A State Space Model for Monitoring of the Dynamic Blast Furnace System

Jinhui CAI,1) Jiusun ZENG1) and Shihua LUO2)

1) College of Metrology & Measurement Engineering, China Jiliang University, Hangzhou, 310018 China. 2) School of Statistics, Jiangxi University of Finance and Economics, Nanchang, 330013 China.

(Received on July 4, 2012; accepted on August 17, 2012)

Condition monitoring of the blast furnace system plays an important role in the safe and efficient production of high quality hot metal. This article proposes to use a state space model for monitoring the dynamic blast furnace system. The blast furnace data is assumed to be generated by some independent non-Gaussian source signals and a state space model is used to extract the source signals. An optimization problem with the objective function of minimum Kullback-Leibler divergence, i.e., maximum independence between the source signals is constructed. The system matrix and non-Gaussian signals are obtained by solving the optimization problem. Based on the extracted signals, the support vector data description (SVDD) is used for constructing monitoring statistics. Operational data collected from a real blast furnace containing both normal and faulty data are analyzed and used to test the proposed monitoring strategy. The proposed method is then compared with the dynamic independent component analysis (DICA) based monitoring strategy. It is shown that the state space model based monitoring strategy is more appropriate for monitoring of blast furnace faults.

KEY WORDS: dynamic blast furnace system; state space model; non-Gaussian; fault detection.

1. Introduction

To maintain efficient production of high quality blast furnace hot metal, it is essential to keep the blast furnace system at a steady production status. Hence, condition monitoring of the blast furnace system becomes an important issue. Many kinds of faults can be observed during the blast furnace ironmaking process, e.g., hanging, low stockline, abnormal gas flow distribution, etc. It is highly desired that a fault detection and diagnosis system is capable of detecting the process fault at an early stage so that there is enough time for corrective operation.

The problem of fault detection and diagnosis in the blast furnace ironmaking process is traditionally handled in expert systems using expert knowledge and fuzzy logic.1–3) In many expert systems, fault detection and diagnosis using knowledge based technique has become a standard configuration.4) One of the main advantages of the knowledge based techniques is that they are easy to understand, however, collection and maintenance of knowledge base is not a trivial task. Another method to deal with fault detection and diagnosis of blast furnace is to formulate it as a kind of fault classification problem, where self organization map and support vector machine (SVM) may be the most commonly used techniques.5–7) The problem of such method is that one needs to collect enough samples for almost all the faulty conditions to successfully distinguish between different faults. As the blast furnace ironmaking process operates under normal conditions most of the time, collection of sufficient faulty samples requires very long time and heavy efforts.

In contrast, multivariate statistical process control (MSPC) techniques use different kinds of control charts to determine whether the new sample is normal or faulty data. It is data based and easy to implement, hence received considerable attention during recent years.8–10) Vanhatalo11) used principal component analysis (PCA) to monitor the thermal state of an experimental blast furnace and reported success. It was found that multivariate approach to monitoring of the blast furnace is a promising method and is capable of providing unbiased information of the blast furnace.11) Despite the success, principal component analysis has its own disadvantages, e.g., it is designed for static systems and applicable only to Gaussian distributed data, while it is well known that the blast furnace ironmaking process is a highly dynamic non-Gaussian system.12)

In this article, a state space model is used to model the dynamic blast furnace system and it is assumed that the collected process data is generated by some non-Gaussian source signals. In the state space model, the observed process variables are used as model input and the non-Gaussian independent signals are used as model output. An optimization problem involving maximizing the Kullback-Leibler divergence of the source signals is formulated and solved to extract the non-Gaussian source signals. The extracted source signals are then used to monitor the blast furnace ironmaking process. The proposed method can well accom-
modulate the dynamics and non-Gaussianity of the blast furnace data and is suitable for monitoring of the blast furnace system.

The remainder of the paper is organized as follows. Section 2 presents the process model and the signal extraction method. Section 3 gives a brief introduction to the support vector data description (SVDD) method used to monitor the non-Gaussian source signals. Section 4 and 5 present the monitoring strategy and the application results to a real blast furnace. Concluding summaries are then given in Section 6.

2. Non-Gaussian Source Signal Extraction Based on State Space Model

2.1. Model Structure

In this article, the process data of blast furnace ironmaking process is assumed to exhibit non-Gaussian and dynamic behaviors. The model structure used is as follows

\[ \begin{align*}
\dot{x}(k+1) &= \hat{A}x(k) + \hat{B}s(k) + \hat{K}\hat{e}(k) \\
u(k) &= \hat{C}\hat{x}(k) + \hat{D}s(k) + \hat{e}(k)
\end{align*} \] (1)

where \( u \in \mathbb{R}^N \) is the observed process data, \( s \in \mathbb{R}^M \) is the non-Gaussian source signal generating the process data, \( \hat{x} \in \mathbb{R}^n \) is the system state vector, \( \hat{e} \in \mathbb{R}^n \) is the innovation sequence. \( \hat{A} \in \mathbb{R}^{nxn}, \hat{B} \in \mathbb{R}^{nxM}, \hat{C} \in \mathbb{R}^{nxM}, \hat{D} \in \mathbb{R}^{nxM} \) are the system matrices of appropriate dimensions and \( \hat{K} \) is the Kalman filter gain. For the sake of simplicity, it is assumed that the number of source signals and observed signals is the same, i.e., \( N = M \) so that \( \hat{D} \) is full rank. The objective is to extract non-Gaussian source signals \( s \) from the observed signals \( u \) from model (1).

Now define

\[ \begin{align*}
A &= \hat{A} - \hat{B}\hat{D}^{-1}\hat{C}, \\
B &= \hat{B}\hat{D}^{-1}, \\
C &= -\hat{D}^{-1}\hat{C}, \\
D &= \hat{D}^{-1}, \\
K &= \hat{K} - \hat{D}^{-1}\hat{e}, \\
e_1 &= -\hat{D}^{-1}\hat{e}1 
\end{align*} \] (2)

Since \( \hat{D} \) is full rank, model (1) can be reformulated as follows

\[ \begin{align*}
\dot{x}(k+1) &= Ax(k) + Bu(k) + Ke(k) \\
\dot{s}(k) &= Cx(k) + Du(k) + e(k)
\end{align*} \] (3)

Reference\textsuperscript{13} proved that using model (3) one can effectively extract the source signals generating the dynamic data, hence the problem of source extraction for dynamic systems can be transferred into the signal separation problem based on model (3).

2.2. Objective Function

To get the non-Gaussian source signals generating the dynamic data from model (3), consider the optimization problem of minimizing mutual information between extracted signals. The mutual information between variables is commonly measured by the Kullback-Leibler (KL) divergence, which is also used in signal extraction algorithms like independent component analysis (ICA)\textsuperscript{14} which finds the independent components by maximizing the statistical independence of the estimated components. Let \( p(s), q(s) \) be two sets of probability density functions (PDF), the KL divergence is defined as

\[ l = \int p(s) \log \frac{p(s)}{q(s)} \, ds \] (4)

The KL divergence is commonly used to measure the difference between two probability density functions (PDF). From Eq. (4) it can be seen that the KL divergence is non-negative. The KL divergence is equal to zero if and only if \( p(s) = q(s) \), hence it can be used to measure the dependence between two sets of signals. In this article, \( q(s) \) is the reference PDF which is assumed to be independent and identically distributed (i. i. d.) and \( p(s) \) is the actual PDF of the source signals. Hence minimizing Eq. (4) is equivalent to maximizing the independence between the extracted source signals.

Let \( p(s_i) \) be the marginal PDF of the \( i \)th component defined as

\[ p_i(s^i) = \int p(s)ds^{i}, \ i = 1, \ldots, N \] (5)

where \( s^i \) is the \( N - 1 \) dimensional vector after deleting the \( i \)th component of \( s \). For the reference PDF \( q(s) \), the following equation holds

\[ q(s) = \prod_{i=1}^{N} p_i(s_i) \] (6)

The following objective for maximizing the independence between the source signals can then be formulated

\[ l = \int p(s) \log \left( \frac{p(s)}{\prod_{i=1}^{N} p_i(s_i)} \right) \, ds \]

\[ = \int p(s) \log(p(s))ds - \int p(s) \log p_i(s_i)ds_i \] (7)

Provided \( p(s) \) and \( p_i(s_i) \) are known, it is possible to obtain the system matrices and further the source signals \( s \) by solving Eq. (7). Since the PDFs are unknown, consider the case of \( L \) samples and arrange the samples into the following matrix form

\[ U(k) = \left[ u^T(1), u^T(2), \ldots, u^T(L) \right]^T \]

\[ S(k) = \left[ s^T(1), s^T(2), \ldots, s^T(L) \right]^T \]

where \( u(k), s(k) \) are the \( k \)th sample of the observed signal and the source signals. If the source signals are independent and identically distributed, then the PDF of matrix \( S \) can be written as

\[ p(S) = \prod_{i=1}^{N} \prod_{k=1}^{L} p_i(s_k(i)) \] (8)

Substituting Eq. (8) into Eq. (7) one obtains

\[ l = \frac{1}{L} \int p(S) \log \frac{p(S)}{\prod_{i=1}^{N} \prod_{k=1}^{L} q_i(s_k(i))} \] (9)

Similar to the case of independent component analysis, \( q_i(\cdot) \) is used to approximate the PDFs of the source signals \( p_i(\cdot) \), where \( q_i(\cdot) \) has the following form\textsuperscript{15}:

\[ q_i(\cdot) = \log\cos x \] (10)

Let the initial state be \( x(1) = 0 \), the following holds

\[ S = \mathcal{W}U \] (11)

\[ \mathcal{W} = \begin{bmatrix} 2195 \end{bmatrix} \] (12)
\[
\gamma = \begin{bmatrix}
H_0 & 0 & \cdots & 0 & 0 \\
H_1 & H_0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
H_{L-2} & H_{L-3} & \cdots & H_0 & 0 \\
H_{L-1} & H_{L-2} & \cdots & 0 & H_0
\end{bmatrix}
\]

where \( H_0 = D, H_i = CA_i^iB, i = 1, \ldots, N \). According to the property of PDF, we have

\[
p(S) = \frac{p(l)}{\det H_0} \quad \text{................. (12)}
\]

where \( \det H_0 \) is the determinant of \( H_0 \). Combining Eqs. (9), (10) and (12), the objective function for extracting the source signals can be formulated as

\[
\min l = -\log \det H_0 - \sum_{i=1}^{N} \frac{1}{L} \sum_{t=1}^{L} \log q_i(s_i(k)) = -\log \det D - \sum_{i=1}^{N} \sum_{t=1}^{L} \log q_i(s_i) \quad \text{................. (13)}
\]

Solving the above optimization problem one can get the system matrices \( A, B, C, D \) and further the source signals. It should be noted that the PDF \( p(l) \) has been neglected in Eq. (13) as for a set of given samples \( p(l) \) is a constant.

### 2.3. Solution Strategy

The system matrices \( C, D \) can be obtained by solution of Eq. (13) through gradient search. The total differential of Eq. (13) is as follows

\[
dl = -\text{tr}(\Delta DD^{-1}) + \varphi(T)(s)dS \quad \text{................. (14)}
\]

where \( \text{tr}(\cdot) \) is the trace of a matrix, \( \varphi(T)(s) \) is a vector with the \( f^\text{th} \) component

\[
\varphi(s_j) = -\frac{d \log q_j(s_j)}{d S_j} = -\tanh(s_j) \quad \text{................. (15)}
\]

On the other hand, the total differential of Eq. (3) is

\[
ds = dCx + dDu + Cdx \quad \text{................. (16)}
\]

Combining Eqs. (14) and (16), the derivative of \( l \) with respect to system matrices \( C, D \) can be obtained

\[
\frac{\partial l}{\partial C} = \varphi(s)x^T \quad \text{................. (17)}
\]

\[
\frac{\partial l}{\partial D} = -D^T + \varphi(s)u^T \quad \text{................. (18)}
\]

So that the we have

\[
\Delta C(k) = -\eta \varphi(s)x^T \quad \text{................. (19)}
\]

\[
\Delta D(k) = \eta(D^T - \varphi(s)u^T) \quad \text{................. (20)}
\]

where \( \eta \) is the learning rate. Using Eqs. (19) and (20) the system matrices \( C, D \) can be solved.

To get the matrices \( A, B \), the Kalman filter based approach(13) is used. According to Eq. (3) we have

\[
x(k + 1) = Ax(k) + Bu(k) + Ke(k) \quad \text{................. (21)}
\]

Once the estimation of innovation sequence \( e(k) \) can be obtained, it is possible to get estimation of \( A, B \) through Kalman filter since \( C, D \) are known. The estimation of \( e(k) \) can be computed as

\[
e(k) = \Delta e(k) = \Delta Cx(k) + \Delta Du(k) \quad \text{................. (22)}
\]

With the estimation of innovation sequence, the state sequence can be obtained from the following steps

Step 1: Get the Kalman filter gain \( K_k = P_k C_k^T(C_k P_k C_k^T + R_k)^{-1} \), where \( R_k \) is the covariance of the estimated innovation sequence \( e(k) \);

Step 2: Obtain the innovation sequence through Eq. (22) and then get the estimation of state sequence \( \hat{x}_k = x_k + K_k e(k) \);

Step 3: Update the error covariance matrix \( \hat{P}_k = (I - K_k C_k) \hat{P}_k \);

Step 4: Compute the state sequence of the next time step \( \hat{x}_{k+1} = A \hat{x}_k + B \hat{u}_k \).

Step 5: Update the covariance matrix of the next time step \( \hat{P}_{k+1} = A \hat{P}_k A^T_k + Q_k \).

Here \( Q_k \) and \( R_k \) are the covariance matrix of the output and measurement noise. With the state sequence obtained, the estimation of system matrices \( A, B \) and hence the source signal \( s \) can be obtained.

### 3. Support Vector Data Description

The source signals generated from observed blast furnace data using model (3) are independent and identically distributed non-Gaussian. Traditional methods like Hotelling’s \( T^2 \) statistic are not suitable for monitoring of blast furnace system due to its non-Gaussianity. Here the support vector data description (SVDD)(6,7) is used to construct statistics for the extracted non-Gaussian signals \( s \). The basic idea of SVDD is to project the original data into high dimensional kernel space and to envelop the data by a hyper sphere. The hyper sphere should have a minimum radius whilst covering as many samples as possible. In the kernel space, normal samples are within the hyper sphere while faulty samples are outside the sphere.

For a set of training set \( \{s_i\}_{i=1}^{L} \), the hyper sphere can be obtained by solving the following optimization problem

\[
\min f(R, \text{a}) = R^2 + C \sum_i \xi_i
\]

s.t. \( \|s_i - \text{a}\| \leq R + \xi_i, \xi_i \geq 0, i = 1, \ldots, N \quad \text{................. (23)} \)

where \( R \) is the radius of the hyper sphere, \( \text{a} \) is the centre. Slack variable \( \xi_i \) denotes the misclassification probability for the \( i^\text{th} \) training sample, constant \( C \) gives the tradeoff between misclassification and the radius of the hyper sphere. Eq. (23) can be transformed into the following dual optimization problem using Lagrange multiplier

\[
\max \sum_{\alpha} \sum_{i=1}^{L} \alpha_i \alpha_i (s_i, s_j) - \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j (s_i, s_j)
\]

s.t. \( \sum_{\alpha} \alpha_i = 1, \alpha_i \in [0, C], i = 1, \ldots, N \quad \text{................. (24)} \)

Solving Eq. (24) the centre can be obtained as \( \text{a} = \sum_{\alpha} \alpha_i s_i \), and the radius

\[
R^2 = (s_k, s_k) - \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j (s_i, s_j) \quad \text{................. (25)}
\]

where \( S_k \) is the support vector on the boundary of the hyper sphere. Now it is possible to judge whether a test sample \( z \) is a normal or faulty sample. If the following conditions holds
Then \( z \) is a normal sample, otherwise it is a faulty sample. The inner product \( \langle z, z \rangle \) is always replaced by a kernel function, which can be chosen as \( K(s_i, s_j) = \exp \left( -\frac{\|s_i - s_j\|^2}{\sigma^2} \right) \).

4. Process Monitoring Strategy Based on State Space Model

With the source signals obtained, the monitoring problem of dynamic blast furnace system can be transferred into the monitoring problem of non-Gaussian source signals. The process monitoring strategy consists of two procedures, i.e., offline modeling and online monitoring. The offline modeling step consists of the following steps

Step 1: Collect normal samples to construct a database for model training;
Step 2: Pre-whiten the training samples using PCA;
Step 3: Obtain the system matrices \( A, B, C, D, K \) from the algorithm in Section 2;
Step 4: Get the non-Gaussian source signals \( s \) from Eq. (3);
Step 5: Compute the confidence limit \( R^2 \) in Eq. (26) using SVDD.

The flowchart is shown in Fig. 1.

![Flowchart of offline modeling](image)

On the other hand, the online monitoring procedure consists of the following steps

Step 1: Collect the sample at the current time step and pre-whiten it;
Step 2: Get the non-Gaussian source signals from Eq. (3);
Step 3: Compute the distance of the current sample from the centre \( D \), if \( D \leq R^2 \), the current sample is normal sample, otherwise, it is a faulty sample.

5. Application to Process Monitoring of a Real Blast Furnace

This section presents the application results of the proposed monitoring strategy to a real blast furnace in China. The blast furnace has an inner volume of 2,500 m³. For the purpose of model training, 2,000 samples of 10 process variables under normal operation are collected. A list of input variables is shown in Table 1.

A test set containing 580 hourly mean values of the process variables are collected under a typical faulty condition, i.e., hanging. The 2,000 normal samples and the 580 faulty samples are hourly mean values of process variables. During the hanging fault, the quantity of blast and the permeability index reduced and the CO and CO₂ concentration in the top gas increased. Figure 2 illustrates the time series graph of the 2,580 samples.

Figure 2 shows that there is clear change in several variables, e.g., CO concentration \((u_9)\), CO₂ concentration \((u_{10})\) and quantity of coal powder \((u_4)\); also the quantity of blast \((u_1)\) and permeability index \((u_6)\) decreases in the last 580 samples. A state space model is constructed using the normal samples and the system order is set as 20. The source signals are extracted from the state space model and SVDD is used to construct monitoring statistics, denoted by \( D_{svdd}^2 \). The parameters are set as \( \sigma = 100, C = 0.01 \). With 99% of confidence limit, the monitoring results are shown in Fig. 3.

It can be seen from Fig. 3 that significant violation of the confidence limit can be observed, indicating there is a fault.

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>quantity of blast</td>
</tr>
<tr>
<td>2</td>
<td>temperature of blast</td>
</tr>
<tr>
<td>3</td>
<td>pressure of blast</td>
</tr>
<tr>
<td>4</td>
<td>quantity of coal powder</td>
</tr>
<tr>
<td>5</td>
<td>pressure of top gas</td>
</tr>
<tr>
<td>6</td>
<td>permeability index</td>
</tr>
<tr>
<td>7</td>
<td>the quantity of oxygen blasted</td>
</tr>
<tr>
<td>8</td>
<td>coke ratio</td>
</tr>
<tr>
<td>9</td>
<td>CO₂ concentration in top gas</td>
</tr>
<tr>
<td>10</td>
<td>CO concentration in top gas</td>
</tr>
</tbody>
</table>

![Original time series data (the horizontal lines being the sample number of the hourly mean value of BF variables)](image)
during the operation of the blast furnace system. The hanging fault is successfully detected.

For comparison, the dynamic independent component analysis (DICA) method proposed in reference\textsuperscript{18}) was tested. The DICA method applies ICA to the augmenting matrix with the time lagged variables and construct 2 statistics, i.e., $D_{dica}^2$ for extracted non-Gaussian components and SPE statistics for the residual space. For the blast furnace data, the time lag for all the 10 variables are set as 3, hence a variable set with 30 variables can be obtained. ICA is first applied to extract 10 sets of non-Gaussian signals from the 30 variable sets. Similar to the state space model, SVDD is used to get the confidence limit for process monitoring. The parameters of SVDD are set as $\sigma = 10$, $C = 0.01$ and the confidence level is also set as 99%. The monitoring results of DICA are shown in Fig. 4.

On the other hand, the SPE statistic for the residual space of DICA is constructed for monitoring of the hanging fault and the results are shown in Fig. 5.

Comparing Figs. 3, 4 and 5, it can be seen that $D_{dica}^2$ statistic has the best detection results. For $D_{dica}^2$, although the fault can also be detected, there are many missed detection. In contrast, the SPE statistic performs poorly with significant number of missed detection and false alarms. To further compare the proposed method with the DICA monitoring strategy, Table 2 gives the comparison of detection rate between different statistics.

It is shown in Table 2 that the detection rates for $D_{dica}^2$ under 95% and 99% confidence level are 90.72% and 86.72% respectively. In contrast, the detection rates for $D_{dss}^2$ statistic of the DICA approach are 44.31% and 21.9% respectively. The SPE statistic for DICA has the lowest detection rates of 3.25% and 1.38%. Hence the proposed monitoring strategy can better accommodate the dynamic non-Gaussian blast furnace data. Comparing to the dynamic ICA approach, the proposed state space model can better account for the dynamic and non-Gaussian characteristics of the blast furnace system. The dynamic ICA considers the augmenting matrix of time lagged variables, which tends to significantly increase the scale of the problem. Moreover, virtually all ICA algorithms assume the observed data to be independently and identically distributed. The dynamic ICA may be unable to recover the sources in such an environment.\textsuperscript{19} Different from the dynamic ICA approach, the state space model does not increase the dimension of the system and converts the source signal extraction problem to a deconvolution problem. Hence the state space model can avoid the problem confronted by dynamic ICA.

6. Conclusions

This article proposes to use a state space model to monitor the dynamic blast furnace ironmaking process. The state space model is formulated into an optimization problem, which can be solved by a combination of gradient search and Kalman filter. The model is then used to extract non-Gaussian source signals and SVDD is used to construct monitoring statistics for the non-Gaussian components. Application to a real blast furnace shows that the proposed state space model successfully detects the hanging fault. Comparing to a real blast furnace shows that the proposed state space model based monitoring method has lower missing rates. The proposed method well accounts for the dynamics and non-Gaussianity of the blast furnace ironmaking process.
and hence are suitable for fault detection and diagnosis of the blast furnace ironmaking process.

Acknowledgements

The authors wish to thank sponsor and financial support from Natural Science Foundation of China under grant No. 61203088, 61263014 and Zhejiang Provincial Natural Science Foundation under grant No. LQ12F03015 and Y1110171.

REFERENCES