Incorporating Setting Information for Maintenance-Free Quality Modeling of Batch Processes

Zhiqiang Ge and Zhihuan Song
State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Department of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China

Furong Gao
Dept. of Chemical and Biomolecular engineering, The Hong Kong University of Science and Technology, Hong Kong, China

Center for Polymer Processing and Systems, Fok Ying Tung Graduate School, Hong Kong University of Science and Technology, Hong Kong, China

DOI 10.1002/aic.13864
Published online June 25, 2012 in Wiley Online Library (wileyonlinelibrary.com).

Typically, the operation condition of the batch process is changed frequently, following different recipes or manufacturing various production grades. For quality prediction purpose, the prediction model should also be updated or rebuilt, which leads to a significant model maintenance effort, especially for those processes which have various phases. To reduce such effort, a maintenance-free method is proposed in this article, which incorporates the setting information of the batch process for modeling. The whole process variations are separated into two parts: setting information related and other quality related variations. By constructing a relationship between setting variables and other process variables, the data variations explained by the setting information can be efficiently removed. Then, a robust regression model connecting process variables to the quality variable is developed in different phases of the batch process. The feasibility and effectiveness of the proposed method is evaluated through an industrial injection molding process. © 2012 American Institute of Chemical Engineers *AIChE J,* 59: 772–779, 2013
Keywords: quality prediction, multiphase batch process, maintenance-free, setting information, robust modeling

Introduction

With the increased requirements of producing low volume and high value-added products, batch processes have played an important role in modern industrial manufacturing processes. Particularly, in the past years, data-based modeling methods such as multivariate statistic analysis methods have been widely used for batch process monitoring and quality prediction, such as multiway principal component analysis (PCA), multiway partial least squares (PLS), and multiway independent component analysis.[1–7] Later, for those batch processes which have multiple phases, the traditional modeling methods have been successfully extended to phase-based approaches, based on which various monitoring and quality prediction methods have also been developed.[8–20]

To our best knowledge, however, most existing methods for quality prediction have assumed that the batch process is operated under a stable condition. In practice, the operation condition of batch process often changes, such as the load of raw materials, product grade transitions, catalyst deactivation, seasonal effects, equipment aging, and so forth. Because of these changes, if the prediction model remains unchanged, the quality prediction performance will be deteriorated under those operation regions which have not been incorporated for modeling. Therefore, it is necessary to maintain the model adaptively to track the condition change of the process.

In the past years, adaptive modeling approaches have been developed to handle the time-varying behavior of the process, such as recursive and adaptive PCA, recursive PLS, moving window-based PCA and PLS, Exponentially weighted moving average-based PCA, and so on.[21–26] For processes with multiple operation modes, the model library-based method has been developed. In each of the operation mode, a corresponding model was developed for process monitoring.[27–30] Recently, the local model-based approach has been developed, which are trained on a partition of the historical dataset, corresponding to a particular region of the process.[31–36] Based on those local models, both of the time-varying and operation mode change behaviors can be well addressed. However, the adaptive method is often carried out blindly, which means that continuous updating is performed whether a process change has been identified. Besides, there is also a risk that the adaptive model will be accommodated to faulty process conditions if some abnormal data samples have been used for model updating. A common drawback of the model library based method and the local model based approach is due to their limitation

Correspondence concerning this article should be addressed to Z. Ge at zqge@iipc.zju.edu.cn or F. Gao at kefgao@ust.hk.

for new process changes. This is because their models are both built on existing operation regions. Therefore, both of these two methods are not able to provide a satisfactory modeling performance.

It can be noticed that to track the change of the process, all of the existing methods need long-term periodical maintenance of the model. Although it is easy for those processes which have a short running time, some other processes may have a complex operation procedure and the model updating is computational expensive. Therefore, to minimize the effort of long-term model maintenance, it will be very useful if the quality prediction model is robust to the change of the process. Kano et al.[37] developed an external analysis method for monitoring processes with various changes, which was later extended to the nonlinear case by Ge et al.[38] In this method, the whole set of process variables were divided into two classes: external variables (which are related to the change of the process) and main variables. By removing the influence of the external variables on the main variables, the model was constructed upon the residual information of the main variables, which is robust to the change in the process. However, for most industrial processes, it is difficult to identify the external variables, especially when the change of the process is complex and subjected to many variables. Furthermore, this external analysis method only considered the data information of ordinary process variables, the quality information has not been incorporated.

In batch processes, typically, most production grade transitions or characteristic changes of different batches are subjected to process/recipe settings, which are very easy to obtain and identify. In other words, the setting information of the batch process is the key ingredient to manipulate the change of different batches/products. If the setting information can be incorporated for quality modeling, a robust model can be developed. In this case, the whole variation of the batch process can be separated into two parts: (1) those controlled or explained by the setting information; (2) other variations. Therefore, if the part explained by the setting information is removed from the data variation, the quality prediction model can be constructed between the remaining variation of the measured process variables and the quality variables, which is robust to the change of the batch conditions.

The remainder of this article is organized as follows. In ''Methodology'' Section, a detailed demonstration of the proposed method is provided, followed by an industrial case study of the injection molding process. Finally, conclusions are made.

Methodology

In this section, the detailed description of the maintenance-free quality modeling method is provided. First, the setting information of the batch process is incorporated for quality modeling, based on which a robust relationship between the process variables and the quality variable can be built. Second, based on this robust modeling method, a maintenance-free quality prediction model is built and applied in multiphase batch processes.

Incorporating setting information for robust modeling

In batch processes, the setting information is always reloaded when the operation condition or the production grade changes. With the incorporation of the setting information for process modeling, the variation of the whole measured data information can be partitioned into two parts: setting information explained and unexplained parts. By constructing a relationship between the setting information and process variables, the data variations which explained by the setting information can be efficiently removed. Therefore, the rest data variation is not sensitive to changes of the batch process. This robust modeling method is illustrated in Figure 1.

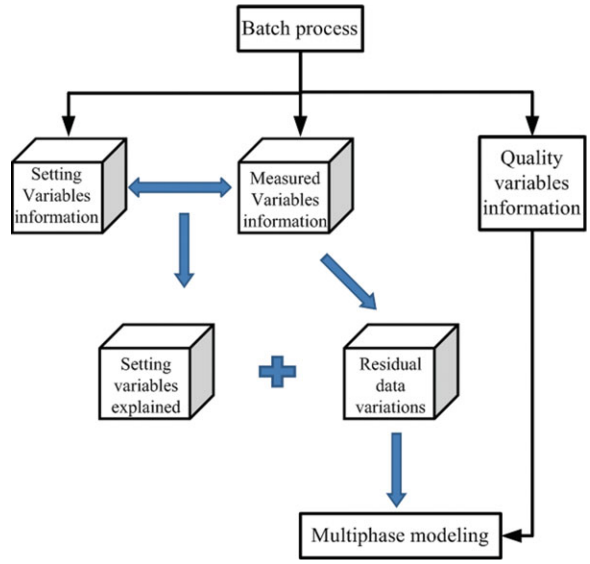

Figure 1. Illustration of the robust modeling method.

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

Representing the data matrix of the process variable as $\mathbf{X}$, and the setting data information as $\mathbf{S}$. According to the robust modeling method, to remove the data variation explained by the setting information, a regression model is intended to be built between process and setting variables. Generally, various regression analysis tools can be used, such as (partial) least squares model, principal component regression model, artificial neural network, support vector regression model, among others.

Maintenance-free quality prediction model for multiphase batch processes

In the present article, the robust method is used for quality modeling of multiphase batch processes. Precisely, a maintenance-free model is developed for quality prediction purpose in each phase of the batch process. Therefore, despite of the operation condition change of the process, the quality prediction model updating is not desired. Instead, we only need to remove the data variation which is explained by the setting variables from the process variables, and then use the quality prediction model in each phase of the batch process.

Beforehand, it is assumed that the batch process has already been divided into different phases. Although this issue is important in multiphase quality prediction, it is not the main focus of the present work. Given the three-way batch process dataset $\mathbf{X}(I \times J \times K)$, the setting variable dataset $\mathbf{S}(I \times J_s \times K)$, and the corresponding quality dataset $\mathbf{Y}(I \times J_y)$, where I is the number of batches, J, J_s, and J_y are numbers of process, setting and quality variables, and the

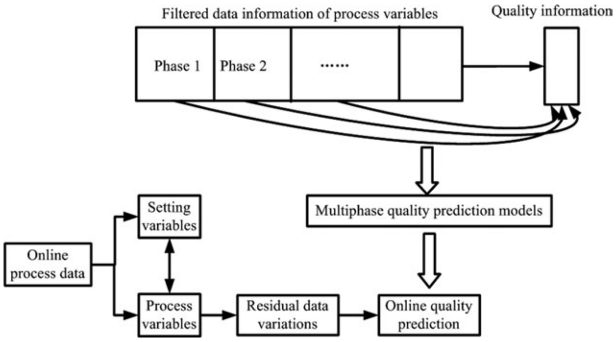

Figure 2. Illustration of multiphase modeling and online quality prediction.

duration of each batch is K. Through the batch direction, the three-way data matrices can be unfolded into two-dimensional (2-D) data matrices $\mathbf{X}(I \times JK)$, and $\mathbf{S}(I \times J_s K)$. Suppose, the whole batch process has been divided into Ph phases, $\mathbf{X}(I \times JK)$ and $\mathbf{S}(I \times J_s K)$ can be further partitioned as follows

$$\mathbf{X}(I \times JK) = [\mathbf{X}_1(I \times JK_1) \;\; \mathbf{X}_2(I \times JK_2) \;\; \cdots \;\; \mathbf{X}_{Ph}(I \times JK_{Ph})] \tag{1}$$

$$\mathbf{S}(I \times J_s K) = [\mathbf{S}_1(I \times J_s K_1) \;\; \mathbf{S}_2(I \times J_s K_2) \;\; \cdots \;\; \mathbf{S}_{Ph}(I \times J_s K_{Ph})] \tag{2}$$

where $K_1, K_2, \cdots, K_{Ph}$ are the numbers of time slices in different phases

Suppose, the batch process has various operation conditions or producing various production grades, the 2-D process and setting variable matrices can be further extended as follows

$$\mathbf{X}(I \times JK) = \begin{bmatrix} \mathbf{X}_1^1(I_1 \times JK_1) & \mathbf{X}_2^1(I_1 \times JK_2) & \cdots & \mathbf{X}_{Ph}^1(I_1 \times JK_{Ph}) \\ \mathbf{X}_1^2(I_2 \times JK_1) & \mathbf{X}_2^2(I_2 \times JK_2) & \cdots & \mathbf{X}_{Ph}^2(I_2 \times JK_{Ph}) \\ \vdots & \vdots & \mathbf{X}_{ph}^c(I_c \times JK_{ph}) & \vdots \\ & & \vdots \quad \ddots & \\ \mathbf{X}_1^C(I_C \times JK_1) & \mathbf{X}_2^C(I_C \times JK_2) & \cdots & \mathbf{X}_{Ph}^C(I_C \times JK_{Ph}) \end{bmatrix} \tag{3}$$

$$\mathbf{S}(I \times J_s K) = \begin{bmatrix} \mathbf{S}_1^1(I_1 \times J_s K_1) & \mathbf{S}_2^1(I_1 \times J_s K_2) & \cdots & \mathbf{S}_{Ph}^1(I_1 \times J_s K_{Ph}) \\ \mathbf{S}_1^2(I_2 \times J_s K_1) & \mathbf{S}_2^2(I_2 \times J_s K_2) & \cdots & \mathbf{S}_{Ph}^2(I_2 \times J_s K_{Ph}) \\ \vdots & \vdots & \mathbf{S}_{ph}^c(I_c \times JK_{ph}) & \vdots \\ & & \vdots \quad \ddots & \\ \mathbf{S}_1^C(I_C \times J_s K_1) & \mathbf{S}_2^C(I_C \times J_s K_2) & \cdots & \mathbf{S}_{Ph}^C(I_C \times J_s K_{Ph}) \end{bmatrix} \tag{4}$$

where we have assumed that I_c, $c = 1,2, \cdots ,C$ batches are obtained under each operation condition of the batch process, hence, $\sum_{c=1}^{C} I_c = I$, $ph = 1,2, \cdots ,Ph$ is the phase number. Based on the robust modeling method, a simply least squares regression model is used to build the relationship between the process variables and the setting variables in each time slice of the batch process, which is given as follows

$$\mathbf{R}_k = (\mathbf{S}_k^T \mathbf{S}_k)^{-1} \mathbf{S}_k^T \mathbf{X}_k, \quad k = 1, 2 \cdots, K \tag{5}$$

Therefore, the data variation of the process variables can be partitioned into two parts

$$\mathbf{X}_k = \hat{\mathbf{X}}_k + \tilde{\mathbf{X}}_k = \mathbf{R}_k^T \mathbf{S}_k + \tilde{\mathbf{X}}_k \tag{6}$$

where $\hat{\mathbf{X}}_k = \mathbf{R}_k^T \mathbf{S}_k$ corresponds to the variation part which is explained by the setting variables, and $\tilde{\mathbf{X}}_k$ is the unexplained data variation, which is not sensitive to the process change any more.

Therefore, after the nonsensitive data variation information of each time slice has been separated, the relationship between $\tilde{\mathbf{X}}_k$ and the quality data $\mathbf{Y} \in \mathbf{R}^{I \times J_y}$ can be modeled by the PLS model, which is given as[2]

$$\begin{aligned} \tilde{\mathbf{X}}_k &= \mathbf{T}_k \mathbf{P}_k^T + \mathbf{E}_k \\ \mathbf{Y} &= \mathbf{U}_k \mathbf{Q}_k^T + \mathbf{F}_k \\ \mathbf{R}_k^{pls} &= \mathbf{W}_k (\mathbf{P}_k^T \mathbf{W}_k)^{-1} \mathbf{Q}_k \end{aligned} \tag{7}$$

where $k = 1,2, \cdots ,K$, $\mathbf{P}_k$ and $\mathbf{Q}_k$ are the loading matrices, $\mathbf{T}_k$ and $\mathbf{U}_k$ are the latent variable matrices, $\mathbf{E}_k$ and $\mathbf{F}_k$ are the

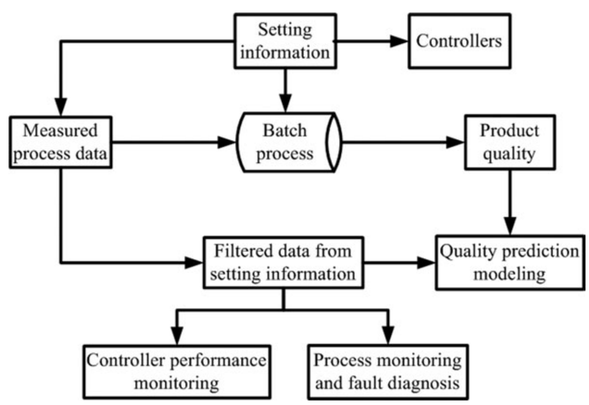

Figure 3. A general data-based modeling and application framework with incorporation of setting information.

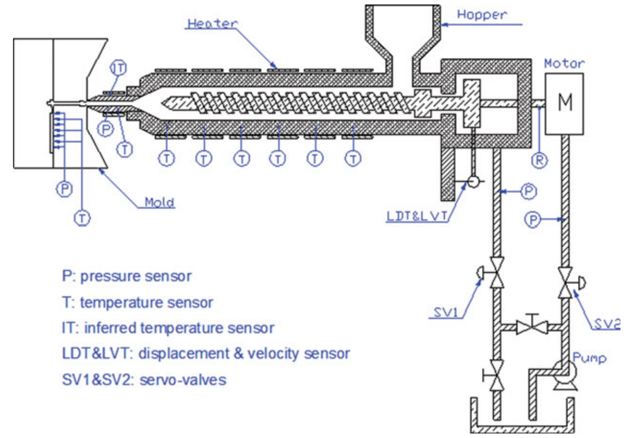

Figure 4. Simplified schematic flowchart of the injection molding machine.10.

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

residual matrices, $\mathbf{W}_k$ is the weighting matrix of each PLS model, and $\mathbf{R}_k^{\text{PLS}}$ is the regression matrix for the PLS model.

Finally, the representative PLS model in each phase of the batch process can be constructed, the regression matrix of which is determined as follows

$$\mathbf{R}_{ph}^{\text{PLS}} = \frac{1}{K_{ph}} \sum_{k=K_{ph-1}+1}^{K_{ph}} \mathbf{R}_k^{\text{PLS}} \tag{8}$$

where K_{Ph}, $Ph = 1{,}2, \cdots, Ph$ is the number of time slices in each phase of the batch process.

Online quality prediction algorithm for new batches

Based on the developed maintenance-free model, an online quality prediction algorithm is formulated here. Given a new batch with its data information up to the kcth time point, and denote the process and setting variable vectors as $\mathbf{x}_{\text{new},kc}$ and $\mathbf{s}_{\text{new},kc}$. In the first step, the influence of the setting variables on the process variables should be removed. Therefore, a least squares model between the process data and setting data is built, based on which and the variation of the new data sample can be partitioned into two parts, which are given as follows

$$\begin{aligned}
\mathbf{S}_{\text{new},kc} &= \begin{cases} [\mathbf{s}_{\text{new},1}^T \ \ \mathbf{s}_{\text{new},2}^T \ \ \cdots \ \ \mathbf{s}_{\text{new},kc}^T] & 1 \le kc \le K_1 \\ [\mathbf{s}_{\text{new},K_1+1}^T \ \ \mathbf{s}_{\text{new},K_1+2}^T \ \ \cdots \ \ \mathbf{s}_{\text{new},kc}^T] & K_1+1 \le kc \le K_1+K_2 \\ \vdots & \\ [\mathbf{s}_{\text{new},K_1+\cdots+K_{Ph-1}+1}^T \ \ \mathbf{s}_{\text{new},K_1+\cdots+K_{Ph-1}+2}^T \ \ \cdots \ \ \mathbf{s}_{\text{new},kc}^T] & \sum_{i=1}^{Ph-1} K_i + 1 \le kc \le \sum_{i=1}^{Ph} K_i \end{cases} \\
\mathbf{X}_{\text{new},kc} &= \begin{cases} [\mathbf{x}_{\text{new},1}^T \ \ \mathbf{x}_{\text{new},2}^T \ \ \cdots \ \ \mathbf{x}_{\text{new},kc}^T] & 1 \le kc \le K_1 \\ [\mathbf{x}_{\text{new},K_1+1}^T \ \ \mathbf{x}_{\text{new},K_1+2}^T \ \ \cdots \ \ \mathbf{x}_{\text{new},kc}^T] & K_1+1 \le kc \le K_1+K_2 \\ \vdots & \\ [\mathbf{x}_{\text{new},K_1+\cdots+K_{Ph-1}+1}^T \ \ \mathbf{x}_{\text{new},K_1+\cdots+K_{Ph-1}+2}^T \ \ \cdots \ \ \mathbf{x}_{\text{new},kc}^T] & \sum_{i=1}^{Ph-1} K_i + 1 \le kc \le \sum_{i=1}^{Ph} K_i \end{cases} \\
\mathbf{R}_{\text{new},kc} &= (\mathbf{S}_{\text{new},kc}^T \mathbf{S}_{\text{new},kc})^{-1} \mathbf{S}_{\text{new},kc}^T \mathbf{X}_{\text{new},kc} \\
\mathbf{x}_{\text{new},kc} &= \hat{\mathbf{x}}_{\text{new},kc} + \tilde{\mathbf{x}}_{\text{new},kc} = \mathbf{R}_{\text{new},kc}^T \mathbf{s}_{\text{new},kc} + \tilde{\mathbf{x}}_{\text{new},kc}
\end{aligned} \tag{9}$$

As a result, the unexplained part of the new data information $\tilde{x}_{\text{new},kc}$ is robust to the change of the process. In the second step, the PLS regression model is used for calculating the prediction value based on the new data sample, which is given as

$$\hat{\mathbf{y}}_{\text{new},kc} = \mathbf{R}_{ph}^{\text{PLST}} \tilde{\mathbf{x}}_{\text{new},kc} \tag{10}$$

where "ph" represents the current phase which the new data sample belongs to.

Detailed multiphase modeling and online quality prediction algorithms are illustrated in Figure 2.

Remarks and discussions

With the incorporation of the setting information in the batch process, a robust regression model has been constructed for quality prediction in different phases. As a result, when the operation condition of the batch process has been changed, based on the robust model, the relationship between the process variables which are uncorrelated with the setting information and the quality variables remains stable. Therefore, the change of process condition will have little impact on the quality prediction performance. Compared to existing quality prediction models, the proposed method is particularly useful when the change of the batch process is new and has not been incorporated for modeling. However, if an abnormal change of the batch process has been

Table 1. Select Variables for Quality Prediction

No.	Process Variables	Unit	No.	Setting Variables	Unit
1	Valve 1	%	1	Temperature 1	°C
2	Valve 2	%	2	Temperature 2	°C
3	Screw stroke	Mm	3	Temperature 3	°C
4	Screw velocity	Mm/s			
5	Ejector stroke	Mm			
6	Mold stroke	Mm			
7	Mold velocity	Mm/s			
8	Injection press	Bar			

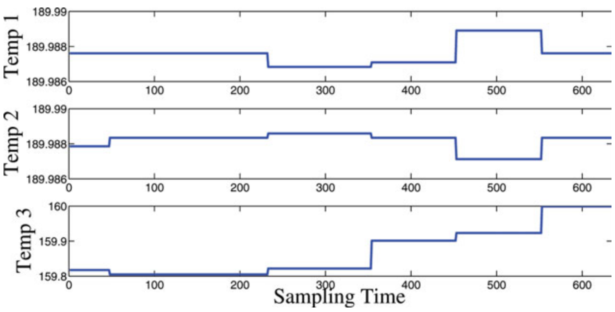

Figure 5. The characteristics of the three setting variables under base operation condition.

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

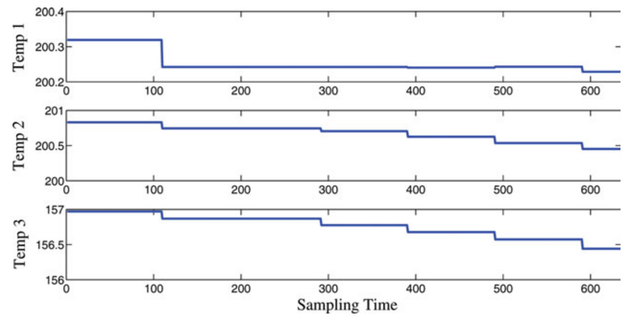

Figure 6. The characteristics of the three setting variables under new operation condition.

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

included, the model may be adapted to some fault which is not expected in the process. A promising method has been developed to avoid adapting abnormal changes of the process in Wang et al.[26] Besides, incorporations of available process knowledge and experiences of experts are also useful to address this problem.

Although most prediction models need to update their model parameters or even reconstruct the whole model structure, the new method remains unchanged under different operation conditions, thus is maintenance-free. However, for compromise, the online implementation of the proposed method is a little more complicated, this is because an additional relationship between the setting information and the measured data information should be built before the main quality prediction procedure. However, compared to the periodical maintenance of the prediction model, this online implementation step is much easier, and also computational inexpensive.

For simplicity, we have used the ordinary least squares model to characterize the relationship between setting variables and the measured process variables. In general, however, this model can be extended to other forms, for example, some nonlinear models such as artificial neural networks, support vector regression models, some dynamic modeling approaches such as time series models, state-space identification models, and so forth. Similarly, the robust regression model between the residual process variations and the quality data can also be extended to other cases, which always depends on the proper data characteristic of the process.

Although the main work of this study is focused on quality prediction of the multiphase batch process, this setting information incorporation idea is actually a general modeling method. Therefore, it can also be used for other purposes, such as data-based process monitoring, control of processes with various operation conditions. Besides, the setting information is always related to the set-points of various controllers in the batch process. If a proper relationship can be built between the setting information and inputs/outputs of the controllers, the operation performance of each controller can be monitored, and the disabled controllers can thus be separated from the process. A general setting information incorporated modeling and application framework is demonstrated in Figure 3, which is purely data-driven.

Illustrations and Results

In this section, the performance of the proposed method is evaluated through an industrial injection molding process, which is a typical multiphase batch process. A simplified flowchart of a reciprocating-screw injection molding process is shown in Figure 4. To produce various products, the operation condition of this process is required to change frequently. In this process, the weight of the final product is regarded as the quality variable. For prediction of this quality variable, some online measured process variables are used, such as temperatures, pressures, and the screw velocity, all of which are tabulated in Table 1. The setting information of this process is represented by three temperatures, which are also listed in Table 1. Figure 5 shows the detailed information of three setting variables, it can be seen that the setting values of these three variables are changed with phases. With the changes of these three setting variables, the relationships among different process variables could be changed significantly. In this case, if the influence of setting variables has not been removed, the quality prediction model should be rebuilt or adapted to include the new patter of the process. However, with the separation of the three setting variables, a stable relationship among different process variables can be extracted. As a result, periodical maintenance of the quality prediction model can be eliminated. More detailed description and experimental design of the injection molding process can be found in Ref. 10.

A dataset which contains 115 batches has been collected in the injection molding process, among which 100 batches are used for model training and the rest 15 batches are for testing. The duration of each batch is 635 sampling points. Based on the phase division method, this process can be divided into seven different phases, detailed phase division results are given in Table 2. When the process condition has been changed, another 15 batches are collected, which are represented as the new batches and also used for testing purpose. In the new operation condition, the characteristics of the three setting variables are given in Figure 6. Therefore, we have

Table 2. Phase Division Results of the Injection Molding Process

Phases numbers	1#	2#	3#	4#	5#	6#	7#
Time intervals	1–18	19–117	118–119	120–367	368–369	370–609	610–635

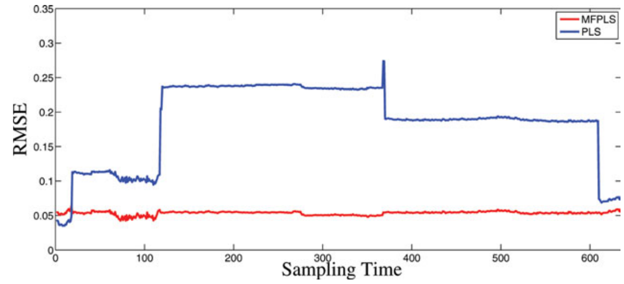

Figure 7. RMSE values of the two methods for the whole testing batches.

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

100 batches which are collected under the base operation condition for modeling training, and other 30 batches generated under two different operation conditions (known condition and new condition) for performance evaluation.

To develop the maintenance-free quality prediction model in each phase of the process, the least squares model is used to remove the explained variations of the setting variables on the process variables. In the second step, the PLS model is developed between the residual data information of the process variables and the quality variable in each phase. For comparison, the conventional phase-based PLS based quality prediction model is also constructed. The numbers of latent variables are selected as 4 in both of the two PLS models. To be clear, the two methods are named as maintenance-free partial least squares (MFPLS) and PLS in this article.

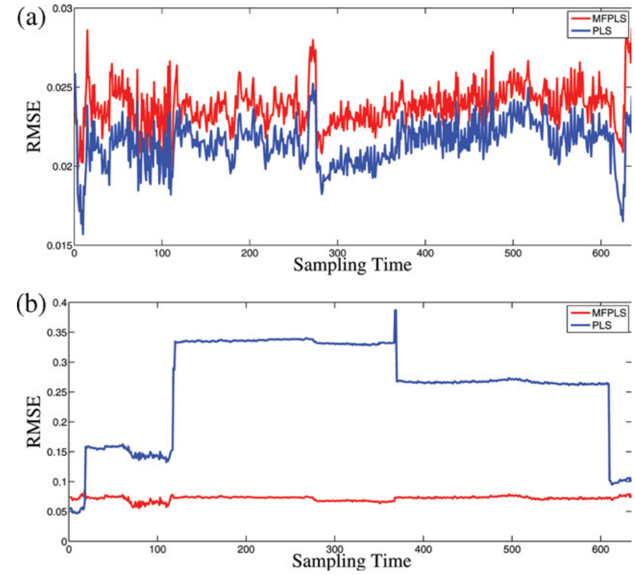

Figure 9. RMSE values of the two methods for testing batches under known and new operation conditions, (a) known operation condition; (b) new operation condition.

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

For quality prediction performance evaluation of these two methods, the root-mean-square error (RMSE) index is defined as follow

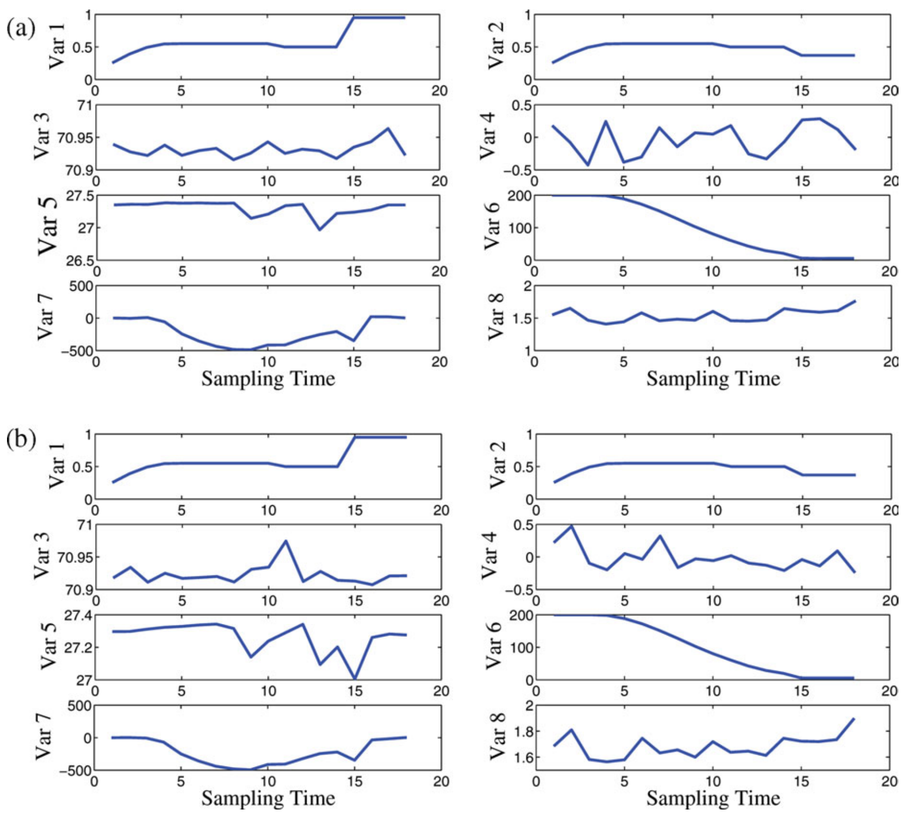

Figure 8. Data characteristics of process variables in the first phase, (a) base operation condition; (b) new operation condition.

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

Table 3. Prediction Results of Two Methods

End-of-phases		1#	2#	3#	4#	5#	6#	7#
Overall	MFPLS	0.0640	0.0581	0.0539	0.0504	0.0544	0.0539	0.0604
	PLS	0.0493	0.1095	0.2038	0.2352	0.2737	0.1871	0.0783
Known operation condition	MFPLS	0.0242	0.0264	0.0239	0.0234	0.0240	0.0238	0.0288
	PLS	0.0197	0.0235	0.0220	0.0208	0.0229	0.0216	0.0245
New operation condition	MFPLS	0.0873	0.0779	0.0724	0.0673	0.0731	0.0724	0.0804
	PLS	0.0668	0.1530	0.2873	0.3320	0.3863	0.2637	0.1079

$$\mathrm{RMSE}(kc)=\sqrt{\frac{\sum_{i=1}^{I_{te}}\left|\mathbf{y}_i-\hat{\mathbf{y}}_i^{kc}\right|^2}{I_{te}}} \tag{11}$$

where I_{te} is the number of testing batches, $kc = 1,2,\cdots, K$, $\mathbf{y}_I$, $I = 1,2, \cdots, I_{te}$ is the measured value of each testing batch, and $\hat{\mathbf{y}}_i^{kc}$ is the predicted value of the testing batch i at time point kc.

After all of the 30 testing batches have been evaluated, the overall RMSE values of the two methods during the whole batch time are shown in Figure 7. It can be seen that the MFPLS model performs much better than the traditional PLS model in almost all time intervals of the batch. The PLS model only show a little superiority in the first phase. This may due to the fact that in the first phase the change of the operation condition has little impact on the process variables, which can be seen in Figure 8. When the setting variables of the process have been changed, most process variables in the first phase remain similar. To examine the quality prediction performance the two methods under the known and new operation conditions of the batch process, the RMSE values are calculated separately, both of which are given in Figure 9. As been expected, the prediction performances of the two methods are comparative for testing batches collected under the operation condition based on which the models are constructed. In contrast, when the process was changed to the new operation condition, the performance of the traditional PLS model has been seriously deteriorated, since the corresponding RMSE values are much bigger than that in the previous operation condition. However, based on the MFPLS model, the influence of the operation condition change can be efficiently reduced. As can be seen from Figure 9b, the quality prediction performance of the testing batches under the new operation condition has been greatly improved by the MFPLS model. Particularly, at the end of each operation phase, the RMSE values of the two methods are given in Table 3. Detailed prediction results of two testing batches are provided in Figure 10, where the first batch is from the known operation condition of the process, and the second one belongs to the new operation condition. Compared to the first batch, the MFPLS model has made a more significant improvement for the second testing batch.

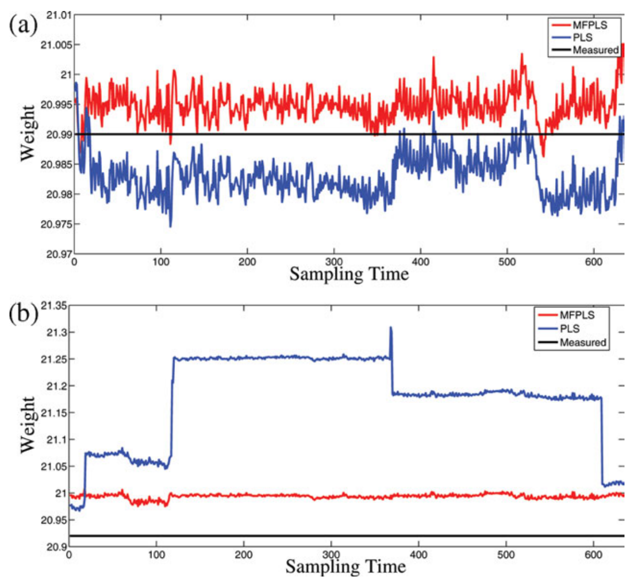

Figure 10. Quality prediction results of two particular testing batches, (a) under known operation condition; (b) under new operation condition.

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

Finally, the computational effort of the new method is evaluated and compared with the traditional PLS model. In the current example, only three operation modes have been used. If the same size of the modeling dataset is used in both methods, the computational effort can be saved around 20%. However, with the incorporation of more and more datasets in different operation conditions, the offline modeling effort can be saved significantly. In contrast, as discussed in "Remarks and discussions" Section, the computational effort of online quality prediction will be slightly increased in the new developed method, due to the calculation of the relationship between setting variables and process variables. Running on the same laptop, the online computation times of the two methods are 1.8408 s and 1.2496 s for the current example.

Conclusions

In the present article, a maintenance-free model has been developed for quality prediction of multiphase batch processes. The main feature of this method is the incorporation of the setting information for modeling. This is due to the fact that most changes of the batch process are subjected to the change of the setting information. Based on the developed model, the variation part explained by the setting variables can be significantly removed from the whole data variation that related to the final product quality. As a result, a robust regression model has been constructed for quality prediction in each phase of the batch process. The main advantage of the proposed method is that the prediction model does not need to be rebuilt or updated when the operation condition of the process is changed. Therefore, for those batch processes which have various phases and the operation condition is changed frequently, the model maintenance effort can be greatly cut by the proposed method.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) (61004134), National Project 973 (2012CB720500), Guangzhou scientific and technological project (11F11140010), and the Fundamental Research Funds for the Central Universities.

Literature Cited

1. Nomikos P, MacGregor JF. Monitoring batch processes using multiway principal component analysis. *AIChE J*. 1994;44:1361–1375.
2. Nomikos P, MacGregor JF. Multi-way partial least square in monitoring batch processes. *Chemom Intell Lab Syst*. 1995;30:97–108.
3. Yoo CK, Lee JM, Vanrolleghem PA, Lee IB. On-line monitoring of batch processes using multiway independent component analysis. *Chemom Intell Lab Syst*. 2004;71:151–163.
4. Ramaker HJ, Sprang ENM van, Westerhuis JA, Smilde AK. Fault detection properties of global, local and time evolving models for batch process monitoring. *J Proc Control*. 2005;15:799–805.
5. 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.
6. Sprang ENM van, Ramaker HJ, Westerhuis JA, Guiden SP, Smilde AK. Critical evaluation of approaches for on-line batch process monitoring. *Chem Eng Sci*. 2008;57:3979–3991.
7. Chen T, Zhang J. On-line multivariate statistic monitoring of batch processes using Gaussian mixture model. *Comput Chem Eng*. 2010;34:500–507.
8. Undey C, Cinar A. Statistical monitoring of multistage, multiphase batch processes. *IEEE Control Syst Mag*. 2002;22:53–63.
9. Lu NY, Gao FR, Wang FL. A sub-PCA modeling and online monitoring strategy for batch processes. *AIChE J*. 2004;50:255–259.
10. Lu NY, Gao FR. Stage-based process analysis and quality prediction for batch processes. *Ind Eng Chem Res*. 2005;44:3547–3555.
11. Liu J, Wang DSH. Fault detection and classification for a two-stage batch process. *J Chemom*. 2008;22:385–398.
12. Camacho J, Pico J, Ferrer A Multi-phase analysis framework for handling batch process data. *J Chemom*. 2008;22:632–643.
13. Yao Y, Gao FR. A survey on multistage/multiphase statistical modeling methods for batch processes. *Ann Rev Control*. 2009;33:172–183.
14. Yu J, Qin SJ. Multiway Gaussian mixture model based multiphase batch process monitoring. *Ind Eng Chem Res*. 2009;48:8585–8594.
15. Zhao CH, Gao FR, Wang FL. Phase-phased joint modeling and spectroscopy analysis for batch process monitoring. *Ind Eng Chem Res*. 2010;49:669–681.
16. Chen J, Liu KC. On-line batch process monitoring using dynamic PCA and dynamic PLS models. *Chem Eng Sci*. 2003;57:63–75.
17. Chen JH, Chen HH. On-line batch process monitoring using MHMT-based MPCA. *Chem Eng Sci*. 2006;61:3223–3239.
18. Tian XM, Zhang XL, Deng XG, Chen S. Multiway kernel independent component analysis based on feature samples for batch process monitoring. *Neurocomputing* 2009;72:1584–1596.
19. Zhang YW, Teng YD, Zhang Y. Complex process quality prediction using modified kernel partial least squares. *Chem Eng Sci*. 2010;65:2153–2158.
20. Reiss R, Wojsznis W, Wojewodka R. Partial least squares confidence interval calculation for industrial end-of-batch quality prediction. *Chem Intel Lab Syst*. 2010;100:75–82.
21. Dayal BS, MacGregor JF. Recursive exponentially weighted PLS and its applications to adaptive control and prediction. *J Process Control*. 1997;7:169–179.
22. Qin SJ. Recursive PLS algorithms for adaptive data monitoring. *Comput Chem Eng*. 1998;22:503–514.
23. Li W, Yue HH, Valle-Cervantes S, Qin SJ. Recursive PCA for adaptive process monitoring. *J Process Control*. 2000;10:471–486.
24. Wang X, Kruger U, Lennox B. Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Eng Pract*. 2003;11:613–632.
25. Jin HD, Lee YH, Lee G, Han CH. Robust recursive principal component analysis modeling for adaptive monitoring. *Ind Eng Chem Res*. 2006;45:696–703.
26. Wang X, Kruger U, Irwin GW. Process monitoring approach usign fast moving window PCA. *Ind Eng Chem Res*. 2005;44:5691–5702.
27. Hwang DH, Han C. Real-time monitoring for a process with multiple operating modes. *Control Eng Pract*. 1999;7:891–902.
28. Zhao SJ, Zhang J, Xu YM. Monitoring of processes with multiple operation modes through multiple principal component analysis models. *Ind Eng Chem Res*. 2004;43:7025–7035.
29. Singhal A, Seborg DE. Evaluation of a pattern matching method for the Tennessee Eastman challenge process. *J Process Control*. 2006;16:601–613.
30. Liu JL. On-line soft sensor for polyethylene process with multiple production grades. *Contol Eng Pract*. 2007;15:769–778.
31. Cheng C, Chiu MS. A new data-based methodology for nonlinear process modeling. *Chem Eng Sci*. 2004;59:2801–2810.
32. Cheng C, Chiu MS. Nonlinear process monitoring using JITL-PCA. *Chem Intel Lab Syst*. 2005;76:1–13.
33. Ge ZQ, Song ZH. Online monitoring of nonlinear multiple mode processes based on adaptive local model approach. *Control Eng Pract*. 2008;16:1427–1437.
34. Fujiwara K, Kano M, Hasebe S, Takinami A. Soft sensor development using correlation-based just-in-time modeling. *AIChE J*. 2009;55:1754–1765.
35. Fujiwara K, Kano M, Hasebe S. Development of correlation-based clustering method and its application to software sensing. *Chem Intel Lab Syst*. 2010;101:130–138.
36. Kadlec P, Gabrys B. Local learning-based adaptive soft sensor for catalyst activation prediction. *AIChE J*. 2010;57:1288–1301.
37. Kano M, Hasebe S, Hashimoto I, Ohno H. Evolution of multivariate statistical process control: application of independent component analysis and external analysis. *Comput Chem Eng*. 2004;28:1157–1166.
38. Ge ZQ, Yang CJ, Song ZH, Wang HQ. Robust online monitoring for multimode processes based on nonlinear external analysis. *Ind Eng Chem Res*. 2008;47:4775–4783.

Manuscript received Dec. 20, 2011, and revision received May 22, 2012.