Concurrent Phase Partition and Between-Mode Statistical Analysis for Multimode and Multiphase Batch Process Monitoring

Chunhui Zhao
State Key Laboratory of Industrial Control Technology, Dept. of Control Science and Engineering, Zhejiang University, Hangzhou 310007, China

DOI 10.1002/aic.14282
Published online November 16, 2013 in Wiley Online Library (wileyonlinelibrary.com)

The exiting automatic phase partition and phase-based process monitoring strategies are in general limited to single-mode multiphase batch processes. In this article, a concurrent phase partition and between-mode statistical modeling strategy (CPPBM) is proposed for online monitoring of multimode multiphase batch processes. First, the time-varying characteristics of batch processes are concurrently analyzed across modes so that multiple sequential phases are simultaneously identified for all modes. The feature is that both time-wise dynamics and mode-wise variations are considered to get the consistent phase boundaries. Then within each phase, between-mode statistical analysis is performed where one mode is chosen for the development of reference monitoring system and the relative changes from the reference mode to each alternative mode are analyzed. From the between-mode perspective, each of the original reference monitoring subspaces, including systematic subspace and residual subspace, are further decomposed into two monitoring subspaces for each alternative mode, which reveal two kinds of between-mode relative variations. The part which shows significant increases represents the variations that will cause alarm signals if the reference models are used to monitor the alternative modes, whereas the part that shows no increases will not issue alarms. By modeling and monitoring different types of between-mode relative variations, the proposed CPPBM method can not only efficiently detect faults but also offer enhanced process understanding. It is illustrated with a typical multiphase batch process with multiple modes.
 AIChE J, 60: 559–573, 2014
Keywords: multimode, between-mode analysis, concurrent phase partition, multiphase, batch process monitoring

Introduction

As one important type of manufacturing, batch and semi-batch processes have been playing an important role in the processing of specialty chemical, semiconductor, food and biology industries for producing high-value-added products to meet today's rapidly changing market. Characterized by finite duration, batch process operation is in general carried out in different steps to produce products of desired quality at the lowest possible cost. Process disturbances which affect both process and product reproducibility may vary with the development of time and from batch to batch. For process safety and quality improvement,[1–7] various process monitoring and diagnosis strategies have drawn increasing attention recently. Multivariate statistical methods such as principal component analysis (PCA)[8] and partial least square (PLS)[9,10] have been successfully used in modeling multivariate continuous processes. Several extensions of the conventional PCA/PLS to batch processes have also been reported, among which most batch process monitoring methods are based on multiway principal component analysis (MPCA) and partial least squares.[11–16] However, conventional MPCA method may be difficult to reveal the time-wise changes of process correlations because it takes the entire batch data as a single object. Also, it is difficult for online application as the entire new batch measurements are not available up to the concerned time so that the unknown future values have to be estimated. Variable-unfolding modeling method[3,17,18] has been developed to overcome the aforementioned problem. This method can be put into online application without fulfilling missing data, which, however, may not well reflect multiplicity of process characteristics across different time regions as it uses single monitoring model for the entire process.

Considering that the multiplicity of operation phases is an inherent nature of many batch processes and each phase exhibits significantly different underlying behaviors, it is desirable to develop multiple phase-based models.[6,18] Then, each model represents a specific phase and explains the local process behaviors, which can effectively enhance process understanding and improve monitoring reliability. Kosanovich et al.[19] and Dong and McAvoy[20] developed two MPCA/nonlinear MPCA models to analyze the phase-specific nature of a two-phase jacketed exothermic batch

Correspondence concerning this article should be addressed to C. Zhao at chhzhao@zju.edu.cn.

chemical reactor, where monitoring results show that the two phase-based models are more powerful than a single model. Their phase models, however, inherit the common weakness of the conventional MPCA model that the unavailable future data in an evolving batch should be estimated for online monitoring which may distort the real dynamics of process correlations. Zhao et al.[21] developed an automatic phase division method based on clustering algebra to capture changes of process characteristics and phase-based sub-PCA model were developed for process monitoring. It is based on the fact that changes of the process correlations may relate to the phase shift in multiphase batch processes. Phase-based sub-PCA model does not require fulfilling missing process observations and preserves the dynamic relationships. Also, it can help to understand the operation process, revealing the similar process characteristics within a certain time region which can be described by the same model structure. Since then, phase-based sub-PCA/PLS modeling methods[22–27] have been widely developed to handle different problems in batch processes with multiphase nature. However, using clustering-based phase division algorithm, segments at different time within a batch may be mixed as a single phase and some points in one phase may be assigned to other phases based on the evaluation of process similarity, which thus may make the phase division results hard to understand. A heavy postprocessing[21] has to be made to treat the clustering results. Moreover, the clustering result is greatly influenced by the similarity evaluation index (i.e., distance), which, however, is not directly related with the purpose of process monitoring.

Moreover, the wide and successful report of phase-based modeling and monitoring strategies, however, are in general limited to single-mode batch processes. In practice, manufacturing processes frequently go through mode changes due to various factors, such as alterations of feedstocks and compositions, fluctuations in the external environment, and various product specifications. Especially, to satisfy fast changing market demands, the manufacturing strategies and operation conditions have to be adjusted frequently which also make multimode a popular and important problem. Traditional modeling methods in general experience the model mismatch problem where the process behaviors to be evaluated operate under a condition other than the one the reference model is built upon. These shifts in operation mode can cause frequent false alarms even though the process is under another nominal operation mode. The statistical modeling and monitoring of multimode processes is a challenging problem, which has drawn increasing attention recently[28–30] and most of which are for continuous processes.[28,29] However, few research work regarding statistical analysis and monitoring of multimode and multiphase batch processes have been reported.

This article proposes a concurrent phase partition and between-mode statistical modeling strategy (CPPBM) for monitoring of multimode and multiphase batch processes. It includes two key features: automatic concurrent phase partition for all modes from a consistent perspective and comprehensive subspace decomposition based on analysis of between-mode relative changes. The objective of concurrent phase partition is to automatically and simultaneously identify data segments in sequence along time direction for different modes of batch processes. Sequential and cross-mode consistent phase partition results provide the basis for the following phase-based between-mode analysis. The relative changes from the reference mode to alternative modes are analyzed in each phase. The underlying process behaviors of each alternative mode can be decomposed into four subspaces, revealing different types of relative variations in comparison with the reference mode. They are then modeled and online monitored, respectively. The proposed method can not only efficiently distinguish normal modes from fault cases but also offers enhanced process understanding. Its feasibility and performance are illustrated with a typical multiphase batch process with multiple modes.

Methodology

Concurrent phase partition algorithm

Considering the time-varying process characteristics and similarity within a certain time region as well, it is a natural idea to separate the batch process into multiple phases and develop different models to capture their different process characteristics. There have been many phase partition methods[21–27] for single-mode batch processes, which revealed the modeling phases from different viewpoints and based on different principles. Assuming that no prior process knowledge is available, the phase information should be automatically identified by capturing changes of underlying characteristics. In previous work,[21–24] a variant k-means clustering algorithm was used to check multiplicity of phases in batch processes. For readability, the clustering-based algorithm is briefly revisited in Appendix. However, as mentioned before, it did not consider the time sequence of process phases, which requires a heavy and complicated postprocessing. Also, the process similarity was simply evaluated by distance index where two models with small distance may not necessarily mean that they can result in similar monitoring performance because the distance index is not directly related with how the monitoring action is conducted. Based on the specific purpose in practical application (process monitoring in the present work), the similarity evaluation should be directly related with its influences on monitoring performance, for example, false alarm and missing alarm. In the present work, considering the multimode case and also to overcome the aforementioned problem, the changes of process characteristics are checked for all modes in a consistent way with respect to their influences on monitoring system.

In each batch run (batch index $i=1,2,...,I$), assume that J process variables are measured online at $k=1,2,...,K$ time instances throughout the operation cycle, forming each regular batch dataset, denoted as $\mathbf{X}(K\times J)$. In the present work, batches are of equal length without special declaration so that the specific process time can be used as an indicator to data preprocessing. For uneven-length batches, they can be synchronized using different methods, such as dynamic time warping[31] or instrumental variable.[3] Therefore, for each mode, a three-way data array is available by collecting data from I batches, $\underline{\mathbf{X}}_m(I\times J\times K)$ $(m = 1,2,...,M)$. There are M modes in all.

The basic analysis and modeling unit is time-slice data matrix in each mode. The basic idea of concurrent phase division is to simultaneously find the phase landmarks for all modes to make sure that they enter and leave each phase at the same time. To achieve this purpose, three rules are followed for automatic phase division, including the similar characteristics along time direction within each mode, the

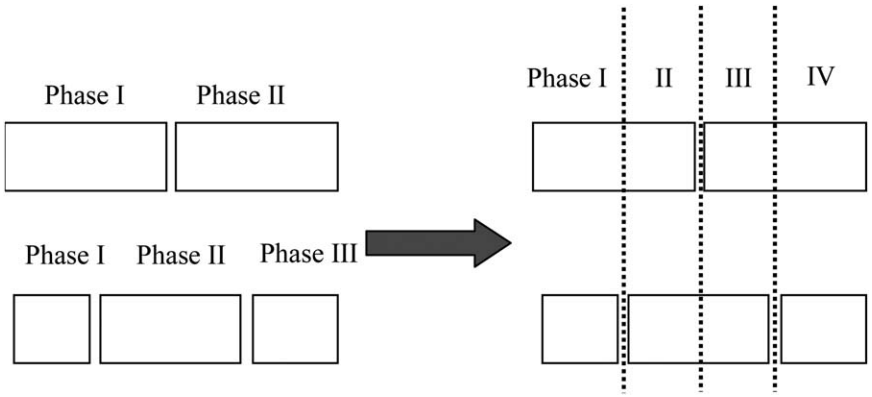

Figure 1. Simple illustration of concurrent phase division for two-mode batch processes.

time sequence, and the concurrent consideration of all modes. In one word, the phase landmarks for all modes are identified by simultaneously and sequentially checking the changes of variable correlations of time-slices from the process beginning step by step within each mode. It is thus termed concurrent multimode step-wise sequential phase partition algorithm here.

Prepare the time-slice data matrix $\mathbf{X}_{m,k}(I\times J)$ at each time k from the process beginning of each mode. The basic procedure is described as follows:

Step 1: Input the normalized time-slice data matrix $\bar{\mathbf{X}}_{m,k}(I\times J)$ which are preprocessed to have zero mean and standard deviation respectively.

Step 2: Perform PCA algorithm on the time-slice data matrices and get the initial time-slice models, $\mathbf{P}_{m,k}\left(J\times R_{m,k}\right)$, where the number of principal components (PCs) $R_{m,k}$ is retained in each time-slice model to keep most of the process variability (90% here). Then, find the number of PCs that occurs most at different time intervals and set it as the unified dimension of PCA models for each mode, R_m. Calculate residual related statistic, squared prediction errors (SPE), after the explanation of time-slice PCA models at each time for each mode and determine the time-slice confidence limit, $Ctr_{m,k}$, by a weighted Chi-squared distribution.[32] It represents the reconstruction power of time-slice PCA models to each time-slice for each mode.

Step 3: From the beginning of batch processes in each mode, add the next time-slice one by one to the existing ones and variable-unfold them, $\mathbf{X}^{v}_{m,k}(Ik\times J)$(where subscript k denotes the current time). Perform PCA on the rearranged data matrix and get the new PCA model, $\mathbf{P}^{v}_{m,k}(J\times R_m)$, which is called time-segment model. Calculate the SPE values of each time-slice data matrix after the explanation of the current time-segment PCA model ($\mathbf{P}^{v}_{m,k}(J\times R_m)$) for each mode and determine the confidence limit, $Ctr^{v}_{m,k}$, by a weighted Chi-squared distribution. It represents the reconstruction power of this time-segment PCA model to each time-slice within the concerned time region for each mode.

Step 4: Compare $Ctr^{v}_{m,k}$ with $Ctr_{m,k}$ for each time slice within the concerned time region of each mode. If it finds that there are three consecutive samples showing $Ctr^{v}_{m,k} > \alpha_m * Ctr_{m,k}$ in one mode, the addition of the current time-slice have imposed great influences on the time-segment PCA monitoring model and the resulting monitoring performance for this mode. α_m is a constant vector attached to $Ctr_{m,k}$ for each mode, termed relaxing factor here, which determines how much the time-segment PCA model in the mth mode is allowed to be less representative (i.e., insufficient reconstruction power) than the same-dimensional time-slice PCA models within the concerned time region. It is clear that different values of relaxing factor (α_m) can be attached to each mode. The time slices before k are denoted as one subphase for all modes.

Step 5: For all modes, remove the time-slice data matrices of the first subphase and the left-batch process data are now used as the new input data in Step 3. Recursively, repeat Steps 3–4 based on the updated process beginning for each mode to find the following subphases.

From the above procedure, to identify phases for each mode, variable-unfolding modeling unit is arranged iteratively by adding new time-slices sequentially and comparing the reconstruction power of resulting time-segment model with that of time-slice models. It is noted that the time-segment model is kept to be of the same dimension as that of time-slice models so that their reconstruction power can be fairly compared. If the process characteristics of some time-slices are similar, they can be explained by a unified time-segment model, revealing similar reconstruction power with that of those same-dimensional time-slice models. The output is a simultaneous phase partition along time direction for all modes, which presents two features: one is that the alternation of different phases is consecutive along time direction and the other is that the phase boundaries are consistent across modes, which is different from mode-specific separate phase division where the phase landmarks may disagree from mode to mode. One simple illustration is shown in Figure 1 for two-mode batch processes to explain the concurrent phase division. Four phases are obtained by simultaneously considering the two modes which are composed of three phases and two phases, respectively. By considering multiple modes at the same time, the division results tend to reflect faster process dynamics among all modes within a certain time region. From another viewpoint, single-mode separate phase division can be regarded as one extreme case of the proposed concurrent division algorithm.

It is noted that α_m is an important parameter associated with mode m. It reflects the compromise between model accuracy and model complexity. In general, it can be set by trial and error so that each phase-representative regression model will not cover too many process patterns to ensure sensitiveness to changes of process characteristics. However, up to now, there is no definite criterion or uniform standard

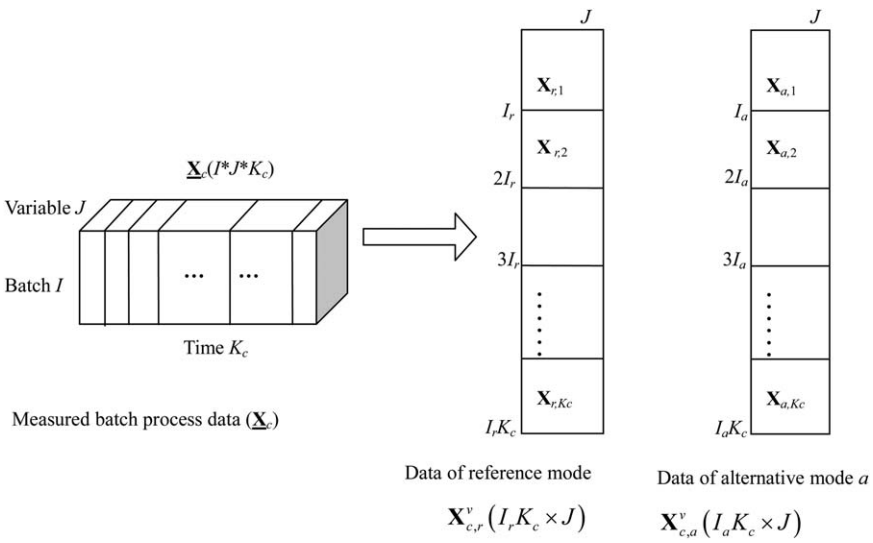

Figure 2. Illustration of phase-based variable-unfolding data arrangement for between-mode analysis.

to strictly quantify it. Therefore, its determination is inevitably affected more or less by artificial subjectivity factors. The investigation on the determination of α_m is meaningful and deserves further devotions in future.

Between-mode statistical modeling

For multimode processes, instead of mode-specific separate modeling method, between-mode analysis can reveal more interesting information, where, the general idea is to analyze the relative changes from one mode to another. One important question is: what kinds of relative changes should be focused on in between-mode analysis for the purpose of online monitoring? In the present work, one reference mode is chosen with random and the others as alternative modes are related with the reference one to reveal the between-mode relative changes regarding their different influences on monitoring results. For each alternative mode, four subspaces are decomposed based on whether the variations in alternative modes increase in comparison with those in the reference mode in principal component subspace (PCS) and residual subspace (RS) of PCA system, respectively. It is based on the fact that the increased between-mode relative variations in each alternative mode are responsible for alarm signals when they are monitored by reference models. The specific modeling procedure is described as follows.

First, as shown in Figure 2, two phase-representative datasets are prepared using variable-unfolding in each phase for the reference mode and one alternative mode, $\mathbf{X}^v_{c,r}(I_rK_c\times J)$ and $\mathbf{X}^v_{c,a}(I_aK_c\times J)$ (where subscript r means the reference mode and a denotes one alternative mode; c indicates phase, and superscript v means variable-unfolding is used for data arrangement; K_c denotes the phase duration for each mode, I_r and I_a are the number of batches in the reference mode and alternative mode, respectively), each being composed of the same number of variables and maybe different number of samples. The data in $\mathbf{X}^v_{c,r}$ and $\mathbf{X}^v_{c,a}$ have been centered and scaled to be of zero mean and unit standard deviation, respectively.

In PCS of PCA monitoring system:

Step (1): Perform PCA algorithm on $\mathbf{X}^v_{c,r}$ to get monitoring models

$$\begin{cases}\mathbf{T}^v_{c,r}=\mathbf{X}^v_{c,r}\mathbf{P}_{c,r}\\ \mathbf{E}^v_{c,r}=\mathbf{X}^v_{c,r}\mathbf{P}^e_{c,r}\mathbf{P}^{e\ \mathrm{T}}_{c,r}\end{cases} \tag{1}$$

where $\mathbf{T}^v_{c,r}\left(I_rK_{c,r}\times R_{c,r}\right)$ and $\mathbf{P}_{c,r}\left(J\times R_{c,r}\right)$ are PCs and the corresponding principal loadings in PCS; $\mathbf{E}^v_{c,r}\left(I_rK_{c,r}\times J\right)$ and $\mathbf{P}^e_{c,r}\left(J\times R^e_{c,r}\right)$ are PCA residuals and the corresponding residual loadings. $R_{c,r}$ is the number of retained PCs determined by cumulative explained variance rate.[33] This reveals the first largest variation directions of normal process data which are also the monitoring directions of T^2 statistic. $R^e_{c,r}$ is the number of retained directions in RS, $R^e_{c,r}=J-R_{c,r}$.

Step (2): Project $\mathbf{X}^v_{c,a}$ onto $\mathbf{P}_{c,r}$ to get scores $\mathbf{T}^v_{c,a}=\mathbf{X}^v_{c,a}\mathbf{P}_{c,r}$, that is, the reference model $\mathbf{P}_{c,r}$ is used to monitor alternative mode. Define the ratio of variation between each alternative mode and the reference mode along different monitoring directions in PCS as

$$\mathrm{Ratio}_{c,a,i}=\frac{\mathrm{var}\left(\mathbf{T}^v_{c,a}(:,i)\right)}{\mathrm{var}\left(\mathbf{T}^v_{c,r}(:,i)\right)}\left(i=1,2,...,R_{c,r}\right) \tag{2}$$

where $\mathrm{var}()$ denotes the score variance around the center (zero here) and $(:,i)$ denotes the ith column vector in a matrix. So, $\mathrm{Ratio}_{c,a}$ is a R-dimensional vector composed of $\mathrm{Ratio}_{c,a,i}$.

Step (3): Sort the values of $\mathrm{Ratio}_{c,a}$ index. If the $\mathrm{Ratio}_{c,a,i}$ value is larger than one, it means that the corresponding variation in the alternative mode is larger than that in the reference mode, which thus is responsible for out-of-control T^2 when the reference model $\mathbf{P}_{c,r}$ is used to monitor the alternative mode. Keep the directions corresponding to $\mathrm{Ratio}_{c,a,i}$ values of larger than one, composing $\mathbf{P}_{c,a}{}^*\left(J\times R_{c,a}{}^*\right)$, and the number of retained principal directions is $R_{c,a}{}^*$. $\mathbf{P}_{c,a}{}^n$ are composed of the left directions in $\mathbf{P}_{c,r}\left(J\times R_{c,r}\right)$ after the extraction of $\mathbf{P}_{c,a}{}^*$, along which no increase of variation is found in alternative mode relative to the reference mode. Also, it is noted that the orthogonality of $\mathbf{P}_{c,a}{}^*$ is kept as well as $\mathbf{P}_{c,a}{}^n$.

Step (4): The process variations explained along $\mathbf{P}_{c,a}{}^{*}$ and $\mathbf{P}_{c,a}{}^{n}$, respectively, can be modeled as

$$\begin{cases} \widehat{\mathbf{X}}_{c,a,f}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}{}^{*}\mathbf{P}_{c,a}{}^{*\mathrm{T}} \\ \widehat{\mathbf{X}}_{c,a,o}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}{}^{n}\mathbf{P}_{c,a}{}^{n\mathrm{T}} \end{cases} \tag{3}$$

Clearly, they represent two parts of variations separated from PCS, which make contribution and no contribution to out-of-control T^2 monitoring statistic in the alternative mode, respectively, under the supervision of reference model $\mathbf{P}_{c,r}$. PCA is then performed on $\widehat{\mathbf{X}}_{c,a,f}$ and $\widehat{\mathbf{X}}_{c,a,o}$ to explain the major systematic variations in order with $R_{c,a,f}$ and $R_{c,a,o}$ components, respectively

$$\begin{cases} \widehat{\mathbf{X}}_{c,a,f}=\mathbf{T}_{c,a,f}\mathbf{P}_{c,a,f}{}^{\mathrm{T}} \\ \widehat{\mathbf{X}}_{c,a,o}=\mathbf{T}_{c,a,o}\mathbf{P}_{c,a,o}{}^{\mathrm{T}} \end{cases} \tag{4}$$

where $R_{c,a,f}=\mathrm{rank}\,(\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}{}^{*})$ and $R_{c,a,o}=\mathrm{rank}\,(\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}{}^{n})$. In this way, the increased relative variations in the alternative mode are separated from those nonincreased ones and modeled orderly by PCA.

In RS of PCA monitoring system:

Step (1): The same as Step (1) in PCS of PCA monitoring system.

Step (2): Project $\mathbf{X}_{c,a}^{v}$ onto the column space of $\mathbf{P}_{c,r}^{e}$ to get the variations along each residual direction in $\mathbf{P}_{c,r}^{e}$

$$\mathbf{E}_{c,a}^{v}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,r}^{e}\mathbf{P}_{c,r}^{e\ T}=\sum_{j=1}^{R_{c,r}^{e}}\mathbf{X}_{c,a}^{v}\mathbf{p}_{c,r,j}^{e}\mathbf{p}_{c,r,j}^{e\ \mathrm{T}} \tag{5}$$

The variation difference between each alternative mode and the reference mode along different monitoring directions in RS is defined as

$$\Delta_{c,a,i}=\|\mathbf{X}_{c,a}^{v}\mathbf{p}_{e,i}\mathbf{p}_{e,i}{}^{\mathrm{T}}\|^2-\|\mathbf{X}_{c,a}^{v}\mathbf{p}_{e,i}\mathbf{p}_{e,i}{}^{\mathrm{T}}\|^2\left(i=1,2,...,R_{c,r}^{e}\right) \tag{6}$$

where $\|\ \|$ denotes the Euclidean length. So, $\Delta_{c,a}$ is a $R_{c,r}^{e}$-dimensional vector composed of $\Delta_{c,a,i}$.

Step (3): Sort the values of $\Delta_{c,a}$ index. If the $\Delta_{c,a,i}$ index is larger than zero, it means that the variation in each alternative mode is larger than that in the reference mode, which thus may be responsible for out-of-control SPE monitoring statistic when reference model $\mathbf{P}_{c,r}^{e}$ is used to monitor the alternative mode. Keep the directions corresponding to $\Delta_{c,a,i}$ values of larger than zeros, composing $\mathbf{P}_{c,a}{}^{e*}\left(J\times R_{c,a}^{e\ *}\right)$, and the number of retained principal directions is $R_{c,a}^{e\ *}$. $\mathbf{P}_{c,a}^{e\ n}$ are composed of the left directions in $\mathbf{P}_{c,r}\left(J\times R_{c,r}\right)$ after the extraction of $\mathbf{P}_{c,a}{}^{*}$, along which no increase of variation is found in the alternative mode relative to the reference mode. Also, it is noted that the orthogonality of $\mathbf{P}_{c,a}^{e\ *}$ is kept as well as $\mathbf{P}_{c,a}^{e\ n}$.

Step (4): The process variations explained along $\mathbf{P}_{c,a}^{e\ *}$ can be modeled as

$$\widehat{\mathbf{X}}_{c,a,f}^{e}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}^{e\ *}\mathbf{P}_{c,a}^{e\ *\mathrm{T}} \tag{7}$$

It is clear out-of-control SPE monitoring statistic is only caused by $\widehat{\mathbf{X}}_{c,a,f}$. Also, considering that there are also systematic information in $\widehat{\mathbf{X}}_{c,a,f}$ besides residuals, PCA is then performed on $\widehat{\mathbf{X}}_{c,a,f}^{e}$ to extract the major systematic variations in order with $R_{c,a,f}^{e}$ components

$$\hat{\widehat{\mathbf{X}}}_{c,a,f}^{e}=\mathbf{T}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}} \tag{8}$$

where $R_{c,a,f}^{e}$ is determined by cumulative explained variance rate.[33] In this way, the major increased relative variations along $\mathbf{P}_{c,a}^{e\ *}$ in the alternative mode are separated from those nonincreased ones and modeled orderly by PCA. The final residuals are then calculated as

$$\mathbf{E}_{c,a}^{f}=\mathbf{E}_{c,a}^{v}-\hat{\widehat{\mathbf{X}}}_{c,a,f}^{e}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,r}^{e}\mathbf{P}_{c,r}^{e\ T}-\widehat{\mathbf{X}}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}} \tag{9}$$

It is clear that $\widehat{\mathbf{X}}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a}^{e\ *}\mathbf{P}_{c,a}^{e\ *\mathrm{T}}\mathbf{P}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}}=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}}$ as $\mathbf{P}_{c,a,f}^{e}$ are in fact linear combinations of $\mathbf{P}_{c,a}^{e\ *}$.

Based on the above between-mode subspace decomposition, two different subspaces are separated from the original PCS and RS, respectively. There are three systematic subspaces for each alternative mode by checking the between-mode relative changes, two from PCS and the third from RS. $\mathbf{P}_{c,a,f}$ and $\mathbf{P}_{c,a,f}^{e}$ capture those major increased variations in each alternative mode relative to the reference mode in PCA systematic and RSs, respectively.

CPPBM-based online monitoring

Based on the phase-based between-mode subspace decomposition, different monitoring statistics are calculated for the reference mode and each alternative mode in phase c by projecting the current phase data onto different subspaces.

For the reference mode

$$\begin{aligned} \mathbf{T}_{c,r}^{v}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,r} \\ \mathbf{E}_{c,r}^{v}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,r}^{e}\mathbf{P}_{c,r}^{e\ T} \end{aligned} \tag{10}$$

where $\mathbf{T}_{c,r}^{v}$ are the systematic scores and $\mathbf{E}_{c,r}^{v}$ are the residuals corresponding to two different monitoring subspaces spanned by $\mathbf{P}_{c,r}$ and $\mathbf{P}_{c,r}^{e}$, respectively.

For each alternative mode

$$\begin{aligned} \mathbf{T}_{c,a,f}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a,f} \\ \mathbf{T}_{c,a,o}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a,o} \\ \mathbf{T}_{c,a,f}^{e}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a,f}^{e} \\ \mathbf{E}_{c,a}^{f}&=\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,r}^{e}\mathbf{P}_{c,r}^{e\ T}-\mathbf{X}_{c,a}^{v}\mathbf{P}_{c,a,f}^{e}\mathbf{P}_{c,a,f}^{e\ \mathrm{T}} \end{aligned} \tag{11}$$

where, $\mathbf{T}_{c,a,f}$, $\mathbf{T}_{c,a,o}$, and $\mathbf{T}_{c,a,f}^{e}$ are the scores in systematic subspaces spanned by $\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}_{c,a,f}^{e}$, respectively; $\mathbf{E}_{c,a,o}^{f}$ are the final residuals. They represent different types of relative variations in each alternative mode from the between-mode perspective. The corresponding time-slice statistics are then separated from the variable-unfolding phase statistics based on the indication of process time. For example, $\mathbf{T}_{a,f,k}$ means the kth time-slice in alternative mode a separated from the phase-representative scores $\mathbf{T}_{c,a,f}$ where based on the indication of process time, the current phase c is known.

Based on the calculation of different time-slice statistics in different subspaces, the monitoring statistics are calculated

Table 1. Definitions of the Different Monitoring Subspaces in Each Phase *c* for Each Alternative Mode *a*

Subspace	Description	Dimension	Monitoring Statistics
$\mathbf{P}_{c,a,f}$	Major systematic variability that contributes to the increased deviations in alternative mode relative to the reference mode in PCS of PCA monitoring system	$R_{c,a,f}$	$T_{a,f}^2$
$\mathbf{P}_{c,a,o}$	Major systematic variability that makes no contribution to the increased deviations in alternative mode relative to the reference mode in PCS of PCA monitoring system	$R_{c,a,o}$	$T_{a,o}^2$
$\mathbf{P}_{c,a,f}^e$	Major systematic variability that contribute to the increased deviations in alternative mode relative to the reference mode in RS of PCA monitoring system	$R_{c,a,f}^e$	$T_{a,f}^{e\,2}$
$\mathbf{E}_{c,a}^f$	Final residual subspace	$R^{\#}-R_{c,a,f}-R_{c,a,o}-R_{c,a,f}^e$	SPE_a^f

$R^{\#}$ denotes the original full dimension.

for reference mode and each alternative mode at each time k. For the reference mode

$$T_{r,k,i}^2=(\mathbf{t}_{r,k,i}-\bar{\mathbf{t}}_{r,k})^{\mathrm{T}}\boldsymbol{\Sigma}_{r,k}^{-1}(\mathbf{t}_{r,k,i}-\bar{\mathbf{t}}_{r,k})$$
$$\mathrm{SPE}_{r,k,i}=\mathbf{e}_{r,k,i}^{\mathrm{T}}\mathbf{e}_{r,k,i} \tag{12}$$

where, subscript i denotes the ith batch in each time slice. $\bar{\mathbf{t}}_{r,k}$ denotes the mean vector of $\mathbf{T}_{r,k}$ which is zero due to data preprocessing at each time. $\boldsymbol{\Sigma}_{r,k}$ is a diagonal matrix with elements being the variance of each PC in time-slice score matrix $\mathbf{T}_{r,k}$.

For each alternative mode

$$T_{a,f,k,i}^2=(\mathbf{t}_{a,f,k,i}-\bar{\mathbf{t}}_{a,f,k})^{\mathrm{T}}\boldsymbol{\Sigma}_{a,f,k}^{-1}(\mathbf{t}_{a,f,k,i}-\bar{\mathbf{t}}_{a,f,k})$$
$$T_{a,o,k,i}^2=(\mathbf{t}_{a,o,k,i}-\bar{\mathbf{t}}_{a,o,k})^{\mathrm{T}}\boldsymbol{\Sigma}_{a,o,k}^{-1}(\mathbf{t}_{a,o,k,i}-\bar{\mathbf{t}}_{a,o,k})$$
$$T_{a,f,k,i}^{e\,2}=(\mathbf{t}_{a,f,k,i}^e-\bar{\mathbf{t}}_{a,f,k}^e)^{\mathrm{T}}\boldsymbol{\Sigma}_{a,f,k}^{e\,-1}(\mathbf{t}_{a,f,k,i}^e-\bar{\mathbf{t}}_{a,f,k}^e) \tag{13}$$
$$\mathrm{SPE}_{a,k}^f=\mathbf{e}_{a,k,i}^{f\,\mathrm{T}}\mathbf{e}_{a,k,i}^f$$

where, $\bar{\mathbf{t}}_{a,f,k}$, $\bar{\mathbf{t}}_{a,o,k}$, and $\bar{\mathbf{t}}_{a,f,k}^e$ denote the mean vectors of $\mathbf{T}_{a,f,k}$, $\mathbf{T}_{a,o,k}$, and $\mathbf{T}_{a,f,k}^e$, respectively, which are all zero

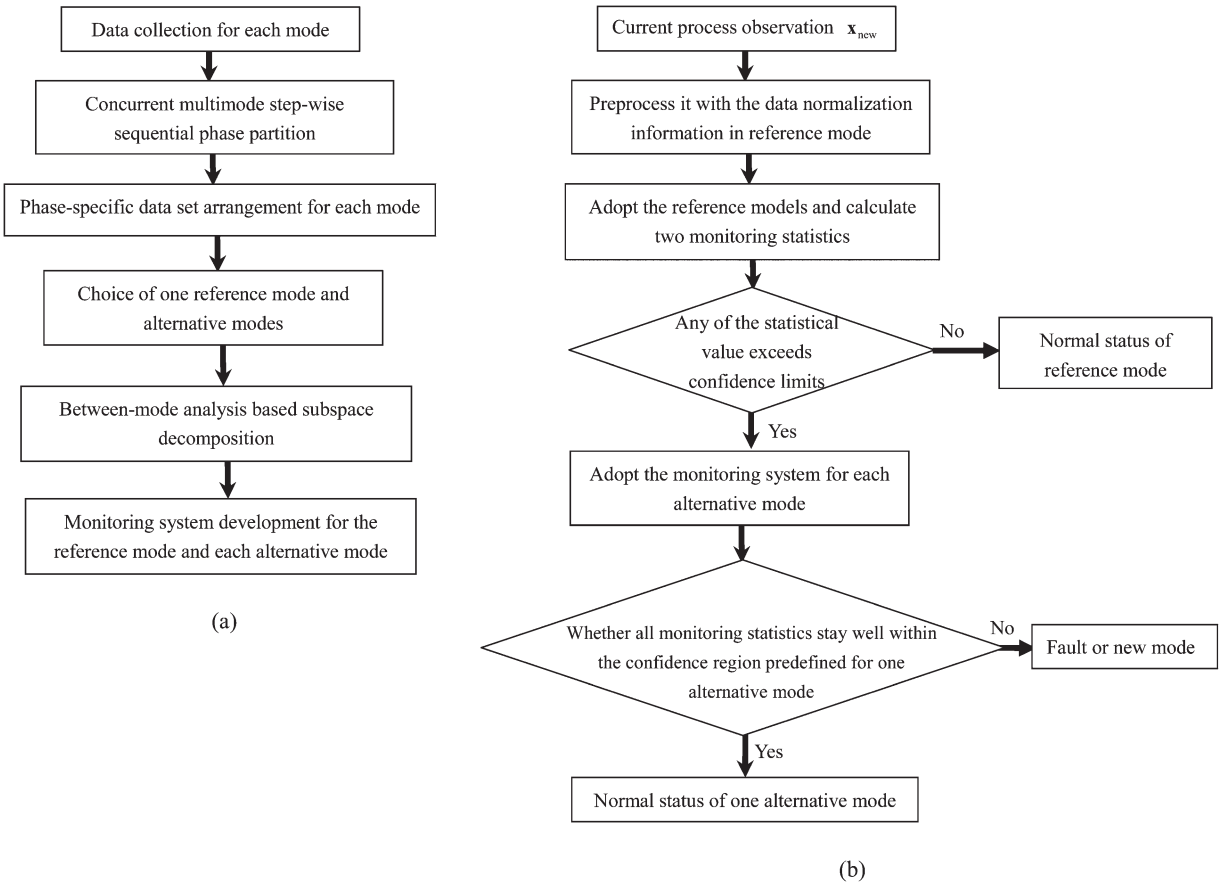

Figure 3. Flow chart of the proposed algorithm (a) model development and (b) online monitoring strategy.

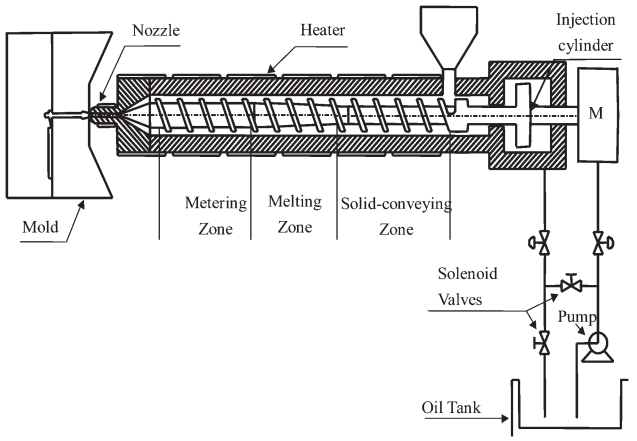

Figure 4. A simplified schematic diagram of the injection molding machine.

vectors due to the data preprocessing at each time. $\mathbf{\Sigma}_{a,f,k}$, $\mathbf{\Sigma}_{a,o,k}$, and $\mathbf{\Sigma}^{e}_{a,f,k}$ are diagonal matrices with elements being the variance of each PC in time-slice scores $\mathbf{T}_{a,f,k}$, $\mathbf{T}_{a,o,k}$, and $\mathbf{T}^{e}_{a,f,k}$, respectively. The meanings of different subspaces are summarized in Table 1 regarding each phase c for each alternative mode a.

The control limits in the systematic subspace for each mode are defined by the F-distribution with α as the significance factor[13] and in the RS, the representative confidence limit of SPE for each mode can be approximated by a weighted Chi-squared distribution.[32]

Whenever new observation at the kth time interval, $\mathbf{x}_{\text{new}}\,(J\times 1)$, is available, it is first preprocessed using the data normalization information from the reference mode. Based on the indication of process time, the current phase c is known. The normalized new observation is then projected onto the phase-specific reference monitoring system using Eq. 10 and the monitoring statistics are calculated using Eq. 12 where the related phase-representative monitoring models are adopted based on the indication of process time k

$$\begin{aligned}\mathbf{t}_{r,\text{new}}{}^{T}&=\mathbf{x}_{\text{new}}{}^{T}\mathbf{P}_{c,r}\\ \mathbf{e}_{r,\text{new}}{}^{T}&=\mathbf{x}_{\text{new}}{}^{T}\mathbf{P}^{e}_{c,r}\mathbf{P}^{e}_{c,r}{}^{T}\end{aligned}\tag{14}$$

$$\begin{aligned}T_{r,\text{new}}{}^{2}&=(\mathbf{t}_{r,\text{new}}-\bar{\mathbf{t}}_{r,k})^{\mathrm{T}}\mathbf{\Sigma}_{r,k}{}^{-1}(\mathbf{t}_{r,\text{new}}-\bar{\mathbf{t}}_{r,k})\\ \text{SPE}_{r,\text{new}}&=\mathbf{e}_{\text{new}}{}^{T}\mathbf{e}_{\text{new}}\end{aligned}\tag{15}$$

Compare the values of two monitoring statistics ($T_{r,\text{new}}{}^{2}$ and $\text{SPE}_{r,\text{new}}$) with the predefined control limits in PCS and RS, respectively. If both monitoring statistics stay well within the predefined normal regions for the reference mode,

Table 2. Eleven Process Variables Used in IM Process

No.	Variable's Descriptions	Unit
1	Valve 1	%
2	Valve 2	%
3	Screw stroke	mm
4	Screw velocity	mm/sec
5	Ejector stroke	mm
6	Mold stroke	mm
7	Mold velocity	mm/sec
8	Injection press	Bar
9	Barrel temperature zone 3	°C
10	Barrel temperature zone 2	°C
11	Barrel temperature zone 1	°C

Table 3. The Setting of Operation Conditions for Three-Mode IM Process

	Operation Conditions	
Mode #	Barrel Temperature (BT) (°C)	Packing Pressure (PP) (Bar)
1	180	25
2	200	30
3	220	35

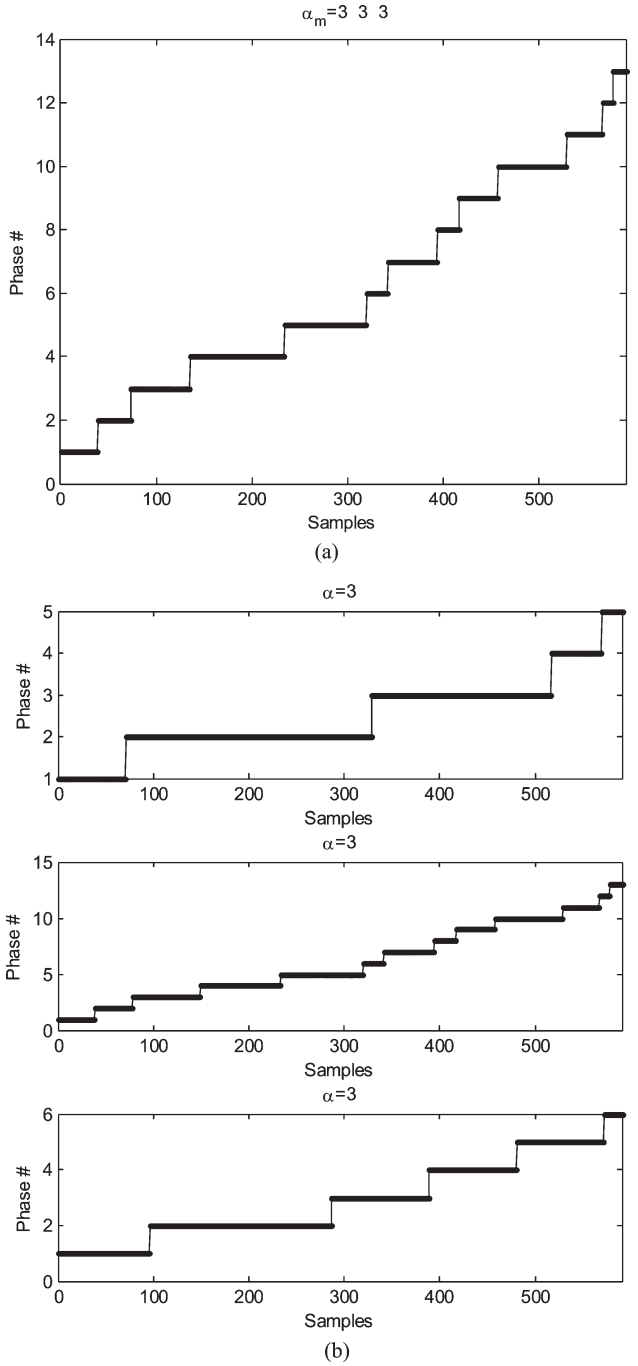

Figure 5. Phase partition results for multimode injection molding process using (a) the proposed concurrent partition algorithm and (b) the mode-specific separate partition algorithm (from top to bottom: Mode 1 to Mode 3).

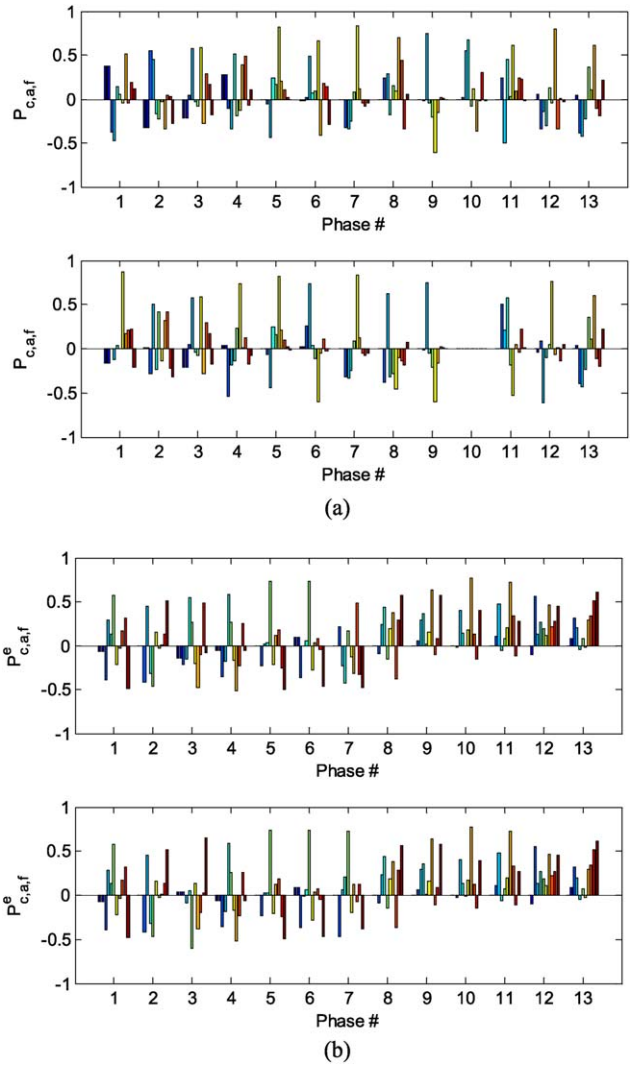

Figure 6. Coefficients of the first principal direction of (a) $\mathbf{P}_{c,a,f}$ and (b) $\mathbf{P}^e_{c,a,f}$ for alternative Mode 1 (top) and alternative Mode 2 (bottom) across different phases.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the current sample can be deemed to be operating according to the normal operation rule of the reference mode. Also, considering that the operation mode will not change abruptly and the mode membership should be consistent with its previous one, the mode affiliation can be readily determined by combining the current monitoring result and previous ones. On the contrary, if alarm signals are issued, the current process may be operating at some fault or some alternative mode status. To distinguish this, the four-model system defined for different alternative modes will be tested in turn to see which alternative mode can best match the current operation pattern

$$\begin{aligned}
\mathbf{t}_{a,f,\mathrm{new}}{}^T &= \mathbf{x}_{\mathrm{new}}{}^T\mathbf{P}_{c,a,f}\\
\mathbf{t}_{a,o,\mathrm{new}}{}^T &= \mathbf{x}_{\mathrm{new}}{}^T\mathbf{P}_{c,a,o}\\
\mathbf{t}^e_{a,f,\mathrm{new}}{}^T &= \mathbf{x}_{\mathrm{new}}{}^T\mathbf{P}^e_{c,a,f}\mathbf{P}^e_{c,a,f}{}^T\\
\mathbf{e}^f_{a,\mathrm{new}}{}^T &= \mathbf{x}_{\mathrm{new}}{}^T\mathbf{P}^e_{c,r}\mathbf{P}^e_{c,r}{}^T - \mathbf{x}_{\mathrm{new}}{}^T\mathbf{P}^e_{c,a,f}\mathbf{P}^e_{c,a,f}{}^T
\end{aligned} \tag{16}$$

$$\begin{aligned}
T_{a,f,\mathrm{new}}{}^2 &= (\mathbf{t}_{a,f,\mathrm{new}} - \bar{\mathbf{t}}_{a,f,k})^{\mathrm{T}}\boldsymbol{\Sigma}_{a,f,k}{}^{-1}(\mathbf{t}_{a,f,\mathrm{new}} - \bar{\mathbf{t}}_{a,f,k})\\
T_{a,o,\mathrm{new}}{}^2 &= (\mathbf{t}_{a,o,\mathrm{new}} - \bar{\mathbf{t}}_{a,o,k})^{\mathrm{T}}\boldsymbol{\Sigma}_{a,o,k}{}^{-1}(\mathbf{t}_{a,o,\mathrm{new}} - \bar{\mathbf{t}}_{a,o,k})\\
T^e_{a,f,\mathrm{new}}{}^2 &= (\mathbf{t}^e_{a,f,\mathrm{new}} - \bar{\mathbf{t}}^e_{a,f,k})^{\mathrm{T}}\boldsymbol{\Sigma}^e_{a,f,k}{}^{-1}(\mathbf{t}^e_{a,f,\mathrm{new}} - \bar{\mathbf{t}}^e_{a,f,k})\\
\mathrm{SPE}^f_{a,\mathrm{new}} &= \mathbf{e}^f_{a,\mathrm{new}}{}^T\mathbf{e}^f_{a,\mathrm{new}}
\end{aligned} \tag{17}$$

Compare the updated monitoring statistics with the predefined confidence limits in each alternative mode. If all new monitoring statistics stay well within the confidence regions of one alternative mode, the current operation mode is judged. On the contrary, if no modes are found to match the new sample, it is possible that one process fault is occurring or it is operating in a new mode which is not included in the library of modeling data. Figure 3 presents the flow diagram of the proposed model development and online monitoring procedure for multimode and multiphase batch processes.

Illustration and Discussion

Injection molding process

Injection molding as illustrated in Figure 4, a key process in polymer processing for manufacturing industry, transforms polymer materials into various shapes and types of products. A typical injection molding process consists of three major operation phases, injection of molten plastic into the mold, packing-holding of the material under pressure, and cooling of the plastic in the mold until the part becomes sufficiently rigid for ejection. Besides, plastication takes place in the barrel in the early cooling phase, where polymer is melted and conveyed to the barrel front by screw rotation, preparing for next cycle. It is a typical multiphase batch process and has been widely used for process monitoring and quality analysis.[23,34] It can be readily implemented for experiments, in which, all key process conditions such as the temperatures,

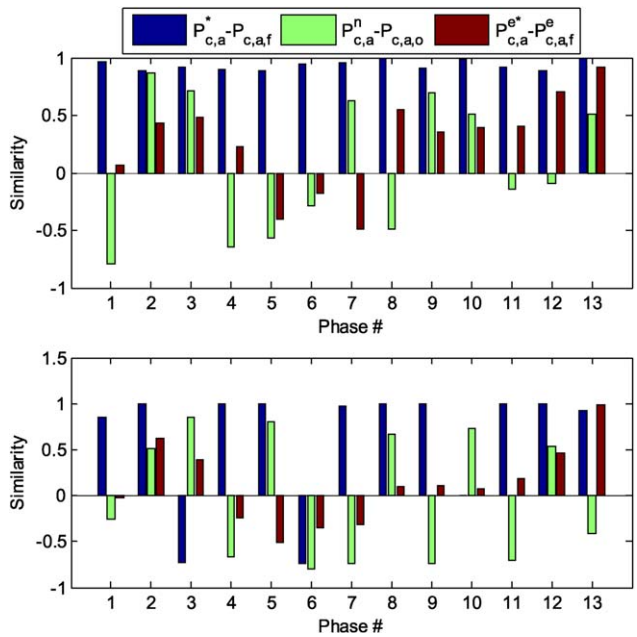

Figure 7. Similarity of the first principal directions for alternative Mode 1 (top) and alternative Mode 2 (bottom).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

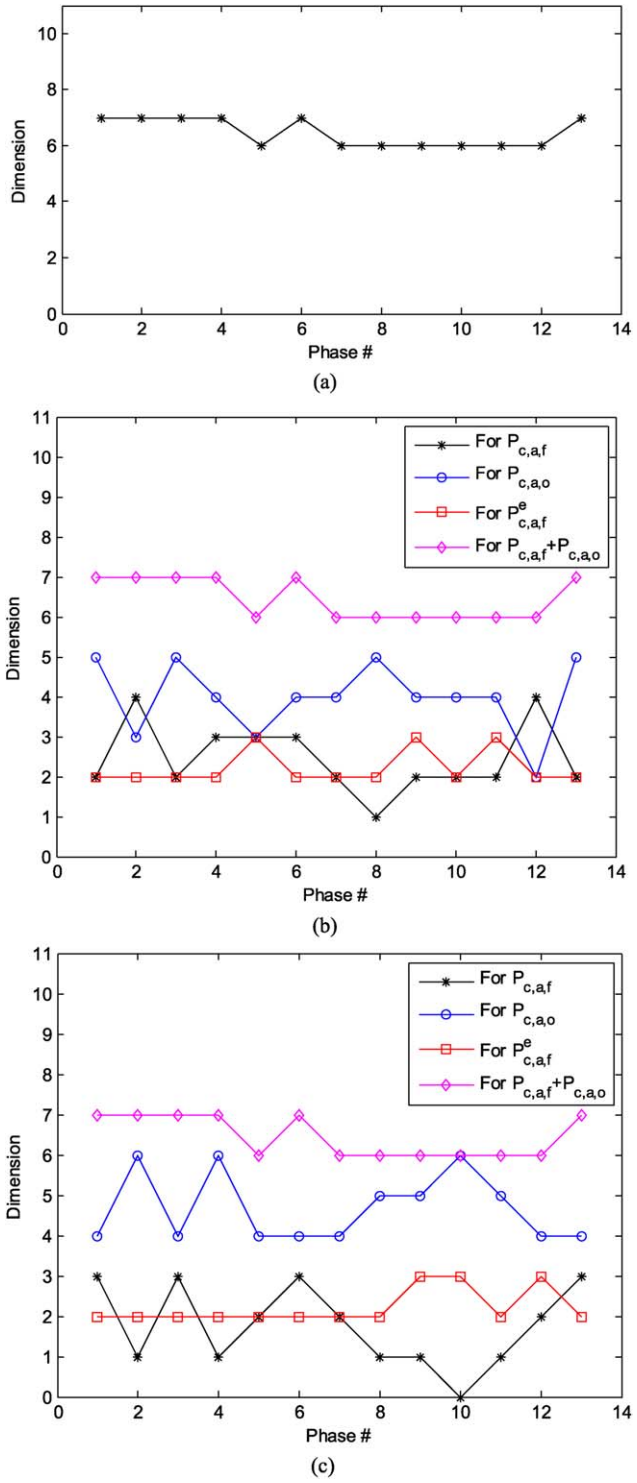

Figure 8. The dimensions of subspaces decomposed using the proposed between-mode analysis algorithm for (a) the reference mode, (b) alternative Mode 1, and (c) alternative Mode 2 in different phases.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

pressures, displacement, and velocity can be online measured by various transducers, providing abundant process information.

The material used in this work is high-density polyethylene. Eleven process variables are selected for modeling as shown in Table 2, which can be collected online from measurements with a set of sensors. Three modes are considered as shown in Table 3 by changing barrel temperature and

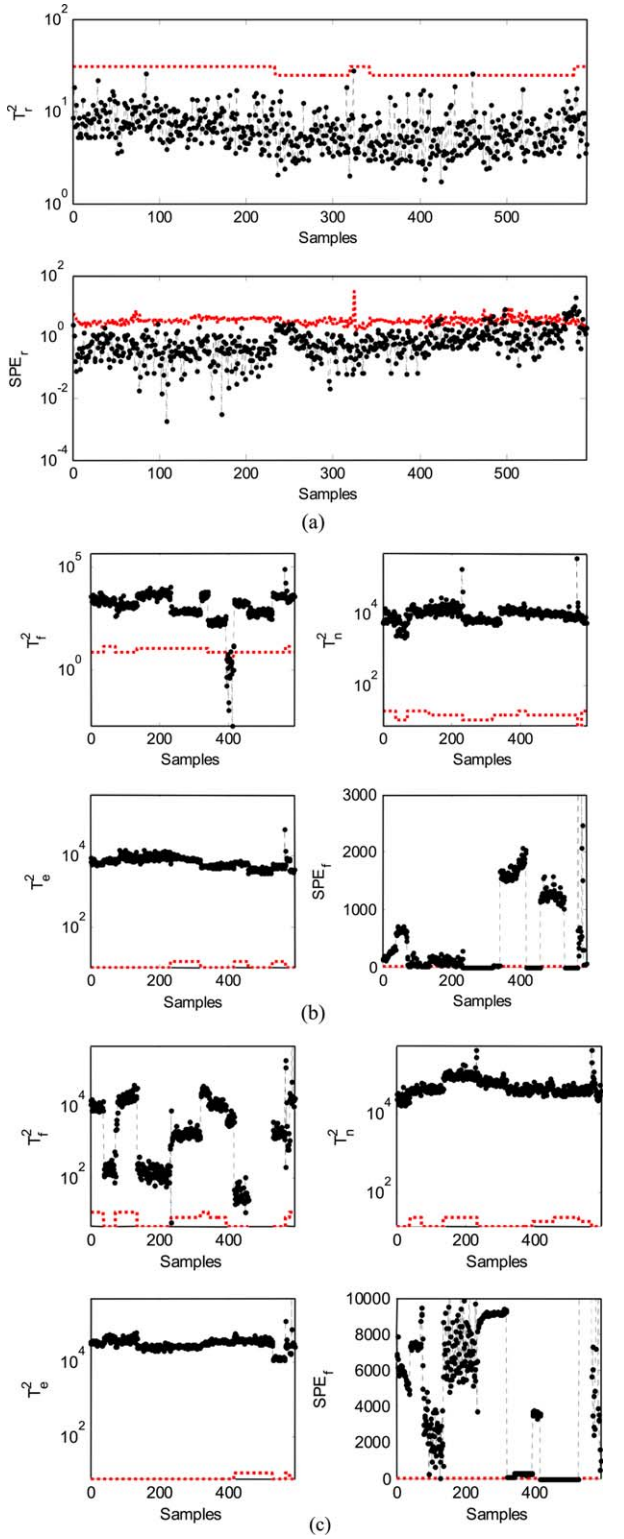

Figure 9. Online monitoring for one testing batch in the reference mode using monitoring models developed from (a) the reference mode, (b) alternative Mode 1, and (c) alternative Mode 2, respectively.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

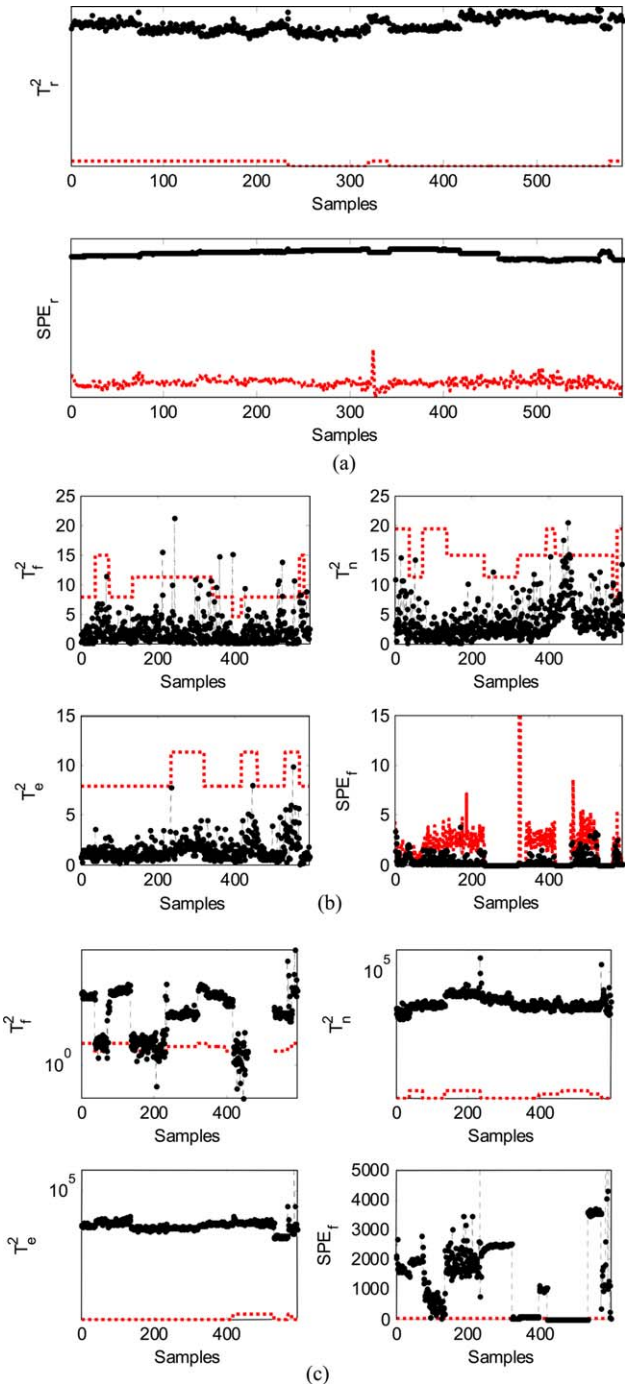

Figure 10. Online monitoring for one testing batch in alternative Mode 1 using monitoring models developed from (a) the reference mode, (b) alternative Mode 1, and (c) alternative Mode 2, respectively.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

packing pressures while keeping the other operation conditions invariable. Twenty-eight batches with even duration (591 samples in this experiment) are collected for each mode, which, thus, result in the descriptor array $\underline{\mathbf{X}}_m(28\times11\times591)(m=1,2,3)$. The first 20 batches are used for model development, whereas the other eight cycles are used for model testing.

Concurrent phase division and between-mode modeling

First, prepare the time-slice data matrix $\mathbf{X}_{m,k}(I\times J)$ at each time from the process beginning of each mode. Using the proposed concurrent phase partition algorithm, three modes are considered simultaneously to check the changes of underlying process characteristics. Figure 5a shows the phase partition results for three-mode injection molding process using the proposed concurrent partition algorithm where the relaxing factors for all modes are set to be 3. In comparison,

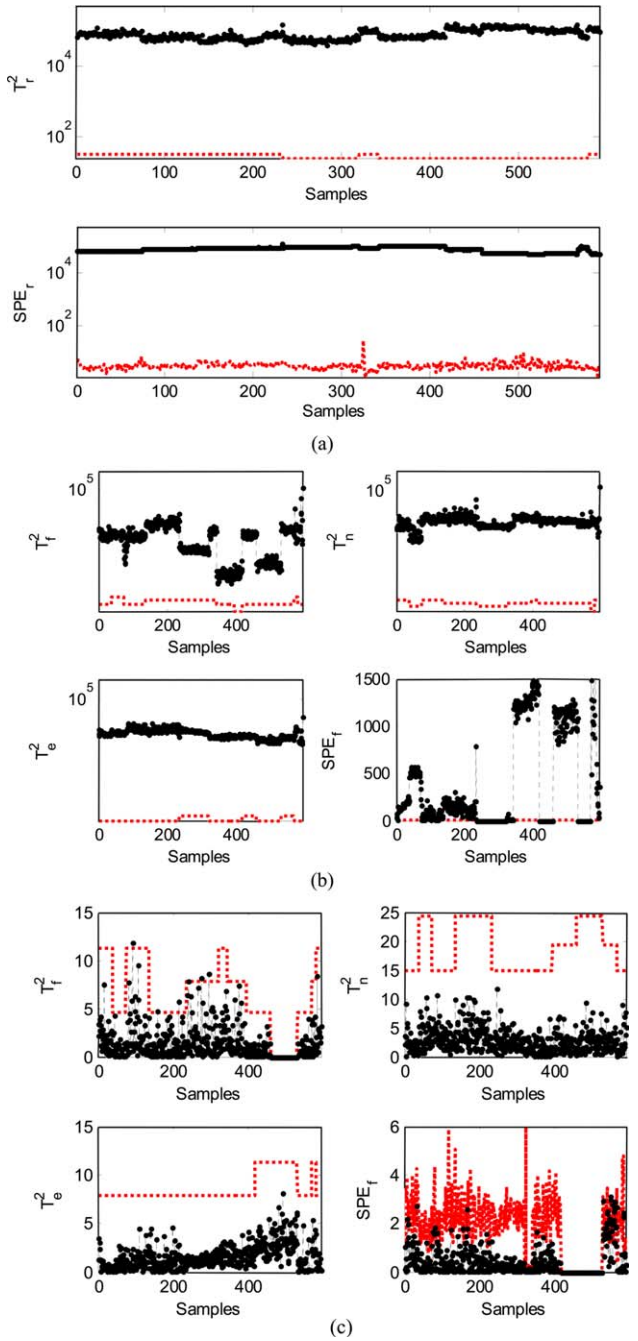

Figure 11. Online monitoring results for one testing batch in alternative Mode 2 using monitoring models developed from (a) the reference mode, (b) alternative Mode 1, and (c) alternative Mode 2, respectively.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 4. Same-Mode and Cross-Mode Online Monitoring Results (Mean ± MAD[a]) for Training Batches and Normal Testing Batches Using the Proposed Algorithm with Mode 1 as Reference Mode

Model / Data	The Reference Mode		Alternative Mode 1				Alternative Mode 2			
	T_r^2	SPE_r	T_f^2	T_o^2	T_e^2	SPE	T_f^2	T_o^2	T_e^2	SPE
Training Batches										
Reference data	0.02 ± 0.03[b]	4.02 ± 4.93	2.47 ± 0.41	0 ± 0	0 ± 0	1.91 ± 0.28	0.10 ± 0.08	0 ± 0	0 ± 0	0 ± 0
Alternative data 1	0 ± 0	0 ± 0	0.85 ± 0.84	0.05 ± 0.07	0.30 ± 0.41	3.16 ± 2.40	5.42 ± 1.67	0 ± 0	0 ± 0	0.18 ± 0.05
Alternative data 2	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.92 ± 0.39	1.23 ± 0.70	0.09 ± 0.16	0.27 ± 0.36	3.13 ± 2.39
Normal Testing Batches										
Reference data	1.71 ± 2.41	9.67 ± 9.46	2.71 ± 0.30	0 ± 0	0 ± 0	1.92 ± 0.38	0.13 ± 0.06	0 ± 0	0 ± 0	0 ± 0
Alternative data 1	0 ± 0	0 ± 0	4.82 ± 1.86	4.46 ± 1.88	1.50 ± 1.87	8.46 ± 5.54	4.59 ± 0.82	0 ± 0	0 ± 0	0.17 ± 0.00
Alternative data 2	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.63 ± 0.25	7.17 ± 2.75	21.64 ± 15.08	43.30 ± 29.29	39.42 ± 16.12

[a]MAD: mean absolute deviation, which is calculated as $\frac{1}{I}\left(\Sigma_{q=1}^{N}\left|Z_q-\frac{\Sigma_{i=1}^{I}Z_i}{I}\right|\right)$ where Z denotes the values of FAR or MAR index for different batches. I denotes the number of training batches or testing batches. Mean is calculated to evaluate the batch-wise "average" monitoring performance for each monitoring statistic; MAD can evaluate the batch-wise variability of monitoring performance for each monitoring statistic.
[b]The shaded values indicate same-mode monitoring results as evaluated by FAR (%) index and the unshaded values indicate the cross-mode monitoring results as evaluated by MAR (%) index.

Table 5. Same-Mode and Cross-Mode Online Monitoring Results (Mean ± MAD) for Training Batches and Normal Testing Batches Using Mode-Specific Separate Modeling Method

Model / Data	Mode 1		Mode 2		Mode 3	
	T^2	SPE	T^2	SPE	T^2	SPE
Training Batches						
Mode 1	0.02 ± 0.03	4.02 ± 4.93	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Mode 2	0 ± 0	0 ± 0	0.01 ± 0.02	3.57 ± 3.44	0 ± 0	0 ± 0
Mode 3	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.01 ± 0.02	3.37 ± 2.50
Normal Testing Batches						
Mode 1	1.71 ± 2.41	9.67 ± 9.46	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Mode 2	0 ± 0	0 ± 0	1.54 ± 0.99	26.92 ± 21.30	0 ± 0	0 ± 0
Mode 3	0 ± 0	0 ± 0	0 ± 0	0 ± 0	5.16 ± 4.99	70.20 ± 24.48

The shaded values indicate same-mode monitoring results as evaluated by FAR (%) index and the unshaded values indicate the cross-mode monitoring results as evaluated by MAR (%) index.

the mode-specific separate partition results are also shown in Figure 5b. It is deemed to be an extreme case for the proposed concurrent phase division algorithm where only one mode is considered. For fair comparison, in each separate mode, the retained dimension of PCA models is unified as the number of PCs that occurs most throughout the batch

Table 6. Same-Mode and Cross-Mode Online Monitoring Results (Mean ± MAD[a]) for Training Batches and Normal Testing Batches Using the Proposed Algorithm with Mode 2 as Reference Mode

Model / Data	The Reference Mode		Alternative Mode 1				Alternative Mode 2			
	T_r^2	SPE_r	T_f^2	T_o^2	T_e^2	SPE	T_f^2	T_o^2	T_e^2	SPE
Training Batches										
Reference data	0.01 ± 0.02[b]	3.57 ± 3.44	0 ± 0	0 ± 0	0.08 ± 0.09	1.81 ± 0.16	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Alternative data 1	0 ± 0	0 ± 0	1.08 ± 0.67	0.14 ± 0.20	0.09 ± 0.15	3.21 ± 3.03	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Alternative data 2	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.02 ± 0.14	0.47 ± 0.44	0.42 ± 0.46	0.45 ± 0.51	2.38 ± 2.96
Normal Testing Batches										
Reference data	1.54 ± 0.99	26.92 ± 21.30	0 ± 0	0 ± 0	0.13 ± 0.13	1.71 ± 0.32	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Alternative data 1	0 ± 0	0 ± 0	3.26 ± 2.89	4.65 ± 6.51	1.82 ± 1.72	8.16 ± 5.04	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Alternative data 2	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.08 ± 0.22	18.85 ± 11.73	23.03 ± 19.65	27.43 ± 21.68	61.44 ± 28.19

[a]MAD: mean absolute deviation, which is calculated as $\frac{1}{I}\left(\Sigma_{q=1}^{N}\left|Z_q-\frac{\Sigma_{i=1}^{I}Z_i}{I}\right|\right)$ where Z denotes the values of FAR or MAR index for different batches. I denotes the number of training batches or testing batches. Mean is calculated to evaluate the batch-wise "average" monitoring performance for each monitoring statistic; MAD can evaluate the batch-wise variability of monitoring performance for each monitoring statistic.
[b]The shaded values indicate same-mode monitoring results as evaluated by FAR (%) index and the unshaded values indicate the cross-mode monitoring results as evaluated by MAR (%) index.

Table 7. Comparison of Online Fault Detection Performance (FDT) (Mean ± MAD) in Same-Mode Analysis for Artificial Faults Using the Proposed Between-Mode Modeling Algorithm (BM) with Mode 1 as Reference Mode and Mode-Specific Separate Modeling Method (MS)

Methods / AF #	Alternative Mode 1		Alternative Mode 2	
	BM	MS	BM	MS
AF 1	84.05 ± 14.05	119.40 ± 44.04	55.45 ± 13.69	125.35 ± 87.72
AF 2	147.45 ± 48.12	173.65 ± 69.15	188.55 ± 61.79	268.25 ± 40.50
AF 3	170.00 ± 122.30	306.50 ± 92.00	197.65 ± 145.68	252.20 ± 124.74

process. From the results, it is clear that using the proposed algorithm, the operation processes for all modes can be automatically partitioned into different subphases in time order with consistent phase landmarks across modes. For comparison, mode-specific separate phase division shows different phase landmarks for each mode which makes it inconvenient for the following phase-based between-mode relative analysis. It is noted that those small time segments between long phases are in fact transition patterns, which are separated from steady phases resulting from faster changes of process characteristics. For simplicity, the phase division results using clustering algorithm shown in Appendix are not shown here, where the affiliation of different time intervals shows quite frequent and chaotic changes. Also, different time regions are mixed together frequently, requiring a heavy postprocessing. The concurrent phase division algorithm jointly considers different time-varying characteristics for all modes so that a unified division result is available for all modes. In this case, clearly, Mode 2 presents more dynamic characteristics, whereas the other two modes have more stable characteristics when the same value of relaxing factor is used. By the proposed algorithm, the original phases in Mode 1 and Mode 3 shown in Figure 5b can be further divided into different time segments to be consistent with the phase information in Mode 2.

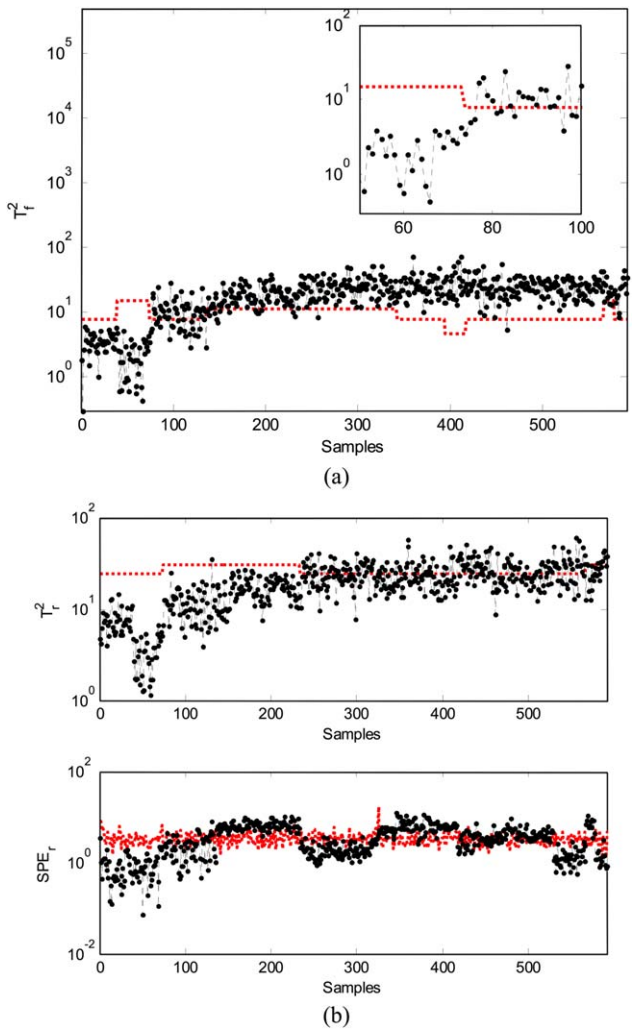

Figure 12. Online fault detection results for AF 1 generated from one batch in alternative Mode 1 using (a) the proposed algorithm and (b) the mode-specific separate modeling method (dashed line, 95% control limits; dot line, the monitoring statistics).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Based on the phase division result, phase-based between-mode analysis is performed. Mode 1 is chosen as the reference mode from which PCA monitoring system is developed. Then, variations in two alternative modes are decomposed, respectively, relative to the reference mode. Three systematic subspaces, $\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}^e_{c,a,f}$, are decomposed from PCS and RS of PCA monitoring system. Figure 6 shows the coefficients of the first principal direction of $\mathbf{P}_{c,a,f}$ and $\mathbf{P}^e_{c,a,f}$ for two alternative modes across different phases. Comparatively, in different phases, the dominative variables as indicated by larger coefficient values are different for each mode. Also, they are different between the two alternative modes in the same phase. To demonstrate the use of PCA in Step (4) of the proposed algorithm, as shown in Figure 7, similarity of the first principal directions in each subspace is evaluated for each mode by correlation analysis. For some phases, the similarity is quite low, such as that for $\mathbf{P}^e_{c,a}{}^* - \mathbf{P}^e_{c,a,f}$ in Phase 1 for both alternative modes. It demonstrates that PCA can integrate the concerned between-mode relative variations which have been separated from the original PCA monitoring system. Figure 8 shows the dimensions of different systematic subspaces decomposed using the proposed between-mode analysis algorithm for the reference mode and each alternative mode in different phases. In general, for both alternative modes, the dimension of $\mathbf{P}_{c,a,o}$ is larger than that of $\mathbf{P}_{c,a,f}$. It is noted that no $\mathbf{P}_{c,a,f}$ subspace is separated in Phase 10 for alternative Mode 2.

CPPBM-based online monitoring

Based on the developed models, online monitoring is implemented to check the performance of the proposed algorithm. First, to check the online mode identification ability

Table 8. Comparison of Online Fault Detection Performance (FDT) (Mean ± MAD) in Same-Mode Analysis for Artificial Faults Using the Proposed Between-Mode Modeling Algorithm (BM) with Mode 2 as Reference Mode and Mode-Specific Separate Modeling Method (MS)

Methods / AF #	Alternative Mode 1		Alternative Mode 2	
	BM	MS	BM	MS
AF 1	64.15 ± 16.44	104.80 ± 41.52	173.35 ± 53.02	183.25 ± 89.55
AF 2	188.50 ± 34.15	235.30 ± 15.94	98.60.30 ± 40.22	171.25 ± 128.68
AF 3	39.70 ± 4.77	191.05 ± 53.24	70.15 ± 17.30	88.80 ± 39.62

of the proposed algorithm, two types of monitoring actions are taken:

a. Same-mode monitoring where monitoring models are developed based on between-mode analysis of the reference mode and one alternative mode, and then used to monitor the same modes.

b. Cross-mode monitoring where monitoring models are developed based on between-mode analysis of the reference mode and one alternative mode, and then used to monitor other modes.

In both types of monitoring actions, normal batches are used. For accurate mode identification, it is expected to issue no alarm signals in same-mode monitoring. In comparison, as the wrong models are used to monitor the operation modes in cross-mode analysis, it is expected to issue alarm signals as much as possible even for the normal status. Therefore, false alarm rate (FAR) can be used to evaluate the performance of mode identification in same-mode analysis and missing alarm rate (MAR) can be used in cross-mode analysis. Figure 9 shows the online monitoring results of one normal testing batch in the reference mode using monitoring models developed for the reference mode, alternative Mode 1 and alternative Mode 2, respectively. Clearly, the monitoring statistics stay well below the confidence limits when the right models (reference models) are used as shown in Figure 9a. In comparison, when wrong models are used which are developed for alternative modes, the monitoring statistics are clearly out-of-control as shown in Figures 9b,c. In this way, the operation mode is identified correctly as only models developed for the reference mode can match the current batch. Figures 10 and 11 show the online monitoring of one normal testing batch in alternative Mode 1 and alternative Mode 2, respectively, using monitoring models developed from the reference mode, alternative Mode 1 and alternative Mode 2. For same-mode analysis, the online monitoring statistics should stay well within the confidence regions, whereas for cross-mode analysis, the online monitoring statistics should go beyond the confidence limits, which well agree with the monitoring results in the two figures.

In Tables 4 and 5, the same-mode and cross-mode online monitoring results are summarized for both training batches and normal testing batches using the proposed algorithm in comparison with the mode-specific separate modeling algorithm. FAR is calculated to evaluate the monitoring performance in same-mode analysis as indicated by the shadowed values. MAR is used to evaluate the monitoring performance in cross-mode analysis where the wrong monitoring models are used to monitor normal batches. In general, the two methods show comparable monitoring performance, that is, for normal batches, the two methods can identify the current operation mode with similar accuracy.

It is noted that the results in Table 4 are obtained with Mode 1 as reference mode. To reveal the effects of the choice of reference mode, Mode 2 is chosen as new reference mode and CPPBM-based online monitoring is performed again. The same-mode and cross-mode online monitoring results are summarized in Table 6 for both training batches and normal testing batches with FAR and MAR as evaluation indices. For training data, the results are similar with those shown in Table 4. For testing data, most of the results also show similar accuracy with those with Mode 1 as reference mode. In general, it is hard to say which mode is better to serve as reference mode from the results.

Fault detection for artificial faults

Here, three artificial faults (AFs) are generated to illustrate the superiority of the proposed algorithm, where the disturbances are imposed in each of three systematic subspaces ($\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}^{e}_{c,a,f}$). For simplicity, only an exponential curve ($\mathbf{f}(K\times 1)$) is considered and imposed on each batch. The exponential curve is defined by $K\left(1-e^{-t/\tau}\right)$ where time constant $\tau = 100$ and magnitude $K = 5$. To get one AF which influences one specific monitoring subspace, the imposed disturbance is programmed as

$$\mathbf{x}_i^{f\mathrm{T}} = \mathbf{x}_i^{\mathrm{T}} + \Delta_k^{\mathrm{T}} = \mathbf{x}_i^{\mathrm{T}} + \boldsymbol{\theta}_{c,1}^{\mathrm{T}} f_k \tag{18}$$

where, $\mathbf{x}_i\left(J_{x,i}\times 1\right)$ represents the original measurement data vector at each time under the normal situation; $\Delta_k(J\times 1)$ represents the imposed disturbance; f_k denotes the kth value of exponential curve. $\boldsymbol{\theta}_{c,1}$ is the first direction in one systematic subspace which can be $\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, or $\mathbf{P}^{e}_{c,a,f}$ obtained using the proposed algorithm. So, three AFs can be obtained based on the choice of $\boldsymbol{\theta}_{c,1}$. As three systematic subspaces ($\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}^{e}_{c,a,f}$) are orthogonal with each other, therefore, the disturbance imposed on one subspace will not influence the systematic variability in the other subspaces.

Table 7 summarizes the comparison of online fault detection performance as evaluated by first detection time (FDT) in same-mode monitoring for AFs using the proposed between-mode modeling algorithm (BM) and mode-specific separate modeling method (MS) which are both developed based on the concurrent phase division results. So, the comparison can focus on revealing the role of between-mode analysis. It is noted that Mode 1 is used as reference mode for the extraction of different between-mode subspaces and thus generation of artificial disturbances. FDT indicates the first time when the fault is detected as indicated by three consecutive out-of-control monitoring statistics. For AFs 1 through 3, the disturbances are imposed in the systematic subspaces obtained by BM algorithm, $\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}^{e}_{c,a,f}$, respectively. Clearly, the proposed BM algorithm shows superiority over MS method where FDT values by BM

algorithm are smaller than those by MS algorithm for all AFs, revealing faster fault detection. Figure 12 presents online fault detection of AF 1 generated from one batch in alternative Mode 1 using the proposed algorithm in comparison with MS method. For simplicity, only the monitoring results by $\mathbf{P}_{c,a,f}$ are shown for the proposed algorithm. From the zoomed plot, the proposed algorithm clearly indicates the occurrence of abnormality around the 80th sampling time while MS method shows a great time delay.

In comparison, Mode 2 is chosen as new reference mode to extract between-mode subspaces ($\mathbf{P}_{c,a,f}$, $\mathbf{P}_{c,a,o}$, and $\mathbf{P}^{e}_{c,a,f}$) and then generate three AFs. With the new reference mode, new systematic subspaces are extracted to reveal between-mode relative changes. Then, the artificial disturbances will be different from those in Table 7 using the generation way shown in Eq. 18. As shown in Table 8, BM and MS-based monitoring results are summarized and compared. Still, the proposed BM algorithm shows superiority over MS method where FDT values by BM algorithm are smaller than those by MS algorithm for all AFs, revealing faster fault detection.

Conclusions

In this article, a concurrent phase partition and between-mode analysis algorithm (CPPBM) is proposed for multimode and multiphase batch process monitoring. By the automatic concurrent phase partition algorithm, the phase landmarks are identified simultaneously for all modes in a consistent way, which reflects the time sequence of operation phases within each batch and requires no postprocessing. By further decomposition of different subspaces from the between-mode viewpoint, the relative changes from one mode to another can be clearly revealed and captured for online monitoring, which also provides enhanced process understanding. The case study demonstrates the superior performance of the proposed algorithm in fault detection as more comprehensive fault detection subspaces are available.

Acknowledgment

This work is supported by Program for New Century Excellent Talents in University (NCET-12-0492), Zhejiang Provincial Natural Science Foundation of China (LR13F030001), Project of Education Department of Zhejiang Province (Y201223159), and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT1320).

Literature Cited

1. Kourti T, MacGregor JF. Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemom 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. *Chemom 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 AK. 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 Chemom.* 2003;17:93–109.
8. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemom Intell Lab Syst.* 1987;2:37–52.
9. Burnham AJ, Viveros R, MacGregor JF. Frameworks for latent variable multivariate regression. *J Chemom.* 1996;10:31–45.
10. Dayal BS, MacGregor JF. Improved PLS algorithms. *J Chemom.* 1997;11:73–85.
11. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemom Intell Lab Syst.* 1987;2:37–52.
12. Nomikos P, MacGregor JF. Monitoring of batch processes using multiway principal component analysis. *AIChE J.* 1984;40:1361–1375.
13. Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. *Technometrics.* 1995;37:41–59.
14. Nomikos P, MacGregor JF. Multiway partial least squares in monitoring batch processes. *Chemom Intell Lab Syst.* 1995;30:97–108.
15. Wise BM. A comparison of multiway principal components analysis, tri-linear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. *ICRM98, International Chemometrics Research Meeting. Veldhoven*, 1998.
16. Smilde AK. Comments on three-way analyses used for batch process data. *J Chemom.* 2001;15:19–27.
17. Smilde AK. Three-way analysis. *Problems and prospects. Chemom Intell Lab Syst.* 1992;15:143–157.
18. Ündey C, Ertunc S, Cinar A. Online batch/fed-batch process performance monitoring, quality prediction, and variable-contribution analysis for diagnosis. *Ind Eng Chem Res.* 2003;42:4645–4658.
19. Kosanovich KA, Piovos MJ, Dahl KS. Multi-way PCA applied to an industrial batch process. Process of American Control Conference, Baltimore, MD. 1994;1294–1298.
20. Dong D, McAvoy TJ. Multi-stage batch process monitoring. Process of American Control Conference, Seattle, WA. 1995;1857–1861.
21. Zhao CH, Wang FL, Lu NY, Jia MX. Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes. *J Process Control.* 2007;17:728–741.
22. Zhao CH, Wang FL, Mao ZZ, Lu NY, Jia MX. Improved batch process monitoring and quality prediction based on multiphase statistical analysis. *Ind Eng Chem Res.* 2008:47;835–849.
23. Zhao CH, Wang FL, Gao FR, Lu NY, Jia MX. Adaptive monitoring method for batch processes based on phase dissimilarity updating with limited modeling data. *Ind Eng Chem Res.* 2007;46:4943–4953.
24. Zhao CH, Wang FL, Gao FR. Improved calibration investigation using phase-wise local and cumulative quality interpretation and prediction. *Chemom Intell Lab Syst.* 2009;95:107–121.
25. Zhao CH, Sun YX. Step-wise sequential phase partition (SSPP) algorithm based statistical modeling and online process monitoring, *Chemom Intell Lab Syst.* 2013;125:109–120.
26. Yao Y, Gao FR. Phase and transition based batch process modeling and online monitoring. *J Process Control.* 2009;19:816–826.
27. Yao Y, Gao FR. A survey on multistage/multiphase statistical modeling methods for batch processes. *Annu Rev Control.* 2009;33:172–183.
28. Lee YH, Jin HD, Han CH. On-line process state classification for adaptive monitoring. *Ind Eng Chem Res.* 2006;45(9):3095–3107.
29. Zhao SJ, Zhang J, Xu YM. Monitoring of processes with multiple operating modes through multiple principal component analysis models. *Ind Eng Chem Res.* 2004;43(22):7025–7035.
30. Yoo CK, Villez K, Lee I, Rosen C, Vanrolleghem PA. Multi-model statistical process monitoring and diagnosis of a sequencing batch reactor. *Biotechnol Bioeng.* 2007;96(4):687–701.
31. Kassidas A, MacGregor JF, Taylor PA. Synchronization of batch trajectories using dynamic time warping. *AIChE J.* 1998;44:864–875.
32. Lowry CA, Montgomery DC. A review of multivariate control charts. *IIE Trans.* 1995;27(6):800–810.
33. Jackson JE. *A User's Guide to Principal Components*. New York: Wiley, 1991.
34. Zhao CH, Wang FL, Mao ZZ, Lu NY, Jia MX. Quality prediction based on phase-specific average trajectory for batch processes. *AIChE J.* 2008;54:693–705.

Appendix
Clustering-Based Phase Division Algorithm

Inputs: the patterns to be partitioned, $\{\breve{\mathbf{P}}_1,\breve{\mathbf{P}}_2,\ldots,\breve{\mathbf{P}}_k\}$, and the threshold θ for cluster elimination, the minimal phase length, $L_{\min}$.

Outputs: the number of clusters C, the cluster centers $\{\mathbf{W}_1, \mathbf{W}_2, ..., \mathbf{W}_C\}$, and the strict membership of $\breve{\mathbf{P}}_k$ to C centers, $m(k)$.

The index variables are the iteration index i, and the pattern index, k.

1. Choose $C^0(i=0)$ cluster centers $\mathbf{W}_c^0(c=1,2,\ldots,\mathrm{C}^0)$ from the K patterns along the time direction. Practically, the initial cluster centers can be assumed to be uniformly distributed in the pattern set.
2. Merge pairs of clusters whose intercenter distance, dist $(\mathbf{W}_{c1}^{i-1}, \mathbf{W}_{c2}^{i-1})$, is below the predetermined threshold θ.
3. Calculate the distances from each pattern $\breve{\mathbf{P}}_k$ to all of the centers, dist $(\breve{\mathbf{P}}_k, \mathbf{W}_c^{i-1})$, assign $\breve{\mathbf{P}}_k$ to the nearest center $\mathbf{W}_{c^*}^{i-1}$, and denote its membership as $m(k)=c^*$.
4. Eliminate the clusters that catch few patterns (fewer than the minimal phase length, $L_{\min}$) after a set number of iterations to avoid singular clusters.
5. Update the number of cluster centers to be C^i, recompute the new cluster centers $\mathbf{W}_c^i(c=1,2,\ldots,C^i)$, using the current cluster membership, $m(k)$.
6. Go back to Step 2 if a convergence criterion is not met. Typical convergence criteria are minimal changes in the cluster centers and/or minimal rate of decrease in squared errors.

Manuscript received Mar. 24, 2013, and revision received Oct. 18, 2013.