A variational autoencoders approach for process monitoring and fault diagnosis

Peng Tang, Kaixiang Peng*, Jie Dong,
Kai Zhang and Ruihua Jiao

Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education,
School of Automation and Electrical Engineering,
University of Science and Technology Beijing,
Beijing, 100083, China
Email: gnepgnat@163.com
Email: kaixiang@ustb.edu.cn
Email: dongjie@ies.ustb.edu.cn
*Corresponding author

Abstract: Probabilistic models, which can model the process noise and can handle the problem of missing data in the probabilistic framework, recently have been got much attention in process monitoring and fault diagnosis area. This paper presents a new probabilistic methodology for fault detection and diagnosis in nonlinear processes using a variational autoencoders (VAEs) models. Two statistic index, based on the probability density distribution of measure variables and latent structure variable, are built to monitoring fault. Then a probabilistic contribution analysis method, based on the concept of missing variable estimation, is proposed for fault diagnosis. The performance of fault detection and diagnosis is demonstrated through its application for the monitoring of Tennessee Eastman (TE) industrial process, and the effectiveness is verified.

Keywords: VAE; variational autoencoder; process monitoring and fault diagnosis; probabilistic contribution analysis; nonlinear processes; TE.

Reference to this paper should be made as follows: Tang, P., Peng, K., Dong, J., Zhang, K. and Jiao, R. (2021) ‘A variational autoencoders approach for process monitoring and fault diagnosis’, *Int. J. System Control and Information Processing*, Vol. 3, No. 3, pp.229–245.

Biographical notes: Peng Tang received his BE in Automation from Changsha University of Science and Technology, Changsha, China, in 2013, and received his ME in Control Science and Engineering from North China University of Technology, Beijing, China, in 2016, respectively. He is currently working toward PhD at the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China. His research interests include industrial process monitoring, fault diagnosis, machine learning and big data analysis.

Kaixiang Peng received his BE in Automation and the ME and PhD in Control Theory and Control Engineering from the University of Science and Technology Beijing, Beijing, China, in 1995, 2002, and 2007, respectively. He is a Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His research interests include fault diagnosis and fault tolerant control, prognosis, and maintenance of complex industrial processes.

Jie Dong received his BE in Automation and the ME and PhD in Control Theory and Control Engineering from the University of Science and Technology Beijing, Beijing, China, in 1995, 1997, and 2007, respectively. She is currently an Associate Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing. From July to December 2004, she was a Visiting Scholar with the University of Manchester, Manchester, UK. Her research interests include intelligent control theory and application, PM-FD, and complex system modelling and control.

This paper is a revised and expanded version of a paper entitled 'A variational autoencoders approach for process monitoring and fault diagnosis' presented at the *30th Chinese Process Control Conference*, Kunming, China, 1 August, 2019.

1 Introduction

Industrial process monitoring and fault diagnosis is the essential measure to ensure process safety and product quality stability, and is widely used in industrial production. With the rapid development of industrial automation and informatisation, large-scale data has been accumulated in industrial production, which provide strong support for data-driven methods. Data-driven process monitoring methods are currently receiving considerably increasing attention both in application and in research domains, especially the one based on multivariate statistical process control (MSPC), such as principal component analysis (PCA) (Nomikos and MacGregor, 1994), partial least squares (PLS) (Muradore and Fiorini, 2011), etc. And some extended methods are studied to resolve the issues like nonlinear, non-Gaussian, dynamic in academic communities and applied in wide range of industrial applications for fault diagnosis (Peng et al., 2017, 2016; Yin et al., 2014; Choi et al., 2005; Yin et al., 2016; Ding et al., 2009; Zhou et al., 2016; Zhou and Qin, 2008; Chen et al., 2017). However, most traditional MSPC methods are lack of proper probabilistic mechanism for modelling process uncertainties. To solve this problem, some traditional MSPC methods have been extended to their probabilistic model, and are used for process monitoring. Probabilistic PCA (PPCA) is proposed by Tipping and Bishop (1999) and used for process monitoring by Kim and Lee (2003). Then the PPCA has been extended to the PPCA mixture model to deal with multimode data in industrial process monitoring (Ge and Song, 2010). A probabilistic kernel PCA method is proposed for nonlinear process monitoring (Ge and Song, 2010). To realise non-Gaussian process modelling and monitoring, ICA has been extended to probabilistic ICA (PICA) (Zhu et al., 2016).

In this paper, variational autoencoders (VAE) is introduced to into nonlinear process monitoring in the form of probability. VAE, proposed by Kingma and Welling (2013)), is a probabilistic generative model that combines variational inference with deep learning. As a special kind of autoencoders, VAE can reduce dimensions in a probabilistically sound way, and provide the reconstruction probability. Through the probability density distribution of measure variables and latent structure variables, the monitoring index can be constructed for process monitoring.

A further contribution of this paper is due to the fault diagnosis method. In Chen and Sun (2009)), the authors have developed a probabilistic contribution analysis method based on missing variable approach. Once a fault is detected, the monitoring index will be recalculated with on variable being missing. This will be repeated for all variables. The variable corresponding to the smallest recalculated index will be denoted as the risky variable. The proposed idea has been extended to PPCA mixture model for fault detection and diagnosis in multimode processes.

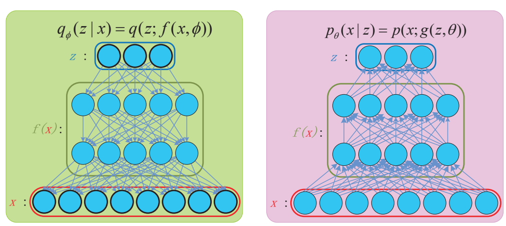
Motivated by the above mentioned works, a fault diagnosis based on VAE is introduced in this paper. Once a fault is detected, the monitoring index of each variables will be recalculated by probability estimate of missing variables.

The rest of this paper is organised as follows. In Section 2, the VAE is briefly introduced . Section 3 describes the proposed monitoring approach. In Section 3, the method of missing date estimate is inferred and the fault diagnosis approach is proposed. In Section 4, the study on the Tennessee Eastman (TE) benchmark case is provided to evaluate the efficiency of the proposed method. Finally, a summary of the paper is made.

1.1 Variational Autoencoders Analysis

The VAE is a directed probabilistic graphical model (DPGM) with certain types of latent variables, which forming an autoencoder-like architecture, as show in Figure 1.

Figure 1 Encoder and decoder of a directed probabilistic graphical model (see online version for colours)



The most essential of VAE is the evaluation of probability densities of all variables in the VAE model, which include $p(\boldsymbol{z})$, and $p(\boldsymbol{z}|\boldsymbol{x})$. The $\boldsymbol{z} \sim N(0, \mathbf{I})$ is already assumed. From a coding theory perspective, The generative model $p_\theta(\boldsymbol{x}|\boldsymbol{z})$ is the probabilistic decoder, where the data $\boldsymbol{x}$ is generated by the generative distribution $p_\theta(\boldsymbol{x}|\boldsymbol{z})$ conditioned on $\boldsymbol{z} : \boldsymbol{z} \sim p_\theta(\boldsymbol{z})$, $\boldsymbol{x} \sim p_\theta(\boldsymbol{x}|\boldsymbol{z})$. For the continuous value of $\boldsymbol{x}$, a typical choice for the parameterised distribution is to use a neural network where the input is $\boldsymbol{z}$ and the output is a Gaussian distribution over $\boldsymbol{x}$. Due to the complex nonlinearity of the neural network, $p_\theta(\boldsymbol{z}|\boldsymbol{x})$ is intractable. The recognition model $q_\phi(\boldsymbol{z}|\boldsymbol{x})$ is introduced as the probabilistic encoder to approximate the intractable true posterior $p_\theta(\boldsymbol{z}|\boldsymbol{x})$ by a neural network with $\boldsymbol{x}$ and $\boldsymbol{x}$ as its input and output, respectively. A variational inference method is introduced for learning the recognition model parameters ϕ jointly with the generative model parameters θ.

The Kullback-Leibler divergence (KL-divergence) between $p_\theta(\boldsymbol{z}|\boldsymbol{X})$ and $q_\phi(\boldsymbol{z})$ is defined as:

$$\begin{aligned}\mathrm{D}_{KL}[q_\phi(\boldsymbol{z}|\boldsymbol{X})||p_\theta(\boldsymbol{z}|\boldsymbol{X})] &= E_{\boldsymbol{z}\sim q}[\log q_\phi(\boldsymbol{z}|\boldsymbol{X}) - \log p_\theta(\boldsymbol{z}|\boldsymbol{X})] \\ &= E_{\boldsymbol{z}\sim q}[\log q_\phi(\boldsymbol{z}|\boldsymbol{X}) - \log p_\theta(\boldsymbol{X}|\boldsymbol{z}) \\ &\quad - \log p_\theta(\boldsymbol{z}) + \log p_\theta(\boldsymbol{X})]\end{aligned} \tag{1}$$

Because $\log p_\theta(\boldsymbol{X})$ does not depend on $\boldsymbol{z}$, then the formula can be rewritten as:

$$\begin{aligned}\log p_\theta(\boldsymbol{X}) - \mathrm{D}_{KL}[q_\phi(\boldsymbol{z}|\boldsymbol{X})||p_\theta(\boldsymbol{z}|\boldsymbol{X})] = E_{\boldsymbol{z}\sim q}[\log p_\theta(\boldsymbol{X}|\boldsymbol{z})] \\ - D_{KL}[q_\phi(\boldsymbol{z}|\boldsymbol{X})||p_\theta(\boldsymbol{z})]\end{aligned} \tag{2}$$

The left hand side is something what we want to maximise, which maximise $\log p_\theta(\boldsymbol{X})$, and minimise $\mathrm{D}_{KL}[q_\phi(\boldsymbol{z}|\boldsymbol{X})||p_\theta(\boldsymbol{z}|\boldsymbol{X})]$. The right hand side is something we can optimise via stochastic gradient descent.

The objective of VAE is to maximise the following variational lower bound with respect to the parameters θ and ϕ:

$$\mathcal{L}(\theta,\phi;\boldsymbol{x})) = -D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x})||p_\theta(\boldsymbol{z})) + \mathrm{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\log p_\theta(\boldsymbol{x}|\boldsymbol{z})] \tag{3}$$

Here we give the solution when both the prior $p_\theta(\boldsymbol{z}) \sim N(0,\mathbf{I})$ and the posterior approximation $q_\phi(\boldsymbol{z}|\boldsymbol{x})$ are Gaussian. Let J be the dimensionality of $\boldsymbol{z}$. Let μ and σ denote the variational mean an std. evaluated at datapoint $\boldsymbol{x}$, and Let μ_j and σ_j simply denote the jth element of these vectors. Then:

$$\begin{aligned}-D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x})||p_\theta(\boldsymbol{z})) &= \int q_\theta(\boldsymbol{z})(\log p_\theta(\boldsymbol{z}) - \log q_\theta(z))\mathrm{d}z \\ &= \frac{1}{2}\sum_{j=1}^{J}(1+\log((\sigma_j)^2) - (u_j)^2 - (\sigma_j)^2)\end{aligned} \tag{4}$$

Then the variational lower bound can be represented as:

$$\begin{aligned}\mathcal{L}(\theta,\phi;\boldsymbol{x}) \simeq &\frac{1}{2}\sum_{j=1}^{J}(1+\log((\sigma_j)^2) - (u_j)^2 - (\sigma_j)^2) \\ &+ \frac{1}{L}\sum_{l=1}^{L}\log p_\theta(\mathrm{x}|\mathrm{z}^l)\end{aligned} \tag{5}$$

where $\boldsymbol{z}^l = u + \sigma \odot \varepsilon$ and $\varepsilon^{(l)} \sim N(0,\mathbf{I})$.

The algorithm for training the VAE is shown in Algorithm 1.

Algorithm 1 Minibatch version of the Variational autoencoder training algorithm

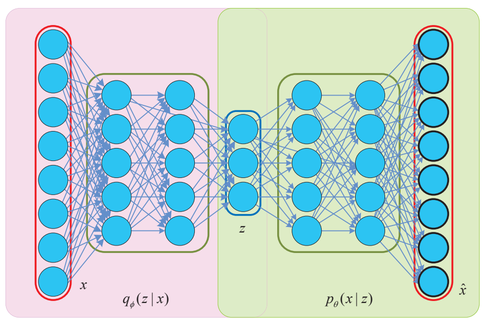
$\theta,\phi \leftarrow$ Initialise parameters
Repeat
  $\boldsymbol{X}^M \leftarrow$ Random minibatch of M datapoints (drawn from full dataset)
  $\varepsilon \leftarrow$ Random samples from noise distribution $p(\varepsilon)$
  $g \leftarrow \nabla_{\theta,\phi}\widetilde{\mathcal{L}}^M(\theta,\phi;\boldsymbol{X}^M,\varepsilon)$ (Gradients of minibatch estimator)
  $\theta,\phi \leftarrow$ Update parameters using gradients of g
Until convergence of parameters (θ,ϕ)
Return θ,ϕ

1.2 The VAE-based fault detection method

A fault detection method that use VAE is proposed to calculate the probability score of expectation $E(\boldsymbol{z}|\boldsymbol{x})$ in the prior distribution $p_\theta(\boldsymbol{z})$ and the probability score of observed data $\boldsymbol{x}$ in the probability distribution$p(\boldsymbol{x}|\hat{\boldsymbol{x}})$, which $\hat{\boldsymbol{x}}$ denotes as the reconstruction of $\boldsymbol{x}$. The calculation of $\boldsymbol{z}$ and $\hat{\boldsymbol{x}}$ is show in Figure 2.

The fault detection based VAE is constructed as an unsupervised learning framework, using only normal data to train the VAE. In the training process, the conditional probability distribution of $q_\phi(\boldsymbol{z}|\boldsymbol{x})$ in latent space and $p_\theta(\boldsymbol{z}|\boldsymbol{x})$ in original input space were learned by optimising the variational lower bound. In the detection process, the detection data is mapped to latent space by recognition model, and a number of samples are drawn from conditional probability distribution in latent space. For each sample, the reconstruction probability distribution is calculated by generative model. Finally, the anomaly score were calculated to predicate fault in latent space and original input space, respectively.

Figure 2 The calculation of z and $\hat{x}$ (see online version for colours)



1.2.1 Monitoring chart of latent variable

Since prior probability $\boldsymbol{z} \sim N(0, \mathbf{I})$ is assumed, the squared Mahalanobis norm of $\boldsymbol{z}$ follows $\chi^2(J)$ distribution, which J denotes as the dimensions of laten variable. Given test data $\boldsymbol{x}_i$, $\underline{\boldsymbol{z}}$ is the estimate of $\boldsymbol{z}|\boldsymbol{x}^i$:

$$\underline{\boldsymbol{z}} \equiv \mathrm{E}[\boldsymbol{z}|\boldsymbol{x}_i] = u_{\boldsymbol{z}^{(i)}} \tag{6}$$

where μ_{z^i}, σ_{z^i} is the mean and s.d., calculated by recognition model $q_\phi(\boldsymbol{z}|\boldsymbol{x})$.

Then Hotelling's T^2 test statistic is denoted as:

$$T^2 = \|\underline{\boldsymbol{z}}\|^2 \tag{7}$$

The thresholds for the fault detection in latent space can be determined with a given confidence level as follows:

$$J_{th,T^2} = \chi^2_\alpha(J) \tag{8}$$

where $\chi^2_\alpha(J)$ denotes the χ^2-distribution with J degrees of freedom and α is user-specified significance level.

We regard the process is normal when T^2 fall into in-control with α LOS:

$$T^2 = \|\underline{\boldsymbol{z}}\|^2 \leq \chi^2_\alpha(J) \tag{9}$$

1.2.2 Monitoring chart of original input space

Given test data $\boldsymbol{x}_i$, then the posterior probability distribution $\boldsymbol{z}|\boldsymbol{x}^{(i)} \sim N(u_{\boldsymbol{z}^{(i)}}, \sigma_{\boldsymbol{z}^{(i)}})$ can be calculated.

Then L samples is drawn from $N(u_{z^{(i)}}, \sigma_{z^{(i)}})$. For each sample $\boldsymbol{z}^{(i,j)}$, the probability distribution $\widehat{\boldsymbol{x}}^{(i,l)}|\mathrm{z}^{(i,l)} \sim N(u_{\widehat{\boldsymbol{x}}^{(i,l)}}, \sigma_{\widehat{\boldsymbol{x}}^{(i,l)}})$, where $u_{\widehat{x}^{(i,l)}}$, $\sigma_{\widehat{x}^{(i,l)}}$ is the mean and s.d., calculated by generative model $p_\theta(\boldsymbol{z}|\boldsymbol{x})$.

Then input variable $\boldsymbol{x}_i$ can be transformed to the standard Gaussian via whitening:

$$\boldsymbol{x}_{(i,l)}^{w} = \frac{\boldsymbol{x}_i - u_{\widehat{\boldsymbol{x}}^{(i,l)}}}{\sigma_{\widehat{x}^{(i,l)}}} \tag{10}$$

And $\boldsymbol{x}_{(i,l)}^{w} \sim N(0, \mathbf{I})$.

SPE (squared prediction error) test statistic denotes as:

$$SPE = \left\| \frac{1}{L} \sum_{l=1}^{L} \boldsymbol{x}_{(i,l)}^{w} \right\|^2 \tag{11}$$

The thresholds for the fault detection in original input space can be determined with a given confidence level as follows:

$$J_{th,SPE} = \chi_{\alpha}^{2}(n) \tag{12}$$

where $\chi_{\alpha}^{2}(n)$ denotes the χ^2-distribution with n degrees of freedom and α is user-specified significance level.

Therefore, the detection logic is:

$$\begin{cases} SPE \le J_{th,SPE} \mathrm{and} \mathrm{T}^2 \le J_{th,T^2}, \mathrm{fault-free} \\ SPE > J_{th,SPE} \mathrm{or} \mathrm{T}^2 > J_{th,T^2}, \mathrm{faulty}. \end{cases} \tag{13}$$

The calculate process of VAE-based fault detection algorithm is shown in Algorithm 2.

Algorithm 2 Variational autoencoder based fault detection

$\theta, \phi \leftarrow$ train a variational autoencoder using the normal dataset $\boldsymbol{X}$
Input: sample data $\boldsymbol{x}^{(i)}$
$u_{\boldsymbol{z}^{(i)}}, \sigma_{\boldsymbol{z}^{(i)}} = f_\theta(z|x^{(i)})$
$T^2 = \|u_{\boldsymbol{z}^{(i)}}\|^2$
$\boldsymbol{z}^{(i,l)} \leftarrow$ Draw L samples from $\boldsymbol{z}|\boldsymbol{x}^{(i)} \sim N(u_{\boldsymbol{z}^{(i)}}, \sigma_{\boldsymbol{z}^{(i)}})$
for l=1 to L do
  $u_{\widehat{\boldsymbol{x}}^{(i,l)}}, \sigma_{\widehat{\boldsymbol{x}}^{(i,l)}} = g_\phi(\boldsymbol{x}|\boldsymbol{z}^{(i,l)})$
end for
$SPE = \left\| \frac{1}{L} \sum_{l=1}^{L} \frac{\boldsymbol{x} - u_{\widehat{\boldsymbol{x}}^{(i,l)}}}{\sigma_{\widehat{\boldsymbol{x}}^{(i,l)}}} \right\|^2$
if $SPE \le J_{th,SPE} \mathrm{and} \mathrm{T}^2 \le J_{th,T^2}$ **then**
  $\boldsymbol{x}^{(i)}$ is an anomaly
else
  $\boldsymbol{x}^{(i)}$ is an faulty
end for

2 The VAE-based fault diagnosis method

Once the fault was successfully detected by the two monitoring statistics, the root cause of the fault should be find by fault diagnosis. A missing variable estimation-based contribution analysis method is proposed for fault diagnosis. The objective of contribution analysis is to identify which variables are the most responsible for the occurrence of the process fault. In general contribution analysis may not explicitly reveal the root-cause of the onset of faults, but it is undoubtedly helpful in pinpointing the inconsistent variables that should undergo further diagnosis procedures. Assume that a fault is caused by the change of one variable, if the variable is changed by the estimation of other variables, the fault should disappear. Probabilistic model is used to estimate one variable by other variables, then the estimated variable is used to compute SPE and T^2 test statistic. The estimated variable, which caused fault, will reduce the SPE and T^2 test statistic. In the diagnosis process, the key variables can be observed from the chart which original signal test statistic subtracted reconstruction signal test statistic.

For n dimensional data $\boldsymbol{x}$, each time one variable is regard as missing. Let $\boldsymbol{x}_d(d = 1, 2, \ldots, n)$ be the missing variable, and $\boldsymbol{x}_{-d}$ denotes as the vector of other observed variables in $\boldsymbol{x}$. For $d = 1 : n$, reconstructed variable $\tilde{\boldsymbol{x}}_d$ can be calculated by the expected value of missing variable $E_{(\boldsymbol{x}_{-d})}(\boldsymbol{x}_d)$ given the other variables. Then the monitoring statistic SPE_d and T_d^2 with $\tilde{\boldsymbol{x}}_d$ is re-calculated. If the dth variable contributes significantly to the data being detected as faulty, then the re-calculated statistic will be much small than the original monitoring statistic. Therefore, the difference between re-calculated statistic and original statistic can measure the impact of missing variable on the fault.

The contribution value of each variable in this work is defined as:

$$\begin{aligned} RBC_d^{T^2} &= T^2 - T_d^2 \\ RBC_d^{SPE} &= SPE - SPE_d \end{aligned} \tag{14}$$

The rest of this section will discusses how to estimate the missing variable.

The computation for $p(\boldsymbol{x}_d|\boldsymbol{x}_{-d})$ requires the marginalisation of the latent variables $\boldsymbol{z}$ from the joint distribution $p(\boldsymbol{x}_d, z|\boldsymbol{x}_{-d})$.

$$\begin{aligned} p(\boldsymbol{x}_d|\boldsymbol{x}_{-d}) &= \int p(\boldsymbol{x}_d, \boldsymbol{z}|\boldsymbol{x}_{-d})\mathrm{d}z \\ &= \int p(\boldsymbol{x}_d|z)p(\boldsymbol{z}|\boldsymbol{x}_{-d})\mathrm{d}z \end{aligned} \tag{15}$$

Using Bayes formula, $p(\boldsymbol{z}|\boldsymbol{x}_{-d})$ can be formulated as:

$$p(\boldsymbol{z}|\boldsymbol{x}_{-d}) = \frac{p(\boldsymbol{x}_{-d}|\boldsymbol{z})p(\boldsymbol{z})}{\int p(\boldsymbol{x}_{-d}|\boldsymbol{z})p(\boldsymbol{z})\mathrm{d}z} \tag{16}$$

The conditional distribution $p(\boldsymbol{x}_d|\boldsymbol{x}_{-d})$ can be re-written as:

$$\begin{aligned} p(\boldsymbol{x}_d|\boldsymbol{x}_{-d}) &= \int p(\boldsymbol{x}_d|\boldsymbol{z})p(\boldsymbol{z}|\boldsymbol{x}_{-d})\mathrm{d}z \\ &= \frac{\int p(\boldsymbol{x}_d|\boldsymbol{z})p(\boldsymbol{x}_{-d}|\boldsymbol{z})p(\boldsymbol{z})\mathrm{d}z}{\int p(\boldsymbol{x}_{-d}|\boldsymbol{z})p(\boldsymbol{z})\mathrm{d}z} \end{aligned} \tag{17}$$

Then the expectation of $\boldsymbol{x}_d$ can be expressed as:

$$
\begin{aligned}
E_{(\boldsymbol{x}_{-d})}(\boldsymbol{x}_d) &= \int \boldsymbol{x}_d p(\boldsymbol{x}_d|\boldsymbol{x}_{-d}) d\boldsymbol{x}_d \\
&= \frac{\int\int \boldsymbol{x}_d p(\boldsymbol{x}_d|\boldsymbol{z}) p(\boldsymbol{x}_{-d}|\boldsymbol{z}) dx_d \mathrm{d}z}{\int P(\boldsymbol{x}_{-i}|\boldsymbol{z}) P(\boldsymbol{z}) \mathrm{d}z} \\
&= \frac{\int u_{\boldsymbol{x}_d|\boldsymbol{z}} p(\boldsymbol{x}_{-d}|\boldsymbol{z}) p(\boldsymbol{z}) \mathrm{d}z}{\int p(\boldsymbol{x}_{-d}|\boldsymbol{z}) p(\boldsymbol{z}) \mathrm{d}z} \\
&= \frac{\sum_{k=1}^{K} u_{\boldsymbol{x}_d|\boldsymbol{z}_k} p(\boldsymbol{x}_{-d}|\boldsymbol{z}_k)}{\sum_{k=1}^{K} p(\boldsymbol{x}_{-d}|\boldsymbol{z}_k)}
\end{aligned}
\tag{18}
$$

where $\mu_{\boldsymbol{x}_d|\boldsymbol{z}}$ is the mean of $p(\boldsymbol{x}_d|\boldsymbol{z})$. Because of the assumption that all variables in data $\boldsymbol{x}$ are conditionally independent given $\boldsymbol{z}$, the conditional distribution $p(\boldsymbol{x}_{-d}|\boldsymbol{z}_k)$ and $p(\boldsymbol{x}_d|\boldsymbol{z})$ can be calculated by the generative model $p_\theta(\boldsymbol{x}|\boldsymbol{z})$. And the samples of $\boldsymbol{z}$ can be drawn from the recognition model $q_\phi(\boldsymbol{z}|\boldsymbol{x})$. To improve estimation accuracy of variable $\boldsymbol{x}_d$, a k-step iterative method is used in the estimation of missing variable, which the estimated variable replaces the original variable $\boldsymbol{x}_d$ to update the calculated value of missing variable by k times.

3 Case study

In this section, fault detection and diagnosis based on the proposed VAE method is demonstrated on the Tennessee Eastman benchmark problem (Downs and Vogel, 1993)).

Figure 3 illustrates the flowchart of the TE process with the plant-wide control structure in ref (Nomikos and MacGregor, 1995)). The model is widely accepted as a challenging benchmark for control and monitoring studies. Five major units, i.e., reactor, condenser, separator, compressor and stripper, constitute the whole process. The process has 41 measurements and 12 manipulated variables. The measurements include 22 continuous process measurements and 19 sampled process measurements. Totally 21 different faults has been designed and the dataset used in this paper are given in Chiang and Russell (2000)) and is widely accepted for process monitoring and fault diagnosis, which can be downloaded from http://web.mit.edu/braatzgroup/links.html. The dataset includes 22 training sets and 22 testing sets. Except that one testing set was obtained under normal operational condition, the other 21 sets were collected under 21 different faulty conditions for 48 operation hours and 960 samples are obtained for each testing set. For each of 21 faulty testing sets, the fault was introduced in at 8th operation hour (161th sample). Similarly, the training sets are also composed of one normal operational set and 21 faulty operational sets. These 21 process faults include 7 step faults, 5 random faults, 3 sticking and slow change fault, and 6 unknown process faults.

3.1 Simulation and analysis

In this paper, a total of 33 process measurements are selected as measured variables, listed in Table 1. The testing set and training set collected under the normal situation are used as training dataset. The 21 faulty testing sets are used to validate the performances of the models.

Figure 3 Flowchart of the TE process

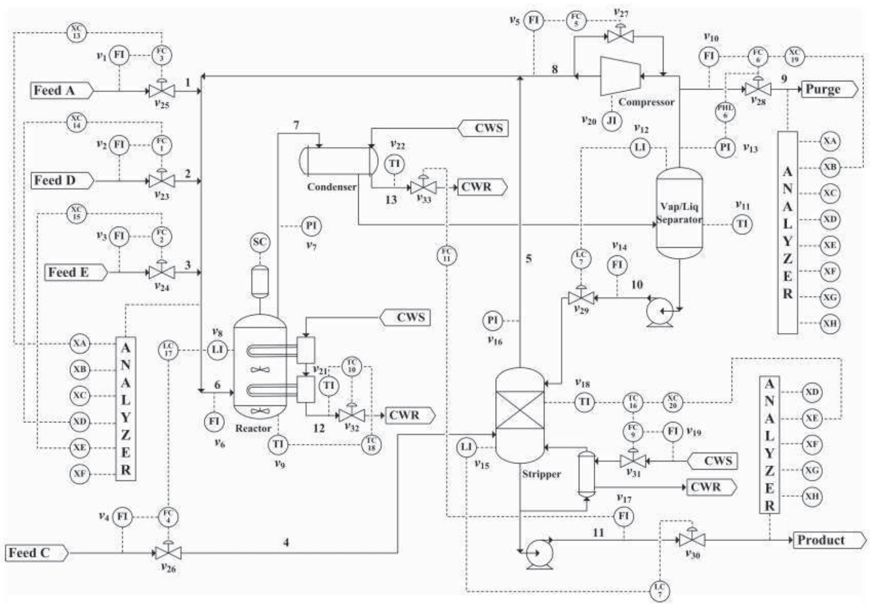
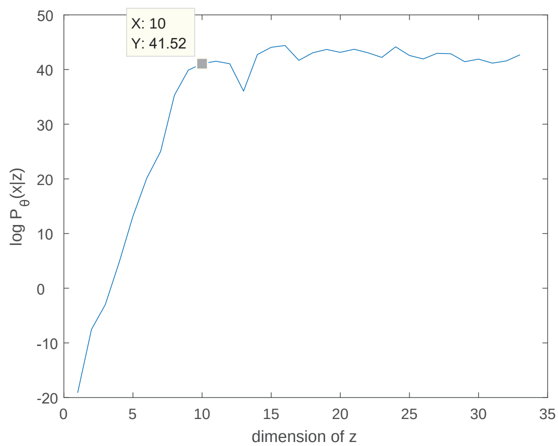


Table 1 Monitored variables in the Tennessee Eastman process

ID	*Variable description*	*ID*	*Variable description*
x1	A feed (Stream 1)	x18	Stripper temperature
x2	D feed (Stream 2)	x19	Stripper steam flow
x3	E feed (Stream 3)	x20	Compressor work
x4	A and C feed (Stream 4)	x21	Reactor cooling water outlet temperature
x5	Recycle flow (Stream 8)	x22	Separator cooling water outlet temperature
x6	Reactor feed rate (Stream 6)	x23	MV to D feed flow (Stream 2)
x7	Reactor pressure	x24	MV to E feed flow (Stream 3)
x8	Reactor level	x25	MV to A feed flow (Stream 1)
x9	Reactor temperature	x26	MV to total feed flow (Stream 4)
x10	Purge rate (Stream 9)	x27	Compressor recycle valve
x11	Product separator temperature	x28	Purge valve (Stream 9)
x12	Product separator level	x29	Separator pot liquid flow (Stream 10)
x13	Product separator pressure	x30	Stripper liquid product flow (Stream 11)
x14	Product separator underflow (Stream 10)	x31	Stripper steam valve
x15	Stripper level	x32	Reactor cooling water flow
x16	Stripper pressure	x33	Condenser cooling water flow
x17	Stripper underflow (Stream 11)		

The detection process of proposed VAE method is applied to monitor the TE process in comparison with the traditional PPCA-based methods. Both the generative model and the recognition model are set to 3-tier structure, which have one hidden layer, and the size of hidden layer is 50. The dimension of latent space is determined by 5-fold cross validation. The relationship between $\mathrm{E}_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\log p_\theta(\boldsymbol{x}|\boldsymbol{z})]$ and the dimension of $\boldsymbol{z}$ is shown in Figure 4. Considering the complexity and performance of model, 10 is selected for the size of latent variables. In order to be fair, the component numbers of PPCA models is also selected as 10. The confidence level of thresholds of SPE and T^2 is set as 99%. Tow types of fault are enumerated to demonstrate the efficiency of VAE.

Figure 4 The relationship between and the dimension of latent space (see online version for colours)



Fault 5 is a step change in the internal temperature of condenser cooling water, and it causes a step change in the flow rate of the outlet stream of the condenser, which can affect the temperature in the separator as well as the cooling water outlet temperature of the separator. The monitoring results of VAE and PPCA are show in Figure 5. The fault was detected by PPCA at the 161th sample, but it was soon compensated by the control system and its SPE test statistic went below the confidence limit at the 394th sample. The time that fault was detected by VAE is 161th. And its test statistic SPE was still greater than the confidence limit after the 394th sample. In Figure 6, the fault kept affecting variable 33. The other variables were pulled back to their normal values after 394th sample with system compensation. The low value of variable 33 given a low detection rate of PPCA. In contrast, the method of VAE can detect the low change of variable 33.

Fault 10 is fault where a random variation is introduced in C feed temperature, which may cause a change to the condition of stripper and condenser. The monitoring results of VAE and PPCA are show in Figure 7. Fault 10 can be detected by PPCA with a low detection rate (34.1% for T^2 and 39.1% for SPE). The monitoring result of VAE is show in Figure 7(a) and (b). It is obvious that the SPE statistic is much better than PPCA. The fault detection rate of VAE is 82.4% for T2 and 32.0% for SPE. Part variables are shown in Figure 8. It can be seen that most variables changed slightly between 350th sample and 600th sample. It cause the low detection rate of PPCA. But the fault can still reflected in SPE statistic of VAE.

Figure 5 Monitoring result of VAE (a) (b) and PPCA (c) (d) for fault 5 (see online version for colours)

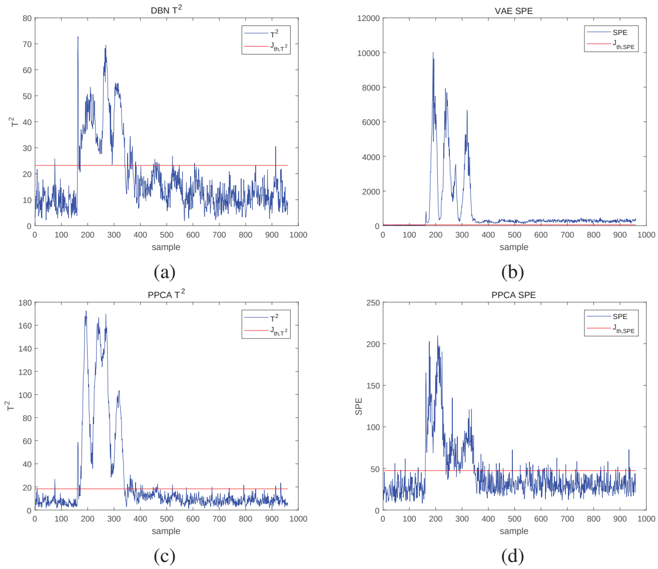
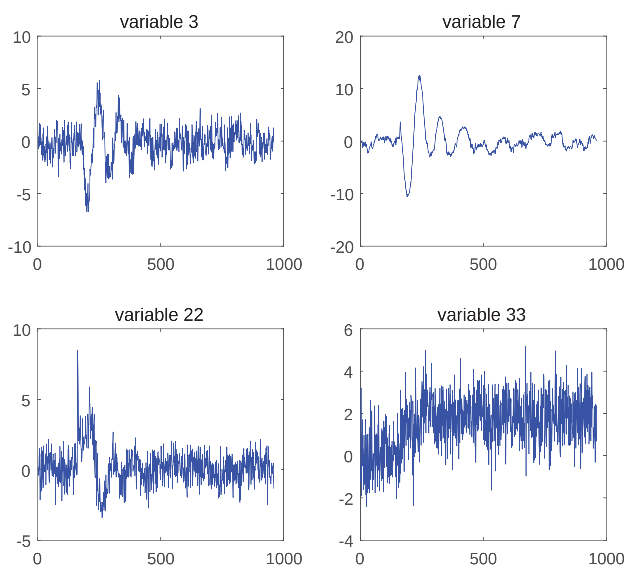


Figure 6 Process variables in fault 5 (see online version for colours)



All monitoring results (FAR and FDR) on 21 types of fault by VAE and PPCA are listed in Table 2. Faults 3, 9, 15 are very subtle faults that are hard to detect. Both PPCA and VAE are failed to detect these faults. Except these faults, the detection rates of proposed method are over 50%. Especially, the detection rates of VAE for fault 5, 10, 16, 19 and 20 are much higher than PPCA.

Figure 7 Monitoring result of VAE (a) (b) and PPCA (c) (d) for fault 10 (see online version for colours)

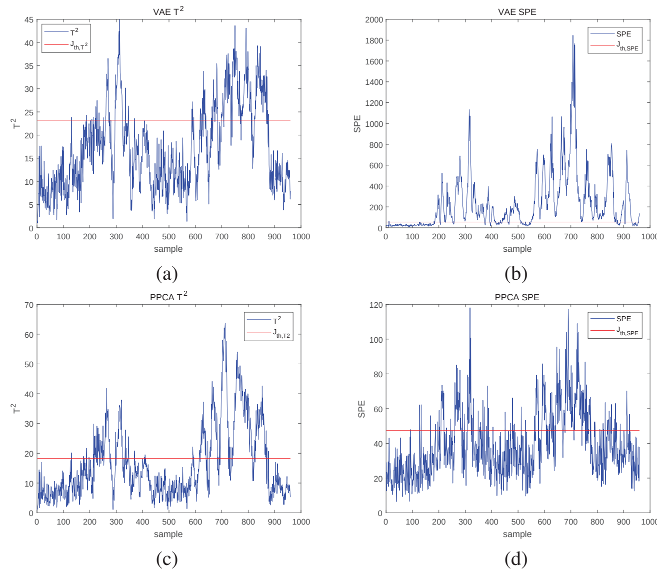
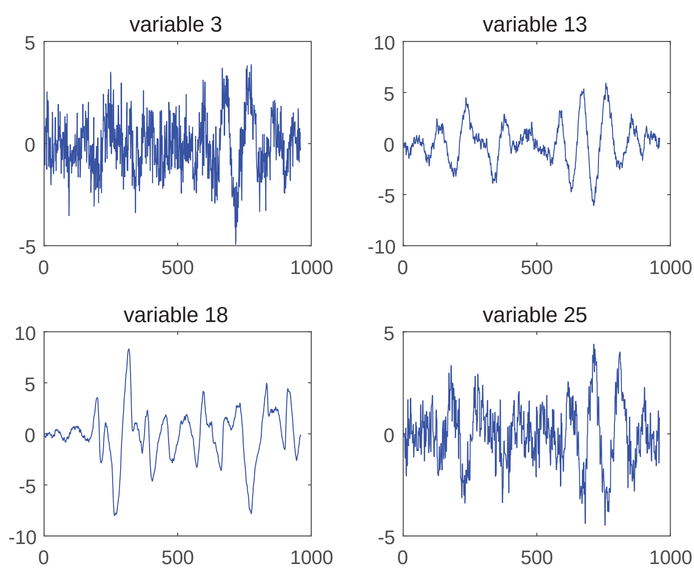


Figure 8 Process variables in fault 10 (see online version for colours)



Fault 2 is a step change in the B composition and the A/C ratio constant (stream 4). B is the inert component, which will cause a change in the purge rate, i.e., variable No. 8 (Purge rate) in stream 9. Furthermore, the A/C ratio constant will be reflected in stream 4, and the action of some control loops will also change other process variables. Figure 9 shows part of variables (Original variables and reconstruction variables) for fault 2. The original variables are represented by bule lines, and the reconstruction variables are represented by red lines. Fault 2 were triggered after the 160th sample. Before the fault occurred, reconstruction variables can flow original variables well. While original variables ware influenced by fault, reconstruction variables can reduce the influence by estimating with other variables. And reconstruction variables can also reduce signal nosie. Therefore, the approach, that missing variables probability estimate of by VAE, is reasonable to fault reconstruct. Figure 10 shows the result of fault detection and diagnosis for fault 2. NO. 10 (Purge rate (Stream 9)) and NO. 28 (Purge valve (Stream 9)) are identified as the primary variables.

Table 2 FDRs and FASs of the 21 faults in the TE benchmark

	VAE(%)				*PPCA(%)*			
Fault no.	*FAR* (*SPE*)	*FDR* (*SPE*)	*FAR* (T^2)	FDR (T^2)	*FAR* (*SPE*)	*FDR* (*SPE*)	*FAR* (T^2)	*FDR* (T^2)
1	0.6	99.8	0.0	99.8	2.5	99.9	0.6	99.1
2	0.0	98.4	0.6	98.6	0.6	98.1	1.3	98.5
3	3.8	5.4	0.6	1.5	1.9	5.0	0.6	3.0
4	0.6	98.5	0.6	2.3	3.1	100.0	1.9	10.6
5	0.6	100.0	0.6	25.0	3.1	28.6	1.9	25.6
6	0.6	100.0	0.0	96.5	1.3	100.0	0.6	99.4
7	0.0	100.0	0.0	35.9	1.9	100.0	0.0	100.0
8	0.0	97.8	0.0	97.8	1.3	96.5	0.0	96.9
9	8.8	3.4	4.4	1.1	5.6	4.0	5.0	2.9
10	0.6	82.4	0.6	32.0	1.9	34.1	0.6	39.1
11	0.0	73.5	1.3	4.5	3.8	79.9	0.6	27.1
12	1.9	99.6	1.3	97.5	2.5	96.4	0.6	98.3
13	0.0	95.1	0.0	94.6	1.3	95.3	0.0	93.9
14	0.6	100.0	0.0	56.4	2.5	100.0	1.3	85.1
15	0.6	9.0	0.6	3.1	2.5	5.0	0.6	3.8
16	11.9	86.0	3.1	16.0	4.4	34.5	7.5	21.5
17	0.6	95.9	0.0	64.5	5.0	95.0	0.0	76.5
18	0.6	90.0	0.0	87.9	4.4	90.1	0.0	89.3
19	0.0	80.6	0.0	0.5	1.3	48.9	0.0	1.3
20	0.0	71.4	0.0	54.8	3.1	54.3	0.0	38.0
21	3.1	56.0	0.6	51.1	5.0	52.9	0.0	33.1

Fault 2 and 11 are used to test the performance of fault diagnosis, and show in Table 3.

Table 3 Descriptions of process faults

Fault no.	*Process variables*	*Type*
2	B composition,A/Crationconstant	Step
11	Reactor cooling water inlet temperature	variation

Fault 11 is the fault where a random variation is introduced in the reactor cooling water inlet temperature, and the temperature of the reactor fluctuates. And it will cause NO. 9 (Reactor temperature) and No. 32 (Reactor cooling water flow) fluctuate sharply. Figure 11 shows the changed tend of two variables. The reconstruction variables eliminated fluctuate during fault period. The fault detection and diagnosis result are show in Figure 12. Fault

11 can be detected by statistic SPE after the 167th sample, as shown in Figure 12(a). So the SPE contribution plot is used to identify fault variables. Figure 12(b) shows No. 9 and No. 32 are the primary variables, in conformity with mechanism analysis of fault 11.

Figure 9 Original and reconstruction variables for fault 2 (see online version for colours)

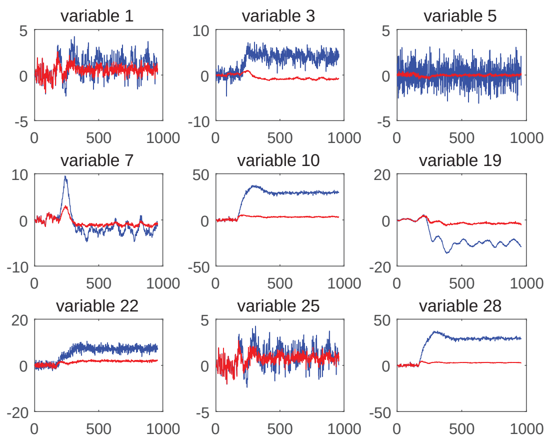
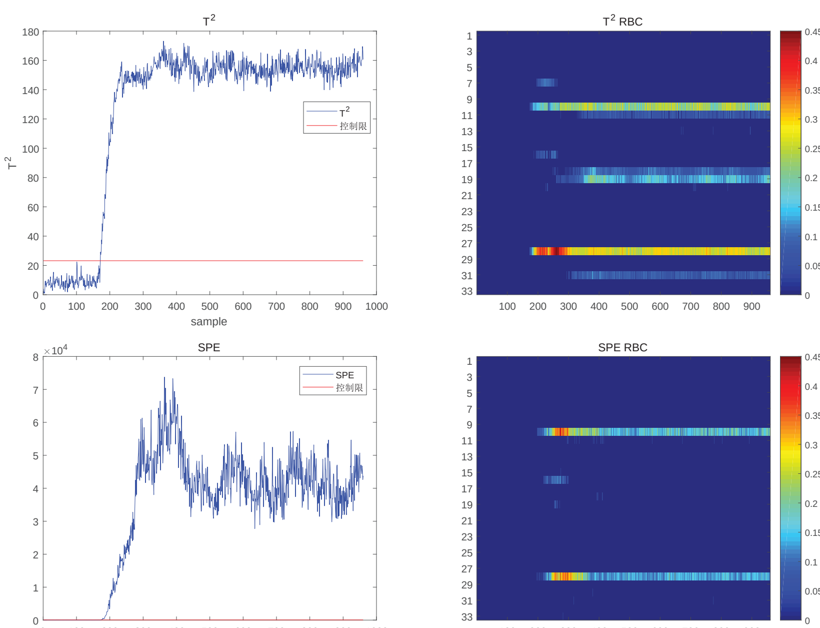


Figure 10 Fault detection and diagnosis for fault 2 based on VAE: (a) fault detection result and (b) fault diagnosis result (see online version for colours)



(a) (b)

Figure 11 Original and reconstruction variables of no. 9 and no. 32 for fault 11 (see online version for colours)

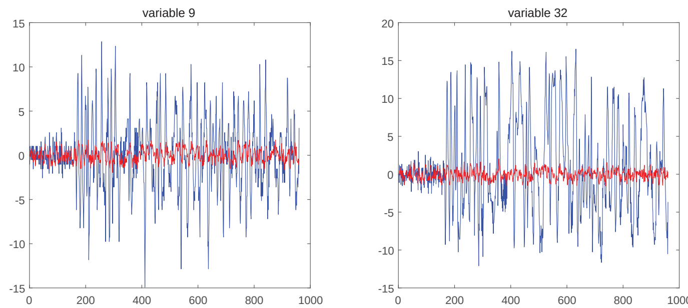
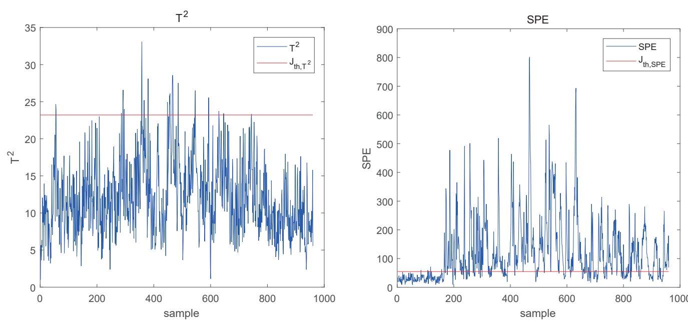
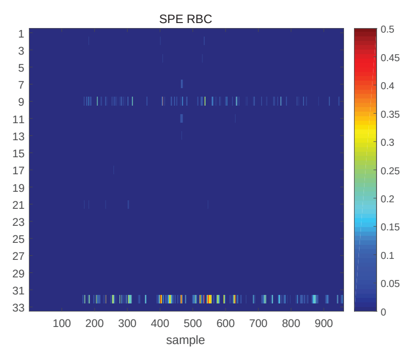


Figure 12 Fault detection and diagnosis for fault 11 based on VAE: (a) fault detection result and (b) fault diagnosis result (see online version for colours)



(a)



(b)

4 Conclusion

VAE, one kind of deep learning algorithm, is used in process monitoring and fault diagnosis. In monitoring process, the probability distribution of original space and latent space are calculated, then test statistics on the two space are designed to monitoring fault. In diagnosis process, a missing variable based contribution analysis methodology is proposed. The probabilistic framework of VAE provides a natural way of handling the missing variables that form the basis of the contribution analysis. The application of PPCA and VAE to TE process is discussed in detail. The case studies demonstrate that compare with PPCA, the detection perform of VAE is greatly improved. And the proposed contribution analysis can provide significant information to facilitate process fault diagnosis.

Future work is focused on extending the proposed contribution analysis for the monitoring of plant-level process.

Acknowledgement

We would also like to thank the National Key R&D Program of China (No.2017YFB0306403) for funding. Dr. Kai Zhang and Prof. Kaixiang Peng would like to thank the Natural and Science Foundation of China (NSFC) (Grant #61703036,#61473033 and #61773053 for funding.

References

Chen, T. and Sun, Y. (2009) ‘Probabilistic contribution analysis for statistical process monitoring: a missing variable approach’, *Control Engineering Practice*, Vol. 17, No. 4, pp.469–477.

Chen, Z., Ding, S.X., Peng, T., Yang, C. and Gui, W. (2017) ‘Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms’, *IEEE Transactions on Industrial Electronics*, Vol. 65, No. 2, pp.1559-1567.

Chiang, L.H., Russell, E.L. and Braatz, R.D. (2000) *Fault Detection and Diagnosis in Industrial Systems*, Springer Science & Business Media, pp.103-120.

Choi, S.W., Lee, C., Lee, J.M., Park, J.H. and Lee, I.B. (2005) ‘Fault detection and identification of nonlinear processes based on kernel PCA’, *Chemometrics and Intelligent Laboratory Systems*, Vol. 75, No. 1, pp.55–67.

Ding, S.X., Zhang, P., Naik, A., Ding, E.L. and Huang, B. (2009) ‘Subspace method aided data-driven design of fault detection and isolation systems’, *Journal of Process Control*, Vol. 19, No. 9, pp.1496–1510.

Downs, J.J. and Vogel, E.F. (1993) ‘A plant-wide industrial process control problem’, *Computers and Chemical Engineering*, Vol. 17, No. 3, pp.245–255.

Ge, Z. and Song, Z. (2010) ‘Kernel generalization of PPCA for nonlinear probabilistic monitoring’, *Industrial and Engineering Chemistry Research*, Vol. 49, No. 22, pp.11832–11836.

Ge, Z. and Song, Z. (2010) ‘Mixture Bayesian regularization method of PPCA for multimode process monitoring’, *AIChE Journal*, Vol. 56, No. 11, pp.2838–2849.

Kim, D. and Lee, I.B. (2003) ‘Process monitoring based on probabilistic PCA’, *Chemometrics and Intelligent Laboratory Systems*, Vol. 62, No. 2, pp.109–123.

Kingma, D.P. and Welling, M. (2013) ‘Auto-encoding variational bayes’, *arXiv preprint*, arXiv:1312.6114.

Muradore, R. and Fiorini, P. (2011) ‘A PLS-based statistical approach for fault detection and isolation of robotic manipulators’, *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 8, pp.3167–3175.

Nomikos, P. and MacGregor, J.F. (1994) ‘Monitoring batch processes using multiway principal component analysis’, *AIChE Journal*, Vol. 40, No. 3, pp.1361–1375.

Nomikos, P. and MacGregor, J.F. (2009) ‘Multivariate SPC charts for monitoring batch processes’, *Technometrics*, Vol. 37, No. 1, pp.41–59.

Peng, K., Zhang, K., You, B., Dong, J. and Wang, Z. (2016) 'A quality-based nonlinear fault diagnosis framework focusing on industrial multimode batch processes', *IEEE Transactions on Industrial Electronics*, Vol. 63, No. 4, pp.2615–2624.

Peng, K.X., Ma, L. and Zhang, K. (2017) 'Review of quality-related fault detection and diagnosis techniques for complex industrial processes', *Acta Automatica Sinica*, Vol. 43, No. 3, pp.349–365.

Tipping, M.E. and Bishop, C.M. (1999) 'Probabilistic principal component analysis', *Journal of the Royal Statistical Society: Series B (Statistical Methodology*, Vol. 61, No. 3, pp.611–622.

Yin, S., Ding, S.X., Xie, X. and Luo, H. (2014) 'A review on basic data-driven approaches for industrial process monitoring', *IEEE Transactions on Industrial Electronics*, Vol. 63, No. 11, pp.6418–6428.

Yin, S., Gao, H., Qiu, J. and Kaynak, O. (2016) 'Fault detection for nonlinear process with deterministic disturbances: A just-in-time learning based data driven method', *IEEE Transactions on Cybernetics*, Vol. 47, No. 11, pp.3649–3657.

Yu, J. and Qin, S.J. (2008) 'Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models', *AIChE Journal*, Vol. 54, No. 7, pp.1811–1829.

Zhou, L., Li, G., Song, Z. and Qin, S.J. (2016) 'Autoregressive dynamic latent variable models for process monitoring', *IEEE Transactions on Control Systems Technology*, Vol. 25, No. 1, pp.366–373.

Zhu, J., Ge, Z. and Song, Z. (2016) 'Non-Gaussian industrial process monitoring with probabilistic independent component analysis', *IEEE Transactions on Automation Science and Engineering*, Vol. 14, No. 2, pp.1309–1319.