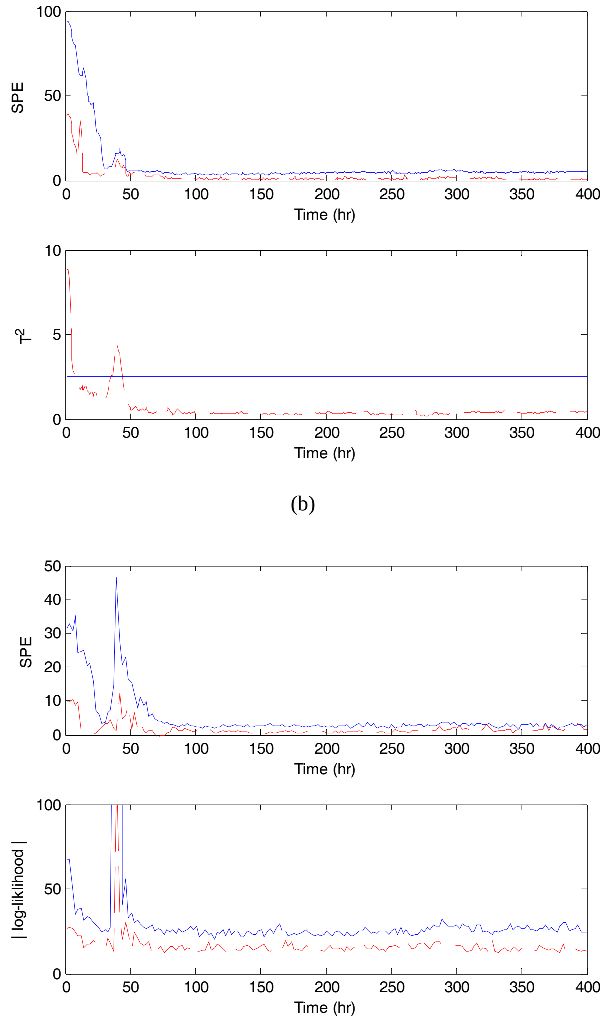
IV. EXAMPLES

A benchmark simulation of fed-batch penicillin production is used here to demonstrate the application of the proposed method. The system has nonlinear dynamics and multi-phase characteristics. During the initial pre-culture phase, the necessary cell mass is generated. Then penicillin starts to be generated at the exponential growth rate and continues to be produced until the stationary phase. The simulation condition and the relevant parameters are referred to Birol, et al. [12]. The system switching itself from batch modes to the fed-batch modes of operation depends on whether glucose concentration reaches its threshold value ($0.4\,g/L$). Due to input variations, the batch/fed-batch switch times are quite different for different batch runs. In the iterative operation, the initial condition of each batch may not be reset to the desired initial condition or inside a neighborhood of the desired initial condition, because small amounts of biomass and substrate in the previous batch run are remained in the tank. Thus, the initial concentrations of biomass ($C_m(i,k=1)$) and substrate ($C_s(i,k=1)$) at the starting point ($k=1$) are the mixing of the biomass in the feed ($C_s^f(i,k=1)$) and the previous batch ($C_s(i-1,k=K)$) as well as the mixing of the substrate in the feed ($C_m^f(i,k=1)$) and the previous batch ($C_m(i-1,k=K)$), respectively.

To build up a DMPCA-HSMM based monitoring system, the normal operating condition of the penicillin fermentation is simulated in this study. In the normal condition, a total of 11 process variables (J) with six states (L) and two Gaussian functions (V) are considered to train DMPCA-HSMM. Two additional batches, comprising one normal batch and one abnormal batches, are generated for testing. In the abnormal batch, it is assumed that the substrate feed rate is linearly decreased from 0.04 to 0.028 from the 60th hour till the end of the batch.

The first batch (normal) is monitored at each time point. Two different models, 2DPCA and DMPCA-HSMM, are used to monitor the operating batch for a comparison. In 2DPCA, the dynamic batch data of time-wise dimension and batch-wise dimension are used for batch process modeling. Appling 2DPCA gives the results presented in Fig. 4(a). There is still some false detection between the 40th hour and the 50th hour when the transition of two different phases occur. Again, DMPCA-HSMM is applied to the same test batch. In Fig. 4(b), the SPE and $\log P(\mathbf{t}_k \mid s_k, \lambda)$ charts at each time point are obtained from the proposed DMPCA-HSMM. It is evident from the charts that DMPCA-HSMM can capture the dominant behavior without getting wrong conclusions. In 2DPCA, Fig. 5 also shows the normal probability plot of the first scores of the stage PCA model. It indicates that there is a problem with the normality assumption. As a result, 2DPCA (Fig. 4(a)) has false detections when applying the T^2 and SPE test statistics.

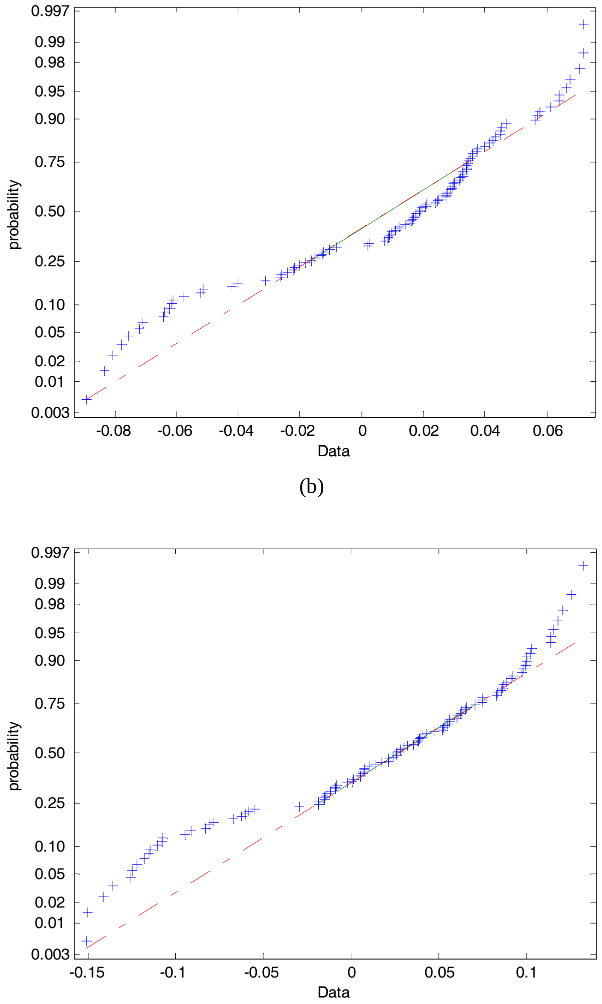
(a)



(b)

Fig. 4. Control charts for on-line monitoring of the normal batch in the example: (a) 2DPCA, and (b) DMPCA-HSMM. Each chart contains 95% (solid line) control limit. The dashed line represents the normal batch

(a)



(b)

Fig. 5. Normal distribution plot of (a) score 1 and (b) score 2 from 2DPCA.

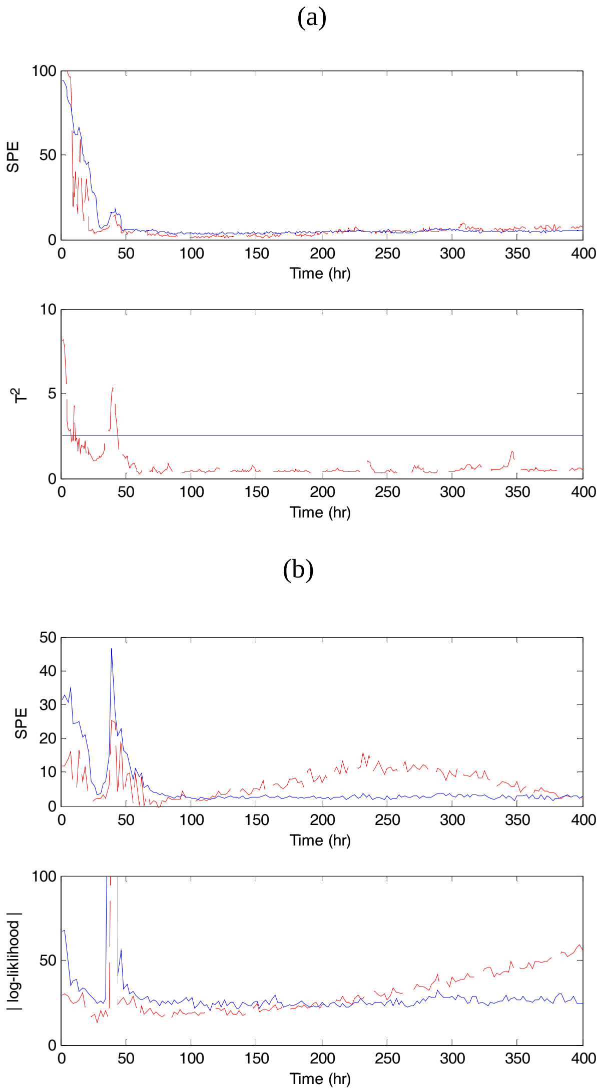


Fig. 6. Control charts for on-line monitoring of the normal batch in the example: (a) 2DPCA, and (b) DMPCA-HSMM. Each chart contains 95% (solid line) control limits. The dashed line represents the abnormal batch.

For the first abnormal batch, the monitoring outcomes of three models are shown in Fig. 6. The initial time of the fault is induced at the 60th hour and until the end of cultivation. The decrease in the feed rate results in a reduction of penicillin production. As shown in Fig. 6 (a), the SPE or T^2 monitoring charts of 2DPCA cannot detect this fault in the whole operating duration. Moreover, there are false detections of 2DPCA around time point 40. The detection abilities of DMPCA-HSMM have been significantly improved and the fault is able to be detected earlier after applying DMPCA-MSMM to characterizing the temporal dynamics between batches. In Fig. 6(b), after the 150th hour SPE has increased remarkably and fallen outside of the 95% confidence limit and $\log P(\mathbf{t}_k \mid s_k, \lambda)$ also fallen outside of the 95% confidence limit after the 220th hour.

V. CONCLUSION

On-line process monitoring has been attracting increasing interest in the field of process safety and quality control. In the past research, most of the tools used for the detection of disturbance assumed that the process data follow a Gaussian distribution. The assumption is not necessarily satisfied in batch operation processes with complex and dynamic characteristics.

In this research, the conventional multivariate SPC based on MPCA are extended by incorporating the HSMM structure to solve the detection problem for the two-dimensional batch operation. At the first stage, the current and the past batch measurements are stacked together to take into account the batch-to-batch serial correlations. In this stage, the temporal evolution of the two-dimensional batch operation behavior can be tracked. At the extracting feature stage, the MPCA scheme is used to determine the correlation between the measurements within the adjacent batch runs as well as among other non-adjacent batch runs. The objective of the feature extraction stage is to increase the robustness of the two-dimensional pattern by reducing the dimensionality of the data and retains most of the essential features. In order to quantitatively discriminate the probability distribution of multiphase operations, HSMM is applied at the modeling probability distribution stage. The two-dimensional patterns can be trained. In comparison with conventional 2DPCA model, the proposed MPCA-HSMM model shows negligible erroneous judgment on the normal operation condition and efficient monitoring capability to detect the time of occurrence of the process fault.

REFERENCES

[1] P. Nomikos and J. F. MacGregor, “Monitoring Batch Processes Using Multiway Principal Component Analysis,” *AIChE Journal*., 40, pp. 1361, 1994.

[2] J. Chen and K.-C. Liu, „On-Line Batch Process Monitoring Using Dynamic PCA and Dynamic PLS Models,” *Chemical Engineering Science*, 57, pp. 63, 2002**.**

[3] B.Lennox, H.G. Hiden, G.A. Montague, G. Kornfeld, P.R. Goulding, Process Monitoring of an Industrial Fed-Batch Fermentation, *Biotechnology and Bioengineering*, 74, pp. 125, 2001.

[4] J. Flores-Cerrillo and J. F. MacGregor, “Multivariate Monitoring of Batch Processes Using Batch-to-Batch Information,” *AIChE J*., 50, pp. 1219, 2004.

[5] N. Lu, Y. Yao, F. Gao, and F. Wang, “Two-dimensional Dynamic PCA for Batch Process Monitoring,” *AIChE J*., 51, pp. 3300, 2005**.**

[6] L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” *Proc. IEEE*., 76, pp. 257, 1989.

[7] Y. Guédon, “Estimating Hidden Semi-Markov Chains from Discrete Sequences,” *Journal of Computational and Graphical Statistics*, 12, pp. 604, 2003**.**

[8] M. Dong and D. He, “Hidden Semi-Markov Models for Machinery Health Diagnosis and Prognosis,” *Trans NAMRI/SMEXXXII*., 32, pp. 199, 2004.

[9] J. D. Ferguson, “Variable Duration Models For Speech,” *in Symp. on the Application of Hidden Markov Models to Text and Speech*, Princeton, NJ, pp. 143–179, 1980.

[10] M. D. Moore, “Most Likely State Sequence Speech Reconstruction Using a Generalized Hidden Semi-Markov Model with Two Distinct Regeneration Times Applied to English,” Ph.D. Dissertation, Rensselear Polytechnic Institute, 2004.

[11] C.M. Bishop, *Neural Networks for Pattern Recognition,* Clarendon Press: New York, 1995.

[12] G. Birol,; C.Ündey, , and A. Cinar, “A Modular Simulation Package for Fed-Batch Fermentation: Penicillin Production,” *Computers and Chemical Engineering*, 26, pp. 1553, 2002.