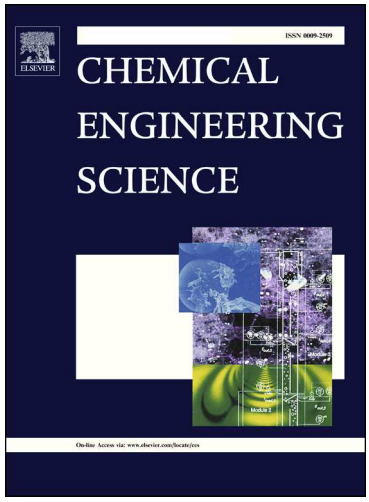
Article:
Qin, Y, Zhao, C, Wang, X et al. (1 more author) (2017) Subspace decomposition and critical phase selection based cumulative quality analysis for multiphase batch processes. Chemical Engineering Science, 166. pp. 130-143. ISSN 0009-2509

https://doi.org/10.1016/j.ces.2017.03.033



Subspace decomposition and critical phase selection based cumulative quality analysis for multiphase batch processes

Yan Qin, Chunhui Zhao, Xuezhong Wang, Furong Gao

PII: S0009-2509(17)30201-4
DOI: http://dx.doi.org/10.1016/j.ces.2017.03.033
Reference: CES 13501

To appear in: *Chemical Engineering Science*

Received Date: 17 October 2016
Revised Date: 7 January 2017
Accepted Date: 13 March 2017

Please cite this article as: Y. Qin, C. Zhao, X. Wang, F. Gao, Subspace decomposition and critical phase selection based cumulative quality analysis for multiphase batch processes, *Chemical Engineering Science* (2017), doi: http://dx.doi.org/10.1016/j.ces.2017.03.033

Subspace decomposition and critical phase selection based cumulative quality analysis for multiphase batch processes

Yan Qin[1], Chunhui Zhao[1], Xuezhong Wang[2], Furong Gao[3]

1 State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China

2 Institute of Particle Science and Engineering, School of Process, Environmental and Materials Engineering, The University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

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

Abstract—Quality analysis and prediction have been of great significance to ensure consistent and high product quality for chemical engineering processes. However, previous methods have rarely analyzed the cumulative quality effect which is of typical nature for batch processes. That is, with time development, the process variation will determine the final product quality in a cumulative manner. Besides, they can not get an early sense of the quality nature. In this paper, a quantitative index is defined which can check ahead of time whether the product quality result from accumulation or the addition of successive process variations and cumulative quality effect will be addressed for quality analysis and prediction of batch processes. Several crucial issues will be solved to explore the cumulative quality effect. First, a quality-relevant sequential phase partition method is proposed to separate multiple phases from batch processes by using fast search and find of density peaks clustering (FSFDP) algorithm. Second, after phase partition, a phase-wise cumulative quality analysis method is proposed based on subspace decomposition which can explore the non-repetitive quality-relevant information (NRQRI) from the process variation at each time within each phase. NRQRI refers

[1] Corresponding author: Tel: 86-571-87951879, Fax: 86-571-87951879, E-mail address: chhzhao@zju.edu.cn

to the quality-relevant process variations at each time that are orthogonal to those of previous time and thus represents complementary quality information which is the key index to cumulatively explain quality variations time-wise. Third, process-wise cumulative quality analysis is conducted where a critical phase selection strategy is developed to identify critical-to-cumulative-quality phases and quality predictions from critical phases are integrated to exclude influences of uncritical phases. By the two-level cumulative quality analysis (i.e., phase-wise and process-wise), it is feasible to judge whether the quality has the cumulative effect in advance and thus proper quality prediction model can be developed by identifying critical-to-cumulative-quality phases. The feasibility and performance of the proposed algorithm are illustrated by a typical chemical engineering process, injection molding.

Index—cumulative quality analysis, subspace decomposition, critical phase selection, batch processes.

1. Introduction

Batch processes have experienced rapid development and become an important manufacturing mode of producing high-value-added products through process repetition in chemical engineering industry[1]. With the nature of quickly responding to changing market demand and customer requirement, batch processes have been widely applied in specialty chemical, biomedical, semiconductor, etc[2]. Caused by process disturbances and batch to batch variations, low reproducibility brings a great challenge to ensure the consistency of product quality. Besides, for many batch processes, measurements of product quality are in general not available until the end of a batch. Thus, online quality prediction technology is indispensable to guarantee high product quality and improve process efficiency for batch processes.

Benefiting from the development of data acquisition and storage technologies, data-driven multivariate statistical methods[3-9] have attracted increasing attention. Only using process

data, multivariate statistical methods have been widely used to develop prediction models, including partial least squares (PLS)[7], canonical correlation analysis (CCA)[8], etc. These methods overcome the disadvantages of the first principle models, which are time-consuming and require in-depth process knowledge. As an extension of PLS algorithm, multi-way partial least squares (MPLS)[9] was proposed for batch processes and it was commonly regarded as a milestone in quality prediction. In MPLS, three-dimensional process data matrix is unfolded batch-wise, which keeps the dimension of batch direction unchanged, and then a prediction model is developed by performing PLS between the batch-wise unfolded process data and corresponding quality data. However, there are two obvious drawbacks in MPLS as pointed out in previous work[10-12]. On the one hand, the entire batch-wise unfolded process data are employed for process modeling so that the prediction accuracy heavily relies on the estimation of unavailable future data for the online purpose. On the other hand, it neglects the multiphase characteristics of batch processes, which brings the difficulty for process understanding and may lead to inaccurate prediction. Sequentially, a series of improvements[13-15] have been developed. In order to avoid estimating future data, Randolf et al.[13] suggested the separation of time-slice loading and weight coefficients from batch-wise unfolding based model for online prediction. Aiming at the same problem, Chiu et al.[14] proposed a method by performing elastic net on the batch-wise unfolded process data and quality data. The values of calculated coefficients are used to evaluate the importance of sampling times and process variables. However, batch processes, in general, operate in a sequence of physical phases and each phase may have its specific characteristic[15]. It is noticed that process variable correlations, as well as the influences on quality, keep similar within a phase while may have a significant difference between different phases. However, the above mentioned methods[13, 14] treat entire batch as a single subject without exploring the changes of process characteristics over different phases.

As an important feature of batch processes, the multiphase characteristic has attracted increasing attention[16-26]. A class of phase partition methods[16-21] based on expert knowledge or process analysis were proposed, such as indicator variable method[16], multiblock modeling technique[17] etc. In order to overcome the dependence on process knowledge and partition the phases automatically, data-driven phase partition algorithms have been paid special attention[22-30]. Lu et al.[22] put forward the sub-PCA algorithm by clustering those sampling points that have similar variable correlations into one phase. Sequentially, sub-PLS[23] was proposed for online quality prediction. Considering the between-phase transition patterns, the soft-transition multiple PCA method[24] and an angle based phase partition method[25] were proposed. For these methods, variable correlations or process correlations are clustered by k-means algorithm, which does not take the time sequential property into consideration. Therefore, initial phase partition results may be discontinuous, which bring heavy burdens of post-processing. Camacho et al.[26] recursively divided the batch cycle into phases at the points where the prediction error achieves minimum. Yu et al.[27] employed Gaussian mixture model to cluster the sampling times into different classes. As pointed by Zhao et al.[24], the influences on monitoring performance should also be considered during phase partition as well as changes of variable correlations. Thus, a step-wise sequential phase partition algorithm was proposed for fault detection by checking the changes of monitoring statistics. A quality-relevant sequential phase partition method (QSSPP) was also developed for quality prediction[28]. However, the results of phase partition are greatly influenced by a tunable parameter, i.e. relaxing factor. Besides, the phase-based prediction model isolates the influences of each time on product quality without considering the cumulative quality effect. Influences of process variations on quality are in general increased cumulatively with time evolution which is termed as cumulative quality effect. Zhao et al.[29] used each phase as the basic analysis object from which the local contributions of

different phases are stacked to explore the cumulative quality effect. However, they did not evaluate the significance of different phases. In fact, the significance of different phases is not necessarily in accord with their operation sequence. Uncritical phases should be removed from model development to avoid introducing undesirable disturbances into the model if they have no significant influences on product quality. Zhao et al.[30] treated each phase as a single block and analyzed their priority for quality interpretation by checking their different contributions to qualities. However, both work[29, 30] did not explore how the cumulative effects change time-wise within the same phase. Besides, they did not consider the problem of online application by using the measurement of the entire phase.

From the above analysis, the cumulative quality effect has not been well analyzed time-wise within each phase and explored for online quality prediction. Besides, several problems are noticed. First, phase partition results of QSSPP[28] are subject to a tunable parameter (relaxing factor), which directly leads to uncertainty of phase partition. Second, the process variations at each time present both similar and dissimilar influences on quality in comparison with previous variations. They explore different quality information and should be well decomposed and separated from each other for analysis. Third, different phases may contribute differently to cumulative quality effects in which only the critical-to-cumulative-quality phases can reliably provide cumulative quality information and thus should be separated from those uncritical ones. In order to solve the above mentioned problems, a subspace decomposition and critical phase selection method is proposed for cumulative quality analysis for multiphase batch processes. First, multiple phases are separated from the perspective of quality analysis in which the tunable parameter is determined by using the fast search and find of density peaks clustering (FSFDP) algorithm[31]. Second, after phase partition, the measurement space at each sampling time is decomposed into several parts to explore their different influences on quality. Here, a quantitative index, termed non-repetitive quality-relevant information (NRQRI), is defined to

explore the quality-relevant data variations at each time that are orthogonal to previous variations. This part of variations can explain completely different quality variation in comparison with previous process variations, from which the qualities will thus be interpreted accumulatively along time direction. Third, a recursive critical phase identification strategy is developed from the perspective of cumulative quality analysis which can sort critical phases that are of significant cumulative influences on quality and then the final cumulative quality prediction is made by integrating the results from critical phases.

The proposed algorithm has the following advantages:

(1) It can get sequential quality-relevant phase partition results and properly determine the value of tunable parameters to avoid the uncertainty of phase partition results.

(2) It can judge the type of quality index in advance to determine whether it is of cumulative type and properly decompose the data space at each time to probe into the variation that contributes to cumulative quality effect.

(3) It focuses on the critical information for cumulative quality analysis and excludes the influences of unimportant phases by identifying critical phases from the whole batch run.

The remainder of this paper is organized as follows. The details of the proposed method are described in Section 2, including quality-relevant sequential phase partition, time-wise subspace decomposition and the selection of critical phases for cumulative quality analysis. In Section 3, the feasibility and effectiveness of the proposed method are illustrated through a typical multiphase batch process, injection molding process. Conclusions are drawn in the last section.

2. Methodology

2.1 A quality-relevant sequential phase partition method

Here, QSSPP[28] is used as the basis for quality-relevant phase partition. Besides, the undesirable influences of relaxing factor on phase partition in QSSPP method is overcome by

employing the fast search and find of density peaks clustering algorithm (FSFDP). In fact, in our previous work[32], the adjustment of the relaxing factor has already been studied and the concrete steps are given in Appendix for readability. The main difference between this work and the previous one is that the process is partitioned from the perspective of quality analysis here instead of fault detection. Before introducing the phase partition algorithm, data preparation is needed.

In each batch run, assume that J_x process variables are measured online at $k = 1, 2, \ldots K$ time instances throughout the operation cycle and J_y quality variables are obtained offline. After I batches, a three-dimensional array of process data $\mathbf{X}(I \times J_x \times K)$ and a two-dimensional quality matrix $\mathbf{Y}(I \times J_y)$ are formed, as shown in Fig. 1. Split the data array $\mathbf{X}$ along the time dimension and K time-slice data matrices $\mathbf{X}_k(I \times J_x)$ are obtained, where subscript k is the index of sampling time. Then, each time-slice data matrix $\mathbf{X}_k$ and $\mathbf{Y}$ form the time-slice regression data pair $\{\mathbf{X}_k, \mathbf{Y}\}$ $k \in [1, K]$. $\{\mathbf{X}_k, \mathbf{Y}\}$ is the basic analysis unit, which covers multiple batches observed at each sampling time. Further, time-slice regression model refers to a model developed by $\{\mathbf{X}_k, \mathbf{Y}\}$ using basic PLS algorithm. The variables at each time are then preprocessed to have zero mean and standard deviation and the normalized data are still noted as $\{\mathbf{X}_k, \mathbf{Y}\}$ for simplicity.

In the present work, an assumption is that all batches are synchronized, and key features occur at the same time. Therefore, observations with the same time stamp are comparable and can be used for model development. If the process data are equal length, the proposed method can be directly used. On the contrary, several methods for batch trajectory synchronization are needed as pretreatments to solve the uneven length problem, such as instrumental variable[33], dynamic time warping[34], and multisynchro[35]. Instrumental variable method is suitable

when indicator variables that indicate process progress instead of process time are available. If such process knowledge is not available, dynamic time warping can be used when the beginning and end points of phases in the concerned batches are similar. Multisynchro method is suggested when uneven-length problem is complicated in which the key process events may occur at the different time of process evolution. Therefore, this method can achieve better performance for unequal batch processes with multiphase characteristics.

The specific phase partition procedure with adjustment of relaxing factor is described as below.

Step 1: Time-slice based PLS modeling

Perform PLS algorithm on $\{\mathbf{X}_k, \mathbf{Y}\}\ k \in [1, K]$ and the regression coefficients of each time-slice regression model are $\boldsymbol{\theta}_k$. Evaluate the prediction performance of time-slice regression model at time k by the index of root mean square error (RMSE), which is calculated as follows,

$$\begin{aligned} \hat{\mathbf{Y}}_k &= \mathbf{X}_k \boldsymbol{\Theta}_k \\ \mathrm{RMSE}_k &= \sum_{i=1}^{I} \sum_{j=1}^{J_y} \left(\hat{y}_{k,i,j} - y_{i,j} \right)^2 \end{aligned} \tag{1}$$

where $\hat{y}_{k,i,j}$ is the prediction of the j^{th} quality index of the i^{th} batch at time k which is the i^{th} row and j^{th} column of $\hat{\mathbf{Y}}_k$, $y_{i,j}$ is the real value of the j^{th} quality variable of the i^{th} batch.

Step 2: Time-segment based PLS modeling

From the beginning of batch process, add the next time-slice data matrix to the existing ones and unfold them variable-wise within the concerned time region, i.e. $\mathbf{X}_{v,k}\left(Ik \times J_x\right)$, where subscript v indicates variable-wise data unfolding. Then, arrange the quality data $\mathbf{Y}_k\left(Ik \times J_y\right)$ in the similar way by duplicating the quality matrix $\mathbf{Y}\left(I \times J_y\right)$ to have the same row dimension as $\mathbf{X}_{v,k}$. Perform PLS on $\left\{\mathbf{X}_{v,k}, \mathbf{Y}_k\right\}$ and obtain the regression coefficient $\boldsymbol{\theta}_{v,k}$. Quality can then

be predicted for $\mathbf{X}_k$ at each time up to k using the time-segment model ($\boldsymbol{\theta}_{v,k}$). At each time, the time-wise RMSE index defined in Eq. (2) is then updated as $\mathrm{RMSE}_{v,k}$ which can evaluate the performance of the time-segment model.

$$\begin{aligned} \hat{\mathbf{Y}}_{v,k} &= \mathbf{X}_k \boldsymbol{\theta}_{v,k} \\ \mathrm{RMSE}_{v,k} &= \sum_{i=1}^{I} \sum_{j=1}^{J_y} \left(\hat{y}_{v,k,i,j} - y_{i,j} \right)^2 \end{aligned} \tag{2}$$

where $\hat{y}_{v,k,i,j}$ is the prediction value of the jth quality index of the ith batch at time k using time-segment model which is the element of $\hat{\mathbf{Y}}_{v,k}$.

Step 3: Comparison of prediction error

Compare RMSE_k with $\mathrm{RMSE}_{v,k}$ within the concerned time region. If $\mathrm{RMSE}_{v,k} < \alpha \mathrm{RMSE}_k$ in which α is the relaxing factor reflecting the loss tolerance of prediction accuracy of time-segment in comparison with time-slice models, it means that time-segment model has the similar prediction performance with time-slice model. Update k as k+1 and repeat Steps 2 and 3. Otherwise, if there are three consecutive samples satisfying the relationship $\mathrm{RMSE}_{v,k} \geq \alpha \mathrm{RMSE}_k$ from time k, it means the prediction relationship has significantly changed. The current time k is denoted as the phase partition time k^* from which we can initially get a new phase. To avoid the influence of relaxing factor, go to Step 4 to search for the optimal value of α.

Step 4: The updating of relaxing factor

The tunable parameter (α) has a great influence on the phase partition results. If the value of α is large, the time-segment model is allowed to be less accuracy as calculated by $\mathrm{RMSE}_{v,k} \geq \alpha \mathrm{RMSE}_k$ and more time-slice data matrixes will be allocated into the same phase. On the contrary, if the value of α is small, the process correlations will be separated into more different phases. Therefore, its value should be properly determined to avoid the uncertainty.

The FSFDP algorithm[31] shown in Appendix is adopted here to determine whether the current α can be increased. The initial value of α is set to be one. Regression coefficient $\boldsymbol{\theta}_k$ ($k \in [1, k^*]$) that presents process correlations at each time is employed as the analysis object. The number of clusters is determined based on two indices, density and distance where cluster centers are recognized as the points that have anomalously large distance and a high local density. More than one cluster reveals that different process correlations are presented in the time-segment. If there is only one cluster is identified, it reveals that the prediction relationships stay the same within the concerned time-segment. Increase the value of relaxing factor by $\alpha=\alpha+\Delta$ where the parameter Δ is the step length. Go back to Step 2 and update the phase partition result. Repeat the Steps 2 through 4 until the number of clusters is larger than one which means different prediction relationships are covered in the current time segment. Thus, the previous value of α should be chosen, denoted as α^*, and the time-slices before k^* are assigned to a phase.

Step 5. Data updating and recursive implementation

Remove the first phase and employ the remaining time-slice data as the new input. Recursively repeat Steps 2 through 5 to find the remaining phases and the corresponding values of α^*.

Based on the proposed phase partition method, multiple phases are separated from the batch cycle and the proper value of α is determined for each phase. It is noted that for different phases, the relaxing factor may have different values.

2.2 Subspace decomposition for phase-wise cumulative quality analysis

Considering different phases may act differently and cumulatively on quality variability, the phase-based cumulative analysis can help to improve process understanding and capture how these phases contribute to the quality variability. Based on phase partition results, cumulative quality analysis is first implemented within each phase to explore different cumulative quality effects over different phases. Here, a subspace decomposition algorithm is proposed to explore

the process variations of each time-slice data matrix based on the following recognition: (1) the process variations at each time present time-varying influences on quality although they have similar quality effects; (2) the time-varying quality-relevant variations reveal different influences on quality which should be well extracted and used for cumulative quality analysis. The non-repetitive quality-relevant information (termed as NRQRI), is extracted at each time to explore the time-varying influences on quality. This part of variations can explain different quality information in comparison with previous process variations, from which the quality will thus be interpreted accumulatively along time direction. This part of information can also be used to indicate the quality type. If NRQRI is zero, the quality-relevant process variations are quite similar with each other and present similar quality interpretability which do not tell any new quality information along time direction. Otherwise, the process variations can explain new quality information with time evolution and thus the quality index is cumulative.

To extract NRQRI, time-slice data matrix $\mathbf{X}_{c,k}$ is divided into four different subspaces in each phase using the proposed subspace decomposition algorithm as shown in Fig. 2, where the subscript c indicates the phase number. First, PLS algorithm is performed on each regression data pair $\{\mathbf{X}_{c,k}, \mathbf{Y}\}$ to initially extract the quality-relevant variation ($\hat{\mathbf{X}}_{c,k}$) from the time-slice data space at each time. Second, similar quality-relevant variations ($\mathbf{X}_{s,c,k}$) are extracted within each phase by data reconstruction, which imposes repetitive influences on quality. Third, the time-specific variations ($\mathbf{X}_{d,c,k}$) that reveals dissimilar influences on quality are explored by removing the similar variations ($\mathbf{X}_{s,c,k}$) from $\hat{\mathbf{X}}_{c,k}$, from which, the final quality-relevant part is extracted by post-processing PLS results using CCA algorithm. The specific subspace decomposition is described below.

(1). Extraction of initial quality-relevant variation

Regress $\mathbf{X}_{c,k}$ against $\mathbf{Y}$ using PLS algorithm to develop time-slice regression model.

$$\begin{aligned}\mathbf{T}_{c,k} &= \mathbf{X}_{c,k}\mathbf{R}_{c,k} \\ \mathbf{P}_{c,k} &= (\mathbf{T}_{c,k}{}^{T}\mathbf{T}_{c,k})^{-1}\mathbf{T}_{c,k}{}^{T}\mathbf{X}_{c,k} \\ \hat{\mathbf{X}}_{c,k} &= \mathbf{T}_{c,k}\mathbf{P}_{c,k}{}^{T} \\ \mathbf{X}_{c,k} &= \hat{\mathbf{X}}_{c,k} + \mathbf{E}_{c,k}\end{aligned} \tag{3}$$

where $\mathbf{R}_{c,k}(J_x \times R_c)$ is the weight matrix to calculate latent variables(LVs) from $\mathbf{X}_{c,k}$, $\mathbf{T}_{c,k}\left(I \times R_c\right)$ is the matrix containing the LVs which are extracted from the process variables and used to explain qualities at each time, $\mathbf{P}_{c,k}(J_x \times R_c)$ is the loading matrix for $\mathbf{X}_{c,k}$, and R_c is the retained number of latent variables.

For PLS algorithm, the initial quality-relevant variations ($\hat{\mathbf{X}}_{c,k}$) are extracted from the time-slice data space ($\mathbf{X}_{c,k}$) and the residual matrix is $\mathbf{E}_{c,k}$.

(2). Extraction of similar quality-relevant variations

Similar process variations describe the repetitive prediction information at the current time in comparison with the previous process variations before time k which can be extracted from the initial quality-relevant part $\hat{\mathbf{X}}_{c,k}$.

Use $\mathbf{T}_{c,k}$ to explain previous quality prediction $\hat{\mathbf{Y}}_{c,k-1}$ using the ordinary least square (OLS):

$$\begin{aligned}\mathbf{P}^{*}_{c,k}{}^{T} &= (\mathbf{T}_{c,k}{}^{T}\mathbf{T}_{c,k})^{-1}\mathbf{T}_{c,k}{}^{T}\hat{\mathbf{Y}}_{c,k-1} \\ \hat{\mathbf{Y}}^{*}_{c,k-1} &= \mathbf{T}_{c,k}\mathbf{P}^{*}_{c,k}{}^{T}\end{aligned} \tag{4}$$

where $\mathbf{P}^{*}_{c,k}$ is the regression coefficient that used for calculating previous quality prediction that can be explained at the current time.

$\hat{\mathbf{Y}}^{*}_{c,k-1}$ represents the quality information that can be explained by process variations at both times, k-1 and k, revealing repetitive quality-relevant variations. Then perform PCA on $\hat{\mathbf{Y}}^{*}_{c,k-1}$ to get orthogonal scores ($\hat{\mathbf{T}}^{*}_{c,k-1} = \hat{\mathbf{Y}}^{*}_{c,k-1}\mathbf{P}_{l,c,k}$) with all principal components with nonzero

variance retained. The loading matrix is noted as $\mathbf{P}_{l,c,k}$, where the subscript l indicates the LV decomposition.

The scores $\hat{\mathbf{T}}^{*}_{c,k-1}$ are then used to explain the repetitive part of the initial quality-relevant variation ($\hat{\mathbf{X}}_{c,k}$):

$$\begin{aligned}\mathbf{P}_{s,c,k}{}^{T} &= (\hat{\mathbf{T}}^{*}_{c,k-1}{}^{T}\hat{\mathbf{T}}^{*}_{c,k-1})^{-1}\hat{\mathbf{T}}^{*}_{c,k-1}\hat{\mathbf{X}}_{c,k} \\ \mathbf{X}_{s,c,k} &= \hat{\mathbf{T}}^{*}_{c,k-1}\mathbf{P}_{s,c,k}{}^{T}\end{aligned} \tag{5}$$

where $\mathbf{X}_{s,c,k}$ reveals the repetitive information in $\hat{\mathbf{X}}_{c,k}$ that exists in previous process variations described by $\hat{\mathbf{X}}_{c,k-1}$, $\mathbf{P}_{s,c,k}$ is the coefficient used to regress the similar influence between the initial quality-relevant prediction process variations and the prediction value $\hat{\mathbf{Y}}^{*}_{c,k-1}$, and the subscript s is used to indicate the similar quality-relevant influence.

(3). Extraction of dissimilar quality-relevant variations

Subtracting $\mathbf{X}_{s,c,k}$ from $\hat{\mathbf{X}}_{c,k}$, the remaining part is denoted as $\mathbf{X}_{n,c,k}$, which does not contain any repetitive information in comparison with process variations before k and the subscript n stands for the non-repetitive quality-relevant information. Considering that PLS scores may cover quality-irrelevant variations since it maximizes the covariance[7], PLS-CCA algorithm[8] is employed to extract the final quality-relevant part from $\mathbf{X}_{n,c,k}$ according to Eq.(6).

$$\begin{aligned}\mathbf{T}_{n,c,k} &= \mathbf{X}_{n,c,k}\mathbf{R}_{n,c,k} \\ \mathbf{T}_{d,c,k} &= \mathbf{T}_{n,c,k}\mathbf{W}_{d,c,k}\end{aligned} \tag{6}$$

where $\mathbf{R}_{n,c,k}$ is the weight matrix to calculate LVs from $\mathbf{X}_{n,c,k}$ by using PLS algorithm, $\mathbf{T}_{d,c,k}$ is the score matrix used to explain the process variations and $\mathbf{W}_{d,c,k}$ is the weight matrix to extract the strict quality-relevant process variations by using CCA algorithm, respectively, and the subscript d presents the dissimilar quality-relevant influence.

Then $\mathbf{T}_{d,c,k}$ is used to explain $\mathbf{X}_{n,c,k}$ by using OLS algorithm.

$$\begin{aligned}\mathbf{P}_{d,c.k}{}^{T} &= (\mathbf{T}_{d,c.k}{}^{T}\mathbf{T}_{d,c.k})^{-1}\mathbf{T}_{d,c.k}{}^{T}\mathbf{X}_{n,c,k} \\ \mathbf{X}_{d,c,k} &= \mathbf{T}_{d,c,k}\mathbf{P}_{d,c.k}{}^{T}\end{aligned} \tag{7}$$

where $\mathbf{X}_{d,c,k}$ denotes dissimilar quality-relevant variations in comparison with its previous time-slice that tell the NRQRI.

(4). Cumulative quality prediction

Up to the current time k, the cumulative process matrix is obtained by $\mathbf{C}_{c,k} = [\mathbf{T}_{p,c,k-1}, \mathbf{T}_{d,c,k}]$ and CCA algorithm is then used for cumulative quality prediction as follows:

$$\begin{aligned}\mathbf{T}_{p,c,k} &= \mathbf{C}_{c,k}\mathbf{W}_{p,c,k} \\ \mathbf{Q}_{p,c,k} &= (\mathbf{T}_{p,c,k}{}^{T}\mathbf{T}_{p,c,k})^{-1}\mathbf{T}_{p,c,k}{}^{T}\mathbf{Y} \\ \hat{\mathbf{Y}}_{c,k} &= \mathbf{T}_{p,c,k}\mathbf{Q}_{p,c,k}{}^{T}\end{aligned} \tag{8}$$

where $\mathbf{W}_{p,c,k}$ is the weight matrix used to derive the quality-relevant process information ($\mathbf{T}_{p,c,k}$) from the cumulative process matrix by using CCA algorithm and subscript p stands for prediction, and $\mathbf{Q}_{p,c,k}$ is the regression coefficient to calculate the final cumulative quality prediction $\hat{\mathbf{Y}}_{c,k}$. It is easy to derive that $\mathbf{T}_{d,c,k}$ is orthogonal with $\mathbf{T}_{p,c,k-1}$ which thus tells the non-repetitive LVs in comparison with previous time.

The above procedure is implemented at each time until it arrives at the end of the current phase. In each phase, to get the cumulative quality prediction, the key points are to derive NRQRI and continuously combine it with this part of information at the previous time. In this way, only the new quality information will be further explained from k-1 to k. The phase-wise cumulative quality prediction is thus obtained at the end of each phase. At the end of each phase, the quality-relevant process information ($\mathbf{T}_{p,c,K_c}$) is available to describe the cumulative effects on quality of each phase.

It is recognized that different phases may have different significances and only some of them are critical to cumulative quality prediction. The effects of uncritical phases should be removed from model development to avoid undesirable disturbance since they have no significant influences on product quality. Index $R^2_{c,k}$ is employed to evaluate the prediction accuracy at each time within each phase,

$$R^2_{c,k} = \frac{\sum_{i=1}^{I}\left(\hat{y}_{i,c,k} - \bar{y}\right)^2}{\sum_{i=1}^{I}\left(y_i - \bar{y}\right)^2} \quad k \in [1, K_c] \tag{9}$$

where $\hat{y}_{i,c,k}$ is the prediction of the i^{th} batch in the c^{th} phase at time k, y_i is the real value of the i^{th} batch, $\bar{y}$ is the mean value over batches and K_c is the total sampling times in phase c.

$R^2_{c,k}$ is an index to measure the prediction accuracy at time k and it varies from zero to one. Within the normal region, a large value of $R^2_{c,k}$ indicates high prediction accuracy. Besides, calculate the index $\Delta R^2_c = R^2_{c,K_c} - R^2_{c,1}$ which presents the changes of prediction accuracy in phase c under the cumulative quality analysis, where $R^2_{c,1}$ denotes the prediction accuracy at the phase beginning and R^2_{c,K_c} denotes the prediction accuracy at the end of the phase. If $\Delta R^2_c > 0$, it means that there are time-varying process variation in Phase c to improve prediction value and cumulative quality effect exists in this phase. Otherwise, there are no process variations related to cumulative quality effect in phase c. R^2_{c,K_c} and ΔR^2_c are combined to evaluate the importance of each phase and judge the critical phases. If R^2_{c,K_c} is large than the predefined threshold value and $\Delta R^2_c > 0$, this phase is determined to be a critical-to-cumulative-quality phase. Otherwise, it is not critical and should be removed from

the following cumulative analysis. Here, the threshold value for R^2_{c,K_c} is defined using the F-statistics[23] with different significance factors.

2.3 Process-wise cumulative quality analysis

Based on the phase-wise cumulative analysis, critical phases have been identified in which $\mathbf{T}_{p,c,K_c}$ describes the phase-wise cumulative effects and R^2_{c,K_c} reveals the cumulative phase prediction accuracy for the concerned phase. A strategy that can combine the cumulative effects of different critical phases will have to be developed. For the critical-to-cumulative-quality phases that have been identified in Subsection 2.2, the following cumulative phase fusion strategy is described as below:

Step (1): The first critical phase is used as the initial analysis unit. Extract the non-repetitive quality-relevant information (NRQRI) from the next critical phase using the same subspace decomposition algorithm as that in Subsection 2.2 in which the analysis unit is the quality-relevant process information ($\mathbf{T}_{p,c,K_c}$) instead of the time-slice used in phase-wise cumulative analysis. The quality prediction accuracy is evaluated by R^2_f for the integrated critical phases and the fused quality-relevant process information ($\mathbf{T}_f$) are extracted using Eq. (8) revealing the quality-relevant process information from the cumulative process matrix.

Step (2): Compare R^2_f with R^2 index for the concerned two critical phases using the following judgment rule:

Case I: If R^2_f is larger than both values, it means that the two critical phases can be combined to improve quality prediction. The combined phase information should replace the initial analysis unit with the fused quality-relevant process information ($\mathbf{T}_f$).

Case II: If R^2_f is smaller than R^2 index of the second critical phase but larger than R^2 index of the first critical phase, the performance of cumulative quality analysis after phase fusion is not

as good as that of the second critical phase. The two phases should not be fused and the second critical phase should replace the initial analysis unit.

Step (3): Update the initial analysis unit and repeat the above two steps until all critical phases have been considered.

The output is the selected critical phases that should be integrated for process-wise cumulative analysis to get the final quality prediction. Besides, how to combine these phases are also determined in which the subspace decomposition is performed for the selected phases using the quality-relevant process information instead of time-slice analysis unit to integrate the cumulative effects.

It is noted that non-critical phases are not used for process-wise cumulative quality analysis here. For the processes that are dominated by critical phases, the proposed method can be directly used for cumulative quality prediction. However, for other processes, all phases may contribute to final product, that is, operation at a succeeding phase may remedy the run and final product. Without changing the basic idea, the proposed method can be readily extended to such processes by employing all phases for process-wise cumulative quality analysis instead of only considering the critical phases.

2.4 Online cumulative quality prediction

Online cumulative quality prediction is only made in each critical phase by adopting the predefined cumulative quality prediction models. Here, the normalized new sample at the k^{th} sampling time of one critical-to-cumulative-quality phase (c) is denoted as $\mathbf{x}_{c,k}(J_x \times 1)$. The specific steps are given below:

a) Calculate the initial quality-relevant process variation $\hat{\mathbf{x}}_{c,k}$

$$\begin{aligned} \hat{\mathbf{x}}_{c,k} &= \mathbf{t}_{c,k}\mathbf{P}_{c,k}^{T} \\ \mathbf{t}_{c,k} &= \mathbf{x}_{c,k}\mathbf{R}_{c,k} \end{aligned} \tag{10}$$

where $\mathbf{R}_{c,k}$ is the weight matrix to calculate LVs and $\mathbf{P}_{c,k}{}^{T}$ is the loading matrix and they are given in Eq.(3). Then the scores $\mathbf{t}_{c,k}$ will be used in Step (b).

b) Explain the quality prediction $\hat{\mathbf{y}}_{c,k-1}$ using $\mathbf{t}_{c,k}$ as follows

$$\hat{\mathbf{y}}^{*}_{c,k-1} = \mathbf{t}_{c,k}\mathbf{P}^{*}_{c,k}{}^{T} \tag{11}$$

where $\mathbf{P}^{*}_{c,k}$ is the regression coefficient in Eq.(4).

c) Reconstruct the similar quality-relevant process variations $\mathbf{x}_{s,c,k}$

$$\mathbf{x}_{s,c,k} = \hat{\mathbf{y}}^{*}_{c,k-1}\mathbf{P}_{l,c,k}\mathbf{P}_{s,c,k}{}^{T} = \hat{\mathbf{t}}^{*}_{c,k-1}\mathbf{P}_{s,c,k}{}^{T} \tag{12}$$

where $\mathbf{P}_{l,c,k}$ is the loading matrix used to extract the main variations of $\hat{\mathbf{y}}^{*}_{c,k-1}$, $\mathbf{P}_{s,c,k}$ is the coefficient used to regress the similar influence between the initial quality-relevant prediction process variations which is given in Eq. (5).

d) Subtract $\mathbf{x}_{s,c,k}$ from $\hat{\mathbf{x}}_{c,k}$, the remaining part is denoted as $\mathbf{x}_{n,c,k}$;

e) The non-repetitive quality-relevant information (NRQRI) is calculated below,

$$\begin{aligned} \mathbf{t}_{n,c,k} &= \mathbf{x}_{n,c,k}\mathbf{R}_{n,c,k} \\ \mathbf{t}_{d,c,k} &= \mathbf{t}_{n,c,k}\mathbf{W}_{d.c,k} \\ \mathbf{x}_{d,c,k} &= \mathbf{t}_{d,c,k}\mathbf{P}_{d,c.k}{}^{T} \end{aligned} \tag{13}$$

where $\mathbf{R}_{n,c,k}$ is the weight matrix to calculate LVs from $\mathbf{x}_{n,c,k}$, $\mathbf{T}_{n,c,k}$ is the score matrix used to explain the process variations using PLS algorithm, and $\mathbf{W}_{d.c,k}$ is the weight matrix to extract the strict quality-relevant process variations using CCA algorithm. These coefficients are given in Eqs. (6) and (7).

f) Update the cumulative matrix as $\mathbf{c}_{c,k} = [\mathbf{t}_{p,c,k-1}, \mathbf{t}_{d,c,k}]$;

g) Online cumulative quality prediction is calculated,

$$\begin{aligned} \mathbf{t}_{p,c,k} &= \mathbf{c}_{c,k}\mathbf{W}_{p,c,k} \\ \hat{\mathbf{y}}_{c,k} &= \mathbf{t}_{p,c,k}\mathbf{Q}_{p,c,k}{}^{T} \end{aligned} \tag{14}$$

where $\mathbf{W}_{p,c,k}$ is the weight matrix used to compress the quality-relevant information into score matrix and $\mathbf{Q}_{p,c,k}$ is the regression coefficient to calculate the final cumulative quality prediction $\hat{\mathbf{y}}_{c,k}$ that has been given in Eq. (8).

3. Illustration results

3.1 Process description

In this section, the efficiency of the proposed algorithm is illustrated through injection molding (IM), which is a typical multiphase batch process. A complete production cycle of IM involves the stages of mold-close, filling, packing-holding, plastication, cooling and mold-open. Through process repetitions, IM produces various quantities of plastic products and it plays a significant role in polymer processing for the manufacturing industry. Before production, the IM machine will be pre-heated to provide a high-temperature environment for barrel, which transforms raw material into melt state. Mold closes during mold-close stage and plastic melt is injected into the mold cavity in filling stage. After the mold is filled with the plastic melt, the packing-holding stage is initiated. During this stage, molten flow is prevented out of the mold and the additional material is compacted into the mold to make up the shrinkage associated with cooling. Sequentially, the molten polymer is conveyed to the front of the barrel by screw rotation in plastication stage, which is the initial period of cooling stage, during which the material is cooled inside the mold until it is rigid enough to be ejected. Finally, the mold is opened and the plastic product is taken out[36].

Process data of IM can be online collected from the transducer placed in IM machine and mold. Quality indexes, such as part weight, are readily measured offline at the end of each batch. A series of studies have been conducted on IM to prove their effectiveness[37] and IM is regarded as an ideal object for application and validation of the proposed quality-relevant phase division and quality analysis strategy. Therefore, it is convenient to analyze the results

using IM because prior process knowledge can be used as aids and the effectiveness of the proposed algorithm is easy to be proved.

In this study, eleven process variables are measured as predictor variables, which are given in Table 1. The quality index is part weight, which is offline measured at the end of each batch using a high precision electronic balance with a resolution of 0.01g. The material used in the experiments is high-density polyethylene (HDPE). In total, eighty-four normal batches are carried out for modeling. The batches are of even duration (634 samples in this experiment) and the process variables data are $\mathbf{X}(84\times11\times634)$. Correspondingly, the quality data are $\mathbf{Y}(84\times1)$. Fifty of them are used as training data for phase partition modeling and prediction model is developed. The next twenty of them are used as validation data to choose the best parameters for prediction model, and the remaining batches are employed as testing data to evaluate the prediction accuracy.

3.2 The Results of Phase Partition

Fifty batches are used as training data and the modeling data are process data $\mathbf{X}(50\times11\times634)$ and quality data $\mathbf{Y}(50\times1)$. By splitting the three-dimensional matrix along time direction, 634 time-slice data matrixes are formed. Normalize these time-slice data matrixes as well as the quality data and 634 time-slice data regression pairs are prepared. Then perform the PLS on each time-slice regression pair and time-slice prediction models are developed. The number of LVs at different times is determined by cross-validation to obtain the best quality prediction. The number of LVs used for the proposed phase partition at each sampling time is unified to be four which appears most.

In the proposed method, time-slice data regression pair is added variable-wise to form the time-segment data regression pair and develop regression models. At each time, the prediction errors of time-slice models and time segment models are compared, which are represented by

$RMSE_k$ and $RMSE_{v,k}$ respectively. The time segments are partitioned at the sampling point k, where the relationship $RMSE_{v,k} \geq \alpha RMSE_k$ is consecutively satisfied according to the rule given in Step 3 of the Subsection 2.1. However, the partition phase results greatly depend on the value of the tunable parameter, α . In the proposed phase partition method, a proper α can be determined using the adjustment strategy based on the FSFDP. Here, α is initialized to be one for convenience. After initial phase partition, α is adjusted by the step length (Δ=0.1) and the value of Δ yields an accurate phase partition results[32]. Finally, the phases are partitioned at the critical point where the number of cluster centers changes from one to more with the adjustment of the relaxing factor, suggesting that the process correlations change from one to more.

For the adjustment strategy based on the FSFDP, cluster center is identified as the point that simultaneously has a large distance and a large density[31]. Using the proposed method, it is observed that the number of cluster centers changes from one to two when the value of α changes from 1.9 to 2. In Fig. 3(a), two points have large distance in the decision graph for the first phase when α is 1.9. However, one of them (marked by red rectangles) has very small density values. Thus, only one point (marked by a red ellipse) simultaneously exhibits large distance and density values. Therefore, one cluster center is observed using the decision graph. If we increase α by 0.1, two points simultaneously exhibit large density and distance values in the decision graph of Fig. 3(b). Based on the rule of the phase partition procedure, the proper value of α is 1.9 for the first phase. Moreover, the index γ is also helpful to determine the cluster number, which is defined as the product of distance and density. Cluster centers have the largest γ values and will be far away from other data points. Comparing the values of γ in Figs. 3(a) and (b), the number of cluster centers changes from one to two, as marked by the red ellipse, which also indicates that 1.9 is the proper value of α. After the separation of the first phase, the remaining phase can be identified similarly, and the specific phase partition results

are presented in Fig. 4. Besides the first phase, the remaining process data are divided into five phases. In each phase, there is only one cluster center, which means covering one kind of process correlation. From the above analysis, the entire process is automatically partitioned into definite phases using the proposed method.

According to the proposed method, six phases are partitioned and the value of the relaxing factor is properly determined, as plotted in Fig. 5(a). In the previous work[32], the transition patterns between adjacent phases are analyzed by the index Tr_k. Index Tr_k is defined as ρ_k/δ_k, where ρ_k is the density of the current sampling time calculated in Eq. (A2), and δ_k denotes the distance in Eq. (A3). Transition patterns are identified as sampling times located at the phase edge and the value of index Tr_k is less than a predefined threshold (0.0001 is used here and a denary logarithm is adopted to make the values easy to see). In Fig. 5(b), several points around the edge of Phases 1, 2 and 3 are lower than the threshold. According to the previous analysis, transition patterns are identified if their values of Tr_k below the threshold and they locate in the edge of each phase. Therefore, quick transition regions exist in IM because only one or two transition patterns are observed between adjacent phases, which are not addressed in previous research works.

For comparison, the results of QSSPP algorithm[28] are presented in Fig. 6 where the partition results are influenced by the value of relaxing factor which however can not be properly determined. When the value of relaxing factor is small, more phases are partitioned. On the contrary, a larger value of relaxing factor will lead to fewer phases. Thus, the determination of the value of the relaxing factor is an important issue that is directly related to the phase partition results. Process knowledge is necessary for QSSPP algorithm to help the selection of this parameter when one wants to obtain reasonable results. However, it may be impractical for an unfamiliar process.

3.3 Cumulative quality analysis results

Phase-wise cumulative quality analysis is conducted as shown in Fig. 7. The number of retained LVs is determined by validation data and the best prediction accuracy is achieved when it is four. The prediction accuracy at each time is evaluated by the index $R^2_{c,k}$ given in Eq. (9). Combining the indexes R^2_{c,K_c} and ΔR^2_c, Phases 1, 2, 3 and 6 are judged to be critical-to-cumulative-quality phases. The threshold of R^2_{c,K_c} is 0.657 when the significant factor is 0.1[23]. The values of R^2_{c,K_c} are 0.7676 for Phase 1, 0.9562 for Phase 2, 0.8966 for Phase 3, and 0.9377 for Phase 6. The values of ΔR^2_c for the four phases are calculated as 0.0876 (Phase 1), 0.1903 (Phase 2), 0.0481 (Phase 3) and 0.2627 (Phase 6), respectively. All the values of ΔR^2_c are larger than zero, revealing that non-repetitive quality-relevant information (NRQRI) can be continuously extracted to improve the prediction accuracy and the quality index, i.e. part weight, is of cumulative-quality type. It is noted that the values of the index $R^2_{c,k}$ are larger than one after the 17^{th} sample in Phase 4 and the 75^{th} sample in Phase 5. Since the normal region of $R^2_{c,k}$ is from zero to one, the cumulative-quality prediction results become unreliable in Phases 4 and 5.

The critical phases are then fused using the proposed algorithm as shown in Fig. 8. First, the first two critical phases are combined in Fig. 8(a), revealing the cumulative-quality effect of the two phases. Using the judgment rule in Subsection 2.3, Phase 1 is excluded from phase integration since the prediction accuracy using the combined phase information satisfies the rule in Case II described in Subsection 2.3. Phase 2 is used for the following analysis. In Phases 8(b) and (c), the critical phase fusion is conducted for the remaining critical phases and Phases 2 and 3 are selected finally and should be fused for process-wise cumulative analysis. According to process knowledge, Phases 2 and 3 are in fact the filling stage and the

packing-holding stage, respectively. In the real case, high injection pressure and large ejector stroke (i.e., large packing pressure) will push more molten flow into the mold cavity and lead to heavy part weight. Thus, injection pressure (Variable #8) will give the most significant contribution in Phase 2 and ejector stroke (Variable #5) plays the most important role in Phase 3 which reveals the changes of packing pressure. That is, injection pressure in Phase 2 and ejector stroke in Phase 3 determine the quality variability. Reflecting the overall importance of each variable on all of the response variables based on principal components, variable importance in the projection(VIP)[38] method was employed to check the obtained critical-to-cumulative-quality phases. VIP method was initially used to select critical variables of which the average values of the squared VIP scores are large one[39]. Sequentially, VIP method has been used to identify important phases with MPLS algorithm. Similarly, if the values of squared VIP scores within a phase are large one, this phase is regarded as an importance phase. Therefore, with four principal components, VIP plots for Variables 5 and 8 are given by applying MPLS algorithm as shown in Fig. 9. In both Fig. 9 (a) and (b), values of squared VIP are large than one in Phases 2 and 3 presenting these two phases are important phases. On the contrary, Phases 1, 4, 5 and 6 are unimportant phases because the values of squared VIP are small than one in these phases. Thus, the critical-to-cumulative-quality phases of the proposed method agree well with the results from VIP method. In practice, filling stage (Phase 2) determines the amount of molten materials injected into the mold, which has a direct relationship with the part weight. Because of material shrinkage, packing-holding stage (Phase 3) will pack more molten flow into the mold. Thus, the results of the proposed method also meet well the process knowledge, which indicating the Phases 2 and 3 are critical phases.

Fourteen batches are used as testing data to show the online prediction performance in critical phases, including Phases 1, 2, 3 and 6. In Fig. 10, the prediction results of a batch are given in different phases. It is easy to see that prediction accuracy in Phase 2 as shown in Fig.

10(a) is continuously improved and the last value almost approaches the real value of the part weight. However, the variation of the prediction value is large. In Phase 3, as shown in Fig. 10(b), the variation of prediction is much smaller than that of Phase 2 and the prediction value keeps steady in the late period. The similar scenario can be observed for Phases 1 and 6. That is, by cumulative quality analysis, the quality prediction accuracy is improved with time evolution.

Besides, the process-wise cumulative analysis results are presented in Fig. 11 where Phases 2 and 3 that have been identified in Fig. 8 are fused to get the final quality prediction. For comparison, MPLS[9] and sub-PLS[19], are also used to get the end-of-batch quality prediction. In Fig. 11, the squared prediction errors are plotted for three methods and fourteen testing batches. The prediction error is the difference between the prediction value and the real value. It is observed that the proposed method yields smaller prediction errors in comparison with the other two methods. In contrast, MPLS presents very large errors for some testing batches, revealing worse generalization ability. It may result from the fact that it includes all phases for modeling which thus introduces some undesirable quality-irrelevant variations. For sub-PLS, it presents better accuracy than that of MPLS but worse accuracy than that of the proposed algorithm since the critical phases are simply summed with different weights.

Moreover, the proposed algorithm can help us to further understand the process. Here, different process variations are accounted. The first index $\mathrm{R}_{\mathrm{s,c,k}}$ is defined as below:

$$\mathrm{R}_{\mathrm{s,c,k}} = \frac{\sum \mathrm{trac}\left(\mathbf{X}_{\mathrm{s,c,k}}{}^{\mathrm{T}}\mathbf{X}_{\mathrm{s,c,k}}\right)}{\sum \mathrm{trac}\left(\hat{\mathbf{X}}_{\mathrm{c,k}}{}^{\mathrm{T}}\hat{\mathbf{X}}_{\mathrm{c,k}}\right)} \tag{15}$$

where $\hat{\mathbf{X}}_{\mathrm{c,k}}$ and $\mathbf{X}_{\mathrm{s,c,k}}$ is given in Eq. (3) and Eq. (5), respectively; the symbol trac indicates the calculation of the trace of a matrix. $\mathbf{X}_{\mathrm{s,c,k}}$ reveals the repetitive information in $\hat{\mathbf{X}}_{\mathrm{c,k}}$ that is

covered in the previous time and thus the index $R_{s,c,k}$ represents the ratio between the repetitive information and the total quality-relevant information at the sampling time k.

In a similar way, $R_{d,c,k}$ is also defined by replacing $\mathbf{X}_{s,c,k}$ by $\mathbf{X}_{n,c,k}$, which reveals the amount of dissimilar quality-relevant variations in $\hat{\mathbf{X}}_{c,k}$ in comparison with its previous time. That is, it tells the new cumulative quality information. In Fig. 12, the two indexes are calculated for Phases 2 and 3. Index $R_{s,c,k}$ decreases when the process enters into Phase 3 from Phase 2 which agrees with the real case. That is, Phase 3 has different influences on quality in comparison with Phase 2, revealing less similar information with that of Phase 2. In contrast, the index $R_{d,c,k}$ stays invariable throughout the two phases, revealing that there are dissimilar quality-relevant variations that can be extracted consecutively to further explain the quality information. Instead, using PLS algorithm, we can only evaluate the sum of $R_{s,c,k}$ and $R_{d,c,k}$ which can not reveal whether the process is cumulative and how the cumulative effect change with the process evolution for the final quality prediction.

4. Conclusions

In this work, a cumulative quality prediction method is proposed for batch processes. It provides definite phase division results which overcome the influence of tunable parameter. And the phase partition results are sequential within each batch and require no post-processing. Based on phase partition results, a mechanism for the analysis of cumulative quality effect is proposed. A subspace decomposition algorithm is developed to probe into the non-repetitive quality-relevant information from time-slice data matrix for model development. Phases are quantitatively classified into two types: the phases that are critical to cumulative quality effect and the ones that are not. The final quality prediction is thus conducted by integrating the critical phases. The proposed algorithm is illustrated by an important chemical engineering process, injection molding.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61422306 and No. 61433005), and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (ICT1600197).

Appendix

Adjustment of relaxing factor based on the FSFDP

The number of clusters is identified through the fast search and find of density peaks (FSFDP) clustering algorithm[27], which can check how many cluster centers are covered in the data obtained for the current α. Revealing the process correlations, the regression coefficients $\boldsymbol{\theta}_k$ $k \in [1, k^*]$ of time-slice regression data pair can be used as the basic analysis unit and the input of the clustering algorithm. The specifics are given below:

Step 1: Calculation of the distance matrix **D**

Input the regression coefficients $\boldsymbol{\theta}_k$ $k \in [1, k^*]$, and calculate the Euclidean distance d_{mn} between two process correlations $\boldsymbol{\theta}_m$ $m \in [1, k^*]$ and $\boldsymbol{\theta}_n$ $n \in [1, k^*]$:

$$d_{mn} = \sqrt{\sum_{r=1}^{R_k} \sum_{j=1}^{J_x} (\boldsymbol{\theta}_m(r, j) - \boldsymbol{\theta}_n(r, j))^2} \quad m, n \in [1, k^*] \tag{A1}$$

where k, m and n are the indexes of sampling times, k^* is the terminal time of the current phase, R_k is the retained number of scores, and J_x denotes the number of process variables. The distance matrix **D** is then formed by all of the distances d_{mn}

$$\mathbf{D} = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1k^*} \\ d_{21} & d_{22} & \cdots & d_{2k^*} \\ \vdots & \vdots & \vdots & \vdots \\ d_{k^*1} & d_{k^*2} & \cdots & d_{k^*k^*} \end{pmatrix}$$

where $d_{mn}=d_{nm}$.

Step 2: Calculation of the density ρ_k

Calculate the density of the kth time-slice regression data pair by

$$\rho_k = \sum_{m=1}^{k^*} e^{-(d_{km}/d_c)^2} \tag{A2}$$

where d_c is the cutoff distance and it can be determined according to the rules given in the FSFDP algorithm.

Step 3: Calculation of the distance δ_k

Based on the density value ρ_k, compute the distance (δ_k) of each time-slice data pair according to Eq. (A3).

$$\delta_k = \begin{cases} \min\limits_{n:\rho_n>\rho_k}(d_{kn}), \rho_k \neq \max\limits_{n=1,\ldots,N}(\rho_n) \\ \max(d_{kn}), \rho_k = \max\limits_{n=1,\ldots,N}(\rho_n) \end{cases} \quad \text{(A3)}$$

For the point with the highest density, δ_k is the maximum of d_{mn}. Otherwise, δ_m is the minimal distance from the points that have a higher density.

Step 4: Cluster center identification

Arrange the set data $\{\rho_k, \delta_k\} k \in [1, k^*]$ and the number of clusters can be identified in two way: 1) decision graph, which plots δ_m as a function of ρ_m for each point; 2) index $\gamma_m = \rho_m \delta_m$. In decision graph, cluster centers are recognized as the points that have anomalously large d_m and a high local density ρ_m. Ranking the index $\gamma_m = \rho_m \delta_m$ in decreasing order, when the value of γ_m becomes anomalously small from a rank order, the number of clusters is identified[33].

Step 5: Relaxing factor updating

If the number of clusters is one, increase the value of relaxing factor by $\alpha=\alpha+\Delta$ and output the number of cluster center and the new α, where the parameter Δ is the step length. On the contrary, output the number of clusters and the value of α without updating.

References

[1] Chen JH, Liu KK. On-line batch process monitoring using dynamic PCA and dynamic PLS models. Chemical Engineering Science. 2002; 57(1): 63-75.

[2] Choi SW, Morris J, Lee IB. Dynamic model-based batch process monitoring. Chemical Engineering Science. 2008; 63(3): 622-636.

[3] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987; 2: 37-52.

[4] Zhang SM, Zhao CH, Wang S, Wang FL, Pseudo time-slice construction using variable moving window-k nearest neighbor (VMW-kNN) rule for sequential uneven phase division and batch process monitoring. DOI: http://dx.doi.org/10.1021/acs.iecr.6b03743.

[5] Zhang Y, An J, Zhang H. Monitoring of time-varying processes using kernel independent component analysis. Chemical Engineering Science. 2013; 88: 23-32.

[6] Yao Y, Gao Furong. Subspace identification for two-dimensional dynamic batch process statistical monitoring. Chemical Engineering Science. 2008; 63(13): 3411-3418.

[7] Dayal BS, Macgregor JF. Improved PLS algorithm. Journal of Chemometrics. 1997; 11: 73-85.

[8] Yu HL, Macgregor JF. Post processing methods (PLS-CCA): simple alternative to preprocessing methods (OSC-PLS). Chemometrics and Intelligent Laboratory Systems. 2004; 73: 199-205.

[9] Nomikos P, MacGregor JF. Multi-way partial least squares in monitoring batch processes. Chemometrics and Intelligent Laboratory Systems. 1995; 30: 97-108.

[10] Xiao ZB, Wang HG. Sparse representation residual space analysis and its application to multimode batch process monitoring. Industrial & Engineering Chemistry Research. 2016; 55 (1): 187-196.

[11] Zhao CH, Wang FL, Mao ZZ, Lu NY, and Jia MX. Quality prediction based on phase-specific average trajectory for batch processes. AIChE Journal. 2008; 54(3): 693-705.

[12] Undey C, Cinar A. Statistic monitoring of multistage, multiphase batch processes. Chemometrics and Intelligent Laboratory Systems. 2002; 22: 40-52.

[13] Randolf R, Willy W, Robert W. Partial least square confidence interval calculation for industrial end-of-batch quality prediction. Chemometrics and Intelligent Laboratory Systems. 2010; 100: 75-82.

[14] Chiu CC, Yao Y. Multiway elastic net(MEN) for final product quality prediction and quality-related analysis of batch process. Chemometrics and Intelligent Laboratory Systems. 2013; 125: 153-165.

[15] Yao Y, Gao FR. Subspace identification for two-dimensional dynamic batch process statistical monitoring. Chemical Engineering Science. 2008; 63(13): 3411-3418.

[16] Undey C, Cinar A. Statistical monitoring of multistage, multiphase batch processes. IEEE Control Systems Magazine. 2002; 22: 40-52.

[17] Smilde, AK., Westerhuis JA., de Jong S. A framework for sequential multiblock component methods. Journal of Chemometrics. 2003; 17: 323-337.

[18] Reinikainen, SP, Hoskuldsson, A. Multivariate statistical analysis of a multistep industrial processes. Analytica Chinica Acta, 2007; 595: 248-256.

[19] Liu J., Wong DSH. Fault detection and classification for a two-stage batch process. Chemometrics, 2008; 22: 385-398.

[20] Facco P, Olivi M, Rebuscini C, Bezzo F, Barolo M. Multivariate statistical estimation of product quality in the industrial batch production of a resin. Dycops, 2007; 2: 93-98.

[21] Doan XT, Srinivasan R. Online monitoring of multi-phase batch processes using phase-based multivariate statistical process control. Computers and Chemical Engineering, 2008; 32: 230-243.

[22] Lu, N., Gao, F., & Wang, F. Sub-PCA modeling and on-line monitoring strategy for batch processes. AIChE Journal. 2004; 50: 255-259.

[23] Lu NY, Gao FR. Stage-based process analysis and quality prediction for batch processes. Industrial and Engineering Chemistry Research. 2005; 44: 3547-3555.

[24] Zhao CH, Wang FL, Lu NY, Jia MX. Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes. Journal of Process Control. 2007; 17: 728-741.

[25] Yao Y, Gao FR. Phase and transition based batch process modeling and online monitoring. Journal of Process Control. 2009; 19: 816-826.

[26] Camacho J, Pico J. Multi-phase principal component analysis for batch processes modelling. Chemometrics and Intelligent Laboratory System. 2006; 81: 127-136.

[27] Yu J, Qin SJ. Multiway Gaussian mixture model based multiphase batch process monitoring. Industrial and Engineering Chemistry Research. 2009; 48: 8585-8594.

[28] Zhao CH. A quality-relevant sequential phase partition approach for regression modeling and quality prediction analysis in manufacturing processes. IEEE Transaction on Automation Science and Engineering. 2013; 11(4): 983-991.

[29] Zhao CH, Wang FL, Gao FR. Improved calibration investigation using phase-wise local and cumulative quality interpretation and prediction. Chemometrics and Intelligent Laboratory Systems. 2009; 95: 107-121.

[30] Zhao CH, Gao FR. Multiphase calibration modeling and quality interpretation by priority sorting. Chemical Engineering Science. 2011; 66: 5400-5409.

[31] Rodriguez A, Laio A. Clustering by fast search and find of density peaks. Science. 2014; 344: 1492-1496.

[32] Qin Y, Zhao CH, Gao FR. An Iterative Two-Step Sequential Phase Partition (ITSPP) Method for Batch Process Modeling and Online Monitoring. AIChE Journal. 2016, DOI 10.1002/aic.15205.

[33] Wasito I, Mirkin B. Nearest neighbour approach in the least-squares data imputation algorithms. Information Sciences. 2005; 169: 1-25.

[34] Kassidas A, MacGregor JF, Taylor PA. Synchronization of batch trajectories using dynamic time warping. AIChE Journal. 1998; 44: 864-875.

[35] González-Martínez JM, Noord OE, Ferrer A. Multisynchro: a novel approach for batch synchronization in scenarios of multiple asynchronisms. Journal of Chemometrics. 2014; 28(5): 462-475.

[36] Yang, Y. Injection molding: from process to quality control. Ph.D. Thesis, The Hong Kong University of Science & Technology, 2004.

[37] Liu T, Gao FR. Robust two-dimensional iterative learning control for batch processes with state delay and time-varying uncertainties. Chemical Engineering Science. 2010; 65: 6134-6144.

[38] Chu YH, Lee YH, and Han C. Improved quality estimation and knowledge extraction in a batch process by bootstrapping-based generalized variable selection. 2004; 43: 2680-2690.

[39] Mireia F, Stefan P, Stefan T, Romà T. Comparison of the variable importance in projection (VIP) and of the selectively ratio (SR) methods for variable selection and interpretation. Journal of Chemometrics, 2015; 29: 528-536.

List of Figure Captions

Figure 1 Three-way data analysis and time-slice regression data pairs.

Figure 2 Procedures of decomposing the time-slice data matrix into four subspaces.

Figure 3 Phase partition results (a) the first phase when α is equal to 1.9 (b) the first phase when α is equal to 2.0 (Cluster centers are marked by red ellipses).

Figure 4 Phase partition results (a) the second phase when α is equal to 1.2 (b) the third phase when α is equal to 1.3 (c) the fourth phase when α is equal to 2.5 (d) the fifth phase when α is equal to 1.3 and (e) the sixth phase when α is equal to 2.4 (Cluster centers are marked by red ellipses).

Figure 5 Phase partition results and transition analysis of the proposed method (a) phase partition results and (b) transition patterns analysis.

Figure 6 Phase partition results of QSSPP algorithm with different values of relaxing factor.

Figure 7 Phase-wise cumulative quality prediction results of each phase using the index $R^2_{c,k}$

Figure 8 Process-wise cumulative-quality effect analysis to determine critical-to-cumulative-quality phases for (a) Phases 1 and 2 (b) Phases 2 and 3 (c) Phases 2, 3 and 6 using the index $R^2_{c,k}$ (the red dashed line denotes the upper limit of $R^2_{c,k}$).

Figure 9 Variable importance in the projection plots for (a) Variable 5 and (b) Variable 8 (the black line is the criterion to determine the quality-determined phases).

Figure 10 Analysis of cumulative quality prediction for (a) Phase 2 (b) Phase 3 and (c) Phases 2 and 3.

Figure 11 Comparisons of prediction accuracy for different methods and testing batches.

Figure 12 Different process variations evaluated by indexes $R_{s,c,k}$ and $R_{d,c,k}$ ($R_{s,c,k}$ reveals the repetitive quality-relevant information and $R_{d,c,k}$ reveals the dissimilar quality-relevant variations).

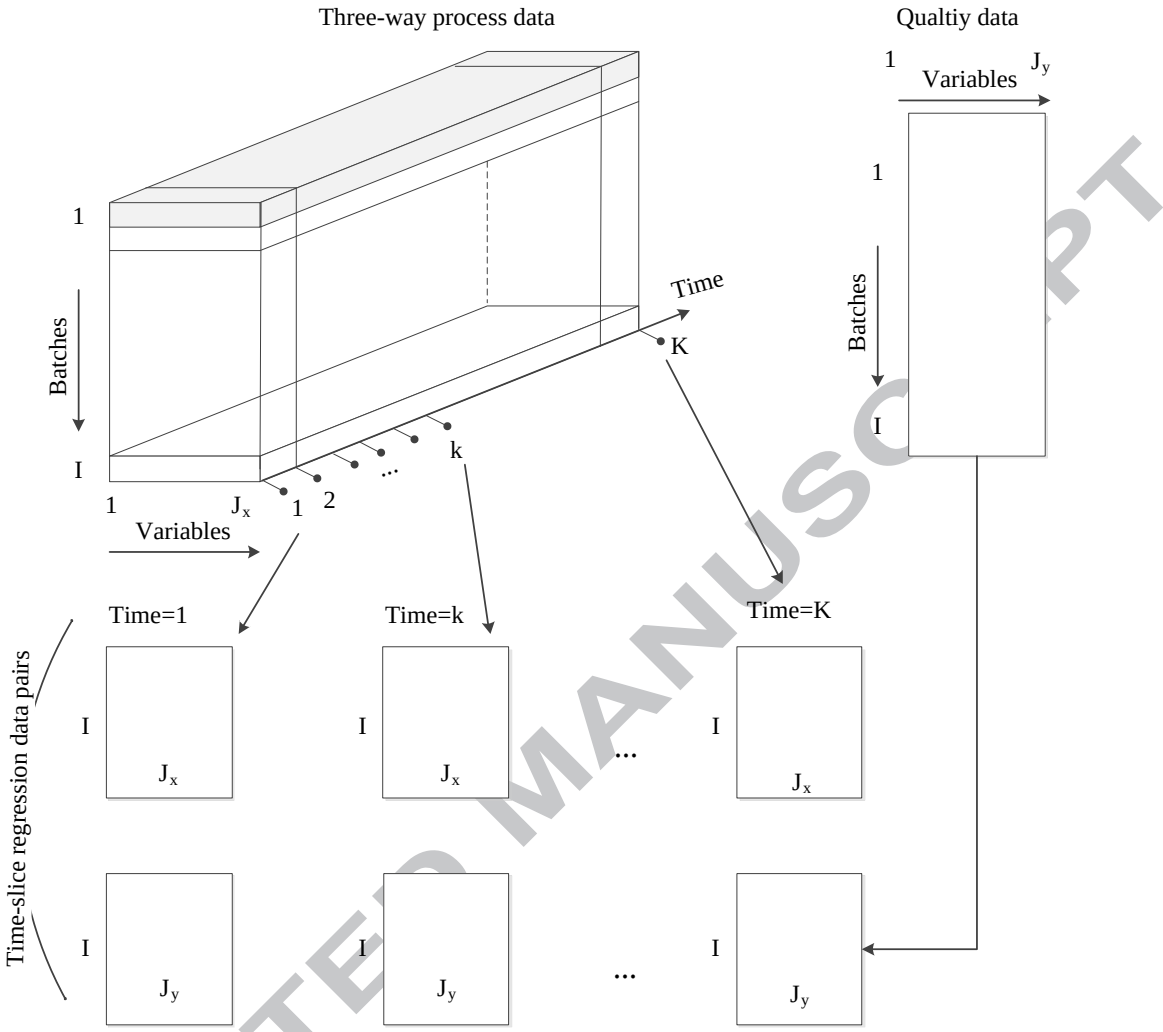


Figure 1 Three-way data analysis and time-slice regression data pairs.

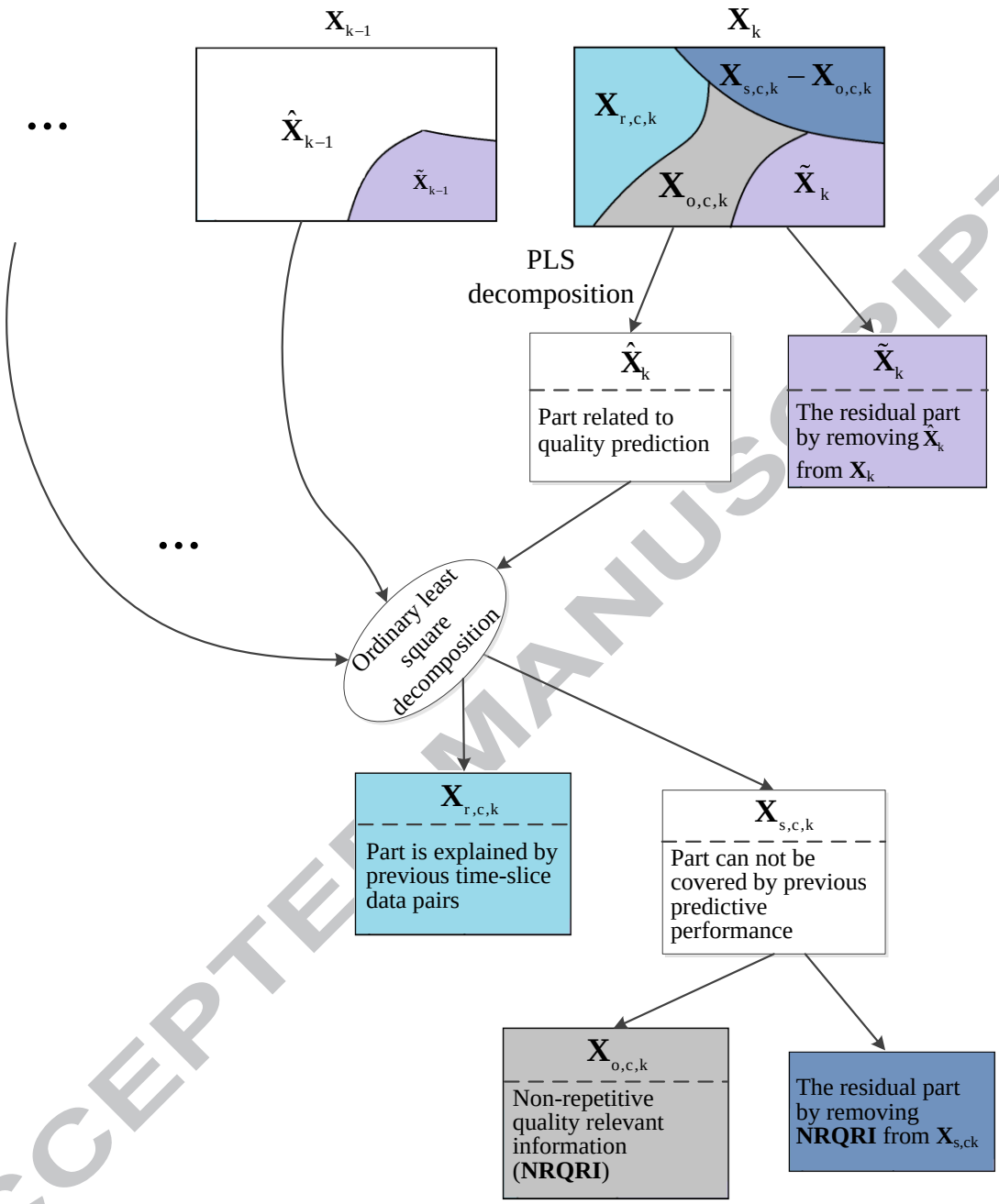


Figure 2 Procedures of decomposing the time-slice data matrix into four subspaces.

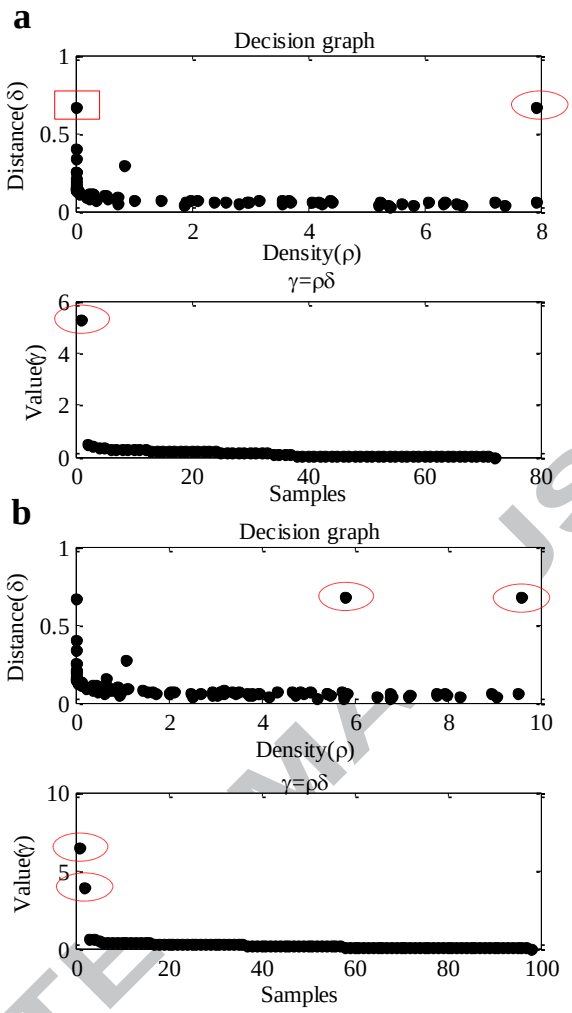


Figure 3 Phase partition results (a) the first phase when α is equal to 1.9 (b) the first phase when α is equal to 2.0 (Cluster centers are marked by red ellipses).

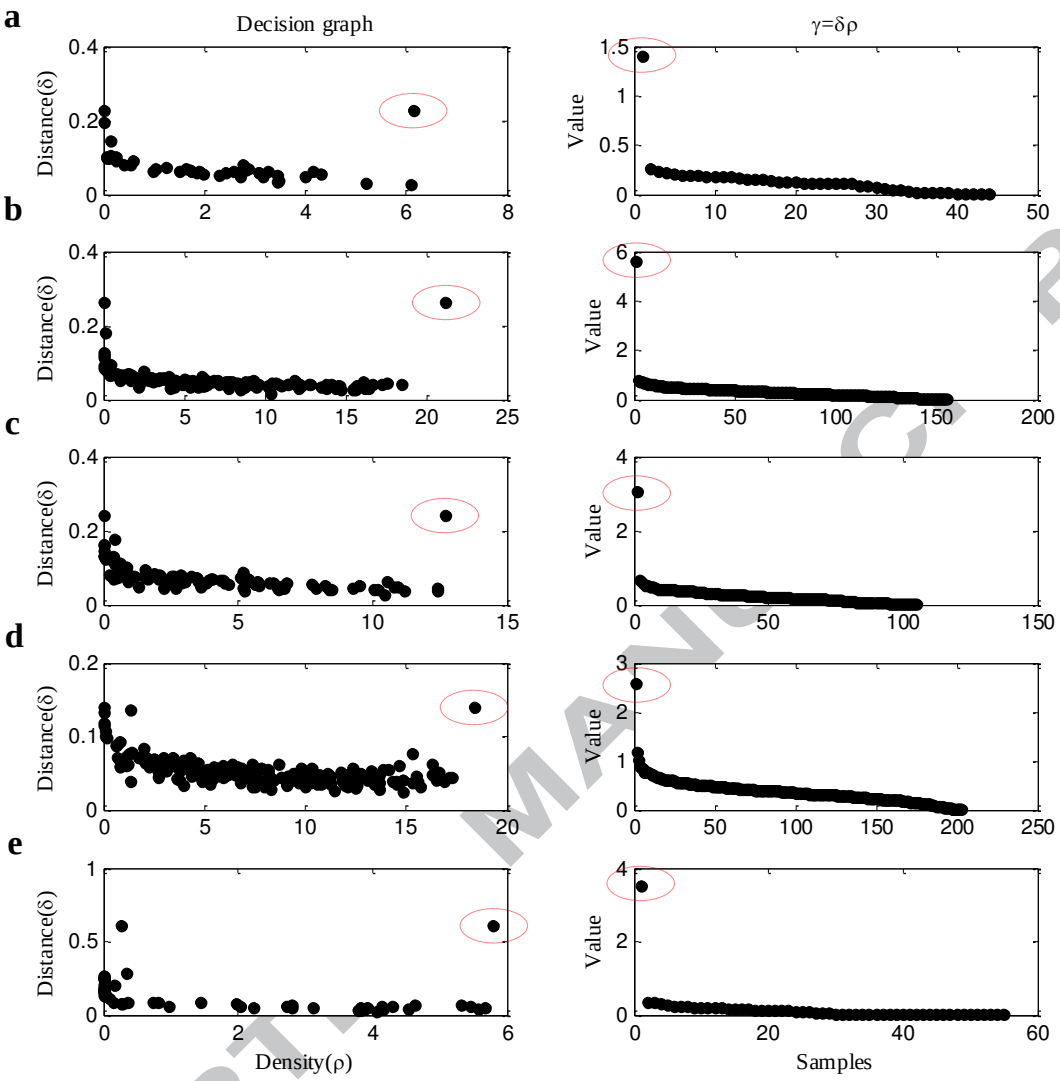


Figure 4 Phase partition results (a) the second phase when α is equal to 1.2 (b) the third phase when α is equal to 1.3 (c) the fourth phase when α is equal to 2.5 (d) the fifth phase when α is equal to 1.3 and (e) the sixth phase when α is equal to 2.4 (Cluster centers are marked by red ellipses).

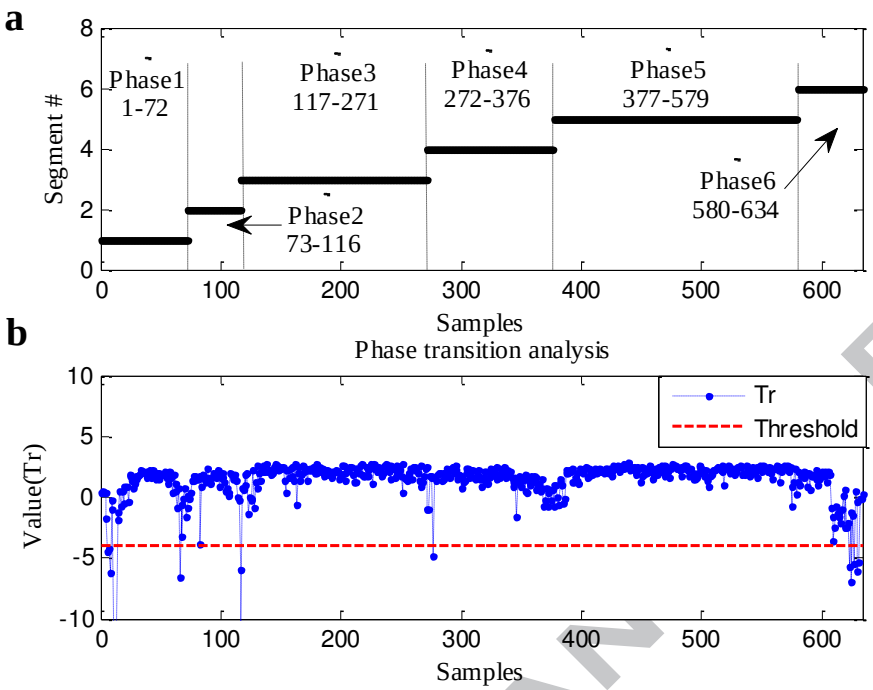


Figure 5 Phase partition results and transition analysis of the proposed method (a) phase partition results and (b) transition patterns analysis.

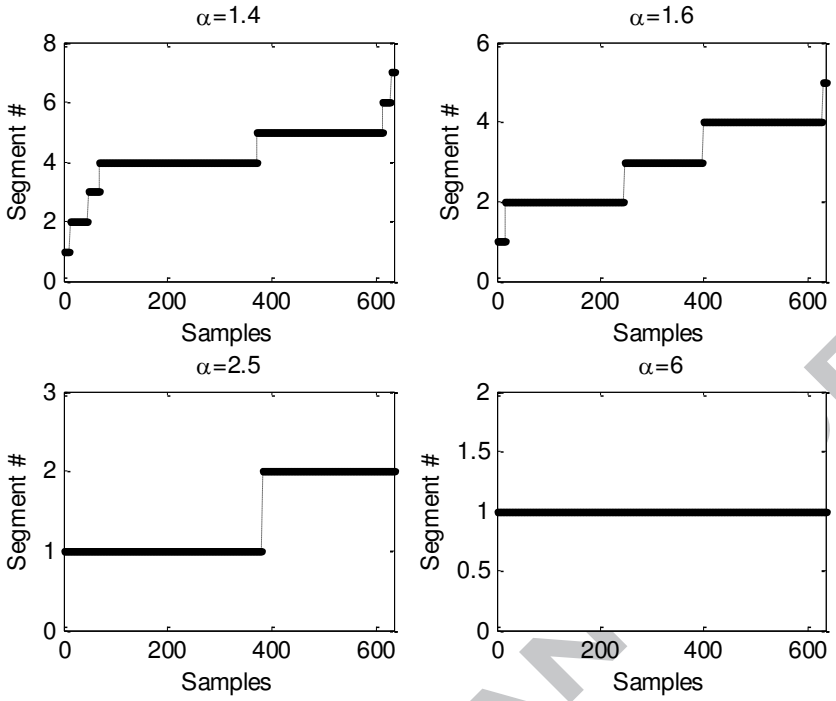


Figure 6 Phase partition results of QSSPP algorithm with different values of the relaxing factor.

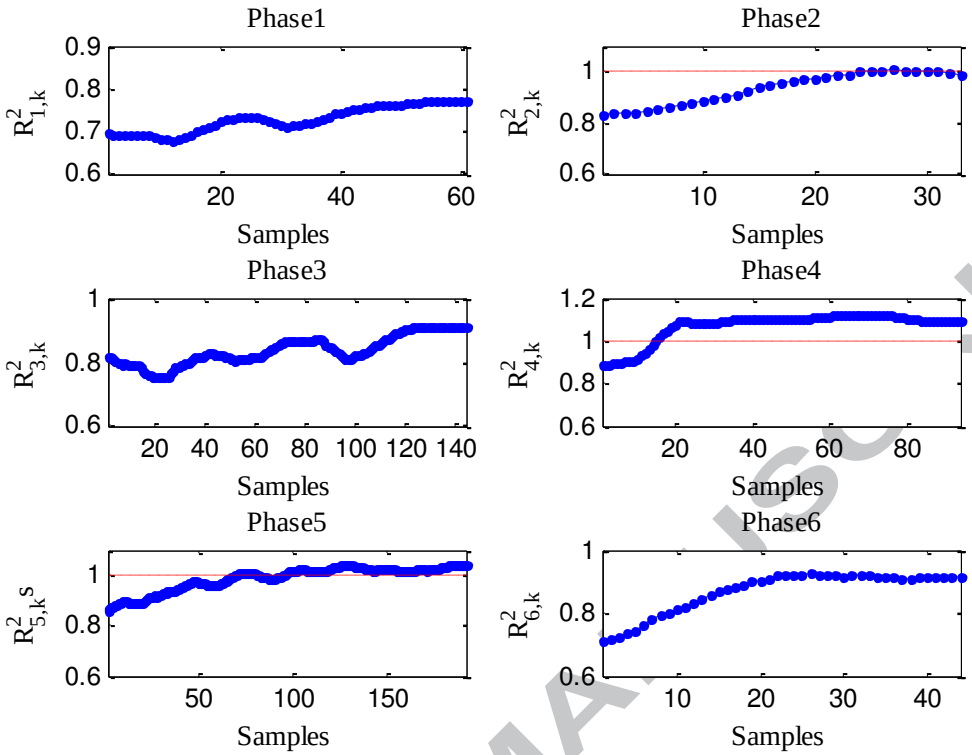


Figure 7 Phase-wise cumulative quality prediction results of each phase using the index $R^2_{c,k}$.

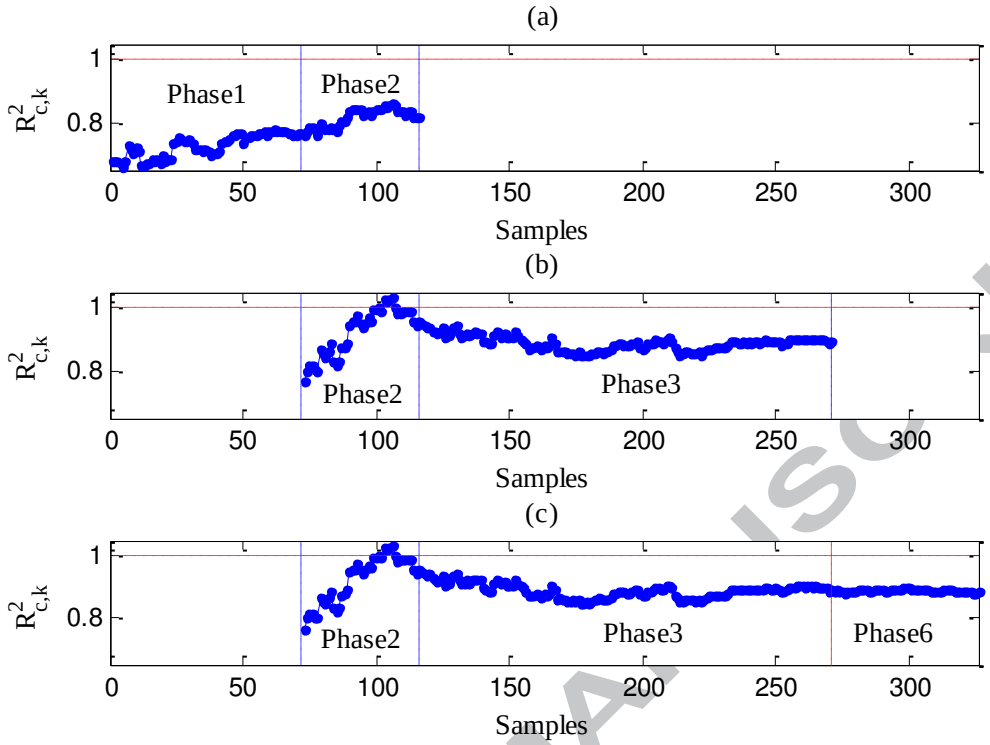


Figure 8 Process-wise cumulative-quality effect analysis to determine critical-to-cumulative-quality phases for (a) Phases 1 and 2 (b) Phases 2 and 3 (c) Phases 2, 3 and 6 using the index $R^2_{c,k}$ (the red dashed line denotes the upper limit of $R^2_{c,k}$).

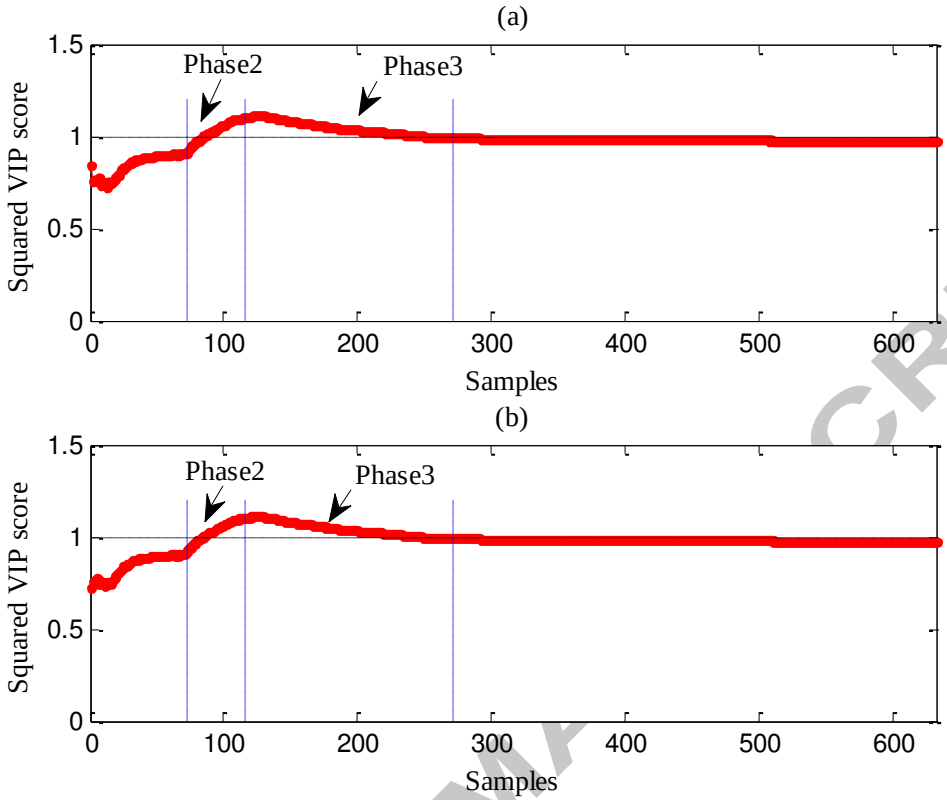


Figure 9 Variable importance in the projection plots for (a) Variable 5 and (b) Variable 8 (the black line is the criterion to determine the quality-determined phases).

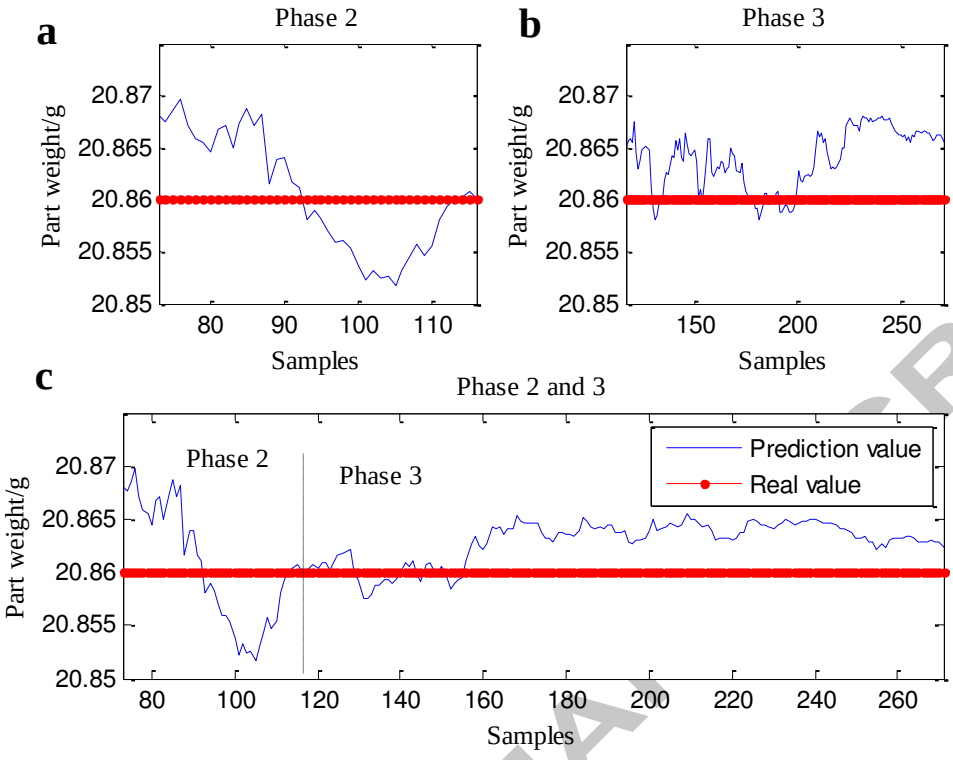


Figure 10 Analysis of cumulative quality prediction for (a) Phase 2 (b) Phase 3 and (c) Phases 2 and 3.

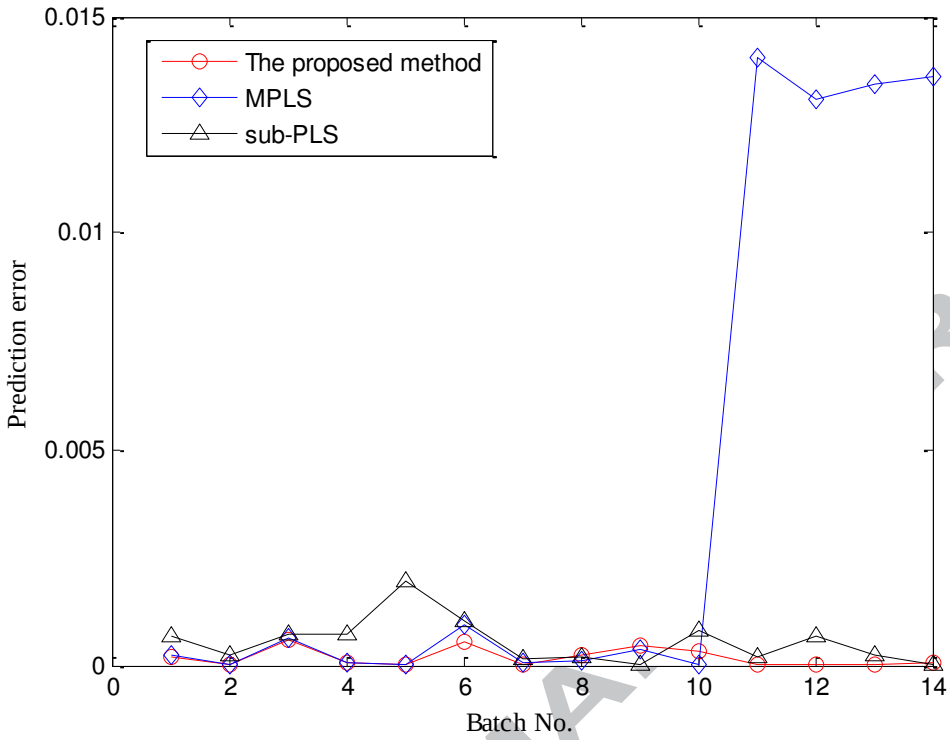


Figure 11 Comparisons of prediction accuracy for different methods and testing batches.

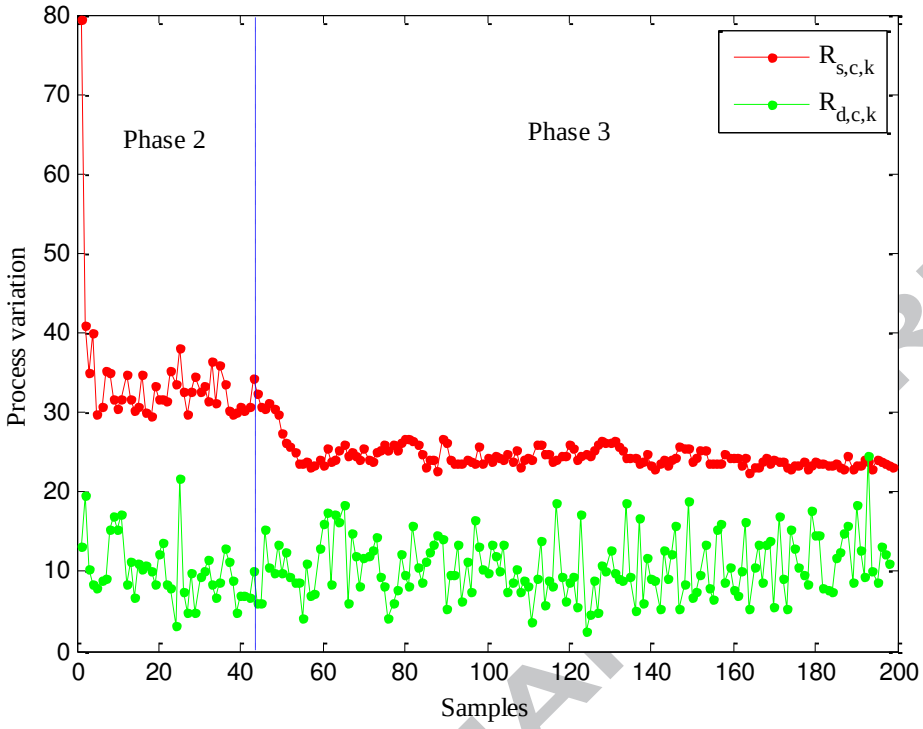


Figure 12 Different process variations evaluated by indexes $R_{s,c,k}$ and $R_{d,c,k}$ ($R_{s,c,k}$ reveals the repetitive quality-relevant information and $R_{d,c,k}$ reveals the dissimilar quality-relevant variations).

Table 1 Eleven process variables of 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 pressure	Bar
9	Barrel temperature zone 3	°C
10	Barrel temperature zone 2	°C
11	Barrel temperature zone 1	°C

- Sequential quality-relevant phases are obtained and values of tunable parameters are properly determined to avoid the uncertainty.
- The type of quality index is judged in advance to determine whether it is of cumulative type.
- Subspace at each time is decomposed to probe into the variation that contributes to cumulative quality effect.
- Critical-to-cumulative-quality phases are identified for inter-phase cumulative quality analysis.