Soft Computing
https://doi.org/10.1007/s00500-022-07022-x

FOCUS

Neural representations for quality-related kernel learning and fault detection

Shifu Yan[1] · Lihua Lv[2] · Xuefeng Yan[1]

Accepted: 18 March 2022

Abstract
Quality-related modeling and monitoring which aim at the key performance indicators have received wide attention in the research community. The widely used kernel-based methods mainly map process variables into kernel space without considering the relationship between the high-dimension features and quality indicators; therefore, the modeling performance of such transform cannot be guaranteed. For quality-related kernel learning, we propose a framework consisting of flexible neural transform and fixed kernel mapping. In this framework, neural network is used to learn representations for predicting quality indicators in the following kernel regression models. For monitoring the quality-related and quality-independent information, we present a solution for relevant subspaces decomposition and the diagnostic logic is summarized based on the quality-related and quality-independent statistics. The effectiveness of the proposed method is evaluated by simulations and real industrial-scale process.

Keywords Neural network · Representation learning · Quality-related · Kernel learning · Fault detection · Process monitoring

1 Introduction

Intelligent manufacturing techniques enable the industrial processes more and more integrated and automated where a large volume of data can be collected. For monitoring such plant-wide systems, model-based methods are gradually limited due to the complex mechanisms of the interact units. Instead, there are more researchers or engineers looking at the data-driven methods (Ge et al. 2013; Ding 2014; Yin et al. 2015; Wang et al. 2018; Yin and Yan 2021). In the area of process monitoring, there is an important subtask, namely, quality-related fault detection, aiming to reveal whether the abnormal state affects quality indicators that are closely related to the economic benefits of companies (Zhang et al. 2015; Jiang and Yin 2019). However, the accurate mining of quality-related information from process variables remains a challenge.

Generally, quality-related process monitoring focuses on the special quality indicators that are measured with high cost, time delay and uncertainties. Therefore, it is necessary to predict the relevant information for the safety and profit. To achieve that, regression models are usually adopted such as least squares (LS) and partial LS (PLS). The information in the process variables is not all contribute to the quality indicators and thus can be divided into quality-related and quality-independent parts, which, respectively, indicate the quality-related and quality-independent faults. Zhou et al. presented total PLS (TPLS) which firstly derived four subspaces based on latent and residual subspaces of PLS (Li et al. 2010; Zhou et al. 2010) since these two subspaces somehow overlap in PLS. In a similar way, Qin et al. proposed three subspace-based concurrent monitoring method (CPLS) to reduce the storage of subspace parameters (Qin and Zheng 2013). Given the redundant information in process variables, Wang et al. consecutively performed principal component analysis (PCA) on process variables and LS with quality indicators (Wang et al. 2016). Jiao et al. presented an efficient

Communicated by Priti Bansal.

✉ Xuefeng Yan
xfyan@ecust.edu.cn

[1] Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Mei Long Road No. 130, P.O. Box 293, Shanghai 200237, China

[2] Baoshan Iron and Steel Co., Ltd., Shanghai 201900, China

Published online: 11 April 2022

subspace decomposition method based on singular vector decomposition (SVD) for the dynamic processes (Jiao et al. 2016).

Obviously, the above linear models may be insufficient when there are nonlinearities in current industrial data. To address this issue, researchers firstly introduce the kernel functions to map low-dimension data into high-dimensional space. In this way, Peng et al. proposed the kernel version of TPLS monitoring scheme (TKPLS) for the hot strip mill process (Peng et al. 2013). Jiao et al. and Wang et al., respectively, proposed modified kernel PLS (MKPLS) and modified kernel LS (MKLS) based on the projections obtained by SVD (Jiao et al. 2017; Wang and Jiao 2017). However, these kernel-based methods are all consisting of two separate steps, i.e., kernel mapping and regression. Since the selection of kernels does not consider the relationship between high-dimension features and quality indicators, such kernel representations may not be optimal for quality-related predictions which may degrade the monitoring performance. At the same time, linear and nonlinear decomposition are, respectively, used for quality-related and quality-independent subspaces which will lead to information loss in quality-independent subspace for fault detection.

In this case, algorithms for kernel learning are solutions to optimize the parameters like maximum variance unfolding (Weinberger and Saul 2006; Xiong et al. 2005). However, the procedure is inevitably time-consuming when facing the industrial big data and without considering quality indicators. Based on the representation learning ability of neural networks (NNs), we propose a novel framework consisting of flexible NN and fixed kernel mapping to learn neural representations for quality-related kernel learning (NRKL). In such transformation, parameters are trained by an objective which both considers kernel mapping and quality-related information. After that, relationship between the trained neural representations with quality indicators can be modeled by kernel least squares regression. For monitoring the quality-related and quality-independent information in the industrial processes, we present a solution for corresponding subspace decomposition and calculation of quality-related, quality-independent and residual statistics. Finally, the diagnostic logic is based on a quality-related statistic and an integrated statistic derived by the Bayesian inference. The contributions of this work lies in the following aspects:

(1) the representations learned by NN are appropriate for quality-related kernel-based models and does not rely on the kernel parameters;
(2) the decomposition of quality-related and quality-independent subspaces is mathematically orthogonal for monitoring scheme;
(3) the method of training, subspace decomposition and calculation of statistics are derived.

2 Background

2.1 Neural network

Current NNs generally consist of more than two layers, which can learn deep and more useful representations with the pre-trained methods or other training tricks. In this way, the growing availability of industrial big data exhibits more and more value to be extracted based on these models. For processing the multidimensional data in industrial processes, fully connected neural network (FCNN) is considered in this paper. To reveal the transformation of NNs, firstly collect m process variables with n instances for training as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n] \in \Re^{m \times n}$.

Let $\boldsymbol{x}$ be one of the instances, the output of first hidden layer consisting l_1 nodes is computed as $\mathbf{h}_1 = s(\mathbf{W}_1 x + \mathbf{b}_1) \in \Re^{l_1}$ where $\mathbf{W}_1, \mathbf{b}_1$ are the weights and bias in this transform, and $s : \Re \rightarrow \Re$ is the activation function which is applied component-wise. Then this first hidden representations $\mathbf{h}_1$ can be used for the transformation in next layer with l_2 nodes as $\mathbf{h}_2 = s(\mathbf{W}_2 x + \mathbf{b}_2) \in \Re^{l_2}$, and until the final layer. Based on these layers, neural representations of such multilayer NN can be expressed as $\mathbf{h} = g(\mathbf{x}, \boldsymbol{\theta})$ where $\boldsymbol{\theta}$ indicates the sets of weights and bias that need to be optimized.

Given the rapid development of deep learning algorithms, neural networks have recently applied in industrial practice such as electric vehicle motors (Ghahfarokhi et al. 2021), remote sensing images (Li et al. 2021) and high-speed train vibration prediction (Sahoo et al. 2021; Hussain et al. 2021). In the predictive task, Yuan et al. applied a variable-wise weighted stack autoencoder (SAE) to model refinery process (Yuan et al. 2018). Yao et al. used hierarchical extreme learning machine for the semi-supervised data modeling (Yao and Ge 2018). For fault diagnosis, Yu et al. proposed a broad convolutional NN with incremental method to improve the online learning ability (Yu and Zhao 2019). These tasks are all focusing on predicting the quality variables in industrial applications. Quality-related fault detection detect whether the fault affects the quality rather than accurate prediction which is challenge for the faulty scenarios.

2.2 Kernel function

Kernel-based methods are commonly used for nonlinear tasks by mapping data into kernel space. Let $\boldsymbol{x}$ be transformed into $\varphi(\mathbf{x}) \in \Re^f$, where f is the dimension in kernel

space. Given the inexplicit of such kernel mapping, kernel matrix $\mathbf{K} = \Phi\Phi^T \in \Re^{n\times n}$ is constructed for implementation, where $\mathbf{\Phi}(\mathbf{X}) = [\varphi(x_1), \varphi(x_2), ..., \varphi(x_n)] \in \Re^{f\times n}$, and the element in this kernel matrix is computed as $K(i,j) = \langle \varphi(x_i), \varphi(x_j) \rangle$, where $\langle \cdot, \cdot \rangle$ means inner product operation. Based on the assumption that the distance between samples in the lower dimensional space is also satisfied in the kernel space, the Gaussian function is widely used for calculating the elements in $\boldsymbol{K}$.

$$K(i,j) = \exp\left(-\frac{\left\| x_i - x_j \right\|^2}{c} \right), \tag{1}$$

where c is the kernel parameter.

3 Methodology

3.1 Neural representations for kernel learning

For quality-related fault detection, the most important step is to separate and predict the quality-related information. Given that traditional kernel-based methods project the original variables into the kernel space without considering the correlation with quality indicators, this will not contribute to an ideal predictive performance using kernel regression. Here we proposed a framework based on multilayer NN which learns representations for quality-related kernel mapping, and thus the quality indicators can be better modeled in the kernel space based on such neural representations. Although FCNN is considered in this work due to the property of industrial data, it is noted that this framework is flexible to other nets and the overall schematic is plotted in Fig. 1.

Collect offline pairwise data $\{\mathbf{X}, \mathbf{y}\}$ with one quality indicator y, FCNN is used to transform original $\boldsymbol{X}$ into representations $\mathbf{H} = g(\mathbf{X}, \boldsymbol{\theta}) \in \Re^{l\times n}$ which is following mapped into the kernel space $\mathbf{\Phi}(\mathbf{H})$. Therefore, we learn $\mathbf{H}$ by maximizing the following covariance between $\mathbf{\Phi}(\mathbf{H})$ and quality indicator in a linear way.

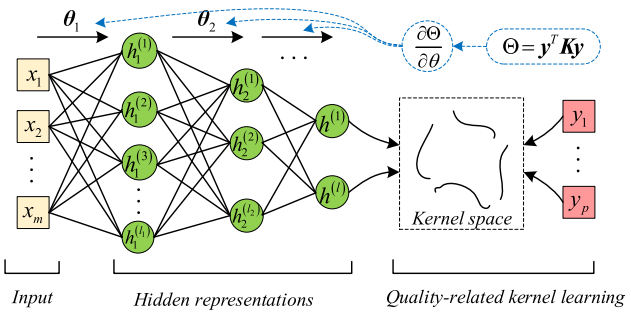

Fig. 1 Schematic of optimizing neural representations for quality-related kernel learning based on fully connected neural networks

$$\Theta = (\Phi\mathbf{q})^T\mathbf{y} = \sum_{i=1}^{f} (\mathbf{\Phi}_i\mathbf{q}_i)^T\mathbf{y}, \tag{2}$$

where $\mathbf{q} = \left[q_1, q_2, ..., q_f\right]^T$ is used to project the kernel space into the quality-related space and thus correlation coefficient can be calculated. Since $\mathbf{\Phi}(\mathbf{H})$ can be learned by adjusting parameters in NNs, the calculation of q_i can be combined into the optimization of $\mathbf{\Phi}(\mathbf{H})$ and the above objective can be further regarded as $\Theta = \sum_{i=1}^{f} \mathbf{\Phi}_i^T\mathbf{y}$. In this way, considering the unknowingness of $\mathbf{\Phi}(\mathbf{H})$, this objective is optimized in the following way.

$$\Theta = \sum_{i=1}^{f} \left(\mathbf{\Phi}_i^T\mathbf{y}\right)^T\mathbf{\Phi}_i^T\mathbf{y} = \mathbf{y}^T\sum_{i=1}^{f} \mathbf{\Phi}_i\mathbf{\Phi}_i^T\mathbf{y} = \mathbf{y}^T\mathbf{K}\mathbf{y}. \tag{3}$$

Based on (3), the loss function of such network is described as follows.

$$\min_{\boldsymbol{\theta}} -\mathbf{y}^T\left[\mathbf{\Phi}(g(\mathbf{X},\theta)) \cdot \mathbf{\Phi}^T(g(\mathbf{X},\theta))\right]\mathbf{y}, \tag{4}$$

where parameters $\boldsymbol{\theta}$ are solved by the gradient-based backpropagation method. To calculate the gradient, chain rule is used as $\frac{\partial\Theta}{\partial\theta} = \sum_{ab}\sum_{cd} \frac{\partial\mathbf{\Theta}}{\partial\mathbf{K}_{ab}} \frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{cd}} \frac{\partial\mathbf{H}_{cd}}{\partial\theta}$. Firstly, the derivative of such object with respective to the kernel matrix can be calculated as $\frac{\partial\mathbf{\Theta}}{\partial\mathbf{K}_{ab}} = [-\mathbf{y}\mathbf{y}^T]_{ab}$; then, $\frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{cd}}$ is the derivatives of the kernel matrix with respect to kernel functions conditioned on the transformation $\boldsymbol{H}$ which holding $\boldsymbol{\theta}$ fixed and the calculation is discussed in Remark 1; finally, $\frac{\partial\mathbf{H}_{cd}}{\partial\boldsymbol{\theta}}$ can be realized in a standard way easily. Furthermore, for q quality indicators $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n] \in \Re^{q\times n}$, the objective can be computed as the sum of each quality indicator.

Remark 1 There exist three situations in $\frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{cd}}$. If $a = c$, $\frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{ad}} = \frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{h}_a} \frac{\partial\mathbf{h}_a}{\partial\mathbf{H}_{ad}} = [-2\mathbf{K}_{ab}\mathbf{h}_a]_d$; else if $b = c$, $\frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{bd}} = \frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{h}_b} \frac{\partial\mathbf{h}_b}{\partial\mathbf{H}_{bd}} = [-2\mathbf{K}_{ab}\mathbf{h}_b]_d$; otherwise, $\frac{\partial\mathbf{K}_{ab}}{\partial\mathbf{H}_{cd}} = 0$.

4 Monitoring scheme

Based on the representations obtained by NRKL, we can construct the following regression model.

$$\mathbf{Y} = \Phi\mathbf{M} + \tilde{\mathbf{Y}}, \tag{5}$$

where $\boldsymbol{M}$ can be solved by the least squares $\mathbf{M} = \left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^{\dagger}\mathbf{\Phi}^T\mathbf{Y} = \mathbf{\Phi}^T\left(\Phi\Phi^T\right)^{\dagger}\mathbf{Y}$ in which $\dagger$ means the Moore–Penrose inverse. The inverse can be realized by performing singular vector decomposition (SVD) on $\mathbf{K} = \Phi\Phi^T = [\mathbf{P}_1 \quad \mathbf{P}_2] \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \mathbf{P}_1^T \\ \mathbf{P}_2^T \end{bmatrix}$ and thus

$\left(\Phi\Phi^T\right)^{\dagger} = \mathbf{K}^{\dagger} = \mathbf{P}_1\mathbf{\Lambda}_1^{-1}\mathbf{P}_1^T$ (Wang et al. 2016).

For designing the quality-related monitoring scheme, the key step is to divide the quality-related and quality-independent information. Therefore, by performing SVD on $\mathbf{M}\mathbf{M}^T = \mathbf{\Phi}^T\mathrm{K}^{\dagger}\mathbf{Y}\mathbf{Y}^T\mathbf{K}^{\dagger}\mathbf{\Phi}$, we can get two projections as follows.

$$\mathbf{M}\mathbf{M}^T = [\mathbf{W}_1 \quad \mathbf{W}_2]\begin{bmatrix}\mathbf{\Sigma}_1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{W}_1^T\\ \mathbf{W}_2^T\end{bmatrix}, \tag{6}$$

where $\mathbf{W}_1 = \left[\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_q\right]$, $\mathbf{W}_2 = \left[\mathbf{w}_{q+1}, \mathbf{w}_{q+2}, ..., \mathbf{w}_n\right]$, and $\mathbf{\Sigma}_1 = diag\left\{\lambda_1, \lambda_2, ..., \lambda_q\right\}$. The quality-related and quality-independent subspaces in the kernel representations can be formed as

$$\mathbf{\Phi}_y = \Phi\mathbf{W}_1\mathbf{W}_1^T, \tag{7}$$

$$\mathbf{\Phi}_o = \Phi\mathbf{W}_2\mathbf{W}_2^T, \tag{8}$$

where $\mathbf{W}_1\mathbf{W}_1^T + \mathbf{W}_2\mathbf{W}_2^T = \mathbf{I}$, $\mathbf{W}_1^T\mathbf{W}_2 = 0$ and $\mathbf{W}_2^T\mathbf{M} = 0$.

Theorem 1 *In the above decomposition, $\mathbf{\Phi}_y$ is the quality-related part in $\mathbf{\Phi}$ while $\mathbf{\Phi}_o$ is quality-unrelated.*

Proof. Firstly, the orthogonal property of $\mathbf{\Phi}_y$ and $\mathbf{\Phi}_o$ can be guaranteed as $\mathbf{\Phi}_y\mathbf{\Phi}_o^T = \Phi\mathbf{W}_1\mathbf{W}_1^T\mathbf{W}_2\mathbf{W}_2^T\mathbf{\Phi}^T = 0$.

Then quality indicators can be predicted as $\hat{\mathbf{Y}} = \Phi\mathbf{M} = \mathbf{\Phi}\left(\mathbf{W}_1\mathbf{W}_1^T + \mathbf{W}_2\mathbf{W}_2^T\right)\mathbf{M} = \mathbf{\Phi}_y\mathbf{M}$.

Therefore, $\mathbf{\Phi}_y \in span\{\mathbf{M}\}$ is quality-related while $\mathbf{\Phi}_o \in span\{\mathbf{M}\}^{\perp}$ is quality-unrelated.

For implementing the above division, we firstly multiply $\mathbf{\Phi}$ on the left side of $\mathbf{M}\mathbf{M}^T\mathbf{w}_i = \lambda_i\mathbf{w}_i$ and have (9).

$$\Phi\mathbf{M}\mathbf{M}^T\mathbf{w}_i = \mathbf{Y}\mathbf{Y}^T\mathbf{K}^{\dagger}\Phi\mathbf{w}_i = \lambda_i\Phi\mathbf{w}_i. \tag{9}$$

Let $\Phi\mathbf{w}_i = \mathbf{u}_i$ and thus we can have $\mathbf{w}_i = \Phi\mathbf{K}^{\dagger}\mathbf{u}_i$ where $\mathbf{u}_i$ can be calculated by performing the eigenvalue decomposition on $\mathbf{Y}\mathbf{Y}^T\mathbf{K}$. Using $\mathbf{U} = \left[\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_q\right]$, then the subspaces can be computed as follows.

$$\mathbf{\Phi}_y = \Phi\Phi^T\mathbf{K}^{\dagger}\mathbf{U}\mathbf{U}^T\mathbf{K}^{\dagger}\mathbf{\Phi} = \mathbf{Z}_y\mathbf{\Phi}, \tag{10}$$

$$\mathbf{\Phi}_o = \mathbf{\Phi} - \mathbf{\Phi}_y = \left(\mathbf{I} - \mathbf{Z}_y\right)\mathbf{\Phi}. \tag{11}$$

Then in the quality-related and quality-unrelated subspaces, quality-related and quality-independent kernel matrixes are calculated as

$$\mathbf{K}_y = \mathbf{\Phi}_y\mathbf{\Phi}_y^T = \mathbf{Z}_y\mathbf{K}\mathbf{Z}_y^T, \tag{12}$$

$$\mathbf{K}_o = \mathbf{\Phi}_o\mathbf{\Phi}_o^T = \left(\mathbf{I} - \mathbf{Z}_y\right)\mathbf{K}\left(\mathbf{I} - \mathbf{Z}_y\right)^T. \tag{13}$$

For the purpose of monitoring, principal component analysis is performed on the kernel matrix and get relevant scores in (14) and (15).

$$\mathbf{T}_y = \mathbf{\Phi}_y\mathbf{P}_y = \mathbf{\Phi}_y\mathbf{\Phi}_y^T\mathbf{W}_y, \tag{14}$$

$$\mathbf{T}_o = \mathbf{\Phi}_o\mathbf{P}_o = \mathbf{\Phi}_o\mathbf{\Phi}_o^T\mathbf{W}_o, \tag{15}$$

where $\mathbf{W}_y$ and $\mathbf{W}_o$ are the eigenvectors of $\mathbf{K}_y$ and $\mathbf{K}_o$.

Based on the above offline modeling procedure, the quality-related and quality-independent part of a new sample $\mathbf{x}_{new}$ to be monitored can be computed as

$$\mathbf{t}_y = W_y^T\mathbf{\Phi}_y\varphi_y(\mathbf{x}_{new}) = \mathbf{W}_y^T\mathbf{Z}_y\mathbf{Z}_y\mathbf{k}_{new}, \tag{16}$$

$$\mathbf{t}_o = \mathbf{W}_o^T\mathbf{\Phi}_o\varphi_o(x_{new}) = \mathbf{W}_y^T\left(\mathbf{I} - \mathbf{Z}_y\right)\left(\mathbf{I} - \mathbf{Z}_y\right)\mathbf{k}_{new}. \tag{17}$$

Thus, the new scores are used for calculating statistics to monitor these subspaces as (18) and (19).

$$T_y^2 = \mathbf{t}_y^T\left(\frac{\mathbf{T}_y^T\mathbf{T}_y}{n-1}\right)^{-1}\mathbf{t}_y, \tag{18}$$

$$T_o^2 = \mathbf{t}_o^T\left(\frac{\mathbf{T}_o^T\mathbf{T}_o}{n-1}\right)^{-1}\mathbf{t}_o. \tag{19}$$

The above two statistics are both monitoring the latent space represented by $\boldsymbol{H}$. Since these representations are calculated by the quality indicators. The residual subspace $\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{\dagger}\mathbf{H}^T\mathbf{X}$ indicating the reconstruction errors of process variables should be further analyzed. Therefore, the *SPE* statistic for $\mathbf{x}_{new}$ can be computed as $SPE = \left\|\mathbf{x}_{new}^T - g(\mathbf{x}_{new})^T(\mathbf{H}^T\mathbf{H})^{\dagger}\mathbf{H}^T\mathbf{X}\right\|^2$. Note that the residual subspace may contain the unpredictable information of quality indicators, *SPE* is only calculated to indicate whether there is a fault.

After nonlinear mapping, the distribution of the above statistics T_y^2, T_o^2 and *SPE* need to be estimated for fault detection using the offline data. Take one of these three statistics as an example, the statistics calculated by the training data are indicated as $\{J_1, J_2, ..., J_n\}$ and the distribution on these instances are modeled using the mixture of Gaussian functions.

$$\hat{\rho}(J) = \frac{1}{nd}\sum_{i=1}^{n}\exp\left(-\frac{(J - J(i))^2}{2\sigma^2}\right), \tag{20}$$

where σ is the bandwidth (Botev et al. 2010). Given the significant level α, the threshold can be determined as $prob(J \le J_{th}) = \int_0^{J_{th}}\hat{\rho}(J)dJ = \alpha$ for fault detection. Therefore, thresholds for T_y^2, T_o^2, *SPE* can be determined as $T_{y,th}^2$, $T_{o,th}^2$ and SPE_{th}.

Actually, the fault magnitude of process variables is partially contributing to the quality indicators, therefore, the *SPE* or T_o^2 can be regarded in the same way. We further aggregate T_o^2 and *SPE* into B_o based on Bayesian inference (Jiang et al. 2020). Here, the conditioned probability under normal (N) and faulty (F) can be calculated as

$$prob(J|N) = \exp\left\{-\frac{J}{J_{th}}\right\}, \tag{21}$$

$$prob(J|F) = \exp\left\{-\frac{J_{th}}{J}\right\}, \tag{22}$$

$$prob(F|J) = \frac{prob(J|F)prob(F)}{prob(J|N)prob(N) + prob(J|F)prob(F)}, \tag{23}$$

where $prob(N)$ and $prob(F)$ can be regarded as $1-\alpha$ and α. Therefore, the statistic can be defined in (24) and the threshold is set as α.

$$B_o = \sum_{i=1}^{2}\left\{\frac{prob(J_i|F)prob(F|J_i)}{\sum_{i=1}^{2} prob(J_i|F)}\right\}. \tag{24}$$

After the proposed model has been constructed, statistics of a new sample $\mathbf{x}_{new}$ can be obtained, and the detective logic can be described as: if $B_{o,new} > \alpha$ and $T^2_{y,new} > T^2_{y,th}$, the abnormality which affects quality indicators is detected; else if $B_{o,new} > \alpha$ and $T^2_{y,new} \le T^2_{y,th}$, the abnormality is detected while quality indicators are not influenced; otherwise, the operation is safe. Based on the above descriptions, the overall monitoring scheme can be found in Fig. 2.

5 Case study

In this section, the efficiency of the proposed NRKL is verified by Tennessee–Eastman process (TEP) and multiphase flow process (MFP). In general, fault detection rate (FDR) and false alarm rate (FAR) are used. FDR indicates the rate of successful detection of faults in the faulty samples while the FAR means the rate of false alarms in the fault-free samples. For fair comparison, following experimented methods are all designed for quality-related monitoring including KPLS (Peng et al. 2013), KPCR (Wang et al. 2016), TKPLS (Peng et al. 2013), MKPLS (Jiao et al. 2017), MKLS (Wang and Jiao 2017) and SAE (Jiang et al. 2020). The experiments are performed on a computer with Intel Xeon Silver 4110 2.1 GHz and 64 GB of RAM.

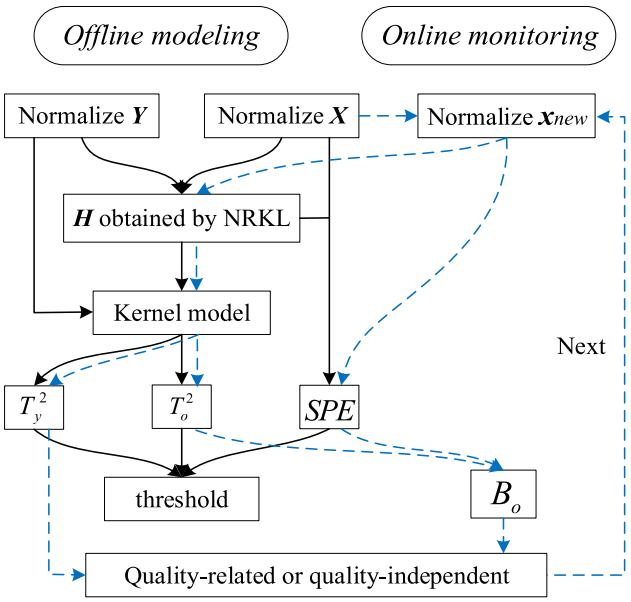

Fig. 2 The monitoring schematic

5.1 Tennessee–Eastman process

Refer to the real industrial process of Eastman Chemical Company, TEP given by Downs and Vogel (1993) is widely used as a benchmark problem in area of process control and monitoring. To better understand this platform, two sets of variables including 41 measurements XMEAS (1)–(41) and 11 manipulations XMV (1)–(11) are designed in five units, namely, reactor, condenser, separator, compressor and stripper. In this work, we consider XMEAS (1)–(21) and XMV (1)–(11) to be process variables (i.e., $x_1, x_2, ..., x_{33}$), and XMEAS (35) which is the composition of product to be the quality indicator (i.e., y). In the offline stage, 960 fault-free samples are collected in the training dataset (i.e., IDV (0)) and used for modeling. For monitoring online, 21 testing datasets involving sixteen known faults (i.e., IDV (1)–(16)) and five unknown faults (i.e., IDV (17)–(21)) are programmed for various experiments. In these faulty cases, anomalies are introduced in the last 800 samples and the former 160 samples are still fault-free. Besides, the source of these datasets is available at http://web.mit.edu/braatzgroup/links.html.

In the following experiments, $c = 50$ is set for the kernel-based methods and the number of latent variables A is determined by cross-validation using mean square error (MSE) on the prediction. The results of five-fold experiments based on the offline data are summarized in Table 1. We can find that the proposed NRKL outperforms other methods and it is more accurate for quality-related modeling. To be more specific, for KPLS-based model, $A = 6$ is set and $A_o = 7$, $A_r = 6$ are further determined in the residual subspace of TKPLS by the cumulative percent variance. For KPCR and MKLS, $A = 23$, $A_o = 156$ and

Table 1 Quality-related modeling performance by different methods

Method	KPLS	KPCR	KLS	NN	NRKL
MSE	0.433	0.695	0.348	0.265	**0.181**

$A_o = 20$ are, respectively, chosen. For the proposed method, NN with structure FC(33)-FC(33)-FC(33)-FC(40) is trained by Adam optimizer.

To provide an overall monitoring result in Fig. 3, FARs are controlled at a similar level which can be seen in Table 2, and best FDRs of statistics in each method are summarized together with the average performance and IDV (3), (9), and (15) are excluded in Fig. 3. In the step-fault cases in Fig. 3a and non-step cases in Fig. 3b, NRKL both exhibits a competitive performance. Furthermore, we can obviously find that NRKL outperforms other methods for the unknown cases in Fig. 3c and the average result in Fig. 3d. Besides, the average FARs of these methods are given in Table 2.

In this way, we analyzed IDV (1) and (19) to show the details. IDV (1) is introduced by a step change in the ratio of input feed composition and the real value of y is presented in Fig. 4a. We can find that the fault affects this quality indicator at the first time of the fault and it recovered to the fault-free state since the closed-loop control in this platform. The monitoring results of IDV (1) are given in Fig. 4b–f. IDV (19) is an unknown fault while we can find that it is a quality-independent in Fig. 5a. The monitoring results of IDV (19) are given in Fig. 5b–f. Generally, quality indicators are automatically monitored or controlled thus it would not be affected all the time. Therefore, the FAR of the quality indicators are more important. The proposed NRKL presents a better monitoring results both in detecting faults and revealing the quality indicators.

The above compared kernel-based methods commonly focus on the nonlinearities among the process variables. For KPLS, although quality-related latent variables are extracted, further relationship with the quality predictions is not analyzed and thus the quality-related monitoring result remains in the process space. Figure 4b clearly indicates this situation where both quality-related and quality-independent statistics alarm during the sampling time. TKPLS and MKPLS, respectively, improve the predictive monitoring performance of quality indicators based on different decomposition methods. Therefore, the monitoring results of quality-related statistics are closer to the real values. While there remains much false alarms in Fig. 4d, KPCR and KLS are based on the LS regression model, we can find that it can more clearly predict the trend of quality indicator. However, it is sensitive to the small extremums during the fluctuations seen in Fig. 4c, e. Compared with these models in Fig. 5, NRKL provides an overall superior monitoring results especially in the quality-independent statistic which indicate the detective ability of faults. Therefore, it is necessary to optimize the kernel representations for quality-related indicators.

5.2 Multiphase flow process

MFP provided by a real industrial facility in Cranfield University is used to process the mixture of water, oil, and air before the pressurized system (Ruiz-Cárcel et al. 2015). Different rates of the three-phase compositions indicate for different operational conditions. Totally 24 variables are collected under a sampling rate of 1 Hz including pressure, density, temperature and so on. Details and link can be found in the reference. In this work, 17 variables are chosen as process variables and the density of top separator output is selected as quality indicator. In the offline stage, 1000 fault-free samples are collected under the operation where the air flow rate is 0.0208 m^3/h and water flow rate is 3.5 kg/s in the input unit. In the same operated condition, three datasets are also collected for online testing.

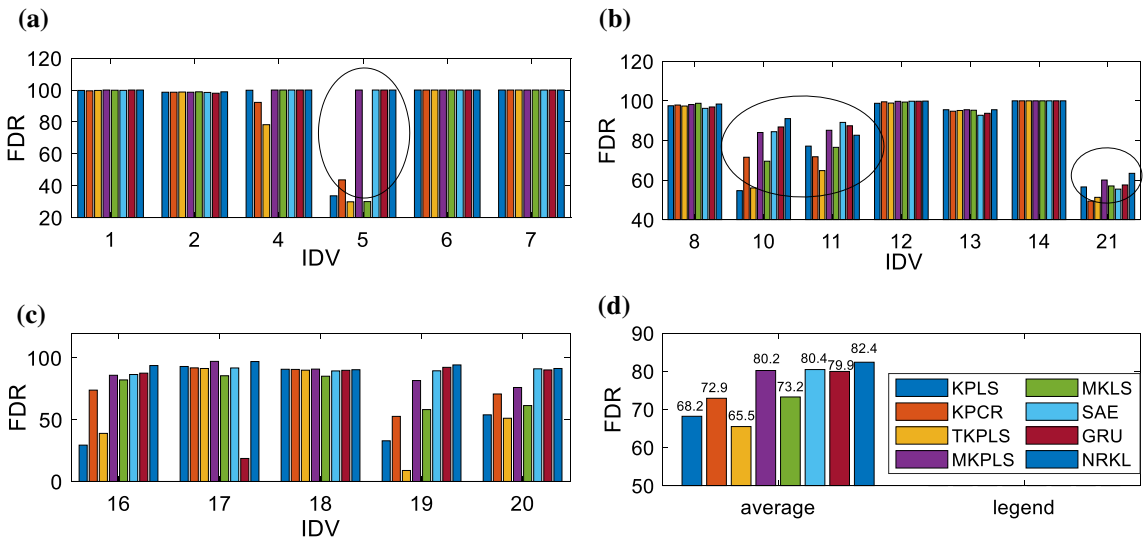

Fig. 3 Overall monitoring results of 21 cases in TEP: **a** step-fault cases; **b** non-step cases; **c** cases of unknown faults; **d** average result

Table 2 Average FARs (%) of different method for TEP

Method	KPLS	KPCR	TKPLS	MKPLS	MKLS	SAE	GRU	NRKL
FAR	2.93	2.98	2.22	3.01	3.18	2.28	2.01	**1.58**

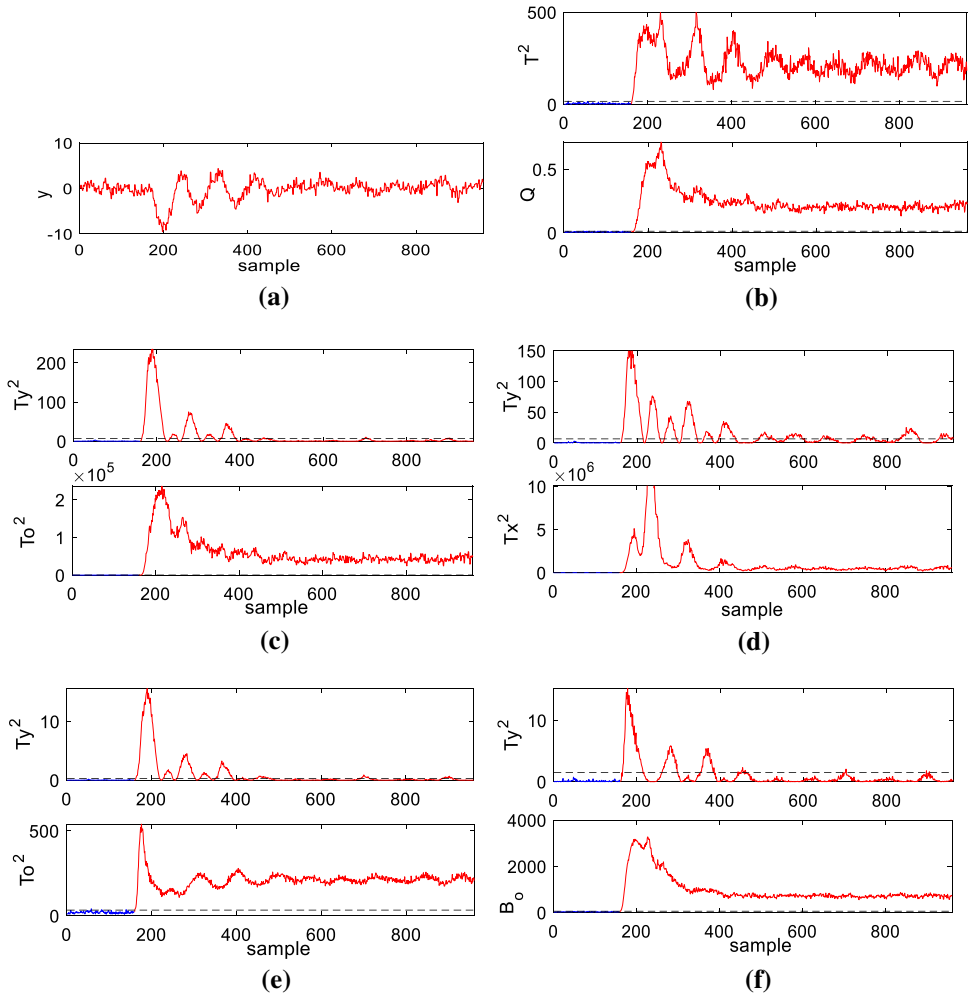

Fig. 4 Monitoring results on IDV (1): **a** real values of y; **b** KPLS; **c** KPCR; **d** MKPLS; **e** MKLS; **f** NRKL

The monitoring results for cases in MFP are summarized in Table 3. The case 1 is caused by airline line blockage in which 3691 samples are collected and the fault is introduced at 691st sample. In this case, rate of quality indicator affected is around 20%. The case 2 is caused by the input blockage on the top separator in which 9566 samples include and the fault is introduced at 596th sample. Quality-related rate is up to 100% which means this fault affects quality indicators as soon as it is introduced. The case 3 is caused by open direct bypass in which 3241 samples are involved and fault is introduced at 241st sample. The quality indicator is affected at 90%. We can find that NRKL provides a better monitoring performance in this real industrial process.

6 Conclusion

Quality-related monitoring is an important subtask in the area of process monitoring. This paper studies a new framework to learn neural representations for quality-related kernel learning. Based on such neural representations, kernel-based models can be better constructed for quality-related predictions. In this way, it can improve the generalized ability in current kernel mappings and loose the constraint on selection of kernel parameters. For monitoring online, quality-related and quality-independent subspaces can be separated in a direct way which is further verified by TEP and MFP where the proposed NRKL exhibits an excellent performance in detecting faults and revealing the trends of quality indicators.

Generally, the learning ability of NNs can be well incorporated while the determination of the neural structure

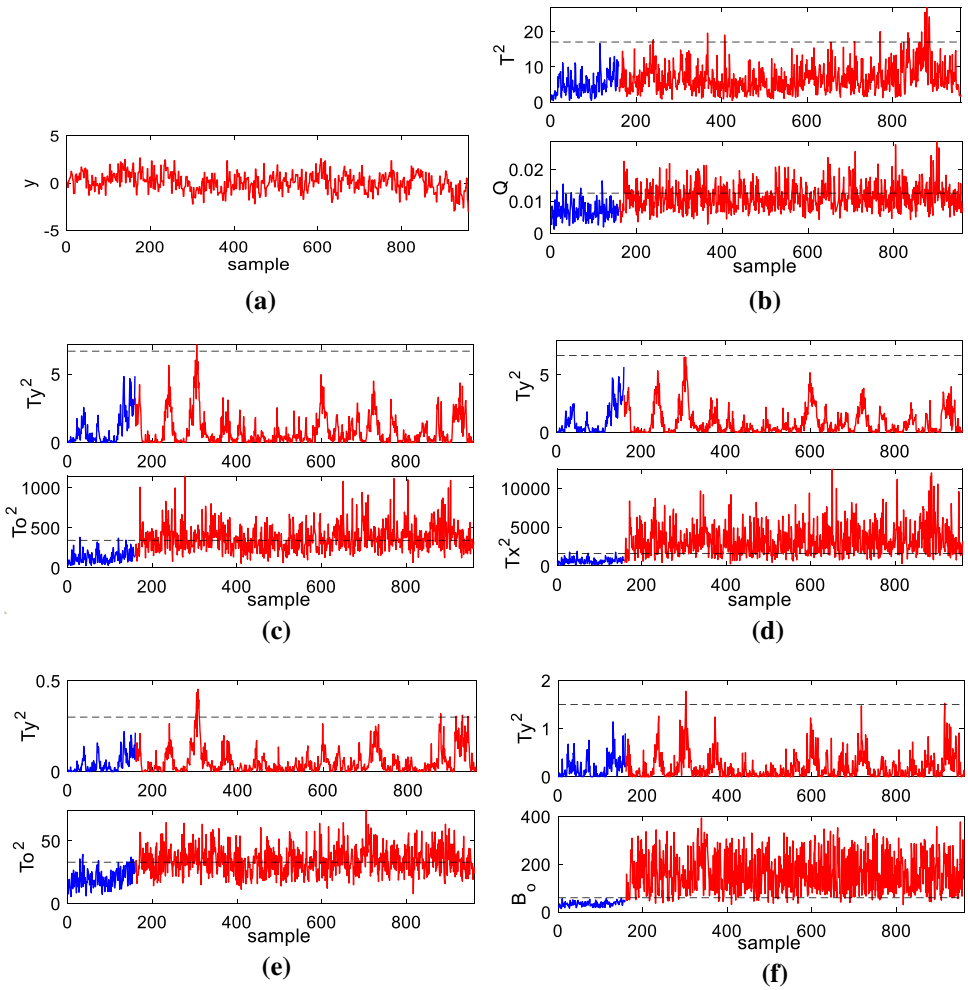

Fig. 5 Monitoring results on IDV (1): **a** real values of y; **b** KPLS; **c** KPCR; **d** MKPLS; **e** MKLS; **f** NRKL

Table 3 FDRs (%) of three cases in MFP

Case no	KPLS		KPCR		MKPLS		MKLS		SAE		GRU		NRKL	
	T^2	Q	T_y^2	T_o^2	T_y^2	T_x^2	T_y^2	T_o^2	T_y^2	T_o^2	T_y^2	T_o^2	T_y^2	B_o
1	50.3	40.7	36.0	44.8	29.9	50.1	31.1	48.5	35.5	50.0	32.5	52.0	25.3	**55.9**
2	99.2	99.5	99.3	99.7	97.8	99.5	99.5	97.9	98.8	98.9	99.2	99.4	99.8	**100**
3	79.1	89.5	59.9	89.8	49.4	90.1	49.3	89.9	50.7	91.2	53.4	90.2	59.4	**93.5**

is remained unsolved. Although more and more researchers have designed various efficient neural architectures, the optimal parameters for training a NN depends on the artificial experience and tricks. Such combination of NN and kernels can release this pressure to a certain degree. Further works can focus on other complex properties other than nonlinearity and the automatic learning of neural structures.

Author contributions All authors contributed to the study conception and design. Data collection, analysis and experiments were performed by SY. The first draft of the manuscript was written by SY, and all authors commented on previous versions of the manuscript.

Funding The authors are grateful for the support of National key research and development program of China (2021YFC2101100), and National Natural Science Foundation of China (21878081).

Data availability The datasets used during the current study are available at http://web.mit.edu/braatzgroup/links.html.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Botev Z, Grotowski J, Kroese D (2010) Kernel density estimation via diffusion. Annals of Stat 38(5):2916–2957

Ding SX (2014) Data-driven design of monitoring and diagnosis systems for dynamic processes: a review of subspace technique based schemes and some recent results. J Process Contr 24(2):431–449

Downs JJ, Vogel EF (1993) A plant-wide industrial process control problem. Comput Chem Eng 17(3):245–255

Ge Z, Song Z, Gao F (2013) Review of recent research on data-based process monitoring. Ind Eng Chem Res 52(10):3543–3562

Ghahfarokhi PS, Podgornovs A, Cardoso AJM, Kallaste A, Belahcen A, Vaimann T (2021) AC losses analysis approaches for electric vehicle motors with hairpin winding configuration. In: IECON 2021–47th annual conference of the ieee industrial electronics society, pp 1–4

Hussain R, Karbhari Y, Ijaz MF, Woźniak M, Singh PK, Sarkar R (2021) Revise-Net: exploiting reverse attention mechanism for salient object detection. Remote Sens 13(23):4941

Jiang Y, Yin S (2019) Recent advances in key-performance-indicator MATLAB toolbox. IEEE Trans Ind Inf 15(5):2849–2858

Jiang Q, Yan S, Cheng H, Yan X (2020) Local-global modeling and distributed computing framework for nonlinear plant-wide process monitoring with industrial big data. IEEE Trans Neural Netw Learn Syst 32(8):3355–3365

Jiao J, Yu H, Wang G (2016) A quality-related fault detection approach based on dynamic least squares for process monitoring. IEEE Trans Ind Electron 63(4):2625–2632

Jiao J, Zhao N, Wang G, Yin S (2017) A nonlinear quality-related fault detection approach based on modified kernel partial least squares. ISA Trans 66:275–283

Li G, Qin SJ, Zhou D (2010) Geometric properties of partial least squares for process monitoring. Automatica 46(1):204–210

Li X, Du Z, Huang Y, Tan Z (2021) A deep translation (GAN) based change detection network for optical and SAR remote sensing images. ISPRS J Photogramm Remote Sens 179:14–34

Peng K, Zhang K, Li G (2013) Quality-related process monitoring based on total kernel PLS model and its industrial application. Math Probl Eng 2013:707953

Qin SJ, Zheng Y (2013) Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures. AIChE J 59(2):2496–2504

Ruiz-Cárcel C, Cao Y, Mba D, Lao L, Samuel RT (2015) Statistical process monitoring of a multiphase flow facility. Contr. Eng. Pract. 42:74–88

Sahoo KK, Dutta I, Ijaz MF, Woźniak M, Singh PK (2021) TLEFuzzyNet: fuzzy rank-based ensemble of transfer learning models for emotion recognition from human speeches. IEEE Access 9:166518–166530

Wang G, Jiao J (2017) A kernel least squares based approach for nonlinear quality-related fault detection. IEEE Trans Ind Electron 64(4):3195–3204

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

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

Weinberger K, Saul L (2006) Unsupervised learning of image manifolds by semidefinite programming. Int J Comp vis 70(1):77–90

Xiong H, Swamy M, Ahmad M (2005) Optimizing the kernel in the empirical feature space. IEEE Trans Neural Netw 16(2):460–474

Yao L, Ge Z (2018) Deep learning of semisupervised process data with hierarchical extreme learning machine and soft sensor application. IEEE Trans Ind Electron 65(2):1490–1498

Yin J, Yan X (2021) Stacked sparse autoencoders monitoring model based on fault-related variable selection. Soft Comput 25(5):3531–3543

Yin S, Li X, Gao H, Kaynak O (2015) Data-based techniques focused on modern industry: an overview. IEEE Trans Ind Electron 62(1):657–667

Yu W, Zhao C (2019) Broad convolutional neural network based industrial process fault diagnosis with incremental learning capability. IEEE Trans Ind Electron 67(6):5081–5091

Yuan X, Huang B, Wang Y, Yang C, Gui W (2018) Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE. IEEE Trans Ind Inf 14(7):3235–3243

Zhang K, Hao H, Chen Z, Ding SX, Peng K (2015) A comparison and evaluation of key performance indicator-based multivariate statistics process monitoring approaches. J Process Contr 33:112–126

Zhou D, Li G, Qin SJ (2010) Total projection to latent structures for process monitoring. AIChE J 56(1):168–178