

Article

Not peer-reviewed version

Fault Detection based Multiple Local Manifold Learning and Its Application to Blast Furnace Ironmaking Process

[Ke Wang](#), [Ping Wu](#)^{*}, Siwei Lou, [Haipeng Pan](#), [Jinfeng Gao](#)

Posted Date: 8 October 2023

doi: 10.20944/preprints202310.0439.v1

Keywords: Data-driven method; Fault detection; Manifold Learning; Blast Furnace Ironmaking Process



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Fault Detection Based Multiple Local Manifold Learning and Its Application to Blast Furnace Ironmaking Process

Ke Wang ¹, Ping Wu ^{1,*}, Siwei Lou ², Haipeng Pan ¹ and Jinfeng Gao ¹

¹ School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China

² College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China

* Correspondence: pingwu@zstu.edu.cn

Abstract: Process safety plays a vital role in the modern process industry. To prevent undesired accidents caused by malfunctions or other disturbances in complex industrial processes, considerable attention has been paid to data-driven fault detection techniques. To explore the underlying manifold structure, manifold learning methods including Laplacian eigenmaps, locally linear embedding, and Hessian eigenmaps have been utilized in data-driven fault detection. However, only the partial local structure information is extracted from the aforementioned methods. In this paper, these typical manifold learning methods are synthesized to find the underlying manifold structure from different angles. A more comprehensive local structure is discovered under a unified framework by constructing an objection optimization function for dimension reduction of the process data. The proposed method takes advantage of different manifold learning methods. Based on the proposed dimension reduction method, a new data-driven fault detection method is developed. Hotelling's T^2 and Q statistics are established for the purpose of fault detection. Experiments on an industrial benchmark Tennessee Eastman process and a real blast furnace ironmaking process are carried out to demonstrate the superiority and effectiveness of the proposed method.

Keywords: data-driven method; fault detection; manifold learning; blast furnace ironmaking process

1. Introduction

Industrial processes are becoming more complex and their hazards to the environment are receiving increasing attention. Process safety is a non-negligible component of industrial processes, which comprises several steps such as hazard identification and analysis [1]. In particular, the identification and analysis of hazards is a key step in the prevention and mitigation of major process accidents.

In industrial processes, timely and accurately identifying abnormal operating conditions can prevent major accidents and improve operational efficiency, thus achieving compliance with environmental and safety regulations. Dynamic process monitoring for hazard/fault identification is already a trend in the future development of process safety and risk management [2]. Real-time monitoring of process operations to ensure safety measures is an essential step in the modern process industry [3].

Fault detection plays a pivotal role in guaranteeing operation safety and reducing downtime in complex industrial processes [4,5]. Broadly speaking, fault detection techniques can be categorized into three classes, model-based, knowledge-based, and data-driven based methods [6]. Model-based methods rely on the mathematical model. However, the mathematical model is often difficult or time-consuming to establish for complex industrial processes such as the blast furnace ironmaking process. For knowledge-based methods, the model is built from expert knowledge or qualitative information, which limits its applications for complex industrial processes. Conversely, only the measured process variables are required for data-driven fault detection methods. Thus, data-driven techniques are more suitable and efficient for the fault detection of complex industrial processes [7].

With the advance of the sensor, communication, and computing technologies, large amounts of data are collected in modern industrial processes. Under such circumstances, data-driven fault detection has gained an explosive amount of attention in recent years both from academia and industry [8].

To handle the highly correlated high-dimensional process data, multivariate analysis (MVA) has been widely employed in industrial processes [9]. In MVA, the process behavior is modeled by transforming the high-dimensional data into a lower-dimensional space. The features are extracted for establishing monitoring statistics. Among the MVA based fault detection methods, principal component analysis (PCA) has gained widespread popularity in process monitoring and fault diagnosis in the past decades [10]. In PCA, process data is projected into a lower-dimensional space to preserve the significant variability information as much as possible. Due to its efficiency and simplicity, PCA has been successfully applied in a large number of industrial processes [11]. Despite this, PCA is regarded as a kind of globality-based linear dimensionality reduction technique. However, the process data mostly lie on or close to a low-dimensional manifold. Compared to globality-based methods, manifold learning is an approach to nonlinear dimensionality reduction through discovering the manifold structure of data. In manifold learning, the input data is assumed to be sampled from a low-dimensional manifold. Representative manifold learning methods include Isomap [12], Locally Linear Embedding (LLE) [13], Laplacian eigenmaps (LE) [14], Local Tangent Space Alignment (LTSA) [15], Locality Preserving Projections (LPP) [16], Neighborhood preserving embedding (NPE) [17], and Hessian eigenmaps (or called Hessian LLE) [18].

In NPE, each data point is represented as a linear combination of the neighboring data points. Then, an optimal embedding is found to preserve the neighborhood structure in the dimensionality reduced space [17]. Chen *et al.* [19] applied eigenvalue decomposition and generalized eigenvalue decomposition to solve the unstable problem caused by singularity problem in NPE, and developed an NPE-based incipient fault detection method for small-scale cyber-physical systems. Since the NPE method can preserve the local manifold structure of different modes, Song *et al.* [20] performed NPE on the time-lagged variables for multimode dynamic process monitoring. LPP is designed to find the optimal linear approximations to the eigenfunctions of the Laplace Beltrami operator on the manifold through the nearest neighbor search in the low-dimensional space [16]. Duan *et al.* [21] employed LPP to preserve the local structure of process data, then adopted least squares support vector machine to predict the key-performance-indicator. Zhang *et al.* [22] combined LPP and PCA to preserve both global and local structures of the data set and developed fault detection and identification method by utilizing the extracted features. LLE attempts to discover nonlinear structure in high-dimensional data by exploiting the local symmetries of linear reconstructions [18]. Wu *et al.* learned structure information by LLE and incorporated the extracted local information into canonical correlation analysis (CCA) for quality-relevant nonlinear process monitoring [23]. Li and Zhang implemented the supervised locally linear embedding projection method for bearing fault diagnosis and illustrated its validity using the experimental data [24]. Different manifold learning methods focus on uncovering the manifold structure with different criteria. It relies on the knowledge and experience of experts for their own purposes. Therefore, only partial information from the underlying manifold is learned by each existing local manifold learning method. To take advantages of different manifold learning methods for better discovering the underlying manifold structure, Xing *et al.* [25] provided a common framework to synthesize the partial information extracted from different local manifold learning methods under local tangent coordinates.

Motivated by the above discussions, a novel data-driven fault detection based on fused local manifold learning (FLML) is proposed in this paper. In the proposed FLML, the partial information on the geometric structure of the underlying manifold is firstly extracted from LE, LLE, and Hessian Locally Linear Embedding (HLLE) methods, respectively. A novel objective function is formulated to fuse the extracted partial information. On the basis of the optimization results, FLML can learn the geometric information from different local methods. The geometric structure of the underlying manifold is more thoroughly explored by the proposed FLML, compared to LE, LLE, and HLLE.

Like the PCA-based fault detection method, two monitoring statistics including Hotelling's T^2 and Q statistics are established. The effectiveness and advantages of the proposed FLML-based fault detection are illustrated by an industrial Tennessee Eastman process benchmark and a real blast furnace ironmaking process.

The rest of this paper is organized as follows. Section II briefly introduces the ideas of LE, LLE, and HLLE. Section III illustrates the proposed FLML method and its application in fault detection in detail. In Section IV, the proposed FLML-based fault detection approach is verified through an industrial Tennessee Eastman (TE) process benchmark and a real blast furnace ironmaking process. Finally, Section V provides the conclusion.

2. Brief Review of LE, LLE and HLLE

2.1. LE

LE is a well-known manifold learning method to extract local structure features in the original sample space. The main idea behind LE is that the corresponding projections of neighboring points in the low-dimensional space should be close if the neighboring points in the high-dimensional space are close [14].

Given a set $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T \in \mathbb{R}^{n \times D}$ of n sample points in \mathbb{R}^D , where D is the number of variables. The goal of LE is to find a set of points $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T \in \mathbb{R}^{n \times d}$ ($d \ll D$) in the low-dimensional space to have the neighbor relations between sample points \mathbf{X} in the high-dimensional space.

In LE, the first step is to construct the adjacency graph. k -nearest neighbors (KNN) is a widely used neighbors selection strategy, due to its simplicity. To model the neighborhood relations between sample points, an adjacency matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ is employed, where each element S_{ij} represents the neighborhood relations between \mathbf{x}_i and its neighbor \mathbf{x}_j . Using the Gaussian heat kernel, the adjacency matrix \mathbf{S}_{ij} can be formed as follows,

$$S_{ij} = \begin{cases} e^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\delta^2} & , \mathbf{x}_i \text{ and } \mathbf{x}_j \in \text{neighbors} \\ 0 & , \text{otherwise} \end{cases} \quad (1)$$

where δ is the kernel width.

The objective function of LE can be cast as,

$$\begin{aligned} \min_{\mathbf{y}_i, \mathbf{y}_j} J_{LE} &= \sum_{i,j=1}^n \|\mathbf{y}_i - \mathbf{y}_j\|^2 S_{ij} = \mathbf{Y}^T \mathcal{L}_e \mathbf{Y} \\ \text{s.t. } &\mathbf{Y}^T \mathbf{D} \mathbf{Y} = \mathbf{I} \end{aligned} \quad (2)$$

where $\mathbf{D}(\mathbf{D}_{ii} = \sum_{j=1}^n S_{ij})$ is a diagonal matrix, and $\mathcal{L}_e(\mathcal{L}_e = \mathbf{D} - \mathbf{S})$ is defined as a Laplacian matrix.

The optimization problem (2) is equivalent to generalized eigenvalue problem,

$$\mathcal{L}_e \mathbf{Y} = \lambda \mathbf{D} \mathbf{Y} \quad (3)$$

It can be readily solved through eigenvalue decomposition.

2.2. LLE

In LLE, the key assumption is that each data point and its neighbors are lied on or closed to a locally linear patch. A sample can be represented as the linear combination of multiple samples from its neighborhood [13]. The optimization problem of LLE can be formulated as,

$$\begin{aligned} \min_{\theta_{ij}} \quad & J_{LLE} = \sum_{i=1}^n \|x_i - \sum_{j=1}^n \theta_{ij} x_j\|^2 \\ \text{s.t.} \quad & \sum_{j=1}^n \theta_{ij} = 1 \end{aligned} \quad (4)$$

where θ_{ij} is the weight coefficient and only when x_i and x_j are neighbors the corresponding θ_{ij} has a value, otherwise it is 0. Similarly, the neighbors selection can be determined by KNN method.

Denoted the low-dimensional features as $\Phi = [\phi_1^T, \phi_2^T, \dots, \phi_n^T]^T$, the minimization problem (4) can be formulated as,

$$\begin{aligned} \min_{w_{ij}} \quad & J_{LLE} = \sum_{i=1}^n \|\phi_i - \sum_{j=1}^n \theta_{ij} \phi_j\|^2 = \text{tr}(\Phi^T \mathcal{L}_l \Phi) \\ \text{s.t.} \quad & \Phi^T \Phi = I \end{aligned} \quad (5)$$

where I is the identity matrix, $\Theta = [\theta_{ij}]_{i=1, \dots, n, j=1, \dots, n}$ is the weight coefficient matrix and $\mathcal{L}_l = (I - \Theta)^T (I - \Theta)$.

Similarly, the optimization problem (5) can be solved through eigenvalue decomposition.

2.3. HLLE

HLLE is regarded as a variant of LLE. Assumed that the low-dimensional data representation is locally isometric to an open and connected subset, the idea behind HLLE is to minimize the curviness of the high-dimensional manifold while embedding it into a low-dimensional space [18].

Assumed that the set $X = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^{n \times D}$ is located on a smooth manifold $\mathcal{M} \subset \mathbb{R}^D$ with an intrinsic dimension $d \ll D$, PCA is firstly performed on X to obtain d eigenvectors $V = [v_1, v_2, \dots, v_d] \in \mathbb{R}^{D \times d}$.

It is assumed that the number of neighbors of the sample x_i is k . Then, the local tangent coordinate $u_j, j = 1, \dots, k$ of the sample x_i can be calculated by projecting the local neighborhood into the tangent subspace,

$$u_j^i = \langle V, x_j - x_i \rangle = V^T (x_j - x_i) \quad (6)$$

Then, the Hessian matrix H^i containing local information will be calculated by the projection of neighbors x_j in the tangent coordinate u_j . Through defining $U^i = [U_1^i, U_2^i, \dots, U_k^i]$ constructing by the u_j^i [26], H^i can be constructed by the last $(U^i)^\dagger$ where $(\cdot)^\dagger$ is the pseudo-inverse symbol. Therefore, the local objective of HLLE can be estimated with,

$$\begin{aligned} \min_{f^i} \quad & J_{HLLE} = \frac{1}{N} \sum_{i=1}^N f^{iT} H^{iT} H^i f^i \\ \text{s.t.} \quad & f^{iT} f^i = 1 \end{aligned} \quad (7)$$

where $f^i = [f(x_1), f(x_2), \dots, f(x_k)]^T$ is the local projection and $f(x_j), j = 1, \dots, k$ is the smooth function used to estimate the local neighborhood information at a fixed point x_i . Then, extending the local

neighborhood information to the all samples, we can get the global projection $f = [f^1, f^2, \dots, f^N]^T$. The optimization objective of HLLE can be rewritten as,

$$\begin{aligned} \min_f \quad J_{HLLE} &= \frac{1}{N} \sum_{i=1}^N f^T S^{iT} H^{iT} H^i S^i f \\ &= \text{tr}(f^T \mathcal{L}_h f) \\ \text{s.t.} \quad f^T f &= 1 \end{aligned} \quad (8)$$

where $\mathcal{L}_h = \frac{1}{N} \sum_{i=1}^N S^{iT} H^{iT} H^i S^i$ is the local objective and the neighboring selection matrix S^i can convert local projection to global projection $f^i = S^i f$ and each entry S_{ij}^i of S^i can be obtained as,

$$S_{ij}^i = \begin{cases} 1, & x_i \text{ and } x_j \in \text{neighbors} \\ 0, & \text{Otherwise} \end{cases} \quad (9)$$

For solving the above optimization problem (8), eigenvalue decomposition is also utilized. Details of HLLE can be found in [18].

3. Proposed Method

3.1. FLML: Fused Local Manifold Learning

From the previous section, it can be found that LE, LLE, and HLLE can explore the local geometric structure information from different perspectives. To synthesize these local information, Xing *et al.* [25] fused the local information obtained by multiple manifold learning methods including LE, LLE, HLLE and LTSA by reformulating the different local manifold algorithms under the local tangent coordinate system to reveal the underlying manifold of the dataset. Similar to [25], the fused local objective \mathcal{F}_f is defined as

$$\mathcal{F}_f = c_1 \mathcal{L}_e + c_2 \mathcal{L}_l + (1 - c_1 - c_2) \mathcal{L}_h \quad (10)$$

to integrate LE, LLE, and HLLE. Here, $c_1 + c_2 \leq 1$, $c_1, c_2 \in [0, 1]$ are the fusion coefficients.

Remark 1. In [25], the local objectives are optimized to determine the fusion coefficients and global embedding coordinates simultaneously. In [27], the selection of c is determined by employing the alternating optimization method which iterative updates c and Y in an alternating fashion. For simplicity, only the global embedding coordinates are obtained in this study. The fusion coefficients are considered fixed values.

Similar to LE, LLE, and HLLE, the proposed FLML method is a kind of linear projection method. To facilitate the online fault detection, an explicit linear mapping from the original space to the low-dimensional space is provided. Thus, the main goal of FLML is to seek a transformation matrix that maps the high-dimensional data to low-dimensional data. Supposed that $w \in \mathbb{R}^{D \times 1}$ is the transformation vector from X to y , therefore, the projections in the low-dimensional space can be represented as $y = w^T X$.

To synthesize the local geometric structure information and impose constraints to prevent multiple solutions, the objective function of FLML is formulated as follows,

$$\begin{aligned} \min_w \quad J_{FLML} &= w^T X^T \mathcal{F}_f X w \\ \text{s.t.} \quad w^T X^T X w &= 1 \end{aligned} \quad (11)$$

Remark 2. As shown in (10) and (11), it is noticed that the corresponding local structure will be extracted from LE, LLE and HLLE while c_1 and c_2 are selected to be $(c_1 = 1, c_2 = 0)$, $(c_1 = 0, c_2 = 1)$ and $(c_1 = 0, c_2 = 0)$, respectively.

To solve the optimization problem (11), we use the technique of Lagrange multipliers as follows,

$$J_{\text{FLML}} = \mathbf{w}^T \mathbf{X}^T \mathcal{F}_f \mathbf{X} \mathbf{w} + \lambda/2(1 - \mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w}) \quad (12)$$

where λ is the Lagrange multiplier.

While $\frac{\partial J_{\text{FLML}}}{\partial \mathbf{w}} = 0$, it results in

$$\mathbf{X}^T \mathcal{F}_f \mathbf{X} \mathbf{w} = \lambda \mathbf{X}^T \mathbf{X} \mathbf{w} \quad (13)$$

Hence, we can use generalized eigenvalue decomposition to obtain the transformation vector \mathbf{w} from (13). Finally, the transformation matrix $\mathbf{W} \in \mathbb{R}^{D \times d}$ can be assembled by the eigenvectors corresponding to the smallest d eigenvalues derived from the result of generalized eigenvalue decomposition.

3.2. FLML based Fault Detection

Generally, data-driven fault detection methods contain two steps, offline modeling and online monitoring. In the offline modeling step, the process data $\mathbf{X} \in \mathbb{R}^{n \times D}$ are collected under normal operating condition for training. Using the training data \mathbf{X} , an FLML model is established. Subsequently, the low-dimensional data $\mathbf{Y} \in \mathbb{R}^{n \times d}$ is obtained. With the transformation matrix \mathbf{W} , the relation between \mathbf{X} and \mathbf{Y} can be represented,

$$\mathbf{Y} = \mathbf{X} \mathbf{W} \quad (14)$$

The residual matrix $\mathbf{E} \in \mathbb{R}^{n \times D}$ is,

$$\mathbf{E} = \mathbf{X} - \mathbf{Y} \mathbf{W}^T = \mathbf{X}(\mathbf{I} - \mathbf{W} \mathbf{W}^T) \quad (15)$$

In the online monitoring step, the low-dimensional data point \mathbf{y}_{new} and residual \mathbf{e}_{new} of a new standardized sample \mathbf{x}_{new} is obtained,

$$\begin{cases} \mathbf{y}_{\text{new}} = \mathbf{W}^T \mathbf{x}_{\text{new}} \\ \mathbf{e}_{\text{new}} = (\mathbf{I} - \mathbf{W} \mathbf{W}^T) \mathbf{x}_{\text{new}} \end{cases} \quad (16)$$

The Hotelling's T^2 and squared prediction error (SPE) statistics (also called Q statistic) are often used for monitoring. For T^2 statistic, the Mahalanobis distance is used to evaluate the variations of \mathbf{y} . The Hotelling's T^2 statistic is defined as,

$$T^2 = \mathbf{y}_{\text{new}}^T \mathbf{\Lambda}^{-1} \mathbf{y}_{\text{new}} \quad (17)$$

where $\mathbf{\Lambda}$ is the covariance matrix of \mathbf{Y} which is extracted from the normal operating condition data \mathbf{X} .

For Q statistic, the Euclidean distance is adopted to evaluate the magnitude of vector in the residual space as follows,

$$Q = \mathbf{e}_{\text{new}}^T \mathbf{e}_{\text{new}} \quad (18)$$

Under the assumption that all operating parameters and prediction errors have a Gaussian distribution, the upper control limits (UCLs) of Hotelling's T^2 and Q statistics are determined by means of F distribution and χ^2 distribution, respectively. Thus, with a level of significance α , the UCLs are calculated as follows,

$$\begin{cases} T_{\text{UCL}}^2 \sim \frac{d(n-1)}{n-d} F(d, n-d, \alpha) \\ Q_{\text{UCL}} \sim g \chi_{h, \alpha}^2 \end{cases} \quad (19)$$

where $g = \sigma^2/2\mu$ and $h = 2\mu^2/\sigma^2$, μ is the sample mean and σ^2 is the sample variance of the statistic Q . g and h are calculated using normal operating condition data.

In FLML, the parameters such as the number of the neighbors k , the bandwidth δ of Gaussian heat kernel, the fusion weights c_1 and c_2 , and the number of latent variables d are determined as follows,

- k : The finding of neighborhood relation is related to the selection of k [23]. To balance the computation complexity and generalization capability, we choose k in the range $[1 : 10]$ with the smallest mean false alarm rate (FAR). The definition of FAR can be found in next section.
- δ : If the bandwidth of the Gaussian heat kernel function is too small, the kernel will be sensitive to noise. A large bandwidth may create an overly smooth mapping [28]. Empirically, the bandwidth δ is chosen as $\delta = bm\sigma^2$ where m is the size of variables, b is a constant, and σ^2 represents the variance of the data, which is 1 as the original data is normalized [22]. In the case studies, $b = 50$ are selected.
- c_1 and c_2 : It is noticeable that the hyper-parameters c_1 and c_2 have a important influence on the performance of the proposed FLML method. However, it is a challenging work to choose a set of optimal hyper-parameters. As a traditional way of performing hyper-parameter optimization, the grid search method is employed. For this purpose, a finite set of $c_1 \in \{0, 0.05, \dots, 0.95\}$ and $c_2 \in \{0, 0.05, \dots, 0.95\}$ are explored by minimizing the mean FAR.
- d : Similar to NPE-based and LPP-based methods, the number of latent variables d is selected by searching for eigenvalues similar to the smallest non-zero eigenvalue.

In summary, the procedure of the FLML based fault detection is described as follows,

Offline training:

- Step 1. Standardize the collected dataset X under normal operating condition.
- Step 2. Extract the local structure information by calculating \mathcal{L}_e , \mathcal{L}_l and \mathcal{L}_h from (2), (5), and (8).
- Step 3. Establish the fused local structure \mathcal{F} defined in (10).
- Step 4. Perform eigenvalue decomposition as (13) to solve the optimization problem (11) and retain the eigenvectors corresponding to the smallest d eigenvalues as the transformation matrix W .
- Step 5. Compute the UCLs T_{UCL}^2 and Q_{UCL} using (19).

Online monitoring:

- Step 1. Standardize a new collected sample x_{new} .
- Step 2. Obtain the score y_{new} and residual e_{new} from (16).
- Step 3. Compute the online statistics T^2 and Q according to (17) and (18), respectively.
- Step 4. If any of T^2 and Q statistics violate the confident limit, a fault is declared.

4. Case Studies

In this section, the proposed FLML-based fault detection method is verified by conducting experiments on the Tennessee Eastman process and a real-world blast furnace ironmaking process.

4.1. Tennessee Eastman Process

The TE process is a well-known and widely used industrial benchmark for comparing the performance of process monitoring and control [29]. In the TE process, there are five main units including a separator, a compressor, a reactor, a vapour/liquid separator, a stripper and a condenser, and eight components A-H. It also has 12 manipulated variables (XMV1-11) and 41 measured variables (XMEAS1-41). Among these variables, XMEAS(23-41) are the composition of A-H measured with 6 min sampling intervals in different positions. Other variables are collected with 3 min sampling intervals. There are 21 preprogrammed faults in TE process. Of 21 faults, due to the absence of observable change in the mean and standard deviation between their corresponding faulty and normal operation, faults 3, 9 and 15 are very difficult to detect [30]. In this study, faults 3, 9 and 15 are ignored. The descriptions of the faults considered in this study are listed in Table 1.

Table 1. Faults description of the TE process.

No.	Fault description	Fault type
IDV(0)	Normal situations	-
IDV(1)	A/C feed ratio, B composition constant (Stream 4)	Step
IDV(2)	B composition, A/C ratio constant (Stream 4)	Step
IDV(4)	Reactor cooling water inlet temperature	Step
IDV(5)	Condenser cooling water inlet temperature	Step
IDV(6)	A feed loss (Stream 1)	Step
IDV(7)	C header pressure loss-reduced availability (Stream 4)	Step
IDV(8)	A, B, C feed composition (Stream 4)	Random variation
IDV(10)	C feed temperature (Stream 4)	Random variation
IDV(11)	Reactor cooling water inlet temperature	Random variation
IDV(12)	Condenser cooling water inlet temperature	Random variation
IDV(13)	Reaction kinetics	Slow drift
IDV(14)	Reactor cooling water valve	Sticking
IDV(16)	Unknown	Unknown
IDV(17)	Unknown	Unknown
IDV(18)	Unknown	Unknown
IDV(19)	Unknown	Unknown
IDV(20)	Unknown	Unknown
IDV(21)	Valve fixed at steady-state position	Constant position

A widely used dataset of TE process can be found in <http://web.mit.edu/braatzgroup/links.html>. In this study, we also adopt this dataset. Specifically, 22 measured variables XMEAS(1:22) and 11 manipulated variables XMV(1:11) are selected as $\mathbf{x} \in \mathbb{R}^{33}$. 500 samples collected under the normal operating condition (IDV(0)) are used as a training dataset. For each faulty dataset, it has 960 samples in total where the fault is injected from the 160th sample.

To assess the fault detection performance, three indices including Missed detection rate (MDR), Detection delay (DD) and false detection rate (FAR) are used. DD is defined as the time interval from the start of fault to the detection time, which is expressed as the first time of 5 consecutive rises. MDR and FAR with a 99.9% confidence level can be calculated as follows,

$$\text{MDR}(\%) = \frac{\text{No. of samples}(J \leq J_{UCL} | \text{Fault})}{\text{total samples}(\text{Fault})} \times 100\%$$

$$\text{FAR}(\%) = \frac{\text{No. of samples}(J > J_{UCL} | \text{Normal})}{\text{total samples}(\text{Normal})} \times 100\%$$

where $J \in \{T^2, Q\}$ and $J_{UCL} \in \{T_{UCL}^2, Q_{UCL}\}$.

For comparative study, several typical fault detection methods including PCA, NPE, LPP, Principal Component Pursuit (PCP)[31], kernel PCA (KPCA)[32], mixed KPCA (MKPCA) and structured joint sparse PCA (SJSPCA)[33] are employed. For PCA, $d = 19$ is set according to the 95% cumulative percentage of variance (CPV). For the NPE, $k = 5$ and $d = 19$ is selected. For LPP, $k = 5$, $\delta = 1650$, and $d = 19$ are selected. For PCP, $d = 19$ is selected. For KPCA, the kernel width $C = 1000$ and $d = 22$ is selected. For MKPCA, the kernel widths C_1, C_2 are set as 500 and 1000 respectively, and $d = 28$ is selected. For SJSPCA, $k = 5$, $\lambda_1 = 600$, $\lambda_2 = 50$, and $d = 19$ are selected. As displayed in Figure 1, $k = 5$ is selected for the TE case. Similarly, As shown in Figure 2, $c_1 = 0.25$ and $c_2 = 0.25$ are chosen. For FLML, $\delta = 1650$, $c_1 = 0.25$, $c_2 = 0.25$, and $d = 19$ are selected.

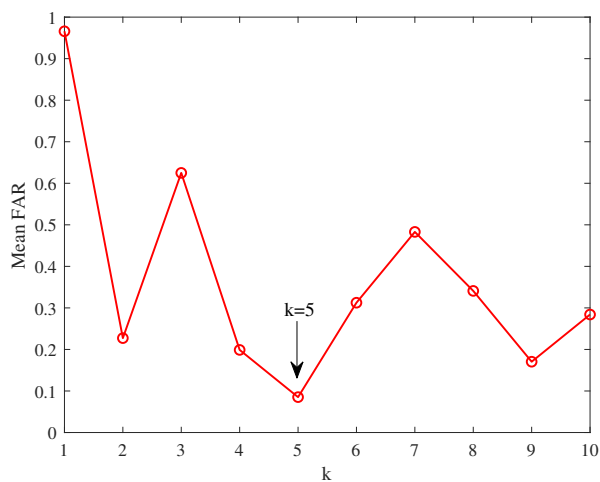


Figure 1. Mean FAR with different k for the TE process.

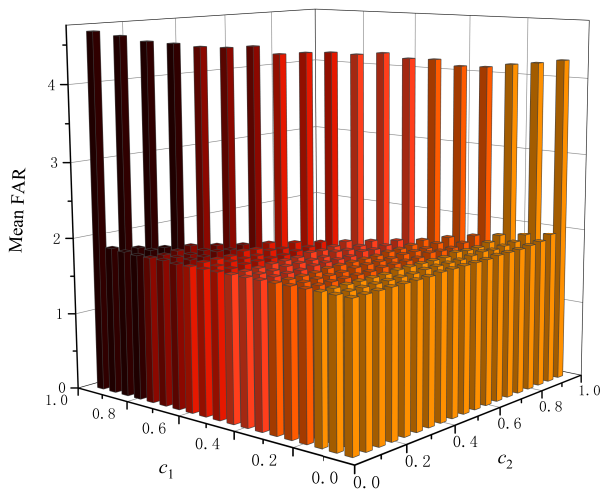


Figure 2. Grid search result of the fusion weights c_1 and c_2 for the TE process.

Tables 2 and 3 list the MDRs and FARs of 18 faults. It can be observed that the NPE and LPP can provide superior performance over PCA. The utilization of local structure information can enhance fault detection performance. On the other hand, FLML T^2 statistic offers the lowest MDRs among all the comparative methods. The average of MDRs of FLML T^2 statistic is 11.49%. Compared to NPE and LPP methods, the average of MDRs of FLML T^2 statistic is increasingly reduced. Furthermore, the FARs of FLML T^2 statistic reach the same level of NPE and LPP. The average of FARs of FLML T^2 statistic is 0.45%. The DD index represents the sensitivity of monitoring statistics. A smaller DD means the monitoring statistic can detect the fault earlier. Table 4 lists the DDs for all methods, where the DD is indicated in the unit of the hour. As displayed in Table 4, FLML T^2 statistic can derive the smallest DD among the all methods.

Table 2. Comparison results of MDR values for TE process.

No.	PCA (%)		NPE (%)		LPP (%)		PCP (%)		KPCA (%)		MKPCA (%)		SJSPCA (%)		FLML (%)	
	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q
1	0.75	0.25	0.13	0.38	0.13	0.50	0.13	1.63	16.50	0.75	0.25	0.88	0.38	0.25	0.38	0.50
2	2.00	0.75	1.38	1.75	1.50	1.75	1.00	14.88	93.38	2.75	1.63	2.00	1.63	1.50	1.50	2.50
4	58.00	0.13	25.25	33.38	30.88	29.25	14.50	5.75	0.00	95.38	0.75	94.50	27.50	2.88	6.38	97.13
5	76.75	75.75	0.00	0.00	0.00	70.25	69.63	93.88	77.38	77.63	72.88	77.50	71.63	71.75	0.00	0.38
6	0.63	0.00	0.00	0.00	0.00	0.00	0.50	0.00	96.63	0.00	0.00	0.00	0.50	0.00	0.00	0.00
7	0.00	6.00	0.00	0.00	0.00	0.00	0.00	0.00	16.13	14.25	0.00	6.63	0.00	0.00	0.00	24.00
8	2.63	9.38	2.13	2.00	2.25	2.00	1.88	49.00	18.38	5.38	1.50	3.63	2.25	2.13	2.25	13.00
10	67.13	45.63	42.13	41.63	43.13	47.88	40.38	91.63	57.13	72.00	53.75	43.88	54.88	45.88	20.50	43.13
11	48.88	39.88	38.25	38.63	38.38	38.63	31.00	60.50	18.88	77.50	32.50	72.13	37.63	30.50	38.13	91.75
12	1.75	7.38	0.38	0.25	0.38	1.00	1.00	38.25	30.25	2.50	0.75	2.38	1.25	1.00	0.25	5.25
13	5.88	4.63	4.88	5.50	5.13	5.50	4.75	17.50	54.75	7.13	5.13	6.13	5.38	5.13	4.75	12.13
14	0.13	11.88	0.00	0.13	0.00	0.00	0.00	0.38	0.00	0.13	0.00	0.38	0.00	0.00	0.13	30.00
16	84.38	48.75	63.38	59.50	51.13	60.13	59.75	83.25	76.13	87.75	74.13	58.38	69.50	57.38	15.88	90.63
17	21.00	3.63	16.88	11.63	12.88	11.63	9.88	11.00	35.88	8.38	5.25	24.13	13.13	9.38	6.25	37.50
18	10.88	10.00	10.75	10.50	10.88	10.38	9.50	10.38	94.00	10.13	10.13	11.38	10.38	10.38	9.88	10.88
19	94.38	71.63	92.25	95.38	88.50	96.00	75.00	98.63	84.75	96.00	85.00	99.63	82.25	91.38	24.13	97.63
20	62.63	40.63	50.63	44.00	48.50	46.88	31.88	72.13	45.38	46.75	45.00	65.75	48.13	43.63	20.50	44.50
21	60.63	44.13	54.00	55.63	56.25	57.00	40.00	89.00	46.63	66.75	59.75	69.25	55.63	55.88	56.00	84.38
Aver.	33.24	23.35	22.35	22.24	21.66	26.60	21.71	40.99	47.90	37.28	24.91	35.47	26.78	23.83	11.49	38.07

Table 3. Comparison results of FAR values for TE process.

No.	PCA (%)		NPE (%)		LPP (%)		PCP (%)		KPCA (%)		MKPCA (%)		SJSPCA (%)		FLML (%)	
	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q
1	0.00	1.25	0.00	0.00	0.00	0.00	0.50	0.00	0.50	0.00	0.00	0.00	0.50	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.50	0.00	0.00	0.00	0.50	0.50	0.63	0.63
4	0.63	0.00	0.63	0.63	0.63	0.63	2.00	0.00	1.00	0.00	0.50	0.00	1.00	1.00	0.00	0.00
5	0.63	0.00	0.63	0.63	0.63	0.63	2.00	0.00	1.00	0.00	0.50	0.00	1.00	1.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	1.25	0.00	0.00	0.00	0.00	0.50	0.00	0.50	0.00	0.00	0.00	1.50	0.50	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.50	1.00	0.50	0.00	0.00	0.00	0.50	0.00	0.00	0.00
10	0.00	0.63	0.00	0.63	0.00	0.00	1.00	0.00	1.50	0.00	0.00	0.00	1.50	1.00	0.63	0.00
11	0.00	1.88	0.00	1.25	0.00	1.25	2.50	0.00	1.50	0.00	1.50	0.00	1.50	1.00	0.00	0.00
12	0.63	0.00	0.63	4.38	0.63	5.63	3.00	0.00	1.50	0.00	1.50	1.00	2.00	1.00	0.63	1.25
13	0.00	0.00	0.00	0.63	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	1.25
14	0.00	0.00	0.63	0.00	0.63	0.00	1.00	0.00	1.50	0.00	0.00	0.00	1.50	1.00	0.00	0.63
16	0.00	2.50	3.75	19.38	2.50	19.38	15.50	0.00	6.50	0.50	4.50	3.00	10.50	10.00	3.13	2.50
17	0.00	1.25	0.00	0.00	0.00	0.00	2.50	0.00	0.50	0.00	0.00	0.00	1.00	0.50	0.00	0.63
18	0.63	0.00	0.63	0.63	0.63	0.00	2.50	0.50	2.00	0.00	0.00	0.00	1.00	0.00	0.63	0.00
19	0.00	0.63	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0.00	0.63	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00
21	0.00	1.25	0.63	1.88	0.00	3.13	6.00	0.00	3.50	0.00	0.00	0.00	2.50	1.00	0.63	1.25
Aver.	0.15	0.65	0.42	1.96	0.36	2.08	2.33	0.03	1.25	0.03	0.47	0.22	1.53	1.03	0.45	0.48

Table 4. Comparison results of DD values for TE process.

No.	PCA (h)		NPE (h)		LPP (h)		PCP (h)		KPCA (h)		MKPCA (h)		SJSPCA (h)		FLML (h)	
	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q	T ²	Q
1	0.35	0.10	0.10	0.20	0.10	0.25	0.10	0.50	0.05	0.35	0.15	0.40	0.20	0.15	0.20	0.10
2	0.85	0.40	0.60	0.75	0.70	0.75	0.55	5.80	0.55	1.25	0.75	0.85	0.75	0.75	0.75	0.40
4	3.20	0.10	0.75	0.15	0.40	0.15	0.15	0.25	0.05	39.80	0.15	39.80	1.95	0.15	0.20	0.10
5	0.35	0.15	0.05	0.05	0.05	0.05	0.05	0.90	0.05	0.10	0.05	0.65	0.05	0.05	0.05	0.15
6	0.30	0.05	0.05	0.05	0.05	0.05	0.25	0.05	0.45	0.05	0.05	0.05	0.25	0.05	0.05	0.05
7	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
8	1.05	0.95	1.00	1.05	1.00	1.05	1.00	3.25	1.00	1.30	0.60	1.35	1.05	1.05	1.00	0.95
10	4.85	1.65	1.75	1.25	1.80	1.25	1.15	24.75	2.40	3.35	2.90	1.70	3.05	1.80	1.25	1.65
11	0.55	0.35	0.30	0.55	0.50	0.55	0.30	2.85	0.30	4.80	0.30	4.80	0.55	0.30	0.50	0.35
12	1.10	1.15	0.15	0.15	0.35	0.35	0.15	4.00	0.35	0.20	0.35	1.10	0.35	0.35	0.15	1.15
13	2.40	1.85	2.05	2.25	2.25	2.25	2.05	3.30	2.05	2.45	2.15	2.45	2.35	2.15	2.10	1.85
14	0.05	0.10	0.05	0.10	0.05	0.05	0.05	0.15	0.05	0.10	0.05	0.10	0.05	0.05	0.10	0.10
16	15.40	0.75	9.65	1.70	1.80	1.80	1.80	9.90	9.80	15.55	11.20	1.70	1.80	1.80	0.40	0.75
17	1.45	1.10	1.40	1.25	1.25	1.20	1.15	1.20	1.10	1.10	1.45	1.25	1.20	1.10	1.10	1.10
18	4.60	4.20	4.40	4.30	4.45	4.35	4.25	4.25	4.25	4.50	4.20	4.65	4.40	4.25	4.15	4.20
19	39.80	9.45	39.80	39.80	39.80	39.80	26.90	39.80	39.80	39.80	7.20	39.80	26.80	39.80	0.55	9.45
20	4.35	4.05	4.25	3.95	4.20	3.95	4.05	4.75	4.35	4.25	4.05	4.25	4.20	4.05	3.50	4.05
21	25.35	12.80	23.75	21.15	23.85	22.70	12.80	37.10	12.80	27.80	25.05	28.10	23.20	23.20	23.60	12.80
Aver.	5.89	2.18	5.01	4.38	4.59	4.48	2.33	0.03	4.41	8.14	3.35	7.40	4.02	4.51	2.21	2.18

To further illustrate the superiority of the proposed FLML method, the monitoring results are depicted in Figures 3 and 4 by PCA, NPE, LPP, PCP, KPCA, MKPCA, SJSPCA, and FLML methods for IDV (10) and IDV (19), respectively. Fault 10 is designed by adding a random disturbance on the C

feed temperature (Stream 4). Compared to the step type fault scenarios such as fault 1, fault 2, fault 10 is more difficult to detect. As shown in Figure 3, it is observed that FLML T^2 can detect most of the faulty samples. In contrast, other monitoring statistics such as PCA T^2 , NPE T^2 , PCP T^2 , KPCA T^2 , MKPCA T^2 , SJSPCA T^2 , and LPP T^2 can not effectively detect the fault 10, since most of the monitoring statistics are below the corresponding UCLs. Fault 19 is an unknown faulty type. The monitoring results of fault 19 is displayed in Figure 4. By comparison, it can be observed that PCA, NPE, LPP, PCP, KPCA, MKPCA, and SJSPCA fail to detect the occurrence of fault 19. The online statistics are almost lower than the UCLs after 161th sample. A promising result can be obtained using FLML T^2 as shown in Figure 4(h).

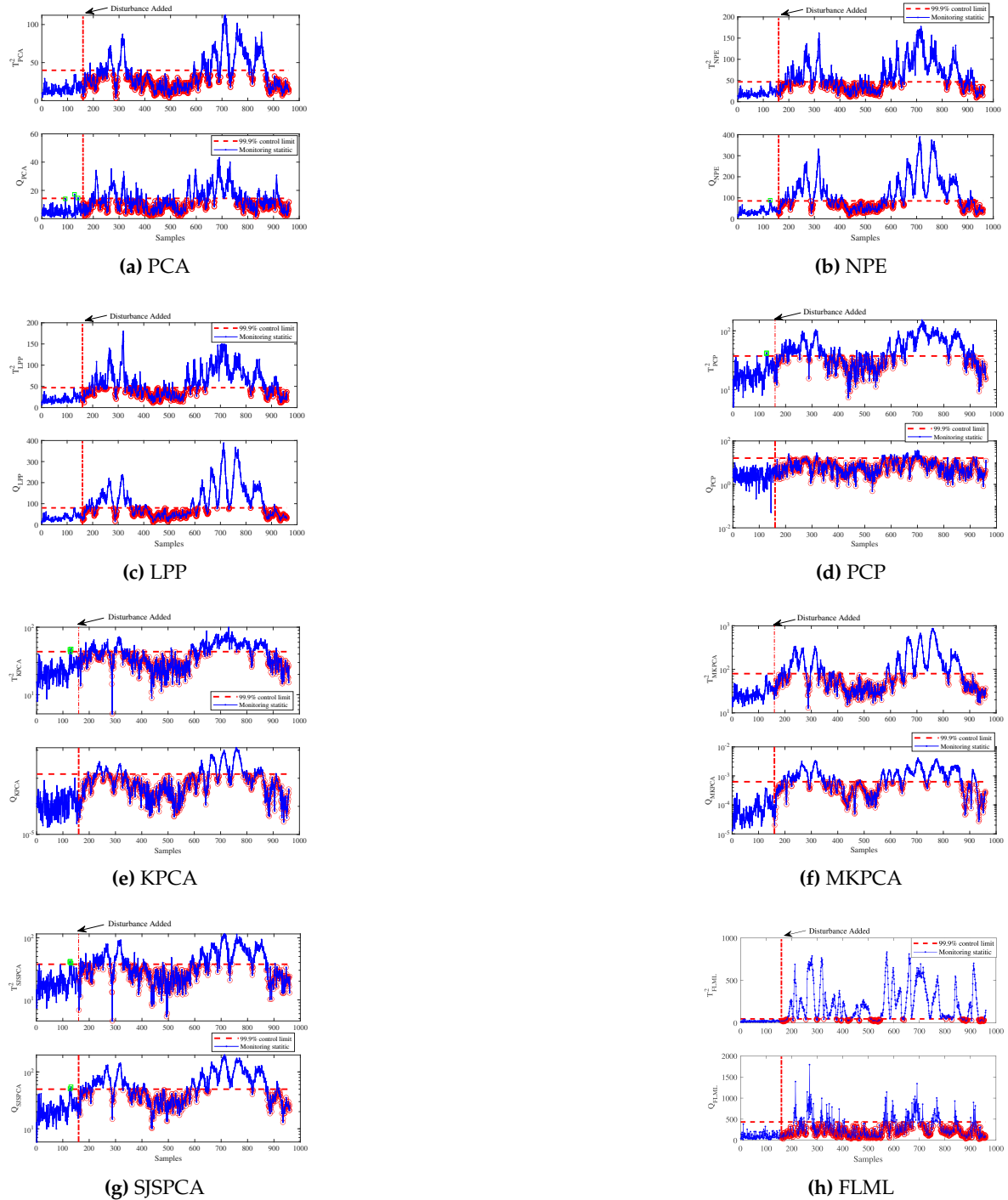
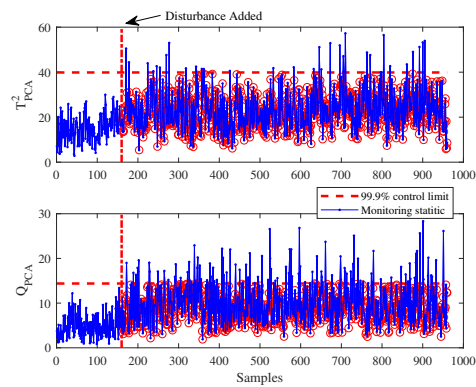
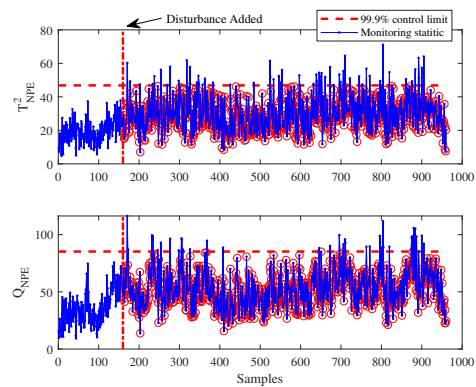


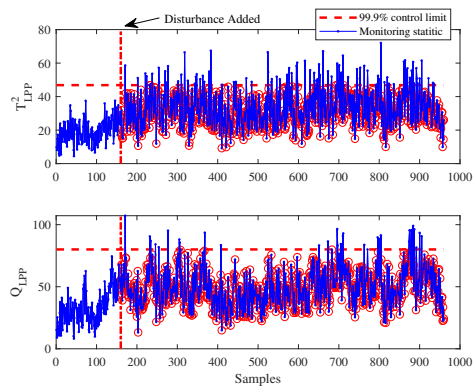
Figure 3. Monitoring results for the fault 10: TE process.



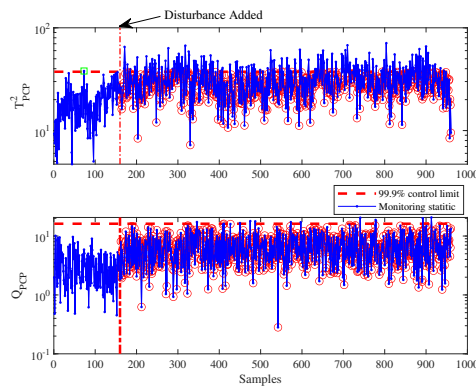
(a) PCA



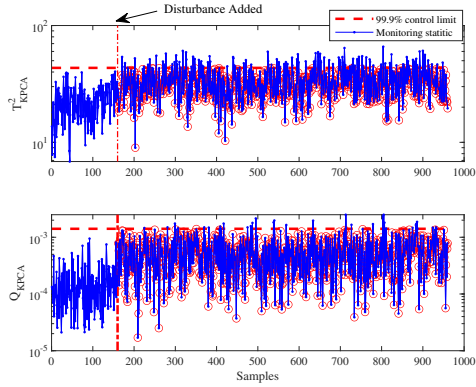
(b) NPE



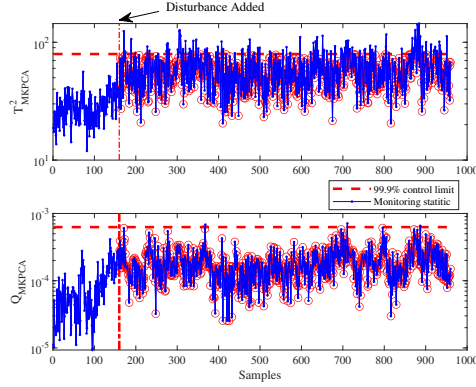
(c) LPP



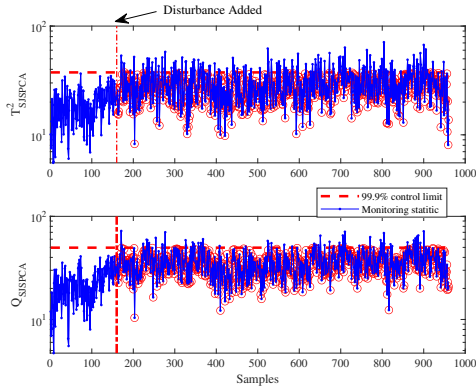
(d) PCP



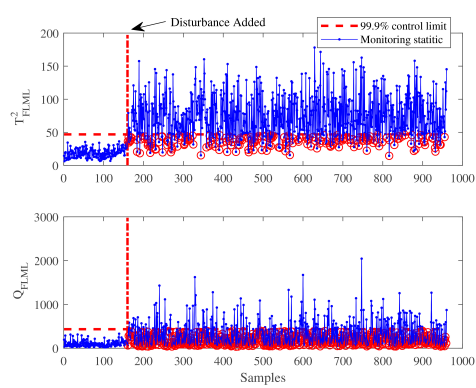
(e) KPCA



(f) MKPCA



(g) SJSPCA



(h) FLML

Figure 4. Monitoring results for the fault 19: TE process.

4.2. Blast Furnace Ironmaking Process

In this section, the proposed method is verified through a real blast furnace (BF) ironmaking process at a steel company in South China. In the steel manufacturing process, the blast furnace ironmaking process plays a vital role. The blast furnace ironmaking process is considered one of the most complex industrial processes. The basic units of the blast furnace ironmaking process are depicted in Figure 5. As it can be seen from Figure 5, the ironmaking process can be mainly divided into several sub-systems including the charging system, gas processing system, hot air system, the pulverized coal injection system, the iron system, and the BF body. The inner structure of the blast furnace is vertical. In the BF, the iron ore and coke are fed from the top along the vertical direction. The 1000°C hot air and coal powder are blown into the furnace from the bottom. Through complex chemical reactions, the molten iron and slag are generated and accumulated in the hearth. In a periodical way, the molten iron and slag are discharged from the bottom of the furnace through the tap hole. As a byproduct, the flux gas escapes from the top of the furnace.

In the ironmaking process, it often suffers from abnormal furnace conditions, due to the effects of unreasonable daily operation and various disturbances. If these abnormal furnace conditions are not timely detected, the product quality will be degraded, and even the safety of the plant would not be ensured. Thus, the effective detection of abnormalities becomes an indispensable component of the operation of a blast furnace. In this study, we consider the detection of the channeling fault. The channeling fault may be caused by several reasons such as the low-quality coke pulverized coal, the inappropriate adjustment of air volume. In a channeling accident, the high-temperature furnace gas passes in the path of least resistance at a high velocity. The furnace gas increases the heat load at the wall and top of the furnace, resulting in possible equipment damage such as burning of the bag dust catcher [34].

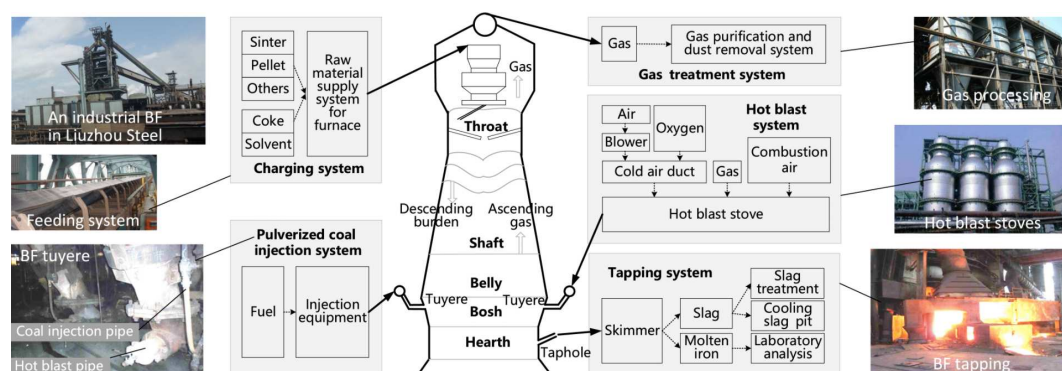


Figure 5. Diagram of Blast Furnace Ironmaking Process [35].

In this case, a data set was collected from 2013 12/21 to 2014 01/05. The sampling interval is 10s. As demonstrated in [36], we select 7 variables that are the most sensitive to faults. These variables are listed in Table 5. A total of 1000 samples are collected from 12/21 00:20 to 2012/28 09:40 under the normal operating conditions as the training dataset. According to the accident report of the operation personnel, from 2013 12/30 23: 44 to 2013 12/31 05: 18, the ironworks accident report has recorded the occurrence of the channeling. To facilitate the verification of fault detection performance, a testing dataset are generated, where 1200 samples were collected where the fault occurred from the 200th sample.

Table 5. Variable description of the Blast furnace ironmaking process.

No.	Variable description	Unit
1	Oxygen enrichment rate	%
2	Enriching oxygen flow	(m^3/s)
3	Hot blast temperature	$^{\circ}C$
4	Top temperature(1)	$^{\circ}C$
5	Top temperature(2)	$^{\circ}C$
6	Top temperature(3)	$^{\circ}C$
7	Downcomer temperature	$^{\circ}C$

For PCA, $d = 3$ is selected. For NPE, the neighbor parameter k and number of principle components d are selected to be 5 and 3, respectively. For LPP, $k = 5$, $d = 3$ and $\delta = 7 * 50$ are determined. For PCP, $d = 4$ is selected. For KPCA, $C = 200$ and $d = 3$. For MKPCA, $C_1 = 150$, $C_2 = 200$ and $d = 3$. For SJSPCA, $k = 5$, $\lambda_1 = 600$, $\lambda_2 = 50$, and $d = 3$. For fair comparison, $k = 5$, $d = 3$, $\delta = 7 * 50$ are selected for FLML. The result of grid search for selecting c_1 and c_2 is plotted in Figure 6. On the basis of grid search results, $c_1 = 0.3$ and $c_2 = 0.3$ are determined.

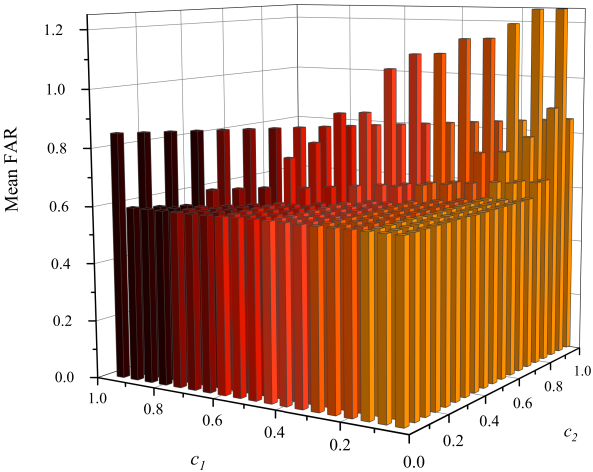
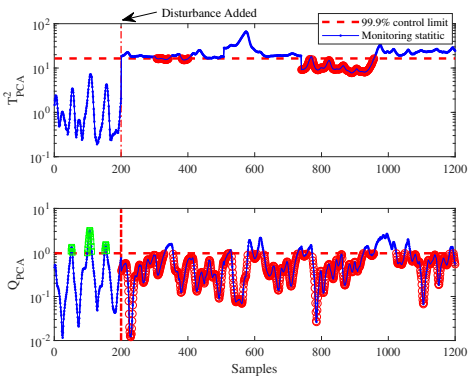
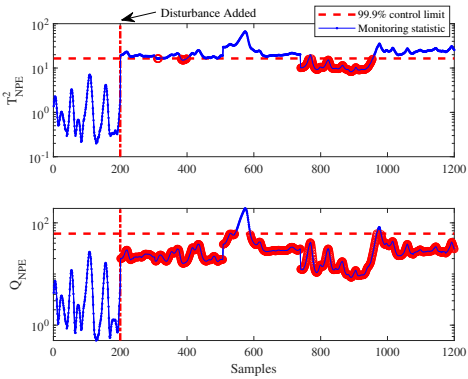


Figure 6. Grid search result of the fusion weights c_1 and c_2 for the blast furnace ironmaking process.

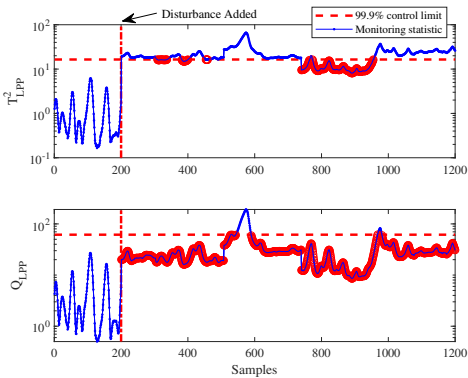
The monitoring results are plotted in Figure 7. It can be observed that NPE and LPP T^2 statistics can detect more faulty samples than PCA T^2 statistic. But, the improvement is limited as shown in Figure 7. The reason may be that NPE and LPP only extract partial local structure information. It is also noticed that there are more false alarms for PCA Q and PCP Q statistics. Compared to the other methods, the proposed FLML method can provide much better performance as shown in Figure 7(h). Only a few faulty samples are missed. Table 6 lists the MDRs, FARs, and DDs. As shown in Table 6, the MDR and FAR of FLML T^2 are 2.80% and 0.00%, respectively. The DD of FLML T^2 is 0.17 mins. The channeling condition can be timely and accurately detected by FLML T^2 . Among these comparative methods, the proposed FLML method can offer the best monitoring performance.



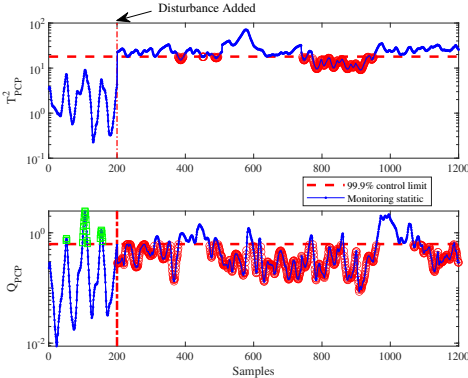
(a) PCA



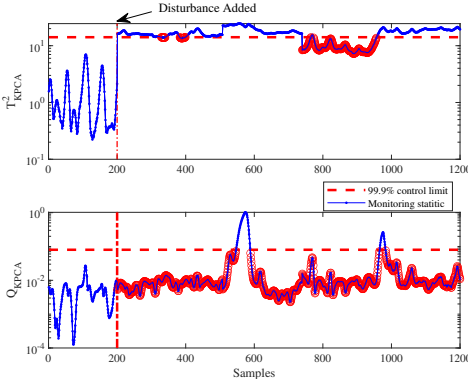
(b) NPE



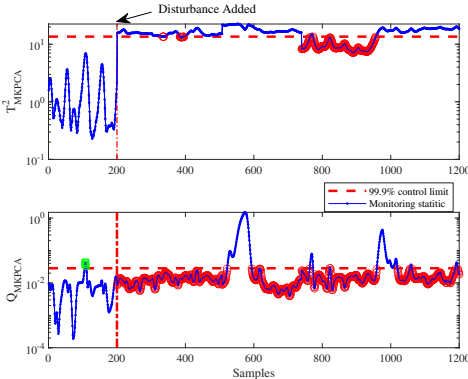
(c) LPP



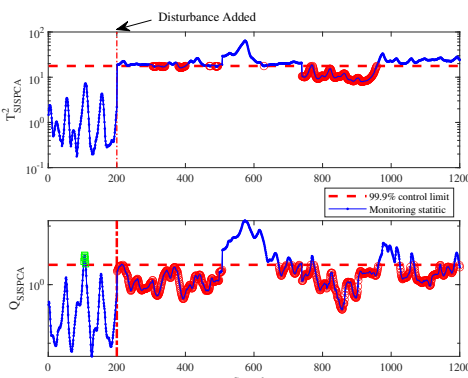
(d) PCP



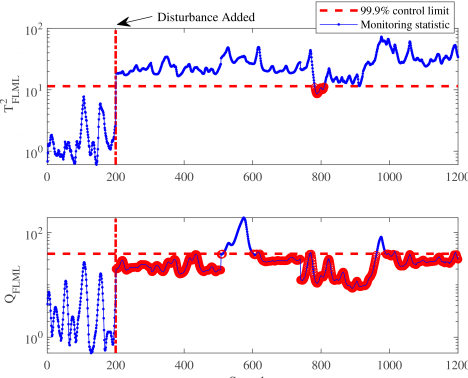
(e) KPCA



(f) MKPCA



(g) SJSPCA



(h) FLML

Figure 7. Monitoring results for the channeling condition: blast furnace ironmaking process.

Table 6. Comparison results of MDR, FAR and DD values for blast furnace ironmaking process.

PCA		NPE		LPP		PCP		KPCA		MKPCA		SJSPCA		FLML	
T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q	T^2	Q
26.70a	78.80	23.10	93.80	26.20	93.80	21.80	71.20	23.90	94.10	22.60	84.70	28.10	71.10	2.80	85.70
0.00b	13.00	0.00	0.00	0.00	0.00	0.00	13.00	0.00	0.00	0.00	2.50	0.00	3.00	0.00	0.00
0.17c	21.67	0.17	6.40	0.17	6.40	0.17	18.33	0.17	57.83	0.17	53.83	0.17	2.17	0.17	5.89

a First row: Missed Detection Rates (MDR, %). b Second row: False Alarm Rates (FAR, %). c Third row: Detection delays(DD, mins).

5. Conclusions

In this paper, a novel data-driven fault detection technique based on fusing local manifold learning methods is proposed for complex industrial processes. The proposed method aims to explore a more comprehensive local structure by synthesizing partial information learned from LE, LLE, and HLL. With the exploit of local geometric structure, the process data is projected into a lower-dimensional space. Hotelling’s T^2 and Q statistics are employed for fault detection. Case studies on the widely used TE process benchmark and a real blast furnace ironmaking process show the superior monitoring performance of the proposed methods, by comparison with other related methods. Future work will investigate the application of the proposed method on fault identification and classification.

Author Contributions: Conceptualization, P.W.; Methodology, P.W., K.W.; Investigation P.W.; Writing– original draft, P.W., K.W. and S.L.; Writing – review & editing, P.W., H.P, J.G.; Data Curation, K.W. and J.G.

Acknowledgments: This work was supported in part by the Zhejiang Province Public Welfare Technology Application Research Project (Grant No.LGF19F030004, LGG21F030015), and Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT2023B19).

References

1. Bahr, N.J. *System safety engineering and risk assessment: a practical approach*; CRC press, 2014.
2. Khan, F.; Rathnayaka, S.; Ahmed, S. Methods and models in process safety and risk management: Past, present and future. *Process safety and environmental protection* **2015**, *98*, 116–147.
3. Nan, C.; Khan, F.; Iqbal, M.T. Real-time fault diagnosis using knowledge-based expert system. *Process safety and environmental protection* **2008**, *86*, 55–71.
4. Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S.N. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers & chemical engineering* **2003**, *27*, 293–311.
5. Tidriri, K.; Chatti, N.; Verron, S.; Tiplica, T. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annual Reviews in Control* **2016**, *42*, 63–81.
6. Dai, X.; Gao, Z. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Transactions on Industrial Informatics* **2013**, *9*, 2226–2238.
7. Yin, S.; Ding, S.X.; Xie, X.; Luo, H. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics* **2014**, *61*, 6418–6428.
8. Qin, S.J.; Chiang, L.H. Advances and opportunities in machine learning for process data analytics. *Computers & Chemical Engineering* **2019**, *126*, 465–473.
9. Yin, S.; Ding, S.X.; Haghani, A.; Hao, H.; Zhang, P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of process control* **2012**, *22*, 1567–1581.
10. Dong, Y.; Qin, S.J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *Journal of Process Control* **2018**, *67*, 1–11.
11. Yin, S.; Li, X.; Gao, H.; Kaynak, O. Data-based techniques focused on modern industry: An overview. *IEEE Transactions on Industrial Electronics* **2014**, *62*, 657–667.
12. Tenenbaum, J.B.; De Silva, V.; Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *science* **2000**, *290*, 2319–2323.
13. Roweis, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *science* **2000**, *290*, 2323–2326.

14. Belkin, M.; Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation* **2003**, *15*, 1373–1396.
15. Zhang, Z.; Zha, H. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM journal on scientific computing* **2004**, *26*, 313–338.
16. He, X.; Niyogi, P. Locality preserving projections. *Advances in neural information processing systems* **2004**, *16*, 153–160.
17. He, X.; Cai, D.; Yan, S.; Zhang, H.J. Neighborhood preserving embedding. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. IEEE, 2005, Vol. 2, pp. 1208–1213.
18. Donoho, D.L.; Grimes, C. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences* **2003**, *100*, 5591–5596.
19. Chen, H.; Wu, J.; Jiang, B.; Chen, W. A modified neighborhood preserving embedding-based incipient fault detection with applications to small-scale cyber–physical systems. *ISA transactions* **2020**, *104*, 175–183.
20. Song, B.; Ma, Y.; Shi, H. Multimode process monitoring using improved dynamic neighborhood preserving embedding. *Chemometrics and Intelligent Laboratory Systems* **2014**, *135*, 17–30.
21. Duan, Y.; Liu, M.; Dong, M. A Metric-Learning-Based Nonlinear Modeling Algorithm and Its Application in Key-Performance-Indicator Prediction. *IEEE Transactions on Industrial Electronics* **2019**, *67*, 7073–7082.
22. Zhang, M.; Ge, Z.; Song, Z.; Fu, R. Global–local structure analysis model and its application for fault detection and identification. *Industrial & Engineering Chemistry Research* **2011**, *50*, 6837–6848.
23. Wu, P.; Lou, S.; Zhang, X.; He, J.; Gao, J. Novel Quality-Relevant Process Monitoring based on Dynamic Locally Linear Embedding Concurrent Canonical Correlation Analysis. *Industrial & Engineering Chemistry Research* **2020**, *59*, 21439–21457.
24. Li, B.; Zhang, Y. Supervised locally linear embedding projection (SLLEP) for machinery fault diagnosis. *Mechanical Systems and Signal Processing* **2011**, *25*, 3125–3134.
25. Xing, X.; Wang, K.; Lv, Z.; Zhou, Y.; Du, S. Fusion of local manifold learning methods. *IEEE Signal Processing Letters* **2014**, *22*, 395–399.
26. Xing, X.; Du, S.; Wang, K. Robust hessian locally linear embedding techniques for high-dimensional data. *Algorithms* **2016**, *9*, 36.
27. Bezdek, J.C.; Hathaway, R.J. Some notes on alternating optimization. In Proceedings of the AFSS international conference on fuzzy systems. Springer, 2002, pp. 288–300.
28. Bernal-de Lázaro, J.; Llanes-Santiago, O.; Prieto-Moreno, A.; Knupp, D.; Silva-Neto, A. Enhanced dynamic approach to improve the detection of small-magnitude faults. *Chemical Engineering Science* **2016**, *146*, 166–179.
29. Downs, J.J.; Vogel, E.F. A plant-wide industrial process control problem. *Computers & chemical engineering* **1993**, *17*, 245–255.
30. Chiang, L.; Russell, E.; Braatz, R. Fault detection and diagnosis in industrial systems, 2001.
31. Candès, E.J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *Journal of the ACM (JACM)* **2011**, *58*, 1–37.
32. Lee, J.M.; Yoo, C.; Choi, S.W.; Vanrolleghem, P.A.; Lee, I.B. Nonlinear process monitoring using kernel principal component analysis. *Chemical engineering science* **2004**, *59*, 223–234.
33. Liu, Y.; Zeng, J.; Xie, L.; Luo, S.; Su, H. Structured joint sparse principal component analysis for fault detection and isolation. *IEEE Transactions on Industrial Informatics* **2018**, *15*, 2721–2731.
34. Zhou, B.; Ye, H.; Zhang, H.; Li, M. Process monitoring of iron-making process in a blast furnace with PCA-based methods. *Control engineering practice* **2016**, *47*, 1–14.
35. Zhou, P.; Zhang, R.; Xie, J.; Liu, J.; Wang, H.; Chai, T. Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: An integrated PCA-ICA method. *IEEE Transactions on Industrial Electronics* **2020**, *68*, 622–631.
36. Wang, L.; Yang, C.; Sun, Y.; Zhang, H.; Li, M. Effective variable selection and moving window HMM-based approach for iron-making process monitoring. *Journal of Process Control* **2018**, *68*, 86–95.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.