2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Isolating Faulty Variables for Fault Propagation Using Bayesian Decision Theory

Jialin Liu, David Shan Hill Wong, and Ding-Sou Chen

***Abstract*—Isolating fault variables is a crucial step to provide the information that which variables are responsible for the fault for diagnosing the root causes of a process fault. In chemical processes, process faults rarely show a random behavior; on the contrary, they will be propagated to varying variables due to the actions of the process controllers. During the evolution of a fault, the task of isolating faulty variables needs to be concerned with the faulty variables decided in the previous data; in addition, the current decisions should influence the isolation results for the next sample when the fault is constantly occurring. In the presented work, an unsupervised data-driven fault isolation method was developed based on Bayesian decision theory. The proposed approach successfully located the faulty variables that were individually responsible for the simultaneous occurrence of multiple sensor faults and a process fault.**

I. Introduction

The statistical process monitoring (SPM) is a popular tool for extracting process information from vast amounts of operating data. Qin [1] reviewed the use of multivariate statistics T2 and Q, which were calculated using principal component analysis (PCA) or partial least squares (PLS) from the normal operating data, for the detection of an abnormal event. The contribution plots are a popular tool for identifying which variables are pushing the statistics out of their control limits. Several industrial applications [2]-[4] were reported in which a group of variables with the larger contributions were identified as the faulty variables. Westerhuis et al. [5] introduced the confidence limits of the contribution plots to enhance the capability of identifying the behaviors of faulty variables departing from the normal operating condition (NOC). They reported that there must be a careful interpretation of the contribution plots, since the residuals of the PCA will be smeared out over the other variables. More recently, Alcala and Qin [6] unified the forms of the confidence limits to the statistics Q, T2 and the combined index [7] according to the result of Box [8]. However, it is not always true that the corresponding statistic must be over the control limit whenever the contributions exceed the confidence limits. Therefore, the fault isolation results from using the contributions and the corresponding confidence limits may be misled, not only due to the smearing effect but also due to the definition of the confidence limits. Alcala and Qin [6] derived the reconstruction-based contribution (RBC) using the reconstruction method [9] to isolate the faulty variables, and it was reported that the RBC still suffers the smearing effect, as the contribution plots of the PCA are enduring. Therefore, the magnitude of RBCs was used to isolate the faulty variable for a single sensor fault. Kariwala et al. [10] integrated the branch and bound (BAB) method with the missing variable approach of probabilistic PCA (PPCA) to locate faulty variables. Although the BAB method is an efficient tool for solving a combinatorial optimization problem, it is still a challenge task to locate the faulty variables from a process with the massive number of variables. Liu [11] derived a contribution plot from the reduced statistics based on the reconstruction approach. Since the contributions are confined to the candidates of faulty variables, the fault cannot smear out over to the non-faulty ones. These unsupervised approaches can locate the faulty variables sample-wisely, which allows the problem of varying faulty variables to be dealt with when a process fault evolves; however, the information of the previous faulty data is not taken into consideration, which implies that these approaches suffer an efficiency problem during the search for the correct solution.

For a continuous process, the normal operating data can be assumed to follow the behavior of independent, identically distributed (i.i.d.) random variables. However, this assumption is not sustained when encountering a process fault, i.e., the information of the previous faulty variables needs to be taken into concern for isolating the current faulty variables when a fault evolves. The presented work developed a new fault isolation method based on Bayesian decision theory, which was used to incorporate the previous decisions of faulty variables. The remainder of this paper is organized as follows. Section 2 gives an overview of PCA and RBC. In addition, it addresses the situation in which the results of fault isolation may be misled by the confidence limits of the contributions. The proposed approach is detailed in section 3. In section 4, an industrial approach is utilized to demonstrate the effectiveness of the proposed approach. Finally, conclusions are given.

II. Basic Theory

A. Principal Component Analysis

Consider the data matrix $\mathbf{X} \in R^{N \times M}$ with N rows of variables and M columns of observations. Each row is normalized to zero mean and unit variance. The covariance of the reference data can be estimated as:

$$\mathbf{S} \approx \frac{1}{M-1}\mathbf{X}\mathbf{X}^{\mathrm{T}} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{\mathrm{T}} + \text{[illegible]} \tag{1}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with the first k terms of the significant eigenvalues and $\mathbf{P}$ contains the respective eigenvectors. The [illegible] and [illegible] are the residual eigenvalues and

F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (corresponding author to provide phone: 303-555-5555; fax: 303-555-5555; e-mail: author@ boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar. colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

978-3-952-41734-8/©2013 EUCA

eigenvectors respectively. The statistic Q is defined as a measure of the variations of the residual parts of data:

$$Q=\left(\mathbf{x}-\hat{\mathbf{x}}\right)^{T}\left(\mathbf{x}-\hat{\mathbf{x}}\right)=\mathbf{x}^{T}\mathbf{C}\mathbf{x} \quad (2)$$

where $\mathbf{C}\equiv\tilde{\mathbf{P}}\tilde{\mathbf{P}}^{T}$. In addition, another measure for the variations of systematic parts of the PC subspace is the statistic T^2:

$$T^{2}=\mathbf{t}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{t}=\mathbf{x}^{\mathrm{T}}\mathbf{D}\mathbf{x} \quad (3)$$

where $\mathbf{D}\equiv\mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}^{\mathrm{T}}$ and $\mathbf{t}$ are the first K term scores. Yue and Qin [7] combined statistics Q and T^2 for monitoring the variations of processes.

$$\varphi=\mathbf{x}^{\mathrm{T}}\mathbf{\Phi}\mathbf{x} \quad (4)$$

where $\mathbf{\Phi}=\mathbf{C}/\gamma_{\alpha}^{Q}+\mathbf{D}/\gamma_{\alpha}^{T}$ in which γ_{α}^{Q} and γ_{α}^{T} respectively are the (1-α) confidence limits of statistics Q and T^2. According to the results of Box [8], the control limits for statistics Q, T^2 and φ can be written as:

$$\gamma_{\alpha}^{m}=g^{m}\chi_{\alpha}^{2}\left(h^{m}\right), m=Q,T \text{ and } \varphi$$

$$g^{m}=\frac{tr\left(\mathbf{SM}\right)^{2}}{tr\left(\mathbf{SM}\right)}, h^{m}=\frac{\left[tr\left(\mathbf{SM}\right)\right]^{2}}{tr\left(\mathbf{SM}\right)^{2}}, \mathbf{M}=\mathbf{C}, \mathbf{D} \text{ and } \mathbf{\Phi} \quad (5)$$

where γ_{α}^{Q}, γ_{α}^{T} and $\gamma_{\alpha}^{\varphi}$ respectively are the (1-α) confidence limits of statistic Q, T^2 and φ, and $\chi_{\alpha}\left(h^{m}\right)$ is the (1-α) confidence level of a Chi-squared distribution with h^m degrees of freedom.

When a fault is detected by any one of above-mentioned statistics, the contribution plots provide a preliminary tool to isolate faulty variables without any prior knowledge of the fault. From (2), the contribution of the ith variable to the Q statistic can be written as:

$$c_{i}^{Q}=\left(\mathbf{x}^{\mathrm{T}}\mathbf{C}\xi_{i}\right)^{2} \quad (6)$$

where ξ_i is a column vector in which the ith element is one and the others are zero. Qin [1] derived the contribution of the ith variable to T^2 and φ as:

$$c_{i}^{T}=\left(\mathbf{x}^{\mathrm{T}}\mathbf{D}^{0.5}\xi_{i}\right)^{2} \quad (7)$$

$$c_{i}^{\varphi}=\left(\mathbf{x}^{\mathrm{T}}\mathbf{\Phi}^{0.5}\xi_{i}\right)^{2} \quad (8)$$

These confidence limits of the contributions can be derived based on the results of Box [8]. However, since the contributions of the statistics are transformed from the process variables through a matrix multiplication, the faulty variables may smear out over the other variables, which will mislead a diagnosis of the correct root causes of the faults [5], [6].

B. Reconstruction-based Contributions

Alcala and Qin [6] proposed reconstruction-based contributions (RBCs) for process monitoring. The ith RBCs for statistics Q, T^2 and φ can be written as follows.

$$RBC_{i}^{m}=\frac{1}{\xi_{i}^{\mathrm{T}}\mathbf{M}\xi_{i}}\left(\mathbf{x}^{\mathrm{T}}\mathbf{M}\xi_{i}\right)^{2} \quad (9)$$

where $m=Q,T$ and φ; $\mathbf{M}=\mathbf{C},\mathbf{D}$ and $\mathbf{\Phi}$. They applied the Cauchy–Schwarz inequality to prove that the RBC of the faulty variable is larger than those of the non-faulty ones. The control limits of the RBCs were also provided in their work. Although the faulty variables have the larger RBCs, it cannot be guaranteed that the smearing effect of the RBCs to the non-faulty variables would be under the corresponding control limits. Therefore, the smearing effect is inevitable as long as the fault isolation task is to implement control limits onto any kind of contributions. Since the control limits of the contributions were obtained from the normal operating data where any information about the smearing effect of the non-faulty variables was absent, the smearing effect and the fault behavior cannot be differentiated when a contribution was out of its control limit.

C. Bayesian Decision Theory

In this subsection, the Bayesian decision theory is briefly described. The details can be found in Bishop [12]. Consider a rule will divide the input data $\mathbf{x}$ into k regions, such that the sample in R_k is assigned to class C_k. A mistake occurs when a sample belongs to class C_k is assigned to class $C_{j\neq k}$. The probability of the mistake can be written as:

$$P\left(mistake\right)=\sum_{k=1}^{L}\int_{R_k}\sum_{j\neq k}^{L}P\left(\mathbf{x},C_{j}\right)d\mathbf{x} \quad (10)$$

where L is the number of classes and $P\left(\mathbf{x},C_{j}\right)$ is the joint probability of $\mathbf{x}$ and C_j. Therefore, the decision rule, i.e., R_k, k=1…L, should be designed by minimizing the misclassification rate or maximizing the probability of correct assignment.

$$P\left(correct\right)=\sum_{k=1}^{L}\int_{R_k}P\left(\mathbf{x},C_{k}\right)d\mathbf{x}=\sum_{k=1}^{L}\int_{R_k}P\left(C_{k}|\mathbf{x}\right)P\left(\mathbf{x}\right)d\mathbf{x} \quad (11)$$

in which $P\left(\mathbf{x}\right)$ is the probability of $\mathbf{x}$, $P\left(C_{k}|\mathbf{x}\right)$ is the posterior probability of C_k, and the product rule is used. Since the factor of $P\left(\mathbf{x}\right)$ is common to all regions, the decision should assign $\mathbf{x}$ into the class with the highest posterior probability $P\left(C_{k}|\mathbf{x}\right)$.

III. PROPOSED APPROACH

It should be pointed out that the control limits of the contributions are obtained from the normal operating data; they can be used for monitoring purposes but not for isolation. For example, the statistic T^2 normalized with the confidence limit can be expressed as follows:

$$\frac{T^{2}}{\gamma_{\alpha}^{T}}=\sum_{i=1}^{N}\left(\frac{\gamma_{\alpha,i}^{T}}{\gamma_{\alpha}^{T}}\right)\left(\frac{c_{i}^{T}}{\gamma_{\alpha,i}^{T}}\right)=\frac{\chi_{\alpha}^{2}\left(1\right)}{\chi_{\alpha}^{2}\left(K\right)}\sum_{i=1}^{N}g_{i}^{T}\left(\frac{c_{i}^{T}}{g_{i}^{T}\chi_{\alpha}^{2}\left(1\right)}\right) \quad (12)$$

where K is the number of PCs, and the confidence limits of T^2 and the contributions respective are $\gamma_{\alpha}^{T}=\chi_{\alpha}^{2}\left(K\right)$ and $\gamma_{\alpha,i}^{T}=g_{i}^{T}\chi_{\alpha}^{2}\left(1\right)$ where $g_{i}^{T}=\sum_{j=1}^{K}p_{i,j}^{2}<1$ and $p_{i,j}$ is an element of $\mathbf{P}$. In general, the number of PCs is greater than one;

therefore, $\chi^2_\alpha(1)/\chi^2_\alpha(k)$ is less than one. It is possible that the statistic T^2 will still be under the control limit, even if individual variable contributions exceed the corresponding control limits, i.e., $\frac{c_i^T}{g_i^T \chi^2_\alpha(1)} > 1$. On the other hand, when a fault is detected, using T^2, it is possible that none of the contributions will exceed the corresponding control limit. Therefore, the control limits of the contributions cannot provide any useful information about fault isolation.

Although variables with larger contributions or RBCs are likely to be responsible for the detected faults, the decision of faulty variables cannot be based on control limits derived from normal operations. In addition, a process fault will not be a random action; therefore, the faulty variables will be highly auto-correlated with the previous samples. The decision of the previous faulty sample should be taken into account for evaluating the current decision. Hence, an unsupervised approach for isolating faulty variables was developed based on the Bayesian decision theory and a likelihood function estimated by the RBCs.

When the system signals a fault in the monitoring index at time t, denoted by the event, there will be two decisions for the ith variable: in fault (F_i) or normal (N_i). Let $P(N_i)$ and $P(F_i)$ be the prior probabilities of the two classes, and let $P(\mathbf{x}_t|F_i)$ and $P(\mathbf{x}_t|N_i)$ be the conditional probabilities of observing a sample $\mathbf{x}_t$. According to Bayes' theorem, given that the overall system is in fault, the posterior probabilities for the ith variable in classes F_i and N_i can be given by:

$$P(F_i|\mathbf{x}_t) = \frac{P(\mathbf{x}_t|F_i)P(F_i)}{P(\mathbf{x}_t|F_i)P(F_i) + P(\mathbf{x}_t|N_i)P(N_i)} \quad (13)$$
$$P(N_i|\mathbf{x}_t) = 1 - P(F_i|\mathbf{x}_t)$$

The risk of misjudging the variable as faulty is therefore proportional to the identification of $P(N_i|\mathbf{x}_t)$, and the risk of misjudging the variable as normal is therefore proportional to the identification of $P(F_i|\mathbf{x}_t)$. Hence, a set of decision rules that minimize misjudgment is given by:

$$\alpha_i(\mathbf{x}_t) = \begin{cases} F_i & P(F_i|\mathbf{x}_t) \ge P(N_i|\mathbf{x}_t) \\ N_i & otherwise \end{cases} \quad (14)$$

where $\alpha_i(\mathbf{x}_t)$ is the decision for the ith variable at time t.

The crucial step in the above inference mechanism is the estimation of the conditional probability of observing $\mathbf{x}_t$, given that the ith variable is faulty. The practical rule of the larger the RBC of the ith variable, the more likely that the ith variable will be faulty was employed in the proposed approach. Sigmoidal filtering was used to ensure that variables with an RBC of less than $0.5 \times RBC_{max}(\mathbf{x}_t)$ were quickly filtered.

$$P(\mathbf{x}_t|F_i) = \frac{1}{1+e^{-z_{i,t}}}$$
$$z_{i,t} = s \times \left(RBC_i(\mathbf{x}_t)/RBC_{max}(\mathbf{x}_t) - 0.5\right) \quad (15)$$

where $RBC_i(\mathbf{x}_t)$ and $RBC_{max}(\mathbf{x}_t)$ respectively are the ith RBC of $\mathbf{x}_t$ and the maximum of the RBCs for the faulty sample, and s is a scaling factor used to adjust the confidence level of the RBCs not being due to the smearing effect.

TABLE I
THE ALGORITHM OF THE PROPOSED APPROACH

Get $\mathbf{x}_t$
Calculate $\varphi_t = \mathbf{x}^T \mathbf{\Phi} \mathbf{x}$
While $\varphi_t > \gamma^\varphi_\alpha$
  For all i
    $RBC_i^\varphi(\mathbf{x}_t) = \frac{1}{\xi_i^T \mathbf{\Phi} \xi_i}\left(\mathbf{x}_t^T \mathbf{\Phi} \xi_i\right)^2$
    $z_{i,t} = s \times \left(\frac{RBC_i^\varphi(\mathbf{x}_t)}{\max_j\{RBC_j^\varphi(\mathbf{x}_t)\}} - 0.5\right)$
    $P(\mathbf{x}_t|F_i) = \frac{1}{1+e^{-z_{i,t}}}$
  End
  For all i
    If $\varphi_{t-1} > \gamma^\varphi_\alpha$ then
      $P(F_i) = \begin{cases} P(F_i|\mathbf{x}_{t-1}) & P_{LB} \le P(F_i|\mathbf{x}_{t-1}) \le P_{UB} \\ P_{LB} & P(F_i|\mathbf{x}_{t-1}) \le P_{LB} \\ P_{UB} & P_{UB} \le P(F_i|\mathbf{x}_{t-1}) \end{cases}$
      $P(N_i) = 1 - P(F_i)$
    Else
      $P(F_i) = P(N_i) = 0.5$
    End
    $P(F_i|\mathbf{x}_t) = \frac{P(\mathbf{x}_t|F_i) \times P(F_i)}{P(\mathbf{x}_t|F_i) \times P(F_i) + \left[1 - P(\mathbf{x}_t|F_i)\right] \times P(N_i)}$
    $P(N_i|\mathbf{x}_t) = 1 - P(F_i|\mathbf{x}_t)$
    $\alpha_i(\mathbf{x}_t) = \begin{cases} F_i & P(F_i|\mathbf{x}_t) \ge P(N_i|\mathbf{x}_t) \\ N_i & otherwise \end{cases}$
  End
End

Table 1 lists the algorithm of using Bayesian decision theory to locate the faulty variables. First, the fault isolation task was only conducted when a fault was detected at time t, i.e., $\varphi_t > \gamma^\varphi_\alpha$. Then, the likelihood of each variable being faulty was obtained using (15). If the previous data was in fault, the prior probabilities were set according to the corresponding posterior probabilities of the previous sample where P_{UB} and P_{LB} are the upper and lower bounds for the posterior probability in order to avoid arriving at an entrenched prior to allow for sensor variables returning to normal by control actions or propagation of fault to other variables. In this study, the bounds were respectively set to 0.99 and 0.01. If the previous sample was in a normal operating condition, all prior probabilities of the variables being faulty were set to 0.5. This allowed the decision to be reinforced if a fault persisted and the isolation to be reset when the fault disappeared. The posterior probabilities for each variable could be obtained using (13); after that, the decision was made according to the posterior probabilities. The advantage of the proposed approach was that the decision of the current sample would be affected by the decision of the previous faulty sample; in addition, it would also influence the decision of the next faulty sample. Since the faulty variables may be different for each sampling data when a fault evolves, isolating the faulty variables should be conducted sample-wisely, and the decision of the previous faulty sample should be taken into account.

IV. Industrial Application

The compression process used a four-stage centrifugal compressor that was equipped with an intercooler between stages to cool down the compressed air, as Fig. 1 shows. A detailed description of the process can be found in the previous study [13]. The measured variables are listed in Table II. PCA was applied to the training dataset, in which 22 monitored variables were collected every five minutes for ten days, and two PCs were retained by cross-validation, which captured about 92% of the total variance. The fault detection and isolation results are shown in Fig. 2, in which the abnormal events were detected after the first day, as Fig. 2(a) shows. Fig. 2(b) reveals faulty variables as follows: firstly, x_2 was identified as the faulty variable when the fault was detected, and x_{15} was located around day 1.5. Several more faulty variables were reported around day 3.5. The normalized RBCs, which were divided by the corresponding control limits, are shown in Fig. 2(c) where the smearing effect can be observed significantly. Alcala and Qin [6] mentioned that the control limits of RBCs cannot be used to isolate faulty variables; therefore, only the magnitude of RBCs can be used to isolate the faulty variable for a single sensor fault. However, the RBC approach did not address the method to tackle the fault isolation issue, when encountering multiple sensor faults or process faults.

TABLE II
MEASURED VARIABLES FOR THE COMPRESSION PROCESS

Variable	Description
F_a (x_1)	Feed flow rate of air
$P_{in,\,i}$ (x_2-x_5)	Inlet pressure for the ith compression stage, i = 1…4
$P_{out,\,i}$ (x_6-x_9)	Outlet pressure for the ith compression stage, i = 1…4
$T_{in,\,i}$ (x_{10}-x_{13})	Inlet temperature for the ith compression stage, i = 1…4
$T_{out,\,i}$ (x_{14}-x_{17})	Outlet temperature for the ith compression stage, i = 1…4
F_c (x_{18})	Feed flow rate of cooling water
T_c (x_{19})	Inlet temperature of the cooling water
$T_{c,\,i}$ (x_{20}-x_{22})	Outlet temperature of the ith intercooler, i = 1…3

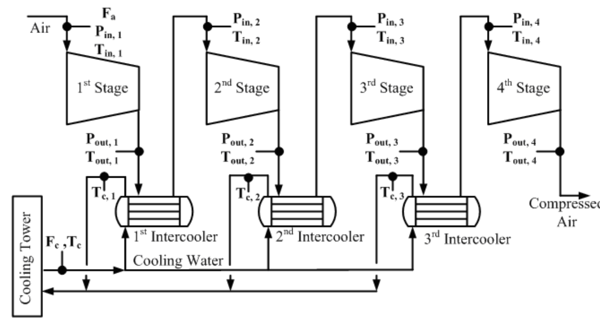


Fig. 1. Air compression process flow diagram.

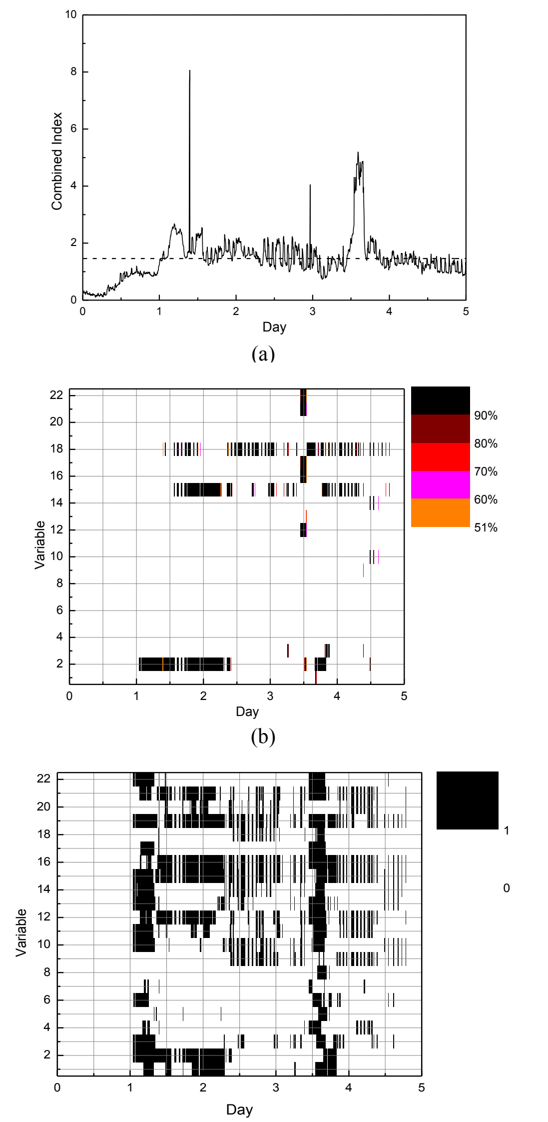


Fig. 2. Process monitoring using the combined index of PCA, (a) fault detection result, (b) posterior probabilities of Bayesian inference, (c) the normalized RBCs.

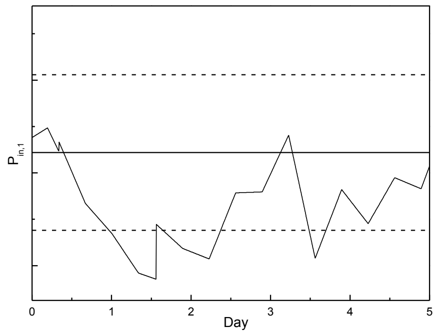
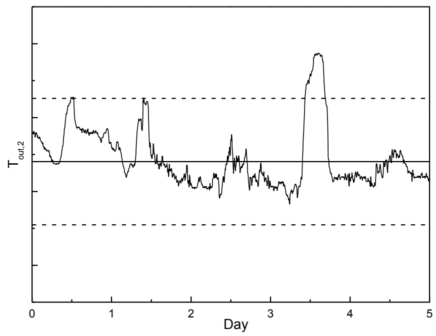
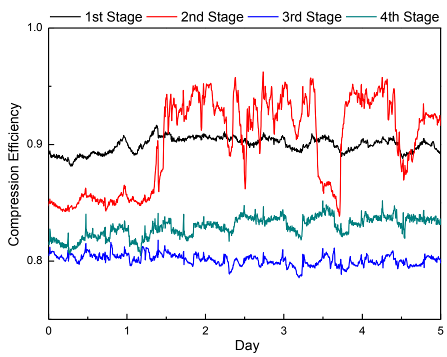


Fig. 3. Fault diagnosis for the first abnormal event.

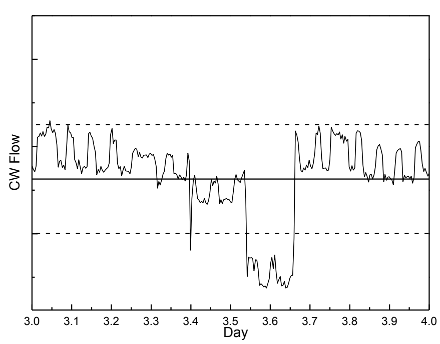


(a)

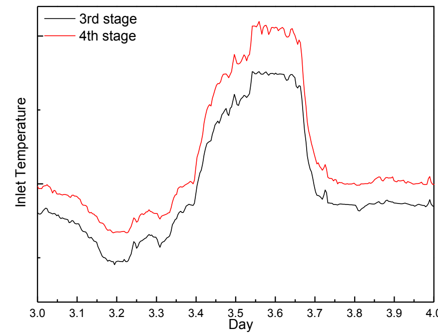


(b)

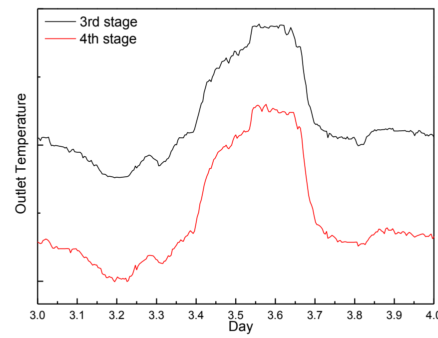
Fig. 4. Fault diagnosis for the second abnormal event, (a) comparison of the test data of x_{15} and the three sigma control limits, (b) comparison of the compression efficiencies of all stages.



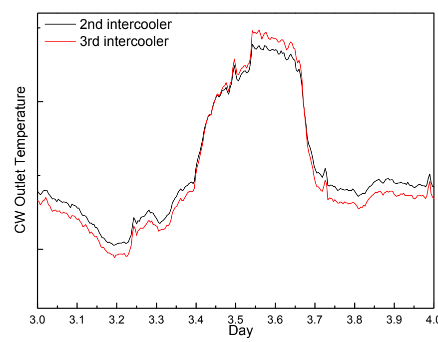
(a)



(b)



(c)



(d)

Fig. 5. Fault diagnosis for the third abnormal event, (a) cooling water flow rate, (b) inlet temperatures of the third and fourth stages, (c) outlet temperatures of the third and fourth stages, (d) outlet temperatures of the second and third intercoolers.

Fig. 3 compares the measurements of x_2 with the central line and the three sigma control limits estimated from the training dataset. The sensor drift can be observed from this, since nearly the entire test data were below the central line. The field operators requested calibration of the sensor after they were informed about this abnormality. The outlet temperature of the second compression stage, x_{15}, was responsible for the second abnormal event. A comparison of the test data with the three sigma control limits is shown in Fig. 4(a), which demonstrates the difficulty of convincing the operators that the measurements of the sensor were questionable, since the data after day 1.5 remained near the central line, except for the data around day 3.5 that will be addressed later. However, they were convinced after the comparison of the compression efficiencies for all stages was displayed, as Fig. 4(b) shows. The second stage efficiency dramatically surged around day 1.5, which coincided with that shown in Fig. 2(b), and then it fluctuated dramatically; meanwhile, the other stages' efficiencies were maintained in stable ranges. Since the compression efficiencies of all stages were highly correlated according to the operators' experiences, some sensors in the second stage (most likely x_{15}) were questionable.

Fig. 2(b) also indicates that several sensors were responsible for an abnormal event that occurred around day 3.5. These variables were x_{12}, x_{13}, x_{16}, x_{17}, x_{18}, x_{21}, and x_{22}, i.e., the inlet and outlet temperatures of the third and fourth compression stages, the cooling water flow rate and the outlet temperatures of the second and third intercoolers. Fig. 5(a) shows that the cooling water flow rate suddenly dropped at day 3.4, bounced back to a lower flow rate, and then dropped again around day 3.5. Since the intercoolers were arranged in a series, as shown in Fig. 1, the decreasing cooling water flow rate apparently did not affect the function of the first intercooler; contrarily, the performances of the second and third intercoolers deteriorated due to the insufficient amount of cooling water. Fig. 5(b) shows that the inlet temperatures of the third and fourth compression stages were higher due to the performance deterioration of the second and third intercoolers; therefore, the outlet temperatures of the third and fourth stages were higher, as shown in Fig. 5(c). Meanwhile, the outlet temperatures of the second and third intercoolers were higher, as shown in Fig. 5(d). The results showed that the proposed approach was not only capable of locating the faulty variables due to the sensor faults but was also capable of isolating the variables affected by the process fault.

V. Conclusion

In the presented work, an unsupervised fault isolation method was developed based on Bayesian decision theory. The decisions of the faulty variables for the previous faulty data were incorporated into the decision making for the current data when the fault was continuously occurring; meanwhile, the current decisions would influence the fault isolation results of the next faulty sample. In addition, the smearing effect of both the traditional contributions and the RBCs was also clarified in this study. It was proven that the two types of contributions suffered the exact same smearing effect when the corresponding confidence limits were respectively implemented onto the contributions. Since the confidence limits of the contributions were derived based on normal operating data, it could not be guaranteed that a variable was faulty when the corresponding contribution exceeded its control limit. Therefore, the control limits of the contributions could not be the only item relied upon when applying the contribution analysis for isolating faulty variables.

References

[1] S. J. Qin, "Statistical process monitoring: basics and beyond," *J. Chemom*., vol. 17, pp. 480-502, 2003.

[2] J. F. MacGregor and C. Jaeckle, "Process monitoring and diagnosis by multiblock PLS methods," *AIChE J*., vol. 40, pp. 826-838, 1994.

[3] T. Kourti and J. F. MacGregor, "Multivariate SPC methods for process and product monitoring," *J. Qual. Technol.*, vol. 28, pp. 409-428, 1996.

[4] T. Kourti, J. Lee, and J. F. MacGregor, "Experiences with industrial applications of projection methods for multivariate statistical process control," *Comput. Chem. Eng.,* vol. 20, pp. S745-S750, 1996.

[5] J. A. Westerhuis, S. P. Gurden SP, and A. K. Smilde, "Generalized contribution plots in multivariate statistical process monitoring," *Chemom. Intell. Lab. Syst.,* vol. 51, pp. 95-114, 2000.

[6] C. F. Alcala and S. J. Qin, "Reconstruction-based contribution for process monitoring," *Automatica*, vol. 45, pp. 1593-1600, 2009.

[7] H. H. Yue and S. J. Qin, "Reconstruction-based fault identification using a combined index," *Ind. Eng. Chem. Res.,* vol. 40, pp. 4403-4414, 2001.

[8] G. E. P. Box, "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 Mathematics and Statistics*, vol. 25, pp. 290-302, 1954.

[9] R. Dunia and S. J. Qin, "Subspace approach to multidimensional fault identification and reconstruction," *AIChE J.,* vol. 44, pp. 1813-1831, 1998.

[10] V. Kariwala, P. E. Odiowei, Y. Cao, T. Chen, "A branch and bound method for isolation of faulty variables through missing variable analysis," *J. Proc. Cont.,* vol. 20, pp. 1198-1206, 2010.

[11] J. Liu, "Fault diagnosis using contribution plots without smearing effect on non-faulty variables," *J. Proc. Cont.,* vol. 22, pp. 1609-1623, 2012.

[12] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[13] J. Liu and D. S. Chen, "Fault detection and identification using modified Bayesian classification on PCA subspace," *Ind. Eng. Chem. Res.,* vol. 48, pp. 3059-3077, 2009.