# Sequential Time Slice Alignment Based Unequal-Length Phase Identification and Modeling for Fault Detection of Irregular Batches

**ARTICLE** *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · SEPTEMBER 2015

Impact Factor: 2.59 · DOI: 10.1021/acs.iecr.5b01405



Chunhui Zhao

Zhejiang University



Furong Gao

The Hong Kong University of Science and Tec…

I&EC research
Industrial & Engineering Chemistry Research
Article
pubs.acs.org/IECR

# Sequential Time Slice Alignment Based Unequal-Length Phase Identification and Modeling for Fault Detection of Irregular Batches

Wenqing Li,[†] Chunhui Zhao,[*,†,‡] and Furong Gao[§]

[†]State Key Laboratory of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China

[‡]Key Laboratory of Advanced Control and Optimization for Chemical Processes, Shanghai, 200237 China

[§]Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR

**ABSTRACT:** In practice, batch processes may have different durations, revealing unsynchronized operation events, due to the changes in operation conditions or control objectives. For batches with multiple operation phases, batch-wise process characteristics are irregular, and at the same time, phases are misaligned over batches, which have caused problems in phase analysis and modeling as well as online process monitoring. To solve the uneven-length problem in multiphase batch processes, this paper proposes a sequential time slice alignment based unequal-length phase identification and modeling method for fault detection of irregular batches. In comparison with previous work, the major contribution of the proposed method is as follows: (1) The irregular process characteristics are evaluated in sequence and directly related with the monitoring performance. (2) Multiple irregular phases are readily identified and modeled using sequential time slice alignment which avoids cumbersome postprocessing. (3) The sequential nature provides an easy way to real-time judge the phase affiliation of each new sample for online fault detection. Also, comparison is conducted between the proposed algorithm and clustering-based uneven-length phase division and process monitoring algorithm. The application to a typical batch process with varying durations illustrates the online monitoring performance of the proposed method.

## 1. INTRODUCTION

Batch processes play a significant role in modern industrial manufacturing such as specialty polymers, pharmaceuticals, and biochemicals. The work of monitoring these batch processes is of great importance to ensure safe production and consistent high-quality products. However, due to the high variable dimensionality and complexity characteristics of batch processes, it is difficult to create first-principles based models. Multivariate statistical analysis methods,[1−5] which require only process history data, thus have attracted rising attentions. Among them, multiway principle analysis (MPCA) and multiway partial least regression (MPLS), which were proposed by Nomikos and Macgregor,[3,4] extend the applications of traditional multivariate statistical analysis methods from continuous processes to batch processes. Since then, batch processes have been comprehensively studied[6−10] by statistical analysis and modeling methods.

However, most researches[6−10] are based on the idealized assumption that all batches have the same duration and batch trajectories are synchronized. In practice, the above assumption cannot be well satisfied as a result of batch-to-batch quality variations of raw materials, changes in operation conditions and different control objectives. What's more, for multiphase batch processes, the uneven length problem becomes more complex that it may occur within each phase and the resulting irregular phase data cannot be directly used for statistical analysis and modeling. Since uneven batch length problem commonly arises in batch processes, which brings great difficulty in process modeling and monitoring, many studies have been conducted to handle it.[11−28] Previous work can be divided into two types: direct signal synchronization methods[11−25] and irregular phase partition based modeling methods.[26−28]

Direct signal synchronization means batch trajectories are synchronized by performing some signal processing strategies. Within this category, batches were first synchronized using simple ideas, such as cutting the trajectories of all batches to the shortest one[12] or expanding the trajectories by missing-data-estimation methods.[12] Unfortunately, these methods were only suitable for the situation that batch trajectories overlapped in the common part. Besides, the missing-data-estimation methods needed enough long batches for identifying model to estimate missing data, which is sometimes impractical in industrial manufacturing. Jian et al.[19] recently proposed an improved missing-data-estimation method for handling data synchronization issue, overcoming the limitation that sufficient long batches were required. However, this method could not be applied to multiphase batch processes because a single model was incapable of describing the multiphase characteristics. Nomikos and MacGregor proposed an alternative solution[13] for batch alignment, in which, time index was replaced by an indicator variable (IV) which had the same starting and ending values and progresses monotonically for each batch. Then according to the indicator variable, process trajectories could be rescaled to the same length by resampling and interpolation methods. IV was a simple and useful method when a

**Received:** April 14, 2015
**Revised:** September 21, 2015
**Accepted:** September 27, 2015
**Published:** September 27, 2015




DOI: 10.1021/acs.iecr.5b01405
*Ind. Eng. Chem. Res.* 2015, 54, 10020−10030

monotonous variable existed in a batch, but in general case, process maturity might not be reflected by such a variable in particular for multiphase batch processes. What's more, during the resampling or interpolation procedure, the original process variable correlations might be distorted. Since batch process data was a kind of temporal signal, methods based on warping techniques[20−25] were also introduced to aligning the batches with uneven-length trajectories, such as dynamic time warping (DTW)[20] and its improved methods,[21−25] which synchronized batches by translating, expanding, and compressing local segments of trajectories to decrease their dissimilarities. These methods effectively synchronized uneven batch profiles, however, might also distort process correlations during the synchronization procedure. Moreover, multiphase characteristics was ignored except for the methods[24,25] based on dynamic locus analysis (DLA), which first identified each time segment in a batch via reference landmarks before synchronizing each segment. In addition, correlation optimizing warping (COW)[23] was another major approach for synchronizing uneven batches, which behaved similarly with DTW. In general, these techniques effectively synchronized the total batch profiles or each phase of batches and allowed for the subsequently process modeling strategies, however, process correlations might be distorted during the synchronization procedure which would jeopardize online application performance.

To overcome the limitations of direct signal synchronization methods and considering the multiphase nature of batch processes, Lu et al.[26] proposed the subphase division method for handling uneven-length problem of multiphase batch processes, where there is no need of aligning the uneven data to the same length and process correlations would not be distorted. In this method, irregular phases were divided by clustering based algorithm, and monitoring models were then developed based on the phase partition results. While Lu's method might be improper when the uneven problem is serious, Zhao et al.[27] distinguished and analyzed two cases of uneven-length problems: moderate uneven case and serious uneven case. For serious uneven case, irregular batches have different process characteristics even in the same phase, which thus cannot be described by one specific phase model. They proposed a group division and subspace separation based method to model the different characteristics of irregular batches. Batches with the dissimilar characteristics were divided into the different group, and then group-common and group-specific information were modeled respectively by proper subspace separation method for developing high-resolution models and close confidence regions. While the clustering-based methods[26−28] have been demonstrated effective and applied successfully, there might exist some drawbacks during the irregular phase partition. First, the clustering algorithm did not consider the time sequence of process operation and might lead to disordered phase partition results where samples from different time regions would be mixed together. Moreover, distance index was used to evaluate process characteristics, which was not directly related to the monitoring performance. It might result in improper phase models, which cannot well describe and supervise the distance-based similar time-slices. In addition, these phase partition methods took the whole time-slice as the basic analysis unit, which thus might not correctly reveal the unsynchronized characteristics of some specific batches in some time-slices, in particular for phase shift. For example, in some time-slices, only one batch enters the next phase while the others still lie in the current phase. Since its impact on the whole time-slice is covered by other batches, it will be classified into the same phase with other batches in the same time-slice according to the analysis of the whole time-slice. Thus, the phase shift of this specific batch cannot be correctly judged.

For multiphase batch processes with uneven-length phases, different operation patterns may be observed at the same time interval. Thus, process characteristics cannot be well reflected if samples with significantly different characteristics are described by a unified model. So it is necessary to distinguish irregular phases with different operation patterns before implementing statistical modeling. In order to automatically identify each irregular phase in sequence along time direction, the present work proposes a sequential time slice alignment based unequal-length phase identification and modeling method for fault detection of irregular batches, which extends the stepwise sequential phase partition (SSPP) algorithm[10] to the uneven batch case. However, the basic idea and specific implementation procedure of the proposed method are almost different from the SSPP algorithm. For readability, the SSPP algorithm is presented in the Appendix. The major contribution of the method proposed here is specified as follows: (1) The irregular process characteristics are evaluated in sequence directly related to the monitoring performance, where time slices with similar process characteristics are sequentially collected to develop an iteratively updated model for checking the changes of variable correlations of each batch regarding their influences on model accuracy and monitoring performance. (2) Based on the process characteristics evaluation results, multiple irregular phases are automatically identified in order without any postprocessing and then irregular phase models can be designed by sequential time slices alignment to explain the similar local process behaviors. (3) The sequential nature provides an easy and effective way to real time locate new samples to their specific phases and check their operation statuses, in which procedure two regions are distinguished according to the irregular phase landmarks and then different online monitoring strategies can be implemented for them. One is the within-phase region, where samples belong to the current phase and their operation statuses can be checked by the current phase model. Another termed as the between-phase shifting region refers to the samples located at the border of the adjacent phases, in which the phase affiliation of samples is uncertain and the operation statuses of samples should be checked by adopting proper phase models. In general, during the procedure of irregular phase division, the process characteristics of each batch are analyzed in detail. Irregular time-varying underlying process characteristics are effectively tracked and different underlying patterns are clearly separated. Samples with similar underlying characteristics are collected within the same phase so that a high-resolution unified model can be developed. The feasibility and superior performance of the proposed algorithm are illustrated by a typical multiphase batch process with uneven batch durations.

## 2. METHODOLOGY

**2.1. Data Preparation.** Since PCA will be performed to model the process, data should be first normalized. In general, there are two common approaches for normalizing the three-dimensional data of batch processes: variable-unfolding based method and batch-unfolding based method. Assume that $\mathbf{X}(I \times K \times J)$ is a three-dimension batch data, where $I$ is the number of batches, $J$ is the number of variables, and $K$ represents the


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

number of samples. For variable-unfolding based method, variable dimension is kept unchanged and data is arranged from top to bottom in columns, forming a $(KI \times J)$-dimensional matrix. It can be seen that batch-wise and time-wise data is merged into one dimension, so that variations along time direction are mainly focused on by this method. While for batch-unfolding based method, batch dimension is retained and time-slices are placed side by side from left to right, giving a $(I \times KJ)$-dimensional matrix. Average trajectory is calculated over batches and batch-wise variations along time direction is revealed after normalization. Since variations along time direction are larger and more complicated than that of batch-wise, models developed based on variable-unfolding normalization method will cover more variations simultaneously, and thus the resulting confidence region can be very wide, which makes the model insensitive to small variations. In contrast, batch-unfolding based method which focuses on batch-wise variations, is suitable for monitoring system development. However, it cannot be directly applied to multiphase batch processes with varying durations for the misaligned operation events over batches. In order to divide the irregular phases and develop a reliable monitoring system, the advantages of the two methods can be combined and utilized in the proposed algorithm. Variable-unfolding based method is only chosen for irregular phases partition, and batch-unfolding based approach will be adopted later for monitoring system development.

In each batch run (batch index $i = 1, 2, ..., I$), assume that $J$ process variables are measured online at $k = 1, 2, ..., K$ time intervals throughout the operation cycle where the duration is not fixed in length, forming each batch set, denoted as $\mathbf{X}_i^*(K_i \times J)$. As is shown in Figure 1, almost each batch $\mathbf{X}_i^*(K_i \times J)$ has

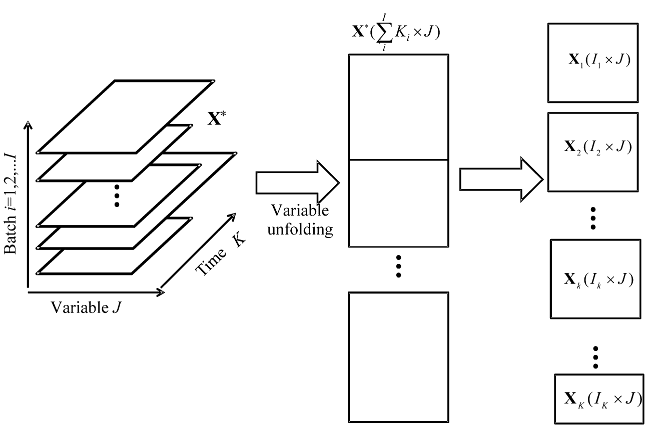



**Figure 1.** Illustration of irregular data arrangement.

different time duration. Then as mentioned above, we arrange batch set $\mathbf{X}_i^*(K_i \times J)$ from top to bottom and keep the dimension of variable unchanged, giving a $\sum_{i=1}^{I} K_i \times J$ matrix $\mathbf{X}^*(\sum_{i=1}^{I} K_i \times J)$. Subsequently, $\mathbf{X}^*(\sum_{i=1}^{I} K_i \times J)$ is normalized by subtracting the mean and dividing by its deviation. Then time-slices $\mathbf{X}_k(I_k \times J)(k = 1,2,...,K)$ can be separated from $\mathbf{X}^*(\sum_{i=1}^{I} K_i \times J)$.

**2.2. Sequential Time Slice Alignment Based Unequal-Length Phase Partition.** The basic idea of the proposed phase partition algorithm is to evaluate samples of each batch sequentially along time direction starting from the initial phase model by checking the changes of variable correlations. The initial phase model is developed based on the first time-slice where batches have similar characteristics, and then it can be used to judge whether samples of next time-slice are in the same phase as the current one or not. The specific procedure is presented as follows.

*Step 1. Data Input.* Input the time-slice data matrix $\mathbf{X}_k(I_k \times J)$. It is noted that each of the time slices does not have zero mean and standard deviation since variable-unfolding based normalization method are performed to normalize the whole variable-unfolding batch data rather than each time slice.

*Step 2. Initial Phase Model Development.* Develop an initial monitoring model $\mathbf{P}_{v,1}(J \times R)$ by performing PCA algorithm on the first time-slice data matrix, here termed as $\mathbf{X}_{v,1}(I_1 \times J)$.

$$\mathbf{X}_{v,1} = \mathbf{T}\mathbf{P}_{v,1}^{\mathrm{T}} + \mathbf{E} = \sum_{r}^{R} \mathbf{t}_r \mathbf{p}_r^{\mathrm{T}} + \mathbf{E}$$
$$\mathbf{T} = \mathbf{X}_{v,1}\mathbf{P}_{v,1} \tag{1}$$

where $v$ means that time-slice data is variable-unfolded. If there is only one time-slice, $\mathbf{X}_{v,k}(\sum_1^k I_k \times J)$ equals to $\mathbf{X}_k(I_k \times J)$. $\mathbf{P}_{v,1}$ will work as the initial phase model to evaluate the variable corrections of the following samples for each batch. The number of principal components (PCs) $R$ is determined by cumulative explained variance rate to keep most of the process variability (95% here). Then calculate the monitoring statistic values of squared prediction errors (SPE) and determine the preliminary confidence limit $\mathrm{Ctrl}_{v,1}^*$ of SPE by a weighted Chi-squared distribution.[29] SPE $\approx g{\cdot}\chi^2_{h,\alpha^*}$, where $g = v/2m$ and $h = 2(m)^2/v$, in which $m$ is the average of SPE values, and $\boldsymbol{v}$ is the corresponding variance. Since the time-wise variations are larger and more complex than that of batch-wise, a looser confidence limit should be set up to construct a comprehensive confidence region for evaluating process characteristics along time direction. Thus, the final control limit of SPE $\mathrm{Ctrl}_{v,1} = \alpha \mathrm{Ctrl}_{v,1}^*$, $\alpha$ is an adjustable coefficient, which determines how much difference is allowed for variable correlations of the next time-slice in comparison with the existing ones. A simple analysis regarding the adjustable parameter $\alpha$ can be seen the work for phase partition of regular batches.[10]

*Step 3: Variable Correlation Evaluation.* The current model $\mathbf{P}_{v,1}(J \times R)$ is adopted to monitor each batch of the next time-slice and *SPE* statistic is calculated for evaluating variable correlations.

$$\mathbf{t}_i = \mathbf{P}_{v,1}\mathbf{x}_i$$
$$\mathbf{e}_i = \mathbf{x}_i - \mathbf{P}_{v,1}^{\mathrm{T}}\mathbf{t}_i$$
$$\mathrm{SPE}_i = \mathbf{e}_i^{\mathrm{T}}\mathbf{e}_i \tag{2}$$

where subscript $i$ represents the $i$th batch of the next time-slice. For each batch, if three consecutive SPE exceed the control limit by adopting the current monitoring model $\mathbf{P}_{v,1}$, it is "abnormal", which means that its variable correlations cannot be well described by the current phase model and it may enter the next phase. Then record the first alarming time $k_i$ and the batch data before $k_i$ is denoted as one subphase. Otherwise, it means the batch is operating in the "normal" condition. Here, "normal" means the samples are in the current phase.

*Step 4: Phase Model Updating.* Add the "normal" batches of the next time-slice into the modeling data of current monitoring model and variable-unfold them, $\mathbf{X}_{v,2}(\sum_1^2 I_k \times J)$. Perform PCA on the variable-unfolded data $\mathbf{X}_{v,2}(\sum_1^2 I_k \times J)$ to obtain the updated monitoring model $\mathbf{P}_{v,2}(J \times R)$ and


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

confidence limit $\mathrm{Ctrl}_{v,2}$. The updated model is then utilized to monitor batches of the next time-slice.

*Step 5: Phase Landmark Determination.* Since phases are irregular over batches, they cannot be identified simultaneously and iterative steps should be carried out. Repeat steps 3 and 4 until the first subphase is identified for every batch.

*Step 6: Recursive Implementation.* Remove the first irregular subphase, and the remaining data is aligned and employed as the new input in step 2. Then repeat steps 2−5 to determine the following phases.

The flowchart of the proposed phase partition algorithm is presented in Figure 2. Through the above partition procedure,

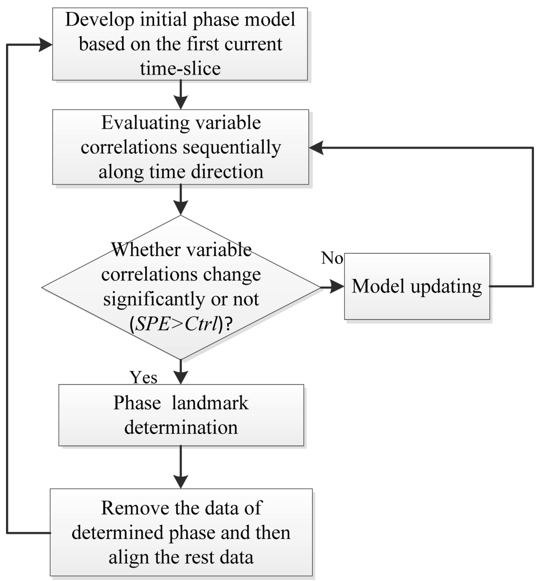



**Figure 2.** Flowchart of the proposed sequential time slice alignment based unequal-length phase partition algorithm.

irregular time-varying process characteristics is effectively traced and thus irregular phases are automatically identified for each batch in order. Those samples with similar variable correlations denoting similar operation patterns are collected together so that they can be modeled by a representative model while those different ones will be characterized by different models. What's more, two issues should be noted. First, moderate uneven problem is focused on in this work, where process characteristics over irregular batches do not change significantly and can be modeled by a representative model within the same phase. Otherwise, when the uneven problem is serious, the process characteristics over irregular batches are of great difference where a unified phase model may lose high resolution to each batch. Second, when the process characteristics slowly changes between phases, that is, between-phase transition patterns may exist, the proposed algorithm does not distinguish phases and between-phase transitions. It determines the phase landmark until the changes of process characteristics is significant. The between-phase transition problem has been an important issue and can cause serious phase partition problem for both regular batches and irregular batches, which deserves further studied in our future work.

**2.3. Discussion and Analysis.** (1) During the uneven-length batch situation, batches may present significantly different process characteristics (e.g., variable correlations) in the same time-slice, which is different from the even-length batch case where similar operation events happen within the same time-slice and the time-slice can be used as the basic unit for phase partition. Correspondingly, each batch should be analyzed separately to identify the phase landmarks. Moreover, though phases are irregular over batches, samples within the same irregular phase have the similar variable correlations while different phases have different variable correlations. By sequentially checking the changes of variable correlations for each batch, the irregular phases can be identified in order.

(2) As the confidence limit is enlarged by a relaxing coefficient $\alpha$, the irregular phase partition results are closely related with the parameter $\alpha$. It determines the extent of dissimilarity of variable correlations for the neighboring samples within the same phase. Larger $\alpha$ value means more difference of variable correlations is allowed and the accuracy of phase model will be decreased since different variations should be simultaneously captured by one representative model; smaller $\alpha$ value means more accurate monitoring models to describe each sample, resulting in more monitoring models.

(3) It is noted that different normalization methods are performed for phase partition and subphase modeling. First, for phase partition, variable-unfolding based method are performed before phase alignment, real batch-wise variations cannot be well reflected and more variations are thus simultaneously covered by the resulting PCA models. However, although the PCA model is influenced by the normalization, the changes of process characteristics remain the same. The irregular phase landmarks can thus be determined by checking the changes of variable correlations evaluated by SPE. Second, for subphase modeling, since irregular phases are aligned and operation events are synchronized within the same phase, batch-wise unfolding based method is performed to normalize each time slice and real batch-wise variations are captured to establish monitoring models.

**2.4. Time Slice Rearrangement for Subphase Model Development.** Since uneven-length batch process has been divided into several irregular phases, the uneven batch information on each phase such as the shortest length $k_s$ as well as the longest length $k_l$ have been determined. It is noted that time slices are regular during the region $[1, k_s]$ and irregular within the region $[k_s+1, k_l]$. In order to perform reliable statistical analysis, a time slice $\mathbf{X}(I \times J)$ should meet the constraint that $I \geq (2{-}3)J$ as suggested by Johnson et al.[30] For regular time slices, the constraint is satisfied. For those irregular time slices, since each of them contains insufficient batches, a generalized time-slice $\mathbf{X}^w(\sum_{k_s+1}^{k_l} I_k \times J)$ is constructed by variable-wise unfolding them. Thus, for samples during $[k_s+1, k_l]$, they are all represented by the generalized time-slice $\mathbf{X}^w$. After aligning the time slices, batch-unfolding based method is then performed to normalize the regular time slices and generalized time slice.

Then, the subphase data $\mathbf{X}_c(\sum_1^{K_c} I_k \times J)$ are arranged by variable-unfolding the normalized time-slices $\mathbf{X}_k(I_k \times J)(k = 1,2,...,K_c)$ and $\mathbf{X}^w$ within each phase $c$. The phase-specific model can be obtained by performing PCA on the subphase data $\mathbf{X}_c$.

$$\begin{aligned}
\mathbf{T}_c &= \mathbf{X}_c\mathbf{P}_{c,s} \\
\mathbf{X}_c &= \mathbf{T}_c\mathbf{P}_{c,s}^{\mathrm{T}} + \mathbf{E}_c = \mathbf{X}_c\mathbf{P}_{c,s}\mathbf{P}_{c,s}^{\mathrm{T}} + \mathbf{E}_c \\
\hat{X}_c &= \mathbf{T}_c\mathbf{P}_{c,s}^{\mathrm{T}} = \mathbf{X}_c\mathbf{P}_{c,s}\mathbf{P}_{c,s}^{\mathrm{T}} \\
\mathbf{E}_c &= \mathbf{X}_c - \hat{\mathbf{X}}_c = \mathbf{X}_c\mathbf{P}_{c,e}\mathbf{P}_{c,e}^{\mathrm{T}}
\end{aligned} \tag{3}$$


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

where $\mathbf{P}_{c,s}(J \times R_c)$ are loadings revealing the major variation directions captured within the current phase. The major process variations are captured by $\mathbf{T}_c(\sum_1^{K_c} I_k \times R_c)$, which is the scoring matrix extracted from process variables. $R_c$ is the PCs number retained in the current phase. $\hat{\mathbf{X}}_c$ is the reconstructed subphase data by $\mathbf{T}_c$, which contains the systematic information on $\mathbf{X}_c$. $\mathbf{E}_c$ are residuals after the reconstruction, where variable correlation information is included. $\mathbf{P}_{c,e}(J \times R_r)$ $(R_r = J - R_c)$ reflects the minor variation directions.

The subspaces spanned by $\mathbf{P}_{c,s}$ and $\mathbf{P}_{c,e}$ are called systematic subspace and residual subspace, respectively. Then, two monitoring statistics are calculated in each subspace, $T^2$ for systematic subspace and SPE for residual subspace.

$$T_k^{\,2} = (\mathbf{t}_k - \bar{t}_k)^{\mathrm{T}} \sum\nolimits_c^{-1} (\mathbf{t}_k - \bar{\mathbf{t}}_k)$$
$$\mathrm{SPE}_k = \mathbf{e}_k^T \mathbf{e}_k \tag{4}$$

where $\mathbf{t}_k(R_c \times 1)$ is the PC vector separated from $\mathbf{T}_c(\sum_1^{K_c} I_k \times R_c)$ and $\bar{\mathbf{t}}_k$ is the mean vector of $\mathbf{T}_k(I_k \times R_c)$ which are time-slice scores separated from $\mathbf{T}_c(\sum_1^{K_c} I_k \times R_c)$. It is noted that time-slices $\mathbf{X}_k$ during $[1, k_s]$ are in certain normalized to have zero mean and standard deviation, resulting in zero vector $\bar{\mathbf{t}}_k$. For the irregular time-slices at the end of each phase, since less process information is included in each of them, $\bar{\mathbf{t}}_k$ is in fact zero vector. $\sum_c$ is the covariance matrix of $\mathbf{T}_c(\sum_1^{K_c} I_k \times R_c)$. $\mathbf{e}_k(J \times 1)$ is the residual vector which is obtained from $\mathbf{E}_c(\sum_1^{K_c} I_k \times J)$. In this way, $T^2$ describes systematic variations captured by the monitoring model $\mathbf{P}_{c,s}$ and SPE reveals variations occupied by $\mathbf{P}_{c,e}$ in the residual part which reflects variable correlations.

For each irregular phase, confidence limit can be defined at each time for the two monitoring statistics. Assuming the process data follows a multivariate normal distribution, so confidence limit can be determined by *F*-distribution and weighted chi-squared distribution[29] for $T^2$ and SPE respectively. However, considering the confidence limit based on $T^2$ established by *F*-distribution is not very sensitive, another way can be utilized to define confidence limit: arrange the values of monitoring statistic in a descending order at each time and choose the value at 95% (or 99%) percentile of the ordered data as the confidence limit.

### 2.5. Online Phase Judging and Process Monitoring.

After the monitoring models have been constructed for each irregular phase, online monitoring strategy can be implemented for new samples. For multiphase batch processes with irregular phases, an important issue is to judge the phase affiliation of each new sample because process time index fails to reflect the phase affiliations of them. Only after new samples have been located to their correct phases, the operation statuses of them can be correctly checked by utilizing proper monitoring model. Since monitoring models are developed in a sequential way, an easy but effective way to judge phase affiliation and check the operation statuses of new samples is to distinguish different regions with different online monitoring strategies according to the phase landmarks. Here, two regions are distinguished: within-phase region and between-phase shifting region.

*a. Within-Phase Region.* The region $[1, k_s]$ ($k_s$ is the shortest phase length of the current phase) is defined as within-phase region because the phase affiliation of the data during this region is certain and process time is sufficient to judge their phase affiliation. Then according to the process time index, proper phase model will be adopted and two statistics are calculated to check their operation statuses. If both statistics stay well in the confidence region, it means the new sample operates normally; otherwise, there may be process fault.

*b. Between-Phase Shifting Region.* The region $[k_s + 1, k_l]$ ($k_l$ is the longest phase length of the current phase) is termed as between-phase shifting region, in which samples may still lie in the current phase or shift to the next phase. Since phase landmarks are misaligned over batches in this region and phase affiliation of each new sample cannot be determined only by process time, other online monitoring strategy should be implemented. First, the current phase model is employed. If there is no alarm issued by the current phase model, it means the new data operate normally in the current phase. Otherwise, the current phase model fails to explain their underlying variable correlations and thus the next phase models are adopted to check their operation statuses. For normal data shifting to the next phase, they can be well described by the next phase model, so that alarms can be eliminated. For process fault, alarms still exist by adopting the next phase model and neither of the current phase model and the next phase model cannot well describe their underlying characteristics.

## 3. ILLUSTRATION AND DISCUSSION

### 3.1. Process Description.

In this section, a typical multiphase batch process, injection molding,[31,32] is used to illustrate the performance of the proposed algorithm. Injection molding is a critical process in polymer processing, it can be used to form a wide variety of products, whose complexity is unlimited and sizes may range from very small to very large. A typical injection molding process can be divided into three major operation phases: first, injection of the plastic melt into the mold, followed by the packing holding of the plastic in the mold under pressure, and finally, cooling of the plastic in the mold. What's more, during the early stage of the cooling phase, plastication happens, the material is conveyed forward from the feed hopper through the barrel by a rotating screw. When the cooling is sufficiently completed, the mold opens and parts are ejected from the mold, preparing for the next cycle.

**Table 1. Description of the Used Process Variables**

| no. | description | units |
|---|---|---|
| 1 | nozzle temperature | °C |
| 2 | nozzle pressure | bar |
| 3 | screw stroke | Mm |
| 4 | screw velocity | mm/s |
| 5 | injection pressure | bar |
| 6 | plastication pressure | Var |
| 7 | SV1 opening | % |
| 8 | SV2 opening | % |
| 9 | cavity pressure | bar |

Injection molding is a typical multiphase batch process, in which many factors such as operation conditions and control objectives may cause the different durations for each operation cycle. It can be easily set as a typical uneven-length multiphase batch process for experiments to verify the proposed phase analysis and modeling method for process monitoring, where the injection phase duration is not fixed but rather depends on the injection velocity. Obviously, a lower injection velocity results in more injection time and thus a longer batch with more process data. Here, the injection velocity is artificially set to change from 22 to 26 mm/s, involving three typical velocity values: 22, 24, 26 mm/s. The injection phase ranges from 96 to


DOI: 10.1021/acs.iecr.5b01405
*Ind. Eng. Chem. Res.* 2015, 54, 10020−10030

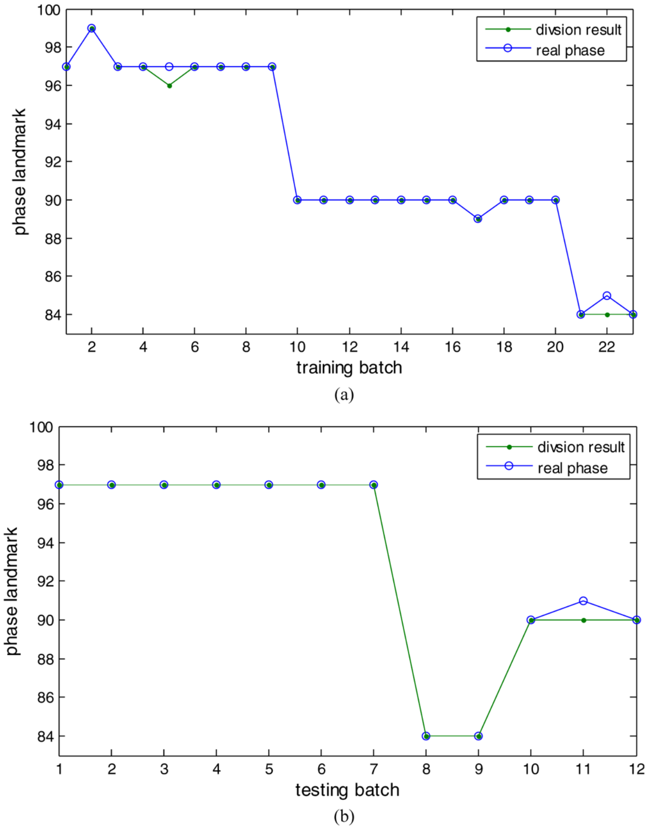



**Figure 3.** Phase partition results of injection phase for (a) training batches and (b) normal testing batches (blue cycle line: real phase; green dot line: partition phase).

84 samples, where maximal injection phase duration is 99 sample intervals corresponding to injection velocity 22 mm/s and the minimal filling duration is 84 associated with velocity 26 mm/s. Therefore, the difference of filling duration is 15 samples. It is apparently that moderate uneven-length problem has been simulated, where the underlying characteristics of each uneven batch are similar within the same phase. Here for simplicity, except the injection phase, the other phases are controlled to have exactly the same duration.

The material used in this work is high-density polyethylene (HDPE). Nine process variables are selected for modeling, which can well reveal the process characteristics of different operation statuses. They are shown in Table 1. In total, 35 normal batches are collected under normal operation conditions, where 23 batches are used for modeling. Besides, two types of faults are simulated.

(1) Injection phase fault simulation, where the screw does not start at the original point, causing the mold to fill incompletely. Normal screw stroke is about 42.5 mm while it is round 20 mm under this fault condition.

(2) Packing-holding pressure fault simulation, in which the holding pressure cannot maintain at its steady value, the real holding pressure falls down when the packing-holding stage begins.

**3.2. Irregular Phase Partition Results.** First, 23 uneven-length batches are obtained, in which batch duration ranges from 532 samples to 545 samples. Then, 545 irregular time-slices $\mathbf{X}_k(I_k \times J)$ are preprocessed for phase partition. The proposed algorithm is then performed to determine the irregular phases. As shown in Figure 3, the uneven injection phases for each batch have been automatically identified by the proposed algorithm regarding $\alpha$ = 1.2. For comparison, real injection physical phases, which are indicated by the variable called screw stroke here, have also been plotted. It is noted that

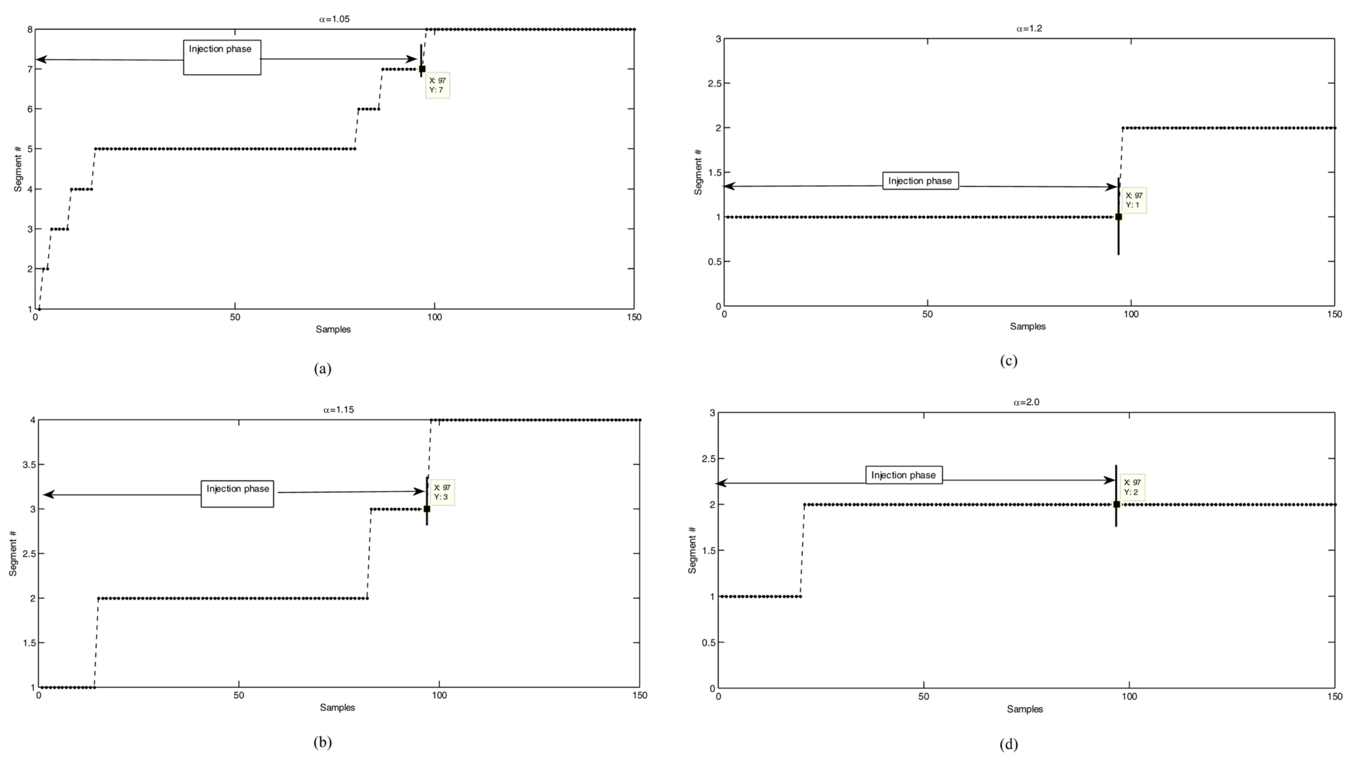



**Figure 4.** Phase partition results regarding different $\alpha$ of one batch ((a) $\alpha_1$, (b) $\alpha_2$, (c) $\alpha_3$, and (d) $\alpha_4$).


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

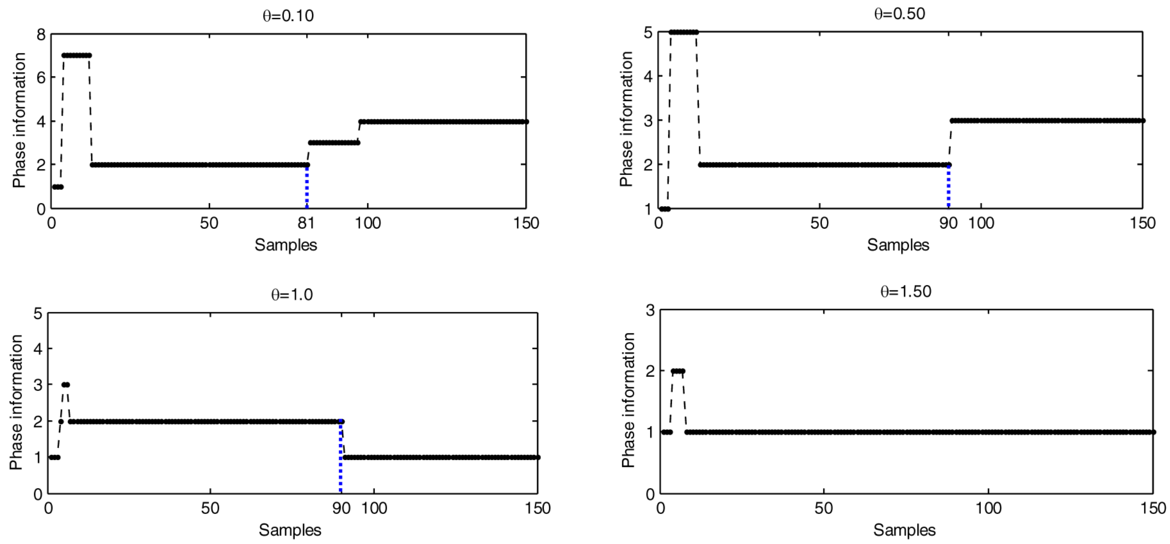



**Figure 5.** Initial phase partition results using clustering-based algorithm regarding different $\theta$.

**Table 2. Online Monitoring Performance Comparison between the Proposed Method and Clustering Based Method**

| method | $\alpha$ | | 1.00 | 1.15 | 1.20 | 2.00 |
|---|---|---|---|---|---|---|
| proposed algorithm | FAR% | $T^2$ | 3.00 ± 2.65 | 2.84 ± 1.67 | 1.18 ± 1.01 | 6.02 ± 5.52 |
| | | SPE | 4.56 ± 3.89 | 4.02 ± 3.47 | 2.70 ± 2.81 | 8.95 ± 8.48 |
| | $\Delta T$ | $T^2$ | 2.28 ± 2.69 | 1.93 ± 1.38 | 0.57 ± 0.98 | 5.39 ± 5.72 |
| | | SPE | 5.24 ± 4.69 | 4.96 ± 3.54 | 4.00 ± 1.96 | 7.78 ± 7.32 |
| method | $\theta$ | | 0.10 | 0.50 | 1.00 | 1.50 |
| clustering based algorithm | FAR% | $T^2$ | 5.68 ± 3.12 | 8.04 ± 6.67 | 10.02 ± 9.52 | 11.18 ± 10.02 |
| | | SPE | 7.86 ± 4.68 | 9.82 ± 5.47 | 12.73 ± 9.96 | 15.95 ± 10.48 |
| | $\Delta T$ | $T^2$ | 3.80 ± 3.76 | 6.93 ± 4.38 | 8.39 ± 8.72 | 9.59 ± 8.13 |
| | | SPE | 5.28 ± 3.18 | 7.96 ± 5.54 | 10.78 ± 10.32 | 12.71 ± 10.78 |

**Table 3. Online Monitoring Performance (FAR%) for Normal Case for Two Phases ($\alpha$ = 1.2)**

| | injection phase | packing holding phase |
|---|---|---|
| SPE | 2.13 ± 2.84 | 4.14 ± 2.86 |
| $T^2$ | 1.29 ± 1.93 | 1.07 ± 1.07 |

**Table 4. Online Monitoring Performance ($\Delta T$ (Sample)) for Two Fault Cases ($\alpha$ = 1.2)**

| fault type | injection fault | | | | | holding pressure fault | | mean ± MAD |
|---|---|---|---|---|---|---|---|---|
| batch no. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| $T^2$ | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0.57 ± 0.98 |
| SPE | 3 | 5 | 6 | 4 | 9 | 0 | 1 | 4.00 ± 2.29 |

using the proposed algorithm, the uneven injection phases are clearly identified and agree well with real physical phases. What's more, a measurement index $\Delta t = |\text{phase}_{\text{real}} - \text{phase}_{\text{partition}}|$, has been defined to evaluate the accuracy of phase division results. The mean and mean absolute deviation (MAD) values of $\Delta t$ are calculated for 23 modeling batches and 12 testing normal batches, respectively. For modeling batches, $\Delta t = 0.13 \pm 0.23$ (mean ± MAD); for testing data, $\Delta t = 0.08 \pm 0.15$ (mean ± MAD). It is noted that both $\Delta t$ have small means and deviations, also revealing a good phase partition results which agree well with the real phases, so that reliable phase information can be derived for the following statistical analysis and modeling. Since $\alpha$ is an important parameter, how $\alpha$ influences the irregular phase partition results is discussed in this section. In general, the uneven-length injection phase will be divided into several batch-wise irregular subphases in order when $\alpha$ is small; If $\alpha$ becomes large, less phases will be partitioned and then the irregular injection phase may not be identified correctly. The phase partition results of one batch regarding different $\alpha$ are presented in Figure 4 to illustrate how $\alpha$ influences the irregular phase partition results. First, it is obvious that the phase partition results are easy to read and need no post processing. Then, for $\alpha_1$ and $\alpha_2$, the injection phase is divided into several subphases; for $\alpha_3$, injection phase can be correctly identified while during the last case, injection phase is mixed with the following phase. The larger $\alpha$ becomes, the more differences of variable correlations is allowed for samples within the same phase. Thus, the model accuracy will be decreased and the irregular phase characteristics may not be well reflected.

For comparison, the initial phase partition results using clustering-based algorithm regarding different $\theta$[26−28] without post processing are shown in Figure 5. First, it is noted that phases are disordered that samples belonging to different time regions are classified into the same phase. The results are hard to understand and thus post processing should be implemented. Second, compared to the real physical phases, samples belonging to the packing-holding phase are divided into the common part[26,27] of the irregular injection phase regarding $\theta_2 \sim \theta_4$, which may decrease the accuracy of the monitoring model for that samples with significantly different characteristics are mixed together. It is because that time-slice is treated as the


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

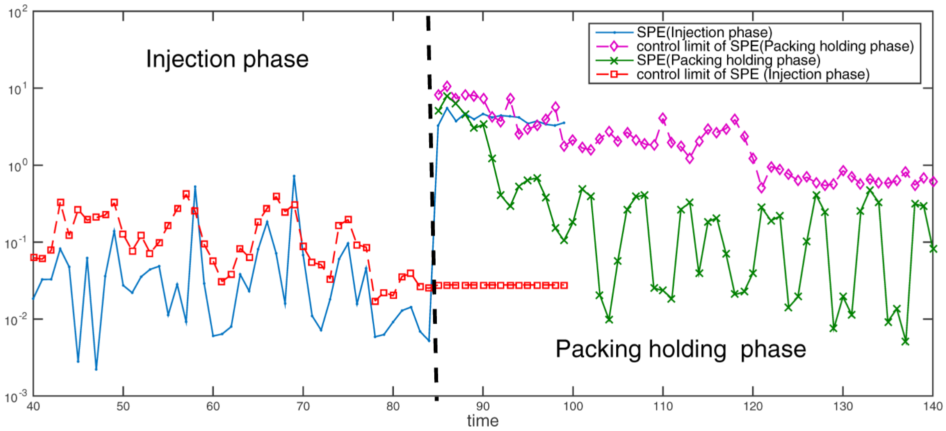



**Figure 6.** SPE monitoring results for normal phase switch (square line: *SPE* control limit of injection phase; diamond line: SPE control limit of packing-holding phase; dot line: SPE statistic in injection phase; cross line: SPE statistic in packing-holding phase).

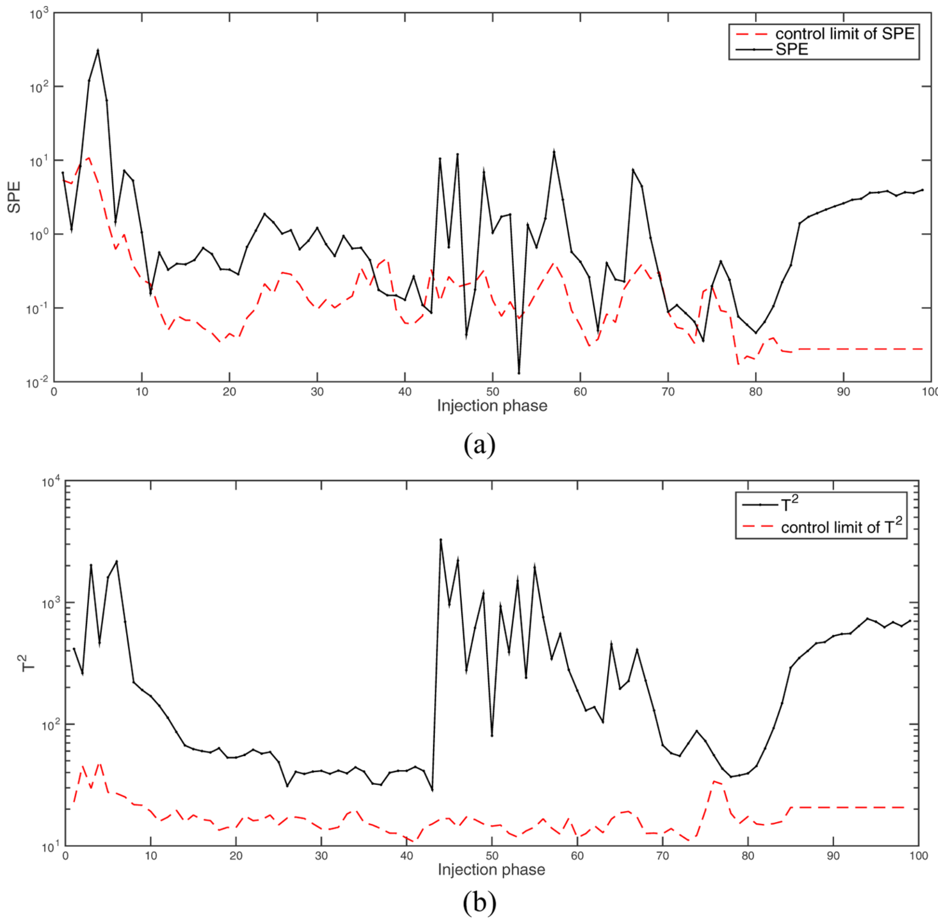
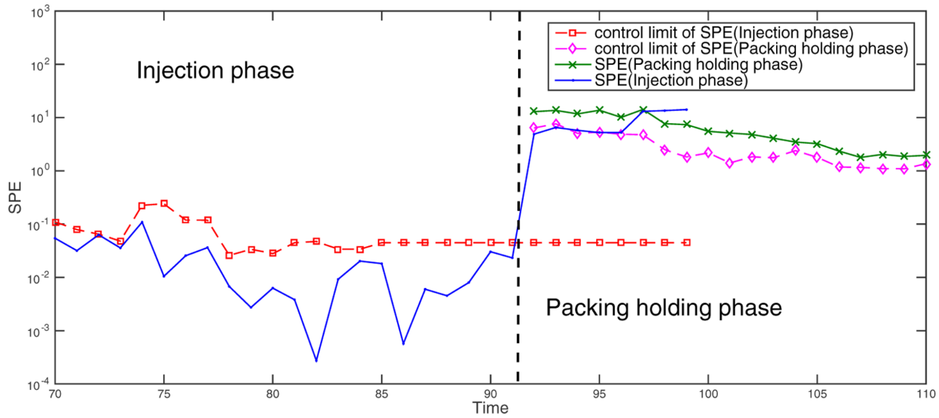



**Figure 7.** Monitoring results of injection fault by (a) SPE and (b) $T^2$ (dash line: control limit; dot line: monitoring statistic).

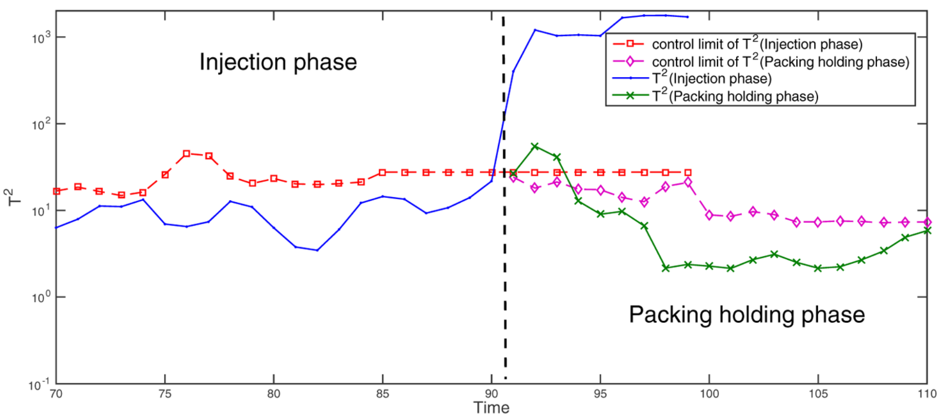
analysis unit during the clustering-based method and the characteristics of specific batches may not be correctly reflected.

**3.3. Online Monitoring Performance.** Based on the phase division results, different monitoring models are developed for each phase. First, monitoring performance of the proposed method and the clustering-based methods[26,27] regarding different parameters are summarized in Table 2. For illustrating the online monitoring performance, monitoring results of both irregular injection phase and the following packing-holding phase are presented. FAR and $\Delta T$ are calculated for evaluating the monitoring performance. FAR is false alarming rate, one FAR is calculated for 12 batches where the total number of false alarming signals is divided by the total number of samples from 12 batches. $\Delta T$ = FAT − FOT, revealing the fault detection ability (FAT is termed as first alarming time, FOT is called first occurring time). The mean and mean absolute deviation (MAD) values of FAR and $\Delta T$ are calculated, respectively. It is clear that the best monitoring performance are obtained with respective to $\alpha$ = 1.2 for the proposed algorithm and $\theta$ = 0.1 for the clustering-based algorithm. It can be seen that the proposed algorithm generates better results than those by clustering-based method, which is


DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030




(a)




(b)

**Figure 8.** Monitoring results of holding pressure fault by (a) SPE and (b) $T^2$ (square line: injection phase control limit; diamond line: packing holding phase control limit; dot line: monitoring statistics in injection phase; cross line: monitoring statistics in packing holding phase).

evaluated by paired $t$ test ($\alpha$ = 0.05). That may result from the fact that the clustering-based method, which uses distance to evaluate the similarity, does not directly consider the monitoring performance. That is, the samples in phase A may incorrectly judged to be similar to those in phase B by distance index.

Conclusion can also be drawn from Table 2 that online monitoring performance shows no statistical significantly different regarding different $\alpha$ except $\alpha$ = 2.0, based on paired $t$ test ($\alpha$ = 0.05). Then, based on the phase partition results by $\alpha$ = 1.2, the online strategy of the proposed method is illustrated in detail. For samples during the within-phase region [1, 84], their operation statuses can be checked by adopting the injection phase model. Two statistics $T^2$ and SPE are also calculated. If both of them stay well within the confidence region, it means the current phase model can well describe the process characteristics and thus process operates normally in the current phase; otherwise, process fault may happen. For data during between-phase shifting region, which is [85,99] here, how to distinguish normal phase shift and process fault is critical. Different online monitoring strategy should be implemented. First, injection phase model is employed to check whether there are any alarms for $T^2$ or SPE. If alarms occur, packing-holding phase model is then utilized to explore the cause of the alarms, the real process fault or normal phase shift. On condition that alarming signal can be eliminated by the packing-holding phase model, process is regarded as normal operation and it just shifts from injection phase to packing-holding phase; otherwise, process fault happens and fault diagnosis methods should be carried out which is not discussed here. What's more, with regarding to data after 99, they are in certain beyond the irregular injection phase and corresponding monitoring model can be adopted according to time index.

Table 3 shows the monitoring performance results of injection phase and packing-holding phase respectively with $\alpha$ = 1.2. FAR (false alarming rate) of $T^2$ and SPE are calculated for normal testing batches. Obviously, FAR of SPE range from 5% to 8%, which agree with the 95% confidence limit. While for $T^2$, FAR are much smaller for the reason that phases are identified according to variable correlations evaluated by SPE, so $T^2$ is not so sensitive as SPE for online monitoring. What's more, it is noted that the FAR results do not show no statistical significantly differences between injection phase and packing-holding phase, based on paired $t$ test ($\alpha$ = 0.05), indicating that the irregular injection phase is correctly identified so that the injection phase model and the following phase model can well capture the corresponding process characteristics.

Table 4 summarizes the specific monitoring performance results of two types of faults, evaluated by $\Delta T$ index. As illustrated in Table 4, batches 1−5 are collected under injection fault, batches 6 and 7 belong to holding pressure fault. It is noted that for injection fault, both statistics are effective to detect fault with small time delays. Besides, $T^2$ has better detection ability than SPE, as the fault is mainly caused by



DOI: 10.1021/acs.iecr.5b01405
Ind. Eng. Chem. Res. 2015, 54, 10020−10030

reducing screw stroke which is readily to be reflected on $T^2$ in principle component space. While for holding pressure fault mentioned above, SPE is as sensitive as $T^2$ for fault detection because when holding pressure falls down, variable correlation is also changed significantly.

As mentioned above, it is a key problem to distinguish process fault and phase shift for those samples during between-phase shifting region. Figure 6 demonstrates SPE monitoring chart for phase shift. Real phase shift occurs at the 86th sampling time, from injection phase to packing–holding phase. It is clearly that SPE index continuously exceeds the confidence limit starting from 86th sample by adopting the injection phase model and then alarming signal has been brought back to the normal region using packing-holding phase model. Thus, phase shift of the testing batch can be correctly pointed out. For testing batches of two types of faults, the fault detection charts are presented in Figures 7 and 8 respectively. For the injection fault, since it happens at the beginning of the process, it can be readily distinguished from normal phase shift by the proposed algorithm. As shown in Figure 7, both the SPE chart and $T^2$ chart can detect the injection fault happening at the beginning of the process almost in time. It can also be concluded from Figure 7 that $T^2$ chart has better fault detection performance than SPE with regard to this specific fault. While for packing-holding fault which is introduced during the between-phase shifting region, it can be seen from Figure 8 that the alarming signal cannot be brought back to the confidence region for both two statistics. Thus, packing-holding fault is distinguished from normal phase shift. However, for packing-holding fault, alarms last for several sampling time and then come back to the confidence region with regard to $T^2$, which may result from interactive effects among variables that the influences on $T^2$ caused by falling down holding pressure are offset by some other variables.

## 4. CONCLUSION

In the present work, a sequential time slice alignment based unequal-length phase identification and modeling method is proposed for fault detection of irregular batches. By sequentially evaluating changes of variable correlations of each batch along time direction regarding their influences on monitoring performance, different underlying process characteristics are distinguished and irregular phases are automatically identified without post processing. Based on the irregular phase partition results, time slices are realigned and generalized time-slices are constructed for those irregular data to develop monitoring models. For online monitoring, phase affiliation can be real-time judged and phase shift can be well distinguished from process fault. The proposed method is applied to process monitoring in injection molding process compared with the clustering-based methods. The monitoring results demonstrate that the proposed approach has shown preferable monitoring performance.

## ■ APPENDIX

### SSPP Algorithm

The specific procedure of SSPP algorithm is described as below.

*Step 1: Data Preparation.* Prepare the time-slice data matrix $\mathbf{X}_k(I \times J)$ from the process beginning. The variables at each time are assumed to have been normalized to have zero mean and standard deviation.

*Step 2: Time-Slice Based PCA Modeling.* Perform PCA algorithm on the time-slice data matrices and get the initial time-slice models. The number of principal components (PCs) is determined by cumulative explained variance rate to keep most of the process variability (90% here). Then find the number of PCs that occurs most throughout the batch process and set it as the unified dimension of time-slice PCA models, $R$. Calculate the monitoring statistic values of squared prediction errors (SPE) after the explanation of PCA model at each time and determine the confidence limit, $\mathrm{Ctr}_k$, by a weighted Chi-squared distribution. It represents the reconstruction power of time-slice PCA model.

*Step 3: Time-Segment Based PCA Modeling.* From the beginning of batch processes, add the next time-slice to the existing ones and variable-unfold them within the current time region, $\mathbf{X}_{v,k^*}(Ik^* \times J)$ (where subscript $v$ denotes variable-unfolding data arrangement and subscript $k^*$ denotes the current time). Perform PCA on the rearranged data matrix and get the time-segment PCA model, $\mathbf{P}_{v,k^*}(J \times R)$. Calculate the SPE values for each time-slice data matrix up to $k^*$ by the explanation of the current time-segment PCA model ($\mathbf{P}_{v,k^*}(J \times R)$) and determine the confidence limit at each time, $\mathrm{Ctr}_{v,k}$, by a weighted Chi-squared distribution. It represents the reconstruction power of this time-segment PCA model to each time-slice up to $k^*$.

*Step 4: Compare Model Accuracy.* Compare $\mathrm{Ctr}_{v,k}$ with $\mathrm{Ctr}_k$ for each time slice up to $k^*$. If more than three consecutive samples show $\mathrm{Ctr}_{v,k} > \alpha^*\mathrm{Ctr}_k$, where $\alpha$ is a constant attached to $\mathrm{Ctr}_k$, termed relaxing factor here. It means that the addition of the current time-slice have imposed great influences on the time-segment PCA monitoring model and the resulting monitoring performance. The accuracy of time-segment model is thus significantly worse than that of time-slice models. So $\alpha$ determines how much the time-segment PCA model is allowed to be less representative than time-slice PCA models. The time slices before $k^*$ are thus denoted as one phase.

*Step 5: Data Updating and Recursive Implementation.* Remove the first subphase and the left batch process data are now employed as the new input data in step 3. Recursively repeat steps 3–4 from the updated beginning of the batch process to find the following segments.

The output is a partition of process trajectory along time direction and the alternation of different subphases along time direction.

## ■ AUTHOR INFORMATION

### Corresponding Author

*Tel: 86-571-87951879. Fax: 86-571-87951879. E-mail: chhzhao@zju.edu.cn.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work is supported by Zhejiang Provincial Natural Science Foundation of China (LR13F030001), the National Natural Science Foundation of China (No. 61422306 and No. 61433005), the Fundamental Research Funds for the Central Universities, Zhejiang University K.P. Chao's High Technology Development Foundation, and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT1522).



DOI: 10.1021/acs.iecr.5b01405
*Ind. Eng. Chem. Res.* 2015, 54, 10020–10030

## REFERENCES


(1) Jackson, J. E. *A User's Guide to Principal Components*; Wiley: New York, 1991,.

(2) Wang, X. Z. *Data mining and knowledge discovery for process monitoring and control*; Springer: London, 1999.

(3) Nomikos, P.; MacGregor, J. F. Monitoring batch processes using multiway principal component analysis. *AIChE J.* **1994**, *40*, 1361–1375.

(4) Nomikos, P.; MacGregor, J. F. Multi-way partial least squares in monitoring batch processes. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 97–108.

(5) García-Munoz, S.; Kourti, T.; MacGregor, J. F. Model predictive monitoring for batch processes. *Ind. Eng. Chem. Res.* **2004**, *43*, 5929–5941.

(6) Dong, D.; McAvoy, T. J. Multistage batch process monitoring. *Proceedings of American control conference* **1995**, 1857–1861.

(7) Neogi, D.; Schlags, C. E. Multivariate statistical analysis of an emulsion batch process. *Ind. Eng. Chem. Res.* **1998**, *37*, 3971–3979.

(8) Ündey, C.; Cinar, A. Statistical monitoring of multistage, multiphase batch processes. *IEEE Contr. Syst. Mag.* **2002**, *22*, 40–52.

(9) Lu, N. Y.; Gao, F. R.; Wang, F. L. Stage-based process analysis and quality prediction for batch processes. *Ind. Eng. Chem. Res.* **2005**, *44*, 3547–3555.

(10) Zhao, C. H.; Sun, Y. X. Step-wise sequential phase partition (SSPP) algorithm based statistical modeling and online process monitoring. *Chemom. Intell. Lab. Syst.* **2013**, *125*, 109–120.

(11) Rothwell, S. G.; Martin, E. B.; Morris, A. J. Comparison of methods for dealing with uneven length batches. *Proc. 7th Int. Conf. Comput. Appl. Biotechnol.* **1998**, 387–392.

(12) 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.

(13) Nomikos, P.; MacGregor, J. F. Multivariate SPC charts for monitoring batch processes. *Technometrics* **1995**, *37*, 41–59.

(14) Wold, S.; Kettaneh, N.; Fridén, H.; et al. Modelling and diagnostics of batch processes and analogous kinetic experiments. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 331–340.

(15) Itakura, F. Minimum prediction residual principle applied to speech recognition. *IEEE Trans. Acoust., Speech, Signal Process.* **1975**, *23*, 67–72.

(16) Kassidas, A.; MacGregor, J. F.; Taylor, P. A. Synchronization of batch trajectories using dynamic time warping. *AIChE J.* **1998**, *44*, 864–875.

(17) Niels-Peter, V. N.; Jens, M. C.; Jørn, S. Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimized warping. *J. Chromatogr.* **1998**, *805*, 17–35.

(18) Westerhuis, J. A.; Kourti, T.; MacGregor, J. F. Comparing alternative approaches for multivariate statistical analysis of batch process data. *J. Chemom.* **1999**, *13*, 397–413.

(19) Jian, W.; Ognjen, M.; Barry, L. Uneven batch data alignment with application to the control of batch end-product quality. *ISA Trans.* **2014**, *53*, 584–590.

(20) Ramaker, H.; van Sprang, E.; Westerhuis, J.; Smilde, A. K. Dynamic time warping of spectroscopic batch data. *Anal. Chim. Acta* **2003**, *498*, 133–153.

(21) González-Martínez, J. M.; Ferrer, A.; Westerhuis, J. A. Real-time synchronization of batch trajectories for on-line multivariate statistical process control using dynamic time warping. *Chemom. Intell. Lab. Syst.* **2011**, *105*, 195–206.

(22) Gins, G.; Van den Kerkhof, P.; Van Impe, J. F. M. Hybrid derivative dynamic time warping for online industrial batch-end quality estimation. *Ind. Eng. Chem. Res.* **2012**, *51*, 6071–6084.

(23) Giorgio, T.; Frans van den, B.; Claus, A. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *J. Chemom.* **2004**, *18*, 231–241.

(24) Srinivasan, R.; Qian, M. S. Online fault diagnosis and state identification during process transitions using dynamic locus analysis. *Chem. Eng. Sci.* **2006**, *61*, 6109–6132.

(25) Srinivasan, R.; Qian, M. S. Online temporal signal comparison using singular points augmented time warping. *Ind. Eng. Chem. Res.* **2007**, *46*, 4531–4548.

(26) Lu, N. Y.; Gao, F. R.; Yang, Y.; et al. PCA-based modeling and on-line monitoring strategy for uneven-length batch processes. *Ind. Eng. Chem. Res.* **2004**, *43*, 3343–3352.

(27) Zhao, C. H.; Lu, N. Y. *Statistical monitoring and quality analysis for batch processes*; Science Press: Beijing, 2014.

(28) Zhao, L. P.; Zhao, C. H.; Gao, F. R. Inner-Phase Analysis Based Statistical Modeling and Online Monitoring for Uneven Multiphase Batch Processes. *Ind. Eng. Chem. Res.* **2013**, *52*, 4586–4596.

(29) Lowry, C. A.; Montgomery, D. C. A review of multivariate control charts. *IIE Trans.* **1995**, *27*, 800–810.

(30) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*; Prentice Hall: Englewood Cliffs, NJ, 2002.

(31) Thyregod, P.; Spliid, H.; Melgaard, H.; et al. Modelling and monitoring in injection molding. Technical University of Denmark-Danmarks Tekniske Universitet, Administration, Office for Study Programmes and Student Affairs Afdelingen for Uddannelse og Studerende, 2001.

(32) Yang, Y.; Gao, F. R. Adaptive control of the filling velocity of thermoplastics injection molding. *Control Eng. Pract.* **2000**, *8*, 1285–1296.




DOI: 10.1021/acs.iecr.5b01405
*Ind. Eng. Chem. Res.* 2015, 54, 10020–10030