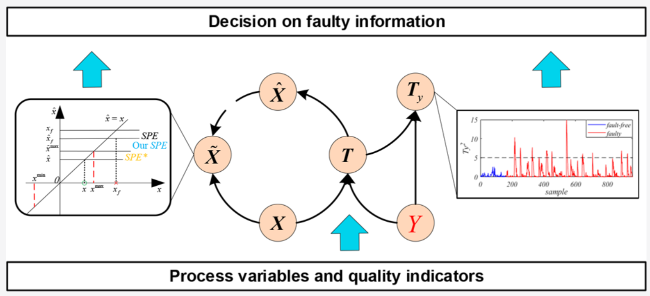
I&EC research
Industrial & Engineering Chemistry Research
pubs.acs.org/IECR
Article

Quality-Driven Autoencoder for Nonlinear Quality-Related and Process-Related Fault Detection Based on Least-Squares Regularization and Enhanced Statistics

Shifu Yan and Xuefeng Yan*

Cite This: *Ind. Eng. Chem. Res.* 2020, 59, 12136–12143
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Although many kernel-based quality-related monitoring methods have been developed for nonlinear processes, the nonlinearity between process variables and quality indicators is not well interpreted by kernel mapping and subsequent regression. To monitor a nonlinear quality-related latent space, a novel framework that consists of quality-related and process-related statistics rather than quality-related and quality-independent statistics is proposed. First, we train a quality-driven autoencoder (QdAE) with least-squares regularization through the gradient descent algorithm using quality indicators. Quality-related information can be predicted using latent variables through the auxiliary supervision of the quality indicators. Second, quality-related statistic T_y^2 is constructed to monitor the quality indicators. In the residual subspace derived by the QdAE, we compute the *SPE* statistic, which contains quality-related and quality-independent information. Furthermore, we present a strategy to enhance the *SPE* statistic to improve performance. Considering the quality-related and process-related monitoring using T_y^2 and SPE_{new}, we can also provide a reliable decision about whether the fault is quality-related or quality-independent. Finally, the proposed method is evaluated using the cases in the Tennessee Eastman process.



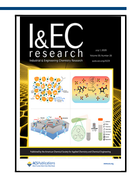
1. INTRODUCTION

Given that industrial processes are aiming toward becoming large scale and intelligent, the data recorded by plant sensors are indispensable. An increasing number of engineers focus on the characteristics of data, rather than analyzing the complicated mechanism of the processes. Therefore, data-driven modeling and monitoring methods are widely developed because of their safety and economic benefits in industries.[1−6] Consequently, quality-related process monitoring has received widespread attention over the past five years.[7,8] In addition to monitoring process variables in industries, this task aims to reveal the abnormal state of quality variables or key performance indicators upon fault detection. Engineers can make optimal decisions to avoid unnecessary losses by judging whether the occurring faults are quality-related or quality-independent.

The difference between quality- and process-related monitoring mainly lies in the process variables that can be collected in a timely manner while encountering the time lag and high costs in the acquisition of quality variables. Quality information can be predicted by process variables given that quality variables are related to process variables. Least square regression (LSR) and projection to latent structures (PLS) are widely used in data-driven predictive methods. Wang et al. proposed principal component regression (PCR), in which principal component analysis (PCA) and LSR between the process variables and the quality-related indicators were performed successively to predict quality-related information.[9] Given that the latent variables obtained by the PLS cannot distinguish quality-related and quality-independent information accurately, Zhou et al. further developed total PLS (TPLS), which extended PLS to four subspaces to indicate the specific faults that occurred in the processes.[10] Considering the dynamic properties of quality indicators, Li et al. and Jiao et al. proposed dynamic TPLS and LSR to achieve a better result in quality-related fault detection, respectively.[11,12]

Although quality indicators are affected by various factors, nonlinearities inevitably exist between the process variables and the quality indicators. Therefore, traditional methods have been transformed into nonlinear versions, such as kernel PCR (KPCR) and kernel TPLS (TKPLS).[9,13] Yan et al. presented a multiblock monitoring scheme based on the division of relevant and irrelevant variables to quality.[14] Jiao et al. and Wang et al. introduced singular value decomposition (SVD) based on kernel regression methods to divide the original

Received: February 12, 2020
Revised: May 6, 2020
Accepted: June 5, 2020
Published: June 5, 2020



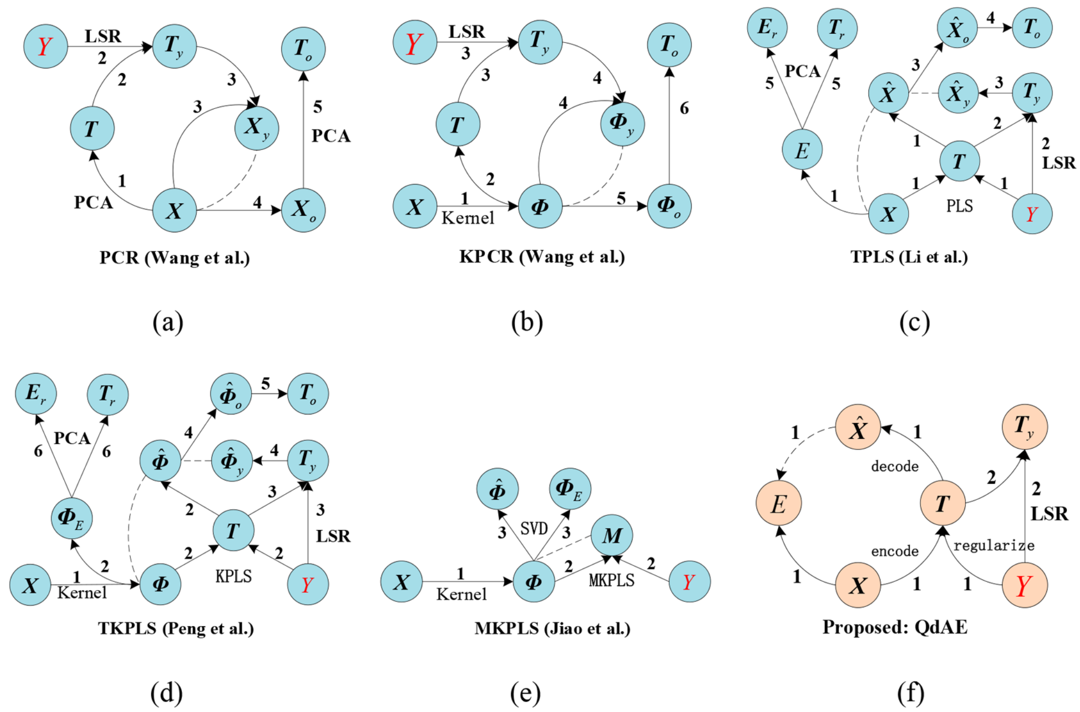


Figure 1. Comparison of quality-related models: (a) PCR,[9] (b) KPCR,[9] (c) TPLS,[10] (d) TKPLS,[13] (e) MKPLS,[15] (f) Proposed: QdAE. The serial numbers indicate the modeling steps of each method.

kernel spaces into quality-related and quality-independent subspaces.[15,16] However, their methods process the non-linearities among the process variables rather than the nonlinearities between process variables and quality indicators. Moreover, the establishment of the kernel spaces was not guided by the quality indicators. For nonlinear feature engineering, neural networks (NNs) are gradually becoming the mainstream approach, and many NN-based works on the industrial process have been conducted.[17−22] Yu et al. presented an incremental method for fault diagnosis based on broad convolutional NN.[23] Zhao et al. used NNs to learn the global and local information on the original process data.[24] For the prediction of quality indicators, Qiao et al. developed a nonlinear modeling method based on the self-organizing deep belief network.[25] Yan et al. and Yuan et al. respectively applied a denoising autoencoder (AE) and quality-relevant data in the application of soft sensors.[26−28]

For nonlinear quality-related fault detection, the accurate modeling of quality indicators and division of quality-related and quality-independent subspaces are still challenging in nonlinear processes. Currently, works on the nonlinear decomposition of quality-related and quality-independent subspaces are limited. Therefore, to avoid the inaccurate decomposition of subspaces, this study proposes a new framework that combines quality- and process-related statistics rather than the establishment of quality-related and quality-independent subspaces based on a quality-driven AE (QdAE). In the modeling stage, an AE is used to extract the latent variables of the process data. Meanwhile, we impose least-squares regularization between latent variables and quality indicators. In this way, the extraction of process- and quality-related information can be integrated. Given that the process variables are related to the quality indicators, such regularization will not affect the feature extraction and can be convenient for subsequent prediction. In the monitoring stage, quality indicators can be predicted and monitored by LSR and Hotelling's T^2 statistic in the latent subspace derived by the QdAE. For process-related monitoring, a new strategy is presented to enhance the sensitivity of the traditional SPE statistic. This new statistic aims to eliminate faulty information when reconstructing the original data and thus results in a larger reconstructive error. The contribution of this study mainly relies in the following aspects: (1) The least-squares regularization focuses on the trend of quality indicators, and the fusion of nonlinear and linear modeling enables the direct prediction of quality indicators using the latent variables. (2) The transfer of faulty information in the latent space into the residual subspace using the new *SPE* statistic can result in enhancement. (3) The combination of quality- and process-related monitoring can avoid traditional subspace decomposition because traditional methods mainly divided such subspaces linearly, which is inappropriate for nonlinear processes.

The remainder of this article is arranged as follows. Section 2 describes the related works. Section 3 presents the details of the proposed methodology. Section 4 provides the experiment and discussions. Finally, Section 5 elaborates the conclusions.

2. RELATED WORK AND PROBLEM FORMULATION

Fault detection is mainly based on hypothesis testing. Generally, the current methods for fault detection involve feature extraction and process reconstruction. Thus, the process space can be divided into latent and residual subspaces. The variation in latent space is mainly measured by the Hotelling's T^2 statistic, whereas the *SPE* statistic calculated by the reconstructive error is for the residual space or other detectors.[29−31] The quality-related task focuses on dividing the process space into quality-related and quality-independent subspaces, and then, the corresponding statistics can be calculated to reveal quality-related or quality-independent faults based on the process monitoring procedure. Meanwhile, obtaining quality-related and quality-independent information accurately remains a challenge.

Current methods are compared in Figure 1. On the basis of the latent variables extracted by PCA, PCR performs LSR between the latent variables to obtain the quality-related subspace ($\boldsymbol{T}_y$) and the quality-independent subspace ($\boldsymbol{T}_o$), and these two subspaces are monitored by T^2 statistics. Given that PLS performs an oblique decomposition, the latent and residual subspaces are not orthogonal to each other and thus contain quality-related and quality-independent information. Therefore, TPLS separates the latent variables further into quality-related $\boldsymbol{T}_y$ and quality-independent $\boldsymbol{T}_o$, which are monitored by T^2. Meanwhile, the residual subspace is also processed by PCA and monitored by T^2 and *SPE*. Considering the nonlinearity issues, KPCR and TKPLS use kernel functions to map the original data into a high-dimensional space, and subsequent procedures are executed in this space. Furthermore, the aforementioned methods extract quality-related information using latent variables. After establishing the regression models ($\hat{Y} = \boldsymbol{\Phi M}$), MKPLS performs SVD using $\boldsymbol{M}$ to divide the original space into quality-related and quality-independent subspaces, which are subsequently monitored by T^2 and *SPE*.

Evidently, such kernel-based methods only focus on the nonlinearity among the process variables instead of the nonlinearity between quality indicators and process variables, which may result in inefficiency. However, if nonlinear quality-related information is extracted as $\boldsymbol{T}_y = f_y(X)$ by the function $f_y(.)$, then the quality-independent subspace ($\boldsymbol{T}_o = f_o(X)$ obtained by $f_y(.)$) that follows $\boldsymbol{T}_y \perp \boldsymbol{T}_o$ is difficult to construct because of two reasons. First, we aim to find $\boldsymbol{X}_y$, which satisfies $\boldsymbol{T}_y = f_y(X) = f_y(\boldsymbol{X}_y)$; therefore, reconstructing $\boldsymbol{X}_y$ using $\boldsymbol{T}_y$ from $\boldsymbol{X}$ as $\boldsymbol{X}_y = \boldsymbol{T}_y\boldsymbol{P}_y^T$, where $\boldsymbol{P}_y^T = (\boldsymbol{T}_y^T\boldsymbol{T}_y)^{-1}$, can hardly satisfy $\boldsymbol{X}_y = \hat{\boldsymbol{X}}_y$,although $\boldsymbol{T}_y \perp \boldsymbol{T}_o$ can be satisfied in this manner. Second, if we reconstruct $\boldsymbol{X}_y$ using $\boldsymbol{T}_y$ in a nonlinear manner, then the quality-independent information can be also included because the reconstruction is supervised by the original $\boldsymbol{X}$. Furthermore, $\boldsymbol{T}_y \perp \boldsymbol{T}_o$ is not satisfied in the nonlinear approach.

Therefore, we adopt a new framework to avoid this problem. Instead of computing quality-related and quality-independent statistics to detect relevant faults, we consider the quality-related T^2 statistic and the process-related SPE statistic in this study (Figure 1(f)). The latter contains quality-related and quality-independent information. In fact, if a fault affects the quality indicators, then the quality-independent information will also detect the abnormal state unless the faulty variables are completely separated into the quality-related subspace. Through these two statistics, we can also specify whether the faults are quality related or quality independent. The details are presented in the following sections.

3. METHODOLOGY

3.1. QdAE. An AE is a three-layer NN that uses two-step learning to extract latent variables by minimizing reconstructive errors. Let $\mathbf{X} \in \mathcal{R}^{n\times m}$ indicate n samples of m input variables $\mathbf{x} \in \mathcal{R}^m$. First, the latent variables ($\mathbf{h} \in \mathcal{R}^d$) with d nodes are estimated by the encoder function ($\boldsymbol{h} = \sigma_1\ (\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{b}_1)$) in the hidden layer. Second, the latent representations are transformed back into the original space by the decoder function ($\hat{\boldsymbol{x}} = \sigma_2\ (\boldsymbol{W}_2\boldsymbol{h} + \boldsymbol{b}_2)$). $\sigma_1(\cdot)$ and $\sigma_2(\cdot)$ are the nonlinear activation functions, and the parameters ($\boldsymbol{\theta} = \{\boldsymbol{W}_1,\boldsymbol{W}_2,\boldsymbol{b}_1,\boldsymbol{b}_2\}$) in the encoder and the decoder are trained by using the following reconstruction errors.

$$\boldsymbol{\theta} = \arg\min \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2 \tag{1}$$

where the solution can be achieved by using the back-propagation algorithm. For complex data, stacked AE (SAE), which consists of multilayer hidden layers, is widely used to extract abstract representations.

Traditional kernel methods map variables into a high-dimensional space to predict quality-related information. Then, linear regression is performed between the high-dimension variables and the quality indicators. Nonlinear quality information cannot be guaranteed given that mapping and regression are performed step by step, and the mapping is unsupervised by quality indicators. Therefore, the QdAE uses the quality indicators to supervise the feature extraction and imposes a least-squares regularization on the latent variables. In this model, the original variables are nonlinearly transformed into latent variables, which are assumed to be linear to the quality indicators. If l quality indicators are collected as $\mathbf{Y} \in \mathcal{R}^{n\times l}$ and d-dimensional hidden features are indicated as $\mathbf{H} = g_1(\mathbf{X}) \in \mathcal{R}^{n\times d}$, then the overall objective can be described as

$$\boldsymbol{\theta} = \arg\min \frac{1}{n}\left(\sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2 + \gamma\ \|\boldsymbol{HQ} - \boldsymbol{Y}\|_2^2\right) \tag{2}$$

where $\mathbf{Q} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{Y} \in \mathcal{R}^{l\times d}$ is determined by least-squares, and γ is the weight of the regularized term.

Remark 1. Once the features ($\boldsymbol{H}$) are obtained, $\mathbf{Q} = [\boldsymbol{q}_1,\boldsymbol{q}_2, ...,\boldsymbol{q}_l]$ is the solution of the regularized term. Unlike the parameters in NNs, $\mathbf{Q}$ can be determined in each iteration optimally, thereby accelerating the training process.

The proposed model contains two loss terms. Thus, the optimization can be realized in an alternating manner. Given that parameter θ_2 of the decoder is only related to $\psi_1 = \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2$ and unrelated to regularized term $\psi_2 = \frac{\alpha}{n}\ \|\boldsymbol{HQ} - \boldsymbol{Y}\|_2^2$, the gradient $(\frac{\partial\psi_1}{\partial\theta_2})$ can be calculated directly. For θ_1 in the encoder, the gradient can be described as $\frac{\partial\psi_1}{\partial\theta_1} + \frac{\partial\psi_2}{\partial\theta_1}$. On the basis of the chain rule, $\frac{\partial\psi_2}{\partial\theta_1} = \frac{\partial\psi_2}{\partial\boldsymbol{H}}\frac{\partial\boldsymbol{H}}{\partial\theta_1}$ where $\frac{\partial\psi_2}{\partial\boldsymbol{H}} = 2\frac{\alpha}{n}(\boldsymbol{HQ} - \boldsymbol{Y})\boldsymbol{Q}^T$. The computation of $\frac{\partial\psi_1}{\partial\theta_1}$ and $\frac{\partial\boldsymbol{H}}{\partial\theta_1}$ are the same as that of $\frac{\partial\psi_1}{\partial\theta_2}$. Owing to the regularized term, the optimization with batch samples are adopted, and quality indicators are modeled by piecewise linear fitting. The optimized procedure is summarized in Algorithm 1.

https://dx.doi.org/10.1021/acs.iecr.0c00735
Ind. Eng. Chem. Res. 2020, 59, 12136−12143

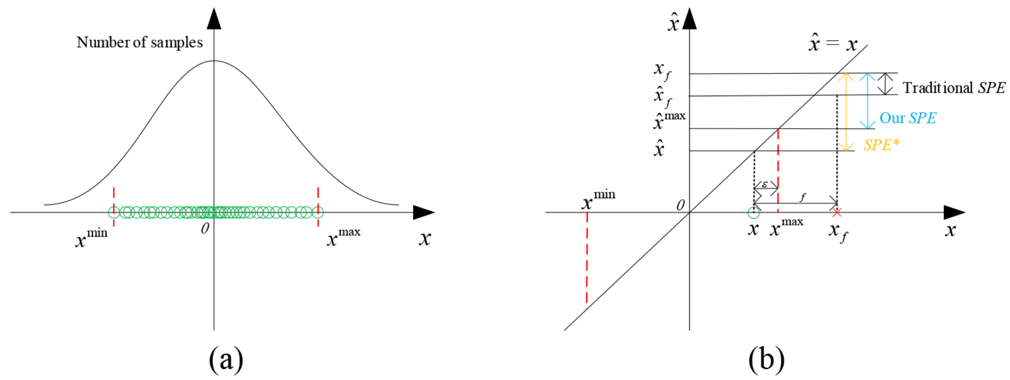


Figure 2. One-dimensional calculation of *SPE*: (a) storage of the limits and (b) effectiveness of the new *SPE*.

Algorithm 1: Algorithm for training a QdAE.

Require: process variables $\boldsymbol{X}$, quality indicators $\boldsymbol{Y}$.

Require: epoch T, batch size N, learning rate η.

Initialize the parameters ($\boldsymbol{\theta}$) of QdAE.

For $i = 1$ to T **do**

  For $j = 1$ to n/N **do**

    1) draw N samples $\{(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2), \ldots, (\boldsymbol{x}_N, \boldsymbol{y}_N)\}$;

    2) $\boldsymbol{H} = g_1(\boldsymbol{X})$;

    3) $\boldsymbol{Q} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{Y}$;

    4) $J = \frac{1}{B}\left(\sum_{i=1}^{B}\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2 + \gamma\|\boldsymbol{HQ} - \boldsymbol{Y}\|_2^2\right)$

    5) update parameters $\theta \leftarrow \theta - \eta\nabla_\theta J$;

  End

End

Lemma 1. The optimization method in Algorithm 1 can converge to a local optimum.

Proof: In the tth iteration, the loss of the model can be computed as

$$J(\boldsymbol{\theta}^t, \boldsymbol{Q}^t) = \frac{1}{n}\left(\sum_{i=1}^{n}\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i^t\|_2^2 + \alpha\,\|\boldsymbol{H}^t\boldsymbol{Q}^t - \boldsymbol{Y}\|_2^2\right) \quad (3)$$

where $\boldsymbol{\theta}^t$ $\boldsymbol{Q}^t$ are the trainable parameters after t iterations.

By the backpropagation method to update $\boldsymbol{\theta}^t$ where $\boldsymbol{Q}^t = \{\boldsymbol{q}_1^t, \boldsymbol{q}_2^t, \ldots, \boldsymbol{q}_l^t\}$ is fixed, we can have $J(\boldsymbol{\theta}^{t+1}, \boldsymbol{Q}^t) \le J(\boldsymbol{\theta}^t \boldsymbol{Q}^t)$. In this way, the updated latent variables are encoded by the $(t + 1)$th parameters $(\boldsymbol{\theta}^{t+1})$ as $\boldsymbol{H}^{t+1}$. Since $\boldsymbol{Q}^t$ is optimal for $\boldsymbol{H}^t$ not $\boldsymbol{H}^{t+1}$, $\boldsymbol{Q}^{t+1} = ((\boldsymbol{H}^{t+1})^T\boldsymbol{H}^{t+1})^{-1}\boldsymbol{Y}$ is subsequently obtained, which can have $\|\boldsymbol{H}^{t+1}\boldsymbol{Q}^{t+1} - \boldsymbol{Y}\|_2^2 \le \|\boldsymbol{H}^{t+1}\boldsymbol{Q}^t - \boldsymbol{Y}\|_2^2$ and $J(\boldsymbol{\theta}^{t+1}, \boldsymbol{Q}^{t+1}) \le J(\boldsymbol{\theta}^{t+1}, \boldsymbol{Q}^t)$. Therefore, $J(\boldsymbol{\theta}^{t+1}, \boldsymbol{Q}^t) \le J(\boldsymbol{\theta}^t \boldsymbol{Q}^t)$, and the optimized algorithm can converge after a certain number of iterations.

Remark 2. The time complexity for training an AE is $O(nm^2)$, and the time complexity of the regularized term is $O(nd^2 + d^3)$. The total time complexity of the QdAE is $O(nm^2 + d^3)$ that the dimension in the latent subspace is usually less than that at the input space $O(nm^2 + d^3)$. In this way, we can find that the calculation of QdAE is nearly equal to traditional AE when n is large.

3.2. Quality-Related Monitoring.

The latent features of the original data $\boldsymbol{X}$ can be calculated as $\boldsymbol{H}$ based on the encoder of the QdAE. Thus, the quality information can be predicted as $\boldsymbol{T}_y = \boldsymbol{HQ}$ where $\boldsymbol{Q} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{Y})$. For a new sample $(\boldsymbol{x}_{new})$, the corresponding feature vector can be calculated as $\boldsymbol{h}_{new}$, and the quality indicator can be predicted as $\boldsymbol{t}_{ynew}$. Therefore, Hotelling's T^2 statistic for this sample can be calculated as

$$T_y^2 = \boldsymbol{t}_{ynew}^T\left(\frac{\boldsymbol{T}_y^T\boldsymbol{T}_y}{n - 1}\right)^{-1}\boldsymbol{t}_{ynew} \quad (4)$$

3.3. Process-Related Monitoring.

To detect process-related faults that contain quality-related and quality-independent information, traditional models monitor the latent and residual subspaces using T^2 and *SPE*, respectively. These two subspaces complement each other. Therefore, we consider integrating the faulty information to the residual subspace and propose a new way of calculating SPE_{new}. We first take PCA as an example as AE can be regarded as a nonlinear extension of PCA. Given a faulty sample ($\boldsymbol{x}_f = \boldsymbol{x} + \boldsymbol{f}$, where $\boldsymbol{x}$ and $\boldsymbol{f}$ are the normal part and the faulty magnitude, respectively), the traditional *SPE* can be calculated using eq 5 under the PCA model, $\boldsymbol{X} = \boldsymbol{TP}^{\mathrm{T}} + \boldsymbol{E}$.

$$SPE = \left\|\boldsymbol{x}_f - \boldsymbol{PP}^T\boldsymbol{x}_f\right\|^2 = \|(\boldsymbol{I} - \boldsymbol{PP}^T)\boldsymbol{x} + (\boldsymbol{I} - \boldsymbol{PP}^T)\boldsymbol{f}\|^2 \quad (5)$$

where *SPE* contains the reconstructive errors of $\boldsymbol{x}$ and $\boldsymbol{f}$. For fault detection using *SPE*, models aim to reconstruct the normal part of the samples. Thus, the optimal statistic is

$$SPE^* = \left\|\boldsymbol{x}_f - \boldsymbol{PP}^T\boldsymbol{x}_f\right\|^2 = \|(\boldsymbol{I} - \boldsymbol{PP}^T)\boldsymbol{x} + \boldsymbol{f}\|^2$$

For the AE, *SPE* can be calculated using eq 6 based on encoder $g_1(\cdot)$ and decoder $g_2(\cdot)$.

$$SPE = \left\|\boldsymbol{x}_f - g_2g_1(\boldsymbol{x}_f)\right\|^2 = \left\|\boldsymbol{x} + \boldsymbol{f} - g_2g_1(\boldsymbol{x} + \boldsymbol{f})\right\|^2 \quad (6)$$

In this manner, the Taylor expansion can be executed at $\boldsymbol{x}$, thereby obtaining

$$SPE = \left\|\boldsymbol{x} - g_2g_1(\boldsymbol{x}) + \boldsymbol{f} - \Xi\boldsymbol{f}\right\|^2 \quad (7)$$

where

$$\Xi = \left.\frac{\partial g_2g_1}{\partial x}\right|_{x=\boldsymbol{x}}$$

https://dx.doi.org/10.1021/acs.iecr.0c00735
Ind. Eng. Chem. Res. 2020, 59, 12136−12143

The first term in eq 7 indicates the normal reconstructive error, while the next term refers to the reconstruction of faulty information. To increase the significance of the traditional SPE and approach the optimal $SPE^* = \|\boldsymbol{x} - g_2 g_1(\boldsymbol{x}) + \boldsymbol{f}\|^2$

$$SPE^* = \left\|\boldsymbol{x} - g_2 g_1(\boldsymbol{x}) + \boldsymbol{f}\right\|^2$$

two strategies are designed.

Strategy 1. Given that training data are collected under normal conditions, the normal values of the input signals can be roughly determined by these samples. In each sensor, samples are collected under the assumption of normal distribution (Figure 2(a)). We can determine the normal range of the corresponding signal based on the extreme values (i.e., $x^{\max}$ and $x^{\min}$). For m-dimensional input signals, the ith attribute of the faulty sample can be specified as $x_{f,i} = x_i + f_i$. Generally, retrieving x_i from the faulty sample to realize optimal detection performance is difficult. Thus, we can only estimate the value. If $x_{f,I}$ is outside $[x_i^{\min}, x_i^{\max}]$, then we can replace $x_{f,i}$ using the nearest limit value.

For samples in the normal area, the model can learn the relationship where $\hat{x} \approx x$, and the faulty sample beyond the area produces a larger reconstructive error, which is the traditional SPE. Our strategy used the estimated normal part of the sample, which is indicated as $x_f^* = x + \varepsilon$, to reconstruct the normal output, which can yield a larger reconstructive error than the traditional SPE (Figure 2(b)). Furthermore, the optimal solution can only be realized theoretically. Operating on all dimensions, the faulty sample can be adjusted as $\boldsymbol{x}_f^* = \boldsymbol{x} + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon}$ is the deviation of the estimated sample and the real normal part. In this manner

$$SPE_{new} = \left\|\boldsymbol{x} - g_2 g_1(\boldsymbol{x}) + \boldsymbol{f} - \Xi\varepsilon\right\|^2$$

Given that $\boldsymbol{\varepsilon}$ can be regarded as a part of $\boldsymbol{f}$, the reconstructive error of $\boldsymbol{\varepsilon}$ is less than $\boldsymbol{f}$ whether the values are negative or positive. Thus, $SPE_{new} \geq SPE$. Notably, the new statistics will increase the false alarm rate because SPE_{new} is the same as SPE if the sample falls into the normal area.

Remark 3. $\boldsymbol{\varepsilon} = 0$ indicates that our SPE is the same as the optimal one, whereas $\boldsymbol{\varepsilon} = \boldsymbol{f}$ indicates that the disturbance will have an insignificant effect.

Strategy 2. Apart from operating in the original space, we can also focus on the latent variables extracted by the AE. For the training samples, we can have the limits in latent subspaces $h_i^{\min}$ and $h_i^{\max}$, where $i = 1,2, ...d$. For faulty sample $\boldsymbol{x}_f = \boldsymbol{x} + \boldsymbol{f}$, the corresponding latent representation is $g_1(\boldsymbol{x}_f) = g_1(\boldsymbol{x}) + f_{latent}$. Similar to Strategy 1, we can construct $g_1(\boldsymbol{x}_f)^* = g_1(\boldsymbol{x}) + \boldsymbol{\varepsilon}$ given the normal area $[\boldsymbol{h}^{\min}, \boldsymbol{h}^{\max}]$. Therefore,

$$\begin{aligned} SPE_{new} &= \left\|\boldsymbol{x}_f - g_2(g_1(\boldsymbol{x}_f)^*)\right\|^2 \\ &= \left\|\boldsymbol{x} + \boldsymbol{f} - g_2(g_1(\boldsymbol{x}) + \boldsymbol{\varepsilon})\right\|^2 \end{aligned}$$

Obviously, Strategy 1 focuses on the original space, while Strategy 2 operates on the latent subspace. Compared with Strategy 1, Strategy 2 has two advantages: (1) The storage of the normal limit in the latent subspace is less than that on the original space because $d < m$, and thus, the computational complexity and storage of Strategy 2 is less than those of Strategy 1 in this way. (2) The model is trained to reconstruct the original data, thereby eliminating the faulty information on some dimensions in latent subspace.

3.4. Overall Monitoring Scheme.

Fault detection is implemented by determining whether the statistics of the test samples are outside the distribution. Kernel density estimation is used to learn the distributions of the aforementioned statistics, including T^2 and SPE. Given that the statistics of n samples are computed as $\{J_1, J_2, ...,J_n\}$, the probability density function can be estimated on the basis of Gaussian functions as

$$\hat{\phi}(J) = \frac{1}{nd}\sum_{i=1}^{n}\exp\left(-\frac{1}{2}\left(\frac{J - J_i}{d}\right)^2\right) \tag{8}$$

where d is the bandwidth.[32] Therefore, the normal threshold of this statistic can be determined as

$$prob(J \leq J_{th}) = \int_0^{J_{th}} \hat{\phi}(J)d(J) = \alpha$$

given significance level α.

The overall monitoring scheme based on the QdAE can be described as follows.

Offline modeling.

(1) Normalize offline samples $\boldsymbol{X}$ and $\boldsymbol{Y}$.
(2) Train a QdAE, including encoder $g_1(\cdot)$ and decoder $g_2(\cdot)$.
(3) On the basis of the latent variables ($\boldsymbol{H} = g_1(\boldsymbol{X})$), calculate the quality-related projection ($\boldsymbol{Q} = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Y})$).
(4) Determine thresholds T_{th}^2 and SPE_{th} (set 99% significance level in the following experiments).

Online monitoring.

(1) For new sample x_{new}.
(2) Calculate latent representation $\mathbf{h}_{new} g_1 = (\boldsymbol{x}_{new})$ and reconstruction $\hat{\boldsymbol{x}}_{new} = g_2(g_1(\boldsymbol{x}_{new}))$.
(3) Calculate statistics $T_y^2 = \boldsymbol{t}_{ynew}^T[\boldsymbol{T}_y^2\boldsymbol{T}_y/(n-1)]^{-1}\ \boldsymbol{t}_{ynew}$ and

$$SPE_{new} = \left\|\boldsymbol{x}_{new} - g_2(g_1(\boldsymbol{x}_{new})^*)\right\|^2$$

(4) The decision logic can be described as follows: if $T_y^2 > T_{th}^2$, then a quality-related fault exists; if $T_y^2 \leq T_{th}^2$ and $SPE_{new} > SPE_{th}$, then the fault does not affect the quality indicators; otherwise, no fault exists.

4. EXPERIMENT

4.1. Data Description.

We evaluate the proposed method on the data sets based on the Tennessee Eastman process (TEP). This platform, which is designed by Eastman Chemical Company is based on the real industrial process and considers many complex properties, such as nonlinear, integrated, and dynamic properties.[33] Overall, five units, namely, the reactor, condenser, separator, compressor, and stripper units, exist. To understand the process and implement control and monitoring strategies, two sets of variables, including 41 measurements (XMEAS (1)−(41)) and 11 analyzed components (XMV (1)−(11)), are collected. More details about this process can be found in ref 34.

In this study, XMEAS (1)−(22) and XMV (1)−(11) are selected as the process variables (i.e.,x_1,x_2, ...,x_{33}). Quality indicator y is the component (G) of purge gas XMEAS (35). A total of 22 data sets (i.e., IDV (0)−(21)) that contain 960 sample points in each set are adopted. In these data sets, the samples in IDV (0) are fault free, and those in IDV (1)−(21) are collected under 21 faulty conditions. In each faulty case,

https://dx.doi.org/10.1021/acs.iecr.0c00735
Ind. Eng. Chem. Res. 2020, 59, 12136−12143

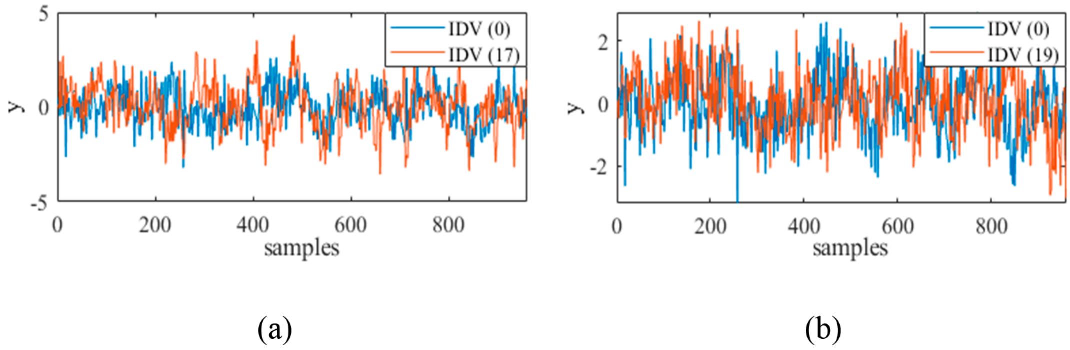


Figure 3. Values of quality indicator: (a) comparison of IDV (0) and (17) and (b) comparison of IDV (0) and (19).

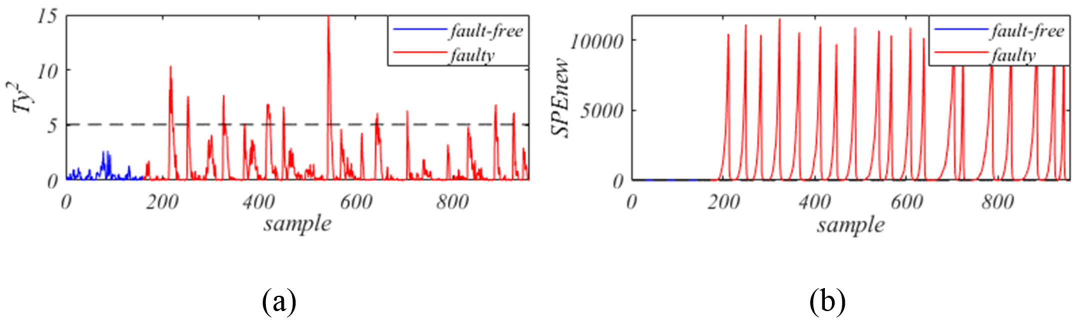


Figure 4. Monitoring result of IDV (17) by QdAE: (a) T_y^2 statistic and (b) SPE_{new} statistic.

the first 160 samples are fault free, and the 800 other samples are faulty.

4.2. Results and Discussion. The experiments are performed on a computer with Intel Xeon Silver 4110 2.1 GHz and 64 GB of RAM. The parameters for training the QdAE are indicated as follows: (1) γ = 0.1, (2) activation function = "tanh," (3) optimizer = "Adam," (4) epoch = 3000, and (5) batch size = 20. The structure is determined as 33−30−23−30−33 using the greedy layer-wise method by several AEs according to the reconstructive errors between original and reconstructive variables. To be more specific, in the first layer, we try AE with hidden nodes descending and determine the number when the error abruptly increases. The fault detection rates (FDR) of the proposed methods compared with those of some other methods are presented in Table S1 in the Supporting Information. The parameters of the other models are specified as follows: KPCR (A = 9, A_o = 156), QKPCA (A = 27), KPLS (A = 6), MKPLS (A = 6), TKPLS (A = 6, A_o = 6, A_r = 6). To compare the performance of these methods at the same level, the average false alarm rate (FAR) of each method calculated using the first 160 fault-free samples in 21 cases is below 2%. The best FDRs for faulty cases are in bold type. Overall, the QdAE provides better results in most cases, especially for IDV (16), (19), and (20). The average results are provided at the bottom of the table.

However, the FDRs of the quality-related statistics are very low because the quality indicator is unaffected all the time. Except for the quality-unrelated faults, the closed-loop control strategy will also eliminate the abnormal state of the quality indicator. Given the real values of the quality indicator that are absent when testing online, we compute quality-related statistic T_y^2 as follows to provide a reference.[35]

$$T_y^2 = \left(\frac{y - \bar{y}}{\sigma_y}\right)^2 \sim \frac{(n+1)}{n} F_\alpha(1, n-1) \tag{9}$$

where $\bar{y}$ and σ_y are the mean and standard deviation of the values in IDV (0), respectively. Given the assumption of normal distribution, T_y^2 in eq 9 follows the F-distribution with significance level α and freedom 1, n −1. The results are provided at the rightmost column of Table S1, which reflects the quality indicator information accurately. Here, we assume that the results are better when they are close to those of these results. Compared with other methods, the QdAE has a more reliable performance in terms of the quality-related statistic, especially in IDV (1), (7), (8), (12), (13), and (20). The results of these methods indicate that the quality-related statistic can be sensitive when a fault is introduced, regardless of whether the FDR is high or low. However, the inappropriate decomposition of quality-related and quality-unrelated subspaces will result in the misjudgment of quality-unrelated faults. Specifically, IDV (17) and (19) are selected for further analysis as follows.

Although the faulty reasons of IDV (17) and (19) are unknown, these two are quality-unrelated faults compared with the real values of the quality indicator shown in Figures 3(a) and (b). Figure 4 shows the monitoring results of QdAE. The results of KPCR and TKPLS for IDV (17) are presented in Figures S1 and S2. QdAE and KPCR provide a better quality-related statistic result, while the performance of other methods with higher FDRs is not ideal. Considering the process-related statistics, the QdAE outperforms other methods in IDV (14). For another quality-unrelated fault (i.e., IDV (19)), although the results of quality-related statistics are similar among the compared algorithms, the process-related performance of these methods is much lower than that of the QdAE, which is up to 88% which can be seen in Figure 5. The results of KPCR and

https://dx.doi.org/10.1021/acs.iecr.0c00735
Ind. Eng. Chem. Res. 2020, 59, 12136−12143

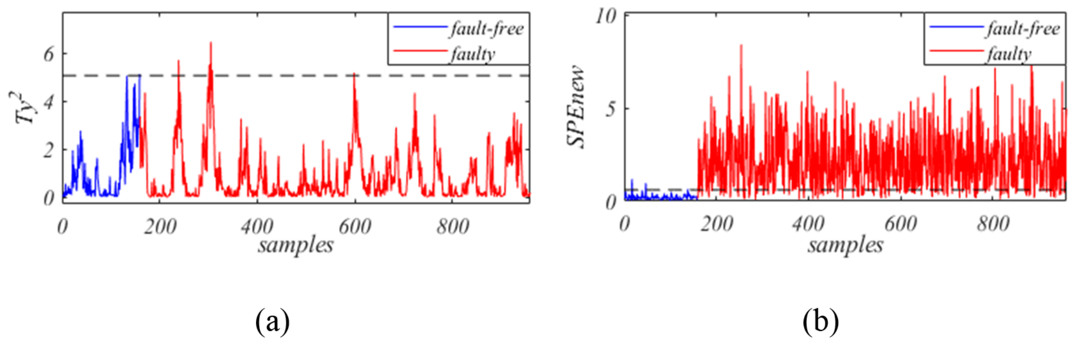


Figure 5. Monitoring result of IDV (19) by QdAE: (a) T_y^2 statistic and (b) SPE_{new} statistic.

TKPLS for IDV (19) are presented in Figures S3 and S4. Therefore, the proposed QdAE can provide a more reliable quality-related fault detection result than the other methods.

In this study, we propose a strategy to enhance the *SPE* statistic and perform some ablation studies to show the effectiveness of using different AE or SAE structures. The average FDRs of the 21 cases are presented in Table S2. The results indicate that the enhancing strategy can improve the performance of *SPE* for fault detection, especially for cases in which detection is difficult.

5. CONCLUSIONS

In this article, a novel framework which adopts quality-related and process-related statistics for nonlinear quality-related fault detection is proposed. A new process-related statistic is calculated by a new strategy to improve the monitoring performance of traditional SPE. The superiority of the proposed method is verified by the cases of TEP. In this way, the process-related statistic detects the faults more sensitively, and the quality-related statistic provides a more accurate result for alarms of quality indicators. This study is presented as a complement for nonlinear reconstruction since linear reconstruction is commonly used. Therefore, future works can focus on the development of a nonlinear decomposition method for quality-related and quality-independent subspaces and the accurate estimation of the normal part of a faulty sample, which is the key to improving fault detection performance.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.0c00735.

Tables S1 and S2 and Figures S1−S4(PDF)

■ AUTHOR INFORMATION

Corresponding Author

Xuefeng Yan − *Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education, East China University of Science and Technology, Shanghai 200237, P. R. China;* orcid.org/0000-0001-5622-8686; Phone: 0086-21-64251036; Email: xfyan@ecust.edu.cn

Author

Shifu Yan − *Key Laboratory of Advanced Control and Optimization for Chemical Processes of Ministry of Education, East China University of Science and Technology, Shanghai 200237, P. R. China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.iecr.0c00735

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful for the support of the National Natural Science Foundation of China (21878081) and Fundamental Research Funds for the Central Universities under Grant of China (222201917006).

■ REFERENCES

(1) Qin, S. J. Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **2012**, *36* (2), 220−234.

(2) Ge, Z.; Song, Z.; Gao, F. Review of Recent Research on Data-Based Process Monitoring. *Ind. Eng. Chem. Res.* **2013**, *52* (10), 3543−3562.

(3) Yin, S.; Li, X.; Gao, H.; Kaynak, O. Data-Based Techniques Focused on Modern Industry: An Overview. *IEEE Trans. Ind. Electron.* **2015**, *62* (1), 657−667.

(4) Ding, S. X. Data-driven design of monitoring and diagnosis systems for dynamic processes: A review of subspace technique based schemes and some recent results. *J. Process Control* **2014**, *24* (2), 431−449.

(5) Wang, Y.; Si, Y.; Huang, B.; Lou, Z. Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008−2017. *Can. J. Chem. Eng.* **2018**, *96* (10), 2073−2085.

(6) Jiang, Q.; Yan, X.; Huang, B. Review and Perspectives of Data-Driven Distributed Monitoring for Industrial Plant-Wide Processes. *Ind. Eng. Chem. Res.* **2019**, *58* (29), 12899−12912.

(7) Zhang, K.; Hao, H.; Chen, Z.; Ding, S. X.; Peng, K. A comparison and evaluation of key performance indicator-based multivariate statistics process monitoring approaches. *J. Process Control* **2015**, *33*, 112−126.

(8) Jiang, Y.; Yin, S. Recent Advances in Key-Performance-Indicator Oriented Prognosis and Diagnosis With a MATLAB Toolbox: DB-KIT. *IEEE Trans. Ind. Informat.* **2019**, *15* (5), 2849−2858.

(9) Wang, G.; Luo, H.; Peng, K. Quality-related fault detection using linear and nonlinear principal component regression. *J. Franklin Inst.* **2016**, *353* (10), 2159−2177.

(10) Zhou, D.; Li, G.; Qin, S. J. Total Projection to Latent Structures for Process Monitoring. *AIChE J.* **2010**, *56* (1), 168−178.

(11) Li, G.; Liu, B.; Qin, S. J.; Zhou, D. Quality relevant data-driven modeling and monitoring of multivariate dynamic processes: the dynamic T-PLS approach. *IEEE Trans. Neural Netw.* **2011**, *22* (12), 2262−2271.

(12) Jiao, J.; Yu, H.; Wang, G. A Quality-Related Fault Detection Approach Based on Dynamic Least Squares for Process Monitoring. *IEEE Trans. Ind. Electron.* **2016**, *63* (4), 2625−2632.

(13) Peng, K.; Zhang, K.; Li, G. Quality-Related Process Monitoring Based on Total Kernel PLS Model and Its Industrial Application. *Math. Probl. Eng.* **2013**, *2013*, 1−14.
(14) Yan, S.; Huang, J.; Yan, X. Monitoring of quality-relevant and quality-irrelevant blocks with characteristic-similar variables based on self-organizing map and kernel approaches. *J. Process Control* **2019**, *73*, 103−112.
(15) Jiao, J.; Zhao, N.; Wang, G.; Yin, S. A nonlinear quality-related fault detection approach based on modified kernel partial least squares. *ISA Trans.* **2017**, *66*, 275−283.
(16) Wang, G.; Jiao, J. A Kernel Least Squares Based Approach for Nonlinear Quality-Related Fault Detection. *IEEE Trans. Ind. Electron.* **2017**, *64* (4), 3195−3204.
(17) Yan, S.; Yan, X. Design teacher and supervised dual stacked auto-encoders for quality-relevant fault detection in industrial process. *Appl. Soft Comput.* **2019**, *81*, 105526.
(18) Yan, S.; Yan, X. Using Labeled Autoencoder to Supervise Neural Network Combined with k-Nearest Neighbor for Visual Industrial Process Monitoring. *Ind. Eng. Chem. Res.* **2019**, *58* (23), 9952−9958.
(19) Wang, Y.; Pan, Z.; Yuan, X.; Yang, C.; Gui, W. A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network. *ISA Trans.* **2020**, *96*, 457−467.
(20) Yuan, X.; Li, L.; Wang, Y. Nonlinear Dynamic Soft Sensor Modeling With Supervised Long Short-Term Memory Network. *IEEE Trans. Ind. Inform.* **2020**, *16* (5), 3168−3176.
(21) Yu, J.; Yan, X. Whole Process Monitoring based on Unstable Neuron Output Information in Hidden Layers of Deep Belief Network. *IEEE Trans. Cybern.* **2020**, *2020*, 1.
(22) Zhu, J.; Shi, H.; Song, B.; Tan, S.; Tao, Y. Deep neural network based recursive feature learning for nonlinear dynamic process monitoring. *Can. J. Chem. Eng.* **2020**, *98*, 919−933.
(23) Yu, W.; Zhao, C. Broad Convolutional Neural Network based Industrial Process Fault Diagnosis with Incremental Learning Capability. *IEEE Trans. Ind. Electron.* **2020**, *67* (6), 5081−5091.
(24) Zhao, H.; Lai, Z.; Chen, Y. Global-and-local-structure-based neural network for fault detection. *Neural Networks* **2019**, *118*, 43−53.
(25) Qiao, J.; Wang, G.; Li, X.; Li, W. A self-organizing deep belief network for nonlinear system modeling. *Appl. Soft Comput.* **2018**, *65*, 170−183.
(26) Yan, W.; Tang, D.; Lin, Y. A Data-Driven Soft Sensor Modeling Method Based on Deep Learning and its Application. *IEEE Trans. Ind. Electron.* **2017**, *64* (5), 4237−4245.
(27) Yuan, X.; Zhou, J.; Huang, B.; Wang, Y.; Yang, C.; Gui, W. Hierarchical Quality-Relevant Feature Representation for Soft Sensor Modeling: A Novel Deep Learning Strategy. *IEEE Trans. Ind. Inform.* **2020**, *16* (6), 3721−3730.
(28) Yuan, X.; Ou, C.; Wang, Y.; Yang, C.; Gui, W. Deep quality-related feature extraction for soft sensing modeling: A deep learning approach with hybrid VW-SAE. *Neurocomput.* **2020**, *396*, 375.
(29) Jiang, Q.; Yan, S.; Yan, X.; Yi, H.; Gao, F. Data-Driven Two-Dimensional Deep Correlated Representation Learning for Nonlinear Batch Process Monitoring. *IEEE. Trans. Ind. Inoform.* **2020**, *16* (4), 2839−2848.
(30) Jiang, Q.; Yan, S.; Yan, X.; Chen, S.; Sun, J. Data-driven individual−joint learning framework for nonlinear process monitoring. *Contr. Eng. Pract.* **2020**, *95*, 104235.
(31) Lv, Z.; Yan, X.; Jiang, Q. Batch process monitoring based on self-adaptive subspace support vector data description. *Chemom. Intell. Lab. Syst.* **2017**, *170*, 25−31.
(32) Botev, Z. I.; Grotowski, J. F.; Kroese, D. P. Kernel density estimation via diffusion. *Ann. Stat.* **2010**, *38* (5), 2916−2957.
(33) Downs, J. J.; Vogel, E. F. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17* (3), 245−255.
(34) Ricker, N. L.; Lee, J. H. Nonlinear model predictive control of the Tennessee Eastman challenge process. *Comput. Chem. Eng.* **1995**, *19* (9), 961−981.
(35) Huang, J.; Yan, X. Quality Relevant and Independent Two Block Monitoring Based on Mutual Information and KPCA. *IEEE Trans. Ind. Electron.* **2017**, *64* (8), 6518−6527.