

Research Article

A Combined Markov Chain Model and Generalized Projection Nonnegative Matrix Factorization Approach for Fault Diagnosis

Niu Yuguang,¹ Wang Shilin,² and Du Ming²

¹State Key Laboratory for Alternate Electric Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, China

²School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

Correspondence should be addressed to Wang Shilin; wslncepu@163.com

Received 9 March 2017; Revised 9 May 2017; Accepted 22 May 2017; Published 13 June 2017

Academic Editor: Honglei Xu

Copyright © 2017 Niu Yuguang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The presence of sets of incomplete measurements is a significant issue in the real-world application of multivariate statistical process monitoring models for industrial process fault detection. Since the missing data in the incomplete measurements are usually correlated with some of the available variables, these measurements can be used if an efficient algorithm is presented. To resolve the problem, a novel method combining Markov chain model and generalized projection nonnegative matrix factorization (MCM-GPNMF) is proposed to detect and diagnose the faults in industrial process. The basic idea of the approach is to use MCM-GPNMF to extract the dominant variables from incomplete process data and to combine them with statistical process monitoring techniques. T_G^2 and SPE_G statistics are defined as online monitoring quantities for fault detection and corresponding contribution plots are also considered for fault isolation. The proposed method is applied to a 1000 MW unit boiler process. The simulation results clearly illustrate the feasibility of the proposed method.

1. Introduction

As industrial process becomes more and more complex, process monitoring and diagnosis techniques are gaining importance for plant safety, maintenance cost, and profit margins. Multivariate statistical process monitoring (MSPM) techniques have been widely used to build statistical models for some unmeasured variables and for establishing online monitoring schemes for industrial process [1, 2]. These models extract a small number of latent variables, which in a manner better summarize the properties contained in the original variables. Monitoring and diagnosis using these latent variables are both simpler and more powerful than using the original variables [3].

It is well known that the traditional MSPM models usually require that the process data on all variables must be complete. In practice, however, one of the challenges in applying these models is to deal with the process data sets that contain some missing observations. Sometimes, more than 50% of industrial process data would contain the missing data [4]. Since the missing data in the incomplete measurements

is usually correlated with some of the available variables, the conventional MSPM methods generally eliminate the sample data in the data matrix that contain them, but doing so will leave the corresponding nature of process unknown.

It is of great importance to determine how to use the incomplete process data sets to build the normal operating model. The well-known Markov chain model (MCM), a typical stochastic process model, is one of commonly used prediction approaches. The most basic feature of MCM is “Markov property,” also known as “no aftereffect.” The next process variable state is predicted by the transition probability matrix obtained using MCM to predict the mobility of measurements if the original time series satisfies the conditions of the Markov chain (MC) [5, 6]. The MC prediction method has been widely used in various fields. In Jeon and Lee’s research, MC-based prediction routing methods are proposed to select the optimal behavior nodes [7]. In order to keep balance of electric power system, Yoder et al. use the MCM to predict short-term wind power [8].

Nonnegative matrix factorization (NMF) is a novel multivariate data analysis and dimension reduction technique

that has many applications in spectroscopy, data mining, and pattern recognition [9]. NMF and its variant methods are typically applied to high-dimensional data where each element has a non-negative value, and they find a low rank approximation from the historical process data sets. Unlike the traditional MSPM methods, the NMF-based algorithms do not have any assumption about the nature of the process variables except for nonnegativity. The nonnegativity restriction lets only additions in the factorization process. This property makes NMF obtain sparse and part-based subspace representations of the original data sets [10, 11]. Therefore, NMF-based methods have potential superiority to solve the monitoring and diagnosis problem of complex industrial process.

Normally, the NMF-based algorithms require that the measurement data be nonnegative. In practice, however, the process data of industrial process may be not fulfilling this constraint. Due to the difference in unit selection, the collected data is likely to contain negative numbers. Although it is possible by adjusting the unit to make negative data satisfy the nonnegative constraints, in order to make NMF method have a wider range of applications, we want to relax the nonnegative constraints on the original data sets. In this research, we will propose a new variant of NMF to solve the above problem. It can be called generalized projection nonnegative matrix factorization (GPNMF). Then, MCM and GPNMF are combined to extract useful information from incomplete process data and to combine them with statistical process monitoring techniques.

The rest of the article is organized as follows: Section 2 introduces MCM-based prediction method. Section 3 proposes the GPNMF method. In Section 4, the MCM-GPNMF-based process monitoring method is introduced. In Section 5, a case study in a 1000 MW unit boiler process is shown and discussed. Section 6 contains the conclusions.

2. MCM-Based Prediction Method

2.1. Markov Chain Model. Markov chain model is a stochastic process which can be used to estimate the transition probability between the discrete states of the system, and it can be expressed by some parameters. In the first-order Markov chain model, the state at a certain moment only depends on the state of its previous moment [12]. In the further research, the second-order or even higher-order Markov chain process was proposed [12, 13]. In these second-order or higher-order Markov chain processes, the state of a moment depends on two or more states before it. In this paper, we will use the second-order Markov chain model to estimate the missing values in the data set.

For a second-order Markov chain model, let X_t be a stochastic process. The set of all possible states in the stochastic process is called state space. The state space is defined as $S = \{1, 2, \dots, M\}$.

For a given time series $t_1 < t_2 < \dots < t_{n-1} < t_n$, its conditional probability distribution can be expressed as

$$\begin{aligned} P \{X_{t_{n+1}} = i_{n+1} \mid X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i_n\} \\ = P \{X_{t_{n+1}} = i_{n+1} \mid X_{t_{n-1}} = i_{n-1}, X_{t_n} = i_n\}. \end{aligned} \quad (1)$$

Markov transition probability $p_{i,j,k}$ is defined as follows:

$$p_{i,j,k} = P \{X_{t+1} = k \mid X_{t-1} = i, X_t = j\}, \quad (2)$$

where j represents the current state of the moment and i and k represent the previous and next states of i , respectively.

If a stochastic process contains M states, the matrix $P \in \mathbb{R}^{M \times M \times M}$ consisting of all transition probabilities is called the state transition matrix of the second-order Markov chain. The state transition matrix P can be expressed as follows:

$$P = \begin{bmatrix} p_{1,1,1} & p_{1,1,2} & \cdots & p_{1,1,m} \\ p_{1,2,1} & p_{1,2,2} & \cdots & p_{1,2,m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{1,m,1} & p_{1,m,2} & \cdots & p_{1,m,m} \\ p_{2,1,1} & p_{2,1,2} & \cdots & p_{2,1,m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{2,m,1} & p_{2,m,2} & \cdots & p_{2,m,m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m,m,1} & p_{m,m,2} & \cdots & p_{m,m,m} \end{bmatrix}. \quad (3)$$

The second-order Markovian transfer matrix has the following property. For any state i , the sum of the probability of its transition to all states is 1. Consider

$$\sum_{k=1}^m p_{i,j,k} = 1 \quad \forall i, j. \quad (4)$$

The maximum likelihood estimation of transfer probability $p_{i,j,k}$ for second-order Markov chains is defined as follows:

$$p_{i,j,k} = \frac{n_{i,j,k}}{\sum_{k=1}^m n_{i,j,k}}, \quad (5)$$

where $n_{i,j,k}$ is the transition frequency that the state is transferred from the previous moment state i and the current time state j to the next moment state k .

The cumulative distribution function (CDF) P_{CDF} is also used in the process of generating the time series. The CDF of the second-order Markov chain model is calculated as follows:

$$P_{\text{CDF}i,j,k} = \sum_{l=1}^k p_{i,j,l}. \quad (6)$$

The second-order Markov chain model is used to generate the process variable time series, where the initial state is completely random. Generate a random number ε between 0 and 1 by the random number generator. For the second-order Markov chain model, the current time state j and the previous time state i are known. If ε satisfies $P_{\text{CDF}i,j,k-1} < \varepsilon < P_{\text{CDF}i,j,k}$, then the next time state is k .

Only a time series of state is obtained by the above steps. Next, it is necessary to transform the resulting time series into

the time series of actual process variable. In addition to the fact that the first and last states do not need to be transformed, all other states need to use formula (7) to transform.

$$P = P_l + Y_i \times (P_u - P_l), \quad (7)$$

where P_u and P_l represent the upper and lower limits of a state, respectively. Y_i represents a random number evenly distributed between 0 and 1. P represents the actual value of the process variable.

2.2. Markov Property Test. When using the Markov chain to predict the value of a variable, it is not necessary to consider the change of the variable in the past. If the current state of the variable is known, the state of the next moment can be predicted. The Markov property test of the original time series should be carried out before the establishment of the Markov chain model. The specific method of Markov property test is given as follows: we assume that all states of the original time series are m . The marginal probability $p_{\cdot,j}$ is defined as follows:

$$p_{\cdot,j} = \frac{\sum_{i=1}^m n_{i,j}}{\sum_{i=1}^m \sum_{j=1}^m n_{i,j}}, \quad (8)$$

where $n_{i,j}$ stands for the frequency that the variable transitions from state i to state j .

When the amount of data is large enough, the statistic ξ obeys the χ^2 distribution and the degree of freedom of χ^2 distribution is $(m - 1)^2$. The statistic ξ can be calculated as follows:

$$\xi = 2 \sum_{i=1}^m \sum_{j=1}^m n_{i,j} \left| \log \frac{p_{i,j}}{p_{g,j}} \right|, \quad (9)$$

where $p_{i,j}$ stands for the probability that the variable transitions from state i to state j .

When the significance level α is given, the value of $\chi_{\alpha}^2((m-1)^2)$ can be obtained by looking up the table. The Markov chain model can be used to process the variable if $\xi > \chi_{\alpha}^2((m-1)^2)$ is satisfied.

2.3. Numerical Example. In this section, the active power of a 1000 MW unit is chosen for numerical experiment. Under the influence of random noise, the power measurement is randomly fluctuating near the set point. Therefore, the active power can be regarded as a stochastic process. The normal operation data of the active power is used as experiment data vector with 500 samples. The maximum and minimum values of power in all samples are 743.86 MW and 749.94 MW, respectively. Therefore, the stochastic process is divided into 16 states. In these states, the power values 743 MW and 750 MW are defined as state 1 and state 16, respectively. The other states are evenly distributed between state 1 and state 16, and the power interval between each two states is 500 kW. The significance level is given by 0.05 in the present work. The conclusion can be easily obtained through the calculation; that is, $\xi \gg \chi_{0.05}^2(225)$. Therefore, The Markov chain model can be used to process this time series.

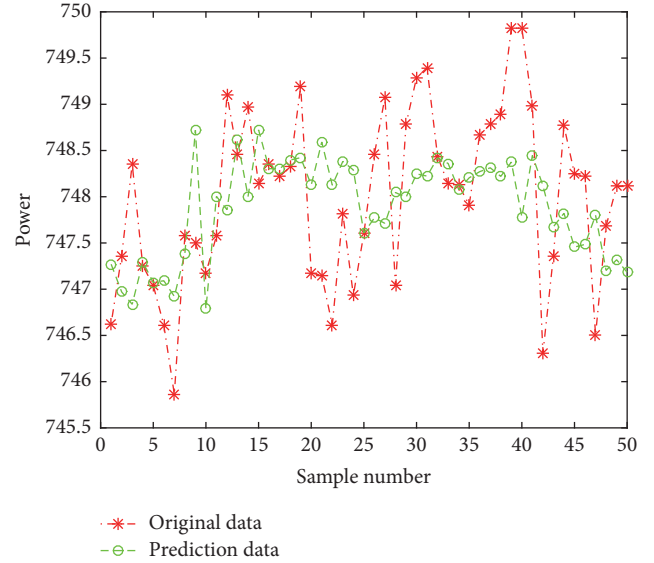


FIGURE 1: Comparison of predicted and real values.

Next, we will select 50 samples from the experimental data randomly and their values will be set to zero. These 50 samples represent the missing data in experiment data vector. Then, the second-order Markov chain model is used to deal with the missing data in experiment data vector. The incomplete data can be replaced by the predicted value, which is shown in Figure 1.

As seen in Figure 1, the predicted data are very close to the corresponding original data. In other words, the prediction value by second-order MCM is very effective. Therefore, this example can illustrate the feasibility of the second-order MCM algorithm.

3. Generalized Projection Nonnegative Matrix Factorization

3.1. NMF Algorithm. The mathematical formulation of NMF is expressed as follows. Given a process data matrix $X = [x_1, x_2, \dots, x_n] \in R^{m \times n}$ (each column of X is a sample vector), where all elements are nonnegative and a natural number $k < \{m, n\}$, NMF aims to find two low rank nonnegative matrices $W \in R^{m \times k}$ and $H \in R^{k \times n}$ such that $X \approx WH$. Here, each column of W is called basis vector. Each column of H is called an encoding and is in one-to-one correspondence with a sample vector in X [10]. In other words, each sample vector x_i is approximated by a linear combination of the columns of W , weighted by the components of h_i [11].

The approximate factorization $X \approx WH$ problem can be formulated as an optimization problem that uses a cost function to measure the quality of the approximation. Lee and Seung constructed a simple objective function which utilizes the square of the Euclidean to measure the distance between X and WH [11]. The objective function of the optimization problem can be expressed as follows:

$$\begin{aligned} \min_{W, H} \quad & F = E(X \| WH) = \frac{1}{2} \|X - WH\|^2 \\ \text{s.t.} \quad & W \geq 0, H \geq 0. \end{aligned} \quad (10)$$

It is well known that the function $\|X - WH\|^2$ is convex in W only or H only but is not convex in both variables meanwhile. Therefore, it is realistic to find local minima by solving the above optimization problem. Lee and Seung presented an iterative algorithm to obtain the local minima. The objective function $\|X - WH\|^2$ is monotonically nonincreasing under the following rules [11]:

$$\begin{aligned} H_{kj} &\leftarrow H_{kj} \frac{(W^T X)_{kj}}{(W^T WH)_{kj}} \\ W_{ik} &\leftarrow W_{ik} \frac{(XH^T)_{ik}}{(WHH^T)_{ik}}. \end{aligned} \quad (11)$$

Heretofore, the multiplicative iterative algorithm is the most classic and widely used monotone algorithm. The contradiction of the optimization algorithm between convergence speed and easy use can be better coordinated by using above iterative rules.

3.2. GPNMF Algorithm. Many improvements of NMF have been presented based on the objective function of the basic NMF algorithm by adding different regularization terms so as to increase the constraint conditions to NMF from different perspectives [14, 15]. On the contrary, Yuan and Oja proposed an improved NMF algorithm based on linear projection [16]. It can be called projective nonnegative matrix factorization (PNMF). However, the PNMf algorithm does not guarantee that the objective function is monotonically decreasing during the iteration, and there may be oscillations. In order to solve this problem and make the NMF algorithm better adapt to the actual industrial process, we propose a new variation of NMF method named generalized projection nonnegative matrix factorization (GPNMF). This method will not only guarantee the objective function monotone nonincreasing under the new iterative rule but also make the process data not necessarily constrained to be nonnegative. In this approach, the approximate factorization $X \approx WH$ will be rewritten as $X \approx XH^T H$ and the optimization problem can be expressed as follows:

$$\begin{aligned} \min_H \quad & F = E(X \| WH) = \frac{1}{2} \|X - XH^T H\|^2 \\ \text{s.t.} \quad & H \geq 0. \end{aligned} \quad (12)$$

The above constrained optimization problem can be regarded as a typical application of regularized least squares problem proposed by Tikhonov in 1963 [17]. The objective function of (12) can be expanded to a function about the variable H which ignores the constant term $X^T X$. Hence, (12) can be rewritten as follows:

$$F(H) = \frac{1}{2} \text{Tr}(-2H^T HX^T X + H^T HX^T XH^T H). \quad (13)$$

The gradient matrix of (13) can be obtained by using matrix differential:

$$\frac{\partial F(H)}{\partial H} = (-2HX^T X + 2HX^T XH^T H). \quad (14)$$

When the value of (14) is equal to 0, the solution of the least squares problem in (12) is given by

$$HX^T X = HX^T XH^T H. \quad (15)$$

Since the nonnegative constraint for the original data matrix X is relaxed in the GPNMF algorithm, now consider the case where X contains positive and negative elements. The factors X_+ and X_- are defined as the absolute values of all positive elements and negative elements in X , respectively. X_+ and X_- are calculated separately as follows:

$$\begin{aligned} X_+ &= \frac{1}{2} (|X| + X) \\ X_- &= \frac{1}{2} (|X| - X), \end{aligned} \quad (16)$$

where $|X|$ represents the absolute value for all the elements of the matrix X . Then the original data matrix X can be expressed as $X_{\pm} = X_+ - X_-$ and (15) can be adapted as follows:

$$\frac{(H[X^T X]_+ + H[X^T X]_- H^T H)_{ij}}{(H[X^T X]_- + H[X^T X]_+ H^T H)_{ij}} = 1. \quad (17)$$

It can be seen from (13) that the objective function F is a quartic function with respect to the coefficient matrix H . Here, a new iterative rule for the optimization problem in (12) is presented as follows:

$$H_{ij} = H_{ij} \sqrt[4]{\frac{(H[X^T X]_+ + H[X^T X]_- H^T H)_{ij}}{(H[X^T X]_- + H[X^T X]_+ H^T H)_{ij}}}. \quad (18)$$

The objective function in (12) is invariant under this update if and only if H is at local minima of the constrained optimization problem.

4. Fault Diagnosis Algorithm Based on MCM-GPNMF

4.1. Initialization of GPNMF Algorithm. For the moment, NMF and its improved algorithms are solved by iteration. It is well known that a good iteration initial value can improve the convergence speed and accuracy of the NMF-based algorithm. Although many researchers have pointed out the importance of a good initial value for the NMF algorithm in their article, most people still use the random method to initialize the NMF algorithm in the practical application. Since the NMF can only converge to the local optimal solution, the different initial values will result in different results. Langville et al. compared several commonly used initialization methods [18]. For the presented GPNMF algorithm, there is only one unknown factor, that is, coefficient matrix H . Besides, the approximate factorization $X \approx XH^T H$ can clearly indicate that the coefficient matrix H is calculated to satisfy $H^T H \approx I$, where I is identity matrix of size k . In this research, therefore, the coefficient matrix H is initialized with a k -by- n -dimensional matrix with ones on the diagonal and zeros elsewhere.

4.2. *Statistical Monitoring Model Based on GPNMF.* The GPNMF-based statistical process monitoring model can be described as

$$X = W\widehat{H} + E, \quad (19)$$

where the matrix $W\widehat{H}$ and E represent the eigensubspace and the residual subspace, respectively. The matrix $\widehat{H} = (W^T W)^{-1} W^T X$ is the reconstructed value of the coefficient matrix H . It reflects the changes in process variables. Due to the fact that the eigensubspace and the residual subspace of GPNMF algorithm are similar to the principal component subspace and residual subspace of PCA-based method, according to the definition of monitoring statistic T^2 and SPE in PCA method, we construct two new monitoring statistics based on GPNMF statistical monitoring model to monitor the change of eigensubspace and residual subspace. They are calculated as follows:

$$T_G^2 = \widehat{H}(i)^T \widehat{H}(i) \quad (20)$$

$$\text{SPE}_G = (x(i) - \widehat{x}(i))^T (x(i) - \widehat{x}(i)),$$

where the vector $\widehat{H}(i)$ and $x(i)$ represent the i th column of matrices \widehat{H} and X , respectively. The column vector $\widehat{x}(i)$ is the reconstructed value of $x(i)$. It can be calculated as follows:

$$\widehat{x}(i) = W\widehat{H}(i) = WW^T x(i). \quad (21)$$

For the two new monitoring statistics proposed in this section, it is obviously unreasonable to assume that these two metrics are subject to a particular distribution and calculate their upper control limits like traditional monitoring algorithm. Therefore, we use kernel density estimation (KDE) which is a more general method to calculate the actual distribution information of the monitoring statistics and then determine their upper control limits. The confidence interval θ is given by 99% during the calculation process.

4.3. *Contribution Plot Method.* It is well known that contribution plot method has been widely used in fault isolation research field. Once a fault is detected by the fault detection method, this method will be used to isolate the variables that are most likely to cause the failure. Contribution plot method is a commonly used fault separation method in PCA-based fault diagnosis algorithm. The monitoring statistics T^2 and SPE are often used in the contribution plot method. The corresponding contributions to the T^2 statistic and SPE statistic can be expressed using the following equations, respectively:

$$\text{Cont}_i^{T^2} = (\xi_i^T D^{1/2} x)^2 = x^T D^{1/2} \xi_i \xi_i^T D^{1/2} x \quad (22)$$

$$\text{Cont}_i^{\text{SPE}} = (\xi_i^T \widetilde{C} x)^2,$$

where $D = PA^{-1}P^T$ and $\widetilde{C} = I - PP^T$. The vector ξ_i represents the i th column of unit matrix I_m . The integer

m represents the number of process variables in PCA. The parameters $\text{Cont}_i^{T^2}$ and $\text{Cont}_i^{\text{SPE}}$ represent the contribution of each process variable to the monitoring statistics T^2 and SPE, respectively.

Alcala and Qin made a detailed analysis and summary of the contribution plot method [19]. Besides, a basic framework to construct contribution graph method was given in their research. Based on this framework, the corresponding contributions to the T_G^2 statistic and SPE_G statistic can be designed as follows:

$$\text{Cont}_j^{T_G^2} = \left(\delta_j^T (WW^T)^{1/2} x \right)^2 \quad (23)$$

$$\text{Cont}_j^{\text{SPE}_G} = \left(\delta_j^T (I - WW^T) x \right)^2,$$

where the parameters $\text{Cont}_i^{T_G^2}$ and $\text{Cont}_i^{\text{SPE}_G}$ represent the contribution of each process variable to the monitoring statistics T_G^2 and SPE_G , respectively. The vector δ_i represents the i th column of unit matrix I_m . The natural number m represents the number of process variables in GPNMF method.

5. Monitoring Performance

5.1. *Fault Detection.* In this section, the proposed method is applied to a 1000 MW unit boiler process. The boiler process for monitoring experiment contains three control systems: main steam temperature control system, main steam pressure control system, and feedwater control system. This process mainly consists of 33 consecutive process measurements variables, including 14 temperature signals, 9 pressure signals, 9 flow signals, and 1 power signal. All the 33 process variables are used for fault detection in this experiment. The normal operation history data of the boiler process is used as training data set which has 500 samples. For the fault data, the corresponding testing data set consists of 500 observations. Remarkably, the testing data set begins with normal operation data and the abnormal data is introduced from sample 51. All these process data are sampled every 3 sec.

In order to display the monitoring performance of MCM-GPNMF-based method, a bias fault in the main steam temperature A is taken into account. The magnitude of the fault is 1% of the actual measured value. Next, we consider the condition that 10% of the measurement values in testing matrix are missing randomly. The testing matrix contains 33 variables and each variable has 500 samples. In other words, 1650 values of testing matrix are missing. The monitoring result of bias fault when 10% of testing data are missing is shown in Figure 2.

Figures 2(a) and 2(d) show the monitoring result based on the GPNMF algorithm when the testing data set is complete. Figures 2(b) and 2(e) represent the detection result when the same method is used and 10% of the measurement values in testing matrix are missing. Figures 2(c) and 2(f) represent the monitoring result of MCM-GPNMF-based algorithm when 10% of the measurement values in testing matrix are missing. As shown in Figure 2, the GPNMF-based monitoring algorithm has an excellent monitoring

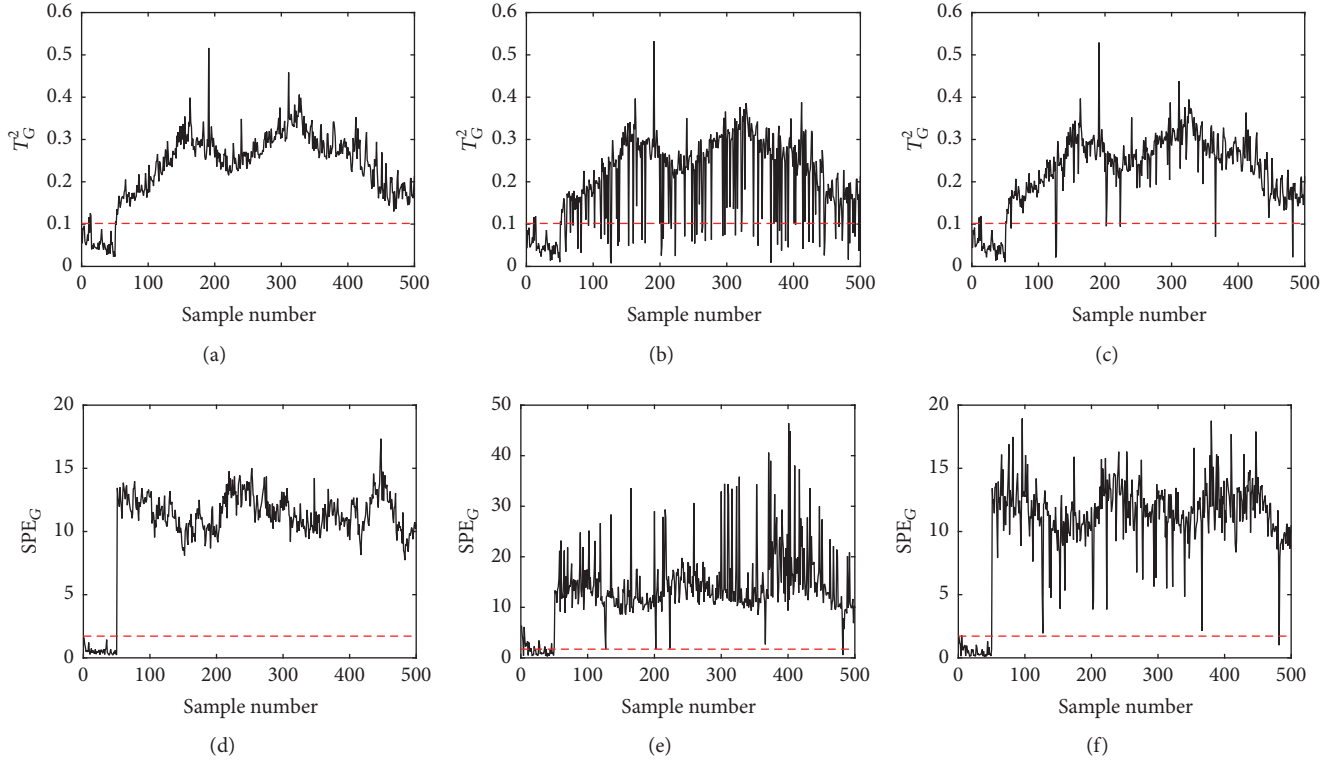


FIGURE 2: The monitoring result of bias fault.

performance, if all the process data is complete. However, the monitoring performance of the method has a serious decline when the testing data set contains missing data. When the fault occurs, the localized features of missing values will have a little change. So it is a huge challenge for T_G^2 statistic of GPNMF, which is shown in Figure 2. Many fault samples are error-detected by T_G^2 statistic. For the SPE_G statistic, although it can detect almost all of the faulty samples, many normal samples are error-detected. On the contrary, the T_G^2 statistic and SPE_G statistic of MCM-GPNMF-based method can detect almost all the faulty samples (over 98% actually). Meanwhile, the MCM-GPNMF-based method has the lower false alarm rate of SPE_G statistic than GPNMF, when the testing matrix is incomplete. However, in all three cases, it should be mentioned that several normal samples are error-detected, but the monitoring result of normal samples is still acceptable. In summary, the presented MCM-GPNMF-based monitoring method can deal with the fault detection problem, which contains incomplete data.

5.2. Fault Isolation. Based on the monitoring result of the previous section, in this part, we will use the contribution plot based on T_G^2 statistic and SPE_G statistic to identify the variable that most likely leads to the fault. Due to the fact that the accuracy of fault variable identification is closely related to the accuracy of fault detection, it is meaningful to calculate the contribution value of each variable after the system fault

is detected accurately. The most obvious change when bias fault occurs is that the value of main steam temperature A suddenly changes by the reason of a step temperature change at itself. The main steam temperature A is the 13th variable of the chosen boiler process. Both of the contribution plots corresponding to T_G^2 statistic and SPE_G statistic, shown in Figure 3, identify this variable accurately. It could illustrate the effectiveness of the presented algorithm.

6. Conclusions

In recent years, the NMF-based algorithm, which is still under development, has shown great potential in the field of industrial process fault detection. However, there are still some weaknesses for these methods in dealing with the problem of missing data. This paper proposed a new variant of NMF named generalized projection nonnegative matrix factorization which combines with the second-order Markov chain model for the situation that the process data contain missing data, which significantly improves the detection rate of industrial process. Meanwhile, two types of monitoring indices, T_G^2 statistic and SPE_G statistic, and the corresponding contribution plots were designed, respectively. As a result, the simulation in a 1000 MW unit boiler process showed that the proposed method yielded better results for fault detection and fault isolation. In other words, the monitoring method based on MCM-GPNMF proposed in this work is valuable for research and application.

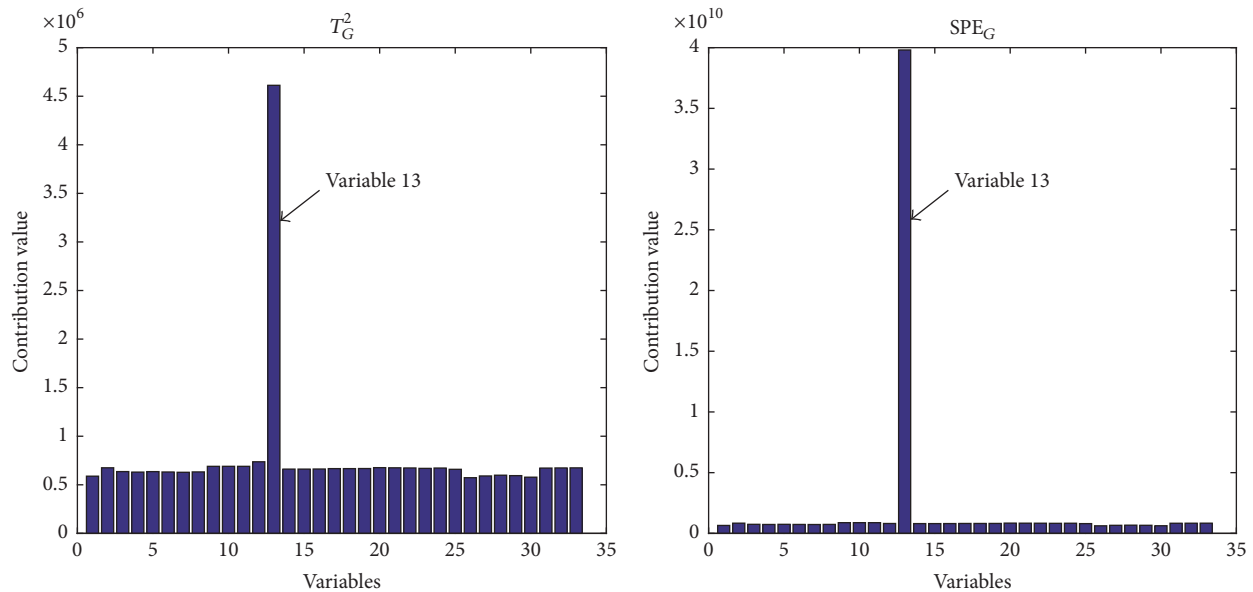


FIGURE 3: The contribution plots when bias fault occurs.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] M. Kano and Y. Nakagawa, "Data-based process monitoring, process control, and quality improvement: recent developments and applications in steel industry," *Computers & Chemical Engineering*, vol. 32, no. 1, pp. 12–24, 2008.
- [2] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annual Reviews in Control*, vol. 36, no. 2, pp. 220–234, 2012.
- [3] P. R. C. Nelson, J. F. MacGregor, and P. A. Taylor, "The impact of missing measurements on PCA and PLS prediction and monitoring applications," *Chemometrics and Intelligent Laboratory Systems*, vol. 80, no. 1, pp. 1–12, 2006.
- [4] K. Lakshminarayan, S. A. Harp, and T. Samad, "Imputation of missing data in industrial databases," *Applied Intelligence*, vol. 11, no. 3, pp. 259–275, 1999.
- [5] A. A. Nudelman, *The Markov Moment Problem And Extremal Problems: Ideas And Problems of PL Čebyšev And AA Markov And Their Further Development*, American Mathematical Society, 1977.
- [6] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bulletin of the American Mathematical Society*, vol. 73, pp. 360–363, 1967.
- [7] I.-K. Jeon and K.-W. Lee, "A dynamic Markov chain prediction model for delay-tolerant networks," *International Journal of Distributed Sensor Networks*, vol. 12, no. 9, pp. 1–7, 2016.
- [8] M. Yoder, A. S. Hering, W. C. Navidi, and K. Larson, "Short-term forecasting of categorical changes in wind power with Markov chain models," *Wind Energy*, vol. 17, no. 9, pp. 1425–1439, 2014.
- [9] Z. Li, X. Wu, and H. Peng, "Nonnegative Matrix Factorization on Orthogonal Subspace," *Pattern Recognition Letters*, vol. 31, no. 9, pp. 905–911, 2010.
- [10] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [11] D. D. Lee and H. S. Seung, *Algorithms for Non-Negative Matrix Factorization*, Advances in Neural Information Processing Systems, 2001.
- [12] G. D'Amico, F. Petroni, and F. Prattico, "First and second order semi-Markov chains for wind speed modeling," *Physica A. Statistical Mechanics and its Applications*, vol. 392, no. 5, pp. 1194–1201, 2013.
- [13] A. E. Raftery, "A model for high-order Markov chains," *Journal of the Royal Statistical Society, Series B: Methodological*, pp. 528–539, 1985.
- [14] D. Cai, X. He, J. Han, and T. S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548–1560, 2011.
- [15] H. Lee, J. Yoo, and S. Choi, "Semi-supervised nonnegative matrix factorization," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 4–7, 2010.
- [16] Z. Yuan and E. Oja, "Projective nonnegative matrix factorization for image compression and feature extraction," in *Proceedings of the Scandinavian Conference on Image Analysis*, Springer, Berlin, Germany, 2005.
- [17] A. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Mathematics Doklady*, vol. 5, 1963.
- [18] A. N. Langville et al., "Algorithms, initializations, and convergence for the nonnegative matrix factorization," <https://arxiv.org/abs/1407.7299>.
- [19] C. F. Alcalá and S. J. Qin, "Reconstruction-based contribution for process monitoring," *Automatica. A Journal of IFAC, the International Federation of Automatic Control*, vol. 45, no. 7, pp. 1593–1600, 2009.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

