Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A LASSO-based batch process modeling and end-product quality prediction method

Zhengbing Yan*. Chih-Chiun Chiu*, Weiwei Dong**. Yuan Yao*

*Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
**National Engineering Research Center for Industrial Automation (South China), Guangzhou 511458, China

Abstract: A least absolute shrinkage and selection operator (LASSO) based batch process modeling and end-product quality prediction method is developed in this paper, which overcomes the shortcomings of both multiway partial least squares (MPLS) and the phase-based methods. The proposed phase-LASSO approach models not only the phase characteristics but also the within-phase and between-phase time-dependent information. In the meantime, it automatically selects the critical-to-quality phases via LASSO-type regularization. As a result, phase-LASSO has good prediction accuracy and model interpretation. The effectiveness of the proposed method is illustrated by the case study on injection molding.

1. INTRODUCTION

Batch processes have been widely applied for producing low-volume and high value-added products to meet rapidly changing market demands. Online measurement or prediction of end-product quality is crucially important for achieving good quality control of batch processes. However, real-time measurements of product quality are usually unavailable during batch processing and cannot be achieved through laboratory analysis until the completion of the entire batch. Hence, predictive models have received considerable research interests for online quality prediction.

Multivariate statistical models, such as principal component analysis (PCA) and partial least squares (PLS), have been widely used in process modeling, as they can be derived directly from historical data with little prior process knowledge. Many extensions of the conventional PCA/PLS methods to batch processes have been reported (Kourti, 2003). For end-product quality prediction, multiway partial least squares (MPLS) (Nomikos and MacGregor, 1995a) is the most famous method with good applications in industry (Gunther *et al.*, 2009), which takes the entire batch data as a single object. However, many batch processes contains multiple operation phases, implying that process variables may have different impacts on the final product quality at different phases in a batch run. MPLS does not consider such issue explicitly. Furthermore, MPLS involves all process variables and sampling intervals into the model, no matter they are critical to end-product quality or not. Such characteristic affects both prediction and interpretation of the MPLS model.

Considering that the multiplicity of phases is an inherent nature of most batch processes, multiphase models have been proposed to capture various underlying behaviors of different phases (Yao and Gao, 2009). For online quality prediction, the phase-based PLS modeling method was developed by Lu and Gao (2005), where all measurements within the same phase share the same regression model, i.e. the phase model. Following their work, Zhao *et al.* (2013) recently proposed a between-phase calibration modeling strategy. In their method, different predictive models are developed for different steady phases and transition regions. Although these methods can identify the phases critical to quality improvement and enhance process understanding, there is a major drawback, i.e. the within-phase and between-phase time-dependence information is missing in phase model building.

In this paper, a LASSO-based batch process modeling method named phase-LASSO is introduced to overcome the problems of the existing methods. The full name of LASSO is least absolute shrinkage and selection operator that is a regularization method originally proposed for variable selection (Tibshirani, 1996). To capture the changing process characteristics, the process is divided into several phases with a MPLS-based phase division algorithm. Comparing to the approaches only utilizing the time-slice information, the developed division method better recognizes the changes in process dynamics. After phase division, the process behaviors are well approximated by the proposed phase-LASSO model. Unlike the conventional phase-based regression models, this model takes both the phase characteristics and the time-dependence information into consideration, while the model interpretation is also improved.

The paper is organized as follows. The phase-LASSO batch process modeling method is developed in section 2, followed by the procedure of online quality prediction based on phase-LASSO presented in section 3. In section 4, a case study on an injection molding process is utilized to illustrate the feasibility of the proposed method. Finally, conclusions are drawn in section 5.

2. PHASE-LASSO MODELING OF BATCH PROCESSES

2.1 Data pretreatment

This work was supported in part by the Ministry of Economic Affairs, Taiwan under Grant No. 103-EC-17-A-09-S1-198.
Corresponding author: Dr. Yuan Yao. Tel: 886-3-5713690, Fax: 886-3-5715408, Email: yyao@mx.nthu.edu.tw.

Batch process data are usually stored in a three-way matrix $\underline{\mathbf{X}}(I \times J \times K)$ that should be transformed to a two dimensional form before regression, where I is the number of batches, J is the number of process variables, and K is the number of sampling intervals in each batch. The batch-wise unfolding is the most widely applied transformation method, where $\underline{\mathbf{X}}$ is unfolded into a two-dimensional matrix with I rows and $P = J \times K$ columns. Each row of the unfolded matrix contains all the measurements within a batch. After normalization, the mean of each column is subtracted from the unfolded matrix and the variance of each column is scaled to unit. As a result, the major nonlinearities contained in variable trajectories are removed, while the between-batch variation information is highlighted. The prediction model based on such pretreatment, e.g. MPLS, has chance to capture both the cross- and auto-correlation information among process variables, and reflects the time-dependent information contained in variable trajectories. The phase-LASSO model presented in below will also be based on such data structure.

2.2 Phase-LASSO modeling

A linear prediction model for batch process end-quality prediction is usually formulated as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{X}(I \times P)$ is the batch-wise unfolded and normalized matrix of process variables, $\mathbf{y}(I \times 1)$ is the normalized vector of quality measurements, $\boldsymbol{\beta}(P \times 1)$ is the vector of regression coefficients, and $\boldsymbol{\varepsilon}$ is the residual vector.

The ordinary lease square (OLS) estimate can be obtained by minimizing the residual squared error.

$$\min_{\beta} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2, \tag{2}$$

where $\|\cdot\|_2$ stands for the L2-norm. As well known, OLS is not suited to batch process modeling where the number of rows in **X** is usually far less than the number of columns (*I*<<*P*) and the columns in **X** are highly correlated. That is the reason why PLS instead of OLS is usually adopted in batch process monitoring. MPLS is a typical example. However, in MPLS, all variables and all sampling intervals are involved in the model, no matter they are relevant to product quality or not. As discussed by Lu and Gao (2005), usually not all operation phases are critical to certain product quality. In such situation, variable selection is desired to enhance the prediction accuracy and the model interpretation.

Different from OLS, the regression coefficients in LASSO (Tibshirani, 1996) can be calculated using (3):

$$\min_{\beta} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1, \tag{3}$$

where λ is a tuning parameter, and $\|\cdot\|_1$ represents the L1-norm. The involvement of the L1-norm penalty induces sparsity in the solution, which means that a number of terms in $\boldsymbol{\beta}$ are shrunk to be 0. Such shrinkage is equivalent to variable selection in a sense, and often improves the prediction accuracy due to the bias–variance trade-off. Although LASSO enjoys great computational advantages and excellent performance comparing to conventional variable selection methods, it is not suited to batch process modeling either. The reasons are of two folds. First, in the situation of (*I*<*P*), LASSO selects at most *I* variables. Second, if there are groups of highly correlated variables, LASSO tends to arbitrarily select one variable from each group and results in models difficult to interpret.

In multiphase batch processes, the trajectory points of a variable within the same phase usually contribute similarly to the end-product quality. Therefore, they should be selected into or kept out of the model simultaneously. Based on such finding and motivated by the group LASSO technique (Yuan and Lin, 2006), a phase-LASSO method for batch process modeling is proposed as following.

Supposing the batch process has been divided into S phases, the phase-LASSO model can be built by solving the following optimization problem:

$$\min_{\beta} \frac{1}{2}\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \sum_{j=1}^{S} \sqrt{p_j}\, \|\boldsymbol{\beta}_j\|_2 \tag{4}$$

where $\boldsymbol{\beta}_j\,(j = 1, \dots, S)$ contains the regression coefficients corresponding to the variable trajectory points in the *j*th phase, and p_j is the number of variable trajectory points in the *j*th phase, which takes the different lengths of phases into consideration. The L2-norm term in equation (4) encourages a grouping effect and tends to select or reject the entire phase data together, while the L1 penalty expressed by the sigma notation encourages sparsity at the phase level.

By using phase-LASSO and specifying a proper tuning parameter λ, all regression coefficients belonging to the irrelevant phases can be shrunk to zero at the same time, while the variable trajectory information in the critical-to-quality phases is retained in the regression model. As a result, the model interpretation is improved, while the prediction accuracy is also enhanced due to the noise suppression by selecting meaningful phases.

The optimization problem expressed in (4) can be reformulated as two equivalent smooth convex optimization problems solved via the Nesterov's method. For more detailed algorithm, please refer to (Liu *et al.*, 2009). To select a proper value of the tuning parameter λ, the cross-validation procedure is utilized.

2.3 Phase division based on MPLS

The performance of the proposed phase-LASSO highly depends on the results of phase division. In recent years, a number of phase identification methods have been developed (Yao and Gao, 2009). Most of these methods are based on time-slice data structure and ignore the time-dependence information. To solve this problem, a MPLS-based phase division algorithm is developed as follows.

First, an MPLS model is trained as below, by applying the PLS algorithm to {$\mathbf{X}$, $\mathbf{y}$}:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} , \tag{5}$$

$$\mathbf{y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} . \tag{6}$$

The above MPLS model can be written in a compact form as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{F}^* , \tag{7}$$

where $\mathbf{T}$ and $\mathbf{U}$ are scores matrices, $\mathbf{P}$ and $\mathbf{Q}$ are loading matrices, $\mathbf{E}$, $\mathbf{F}$ and $\mathbf{F}^*$ are residual matrices, and $\boldsymbol{\theta}$ is a vector containing regression coefficients.

In the regression coefficient vector $\boldsymbol{\theta}(P \times 1)$, every J elements represent the contributions of process variables to end-product quality at a certain sampling time interval. Accordingly, $\boldsymbol{\theta}$ can be cut to pieces along the time axis, resulting in K subvectors each of which has dimensions of $(J \times 1)$, i.e.:

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1^T & \boldsymbol{\theta}_2^T & \cdots & \boldsymbol{\theta}_k^T & \cdots & \boldsymbol{\theta}_K^T \end{bmatrix}^T , \tag{8}$$

where, $\boldsymbol{\theta}_k$ contains the regression coefficients corresponding to the kth sampling interval.

In a multiphase batch process, the process data belonging to the same phase usually contribute similarly to product quality, while different phases have different statistical features. As a result, it can be inferred that $\boldsymbol{\theta}_k$ and $\boldsymbol{\theta}_{k+1}$ should have a high degree of similarity, if the time intervals k and k+1 are in the same phase; otherwise, the similarity between these two subvectors is relatively low. Therefore, the range of each phase can be identified by clustering $\boldsymbol{\theta}_k$, $k = 1, 2, \cdots, K$.

Here, different clustering algorithms can be applied. In this paper, the popular k-means clustering algorithm (Jain *et al.*, 1999) is adopted for phase division. The detailed steps are described as following. 1. Set $k = 1$, $m = 1$. Choose $\boldsymbol{\theta}_k$ as the initial cluster center $\mathbf{w}_m$. 2. Calculate the Euclidean distance between $\mathbf{w}_m$ and $\boldsymbol{\theta}_{k+1}$. 3. If the distance is less than the pre-defined threshold δ , assign $\boldsymbol{\theta}_{k+1}$ to class m, and update $\mathbf{w}_m$ by making $\mathbf{w}_m$ equal to the average of all subvectors belonging to class m. Otherwise, add a new cluster center $\mathbf{w}_{m+1} = \boldsymbol{\theta}_{m+1}$, and let $m = m + 1$. 4. Let $k = k + 1$, and go back to step 2 until $k > K$.

As in any clustering algorithm, the threshold δ makes a trade-off between the model complexity and accuracy. Specifically, a large threshold results in few classes, but less accurate modeling.

3. PROCEDURE OF ONLINE QUALITY PREDICTION

After phase-LASSO modeling, quality prediction for a new batch can be conducted as:

$$y_{new} = \mathbf{x}_{new}^T \boldsymbol{\beta} , \tag{9}$$

where y_{new} is the prediction of end-product quality, and $\mathbf{x}_{new}^T (1 \times P)$ contains all the normalized measurements in a new batch, which is prepared in the following way. After collecting the data matrix of a new batch $\mathbf{X}_{new}(J \times K)$, the matrix is unfolded into a vector with dimensions $(1 \times P)$ and then normalized to $\mathbf{x}_{new}^T$ based on the means and variances that have been calculated in the model training steps.

However, such procedure cannot be applied directly to online prediction of end-product quality, since $\mathbf{x}_{new}^T$ contains all process data in the entire cycle, which is unavailable until the completion of the batch. Therefore, future data estimation should be conducted for online application of the prediction model. In previous research, different types of estimation approaches have been proposed, including the "PCA/PLS projection" approach, the "current deviation" approach, and the "zero deviation" approach. (Nomikos and MacGregor, 1995a, b). Most recently, a k-nearest neighbor (kNN) based future data estimation approach was proposed and showed its effectiveness comparing to the traditional methods (Chiu and Yao, 2013). In the case study in the next section, the kNN based approach is adopted for the online estimation of the future unavailable measurements.

4. CASE STUDY

4.1 Injection molding

In the follows, an injection molding process is utilized to illustrate the feasibility of the phase-LASSO method. Injection molding is a typical batch process. Each batch in injection molding consists of several sequential operation phases, which are filling, packing-holding, plastication, and cooling. Each phase has different characteristics and contributes to the end-product quality in different ways. During the first phase, filling, the screw moves forward and injects the melt into the mold cavity. After that, the process switches to the packing-holding phase during which additional melt is packed into the cavity at a high pressure. By doing so, the shrinkage caused by material solidification is compensated. The packing-holding phase lasts for several seconds, followed by the plastication phase. During plastication, the screw rotates and moves back, shearing and melting the material in the barrel and conveying the melt to the front of the screw. In the meantime, the material in the mold cavity is cooled down. Hence, the product can be solidified completely and ejected from the mold easily. Usually, cooling takes longer time than plastication. Fig. 1 illustrates a typical injection molding machine.

In the experiments, the feed material is high-density polyethylene (HDPE). As listed in Table 1, there are eight key process variables measured online, including valve opening, stroke displacement, velocity, pressure, and temperatures. In addition, a quality variable, i.e. product weight, can be measured after the completion of each batch. The operating conditions are listed as following. The set point of the barrel temperature in the third zone is 160 °C. The packing-holding time is fixed to be 5 s. The sampling time interval is 50 ms. For exciting the process and generating necessary regression information for predictive modeling, a 3^3 full design factorial design of experiments (DOE) is conducted, where the three factors to be designed are injection velocity, packing pressure and barrel temperature. The three levels of injection velocity are 25, 35 and 45 mm/s, respectively. For packing pressure, the three levels are set to be 25, 30 and 35 bar. The barrel temperatures in the first two zones are chosen to be the third factor in DOE, which are set to the same values in each level. The three levels are 190, 200 and 210 °C, respectively.

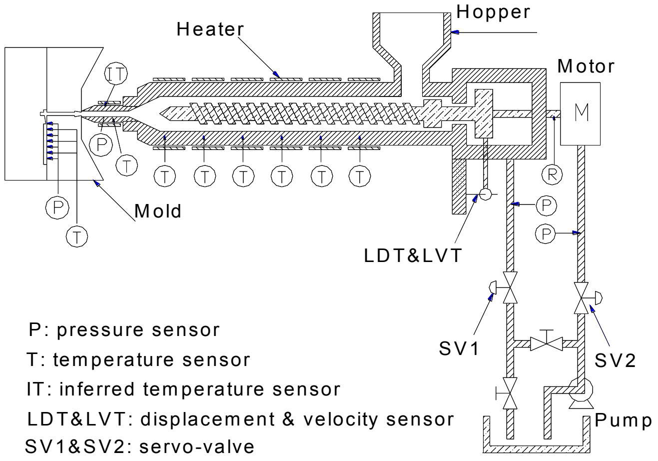

Fig. 1 Simplified schematic of an injection molding machine

There exist 27 different treatment combinations generated by DOE, each with 10 repetitive experiments. Hence, there are totally 270 recorded batches included in the training and testing sets, among which 135 batches are used as the training data for modeling, while the other 135 batches are used to check the effectiveness of the phase-LASSO method. The different injection velocities lead to variations in the filling time. Therefore, data alignment is necessary to be conducted in the filling phase. Here, the process variables are re-sampled in this phase with respect to an indicator variable, the displacement of the screw stroke, instead of operation time. After data alignment, there are totally 503 measurements for each variable in each batch.

Table 1. Process variables

No.	Variable	Units
1	Valve opening 1	%
2	Valve opening 2	%
3	Screw stroke	mm
4	Injection velocity	mm/s
5	Back pressure	bar
6	Barrel temperature 1	^oC
7	Barrel temperature 2	^oC
8	Barrel temperature 3	^oC

4.2 Phase division

Before regression modeling using phase-LASSO, phase division is performed based the proposed method. Fig. 2 shows the division results together with the variable trajectories of valve opening 1 from all 135 batches in the training set. It can be seen that the process with 4 operation phases has been divided into 7 modeling phases. The first modeling phase only contains several sampling intervals and corresponds to the machine start-up in each batch. The second phase represents the filling behavior. The packing-holding operation phase is partitioned into two modeling phases at sampling interval 137, implying that the process characteristics change at this time point even though the operation conditions are set to be unchanged. More details of these two modeling phases will be discussed in section 4.3. The fifth phase corresponds to plastication, during which the screw moves backward at a constant speed from a position near the nozzle to a pre-determined position. Since the screw moves forward by different distances during filling, the durations of plastication is also different. Therefore, the time intervals from 296 to 299 are clustered as the sixth phase, indicating the uneven tails of plastication in different batches. The seventh modeling phase represents the cooling period after plastication. In this phase, no more manipulation is applied to the process except the circulation of the cooling water.

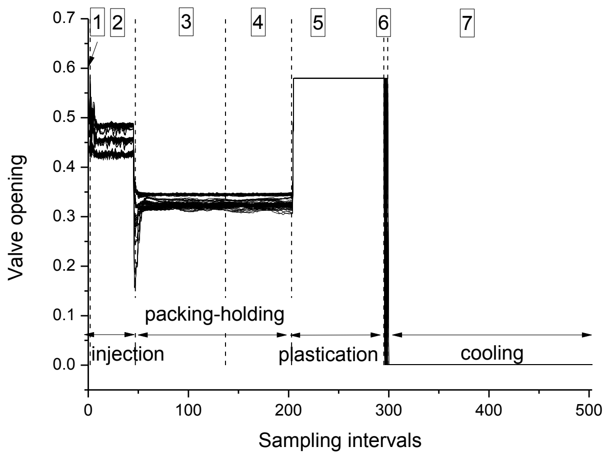

Fig.2 Phase division results based on the proposed method

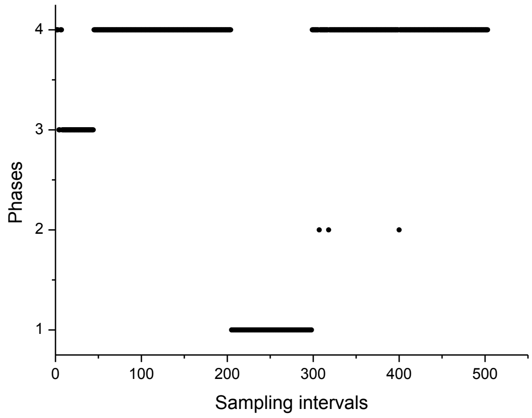

Fig. 3 Phase division results based on time-slice PLS

The results from the time-slice PLS based phase division published by Lu and Gao (2005) are also plotted in Fig. 3 for comparison. It is clearly that the time intervals in the packing-holding phase and the cooling phase are not discriminated, because such method omits the time-dependent information contained in process data. In addition, the changing process characteristics during packing-holding is not revealed.

4.3 *Process modeling, model interpretation, and quality prediction*

In this section, different types of batch process models, including MPLS, phase-based PLS, and phase-LASSO, are compared. The parameters in the phase-LASSO model are determined based on cross-validation.

The regression coefficients of MPLS and phase-LASSO are plotted in Fig. 4 and Fig. 5, respectively. In these figures, the horizontal coordinate corresponds to the sampling intervals, while the vertical coordinate represents the regression coefficients. Four of eight process variables are plotted. Since the valve opening 2 synchronizes with the valve opening 1, only one of them is shown in the figures. The barrel temperature variables are closed-loop controlled throughout the entire batch operation. As a result, the values of these variables and the corresponding regression coefficients are almost constant. Therefore, these variables are not plotted.

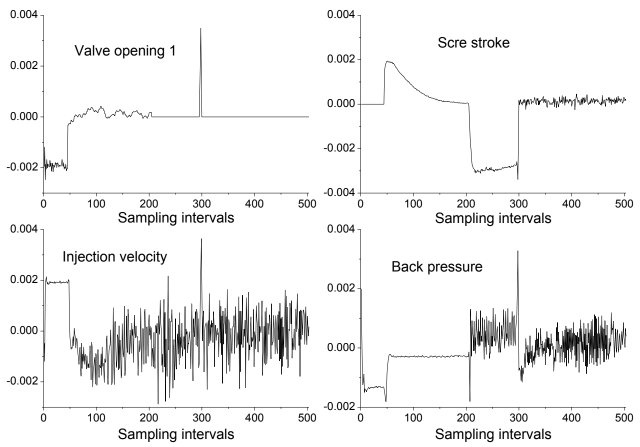

Fig. 4 Regression coefficients of MPLS

In Fig. 4, it is noted that the coefficients corresponding to injection velocity and back pressure looks very "noisy". Especially, in the cooling period, most MPLS coefficients are not 0, indicating the process variables still have chance to influence the product quality during such phase. Obviously, such implication does not confirm the process knowledge. In contrast, the phase-LASSO coefficients in Fig. 5 provide better interpretation.

In Fig. 5, most regression coefficients in the filling phase (i.e. the first and second modeling phases) are not 0, indicating that such phase has critical influences on the product weights. The sign of each coefficient depends upon the direction of the measurement system, where the direction towards the nozzle is defined as the positive direction of both injection velocity and screw stroke displacement. At a first glance, it seems strange that the displacement of screw stroke has little impact on the quality. The reason is as following. As introduced in previous, the process data have been re-sampled with respect to the screw stroke during the filling phase. After normalization, all values of screw stroke become 0. Therefore, the corresponding regression coefficients are also equal to 0. During the third modeling phase, which is the forepart of packing-holding, the regression coefficients show that this phase is also critical to quality. The regression coefficients of screw stroke have large positive absolute values. Taking the coordinate of the measurement system into consideration, this means that larger displacement towards the nozzle indicates lighter final product. Usually, people think that, if the screw moves forward longer, more melt can be conveyed into the mold cavity and the final product should become heavier. In fact, such inference is only half correct. Large values of the screw stroke also indicate small melt density. Since weight is equal to volume multiplied by density, more melt in volume does not necessary mean larger weight. The back pressure shows little contribution to the product weight, which means that the main driving force for packing the melt into the mold cavity is the shrinkage of the material in the cavity instead of the value of back pressure. In the latter part of packing-holding, all regression coefficients become 0. The fundamental behind such phenomenon is that there is no more space in the mold cavity for packing. In other words, the packing has been completed, and the process switches to holding. This is the reason why the packing-holding operation phase is divided into two modeling phases. During the fifth phase, corresponding to plastication, the valve opening is kept constant. Hence, the velocity and back pressure are also constant, and the screw stroke moves backward at a constant speed. Due to the material property of HDPE and the limited packing time in the experiments, the gate between the runner and the cavity does not entirely freeze at the end of packing-holding. When the screw stroke moves backward, the melt in the cavity may reflux, causing loss of weight in the end-products. Therefore, the regression coefficients cannot be regularized to zero in such phase. Since the gate has been completely frozen in the fifth phase, no melt can flow into or out of the mold cavity in the last two phases that corresponds to cooling. The phase-LASSO model well reflects such process knowledge, and sets all the coefficients in these two phases to 0.

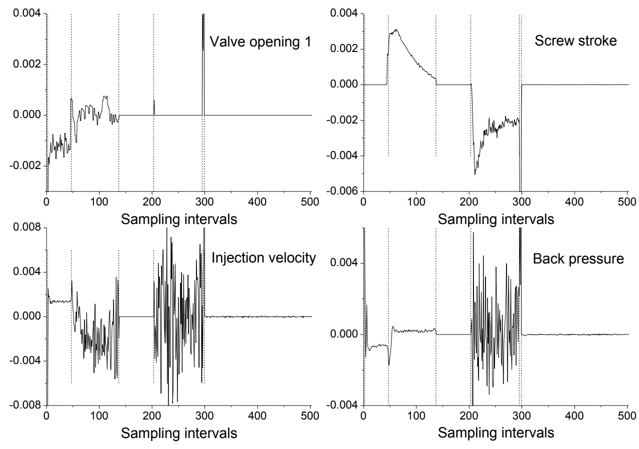

Fig. 5 Regression coefficients of phase-LASSO

In order to evaluate the prediction accuracy of each process model, the root mean squared error (RMSE) is employed. The offline prediction results of the testing batches are summarized in Table 2, showing that the phase-LASSO model outperforms MPLS and stage-based PLS.

Table 2 Offline prediction results

Models	RMSE
MPLS	0.0215
Stage-based PLS	0.0230
phase-LASSO	0.0205

For online quality prediction, the number of nearest neighbors in the kNN algorithm for future data estimation is selected to be 5. Again, phase-LASSO performs best in such online application. Fig. 6 shows the online prediction results for an arbitrary testing batch. It is clear that the proposed phase-LASSO method not only provides more accurate predictions throughout the batch operation, but also makes an early final prediction at the beginning of the cooling phase rather than at the end of the entire batch.

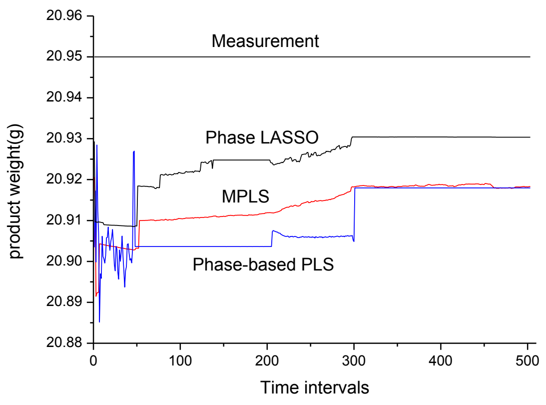

Fig. 6 Online quality prediction for an arbitrary testing batch

5. CONCLUSIONS

For end-product quality prediction in batch processes, MPLS and phase-based PLS are two typical types of modeling methods. However, each method has certain shortcomings limiting their performance in prediction and model interpretation. In this paper, a phase-LASSO approach is proposed to solve the problems of the existing methods. Both phase characteristics and time-dependent information are utilized in phase-LASSO modeling. Especially, by solving a LASSO-type optimization problem, phase-LASSO is able to select critical-to-quality phases, while the regression coefficients corresponding to the irrelevant phases are set to 0 automatically. An MPLS-based phase division method is also proposed. The case study on an injection molding process shows that the phase-LASSO approach outperforms the conventional end-quality prediction methods on both prediction accuracy and process understanding. However, the phase-LASSO approach can still be improved. As seen in Fig. 5, although phase-LASSO removes the information of irrelevant phases from the model automatically, it cannot deal with the noise contained in critical-to-quality phases. As a result, the model interpretation in the noisy phases may be affected. Such problem will be solved in the future research.

REFERENCES

Chiu, C.-C. and Y. Yao (2013). Multiway elastic net (MEN) for final product quality prediction and quality-related analysis of batch processes. *Chemometrics and Intelligent Laboratory Systems,* **125,** 153-165.

Gunther, J.C., J.S.Conner and D.E. Seborg (2009). Process monitoring and quality variable prediction utilizing PLS in industrial fed-batch cell culture. *Journal of Process Control,* **19,** 914-921.

Jain, A., M. Murty and P. Flynn (1999). Data clustering: a review. *ACM computing surveys,* **31,** 264-323.

Kourti, T. (2003). Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions. *Journal of Chemometrics,* **17,** 93-109.

Liu, J., S. Ji and J. Ye (2009). Multi-task feature learning via efficient $l_{2,1}$-norm minimization. *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence.* Montreal, Quebec, Canada: AUAI Press.

Lu, N. and F. Gao (2005). Stage-based process analysis and quality prediction for batch processes. *Industrial & Engineering Chemistry Research,* **44,** 3547-3555.

Nomikos, P. and J. MacGregor (1995a). Multi-way partial least squares in monitoring batch processes. *Chemometrics and Intelligent Laboratory Systems,* **30,** 97-108.

Nomikos, P. and J. MacGregor (1995b). Multivariate SPC charts for monitoring batch processes. *Technometrics,* **37,** 41-59.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* **58,** 267-288.

Yao, Y. and F. Gao (2009). A survey on multistage/multiphase statistical modeling methods for batch processes. *Annual Reviews in Control,* **33,** 172-183.

Yuan, M. and Y. Lin (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **68,** 49-67.

Zhao, C., F. Gao and Y. Sun (2013). Between-phase calibration modeling and transition analysis for phase-based quality interpretation and prediction. *AIChE Journal,* **59,** 108-119.