



Article

Online complex nonlinear industrial process operating optimality assessment using modified robust total kernel partial M-regression☆

Fei Chu^{1,*}, Wei Dai¹, Jian Shen¹, Xiaoping Ma¹, Fuli Wang^{2,3}

¹ School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China

² State Key Laboratory of Integrated Automation for Process Industries, Northeastern University, Shenyang 110819, China

³ College of Information Science and Engineering, Northeastern University, Shenyang 110819, China



ARTICLE INFO

Article history:

Received 19 November 2016

Received in revised form 31 December 2016

Accepted 28 June 2017

Available online 4 July 2017

Keywords:

Performance assessment

Optimization

Model

Economics

T-KPRM

Robust

ABSTRACT

Although industrial processes often perform perfectly under design conditions, they may deviate from the optimal operating point owing to parameters drift, environmental disturbances, etc. Thus, it is necessary to develop efficacious strategies or procedure to assess the process performance online. In this paper, we explore the issue of operating optimality assessment for complex industrial processes based on performance-similarity considering nonlinearities and outliers simultaneously, and a general enforced online performance assessment framework is proposed. In the offline part, a new and modified total robust kernel projection to latent structures algorithm, T-KPRM, is proposed and used to evaluate the complex nonlinear industrial process, which can effectively extract the optimal-index-related process variation information from process data and establish assessment models for each performance grades overcoming the effects of outlier. In the online part, the online assessment results can be obtained by calculating the similarity between the online data from a sliding window and each of the performance grades. Furthermore, in order to improve the accuracy of online assessment, we propose an online assessment strategy taking account of the effects of noise and process uncertainties. The Euclidean distance between the sliding data window and the optimal evaluation level is employed to measure the contribution rates of variables, which indicate the possible reason for the non-optimal operating performance. The proposed framework is tested on a real industrial case: dense medium coal preparation process, and the results shows the efficiency of the proposed method comparing to the existing method.

© 2017 The Chemical Industry and Engineering Society of China, and Chemical Industry Press. All rights reserved.

1. Introduction

In modern industrial processes, optimization of process operating performance has become a crucial issue and has received tremendous attention from both academia [1–4] and practice [5] recently because, in the face of growing competition, it represents the natural advantage for maximizing the economic profits. However, it is well known that the industrial process operating performance may deviate from the preliminary designed optimal operating point due to the disturbances and uncertainties. Therefore, process operating optimality assessment becomes important for industrial processes.

Operating optimality assessment is an issue to answer the question “How optimal the current operating performance is?” [6], in other words it is to decide whether the operating performance is optimal or non-optimal and to get a measure on how far the current operating

performance is from the optimum under the normal operating conditions [7]. Namely, based on the assessment results, managers and operators can further adjust the industrial production and improve process operating performance.

In recent years, some works have been published on the operating optimality assessment gradually [5–8]. Ye *et al.* [5] proposed a probabilistic framework of operating safety and optimality assessment for multi-mode industrial processes, where an optimality assessment index is defined and used to evaluate the process operating optimality. However, the assessment index is constructed based on the value of optimized objective function, such as plant cost, profits, product quality, the difficulty of calculating plant cost/profits online, and delay of production quality test, which makes the approach not suitable for the practical industrial application online, especially for the time-consuming and large-scale complex industrial processes. Recognizing that the output-related process variation information is different between different performance grades but are maintained consistent in one performance grade, Liu *et al.* [7,8] proposed a performance-similarity-based online assessment approach. They divided the training data into several steady performance grade based on the levels of an economic index, and then the process variation information of each steady performance grade was extracted by the featured extraction

☆ Supported by the National Natural Science Foundation of China (61503384, 61603393), Natural Science Foundation of Jiangsu (BK20150199, BK20160275), the Foundation Research Funds for the Central Universities (2015QNA65), the Postdoctoral Foundation of Jiangsu Province (1501081B).

* Corresponding author.

E-mail addresses: chuifeizhuifei@sina.com, chuifei@cumt.edu.cn (F. Chu).

methods, such as principal component analysis (PCA) and total projection to latent structures (T-PLS) [10,11], and used for operating optimality assessment. Although the PCA-based and T-PLS-based assessment methods are similar to some extent, the latter can overcome the low anti-interference ability and weak sensitivity drawbacks presented by the former, by removing the unrelated process variation with respect to the operating performance. Further in, Liu *et al.* proposed several improved versions of the performance-similarity and T-PLS-based online assessment approach to handle the nonlinearity issue [12], non-Gaussian multimode processes [13], and the misalignment between the actual process measurements and corresponding economic index [5].

However, for real industrial process, the obtained process data may be subjected to outliers [14], which originates from sensor faults, missing data, process disturbances, heavy tails of noise distribution functions, malfunction of instruments, and so on; since PLS is known to be very sensible to outliers [15], the existence of a very small amount of outliers may lead to deceptive results of T-PLS-based operating optimality assessment methods. The approaches for operating optimality assessment listed above did not consider this issue, despite Liu *et al.* mentioning the outliers in reference [7]. Based on the assumption that the data corresponding to a performance grade mainly distribute in an ellipse center area of the dataset, while the edge area of the ellipse is occupied by the outliers, the outliers were merely detected and removed from the modeling datasets by calculating the sample-similarity between each sample and the ellipse center of its data set before the assessment model for each performance grade were established. Nevertheless, for process with strong nonlinearities, the above assumption is usually not valid, and outliers are hard to be detected absolutely especially for processes with heavy-tail noise. Further, if tight sample-similarity threshold is used, too much “outliers” will be detected and removed, leading to insufficient samples especially for some performance grades which naturally lack samples. Meanwhile, outliers may contain some important process variations related to the operating performance, which should be extracted to establish exhaustive assessment models for each performance grade.

To minimize the adverse effects of outliers, several robust PLS methods have been proposed [14]. Among different robust PLS methods, a version of iteratively reweighted least squares algorithm for PLS has been obtained by Cumms and Andrew, called iteratively reweighted PLS. This method uses regression residuals to examine whether an object is an outlier or not, and calculates weight values to suppress the outliers in the model building step. Compared with other robust PLS methods, reweighted least squares algorithms possess low computational effort [15]. Yet, this method is not resistant to bad leverage points (outliers in the space of input variables) [16]. To solve this problem, partial robust M-regression (PRM) is proposed. With an appropriate choosing–weighting scheme, bad leverage points are also down-weighted at the same time. Further, a nonlinear version PRM [15], kernel Partial Robust M-regression, KPRM, is proposed which can cope with nonlinear robust problem. To the best of the author's knowledge, however, there have been few reports concerning robust nonlinear T-PLS method despite that the nonlinear robust problem is much more important in practical industrials.

In this paper, we explore the issue of operating optimality assessment for complex industrial process based on performance-similarity considering nonlinearities and outliers simultaneously, and a general enforced online performance assessment framework is proposed, which consists of offline training, online assessment and non-optimal cause identification. In practical, the process operating optimality performance is closely associated with optimal index (such as plant cost, profits or product quality), and in the offline training, the training samples can be divided into different data sets according to the level of optimal index, and each data set roughly represents one performance grade; then, total projection to latent structures algorithm will be used to extract the process variations with respect to the optimal index to establish the performance assessment models for each performance

grade. Considering the nonlinearities and robustness simultaneously, the primary version of T-PLS is not a good choice. A more powerful T-PLS algorithm is needed, which can cope with nonlinearities and minimize the adverse effects of outliers concurrently, establishing exhaustive assessment models for each performance grade. In the present study, we propose a novel modified robust nonlinear T-PLS method, called Total Kernel Partial M-regression (T-KPRM), by extending the choosing–weighting scheme to the kernel based nonlinear T-PLS algorithm. T-KPRM algorithm can also be treated as a post-processing method to decompose the KPRM kernel principal space and residual space, in which, with an appropriate choosing–weighting scheme, both bad leverage points and high residual points have been down-weighted.

In the online assessment, based on the understanding that the optimal index related process variation information differ between different performance grades, but maintain consistent among the same performance grades, which is also known as performance-similarity [8], the current process operating performance can be evaluated by calculating the performance-similarities between online data and each performance grade modeling data. For the conversion processes between performance grades, the optimal index related process variation would change gradually from one performance grade to another, thus performance-similarities with an ascending rule can also be used to distinct the performance grade conversion. However, the existence of process disturbances and outliers online often make the assessment results not credible. Thus, an enforced online assessment procedure is proposed, and for the non-optimal operating performance, the responsible variables can be identified by the variable contribution-based non-optimal cause identification method.

The contribution of the present paper are summarized as follows: (1) a robust and nonlinear version of total projection to latent structure is proposed, and based on it, a novel and robust operating optimal assessment approach is developed which considers nonlinearities and outliers simultaneously; (2) an enforced framework for online performance assessment considering the existence of process disturbance and outliers is proposed, and the online assessment results contain both steady performance grades and the conversations between performance grades.

The remainder of this paper is organized as follows: Section 2 gives a brief introduction of the algorithms of T-PLS and KPRM, and then the details of the proposed method based on T-KPRM for robust nonlinear process operating performance assessment that are presented in Section 3. The effect of the proposed method is demonstrated by a real case dense-medium coal preparation process, compared with the method proposed by Liu *et al.* [8]. Finally, the paper ends with concluding remarks in Section 5 and some acknowledgements.

2. Review of T-PLS and KPRM

2.1. T-PLS algorithm

T-PLS proposed by Zhou [9] is considered as an improved version of standard PLS. Through further decomposition, the output orthogonal and correlated part can be separated in the principal space, and large process variation can be separated from noise in residual space. Generally, compared with other improved PLS algorithms, *e.g.*, orthogonal signal correction based PLS and orthogonal PLS, T-PLS have been proven to be more effective in providing more accurate process information for those who are more concerned with certain aspects of the whole process information [17].

Let $\mathbf{X} \in R^{n \times m}$ be the input matrix consisting of n samples with m process variables per sample, and $\mathbf{Y} \in R^{n \times v}$ be the output matrix with v single quality variables per sample. It should be noted that in this paper only the case with single output quality variable is considered, *i.e.*, the ash of the product coal, and T-PLS algorithm for a single output $y \in R^{n \times 1}$ be the output matrix with v single quality variables per sample. It should be noted that in this paper only the case with single output

quality variable is considered, i.e., the ash of the product coal, and T-PLS algorithm for a single output $\mathbf{y} \in R^{n \times 1}$ is used; for the multiple output case, one just need to treat all of the output variables and use T-PLS algorithm for multiple outputs of $\mathbf{Y}$ simultaneously. It is assumed that $\mathbf{X}$ and $\mathbf{y}$ are centered to zero and scaled to unit variance. PLS decomposes $\mathbf{X}$ and $\mathbf{y}$ as follows [9],

$$\begin{cases} \mathbf{X} = \mathbf{T}\mathbf{P} + \mathbf{E} \\ \mathbf{y} = \mathbf{T}\mathbf{q} + \mathbf{f} \end{cases} \quad (1)$$

where $\mathbf{T} \in R^{n \times A}$ is the score matrix, and $\mathbf{P} \in R^{m \times A}$ and $\mathbf{q} \in R^{p \times A}$ are the loading matrices for $\mathbf{X}$ and $\mathbf{y}$, respectively. A is the number of PLS components, which is usually determined by cross validation. $\mathbf{E}$ and $\mathbf{f}$ are the residual matrices of $\mathbf{X}$ and $\mathbf{y}$, respectively. In the PLS procedure, to obtain the matrix $\mathbf{T}$, the weight matrix $\mathbf{W}$ is used. Let $R = \mathbf{W}(\mathbf{P}\mathbf{T}\mathbf{W}) - 1$, then $\mathbf{T} = \mathbf{X}\mathbf{R}$.

The PLS algorithm extracts the scores $\mathbf{T}$ to maximize the covariance between $\mathbf{X}$ and $\mathbf{y}$. However, the PLS scores contain components orthogonal to $\mathbf{y}$, which are output-irrelevant parts, and the $\mathbf{X}$ -residuals in PLS still contain a large variability, as PLS does not extract the $\mathbf{X}$ -variance in descending order. In addition, significant variations in the $\mathbf{X}$ -residuals may be affecting the product quality as well. This subspace is not modeled in the PLS model simple because it is not excited in the data that are used to build the PLS model. In summary, variations in both PLS scores and residuals can be relevant to the output $\mathbf{y}$. The T-PLS algorithm is proposed to further decomposes the PLS scores and residuals to sort out components that are relevant to the output $\mathbf{y}$.

Then, based on the T-PLS algorithm, the $\mathbf{X}$ -scores and $\mathbf{X}$ -residuals can be further decomposed as follows,

$$\begin{cases} \mathbf{X} = \mathbf{t}_y\mathbf{p}_y^T + \mathbf{T}_o\mathbf{P}_o^T + \mathbf{T}_r\mathbf{P}_r^T + \mathbf{E}_r \\ \mathbf{y} = \mathbf{t}_y + \mathbf{f} \end{cases} \quad (2)$$

where $\mathbf{t}_y \in R^{n \times 1}$, $\mathbf{T}_o \in R^{n \times (A-1)}$ and $\mathbf{T}_r \in R^{n \times Ar}$ are three score matrices and $\mathbf{P}_y \in R^{m \times 1}$, $\mathbf{P}_o \in R^{m \times (A-1)}$ and $\mathbf{P}_r \in R^{m \times Ar}$ are the corresponding loading matrices. $\mathbf{E}_r$ is the new residual matrix after performing PCA on $\mathbf{E}$ and can be expressed as $\mathbf{E}_r = \mathbf{E}(\mathbf{I} - \mathbf{P}_r\mathbf{P}_r^T)$. A_r is the number of output unrelated components. For a single output $\mathbf{y}$, T-PLS needs only one score vector $\mathbf{t}_y$ to predict the quality variable $\mathbf{y}$.

In Eq. (2), $\mathbf{t}_y$ represents the variations only related to $\mathbf{y}$ in original $\mathbf{T}$ of PLS model, i.e., the output-related variations in $\mathbf{T}$, to represent the variations orthogonal to $\mathbf{y}$ in $\mathbf{T}$, i.e., the output-unrelated variations in $\mathbf{T}$, $\mathbf{T}_r$ is the major part of original $\mathbf{X}$ -residual $\mathbf{E}$, and $\mathbf{E}_r$ is the residual part of $\mathbf{E}$ after $\mathbf{T}_r$ is removed. The T-PLS model decomposes input space $\mathbf{X}$ into four subspaces. They are the range spaces of $\mathbf{P}_y$, $\mathbf{P}_o$, $\mathbf{P}_r$, and $(\mathbf{I} - \mathbf{P}_r\mathbf{P}_r^T)(\mathbf{I} - \mathbf{P}_o\mathbf{P}_o^T)$ with dimensions of 1, $A-1$, Ar , $m-A-Ar$. A new sample vector $\mathbf{x}_{\text{new}}$ is partitioned into four portions,

$$\mathbf{x} = \mathbf{x}_y + \mathbf{x}_o + \mathbf{x}_r + \bar{\mathbf{x}}_r \quad (3)$$

$$\mathbf{x}_y = \mathbf{p}_y\mathbf{q}\mathbf{R}^T\mathbf{x} \quad (4)$$

$$\mathbf{x}_o = (\mathbf{P} - \mathbf{p}_y\mathbf{q})\mathbf{R}^T\mathbf{x} \quad (5)$$

$$\mathbf{x}_r = \mathbf{P}_r\mathbf{P}_r^T(\mathbf{I} - \mathbf{P}\mathbf{R}^T)\mathbf{x} \quad (6)$$

$$\bar{\mathbf{x}}_r = (\mathbf{I} - \mathbf{P}_r\mathbf{P}_r^T)(\mathbf{I} - \mathbf{P}\mathbf{R}^T)\mathbf{x} \quad (7)$$

and the scores of $\mathbf{x}_{\text{new}}$ can be calculated as follows,

$$\mathbf{t}_{y,\text{new}} = \mathbf{q}^T\mathbf{R}^T\mathbf{x}_{\text{new}} \quad (8)$$

$$\mathbf{t}_{o,\text{new}} = \mathbf{P}_o^T(\mathbf{P} - \mathbf{p}_y\mathbf{q}^T)\mathbf{R}^T\mathbf{x}_{\text{new}} \quad (9)$$

$$\mathbf{t}_{r,\text{new}} = \mathbf{P}_r^T(\mathbf{I} - \mathbf{P}\mathbf{R}^T)\mathbf{x}_{\text{new}} \quad (10)$$

From the perspective of process operating performance assessment, only the score $\mathbf{t}_y$ obtained from the offline modeling and $\mathbf{t}_{y,\text{new}}$ of online new sample are needed for they represent the output related process variations which can fully reflect the influence of process variation on the output, and it is appropriate to evaluate the process operating performance based on them [18]. For the detailed procedures for single output T-PLS algorithm please refer to [19].

2.2. KPRM algorithm

The partial robust M-regression (PRM), developed from Iterative weighted PLS algorithm, is a robust version of the PLS1 approach [15]. Through iterative computation, the algorithm can adaptively assign different weights to different sample data (the outliers are assigned close to 0 weight, the normal data are assigned close to 1 weight). The weighted method is used to overcome the influence of outlier on model and retain the part of useful information from outliers. Then, the outliers will be divided into two types by PRM algorithm. One is a high leverage point, which is far from the input data center of the sample points, and the other is the high residual point, which is the sample point of the large difference between the output value and the actual value. The algorithm uses different weighting methods to calculate the two types of outliers. The leverage weight of the i th sample data w_i^x latent is computed as

$$w_i^x = f\left(\frac{\|\mathbf{t}_i - \text{med}_{L1}(\mathbf{T})\|}{\text{med}_i\|\mathbf{t}_i - \text{med}_{L1}(\mathbf{T})\|}, c\right) \quad (11)$$

$$f(z, c) = \frac{1}{(1 + |z/c|)^2} \quad (12)$$

where $c = 4$ represent Euclidean distance, med_{L1} represent the median estimate of $L1$. $\mathbf{t}_i$ is the score vector of the i th sample data, c is a tuning constant, taken as $f(\cdot)$. the weight function $f(\cdot)$ is called the “Fair” function, and is one of several possible weight functions cited in the original IRPLS.

The residual weight of the i th sample data w_i^r is computed as

$$w_i^r = f\left(\frac{\mathbf{r}_i}{\bar{\mathbf{r}}}, c\right) \quad (13)$$

where $\mathbf{r}_i$ is the residual part between the actual value and predictive value of the i th sample data. $\bar{\mathbf{r}}$ is a robust estimator of the sample data.

$$\bar{\mathbf{r}} = \text{med}_i|\mathbf{r}_i - \text{med}_j(\mathbf{r}_j)|, i, j = 1, 2, \dots, n \quad (14)$$

Using both types of outliers, we can obtain the comprehensive weights of the sample data,

$$w_i = \sqrt{w_i^x w_i^r} \quad (15)$$

the weights of the sample data are updated continuously until the algorithm satisfies the convergence conditions [20].

KPRM algorithm is developed on the basis of the PRM algorithm. The output data matrix $\mathbf{y}$ discussed in this paper is a single output variable. In order to solve the nonlinear problem of process data, a kernel function ϕ is introduced to map the process data from original low dimensional space to high dimensional feature space F . In the high dimensional feature space F , the relationship of the process variables is linear. After the kernel function mapping, the sample data matrix will be changed to the characteristic data matrix $\Phi(\mathbf{X}) = [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots, \phi(\mathbf{x}_N)]^T \in R^{N \times M}$. The dimension M of the feature space can be any value, and even can be set to infinity. Define kernel matrix $\mathbf{K} = \Phi(\mathbf{X})\Phi^T(\mathbf{X})$. When the kernel function is introduced to calculate, we do not need to know the concrete form of the nonlinear mapping function $\Phi(\cdot)$, and also there is no need for inner product

computation in high dimensional space [12]. The kernel matrix of high dimensional space can be obtained by calculating the kernel function $K_{i,j}$ of the original low dimensional space.

The steps of the KPRM algorithm are as follows:

- Step 1: The input data matrix $\mathbf{X}$ is projected to the high dimensional feature space F by nonlinear mapping function $\Phi(\cdot)$.
- Step 2: Standard kernel matrix $\mathbf{K}$, then the input kernel matrix $\mathbf{K}$ and output matrix $\mathbf{y}$ are processed by PRM algorithm.
- Step 3: Initialize the weight $\mathbf{W}_i$ by using the formulas (11), (13), and (15).
- Step 4: Compute weighted input kernel matrix $\mathbf{K}_W$ and output data matrix $\mathbf{y}_W$ which obtained by weighting the original input kernel matrix $\mathbf{X}$ and output data $\mathbf{y}_W$. Then the PLS1 regression model is established after the weighted Input and output data, and the score vector is modified.
- Step 5: Calculate the residual $\mathbf{r}_i$ of each sample data, and update the weight $\mathbf{W}_i$ by using the formulas (11), (13), and (15).
- Step 6: Through consecutive calculation twice, the relative difference between $\mathbf{q}$ is less than a certain threshold (10^{-2}), then the convergence condition is satisfied, and go to step 5, Otherwise go back to step 4.

3. Proposed Approach for Robust Nonlinear Online Operating Optimal Assessment

3.1. T-KPRM algorithm

In this paper, the T-KPRM algorithm is proposed by combining the advantages of T-PLS and KPRM. T-KPRM algorithm can further decompose the high dimensional principal component subspace and the residual subspace of the KPLS model. In the principal component subspace, the output related part and the output independent part will be separated. Meanwhile, in the residual subspace, the larger residual and the final noise will be separated. Thus, T-KPRM can accurately extract the variable information related to the output from process data and overcome the influence of outlier by adaptively assigning different weights to each sample data, obtaining exhaustive assessment models for each performance grade.

Assuming that the process data matrix $\mathbf{X} \in R^{N \times J}$ consists of N samples with J process variables, and the output data matrix $\mathbf{y} \in R^N$ only contains one output process variable. The radial basis function is selected as the kernel function $\mathbf{K}$: $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right)$, where σ is the standard deviation parameter. The steps of the T-KPRM algorithm are summarized as follows,

- Step 1 The columns of the input data matrix $\mathbf{X}$ are processed by zero mean and unit variance. Similarly, the output data matrix $\mathbf{y}$ is standardized.
- Step 2 The input data matrix $\mathbf{X}$ is projected to the high dimensional feature space F by nonlinear mapping function $\Phi: \mathbf{x}_i \in R^J \rightarrow \Phi(\mathbf{x}_i) \in F$, and obtain the kernel matrix $\mathbf{K}$, which is standard.
- Step 3 Input kernel matrix $\mathbf{K}$ and output matrix $\mathbf{y}$ are processed by PRM algorithm. Initialize the weight $\mathbf{W}_i$ by using the formulas (11), (13), and (15) and compute weighted input kernel matrix $\mathbf{K}_W$ and output data matrix $\mathbf{y}_W$. Then the PLS1 regression model is established after the weighted $\mathbf{K}_W$ and $\mathbf{y}_W$, and the score vector is modified. Calculate the residual $\mathbf{r}_i$ of each sample data, and update the weight $\mathbf{W}_i$ by using the formulas (11), (13) and (15) until the convergence condition is satisfied.
- Step 4 The input data matrix is $\mathbf{K}_W$, the output data matrix is $\mathbf{y}_W$. Extract convergence of $\mathbf{u}_i$ from output matrix $\mathbf{y}_W$, where $i = 1, \mathbf{K}_{Wi} = \mathbf{K}_W$, $\mathbf{y}_{Wi} = \mathbf{y}_W$. (i) Make any column in the $\mathbf{y}_{Wi}$ equal to $\mathbf{u}_i$. (ii) Calculate the score vector of $\mathbf{K}_W$, $\mathbf{t}_i = \mathbf{K}_{Wi}\mathbf{u}_i$, $\mathbf{t}_i \leftarrow \mathbf{t}_i / \|\mathbf{t}_i\|$. (iii) $\mathbf{q}_i = \mathbf{y}_{Wi}^T \mathbf{t}_i$;

- (iv) Calculate the score vector of $\mathbf{y}_{Wi}$, $\mathbf{u}_i = \mathbf{y}_{Wi} \mathbf{q}_i$, $\mathbf{u}_i \leftarrow \mathbf{u}_i / \|\mathbf{u}_i\|$.
- (v) Determine whether $\mathbf{u}_i$ meet the convergence condition, if the condition is satisfied, then turn to step 5, otherwise go back to (i);

- Step 5 Calculate the load matrix of $\mathbf{K}_{Wi}$, $\mathbf{p}_i = \mathbf{K}_{Wi}^T \mathbf{t}_i$. Extract all the main elements, compute $\mathbf{T}$, $\mathbf{P}$, $\mathbf{U}$, $\mathbf{Q}$. (i) Make $\mathbf{K}_{Wi+1} = \mathbf{K}_{Wi} - \mathbf{t}_i \mathbf{p}_i^T$, $\mathbf{y}_{Wi+1} = \mathbf{y}_{Wi} - \mathbf{u}_i \mathbf{q}_i^T$. (ii) Command $i = i + 1$, repeat steps 5 and 6 until the main element $\mathbf{A}$ is extracted [4], the number of principal components can be determined by cross validation method. (iii) $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_A]$, $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_A]$, $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_A]$, $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_A]$. Then, $\mathbf{K}_W = \mathbf{T} \mathbf{P}^T + \mathbf{E}$, $\mathbf{y}_W = \mathbf{U} \mathbf{Q}^T + \mathbf{f}$.
- Step 6 The main element $\mathbf{T} \mathbf{P}^T$ is processed by PRM algorithm: $\mathbf{T} \mathbf{P}^T = \mathbf{t}_y \mathbf{p}_y^T + \mathbf{T}_o \mathbf{P}_o^T$ and the residual $\mathbf{E}$ is processed by PRM algorithm: $\mathbf{E} = \mathbf{T}_r \mathbf{P}_r^T + \mathbf{E}_r$.

If the output data matrix $\mathbf{y}$ is a single output variable, the T-KPRM model is as follows,

$$\begin{cases} \mathbf{K}_W = \mathbf{t}_y \mathbf{p}_y^T + \mathbf{T}_o \mathbf{P}_o^T + \mathbf{T}_r \mathbf{P}_r^T + \mathbf{E}_r \\ \mathbf{y}_W = \mathbf{t}_y + \mathbf{f} \end{cases} \quad (16)$$

where the score $\mathbf{t}_y$ represents the direct parts related to $\mathbf{y}_W$ in $\mathbf{T}$. $\mathbf{t}_y$ is also the most important variable in the evaluation model. The score $\mathbf{t}_y$ represents the orthogonal parts related to $\mathbf{y}_W$, while $\mathbf{T}_r$ is part of a larger variance in $\mathbf{E}$, and $\mathbf{E}_r$ is the final residual part.

When a new sample is introduced, the score matrix will be calculated as

$$\mathbf{t}_y = \mathbf{y}_W^T \mathbf{T} (\mathbf{U}^T \mathbf{K}_W \mathbf{T})^{-1} \mathbf{U}^T \mathbf{k}_{\text{new}} = \mathbf{g}_y \mathbf{k}_{\text{new}} \quad (17)$$

where $\mathbf{k}_{\text{new}}$ is the kernel function of a new sample data $\mathbf{x}_{\text{new}}$ and calculated as

$$\mathbf{k}_{\text{new}} = \Phi(\mathbf{X}) \Phi(\mathbf{x}_{\text{new}}) = [\mathbf{k}(\mathbf{x}_1, \mathbf{x}_{\text{new}}), \dots, \mathbf{k}(\mathbf{x}_n, \mathbf{x}_{\text{new}})]^T \quad (18)$$

and the mean centering of $\mathbf{k}_{\text{new}}$ is

$$\bar{\mathbf{k}}_{\text{new}} = \mathbf{k}_{\text{new}} - \mathbf{K} \mathbf{1}_t - \mathbf{1}_n \mathbf{k}_{\text{new}} + \mathbf{1}_n \mathbf{K} \mathbf{1}_t \quad (19)$$

with $\mathbf{1}_t = 1/n \cdot [1, 1, \dots, 1]^T \in R^n$.

The score vector $\mathbf{t}_y$ obtained by the T-KPRM algorithm can be used to establish an offline assessment model of industrial process [7], which contains the information related to the comprehensive economic benefit.

3.2. Off-line modeling of performance grade

Before establishing the evaluation model, the historical data can be divided into several data sets according to the comprehensive economic benefit, and each data set will represent the corresponding performance grade. For complex industrial processes, it is appropriate to set up 3 to 4 performance grades [5]. Setting too many performance grades, will make the performance grades very complicated, while setting too few performance grades, will reduce the accuracy of the assessment results. In this paper, four kinds of performance grades, i.e., poor, general, suboptimal, and optimal, are used and numbered as 1, 2, 3, and 4 for simplicity, and they correspond to the level of comprehensive economic benefit from low to high respectively. The modeling data sets for each performance grade are denoted as $(\mathbf{X}_c, \mathbf{y}_c)$, $c = 1, 2, 3, 4$, where $\mathbf{X}_c = [\mathbf{x}_c(1), \mathbf{x}_c(2), \dots, \mathbf{x}_c(N_c)]^T \in R^{N_c \times J}$, $\mathbf{y}_c \in R^{N_c}$, and N_c is the number of sample data. The offline evaluation model $\mathbf{g}_y^c$, and $c = 1, 2, 3, 4$. The score vector form offline evaluation model calculated by the T-KPRM algorithm is expressed as follows, $\mathbf{t}_y^c(i) = \mathbf{g}_y^c \mathbf{k}_c(i)$, $i = 1, 2, \dots, N_c$, where c is the number of performance grades.

3.3. Enforced online performance assessment strategy

In the online evaluation part, taking into account that the single sample data is not sufficient to characterize the production process performance and is vulnerable to interference of disturbances and noises, thus an online sliding data window with width H is introduced as the basic analysis unit [8]. The performance-similarities $\gamma_{on,k}^c$ between the sliding data window and the corresponding performance grade are calculated at the current moment, and the similarity will be used to evaluate the process operating performance. In order to distinguish the deterministic performance grade and the performance grade conversion, the performance-similarity threshold ε ($0.5 < \varepsilon < 1$) will be used. When the maximum value of the performance-similarity $\gamma_{on,k}^p = \max_{1 \leq c \leq 4} (\gamma_{on,k}^c)$ is greater than the threshold ε , the process performance grade can be determined as p . The performance grade conversion is the performance-similarity change gradually from the previous performance grade to the next one, and the value of the similarity will show a gradual increasing trend. Then, if the similarity is less than the threshold ε and its value meets the trend of increasing gradually, we can determine the process operates on a performance grade conversion process.

However, despite the sliding window is used, the online assessment results maybe inexplicable or even non-actual due to process disturbances and uncertainties. Therefore, the online assessment strategy should be exhaustive and robust to the process disturbances and uncertainties. In this paper, we proposed a modified and enforced online operating performance assessment strategy based on three predetermined rules. The main idea is based on the continuity of for the process industries, e.g. chemical process, the industries process operation performance may change from one performance grade to another performance grade, e.g. change from the optimal performance grade to sub-optimal performance grade. However, the change cannot be finished in an instant, generally, the industrial process performance changes gradually from one grade to another one, the industrial process will go through a conversion performance grade, thus the adjacent online performance assessment results cannot be two different certain performance grades. Of course, we can't completely rule out individual extreme cases, which rarely appear in the practical industrial process and not be considered in this paper. Generally speaking, we can take all of the assessment results of moment $k-1$, k and $k+1$ to determine the ultimate assessment result of the current moment k .

The enforced online operating performance assessment strategy is summered as follows,

- Step 1 Construct the online sliding data window $\mathbf{X}_{on,k} = [\mathbf{x}_{on,k-H+1}, \dots, \mathbf{x}_{on,k}]^T$, where H is the width of data window, for the initial running moment the data window can be constructed by copying the latest sample data.
- Step 2 Normalize $\mathbf{x}_{on,k}$ with the mean and standard deviation of each performance grade. Calculate the standard online window data matrix $\bar{\mathbf{X}}_{on,k}^c$, $c = 1, 2, 3, 4$ by using offline modeling.
- Step 3 Extract the scores $\mathbf{t}_{on,k}^c = \mathbf{g}_y^c \bar{\mathbf{k}}_{on,k}^c$ from $\bar{\mathbf{X}}_{on,k}^c$, where $\bar{\mathbf{k}}_{on,k}^c$ is the mean centering of $\mathbf{k}_{on,k}^c = [\mathbf{k}(\mathbf{x}_1^c, \bar{\mathbf{x}}_{on,k}^c), \mathbf{k}(\mathbf{x}_2^c, \bar{\mathbf{x}}_{on,k}^c), \dots, \mathbf{k}(\mathbf{x}_{n_c}^c, \bar{\mathbf{x}}_{on,k}^c)]^T$.
- Step 4 Calculate the Euclidean distance between the online data window and the corresponding performance grade c by $d_{on,k}^c = \|\mathbf{t}_{on,k}^c - \bar{\mathbf{t}}_y^c\|^2$, where $\bar{\mathbf{t}}_y^c$ is the mean of score $\mathbf{t}_y^c$ according to T-KPRM algorithm, $\bar{\mathbf{t}}_y^c = 0$, $d_{on,k}^c = \|\mathbf{t}_{on,k}^c\|^2$, and use $d_{on,k}^c$ to compute the similarity between the sliding data window and the corresponding performance grade c , suppose $d_{on,k}^c \neq 0$,

$$\gamma_{on,k}^c = \frac{1/d_{on,k}^c}{\sum_{p=1}^4 1/d_{on,k}^p} \quad (20)$$

where $\sum_{c=1}^4 \gamma_{on,k}^c = 1$, $0 \leq \gamma_{on,k}^c \leq 1$.

Step 5 Then the online assessment result is determined by the following three predetermined rules.

Rule 1 Firstly, determine whether the discriminate formula $\gamma_{on,k}^p = \max_{1 \leq c \leq 4} (\gamma_{on,k}^c) > \varepsilon$ is satisfied, if satisfied, the current process at moment k can be temporarily determined to belong to a certain process performance grade p . then we need to determine whether the assessment result of moment k consist with that of moment $k-1$, if the assessment result of moment k is consist with that moment $k-1$, it can be finally determined that the current process is belong to a certain process performance grade p . Otherwise, we need further determine whether the assessment results of moment $k+1$ and $k-1$ are the same, if the assessment result of moment $k+1$ is the same with moment $k-1$, it can be determined that the assessment result of moment k is unreliable, and the assessment result of moment k should be consist with those of moment $k-1$ and $k+1$. If the assessment result of moment $k+1$ is different with that of moment $k-1$, but consist with that of moment k , it can be determined that the assessment result of moment k and $k+1$ is reliable, and the performance grade of the process is changing, from certain grade to grade conversion or from grade conversion to certain grade. In the rare instances, the process may jump from one certain grade to another for the existence of strong disturbance, which is not considered in this study.

Rule 2 If rule 1 is unsatisfied and the condition $p = \arg \max \{\gamma_{on,k}^p | \gamma_{on,k-l+1}^p < \dots < \gamma_{on,k}^p, 1 \leq p \leq 4\}$ is satisfied, it can be temporarily determined that the current process belongs to the performance grade conversion which is the performance-similarity change gradually from the previous performance grade to the next one. l is a positive integer and can be defined according to production experience. Then we need to determine whether the assessment result of moment k consist with that of moment $k-1$, if the assessment result of moment k consists with that moment $k-1$, it can be finally determined that the above assessment result is reliable. Otherwise, we need further determine whether the assessment results of moment $k+1$ and $k-1$ are same, if the assessment result of moment $k+1$ is same with moment $k-1$, it can be determined that the assessment result of moment k is unreliable, and the assessment result of moment k should consist with that of moment $k-1$.

Rule 3 If both rule 1 and 2 are unsatisfied due to the influence of process disturbances and uncertainties, e.g. the assessment results of moment $k-1$, k and $k+1$ are all different, the final assessment result of moment k and $k+1$ is considered unchanged and consist with the previous assessment results of moment $k-1$.

Through the above three predetermined rules, we can effectively determine the certain performance grade or the performance grade conversion of the current moment. The enforced online assessment procedure is shown in Fig. 1.

3.4. Non-optimal cause identification

When the process operating performance is non-optimal, it is meaningful to find the corresponding causes variables. In the data-based fault diagnosis, the application of the contribution plots [21] method is common. Especially for the linear model, the method can be used to find out the relevant variables which are responsible for fault. However, the traditional contribution plots method is not suitable for the model established by kernel function. Recently, Peng *et al.* [22] proposed a

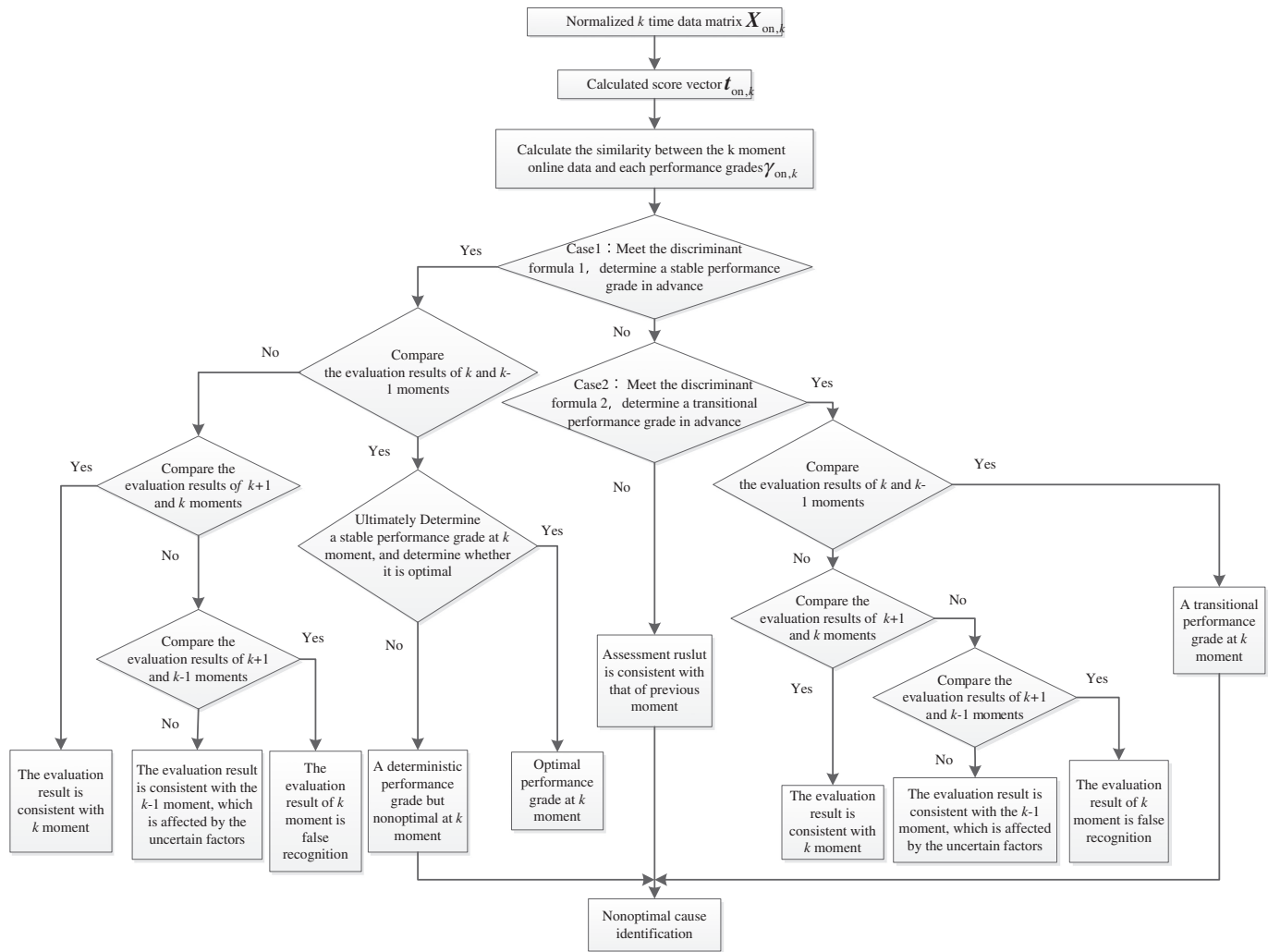


Fig. 1. Procedure of the online assessment and non-optimal cause identification.

contribution plots method based on T-KPLS to identify the faulty variables. In this section, the method of non-optimal cause identification based on T-KPRM is used. By calculating the partial derivative of the Euclidean distance which is between the online sliding data window and optimal performance grade, we can obtain the variable contribution, and the variable with the largest contribution is chosen as the cause variable for non-optimal operating performance.

The distance $d_{on,k}^4$ between the online sliding data window and optimal performance grade can be re-expressed as follow,

$$d_{on,k}^4 = t_{on,k}^{4T} t_{on,k}^4 = \bar{k}_{on,k}^{4T} \mathbf{g}_y^{4T} \mathbf{g}_y^4 \bar{k}_{on,k}^4 = \text{tr}(\mathbf{g}_y^4 \bar{k}_{on,k}^4 \bar{k}_{on,k}^{4T} \mathbf{g}_y^{4T}) \quad (21)$$

where $\mathbf{g}_y^4$ and $\bar{k}_{on,k}^4$ are the offline evaluation model matrix and mean centered kernel vector of online sliding data window with respect to optimal operating performance grade, respectively. Then the contribution of variable j to $d_{on,k}^4$ is calculated as

$$\begin{aligned} \text{Contr}_j &= \left| \frac{\partial d_{on,k}^4}{\partial \bar{x}_{on,k,j}^4} \right| \cdot \bar{x}_{on,k,j}^4 = \left| \frac{\partial}{\partial \bar{x}_{on,k,j}^4} \text{tr}(\mathbf{g}_y^4 \bar{k}_{on,k}^4 \bar{k}_{on,k}^{4T} \mathbf{g}_y^{4T}) \right| \cdot \bar{x}_{on,k,j}^4 \\ &= \left| \text{tr} \left[\mathbf{g}_y^4 \left(\frac{\partial \bar{k}_{on,k}^4}{\partial \bar{x}_{on,k,j}^4} \bar{k}_{on,k}^{4T} \right) \mathbf{g}_y^{4T} \right] \right| \cdot \bar{x}_{on,k,j}^4, j = 1, 2, \dots, m, \end{aligned} \quad (22)$$

where $\bar{x}_{on,k,j}^4$ is the j th variable of the sample data window $\bar{\mathbf{x}}_{on,k}^4$. $\text{tr}(\cdot)$ represents the trace of a matrix.

4. Simulation Study: Dense Medium Coal Preparation Process

Dense medium coal preparation process, also known as heavy medium coal preparation process has been widely used to upgrade run-of-mine coal in the modern coal industry by separating gangue from product coal to produces metallurgical coal or power station coal [23,24]. In order to ensure the stability of product quality and maximum the comprehensive economic benefit, the dense medium coal preparation process is expected to under long-term stable optimal operating performance. However, dense medium coal preparation process always suffers from some process disturbances and uncertainties, such as fluctuation of the raw coal property and deviate from the preliminary designed optimal operating point. Thus, developed suitable online operating performance assessment and non-optimal cause identification methods are crucial for the optimal operation of dense medium coal preparation process. In this study, a dense medium coal preparation process is introduced as the research background, and the actual process data collected from this process is used to illustrate the feasibility and effectiveness of the proposed approach.

4.1. Plant description

The coal preparation plant utilizes the heavy medium cyclone to carry out coal sorting, raw coal, and circulating medium are fed into the mixing barrel at the same time, and the mixture is sorted by two products dense medium cyclone, then the overflow and underflow respectively are removed medium, dehydrated and related treatment. Finally, clean coal, middlings and gangue products are obtained. While the medium is recovered, concentrated, and ultimately be recycled by medium system. The whole sorting process is divided into the coal flow process and the media flow process. The sensors used in the process include the density meter, the magnetic material, the liquid level meter, and so on. The density meter is used to monitor dielectric barrel suspension density, the liquid level gauge is used to monitor the liquid level of dielectric medium barrel, the magnetic content meter is used to monitor the coal slime content of sorting process. In addition, the whole system with all kinds of solenoid valve controls medium, water, and other flow. The density of heavy medium suspension is the main control object in the process of heavy medium separation. The control of the medium density contains density feedback and ash feedback. The density feedback is used to closed-loop control the density of heavy medium suspension, and the stability of the suspension density is ensured by the control method of the water supply, the diversion, and the adding media. Ash feedback is used to monitor the quality of the product, adjust the given value of heavy medium suspension density based on ash results and ensure the stability of the product quality [25,26]. Fast ash test is the original coal ash indexing which is the basic index of coal quality in coal separating plant. The smaller the ash index is, the less the coal dust is, and the better the quality of coal is.

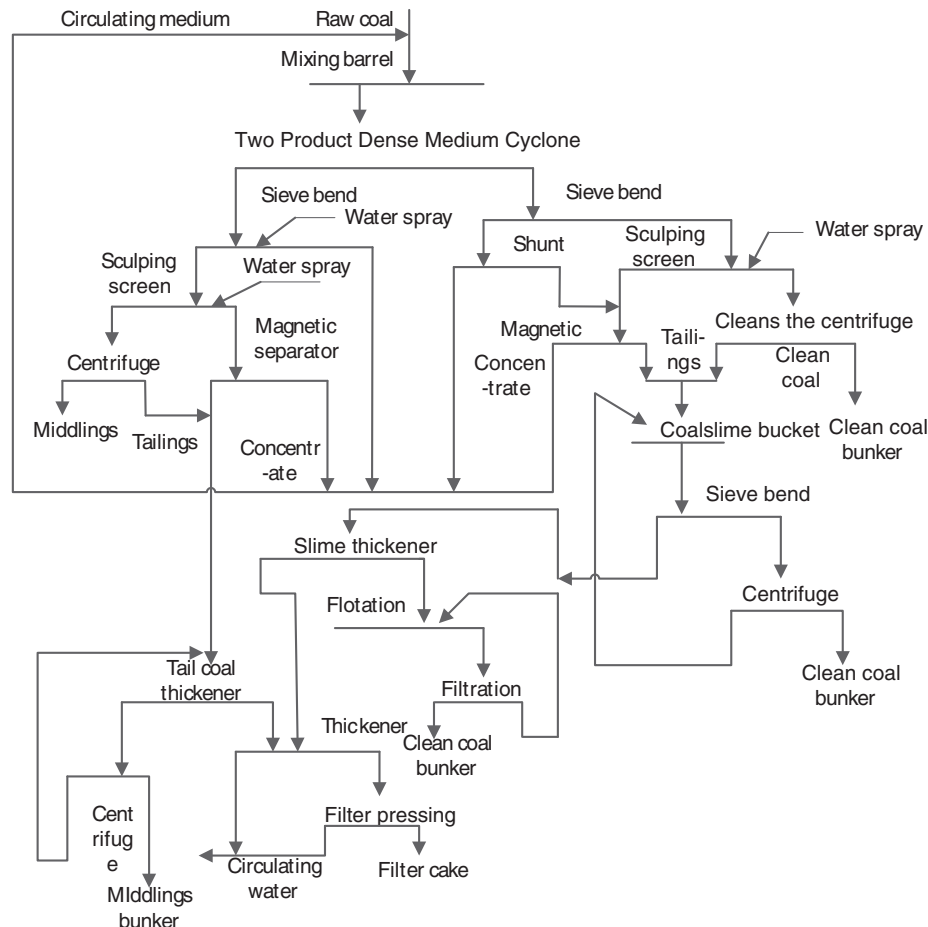
Fig. 2 shows the schematic diagram of dense medium coal preparation process.

4.2. Offline modeling for performance assessment

We select five process variables and one output variable for online operating performance assessment of the dense medium coal preparation process. These five process variables are main density/ $\text{g} \cdot \text{L}^{-1}$, main liquid level/m, main pressure/MPa, magnetic concentration/ $\text{g} \cdot \text{L}^{-1}$, coal slime content/%, respectively. Then, the one output variable is: ash index of product coal/%. A total of 9900 samples are collected for establishing operating performance model. According to the practical production process, the dense medium coal preparation process is divided into four kinds of performance grade and corresponding performance grade conversion. Table 1 lists the division criterion for assessment in dense medium coal preparation process.

Table 1
The division criterion for assessment

Ash index of product coal/%	The process operating performance
6.0–6.5	Optimal
6.5–6.7	Conversion
6.7–7.2	Suboptimal
7.2–7.5	Conversion
7.5–8.0	General
8.0–8.2	Conversion
8.2–9.0	Poor



It is worth to note that the 'Optimal' in Table 1 is defined according to the ash index 6.0–6.5 for the real process under study, which is different with the optimal defined by an optimization problem. The 'Optimal' in the table contains the optimal point and near optimal points, thus the 'Optimal' in Table 1 can also be considered as near optimal. In this section, the historical data of performance grade poor, general, suboptimal and optimal are collected from the actual process and constitute the modeling datasets (X_1, y_1) , (X_2, y_2) , (X_3, y_3) and (X_4, y_4) respectively, and the evaluation model $g_c^*(c = 1, 2, 3, 4)$ of each stable performance grade is established by the proposed T-KPRM algorithm.

4.3. Online assessment without outliers

In online assessment part, another 640 samples including four performance grades and the performance grade conversions are collected to verify the effectiveness of the proposed method. Then, the four performance grades: poor, general, suboptimal, and optimal respectively correspond to the number of test samples 103, 230, 130, and 112. The three performance grade conversions from poor to general, from general to suboptimal, and from suboptimal to optimal respectively correspond to the number of test samples 23, 23, and 19. Set the parameters as $H = 4$, $l = 1$, and $\varepsilon = 0.8$. In the simulation, the process operating performances in the test samples are poor $\rightarrow$ conversion $\rightarrow$ general $\rightarrow$ conversion $\rightarrow$ suboptimal $\rightarrow$ conversion $\rightarrow$ optimal. The online assessment results are shown in Fig. 3.

Fig. 3(a)–(d) shows the performance-similarities between the online sliding data window and each of the performance grades. For the sample moment 1–103, the performance-similarities between the online data window and performance grade poor are greater than 0.8, which indicating that the current process operating performance is poor based on the assessment rule 1. Similarly, for the sampling

moment 127–355, 380–509, and 529–640, the performance-similarities between the online data window and performance grade general, suboptimal, and optimal respectively are greater than 0.8, thus these illustrate that the corresponding process operating performance is general, suboptimal, and optimal separately.

In addition to the four stable performance grades, at the left sampling moment the process runs into performance grade conversions. For instance, for the sampling moment 104–126, the performance-similarities between the online data window and performance grade poor are no greater than 0.8, and the successive ascending conditions $2 = \arg \max_{1 \leq p \leq 3} \{\gamma_{106} | \gamma_{104} < \gamma_{105} < \gamma_{106}\}$ are satisfied. Therefore, it

can be identified that the process is operating on the performance grade conversion from poor to general based on the assessment rule 2. Similarly, at the sampling moment 356–378, 510–528, it can be determined that the process operating performance is conversing from general to suboptimal and from suboptimal to optimal respectively. Fig. 3(e) shows the assessment results based on the similarities, the vertical axis 1,2,3,4 indicate the performance grade poor, general, suboptimal, and optimal in turn 1.5, 2.5, 3.5 represent three performance grade conversions respectively.

4.4. Online assessment with outliers

Process data is often subjected to a variety of disturbances (such as sensor failures, process disturbances, etc.), which leads to the existence of outliers. The existence of outliers may make the modeling analysis method lose its proper generalization ability, leading to the decrease of the accuracy of the assessment models. Thus the robust online assessment method is needed. In this section, in order to simulate the outliers, we choose to introduce the perturbation to the modeling data sets

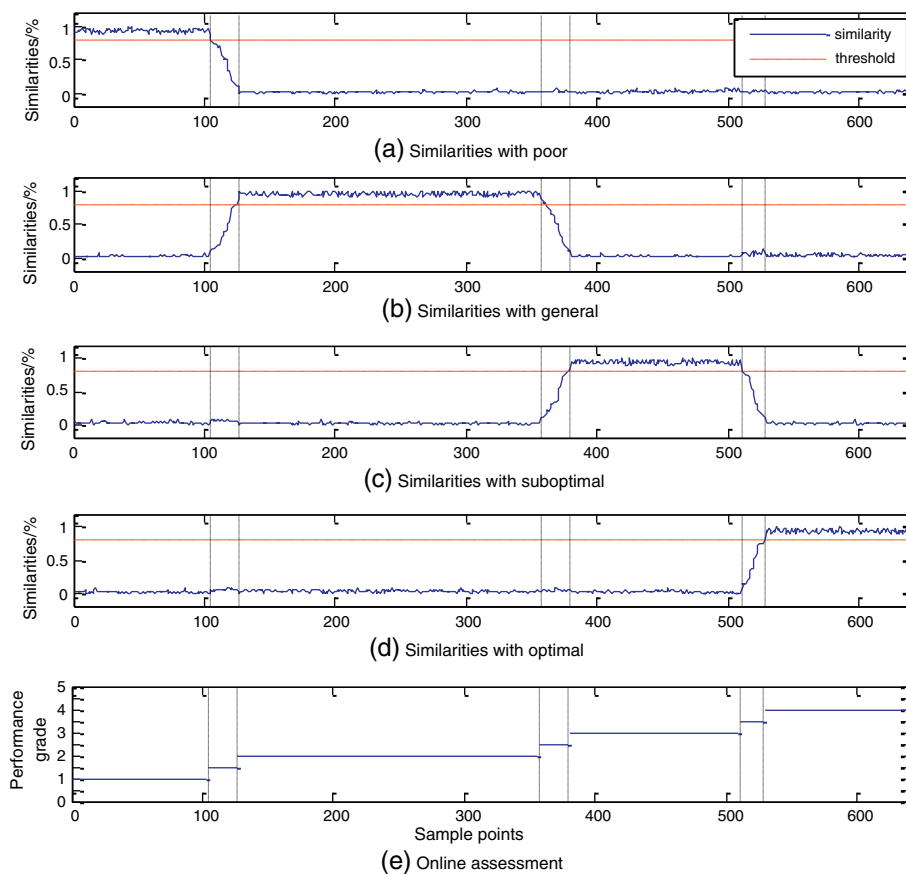


Fig. 3. Similarities and online assessment results of the proposed method.

(X_1, y_1) , (X_2, y_2) , (X_3, y_3) , and (X_4, y_4) . Taking into account that, in the actual production, the number of modeling data sets of optimal and suboptimal performance grades are less than that of poor and general performance grades, 5% data of the former modeling data sets are

randomly selected to be introduced with disturbances. The way of introducing disturbance is that the 5% samples ('outliers') selected are divided into two equal parts, and the input of the first part are introduced with 30% disturbance, and the output of the other part are introduced

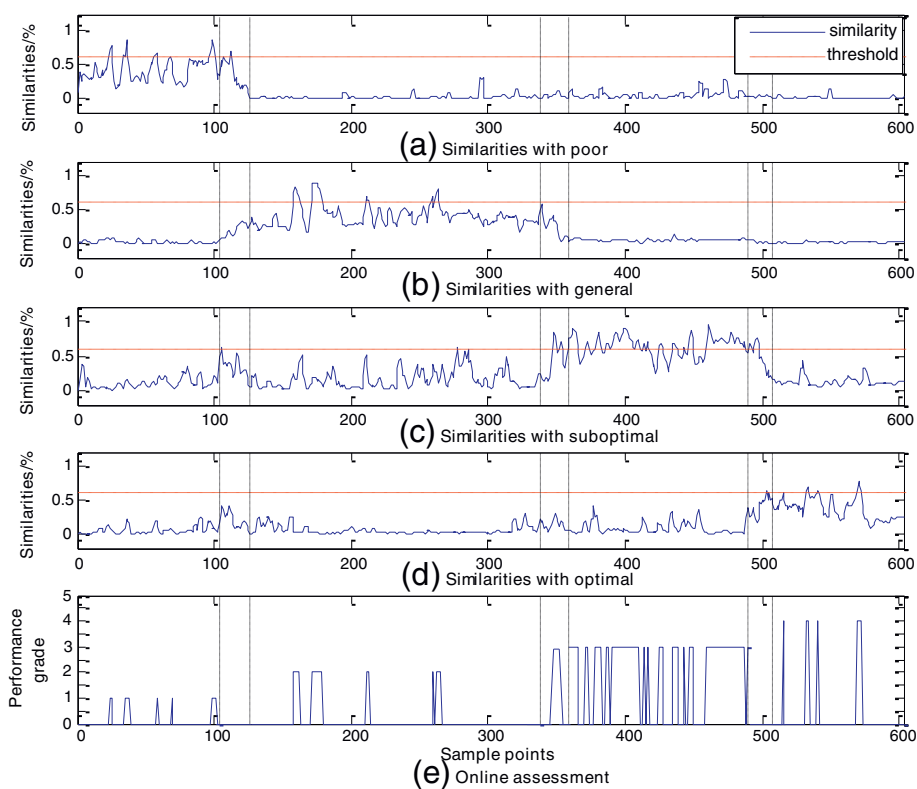


Fig. 4. Similarities and online assessment results with outliers of the existing method (T-KPLS_based).

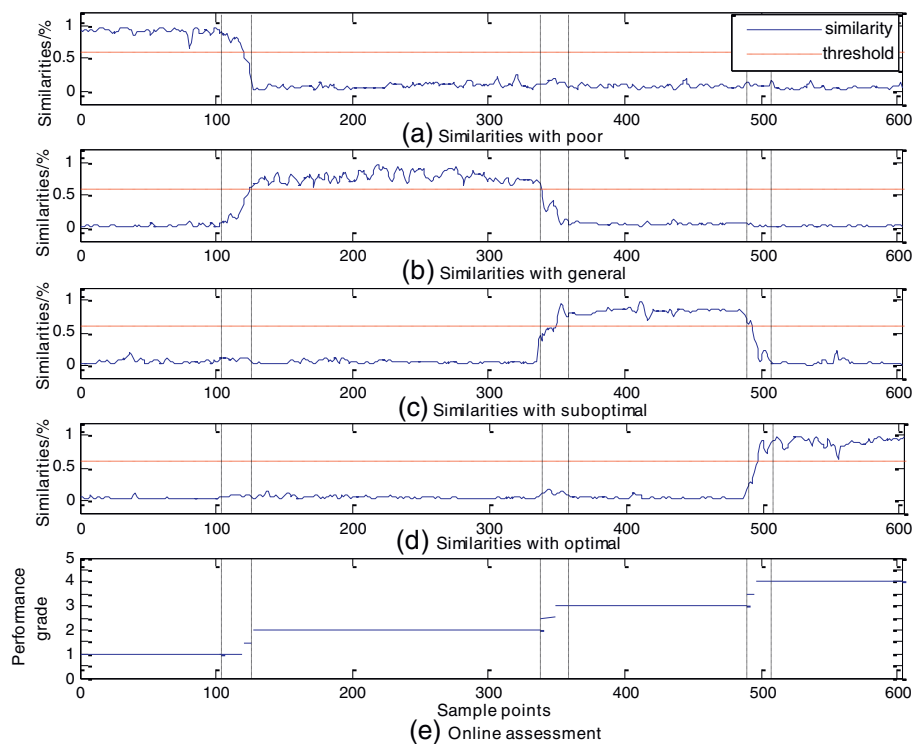


Fig. 5. Similarities and online assessment results with outliers of the proposed method (T-KPRM_based).

with added 30% disturbance. Meanwhile, 10% data of the modeling data of poor and general performance grades are randomly selected to be introduced with disturbance in the similar way as former respectively. The evaluation model $g_y^c(c=1,2,3,4)$ of each stable performance grade is established by the proposed method (noted as T-KPRM_based). Meanwhile, for comparison, the evaluation model $g_y^c(c=1,2,3,4)$ of each stable performance grade is also established by the existing method (noted as T-KPLS_based) mentioned in [8].

Another 603 samples with outliers including four performance grades and the performance grade conversions are collected to verify the robustness of the proposed method. The four performance grades are poor, general, suboptimal, and optimal respectively with the number of test samples 103, 211, 130, and 96. And the three performance grade conversions are from poor to general, from general to suboptimal, from suboptimal to optimal respectively with the number of test samples 23, 21, 19. Figs. 4 and 5 show the evaluation results of the proposed method and the existing method.

It can be seen that the assessment result of the existing method is not available for actual application due to the effect of outliers. On the other hand, the method proposed in this paper can overcome the effect of outliers and shown better robustness. The assessment result of the proposed method is similar with the assessment results without outliers mentioned in Section 4.3, even though there is time delay owing

to the small value of the similarities threshold and use of data window, which is acceptable for actual application.

For further comparison, Tables 2 and 3 present the assessment identification accuracy rate of the existing method and the proposed method for certain performance grades respectively. The value of the performance-similarities threshold ε is chosen as 0.6, 0.7, and 0.8. It can be seen that the assessment identification accuracy rate of the proposed method is much better than the existing method, which indicates that the proposed method is robust to outliers. It can be also seen from Table 3 that bigger similarities threshold may lead to lower identification accuracy of the proposed performance assessment method for the similarities between the online sliding data window and performance grades emerge obvious fluctuations due to the effect of outliers and noises. In the contrary, smaller similarities threshold may lead to time delay as shown in Fig. 5. Therefore, the value of the similarities threshold should be reasonably determined according to actual production process, based on an overall analysis of identification accuracy and time delay. For the case used in this study, we suggest the suitable value of performance-similarities threshold is between 0.6 and 0.8.

4.5. Non-optimal cause identification

The non-optimal performance grade means that the industrial process is operating on the other stable performance grade or the performance grade conversion except optimal. The non-optimal cause is usually unchanging when the process operates on a non-optimal performance grade or a performance grade conversion. Thus, the identification result will not be updated until the assessment result changed.

The corresponding non-optimal cause identification results for all the non-optimal operating performances are shown in Fig. 6. The Z axis represents the variable contribution rate, the X axis indicate five process variables main density, main liquid level, main pressure, magnetic concentration and coal slime content and the six sampling moments on the Y axis represent respectively correspond to the non-optimal operating performances. For example, at 515 sample moment, the variable contribution rate of main pressure and main liquid level is higher than the other process variables, thus the major non-optimal cause is main pressure and main liquid level at this moment. For 472 sample moment, the major non-optimal cause is main pressure and main density. In the same way we can find major non-optimal cause at sample moment 50, 115, 310, and 365 respectively. In summary,

Table 2
The assessment identification accuracy rate of the existing method (T-KPLS_based)

Similarities threshold	Poor	General	Suboptimal	Optimal
$\varepsilon \geq 0.6$	13%	10%	59%	10%
$\varepsilon \geq 0.7$	6%	6%	36%	2%
$\varepsilon \geq 0.8$	2%	3%	8%	0%

Table 3
The assessment identification accuracy rate of the proposed method (T-KPRM_based)

Similarities threshold	Poor	General	Suboptimal	Optimal
$\varepsilon \geq 0.6$	100%	100%	100%	100%
$\varepsilon \geq 0.7$	99%	89%	98%	97%
$\varepsilon \geq 0.8$	97%	83%	82%	93%

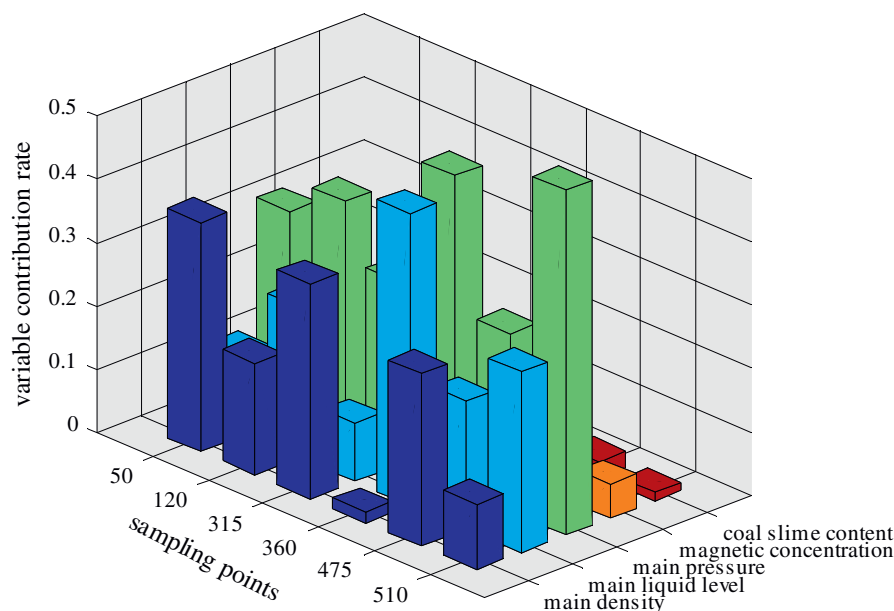


Fig. 6. Non-optimal cause identification results.

the results of online assessment and non-optimal cause identification can guide the operator timely adjust the control strategy of industrial production, improved the performance of the dense medium coal preparation process, and ensure the product coal quality.

5. Conclusions

In this paper, we proposed a novel approach for operating performance assessment of complex nonlinear industrial process based on modified robust total kernel partial M-regression, T-KPRM. The proposed approach T-KPRM can cope with nonlinearities and outliers simultaneously by combing the advantage of T-PLS and KPRM algorithm, and effective extract the optimal-index-related process variation information from process data to establish assessment models for each performance grade. In the online performance assessment part, we propose an enforced online identification framework taking account of the effects of process disturbances and uncertainties, and the process operational performance can be evaluated accurately based on the predetermined assessment rules. Further, the possible responsible cause for non-optimal operating performance can be found by variable contribution-based identification strategy. Finally, the proposed approach is illustrated with a real industrial case of dense medium coal preparation process, and the results show the efficiency of the proposed method comparing to the existing method, which have been confirmed by the actual process data from the dense medium coal preparation process without outliers.

Acknowledgements

The authors are grateful to Qi Wu from China University of Mining and Technology for his kind work on the simulation and associate professor Ruhai Lei from China University of Mining and Technology for sharing the real process data used in the case study.

References

- [1] S. Pashah, A. Moinuddin, S.M. Zubair, Thermal performance and optimization of hyperbolic annular fins under dehumidifying operating conditions—analytical and numerical solutions, *Int. J. Refrig.* 65 (2016) 42–54.
- [2] P. Wang, C. Yang, X. Tian, et al., Adaptive nonlinear model predictive control using an on-line support vector regression updating strategy, *Chin. J. Chem. Eng.* 22 (7) (2014) 774–781.
- [3] Z. Fei, K. Liu, B. Hu, et al., An efficient latent variable optimization approach with stochastic constraints for complex industrial process, *Chin. J. Chem. Eng.* 23 (10) (2015) 1670–1678.
- [4] F. Xu, H. Jiang, R. Wang, et al., Influence of design margin on operation optimization and control performance of chemical processes, *Chin. J. Chem. Eng.* 22 (1) (2014) 51–58.
- [5] Y. Liu, F. Wang, Y. Chang, Operating optimality assessment based on optimality related variations and nonoptimal cause identification for industrial processes, *J. Process Control* 39 (2016) 11–20.
- [6] L. Ye, Y. Liu, Z. Fei, et al., Online probabilistic assessment of operating performance based on safety and optimality indices for multimode industrial processes, *Ind. Eng. Chem. Res.* 48 (24) (2009) 10912–10923.
- [7] Y. Liu, Y. Chang, F. Wang, Online process operating performance assessment and nonoptimal cause identification for industrial processes, *J. Process Control* 24 (10) (2014) 1548–1555.
- [8] Y. Liu, Y. Chang, F. Wang, et al., Complex process operating optimality assessment and nonoptimal cause identification using modified total kernel PLS, The 26th Chinese Control and Decision Conference. IEEE 2014, pp. 1221–1227.
- [9] D. Zhou, G. Li, S.J. Qin, Total projection to latent structures for process monitoring, *AIChE J.* 56 (1) (2010) 168–178.
- [10] H. Li, D.Y. Xiao, Survey on data driven fault diagnosis methods, *Control Decis.* 26 (1) (2011) 1–9.
- [11] Y. Wang, H. Cao, Y. Zhou, et al., Nonlinear partial least squares regressions for spectral quantitative analysis, *Chemom. Intell. Lab. Syst.* 148 (2015) 32–50.
- [12] Y. Liu, F. Wang, Y. Chang, et al., Operating optimality assessment and nonoptimal cause identification for non-Gaussian multimode processes with transitions, *Chem. Eng. Sci.* 137 (2015) 106–118.
- [13] Y. Xu, X. Deng, Fault detection of multimode non-Gaussian dynamic process using dynamic Bayesian independent component analysis, *Neurocomputing* 200 (2016) 70–75.
- [14] Z. Mao, Y. Chang, L. Zhao, Soft-sensor for copper extraction process in cobalt hydrometallurgy based on adaptive hybrid model, *Chem. Eng. Res. Des.* 89 (6) (2011) 722–728.
- [15] Runda Jia, Zhizhong Mao, Yuqing Chang, Shuning Zhang, Kernel partial robust M-regression as a flexible robust nonlinear modeling technique, *Chemom. Intell. Lab. Syst.* 100 (2) (2010) 91–98.
- [16] I. Hoffmann, S. Serneels, P. Filzmoser, et al., Sparse partial robust M regression, *Chemom. Intell. Lab. Syst.* 149 (2015) 50–59.
- [17] L.I. Gang, Q.I.N. Si-Zhao, J.I. Yin-Dong, et al., Total PLS based contribution plots for fault diagnosis, *Acta Automat. Sin.* 35 (6) (2009) 759–765.
- [18] J.J. Lee, C.H. Lim, H.S. Son, et al., In vitro evaluation of the performance of Korean pulsatile ECLS (T-PLS) using precise quantification of pressure-flow waveforms, *ASAIO J.* 51 (5) (2005) 604–608.
- [19] S. Yin, X. Zhu, O. Kaynak, Improved PLS focused on key-performance-indicator-related fault diagnosis, *IEEE Trans. Ind. Electron.* 62 (3) (2015) 1651–1658.
- [20] J. González, D. Pena, R. Romera, A robust partial least squares regression method with applications, *J. Chemom.* 23 (2) (2009) 78–90.
- [21] J. Liu, D.S. Chen, Fault isolation using modified contribution plots, *Comput. Chem. Eng.* 61 (2014) 9–19.
- [22] K. Peng, K. Zhang, G. Li, et al., Contribution rate plot for nonlinear quality-related fault diagnosis with application to the hot strip mill process, *Control. Eng. Pract.* 21 (4) (2013) 360–369.
- [23] J. Bai, Y. Xing, Y. Chen, Application of three products heavy medium coal preparation process in Tangshan Chun'ao coal preparation plant, *Clean Coal Technol.* 19 (3) (2013) 26–29.
- [24] M. Wang, L. Tang, The introduction and selection of the wear-resistant material used for the heavy medium pipeline in coal preparation plant, *Coal Qual. Technol.* 1 (2011) 69–72.
- [25] D. Dou, J. Yang, J. Liu, et al., A novel distribution rate predicting method of dense medium cyclone in the Taixi coal preparation plant, *Int. J. Miner. Process.* 142 (2015) 51–55.
- [26] J. Chen, K.W. Chu, R.P. Zou, et al., Prediction of the performance of dense medium cyclones in coal preparation, *Miner. Eng.* 31 (5) (2012) 59–70.