

Systems Science & Control Engineering

An Open Access Journal

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tssc20>

Dynamic multimode process monitoring using recursive GMM and KPCA in a hot rolling mill process

Gongzhuang Peng, Keke Huang & Hongwei Wang

To cite this article: Gongzhuang Peng, Keke Huang & Hongwei Wang (2021) Dynamic multimode process monitoring using recursive GMM and KPCA in a hot rolling mill process, Systems Science & Control Engineering, 9:1, 592-601, DOI: [10.1080/21642583.2021.1967220](https://doi.org/10.1080/21642583.2021.1967220)

To link to this article: <https://doi.org/10.1080/21642583.2021.1967220>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 18 Aug 2021.



[Submit your article to this journal](#)



Article views: 1014



[View related articles](#)




[View Crossmark data](#)



Citing articles: 5 [View citing articles](#)

Dynamic multimode process monitoring using recursive GMM and KPCA in a hot rolling mill process

Gongzhuang Peng ^a, Keke Huang^b and Hongwei Wang^c

^aNational Engineering Research Center for Advanced Rolling Technology, University of Science and Technology Beijing, Beijing, People's Republic of China; ^bSchool of Information Science and Engineering, Central South University, Changsha, People's Republic of China; ^cZJU-UIUC Institute, Zhejiang University, Haining, People's Republic of China

ABSTRACT

The increasing competitive market has put forward higher demand for iron and steel production process, which is characterized by high-dimensional, nonlinear and multi-scale coupling. The newly rising internet of things (IoT) and advanced communication technologies have promoted the widespread application of data-driven process monitoring methods. To deal with the multimode and non-stationary properties of hot rolling production, a dynamic multimode process monitoring method is proposed based on the recursive Gaussian mixture model (RGMM) and recursive kernel principal component analysis (RKPCA). The proposed approach is applied to the monitoring of the hot-rolled strip thickness oversizing, and comparative experiments are conducted with KPCA, GMM-KPCA on actual production data. Results show that the proposed method shows better performance than conventional methods in terms of fault detection rate and false alarm rate when detecting time-varying multimode faults. The proposed method has also been integrated into an actual system and has been running smoothly in large steel mills in China.

ARTICLE HISTORY

Received 2 July 2021
Accepted 9 August 2021

KEYWORDS

Multimode process monitoring; recursive Gaussian mixture model; recursive kernel principal component analysis; hot rolling mill

1. Introduction

With the increasing fierce market competition, the demand for multi-species, multi-specifications as well as high-quality products from customers is becoming stronger, which makes the process monitoring of steel production an area of both academic and industrial activity (Ma et al., 2017). The steel production process is composed of multiple interdependent physical and chemical processes, characterized by a heterogeneous, high-dimensional, multi-scale coupling, and complex genetic mechanism (Zhang et al., 2020). Taking the hot rolling production process as an example, it consists of a reheating furnace, a roughing mill, a finishing rolling mill and a cooling and coiler sub-process. The final quality of hot-rolled coil products consists of three types: size (e.g. steel strip thickness and width), shape (e.g. crown, wedge and flatness) and surface (e.g. oxidized scale and longitudinal cracks). Most of these quality indices are generally not measurable on-line. The above characteristics make the traditional mechanism model-based fault diagnosis or process monitoring methods difficult to be applied in the dynamic nonlinear hot rolling process.

After decades of application of advanced sensing, communication technology and distributed control system (DCS), most large steel mills have formed a five-level automation and information architecture that includes basic automation system, process control system, manufacturing execution system, manufacturing management system, and business decision-making system. The newly rising industrial internet of things (IIoT) and cyber physical system (CPS) technologies have further broken the barriers between different information systems and promoted the collection, fusion and storage of multi-source heterogeneous data (Cao et al., 2020; Han et al., 2020; Zhang et al., 2016). The IIoT platform and development of data science enable the widespread application of data-driven process monitoring methods (Nkonyana et al., 2019), in which normal operation conditions are modelled with historical process data and the state of the monitored process is then examined by evaluating the deviation of indicators (Jiang et al., 2019; Quiñones-Grueiro et al., 2019; Zhang et al., 2016). The most typical data modelling methods are based on feature extraction, such as principal component analysis (PCA) and independent component

analysis (ICA) (Guo et al., 2019; Jiang et al., 2016; Jiang & Yan, 2018; Zhou et al., 2016). They extract the main features from the sample that reflect the normality or abnormality by dimensionality reduction. Another category is correlation-related methods, such as canonical correlation analysis (CCA) and partial least squares (PLS) (Jiang & Yan, 2019; Liu et al., 2017; Liu et al., 2018). Some machine learning and deep learning methods have also been applied in the process monitoring due to their powerful feature extraction capability in dealing with high-dimensional problems, such as autoencoder (AE) and its variations, and Bayesian networks (Huang et al., 2020; Lee et al., 2019; Song et al., 2020; Yu & Zhao, 2019).

However, a variety of product specifications and multiple system states bring multimode characteristics to the hot rolling production process. This issue has brought challenges to the process monitoring of complicated and changeable industrial process. Several approaches have been proposed to deal with these issues. Regarding the multimode process monitoring, the improved methods can be mainly divided into two categories: global models and multiple models. Zhang et al. proposed a basis vectors-based common and specific features extraction method and used a Kullback–Leibler distance-based metric to measure the changes in both the features (Zhang et al., 2020). Huang et al. divided the multimode process into a common pattern and a mode-specific pattern and applied the sparse coding method in dictionary learning to obtain the reconstruction error (Huang et al., 2020). Another category of methods is to apply clustering algorithms to divide sample data into different modes and then establish a corresponding monitoring model for each single mode. Chen et al. integrated the K -means algorithm with Just-In-Time-Learning-aided CCA in monitoring processes with operating mode change and slowly time-varying behaviour, and conducted comparative experiments with kernel-based methods (Chen et al., 2021). Song et al. proposed a neighbourhood subtractive clustering algorithm to divide the training dataset into several sub-datasets and constructed the process monitoring model in each mode based on the elastic network (EN). For online monitoring, kNN and the voting strategy were applied to select the most suitable monitoring model and the contribution plot-based method was used for fault diagnosis (Song et al., 2020). The Gaussian mixture model (GMM) has been widely applied in multimode process monitoring to determine the number of modes automatically, since it works without any prior knowledge about the mode number. Wang et al. applied the Dirichlet process Gaussian mixed model (DPGMM) to classify the dataset and the t-distributed stochastic neighbour embedding (t-SNE) to obtain the low-dimensional features. Based on the features, the monitoring index

for fault detection was constructed using the support vector data description (SVDD) (Wang et al., 2020). However, most of the existing GMM-based methods ignore the time-varying behaviour of the monitoring process, which subverts the online adaption of the mode classification and identification.

In addition to the multimode characteristics, the non-stationary property of process parameters and equipment states also makes the hot rolling production time-varying. Several adaptive methods, based on the multivariate statistical process monitoring (MSPM) models, have been proposed. Liu et al. developed a moving window KPCA (MWKPCA) monitoring approach, which adapted the data mean and covariance matrix of the feature space in updating procedure and approximated the eigenvalues and eigenvectors of the Gram matrix in the downdating procedure (Liu et al., 2009). To overcome the heavy computation and memory burden of KPCA, Jaffel et al. proposed a reduced KPCA approach in which the reduced kernel matrix was constructed based on the projection value of the transformed data (Jaffel et al., 2016). Zhang et al. improved the dynamical recursive KPCA method based on the singular value decomposition technique and applied it to the fault detection of the continuous annealing process and the penicillin fermentation process (Zhang et al., 2012). Most of the existing adaptive research studies are based on single-mode process monitoring approaches. In this paper, a method that integrates recursive GMM and recursive KPCA is developed for dynamic and multimode fault detection in the hot rolling process. The proposed hybrid framework takes into account multiple characteristics of the hot rolling production process, including multimode, nonlinear and non-stationary. In this framework, GMM is applied to conduct mode identification, and KPCA is used in the fault detection process, the recursive mechanism is combined with both GMM and KPCA to achieve the completely online adaption of the monitoring process.

The rest of this paper is organized as follows. Section 2 introduces the preliminary knowledge of GMM and KPCA, and presents the updating process of RGMM and RKPCA in detail. Section 3 described the mechanism model and parameters' selection of thickness monitoring. Then the proposed RGMM–RKPCA approach is tested in comparative experiments. Finally conclusions are drawn in Section 5, together with some discussion of potential future work.

2. RGMM and RKPCA

2.1. RGMM

Assuming that a set of samples $\mathbf{X} = \{\mathbf{x}_l \in R^D\}_{l=1}^N$ is generated from K multivariate Gaussian components, the

distribution of x can then be expressed as

$$p(x) = \sum_{k=1}^K \omega_k N(x|\mu_k, \Sigma_k) \\ = \sum_{k=1}^K \frac{\omega_k \cdot \exp\{-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k)\}}{(2\pi)^{D/2} |\Sigma_k|^{1/2}} \quad (1)$$

where ω_k is the weight of the k th Gaussian component, satisfying the constraint $0 \leq \omega_k \leq 1$ & $\sum_{k=1}^K \omega_k = 1$; μ_k, Σ_k is the mean and covariance of the k th component. The whole model parameter set is represented by $\Theta = \{\omega_k, \theta_k\}_{k=1}^K = \{\omega_k, \mu_k, \Sigma_k\}_{k=1}^K$. The log-likelihood of the joint density of sample X can be calculated as

$$\log p(X|\Theta) = \log \prod_{i=1}^N p(x_i|\Theta) = \sum_{i=1}^N \log \sum_{k=1}^K \omega_k p(x_i|\theta_k) \quad (2)$$

The EM algorithm can be used to obtain the model parameters.

In the E-step, the conditional expectation of γ_{lk} is calculated, which denotes the posterior responsibility of sample x_l belonging to the k -th component. γ_{lk} is obtained according to Bayes' theorem,

$$\gamma_{lk} = \frac{\omega_k N(x_l|\theta_k)}{\sum_{k=1}^K \omega_k N(x_l|\theta_k)} \quad (3)$$

In the M-step, the parameters of the probability model are estimated by seeking the maximum value of the log-likelihood function. The Lagrange function is used to find the extreme value.

$$L(\omega, \mu, \Sigma) = \log p(X|\omega, \mu, \Sigma) + \lambda \left(\sum_{k=1}^K \omega_k - 1 \right) \quad (4)$$

Then $\omega_k, \mu_k, \Sigma_k$ yields

$$\mu_k = \frac{\sum_{l=1}^N \gamma_{lk} x_l}{\sum_{l=1}^N \gamma_{lk}} \quad (5)$$

$$\Sigma_k = \frac{1}{\sum_{l=1}^N \gamma_{lk}} \sum_{l=1}^N \gamma_{lk} (x_l - \mu_k)(x_l - \mu_k)^T \quad (6)$$

$$\omega_k = \frac{1}{N} \sum_{l=1}^N \gamma_{lk} \quad (7)$$

When the new data x_{N+1} come, in RGMM the model parameter is updated as follows.

$$\omega_k^{(N+1)} = \frac{1}{N+1} \sum_{l=1}^{N+1} \gamma_{lk}^{(N+1)} = \frac{1}{N+1} (\gamma_{(N+1)k}^{(N+1)} + N\omega_k^{(N)}) \\ = \omega_k^{(N)} + \frac{1}{N+1} (\gamma_{(N+1)k}^{(N+1)} - \omega_k^{(N)}) \quad (8)$$

$$\mu_k^{(N+1)} = \mu_k^{(N)} + \frac{1}{N+1} \frac{\gamma_{(N+1)k}^{(N+1)}}{\omega_k^{(N)}} (x_{N+1} - \mu_k^{(N)}) \quad (9)$$

$$\Sigma_k^{(N+1)} = \Sigma_k^{(N)} + \frac{1}{N+1} \frac{\gamma_{(N+1)k}^{(N+1)}}{\omega_k^{(N)}} \left[(x_{N+1} - \mu_k^{(N)}) \right. \\ \left. \times (x_{N+1} - \mu_k^{(N)})^T - \Sigma_k^{(N)} \right] \quad (10)$$

2.2. RKPCA

Let $X^M = \{x_1, x_2, \dots, x_M\} \in \mathbb{R}^{d \times M}$ be the set of zero-mean sample data. The basic idea of the KPCA method is to map the input space to a high-dimensional space through nonlinear mapping Φ .

The covariance matrix C in the feature space is

$$C = \frac{1}{M} \sum_{\mu=1}^M \Phi(x_\mu) \Phi(x_\mu)^T \quad (11)$$

The eigenvalues $\lambda \geq 0$ and eigenvectors v are calculated by

$$\lambda v = C v = \frac{1}{M} \sum_{i=1}^M \Phi(x_i) \Phi(x_i)^T v \quad (12)$$

Considering that all eigenvectors v can be expressed as a linear expansion of $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)$, that is $v = \sum_{i=1}^M \alpha_i \Phi(x_i)$, then (12) becomes

$$\frac{1}{M} \sum_{i=1}^M \alpha_i \left(\sum_{j=1}^M \Phi(x_k) \cdot \Phi(x_j) \right) (\Phi(x_j) \cdot \Phi(x_i)) \\ = \lambda \sum_{i=1}^M (\Phi(x_k) \cdot \Phi(x_i)) \quad (13)$$

where $k = 1, 2, \dots, M$, define the kernel matrix $K_{\mu\nu} := (\Phi(x_\mu) \cdot \Phi(x_\nu))$, (9) can be simplified to

$$M\lambda K\alpha = K^2\alpha \quad (14)$$

For a new testing sample x , its projection in the eigenvector space is

$$t_k = (v^k \cdot \Phi(x)) = \sum_{i=1}^M (\alpha_i)^k (\Phi(x_i), \Phi(x)) \quad (15)$$

where $k = 1, \dots, l$ (l denotes the number of retained KPCs). The two statistics are calculated as

$$T^2 = [t_{\text{new},1}, t_{\text{new},2}, \dots, t_{\text{new},l}] \Lambda^{-1} [t_{\text{new},1}, t_{\text{new},2}, \dots, t_{\text{new},l}]^T \quad (16)$$

$$SPE = \sum_{j=1}^n (t_{\text{new},j})^2 - \sum_{j=1}^l (t_{\text{new},j})^2 \quad (17)$$

In the recursive KPCA, if a new sample x_{M+1} comes, to reduce the computational complexity of KPCA, we retain the first m eigenvalues of X^M . The parameters of the KPCA model are updated according to the following steps.

$$B = \frac{1}{M} 1_M + A_M A_M^T (k(X, x_{M+1}) - \frac{1}{M} K(X) 1_M) \quad (18)$$

$$g = \sqrt{\frac{M}{M^2 - 1}} \times \sqrt{K(x_{M+1}, x_{M+1}) - 2B^T k(X, x_{M+1}) + B^T K(X) B} \quad (19)$$

$$F = \begin{bmatrix} \Sigma_m & A_m^T k(X, x_{M+1}) \\ 0_m^T & g \end{bmatrix} \quad (20)$$

where $\Sigma_m = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})$, F can be represented with $F = U' \Sigma' D'$ by SVD (Singular Value Decomposition). Then the updated KPCA projection direction is calculated as

$$P'_m = \Phi([X, x_{M+1}]) \begin{bmatrix} A_m & -\frac{1}{g} \sqrt{\frac{M}{M^2 - 1}} B \\ 0_m^T & \frac{1}{g} \sqrt{\frac{M}{M^2 - 1}} \end{bmatrix} \quad (21)$$

$$U' = \Phi([X, x_{M+1}]) A'_m$$

Since the number of the PCs will change over time, we use the cumulative percent variance (CPV) method to determine the number of PCs in RKPCA modelling.

$$CPV(r) = \frac{\sum_{j=1}^r \lambda_j}{\sum_{j=1}^M \lambda_j} \times 100\% \quad (22)$$

The r PCs are chosen with a predetermined CPV. The r -dimensional nonlinear principal component t of the new sample in the feature space is

$$t = A'^T \begin{bmatrix} \Phi(X)^T \\ \Phi(x_{M+1})^T \end{bmatrix} \begin{bmatrix} \frac{M}{M+1} \Phi(x_{M+1}) - \frac{1}{M+1} \Phi(X) 1_M \end{bmatrix} \quad (23)$$

The confidence limit for Q is approximated by

$$Q = \eta \chi^2(h) \quad \eta = \frac{\rho^2}{2\mu} \quad h = \frac{2\mu^2}{\rho^2} \quad (24)$$

where μ is the mean of Q statistics and ρ^2 is the variance. The confidence limit for T^2 is approximated by

$$T^2 = \frac{r(M^2 - 1)}{M(M - r)} F_{r, M-r, \beta} \quad (25)$$

where β is the confidence level.

3. Data source and parameters' selection

3.1. Description of the hot-rolled strip thickness oversizing

With the development of the national economy, industries, such as automobile manufacturing, machining, electronics and electrical appliances, have an increasing demand for strip steel, and at the same time, the quality requirements of products are also getting higher. Due to the large width-to-thickness ratio of strip steel, thickness is one of the most important dimensional quality indicators of strip steel. Thickness accuracy is usually divided into strip thickness deviation and head thickness hit rate. This paper focuses on the latter, which requires the difference between the thickness of the strip head and the required thickness of the finished product to be within the allowable accuracy range. Because the adjustment accuracy of the automatic gauge control (AGC) system is generally $30 \sim 50 \mu\text{m}$, the thickness with more than $\pm 50 \mu\text{m}$ deviation from the standard is defined as a product quality problem. Figure 1 shows the examples of the strip thickness oversizing.

To determine the key parameters of the multimode process monitoring model, we first analyse the thickness influencing factors and change law from the perspective of the mechanism model. When the rolled piece bites into the rolling mill, the roll will give the rolled piece a large rolling force, resulting in plastic deformation. At the same time, the roll system is subjected to a force in the opposite direction and equal in magnitude to cause the roll to bend and deform. The resulting loaded roll gap is the thickness of the rolled strip. Therefore, the thickness setting model is based on the elastic deformation theory of the roll and the plastic deformation theory of the rolled product. Figure 2 shows the schematic diagram of the thickness change of the strip rolling, and the classical P-h curve coupling the spring equation and the rolling force formula, where the rolling force P is the ordinate, thickness and roll gap value is the abscissa. The P-h curve can be used to intuitively analyse the various causes for the thickness difference (Figure 2).

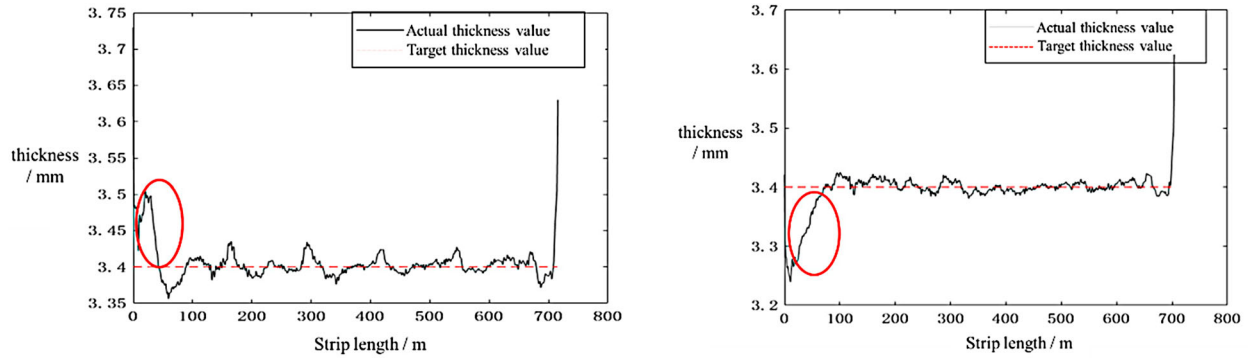


Figure 1. Examples of the strip thickness oversizing.

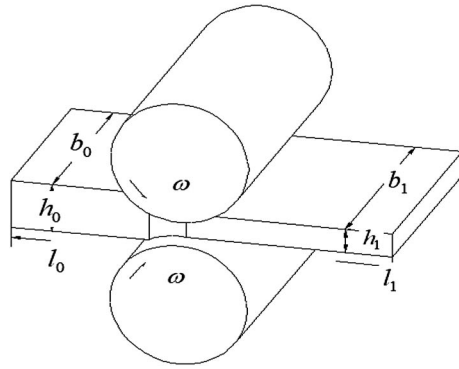


Figure 2. P-h curve of thickness change.

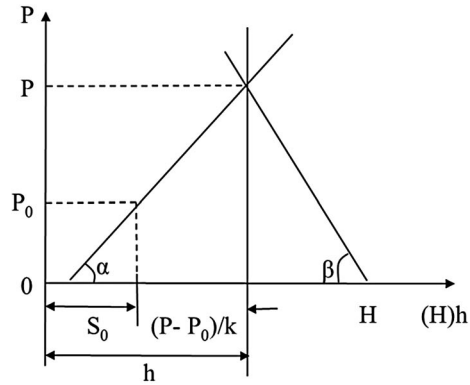
Corresponding to the p-h curve is the spring equation for the rolling mill shown below.

$$h = S_0 + \frac{P - P_0}{K} + S_F + \delta + G \quad (26)$$

where h is the thickness of the rolled piece, S_0 is the no-load roll gap, P is the rolling force, P_0 is the prepressing force shown in Figure 6, K is the stiffness coefficient of the rolling mill, and S_F is the thickness change caused by the bending force, G is the zero position of the roll gap, and δ is the oil film thickness compensation. P is related to strain accumulation, dynamic recrystallization, static recrystallization, phase change and other factors. Existing rolling force models are derived from the concise deduction of relevant influencing factors under specific constraints, and the following expressions can be used.

$$P = B l_c Q_p K K_T \quad (27)$$

where B is the bandwidth, l_c is the horizontal projection length of the contact arc between the roll and the rolled piece, which is related to the radius of the roll. Q_p is the stress status modulus, which is determined by the shape parameter of the deformation zone. K is the metal deformation resistance, which is related to the deformation



temperature, deformation speed and degree of deformation. K_T is the influence coefficient of the front and back tensile stress on the rolling force.

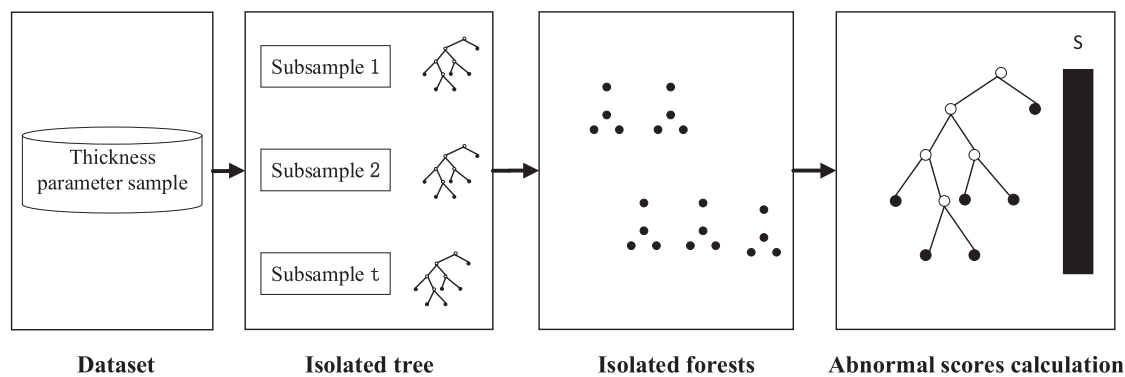
3.2. Parameters' selection

According to the above mechanism model, we selected 48 parameters as inputs of the process monitoring model, which are divided into incoming slab-related, rolling force-related and rolling mill-related parameters. Table 1 shows the range and distribution of some of the variables.

Due to fluctuations in the data acquisition equipment, such as the unstable value of the sensor in a complex and changeable environment, the values of the obtained process parameters will fluctuate. When the fluctuation is large, it will destroy the statistical law of the data set, which is regarded as an outlier; when it is small, this type of fluctuation is regarded as noise. Due to the complex working conditions of the industrial production process and changes in the production environment, the existence of noise is inevitable. Considering the existence of noise during modelling can improve the robustness and stability of the model. Therefore, it is necessary to select an algorithm that can retain noisy data when eliminating outliers. In this paper, the isolated forest (iForest)

Table 1. A part of the selected variables.

Type	Variables	Max	Min	Variance	Mean
Raw material parameters	Intermediate slab temperature°C	1048.7	975.7	11.42	1015.3
	F7 Strip modulus	12.171	7.549	0.715	10.183
Rolling force-related parameters	F7 Strip deformation resistance MPa	232.11	167.59	9.308	203.81
	F7 Head rolling force kN	5.720	4.390	0.233	5.077
	Final rolling temperature deviation°C	16.260	12.400	4.604	−1.098
	Interchange temperature°C	1096.6	1022.5	11.6	1065.1
	temperature drop°C	171.63	87.02	13.31	136.42
	Velocity m/s	9.750	6.890	0.522	8.544
Roller-related parameters	F7 Thickness correction factor	0.012	−0.014	0.004	−0.003
	F7 Elastic deformation mm	0.600	0.460	0.027	0.535
	F7 Roller rigidity MPa	4.393	4.249	0.017	4.344
	F7 Roller wear correction mm	0.260	−0.320	0.134	−0.003

**Figure 3.** Outliers elimination based on iForest.

algorithm is selected to eliminate the abnormal values in the thickness data of the hot-rolled strip production. iForest is a fast outlier detection method based on ensemble learning, which has linear time complexity and high accuracy, and is suitable for continuous numerical data. Unlike other anomaly detection algorithms that use distance and density indicators to describe the degree of alienation between samples, iForest detects outliers by isolating sample points. Compared with traditional algorithms such as LOF (Local Outlier Factor) and K -means, iForest is more robust to high-dimensional data.

The algorithm flow is shown in Figure 3. Given the sample data, to construct an isolated tree, a feature and its segmentation value are randomly selected, and the data set is recursively divided until any one of the following conditions is met: the tree reaches the restricted height; there is only one sample on the node; all features of the samples on the node are the same. The task of anomaly detection is to give a ranking of the degree of abnormality. The commonly used ranking method is based on the path length or anomaly score of the sample points. The path length refers to the number of nodes that the sample travels from the starting node to the final node in the isolated tree.

4. Experimental results

The steel industry IoT platform, independently developed by our team, has been applied to several large- and medium-sized steel companies in China. The big data platform collects the automated control systems of different manufacturers on the hot strip production line and the L2 database management system, and builds a complete factory-level database of quality, process, equipment and energy consumption. To eliminate the influence of steel grade, thickness, width, etc., we set the steel grade as Q195L, the thickness of the intermediate billet is 45 mm, and the thickness of the finished product is 3.4 mm. A total of 2300 pieces of production data were obtained from the field, and 238 abnormal data are eliminated by the isolated forest algorithm. We select 2000 pieces of data from the remaining set as historical data sets for algorithm training and testing. The overall offline learning and online monitoring process is shown in Figure 4.

Two commonly used indices, the FDR (fault detection rate) and the FAR (false alarm rate), are applied in the paper to evaluate the performance of the process monitoring method. FDR reflects the ratio of the number of correctly detected faults to the total number of faults,

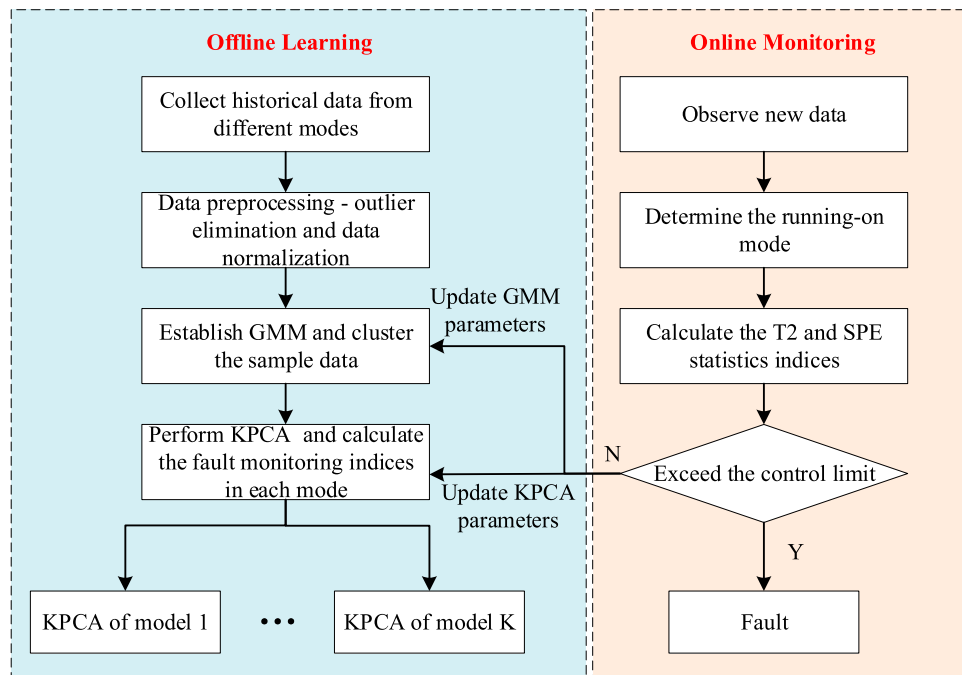


Figure 4. Flowchart of the dynamic multimode monitoring approach.

Table 2. Monitoring results of different methods.

Index		KPCA		GMM-KPCA		RGMM-RKPCA	
		SPE	T ²	SPE	T ²	SPE	T ²
Mode1	FAR	0%	0%	4%	0	1.2%	0.4%
	FDR	82.4%	22%	100%	100%	96%	100%
Mode2	FAR	7.2%	1.2%	2%	2.3%	0.8%	0.4%
	FDR	100%	100%	92.4%	98.2%	98.4%	99.2%

while FAR indicates the proportion of normal samples that are falsely reported as faults in the total number of normal samples. To verify the effectiveness of GMM algorithm in modal recognition and the advantage of RKPCA in dealing with dynamic processes compared with KPCA algorithm, we used three methods for comparative experiments: KPCA, GMM + KPCA, and RGMM + RKPCA. The threshold of the CPV is 90%, and the confidence level of all methods is 99% (Table 2).

Figure 5 shows the results of different methods of monitoring thickness-related variables in the hot rolling production process. In the KPCA method, we first use 1000 normal samples for model training. The test data includes four parts, 250 normal data of mode 1, 250 normal data of mode 2, and 250 faulty data of mode 1, 250 faulty data of mode 2. SPE statistics and T2 statistics of the KPCA can detect the fault of mode 2 well, but the accuracies are low when detecting the fault of mode 1, which can be seen from the FDR value –22%. In addition, FAR of the SPE statistics of KPCA is 7.2% in mode

2, which means that since the KPCA algorithm ignores the data characteristics of different modes, the number of false alarms will increase.

In the latter two algorithms, we first use the GMM algorithm to classify 1000 normal samples, and use the normal samples of each mode to establish the corresponding KPCA model. In the online monitoring process, the GMM algorithm is applied to identify the mode of the new sample, and the corresponding KPCA model is used for diagnosis. The difference between the two models is that the RGMM–RKPCA algorithm will update the model parameters according to the diagnosis results of the samples during the training phase, and the SPE and T2 indicators will also change dynamically. By comparing KPCA with GMM–KPCA, we can see that after mode identification all the fault data in mode 1 are detected with a FAR value of 4%. Although the fault detection rate of mode 2 is reduced by 7.6% and 1.8% with SPE and T2 statistics, respectively, the false alarm rate is also much reduced. It can be inferred that establishing multiple nonlinear monitoring models for different steady-state processes can effectively solve the problem of multi-modality. However, the non-stationary problem is still not well addressed in GMM–KPCA, which can be found by the false alarm rate. It can be seen from Figure 5(e,f) that compared to the KPCA algorithm, the T2 and SPE statistics of GMM–KPCA can better identify fault data in different modes, and RKPCA can further reduce the proportion of false alarms in each mode. By online adapting the sample data of the KPCA

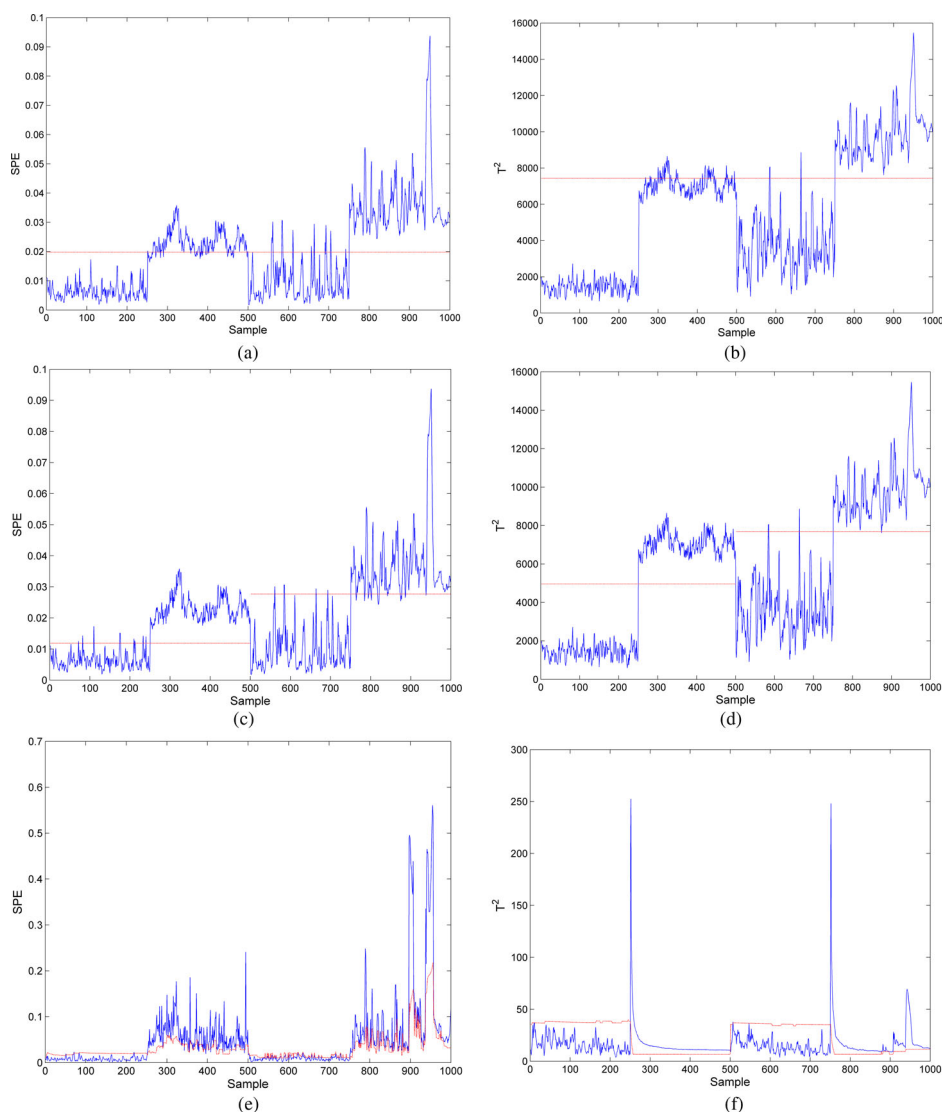


Figure 5. Monitoring results of different methods. (a) and (b) SPE and T2 statistics of the KPCA method. (c) and (d) SPE and T2 statistics of the GMM-KPCA method. (e) and (f) SPE and T2 statistics of the RGMM-RKPCA method.

model, the monitoring statistics in RKPCA also become adaptive and robust to slight faults.

The algorithms proposed in this paper have been running smoothly for six months in the actual system. Figure 6 shows the GUIs of the monitoring software on the site of a steel plant. Figure 6(a) shows the monitoring of different indices in the hot rolling process, such as thickness, width and FDT (finisher delivery temperature). Figure 6(b) illustrates the parameter importance ranking that affects different monitoring indices, and the corresponding reasonable parameter interval recommendation.

5. Conclusion

To better deal with the multimode and time-varying properties of hot rolling production, we propose a

dynamic multimode process monitoring method based on the recursive Gaussian mixture model (RGMM) and recursive kernel principal component analysis (RKPCA) in this study. GMM is applied to classify the dataset and determine the mode for a new sample, and KPCA is used to construct the monitoring index for fault detection in each mode. The proposed approach is applied to the monitoring of the hot-rolled strip thickness oversizing in comparison with KPCA, GMM-KPCA on actual production data. Results show that the proposed method shows better performance than conventional methods in terms of fault detection rate and false alarm rate when detecting time-varying multimode faults.

It is worth mentioning that the proposed method has some shortcomings and there are some research directions for future study. The time complexity of the proposed approach can be improved for real-time



(a) GUI of the process monitoring



(b) GUI of the parameters analysis

Figure 6. Screenshots of the on-site monitoring software interface.

industrial application. As for the nonlinear monitoring methods, incorporation with some emerging deep learning algorithms, such as the recurrent neural network (RNN) and deep belief networks (DBN), could be a promising extension. As the subsequent step of fault diagnosis, the fault location and repair mechanism are necessary in the industrial field to guide the operation of physical entities. Besides, with the higher dimensions of process industrial data, the distributed parallel approach can also be taken into consideration to handle the big process data more efficient.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is supported by the National Natural Science Foundation of China under the grant number 61903031 and the Fundamental Research Funds for the Central Universities

under the grant numbers FRF-TP-18-035A1, FRF-TP-20-050A2. This work was also supported, in part, by the Zhejiang University/University of Illinois at Urbana-Champaign Institute, and was led by Principal Supervisor Prof. Hongwei Wang.

ORCID

Gongzhuang Peng  <http://orcid.org/0000-0002-4395-6323>

References

- Cao, Y., Jia, F., & Manogaran, G. (2020). Efficient traceability systems of steel products using blockchain-based industrial internet of things. *IEEE Transactions on Industrial Informatics*, 16(9), 6004–6012. <https://doi.org/10.1109/tii.2019.2942211>
- Chen, Z., Liu, C., Ding, S. X., Peng, T., Yang, C., Gui, W., & Shardt, Y. A. W. (2021). A just-in-time-learning-aided canonical correlation analysis method for multimode process monitoring and fault detection. *IEEE Transactions on Industrial Electronics*, 68(6), 5259–5270. <https://doi.org/10.1109/tie.2020.2989708>

- Guo, L., Wu, P., Gao, J., & Lou, S. (2019). Sparse kernel principal component analysis via sequential approach for nonlinear process monitoring. *IEEE Access*, 7, 47550–47563. <https://doi.org/10.1109/access.2019.2909986>
- Han, Y., Zhang, C.-J., Wang, L., & Zhang, Y.-C. (2020). Industrial IoT for intelligent steelmaking with converter mouth flame spectrum information processed by deep learning. *IEEE Transactions on Industrial Informatics*, 16(4), 2640–2650. <https://doi.org/10.1109/tii.2019.2948100>
- Huang, K., Wen, H., Zhou, C., Yang, C., & Gui, W. (2020). Transfer dictionary learning method for cross-domain multimode process monitoring and fault isolation. *IEEE Transactions on Instrumentation and Measurement*, 69(11), 8713–8724. <https://doi.org/10.1109/tim.2020.2998875>
- Huang, K., Wu, Y., Yang, C., Peng, G., & Shen, W. (2020). Structure dictionary learning-based multimode process monitoring and its application to aluminum electrolysis process. *IEEE Transactions on Automation Science and Engineering*, 17(4), 1989–2003. <https://doi.org/10.1109/tase.2020.2984334>
- Jaffel, I., Taouali, O., Harkat, M. F., & Messaoud, H. (2016). Moving window KPCA with reduced complexity for nonlinear dynamic process monitoring. *ISA Transactions*, 64, 184–192. <https://doi.org/10.1016/j.isatra.2016.06.002>
- Jiang, Q., & Yan, X. (2018). Parallel PCA–KPCA for nonlinear process monitoring. *Control Engineering Practice*, 80, 17–25. <https://doi.org/10.1016/j.conengprac.2018.07.012>
- Jiang, Q., & Yan, X. (2019). Multimode process monitoring using variational Bayesian inference and canonical correlation analysis. *IEEE Transactions on Automation Science and Engineering*, 16(4), 1814–1824. <https://doi.org/10.1109/tase.2019.2897477>
- Jiang, Q., Yan, X., & Huang, B. (2016). Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and Bayesian inference. *IEEE Transactions on Industrial Electronics*, 63(1), 377–386. <https://doi.org/10.1109/tie.2015.2466557>
- Jiang, Q., Yan, X., & Huang, B. (2019). Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Industrial & Engineering Chemistry Research*, 58(29), 12899–12912. <https://doi.org/10.1021/acs.iecr.9b02391>
- Lee, S., Kwak, M., Tsui, K.-L., & Kim, S. B. (2019). Process monitoring using variational autoencoder for high-dimensional nonlinear processes. *Engineering Applications of Artificial Intelligence*, 83, 13–27. <https://doi.org/10.1016/j.engappai.2019.04.013>
- Liu, Q., Zhu, Q., Qin, S. J., & Chai, T. (2018). Dynamic concurrent kernel CCA for strip-thickness relevant fault diagnosis of continuous annealing processes. *Journal of Process Control*, 67, 12–22. <https://doi.org/10.1016/j.jprocont.2016.11.009>
- Liu, X., Kruger, U., Littler, T., Xie, L., & Wang, S. (2009). Moving window kernel PCA for adaptive monitoring of nonlinear processes. *Chemometrics and Intelligent Laboratory Systems*, 96(2), 132–143. <https://doi.org/10.1016/j.chemolab.2009.01.002>
- Liu, Y., Liu, B., Zhao, X., & Xie, M. (2017). A mixture of variational canonical correlation analysis for nonlinear and quality-relevant process monitoring. *IEEE Transactions on Industrial Electronics*, 65(8), 6478–6486. <https://doi.org/10.1109/TIE.2017.2786253>
- Ma, L., Dong, J., Peng, K., & Zhang, K. (2017). A novel data-based quality-related fault diagnosis scheme for fault detection and root cause diagnosis with application to hot strip mill process. *Control Engineering Practice*, 67, 43–51. <https://doi.org/10.1016/j.conengprac.2017.07.005>
- Nkonyana, T., Sun, Y., Twala, B., & Dogo, E. (2019). Performance evaluation of data mining techniques in steel manufacturing industry. *Procedia Manufacturing*, 35, 623–628. <https://doi.org/10.1016/j.promfg.2019.06.004>
- Quiñones-Grueiro, M., Prieto-Moreno, A., Verde, C., & Llanes-Santiago, O. (2019). Data-driven monitoring of multimode continuous processes: A review. *Chemometrics and Intelligent Laboratory Systems*, 189, 56–71. <https://doi.org/10.1016/j.chemolab.2019.03.012>
- Song, B., Yan, H., Shi, H., & Tan, S. (2020). Multisubspace elastic network for multimode quality-related process monitoring. *IEEE Transactions on Industrial Informatics*, 16(9), 5874–5883. <https://doi.org/10.1109/tii.2019.2959784>
- Wang, B., Li, Z., Dai, Z., Lawrence, N., & Yan, X. (2020). Data-driven mode identification and unsupervised fault detection for nonlinear multimode processes. *IEEE Transactions on Industrial Informatics*, 16(6), 3651–3661. <https://doi.org/10.1109/TII.2019.2942650>
- Yu, W., & Zhao, C. (2019). Online fault diagnosis for industrial processes with Bayesian network-based probabilistic ensemble learning strategy. *IEEE Transactions on Automation Science and Engineering*, 16(4), 1922–1932. <https://doi.org/10.1109/tase.2019.2915286>
- Zhang, F., Liu, M., Zhou, Z., & Shen, W. (2016). An IoT-based online monitoring system for continuous steel casting. *IEEE Internet of Things Journal*, 3(6), 1355–1363. <https://doi.org/10.1109/jiot.2016.2600630>
- Zhang, K., Peng, K., Zhao, S., & Chen, Z. (2020). A novel common and specific features extraction-based process monitoring approach with application to a hot rolling mill process. *Control Engineering Practice*, 104, 104628. <https://doi.org/10.1016/j.conengprac.2020.104628>
- Zhang, Y., Fan, Y., & Du, W. (2016). Nonlinear process monitoring using regression and reconstruction method. *IEEE Transactions on Automation Science and Engineering*, 13(3), 1343–1354. <https://doi.org/10.1109/tase.2016.2564442>
- Zhang, Y., Li, S., & Teng, Y. (2012). Dynamic processes monitoring using recursive kernel principal component analysis. *Chemical Engineering Science*, 72, 78–86. <https://doi.org/10.1016/j.ces.2011.12.026>
- Zhou, B., Ye, H., Zhang, H., & Li, M. (2016). Process monitoring of iron-making process in a blast furnace with PCA-based methods. *Control Engineering Practice*, 47, 1–14. <https://doi.org/10.1016/j.conengprac.2015.11.006>