MULTIVARIATE FORECASTING OF BATCH EVOLUTION FOR MONITORING AND FAULT DETECTION

Salvador Garcia Muñoz, Theodora Kourti, John F. MacGregor

McMaster Advanced Control Consortium, McMaster University, Hamilton, ON, Canada.

Abstract: To monitor a batch process, which is dynamic in nature, it is necessary to consider the time varying relationship of its variables throughout the entire run. MPCA models built with batch wise unfolded data have been used extensively for batch process monitoring, these methods will not only consider the known samples to asses the ongoing batch run, but will also consider a dynamic forecast of the future unknown samples. Such forecast, implicit in the methodology, is uncovered and analyzed in this work; and proven to be a powerful feature of a batch-monitoring scheme built with MPCA and the batch-wise unfolded matrix of batch data.

Keywords: Batch, Monitoring, Multivariate Statistical Process Control, Forecasting,

1. INTRODUCTION

Multivariate methods based-on multi-way PCA and PLS (Wold et al, 1987) have proven their usefulness in batch process monitoring (Kourti et al, 1995; Nomikos, 1996; Nomikos and MacGregor, 1994; Nomikos and MacGregor, 1995a; Nomikos and MacGregor, 1995b). These methods have been widely studied and extensively applied in industry (Kosanovich et al, 1996; Nomikos, 1996; Ramaker et al, 2002; Schlags and Populle, 2001; Tates et al, 1999). They are essentially the only powerful approaches for monitoring batch processes when there is no deterministic model available. Other variation of the multivariate approaches (Boque and Smilde, 1999; Wold et al, 1998) are found in literature, and have already been compared (van Sprang et al, 2002; Westerhuis et al, 1999). This work focuses on the approach proposed by Nomikos and MacGregor (Kourti et al, 1995; Nomikos, 1996; Nomikos and MacGregor, 1994; Nomikos and MacGregor, 1995a; Nomikos and MacGregor, 1995b), and in particular in the estimation of the missing trajectory information needed in the approach when monitoring a new batch.

In the Nomikos-MacGregor method the 3 dimensional batch dataset consisting of J variables, I batches and K time periods is rearranged as shown in Fig. 1. A natural problem to overcome when monitoring a new batch at time=k, where k<K is the fact that there are K-k unknown samples (Fig. 2) and therefore, the need to estimate or "fill in" these unknown samples for the ongoing batch in order to calculate a score value (note that this problem does not exist when building the PCA model, when analyzing historical data, or when monitoring only at the end of each new batch). In their work, Nomikos and MacGregor (Nomikos and MacGregor, 1994; Nomikos and MacGregor, 1995b) acknowledge this issue and propose three different ways to handle this problem: to complete the observation vector with zero deviation from the nominal trajectories (zero option: Z); complete the observation assuming that the mean centered and scaled current deviation from the nominal trajectory will be maintained throughout the rest of the batch (i.e. current deviations option: CD), or estimate the "missing" values using the PCA model (missing data option: MD).

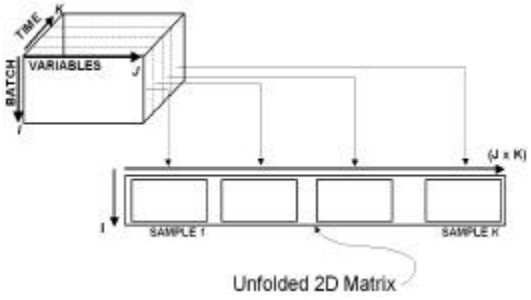

Fig. 1. Unfolding of the 3D batch data matrix.

The control limits of the resulting monitoring charts are dependent upon the used option (Nomikos and MacGregor, 1995b). This "filling in" need of the method has been criticized by several authors (Cho and Kim, 2003; Meng et al, 2003; Undey et al, 2003; van Sprang et al, 2002) without any convincing analysis to support their arguments.

In this work we demonstrate that monitoring a batch process with the Nomikos-MacGregor method, and using the MD approach to fill in the unknown samples when monitoring a new batch, is in fact a powerful model predictive monitoring scheme where the future samples of the trajectories are predicted with good accuracy even at the very beginning of the batch run.

The missing data in the future trajectory are predicted much in the same way as an adaptive multivariate time series forecast, but with optimal use of all the available measurements up to current time and detailed knowledge about the time varying correlation structure among the variables throughout all the rest of the batch.

1.1 Notation

In the following sections, a is used to represent the total number of components in a PCA model; $\mathbf{X}$ represents the original data set used to build the model; $\mathbf{T}$ represents the known scores from the original model; $\mathbf{P}$ represents the loadings of the PCA model, $\mathbf{x}$ is used to represent a vector of data from a new observation; $\hat{\mathbf{t}}$ is used to represent the estimate of the score vector corresponding to $\mathbf{x}$ in the presence of incomplete data and $\mathbf{t}$ is the value of the score vector with the complete data vector.

The dataset $\mathbf{X}$ contains I batches, each batch has J variables sampled K times throughout the batch run. $\mathbf{X} \in \Re^{I \times JK}$, $\mathbf{T} \in \Re^{I \times a}$, $\mathbf{P} \in \Re^{JK \times a}$, $\mathbf{x} \in \Re^{JK \times 1}$, $\boldsymbol{\tau} \in \Re^{a \times 1}$, $\hat{\mathbf{t}} \in \Re^{a \times 1}$.

Lower case bold letters ($\mathbf{x}$) denotes a vector; upper case bold letters ($\mathbf{X}$) denote a matrix. For a certain observation $\mathbf{x}$ with missing elements, we will group all the missing elements at the *end* of the observation, the known samples in the observation will be denoted by $\mathbf{x}^*$, and the missing samples will be denoted by $\mathbf{x}^\#$. So, for a certain observation $\mathbf{x}$ with missing elements all at the end of the vector, we denote as $\mathbf{x}^T = [\mathbf{x}^{*T}\ \mathbf{x}^{\#T}]$.

This notation is applied to other matrices or vectors as well, e.g. $\mathbf{P}^*$ refers to those rows of a loading matrix $\mathbf{P}$ that correspond to the known variables (elements) of $\boldsymbol{x}$ ($\mathbf{x}^*$); and $\mathbf{P}^\#$ refers to those rows of $\mathbf{P}$ corresponding to the missing elements of $\mathbf{x}$ ($\mathbf{x}^\#$). These definitions are illustrated in figure 2.

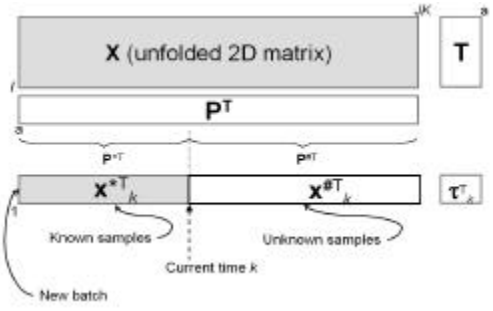

Fig 2. Notation and definitions

2. BATCH TRAJECTORY FORECASTING

2.1 Trajectory Forecasting using AR Models

To represent a batch trajectory with an AR multivariate time series we say that each batch is a collection of samples of the column vector $\boldsymbol{\chi}_j$ such that $\mathbf{x}^T = [\boldsymbol{\chi}_1^T, \boldsymbol{\chi}_2^T \ldots \boldsymbol{\chi}_K^T]$. At time k, the model will be able to make the forecast for the values of the samples at time $k+1$ using Eq. 1 where q is the order of the model.

$$\boldsymbol{\chi}_{k+1} = \mathbf{f}_1 \boldsymbol{\chi}_k + \mathbf{f}_2 \boldsymbol{\chi}_{k-1} + \mathbf{f}_3 \boldsymbol{\chi}_{k-2} + \cdots + \mathbf{f}_n \boldsymbol{\chi}_{k-q} \quad (1)$$

With this fixed model, it is possible to forecast the remainder of the batch at any time, and then compute its scores. If the batch run consists of several phases (or stages) then it is desirable to fit a different model per phase.

Trajectory forecast using the PCA model

In their earlier work, Nomikos and MacGregor (Nomikos and MacGregor, 1995b) clearly state that: when using the MD approach, the missing unknown values will be *replaced* by a prediction of these done with the PCA model. In the following paragraphs this replacement mechanism is uncovered since it not done explicitly but implicitly in the method, starting from the score estimation problem.

The simplest way to estimate the score vector when an observation contains missing data is to use the single component projection method (SCP) (Nelson et al, 1996). This method will calculate each of the scores independently and sequentially as $\hat{\mathbf{t}}_{ki} = \mathbf{z}^* \mathbf{p}_i^* / \mathbf{p}_i^{*T} \mathbf{p}_i^*$ where $\mathbf{z}^*$ is $\mathbf{x}^*$ at time k deflated by all the previous components. This method has proven to be the least performing of the missing data handling methods (Nelson et al, 1996), however, it has been reported to work in some cases (Lennox et al, 2000).

One natural improvement to the method is to use the forecast for the remainder of the trajectory, since it is possible to compute a prediction of the unknown

measurements $\hat{\mathbf{x}}_k^{\#}$ using their corresponding loadings and the score vector estimate (Eq.2);

$$\hat{\mathbf{x}}_k^{\#} = \mathbf{P}_k^{\#}\hat{\mathbf{t}}_k^T \qquad (2)$$

Once the prediction of the unknown measurements at time k is done, a new augmented vector can be created using the known measurements and the new predicted values (Eq.3)

$$\mathbf{x}_{aug}^T = \left[\mathbf{x}^{*T}\ \hat{\mathbf{x}}^{\#T}\right] \qquad (3)$$

This approach was taken by Meng *et al (Meng et al, 2003)*, and Lennox *et al (Lennox et al, 2000)* to forecast the unknown part of the trajectory in their respective studies. However, this is only the first step of a more complete solution which is the *iterative imputation* (II) method proposed by Arteaga and Ferrer (Arteaga and Ferrer, 2002), where the next step to take now is to re-estimate the score vector by multiplying the augmented vector of the known measurements and the predictions, by the loadings matrix (Eq. 4)

$$\hat{\mathbf{t}}_k^T = \mathbf{x}_{aug}^T\mathbf{P} \qquad (4)$$

With this new estimate of the score it is possible to iterate by re-calculating the augmented vector (Eq. 3) and the score estimate (Eq. 4) until convergence is reached in the score estimate. At convergence, the estimate of the score is computed with a full trajectory (that includes the known measurements and the forecast for the rest of the trajectory) and the whole loading matrix.

The *iterative-imputation* (Arteaga and Ferrer, 2002) algorithm is a very illustrative way of understanding how the missing data elements of the trajectories are being replaced by the PCA forecast. This method has been proven (Arteaga and Ferrer, 2002) to be equivalent to projection to the plane method, which was used by Nomikos and MacGregor (1994, 1995a, 1995b) in their early work.

1.4 The PCA forecast model vs. a time series forecast

Since it is clear that the *MD* option uses a forecast of the trajectory to "fill in" the unknown measurements, it is desirable to analyze the properties of the forecast model embedded in the PCA projection. To better understand the PCA forecast model, it can be expressed as a time series forecast model, in order to contrast it with the properties of a conventional multivariate time series models. In the following paragraphs, the embedded one-step a head prediction model is uncovered from the PCA prediction expression to illustrate the nature of the prediction model. However, equivalent expressions for any *f*-step ahead forecast could have been used to illustrate the same point.

Previously, the vector $\boldsymbol{\chi}_k$ was defined as a column vector that contains the J batch variables measured at time k, such that: $\mathbf{x}^{\mathrm{T}}=[\,\mathbf{?}_1^T, \mathbf{?}_2^T \ldots \mathbf{?}_K^T\,]$. Using the notation defined previously, we can express the known batch samples at time k in terms of $\boldsymbol{\chi}_k$:

$$\mathbf{x}_k^{*T} = [\mathbf{?}_1^T \mathbf{?}_2^T \cdots \mathbf{?}_k^T] \qquad (5)$$

From these equations, it is clear that the one step ahead prediction $\mathbf{?}_{k+1}$ corresponds to the forecast of the first J elements of the $\hat{\mathbf{x}}^{\#}$ vector:

$$\mathbf{?}_{k+1} = \hat{\mathbf{x}}_{k[1:J]}^{\#} \qquad (6)$$

Eq. 6 can be expressed as a function of the loadings matrix (Eq. 7), from Eq. (2), and taking only the first J rows of the $\mathbf{P_k}^{\#}$ matrix:

$$\mathbf{?}_{k+1} = \hat{\mathbf{x}}_{k[1:J]}^{\#} = \mathbf{P}_{k[1:J]}^{\#}\,\hat{\mathbf{t}}_k \qquad (7)$$

At time k, and assuming that projection to the model plane (PP_{OLS}) is used to solve the missing data problem, the estimate of the score is,

$$\hat{\mathbf{t}}_k^{\mathrm{T}} = \mathbf{x}_k^{*\mathrm{T}}\mathbf{P}_k^{*}(\mathbf{P}_k^{*\,\mathrm{T}}\mathbf{P}_k^{*})^{-1} \qquad (8)$$

and substituting Eq. 5 and Eq. 8 into Eq. 7:

$$\mathbf{?}_{k+1} = \hat{\mathbf{x}}_{[1:J]}^{\#} = \mathbf{P}_{k[1:J]}^{\#}\left[\mathbf{P}_k^{*}(\mathbf{P}_k^{*\,\mathrm{T}}\mathbf{P}_k^{*})^{-1}\right]^{\mathrm{T}}\begin{bmatrix}\mathbf{?}_1\\ \vdots\\ \mathbf{?}_{k-1}\\ \mathbf{?}_k\end{bmatrix} \qquad (9)$$

The term $\mathbf{P}_{k[1:J]}^{\#}\left[\mathbf{P}_k^{*}(\mathbf{P}_k^{*\mathrm{T}}\mathbf{P}_k^{*})^{-1}\right]^{\mathrm{T}}$ of Eq. 9 is a matrix of J rows and Jk columns; re-expressing each element of this matrix as $\phi_{i,j}$ and grouping now blocks of J columns as new square matrices $\boldsymbol{\Phi}_i$, such that each $\boldsymbol{\Phi}_i$ has J rows and J columns giving place to k number of Φ's,

$$\mathbf{?}_{k+1} = \underbrace{\begin{bmatrix}\phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,J}\\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,J}\\ \vdots & \vdots & & \vdots\\ \phi_{J,1} & \phi_{J,2} & \cdots & \phi_{J,J}\end{bmatrix}}_{\{\Phi_1\}}\underbrace{\begin{bmatrix}\phi_{J+1} & \phi_{1,J+2} & \cdots & \phi_{2,2J}\\ \phi_{2,J+1} & \phi_{2,J+2} & \cdots & \phi_{2,2J}\\ \vdots & \vdots & & \vdots\\ \phi_{J,J+1} & \phi_{J,J+2} & \cdots & \phi_{J,2J}\end{bmatrix}}_{\{\Phi_2\}}\cdots\underbrace{\begin{bmatrix}\cdots & \phi_{1,Jk}\\ \cdots & \phi_{2,Jk}\\ \ddots & \\ \cdots & \phi_{JJk}\end{bmatrix}}_{\{\Phi_K\}}\begin{bmatrix}\mathbf{?}_1\\ \vdots\\ \mathbf{?}_{k-1}\\ \mathbf{?}_k\end{bmatrix} \qquad (10)$$

gives the multivariate time series form for the one-step-ahead forecast embedded in the PCA model (Eq. 11):

$$\mathbf{?}_{k+1} = [\Phi_1\Phi_2\cdots\Phi_k]\begin{bmatrix}\mathbf{?}_1\\ \vdots\\ \mathbf{?}_{k-1}\\ \mathbf{?}_k\end{bmatrix} = \Phi_1\mathbf{?}_1 + \Phi_2\mathbf{?}_2 + \cdots + \Phi_{k-1}\mathbf{?}_{k-1} + \Phi_k\mathbf{?}_k \qquad (11)$$

This one step ahead forecast model has unique characteristics: a) The order of the PCA forecast model (Eq 15) expands, as new measurements are available (k increases); b) An even more important result is that the parameters of this model will adapt as the batch evolves since the elements in $\mathbf{P}^{\#}$ and $\mathbf{P}^{*}$ change as more samples become available.

This means that all the ϕ elements in Eq. 10 and 11 are adapted as time progresses to account for the changing auto-covariance and cross-covariance structure captured by the PCA model, this is possible because the MPCA batch model has embedded within it knowledge of the time varying covariance among all the variables over the entire time history of the batch.

This contrasts with the conventional time series model (Eq. 1) whose coefficients does not change and will always use the same number of past measurements in order to forecast the next sample. As a result, the PCA prediction equation is actually and adaptive, nonlinear function of the data available up to the prediction time k and hence provides a much more powerful prediction than any fixed multivariate time series model; c) The same form of adaptive time series prediction form of the PCA model cam be derived for any f-step ahead prediction with the same conclusions as above.

To further illustrate the adaptive nature of the PCA forecast model, consider the explicit forecast equation for the one step ahead prediction at different times. Figure 3 shows the values of the $\Phi_k^j(2,2)$ parameter for j= 5, 10, 20, 50, 80; and k=j,j-1, j-2…1; for the data set used by Nomikos and MacGregor in earlier work (Nomikos and MacGregor, 1995b).

As seen in this figure, the parameters adapt and expand, as more measurements are available. A result of the adaptive nature of this model is that the behavior of the forecast for the remaining unknown samples changes with time in the batch.

This "adaptive" behavior of the PCA predictive model is unique and results from the fact that the MPCA model built for a training set of complete batches has knowledge of how the auto and cross-covariance structure of all the variables will change over the reminder of the entire batch.

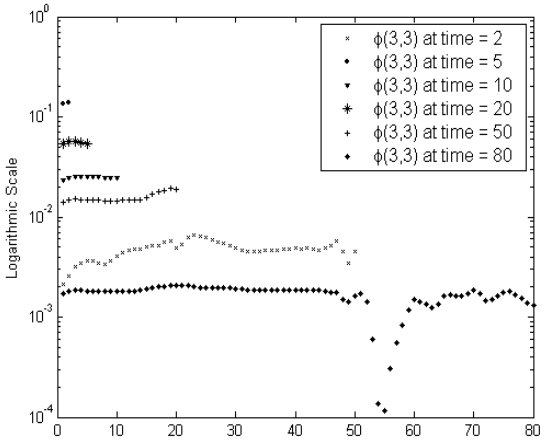

Fig. 3. Coefficient Φ_k^j (3,3) for known measurements as it adapts and expands with time

The tremendous amount of information captured by the MPCA model gives it unique forecasting capabilities, Fig. 4 shows the forecast for variable 4 in batch #7 in the industrial data set used by Nomikos and MacGregor (1995b), the forecast is done with information up to time 30, notice how accurately the prediction follows the true unknown part of the trajectory. The accuracy of the forecast done by the PCA is now contrasted among different options.

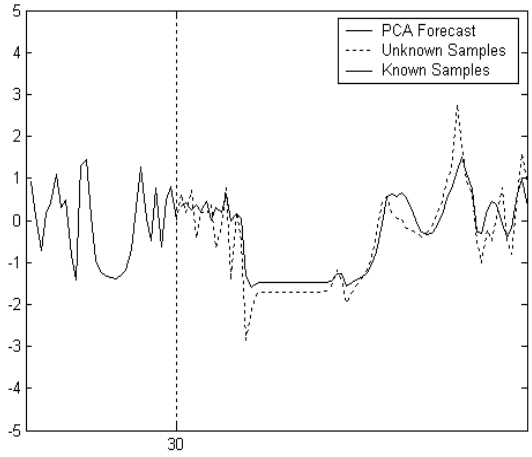

Fig. 4. Forecast done by the PCA model with samples up to time 30

3. ACCURACY COMPARISON

For batch i at time k *(given that* k <K*)*, the forecast for each variable j generates a vector of errors in the future prediction (Eq. 12)

$$\mathbf{e}_k^{ij} = [e_k^{ij}\big|_1 \; e_k^{ij}\big|_2 \; e_k^{ij}\big|_3 \; e_k^{ij}\big|_4 \; e_k^{ij}\big|_5 \cdots \; e_k^{ij}\big|_{K-k}] \quad (12)$$

defined by Eq. 13, and illustrated in figure 5. This (K-k) vector contains the individual errors in the prediction of the future unknown trajectory of variable j, for samples k+1, k+2, k+3…K.

$$\boldsymbol{e}_k^{ij} = \boldsymbol{x}_{ijk}^{\#} - \hat{\boldsymbol{x}}_{ijk}^{\#} \qquad (13)$$

This $\boldsymbol{e}_k^{ij}$ vector will decrease its length as the batch evolves, because there are fewer remaining unknown measurements ($\boldsymbol{x}^{\#}$).

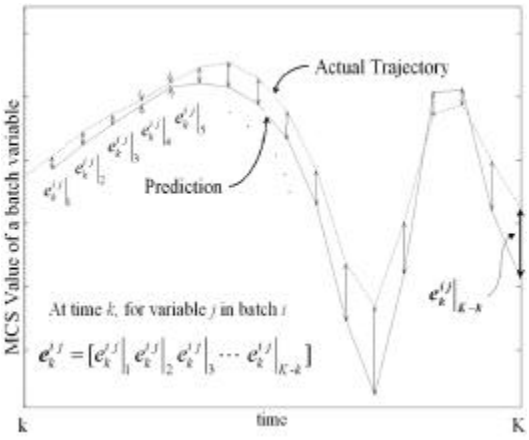

Fig. 5. Future prediction error in batch trajectory forecasting.

The original training batch data consists of I batches, each one with J variables, sampled K times throughout the batch. For each ijk sample in the original data set, we have a $\boldsymbol{e}_k^{ij}$ vector. This set of error vectors can be computed for each option available to treat the unknown part of the trajectory.

To compare all the error vectors among options, it is desirable to summarize the $\boldsymbol{e}_k^{ij}$ vector by one summary statistic in order to quantify the future prediction error as a scalar measurement of accuracy. This accuracy will vary in time and therefore can be plotted as a trajectory for each variable in each batch. To accomplish this, two different approaches are proposed:

a) To consider all individual errors ($[e_k^{ij}\big|_1\ e_k^{ij}\big|_2\ e_k^{ij}\big|_3 \cdots e_k^{ij}\big|_{K-k}]$) equally important and simply square them, and add them in to one number, the *Future Prediction Sum of Squares* (FPRESS) is then defined by Eq. 14

$$\boldsymbol{FPRESS}_k^{ij} = \sum_{l=1}^{K-k} \left(e_k^{ij}\big|_l \right)^2 \qquad (14)$$

b) A second option is to weight each individual error ($\boldsymbol{e}_k^{ij}\big|_l$) by the inverse of the distance to the current time sample, in such way that the one step ahead prediction $\boldsymbol{e}_k^{ij}\big|_1$ becomes more important than the (k-K) step ahead prediction $\boldsymbol{e}_k^{ij}\big|_{K-k}$; the weighted *Future Prediction Mean Square Error* (*FPMSE*) is defined in Eq. 15

$$\boldsymbol{FPMSE}_k^{ij} = \frac{\sum_{l=1}^{K-k} \frac{1}{l} \left(\boldsymbol{e}_k^{ij}\big|_l \right)^2}{\sum_{l=1}^{K-k} \frac{1}{l}} \qquad (16)$$

The *FPRESS* will provide a measure of global forecast accuracy, since it accounts for the error from time $k+1$ to time K, on the other hand, the *FPMSE* provides a measure of local forecast accuracy, considering the immediate forecast more important than the long term forecast.

Figure 6 shows the *mean trajectory* of the ***FPMSE*** for variable 4 in the industrial data set used by Nomikos and MacGregor (1995b), each line corresponds to a different way of handling the unknown part of the trajectory, a multivariate fixed time series model (TS), current deviations (CD), raw current deviations (CDraw), zeros (Z), and MD represents the missing data option handled with any method available, from single component projection to conditional mean replacement (Arteaga and Ferrer, 2002).

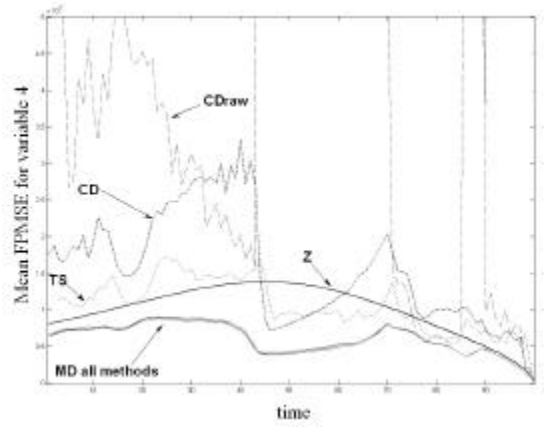

Fig. 6. Mean FPRESS (top) and mean FPMSE (bottom) for variable 1 in simulated SBR dataset.

It is very clear how, from all the methods, the MD options have the least error in the forecast at every time period over the whole trajectory, even from the very initial samples in the batch. This behavior was found to be true in general for all the variables for this data set for both of the two different future error trajectory statistics (*FPRESS* and *FPMSE*), although the CD and TS options on limited occasions showed slightly lower values at a few time intervals.

In batch process monitoring, when the missing data problem is handled correctly changes the initial suggestion of the *missing data* option being unreliable at the beginning of the batch, as illustrated in this work, and practically shown in the industrial application by Zhang, Dudzik and Vaculik (Zhang et al, 2003) where a monitoring system for start-ups is built solving the missing data problem with the PP method with excellent results even at the very

beginning of the start-up – startups can be monitored by using the same modeling technique as in batch monitoring (Kourti, 2003).

4. CONCLUSIONS

From the simulation results obtained in this work, and from a careful analysis of the methods, two general conclusions arise *i*) the *filling in* mechanism is actually a powerful time-varying model predictive *forecasting* feature that uses the PCA model of the time-varying covariance structure in the data, and *ii*) the *filling in* requirement in the Nomikos-MacGregor method for batch monitoring should not be seen as a problem to overcome. The multivariate (PCA/PLS) models obtained using the Nomikos-MacGregor approach for batch systems are dynamic models that capture the time-varying relations among the variables throughout the entire batch history. Therefore, they are capable of producing forecasts of the future unknown trajectories with optimal usage of the available data and knowledge of time-varying covariance structure among the variables. These forecasts are far from being arbitrary assumptions about the trajectories.

REFERENCES

Arteaga, F. and Ferrer, A. (2002) Dealing with missing data in MSPC: several methods, different interpretations, some examples. *Journal of Chemometrics*, **16**, 408-418

Boque, R. and Smilde, A. K. (1999) Monitoring and Diagnosing Batch Processes with Multivariate Covariates Regression Models. *AICHE Journal*, **45**, 1504-1520

Cho, H. and Kim, K (2003) A Method for Predicting Future Observations in the Monitoring of a Batch Process. *Journal of Quality Technology*, **35**, 59-69

Kosanovich, K., Dahl, K. S., and Piovoso, M. J. (1996) Improved Process Understanding Using Multiway Principal Component Analysis. *Industrial & Engineering Chemistry Research*, **35**, 138-146

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

Kourti, T., Nomikos, P., and MacGregor, J. F. (1995) Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *J.Proc.Cont.*, **5**, 277-284

Lennox, B., Hiden, H. G., Montague, G. A., Kornfeld, G., and Goulding, P. R. (2000) Application of multivariate statistical process control to batch operations. *Computers and Chemical Engineering*, **24**, 291-296

Meng, X., Morris, A. J., and Martin, E. B. (2003) On-line monitoring of batch processes using a PARAFAC representation. *Journal of Chemometrics*, **17**, 65-81

Nelson, P. R. C., Taylor, P., and MacGregor, J. F. (1996) Missing Data Methods in PCA and PLS: Score calculations with incomplete observations. *Chemometrics and Intelligent Laboratory Systems*, **35**, 45-65

Nomikos, P. (1996) Detection and diagnosis of abnormal batch operations based on multi-way principal component analysis. *ISA Transactions*, **35**, 259-266

Nomikos, P. and MacGregor, J. F. (1994) Monitoring Batch Processes using Multiway Principal Components. *AICHE Journal*, **40**, 1361-1375

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

Nomikos, P. and MacGregor, J. F. (1995b) Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics*, **37**, 41-58

Ramaker, H. J., van Sprang, E. N. M., Gurden, S. P., Westerhuis, J. A., and Smilde, A. K. (2002) Improved monitoring of batch processes by incorporating external information. *J.Process Contr.*, **12**, 569-576

Schlags, C. E. and Popule, M. P. (2001) Industrial Application of SPC to Batch Polymerization Process. *Proc.ACC (Arlington,VA)*, 25-27

Tates, A. A., Louwerse, D. J., Smilde, A. K., Koot, G. L. M., and Berndt, H. (1999) Monitoring a PVC batch process with multivariate statistical control charts. *Industrial & Engineering Chemistry Research*, **38**, 4769-4776

Undey, C., Ertunc, S., and Cinar, A. (2003) Online Batch/Fed-Batch Process Performance Monitoring, Quality Prediction, and Variable-Contribution Analysis for Diagnosis. *Industrial & Engineering Chemistry Research*, **42**, 4645-4658

van Sprang, E. N. M., Ramaker, H. J., Westerhuis, J., Gurden, S. P., and Smilde, A. K. (2002) Critical evaluation of approaches for on-line batch process monitoring. *Chemical Engineering Science*, **57**, 3979-3991

Westerhuis, J., Kourti, T., and MacGregor, J. F. (1999) Comparing Alternative Approaches for Multivariate Statistical Analysis of Batch Process Data. *Journal of Chemometrics*, **13**, 397-413

Wold, S., Geladi, P., Esbensen, K., and Ohman, J. (1987) Multi-way Principal Components and PLS-analysis. *Journal of Chemometrics*, **1**, 41-56

Wold, S., Kettaneh, N., Friden, H., and Holmberg, A. (1998) Modelling and diagnostics of batch process and analogous kinetic experiments. *Chemometrics and Intelligent Laboratory Systems*, **44**, 331-340

Zhang, Y, Dudzik, M., and Vaculik, V. (2003) Integrated Monitoring Solution to Start-up and run-time operations for continous casting. *Safeprocess2003.IFAC, Washington D.C.June 9-11*.