Article

A Hybrid Cluster Variational Autoencoder Model for Monitoring the Multimode Blast Furnace System

Chenyu Chen and Jinhui Cai *

College of Metrology and Measurement Engineering, China Jiliang University, Hangzhou 310018, China;
chency@cjlu.edu.cn
* Correspondence: caijinhui@cjlu.edu.cn

Abstract: Efficient monitoring of the blast furnace system is crucial for maintaining high production efficiency and ensuring product quality. This article introduces a hybrid cluster variational autoencoder model for monitoring the blast furnace ironmaking process which exhibits multimode behaviors. In contrast to traditional approaches, this method utilizes neural networks to learn data features and effectively handles the diverse feature types observed in different production modes. Through the utilization of a clustering process within the hidden layer of the variational autoencoder, the proposed technique facilitates efficient fault detection in the context of multimodal blast furnace data. Based on the variational autoencoder model, this study further establishes a unified monitoring index and defines a method for computing the control limits. The application of the model to real blast furnace data reveals its proficiency in accurately identifying faults across diverse modes; compared with the probabilistic principal component analysis based on the local nearest neighbor standardization method and the recursive probabilistic principal component analysis, the model shows a reduction in false positives by up to 10.3% and a substantial reduction of 19.2% in the missed detection rate. This method achieves a remarkable false detection rate of only 0.2% and 0 instances of missed detection.

Keywords: multimode blast furnace system; process monitoring; variational autoencoder; Gaussian mixture model

1. Introduction

The iron and steel industry, as a cornerstone of modern manufacturing, plays a vital economic role in modern manufacturing. Blast furnace ironmaking is pivotal within this sector, significantly affecting system stability and productivity. The blast furnace system includes subsystems like feeding, gas treatment, hot air, injection, iron-out, and the furnace body [1]. The furnace body is central to ironmaking, housing reactions yielding molten iron. Here, raw materials enter through the top distributor, while hot air, oxygen-enriched gases, and fuel are injected from the bottom [2]. Liquid iron and slag exit the lower section, and gas is collected from the upper section [3]. Various types of faults occur in blast furnace ironmaking, such as suspension, low feed, cooling, accumulation, and airflow issues, stemming from factors like raw material instability, external fluctuations, manual errors, and equipment glitches. Therefore, detecting and addressing these faults early is crucial for maintaining production quality.

Two primary fault detection methods are employed in blast furnace ironmaking: expert system-based methods and data-driven methods [3]. Expert system-based methods rely on process knowledge and operator expertise to identify abnormal states within the blast furnace [4]. These methods are characterized by their ease of comprehension and widespread utilization, making them the most commonly employed fault detection approach. However, the expert system-based method requires substantial efforts in establishing and maintaining a comprehensive knowledge base. An alternative approach is

data-driven, utilizing historical data; this approach uncovers relationships among variables derived from historical data to construct a fault detection model. Multivariate statistical process control such as principal component analysis (PCA) has been used to monitor the operational state of experimental blast furnaces, achieving significant success [5]. Other machine-learning-based approaches, including techniques like autoencoder [6] and support vector machines [7–9], have also been widely employed. Given that blast furnace ironmaking represents a traditional continuous chemical process, its data typically exhibit characteristics such as nonlinearity, multimodality, and dynamic patterns, etc. Researchers such as Zhou et al. [10] have employed the kernel principal component analysis (KPLS) method to deal with this. And Liu, Zeng et al. [11] developed a novel probabilistic monitoring framework using the probabilistic linear discriminant analysis (PLDA). Additionally, Zhang et al. [2] have demonstrated the efficacy of using denoising autoencoders to extract nonlinear characteristics, exhibiting promising results in detecting cold failures in blast furnaces. Traditional methods employed in the blast furnace ironmaking process have predominantly emphasized stable operating conditions. However, due to fluctuations in raw materials and changes in product proportions, the operational conditions of the blast furnace process often experience variations. In such instances, a multimode process fault detection method for blast furnace ironmaking becomes increasingly significant [12]. Zhu et al. [13,14] have enhanced the algorithm model based on the sliding window technique. Zhou [15] has proposed a fault detection method utilizing an improved independent component analysis (ICA) algorithm to discern the importance of different samples, thereby facilitating more reasonable model updates. Despite promising results obtained in simulation datasets, the application of these methods in blast furnace ironmaking is limited by the fact that sliding window methods discard useful information from long-term historical data.

This article presents a novel fault monitoring method for modeling the multimodal blast furnace system. The proposed method combines the variational autoencoder (VAE) framework with a Gaussian mixture model (GMM) to achieve clustering and fault monitoring objectives. The VAE learns the feature representation of nonlinear data and conducts cluster analysis on potential feature representations using the mixture distribution model. The model learns the latent features and reconstructs them as the observed multimodal blast furnace system data through the hybrid cluster VAE method. For fault detection, the method utilizes normal operational data for training and infers the probability density of reconstructed data by solving the posterior probability. Based on this, a probability monitoring index and its corresponding control limit are derived. The proposed method effectively addresses the challenges of nonlinearity and multimodality, rendering it suitable for fault detection in blast furnace ironmaking systems.

2. Variational Autoencoder and Gaussian Mixture Model

The variational autoencoder (VAE), a type of directed probability graphical model implemented through a variational Bayesian method, has consistently demonstrated excellent performance in uncovering the latent features of data. Figure 1 depicts the structure of the variational autoencoder. Assuming that an observed dataset $\mathbf{X} \in \mathbb{R}^{N \times m}$ has been collected, with n being the number of samples and m the amount of variables, let $\mathbf{x} \in \mathbb{R}^m$ be a data sample and $\mathbf{z} \in \mathbb{R}^d$ the latent feature vector, where d is the number of latent features. The variational autoencoder comprises two primary components: the decoder part, referred to as $p_{\theta}(\mathbf{x}|\mathbf{z})$, and the encoder part, denoted as $q_{\phi}(\mathbf{z}|\mathbf{x})$, where ϕ and θ denote the parameters.

The decoder part, $p_{\theta}(\mathbf{x}|\mathbf{z})$, of the variational autoencoder is commonly referred to as the generative model. It generates the high-dimensional observed variables $\mathbf{x}$ based on the latent variables $\mathbf{z}$, with a prior assumption distribution. Through the trained network structure, the decoder performs the task of data reconstruction. Typically, the prior distribution of the latent variables is chosen by the standard normal distribution

$p(\mathbf{z}) \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$, and the decoder part can be described as a multivariate normal distribution as shown by the following formula:

$$p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{z}; \boldsymbol{\theta}), \mathrm{diag}(\sigma^2(\mathbf{z}; \boldsymbol{\theta}))) \tag{1}$$

Here, $\boldsymbol{\mu}(\mathbf{z}; \boldsymbol{\theta})$ and $\mathrm{diag}(\sigma^2(\mathbf{z}; \boldsymbol{\theta}))$ represent the mean vector and covariance matrix, respectively.
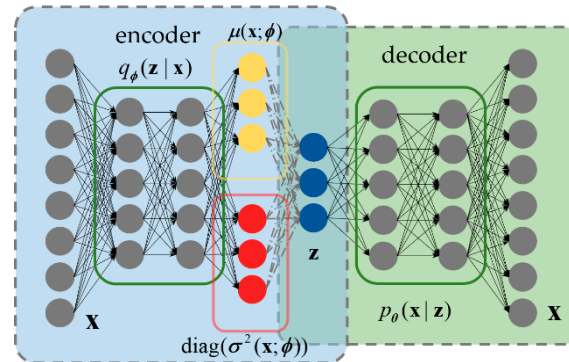

Figure 1. The structure of the variational autoencoder.

The encoder of the variational autoencoder, also referred to as the inferential model or identification model, maps the observed data to a lower-dimensional latent space through nonlinear dimensionality reduction. The encoder operates under similar assumptions as the generative model regarding the distribution of the observed variables. The specific model is as follows:

$$q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{x}; \boldsymbol{\phi}), \mathrm{diag}(\sigma^2(\mathbf{x}; \boldsymbol{\phi}))) \tag{2}$$

To estimate the unknown parameters in the variational autoencoder (VAE) model, the evidence lower bound (ELBO) is used as the cost function. The parameters are then determined by maximizing this variational lower bound:

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \underset{\boldsymbol{\theta}, \boldsymbol{\phi}}{\mathrm{argmax}}(E_{q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})}[\log p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z})] - D_{KL}[q_{\boldsymbol{\phi}}(\mathbf{z}|\mathbf{x})||p(\mathbf{z})]) \tag{3}$$

Here, $D_{KL}(q||p)$ represents the Kullback–Leibler (KL) divergence, which serves as a metric for measuring the difference between two probability distributions. It quantifies how one distribution diverges from another, providing a means to compare the dissimilarity between probability distributions [16].

The Gaussian mixture model (GMM) can be used to classify data into different categories by considering their probability distribution. It assumes that the entire dataset is generated from a mixture of different Gaussian distributions. Assuming that the model consists of *K* Gaussian distributions, and given the parameter $\boldsymbol{\Theta}$, the posterior distribution of observed data $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\} \in \mathbb{R}^{N \times m}$ can be described as follows:

$$p(\mathbf{y}|\boldsymbol{\Theta}) = \sum_{k=1}^{K} p(\boldsymbol{\vartheta}_k) p(\mathbf{y}|\boldsymbol{\vartheta}_k) = \sum_{k=1}^{K} \pi_k f(\mathbf{y}|\boldsymbol{\vartheta}_k) \tag{4}$$

Here, *K* represents the number of Gaussian components, $\pi_k \geq 0$, $\sum_{k=1}^{K} \pi_k = 1$, and π_k is the weight of the *k*-th Gaussian sub-model. $f(\mathbf{y}|\boldsymbol{\vartheta}_k)$ is the posterior distribution of $\mathbf{y}$ under the *k*-th sub-model, with $\boldsymbol{\vartheta}_k = (\boldsymbol{\mu}_k, \Sigma_k)$ and $\boldsymbol{\Theta} = \{\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, \ldots, \boldsymbol{\vartheta}_k\}$. Hence, $f(\mathbf{y}|\boldsymbol{\vartheta}_k)$ can be described as follows:

$$f(\mathbf{y}|\boldsymbol{\vartheta}_k) = \frac{1}{(\sqrt{2\pi})^m |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{(\mathbf{y} - \boldsymbol{\mu}_k)^T \Sigma_k^{-1}(\mathbf{y} - \boldsymbol{\mu}_k)}{2}\right) \tag{5}$$

where m stands for the dimension of the data, and the estimation of the parameters $\mathbf{\Theta}, \mathbf{\Pi} = \{\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_k\}$ in the Gaussian mixture model can be achieved by maximizing the log-likelihood estimation given the training dataset, as follows:

$$\mathbf{\Theta}, \mathbf{\Pi} = \underset{\mathbf{\Theta}, \mathbf{\Pi}}{\text{argmax}} \log L(\mathbf{\Theta}, \mathbf{\Pi}; \mathbf{Y}) \tag{6}$$

Here, the log-likelihood function is

$$\log L(\mathbf{\Theta}, \mathbf{\Pi}; \mathbf{Y}) = \sum_{n=1}^{N} \log p(\boldsymbol{y}_n | \mathbf{\Theta}) = \sum_{n=1}^{N} \log(\sum_{k=1}^{K} \boldsymbol{\pi}_k f(\boldsymbol{y}_n | \boldsymbol{\vartheta}_k)) \tag{7}$$

When dealing with the parameter estimation of the Gaussian mixture model, direct differentiation cannot be easily applied to maximize the log-likelihood function. Therefore, the Expectation–Maximization (EM) algorithm can be employed to solve the maximum likelihood estimation of the parameters in the probabilistic model with hidden variables. The EM algorithm iteratively updates the parameters by alternately performing the E-step (expectation) and M-step (maximization) until convergence is achieved. This iterative process allows for the estimation of the model parameters in the presence of hidden variables.

3. Multimodal Process Monitoring Method Based on Hybrid Cluster Variational Autoencoder

3.1. Model Structure and Parameter Estimation of the Hybrid Cluster Variational Autoencoders

In this section, the details of the hybrid cluster variational autoencoder model are presented. The model introduces a clustering effect on the latent variables, allowing for unsupervised clustering of the hidden-layer variables. This clustering capability facilitates fault monitoring of multimodal data during the generation process of the variational autoencoder. Furthermore, an enhanced version of the variational lower bound is derived for the proposed model. Moreover, a probability index for process monitoring is introduced to further enhance the fault monitoring capability. By considering the clustering aspect of the data, the proposed model effectively addresses the objective of the multimode blast furnace system.

In contrast to the traditional variational autoencoder, the proposed model integrates the variational autoencoder with a Gaussian mixture distribution, specifically designed to handle multimodal data. The aim is to ensure that data of the same mode exhibit proximity in the hidden layer, while data from different modes remain distinct.

In the variational autoencoder model, a common assumption is that the prior distribution $p(\mathbf{z})$ of latent variables follows a standard normal distribution. However, this often restricts variational autoencoders to representing singular distributions, thereby limiting their capability to learn complex distributions with multimodal attributes. To enhance the model's capacity for encompassing various forms of multimodal data, it is proposed to constrain the prior distribution of latent variables using a Gaussian mixture distribution. This adaptation enables the model to combine different modes in accordance with the characteristics of the Gaussian mixture distribution during the latent representation of observed variables. Within the Gaussian mixture distribution, the stratification of distinct data categories is achieved by introducing a parameter-controlled variable $\mathbf{Y}$. Similarly, disparate categories of observation variables can be delineated by introducing parameter c in the HCVAE model.

Within the formulated hybrid clustering variational autoencoder architecture, the latent variables encompass both a discrete variable c, signifying the category of the observed data, and a continuous latent variable $\mathbf{z}$, a constituent also present in the conventional variational autoencoder.

Firstly, we describe the decoding part of the network; in this part, the hidden layer of the model represents the latent variable $\mathbf{z}$. The observed sample $\mathbf{x}$ is generated from $\mathbf{z}$ and c.

The joint probability distribution, also known as the generative model, can be expressed as follows:

$$p(\mathbf{x}, \mathbf{z}, c_k) = p_{\boldsymbol{\theta}_x}(\mathbf{x}|\mathbf{z})p_{\boldsymbol{\theta}_z}(\mathbf{z}|c_k)p(c_k) \tag{8}$$

Since $\mathbf{z}$, $\mathbf{x}$, and c are independent of each other, the probability can be defined in an alternative manner as

$$\begin{aligned} p(c_k) &= Cat(c_k|\boldsymbol{\pi}_k) \\ p_{\boldsymbol{\theta}_z}(\mathbf{z}|c_k) &= \mathcal{N}(\boldsymbol{\mu}_k(c_k; \boldsymbol{\theta}_z), \mathrm{diag}(\sigma_k^2(c_k; \boldsymbol{\theta}_z))) \\ p_{\boldsymbol{\theta}_x}(\mathbf{x}|\mathbf{z}) &= \mathcal{N}(\boldsymbol{\mu}_x(\mathbf{z}; \boldsymbol{\theta}_x), \mathrm{diag}(\sigma_x^2(\mathbf{z}; \boldsymbol{\theta}_x))) \end{aligned} \tag{9}$$

The categories are generated through a categorical distribution $c_k \sim Cat(\boldsymbol{\Pi})$. The potential vector is then generated based on the generated categories, and the original input data are reconstructed using the decoding part of the hybrid cluster variational autoencoder, $p_{\boldsymbol{\theta}_x}(\mathbf{x}|\mathbf{z})$, by sampling the latent variable $\mathbf{z}$. Here, K is a predetermined initial parameter representing the number of modes in the input space. The prior probability of class k ($k = 1, \ldots, K$) is denoted by π_k, where $\boldsymbol{\pi} \in R_+^K$, $\sum_{k=1}^{K} \pi_k = 1$, $Cat(\boldsymbol{\Pi})$ is the distribution controlled by the parameters $\boldsymbol{\Pi}$, and the prior is a uniform distribution of $p(c_k) = K^{-1}$. The distribution $p(\mathbf{z}) = \sum_{k=1}^{K} p_{\boldsymbol{\theta}_z}(\mathbf{z}|c_k)p(c_k)$ can be considered a Gaussian mixture distribution. And the means $\boldsymbol{\mu}_k(c_k; \boldsymbol{\theta}_z)$ and variances $\mathrm{diag}(\sigma_k^2(c_k; \boldsymbol{\theta}_z))$ of each category of the mixture can be given by the networks within the input data. During the training process, the parameters will be updated. Then, by sampling from the distribution $p_{\boldsymbol{\theta}_z}(\mathbf{z}|c_k)$, the latent variable $\mathbf{z}$ can be obtained, and the mean $\boldsymbol{\mu}_x(\mathbf{z}; \boldsymbol{\theta}_x)$ and variance $\mathrm{diag}(\sigma_x^2(\mathbf{z}; \boldsymbol{\theta}_x))$ are calculated by the networks.

With Equation (9), the logarithmic likelihood of the hybrid clustering variational autoencoder can be rederived as follows:

$$\begin{aligned} \log p(\mathbf{x}) &= \log \int_{\mathbf{z}} \sum_{\mathbf{c}} p(\mathbf{x}, \mathbf{z}, c_k) d\mathbf{z} \geq \\ E_{q_{\boldsymbol{\phi}}(\mathbf{z}, c_k|\mathbf{x})} &\left[\log \frac{p(\mathbf{x}, \mathbf{z}, c_k)}{q_{\boldsymbol{\phi}}(\mathbf{z}, c_k|\mathbf{x})}\right] = \mathcal{L}_{ELBO}(\mathbf{x}) \end{aligned} \tag{10}$$

where $\mathcal{L}_{ELBO}(\mathbf{x})$ is the evidence lower bound, and $q_{\boldsymbol{\phi}}(\mathbf{z}, c_k|\mathbf{x})$ is the encoding part of the networks, which is the same as the recognition model.

Furthermore, for the encoding part of the hybrid cluster variational autoencoder, the joint distribution $q_{\boldsymbol{\phi}}(\mathbf{z}, c_k|\mathbf{x})$ can be expressed as

$$q_{\boldsymbol{\phi}}(\mathbf{z}, c_k|\mathbf{x}) = q_{\boldsymbol{\phi}_z}(\mathbf{z}|\mathbf{x})q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x}) \tag{11}$$

The encoding model consists of two parts. $q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x})$ is obtained by the neural network with a Softmax classifier. For the distribution $q_{\boldsymbol{\phi}_z}(\mathbf{z}|\mathbf{x})$, the coding part of the variational autoencoder, whose outputs are the mean $\mu(\mathbf{x}; \boldsymbol{\phi}_z)$ and variance $\sigma^2(\mathbf{x}; \boldsymbol{\phi}_z)$, is used to obtain the latent variables; this process can be described by $\mu(\mathbf{x}; \boldsymbol{\phi}_z), \sigma^2(\mathbf{x}; \boldsymbol{\phi}_z)$. The model's description is illustrated in Figure 2.

$$q_{\boldsymbol{\phi}_z}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mu(\mathbf{x}; \boldsymbol{\phi}_z), \mathrm{diag}(\sigma^2(\mathbf{x}; \boldsymbol{\phi}_z))) \tag{12}$$
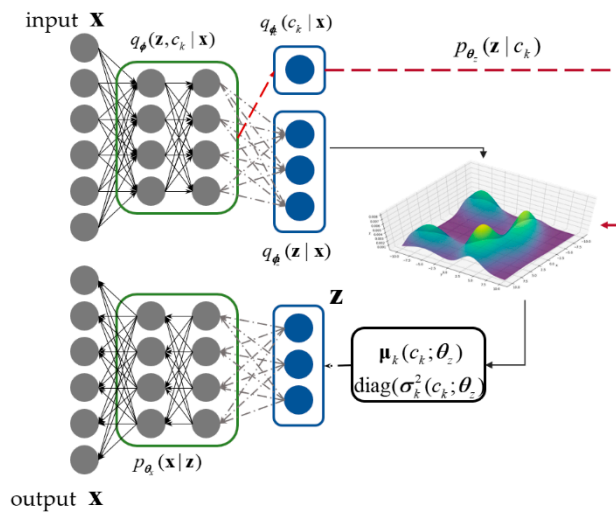
Figure 2. Schematic of the HCVAE model.

Based on Equations (8), (9) and (11), the logarithmic likelihood of the $p(\mathbf{x})$ can be rewritten as

$$
\begin{aligned}
\mathcal{L}_{ELBO}(\mathbf{x}) &= E_{q_{\boldsymbol{\phi}}(\mathbf{z},c_k|\mathbf{x})}[\log \frac{p(\mathbf{x},\mathbf{z},c_k)}{q_{\boldsymbol{\phi}}(\mathbf{z},c_k|\mathbf{x})}] \\
&= E_{q_{\boldsymbol{\phi}}(\mathbf{z},c_k|\mathbf{x})}[\log p(\mathbf{x},\mathbf{z},c_k) - \log q_{\boldsymbol{\phi}}(\mathbf{z},c_k|\mathbf{x})] \\
&= E_{q_{\boldsymbol{\phi}}(\mathbf{z},c_k|\mathbf{x})}[\log p_{\boldsymbol{\theta}_x}(\mathbf{x}|\mathbf{z}) + \log p_{\boldsymbol{\theta}_z}(\mathbf{z}|c_k) \\
&\quad + \log p(c_k) - \log q_{\boldsymbol{\phi}_z}(\mathbf{z}|\mathbf{x}) - \log q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x})]
\end{aligned}
\tag{13}
$$

Then, by employing the stochastic gradient variational estimation and the reparameterization technique, the equation for $\mathcal{L}_{ELBO}(\mathbf{x})$ can be improved as

$$
\begin{aligned}
\mathcal{L}_{ELBO}(\mathbf{x}) = \\
\frac{1}{L}\sum_l^L\sum_m^M x_m \log \boldsymbol{\mu}_x^{(l)}|_m + (1-x_m)\log(1-\boldsymbol{\mu}_x^{(l)}|_m) \\
-\frac{1}{2}\sum_k^K q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x})\sum_{d=1}^D (\log \sigma_k^2|_d + \frac{\sigma^2|_d}{\sigma_k^2|_d} + \frac{(\mu|_d - \mu_k|_d)^2}{\sigma_k^2|_d}) \\
+\sum_k^K q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x})\log\frac{\pi_k}{q_{\boldsymbol{\phi}_c}(c_k|\mathbf{x})} + \frac{1}{2}\sum_{d=1}^D (1+\log\sigma^2|_d)
\end{aligned}
\tag{14}
$$

Here, L is the number of Monte Carlo samples in the SGVB estimation, M is the dimension of data $\mathbf{x}$, x_m represents the m-th element of $\mathbf{x}$, D is the dimension of μ_k, σ_k^2, μ, σ^2, d represents the d-th element, and π_c is the prior probability of class c_k.

Once the network is constructed, an issue arises if the latent space is directly sampled and linked with the decoder as it outputs the mean and variance. This situation can lead to difficulties in the network's backpropagation process, hindering gradient propagation. To address this, a reparameterization trick is employed, which introduces a random noise variable to reparameterize the latent variable sampling process.

In the encoding part of the variational autoencoder, the mean and the variance can be obtained from the networks:

$$
\mu_z^{(l)}, \sigma_z^{2(l)} = f(z^{(l)}; \boldsymbol{\theta}_z)
\tag{15}
$$

where $z^{(l)}$ is the l-th sample obtained from Monte Carlo sampling of the distribution $q_{\boldsymbol{\phi}_z}(\mathbf{z}|\mathbf{x})$. By employing the reparameterization technique, the expression for $z^{(l)}$ can be given as follows:

$$
z^{(l)} = \mu_z^{(l)} + \sigma_z^{(l)} \circ \varepsilon^{(l)}
\tag{16}
$$

where $\varepsilon^{(l)}$ follows the standard normal distribution, and the mean $\pmb{\mu}_z^{(l)}$ and variance parameters $\sigma_z^{(l)}$ are obtained from the encoder.

Furthermore, when dealing with the evidence lower bound, the objective is to maximize it in order to optimize the parameters.

$$
\begin{aligned}
\mathcal{L}_{ELBO}(\mathbf{x}) &= E_{q_{\pmb{\phi}}(\mathbf{z},c_k|\mathbf{x})}[\log \frac{p(\mathbf{x},\mathbf{z},c_k)}{q_{\pmb{\phi}}(\mathbf{z},c_k|\mathbf{x})}] \\
&= \int_z \sum_c q_{\pmb{\phi}_c}(c_k|\mathbf{x})q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})[\log \frac{p_{\pmb{\theta}_x}(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})} + \log \frac{p(c_k|\mathbf{z})}{q_{\pmb{\phi}_c}(c_k|\mathbf{x})}]d\mathbf{z} \\
&= \int_z q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})\log \frac{p_{\pmb{\theta}_x}(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})}d\mathbf{z} \\
&\quad -q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})D_{KL}(q_{\pmb{\phi}_c}(c_k|\mathbf{x})||p(c_k|\mathbf{z}))d\mathbf{z}
\end{aligned}
\tag{17}
$$

In the equation, the first term is independent of *c*, and the second term is non-negative. Therefore, to maximize the evidence lower bound, it is necessary that $D_{KL}(q_{\pmb{\phi}_c}(\mathbf{c}|\mathbf{x})||p(c|\mathbf{z})) \equiv 0$. In this case, $q(c_k|\mathbf{x})$ can be expressed as follows:

$$
q_{\pmb{\phi}_c}(c_k|\mathbf{x}) = p(c_k|\mathbf{z}) \equiv \frac{p(c_k)p_{\pmb{\theta}_z}(\mathbf{z}|c_k)}{\sum\limits_{k=1}^{K} p(c_k)p_{\pmb{\theta}_z}(\mathbf{z}|c_k)}
\tag{18}
$$

The above formula reasonably addresses the problem of mode identification in input data.

When estimating the maximized ELBO, the potential representation $\mathbf{z}$ can be obtained through sampling. At this point, it becomes possible to determine which class the input sample belongs to by examining the distribution *q*. By analyzing the parameters of *q*, one can assign the input data to a specific class or cluster based on the highest probability. This allows for classification or clustering of the data based on the learned representations in the latent space.

3.2. Process Monitoring Method

The process monitoring approach proposed, grounded in the HCVAE model, is constructed within an unsupervised learning framework. During the model training phase, only normal data were employed. This training process iteratively refines the network model parameters by optimizing the variational lower bound. Concerning the data to be processed, the process data are translated into the latent variable space by learning the model $q_{\pmb{\phi}}(\mathbf{z}, c_k|\mathbf{x})$. Subsequently, the likelihood of the input data belonging to each category is computed using Formula (18), followed by sampling from the latent variable space. This sampled data are then processed through the decoder $p(\mathbf{x}, \mathbf{z}, c_k)$ to derive the fault monitoring indicators. This estimated density serves as a statistical index for process fault monitoring. By comparing the estimated density with control limits determined through Monte Carlo sampling, process faults can be detected.

Based on the generation process of the VAE model described earlier, the marginal probability density of the observed variable $\mathbf{x}$ can be calculated as

$$
p(\mathbf{x}) = \int_{\mathbf{z}} \sum_c p(\mathbf{x}, \mathbf{z}, c_k)d\mathbf{z}
\tag{19}
$$

For the normal variable $\mathbf{x}$ of the same mode, the probability of $p_{\theta}(\mathbf{x}|\mathbf{z})$ is concentrated in a region with a greater probability. To sample from this concentrated distribution, we can utilize importance sampling with a concentrated posterior probability distribution. By incorporating importance sampling, the formula can be rewritten as

$$
p(x) = E_{q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})}[\frac{p(\mathbf{x}, \mathbf{z}, c_k)}{q_{\pmb{\phi}_z}(\mathbf{z}|\mathbf{x})}]
\tag{20}
$$

It can subsequently be written as sampling:

$$p(\mathbf{x}) \approx \frac{1}{I} \sum_{i=1}^{I} \frac{p_{\boldsymbol{\theta}_x}(\mathbf{x}|\mathbf{z}_i) p_{\boldsymbol{\theta}_z}(\mathbf{z}_i|c_k) p(c_k)}{q_{\boldsymbol{\phi}_z}(\mathbf{z}_i|\mathbf{x})} \tag{21}$$

In this formulation, I is the number of importance samples, and z_i denotes the i-th sample obtained from the concentrated posterior distribution $q_{\boldsymbol{\phi}_z}(\mathbf{z}_i|\mathbf{x})$. This approximation allows us to estimate the marginal probability density of the observed variable $\mathbf{x}$ by leveraging importance sampling and considering the concentration of the posterior distribution.

Drawing upon the principles of probability and statistics, the smaller the probability $p(\mathbf{x})$, the higher the likelihood of a fault occurrence. Therefore, $p(\mathbf{x})$ can be utilized for fault detection. In the context of monitoring, it is crucial to establish control limits for detecting potential abnormal behavior. By defining confidence limits, we can determine the control limits of probabilities using the following integral [17]:

$$\int_{-\infty}^{h} p(\mathbf{x}) d\mathbf{x} = 1 - \alpha \tag{22}$$

In this equation, x_α represents the threshold value corresponding to the confidence level $1 - \alpha$. By solving this integral equation, we can determine the control limits of probabilities, enabling the identification of potential abnormal behavior in the monitored process.

As for the aforementioned formula, direct integration can be challenging, so we can approximate it by converting the integral into a Markov Chain Monte Carlo (MCMC) sampling method. In the context of the VAE mixture model, both the encoding process and decoding process can be viewed as a simple Markov chain. Through the MCMC sampling process, a large number of samples with complex distributions can be generated [18].

By utilizing the MCMC samples, we can estimate the control limit h through an approximate integral: obtain the S samples from $p(\mathbf{x}^{(j)})$ using MCMC sampling; calculate the likelihood $p(\mathbf{x}^j)$ of each sample $\mathbf{x}^j$; and sort the obtained results in descending order. The control limit can be obtained by $h = p(\mathbf{x}_l)$, where $l = S(1 - \alpha)$.

Since $p(\mathbf{x})$ is a non-negative number between 0 and 1, we can convert it into its negative logarithm to better visualize the monitoring results. This transformation allows us to express the monitoring index as follows:

$$Lp(\mathbf{x}) = -\log p(\mathbf{x}) \tag{23}$$

The control limit is redefined as $h = -\log p(\mathbf{x}_l)$. If $Lp(\mathbf{x}) > h$, the data $\mathbf{x}_n$ are considered faulty.

The process monitoring strategy for blast furnace ironmaking, which employs a hybrid clustering variational autoencoder for multimode data, consists of two distinct phases: offline modeling and online monitoring. The offline modeling phase involves the following steps:

Step 1: Collect a sufficient number of samples with modal labels to construct a database for model training.

Step 2: Standardize the data to ensure the data have zero mean and unit variance.

Step 3: Construct the network proposed in Section 3 and train it using the prepared dataset.

Here are the steps involved in constructing and training the network:

1. Network Architecture: Design the architecture of the hybrid clustering variational autoencoder model. This includes defining the number of layers, the size of each layer, and the activation functions to be used. The model should have separate encoding and decoding parts, with the clustering component integrated into the hidden layer.

2. Training Data Preparation: Split the collected and standardized dataset into training and validation sets. The training set is used to update the model parameters.
3. Validation and Hyperparameter Tuning: learning rate, batch size, and the number of clusters, through techniques like cross-validation or grid search.

Step 4: Compute the monitoring index, and the control limit is determined using Equation (22).

The online monitoring procedure consists of the following steps:

Step 1: Collect the monitored sample and standardize it.

Step 2: Input the data into the network and calculate the monitoring index using Equation (21) and Equation (23). If $Lp(\mathbf{x}) > h$, the data $\mathbf{x}_n$ are considered faulty; otherwise, it is a normal sample.

4. Application to a Real Blast Furnace

This section focuses on the analysis of actual production data collected from a blast furnace in China. The blast furnace under consideration has an inter volume of 2500 m^3. For the training process, 1500 samples of 10 process variables were considered, each collected at a sampling interval of 2 min. These variables include x_1 (quality of blast), x_2 (temperature of blast), x_3 (pressure of blast), x_4 (quantity of coal powder), x_5 (top pressure), x_6 (permeability index), x_7 (quantity of blast oxygen), x_8 (coke ratio), x_9 (CO concentration), and x_{10} (CO_2 concentration) [11]. These samples include process data from three different modes; each training sample is labeled with the corresponding mode, indicating the specific operating state of the blast furnace during that period. The modality number for each sample is listed in Table 1.

Table 1. The number of samples in each modality.

Modes	The Number of Samples
Mode 1	356
Mode 2	357
Mode 3	787

The test set consists of 1000 samples, which were also collected at a sampling interval of 2 min. It is worth noting that the fault occurs after the first 500 sample points in the test set. The abnormality is primarily attributed to the fluctuation of *CO* concentration (x_9) in the flue gas, which is caused by excessive coal powder (x_4) [19]. Figure 3 shows the time series of the 1000 samples.
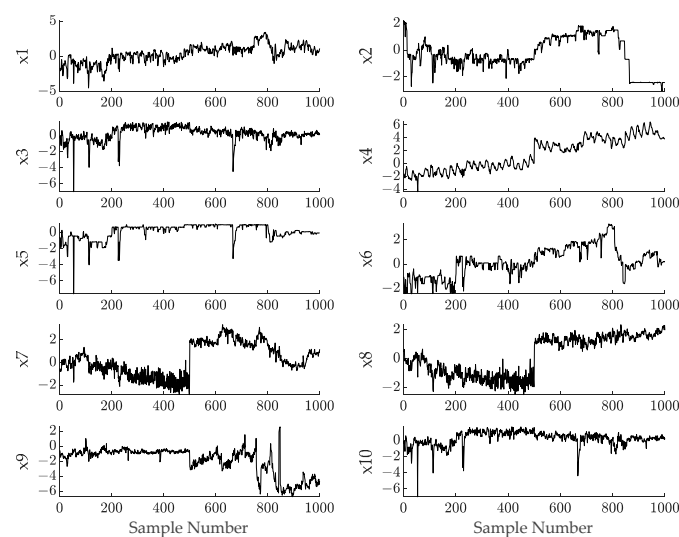

Figure 3. The test series data of 10 variables.

Figure 3 clearly shows a rapid numerical increase in variables x_7 and x_8, while the CO concentration (x_9) decreases in the last 500 points. However, an important observation is that the modal information of the data cannot be solely obtained from the multimode series data.

The proposed HCVAE model adopts a network structure consisting of five layers: an input layer, two intermediate layers, a hidden layer, and an output layer. The size of the intermediate layer is set to 26. The parameters of the latent variable in HCVAE were determined as three by 5-fold cross-validation, which minimizes the reconstruction error. During the model training process, the Adaptive Moment Estimation (ADAM) optimizer was employed with a learning rate of 0.001, the batch size was set to 32, and the activation function used in the model was the hyperbolic tangent (Tanh). After completing the model training process, the mode identification and fault monitoring were performed on a dataset that includes fault samples. The mode identification results are depicted in Figure 4, illustrating the successful identification of three modes. The results indicate that the data belonging to the same mode exhibit concentration and continuity, resembling the production process of blast furnace ironmaking, and the model effectively captures the mode transitions. These findings validate the reasonable and effective mode identification capability of the HCVAE model.
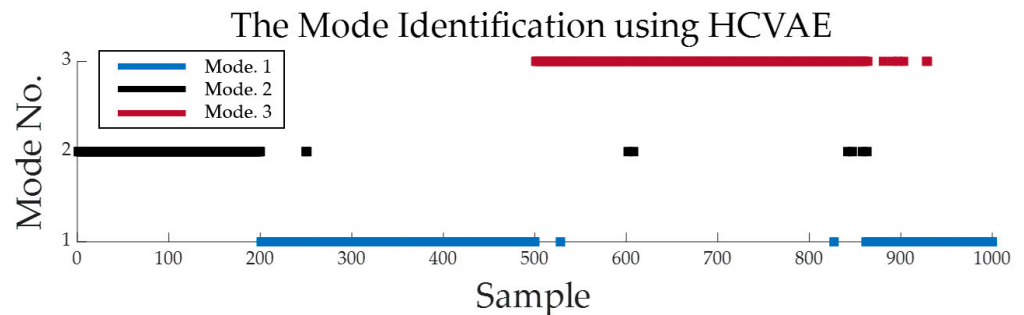

Figure 4. The mode identification using HCVAE for blast furnace data.

Notably, the mode identification results from the 500th to 1000th samples exhibit more pronounced anomalous identification compared with the normal operating state. This suggests that faults influence the data's process characteristics, consequently affecting the mode identification results. Regarding fault monitoring, a unified statistic for fault monitoring was constructed. Figure 5 illustrates the monitoring results with a 95% confidence limit. It can be observed that at the 500th sample point, there is a sudden increase in the statistic, and subsequent sample data consistently exceed the confidence limit, indicating a malfunction in the blast furnace ironmaking system during operation.
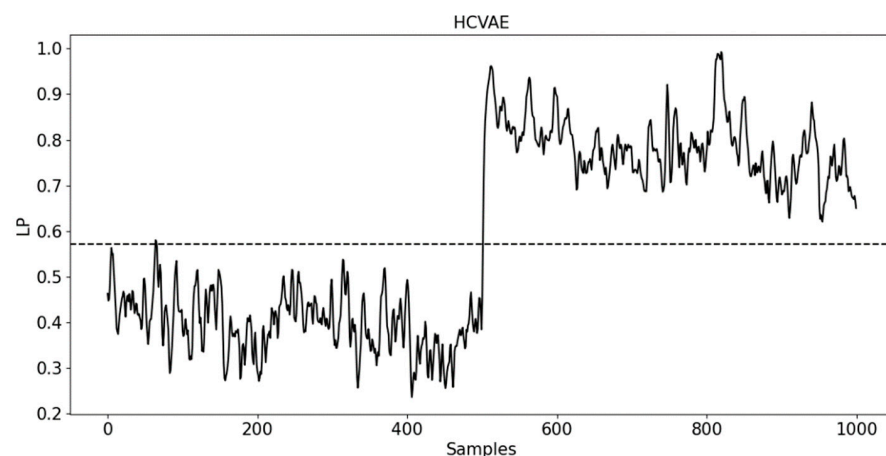

Figure 5. Monitoring results of blast furnace fault using HCVAE.

For comparison, the probabilistic principal component analysis based on the local nearest neighbor standardization (LNSPPCA) method proposed in reference [20] and the recursive probabilistic principal component analysis (RPPCA) method proposed in reference [21] was tested. The LNSPPCA method identifies the *K* nearest neighbors for each sample point in the dataset and forms a neighbor set. The mean and standard deviation of the neighbor set are used to standardize the current sample points, followed by fault monitoring using the PPCA method. For the blast furnace system, the number of principal components in the LNSPPCA method was set to three based on the results of a 5-fold cross-validation. The RPPCA method is grounded in the analysis of the singular values of the historical data matrix. It segments the entire process into distinct steady modes and mode transitions. As the process shifts from one mode to another, the RPPCA algorithm facilitates the recursive update of model parameters [21]. The number of principal components was set to three based on the results of a 10-fold cross-validation. The confidence level of the two methods was set to 95%. The fault monitoring results of LNSPPCA are shown in Figure 6, and the fault monitoring results of RPPCA are shown in Figure 7.
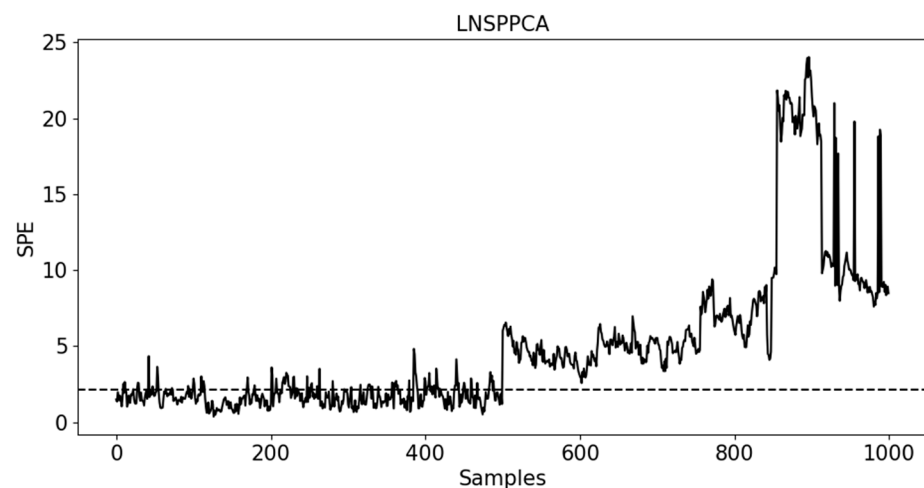

Figure 6. Monitoring results of blast furnace fault using LNSPPCA.
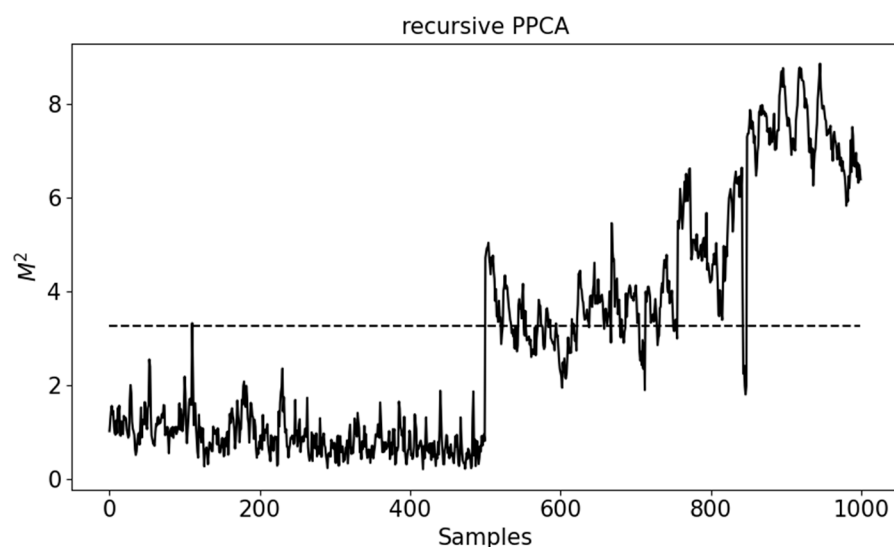

Figure 7. Monitoring results of blast furnace fault using RPPCA.

Comparing Figures 5–7, it is evident that the monitoring of HCVAE yields the best detection results. Although all the methods effectively identify faults after the 500th sample, aligning with the over-coal injection fault indicated by the dataset, a closer examination

of the statistical indicators reveals that LNSPPCA exhibits a higher rate of false positives, with a monitoring error rate of 10.5% at the 95% confidence level. Conversely, the HC-VAE method avoids missed detections and false positives. The proposed hybrid cluster variational autoencoder model outperforms the LNSPPCA method by considering the multimodality and nonlinearity of the blast furnace ironmaking system. The LNSPPCA method assumes a linear subspace for the underlying data distribution, which often leads to inadequate fault detection performance in industrial data. In addition, as shown in Figure 7, the RPPCA method has a low false alarm rate, with only one instance among all normal samples being erroneously identified as faulty. This outcome is similar to the HCVAE method, exhibiting a mere 0.2% false positive rate. However, the RPPCA approach proves ineffective when confronted with faulty samples, as it experiences a considerable number of missed detections. Specifically, the rate of missed detections, whereby faulty samples are incorrectly classified as normal ones, reaches as high as 19.2%. The efficacy of the RPPCA method is heavily reliant on the accuracy of its mode classification results, a dependency that could contribute to a significant error rate. This outcome renders the method unsuitable for practical real-world applications. The hybrid cluster variational autoencoder model employs a nonlinear dimensionality reduction method through neural networks for fault monitoring of multimodal blast furnace data. This fundamentally demonstrates the model's suitability for handling complex industrial process data. The experimental results validate the effectiveness of the proposed monitoring method and strategy in the fault monitoring of the complex multimodal blast furnace ironmaking process.

5. Conclusions

This paper presents a novel fault monitoring method based on a hybrid cluster variational autoencoder model, aiming to address the challenge of multimode process fault monitoring in blast furnace ironmaking. In contrast to traditional methods, we utilize a neural network to learn the data features. With changing production conditions, data from various modes frequently display distinct feature patterns. By harnessing the clustering process within the hidden layer, facilitated by the variational autoencoder, the method effectively monitors faults in multimode data. Based on the proposed model, we established a unified monitoring index and determined the calculation method for the control limit. The application of this method in blast furnace ironmaking demonstrates its effectiveness in identifying faults occurring between different modes in the multimode process. The approach properly accommodates the nonlinearity and temporal variability of the blast furnace ironmaking process, leading to better fault detection capability compared with the comparison method.

References

1. Li, J.; Hua, C.; Yang, Y.; Guan, X. Bayesian Block Structure Sparse Based T–S Fuzzy Modeling for Dynamic Prediction of Hot Metal Silicon Content in the Blast Furnace. *IEEE Trans. Ind. Electron.* **2017**, *65*, 4933–4942. [CrossRef]
2. Zhang, T.; Wang, W.; Ye, H.; Huang, D.; Zhang, H.; Li, M. Fault detection for ironmaking process based on stacked denoising autoencoders. In Proceedings of the 2016 American Control Conference (ACC), Boston, MA, USA, 6–8 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 3261–3267.
3. Yan, F.; Zhang, X.; Yang, C.; Hu, B.; Qian, W.; Song, Z. Data-driven modelling methods in sintering process: Current research status and perspectives. *Can. J. Chem. Eng.* **2022**, *101*, 4506–4522. [CrossRef]

4. Otsuka, K.; Matoba, Y.; Kajiwara, Y.; Kojima, M.; Yoshida, M. A hybrid expert system combined with a mathematical model for blast furnace operation. *ISIJ Int.* **1990**, *30*, 118–127. [CrossRef]

5. Vanhatalo, E. Multivariate process monitoring of an experimental blast furnace. *Qual. Reliab. Eng. Int.* **2010**, *26*, 495–508. [CrossRef]

6. Qian, J.; Song, Z.; Yao, Y.; Zhu, Z.; Zhang, X. A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes. *Chemometr. Intell. Lab. Syst.* **2022**, *231*, 104711. [CrossRef]

7. Saxén, H.; Lassus, L.; Seppänen, M.; Karjalahti, T. Pattern recognition and classification of blast furnace wall temperatures. *Ironmak. Steelmak.* **2000**, *27*, 207–211.

8. Zhang, T.; Ye, H.; Wang, W.; Zhang, H. Fault diagnosis for blast furnace ironmaking process based on two-stage principal component analysis. *ISIJ Int.* **2014**, *54*, 2334–2341.

9. Tian, H.; Wang, A. A novel fault diagnosis system for blast furnace based on support vector machine ensemble. *ISIJ Int.* **2010**, *50*, 738–742. [CrossRef]

10. Zhou, P.; Zhang, R.; Liang, M.; Fu, J.; Wang, H.; Chai, T. Fault identification for quality monitoring of molten iron in blast furnace ironmaking based on KPLS with improved contribution rate. *Control Eng. Pract.* **2020**, *97*, 104354. [CrossRef]

11. Liu, Y.; Zeng, J.; Bao, J.; Xie, L. A unified probabilistic monitoring framework for multimode processes based on probabilistic linear discriminant analysis. *IEEE Trans. Ind. Inform.* **2020**, *16*, 6291–6300. [CrossRef]

12. Zhou, B.; Ye, H.; Zhang, H.; Li, M. Process monitoring of iron-making process in a blast furnace with PCA-based methods. *Control Eng. Pract.* **2016**, *47*, 1–14. [CrossRef]

13. Wang, L.; Yang, C.; Sun, Y.; Zhang, H.; Li, M. Effective variable selection and moving window HMM-based approach for iron-making process monitoring. *J. Process Control* **2018**, *68*, 86–95. [CrossRef]

14. Zhou, P.; Zhang, R.; Xie, J.; Liu, J.; Wang, H.; Chai, T. Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: An integrated PCA-ICA method. *IEEE Trans. Ind. Electron.* **2020**, *68*, 622–631. [CrossRef]

15. Doersch, C. Tutorial on variational autoencoders. *arXiv* **2016**, arXiv:1606.05908. [CrossRef]

16. Chen, T.; Sun, Y. Probabilistic contribution analysis for statistical process monitoring: A missing variable approach. *Control Eng. Pract.* **2009**, *17*, 469–477. [CrossRef]

17. Yang, Y.; Ma, Y.; Song, B.; Shi, H. An aligned mixture probabilistic principal component analysis for fault detection of multimode chemical processes. *Chin. J. Chem. Eng.* **2015**, *23*, 1357–1363. [CrossRef]

18. Tang, P.; Peng, K.; Jiao, R. A process monitoring and fault isolation framework based on variational autoencoders and branch and bound method. *J. Frankl. Inst.* **2022**, *359*, 1667–1691. [CrossRef]

19. Cai, J.; Zeng, J.; Luo, S. A state space model for monitoring of the dynamic blast furnace system. *ISIJ Int.* **2012**, *52*, 2194–2199. [CrossRef]

20. Wang, K.; Forbes, M.G.; Gopaluni, B.; Chen, J.; Song, Z. Systematic development of a new variational autoencoder model based on uncertain data for monitoring nonlinear processes. *IEEE Access* **2019**, *7*, 22554–22565. [CrossRef]

21. Zhang, Z.; Peng, B.; Xie, L.; Peng, L. Process monitoring based on recursive probabilistic PCA for multi-mode process. *IFAC-Pap.* **2015**, *8*, 1294–1299.