Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

A Modified Dynamic PLS for Quality Related Monitoring of Fractionation Processes

Xue Xu Qiang Liu Jinliang Ding

State Key of Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

(e-mail: xuxueneu@126.com; liuq@mail.neu.edu.cn; jlding@mail.neu.edu.cn).

Abstract: The fractionation process is a typical dynamic process, and practitioners highly pay attention to the quality-related abnormal in the real refining processes. In this paper, a modified dynamic PLS (MDPLS) modeling method and the corresponding process monitoring strategy are proposed. The main contributions of the proposed method are in the following. First, a clear dynamic relation is captured between process data and quality indices. Moreover, the process and quality space are comprehensively divided into dynamic quality-related subspace, static quality-unrelated subspace as well as the residual space for improving the performance of monitoring. Finally, the effectiveness of the proposed algorithm is demonstrated with the data from a real fractionation process.

Keywords: Fractionation process, Dynamic modeling, Quality related monitoring

1. INTRODUCTION

Fluid catalysis and cracking unit (FCCU) is an important process for converting heavy oil fractions into more valuable light production in petroleum processing. Fractionation column is an indispensable unit that produces gasoline and diesel oil in fluid catalysis and cracking processes. The two kinds of oil are major products of FCCU. Endpoint is a critical quality index of gasoline and diesel. Therefore, it is playing a dominant role in monitoring endpoint-related faults with the fractionation column to help improve product quality and operation safety.

The fractionation processes with multiple process variables are highly dynamics and correlated that make it be difficult to build first principles model accurately [1-3]. As a result, model-based process monitoring and fault diagnosis approaches are not applicable. Fortunately, due to the extensive application of distributed control system in recent years, a large number of valuable data with production process have been collected. The researchers rely on the available data to achieve data-based operation optimization, process monitoring as well as fault diagnosis [4-10]. Especially, multivariate statistical process monitoring (MSPM) is the most popularly utilized [11-15]. The principal component analysis (PCA), partial least squares (PLS) method project high-dimensional and correlated process variables into lower-dimensional latent space in chemical industry.

As for the fractionation processes, Thomas C et al. [16] employed PCA method for extracting latent variables, with two models based on hourly and minute-by-minute for monitoring distillation column of FCCU. Taking into account the time-varying behavior, such as changes of crude oil property, decay of catalyst activity in reactor, operation transform, Alghazzawi A et al. [17] proposed a recursive multi-block PCA monitoring strategy which can quickly identify and isolate abnormality in the crude distillation unit. In the real industry, auto-correlation and lag cross-correlation or dynamics exist among large-scale process variables. Therefore, Gao et al. [18] proposed indiscernibility dynamic kernel PCA, where two parameters--the indiscernibility and the degree of cross are defined for reducing data dimension. The algorithm that can extract the dynamics is utilized to monitor distillation column with lower missed alarm rate than the conventional method. Process monitoring based on PCA algorithm monitors only the process variables in unique layer, while abnormal variability may not affect product quality, due to compensation from feedback control in the fractionation processes. Unlike PCA, the PLS method is a powerful approach for establishing relationships of data from two layers [19, 20]. PLS-based process monitoring approaches extract latent factors according to the maximum co-variance criterion via nonlinear iterative partial least squares algorithm [21]. The work of [22] developed an improved multiscale PLS method, which combined PLS algorithm and wavelet analysis to build the model, with using generalized likelihood ration (GLR) testing in the residual space. The monitoring performance of this approach outperforms than PLS-based algorithm.

However, the above some methods only consider the static latent variables, which cannot capture the dynamics. Then it will result in high missed alarm in process monitoring. Although a few methods took into account the dynamics in terms of augmented data, which cannot make accurate description of the dynamic and static structure in the data. Therefore, in this paper, a new dynamic PLS method for quality-related monitoring is proposed and applied into fractionation processes. The contributions of the work are in the following. First, the process data space is decomposed into dynamic quality-related and static quality-unrelated subspace, respectively. For quality-unrelated space, it can contain a few large variations that is further divided by performing PCA decomposition to improve the monitoring performance.

The remainder of this paper is organized as follows. In section 2, fractionation process of FCCU is described. In section 3, the PLS method is reviewed briefly and a modified dynamic PLS approach is presented, and the corresponding monitoring strategy is given in later. The proposed algorithm is applied in the real fractionation processes in section 4. Finally, the conclusions are concluded and the following work is discussed in the section 5.

2. FRACTIONATION PROCESS DESCRIPTION

Fractionating column is an important unit in catalytic cracking processes, whose process structure diagram is shown in Figure 1. The high-temperature oil-gas mixture enters the fractionation column from column bottom, and then are separated several fractions according to different boiling temperatures. . Among them, overhead product is gasoline and side-lines product is diesel.

Gasoline and diesel are the main products of FCCU. The endpoint is an important quality index of gasoline and diesel. The major process variables relevant endpoint of gasoline are top tower temperature, top pressure, flow of top reflux, the 1st pump-around reflux flow, that of diesel are top pressure, column bottom temperature, the 1st pump-around reflux temperature, respectively. According to gasoline, diesel oil as the major fuel, endpoint of gasoline is extremely high, which will make it difficult to evaporate and burn the gasoline completely. The endpoint of diesel will affect the mobility under different climate conditions. Further, the endpoint also has an effect on product distribution in the production processes.

Therefore, a great deal of researchers focus on anomalies monitoring of fractionation process. In the early researches, researchers look forward to establish first principle model for solving the abnormal monitoring, whereas several assumptions are made to simplify the model because of difficulties in the solution of quite complex mechanism model. However, the above strategy will lead to a great difference between the built model and the real process [15]. As a result, considering the above drawbacks, multivariate statistical process monitoring method provides a data-driven framework to monitor the fractionation process. The fractionation process is a typical dynamic process, with dynamic characteristics derived from the interactions among complex recycle reflux. Therefore, a quality related modified dynamic PLS monitoring method is proposed for monitoring the fractionation processes.

3. A MODIFIED DYNAMIC PLS MODELING FOR PROCESS MONITORING

3.1 Reviews of PLS

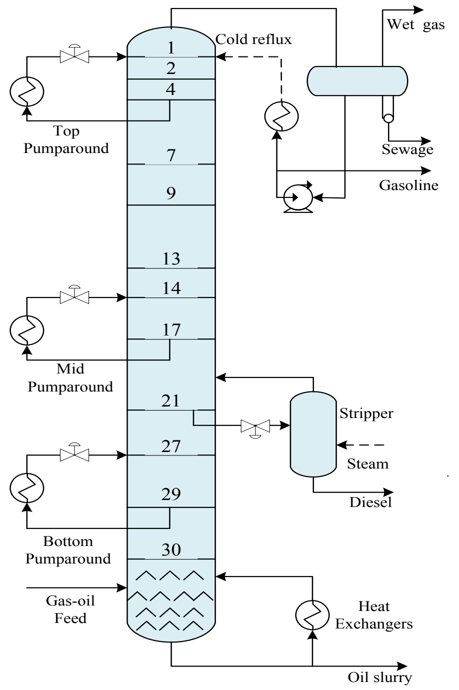


Fig.1 The flowchart around a typical fractionation column

Collect normal process and quality data and construct inputs $\mathbf{X} \in \Re^{n \times m}$ containing n samples with m variables, and outputs $\mathbf{Y} \in \Re^{n \times p}$ containing n samples with p variables. Then, the scaled data is projected to low-dimensional space as the following PLS model [23]:

$$\begin{cases} \mathbf{X} = \sum_i^l \mathbf{t}_i \mathbf{p}_i^T + \mathbf{E} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} = \sum_i^l \mathbf{t}_i \mathbf{p}_i^T + \mathbf{F} = \mathbf{T}\mathbf{Q}^T + \mathbf{F} \end{cases} \tag{1}$$

where $\mathbf{T} \in \Re^{n \times l}$ is the score matrix of inputs $\mathbf{X}$, $\mathbf{P} \in \Re^{m \times l}$ is the loading matrix of inputs $\mathbf{X}$, $\mathbf{Q} \in \Re^{p \times l}$ is the loading matrix of outputs $\mathbf{Y}$, l is the number of latent factors, $\mathbf{E}$ and $\mathbf{F}$ are residuals for inputs and outputs, respectively. The detailed PLS method is described in literatures [24, 25].

The matrix $\mathbf{W}$ is weight matrix of deflation matrix $\mathbf{X}_i = \mathbf{X}_i - \mathbf{t}_{i-1}\mathbf{p}_{i-1}^T$. In order to represent $\mathbf{t}_i$ in terms of original data $\mathbf{X}$, the following equation is utilized.

$$\mathbf{T} = \mathbf{X}\mathbf{R}$$

where $\mathbf{R} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}$ from [26].

3.2 A modified dynamic PLS (MDPLS) modelling algorithm

Generally, both auto-correlations and cross-correlations exist inside real industrial data simultaneously. The above static PLS model cannot capture dynamics in variables. As a result, missed alarm rate may increase when static model is applied to monitor dynamic processes. In this paper, a new dynamic latent variable monitoring algorithm is proposed and applied to quality-related fault detection.

The explicit dynamic and static relations between variables is modeled by the dynamic inner PLS (DiPLS) algorithm[27]:

$$\begin{cases} \mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} = \mathbf{G}_1(z^{-1})\mathbf{t}_1\mathbf{q}_1^T + \ldots + \mathbf{G}_l(z^{-1})\mathbf{t}_l\mathbf{q}_l^T + \mathbf{F} \end{cases} \tag{2}$$

where $\mathbf{G}_i(z^{-1}) = \alpha_i^T z$, $z = [1, z^{-1}, \ldots z^{-s+1}]$ describes the dynamic relationship between the quality data and the dynamic scores $\mathbf{t}_i$, z^{-1} denotes delay operator. l is the number of latent variables, which is obtained by cross-validation method.

However, the above approach focuses on building a regression model, rather than supply a statistical model for process monitoring. Further, the procedures of DiPLS algorithm is similar to that of PLS, thus scores T include variations orthogonal to Y. In order to overcome the aforementioned shortcomings, a modified monitoring strategy is proposed.

According to DiPLS method, the predicted quality variable can be shown [27]:

$$\hat{\mathbf{Y}} = \hat{\mathbf{T}}_s\mathbf{Q}^T = \mathbf{T}_c\mathbf{B}\mathbf{Q}^T = \mathbf{Z}_{s-1}\mathbf{R}\mathbf{B}\mathbf{Q}^T \tag{3}$$

where $\mathbf{B} = [\beta_{s-1}\ \beta_{s-2}\ \ldots\ \beta_1]^T$, $\mathbf{T}_c = [\mathbf{T}_1\ \mathbf{T}_2\ \ldots\ \mathbf{T}_{s-1}]$, with $\mathbf{T}_i = \mathbf{X}_i\mathbf{R}$.

The dynamic predictable quality variable $\hat{\mathbf{Y}}$ is performed singular value decomposition as following:

$$\hat{\mathbf{Y}} = \mathbf{L}_d\mathbf{C}_d\mathbf{H}_d^T \equiv \mathbf{L}_d\mathbf{Q}_d^T \tag{4}$$

where $\mathbf{Q}_d = \mathbf{H}_d\mathbf{C}_d$ contains l_d nonzero singular values in descending order and the corresponding right singular vectors.

$$\mathbf{L}_d = \hat{\mathbf{Y}}\mathbf{H}_d\mathbf{C}_d^{-1} = \mathbf{Z}_{s-1}\mathbf{R}\mathbf{B}\mathbf{Q}^T\mathbf{H}_d\mathbf{C}_d^{-1} = \mathbf{Z}_{s-1}\mathbf{R}_d \tag{5}$$

where $\mathbf{R}_d = \mathbf{R}\mathbf{Q}^T\mathbf{H}_d\mathbf{C}_d^{-1}$ For the remaining static quality-irrelevant process variables via projecting onto the orthogonal of Span $\{\mathbf{R}_d\}^{\perp}$, $\tilde{\mathbf{X}}_s = \mathbf{X} - \mathbf{K}_d\mathbf{R}_d^{\dagger}$,where $\mathbf{R}_d^{\dagger} = (\mathbf{R}_d^T\mathbf{R}_d)^{-1}\mathbf{R}_d^T$. Due to the large variation in the residual subspace, PCA decomposition is further realized on $\tilde{\mathbf{X}}_s$ with l_x components

$$\tilde{\mathbf{X}}_s = \mathbf{T}_x\mathbf{P}_x + \mathbf{E}_x \tag{6}$$

where $\mathbf{T}_x$ is the principal component score matrix that is useless to predict outputs, E_x represents residuals, respectively. According to the proposed MDPLS method, the outer model of input and output spaces are decomposed into the following form:

$$\begin{cases} \mathbf{X} = \mathbf{L}_d\mathbf{R}_d^{\dagger} + \mathbf{T}_x\mathbf{P}_x + \mathbf{E}_x \\ \mathbf{Y} = \hat{\mathbf{T}}_s^T\mathbf{Q}^T + \mathbf{F} \end{cases} \tag{7}$$

3.3 MDPLS-Based Monitoring

Given a new sample $\mathbf{x}_{new}$, it can be calculated as follows:

$$\begin{aligned} \mathbf{t}_d &= \mathbf{R}_d^T\mathbf{x}_{new} \\ \mathbf{t}_x &= \mathbf{P}_x^T\tilde{\mathbf{x}}_s \\ \mathbf{e}_x &= (\mathbf{I} - \mathbf{P}_x\mathbf{P}_x)\tilde{\mathbf{x}}_s \end{aligned} \tag{8}$$

The score vector $\mathbf{t}_x$ can be monitored via Hotelling's T^2 statistic:

$$T_x^2 = \mathbf{t}_x^T\mathbf{\Lambda}_x^{-1}\mathbf{t}_x \tag{9}$$

where $\mathbf{\Lambda}_x$ is the covariance of $\mathbf{t}_x$.

$\mathbf{e}_x$ can be monitored via Q statistic:

$$Q_x = \|e_x\|^2 \tag{10}$$

The $\mathbf{t}_d$ is not time independent series, which is not suitable to monitor. As a consequence, it can be described as a stationary time series model with auto-regressive mode as following:

$$t_d(k) = \gamma_1 t_d(k-1) + \ldots + \gamma_p t_d(k-g) + v(k) \tag{11}$$

where γ_i is the parameters of auto-regressive model, g is model order, it can be obtained by Bayesian Information Criterion. The parameters are solved by the least squares algorithm as following:

$$\hat{\Theta} = \left(\sum_{i=g+1}^{g+n}\varphi(i)\varphi(i)^T\right)^{-1}\left(\sum_{i=g+1}^{g+n}\varphi(i)t_d(i)^T\right) \tag{12}$$

where $\hat{\Theta} = \left[\gamma_1, \ldots \gamma_g\right]^T$, $\varphi(k) = \left[t_d(k-1), \ldots, t_d(k-g)\right]^T$.

As $\mathbf{v}_k$ can be time independent, it can be monitored by statistic T_v^2 .

$$\mathbf{T}_v^2 = \mathbf{v}_k^T\mathbf{\Lambda}_v^{-1}\mathbf{v}_k \tag{13}$$

where Λ_v is covariance of V , $\mathrm{V} = [\mathrm{v}_1, \ldots, \mathrm{v}_n]^{\mathrm{T}}$.

Table 2 Monitoring statistics and control limits

Statistic	Calculation	Control limit
T_x^2	$\mathbf{t}_x^T\mathbf{\Lambda}_x^{-1}\mathbf{t}_x$	$\frac{l_x(n^2-1)}{n(n-l_x)}F_{l_x,n-l_x;\alpha}$
T_v^2	$\mathbf{v}_k^T\mathbf{\Lambda}_v^{-1}\mathbf{v}_k$	$\frac{l_d(n^2-1)}{n(n-l_d)}F_{l_d,n-l_d;\alpha}$
Q_x	$\|e_x\|^2$	$(S_x/2\mu_x)\chi^2(2\mu_x/S_x)$

In this paper, the monitoring statistics and corresponding control limits are listed in the Table 2. Among them, where l_d, l_x are the number of the latent variables, n is the number of training samples, $F_{l_d,n-l_d;\alpha}$ indicates F -distribution with $l_d, n-l_d$ degrees of freedom under α confidence coefficient. $\chi^2(2\mu_x/S_x)$ represents χ^2 -distribution with $2\mu_x, S_x$ degrees of freedom, where S_x means the variance of the training samples, and μ_x represents mean of training samples, respectively [28, 29].

4. THE SCENARIO STUDIES IN QUALITY RELATED MONITORING OF FRACTIONATION PROCESSES

This section applies the proposed MDPLS method to endpoint relevant monitoring in fractionation processes. The collected data contain 28 process variables and 2 quality variables (endpoint of gasoline and endpoint of diesel). In this paper, a number of 600 samples were collected, of which

300 samples under fault-free condition, and additional 300 samples under fault condition. The prediction performance of two quality variables with PLS and dynamic PLS can be seen in the Fig. 3 and Fig. 4, respectively. The dominant process variables are shown in the Fig. 5. It is obvious that the dynamic PLS have better prediction accuracy than PLS.

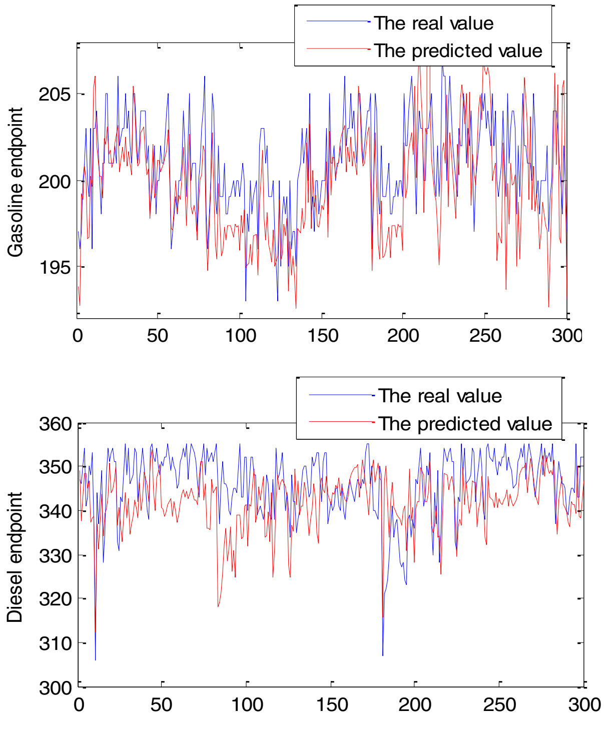


Fig.3. Quality prediction of PLS

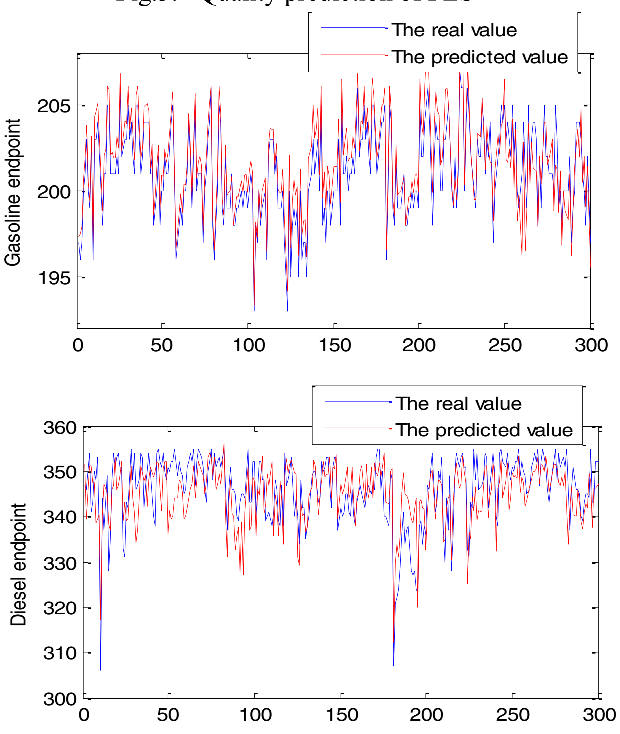


Fig.4 Quality prediction of dynamic PLS

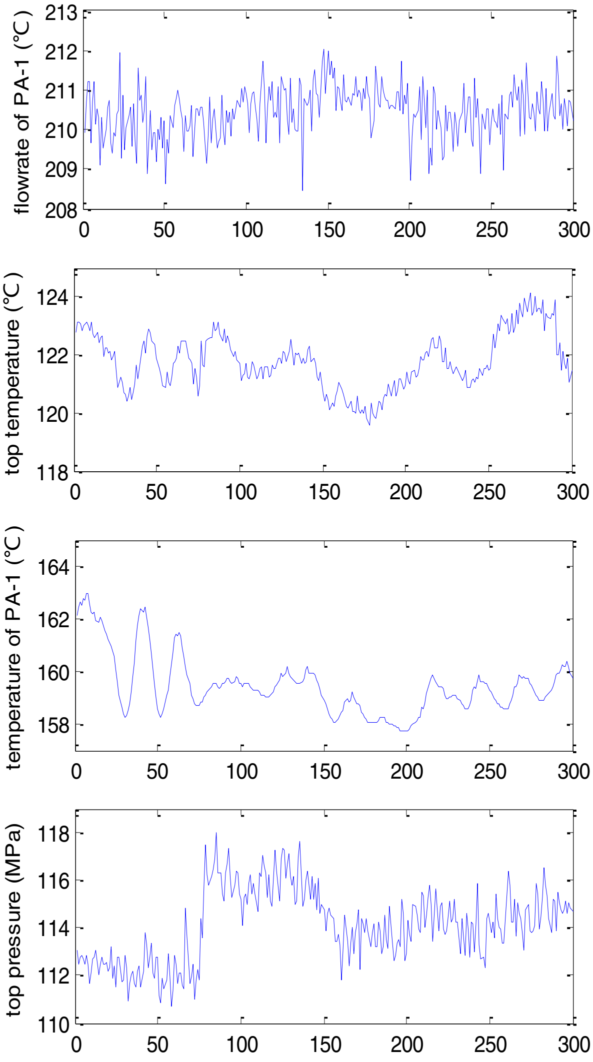


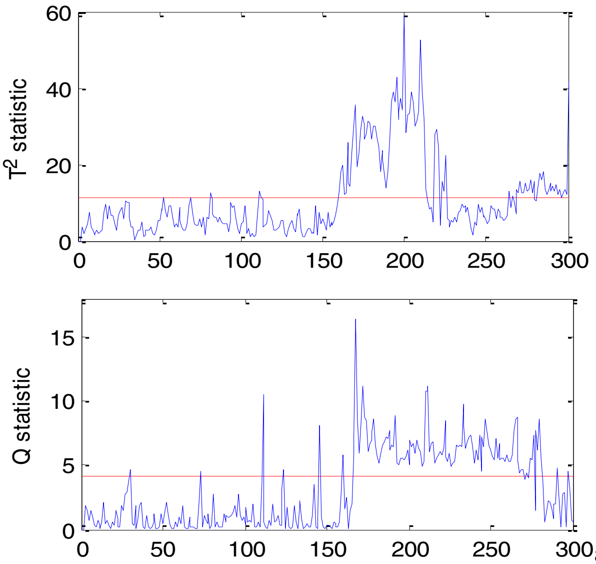
Fig.5 The original dominant process variables

To demonstrate the effectiveness of the algorithm, quality related faults 1 and faults 2 are modeled and monitored, comparing with conventional dynamic PLS (DPLS) methods. In the fractionation processes, 1st pump-around flowrate and the top tower pressure with too low are common faults. The former (fault 1) can lead to a higher temperature with the top and middle of the fractionating column, even further lead to the phenomenon of punching column, which will make the endpoint of gasoline and diesel rise, and also affect the heat source supply for the downstream devices. The latter (fault 2) will also have impact on the product quality.

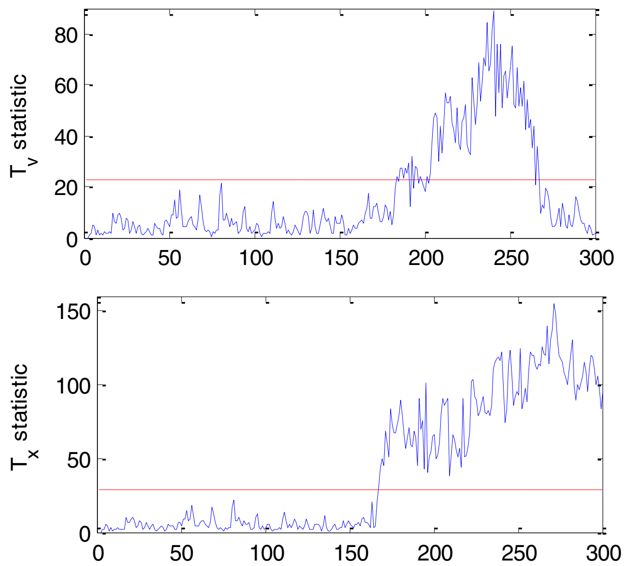
The monitoring performance with traditional DPLS and the MDPLS under fault scenario 1 is described in Fig.6. It is seen from T_y^2 statistic with Fig.6 (b) that the endpoint relevant fault is detected precisely from the 183th to 267th sample and have a trending of returning to the normal condition after the 267th samples because of the closed-loop controller to decrease the impact of the fault on process. Nevertheless, the T^2 statistic from Fig.6 (a) cannot detect the influence of feedback control on quality indexes, which still above the control limit after the 267th sample considering as the false alarms. Even if T_x^2 statistic in Fig 6 (b) above the control limit, which is a process-specific variation considering as normal situation when concentrating on the relevant endpoint fault. Therefore, the effectiveness of proposed monitoring method is demonstrated via the above experimental results.

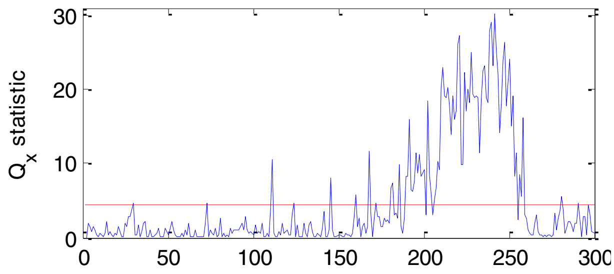
When the fractionation process under fault 2 situation, the monitoring results with DPLS and the MDPLS algorithms are displayed in Fig 7(a) and Fig (b), respectively. The T_v^2 statistic with Fig 7 (b) tends to exceed the control limit from the 145th sample, which denotes the quality-related faults can be dete cted successfully. However, the T statistic with DPLS method from Fig.7 (a) have missed alarms among 165-185, 231-240,267-280 samples.

Fault 1:



(a)…DPLS

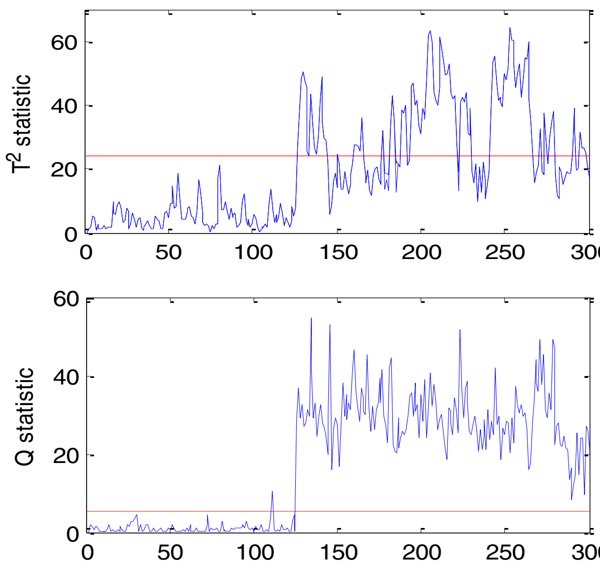




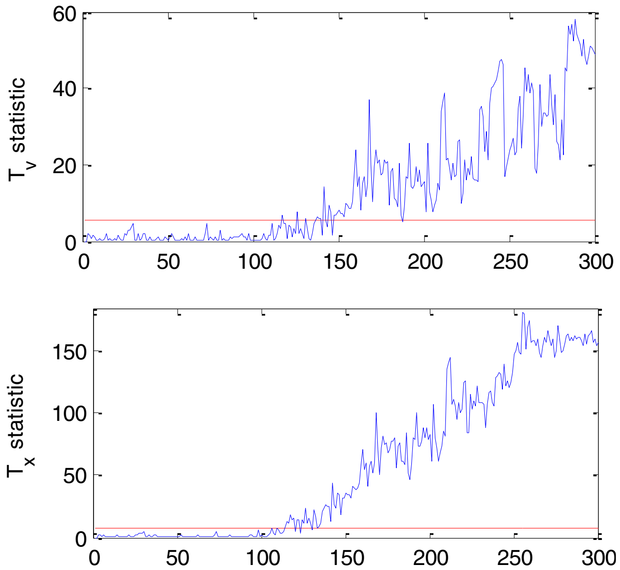
(b) MDPLS

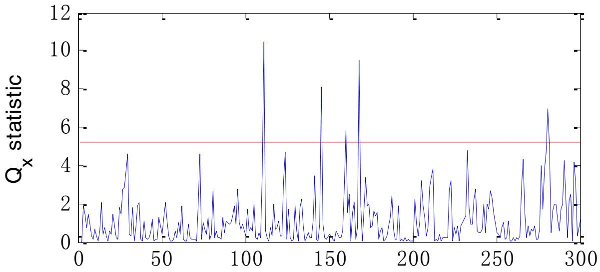
Fig. 6 Fractionation process monitoring results (faulty 1)

Fault 2:



(a) DPLS





(b) MDPLS

Fig.7 Fractionation process monitoring results (faulty 2)

5. CONCLUSIONS

This paper proposed a modified dynamic PLS quality-related monitoring approach for fractionation processes. It provides a simple implementation to achieve comprehensive monitoring. The effectiveness of the proposed method is illustrated by the application results on a real fractionation processes. The corresponding fault diagnosis strategy of the proposed monitoring scheme can be developed for localizing the faults in the future work.

ACKNOWLEDGEMENTS

This work was financially supported in part by the National Natural Science Foundation of China under Grants 61590922, 61525302, 61673097, 61573022, and 61490704, the Project of Ministry of Industry and Information Technology of China under Grand 20171122-6, the Projects of Shenyang under Grand Y17-0-004, and the Fundamental Research Funds for the Central Universities under Grand N160801001, N161608001.

REFERENCES

(1) S. J. Qin, G. Cherry, R. Good, J. Wang, and C. A. Harrison, Semiconductor manufacturing process control and monitoring: A fab-wide framework, Journal Process Control, 2006, 16(3), 179–191.
(2) Yin S, Li X, Gao H, Data-based techniques focused on modern industry: An overview, IEEE Transactions on Industrial Electronics, 2015, 62(1), 657-667.
(3) Z. Ge, Z. Song, and F. Gao, Review of recent research on data-based process monitoring, Industrial & Engineering Chemistry Research, 2013, 52(10), 3543–3562.
(4) Ding, Jinliang, et al. Data-Based Multiobjective Plant-Wide Performance Optimization of Industrial Processes Under Dynamic Environments, IEEE Transactions on Industrial Informatics, 2016, 12(2), 454-465.
(5) Ding J, Yang C, Chai T, Recent Progress on Data-Based Optimization for Mineral Processing Plants, Engineering, 2017, 3(2), 183-187.
(6) Q. Liu, S.J. Qin, T.Y. Chai, Multi-block concurrent PLS for decentralized monitoring of continuous annealing processes, IEEE Transactions on Industrial Electronics, 2014, 61 (11), 6429–6437.
(7) Tong C, Lan T, Shi X, Ensemble modified independent component analysis for enhanced non-Gaussian process monitoring, Control Engineering Practice, 2017, 58, 34-41.
(8) Q. Liu, S.J. Qin, T.Y. Chai, Decentralized fault diagnosis of continuous annealing processes based on multi-level PCA, IEEE Transactions on Automation Science and Engineering, 2013, 10 (3), 687–698.
(9) Tong C, Lan T, Shi X, Statistical process monitoring based on nonlocal and multiple neighborhoods preserving embedding model, Journal of Process Control, 2017, 65, 34-40.
(10) Zhu Y, Zhong Z, Zheng W X, HMM-Based H∞ Filtering for Discrete-Time Markov Jump LPV Systems Over Unreliable Communication Channels, IEEE Transactions on Systems Man & Cybernetics Systems, 2017, 99, 1-12.
(11) Zhao, C. H, Gao, F. R, Niu, D. P, Wang, F. L, Enhanced process comprehension and quality analysis based on subspace separation for multiphase batch processes, AIChE Jouinal, 2011, 57, 388-403.
(12) Wang G, Yin S, Quality-Related Fault Detection Approach Based on Orthogonal Signal Correction and Modified PLS. IEEE Transactions on Industrial Informatics, 2015, 11(2), 398-405.
(13) Li G, Qin S J, Zhou D, A New Method of Dynamic Latent-Variable Modeling for Process Monitoring, IEEE Transactions on Industrial Electronics, 2014, 61(61), 6438-6445.
(14) Chen T, Zhang J, On-line multivariate statistical monitoring of batch processes using Gaussian mixture model, Computers & Chemical Engineering, 2010, 34(4), 500-507.
(15) Peng X, Tang Y, Du W, Multimode Process Monitoring and Fault Detection: A Sparse Modeling and Dictionary Learning Method, IEEE Transactions on Industrial Electronics, 2017, 64(6), 4866-4875.
(16) Thomas C, Wada T, Seborg D E, Principal component analysis applied to process monitoring of an industrial distillation column. IFAC Proceedings Volumes, 1996, 29(1), 5859-5864.
(17) Alghazzawi A, Lennox B, Monitoring a complex refining process using multivariate statistics. Control Engineering Practice, 2008, 16(3), 294-307.
(18) Gao Q, Chang Y, Xiao Z, Monitoring of Distillation Column Based on Indiscernibility Dynamic Kernel PCA. Mathematical Problems in Engineering, 2016, 4, 1-11.
(19) Li G, Liu B, Qin S J, et al. Quality Relevant Data-Driven Modeling and Monitoring of Multivariate Dynamic Processes: The Dynamic T-PLS Approach, IEEE Transactions on Neural Networks, 2011, 22(12), 2262-2271.
(20) Kresta J V, Macgregor J F, Marlin T E. Multivariate statistical monitoring of process operating performance, Canadian Journal of Chemical Engineering, 1991, 69(1), 35-47.
(21) J. F. MacGregor and T. Kourti. Statistical process control of multivariate processes, Control Engineering Practice, 1995, 3(3), 403–414.
(22) Madakyaru M, Harrou F, Sun Y, Improved Data-based Fault Detection Strategy and Application to Distillation Columns, Process Safety & Environmental Protection, 2017, 107, 22-34.
(23) Geladi P, Kowalski B R, Partial least-squares regression: a tutorial. Analytica Chimica Acta, 1986, 185(86), 1-17.
(24) Hoskuldsson A, PLS regression methods. J Chemom. 1988, 2, 211–228.
(25) Kristian HellandHans E. BerntsenOdd S. BorgenHarald Martens, Recursive algorithm for partial least squares regression, Chemometrics & Intelligent Laboratory Systems, 1992, 14(1), 129-137.
(26) De Jong S, SIMPLS: an alternative approach to partial least squares regression, Chemometric and Intelligent Laboratory Systems, 1993, 18, 251–263.
(27) Dong Y, Qin S J, Dynamic-Inner Partial Least Squares for Dynamic Data Modeling, IFAC-Papers OnLine, 2015, 48(8), 117-122.
(28) S. J. Qin, Statistical process monitoring: Basics and beyond, Journal of Chemometrics, 2003, 17(8), 480–502.
(29) Box G E P, Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification, Annals of Mathematical Statistics, 1954, 25(2), 290-302.