

Multirate Partial Least Squares for Process Monitoring

Ya Cong, Zhiqiang Ge*, Zhihuan Song

** State Key Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, Zhejiang, P. R. China
(Tel: +86-87951442; e-mail: gezhiqiang@zju.edu.cn).*

Abstract: In most chemical processes, variables are sampled at different rates which brings great challenges to traditional process monitoring methods that are built upon single sampling rate. In this paper, a multi-rate partial least squares algorithm is proposed. Compared to the traditional PLS method, the proposed algorithm takes use of the incomplete data samples through a modification of both of the covariance matrix of the input dataset and the covariance matrix between the input and output datasets. Iteration is used in the model training step to avoid the calculation of same parameters which requires complete training datasets. Then the fault detection and online prediction strategy is proposed based on this algorithm. A case study on TE process shows that the proposed method had an enhanced performance on both monitoring and online prediction, compared to the traditional PLS method.

Keywords: Multi-rate Process Monitoring, Partial Least Squares, Fault Detection, Online prediction

1. INTRODUCTION

In modern industry, process monitoring has become a key technology for performance improvement of process facilities. In most industrial processes, variables are sampled at different sampling rates, making both process monitoring and control more complex. For example, the values of quality relevant variables which are obtained from laboratory analysis are sampled at a low rate while process variables which are easy to measure are sampled at a fast rate. Those processes with various sampling rates for different variables are known as multi-rate processes (Li, W. et al, 2008). Generally, variables of multi-rate processes have three characteristics: (I) Incomplete. The slow sampling rate variables could not be achievable all the time, which leads to the incompleteness of the process data. (II) Regularly structured. Multi-rate process variables could be divided according to their sampling rates into regular data blocks. Under each data block, the data have the same number of samples. (III) Information asymmetry. Quality relevant variables only constitute a small part of the whole data while other process variables constitute most data samples.

In early works, multi-rate system identification has been widely investigated. For example, a dual-rate system which has only two sampling rates: the fast-rate and the slow-rate was studied (Lu, W. et al, 1988, 1989). A commonly used method called lifting technique was proposed by Li, D. et al (2001). When it comes to the multi-rate system identification problem, the fastest sampling rate is regarded as the base sampling rate and the unavailable data points in the slow sampling rate variables are usually treated as missing data (Raghavan, H. et al, 2006). Ding, F. et al (2004) proposed a FIR model on the multi-rate system to predict the unmeasured data points and then carried out identification between the fast-rate inputs and the predicted outputs.

Multivariate statistical process monitoring (MSPC) methods have also been widely researched since large amount of process data are available (Ge, Z. et al, 2013). Up-sampling methods and down-sampling methods are two typical way to build multi-rate models (Lu, N. et al, 2004). Shao, X. et al. (2011) proposed a Bayesian method for soft sensor model calibration. They proposed a soft sensor model using Bayesian method for the un-sampled data prediction as well as the model calibration. Wu, Y. et al (2010) proposed a Kalman filter based data rectification method using two different Kalman filters to estimate the quality property. Lu, N. et al. (2004) proposed multi-rate dynamic inferential modelling method for multi-rate dynamic system monitoring.

However, it is a common limitation that most of proposed models perform on dual-rate systems, which means that they might not be easily applied to three or more sample rates systems. Particularly, for down-sampling methods, the down-sampled data might contain a large part of fast-rate data points in which the slow-rate data points might be down. In other words, down-sampling methods might ignore some variable autocorrelations due to the high dimension of the down-sampled data. Another limitation of down-sampling methods is that it might encounter some difficulties when applying for online monitoring, because down-sampling substantially reduce the sample rate to a low level while the practical process is still sampled as a fast rate. On the other hand, up-sampling methods use estimated data points to build model, it might lead to a poor performance in online monitoring. Sometimes, the un-sampled data points could also be regarded as missing data. Walczak, B. et al (2001) proposed iterative PCA and PLS models for missing data processing, which made a compromise between the score matrix and the complete data matrix. Nelson, P.R. et al (1996) proposed score calculation method with incomplete observations in PCA and PLS models. Kim, D.S. et al. (2005) proposed a process monitoring method based on Factor

Analysis to deal with incomplete data. However, most missing-data methods assumed that the missing data are randomly dispersed and only consist of a small percentage of the whole data. But in multi-rate systems, the un-sampled data points only exist in slow-rate variables, and the missing percentage is always larger than 50%.

In this paper, a new multi-rate modelling method is developed for process monitoring, which is built on the basic PLS algorithm. In this method, a different way to calculate the covariance matrix is provided. In case that the model training datasets are not complete, an iteration is taken to avoid the calculation of parameters which demands the presence of the full-sampled model training datasets. The rest of paper is organized as follows. In section 2, a short review on PLS algorithm is given. Then the Multi-rate PLS model is proposed in section 3, followed by a case study on TE benchmark process. Finally, conclusions are made.

2. PARTIAL LEAST SQUARES

Partial least squares, also known as projection of latent structure is a dimensionality reduction technique that could extract the latent variables from two different blocks of variables. By extracting the projection directions that could maximize the variation as well as the correlation between the process variables \mathbf{X} and the quality variables \mathbf{Y} , PLS could predict the unmeasured quality variables with measured process variables. Given the input data matrix \mathbf{X} with n process variables and N samples, and the output data matrix \mathbf{Y} with m quality-relevant variables and M samples, they could be decomposed as:

$$\begin{cases} \mathbf{X} = \sum_{i=1}^a \mathbf{t}_i \mathbf{p}_i^T + \mathbf{E} = \mathbf{T} \mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} = \sum_{i=1}^a \mathbf{t}_i \mathbf{q}_i^T + \mathbf{F} = \mathbf{T} \mathbf{Q}^T + \mathbf{F} \end{cases} \quad (1)$$

where $\mathbf{T} = [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \cdots \quad \mathbf{t}_N]^T$ is the score matrix for both input and output data matrix while $\mathbf{P} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_a]$ is the loading matrix of \mathbf{X} and $\mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_a]$ is the loading matrix of \mathbf{Y} .

Every score vector \mathbf{t}_i is the linear combination of input variables. They could be calculated with the following equation:

$$\mathbf{T} = \mathbf{X} \mathbf{R} \quad (2)$$

where:

$$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \cdots \quad \mathbf{r}_a] \quad (3)$$

$$\mathbf{r}_i = \begin{cases} \boldsymbol{\omega}_1, i = 1 \\ \prod_{j=1}^{i-1} (\mathbf{I}_n - \boldsymbol{\omega}_j \mathbf{p}_j^T) \boldsymbol{\omega}_i, i \neq 1 \end{cases} \quad (4)$$

where $\boldsymbol{\omega}_i$ is the eigenvector of covariance matrix $\sum_{\mathbf{X}\mathbf{Y}}^i$ between input matrix \mathbf{X}_i and output matrix \mathbf{Y}_i .

$$\begin{cases} \mathbf{X}_i = \mathbf{X}_{i-1} - \mathbf{t}_{i-1} \mathbf{p}_{i-1}^T \\ \mathbf{Y}_i = \mathbf{Y}_{i-1} - \mathbf{t}_{i-1} \mathbf{q}_{i-1}^T \end{cases} \quad (5)$$

3. MULTI-RATE PARTIAL LEAST SQUARES (MRPLS)

Suppose we encounter a process with input data matrix \mathbf{X} which has α sampling rates and output data matrix \mathbf{Y} which has β sampling rates. The modelling data matrix could be represented as follows

$$\begin{cases} \mathbf{X} = [\mathbf{X}_{r_1} \quad \mathbf{X}_{r_2} \quad \cdots \quad \mathbf{X}_{r_\alpha}] \\ \mathbf{Y} = [\mathbf{Y}_{r_1} \quad \mathbf{Y}_{r_2} \quad \cdots \quad \mathbf{Y}_{r_\beta}] \end{cases} \quad (6)$$

Those sub-matrices might have different sizes from each other, and every two sub-matrices have some sample points at the same interval, as is shown in Fig. 1. We introduce \mathbf{X}_{ij} to represent these sub-matrices in \mathbf{X} , i represents the sub-matrix which the variables belong to and j represents the sub-matrix which is also sampled at these intervals. Parts of \mathbf{X}_{ij} are shown in Fig. 1.

Then, $\mathbf{X}_{ij}^{(\mathbf{Y})}$ and $\mathbf{Y}_{ij}^{(\mathbf{X})}$ are introduced to represent these sub-matrices from \mathbf{X} and \mathbf{Y} as shown in Fig. 2, in which i represents the sub-matrix which the variables belong to and j represents the sub-matrix which is also sampled at these intervals. The details are illustrated in Fig. 2.

Generally, multi-rate PLS involves two main steps, model training step and online prediction step.

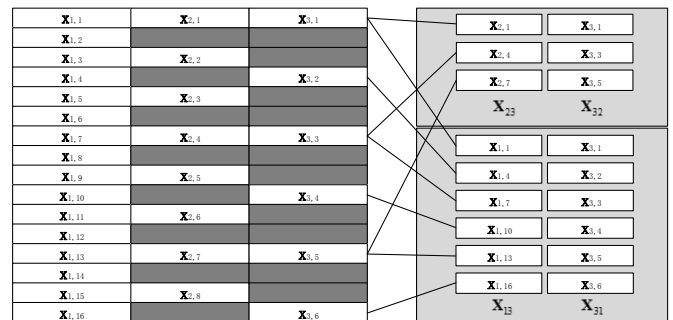


Fig. 1. Sub-matrixes with different sampling rates in the input data matrix.

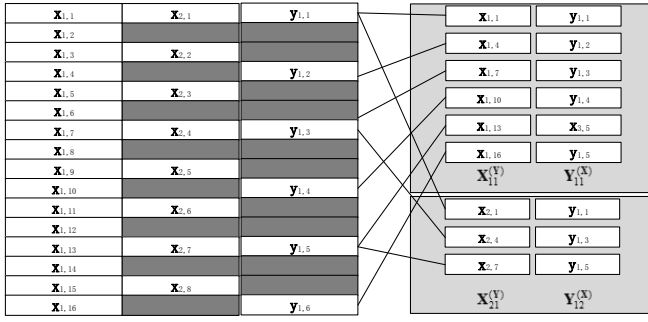


Fig. 2. Sub-matrices with different sampling rates between the input data matrix and the output data matrix.

2.1 Model Training Step

Suppose the training data contain α sampling rates in the input data matrix \mathbf{X} and β sampling rates in the output data matrix \mathbf{Y} . In MRPLS, the modelling step aims to acquire loading matrix \mathbf{P} and \mathbf{Q} . The first step is to centralize the model training matrix. After the centralization, the covariance matrix of the input data matrix \mathbf{X} and output data matrix \mathbf{Y} can be calculated by:

$$\Sigma_{\mathbf{XY}}^1 = \mathbf{X}^T \mathbf{Y} = \begin{bmatrix} \Sigma_{11}^{(XY)} & \Sigma_{12}^{(XY)} & \dots & \Sigma_{1\beta}^{(XY)} \\ \Sigma_{21}^{(XY)} & \Sigma_{22}^{(XY)} & \dots & \Sigma_{2\beta}^{(XY)} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{\alpha 1}^{(XY)} & \Sigma_{\alpha 2}^{(XY)} & \dots & \Sigma_{\alpha \beta}^{(XY)} \end{bmatrix} \quad (7)$$

Each of the items in the correlation matrix can be calculated by:

$$\Sigma_{ij}^{(XY)} = \frac{1}{N_{ij} - 1} (\mathbf{X}_{ij}^{(Y)})^T \mathbf{Y}_{ij}^{(X)} \quad (8)$$

The covariance matrix of the input data matrix \mathbf{X} can be calculated as:

$$\Sigma_{\mathbf{XX}}^1 = \mathbf{X}^T \mathbf{X} = \begin{bmatrix} \Sigma_{11}^{(XX)} & \Sigma_{12}^{(XX)} & \dots & \Sigma_{1\alpha}^{(XX)} \\ \Sigma_{21}^{(XX)} & \Sigma_{22}^{(XX)} & \dots & \Sigma_{2\alpha}^{(XX)} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{\alpha 1}^{(XX)} & \Sigma_{\alpha 2}^{(XX)} & \dots & \Sigma_{\alpha \alpha}^{(XX)} \end{bmatrix} \quad (9)$$

And each sub-matrix in the covariance matrix can be calculated by:

$$\Sigma_{ij} = \frac{1}{N_{ij} - 1} \mathbf{X}_{ij}^T \mathbf{X}_{ji} \quad (10)$$

where N_{ij} represents the sample number of \mathbf{X}_{ij} which is also the sample number of \mathbf{X}_{ji} , as shown in Fig. 1.

Suppose the first iteration is done. Before we move on to the next iteration of PLS, we shall remove the calculated score variation information from the original matrix, according to the procedure of the PLS algorithm.

$$\begin{cases} \mathbf{X}_2 = \mathbf{X}_1 - \mathbf{t}_1 \mathbf{p}_1^T \\ \mathbf{Y}_2 = \mathbf{Y}_1 - \mathbf{t}_1 \mathbf{q}_1^T \end{cases} \Rightarrow \begin{cases} \mathbf{X}_2 = \mathbf{X}_1 - \frac{1}{\xi_1^2} \mathbf{X}_1 \mathbf{w}_1 \mathbf{w}_1^T \mathbf{X}_1^T \mathbf{X}_1 \\ \mathbf{Y}_2 = \mathbf{Y}_1 - \frac{1}{\xi_1^2} \mathbf{X}_1 \mathbf{w}_1 \mathbf{w}_1^T \mathbf{X}_1^T \mathbf{Y}_1 \end{cases} \quad (11)$$

where:

$$\xi_1^2 = \left\| \mathbf{w}_1^T \Sigma_{\mathbf{XX}}^1 \mathbf{w}_1 \right\| \quad (12)$$

and:

$$\Sigma_{\mathbf{XY}}^2 = \left(\mathbf{I} - \frac{1}{\xi_1^2} \Sigma_{\mathbf{XX}}^1 \mathbf{w}_1 \mathbf{w}_1^T \right) \left(\mathbf{I} - \frac{1}{\xi_1^2} \Sigma_{\mathbf{XX}}^1 \mathbf{w}_1 \mathbf{w}_1^T \right) \Sigma_{\mathbf{XY}}^1 \quad (13)$$

$$\Sigma_{\mathbf{XX}}^2 = \left(\mathbf{I} - \frac{1}{\xi_1^2} \Sigma_{\mathbf{XX}}^1 \mathbf{w}_1 \mathbf{w}_1^T \right) \Sigma_{\mathbf{XX}}^1 \left(\mathbf{I} - \frac{1}{\xi_1^2} \mathbf{w}_1 \mathbf{w}_1^T \Sigma_{\mathbf{XX}}^1 \right) \quad (14)$$

With the two formulas we could move on to next iteration.

For the i^{th} iteration, we first calculate the eigenvector \mathbf{w}_i :

$$\Sigma_{\mathbf{XY}}^i (\Sigma_{\mathbf{XY}}^i)^T \mathbf{w}_i = \lambda_i^2 \mathbf{w}_i \quad (15)$$

With the eigenvector \mathbf{w}_i we can get the loading vector \mathbf{p}_i and \mathbf{q}_i :

$$\mathbf{p}_i = \Sigma_{\mathbf{XX}}^i \mathbf{w}_i / \xi_i^2 \quad (16)$$

$$\mathbf{q}_i = (\Sigma_{\mathbf{XY}}^i)^T \mathbf{w}_i / \xi_i^2 \quad (17)$$

Before we move on to the next iteration, we must pre-process the covariance matrix $\Sigma_{\mathbf{XY}}^{i+1}$ and the covariance matrix $\Sigma_{\mathbf{XX}}^{i+1}$:

$$\Sigma_{\mathbf{XY}}^{i+1} = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{\mathbf{XX}}^i \mathbf{w}_i \mathbf{w}_i^T \right) \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{\mathbf{XX}}^i \mathbf{w}_i \mathbf{w}_i^T \right) \Sigma_{\mathbf{XY}}^i \quad (18)$$

$$\Sigma_{xx}^{i+1} = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{xx}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \Sigma_{xx}^i \left(\mathbf{I} - \frac{1}{\xi_i^2} \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \Sigma_{xx}^i \right) \quad (19)$$

and:

$$\xi_i^2 = \left\| \boldsymbol{\omega}_i^T \Sigma_{xx}^i \boldsymbol{\omega}_i \right\| \quad (20)$$

Then we repeat the iteration until the residual information in \mathbf{Y} could be ignored. The loading matrix \mathbf{P} and \mathbf{Q} could also be available by combining the loading vectors:

$$\mathbf{P} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_a] \quad (21)$$

$$\mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_a] \quad (22)$$

where a is the number of latent variables selected in the MRPLS model.

2.2 Online Prediction and Monitoring Step

When an online input sample \mathbf{x}_{new} comes, the variables included in this sample are first recognized. Then, all the data corresponding to those variables are selected as the training input matrix, and all the historical output samples are selected as the training output matrix. Then the MRPLS model is build based on these historical data. The prediction of the new data sample \mathbf{x}_{new} is calculated by:

$$\tilde{\mathbf{y}} = \mathbf{Q} (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{x}_{new} \quad (23)$$

For process monitoring, the traditional T^2 and Q statistics can be calculated by:

$$T^2 = \mathbf{t}_{new}^T \boldsymbol{\Lambda}^{-1} \mathbf{t}_{new} \sim \frac{a(n-1)(n+1)}{n(n-a)} \mathbf{F}(a, n-a) \quad (24)$$

$$Q = \left\| \mathbf{x}_{new} - \mathbf{P} \mathbf{t}_{new} \right\|^2 \sim g \chi_h^2 \quad (25)$$

where

$$\mathbf{t}_{new} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{x}_{new} \quad (26)$$

$$\boldsymbol{\Lambda} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \Sigma_{xx}^1 \mathbf{P} (\mathbf{P}^T \mathbf{P})^{-1} \quad (27)$$

where n is the number of variables in the input data matrix, a is the number of latent variables, T^2 and Q statistics follow F-distribution and χ^2 distribution. The whole algorithm is given as follows.

Model training:

Step 1: according to the sampled vector \mathbf{x}_{new} , choose the variables of the model. Selecting all the historical data samples of these variables as the modelling data set \mathbf{X} and \mathbf{Y} ;

Step 2: centre the training data \mathbf{X} and \mathbf{Y} to zero mean and scale it to unit variance;

Step 3: calculate the covariance matrix Σ_{xx}^1 and Σ_{xy}^1 , set $i = 1$;

Step 4: calculate $\boldsymbol{\omega}_i$ by run singular value decomposition on

$\Sigma_{xy}^i (\Sigma_{xy}^i)^T (i = 1)$, and calculate \mathbf{p}_i and \mathbf{q}_i by

$$\begin{cases} \mathbf{p}_i = \frac{\Sigma_{xx}^i}{\left\| \boldsymbol{\omega}_i^T \Sigma_{xx}^i \boldsymbol{\omega}_i \right\|} \boldsymbol{\omega}_i \\ \mathbf{q}_i = \frac{(\Sigma_{xy}^i)^T}{\left\| \boldsymbol{\omega}_i^T \Sigma_{xx}^i \boldsymbol{\omega}_i \right\|} \boldsymbol{\omega}_i \end{cases};$$

Step 5: calculate Σ_{xy}^{i+1} and Σ_{xx}^{i+1} by

$$\begin{cases} \Sigma_{xy}^{i+1} = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{xx}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{xx}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \Sigma_{xy}^i \\ \Sigma_{xx}^{i+1} = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{xx}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \Sigma_{xx}^i \left(\mathbf{I} - \frac{1}{\xi_i^2} \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \Sigma_{xx}^i \right) \end{cases}$$

Step 6: if the singular value of the covariance matrix is too small then go to step 7, else return to step 3;

Step 7: get the loading matrix \mathbf{P} and \mathbf{Q}

$$\text{by} \begin{cases} \mathbf{P} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_a] \\ \mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_a] \end{cases};$$

Online prediction:

Step 8: calculate the prediction of the output vector $\tilde{\mathbf{y}}$ by

$$\tilde{\mathbf{y}} = \mathbf{Q} (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{x}_{new};$$

Online monitoring:

Step 9: centre the sampled vector \mathbf{x}_{new} , \mathbf{y}_{new} and scale it to unit variance, calculate T^2 statistics and Q statistics by equation (24) and (25);

Step 10: compare the online sample's statistics with the limits and report the process condition.

4. CASE STUDY OF TE PROCESS

The Tennessee Eastman Benchmark process was created by the Eastman Chemical Company to provide a realistic industrial process for evaluating process control and monitoring methods. (Downs, J.J. et al, 1993) This process is based on a simulation of an actual industrial process where the components, kinetics, and operating conditions have been modified for proprietary reasons. The process consists of five major units: a reactor, condenser, compressor, separator, and stripper; and, it contains eight components: A, B, C, D, E, F,

G and H. The gaseous reactants A, C, D, and E and the inert B are fed to the reactor where the liquid products G and H are formed (Chiang, L.H. et al, 2001). The process contains 41 measured and 21 manipulated variables. 22 measured variables are sampled every 3 minutes. 13 component measurements taken from Stream 6 and Stream 9 are sampled every 6 minutes. 5 component measurements taken from Stream 11 are sampled every 15 minutes. Every 30 minutes there will be a full sampled data points with all variables sampled at the same time. In this process, 21 fault cases can be simulated.

For MRPLS modelling, 22 measurement variables and 14 component variables from Stream 6 and Stream 9 are selected as the input data, and 5 component variables from Stream 11 are selected as the output data. A total of 4110 normal samples have been collected from this process, in which the first 3000 normal samples are denoted as training and the rest 1110 samples are used for cross-validation. The number of latent variables is determined as 6 according to the cross-validation method, the results of which are shown in Fig. 4.

Here, only the full sampled data samples are used for the comparison between MRPLS and PLS. Therefore, 96 samples out of 960 multi-rate samples from the first fault case are chosen to test both of the two models. The results of the monitoring and predicting are shown in Fig. 3 and Fig. 4, respectively. It is shown that the result of MRPLS performs almost the same as PLS. MRPLS changes the covariance matrix in the model training step which makes use of more fast-rate sampled data points to describe the correlation between them. However, the PLS model is built on 300 full-sampled samples which could lead to a quite satisfactory linear model. As a result, the model built on MRPLS algorithm has little distinction from PLS. However, it is reasonable to assume that MRPLS could preserve the same performance while PLS could not perform well when the number of available data samples decreases.

In Fig. 4, it is apparently that the sum of RMSE (Residual Mean Squared Error) of 5 output variables in MRPLS is smaller than that in PLS on each latent variable number. When the latent variable number is 6, both PLS and MRPLS have achieved the smallest RMSE values. Fig. 5 provides the detection rate of T2 statistic for both MRPLS and PLS with different amounts of training data, which are between 200 and 2200 with an increasing step as 20. In this figure, it is obvious that MRPLS achieves higher detection rates than PLS when the number of training data points is small.

5. CONCLUSIONS

In this paper, the MRPLS model is developed for monitoring and quality prediction of multi-rate sampling systems. Compared to the traditional PLS model, the MRPLS model can make full use of the process data under different sampling rates. For online monitoring and quality prediction, process data with various sampling rates can be monitored continuously by the MRPLS model, while the traditional PLS model is only valid when the process variables have been fully sampled. The performance of the proposed MRPLS model has been evaluated through the TE benchmark process.

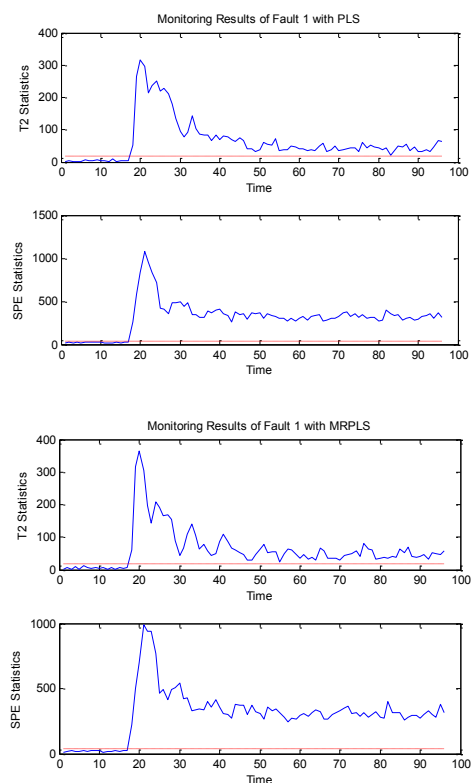


Fig. 3. Online monitoring of the first fault case in TE.

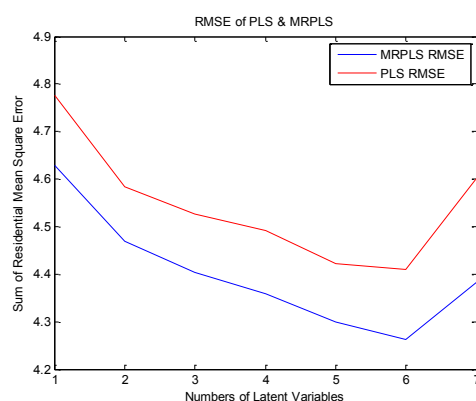


Fig. 4. RMSE of different number of latent variables.

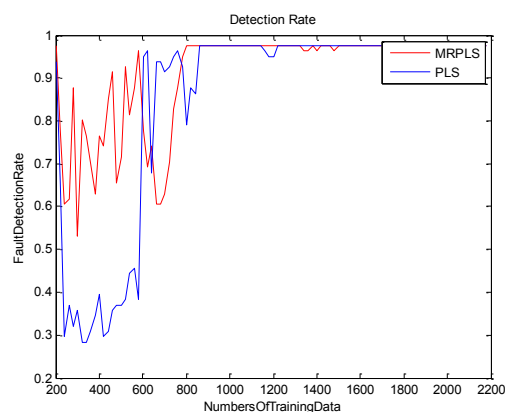


Fig. 5. Detection rate with different amounts of training data.

REFERENCES

- Chiang, L.H., R.D. Braatz, and E.L. Russell, *Fault detection and diagnosis in industrial systems*. 2001: Springer.
- Ding, F. and T. Chen, Identification of dual - rate systems based on finite impulse response models. *International Journal of Adaptive Control and Signal Processing*, 2004. 18(7): p. 589-598.
- Downs, J.J. and E.F. Vogel, A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 1993. 17(3): p. 245-255.
- Ge, Z., Z. Song, and F. Gao, Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 2013. 52(10): p. 3543-3562.
- Kim, D.S., et al., Calibration, prediction and process monitoring model based on factor analysis for incomplete process data. *Journal of chemical engineering of Japan*, 2005. 38(12): p. 1025-1034.
- Li, D., S.L. Shah, and T. Chen, Identification of fast-rate models from multirate data. *International Journal of Control*, 2001. 74(7): p. 680-689.
- Li, W., S.L. Shah, and D. Xiao, Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond. *Automatica*, 2008. 44(1): p. 199-208.
- Lu, N., et al., Multirate dynamic inferential modeling for multivariable processes. *Chemical Engineering Science*, 2004. 59(4): p. 855-864.
- Lu, W. and D. Grant Fisher, Output estimation with multirate sampling. *International Journal of Control*, 1988. 48(1): p. 149-160.
- Lu, W. and D. Grant Fisher, Least-squares output estimation with multirate sampling. *Automatic Control, IEEE Transactions on*, 1989. 34(6): p. 669-672.
- Nelson, P.R., P.A. Taylor, and J.F. MacGregor, Missing data methods in PCA and PLS: Score calculations with incomplete observations. *Chemometrics and intelligent laboratory systems*, 1996. 35(1): p. 45-65.
- Raghavan, H., et al., Identification of chemical processes with irregular output sampling. *Control Engineering Practice*, 2006. 14(5): p. 467-480.
- Shao, X., et al., Bayesian method for multirate data synthesis and model calibration. *AIChE Journal*, 2011. 57(6): p. 1514-1525.
- Walczak, B. and D. Massart, Dealing with missing data: Part I. *Chemometrics and Intelligent Laboratory Systems*, 2001. 58(1): p. 15-27.
- Wu, Y. and X. Luo, A novel calibration approach of soft sensor based on multirate data fusion technology. *Journal of Process Control*, 2010. 20(10): p. 1252-1260.

Appendix A. COVARIANCE MATRIX ITERATION PROOF

From Equation (11), Equation (18) and Equation (19), it is easy to get:

$$\begin{aligned} \begin{cases} \mathbf{X}_{i+1} = \mathbf{X}_i - \mathbf{t}_i \mathbf{p}_i^T = \mathbf{X}_i - \frac{1}{\xi_i^2} \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \mathbf{X}_i \\ \mathbf{Y}_{i+1} = \mathbf{Y}_i - \mathbf{t}_i \mathbf{q}_i^T = \mathbf{Y}_i - \frac{1}{\xi_i^2} \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \mathbf{Y}_i \end{cases} \\ \Sigma_{\mathbf{XY}}^{i+1} = \mathbf{X}_{i+1}^T \mathbf{Y}_{i+1} \\ = \left(\mathbf{X}_i^T - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \right) \left(\mathbf{Y}_i - \frac{1}{\xi_i^2} \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \mathbf{Y}_i \right) \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \mathbf{X}_i^T \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \right) \mathbf{Y}_i \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \left(\mathbf{X}_i^T - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \right) \mathbf{Y}_i \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \mathbf{X}_i^T \mathbf{Y}_i \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{\mathbf{XX}}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{\mathbf{XX}}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \Sigma_{\mathbf{XY}}^i \\ \Sigma_{\mathbf{XX}}^{i+1} = \mathbf{X}_{i+1}^T \mathbf{X}_{i+1} \\ = \left(\mathbf{X}_i^T - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \right) \left(\mathbf{X}_i - \frac{1}{\xi_i^2} \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \mathbf{X}_i \right) \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \mathbf{X}_i^T \mathbf{X}_i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \mathbf{X}_i^T \mathbf{X}_i \left(\mathbf{I} - \frac{1}{\xi_i^2} \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \mathbf{X}_i^T \mathbf{X}_i \right) \\ = \left(\mathbf{I} - \frac{1}{\xi_i^2} \Sigma_{\mathbf{XX}}^i \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \right) \Sigma_{\mathbf{XX}}^i \left(\mathbf{I} - \frac{1}{\xi_i^2} \boldsymbol{\omega}_i \boldsymbol{\omega}_i^T \Sigma_{\mathbf{XX}}^i \right) \end{aligned}$$