

## Research Article

# A Bayesian Approach to Control Loop Performance Diagnosis Incorporating Background Knowledge of Response Information

**Sun Zhou and Yiming Wang**

*Department of Automation, Xiamen University, Xiamen 361005, China*

Correspondence should be addressed to Sun Zhou; [zhousun@xmu.edu.cn](mailto:zhousun@xmu.edu.cn)

Received 20 June 2017; Accepted 3 August 2017; Published 28 September 2017

Academic Editor: Chunhui Zhao

Copyright © 2017 Sun Zhou and Yiming Wang. 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.

To isolate the problem source degrading the control loop performance, this work focuses on how to incorporate background knowledge into Bayesian inference. In an effort to reduce dependence on the amount of historical data available, we consider a general kind of background knowledge which appears in many applications. The knowledge, known as response information, is about what faults can possibly affect each of the monitors. We show how this knowledge can be translated to constraints on the underlying probability distributions and introduced in the Bayesian diagnosis. In this way, the dimensionality of the observation space is reduced and thus the diagnosis can be more reliable. Furthermore, for the judgments to be consistent, the set of posterior probabilities of each possible abnormality that are computed from different observation subspaces is synthesized to obtain the partially ordered posteriors. The eigenvalue formulation is used on the pairwise comparison matrix. The proposed approach is applied to a diagnosis problem on an oil sand solids handling system, where it is shown how the combination of background knowledge and data enhances the control performance diagnosis even when the abnormality data are sparse in the historical database.

## 1. Introduction

Fault diagnosis is a topic of practical significance in process industries. In complex control loop systems, the control performance could be degraded due to various reasons [1]. Sensors and instrumentation problems are usually revealed as the systematic errors in mass or energy balance equations [2]. Poor control loop performance might also be due to changes in the model or modeling error [3].

One challenge of control loop diagnosis is that some similar evidences could be shared among different faults [4]. In a complex industrial control loop system, there may be lots of observations. For example, a large-scale industrial process includes thousands of process measurements [5]. Many diagnostic algorithms are often designed to identify specific components, while the faults may propagate and influence other components which are not to be detected [6]; thus, the methods could be influenced by possible faults in other components [7]. Moreover, all processes may run subject to uncertainty due to missing information, noise, and

so on. Therefore, the occurrence of one fault may lead to flood of abnormal measurements and alarms and make it difficult to distinguish true underlying source.

Fault diagnosis methods for control loop systems may be classified into three categories, qualitative model-based methods, quantitative model-based methods, and data-driven methods [8–13]. However, these methods result in problems when multiple abnormalities have the same influence on the measurements. To deal with these problems, the methods were extended based on qualitative information about signs, magnitudes, and so on, to consider the direction and the magnitude of change [14, 15]. Also, Bayesian method has been proposed; for example, a Bayesian network was constructed from expert knowledge [16]. However, in these methods, the models are assumed to be known. Besides, fuzzy logic methods were proposed [17]. All these previous works rely on prior knowledge only.

To overcome the drawbacks of both quantitative and qualitative model-based diagnosis approaches, data-driven methods are developed [18]. For example, the support

vector machine (SVM) methods were proposed that take the diagnosis problem as a classification one [19]. Besides, multivariate statistical process monitoring methods were suggested [20]. Nevertheless, the problem sources of abnormality may not be explicitly identified by means of the variable contribution methods [21].

Then, a systematic probabilistic approach based on Bayesian inference was proposed that considers all possible abnormal observations. A Bayesian framework for control loop performance diagnosis was developed in [4]. The measurements from plenty of monitors are synthesized to generate a probabilistic result to diagnose the fault. Pernestal [22] proposed using Bayesian approaches to isolate faults in diesel engines. In a similar way, a data-based Bayesian approach is suggested in [23], to diagnose underlying sources of control performance degradation.

The main disadvantage with the data-driven or statistical approaches in diagnosis is that their performance relies heavily on the amount of available historical data. However, the requirement of sufficient training data is hardly met in diagnosis applications since faults are rare in normal processes. On the one hand, in their general form, they require sufficient historical samples from all faulty cases. However, in practical applications, there is only a limited amount of data available. On the other hand, the large number of monitors is a principle challenge for Bayesian diagnosis to be applied in industries. It is required in Bayesian inference to estimate the joint likelihood probability density of the observations from all monitors. In previous works, it is shown that the computational effort in estimating the probabilities grows exponentially with respect to the number of monitors [24, 25]. That phenomenon is also regarded as the *curse of dimensionality*. It makes it difficult to correctly estimate the likelihood probability in more than five dimensions with practical sample sizes. These works using Bayesian methods are based on training data only, and no explicit background knowledge is integrated about the process under diagnosis.

There is also a general type of background knowledge available. In this paper, we consider incorporating the background knowledge together with the training data under the Bayesian framework in order to improve the diagnosis even if the historical data are insufficient with respect to the monitors number. Regarding the process knowledge, it is possibly known that one measurement in observation vector is equally distributed given different abnormalities. This type of knowledge is very general and can be formulated as constraints on the underlying likelihood probability distributions [22, 27]. It can express several types of process knowledge and appear in many diagnosis applications naturally.

In this paper, the background information is expressed in terms of response signature matrix (RSM). With a translation from RSM to the constraints of the marginal probability of the likelihoods, the background knowledge is explicitly taken into account in Bayesian control loop diagnosis. Moreover, we also suggest using a moving window method to consider a sequence of observation rather than a single observation in the diagnosis. In order to evaluate the proposed approach, we applied it to an oil sand solid handling system in such a case where only a few samples from abnormalities are available.

The rest of this paper is organized as follows. A description of the Bayesian control performance diagnosis problem is introduced first, and in Section 2 some terminologies are reviewed. In Section 3, the problem studied in this paper is stated formally, and in Section 4 Bayesian diagnosis evaluating multiple consecutive observations is presented. The computations of the posterior probabilities for different modes considering historical data only are presented first, and, the approach is extended to incorporate process knowledge in Section 5. In Section 6, the proposed approach is evaluated on the diagnosis problem on the oil sand solids handling system, using training and background knowledge. The conclusion is given in Section 7.

## 2. Preliminaries: Bayesian Diagnosis for Control Loop Systems

Before going into the details of the Bayesian diagnosis, some terminologies are introduced based on the definition proposed in [23].

*Component.* Assume that the process under diagnosis consists of some components which is possible to fail or not fail. In a typical control loop, the components can be sensors, actuators, controllers, process models, and so on [28]. Assume there are  $P$  components of interest. Each component may have several different states. For example, a sensor may consist of three states: unbiased, moderate biased, and severe biased.

*Mode.* A mode of the process is defined as an assignment of the states of all the components. It indicates the state of the system. For example, a mode can be as follows: {controller is normal, activator has severe stiction, process model has no error, and sensor has moderate bias}. Each mode represents the status under which the process is operating. It can be normal state or abnormal state. Assume the total number of the modes is  $J$ . If each component has  $n$  states,  $J = n^P$ . The mode can be considered as a random scalar variable described by  $M$  with values  $\{m_j\}$ ,  $j = 1, \dots, J$ .

*Observation.* There are some monitors, sensors, or add-on indirect measurements such as “ad hoc tests” conducted by engineers, model-based diagnostic tests, and monitoring algorithms that are designed to measure certain parameters. They can be represented by a general designation, *monitors*. the  $q$ th monitor is denoted by  $\pi_q$ . Assume there are  $L$  monitors; then  $q = 1, \dots, L$ . Each monitor output or measurement signal can be represented by a discrete value; for example, low, medium, and high are 3 possible values. Define that *observation* is a  $L$ -dimension vector composed of the discrete values of all monitor outputs. These outputs may be preprocessed, for example, in diagnostic tests. The observation vector  $\mathbf{X} = (X_1, \dots, X_L)$  with the domain  $\mathbb{X} = \mathbb{X}_1 \times \dots \times \mathbb{X}_L$ . Denote an assignment of the observation vector by  $\mathbf{X} = x_k$ ,  $k = 1, \dots, K$ , where  $K$  is the number of different observations. If the  $l$ th monitor output has  $K_l$  discrete values,  $K = \prod_{l=1}^L K_l$ . Each value  $x_k$  ( $k = 1, \dots, K$ ) is an  $L$ -dimensional vector, and we write  $x_k = (x_k[1], \dots, x_k[L])$  to explicitly

denote the elements. Consider the observation as a random variable.

*Training Data.* A training sample at time  $t$  consists of simultaneous values of the mode variable  $M$  and the observation vector  $X$  at that time. The value is denoted by  $d^t = (x^t, m^t)$ . All training samples collected from different modes of the system form the training dataset. A realization of training data is denoted by  $D$ . And  $D_{m_j}$  denotes the subset of training data entries where the underlying mode is  $m_j$ .

*2.1. Bayesian Diagnosis.* The Bayesian control loop diagnosis proposed in [23] is briefly reviewed in this section. Each component is possible to suffer from some abnormal operating conditions that may degrade the control performance. Also, any fault in one component may influence the monitors for the other components [4]. Consider there are certain probabilistic interconnections between problem causes and monitor outputs [4]. Bayesian inference is applied to compute the probability of mode variable  $M$  given a current observation  $X$  and the training observation data set  $D$ . The posterior probability of every operating mode can be computed based on Bayes' rule.

$$p(M | X, D) \propto p(X | M, D) p(M), \quad (1)$$

where  $p(X | M, D)$  is the likelihood probability and  $p(M)$  is the probability of mode  $M$  which is typically specified by a priori knowledge. The mode with the highest posterior probability can be determined as the underlying mode based on the *maximum a posteriori* (MAP) principle, and the related abnormality is generally regarded as the fault source.

Thus, the main issue to construct a Bayesian diagnostic system is estimating the likelihood in line with the training observation data  $D$ . Following the results of [22, 23], a Bayesian algorithm is presented for the likelihood estimation for control loop diagnosis.

### 3. Formal Problem Formulation

Consider that there is a general type of background information and multiple consecutive observations available. The task is to determine which fault(s) which has caused the measurements, given consecutive observations  $X$ , training data  $D$ , and background knowledge  $Q$  that is described as follows.

*Background Knowledge in Terms of Probability Constraints.* Background information usually comes from expert or process knowledge. It can be described as  $Q$ . It can be considered to consist of two parts of information. One specifies the prior probability for the modes, and the other defines that there are elements, representing the monitor outputs, in the observation vector which are equally distributed under different modes.

In addition, rather than considering a single observation as [23], assume that  $r$  consecutive observations are recorded and that the same fault is present during collection of these

observations. Now, the fault diagnosis problem studied in this work can be stated formally as to compute

$$p(M^{J_r} = m_j | X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r}, D, Q), \quad (2)$$

that is, to compute the probabilities that each mode  $m_j$  is present at a time instant  $J_r$ , given the training dataset  $D$ , the background knowledge  $Q$ , and the observation values  $X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r}$  from the control loop process under diagnosis. The subscripts on  $J$  are used to denote observation vectors from consecutive instants and those on  $k$  to enumerate the observation values.

In the following, the posterior probabilities of each mode given consecutive observations are calculated with the training data only.

### 4. Bayesian Inference Using Training Data Only

To solve the stated problem (2), a new method is proposed for learning the likelihood probability distribution. Before going into the details, first let us present a previous result on inference based on training data only.

According to Bayes' rule, to compute (2), the likelihood probability needs to be calculated:

$$\begin{aligned} p(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} | M^{J_r} = m_j, D) \\ = p(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} | M^{J_r} = m_j, D_{m_j}), \end{aligned} \quad (3)$$

where  $D_{m_j}$  denotes the training data under mode  $M = m_j$ .

Assume that the likelihood of all possible values of observation under mode  $m_j$  is parameterized by  $\Theta_{m_j}$ ,

$$\begin{aligned} p(x_k | m_j, \Theta, Q) = \theta_{k|m_j}, \\ \Theta_{m_j} = \{\theta_{1|m_j}, \dots, \theta_{K|m_j}\}, \end{aligned} \quad (4)$$

$$\theta_{k|m_j} > 0, \quad \sum_{k=1}^K \theta_{k|m_j} = 1.$$

Let  $\Omega_{m_j}$  denote the space of all the likelihood parameters when the mode  $M = m_j$ . Also, the prior probability of these parameters is Dirichlet distributed

$$f(\Theta_{m_j} | Q) = \frac{\Gamma(\sum_{k=1}^K \alpha_{k|m_j})}{\prod_{k=1}^K \Gamma(\alpha_{k|m_j})} \prod_{k=1}^K \theta_{k|m_j}^{\alpha_{k|m_j}-1}, \quad \alpha_{k|m_j} > 0. \quad (5)$$

It can be shown that the Dirichlet distribution is the only possible choice for  $f(\Theta_{m_j} | Q)$  under certain, not very restrictive assumptions [29]. One attractive property of Dirichlet distribution is that it is conjugate to the multinomial distribution [30], and the distribution for the training samples is proportional to the multinomial distribution. This makes the computations particularly simple. Further, the parameters  $\alpha_{1|m_j}, \dots, \alpha_{K|m_j}$  of the distribution are required.  $\Gamma(\cdot)$  is the gamma function. For real number  $x$ ,  $\Gamma(x) = (x-1)!$ .

By marginalizing over all the likelihood parameters, we have

$$\begin{aligned} & p\left(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} \mid M^{J_r} = m_j, D_{m_j}\right) \\ &= \int_{\Omega_{m_j}} p\left(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} \mid M^{J_r}\right. \\ & \quad \left.= m_j, \Theta_{m_j}, D_{m_j}\right) f\left(\Theta_{m_j} \mid m_j, D_{m_j}\right) d\theta_{m_j}. \end{aligned} \quad (6)$$

For the first factor of the integral (6), given the likelihood parameters  $\Theta_{m_j}$  and assuming that these observations from mode  $m_j$  are independent,

$$\begin{aligned} & p\left(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} \mid M^{J_r} = m_j, \Theta_{m_j}, D_{m_j}\right) \\ &= p\left(x_{k_1} \mid m_j, \Theta_{m_j}\right) \cdots p\left(x_{k_r} \mid m_j, \Theta_{m_j}\right) \\ &= \theta_{k_1|m_j} \times \cdots \times \theta_{k_r|m_j}. \end{aligned} \quad (7)$$

And for the second factor, following the derivation of [23], we can write

$$\begin{aligned} & f\left(\Theta_{m_j} \mid m_j, D_{m_j}\right) \\ &= \frac{p\left(D_{m_j} \mid \Theta_{m_j}, m_j\right) f\left(\Theta_{m_j} \mid m_j\right)}{p\left(D_{m_j} \mid m_j\right)}. \end{aligned} \quad (8)$$

Further, the likelihood of training data subset  $D_{m_j}$  related to the operating mode  $m_j$  can be calculated as

$$p\left(D_{m_j} \mid \Theta_{m_j}, m_j\right) = \prod_{k=1}^K \theta_{k|m_j}^{n_{k|m_j}}, \quad (9)$$

where  $n_{k|m_j}$  is the number of training data samples with the observation  $x_k$  from the mode  $m_j$ .

Then combining (7)–(9) and substituting in (6), likelihood (3) can be obtained.

To introduce the consecutive observations, first some notations are needed. Let  $\varepsilon_{\text{obs}} \subset \mathbb{X}$  denote the set of distinct values present in consecutive observations  $\mathbf{X}^{J:r} = (X^{J_1}, \dots, X^{J_r})$ , and let  $N_{x_k}$  be the total number of observations in  $\mathbf{X}^{J:r}$  with the value  $x_k$ . Following [31], the likelihood probability is given by the expression

$$\begin{aligned} & p\left(X^{J_1} = x_{k_1}, \dots, X^{J_r} = x_{k_r} \mid M^{J_r} = m_j, D, Q\right) \\ &= \frac{\prod_{x_k \in \varepsilon_{\text{obs}}} \prod_{m=0}^{N_{x_k}-1} \left(n_{k|m_j} + \alpha_{k|m_j} + m\right)}{\prod_{m=0}^{r-1} \left(N_{m_j} + A_{m_j} + m\right)}, \end{aligned} \quad (10)$$

where  $N_{m_j} = \sum_{k=1}^K n_{k|m_j}$  is the count of the hypothetical samples and  $A_{m_j} = \sum_{k=1}^K \alpha_{k|m_j}$  is the count of training samples. Theorem A.1 in the appendix can be referred to for the derivation of (10).

TABLE I: A priori response information.

| Monitor | $M = m_0$ (no fault) | $M = m_1$ | $M = m_2$ | $M = m_3$ | $M = m_4$ | ... |
|---------|----------------------|-----------|-----------|-----------|-----------|-----|
| $\pi_1$ | —                    | 0         | 0         | $r$       | 0         |     |
| $\pi_2$ | —                    | $r$       | 0         | 0         | $r$       |     |
| $\pi_3$ | —                    | 0         | $r$       | 0         | 0         |     |

## 5. Bayesian Diagnosis Incorporating RSM and Data

To combine the background knowledge with training data, first, the problem dimensionality needs to be reduced utilizing the probability constraints implied in the available background information, and in the dimension-reduced subspaces, estimate the likelihoods with Bayesian inference. In this way, the estimation accuracy can be improved in the case of small amount of available historical samples that is often encountered in real applications since abnormalities are rare in normal process operations. Then, from the set of posterior probabilities that might be inconsistent as they are computed from different subspaces, derive the partially ordered posteriors that are consistent in the original probability space.

*5.1. Background Knowledge Expressed as RSM.* In many applications, there are only a few historical samples available. Therefore, the process knowledge must be explicitly handled. We consider a general type of process knowledge about what abnormalities can possibly affect each of the monitors. It can be expressed in terms of the following: “observation  $X_k$  has the same but maybe unknown probability distribution under  $M = m_s$  and  $M = m_t$ .”

Table 1 gives an example of such knowledge. “ $r$ ” at the  $l$ th row and the  $j$ th column represents a response signature meaning that the  $l$ th element of the observation that is from the  $l$ th monitor is affected under abnormal mode  $m_j$ , compared with that under the normal mode. The likelihood distribution of the  $l$ th observation element given  $M = m_j$  is different from that under the normal mode  $M = m_j$ . In other words, the  $l$ th monitor output would respond when the operating mode turns into the  $j$ th abnormal operating mode. And “0” in the table indicates that the likelihood probability distribution is the same as that under the fault free mode. Or to say, the  $l$ th monitor measurement shows zero response to the  $j$ th abnormal mode.

The matrix corresponding to the response information table is the Response Signature Matrix (RSM), denoted by  $R = (r_{lj})_{L \times J}$ , and use “1” for “ $r$ ” in the matrix. For example, the RSM according to Table 1 is

$$R = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \end{pmatrix}. \quad (11)$$

**5.2. Dimensionality Reduction.** Using training data only, the Bayesian diagnosis would suffer from the curse of dimensionality. In statistics, the phrase reflects the sparsity of data in multiple dimensions. That phenomenon is an inevitable problem in Bayesian diagnosis. For instance, when a process employs 20 monitors, each one having the same three states that are low, medium, and high, the total number of possible observation values is  $3^{20} = 3.497 \times 10^9$ . This large observation space requires substantially more data to estimate.

Consider any two operating modes  $m_s$  and  $m_t$ , and the domain of discourse consists only of these two modes.

$$p(m_s) + p(m_t) = 1. \quad (12)$$

For the  $l$ th monitor  $\pi_l$ ,  $r_{ls} = r_{lt} = 0$  indicates that the marginal probability distributions of the  $l$ th monitor output under the two modes are equal; that is,

$$p(X[l] = x_k[l] | m_s, Q) = p(X[l] = x_k[l] | m_t, Q). \quad (13)$$

Therefore, the  $l$ th monitor readings can be ignored, and it is possible to reduce the dimension by one. Whereas if  $r_{ls} = 1$  or  $r_{lt} = 1$ , (13) does not hold. Thus, this measurement must be taken into account in the probability computation. Define

$$\gamma_{st} = \{l \in \{1, \dots, L\} : r_{ls} = 1 \text{ or } r_{lt} = 1\} \quad (14)$$

as a set of numbers of monitors whose readings are affected by the  $s$ th or the  $t$ th abnormality, and  $d_{st}$  is the dimension of  $\gamma_{st}$ , or to say, the number of the elements of set  $\gamma_{st}$ . Given the background knowledge in terms of response information,  $d_{st}$  is usually smaller than  $L$ . Take the response information given in Table 1 as an example, we can obtain  $\gamma_{12} = \{2\}$ ,  $\gamma_{13} = \{3\}$ ,  $\gamma_{14} = \{1\}$ ,  $\gamma_{15} = \{2\}$ ,  $\gamma_{23} = \{2, 3\}$ ,  $\gamma_{24} = \{1, 2\}$ ,  $\gamma_{25} = \{2\}$ ,  $\gamma_{34} = \{1, 3\}$ ,  $\gamma_{35} = \{2, 3\}$ , and  $\gamma_{45} = \{1, 2\}$ .

Define  $Z_r^{st} \triangleq (Z[l_1], \dots, Z[l_{d_{st}}])$ , where  $l_1, \dots, l_{d_{st}} \in \gamma_{st}$  is the observation element vector, each element of which represents a monitor output whose distribution is affected under mode  $m_s$  or  $m_t$ .  $Z_r^{st}$  is with the domain  $\mathbb{Z}_{st} = \mathbb{X}_{l_1} \times \dots \times \mathbb{X}_{l_{d_{st}}}$  that is a  $d_{st}$ -dimensional observation space, a subspace of  $\mathbb{X}$ . Also, define  $Z_{-r}^{st}$  as the observation vector whose probability is unaffected. For instance,  $Z_r^{12} = (Z[2])$  and  $Z_{-r}^{12} = (Z[1], Z[3])$ . From (13), we have

$$p(Z_{-r}^{st} | m_s, Q) = p(Z_{-r}^{st} | m_t, Q). \quad (15)$$

Combining (12), it can be obtained that

$$p(Z_{-r}^{st} | Q) p(m_j | Q) = p(Z_{-r}^{st}, m_j | Q), \quad j = s, t. \quad (16)$$

(16) indicates that when only two modes instead of all  $J$  modes are considered,  $Z_{-r}^{st}$  is independent of mode variable. Then, it is easy to prove

$$\frac{p(X | m_s, Q)}{p(X | m_t, Q)} = \frac{p(Z_r^{st} | m_s, Q)}{p(Z_r^{st} | m_t, Q)}. \quad (17)$$

Therefore, while comparing the likelihoods of the observation under two modes, the monitor outputs corresponding

to  $Z_{-r}^{st}$  can be ignored, and only those related to  $Z_r^{st}$  are needed to be taken into the probability computation. In this way, given the background information, the dimension of observation space can be reduced from  $L$  to  $d_{st}$ . Let  $K_{st}$  denote the total count of different observations in  $\mathbb{Z}_{st}$ . Note that in the following,  $Z_r^{st}$  is written as  $Z$  for simplicity.

In the  $d_{st}$ -dimensional subspace  $\mathbb{Z}_{st}$ , the likelihood can also be obtained applying Bayesian inference as follows.

First, consider the likelihood estimation given one observation. We want to compute

$$p(Z = z_k | m_j, D, Q), \quad (18)$$

$$z_k = (z_k[l_1], \dots, z_k[l_{d_{st}}]) \in \mathbb{Z}_{st}, \quad j = s, t.$$

Assume that the likelihood of all possible values of  $d_{st}$ -dimensional observation under mode  $m_j$  is parametrized by a set of parameters  $\Theta_{m_j}$ ,

$$p(z_k | m_j, \Theta, Q) = \theta_{k|m_j}, \quad (19)$$

$$\Theta_{m_j} = \{\theta_{1|m_j}, \dots, \theta_{K_{st}|m_j}\},$$

$$z_k \in \mathbb{Z}_{st}, \theta_{k|m_j} > 0, \sum_{k=1}^{K_{st}} \theta_{k|m_j} = 1, \quad j = s, t.$$

Also, the prior probability of these parameters is assumed to be Dirichlet distributed.

$$f(\Theta_{m_j} | Q) = \frac{\Gamma(\sum_{k=1}^{K_{st}} \alpha_{k|m_j})}{\prod_{k=1}^{K_{st}} \Gamma(\alpha_{k|m_j})} \prod_{k=1}^{K_{st}} \theta_{k|m_j}^{\alpha_{k|m_j} - 1}, \quad (20)$$

$$\alpha_{k|m_j} > 0, \quad z_k \in \mathbb{Z}_{st}, \quad j = s, t.$$

Then, applying Bayesian inference, the likelihood (18) can be obtained as this expression

$$p(Z = z_k | m_j, D, Q) = \frac{n_{k|m_j} + \alpha_{k|m_j}}{N_{m_j} + A_{m_j}}, \quad j = s, t, \quad (21)$$

where  $N_{m_j} = \sum_{k=1}^{K_{st}} n_{k|m_j}$  and  $A_{m_j} = \sum_{k=1}^{K_{st}} \alpha_{k|m_j}$  are the count of hypothetical samples and training samples, respectively.

Now consider we have  $r$  consecutive observations  $z_{k_1}, \dots, z_{k_r} \in \mathbb{Z}_{st}$ , define  $\omega_{\text{obs}} \subset \mathbb{Z}_{st}$  as the set of distinct values presenting in the consecutive  $d_{st}$ -dimensional observations  $Z^{1:r} = (Z^1, \dots, Z^r)$ , and  $N_{z_k}$  is the total number of observations in  $Z^{1:r}$  with value  $z_k$ . The sought likelihood can be obtained as

$$p(Z^1 = z_{k_1}, \dots, Z^r = z_{k_r} | M^r = m_j, D, Q)$$

$$= \frac{\prod_{z_k \in \omega_{\text{obs}}} \prod_{m=0}^{N_{z_k} - 1} (n_{k|m_j} + \alpha_{k|m_j} + m)}{\prod_{m=0}^{r-1} (N_{m_j} + A_{m_j} + m)}, \quad j = s, t. \quad (22)$$

Given the prior probabilities for the modes, the posterior probability can also be computed.

$$p(m_j | Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r}, D, Q) = \frac{p(Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r} | M^{I_r} = m_j, D, Q) p(m_j | Q)}{p(Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r} | D, Q)}, \quad (23)$$

$j = s, t.$

**5.3. Consistent Partially Ordered Posteriors.** For  $J$  modes, there are totally  $(1/2)J(J-1)$  pairs of modes. As discussed in the last section, for each pair, the posteriors of each of the two modes can be obtained. However, the probability space contains only those two modes, not all the  $J$  modes; thus, these pairs of computed posteriors might be inconsistent.

Construct a pairwise comparison matrix  $\mathbf{C}$ . To simplify the notation, write the posterior probability of mode  $m_j$ ; that is,  $p(m_j | Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r}, D, Q)$  as  $p_j$ . Let  $\mathbf{C}$  be the set of  $(J \times J)$  positive reciprocal matrix

$$\mathbf{C} = (c_{st})_{J \times J} = \begin{pmatrix} p_s \\ p_t \end{pmatrix}_{J \times J}, \quad (24)$$

where each entry  $c_{st}$  is the ratio of the posteriors of modes  $m_s$  and  $m_t$ . Therefore, the matrix is of the form

$$\mathbf{C} = \begin{pmatrix} 1 & \frac{p_1}{p_2} & \dots & \frac{p_1}{p_J} \\ \frac{p_2}{p_1} & 1 & \dots & \frac{p_2}{p_J} \\ \vdots & & \ddots & \\ \frac{p_J}{p_1} & \frac{p_J}{p_2} & \dots & 1 \end{pmatrix}, \quad (25)$$

where  $c_{st} = 1/c_{ts}$ ,  $c_{ss} = 1$ ,  $c_{st} > 0$ ,  $s, t \in \{1, \dots, J\}$ . This comparison matrix consists of paired reciprocal comparisons based on (17). By definition (25),  $\mathbf{C}$  is a positive reciprocal matrix.

Combining (22) and (23), we have

$$c_{st} = \frac{p_s}{p_t} = \frac{p(Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r} | M^{I_r} = m_s, D, Q) \times p(m_s | Q)}{p(Z^{I_1} = z_{k_1}, \dots, Z^{I_r} = z_{k_r} | M^{I_r} = m_t, D, Q) \times p(m_t | Q)} \quad (26)$$

or in an equivalent form

$$c_{st} = \frac{\prod_{z_k \in \omega_{\text{obs}}} \prod_{m=0}^{N_{z_k} - 1} (n_{k|m_s} + \alpha_{k|m_s} + m)}{\prod_{z_k \in \omega_{\text{obs}}} \prod_{m=0}^{N_{z_k} - 1} (n_{k|m_t} + \alpha_{k|m_t} + m)} \times \frac{\prod_{m=0}^{r-1} (N_{m_t} + A_{m_t} + m)}{\prod_{m=0}^{r-1} (N_{m_s} + A_{m_s} + m)} \times \frac{p(m_s | Q)}{p(m_t | Q)}. \quad (27)$$

What are the priorities of the modes with respect to the posterior probability? Consider the consistency of the matrix

$\mathbf{C}$ .  $\mathbf{C}$  is consistent if  $c_{ij}c_{jk} = c_{ik}$ ,  $\forall i, j, k$ . The original matrix  $\mathbf{C}$  itself may be inconsistent. In order to determine which mode has the maximum probability, we need to derive a consistent partially ordered relationship set of all the modes  $\{m_1, \dots, m_J\}$  from the paired comparisons of the posteriors (maybe inconsistent) given in  $\mathbf{C}$ .

There is a number of ways to obtain the vector of priorities. With emphasis on consistency, we suggest adopting an eigenvalue formulation [32]. Using this formulation, our problem becomes

$$\mathbf{C}\mathbf{w} = \lambda_{\max}\mathbf{w}, \quad (28)$$

where  $\lambda_{\max}$  is the principal or largest eigenvalue of  $\mathbf{C}$ . The principal eigenvector  $\mathbf{w} = (w_1, \dots, w_J)$ ;  $w_1 > \dots > w_J$  is the partially ordered vector for all  $J$  modes with respect to their posterior probabilities. It is easy to prove that, for any estimate  $\mathbf{x}$ ,

$$\lim_{k \rightarrow \infty} \frac{1}{\lambda_{\max}^k} \mathbf{C}^k \mathbf{x} = a\mathbf{w}, \quad (29)$$

where  $a > 0$  is a constant and  $\mathbf{w}$  is principal eigenvector of  $\mathbf{C}$ . The formula can be interpreted roughly as follows: "if we begin with an estimate and operate on it successively by  $\mathbf{C}/\lambda_{\max}$  to get new estimates, the result converges to a constant multiple of the principal eigenvector."

Therefore, the mode corresponding to the largest element  $w_1$  is the sought operating mode based on the MAP principle.

To sum up, following is the algorithm of the proposed diagnosis method for complex control loops that incorporates training data and background knowledge of response information.

- (a) Based on process knowledge expressed as RSM, for each pair of modes  $m_s$  and  $m_t$ , obtain  $\gamma_{st}$  according to (14).
- (b) For each pair of modes, in an observation subspace  $\mathbb{Z}_{st}$  with respect to  $\gamma_{st}$ , compute the likelihood of each possible observation  $z_k$  under  $m_s$  and  $m_t$ , respectively, with (22).
- (c) Construct the pairwise comparison matrix  $\mathbf{C}$  with (25), (26), or (27).
- (d) Compute the eigenvector  $\mathbf{w}$  using (28), and the mode corresponding to the largest element  $w_1$  is the sought operating mode.

## 6. Evaluation Results for Oil Sand Solids Handling System Diagnosis

**6.1. Diagnostic Settings.** We now consider solids handling system for evaluation. This system is the first stage of the oil sands process, which is a typical setup used for oil sands mining operations. The flowchart is presented in Figure 1 that is based on the industrial application [26]. From the flowchart, the mass of each truckload of the oil sand solid and the time when it was dumped into the dump hopper are available in a database. After being crushed, the solid is transported into the surge pile through the conveyor belt.

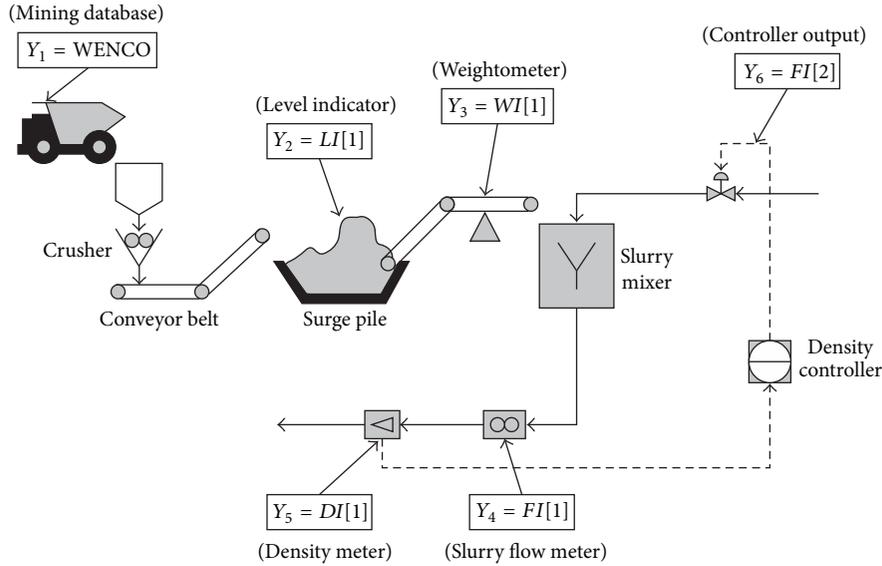


FIGURE 1: Oil sand solid slurry preparing system [26].

TABLE 2: Modes considered.

| Mode  | Mix box density model error | WENCO system instrument bias | Weightometer bias | Slurry flow meter bias | Density meter bias | Water valve stiction |
|-------|-----------------------------|------------------------------|-------------------|------------------------|--------------------|----------------------|
| $m_1$ | —                           | —                            | —                 | —                      | —                  | —                    |
| $m_2$ | *                           | —                            | —                 | —                      | —                  | —                    |
| $m_3$ | —                           | *                            | —                 | —                      | —                  | —                    |
| $m_4$ | —                           | —                            | *                 | —                      | —                  | —                    |
| $m_5$ | —                           | —                            | —                 | *                      | —                  | —                    |
| $m_6$ | —                           | —                            | —                 | —                      | *                  | —                    |
| $m_7$ | —                           | —                            | —                 | —                      | —                  | *                    |

A level indicator gives a reading, from 0% to 100%, of the relative level of the surge pile. A weightometer is on the mixer feed conveyor which feed oil sand from the surge pile to the slurry mixer. The slurry is prepared in this mixer by adding water to the oil sand. The amount of water is controlled by a slurry density controller. The controller output is the volumetric flow rate of water. A slurry flow meter and a density meter give the readings of the volumetric flow rate and the density of the effluent slurry, respectively.

In our simulation, four instruments  $Y_1$  (database),  $Y_3$  (weightometer),  $Y_4$  (slurry flow meter), and  $Y_5$  (density meter) are subject to possible bias. The control valve that is used to manipulate the water flow may suffer from stiction, and due to linearization, the model for slurry density controller is subject to error.

The system is designed to run under seven modes as shown in Table 2. The first mode  $m_1$  is the *No Fault* mode. Each of the other six modes is under a fault of one component. Mode  $m_2$  represents the density model error due to the linearization,  $m_3$ ,  $m_4$ ,  $m_5$ , and  $m_6$  consider bias in each of the four instruments  $Y_1$ ,  $Y_3$ ,  $Y_4$ , and  $Y_5$ , and  $m_7$  considers stiction of the water valve. In this table, a “—” denotes that the

corresponding component is fault free and a “\*” represents that the component has fault. Nine monitors are available for diagnosis, as shown in Table 3. 200 simulation runs were performed for each case, as well as 60 runs used for validation. As there are totally 9 monitors, the generic Bayesian diagnosis using training data only is a 9-dimensional problem. In other words, each likelihood probability given each underlying mode that is needed in the inference is a 9-dimensional joint probability. It is obvious that the available historical samples are very far from sufficient to generate accurate likelihood estimation.

The process knowledge of response information is given (Table 3). From this table, the corresponding RSM can be written that represents the implied probability constraints. According to (14), all  $\gamma_{st}$  sets for each value of  $s$  and  $t$  ( $s, t \in \{1, 2, 3, 4, 5, 6, 7\}$ ) can be obtained. For instance,  $\gamma_{23} = \{2, 4, 5, 6, 7\}$  with a reduced dimension  $d_{st} = 5$ . Then, for each pair of modes, the likelihood of each possible observation under each mode is computed, respectively, in a subspace with respect to  $\gamma_{st}$ . Finally, through the pairwise comparison matrix  $C$ , the underlying operating mode can be determined using the eigenvalue formulation.

TABLE 3: Monitors and a priori response information.

| Monitor | Description   | $m_1$ | $m_2$ | $m_3$ | $m_4$ | $m_5$ | $m_6$ | $m_7$ |
|---------|---|-------|-------|-------|-------|-------|-------|-------|
| $\pi_1$ | Slurry model validation                                   | —     | 0     | 0     | $r$   | $r$   | 0     | $r$   |
| $\pi_2$ | Density flow model validation                             | —     | $r$   | 0     | $r$   | 0     | $r$   | 0     |
| $\pi_3$ | Slurry density disturbance model validation               | —     | 0     | 0     | $r$   | $r$   | $r$   | $r$   |
| $\pi_4$ | Data reconciliation residuals for $Y_1$                   | —     | 0     | $r$   | $r$   | $r$   | 0     | 0     |
| $\pi_5$ | Data reconciliation residuals for $Y_3$                   | —     | 0     | $r$   | $r$   | $r$   | 0     | 0     |
| $\pi_6$ | Data reconciliation residuals for $Y_4$                   | —     | $r$   | 0     | $r$   | 0     | $r$   | 0     |
| $\pi_7$ | Data reconciliation residuals for $Y_5$                   | —     | 0     | $r$   | $r$   | 0     | 0     | 0     |
| $\pi_8$ | Data reconciliation residuals for the control valve input | —     | 0     | 0     | $r$   | $r$   | $r$   | $r$   |
| $\pi_9$ | Valve stiction monitor                                    | —     | 0     | 0     | 0     | 0     | $r$   | 0     |

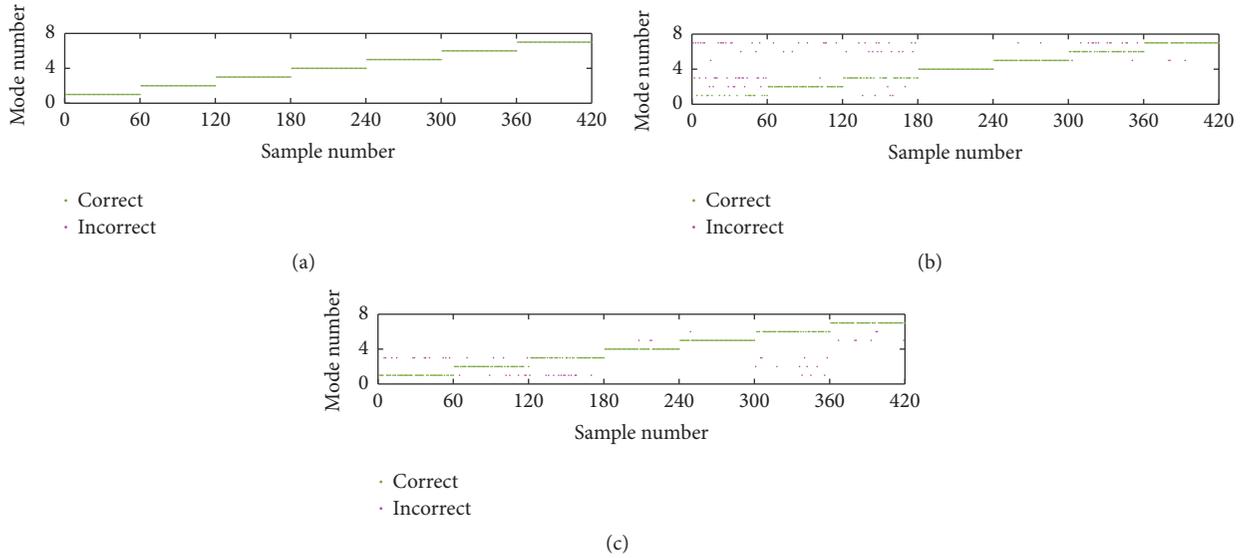


FIGURE 2: The underlying and diagnosed modes for each validating sample. (a) Underlying mode; (b) modes diagnosed using training data only; (c) modes diagnosed incorporating training data and response information.

6.2. *Bayesian Diagnosis Using RSM and Historical Data.* In order to evaluate diagnosis performance, two criteria are used. One is the false negative rate, which can be obtained using the simple quotient

$$F_N = \frac{n_{\text{inc}}}{n_{\text{inc}} + n_{\text{cor}}}, \quad (30)$$

where  $n_{\text{inc}}$  is the number of validating samples that are incorrectly diagnosed and  $n_{\text{cor}}$  is the number of samples that are correctly diagnosed. The misdiagnosis rate is related to mode number. In order to exclude the influence of the mode number, we define a *relative misdiagnosis rate* (RMR), and the aforementioned misdiagnosis rate will be referred to as *absolute misdiagnosis rate* (AMR). Assume that the underlying mode of a validating sample is  $m_j$ ; calculate posteriors of all  $J$  modes. If there are  $t$  posteriors that are less than the posterior of the underlying mode  $p(m_j | e_k)$ , the correct diagnosis number for this sample will be  $t/(J-1)$ , and the incorrect diagnosis number be  $1-t/(J-1)$ . Then  $n_{\text{cor}}$  and  $n_{\text{inc}}$  are obtained by adding up the correct and incorrect

number of all validating samples, respectively. Finally, the RMR is obtained using the same quotient in (30). By such definition, when  $p(m_j | x_k)$  is larger than posteriors of all other modes, it will be counted as one correct diagnosis; also, when  $p(m_j | x_k)$  is larger than some other posteriors, still a positive fraction will be added to  $n_{\text{cor}}$ .

In order to mimic the background knowledge of response information, we obtained 10000 samples through 10000 runs of simulation and established the response information table based on the distribution of these samples.

The diagnosis performance of the proposed approach is evaluated in comparison with diagnosis using training data only.

In Figure 2, the horizontal axis represents the sample number. The underlying mode of the 1–60, 61–120, 121–180, 181–240, 241–300, 301–360, and 361–420 sample is  $m_1$ ,  $m_2$ ,  $m_3$ ,  $m_4$ ,  $m_5$ ,  $m_6$ , and  $m_7$ , respectively. In Figure 2(a), the true underlying modes for each validating sample are shown, while in Figures 2(b) and 2(c), the modes diagnosed without and with incorporating of background knowledge are shown,

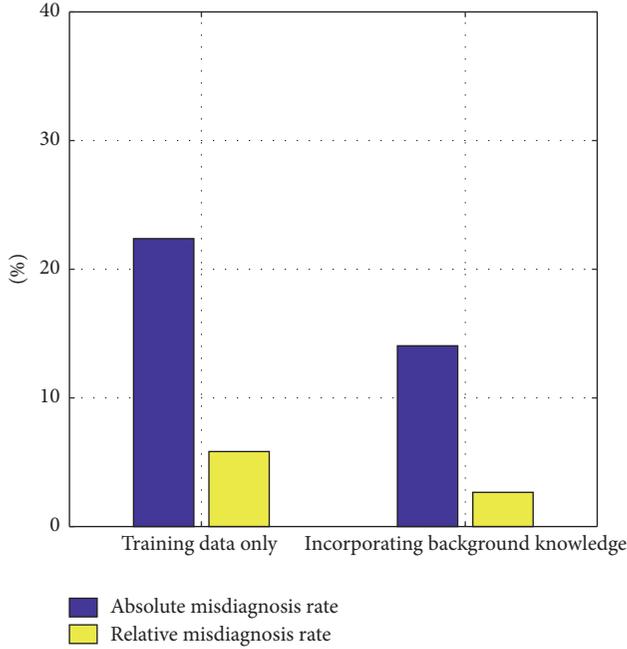


FIGURE 3: Average AMR and RMR from diagnosis with or without incorporating background knowledge.

respectively. Green points represent the samples that are correctly diagnosed, and pink points represent those incorrectly diagnosed. Not surprisingly, the proposed diagnosis approach incorporating background knowledge expressed as response information results in the better performance, while without use of the background knowledge, the diagnosis performs worse in the percentage of mode diagnosed.

The average AMR and the RMR from diagnosis with or without incorporating response information are shown in Figure 3. It is clearly observed that when combining the background knowledge together with the training data, diagnosis results are much better than when only training data is incorporated.

## 7. Conclusions

The objective is to isolate the problem source that is degrading the control performance. In order to reduce dependence on the amount of data available, our approach is to emphasize the use of background information and incorporate the background knowledge of response information into the diagnosis. The knowledge in general terms of RSM can be translated to constraints on the underlying probability distribution. We introduce the constraints in the Bayesian inference such that the dimensionality of the observation space is reduced, and thus the diagnosis can be enhanced. Moreover, for the comparative judgments to be consistent, the set of posterior probabilities computed from different observation subspaces is synthesized by using the eigenvalue formulation on pairwise comparison matrix; therefore, we can obtain the partially ordered posteriors and then determine the state of the process under diagnosis. The approach is applied to a diagnosis problem on an oil sand solids handling

system. The advantage of combining background knowledge and data is achieved even when the amount of training data is limited. To sum up, training data and background knowledge are used for solving different parts of the control performance diagnosis problems. When both are used, the optimal diagnosis is achieved.

## Appendix

**Theorem A.1** (see [31]). Let  $\varepsilon_{obs}$  and  $N_{x_k}$  be defined as in Section 4. Let  $M^{J^l}$  and  $X^{J^l}$ ,  $l = 1, \dots, r$ , be discrete variable, and let  $\{1, \dots, K\}$  be the domain of  $X^{J^l}$ . Let  $D$  denote training data. Introduce parameters  $\Theta$  according to (4), and let the density  $f(\Theta)$  be given by (5). Then it holds that

$$p(X^{J^1} = x_{k_1}, \dots, X^{J^r} = x_{k_r} | M^{J^r} = m_j, D, Q) = \frac{\prod_{x_k \in \varepsilon_{obs}} \prod_{m=0}^{N_{x_k}-1} (n_{k|m_j} + \alpha_{k|m_j} + m)}{\prod_{m=0}^{r-1} (N_{m_j} + A_{m_j} + m)}, \quad (\text{A.1})$$

where  $n_{k|m_j}$  is the number of samples in training data where the observation is  $X^j = x_k$  when  $M^j = m_j$ ,  $N_{m_j} = \sum_{k=1}^K n_{k|m_j}$ , and  $A_{m_j} = \sum_{k=1}^K \alpha_{k|m_j}$ .

## Conflicts of Interest

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

## Acknowledgments

This work was financially supported in part by the National Natural Science Foundation of China, nos. 61304141 and 61573296, the Natural Science Foundation of Fujian Province of China, no. 2014J01252, the Open Project of Key Laboratory of Advanced Process Control in Light Industry by the Ministry of Education of China, no. APCLII602, and the Specialized Research Fund for the Doctoral Program of Higher Education of China, no. 20130121130004.

## References

- [1] K. Zhang, R. Gonzalez, B. Huang, and G. Ji, "Expectation-maximization approach to fault diagnosis with missing data," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 2, pp. 1231–1240, 2015.
- [2] R. Gonzalez, B. Huang, F. Xu, and A. Espejo, "Dynamic bayesian approach to gross error detection and compensation with application toward an oil sands process," *Chemical Engineering Science*, vol. 67, no. 1, pp. 44–56, 2012.
- [3] S. Ahmed, B. Huang, and S. L. Shah, "Validation of continuous-time models with delay," *Chemical Engineering Science*, vol. 64, no. 3, pp. 443–454, 2009.
- [4] B. Huang, "Bayesian methods for control loop monitoring and diagnosis," *Journal of Process Control*, vol. 18, no. 9, pp. 829–838, 2008.
- [5] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and

- process monitoring methods on the benchmark Tennessee Eastman process,” *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.
- [6] H. Jiang, R. Patwardhan, and S. L. Shah, “Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix,” *Journal of Process Control*, vol. 19, no. 8, pp. 1347–1354, 2009.
- [7] F. Qi, *Bayesian approach for control loop diagnosis [Ph.D. dissertation]*, University of Alberta, Alberta, Canada, 2011.
- [8] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, “A review of process fault detection and diagnosis part I: quantitative model-based methods,” *Computers and Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003.
- [9] V. Venkatasubramanian, R. Rengaswamy, and S. N. Kavuri, “A review of process fault detection and diagnosis: part II: qualitative models and search strategies,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.
- [10] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, “A review of process fault detection and diagnosis part III: process history based methods,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, 2003.
- [11] K. Peng, B. Wang, and J. Dong, “An efficient quality-related fault diagnosis method for real-time multimode industrial process,” *Journal of Control Science and Engineering*, vol. 2017, Article ID 9560206, 13 pages, 2017.
- [12] Y. Sun and Z. Liang, “Fault diagnosis and fault tolerant control for non-Gaussian singular time-delayed stochastic distribution systems with disturbance based on the rational square-root model,” *Journal of Control Science and Engineering*, Article ID 9563481, Art. ID 9563481, 11 pages, 2016.
- [13] G. Mani and J. Jerome, “Intuitionistic fuzzy expert system based fault diagnosis using dissolved gas analysis for power transformer,” *Journal of Electrical Engineering and Technology*, vol. 9, no. 6, pp. 2058–2064, 2014.
- [14] B. Pulido, V. Puig, T. Escobet, and J. Quevedo, “A new fault localization algorithm that improves the integration between fault detection and localization in dynamic systems,” in *Proceedings of 16th International Workshop on Principles of Diagnosis*, 2005.
- [15] M. Daigle, X. Koutsoukos, and G. Biswas, “Multiple fault diagnosis in complex systems,” in *Proceedings of 17th International Workshop on Principles of Diagnosis*, 2006.
- [16] M. Askarian, R. Zarghami, F. Jalali-Farahani, and N. Mostoufi, “Fault diagnosis of chemical processes considering fault frequency via Bayesian network,” *Canadian Journal of Chemical Engineering*, vol. 94, no. 12, pp. 2315–2325, 2016.
- [17] X.-X. He, L.-T. Chen, and H.-T. Jia, “Building cognizance rule knowledge for fault diagnosis based on fuzzy rough sets,” *Journal of Intelligent and Fuzzy Systems*, vol. 29, no. 6, pp. 2327–2333, 2015.
- [18] F. Asghar, M. Talha, and S. H. Kim, “Neural network based fault detection and diagnosis system for three-phase inverter in variable speed drive with induction motor,” *Journal of Control Science and Engineering*, vol. 2016, Article ID 1286318, 12 pages, 2016.
- [19] Y. Zhang, H. Wei, R. Liao, Y. Wang, L. Yang, and C. Yan, “A new support vector machine model based on improved imperialist competitive algorithm for fault diagnosis of oil-immersed transformers,” *Journal of Electrical Engineering and Technology*, vol. 12, no. 2, pp. 830–839, 2017.
- [20] R. F. Sales, R. Vitale, S. M. de Lima, M. F. Pimentel, L. Stragevitch, and A. Ferrer, “Multivariate statistical process control charts for batch monitoring of transesterification reactions for biodiesel production based on near-infrared spectroscopy,” *Computers and Chemical Engineering*, vol. 94, pp. 343–353, 2016.
- [21] J. Yu and S. J. Qin, “Statistical MIMO controller performance monitoring. Part II: Performance diagnosis,” *Journal of Process Control*, vol. 18, no. 3-4, pp. 297–319, 2008.
- [22] A. Pernestål, *A Bayesian Approach to Fault Isolation with Application to Diesel Engines [M.S. thesis]*, School of Electrical Engineering, KTH, Stockholm, Sweden, 2007.
- [23] F. Qi, B. Huang, and E. C. Tamayo, “A Bayesian approach for control loop diagnosis with missing data,” *AIChE Journal*, vol. 56, no. 1, pp. 179–195, 2010.
- [24] R. Gonzalez and B. Huang, “Data-driven diagnosis with ambiguous hypotheses in historical data: a generalized Dempster-Shafer approach,” in *Proceedings of the 16th International Conference of Information Fusion (FUSION '13)*, pp. 2139–2144, July 2013.
- [25] R. Gonzalez and B. Huang, “Control loop diagnosis with ambiguous historical operating modes: Part 2, information synthesis based on proportional parametrization,” *Journal of Process Control*, vol. 23, no. 10, pp. 1441–1454, 2013.
- [26] R. Gonzalez and B. Huang, “Control-loop diagnosis using continuous evidence through kernel density estimation,” *Journal of Process Control*, vol. 24, no. 5, pp. 640–651, 2014.
- [27] R. S. Niculescu, T. M. Mitchell, and R. . Rao, “Bayesian network learning with parameter constraints,” *Journal of Machine Learning Research*, vol. 7, pp. 1357–1383, 2006.
- [28] O. Nasri, I. Gueddi, P. Dague, and K. Benothman, “Spacecraft actuator diagnosis with principal component analysis: application to the rendez-vous phase of the mars sample return mission,” *Journal of Control Science and Engineering*, vol. 2015, Article ID 204918, 11 pages, 2015.
- [29] D. Geiger and D. Heckerman, “A characterization of the Dirichlet distribution through global and local parameter independence,” *The Annals of Statistics*, vol. 25, no. 3, pp. 1344–1369, 1997.
- [30] A. O’Hagan and J. Forster, *Kendall’s Advanced Theory of Statistics*, Arnold, London, UK, 2004.
- [31] A. Pernestål, *Probabilistic fault diagnosis-with automotive applications [Ph.D. dissertation]*, Linköping University, Linköping, Sweden, 2009.
- [32] L. Wang, S. Sharkh, A. Chipperfield, and A. Cruden, “Dispatch of vehicle-to-grid battery storage using an analytic hierarchy process,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 2952–2965, 2017.



**Hindawi**

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

