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

Process Monitoring Based on Performance-Triggered Scheme

Ming-shan Zhang, Jian Yang, Shuai Tan, Hong-bo Shi*

* *Corresponding Author. Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education, East China University of Science and Technology, Shanghai 200237, China (e-mail:hbshi@ecust.edu.cn)*

Abstract: In process monitoring, some specific performance indexes need to be paid attention to. Therefore, the performance-triggered process monitoring scheme is proposed. Different from the traditional process monitoring method, the process is considered normal if there is no apparent anomaly happens on the performance index. In order to predict the values of performance indexes that cannot be measured in real time, ridge regression is used. And, the regression coefficients are used to pick the most relevant process variables for subsequent modelling. In this scheme, after the performance index exceeds the control limit, the monitoring of the relevant process variables is triggered to determine whether the prediction is abnormal due to the occurrence of a fault. Then, dictionary learning method and Low rank representation (LRR) are used for feature extraction and construction of the statistic. Finally, the effectiveness of the proposed method is verified by a numerical example and the Tennessee Eastman (TE) process.

Keywords: process monitoring, fault detection, variable selection, dictionary learning, low rank representation.

1. INTRODUCTION

With the rapid development of science and technology, the increasing requirements for modern industry are gaining more and more attentions, especially in terms of process performance. For an industrial process system, several performance indexes are concerned, such as, product quality, cost, and so on. Therefore, performance-driven process monitoring is becoming increasingly popular. Traditional multivariate statistical process monitoring (MSPM) methods use the full variable information of the data. For example, principal component analysis (PCA) and partial least square (PLS) algorithms are modelled using all process data. Some performance-independent process variables can lead to modelling inaccurate and computationally intensive. However, performance driven methods only select part of performance-related variables for modelling. So, several process monitoring methodologies, which based on process performance methods, have been widely proposed.

In the past few decades, least square (LS), principal component regression (PCR) and PLS have been universally studied. Moreover, the process monitoring strategies based on these algorithms are widely proposed. Zhou et al. divided the $\boldsymbol{X}$ -space into four parts and constructed four statistics separately to monitor. Just like total PLS (T-PLS), Wang et al. proposed total kernel PCR (T-KPCR), for quality-related fault detection for nonlinear systems. Wang et al. proposed an enhanced quality-related fault detection approach based on orthogonal signal correction (OSC) and modified-PLS (M-PLS). Compared with T-PLS, the proposed approach is more robust and has lower computational load. The above algorithms are all latent variable regression (LVR) models. LVR models use latent variables to establish regression relationships with performance indexes. In LVR, the input variables that are irrelevant to the performance indexes are used for modelling. When some specific fault with a large magnitude happened in these input variables, the performance indexes would be effected through the regression model. Therefore, it is necessary to select the relevant variables for regression. This paper presents a variable selection method based on the regression coefficient. Ridge regression is used to establish the relationship between input variables and the performance index. In this way, all input variables are divided into two parts, and the related one is used for regression and fault monitoring.

When the relevant variables are selected, the steps of feature extraction are required. Low rank representation (LRR) was initially used for image denoising. Recently, LRR has been widely applied to fault detection. In order to solve the problem of outliers in process monitoring, Pan et al. proposed an improved principal component pursuit method. According to the matrix obtained from the dictionary learning, the dataset can be decomposed into the low rank coefficient and the sparse residual term. The low rank coefficient obtained by LRR can capture the global structure of the dataset. When high-dimensional data can be mapped to low-dimensional linear space, low rank matrices can analyse this kind of data very well. Dictionary learning can find the most pristine features behind the sample. As an improved dictionary learning method, shift-invariant dictionary learning (SIDL) is very suitable to extract the data of periodic impulse information. Thus, dictionary learning can provide a more efficient method to capture the structure from process data.

In this paper, a performance-related variable selection method is proposed firstly. Different from the traditional method based on latent variable, such as PCR and PLS, the proposed approach is based on the regression coefficient. The input variables, whose regression coefficient exceed the threshold, will be considered to be relevant to the performance index. Therefore, it can avoid the impact of irrelevant variables on the performance index. Ridge regression is used to model the regression between the selected correlation variables and the performance index. A fusion of monitoring strategy, performance-triggered process monitoring, is proposed in the process monitoring. When the regression value of the performance index is under the control limit, then the process is considered normal. On the contrary, the process is abnormal. Regression is used for the advance judgment merely. Then we perform the feature extraction for all selected variables. Combined with LRR and dictionary learning, the structure of the dataset can be obtained accurately. Based on LRR, the dataset can be decomposed into low rank coefficients and sparse residuals. Eliminating sparse residuals can reduce the outliers of the dataset to a certain extent. The low rank coefficient is used to construct an S^2- statistic for process monitoring. Finally, the effectiveness of the proposed approach is proved by the Tennessee Eastman (TE) process.

The rest of this paper is organized as follows. In section 2, ridge regression and LRR are briefly illustrated. In section 3, the proposed variable selection method based on the regression coefficient is implemented in detail. Then, a fusion of monitoring strategy is described. In section 4, a numerical simulation and TE process are tested utilizing the proposed scheme. Finally, section 5 is conclusion.

2. PRELIMINARIES

2.1 Ridge regression

For a known matrix $\boldsymbol{X}$ and vector y, the goal of ridge regression is to find a vector w satisfied $\boldsymbol{X}w = y$. A regularization term can be introduced to calculate the minimization:

$$\min\left(\|\boldsymbol{X}w - y\|^2 + \|\boldsymbol{\Gamma}w\|^2\right) \tag{1}$$

where $\boldsymbol{\Gamma}$ is the Tikhonov matrix. Sometimes, $\boldsymbol{\Gamma}$ is defined as $\alpha\boldsymbol{I}$, namely $\boldsymbol{\Gamma}=\alpha\boldsymbol{I}$. And, $\boldsymbol{I}$ is the unit matrix. α is the coefficient. Therefore, equation (2) can be solved directly by:

$$w = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}+\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}y \tag{2}$$

The effect of regularization is changed by adjusting the scale of $\boldsymbol{\Gamma}$.

2.2 Low rank representation (LRR)

Given a set of training data $\boldsymbol{X}=[\boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \cdots, \ \boldsymbol{x}_n]^{\mathrm{T}} \in \boldsymbol{R}^{\mathrm{n\times m}}$, In LRR, $\boldsymbol{X}$ can be decomposed as:

$$\boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{E} \tag{3}$$

where $\boldsymbol{D}$ is a dictionary, which can represent the feature of matrix $\boldsymbol{X}$. $\boldsymbol{Z}$ is the low rank coefficient matrix and $\boldsymbol{E}$ is the sparse residue term. If dictionary $\boldsymbol{D}$ is a global information map for dataset $\boldsymbol{X}$, the low rank coefficient $\boldsymbol{Z}$ can capture the global structures of $\boldsymbol{X}$. The sparse residue term $\boldsymbol{E}$ contains the outliers of the dataset. The coefficient $\boldsymbol{Z}$ can be obtained by solving the problem.

$$\begin{aligned} \min \ & \|\boldsymbol{Z}\|_* + \lambda\|\boldsymbol{E}\|_1 \\ s.t. \ & \boldsymbol{X} = \boldsymbol{D}\boldsymbol{Z} + \boldsymbol{E} \end{aligned} \tag{4}$$

where $\|\boldsymbol{Z}\|_*$ denote the nuclear norm of the matrix $\boldsymbol{Z}$. $\|\boldsymbol{E}\|_1$ denote the l_1 norm of the matrix $\boldsymbol{E}$. Through the LRR algorithm, the outliers can be removed. And, the low rank coefficient is used to construct the statistics for process monitoring.

3. MONITORING SCHEME BASED ON RIDGE REGRESSION AND LRR

In this section, a new process monitoring scheme based on Ridge regression and LRR is proposed, including Ridge regression for related variables selection and performance index prediction, and LRR for feature extraction and process monitoring.

3.1 Variables selection and performance index regression

When focusing on the performance index only, the performance-independent variables will lead to training dataset redundancy. Different form the latent variable model, all the process variables are divided into two parts in the proposed method. In this section, performance-related variables are selected by the Ridge regression algorithm. The regression coefficient is a linear representation of the input variable and the output variable. Through the regression coefficient, the process variables associated with the performance index can be selected.

Given the measured variables $\boldsymbol{X}=[\boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \cdots, \ \boldsymbol{x}_n]^{\mathrm{T}} \in \boldsymbol{R}^{\mathrm{n\times m}}$ with m -dimensional and n samples. And then, the performance index $y \in \boldsymbol{R}^{\mathrm{n\times 1}}$ can be obtained in the training dataset. In order to eliminate the effect of the performance-independent variables in the prediction of the performance index, the original variables $\boldsymbol{X}$ are divided into two parts. And, the part related to the performance index is used for regression and fault detection. The variables are divided by the regression coefficient $w=[\mathrm{w}_1, \ \mathrm{w}_2, \ \cdots, \ \mathrm{w}_m]^{\mathrm{T}}$. Thus, Ridge regression is used to obtain the regression coefficient. Performance index y can be expressed as

$$y = w_1 \bullet x_1 + w_2 \bullet x_2 + \cdots + w_m \bullet x_m \tag{5}$$

Firstly, w is subjected to normalization processing.

$$w(\mathrm{i})=\frac{|w(\mathrm{i})|}{\sum|w|} \tag{6}$$

Then, w is sorted in descend order. Finally, a cumulative percent coefficient (CPC) is proposed to determine the threshold.

$$CPC=\sum_{i=1}^{k} w(\mathrm{i}) \tag{7}$$

When CPC is greater than 0.9, the corresponding variables are considered to be related to the performance index, i.e. performance index related subspace $\boldsymbol{X}_{rel} \in R^{n\times k}$. The remaining variables make up the performance index independent subspace $\boldsymbol{X}_{ind} \in R^{n\times(m\text{-}k)}$, $\boldsymbol{X}=\left[\boldsymbol{X}_{rel} \quad \boldsymbol{X}_{ind}\right]$.

where k is the number of variables related to the performance index.

After the relevant variable $\boldsymbol{X}_{rel}$ is selected, the regression model between $\boldsymbol{X}_{rel}$ and y can be established by Ridge regression.

$$\hat{w}=\left(\boldsymbol{X}_{rel}{}^{\mathrm{T}}\boldsymbol{X}_{rel}+\alpha \boldsymbol{I}\right)^{-1}\boldsymbol{X}_{rel}{}^{\mathrm{T}}y \tag{8}$$

Because the dimension of the process variable is reduced, the computational complexity and the computational memory for calculating the regression coefficient $\hat{w}$ can be saved significantly. Online monitoring, according to the division of the variable space, performance index can be regressed by the relevant variables and $\hat{w}$.

$$\hat{y}_{new}=x_{new}\hat{w} \tag{9}$$

where x_{new} is a new sample with the performance index related variables. $\hat{y}_{new}$ is the predicted value of the performance index.

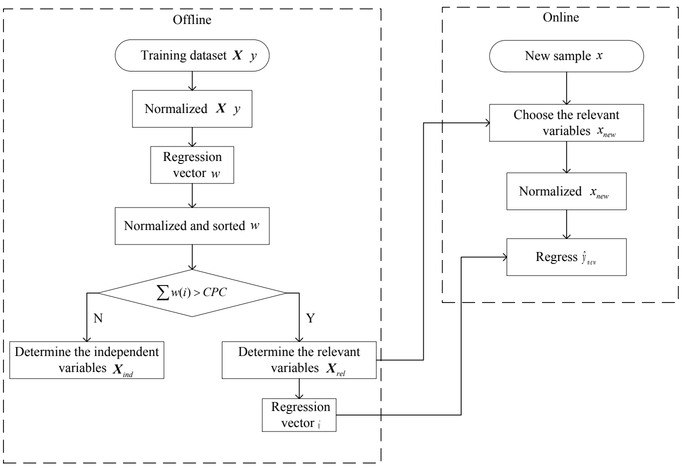

Fig. 1. Variable selection and index regression.

3.2 LRR for feature extraction and process monitoring

As mentioned above, the measured variables $\boldsymbol{X}$ are divided into $\boldsymbol{X}_{rel}$ and $\boldsymbol{X}_{ind}$. In order to detect faults associated with performance index, the part of $\boldsymbol{X}_{rel}$ is used for analysis. In LRR, the low rank coefficient $\boldsymbol{Z}$ contains the low rank information of dataset $\boldsymbol{X}_{rel}$. The relationship between the variables of $\boldsymbol{X}_{rel}$ is represented by the coefficient $\boldsymbol{Z}$. However, dictionary $\boldsymbol{D}$ needs to be determined in LRR. In this section, a dictionary learning method is proposed to solve the problem.

To elaborate, for the dataset $\boldsymbol{X}_{rel}=\left[\boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \cdots, \ \boldsymbol{x}_k\right]^{\mathrm{T}} \in \boldsymbol{R}^{\mathrm{n}\times\mathrm{k}}$, the global structure of the data is extracted. Firstly, in order to maintain the fairness of the data, $\boldsymbol{X}_{rel}$ is normalized. Then, the covariance matrix $\boldsymbol{S}$ is computed.

$$\boldsymbol{S}=\frac{1}{n-1}\boldsymbol{X}_{rel}{}^{\mathrm{T}}\boldsymbol{X}_{rel} \tag{10}$$

By singular value decomposition (SVD), $\boldsymbol{S}$ can be decomposed into:

$$\boldsymbol{S}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} \tag{11}$$

Where $\boldsymbol{\Lambda}$ is a diagonal matrix, and each diagonal element represents a nonnegative eigenvalue λ_i . Each column of the matrix $\boldsymbol{V}$ is the corresponding eigenvector. Finally, the dictionary $\boldsymbol{D}$ can be defined as the projection of the original dataset $\boldsymbol{X}_{rel}$ under the matrix $\boldsymbol{V}$.

$$\boldsymbol{D}=\boldsymbol{X}_{rel}\boldsymbol{V} \tag{12}$$

Once the dictionary $\boldsymbol{D}$ is determined, the low rank coefficient $\boldsymbol{Z}$ can be calculated by (4). Therefore, LRR can be achieved by Pan. In this way, the low rank coefficient matrix $\boldsymbol{Z}$ and the sparse residue term $\boldsymbol{E}$ can be obtained.

In this section, a statistic based on low rank coefficient is constructed for process monitoring. And $\boldsymbol{Z}$ can be regarded as an approximate representation to the original dataset in the lower rank space. Therefore, the projection of a new sample on $\boldsymbol{Z}$ can be calculated.

$$S^2=x_{\mathrm{new}}*\boldsymbol{Z}*\boldsymbol{Z}^{\mathrm{T}}*x_{\mathrm{new}}{}^{\mathrm{T}} \tag{13}$$

In offline modelling, according to the training dataset, a set of statistics can be obtained. Kernel density estimation (KDE) is used for determination of control limit. If the S^2- statistic exceeds the threshold, the fault is considered. On the contrary, the process is considered normal.

3.3 Monitoring scheme

According to the section 3.1 and 3.2, how to deal with the relationship between y_{new} and S^2 is the focus of this section. Thus, performance-triggered process monitoring scheme is proposed to solve the problem.

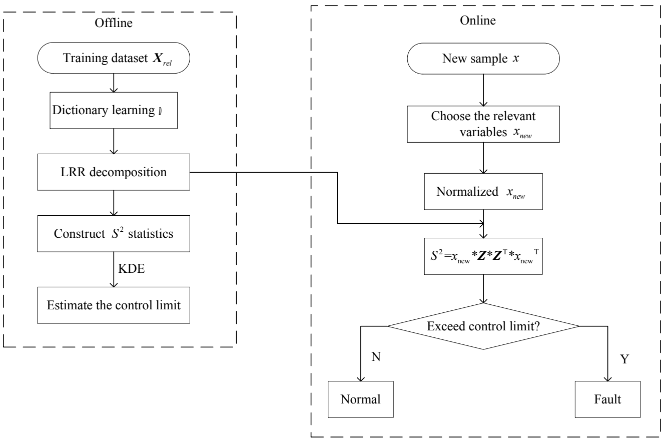

Fig. 2. Process monitoring based on LRR.

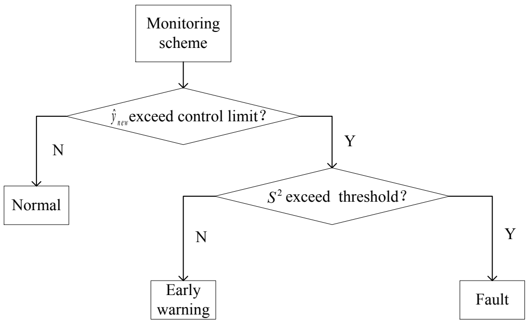

Fig. 3. The monitoring scheme.

In short, the monitoring scheme is divided into two parts. The first part is the prediction of performance index. When the performance index is predicted, we can evaluate the process according to whether the index exceeds the control limit. From the Fig. 3, if the performance index does not exceed the control limit, then the system is considered normal. On the contrary, it is considered abnormal. When the process is judged to be abnormal, further discrimination can be made on the basis of the variables associated with the performance index. And this is the second part of the monitoring scheme. According to this part, we can be determined whether the early warning or fault.

4. CASE STUDIES AND DISUSSION

4.1 *Numerical example*

In this section, a numerical example is used to illustrate the method of selecting the relevant variables for ridge regression.

$$y(\mathrm{t}) = \sum_{i=1}^{m} w_i x_i(\mathrm{t}) + \varepsilon(\mathrm{t}) \tag{14}$$

where w_i is the regression coefficient, and $\varepsilon(\mathrm{t})$ is the noise signal. There are 400 normal samples in the example. $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_m$ are input variables, and y is the performance index. W is set to

$W = [0.032 \quad 2.325 \quad -1.542 \quad 0.212 \quad 0.63 \quad -0.002 \quad 1.23 \quad -0.5 \quad 3.2 \quad 1.3]$

Now, we can select the relevant variables according to the above mentioned method. Table 1. shows the difference between true coefficients and estimated coefficients.

Table 1. True coefficients and estimated coefficients

Variables	True w	Estimated w	Variables	True w	Estimated w
$\boldsymbol{x}_1$	0.032	0.0322	$\boldsymbol{x}_6$	-0.002	-0.002
$\boldsymbol{x}_2$	2.325	2.338	$\boldsymbol{x}_7$	1.23	1.2286
$\boldsymbol{x}_3$	-1.542	-1.5401	$\boldsymbol{x}_8$	-0.5	-0.4997
$\boldsymbol{x}_4$	0.212	0.2121	$\boldsymbol{x}_9$	3.2	3.1973
$\boldsymbol{x}_5$	0.63	0.6293	$\boldsymbol{x}_{10}$	1.3	1.2989

From table we can know that the estimated coefficients are very close to the true coefficients. So, the variables x_9, x_2, x_3, x_{10}, x_7, x_5 are selected to regress the performance index.

To illustrate the regression performance, R^2 is introduced.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{15}$$

in which x_i is the i-th element of the variable, $\hat{x}_i$ is the corresponding regression value and $\bar{x}$ denotes the mean of the real variable. The value of R^2 in this numerical example is 0.9894. Therefore, the method of selecting the relevant variables for regression is effective.

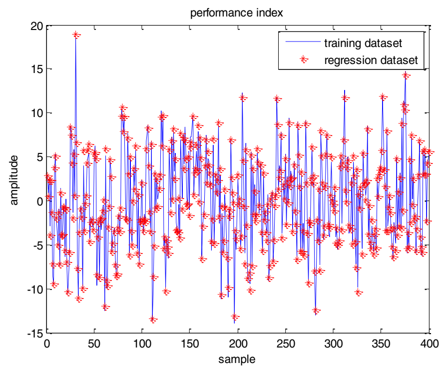

Fig. 4. Comparison of training and regression dataset.

4.2 *Tennessee Eastman (TE) Process*

Tennessee Eastman Process was proposed by Down, and now it is an important benchmark in chemical process. It contains 12 manipulated variables (XMV1-12) and 41 measured variables (XMEAS1-41). The measured variables are divided into 22 process measured variables and 19 component measured variables. And 15 known faults are simulated for algorithm testing. In our work, we select 11 manipulated variables and 22 process measured variables as the input dataset, and the component G in steam 9 (XMEAS35) as the performance index. From Fig.5 we can pick out the most relevant input variables, where the parameters λ and CPC

are set to 0.2 and 0.9. Thereafter the block $\boldsymbol{X}_{rel}$ and $\boldsymbol{X}_{ind}$ are determined. The next step is to regress the performance index. And then, a dictionary is learned through the information of $\boldsymbol{X}_{rel}$. In this way, the global structure of $\boldsymbol{X}_{rel}$ is maintained. Finally, we use the low rank coefficient matrix to construct the statistic for process monitoring. Table shows the different detection results of PCR and the proposed method.

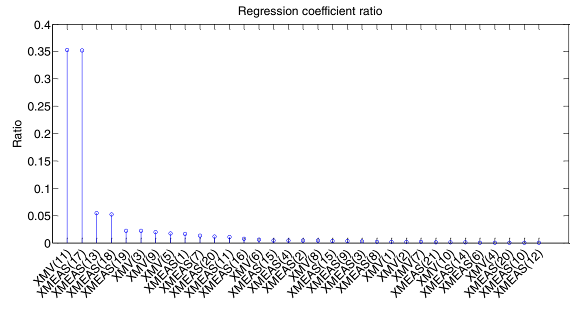

Fig. 5. The regression coefficient ratio of each variable.

Table 2. Fault detection value of PCA PCR and RR_LRR (%)

Fault	T^2(PCA)	SPE(PCA)	T^2(PCR)	S^2
IDV(1)	99.125	99.5	22.75	99
IDV(2)	97.75	98.625	24.5	93
IDV(3)	0.875	1.75	0.625	0.25
IDV(4)	6.375	93.875	0.375	0
IDV(5)	21.5	27.75	10.625	19.75
IDV(6)	98.875	100	97.125	100
IDV(7)	39	100	22.625	35.5
IDV(8)	92.75	97.5	71.5	88.875
IDV(9)	1	2.5	0	0
IDV(10)	28.25	23.375	13	33.75
IDV(11)	21.75	64.625	1.125	0.875
IDV(12)	95.25	97	67.5	87.375
IDV(13)	91.875	95.25	72.5	94
IDV(14)	79.25	100	0.125	0.125
IDV(15)	1.5	1.875	2.125	1.25

Three typical faults, i.e. fault 4, 7 and 13 are used to explain the detection results. Fault 13 is reaction kinetics. When this fault occurs, it will affect the performance index. And the type of this fault is slow drifting. From the Fig. 6. (b), we can know that the fault occurs after 160 sampling points. For this kind of fault, it is very necessary to continue alarming. And the regression of the performance index shocks in the vicinity of the control limit. Frequent alarm switching can cause the operators to work heavily and it is not conducive to the improvement of production efficiency. From the Fig. 6. (c), through the construction of S^2 statistic, it is possible to realize the continuous alarm to the fault. The fault detection rate can reach 94%. And it is higher than the traditional PCR fault detection method. In order to maintain the comparability of the algorithm, the number of selected variables is the same as the latent variables of PCR.

The type of fault 7 is step fault. And it is the C header pressure loss-reduced availability (stream 4). From the performance index, we can know that there is a fault at the 160th sample point and the fault disappears near the 400th sample point. This kind of fault is the fault adjustment. When the fault occurred, the distribution of dataset changes. After a period of time, the fault disappears according to the self-regulation. For the traditional fault detection method, this kind of fault will be alarmed continuously. In Fig. 7. (b), regression of the performance index and the original dataset is very consistent. This method selects the related variables to make up the $\boldsymbol{X}_{rel}$. Moreover, the S^2 statistic is constructed according to $\boldsymbol{X}_{rel}$. Compared with Fig. 7. (a) and (c) can show a higher fault detection rate. In the actual process industry, this type of fault is widespread.

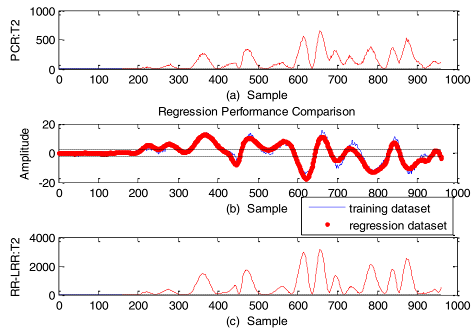

Fig. 6. Monitoring results of Fault 13.

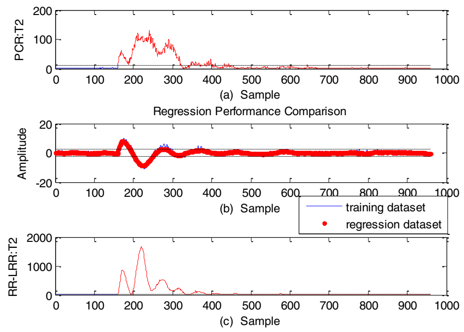

Fig. 7. Monitoring results of Fault 7.

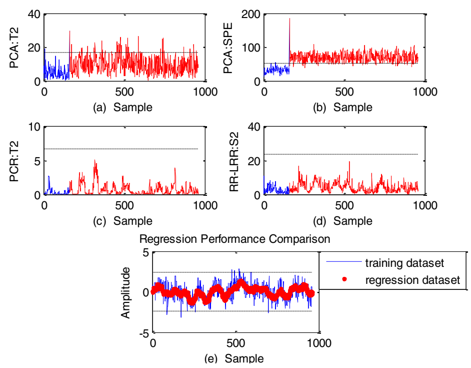

Fig. 8. Monitoring results of Fault 4.

Fault 4 is a step change in reactor cooling water inlet temperature. When the fault occurs, the reactor temperature will suddenly rise. The system compensates for it by closed-loop control. Finally, the performance index does not change.

This type of fault is defined as a failure associated with performance index. The traditional PCA method has higher detection rate for this kind of fault. So the methods like PCA are not suitable for detection this kind of fault. Form Fig. 8. (d) and (e), the proposed method can solve this kind of problem very well.

5. CONCLUSION

In this paper, a performance-triggered process monitoring scheme based ridge regression and LRR is proposed. Different from latent variable selection method, this method divides the original variables directly. i.e. $\boldsymbol{X}_{rel}$ and $\boldsymbol{X}_{ind}$. In this way, the effect of high amplitude failures on the algorithm can be avoided. Then, $\boldsymbol{X}_{rel}$ is used to regress the performance index. In order to maintain the global structure of the dataset, a new dictionary learning method is proposed. Combined with LRR, the low rank coefficient matrix of the dataset is obtained. Similar to the traditional statistic, the S^2-statistic is constructed. A new monitoring scheme is proposed to solve relationship between y_{new} and S^2. Finally, a numerical example is used for related variables selection and the effectiveness of the proposed method is demonstrated by the TE process. However, the related variables can be selected by other methods that contain more information about the relationship between variables and performance index, which can be studied in depth.

REFERENCES

Downs, J.J. and Vogel, E.F. (1993) A plant-wide industrial process control problem. *Computers & Chemical Engineering,* 17(3), 245-255.

Exterkate, P. (2013) Model selection in kernel ridge regression[J]. *Creates Research Papers,* 68(4), 1-16.

Ge, Z.Q., Song, Z.H., and Gao, F.R. (2013). Review of Recent Research on Data-Based Process Monitoring. *Industrial & Engineering Chemistry Research,* 52(10), 3543-3562.

Huang, J.P., and Yan, X.F. (2017). Quality relevant and independent two block monitoring based on mutual information and KPCA. *IEEE Transactions on Industrial Electronics,* 64, 6518-6527.

Jiang, Q.C., Yan, X.F., and Huang, B. (2015). Performance-Driven Distributed PCA Process Monitoring Based on Fault-Relevant Variable Selection and Bayesian Inference. *IEEE Transactions on Industrial Electronics,* 63(1), 377-386.

Kawano, S., Fujisawa, H., Takada, T., and Shiroishi, T. (2015). Sparse principal component regression with adaptive loading. *Computational Statistics & Data Analysis,* 89(C), 192-203.

Li, J., Duan, C., and Fei, Z. (2016) A Novel Variable Selection Approach for Redundant Information Elimination Purpose of Process Control. *IEEE Transactions on Industrial Electronics,* 63(3), 1737-1744.

Lyman, P.R. and Georgakis, C. (1995) Plant-wide control of the Tennessee Eastman problem. *Computers & Chemical Engineering,* 19(3), 321-331.

Madakyaru, M., Nounou, M.N., and Nounou, H.N. (2013). Nonlinear latent variable regression. *IEEE,* 81-88.

Macgregor, J.F., Jaeckle, C., Kiparissides, C., and Koutoudi. M. (1994). Process monitoring and diagnosis by multiblock PLS methods. *Aiche Journal,* 40(5), 826–838.

Peng, X., Tang, Y., Du, W.L., and Qian, F. (2017). Multimode Process Monitoring and Fault Detection: A Sparse Modeling and Dictionary Learning Method. *IEEE Transactions on Industrial Electronics,* 64(6), 4866-4875.

Pan, Y., Yang, C., An, R., and Sun, Y. (2016) Fault detection with improved principal component pursuit method. *Chemometrics & Intelligent Laboratory Systems,* 157(1), 111-119.

Qin, S. J. (2012) Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control,* 36(2), 220-234.

Sedghi, S., Sadeghian, A., and Huang, B. (2017) Mixture Semisupervised Probabilistic Principal Component Regression Model with Missing Inputs. *Computers & Chemical Engineering,* 103, 176-187.

Song, B., Tan, S., and Shi, H.B. (2016). Key principal components with recursive local outlier factor for multimode chemical process monitoring. *Journal of Process Control,* 47, 136-149.

Song, B., Shi, H.B., Ma, Y.X. and Wang, J.P. (2014). Multisubspace Principal Component Analysis with Local Outlier Factor for Multimode Process Monitoring. *Industrial & Engineering Chemistry Research,* 53(42), 16453-16464.

Wang, X., Liang, Y., Wang, Q., and Zhang, Z. (2017) Empirical models for tool forces prediction of drag-typed picks based on principal component regression and ridge regression methods. *Tunnelling & Underground Space Technology,* 62, 75-95.

Wang, F., Tan, S., Yang, Y.W. and Shi, H.B. (2016). Hidden Markov Model-based Fault Detection Approach for Multimode Process. *Industrial & Engineering Chemistry Research,* 55(16), 4613-4621.

Wang, G., Luo, H., and Peng, K. (2016) Quality-related fault detection using linear and nonlinear principal component regression. *Journal of the Franklin Institute,* 353(10), 2159-2177.

Yang, J., Zhang, M.S., Shi, H.B. and Tan, S. (2017) Dynamic learning on the manifold with constrained time information and its application for dynamic process monitoring. *Chemometrics & Intelligent Laboratory Systems,* 167, 179-189.

Zhou, H.T., Chen, J., Dong, G.M., and Wang, R. (2015) Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model. *Mechanical Systems & Signal Processing,* 72–73, 65-79.

Zhou, D.H., Li, G. and Qin, S.J. (2010) Total projection to latent structures for process monitoring. *Aiche Journal,* 56(1), 168–178.