Monitoring and Diagnosing Batch Processes with Multiway Covariates Regression Models

Ricard Boqué and Age K. Smilde
Process Analysis and Chemometrics, Dept. of Chemical Engineering, University of Amsterdam, Amsterdam 1018 WV, The Netherlands

Multivariate statistical procedures for monitoring the behavior of batch processes are presented. A new type of regression, called multiway covariates regression, is used to find the relationship between the process variables and the quality variables of the final product. The three-way structure of the batch process data is modeled by means of a Tucker3 or a PARAFAC model. The only information needed is a historical data set of past successful batches. Subsequent new batches can be monitored using multivariate statistical process control charts. In this way the progress of the new batch can be tracked and possible faults can be easily detected. Further detailed information from the process can be obtained by interrogating the underlying model. Diagnostic tools, such as contribution plots of each of the variables to the observed deviation, are also developed. Finally, on-line predictions of the final quality variables can be monitored, providing an additional tool to see whether a particular batch will produce an out-of-spec product. These ideas are illustrated using simulated and real data of a batch polymerization reaction.

Introduction

In most industrialized countries, chemical industries are increasingly involved in producing high-value, high-quality specialty products that are produced in batch reactors. Typical examples are the manufacturing of pharmaceuticals, biochemicals, pesticides, or polymers. A batch process is a finite-duration process in which a vessel is charged with materials according to a specified recipe. The reaction takes place during a certain period of time, in which some process variables are measured throughout the batch run. Finally, the resultant product is discharged and some of its quality variables are also measured.

To get a good-quality product, the process variables have to follow certain specified trajectories. If this occurs, the batch is said to be in normal operating conditions. However, disturbances in the process may arise, so that the variables deviate from their specified trajectories. These deviations, caused by different sources (such as the level of impurities in the batch), may eventually lead to a bad-quality product. On-line monitoring and statistical process control (SPC) then become very important as indicators of the real state of these batches. Moreover, abnormal deviations can be detected and corrected for a particular batch or for subsequent batches.

Nomikos and MacGregor (1994 and 1995a) have developed a multivariate statistical process control (MSPC) approach for monitoring batch processes, based on multiway principal component analysis of an historical set of successful batches. However, this approach only uses the information available from the measurements taken on the process variables ($\underline{X}$). Nevertheless, the MSPC philosophy can be extended to the case when product quality measurements (Y) are also available. By relating $\underline{X}$ and Y matrices, the model focuses on the variation of the process variables that most affect the final quality of the product. The same authors applied multiway partial least-squares to model data from a polymerization batch reactor (Nomikos and MacGregor, 1995b).

Analyzing an historical set of batches is a typical three-way problem. The $i = 1, 2, \ldots, I$ batch runs are placed along the first axis, the $j = 1, 2, \ldots, J$ process variables are across the

Correspondence concerning this article should be addressed to A. K. Smilde.
Permanent address for Ricard Boqué: Dept. of Analytical and Organic Chemistry, Rovira i Virgili University of Tarragona, Tarragona 43005, Spain.

second axis, and finally, the $k = 1, 2, \ldots, K$ time intervals occupy the third axis. In both multiway principal component analysis and multiway partial least-squares approaches, the three-way problem of modeling the predictor matrix is done by unfolding $\underline{X}$ in a suitable way and then carrying out traditional PCA and PLS on the transformed two-way matrix. Another approach can be used that considers the intrinsic three-way nature of the data. This approach is called multiway covariates regression, and it is a generalization of principal covariates regression. $\underline{X}$ is decomposed by using either a Tucker3 or PARAFAC model, and simultaneously regression of Y onto $\underline{X}$ is performed. The aim of multiway covariates regression, as in multiway partial least squares, is to build a model based on the measurements of a reference batch database that describes the normal operation of the process ($\underline{X}$) when it produces a good-quality product (Y). Note that since multiway partial least squares and multiway covariates regression impose different *models* on the data, the methods do not differ only in the algorithm employed.

Once the model has been built and validated, it can be used to monitor the behavior of new batches. Monitoring is performed by using statistical process control (SPC) charts. In this way, the t-scores, the predicted final quality variables, and the residuals for the new batch (together with their confidence limits) at each time interval can be calculated. The occurrence of an abnormal batch is easily detected by examining these charts. However, most of the time it is important to know which physical variable (or combination of variables) caused the problem. This can be done by examining the underlying model and checking the contribution of each process variable to the detected deviation. Moreover, predictions of the final batch quality variables can be obtained in real-time, providing a powerful tool for detecting whether the process is under control. Finally, the proposed approach has been applied to both simulated and real data from a polymerization batch reactor.

Multiway Covariates Regression Models

In the following, scalars are written as italic lowercase characters, vectors are bold italic lowercase characters, matrices as bold italic uppercase characters, and three-way matrices as bold underlined uppercase characters. More details regarding notation can be found in the Notation section.

The aim of multiway covariates regression is to build a model based on the measurements of a reference batch database that describes the normal operation of the process when it produces a good-quality product. If these data are available for different batches, they can be arranged in two matrices, $\underline{X}$ ($I \times J \times K$) and Y ($I \times M$), I being the number of batches, J the number of process variables, K the number of time intervals where the process variables are measured, and M is the number of quality variables measured in the final product.

Multiway covariates regression models (Smilde, 1997; Smilde and Kiers, 1999) are a generalization of principal covariates regression developed by de Jong and Kiers (1992). Unlike multiway partial least squares, which is a sequential method, principal covariates regression performs both decomposition and regression steps in a simultaneous way. Matrix $\underline{X}$ can be decomposed by using either a Tucker3 (Tucker, 1963; Kroonenberg, 1983) or PARAFAC (Harshman, 1970) model and simultaneous regression of Y onto $\underline{X}$ is performed. The scheme of the procedure is shown in Figure 1. The Tucker3 model can be written in matrix notation as:

$$X = AH(C^T \otimes B^T) + E, \qquad (1)$$

where for notational convenience, $\underline{X}$ has been rearranged to give matrix X ($I \times JK$) by putting each of the K vertical slices ($I \times J$) of matrix $\underline{X}$ side by side to the right, starting with the one corresponding to the first time interval. A ($I \times R_1$), B ($J \times R_2$), and C ($K \times R_3$) are the loading matrices of the batch, variable, and time modes, respectively. Also, the core matrix, H ($R_1 \times R_2 R_3$), has been rearranged from matrix $\underline{H}$ ($R_1 \times R_2 \times R_3$) in Figure 1a and represents the magnitude of and the interactions between the latent components. The squared sum of its values is the percentage of variance in $\underline{X}$ explained by the model. Finally, E ($I \times JK$) is the matrix of residuals.

On the other hand, the PARAFAC model can also be expressed as in Eq. 1, but in this case the core matrix H is a three-way identity matrix (Smilde, 1992), with ones on the main superdiagonal (see Figure 1b). The main difference between Tucker3 and PARAFAC models is that in the former the number of components in each of the modes (R_1, R_2, and R_3) need not be equal. It is also worth remarking that in Tucker3 models, the matrices A, B, and C are columnwise orthonormal, which is not necessarily the case for PARAFAC models. Finally, another important feature of Tucker3 models is that multiple solutions exist, since the A, B, and C matrices can be rotated to produce the same final residual

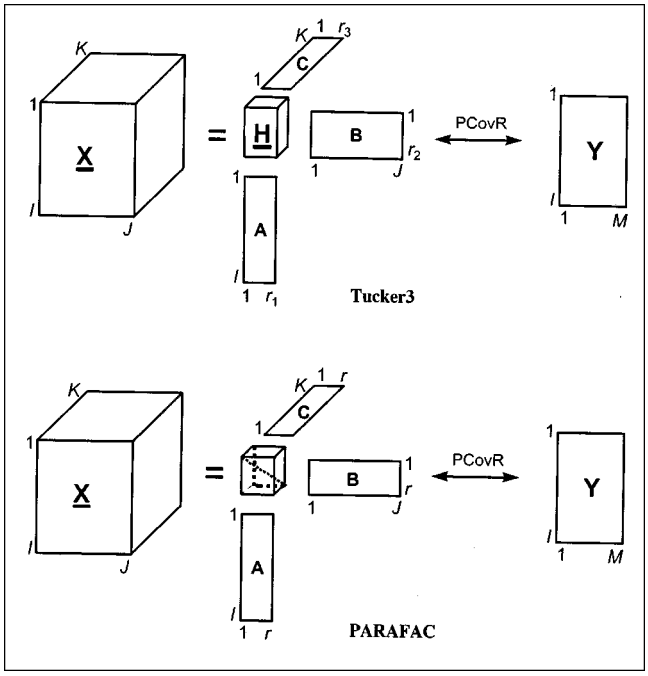

Figure 1. Tucker3 and PARAFAC multiway covariates regression models.

Core matrix in PARAFAC model is a superdiagonal matrix. PCovR is the abbreviation of principal covariates regression.

matrix E. This is not the case for PARAFAC models, since the solution is rotation dependent.

After having defined the Tucker3 and PARAFAC models, the multiway covariates regression model can be summarized in Eqs. 2–5:

$$A = XW \tag{2}$$

$$X = AP^T + E_X \tag{3}$$

$$Y = AQ^T + E_Y \tag{4}$$

$$P^T = H(C^T \otimes B^T), \tag{5}$$

where A is an $(I \times R_1)$ matrix of scores that describes a low-dimensional subspace in the first way (batch domain) of matrix X, and W is a $(JK \times R_1)$ matrix of component weights. The elements in the loading matrices, P $(JK \times R_1)$ and Q $(M \times R_1)$, relate the variables in X and Y, respectively, to the components in A. As can be seen from Eq. 5, P has a special structure that is easily recognized from Eq. 1 and depends on the model (Tucker3 or PARAFAC) that is used. Finally, matrices E_X $(I \times JK)$ and E_Y $(I \times M)$ contain the information in both X and Y that is not correlated with the components in A.

To sum up, the multiway covariates regression model given by Eqs. 2–5 looks for (R_1, R_2, R_3) linear combinations of $\underline{X}$ that explain the maximum amount of variation in both $\underline{X}$ and Y. The model is built by maximizing the following criterion for a given value of α:

$$\max_W \left[\alpha R_X^2 + (1-\alpha) R_Y^2\right] \tag{6}$$

$$R_X^2 = 1 - \|X - XWP^T\|^2 / \|X\|^2 \tag{7}$$

$$R_Y^2 = 1 - \|Y - XWQ^T\|^2 / \|Y\|^2, \tag{8}$$

where R_X^2 and R_Y^2 are the percentages of variance in X and Y, respectively, explained by the model. Both R_X^2 and R_Y^2 can take values between 0 and 1 (1 representing maximal explanation). In batch data, it is a common operation to center or autoscale matrices X and Y. It is also convenient to normalize the data such that $\|X\|^2 = \|Y\|^2 = 1$, to give equal weight when modeling X and Y. By doing this, and substituting Eqs. 7 and 8 in Eq. 6, the maximization criterion can be simplified. Moreover, the model can be built by finding a matrix W that minimizes the following loss function value:

$$\alpha\|X - XWP^T\|^2 + (1-\alpha)\|Y - XWQ^T\|^2. \tag{9}$$

The choice of α is a fundamental step and, as explained in the experimental section, has important implications in monitoring and diagnosis of new batches. A multiway covariates regression model built with a value of $\alpha = 1$ implies (from Eq. 9) that the loss function to minimize is $\|X - XWP^T\|^2$. So, all the effort is put in modeling matrix X (it is the same as performing a normal Tucker3 model on X and then regressing Y onto the score matrix A). This has important consequences, as far as the similarity of W and P matrices is concerned. Since W is wanted to model X, it spans a similar space to loading matrix P. Another important consequence is that when quality data are not available, the method can be easily adapted to model only process data by choosing a value of $\alpha = 1$ and not performing the regression step. On the other hand, if the model is built with a value of $\alpha = 0$, the function value to minimize is in this case $\|Y - XWQ^T\|^2$. This means that the emphasis now lies on fitting Y as much as possible. Note, however, that X has almost always more columns that rows. Consequently, Y is almost always in the column space of X. If Y is in the column space of X, then it can be fitted perfectly. This is clearly a matter of overfitting, and cross-validation prevents this from happening by selecting a proper (and usually high) value of α. As Eq. 9 indicates, the scaling of the data has an influence on the final value of α. If, for instance, autoscaling had been used, the variance in each X or Y block would be proportional to the number of variables of the block. Therefore, X would have more influence on the model and smaller values for α would be expected to compensate for that fact.

The choice of the number of components needed to build a multiway covariates regression model is also highly important. In the case of PARAFAC models, where the number of components is the same for each mode, the problem can be solved by including a cross-validation routine in the developed algorithm. In the cross-validation method, some of the data are left out, a model for each number of components is built with the remaining data, and then a prediction is made using the left-out data. This process is repeated until all the data have been left out once. Finally, the ability of the model to predict is calculated by comparing the predicted vs. the actual values for the validation data and for the different number of components chosen. The optimal model is selected to have the lowest prediction error. For Tucker3 models, the number of components in each mode may be different, and subsequently the cross-validation procedure can be very time-consuming as long as the number of components increases. A quick method to find the optimal model is to perform an eigenvalue decomposition of the squared matrices $X_i X_i^T$ $(I \times I)$, $X_j X_j^T$ $(J \times J)$, and $X_k X_k^T$ $(K \times K)$, where X has been rearranged for each of the three models (Geladi, 1989). The magnitude of the resulting eigenvalues can provide a rough estimate of the number of components in each mode.

Two follow-up remarks are important. First, the goal of a multiway covariates regression model in this application is not primarily predicting the quality variables; it is also important to use the model for monitoring purposes. Nevertheless, it is useful to perform cross-validation to avoid overfitting and to get an idea of the predictive power of the model, that is, whether there is any connection between $\underline{X}$ and Y. Second, in batch processes the process variables and their time histories are usually not independent, that is, they are interacting. It is interesting to see how the different methods deal with this situation. Multiway partial least squares takes all the combinations of process variables and time points, and in that way models the interactions. The multiway covariates regression model that is based on the Tucker3 decomposition deals with the interactions in its core array. Hence, interactions between variables and time histories are modeled on the level of the *manifest* variables by multiway partial least squares,

and on the level of the *latent* variables by multiway covariates regression using a Tucker3 structure. This is an important distinction, and practice shows which works best in which situation. Note that PARAFAC-based multiway covariates regression models can only model interactions to some extent by incorporating extra components.

On-Line Monitoring and Diagnosis

Once the model has been built, it can be used to monitor new batches. The predicted scores and the residuals in the *x*- and *y*-space for a new batch, x_{new}^T $(1\times JK)$, are given by

$$\hat{a}^T = x_{\text{new}}^T \hat{W} \tag{10}$$

$$e^T = x_{\text{new}}^T - \hat{a}^T \hat{P}^T \tag{11}$$

$$f^T = y_{\text{new}}^T - \hat{a}^T \hat{Q}^T, \tag{12}$$

respectively. It is emphasized that the residuals in the *y*-space, f^T, can only be calculated if the quality variables for the new batch are available. By plotting the values of $\hat{a}^T$ and e^T together with the ones obtained for the calibration batches, we can see whether the new batch is out of control. Different types of deviations from normal operation conditions can be observed. If the process is disturbed in the process variables but the covariance structure of the normal operation conditions' data remains the same, then, in the score plot, the value of the new *a*-score will lie far away from the cluster of normal batches, but the residual of the new batch will not show a large value. On the contrary, if a disturbance is produced during the batch that has not been taken into account by the model, the residual for the new batch in the *x*-space will appear significantly different. Furthermore, if the residuals for a new batch in the quality variable space are also large, it means that the relationship between process and quality variables is not valid for that particular batch. There are two possible reasons for that: an instrumental measurement failure has occurred or the process has changed (such as due to deactivation of the catalyst).

Using Eqs. 10–12 for monitoring a new batch has a serious disadvantage: the scores, residuals, and predicted values are obtained postbatch. This means that no action can be taken during the process to correct or even to detect abnormal behavior. On the other hand, the difficulty that one has to face when applying Eqs. 10–12 to monitor the batch process real-time, is that the process variables corresponding to the x_{new} vector are known only up to time k. Different approaches have been proposed to solve this problem (Nomikos and MacGregor, 1995a). The one used in this article consists of filling in the empty values of the process variables for the new batch, from time $k+1$ to time K, with the last deviations from the average trajectories obtained at time k. This procedure was recommended by Nomikos and MacGregor (1994) and assumes that future observations in x_{new} will deviate persistently from their average trajectories at a constant level for the rest of the batch. This is similar to the strategy used in model-predictive control algorithms, where control actions are taken that assume the future values of the disturbances remain constant at their current values over the rest of the process.

If the scores and the residuals are calculated at each time interval, then control charts can be plotted to monitor how the process is behaving. These charts are consistent with the philosophy of statistical process control and are similar to the common and well-known Shewhart charts. The two main control charts that are needed to visualize the process are the *a*-score chart and the squared prediction error of X (SPE_X) chart. The first is based on the significant components that have been chosen to represent the highest variance in the process, and the second one represents the variation that has not been taken into account by the model. The confidence intervals for the control charts are calculated from the reference batches, by applying the same procedure, described earlier, for the new batch to get the scores and the residuals at each time interval (Nomikos and MacGregor, 1995a).

If a new running batch is disturbed in one or more process variables, then the score and/or the SPE value for that batch at time k will be placed outside the control limits in the corresponding chart, and the deviation will be easily detected later. However, the fact that either the score or the SPE charts are out of control is not enough information. It would be interesting to find the process variable (or variables) responsible for this problem. This can be done with the help of the contribution plots (Kourti et al., 1995; Kourti and MacGregor, 1996; Miller et al., 1997). Two different types of contribution can be plotted: the contribution to the scores and the contribution to the squared prediction error. Scores are related to the original variables by Eq. 2, and, for a new batch, the vector of scores, $\hat{a}_k^T$ $(1\times r_1)$, at time k, can be expressed as

$$\hat{a}_k^T = x_{\text{new},k}^T \hat{W}_k = \sum_{j=1}^{J} x_{\text{new},jk} \hat{w}_{jk}^T, \tag{13}$$

where $x_{\text{new},k}$ $(J\times 1)$ is the vector of process variables available at time k, and the scalar $x_{\text{new},jk}$ is the corresponding value of the jth process variable at that time interval; $\hat{W}_k$ $(J\times r_1)$ is the matrix of weights at time interval k for each of the latent variables; and, finally, $\hat{w}_{jk}^T$ is the jth row of matrix $\hat{W}_k$, corresponding to the jth process variable. The individual terms of the right side of Eq. 13 can be seen as contributions of each of the process variables to the final score. The contribution plot is a bar chart of these J contributions at time k for each latent variable.

The squared prediction error for the new batch at time interval k can be expressed as

$$(\text{SPE})_k = \sum_{j=J(k-1)+1}^{kJ} e_j^2 = \sum_{j=J(k-1)+1}^{kJ} \left(x_{\text{new},j} - \hat{x}_{\text{new},j}\right)^2, \tag{14}$$

where e_j^2 is the residual for the jth process variable at time interval k, that is, the difference between the actual value, $x_{\text{new},j}$, and the predicted one, $\hat{x}_{\text{new},j}$. As in the previous case, each of the J terms on the right side of Eq. 14 can be seen as contributions to the SPE of that batch at time k, and can be plotted as bars in a contribution plot to SPE.

Table 1. Process Variables Measured During the Batch Run and Quality Variables Corresponding to the Properties of the Resulting Polymer

Process Variables	Quality Variables
1. Flow rate of styrene	1. Composition of the latex
2. Flow rate of butadiene	2. Particle size
3. Temperature of the feeds	3. Branching
4. Temperature of the reactor	4. Cross-linking
5. Temperature of the cooling water	5. Polydispersity
6. Temperature of the reactor jacket	
7. Density of the latex in the reactor	
8. Total conversion	
9. Instantaneous rate of energy release	

Multiway Covariates Regression Models Applied to Industrial Batch Data

Simulated batch polymerization process

This example is a simulated study of a semibatch reactor for the production of styrene-butadiene rubber (SBR) (Nomikos and MacGregor, 1994). The fact that the simulations are based on a mechanistic model and the disturbances introduced are known and meaningful makes this benchmark data set very useful for testing the methods proposed in this article. The reference data set was chosen to contain fifty batches that gave good-quality products. Nine process variables were measured at 200 time intervals, giving a $(50\times9\times200)$ $\underline{X}$ matrix. Five quality variables that were used to build the matrix Y (50×5) were measured on the final products. Both process and quality variables are listed in Table 1. Additionally, two error batches, with off-spec quality, were simulated. These batches were used to assess the performance of the proposed multiway covariates regression model.

First, matrix $\underline{X}$ $(50\times9\times200)$ was rearranged to X (50×1800), and then the data were autoscaled and normalized so that the total sum of the squares of both the X and Y matrices became one. Both the Tucker3 and PARAFAC models were applied to the data. The number of optimal components as well as the value of α were selected by a leave-one-batch-out cross-validation procedure.

Results from the Tucker3 model are discussed first. From the eigenvalue decomposition of XX^T (X being previously rearranged in each of the three modes) at least two components were seen to be representative for each mode. However, different combinations of components were checked in terms of prediction ability. The results obtained are shown in Table 2.

The joint root-mean-square prediction error of cross-validation (RMSECV) was calculated as:

$$\mathrm{RMSECV}=\sqrt{\frac{\sum_{m=1}^{M}\sum_{i=1}^{I}(y_{mi}-\hat{y}_{mi})^2}{IM}}, \qquad (15)$$

where y_{mi} and $\hat{y}_{mi}$ are the real and predicted values of the mth quality variable in the ith validation sample, I is the total number of validation samples, and M is the number of quality variables. A Tucker3 (2,2,2) model with $\alpha=0.9$ was chosen to be optimal, since including more components did not represent a significant improvement in prediction ability. In Table 3 the percentage of variance explained by the model for the individual quality variables is given for both calibration and cross-validation data. These percentages are calcu-

Table 2. Joint Root-Mean-Square Prediction Error of Cross-Validation of the Five Scaled Variables vs. the Value of α for Different Tucker3 Models

	(2,2,2)	(3,2,2)	(2,3,2)	(2,2,3)	(3,3,2)	(3,2,3)	(2,3,3)	(3,3,3)
$\alpha=0.0$	0.0781	0.0766	0.0781	0.0781	0.0766	0.0766	0.0781	0.0766
$\alpha=0.1$	0.0782	0.0766	0.0782	0.0782	0.0766	0.0767	0.0781	0.0767
$\alpha=0.2$	0.0781	0.0762	0.0781	0.0781	0.0767	0.0764	0.0781	0.0764
$\alpha=0.3$	0.0781	0.0763	0.0782	0.0781	0.0762	0.0767	0.0781	0.0767
$\alpha=0.4$	0.0781	0.0766	0.0781	0.0780	0.0762	0.0771	0.0781	0.0772
$\alpha=0.5$	0.0779	0.0774	0.0777	0.0778	0.0779	0.0778	0.0778	0.0779
$\alpha=0.6$	0.0770	0.0771	0.0767	0.0767	0.0769	0.0760	0.0767	0.0760
$\alpha=0.7$	0.0746	0.0765	0.0742	0.0738	0.0767	0.0733	0.0737	0.0724
$\alpha=0.8$	0.0718	0.0736	0.0718	0.0714	0.0739	0.0689	0.0715	0.0688
$\alpha=0.9$	0.0708	0.0680	0.0708	0.0706	0.0689	0.0678	0.0708	0.0682
$\alpha=1.0$	0.0711	0.0703	0.0711	0.0709	0.0708	0.0699	0.0709	0.0702

Table 3. Percentage of Variance Explained by a Multiway Covariates Regression (MCR) Tucker3 (2,2,2) Model and a Multiway Partial Least Squares (MPLS) Model with Two Factors for the Overall X and Y Matrices and the Individual Quality Variables

		R_Y^2 (%)	y_1	y_2	y_3	y_4	y_5	Y	R_X^2 (%)
MCR Model Tucker 3 (2,2,2)	Calibration set	59.21	27.57	92.09	92.11	52.31	64.66	19.36
	Cross-validation set	52.91	<0	90.78	90.80	43.52	55.10	17.72
MPLS Model (2 factors)	Calibration set	54.30	20.79	91.28	91.29	67.74	65.08	23.02
	Cross-validation set	43.09	<0	88.68	88.69	49.26	53.17	16.22

lated as

$$R^2_{Y,m} = \left(1 - \frac{\|Y_m - AQ_m^T\|^2}{\|Y_m\|^2}\right) \cdot 100 = \frac{\|AQ_m^T\|^2}{\|Y_m\|^2} \cdot 100$$

$$= \frac{\|\hat{Y}_m\|^2}{\|Y_m\|^2} \cdot 100 = \frac{\|Q_m\|^2}{\|Y_m\|^2} \cdot 100, \quad (16)$$

where the subscript m refers to any of the M quality variables, and $\hat{Y}_m$ is the part of Y_m fitted (or explained) by the model. Results are compared with those obtained for a multiway partial least squares model (Nomikos and MacGregor, 1995b), showing that, except for polydispersity, the multiway covariates regression model predicts slightly better. The multiway partial least-squares model fits $\underline{X}$ slightly better than the multiway covariates regression model, but the cross-validation results for $\underline{X}$ do not support this.

Quality variables 3 and 4 (branching and cross-linking of the resulting latex) are very well explained in both cases, whereas variable 2 (particle size) is poorly modeled. It is emphasized that the results in Tables 2 and 3 have been obtained from a model that is based on the optimal joint prediction error. This means that the five quality variables have been modeled at the same time, and different values for the prediction error (and even for the optimal value of the parameter α) would have been obtained had the quality variables been modeled individually. Figure 2 shows a plot of the fitted vs. the measured values for the validation set and for variables 1 and 3, respectively.

The percentage of variance in X explained by the Tucker3 (2,2,2) model, R^2_X, was calculated from Eq. 7, resulting in a value of 20%. This percentage can also be obtained from $\|AP^T\|^2$, $\|P\|^2$, or $\|H\|^2$. Such a low value is normal for this type of data, where each of the variables at each time interval carries a small amount of information. This total variance can be further decomposed to get the percentage of variance explained by the model with respect to variable and time domain (Figures 3a and 3b). These percentages are calculated as:

$$R^2_{X,j} = \frac{\|\hat{X}_j\|^2}{\|X_j\|^2} \cdot 100 = \frac{\|P_j\|^2}{\|X_j\|^2} \cdot 100 \quad (17a)$$

$$R^2_{X,k} = \frac{\|\hat{X}_k\|^2}{\|X_k\|^2} \cdot 100 = \frac{\|P_k\|^2}{\|X_k\|^2} \cdot 100. \quad (17b)$$

The subscripts j and k refer, respectively, to any of the J process variables or the K time intervals. Thus, X_j and X_k are the jth and kth slices of the $\underline{X}$ matrix corresponding to the variable or time modes, and $\hat{X}_j = AP_j^T$ and $\hat{X}_k = AP_k^T$ are, respectively, the parts of X_j and X_k fitted by the model; and P_j ($K \times R_1$) and P_k ($J \times R_1$) are the loading matrices corresponding, respectively, to the jth process variable and the kth time interval. The mean of all $R^2_{X,j}$ (for $j = 1, \ldots, J$) and $R^2_{X,k}$ (for $k = 1, \ldots, K$) gives R^2_X.

By observing Figure 3a it can be clearly seen that variables 7 (density of the latex in the reactor) and 8 (total conversion) are the best explained by the model, while variables 1, 3, and 4 (flow rate of styrene, temperature of the feeds, and temperature of the reactor, respectively) are practically not explained at all. The columns of loading matrix B (Figure 3c) reveal a slight correlation with the explained variance, which seems logical, since a high value in the loading for a particular variable has an effect on the explained part of X. It is seen that variables 7 and 8 have the highest loading values for the first component in the variable mode, while variables 5, 6 (temperatures around the reactor), and 9 (rate of energy release) are the dominant ones in the second component. Regarding the time domain (Figure 3b and 3d), it is noted that the highest percentage of explained variance is concentrated in the second half of the batch process. Matrix C (Figure 3d) shows the same tendency. However, its columns are not directly interpretable in terms of variance, since they are multiplied by matrices B, H, and A to get the predicted matrix X. These results agree with the ones described in the article of Nomikos and MacGregor (1994a).

The results obtained for the PARAFAC models are presented in Table 4. A model with two components and a value of $\alpha = 0.9$ was found to be optimal in terms of prediction. Including another component in the PARAFAC model overfits the data. This can be seen by observing the plots of matri-

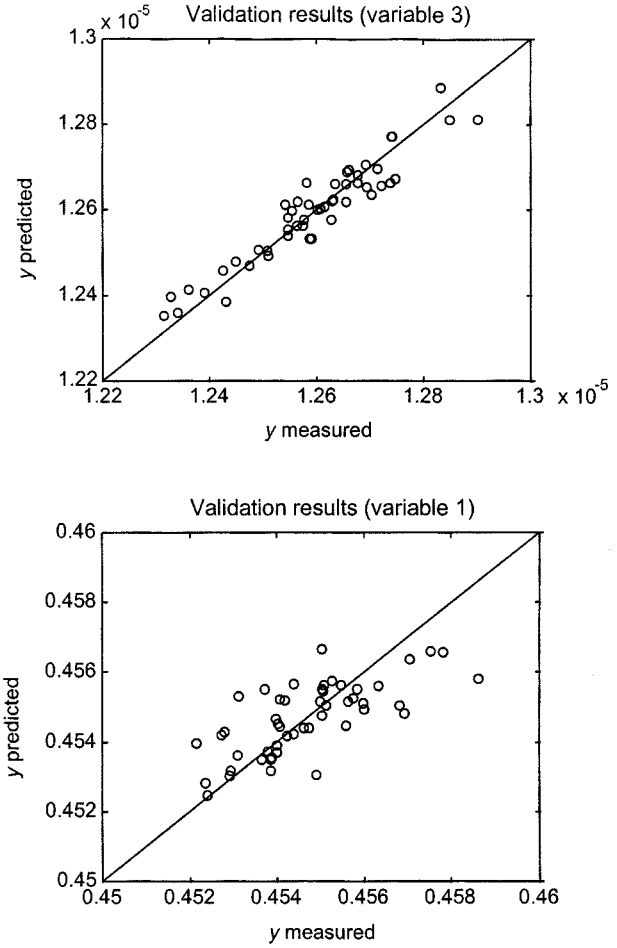

Figure 2. Predicted vs. measured quality variables for the cross-validation set.

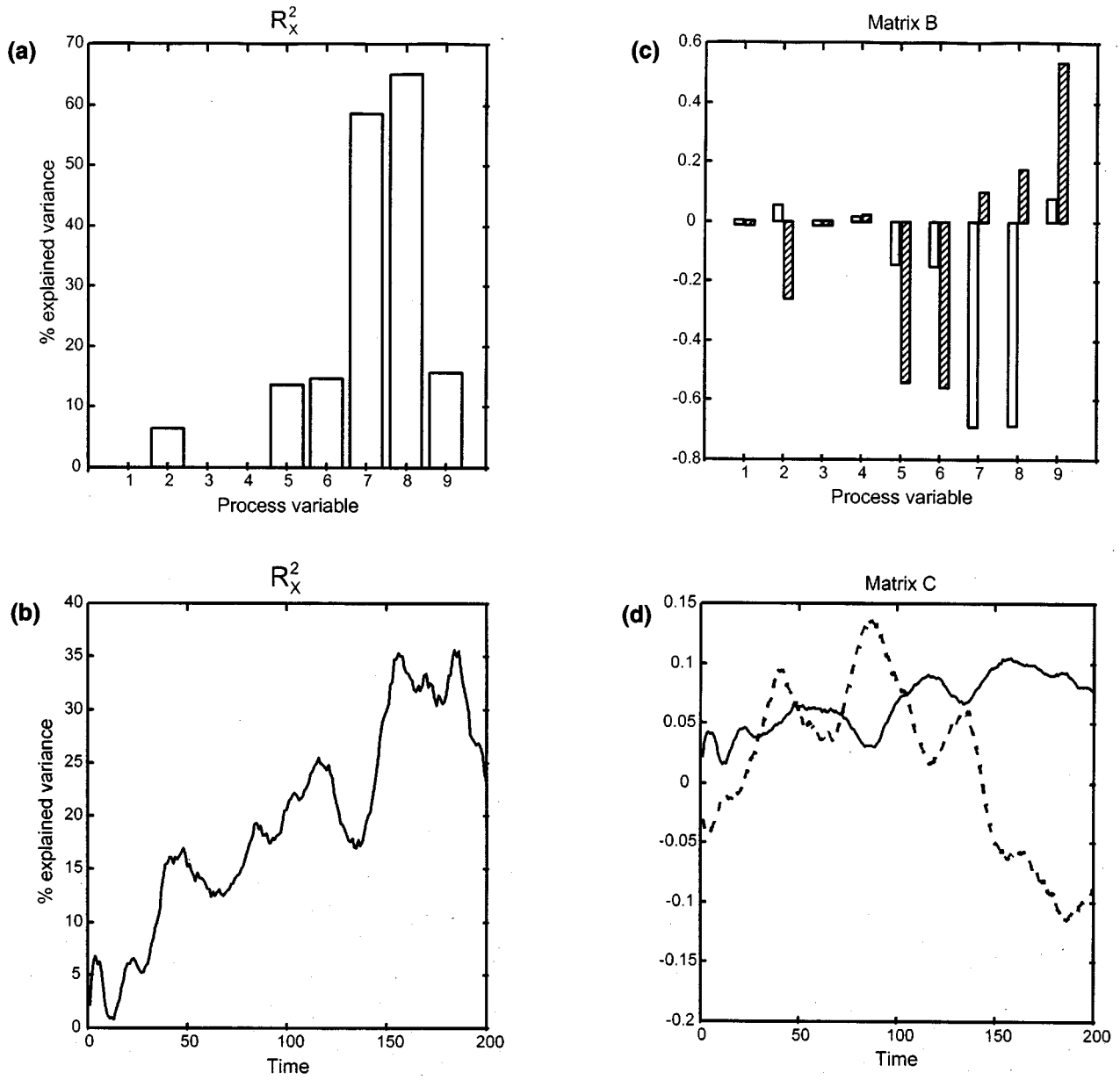

Figure 3. Results from the Tucker3 (2,2,2) model.

(a) Percentage of explained variance for each of the process variables; (b) percentage of explained variance at each time interval; (c) matrix B ($J \times R_2$) for the first (white bars) and second (striped bars) components in the variable domain; (d) matrix C ($K \times R_3$) for the first (solid line) and second (dashed line) components in the time domain.

ces B and C for each of the three components (Figure 4). Note that the third column of each of the matrices is highly correlated with the second one, showing that no substantial information is modeled by this third component.

As can be seen from the results in Tables 2 and 4, the prediction results of both optimal Tucker3 and PARAFAC are practically the same. Moreover, the percentages of variance explained by the PARAFAC (Eq. 2) model for both the calibration and validation sets are practically identical to the ones given in Table 3 for the Tucker3 model. Thus, for the rest of this article, the different results and plots will be shown only for the Tucker3 (2,2,2) model. Modeling the process data with a Tucker3 model gives more freedom with respect to the number of components in each mode than the PARAFAC modeling does. It happened to be in this particular example that a Tucker3 model with the same number of components was found to be optimal. An optimal Tucker3 model with an unequal number of components would have made the difference from PARAFAC clearer.

Figures 5 shows the score plot in the space of the first two latent vectors (a_1, a_2) and the residual sum of squares [$Q_i = \sum_{c=1}^{JK} e(i,c)^2$] for the fifty good batches. Scores for the two error batches, calculated from Eq. 10, have also been included. The plot shows that the two error batches (51 and 52) are placed far away from the main cluster of good batches. However, only one of the abnormal batches (51) is clearly

Table 4. Root-Mean-Square Prediction Error of Cross-Validation of the Five Scaled Variables vs. the Value of α for the Different PARAFAC Models

PARAFAC	1 Comp.	2 Comp.	3 Comp.
$\alpha = 0.0$	0.0722	0.0781	0.0766
$\alpha = 0.1$	0.0721	0.0781	0.0766
$\alpha = 0.2$	0.0721	0.0781	0.0767
$\alpha = 0.3$	0.0720	0.0781	0.0769
$\alpha = 0.4$	0.0719	0.0781	0.0771
$\alpha = 0.5$	0.0718	0.0779	0.0785
$\alpha = 0.6$	0.0717	0.0770	0.0790
$\alpha = 0.7$	0.0716	0.0746	0.0764
$\alpha = 0.8$	0.0715	0.0718	0.0700
$\alpha = 0.9$	0.0716	0.0708	0.0687
$\alpha = 1.0$	0.0722	0.0711	0.0713

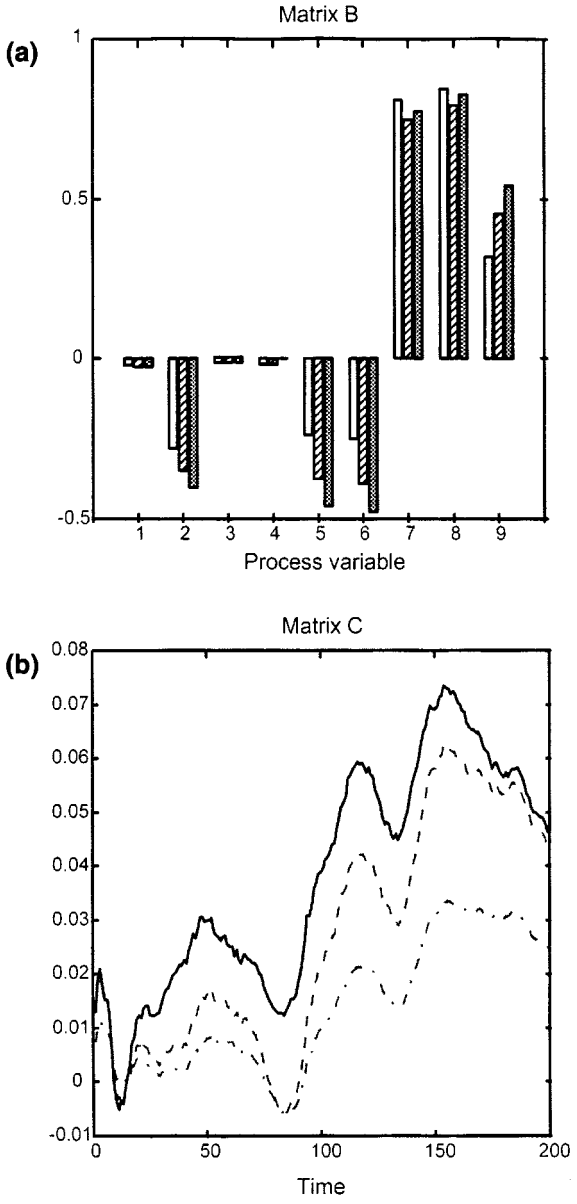

Figure 4. Matrices B $(J \times R)$ and C $(K \times R)$ for a PARAFAC (3) model.

White, striped, and criss-crossed bars, as well as (—), (– –) and (–·), represent first, second, and third component, respectively.

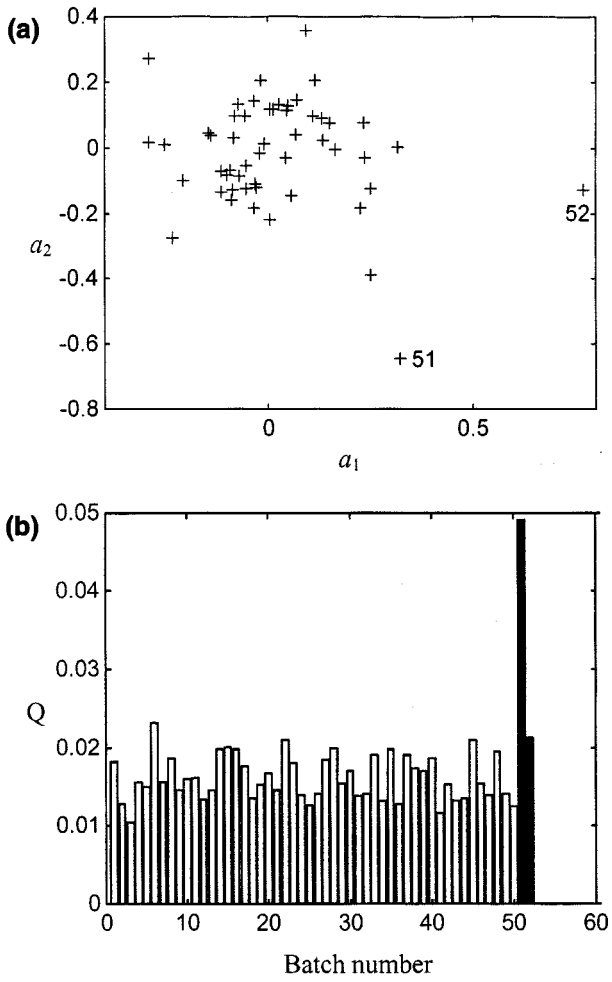

Figure 5. Score plot (a) and Q plot (b) for the 50 reference good batches plus the two error batches.

Results obtained from the Tucker3 (2,2,2) model. Batch 51 had an organic impurity contamination in the butadiene feed started halfway through the batch operation. Batch 52 had the same contamination but at the start of the batch process.

detected as having a large residual in the Q-plot. For the second error batch (52), then, a deviation of the mean trajectories for one or more process variables has occurred, but this change has not affected the correlation structure of the components and therefore its residual is comparable to the ones from the reference batches.

Monitoring and Diagnosing New Batches. To illustrate the usefulness of the developed methodology, the operation of a new running batch was monitored. It is an "error" batch (batch 51) in which the level of impurities in the butadiene feed increased halfway through the process, resulting in a final product whose quality does not meet the specifications. Figure 6 shows the SPC on-line monitoring charts for this abnormal batch. Three different types of chart are given: the scores and SPE_X charts, plus the joint $(a_1 - a_2)$ score plot. The latter chart is not strictly necessary, but it serves as a complement for the other two.

These types of charts are very useful for monitoring and diagnosing new batches. As discussed previously, however, a

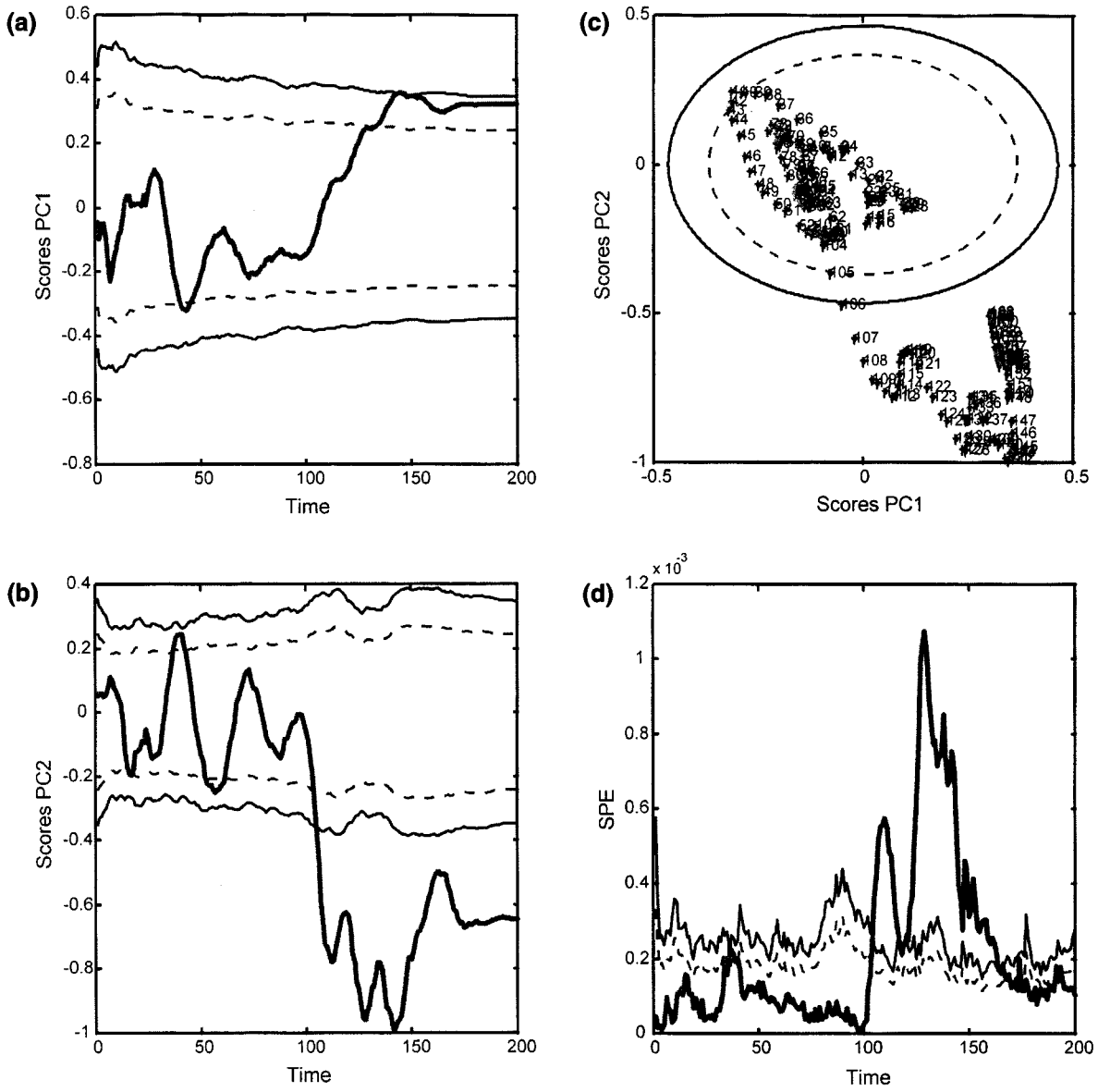

Figure 6. On-line monitoring MSPC charts with their 95% and 99% control limits (dashed and solid lines) for the running error batch.

(a), (b) Scores for first and second component; (c) joint $(a_1 - a_2)$ score plot; (d) SPE chart.

difficulty arises in the calculation of the scores and SPE values, since the measured process variables for the new batch are only available up to time k. In Figure 6, these values were calculated by assuming that the measurements from time $k+1$ to K remain constant at their current values at time k, as previously stated.

The control limits for each plot are built at a 95% and 99% significance level, respectively, based on the distribution of the reference batches. The scores and SPE values for each of the reference batches were calculated at each time interval k by following the same procedure for finding the missing values described earlier.

The confidence intervals for the scores were calculated by assuming that the fifty observation batches of the reference sets at each time interval follow a normal distribution. Normal probability plots of the scores showed that this assumption can be made safely. Control limits associated with the scores were then derived using statistics based on the normal distribution (Hahn and Mecker, 1991).

The squared prediction error for a new batch at each time interval is calculated from Eq. 14. That quadratic form is known to follow a weighted chi-squared distribution, $g\chi^2_{h,\alpha}$. The control intervals for SPE are calculated by matching the moments (mean and variance) between the distribution just mentioned and the distribution of SPE for the reference batches at each time interval (Nomikos and MacGregor, 1995a).

Figures 6a and 6b show the evolution of the scores for the first and second components, respectively. Figure 6c shows the joint plot of the scores for the two components $(a_1 - a_2)$,

while in Figure 6d, the progress of the SPE is monitored. A quick inspection of the plots reveals that something happens at time interval 105, approximately. There is a deviation in the average trajectory of the scores, more pronounced in the second component, and also an increase in the SPE values. It is known that halfway through the batch operation this batch began developing an organic impurity contamination in the butadiene feed that is 50% above normal. The problem is clearly seen in the SPE plot, which means that this variation has not been modeled in the calibration data set. Similar results were obtained by Nomikos and MacGregor (1995b) by applying multiway partial least squares to the same data.

However, the fact that either the score or the SPE chart is out of control is not enough information. It would be quite interesting to find the process variable (or variables) responsible for this problem. This can be determined on-line by looking at the contribution plots for each variable at time interval 105 (Figure 7). Figures 7a and 7b reveal that variables 4, 5, and 6 are the most important ones. Variable 4 (temperature of the reactor) shows the highest contribution, which is somewhat surprising, since this variable, as seen in Figure 3, is not explained at all by the model. However, a deeper insight in the model (Eqs. 2–5) reveals that a contribution to the score for a certain variable is not directly related to its loading. This can be illustrated by looking at Figure 7c and 7d, where the weight matrix W and the loading matrix P at time interval 105 have been plotted. The sum of the squared values of matrix P would give the variance ex-

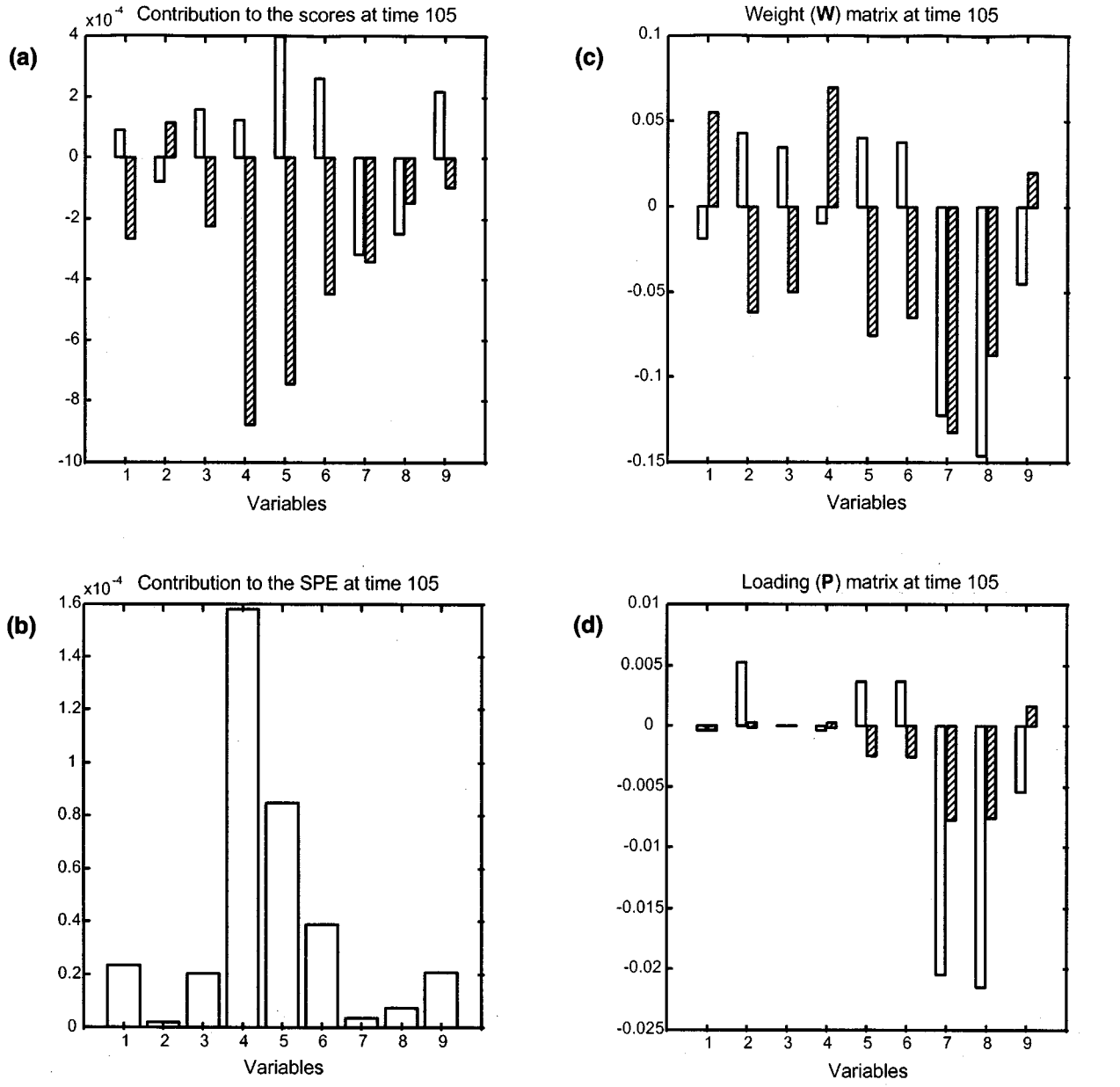

Figure 7. Contribution plots and model matrices for the error batch at time 105.

(a) Contribution to the scores; (b) contribution to the SPE; (c) weight matrix, W ($JK \times R_1$); and (d) loading matrix, P ($JK \times R_1$). White and striped bars in (a), (c), and (d) represent the first and second components, respectively.

plained for each variable at that time interval. It is seen that variable 4 is not explained at all. However, the bar plot of matrix W shows that this variable has a nonnegligible value. Thus, variable 4 (the temperature of the reactor) has no influence in estimating matrix X, because it shows that there is very little variation in the normal operation of the process. In fact, the temperature's intervals spanned by the reactor in a batch process is only about 0.6°C. Any deviation in the latent process variables is mainly reflected in the value of other process variables, but not in the temperature of the reactor. Any small variation in this temperature profile over different batches, however, has some influence on the final quality of the product; this is reflected in the weight matrix. Thus, a certain correlation exists between variable 4 and any or all of the columns in matrix Y.

A closer look at Figure 6 shows that after the first deviation, the process tries to recover normality, but after time interval ~ 120, there is a sudden increase in the SPE value (scores continue beyond specifications). By looking at the contribution plots (Figures 8a and 8b) at time interval 125, we can see that variable 9 is the most important, but that variables 5 and 6 also become significant. The loading and weight matrices (Figure 8c and 8d) show variables 7 and 8 to be the ones that best explain the variation both in X and Y at that time interval. However, this is reflected in neither the contribution to the scores, nor in the contribution to the SPE. A new type of variation (in this case a contamination in the butadiene feed) has been encountered that was not present in the reference batches. This abnormal variation, due to the increase in the reactor impurities, decreases both the poly-

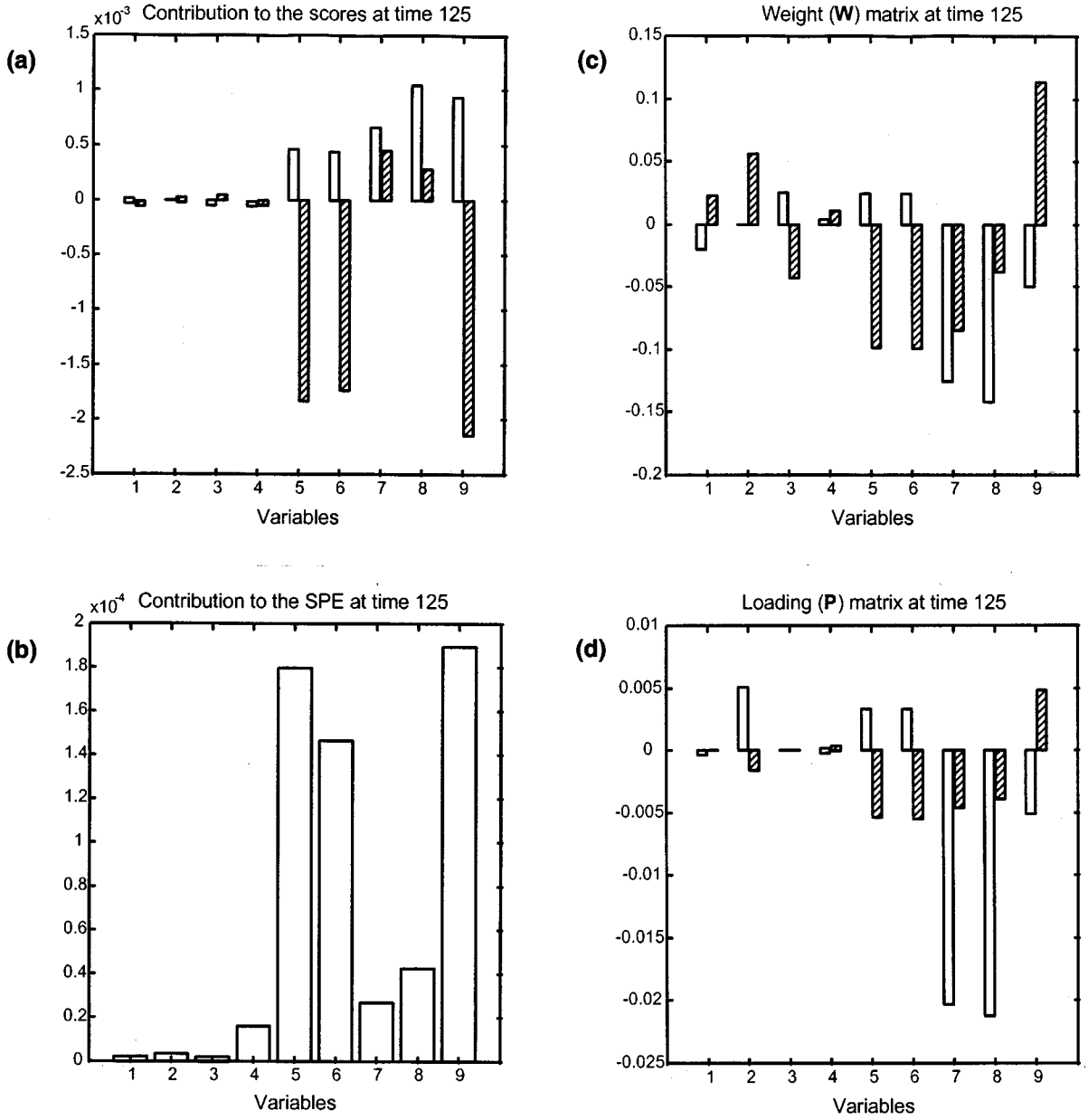

Figure 8. Contribution plots and model matrices for the error batch at time 125.

(a) Contribution to the scores; (b) contribution to the SPE; (c) weight matrix, W ($JK \times R_1$); and (d) loading matrix, P ($JK \times R_1$).

merization rate and the resulting energy release (variable 9). In turn, to maintain constant heat flow, the temperatures of the cooling water and the reactor jacket increase (variables 5 and 6, respectively).

By studying the contribution plots, a deeper knowledge of the process can be achieved than if the usual loading plots are used. This is because contribution plots represent the particular process variables that are unusual at each point for a given batch while the loadings represent the variability across the entire data set. Furthermore, loadings only span the *x*-space, while the contribution to the scores also includes information from the quality variables. In this particular example, however, the differences between the space spanned by W and the one spanned by P are small. Correlation coefficients of 0.92 and 0.83 were found for each component between the W and P columns, respectively. This is due to the fact that a value for $\alpha = 0.9$ has been chosen as optimal, which means that much emphasis is put on modeling X, regardless of its correlation with Y.

Predicting Quality Variables. Apart from on-line monitoring and diagnosing new running batches, the quality variables also can be predicted from the multiway covariates regression model at each time interval. Monitoring these quality variables in time can also help to find abnormal trends and suggest possible causes. Different strategies for monitoring the predicted quality variables are available, which are related to the way the future measurements for x_{new} at time k are filled in. The advantages and drawbacks of these approaches are described by Nomikos and MacGregor (1995b). Although these approaches provide the same Y predicted at the end of the batch, the shape of the curves can vary substantially. The procedure chosen in this article has already been described. Figure 9 shows the Y-predicted monitoring chart for quality variables 2 and 3 of the error batch. Quality variable 2 was very poorly explained by the model and it is also poorly predicted. On the other hand, variable 3, which was explained well, shows a prediction very close to the actual value. It has to be emphasized that, once a fault has been detected by inspecting the SPC charts, and the system is out of normal operation conditions, the plots for the predicted quality variables have to be used with caution. The predicted values may be inaccurate; however, the trends of the curves can be very helpful in showing whether the final product will be off-spec.

It could also happen that a batch showing a significant variation in the residuals results in a good-quality product. At this point, a question arises: Would it be convenient to build a model that also includes error batches? By doing this, the model would span a larger *x*- and *y*-space, and consequently, it would be able to predict (in some cases) the product quality, even when there is an unusual operation. However, imagine that a model is available that takes into account all possible disturbances that can appear in the process. Such a model would probably make good predictions in the *y*-space, but would fail in diagnosing abnormal situations, because every fault would not produce a significant change in the scores or the residual values. When the purpose of the model is monitoring, the null hypothesis, H_0 (the batch is in control), has to be tested against the alternative hypothesis, H_1 (the batch is out of control). In this case, any abnormal situation should be detected, so a reference data set of only good batches is needed. Moreover, a predictive model based on a reference data set that also includes error batches could not be used to monitor, because the assumption of normality of the scores and the residuals would be violated if the error batches are included.

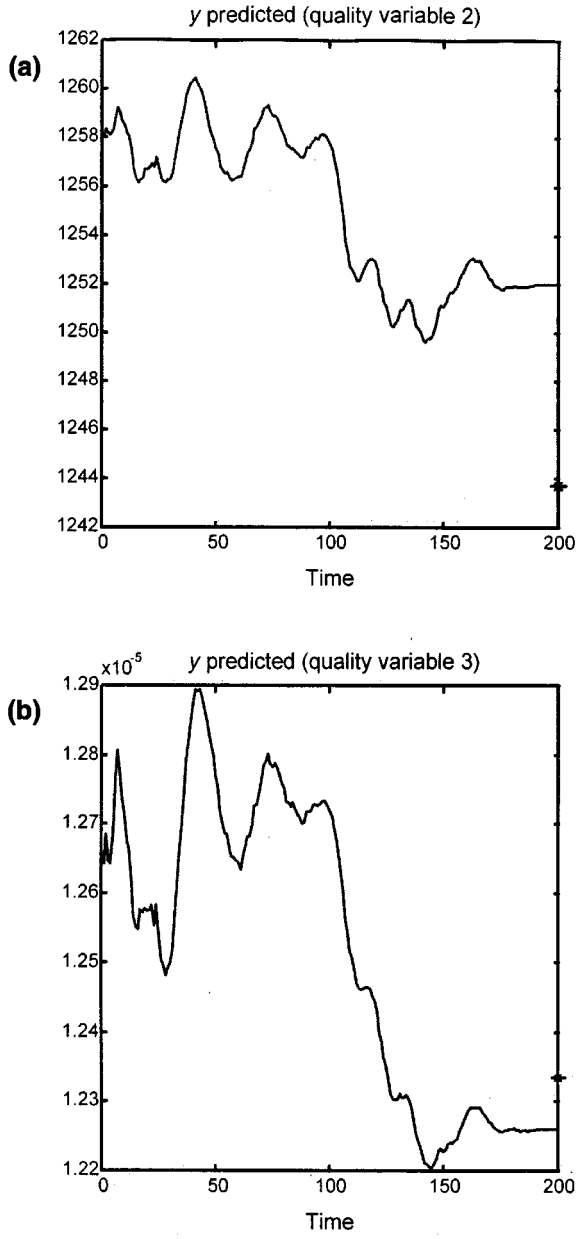

Figure 9. On-line predictions for two of the five quality variables.

Asterisk marks on the right side of each plot correspond to the actual final product quality.

The preceding remarks indicate the following strategy. For a given batch process two models should be made. The first model is completely geared for predicting Y, and the second model is focused on building control charts. Since the goals of these models are very different, the models may be different. In its most extreme form, the second model might not even consider Y. This strategy is the subject of further research.

Table 5. Process Variables Measured During the Batch Run and Quality Variables Corresponding to the Properties of the Resulting Polymer

Process Variables	Quality Variables
1. Polymer center temperature	1. Molecular weight
2. Polymer side temperature	2. Titrated ends
3. Vapor temperature	
4. Autoclave body pressure	
5. Heat-source supply temperature	
6. Heat-source jacket-vent temperature	
7. Heat-source coil-vent temperature	
8. Heat-source supply pressure	
9. Heat-source pressure-control checkpoint	

Commercial batch polymerization process

The second example is a real polymerization batch process. Data were supplied by Dupont and were also previously referred to by Nomikos and MacGregor (1995) and Kosanovich et al. (1996). The data set consists of 50 batches, from which 9 process variables are measured over approximately 120 time intervals. From this set of batches, 2 quality variables of the final product were also available. Both process and quality variables are listed in Table 5. The autoclave in this chemical process converts the aqueous effluent from an upstream evaporator into a polymer product. The recipe specifies reactor and heat-source pressure trajectories through five stages. Since the number of measurements varies at each stage throughout the different batches, the raw data set was linearly interpolated to get, for each batch, the same number of measurements for each process variable at each stage. First, for every batch the number of time points was counted at each stage, and then the mean values for all the batches were calculated. Next, every variable in a particular batch was interpolated by a linear function according to the number of data points at each stage for that batch. Finally, for every batch the same number of data points was extracted from the linear function at each stage. This number was taken to be the mean value for all the batches, but any other fixed number can be chosen. In this way, a total number of 116 time intervals resulted for every batch (9, 43, 22, 20, and 22 time intervals, respectively, at each stage). Other procedures to align batch data exist, namely, taking a surrogate variable, such as the percentage of conversion, that follows a certain function and fitting the measurements to get the same time points for each batch. However, the procedure used was found to produce good results for this type of process data.

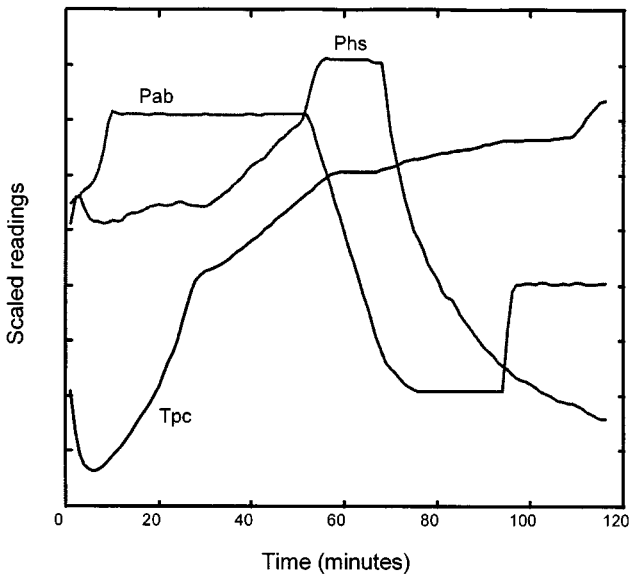

Figure 10. Example of autoclave profiles.

Phs: heat source supply pressure; Pab; autoclave body pressure; and Tpc: polymer center temperature.

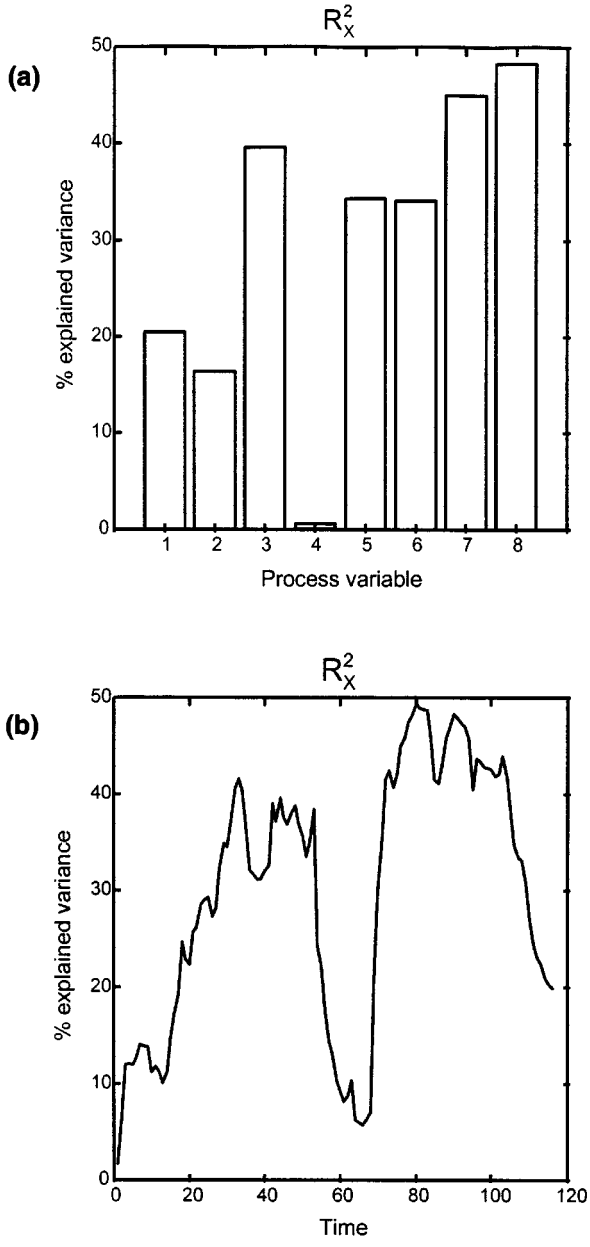

Figure 11. Results from the Tucker3 (3,2,3) model.

(a) Percentage of explained variance for each of the process variables; (b) percentage of explained variance at each time interval.

Figure 10 shows the operating profiles for some of the variables measured in the process. The autoclave body pressure profile clearly shows the five different stages into which the process is divided. Variable 9, the heat-source pressure control checkpoint, was skipped since it was found to be redundant. Batches 46 and 47 were also skipped because their y-values were of borderline quality. An earlier multiway partial least-squares model revealed that these two batches lie far away from the main cluster of normal batches. A multiway covariates regression Tucker3 model was applied to the final $\underline{X}$ $(48\times 8\times 116)$ and Y (48×2) matrices, previously autoscaled and normalized to $\|\underline{X}\|^2 = \|Y\|^2 = 1$. The number of optimal components and the value of α were selected by a leave-five-batches-out cross-validation procedure. A multiway covariates regression Tucker3 (3,2,3) model with $\alpha = 0.9$ was found to be optimal in terms of predicting both quality variables.

Figure 11 shows the percentage of the variance explained by the Tucker3 (3,2,3) model with respect to variable and time domain. It can be seen in Figure 11a that process variables 3, 7, and 8 are the best explained by the model, while variable 4 (autoclave body pressure) is practically not explained at all. In fact, it is a highly controlled variable, and a small variation should be expected from batch to batch. Figure 11b shows that stages 2 and 4 of the process are the ones that show most variability. Further investigation revealed that, in stage 4, the heating medium was turned off during the remaining processing time and, from that time, the system operated adi-

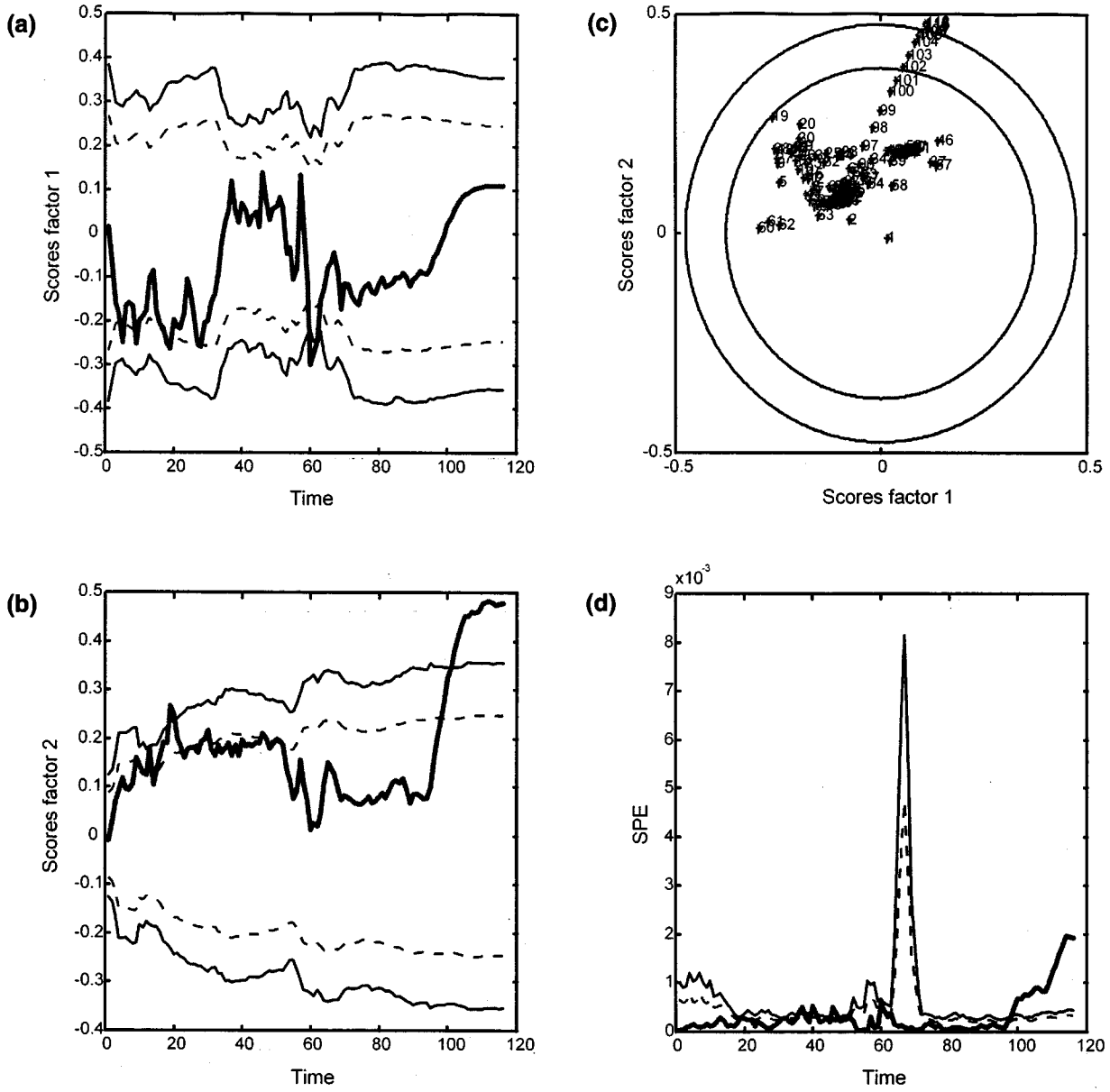

Figure 12. On-line monitoring MSPC charts with their 95% and 99% control limits (dashed and solid lines) for batch 47.

(a), (b) Scores for first and second component; (c) joint $(a_1 - a_2)$ score plot; and (d) SPE chart.

abatically and the reaction was self-sustained. This nonregulated action adds an important source of variability, which is reflected in the sudden increase in the explained variance from time interval of ~ 70.

The behavior of a new batch was monitored based on the model built on the 48 successful batches. Batch 47 was chosen to illustrate this. Although it was not reported as failing to meet product specifications, its quality variables were found to be far from average. The same linear interpolation procedure described earlier was used to generate the 116 time intervals for that batch. Note that this was possible because the measurements for that batch were known in advance. Another procedure has to be adopted for real on-line monitoring of this batch process. Figure 12 shows the SPC on-line monitoring charts for this batch. Figures 12a and 12b show the progression of the scores for the first and second components, respectively (score chart for the third component did not provide additional information). The joint plot of the scores for the two components $(a_1 - a_2)$ is monitored in Figure 12c. Finally, Figure 12d shows the evolution of SPE in time. Apart from slight deviations in the early stages revealed by both scores (for the second component) and SPE charts, there is a clear deviation from the normal operating conditions, starting at time 100 approximately. This corresponds to the beginning of the fifth stage, in which the polymer is discharged by pressurizing the reactor. The sharp peak in the confidence intervals of SPE results from one of the reference batches, whose heating medium pressure reduction began in

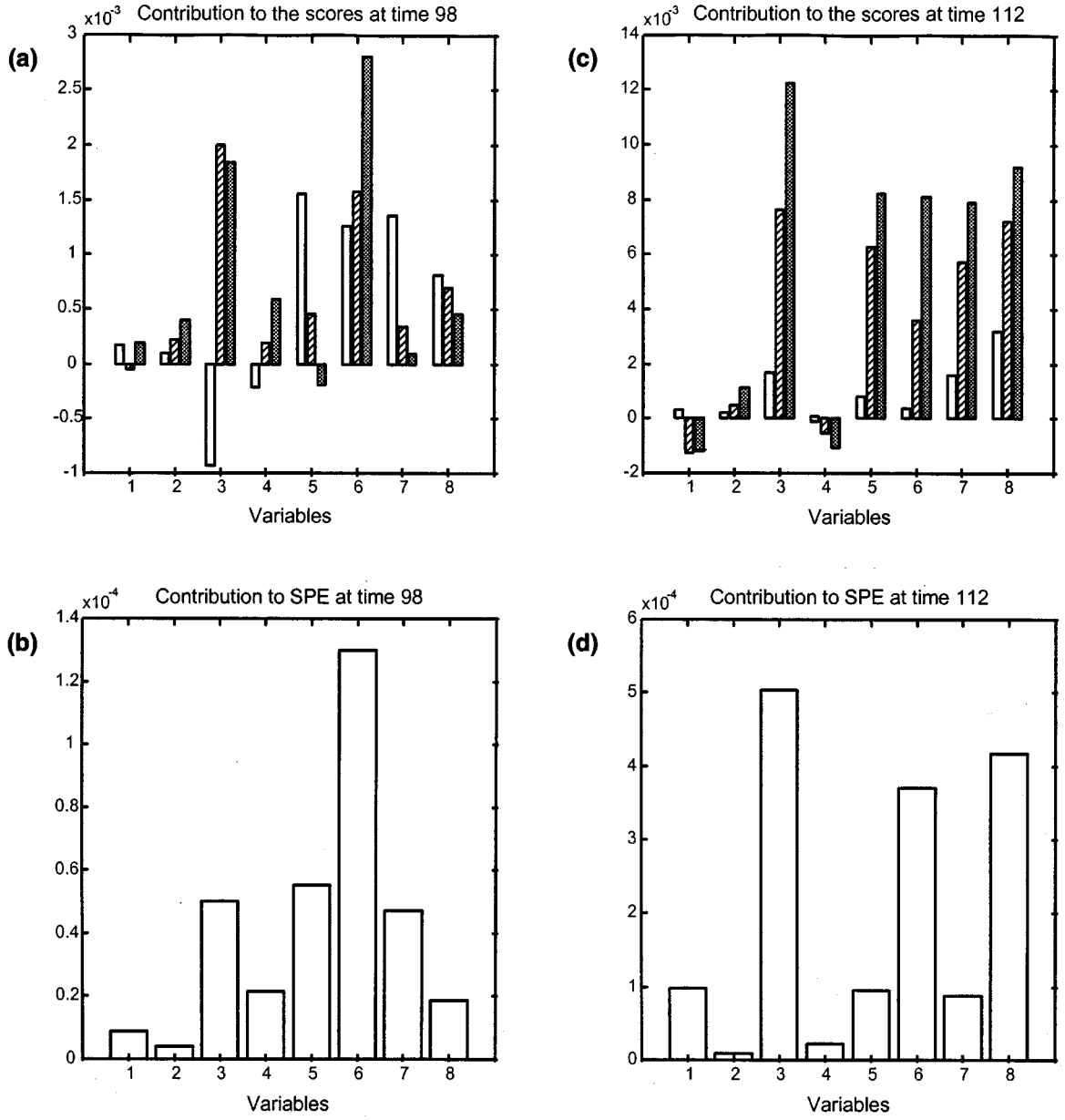

Figure 13. Contribution plots to scores and to the SPE for batch 47 at time intervals 98 and 112.

White, striped, and criss-crossed bars in (a) and (c), represent first, second, and third component, respectively.

stage 3, well before other batches in the group. Although no explanation for this different profile is available, the batch met product quality specifications, and it was also included in the training data set.

By looking at the contribution plots at time interval 98 (Figures 13a and 13b), further information can be obtained from the model. Variables 3 (vapor temperature) and 6 (heat source jacket vent temperature) contributed greatly to the scores (Figure 13a). As was stated earlier, the heating medium was turned off in stage four. The heat transfer to or from the reactor until the end of the batch was then facilitated by the heating medium vapors in the external jacket and internal coils in the autoclave. It thus seems that heat-transfer in the autoclave was not behaving properly, because the temperature in the external jacket was decreasing abnormally. This new source of variation was also detected in the contribution plot to the SPE, where variable 6 had the largest value. Contribution plots at time 112 (Figure 13c and 13d) are also presented to see how the process evolves during the final steps of the process. Variables 3, 5, 6, 7, and 8 show the highest contribution to the scores at that point (Figure 13c). Except vapor temperature inside the autoclave, the rest of these variables start decreasing from stage 4 until the process is finished. Further investigation of the profiles of variables 5, 6, 7, and 8 for this running batch shows that the final values of the temperatures and pressure at the end of the batch

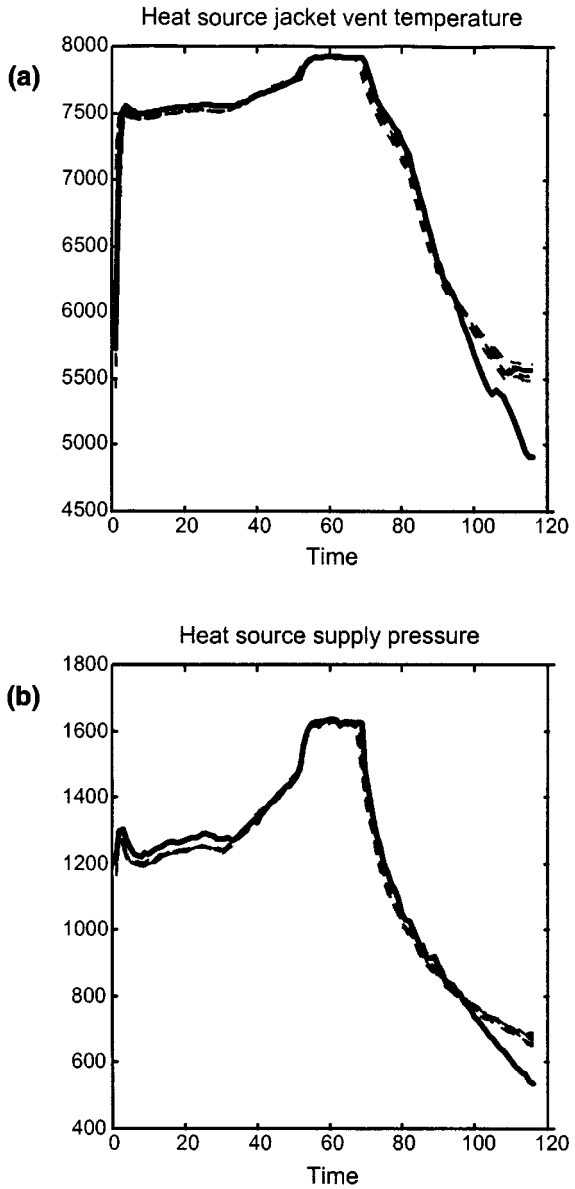

Figure 14. Time trends for variables 6 and 8 for some reference batches (dashed lines) and for the batch under study (solid line).

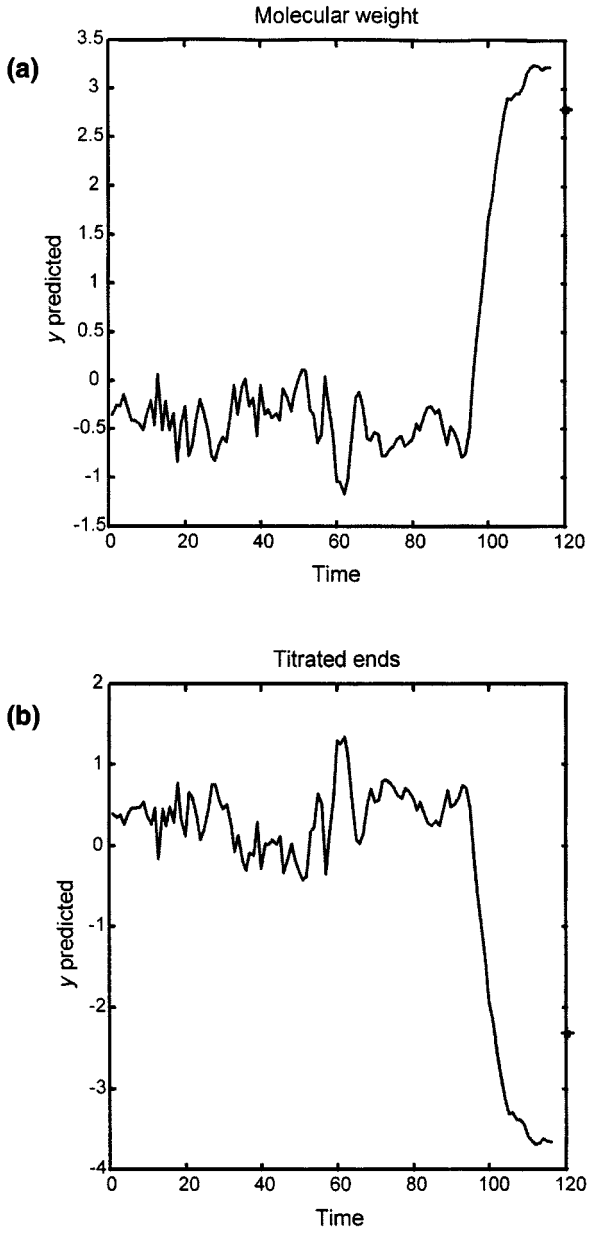

Figure 15. On-line predictions for the two quality variables.

Asterisk marks on the right side correspond to the actual final product quality.

were significantly lower than for the reference batches (see Figure 14). Also, variable 3 was found to show an abnormal profile in the last time intervals, which is reflected in the contribution plot to SPE (Figure 13d).

Figure 15 shows the *y*-predicted monitoring charts for the two quality variables. The predictive ability of the model was low to moderate, yet the evolution of the predicted *y*-values can be very helpful in detecting abnormalities in the process. This is well illustrated in Figures 15a and 15b, where the charts show a sudden change in the predicted properties at time interval 98, just as the MSPC charts detected. Thus, the *Y*-predicted monitoring charts can be a good complement for the traditional scores and SPE plots.

Conclusions

A new multiway covariates regression method for monitoring batch processes and predicting future quality variables has been presented. When a Tucker3 or a PARAFAC model was used to decompose the three-way $\underline{X}$ matrix, information on the variable and time domain was obtained in a clear way through different loading matrices. The process can be monitored by using common SPC charts, which are built from a reference set of past good batches.

Further diagnosis of a new running batch can be achieved by means of contribution plots. Once a fault has been detected at a certain time interval by inspecting the SPC charts, the process can be tracked by plotting the contribution of the process variables to the scores and to the squared prediction error. In this way, the variables responsible for the observed deviation can be easily detected. However, a shortcoming of the MSPC charts is that an out-of-control situation is only detected when either the scores or the SPE values exceed the control limits. Since score and SPE values often follow a certain increasing or decreasing pattern before reaching the control limits, it would be interesting to develop criteria for recognizing these special patterns, as is the case in univariate Shewhart charts. In this way, abnormal variations or faults could be detected earlier.

Predictions of the quality variables can also be monitored on-line. However, predicted values for an error batch cannot be completely trusted. The need for different models, depending on the final purpose (diagnosis, prediction), has also been discussed. However, there are still some unresolved problems. For example, information in most batch processes is also available for some initial variables, which can be arranged in a matrix *Z*. Future work should focus on developing methodologies for modeling the relation between the *X*, *Y*, and *Z* matrices. There is also a need for building reliable confidence intervals for the *Y*-predicted values in the multiway covariates regression models.

Acknowledgments

P. Nomikos (Dupont Canada Inc.), J. F. MacGregor (McMaster University, Canada) and Kenneth Dahl (Dupont, USA) are gratefully acknowledged for providing the batch polymerization data. We also thank Ad Louwerse for his constructive comments.

Notation

x = vector
Y = batch quality data
$\underline{X}$ = three-way array (batch process data)
$\underline{H}$ = core array in Tucker3 model
$m = 1, \ldots, M$ = index for quality variables

Literature Cited

de Jong, S., and H. A. L. Kiers, "Principal Covariates Regression. Part I. Theory," *Chemometrics Intelligent Lab. Syst*., **14**, 155 (1992).

Geladi, P., "Analysis of Multi-Way (Multi-Mode) Data," *Chemometrics Intelligent Lab. Syst*., **7**, 11 (1989).

Hahn, G. J., and W. Q. Mecker, *Statistical Interval. A Guide for Practitioners*, Wiley, New York (1991).

Harshman, R. A., "Foundations of the PARAFAC Procedure: Models and Conditions for an Exploratory Multimodal Factor Analysis," *UCLA Working Papers in Phonetics*, **16**, 1 (1970).

Kosanovich, K. A., K. S. Dahl, and M. J. Piovoso, "Improved Process Understanding Using Multiway Principal Component Analysis," *Ind. Eng. Chem. Res*., **35**, 138 (1996).

Kourti, T., P. Nomikos, and J. F. MacGregor, "Analysis, Monitoring and Fault Diagnosis of Batch Processes using Multiblock and Multiway PLS," *J. Proc. Control*, **5**, 277 (1995).

Kourti, T., and J. F. MacGregor, "Multivariate SPC Methods for Process and Product Monitoring," *J. Qual. Technology*, **28**, 409 (1996).

Kroonenberg, P., *Three-Mode Principal Component Analysis: Theory and Applications*, DSWO Press, Leiden (1983).

Miller, P., R. E. Swanson, and C. E. Heckler, "Contribution Plots: A Missing Link in Multivariate Quality Control," *Appl. Math. Computer Sci*., in press (1999).

Nomikos, P., and J. F. MacGregor, "Monitoring of Batch Processes Using Multi-Way Principal Component Analysis," *AIChE J*., **40**, 1361 (1994).

Nomikos, P., and J. F. MacGregor, "Multivariate SPC Charts for Monitoring Batch Processes," *Technometrics*, **37**, 41 (1995a).

Nomikos, P., and J. F. MacGregor, "Multi-Way Partial Least-Squares in Monitoring Batch Processes," *Chemometrics Intelligent Lab. Syst*., **30**, 97 (1995b).

Smilde, A. K., "Three-Way Analyses. Problems and Prospects," *Chemometrics Intelligent Lab. Syst*., **15**, 143 (1992).

Smilde, A. K., "Comments on Multilinear PLS," *J. Chemometrics*, **11**, 367 (1997).

Smilde, A. K., and H. A. L. Kiers, "Multiway Covariates Regression Models," *J. Chemometrics*, **13**, 19 (1999).

Tucker, L., "Implications of Factor Analysis of Three-Way Matrices for Measurement of Change," *Problems of Measuring Change*, C. Harris, ed., Univ. of Wisconsin Press, Madison, p. 122 (1963).

Manuscript received Feb. 9, 1998, and revision received Mar. 13, 1999.