AIChE WILEY InterScience DISCOVER SOMETHING GREAT

Dissimilarity Analysis Based Batch Process Monitoring Using Moving Windows

Chunhui Zhao, Fuli Wang, and Mingxing Jia
School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning Province, People's Republic of China

DOI 10.1002/aic.11164
Published online March 23, 2007 in Wiley InterScience (www.interscience.wiley.com).

In recent years, a novel MSPC method known as DISSIM, which is based on the dissimilarity of process time series data, has been developed focusing on continuous and stable processes. However, its further application is hindered because of its unsuitability to batch processes which lie at the heart of many industries. In the present work, combined with variable moving windows, an important extension of the novel DISSIM method, termed EDISSIM, is made for its practical application to batch processes monitoring as well as the theoretical analysis basis for the determination of control limit. Moreover, contribution plots of dissimilarity index are used to identify the variables that contribute significantly to the out-of-control state. The applications of the proposed EDISSIM method are illustrated with respect to simulated data collected from both a simple 2 × 2 numerical process and a fed-batch penicillin fermentation industrial batch process. The results clearly show that it functions very well to successfully detect and diagnose fault in batch processes, which implies the significant investigation potential of the proposed method.
AIChE J, 53: 1267–1277, 2007
Keywords: batch processes, EDISSIM, variable moving windows, statistical distribution hypothesis testing, fault detection and diagnosis

Introduction

Batch and semibatch processes play an important role in specialty chemical, semiconductor, food, and biology industry for producing high-value-added products to meet today's rapidly changing market. Characterized in operations by finite duration, batch processing aims to produce products of desired quality at low cost. The nonsteady, time-varying, finite duration, and nonlinear behaviors of batch processes make batch control more difficult than control of a continuous process. Process disturbances, which may vary from batch to batch, affect both process and product reproducibility. Proper process monitoring and diagnosis is important not only to quality improvement but also to process safety.[1–7] Most batch process monitoring methods are based on multiway principal component analysis (MPCA) and partial least squares (MPLS).[8–12] Many successful applications have shown the practicability of MSPC. However, conventional MPCA/MPLS methods do not always function well, because they cannot detect the change of correlation among process variables as long as both T^2 and SPE statistics are inside the control limits.

In recent years, a novel method to monitor statistical processes known as DISSIM has been developed by Kano et al.,[13–16] which is based on the dissimilarity analysis of process data. It detects the change of operating condition in typical continuous processes by quantitatively evaluating and monitoring the distribution of process data with dissimilarity index. A series of successful theory researches and applications have demonstrated that the method can quickly and effectively detect the change of correlations among process variables. However, the method focuses on the monitoring of

Correspondence concerning this article should be addressed to F. Wang at flwang@mail.neu.edu.cn.

continuous processes and does not work well for online monitoring of batch processes. The process characteristics of batches are distinctly different from those of continuous processes because of their obvious dynamic unstable states along time. This means that using the original DISSIM method, the dissimilarity values will show large and ruleless fluctuations. Therefore, it is difficult to select a proper control limit and conduct online monitoring for batch process. In the present work, for its practical application to batch processes monitoring, we investigated an important extension of DISSIM method, termed EDISSIM method, based on dissimilarity analysis of variable moving window data. This work extends and refines the idea of previous work[13–16] and differs from them in that: (1) EDISSIM is applicable to batch processes whereas DISSIM is developed for continuous processes; (2) EDISSIM focuses on the batch-to-batch variation of process trajectories during the same time span, while DISSIM pays attention to the change of operating condition along time direction within one process, which represents the most remarkable distinctness between the two methods; (3) for EDISSIM, multiple reference models are established from one batch by variable moving windows, whereas only one reference model for DISSIM; (4) the control limits in EDISSIM are decided based on Γ-distribution probability statistical rule to capture batch-to-batch dissimilarity variation, whereas those of DISSIM refer to no theoretical analysis.

This article is organized as follows. First, the theoretical basis and details of the proposed method, EDISSIM, are explained. Then the simulation results of a simple 2×2 numerical process with stable nature are given and the performance of the proposed method and that of the conventional method are compared with their applications to simulated data collected from fed-batch penicillin fermentation batch process. Conclusions are presented in the last section.

Monitoring Method Based on Dissimilarity Analysis of Variable Moving Windows

Dissimilarity index

Kano et al.[13–16] proposed a statistical process monitoring method based on the dissimilarity of process data. Their method, termed DISSIM, is based on the idea that a change of operating condition can be detected by monitoring a distribution of process data because the distribution reflects the corresponding operating condition. It is also the important basis of our proposed method, EDISSIM.

To evaluate the difference between distributions of data sets, a classification method based on the Karhunen-Loeve (KL) expansion[17] is used in this work. The KL expansion is a well-known technique for feature extraction or dimensionality reduction in the pattern recognition area. Instead of using distance to measure the systematic variations in the PC subspace as well as those in residual subspace, it emphasizes the analysis of distribution of process data during some period to catch the correlation structures of process data and detect the change of operating condition. Consequently, it monitors different statistic indexes, where PCA employs Hotelling-T^2 and SPE, but the proposed method utilizes eigenvalues of the covariance matrices.

Consider the following two data sets: X_1 and X_2. X_i consists of N_i samples of J variables. The covariance matrices are given by

$$R_i = \frac{1}{N_i} X_i^T X_i \tag{1}$$

And the covariance matrix of the mixture of both data sets is given by R:

$$R = \frac{N_1}{N_1 + N_2} R_1 + \frac{N_2}{N_1 + N_2} R_2 \tag{2}$$

On the basis of the fact that the covariance matrix R can be diagonalized by an orthogonal matrix P_0

$$P_0^T R P_0 = \Lambda \tag{3}$$

The original data matrices X_i are transformed into Y_i.

$$Y_i = \sqrt{\frac{N_i}{N_1 + N_2}} X_i P_0 \Lambda^{-1/2} = \sqrt{\frac{N_i}{N_1 + N_2}} X_i P \tag{4}$$

where P is a transformation matrix defined as $P = P_0\ \Lambda^{-1/2}$.

The covariance matrices of the transformed data matrices

$$S_i = \frac{1}{N_i} Y_i^T Y_i = \frac{N_i}{N_1 + N_2} P^T \frac{X_i^T X_i}{N_i} P = \frac{N_i}{N_1 + N_2} P^T R_i P \tag{5}$$

satisfy the following equations:

$$S_1 + S_2 = I \tag{6}$$

By application of eigenvalue decomposition to the covariance matrices

$$S_i \omega_i^j = \lambda_i^j \omega_i^j \tag{7}$$

Here, λ_i^j and ω_i^j are the eigenvalues and the corresponding eigenvectors, respectively. The superscript j denotes the jth eigenvalue or eigenvector. From Eqs. 6 and 7, the following relationships can be derived:

$$S_2 \omega_1^j = (1 - \lambda_1^j)\omega_1^j \tag{8}$$

$$1 - \lambda_1^j = \lambda_2^j \tag{9}$$

The above two relationships mean the transformed data matrices, Y_1 and Y_2, have the same set of principal components and the principal components are reversely ordered. In other words, the most important correlation for the transformed data set Y_1 is equivalent to the least important correlation for the other transformed data set Y_2, and vice versa.

Finally, the following index D was defined for evaluating the dissimilarity of data sets.

$$D = \ \mathrm{diss}(X_1, X_2) = \frac{4}{J} \sum_{j=1}^{J} (\lambda_j - 0.5)^2 \tag{10}$$

where λ_j denotes the eigenvalues of the covariance matrix of the transformed data matrix and J the number of process variables. When data sets are quite similar to each other, the eigenvalues λ_j must be near 0.5, and then D should be near zero. On the other hand, when data sets are quite different form each other, the largest and the smallest eigenvalues should be near one and zero respectively and D should be near one. Therefore, the index D changes between zero and one.

Variable moving windows

In most previous cases, a fixed-length moving window is employed to track dynamic data, in which it is always important to select proper time span for such a window. If the chosen window is too small, one may capture process changes quickly, but the window does not contain enough information to sufficiently reflect the current process operating condition, thus leading to ambiguous dissimilarity. To obtain a data set representing operation condition appropriately, one needs to use a sufficient number of samples. However, a large window, in the early period, may contain useless data and lead to large diagnostic delays. Therefore, the appropriate selection of time-window size is important for the effective function of monitoring method. However, the determination of window size is still a hot and confusing problem with no uniform quantitatively calculation standard.

Here, a variable moving window is introduced, in which the moving window length grows gradually in length, and contains more valuable data reflecting process operation characteristics to provide the powerful basis for accurate and rapid fault detection using dissimilarity analysis. Different from the fixed-length moving window, the right side of moving window moves forward over time while the left side remains at the initial time. During the early period of process, because there is not enough sampling data, that is, the process distribution information, the initial time-serial window can give a rough estimation. Then with the development of process, new data is added step by step resulting in longer and longer moving window size with time. The more time-serial data, the more process distribution information, which will help to reflect the changing details of process operating conditions and make fault diagnosis reliable and stable. By this way, one can readily obtain the process operation state up to the current time and the effects of disturbances on the distribution trajectory from an overall perspective. Moreover, the choice of window size will not impose great influence on the monitoring effect any longer.

Modeling based on dissimilarity analysis

For continuous process modeling using DISSIM method, one reference data set under normal operating conditions is defined. Then the dissimilarity indices between time series data sets spanning other time regions and the reference one are compared to determine a proper control limit. So it focuses on the changing trend of operation conditions along time direction for online monitoring. However, for batch processes, the means and variances of process measurements will change continuously with time revealing different dynamic process nature in different operation periods. Because of the common multistate, dynamic, time-varying and nonlinear characteristics, only one reference data set as monitoring model is not enough, in which dissimilarity values between different time-serial data sets along time sharply fluctuate and do not follow the normal distribution law. Accordingly, the uniform control limits made during the whole process like DISSIM will be too relaxed for some time regions, and inevitably induce missing alarms in large quantities. Considering the inherent characteristics of batch processes, the batch-to-batch variation information should be emphasized, and its statistical distribution rule provides an important theoretical basis for the determination of control limits.

For applying the proposed monitoring method, reference data sets as monitoring models are determined by the following procedure. The illustration of EDISSIM method is shown in Figure 1.

Step 1: In each batch run, assume that J variables are measured at $k = 1,2,\ldots,K$ time instance throughout the batch forming the time series data matrix $X(K \times J)$. I batch runs, a similar data set for each run ($i = 1,2,\ldots, I$), result in a three-way data array $X(I \times J \times K)$.

Step 2: The three-way batch process data matrix is unfolded to batch-wise form: $X(I \times (K \times J))$ and then normalized along batch in each time-slice matrix to focus on batch-to-batch variations, which is necessary to eliminate the influence of nonlinearity and different measuring scale for further modeling analysis. Subsequently, the mean-centered data matrices are rearranged into time series $\overline{X}(K \times J)$ for each batch.

Step 3: Choose one from I batches and its time-series data $\overline{X}(K \times J)$ is set as the initial reference modeling batch while others as training ones. Determine the original size of time-window, L. Generate series of moving windows, X^w ($w = 1,2,\ldots,K - L + 1$), with variable size of samples from L to

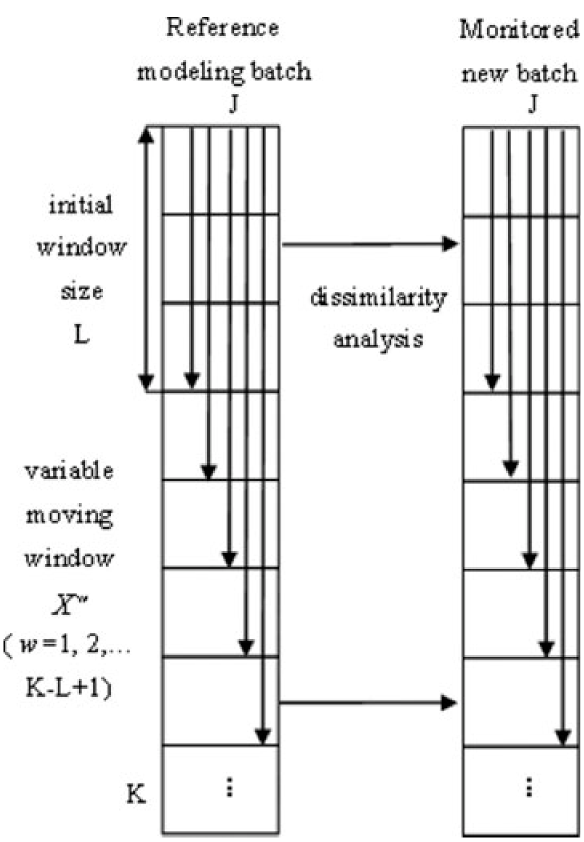

Figure 1. Illustration of the proposed method.

K by moving the time-window forward while the left side remains at the beginning, which will be used as reference data sets. Similarly, the variable moving windows are generated in each training batch.

Step 4: According to the sampling time, a series of dissimilarity index D between training data windows from different batches and the corresponding reference one at the same time region are calculated revealing the batch-to-batch dynamic variations. Then, according to the index values, a proper batch with mid dissimilarity values is set as the final reference batch in place of the initial one.

Step 5: Recalculate the dissimilarity index between the new reference batch and training batches in the same way as the above steps and determine the appropriate initial control limits of D, accordingly using the concept of statistical distribution.

The key to the proposed EDISSIM is that instead of single reference data set in DISSIM, it defines a series of moving windows within one batch as the reference distributions, that is, each moving window represents the normal operating condition corresponding to every time interval. Therefore, at each specific sampling time, there will be a corresponding variable moving window X^w to conduct dissimilarity comparison with the current new batch for online process monitoring. So process time will be used to indicate which model should be employed as the current monitoring tool.

Control limits for fault detection

In DISSIM method,[16] a remaining issue is theoretical analysis of the selection of control limit, where its control limits, 99% and 95%, are determined simply so that the ratio of samples outside the limits to the entire samples is 1% and 5% respectively. Therefore, the confidence limits are invariable throughout the process. From a practical viewpoint, it is important to develop a theoretical basis for selecting appropriate control limits of dissimilarity index. In our study, all the modeling batches should be collected under normal operation conditions. A normal batch implies that the process follows a set of predetermined sequences with acceptable process trajectory variation for a finite duration to convert raw materials to products. A high degree of reproducibility is necessary to obtain successful batches. Batch-to-batch variations from the mean trajectories are caused by stochastic factors, that is, the normal batches in the modeling database are deemed to subject to common-cause variation. In the present article, we can inherit this conclusion and easily deduce that the batch-to-batch variations of process correlation characteristics at the same time also follow normal distribution. Since the process characteristics are represented by the moving windows of process measurements and dissimilarity indices are used to quantitatively evaluate the variations between them, it is obvious that dissimilarity indices, the numerical value representation of the distribution rule, should also follow some statistical distribution related close with normal distribution from batch to batch.

In the real statistical problem, we often see only the sampling data, from which we try to infer the properties of the population that the samplings belong to. To evaluate the distribution rules, we use the statistical distribution hypothesis testing theory in this work. In the practical application, the concept of statistical distribution hypothesis test[18] is often used to infer the complete population distribution based on the information from samples, which is called nonparametic hypothesis test method. There are various hypothesis test methods, such as goodness-of-fit test, Kolmogrov-Smirnov test and independency test. To infer the actual population distribution according to the samples, Kolmogorov hypothesis testing, is used in this work, which is a well-known fitting test technique for statistical distribution function. Not only can it test whether the empirical distribution follows some theoretical distribution but also identify whether two sample sets come from the same population. Here, using Kolmogorov testing, we can checkout that the population follows Γ-distribution. Since the calculated value of Kolmogorov testing for observed samples is smaller than the corresponding critical value with significance level $\alpha = 0.1$, which is found up in the statistical tables, the Γ-distribution hypothesis of the population is accepted. Then, the fundamental probability statistical principle of Γ-distribution, which is a proper description of the probabilistic behaviors of population, can be used as the theory analysis basis to determine the control limits. According to the Γ cumulative probability distribution function,[18] we can readily acquire the confidence interval in which samplings falls with 99% and 95% probability, that is, about 95% and 99% of the entire population, from which the observed data is sampled, lie in the interval from 0 to the corresponding 99% and 95% upper percentile respectively.[19] Here, based on Γ probability distribution, the 99% and 95% control limits of EDISSIM are determined accordingly. In the section of Illustrations and Discussions, the rationality of determining control limits in such a way will be demonstrated illustratively.

Online monitoring and fault diagnosis

For online monitoring of the new batch data, the time window representing current operating condition is updated step by step by adding new sampling data, which is normalized using the same mean and variance obtained from Step 2 of modeling procedure. Then, the corresponding index D is calculated at each step. If the indices are below the control limits, the process is judged to be normal. Otherwise, whenever they go beyond the bounds of normal operation, responding to an abnormality, the contribution plot,[20] a novel diagnosis tool, can be used to identify the variables that contribute significantly to the occurred abnormality.

Since the dissimilarity index is a function of eigenvalues of the covariance matrix of the transformed data matrix, the most contributive eigenvalue to D and the corresponding eigenvector can easily be determined. However, now that the data matrix is transformed, therefore each loading is not directly related to each original variable. For solving the problem, the data matrix has to be inverse-transformed into the original matrix. From Eq. 4, the linear transformation of X_i into Y_i can be simply expressed by

$$Y_i = X_i A \tag{11}$$

Then the score vector related to the most contributive eigenvector is given by

$$t^1 = Y_i \omega_i^1 = X_i A \omega_i^1 = \sum_{j=1}^{J} x_j (A\,\omega_i^1)_j \tag{12}$$

where x_j is the jth column vector in X_i and $(A\omega_i^1)_j$ denotes the jth element of $A\,\omega_i^1$. For evaluating the contribution of each variable to the most contributive score vector t^1, the norm of the transformed variable, that is, the product of each column of the original data matrix and the corresponding transformed loading, is used in the present work. Therefore, a contribution of the jth variable to the dissimilarity index is defined as

$$C_j^{[D]} = \| x_j (A\, \omega_i^1)_j \| \tag{13}$$

In the present work, the above analyses are made based on such a critical implicit assumption that the processes operate with fixed run duration so that the specific process time can be used as an indicator to judge which representative model should be employed to conduct the dissimilarity analysis. However, in practical situations, the total process length may vary normally within a series of batches reflecting the changes of operating conditions or control objectives. A normal batch implies that the process follows a set of predetermined sequences with acceptable process trajectory fluctuation, where the batch-to-batch variations from the mean trace are caused by stochastic causes. On one hand, if the uneven-length problem is not serious, we can assume that the main events have occurred in the common region and the process correlations will not drastically change when they operated beyond the common regions. On such a premise, the proposed method in the present work provides a simple solution to handling batches with unequal duration, which does not require equalizing the batch length beforehand. Firstly, the monitoring models are derived from those normal reference batches with even common duration, where like most statistical modeling and monitoring methods, the modeling data are assumed to cover statistically all normal process variations. When online monitoring, for those new batches with process length shorter than the modeling common duration, there is no problem, where the process time still can be used as an indictor to normalize the coming measurements and calculate the D index values; for those longer batches than the modeling cycles, the real-time data normalization and the determination of monitoring models are the key issues. Here, the mean and variance at the end time of modeling batch are used to normalize the rest process measurements of monitored trajectory beyond the reference common duration. Moreover, the last monitoring model of the reference batch can be employed as the representative monitoring tool for those rest sampling data outside the fixed common duration. Then the control limits can also inherit from the ending value of control limits. On the other hand, if the difference in batch length is significant, the batch duration has to be aligned prior to the dissimilarity analysis. There are several methods for equalizing batch lengths. For industrial processes with varying batch length, Lu et al.[21] have proposed a particular method to examine unequal-duration batch observations. Moreover, indicator variable technique (IVT), dynamic time warping (DTW), curve registration[6] can also be used to adjust the length of process trajectories in the data preprocessing procedure instead of the time axis so as to construct uniform batch structure prior to modeling and monitoring.

Another factor that should be noted is that if the batch has many phases, switching from one phase to another can occur at different time. The variation of the duration of particular phases may result in some reference batches being in one phase and others being within another one at a particular time k. In the present work, however, where the overlap time between different phases is much shorter compared to the phase operating time, the diversity in the phase-length is not so critical; hence its effects on data normalization is negligible and the present approach is applicable. In contrast, if the effect of overlap region is severe, the meaning of the mean and standard deviation may be not clear so that the data normalization method denoted by specific process time may not function ideally. In such a case, the effectiveness of the fault detection system is compromised. So if uneven-phase problem is significant, it could be more appropriate to apply dissimilarity analysis localized in different phases. Research on such case is significative and deserves further investigation.

In conclusion, considering the complexity of unequal-length problem, there may be various circumstantialities should be integrated and paid special attention to. The researches on new methods aiming at the uneven-length cases under different conditions are promising and certainly will arouse extensive academic interest in future.

Illustrations and Discussions

The proposed process monitoring approach is tested with two processes, a simple numerical process and the long-cycle fed-batch penicillin fermentation batch process. Although it is developed for batch processes, the method works equally well with continuous processes with finite duration.

Application 1: 2 × 2 process

In this section, the proposed monitoring method, EDISSIM, as well as DISSIM method by Kano, is applied to online monitoring problems of a simple 2X2 process, which initially is a typical continuous process. For the comparison of monitoring effect with two methods, the numerical process is modified to generate numbers of cycles for EDISSIM whereas the DISSIM is applied to the initial continuous process. The monitoring results initially demonstrate the effectiveness of the proposed method based on dissimilarity analysis of variable windows.

For simplicity, the 2X2 multivariate process[22] is described as follows:

$$x(t) = \begin{bmatrix} 0.118 & -0.191 \\ 0.847 & 0.264 \end{bmatrix} x(t-1) + \begin{bmatrix} 1.0 & 2.0 \\ 3.0 & -4.0 \end{bmatrix} u(t-1) \tag{14}$$

$$y(t) = x(t) + v(t) \tag{15}$$

where u is the correlated input:

$$u(t) = \begin{bmatrix} 0.811 & -0.266 \\ 0.477 & 0.415 \end{bmatrix} u(t-1) + \begin{bmatrix} 0.193 & 0.689 \\ 0.320 & -0.749 \end{bmatrix} w(t-1) \tag{16}$$

The input w_i in w are uncorrelated Gaussian signals with zero mean and unit variance. The output y_i in y is corrupted by uncorrelated Gaussian errors with zero mean and variance 0.1. The input u and the output y are measured, and their measurements are used for monitoring.

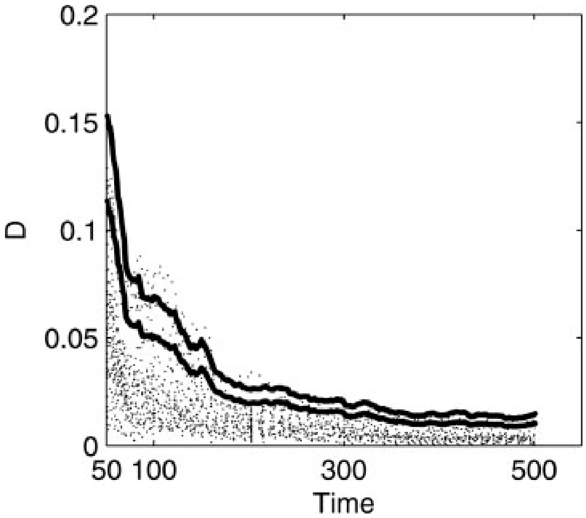

Figure 2. Control limits derived from probability distribution principle (bold dot line, 99% and 95% control limits; dashed line, dissimilarity indices to training data).

Under the normal operating conditions, a time-series data set $X(50500\times4)$ can be obtained from the continuous process with the development of time. Then 101 cycles of time-series data $X(500\times4)$ are generated by cutting off the entire data set. The initial window length is set to be 50 by trial and error and then 451 varying moving windows are generated in each cycle. Using the modeling procedures mentioned in the former section, one cycle is selected properly to build a series of reference models. Then focusing on the cycle-to-cycle variation, the other 100 cycles are used as training data sets to determine the control limits of monitored indexes using probability distribution principle. From Figure 2 we can visually validate the rationality of determining control limits in such a way, which will lay an important foundation for the following fault detection. To verify the performance of the proposed method, mean shifts of w and changes of a coefficient from u_1 to x_2 are investigated and the corresponding process data are generated for each case shown in Table 1, which are used for online monitoring. In the present article, it should be noted that because of the use of moving window, there is a monitoring blank region with initial window length. So the monitoring results of EDISSIM are displayed starting from the end time of the first moving window.

The monitoring results for Case 3 using EDISSIM and DISSIM are shown in Figure 3, where they both detect the violation in the first window since the disturbance is introduced at the very start. However, comparatively, the dissimilarity values of EDISSIM always stay outside the control limits, indicating stably the fault throughout the whole operation duration. On the other hand, DISSIM shows unstable monitoring trend yielding frequently missing alarms whose monitoring values stay below the normal confidence region. Moreover, the monitoring results using EDISSIM and DISSIM methods in each case are compared and summarized in Table 2. For the observations obtained during the whole process duration, the percentage of false monitoring results to all ones is termed "error rate (%)" in the present work and is employed to evaluate the monitoring performance. It is affected by the number of false alarms before the occurrence of fault and missing alarms after the abnormality. Consequently, a monitoring method with smaller error rate is regarded to function better in detecting faults. Conducting more simulations will help to confirm the effectiveness of the proposed method, but we can still get a general and clear impression of the superior monitoring performance of EDISSIM over DISSIM from Table 2. For example, when there is no abnormality in the process (Case 0), the monitoring dissimilarity indices using EDISSIM stay well below the control

Table 1. Settings of Operation Conditions

Case	Type	Size	Introduced Time
0	Normal condition	–	–
1	Mean shift of w	0.0 → 0.5	150
2	Mean shift of w	0.0 → 1.5	25
3	Change of parameter from u_1 to x_2	3.0 → 4.5	0
4	Change of parameter from u_1 to x_2	3.0 → 1.0	100

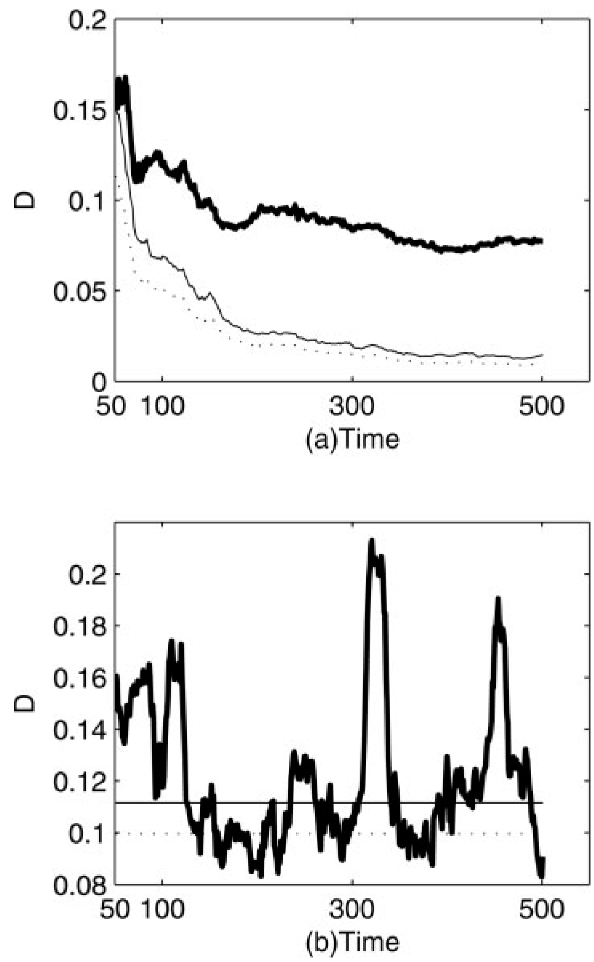

Figure 3. Monitoring results of 2×2 process in Case 3 using (a) EDISSIM and (b) DISSIM (dashed line, 99% and 95% control limits; solid line, on-line dissimilarity values).

Table 2. Error Rate (%) of Alarms: Applications to 2 × 2 Process

Case	Error Rate (%)	
	EDISSIM	DISSIM
0	0	5.1
1	8.2	44.6
2	0	0
3	0	22.6
4	5.7	5.3

limits throughout the process with 0 error rate, indicating that the batch tracks the normal operation trajectory while those using DISSIM go beyond control limits occasionally with 5.1% error rate. For different kinds of abnormal operations with different degrees, the proposed method can fast and correctly detect the fault with small magnitudes, because such abnormal operation conditions investigated in this application affect the correlations among variables and change the distributions of data structure. With the development of process, the monitoring results show a persistent trend for the occurred faults because of the use of variable windows. Contrastively, DISSIM method displays fluctuating results and misses alarms sometimes. Generally, for those processes with continuous states, EDISSIM method achieves the same perfect effects as DISSIM, or even better, which initially proves the reliability of EDISSIM.

Application 2: Fed-batch penicillin fermentation production

In this section, the proposed monitoring method is applied to the monitoring problems of a well-known benchmark process, fed-batch penicillin fermentation industrial process[23,24] to investigate its practicability. A flow diagram of the penicillin fermentation process is given in Figure 4. The production of secondary metabolites such as antibiotics has been the subject of many studies because of its academic and industrial importance. In typical operating procedure for the modeled fed-batch fermentation, most of the necessary cell mass is obtained during the initial preculture period. When most of the initially added substrate has been consumed by the microorganisms, the substrate feed begins. Generation of the penicillin starts in the exponential growth period and continues to be produced until the stationary period. A low substrate concentration in the fermentor is necessary for achieving a high product formation rate due to the catabolite repressor effect. Consequently, glucose is fed continuously during fermentation instead of being added one-off at the beginning.

The Monitoring and Control Group of the Illinois Institute of Technology has developed a simulator (PenSim v 1.0) that is capable of simulating the production of penicillin under various operating conditions (http://www.chee.iit.edu/~cinar). The process variables retained in this work are shown in Table 3. One hundred and one sets of time-series X^i $(400 \times 11)(i = 1,2,\ldots,101)$ are generated from 101 normal batches respectively, and the initial monitoring models are developed based on one normal batch with 50 as the

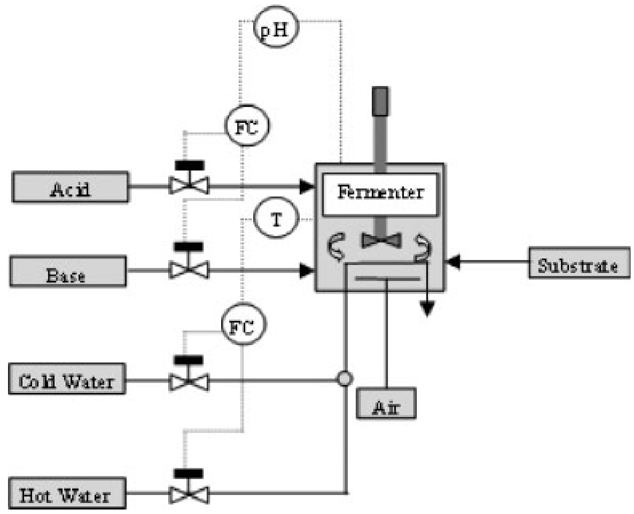

Figure 4. A flow diagram of the penicillin fermentation process.

original window length. The other 100 normal batches cover the mimic normal random variations under the same operation conditions. Then they are employed to train the proper 99% and 95% control limits according to the relevant confidence bounds in Γ probability distribution shown in Figure 5a. The reliability of Γ-distribution hypothesis focusing on the dissimilarity values at each time sampling can be also checked in Figure 5b. Taking the first moving window for example, the D indices more frequently take on values that very close to 0.4 agreeing well with the case shown in Figure 5a. So the density estimate in EDISSIM analysis makes it clear that the batch-to-batch D values at the same time conform to the assumed Γ-distribution. Thus the calculation of control limits based on Γ-distribution statistical rule is reliable and well-founded.

To gain more insight into the performance of the proposed method for batch processes monitoring, the simulator generates fault batches listed in Table 4, which include different types of disturbances introduced at different time. For comparison, conventional MPCA method is also implemented, where we use the first approach by Nomikos and MacGregor[11,12] to fill the missing future observations with zeros. The monitoring and detecting results using the two methods are shown in Figures 6–8 respectively. From an overall view, it can be seen that after the occurrence of fault, the proposed

Table 3. Variables Used in the Penicillin Fermentation Process

No.	Variables
1	Aeration rate (l/h)
2	Agitator power (W)
3	Substrate feed rate (l/h)
4	Substrate feed temperature (K)
5	Dissolved oxygen concentration (g/l)
6	Culture volume (l)
7	Carbon dioxide concentration (g/l)
8	pH
9	Fermentor temperature (K)
10	Generated heat (kcal/h)
11	Cooling water flow rate (l/h)

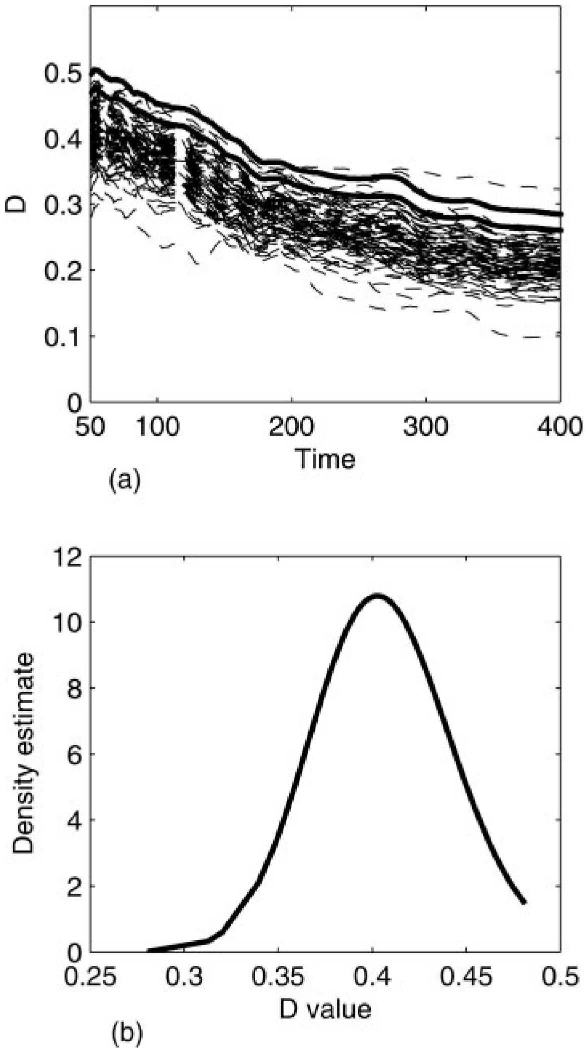

Figure 5. (a) Control limit of EDISSIM method (bold dot line, 99% and 95% control limits; dashed line, dissimilarity indices to training data); (b) probability density estimate of the 1st window.

method shows gradual increase of dissimilarity values and steady fault detection results.

For the first fault occurring at 100 h outside the first time window, it presents fast and correct detection performance using EDISSIM method. From Figure 6a, on the occurrence of fault at time 100 h, the monitoring values go beyond the 95% confidence limits almost immediately, shortly overstep the 99% control region and display a stable monitoring trend since then. In case of MPCA shown in Figure 6b, the T^2 chart does not detect a significant deviation whereas the SPE values indicate obvious false alarms from the beginning and do not clearly exceed the normal boundary until about 150 h with an abrupt rise. The initial false alarms may arise from the less real measurements and more estimated unavailable measurements.

Table 4. A Summary of Fault Types Introduced at Different Time of Fermentation

Fault No.	Fault Type	Occurrence (h)
1	10% step increase in agitator power	100
2	0.05% ramp decrease in substrate feed rate	30
3	0.1% ramp decrease in agitator power	10

Further investigations focusing on various faults in different extents at different introduced time are shown in Figures 7 and 8, which give a more authoritative and all-around demonstration of the proposed method's effectiveness. The two kinds of faults are introduced at the early period of the process. When enough data are collected to form the first time

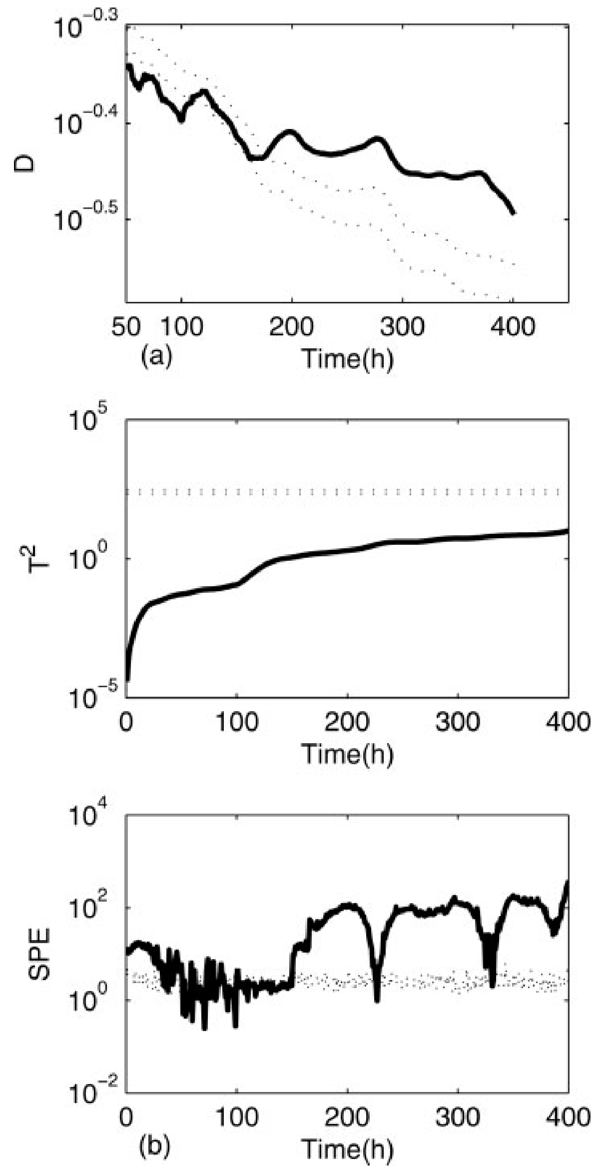

Figure 6. Monitoring results in case of Fault 1 using (a) EDISSIM and (b) MPCA (dashed line, 99% and 95% control limits; solid line, on-line monitoring values).

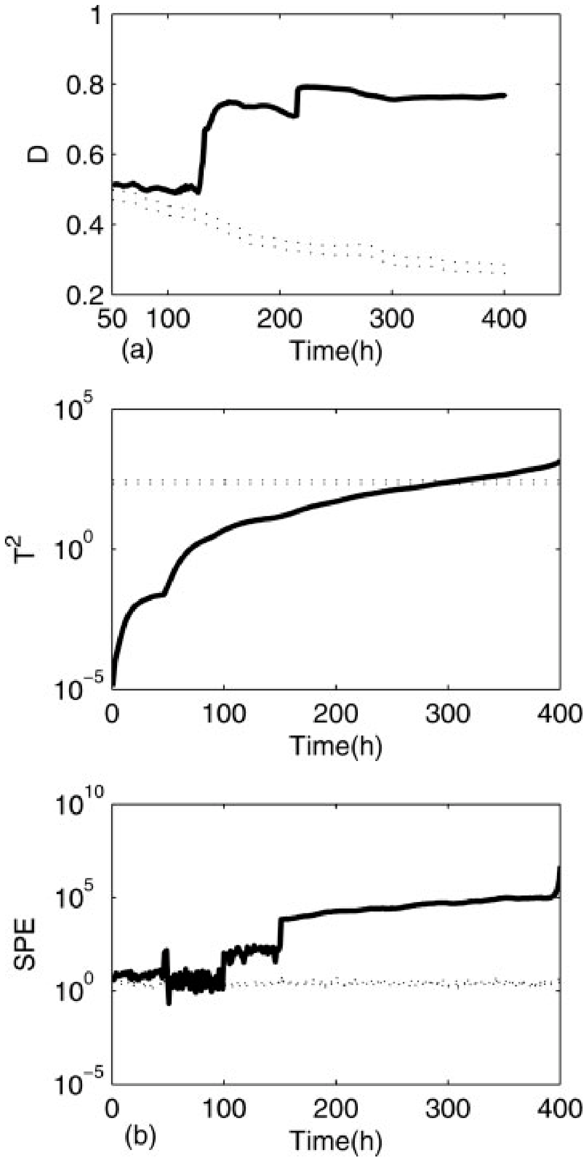

Figure 7. Monitoring results in case of Fault 2 using (a) EDISSIM and (b) MPCA (dashed line, 99% and 95% control limits; solid line, on-line monitoring values).

window, the fault is detected immediately using EDISSIM since the disturbance information is included in the first monitoring window shown in Figures 7a and 8a. Comparatively, for both faults, SPE statistics of MPCA display false alarms at the very start or clear detection delay after the onset of actual upset. Generally speaking, EDISSIM exhibits more steady monitoring results while MPCA shows varying and fluctuating ones. Moreover, in case of Fault 2, the T^2 values exhibit a gradual increase trend around 30 h and do not go outside the confidence limits until time 300 h. It delivers that during the period after the onset of Fault 2, the abnormality first changes the residual part of the process variation and then with the effect propagated, a gradual change is caused in the systematic part of the process failure. Contrastively, the T^2 statistics of Fault 3 stay well within the normal boundary throughout the duration, failing to detect the fault successfully.

Once the abnormal condition is detected by the monitoring charts, it is desirable to analyze the contribution plots to give a reasonable explanation to the fault, which can indeed enhance the process understanding and improve the ability of fault detection and diagnosis. From Figure 9, for Fault 2, where the disturbance is caused by substrate feed rate, the contributions to D index in the initial time windows make it clear that Variable 3 (substrate feed rate) is the dominant, which agrees well with the real situation. However, the processes variables tend to complicatedly related to each other in practical batch industry. With the evolvement of time, if fault detection is delayed, the fault diagnosis will be more

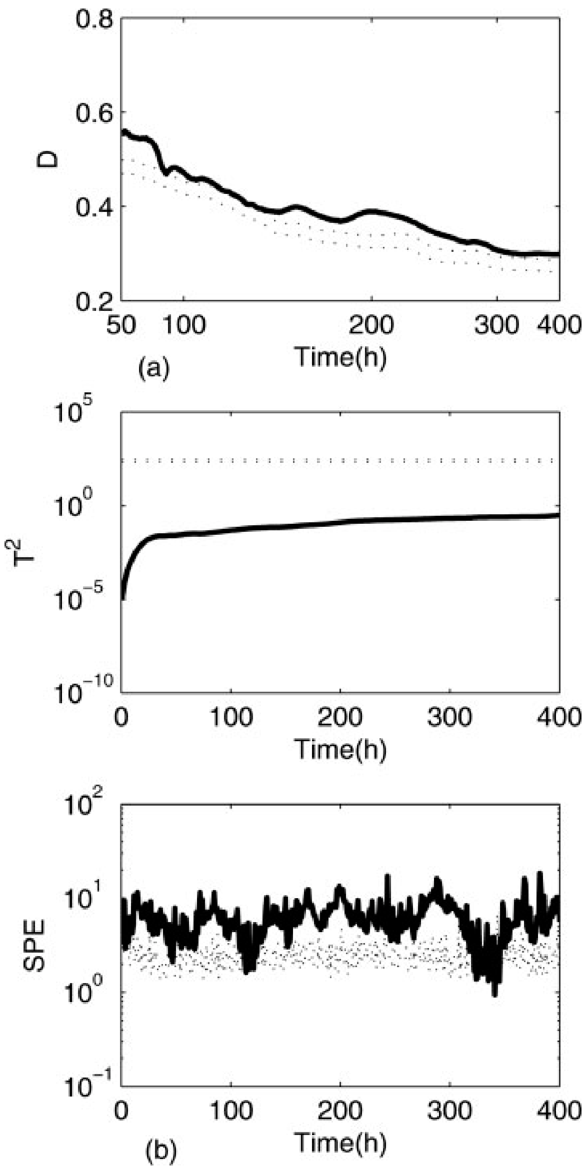

Figure 8. Monitoring results in case of Fault 3 using (a) EDISSIM and (b) MPCA (dashed line, 99% and 95% control limits; solid line, online monitoring values).

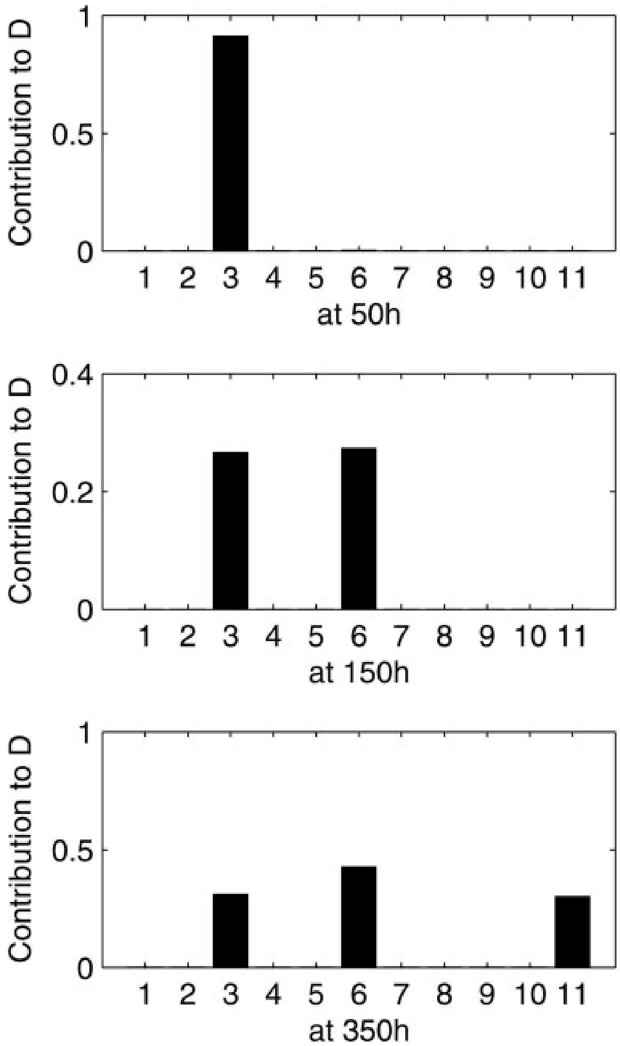

Figure 9. Contribution plot in case of Fault 2 (*x* axis is the process variables, *y* axis is the contribution to dissimilarity index *D*).

complicated and deceptive since all these variables will indicate significant deviations from their expected values. Combined with the contribution plots at 150 and 350 h, it can reveal that the variables with large contribution rate cannot be simply identified as the real cause of the fault only based on the contribution plots. Therefore, to identify the cause correctly, not only the information from contribution plots but also knowledge of the process has to be combined to give a reasonable and accurate explanation to the fault causes. Combined with the abnormality cause analysis using contribution plots, EDISSIM method gives believable fault detecting and diagnosis results when monitoring batch processes.

From the aforementioned, the serials of monitoring results have effectively demonstrated that the proposed method is valid for fault detecting and diagnosis in equal batch processes. Here, to testify the applicability of the proposed method to those batch processes with uncritical duration variation, the simulator generates two unequal batches: one is normal batch with 415h-length and the other is fault data with 410h-length, where 0.05% ramp decrease in substrate feed rate is introduced at 30 h. For the two cases, the differences in number of data points are 15 and 10 respectively compared with the fixed common 400h-duration, about 3.75% and 2.5% of the whole process length, which is deemed to be within the acceptable scope of duration-variation. Figure 10a shows the monitoring result for the normal batch, where the dissimilarity values stay well below the confidence limits. In Figure 10b, in response to the occurrence of fault, the dissimilarity values go beyond the normal trajectory at the very start since the upset is covered in the 1st moving window. Both of the cases prove the reliability of the proposed simple solution to handling unserious unequal-batches.

In conclusion, the simulation illustrations have demonstrated effectively that it is viable to monitor dynamic batch process using EDISSIM method with the control limits derived from Γ-distribution statistics.

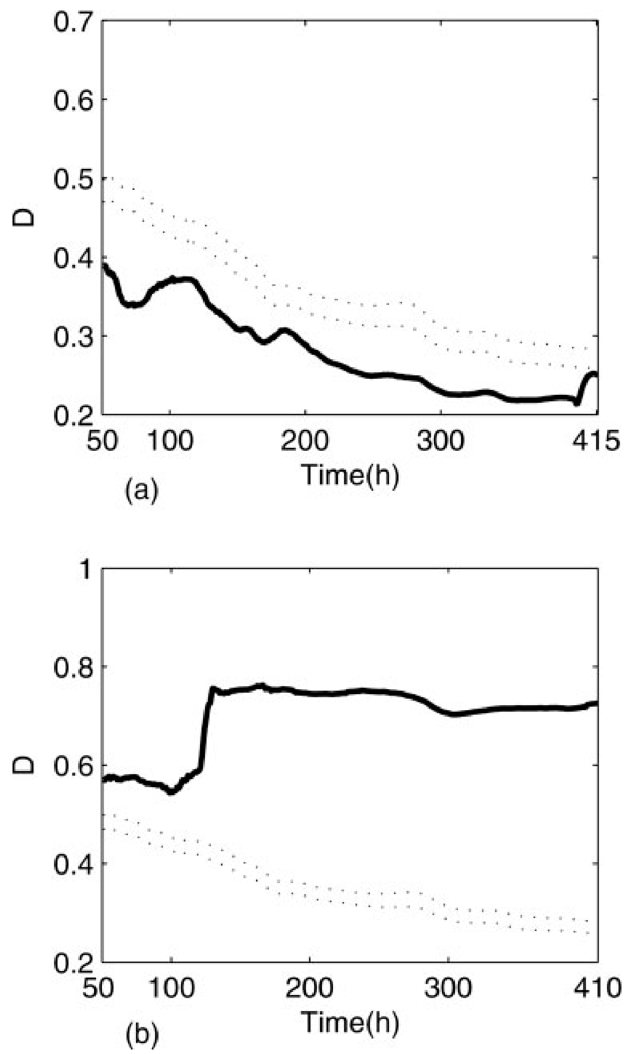

Figure 10. Monitoring results of uneven-length batches using the proposed method in case of (a) normal and (b) fault (dashed line, 99% and 95% control limits; solid line, online dissimilarity values).

Conclusions

A multivariate statistical monitoring method has been proposed in order to improve the performance of batch process monitoring. This method, termed EDISSIM, is based on the dissimilarity analysis of variable moving time-windows, which is introduced to quantitatively evaluate the differences between batch data sets. Different from PCA, it focuses on the correlations among process variables and the data distribution nature. Moreover, the theoretical basis of determining control limits is developed based on the Γ-distribution analysis of batch-to-batch variations. The presented method is such developed under the implicit assumption of even operation duration that process time can be used as an indicator variable to denote which monitoring model should be employed. However, in practice, varying duration is a common feature of industrial batch processes. The proposed method also supplies a simple and promising way to those cases where the problem of uneven-length is not so serious. Both applications to a simple 2×2 process and simulated fed-batch penicillin production demonstrate the feasibility and effectiveness of the proposed method. Therefore, a process monitoring system based on dissimilarity analysis of moving windows will be promising.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 60374003) and Project 973 (No. 2002CB312200), China.

Literature Cited

1. Kourti T, MacGregor JF. Process analysis, monitoring and diagnosis, using multivariate projection methods. Chemometrics Intell. *Lab Syst*. 1995;28:3–21.
2. Kosanovich KA, Dahl KS, Piovoso MJ. Improved process understanding using multiway principal component analysis. *Ind Eng Chem Res*. 1996;35:138–146.
3. Wold S, Kettaneh N, Friden H, Holmberg A. Modelling and diagnosis of batch processes and analogous kinetic experiments. Chemometrics Intell. *Lab Syst*. 1998;44:331–340.
4. Louwerse DJ, Smilde AK. Multivariate statistical process control of batch processes based on three-way models. *Chem Eng Sci*. 2000;55:1225–1235.
5. Sprange ENM, Ramaker HJ, Westerhuis J, Smilde A. Critical evaluation of approaches for on-line batch process monitoring. *Chem Eng Sci*. 2002;57:3979–3991.
6. Ündey C, Cinar A. Statistical monitoring of multistage, multiphase batch processes. *IEEE Control Syst Mag*. 2002;22:40–52.
7. Kourti T. Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions. *J Chemometrics*. 2003;17:93–109.
8. Kresta JV, Marlin TE, MacGregor JF. Development of inferential process models using PLS. *Comput Chem Eng*. 1994;18:597–611.
9. Svante Wold, Paul Geladi, Kim Esbensen. Multi-way principal components- and PLS-analysis. *J Chemometrics*. 1987;1:41–56.
10. Nomikos P, MacGregor JF. Monitoring of batch processes using multi-way principal component analysis. *AIChE J*. 1994;40:1361–1375.
11. Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. *Technometrics*. 1995;37:41–59.
12. Nomikos P, MacGregor JF. Multiway partial least squares in monitoring batch processes. Chemometrics Intell. *Lab Syst*. 1995;30: 97–108.
13. Kano M, Ohno H, Hasebe S, Hashimoto I. Process monitoring based on dissimilarity of time series data. *Kagaku Kogaku Ronbunshu*. 1999;25:1004–1009.
14. Kano M, Nagao K, Ohno H, Hasebe S, Hashimoto I. Dissimilarity of process data for statistical process monitoring. Proceedings of IFAC symposium on advanced control of chemical processes (ADCHEM), Pisa, Italy, June 14–16, Vol. 1, 2000:231–236.
15. Kano M, Hasebe S, Hashimoto I, Ohno H. Fault detection and identification based on dissimilarity of process data. Preprints of European Control Conference (ECC), Porto, Portugal, September 4–7, 2001:1888–1893.
16. Kano M, Ohno H, Hasebe S, Hashimoto I. Statistical process monitoring based on dissimilarity of process data. *AICHE J*. 2002;48: 1231–1240.
17. Fukunaga K, Koontz WLG. Application of the Karhunen-Loeve expansion to feature selection and ordering. *IEEE Trans Comput*. 1970;19:311–318.
18. Zhao X, Xu W, Shi Y. *Mathematical and Statistics*, *2nd ed*. Beijing: Science Press, 2002.
19. He P. *Mathematical Statistics and Multivariate Statistics*. Chengdu: Southwest Jiaotong University Press, 2004.
20. Kano M, Ohno H, Hasebe S, Hashimoto I. Contribution plots for fault identification based on the dissimilarity of process data. AIChE Annual Meeting, paper 255e, Los Angeles, CA, 2000, Nov. 12–17.
21. Lu N, Gao F, Yang Y, Wang F. A PCA-based modeling and online monitoring strategy for uneven-length batch processes. *Ind Eng Chem Res*. 2004;43:3343–3352.
22. Ku W, Storer RH, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics Intell Lab Syst*. 1995;30:179–196.
23. Birol G, Ündey C, Cinar A. A modular simulation package for fed-batch fermentation: penicillin production. *Comput Chem Eng*. 2002;26:1553–1565.
24. Birol G, Ündey C, Parulekar SJ, Cinar A. A morphologically structured model for penicillin production. *Biotechnol Bioeng*. 2002;77: 538–552.

Manuscript received Oct. 25, 2006, and revision received Feb. 6, 2007.