BUTTERWORTH HEINEMANN

J. Proc. Cont. Vol. 5, No. 4, pp. 277–284, 1995

0959-1524/95 $10.00 + 0.00

0959–1524(95)00019–4

Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS

Theodora Kourti*, Paul Nomikos and John F. MacGregor

Chemical Engineering Department, McMaster University, Hamilton, Ontario, L8S 4L7, Canada

Multivariate statistical procedures for the analysis and monitoring of batch processes have recently been proposed. These methods are based on multiway principal component analysis (PCA) and partial least squares (PLS), and the only information needed to exploit them is a historical database of past batches. In this paper, these procedures are extended to allow one to use not only the measured trajectory data on all the process variables and information on measured final quality variables but also information on initial conditions for the batch such as raw material properties, initial ingredient charges and discrete operating conditions. Multiblock multiway projection methods are used to extract the information in the batch set-up data and in the multivariate trajectory data, by projecting them onto low dimensional spaces defined by the latent variables or principal components. This leads to simple monitoring charts, consistent with the philosophy of SPC, which are capable of tracking the progress of new batch runs and detecting the occurrence of observable upsets. Powerful procedures for diagnosing assignable causes for the occurrence of a fault by interrogating the underlying latent variable model for the contributions of the variables to the observed deviation are also presented. The approach is illustrated with databases from two industrial batch polymerization processes.

Keywords: batch processes; statistical process control; principal component analysis

Recent trends in most industrialized countries have been towards the manufacture of higher value added specialty chemicals, which are produced mainly in batch reactors. Examples include specialty polymers, pharmaceuticals and biochemicals. There are also many other batch type operations, such as crystallization and injection moulding, which are very important to the chemical and manufacturing industries. Monitoring these batch processes is very important to ensure their safe operation and to ensure that they produce consistent high quality products. Currently some of the difficulties limiting our ability to provide adequate monitoring include: the lack of on-line sensors for measuring product quality variables, the finite duration of batch processes, the presence of significant nonlinearities, the absence of steady state operation, and the difficulties in developing accurate mechanistic models that characterize all the chemistry, mixing and heat transfer phenomena occurring in these processes. Most of the existing industrial approaches for achieving consistent and reproducible results from batch processes are based on the precise sequencing and automation of all the stages in the batch operation. Monitoring is usually confined to checking that these sequences are followed and that certain reactor variables, such as temperatures and reactant feed-rates, are following acceptable trajectories. In some cases, on-line energy balances are used to keep track of the instantaneous reaction rate, and the conversion or the residual reactant concentrations in the reactor.

Recent research approaches to monitoring batch processes have focused on the use of either fundamental mathematical models (based on state estimation methods[1]), or detailed knowledge based models (using expert systems or artificial intelligence methods to process the data[2]). These methods are reviewed by Nomikos and MacGregor[3] and contrasted with multiway principal component analysis (MPCA). Rather than utilizing detailed engineering knowledge about the process, as in model-based and knowledge-based approaches, MPCA utilizes only the information contained in the historical database of past batches. Such information is readily available for any computer-monitored industrial batch process. Although theoretical models of batch processes and on-line sensors for the quality properties are not usually available, nearly every batch process does have available frequent observations on many easily measured process variables, such as temperatures, pressures, flowrates and agitator power. Measurements on up to 30 or more variables may be available every few seconds throughout the entire history of a batch.

*Author to whom correspondence should be addressed. E-mail: kourtit@mcmail.cis.mcmaster.ca

Nomikos and MacGregor[3,4] suggested approaches to utilize the historical data on the measured process variable trajectories, and extended these approaches to include the final product quality measurements at the end of each batch[5]. These approaches are further extended here to include measured feed-stock properties and other variable initial conditions. Powerful procedures for diagnosing assignable causes for the occurrence of a fault by interrogating the underlying latent variable model for the contributions of the variables to the observed deviation are also presented. Multivariate methods, such as multiway principal components analysis (MPCA), multiway partial least squares (MPLS) and multiblock multiway PLS, are used to extract the information from the trajectories of all the measured batch process variables, quality variables and any other feed-stock properties, and to project it onto a low dimensional space defined by the latent vectors or principal components. The data reduction is tremendous, since all the information from a database of batches is captured in a few vectors and matrices which define the reduced space. A post analysis of past batches enables one to classify similar and dissimilar batches by examining the clustering of their projections onto this hyperplane. New batches can be monitored in real-time, using a sound statistical framework, by tracking their progress in this reduced space. These approaches and monitoring schemes are illustrated with data from two industrial batch polymerization processes.

Nature of batch data

To understand the nature of the data available in a batch monitoring problem, consider a typical batch run in which $j=1,2,...J$ variables (such as temperature and pressure in the reactor, agitator power etc.) are measured at $k=1,2,...K$ time intervals throughout the batch. Similar data will exist on a number of similar batch runs $i=1,2,...I$. All these data can be summarized in the $\mathbf{X}$ (IxJxK) array illustrated in *Figure 1*. The different batch runs have been arranged across the vertical axis, the measurement variables along the horizontal axis and their time evolution occupies the third dimension. Each horizontal slice through this array is a (JxK) matrix containing the trajectories of all the variables

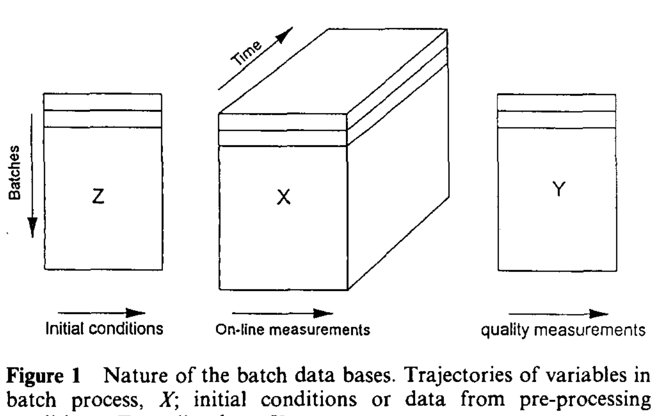

Figure 1 Nature of the batch data bases. Trajectories of variables in batch process, *X*; initial conditions or data from pre-processing conditions, *Z*; quality data, *Y*.

from a single batch. Each of its vertical slices is an (IxJ) matrix representing the values of all the variables for all the batches at a common time interval (k).

Also available may be data describing the final product quality (such as particle size distribution, molecular weight, relative viscosity, etc.); these measurements are taken at the end of each batch, for a few variables $l=1,2,...L$. These are summarized in the (IxL) matrix $\mathbf{Y}$.

Furthermore, for each batch, information usually exists on feed-stock properties, preprocessing and other conditions; for example measured raw material properties and compositions, charges of each ingredient, hold times in charge tanks, and discrete operating conditions such as the operator shift on which batch is produced, raw material suppliers, etc. This information can be summarized in an (IxM) matrix $\mathbf{Z}$.

Projection methods. Multiblock and multiway PCA and PLS

The basic concepts and algorithms of principal components analysis (PCA) and partial least squares (PLS)[6-8] and their use in multivariate monitoring of process operating performance[9-12] have been extensively presented in the literature. A recent review on the use of these methods for statistical process control can be found in MacGregor and Kourti[13]. The projection methods are compared to traditional SPC approaches in Kourti and MacGregor[14]. These procedures are based on projecting the information contained in high dimension data spaces onto low dimension spaces, defined by a small number of latent variables, the scores (t_1, t_2,...t_A). These new latent variables summarize all the important information contained in the original data sets.

In the PCA approach, the original data set ($\mathbf{W}$) of J variables and I observations is projected to an orthogonal structure, which, in general, is of a lower dimensionality[15-17]

$$\mathbf{W} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a^{\mathrm{T}} + \mathbf{E} \tag{1}$$

The location of this A-dimensional space with respect to the original coordinates is given by the loadings ($\mathbf{p}_a$). The location of the projection of an observation onto the A-dimension space is given by the scores ($\mathbf{t}_a$). The squared perpendicular distance of an observation from the projection space, called squared prediction error (SPE), gives a measure of how close the observation is to this A-dimensional space. $\mathbf{E}$ is a residual matrix. Ideally the dimension A is chosen such that there is no significant process information left in $\mathbf{E}$; rather $\mathbf{E}$ should represent random error. SPE for observation i, is

$$SPE = \sum_{j=1}^{J} e_{ji}.$$

Several methods have been suggested for choosing the number of components[17] with cross validation being perhaps the most reliable[18].

Monitoring continuous processes utilizing the PCA approach involves monitoring both the scores and the SPE[10]; here **W** contains observations of the process variables. As in the case of the traditional SPC approaches, a normal operating region is defined from past historical data (both for the SPE and the scores) and each new observation is compared to this normal region.

In the PLS approach, both the process (**W**) and quality (**Y**) variables are used. PLS simultaneously reduces the dimensions of **W** and **Y** spaces, to find latent vectors for **W** and **Y** which have greatest covariance. Utilizing PLS, an empirical model is developed to relate process (**W**) and product (**Y**) variables under normal operating conditions. Then, by monitoring the process variables only, and projecting them to the reduced dimensional space defined by the PLS latent vectors (t_1, t_2, ...t_A), we monitor the variation in the process variables that are more influential on the product quality variables[10].

Multiblock PLS is a technique in which the process variables **W** can be divided into subsets (blocks) of variables ($\mathbf{W}_1$, $\mathbf{W}_2$, ...$\mathbf{W}_n$). These subsets are then related simultaneously to **Y**. The **W** blocks may have equal or different weights. Multiblock PLS is used to either break a process into smaller groups and facilitate the diagnostic procedure or to cluster together variables of equal importance or with similar characteristics and weigh them together[12].

Extension to matrices with higher dimensions

When a three or higher dimensional matrix describes the data to be analysed (as matrix **X** in *Figure 1*), 'multiway' projection methods may be utilized. The relation of PCA and PLS to multiway PCA and PLS (MPCA, MPLS) has been described in a series of articles[19-21]. Nomikos and MacGregor[3,4] applied MPCA to analyse batch data and detailed the procedures for monitoring batch processes. MPCA is equivalent to performing PCA on a large two-dimensional matrix formed by unfolding the three-way array **X** (shown in *Figure 1*) in one of three possible ways. Nomikos and MacGregor[3,4] transformed the three-dimensional array **X** to a two-dimensional array by unfolding **X** in such a way as to put each of its vertical slices ($I \times J$) side by side to the right, starting with the one corresponding to the first time interval. The resulting two-dimensional matrix has size ($I \times JK$). This unfolding allows for analysing the variability among the batches in **X** by summarizing the information in the data with respect both to variables and their time variation. With this particular unfolding, by subtracting the mean of each column prior to performing the MPCA, it is the variation about the mean trajectories of all the variables that is being analysed. Since most of the nonlinearities are removed by removing the average trajectory from each variable, we have found that nonlinear PLS methods offer no improvement in most of the cases we have investigated where the method is used for process monitoring.

In this approach, MPCA classifies batches as good or bad based on their similarity to a group of previous batches that produced an acceptable product. Information from quality measurements is not utilized directly. The approach is based on the basic concepts of statistical process control (SPC), whereby the future behaviour of a process is monitored by comparing it against that observed in the past when the process was performing well, that is in a state of statistical control. Control limits for the monitoring charts are derived from the statistical properties of this historical reference distribution of past 'good' batches. Therefore, the approach relies upon the idea that future 'good' batches should have similar behaviour to past ones.

MPLS may be used to utilize information from the product quality[4]. Once the **X** matrix has been unfolded into a two-dimensional matrix, PLS can be performed between **Y** and this new matrix to relate the quality characteristics to the process conditions. By utilizing the quality measurements the batches may be classified in a way that they are more predictive of **Y** – in this case variables that exhibit high variability but do not affect the quality of the product are weighted less heavily.

When extra information relevant to the batch process is available (in the form of matrix **Z** in *Figure 1*) this information may also be utilized, by performing multiway multiblock PLS. Matrix **Z** and the unfolded **X** matrix may be treated as two blocks, scaled and weighted appropriately. Multiblock PLS can then be applied in the way described in MacGregor *et al.*[12]. The weight of each block may be chosen based on experience with the process. When no prior information exists each block may be given equal weight. It is suggested that several models be derived where the weight ratio of the two blocks changes; as a rule of thumb the best model is the one in which the cross-validated percentage explained in Y is equal to or higher than that explained by a PLS model using each block separately.

Analysis of batch operating records

The use of multiblock multiway PLS to analyse historical data bases of batch processes is illustrated here on data from 61 batches from an industrial polymerization process. For each batch, the trajectories of 10 variables measured during the run at 250 time intervals were provided, together with measurements of 4 quality variables obtained at the end of the run. Measurements of 14 variables that described raw material charges and average conditions in a process that preceded the polymerization were also available. Batches 56, 59, 60 and 61 were characterized by the company as bad, based on the values of the measured quality variables; batch 21 had one quality measurement missing.

Multiblock multiway PLS was performed on all the 61 batches to analyse this data base. The variables that corresponded to preprocessing were grouped as one block $\mathbf{Z}(61 \times 14)$; the second block was assigned to the unfolded three – way array $\mathbf{X}$ ($61 \times 10 \times 250$) that described the polymerization conditions; the quality variables were all assigned in a (61×4) matrix, **Y**.

Figures 2a, 2b and 2c show projections of these 61 batches on the latent variable planes for the pre-processing, polymerization and quality variables, respectively. *Figure 2a* shows the projection of the pre-processing on the first two components (t_1 *vs* t_2). Batches 56, 59, and 61 (characterized by the company as 'bad') are out of the main cluster formed by the rest of the batches for the pre-processing stage. It should be noted, however, that a few more batches appeared out of the main cluster in the plane of the first two components (batches 44, 45, 55) and some in the other components (batch 36, 40). *Figure 2b* shows the projection of the polymerization process data (X) on the plane defined by the first and third latent variables. Batches 56, 59, 60 appear out of the main cluster. Projections on planes defined by the rest of the components showed that batches 36, 40, 44, 50, 56, 59 and 60 are different from the rest of the batches. Although based on the four quality properties only batches 56, 59, 60 and 61 were characterized as bad by the company; the fact that batches 36, 44 and 45 appear in both the preprocessing and the polymerization as outliers shows that these are not random outliers but there is some consistent discrepancy. Projection on the latent variables of the quality space Y (u_1 *vs* u_2) also indicates that batches 56, 59 and 60 are outside of the main cluster of normal batches.

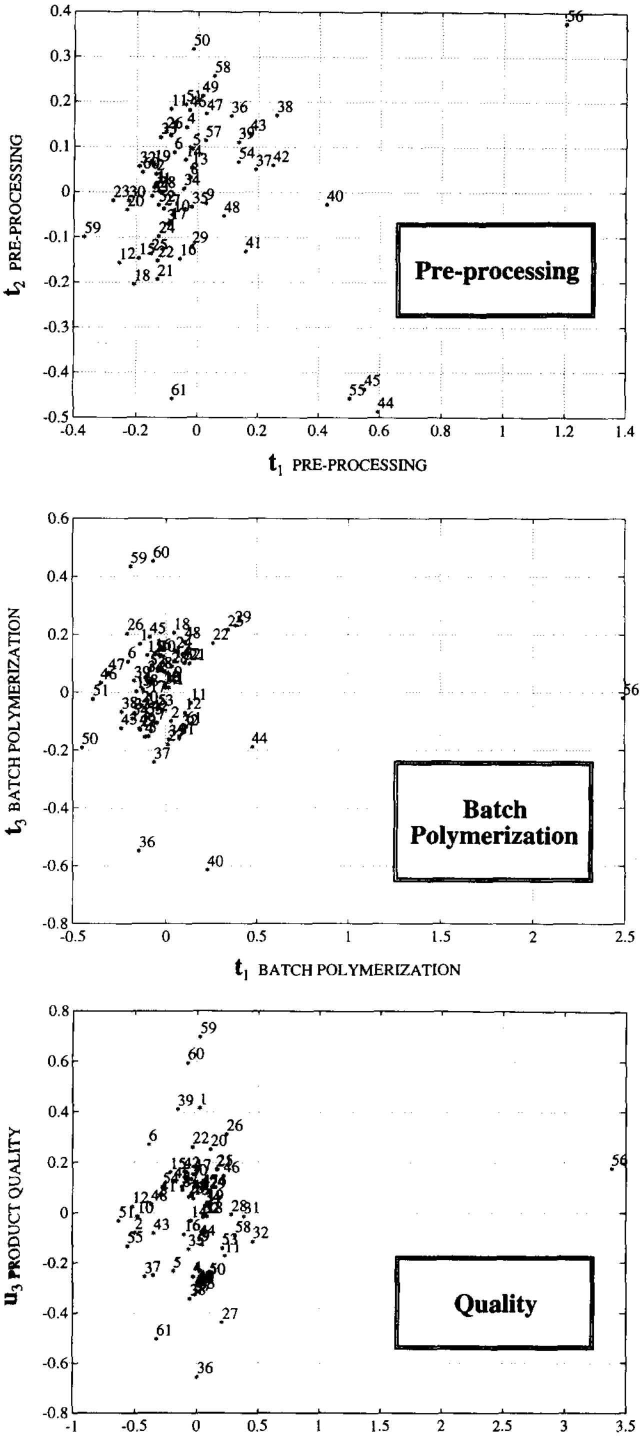

Figure 2 Projection on latent variables determiend by multiblock multiway PLS: **a**, preprocessing; **b**, polymerization; **c**, quality data

At this stage it is established that the method is capable of discriminating between 'good' and 'bad' batches with the available process data; in other words the system is *observable*.

Modelling, post analysis, fault diagnosis

The set of 49 batches (after batches 21, 36, 40, 44, 45, 50, 55, 56, 59, 60 and 61 were excluded) was used to develop a model to relate the preprocessing and polymerization conditions to the product quality properties, using multiblock multiway PLS. The blocks were scaled such that the variance of the second block (X) was four times the variance of the first (Z). The model results are summarized in *Table 1*. Four latent variables explain 56% of the variability in the product quality. The cumulative percentage variation explained per block as well as the weight of each block in explaining Y, is included in the table. Note that in all the components the polymerization conditions (block **X**) have a much higher weight than the pre-processing conditions.

The model built from the 49 batches summarizes information on the normal operating conditions for the batch polymerization. Batches that were detected as 'unusual' (i.e. out of main clusters) in the preliminary analysis can be checked against this model to determine the reason for their differences. *Figures 3a* and *3b* show the SPE per batch, in the preprocessing and polymerization respectively, for all the 61 batches, when the model for the good batches is used; the dashed line in the figures gives the 99% limit of the SPE for the good batches. From these figures it can be seen that batches

Table 1 Multiblock – multiway PLS results, batch process 1

Component	Cum % explain Z	Cum % explain X	Cum % explain Y	% Weight Z	% Weight X
1	29	8	21	31	69
2	40	21	31	9	91
3	54	27	42	17	83
4	62	31	56	24	76

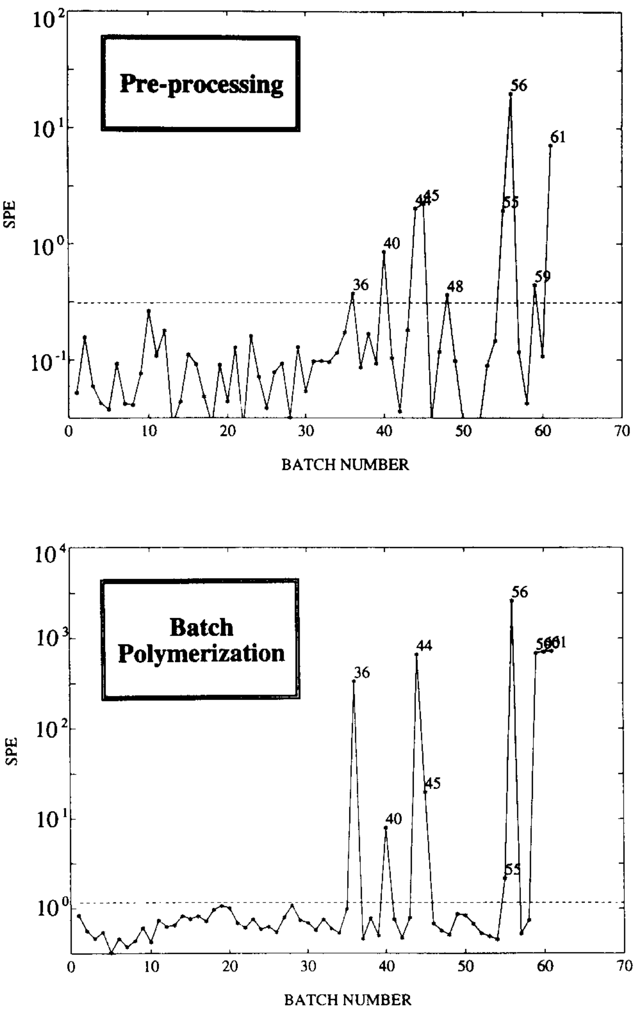

Figure 3 Squared prediction error for 61 batches: **a**, preprocessing; **b**, polymerization

36, 40, 44, 45, 55, 56, 59 and 61 do not follow the pattern of the rest of the batches in both the preprocessing and polymerization. Contribution plots[12,22] can be used to interrogate the model and determine the reason for the difference for each of the 'unusual' batches.

Figure 4a shows the contribution of each variable in the preprocessing (Z) to the SPE for batch 56; variable 9 appears to contribute significantly to this error. Indeed by examining the data base it was revealed that the value of this variable (amount of an ingredient changed) was much lower than usual. *Figure 4b* shows the contribution of each polymerization variable (X) to the SPE of batch 56. Variable 4 has the highest contribution. *Figure 4c* shows the contribution of the polymerization variables at each time interval to the SPE; it is clear that events that took place at time (82–95) and later at times, (135–155), (167–173) and (201–207) have the highest contribution to this error. Looking back at the data base, the trajectory of variable 4 for this batch showed a different trend from the other batches at these time intervals.

On-line monitoring and fault diagnosis

The model built to summarize the information contained in the 49 good batches and the statistical distri-

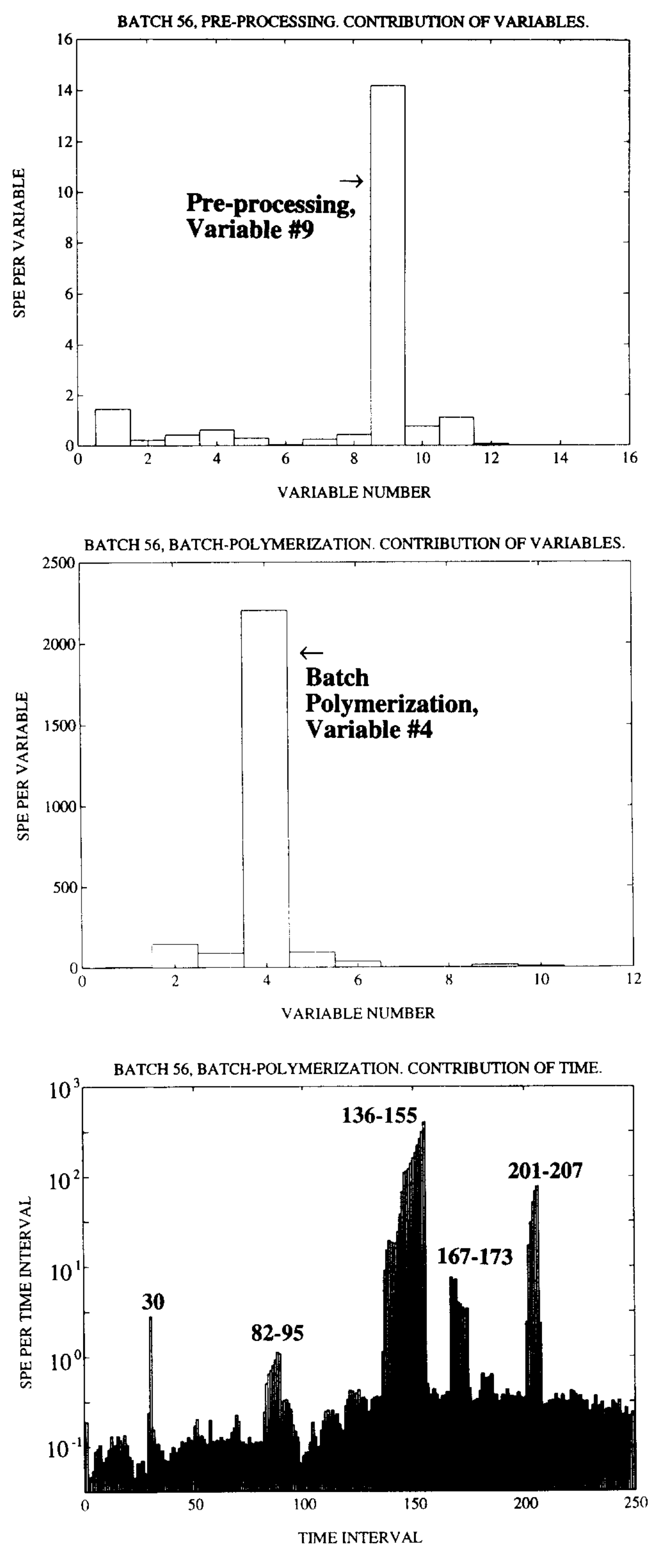

Figure 4 Contribution plots for batch 56: **a**, contribution of preprocessing variables; **b**, contribution of polymerization variables; **c**, contribution per time interval during batch run

bution of the latent variables and SPE can be used as a statistical reference to classify new batches as normal ('good') or abnormal ('bad'), on-line, as the data are collected and the polymerization evolves. As with the traditional SPC approach, this classification is based on the similarity and statistical consistency of the trajectory measurements of a new batch with the historical reference distribution of trajectories from normal operation as summarized by the model.

The procedure is described in Nomikos and MacGregor[3-5]. Limits are calculated for the projection space and the residuals, at each time interval, as the batch evolves. New batches are classified by monitoring the latent variables (scores) and the instantaneous SPE_k, at each time interval k. Both of these quantities are necessary for monitoring: the scores check if each new observation from the new batch remains within the normal operating region in the projection space, and SPE_k at time k checks if the distance of the new observation from the projection space is within normal limits for that time interval.

The problem with monitoring a new batch on-line is that the ($J \times K$) matrix of the new batch is not complete; at each time interval during the batch operation the matrix has all measurements only up to this time interval. Approaches suggested for handling this situation, which consist of filling up the remainder of the matrix with estimates of the future observations or treating them as missing data, are discussed in Nomikos and MacGregor[4]. In this example any future value of each variable is assumed to have a deviation from the target value equal to the current one.

Figure 5a and *5b* show the results that one would have obtained for batch 56, had a multiway PLS model been in use on-line, when the data for this batch were becoming available. *Figure 5a* gives the SPE_k response as a function of time; note that the SPE_k shows excursions out of normal operation at time intervals 80–90 and then 135–155 and 201–207 (compare with results of post analysis for this batch, *Figure 4*). (The solid line corresponds to a 99% limit and the dashed line to a 95% limit). *Figure 5b* gives the projection of the batch evolution on the second and third latent variable space (t_2/t_3); each asterisk on the plot corresponds to an updated estimate of t_2 and t_3 as the batch evolves; numbered asterisks indicate corresponding time intervals. The ellipses give the 95% and 99% confidence intervals in which the value of t_2/t_3 should fall at the end of the reaction. Note that the estimates of t_2/t_3 leave the ellipse at time interval k=82 and stay out of limits until the end of the reaction (k=249). Contribution plots on the SPE_k and the t scores can be constructed to interrogate the model and diagnose an assignable cause for the problem[23].

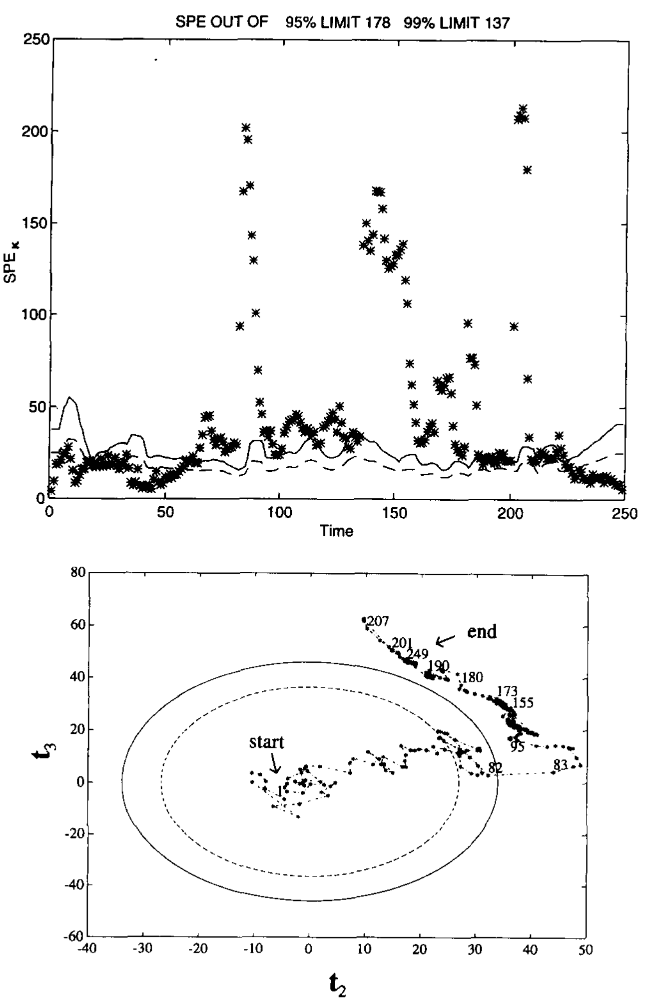

Figure 5 On-line monitoring of batch 56, using multiway PLS: **a**, squared prediction error; **b**, progression of batch as projected on second and third latent variables

Another batch polymerization process

The second example is from another batch polymerization process which consists of two batch stages and for which information on initial feed quality, hold up times and the shift corresponding to a batch was available. Data for 92 batches were provided; no batches were characterized by the company as bad. In this case the data were arranged for the multiblock multiway PLS as follows. Two blocks were assigned to the two batch stages: $\mathbf{X}_1$ for the unfolded threeway array ($92 \times 19 \times 86$) with trajectories of 10 variables for 86 time intervals for the first stage; $\mathbf{X}_2$ for the unfolded three-way array ($92 \times 21 \times 231$) with trajectories of 21 variables for 231 time intervals for the second stage. Measurements of nine variables that described raw material qualities were assigned to block $\mathbf{Z}_1$ (92×9). Data (eight variables) giving the shifts on which each batch was produced and hold up times were assigned to block $\mathbf{Z}_2$ (92×8). Two quality variables were available for each batch and arranged in $\mathbf{Y}$ (92×2).

Multiblock multiway PLS was performed on all the 92 batches. *Figures 6a* to *6d* show the projections of each one of the blocks on the first two latent variable planes. In the projections of block Z_1 (*Figure 6a*) the points fall in groups of three or four, because the same lot of raw material is used for three or four batches. Batches 54–57 had an unusual value for a raw material property that does not seem to have had a significant effect on the product. From the analysis of Block Z_2 it was established that all the operator shifts produced same quality and followed similar operating procedures. Batches 1 and 30 had unusual hold up times. In Block X_1, batch 58 appeared as an outlier; this batch had unusual trajectories but it does not seem to have had an effect on the two product qualities measured. Finally in Block X_2, batches 1 and 30 again appeared as outliers; upon examination of the data base it was discovered

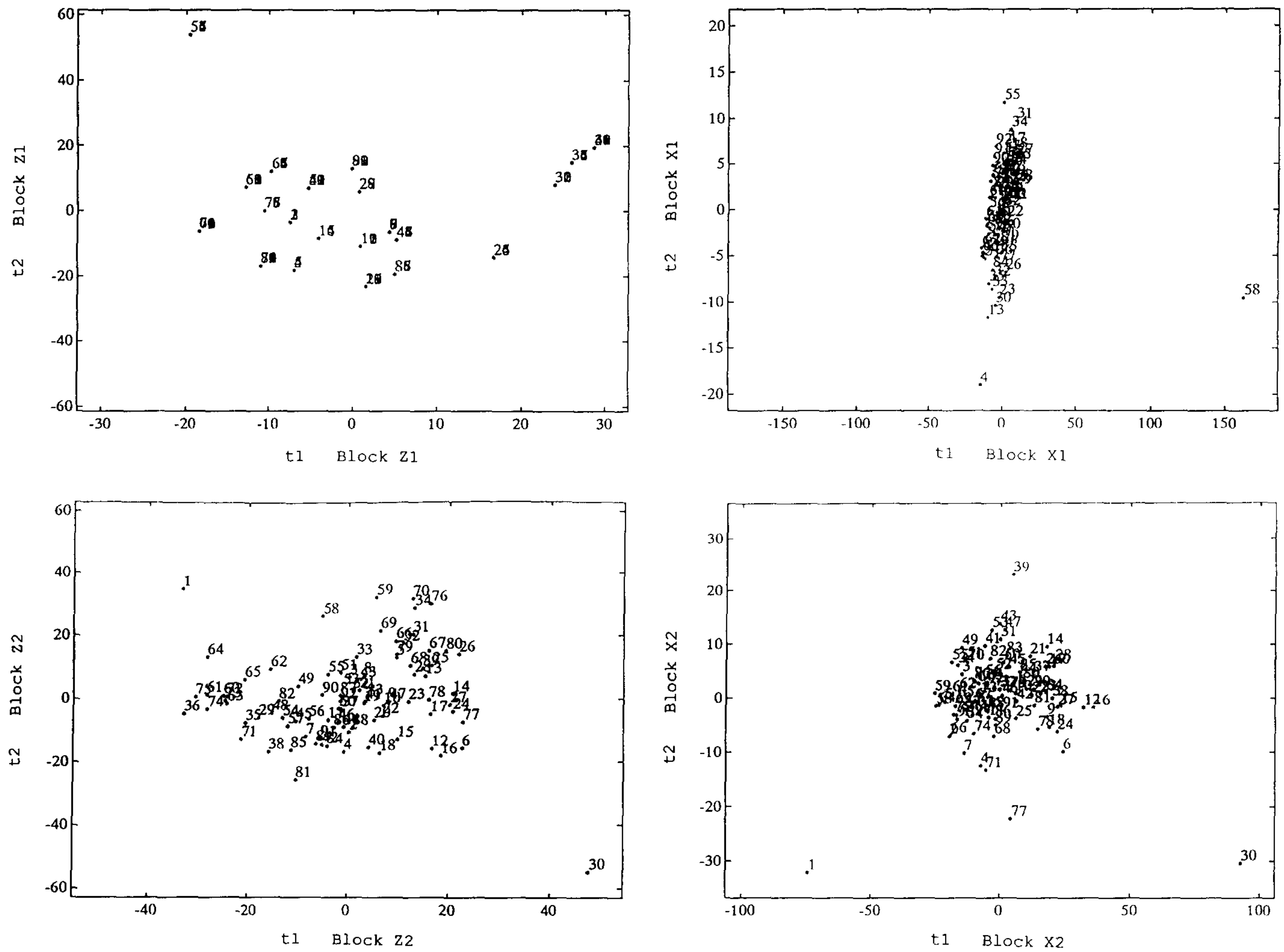

Figure 6 Projection on latent variables determined by mulitway-multiblock PLS: **a**, block Z_1; **b**, block Z_2; **c**, block X_1; **d**, block X_2

that these two batches had different polymerization trajectories from the others. From the quality data it was found that the product of batch 30 had border-line qualities. The abnormal behaviour of batches 1 and 30 in the second batch stage could be related to events described in block Z_2.

Batches 54–57, 58, 1 and 30 were excluded from the data base and a multiblock multiway PLS model was developed. Each block was given equal variance. The model is summarized in *Table 2*. The highest percentage in quality variability is explained by blocks Z_2 and X_2 which were initially judged by the company as the most important ones.

Concluding remarks

Batch analysis and monitoring methods based on MPCA and MPLS have been extended by using multi-block-multiway PLS. This extension allows one to utilize the historical data on the measured process variable trajectories, the measured feed-stock properties and other variable initial conditions and the final product quality measurements at the end of each batch.

Table 2 Multiblock-multiway PLS results, batch process 2

Component	% wieght Z_1	% weight Z_2	%weight X_1	%weight X_2	Cum% Y
1	9.4	45.5	16.5	28.6	13.5
2	28	42.8	18.2	11.0	31.75
3	32	6	32	30	47.68

The proposed monitoring charts are in accordance with the SPC requirements in that they can be easily displayed and interpreted, and they can quickly detect a fault. Furthermore, it is also possible to provide the operators with diagnostic information by interrogating the underlying MPCA, MPLS or multiway-multiblock PLS model.

As in all inferential approaches, the fundamental assumptions of *comparable* runs and *observable* events of interest must hold for the method to work. The first assumption states that the method is valid as long as the reference database is representative of the process operation. If major modifications are made to the process, then one has to build a new database which embodies the changes and re-apply the method. The second assumption expresses the requirement that the events which one wishes to detect must be *observable* from the measurements that are being collected. No monitoring procedure can detect events that do not affect the measurements.

The power of the statistical approach presented here lies in the fact that it utilizes the unsteady state trajectory data on all measured variables in a truly multivariate manner, so as to account not only for the magnitude and trend of the deviations in each measured variable from its average trajectory, but also for the high degree of correlation in both time and among the deviations in all the variables. The objective of the monitoring procedure is to detect and eliminate faults from future appearance, and thereby shrink the control limits and work towards a more consistent production of quality product.

References

1 Frank, P.M. *Automatica* 1990, **26**, 459
2 Konstantinov, K.B. and Yoshida. T. in 'On Line Fault Detection and Supervision in the Chemical Process Industries,' IFAC Symposium, Newark, DE, April 22–24, 1992
3 Nomikos, P. and MacGregor, J.F. *AIChE J.* 1994, **40**, 1361
4 Nomikos, P. and MacGregor, J.F. *Technometrics* 1995, **37** (1)
5 Nomikos, P. and MacGregor J.F. Presented at the First International Chemometrics Internet Conference, September 1994 (To appear in *Chemometri. Intell. Lab. Syst.* 1995
6 Wold, H. 'Research Papers in Statistics' (Ed. F. David) Wiley, New York, 1966, 411
7 Wold, S., Albano, C., Dunn, W.J., Esbensen, K., Hellberg, S., Johansson, E. and Sjöström, M. in 'Food Research and Data Analysis' (Eds H. Martens and Russwurm, H. Jr.) Applied Science, London, 1983
8 Geladi, P. and Kowalski, B. *Analy. Chim. Acta*, 1986, **185**, 1
9 Wise, B.M., Veltkamp, D.J., Davis, B., Ricker, N.L., and Kowalski, B.R. in 'Waste Management '88,' Tuscon, AZ, 1988
10 Kresta, J., MacGregor, J.F. and Marlin, T.E. *CJChE*, 1991, **69**, 35
11 Skagerberg, B., MacGregor, J.F. and Kiparissides, C. *Chemometr. Intell. Lab. Syst.* 1992, **14**, 341
12 MacGregor, J.F., Jaeckle, C. Kiparissides, C. and Koutoudi, M. *AIChE J.* 1994, **40**, 826
13 MacGregor, J.F. and Kourti, T. *Control Eng. Pract.* 1995, **3** (3) 403
14 Kourti, T. and MacGregor, J.F. in 'Fifth International Symposium on Process Systems Engineering (PSE '94),' Kyongju, Korea, 1994
15 Johnson, R.A. and Wichern, D.W. 'Applied Multivariate Statistical Analysis,' Prentice Hall, New Jersey, 1988
17 Jackson, J.E. 'A User's Guide to Principal Components,' John Wiley, New York, 1991
18 Wold, S. *Technometrics* 1978, **20**, 397
19 Lohmöller, J. and Wold, H. Presented at the European Meeting of Psycometrics Society, Gronigen, Netherlands, 1980
20 Wold, S., Geladi, P., Esbensen, K. and Ohman, J. *J. Chemometrics* 1987, **1**, 41
21 Smilde, A.K. and Doornbos, D.A. *J. Chemometrics* 1991, **5**, 345
22 Miller, P., Swanson, R.E. and Heckler, C.F. in '37th Annual Fall Conference, *ASQC*,' Rochester, NY, 1993
23 Nomikos, P. 'Mulitvariate Statistical Process Control of Batch Processes,' PhD Thesis, McMaster University, Canada, 1995