Expert Systems With Applications 319 (2026) 132035

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Progressive anchoring-driven consistent dual reconstruction for cross-domain open-mode process monitoring

Ziqing Deng [a], Xiaofang Chen [a,*], Lipo Wang [b,*], Lihui Cen [a], Yalin Wang [a], Shouli Yu [c]

[a] School of Automation, Central South University, Changsha, 410083, China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 639798, Singapore
[c] Yuncheng Risun Group Energy Co., Ltd, Heze, 274700, China

ARTICLE INFO

Keywords:
Aluminum electrolysis
Coal gasification
Dictionary learning
Multi-mode process monitoring
Pseudo-labeling
Transfer learning

ABSTRACT

Cross-domain multi-mode industrial process monitoring provides effective support for stable operation under diverse operating conditions and across multiple units of similar equipment. However, most existing methods assume that the target-domain data are labeled and fail to identify newly emerging unknown modes, which limits their applicability in real-world scenarios. To address these challenges, this paper proposes a progressive anchoring-driven consistent dual reconstruction (PACDR) method for cross-domain open-mode process monitoring. Specifically, to achieve domain adaptation in the presence of unknown modes, an open-mode dual reconstruction framework based on dictionary pair learning is developed to convert and transfer known mode knowledge across domains while suppressing the interference of unknown information. Besides, an intra- and inter-domain joint graph regularization is proposed to enhance the geometric consistency between the source and target domains and improve the discrimination of the representation coefficients. Meanwhile, a progressive target anchoring strategy is designed to adaptively anchor highly reliable pseudo-labeled samples, boosting the performance of the proposed method. Extensive experiments conducted on the Tennessee Eastman benchmark, a coal gasification industrial process, and an aluminum electrolysis process demonstrate the effectiveness of the proposed approach for anomaly detection, as well as its satisfactory performance in both known-mode identification and unknown-mode discovery.

1. Introduction

As modern industrial production advances toward intelligence and sustainability, the demand for stable and high-quality operation of industrial systems has increased. In practical process industries, system abnormalities may cause equipment damage, reduced production efficiency, substandard product quality, or even serious accidents (Dong et al., 2025; Zhao et al., 2026). Therefore, process monitoring plays a crucial role in real-time evaluation and fault alarming during system operation, which is essential for timely adjustment of control strategies, maintaining safe and efficient production, and preventing cascading failures (Deng et al., 2023a). With the widespread adoption of intelligent information acquisition systems, data-driven industrial process monitoring has become an important research topic (Huang et al., 2025). Representative approaches include statistical-based methods (Zhang et al., 2023; Zheng et al., 2025a) and various machine learning algorithms that have been extended to process monitoring tasks (Chen et al., 2026; Wu et al., 2024; Yin et al., 2025).

In practical industrial production, variations in raw material composition, product quality, and energy consumption often necessitate adjusting the setpoints of operating variables, leading to mode switching. In this context, the multi-mode nature of industrial processes refers to multiple steady operating points under different operating conditions (Chen et al., 2024b). Therefore, multi-mode industrial process monitoring better satisfies the needs of real-world production. In general, existing approaches can be divided into two categories: methods that construct an integrated model for all modes (Fu et al., 2024a, 2025; Xu et al., 2021), and those that build an individual sub-model for each mode (Deng et al., 2021; Lyu et al., 2024; Tan et al., 2023). For example, Xu et al. (2021) designed a manifold joint projection framework for process monitoring, which learns a unified adjacency matrix across multiple modes. Fu et al. (2025) developed a sparse discriminative global

* Corresponding authors.
E-mail addresses: ziqingdeng@csu.edu.cn (Z. Deng), xiaofangchen@csu.edu.cn (X. Chen), elpwang@ntu.edu.sg (L. Wang), lhcen@csu.edu.cn (L. Cen), ylwang@csu.edu.cn (Y. Wang), csuwxysd@outlook.com (S. Yu).

https://doi.org/10.1016/j.eswa.2026.132035
Received 29 November 2025; Received in revised form 25 February 2026; Accepted 9 March 2026
Available online 15 March 2026

model by capturing the collaborative representations among different modes. To more precisely capture the unique characteristics of each mode, Tan et al. (2023) proposed a topological model based on feature-specific separation for multi-mode process monitoring. Lyu et al. (2024) introduced a mixture probabilistic principal component analysis method for industrial process modeling and fault detection, in which each mode is represented by an independent sub-model. However, the parallel construction of sub-models increases computational time as the number of modes grows and may lead to the model over-dispersion problem (Chen et al., 2024c).

In addition, practical industrial production systems and equipment are often affected by various factors such as service time, seasonal variation, and geographical location, which lead to distributional differences in process data collected at different times or sites. As a result, a multi-mode process monitoring model trained under historical operating conditions or on old equipment tends to perform poorly when applied to the current operating conditions or new equipment with inconsistent data distributions (Deng et al., 2025b). Transfer learning provides a breakthrough solution to the challenge of different distributions between training and testing data (Fan et al., 2023; Liu, 2025). Among its various branches, domain adaptation (DA) has shown outstanding performance in overcoming domain discrepancies (Xu et al., 2026). Existing DA methods can generally be categorized according to their focus on samples (Warke et al., 2024; Zhu et al., 2024), models (Chen et al., 2024a; Wei et al., 2021), or features (Li et al., 2025; Ngo et al., 2026; Zheng et al., 2025b). Zhu et al. (2024) approached the problem from the sample perspective and designed a sample-weighted DA framework, which assigns weights to source samples based on their similarity to the target domain, thereby reducing the negative impact of irrelevant or highly divergent samples during model training. Model-based DA typically introduces cross-domain constraints to adapt model parameters or decision boundaries learned from the source domain to the target domain. For instance, Wei et al. (2021) proposed a multi-classifier joint DA method that adjusts the parameters of source classifiers by enforcing consistency between target-domain predictions and local neighborhood predictions. Feature-based DA has been more extensively studied, focusing on feature learning and distribution alignment. To further achieve geometric alignment between source and target domains, Zheng et al. (2025b) developed an adaptive graph-learning-based DA method to extract multi-granularity source features and semantically enhanced target features. Li et al. (2025) proposed a multi-kernel weighted joint DA network that aligns domain distributions through a combination of maximum mean discrepancies. Ngo et al. (2026) refined the DA feature space partitioning and reduced domain discrepancy by adjusting distances between sub-feature spaces. These three categories of DA methods provide effective approaches to address distribution shifts between training and testing data. Therefore, in view of the distributional differences in industrial process data, it is of great significance to develop DA-based models for cross-domain multi-mode industrial process monitoring.

Dictionary learning (DL) has been successfully and widely applied in fields such as image and signal processing (Deng et al., 2025a; Du et al., 2025). According to different dictionary construction mechanisms, DL can be categorized into analytical DL (ADL), synthetical DL (SDL), and dictionary pair learning (DPL) (Tang et al., 2019; Wang et al., 2025b; Zhang et al., 2021). For instance, Tang et al. (2019) incorporated subspace joint learning into ADL and proposed a discriminative structured ADL method for classification tasks. In general, ADL constructs dictionaries through predefined mathematical transformations or analytical formulations, featuring a simple architecture and low computational complexity. To obtain sparser representations and richer dictionary structures, Wang et al. (2025b) proposed an SDL method based on fuzzy sparse discriminative embedding, which jointly optimizes multiple loss terms to exploit more informative features. Considering that the strict sparsity constraint in SDL results in high computational cost, Zhang et al. (2021) improved DPL by introducing an incoherence constraint, enabling fast and effective DL modeling through the simultaneous learning of an analytical-synthetical dictionary pair. Recently, DL-based methods have outperformed traditional multivariate statistical approaches in process monitoring, owing to their stronger feature extraction and data representation capabilities (Fu et al., 2024b; Huang et al., 2020b; Liu et al., 2024). For example, Liu et al. (2024) developed a structured collaborative SDL method for multi-mode process monitoring in blast furnace ironmaking, achieving successful multi-mode identification and anomaly detection. Fu et al. (2024b) proposed a graph-embedded DPL method to capture more realistic low-dimensional representations by leveraging both global and local information, which was effectively applied to monitoring the magnesium smelting process. To address the challenge of distribution discrepancy between training and testing data, cross-domain multi-mode process monitoring methods based on DL have recently attracted increasing attention (Deng et al., 2023b; Huang et al., 2020a; Yang et al., 2021). Huang et al. (2020a) proposed a transfer SDL (TDL) approach for cross-domain monitoring, in which an initial dictionary is learned from the source domain and subsequently fine-tuned using target-domain data. To enhance cross-domain monitoring and mode classification performance, Yang et al. (2021) incorporated linear discriminant analysis into TDL to learn a robust domain-shared dictionary. However, these methods generally assume that the source and target domains share identical known modes while neglecting the potential existence of unknown modes. Moreover, the lack of labeled samples in target processes under new operating conditions or equipment has not been fully investigated, which limits their performance in real industrial applications.

To address the challenges of distribution discrepancy between training and testing samples, the presence of new testing modes, and the lack of labeled data, this paper proposes a progressive anchoring-driven consistent dual reconstruction (PACDR) method for cross-domain open-mode process monitoring. First, considering the potential existence of unknown modes in the target domain, an open-mode dual reconstruction framework based on DPL is developed by introducing an open-set differential mechanism. Through self-reconstruction and cross-reconstruction, domain knowledge is transferred and interacted via a shared dictionary pair, while the negative transfer risk caused by unknown modes is mitigated. Then, an intra- and inter-domain joint graph regularization is proposed, where intra-domain regularization enhances the discriminability of representation coefficients, and inter-domain regularization adaptively improves the geometric consistency between domains. To avoid interference from unreliable pseudo-labels, a progressive target anchoring strategy is designed to gradually select target samples with higher confidence pseudo-labels during each iteration. Finally, for real-time target-domain samples, anomaly detection, known-mode identification, and unknown-mode determination are achieved using the learned classifier parameter matrix and dictionary pair.

The main contributions of this paper can be summarized as follows:

(1) Considering the existence of unknown modes in the target domain, an open-mode dual reconstruction framework is proposed. By performing self- and cross-reconstructions based on a domain-shared dictionary, the framework achieves unified modeling and knowledge transfer across all modes. Moreover, an open-set differential mechanism is embedded to mitigate the misleading effect of unknown-mode information and facilitate better statistical distribution alignment.

(2) To reduce geometric distribution discrepancies between the source and target domains, the intra- and inter-domain graph regularization is designed. From two complementary perspectives, graphs are constructed to adaptively refine the geometric distances of domain representations, thereby enhancing the discriminability of the dual reconstruction and the structural consistency across domains.

(3) To address the lack of target labels, a novel pseudo-labeling strategy is developed, which progressively anchors reliable pseudo-labels according to the confidence levels of known and unknown mode samples during iterative learning. Extensive industrial process experiments demonstrate that the proposed PACDR achieves superior performance in cross-domain open-mode process monitoring.

The paper proceeds as follows. Section 2 reviews the projective double reconstruction together with the open-set difference. Section 3 elaborates on the proposed cross-domain open-mode monitoring framework. Section 4 reports the experimental studies and analyses conducted on both simulated and real industrial processes. Section 5 concludes the study.

2. Preliminaries

2.1. Projective double reconstruction

To alleviate distribution discrepancies across different domains, transfer DL methods have been developed and successfully applied to cross-domain recognition tasks (Han et al., 2020; Li et al., 2019). Among them, the projective double reconstruction (PDR) approach inherits and extends the efficient modeling and low computational cost of DPL, providing a competitive solution for domain adaptation. Specifically, PDR treats the analytical dictionary as a projection matrix and performs self- and cross-reconstructions between the source and target domains by jointly learning domain-shared label information and a synthetical dictionary. This dual reconstruction strategy has been proven effective in aligning the statistical distributions between domains (Fang et al., 2022), thereby endowing the learned synthesis dictionary with strong transferability. The core framework of PDR can be formulated as follows:

$$\begin{aligned}&\min_{D,A}\left\|X_{st}-DAX_{st}\right\|_F^2+\left\|X_{st}-DAX_{ts}\right\|_F^2\\&+\Psi\left\|D,A,X_{st},L\right\|,\quad \text{s.t. } d_j^Td_j=1\end{aligned}\tag{1}$$

where $X_{st}=\left[X_s^1,X_t^1,\ldots,X_s^c,X_t^c,\ldots,X_s^CX_t^C\right]$ and $X_{ts}=\left[X_t^1,X_s^1,\ldots,X_t^c,X_s^c,\ldots,X_t^CX_s^C\right]$ are the data matrices reordered from the source samples $X_s\in\mathbb{R}^{d\times n_s}$ and the target samples $X_t\in\mathbb{R}^{d\times n_t}$, while $X_s^c\in\mathbb{R}^{d\times n_s^c}$ and $X_t^c\in\mathbb{R}^{d\times n_t^c}$ denote the cth class of X_s and X_t, respectively. d is the dimension of samples, n_s and n_t denote the sample numbers of X_s and X_t. $D=\left[d_1,\ldots,d_i,\ldots d_m\right]\in\mathbb{R}^{d\times m}$ and $A=\left[a_1;\ldots;a_i;\ldots;a_m\right]\in\mathbb{R}^{m\times d}$ represent the synthetical and analytical dictionaries, each containing m atoms. $\Psi\left\|D,A,X_{st},L\right\|$ represents the discriminative function term, while L is the label matrix obtained from the label set $\left\{Y_{st}^i\right\}_{i=1}^n$ of X_{st}, $n=n_s+n_t$. $d_j^Td_j=1$ constrains the scale of the dictionary to prevent degeneration of DPL. In Eq. (1), X_s and X_t share the same C classes of labels, without considering the potential emergence of new classes in the target domain.

2.2. Open-set difference

In practice, testing samples may originate from classes that are not necessarily included in the training set (Zhang & Patel, 2017). Conventional DA methods lack the capability to identify unknown classes. Consequently, open-set DA has been explored, with mainstream approaches focusing on detecting and excluding unknown classes to mitigate negative transfer (Fang et al., 2021; Wang et al., 2025a). As a clearly defined open-set recognition paradigm, the open-set difference has been proposed and rigorously validated (Fang et al., 2021), and its definition is given as follows:

$$\Delta_o=\frac{R_{u,C+1}^t(h)}{1-\pi_{C+1}^t}-R_{u,C+1}^s(h)\tag{2}$$

where $R_{u,C+1}^t(h)=\mathbb{E}\ell\left(h\left(X_t\right),y_{C+1}\right)$ and $R_{u,C+1}^s(h)=\mathbb{E}\ell\left(h\left(X_s\right),y_{C+1}\right)$ represent the risks that samples from the target and source domains are identified as unknown classes. The prediction function is denoted by $h(\cdot)$, while $y_{C+1}\in\mathbb{R}^{(C+1)\times1}$ is a one-hot vector representing the unknown class label in the target domain. More specifically, $R_{u,C+1}^t(h)$ is designed as a positive term, which encourages more target samples to be identified as belonging to the unknown class, while $R_{u,C+1}^s(h)$ serves as a negative term to prevent source samples from being misclassified into the unknown class. π_{C+1}^t denotes the prior probability of the unknown class. Inspired by open-set recognition, open-mode process monitoring is defined as the domain-specific instantiation of open-set learning for industrial systems. In this context, the abstract concept of a class corresponds to a specific physical operating mode. Unlike traditional multimode process monitoring, which operates under the closed-set assumption that all operating mode types are known a priori, the open-mode framework specifically addresses the challenge of detecting and identifying novel, unknown operating modes in the target domain.

3. Proposed method

As illustrated in Fig. 1, the proposed cross-domain open-mode process monitoring method based on PACDR consists of two main procedures: offline modeling and online monitoring. In the PACDR modeling stage, domain-shared knowledge of known modes is learned and exchanged through open-mode dual reconstruction, while the interference caused by unknown target-mode information is effectively reduced. Meanwhile, an intra- and inter-class joint graph regularization is employed to simultaneously enhance the distributional consistency between the source and target domains from two complementary perspectives. By the mutual promotion of statistical and structural distribution alignment, the transferability of the proposed PACDR is further strengthened. In addition, to improve the reliability of pseudo-labels, adaptive anchoring strategies are designed separately for known and unknown target-mode samples. During the online monitoring stage, the learned dictionary pair and calculated control thresholds are utilized to perform anomaly detection, unknown-mode discovery, and known-mode identification for real-time target-domain samples. The detailed formulations of the proposed method are presented as follows.

3.1. Open-mode dual reconstruction

Considering the domain distribution discrepancy and the presence of unknown modes in the target domain, an open-mode dual reconstruction framework is designed to achieve distribution alignment and knowledge transfer of domain-shared modes through self- and cross-reconstructions, while the embedded open differential mechanism effectively alleviates the negative transfer caused by unknown modes. First, to reorganize the data matrix $X=\left[X_s,X_t\right]\in\mathbb{R}^{d\times n}$, two permutation matrices, $Q_{st}\in\mathbb{R}^{n\times n}$ and $Q_{ts}\in\mathbb{R}^{n\times n}$, are introduced, where each column is a one-hot vector, $Q_{st}^TQ_{st}=I$, and $Q_{ts}^TQ_{ts}=I$. By applying Q_{st} and Q_{ts} to X, the reordered data matrices $\tilde{X}_{st}=XQ_{st}=\left[X_s^1,X_t^1,\ldots,X_s^c,X_t^c,\ldots,X_s^CX_t^C,X_t^{C+1},\hat{X}_t\right]$ and $\tilde{X}_{ts}=XQ_{ts}=\left[X_t^1,X_s^1,\ldots,X_t^c,X_s^c,\ldots,X_t^CX_s^C,X_t^{C+1},\hat{X}_t\right]$ are obtained. Here, $c(c\le C)$ denotes the modes shared between the two domains, X_t^{C+1} represents the unknown modes in the target domain, and $\hat{X}_t$ corresponds to the unanchored target samples. To further clarify the reordering procedure, an example is provided in Fig. 2.

The dual reconstruction is then formulated as follows:

$$\mathcal{G}_d=\left\|\tilde{X}_{st}-DA\tilde{X}_{st}\right\|_F^2+\left\|\tilde{X}_{st}-DA\tilde{X}_{ts}\right\|_F^2\tag{3}$$

where the first term corresponds to self-reconstruction, and the second term to cross-reconstruction. To reduce the misleading effect of X_t^{C+1} on knowledge transfer in Eq. (3), we modify and embed the open-set difference to separate X_t^{C+1} from the domain-shared known modes. Due to the existence of a high-probability error bound between the true risk and the empirical risk, Eq. (2) can be approximated by its empirical estimate $\hat{\Delta}_o=\hat{R}_{u,C+1}^t(h)/\left(1-\pi_{C+1}^t\right)-\hat{R}_{u,C+1}^s(h)$, which is then used to construct a computable optimization objective (Fang et al., 2021). Here, $\hat{R}_{u,C+1}^t(h)$ and $\hat{R}_{u,C+1}^s(h)$ denote the empirical risks corresponding to $R_{u,C+1}^t(h)$ and $R_{u,C+1}^s(h)$ in Eq. (2), respectively. With the squared loss adopted as the loss function $\ell(\cdot)$, the risk that training samples are regarded as unknown classes can be expressed as follows:

$$\hat{R}_{u,C+1}^t(h)=\frac{1}{n_t}\sum_{x_i\in X_t}\left\|h\left(x_i\right)-y_{C+1}\right\|_2^2\tag{4}$$

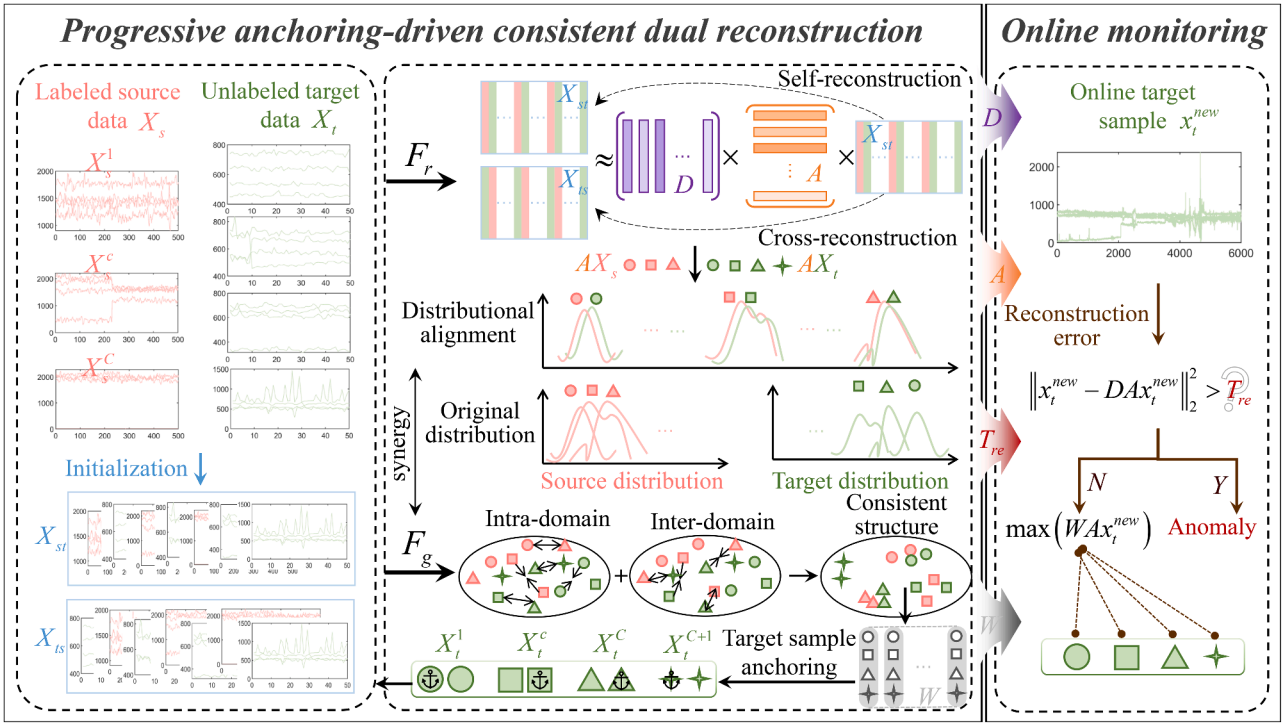


Fig. 1. Overview of cross-domain open-mode process monitoring based on PACDR.

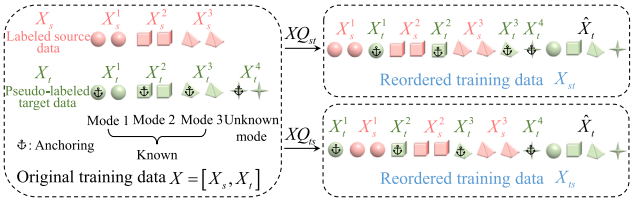


Fig. 2. Illustration of open-mode reordering across source and target domains.

$$\hat{R}^s_{u,C+1}(h) = \frac{1}{n_s}\sum_{x_i \in X_s}\left\|h(x_i) - y_{C+1}\right\|_2^2 \tag{5}$$

In addition, the supervised empirical risk on the source domain is computed as

$$\hat{R}^s(h) = \frac{1}{n_s}\sum_{x_i \in X_s}\left\|h(x_i) - y_i\right\|_2^2 \tag{6}$$

where $y_i \in \{y_c\}_c^C$, and $y_c \in \mathbb{R}^{(C+1)\times 1}$ is a one-hot label vector. Given the classifier parameter matrix W, $WA\tilde{X}_{st}$ represents the prediction matrix for the training samples. Consequently, by combining the supervised empirical risk of the source domain with the empirical open-set difference risk, we obtain the following expression for the overall empirical risk:

$$\begin{aligned}\mathcal{G}_h &= \frac{1}{n_s}\sum_{x_i \in X_s}\left\|h(x_i) - y_i\right\|_2^2 + \frac{\alpha}{n_t}\sum_{x_i \in X_t}\left\|h(x_i) - y_{C+1}\right\|_2^2 \\ &\quad - \frac{\beta}{n_s}\sum_{x_i \in X_s}\left\|h(x_i) - y_{C+1}\right\|_2^2 \\ &= \left\|(H - WA\tilde{X}_{st})\Lambda\right\|_F^2 - \beta\left\|(\bar{H} - WA\tilde{X}_{st})\bar{\Lambda}\right\|_F^2\end{aligned} \tag{7}$$

where $H \in \mathbb{R}^{(C+1)\times(n_s+n_t)}$ and $\bar{H} \in \mathbb{R}^{(C+1)\times(n_s+n_t)}$ are the positive and negative label matrices, respectively, which are defined as follows:

$$H_{ij} = \begin{cases} 1, & x_j \in X_s^i,\ i \le C \\ 1, & x_j \in X_t^i,\ i = C+1 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$\bar{H}_{ij} = \begin{cases} 1, & x_j \in X_s^i,\ i = C+1 \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

In Eq. (7), $\Lambda \in \mathbb{R}^{n\times n}$ and $\bar{\Lambda} \in \mathbb{R}^{n\times n}$ are diagonal matrices. When $x_i \in X_s$, the (i,i)th diagonal entry equals $\Lambda_{ii} = \bar{\Lambda}_{ii} = \sqrt{(1/n_s)}$. If $x_i \in X_t$, then $\Lambda_{ii} = \sqrt{\alpha/n_t}$ and $\bar{\Lambda}_{ii} = 0$. α and β are trade-off parameters that balance the corresponding terms. Accordingly, by jointly considering Eqs. (3) and (7), the proposed open-mode dual reconstruction can be formulated as follows:

$$\begin{aligned}F_r &= \left\|\tilde{X}_{st} - DA\tilde{X}_{st}\right\|_F^2 + \left\|\tilde{X}_{st} - DA\tilde{X}_{ts}\right\|_F^2 \\ &\quad + \left\|(H - WA\tilde{X}_{st})\Lambda\right\|_F^2 - \beta\left\|(\bar{H} - WA\tilde{X}_{st})\bar{\Lambda}\right\|_F^2 \\ &\quad + \delta\left(\|W\|_F^2 + \left\|A^T\right\|_{2,1}\right)\end{aligned} \tag{10}$$

where $\|W\|_F^2$ is a regularization term to prevent overfitting. $\|A^T\|_{2,1}$ imposes an $l_{2,1}$-norm constraint on A, aiming to enhance the robustness of DPL against noise and outliers (Deng et al., 2023a). δ represents the regularization parameter.

3.2. Intra- and inter-domain joint graph regularization

While reducing the statistical distribution discrepancy, a novel joint graph regularization is developed to more precisely align the geometric structures of the two domains and enhance the discriminability of the representation coefficients from both intra-domain and inter-domain perspectives. This design further facilitates effective domain adaptation by jointly optimizing structural consistency and feature separability. In the proposed method, X_s and X_t are mapped into a shared feature space

through A, yielding the encoding coefficients AX_s and AX_t. For intra-domain graph construction, we aim for the encoding coefficients of samples belonging to the same mode within each domain to be closely connected, while those of different modes are encouraged to remain distant. The intra-domain graph regularization can therefore be expressed as follows:

$$\min_A \frac{1}{2}\sum_{i,j=1}^{n}\left\|Ax_i - Ax_j\right\|_2^2 M_w^{ij} \tag{11}$$

where M_w denotes the edge weight matrix of the intra-domain graph, which is defined as follows:

$$M_w^{ij} = \begin{cases} \cos(x_i, x_j), & \text{if } x_i, x_j \text{ belong to the same domain and } y_i = y_j \neq C+1 \\ -\cos(x_i, x_j), & \text{if } x_i, x_j \text{ belong to the same domain and } y_i \neq y_j \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Besides, to further reduce structural distribution discrepancies between domains, multiple mode-specific domain discriminators are employed to obtain the inter-domain edge weight matrix M_b. This allows higher weights to be assigned to the representation coefficients of the same mode that are both geometrically close and structurally similar across different domains, while lower weights are assigned to those of different modes with significant structural discrepancies and larger distances. Specifically, when $x_i \in X_s$, $x_j \in X_t$, $y_i = k$ and $y_j = g$, a linear domain discriminator f_{kg} based on the support vector machine (SVM) is trained using the representation coefficients of the kth mode in the domain of x_i and the gth mode in the domain of x_j. If $k = g$, the weight between the corresponding vertices i and j is calculated as

$$\sigma_w^{ij} = \frac{n}{C\cdot\left(n_s^k + n_t^g\right)} \cdot \frac{1}{1+\exp\left[\left(l_i^2 + l_j^2\right)/2\right]} \tag{13}$$

where l_i and l_j denote the distances from Ax_i and Ax_j to the decision boundary of f_{kg}, respectively. When $k \neq g$, the corresponding weight is set as $\sigma_b^{ij} = -\sigma_w^{ij}$. Accordingly, the inter-domain graph regularization can be formulated as

$$\min_A \frac{1}{2}\sum_{i,j=1}^{n}\left\|Ax_i - Ax_j\right\|_2^2 M_b^{ij} \tag{14}$$

where the adaptive edge weight M_b^{ij} is computed as follows:

$$M_b^{ij} = \begin{cases} \sigma_w^{ij}, & \text{if } x_i, x_j \text{ belong to different domains and } y_i = y_j \\ \sigma_b^{ij}, & \text{if } x_i, x_j \text{ belong to different domains and } y_i \neq y_j \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The inter-domain graph regularization incorporates the similarity across different modes to calculate edge weights, thereby mitigating the unreliability of weights derived solely from Euclidean distance or cosine similarity in the presence of domain shift. Based on the modal label information from different domains, individual linear discriminators are trained for each inter-domain mode. Consequently, this classification-aware metric is more appropriately suited for aligning representation coefficients within a multi-mode manifold structure. By effectively measuring the similarity between the source and target domains, the construction of the inter-domain graph enhances the reliability of labeling for target modes, which is based on the distance to source-domain class prototypes. Therefore, the primary function of the inter-domain graph is to ensure cross-domain consistency. It serves as a complementary mechanism to the intra-domain graph, with both components mutually reinforcing one another to achieve a more refined alignment of the cross-domain geometric distributions. Based on Eqs. (11) and (14), the proposed intra- and inter-domain joint graph regularization can be expressed as follows:

$$\begin{aligned} F_g &= \frac{1}{2}\sum_{i,j=1}^{n}\|Ax_i - Ax_j\|_2^2 M_w^{ij} + \frac{1}{2}\sum_{i,j=1}^{n}\|Ax_i - Ax_j\|_2^2 M_b^{ij} \\ &= \mathrm{Tr}\left(AX_{st}LX_{st}^T A^T\right) \end{aligned} \tag{16}$$

where $L = T - M$ denotes the Laplacian matrix, $T = diag\left(t_1, \ldots, t_i, \ldots, t_n\right)$ is a diagonal matrix, $t_i = \sum_j M^{ij}$, and $M^{ij} = M_w^{ij} + M_b^{ij}$.

3.3. Progressive target-anchoring strategy

At present, most unsupervised DA methods directly use all pseudo-labeled target samples for iterative training. However, some unreliable pseudo-labeled samples may deteriorate model performance. To mitigate this issue, a progressive target anchoring strategy is developed, which gradually anchors target samples with high-confidence pseudo-labels for both known and unknown modes during each iteration of DPL, thereby ensuring robust open-mode reconstruction performance. Specifically, in each iteration, for an unanchored target sample $x_t^i \in X_t$, the mode prediction score is computed as $\mathrm{score}_t^i = WAx_t^i$. Let $\hat{y}_1 = u(u \le C)$ and $\hat{y}_2 = v(v \le C)$ denote the pseudo-labels with the first and second highest scores, respectively, and the corresponding class prototypes of source modes are computed as $p_u = 1/n_s^u \sum_{j=1}^{n_s^u} Ax_s^j\left(x_s^j \in X_s^u\right)$ and $p_v = 1/n_s^v \sum_{h=1}^{n_s^v} Ax_s^h\left(x_s^h \in X_s^v\right)$. x_t^i is anchored to a known mode u when the following condition is satisfied:

$$dist\left(Ax_t^i, p_u\right) < dist\left(Ax_t^i, p_v\right) \tag{17}$$

where $dist(\cdot)$ denotes the Euclidean distance. In addition, let the probability of an anchored known-mode sample x_t^j belonging to mode c be $P_j^c = \exp\left(W_c Ax_t^j\right)/\sum_{c=1}^{C}\exp\left(W_c Ax_t^j\right)$, and W_c is the cth row vector of W. Correspondingly, the self-entropy of x_t^j can be computed as $\mathcal{H}\left(x_t^j\right) = -1/\log C\sum_{c=1}^{C} P_j^c \log P_j^c$. Assume that the pseudo-label of a target sample $x_t^q \in X_t$ is $\hat{y}_q = \arg\max_c\left(WAx_t^q\right) = C+1$. The anchoring threshold for x_t^q is then defined as the maximum self-entropy among all anchored known-mode samples, which is defined as follows:

$$\eta = \max\left\{\mathcal{H}\left(x_t^j\right)\middle| x_t^j \notin \hat{X}_t, \hat{y}_j = c, c \le C\right\}. \tag{18}$$

Accordingly, the condition under which x_t^q is anchored as mode $C+1$ is defined as follows:

$$H\left(x_t^j\right) > \eta. \tag{19}$$

Actually, the self-entropy-based adaptive threshold in Eq. (18) establishes a conservative upper bound on the uncertainty of anchored known-mode samples. This mechanism constrains the selected known-mode samples to low-uncertainty levels while simultaneously anchoring unknown-mode samples with higher uncertainty. Thus, reliable pseudo-labeled samples are anchored by Eqs. (17) and (19), while low-confidence pseudo-labels that remain unanchored are effectively filtered out.

3.4. Modeling and optimization of PACDR

By integrating Eqs. (10) and (16), the objective function of the proposed PACDR can be obtained as follows:

$$\begin{aligned} &\min_{D,A,W}\left\|\tilde{X}_{st} - DA\tilde{X}_{st}\right\|_F^2 + \left\|\tilde{X}_{st} - DA\tilde{X}_{ts}\right\|_F^2 \\ &+\left\|\left(H - WA\tilde{X}_{st}\right)\Lambda\right\|_F^2 - \beta\left\|\left(\bar{H} - WA\tilde{X}_{st}\right)\bar{\Lambda}\right\|_F^2 \\ &+\lambda Tr\left(AX_{st}LX_{st}^T A^T\right) + \delta\left(\|W\|_F^2 + \left\|A^T\right\|_{2,1}\right) \\ &\text{s.t. } d_i^T d_i = 1 \end{aligned} \tag{20}$$

where λ is a weighting parameter. By incorporating the variable matrices $R_{st} \approx A\tilde{X}_{st}$ and $R_{ts} \approx AX_{ts}$, Eq. (20) is relaxed as follows:

$$\begin{aligned}&\min_{D,A,R_{st},R_{ts},W} \|\tilde{X}_{st} - DR_{st}\|_F^2 + \|\tilde{X}_{st} - DR_{ts}\|_F^2 \\ &+\sigma\left(\|A\tilde{X}_{st} - R_{st}\|_F^2 + \|A\tilde{X}_{ts} - R_{ts}\|_F^2\right) \\ &+\left\|(H - WR_{st})\Lambda\right\|_F^2 - \beta\left\|(\bar{H} - WR_{st})\bar{\Lambda}\right\|_F^2 \\ &+\lambda Tr\left(R_{st}LR_{st}^T\right) + \delta\left(\|W\|_F^2 + \|A^T\|_{2,1}\right) \\ &\text{s.t. } d_i^T d_i = 1, R_{st} \geq 0, R_{ts} \geq 0\end{aligned} \tag{21}$$

where σ is defined as a balancing parameter, while $R_{st} \geq 0$ and $R_{ts} \geq 0$ are employed to ensure non-negative coding coefficients. H is defined by Eq. (8) and constructed with source labels and anchored target pseudo-labels. By iteratively updating each variable, the minimization problem of Eq. (21) is solved as follows.

3.4.1. Update of the variable matrix R_{ts}

In updating R_{ts}, the variables D, A, R_{st}, and W are kept fixed, and all terms independent of R_{ts} are omitted. The resulting subproblem for R_{ts} in Eq. (21) can then be expressed as follows:

$$\begin{aligned}&\arg\min_{R_{ts}} \|\tilde{X}_{st} - DR_{ts}\|_F^2 + \sigma\|A\tilde{X}_{ts} - R_{ts}\|_F^2 \\ &\text{s.t. } R_{ts} \geq 0.\end{aligned} \tag{22}$$

For the constraint $R_{ts} = [r_{ts,ij}] > 0$ in Eq. (22), let $\psi_{ts} = [\varphi_{ts,ij}]$ be the corresponding Lagrange multiplier. Then, the Lagrangian function with respect to R_{ts} is formulated as

$$\zeta_{ts} = \|\tilde{X}_{st} - DR_{ts}\|_F^2 + \sigma\|A\tilde{X}_{ts} - R_{ts}\|_F^2 + Tr\left(\psi_{ts}R_{ts}^T\right). \tag{23}$$

Differentiating Eq. (23) with respect to R_{ts}, and subsequently deriving $\varphi_{ts,ij}r_{ts,ij} = 0$ via the Karush-Kuhn-Tucker conditions (Zhang et al., 2021), yields the following update for the (i,j)th element of R_{ts}:

$$r_{ts,ij}^{k+1} \leftarrow r_{ts,ij}^k \frac{\left(D^{Tk}\tilde{X}_{st}^k + \sigma A^k\tilde{X}_{ts}^k\right)_{ij}}{\left(D^{Tk}D^kR_{ts}^k + \sigma R_{ts}^k\right)_{ij}}. \tag{24}$$

3.4.2. Update of the variable matrix R_{st}

After fixing the other variables and eliminating the terms in Eq. (21) that do not depend on R_{st}, the optimization problem for R_{st} is derived as follows:

$$\begin{aligned}&\arg\min_{R_{st}} \|\tilde{X}_{st} - DR_{ts}\|_F^2 + \sigma\|A\tilde{X}_{st} - R_{st}\|_F^2 \\ &+\left\|(H - WR_{st})\Lambda\right\|_F^2 - \beta\left\|(\bar{H} - WR_{st})\bar{\Lambda}\right\|_F^2 \\ &+\lambda Tr\left(R_{st}LR_{st}^T\right), \quad \text{s.t. } R_{st} \geq 0.\end{aligned} \tag{25}$$

Similar to the update of R_{ts}, and considering the constraint $R_{st} = [r_{st,ij}] > 0$, we introduce a Lagrange multiplier $\psi_{st} = [\varphi_{st,ij}]$, leading to the following Lagrangian function with respect to R_{st}:

$$\begin{aligned}\zeta_{st} &= \|\tilde{X}_{st} - DR_{ts}\|_F^2 + \sigma\|A\tilde{X}_{st} - R_{st}\|_F^2 \\ &+\left\|(H - WR_{st})\Lambda\right\|_F^2 - \beta\left\|(\bar{H} - WR_{st})\bar{\Lambda}\right\|_F^2 \\ &+\lambda Tr\left(R_{st}LR_{st}^T\right) + Tr\left(\varphi_{st}R_{st}^T\right).\end{aligned} \tag{26}$$

By differentiating Eq. (26) in terms of R_{st} and applying the condition $\varphi_{st,ij}r_{st,ij} = 0$, the (i,j)th element of R_{st} is obtained as follows:

$$r_{st,ij}^{k+1} \leftarrow r_{st,ij}^k \frac{\begin{pmatrix} D^{Tk}\tilde{X}_{st}^k + \sigma A^k\tilde{X}_{st}^k + W^{Tk}H^k\cdot \\ \Lambda^k\Lambda^{Tk} + \beta W^{Tk}W^kR_{st}^k\bar{\Lambda}^k\bar{\Lambda}^{Tk}\end{pmatrix}_{ij}}{\begin{pmatrix} D^{Tk}D^kR_{st}^k + \sigma R_{st}^k + W^{Tk}W^kR_{st}^k\Lambda^k\cdot \\ \Lambda^{Tk} + \beta W^{Tk}\bar{H}^k\bar{\Lambda}^k\bar{\Lambda}^{Tk} + \lambda R_{st}^kL^k\end{pmatrix}_{ij}}. \tag{27}$$

3.4.3. Update of the analytical dictionary A

By removing the terms in which A is absent, Eq. (21) can be reduced to

$$\arg\min_A \sigma\left(\|A\tilde{X}_{st} - R_{st}\|_F^2 + \|A\tilde{X}_{ts} - R_{ts}\|_F^2\right) + \delta\left\|A^T\right\|_{2,1}. \tag{28}$$

In Eq. (28), according to the definition of $\|\cdot\|_{1,2}$ (Deng et al., 2025b), the third term can be rewritten as $2\delta Tr\left(AVA^T\right)$. Let V be a diagonal matrix whose (i,i)th diagonal element is given by $V_{ii} = 1/2\left\|\left(A^T\right)_i\right\|_2$, where $\left(A^T\right)_i$ denotes the ith row of A^T. To ensure numerical stability, $2\left\|\left(A^T\right)_i\right\|_2 + \varepsilon$ is utilized in practical calculations, where ε is a small constant. Subsequently, the iteratively reweighted least squares method is employed to solve Eq. (28). In the $k+1$th iteration, $V_{ii}^{k+1} = 1/\left(2\left\|\left(A^{Tk}\right)_i\right\|_2 + \varepsilon\right)$ is calculated based on the dictionary A^k updated in the kth iteration. Thus, with V^{k+1} and other variables fixed, a closed-form solution for A^{k+1} is derived as follows:

$$A^* = \left(\sigma R_{st}\tilde{X}_{st}^T + \sigma R_{st}\tilde{X}_{ts}^T\right)\left(\sigma\tilde{X}_{st}\tilde{X}_{st}^T + \sigma\tilde{X}_{ts}\tilde{X}_{ts}^T + 2\delta V\right)^{-1}. \tag{29}$$

3.4.4. Update of the synthetical dictionary D

To optimize D,Eq. (21) can be simplified to the following subproblem:

$$\begin{aligned}&\arg\min_D \|\tilde{X}_{st} - DR_{st}\|_F^2 + \|\tilde{X}_{st} - DR_{ts}\|_F^2 \\ &\text{s.t. } d_i^T d_i = 1.\end{aligned} \tag{30}$$

Let $\tilde{X} = [\tilde{X}_{st}, \tilde{X}_{st}]$, $R = [R_{st}, R_{ts}]$, and $K = \tilde{X} - \sum_{j\neq i} d_j r_j$, where r_j denotes the jth row of R. Then, the minimization problem of Eq. (30) can be rewritten as follows:

$$\arg\min_{d_i} \|K - d_i r_i\|_F^2, \quad \text{s.t. } d_i^T d_i = 1. \tag{31}$$

By employing a Lagrange multiplier φ to handle the constraint $d_i^T d_i = 1$, the corresponding function is obtained as

$$\mathcal{L}(d_i, \varphi) = \|K - d_i r_i\|_F^2 - \varphi\left(d_i^T d_i - 1\right). \tag{32}$$

Differentiating the Lagrangian function $\mathcal{L}(d_i, \varphi)$ with respect to d_i and equating the result to zero yields the following update for the atom of D:

$$d_i^* = \frac{Kr_i^T}{\left\|Kr_i^T\right\|_2}. \tag{33}$$

3.4.5. Update of classifier parameter matrix W

With D, A, R_{st}, and R_{ts} held fixed, the subproblem of W obtained from Eq. (21) can be formulated as follows:

$$\begin{aligned}&\arg\min_W \left\|(H - WR_{st})\Lambda\right\|_F^2 \\ &-\beta\left\|(\bar{H} - WR_{st})\bar{\Lambda}\right\|_F^2 + \delta\|W\|_F^2.\end{aligned} \tag{34}$$

The derivative of Eq. (34) in W, when enforced to be zero, leads to the following solution for W:

$$\begin{aligned}W^* = &\left(H\Lambda\Lambda^T R_{st}^T - \beta\bar{H}\bar{\Lambda}\bar{\Lambda}^T R_{st}^T\right)\cdot \\ &\left(R_{st}\Lambda\Lambda^T R_{st}^T - \beta R_{st}\bar{\Lambda}\bar{\Lambda}^T R_{st}^T + \delta I\right)^{-1}.\end{aligned} \tag{35}$$

To present the proposed PACDR algorithm in a clearer and more organized manner, it is summarized in Algorithm 1. Subsequently, the reconstruction error $re^i = \left\|x_{st}^i - DAx_{st}^i\right\|_2^2$ of the ith sample in X_{st} is computed using the updated dictionary pair D and A. Based on the kernel density estimation method (Huang et al., 2020b), the control threshold T_{re} for anomaly detection is then determined. Thus, the density estimate of the reconstruction error re for the considered sample is expressed as:

$$\hat{f}_e(re) = \frac{1}{nh}\sum_{i=1}^n K\left(\frac{re - re^i}{h}\right) \tag{36}$$

where $K(\cdot)$ typically denotes a Gaussian kernel, and h represents the bandwidth. Given a specified significance level θ, T_{re} is obtained as follows:

$$\int_{-\infty}^{T_{re}} \hat{f}_e(x)dx = 1 - \theta. \tag{37}$$

Algorithm 1 PACDR.

Input: Source data matrix X_s, Source label vector Y_s, target data matrix X_t, dictionary size m, maximum limit $T_{\max}$, and parameters α, β, λ, δ, and σ.

Initialization: Obtain the dictionary pair $D^{(0)}$ and $A^{(0)}$ with random matrices, label X_t as $\hat{Y}_t^{(0)}$ by the open-set nearest neighbor (OSNN) method [45], calculate the classifier parameter matrix $W^{(0)}$ via Eq. (35), and let $k = 0$.

While $k < T_{\max}$ **do**

1: Update $R_{ts}{}^{(k+1)}$ in accordance with Eq. (24);
2: Update $R_{st}{}^{(k+1)}$ in accordance with Eq. (27);
3: Update $A^{(k+1)}$ based on Eq. (29);
4: Update $D^{(k+1)}$ based on Eq. (33);
5: Obtain $W^{(k+1)}$ by computing Eq. (35);
6: Anchor $\hat{Y}_t^{(k+1)}$ via Eqs. (17) and (19);
7: Set $k = k + 1$;

end While

Output: $A = A^{(k)}$, $D = D^{(k)}$, and $W = W^{(k)}$.

3.5. Online monitoring

In the online process monitoring stage, three tasks are involved: anomaly detection, known-mode identification, and unknown-mode determination. For an online target sample x_t^{new}, its reconstruction error is computed using the dictionaries D and A obtained from the offline PACDR model as

$$re^{new} = \left\| x_t^{new} - DAx_t^{new} \right\|_2^2. \tag{38}$$

If $re^{new} > T_{re}$, x_t^{new} is detected as an anomaly. Otherwise, for a normal x_t^{new}, open-mode identification is performed. With the updated W, the identification is conducted as

$$\hat{y}^{new} = \arg\max_c \left(WAx_t^{new} \right). \tag{39}$$

When $re^{new} > T_{re}$, x_t^{new} is identified as belonging to the cth mode. In particular, if $c = C + 1$, x_t^{new} is determined as belonging to a newly emerging unknown mode.

3.6. Computational complexity

The computational complexity of the PACDR-based cross-domain open-mode process monitoring method is analyzed as follows. During the offline modeling procedure, the computational complexities for updating R_{ts}, R_{st}, A, D, and W in each iteration are $\mathcal{O}(dmn)$, $\mathcal{O}\left(dmn + m^2n + mn^2 + n_t^k \bar{n}_t^k + m\left(n_s + C\bar{n}_t^k\right)\right)$, $\mathcal{O}\left(dmn + d^2n + d^3\right)$, $\mathcal{O}\left(dmn + m^2n\right)$, and $\mathcal{O}\left((C+1)mn + mn^2 + m^2n\right)$, respectively. When updating R_{st}, the complexity of constructing the intra- and inter-domain joint graph is $\mathcal{O}\left(n_t^k \bar{n}_t^k + m\left(n_s + C\bar{n}_t^k\right)\right)$, where $\mathcal{O}\left(m\left(n_s + C\bar{n}_t^k\right)\right)$ represents the complexity of training the linear domain discriminator. In our method, the progressive target-anchoring strategy only necessitates calculating edge weights for target domain samples with high-confidence pseudo-labels. Accordingly, n_t^k and $\bar{n}_t^k$ denote the number of target samples anchored in the kth iteration and the cumulative number of anchored samples across k iterations, respectively. Under the condition $n_t^k < n_t \ll n_s < n$, the complexity of graph construction in PACDR is significantly reduced. Given a maximum of $T_{\max}$ iterations, the total complexity of the modeling procedure is $\mathcal{O}\left(T_{\max}\left(mn^2 + \left(dm + m^2 + Cm + d^2\right)n + d^3\right)\right)$. During the online monitoring procedure, the complexity of calculating the reconstruction error for each testing sample is $\mathcal{O}(dm)$. For a total of N_{new} testing samples, the overall monitoring complexity is $\mathcal{O}\left(dmN_{new}\right)$. Consequently, our method leverages the advantages of analytical encoding, eliminating the additional computational overhead typically required for sparse representation. This ensures superior real-time performance for industrial process monitoring.

4. Experimental study

In this section, extensive comparative experiments are conducted on the simulated Tennessee Eastman process (TEP), the industrial coal gasification process (CGP), and the industrial aluminum electrolysis process (AEP). The experimental settings, result analyses, and discussions are presented in detail in the following subsections.

4.1. Process description and data collection

(1) TEP. The TEP simulation platform is an industrial chemical process control model involving two simultaneous gas-liquid exothermic reactions. Under its five major operating units, the system provides 12 manipulated variables and 41 measured variables, offering publicly available data for validating industrial process monitoring methods. The TEP model specifies the parameter settings of six relatively stable operating modes and includes 28 types of process anomalies that can be implemented for evaluation (Deng et al., 2025a). After removing constant variables, 31 process variables are selected for sampling in the TEP experiments (Huang et al., 2020b).

(2) CGP. The CGP industrial system uses a gasifier as its core unit, where a gasification reaction between coal-water slurry and pure oxygen is carried out to produce synthesis gas primarily composed of carbon monoxide and hydrogen. A schematic diagram of the process flow is shown in Fig. 3. The quench chamber is typically located at the outlet of the gasifier and is responsible for rapidly cooling the high-temperature raw syngas. Its operation directly affects the stability of both the gasifier and the scrubbing tower. Consequently, monitoring the quench chamber liquid level and quench-water flow rate is essential for ensuring the efficient and stable operation of the CGP. In the CGP experiments, process data are collected from a coal gasification plant, where each sample contains 27 variables associated with the quench chamber.

(3) AEP. The electrolytic cell is the primary production unit in AEP, where molten alumina reacts with the carbon anode under direct current to produce carbon dioxide and molten aluminum (Deng et al., 2023a). The virtual simulation model of the electrolytic cell is shown in Fig. 4. In current industrial practice, cell conditions under typical operating modes are mainly monitored based on operate consensus and expert experience, which lacks systematic rigor and is unable to identify the emergence of new modes. Therefore, developing intelligent open-mode monitoring techniques provides valuable guidance for more refined operational decision-making and contributes to improving production efficiency. For the AEP experiments, samples are collected from multiple electrolytic cells across different aluminum electrolysis plants, where each sample consists of 24 anode currents representing the local behavior of the reaction cell (Deng et al., 2025b).

In the TEP experiments, the six operating modes are partitioned in two different ways to construct unknown modes in the target domain. In addition, to simulate domain distribution discrepancies, random perturbations following an exponential distribution $E(0.3)$ and the F distribution $0.1F(5,9)$ are applied to the collected process data. For the CGP experiments, data are collected under three operating modes corresponding to different working conditions, including samples acquired under abnormal quench chamber liquid levels and abnormal quench-water flow rates. For the AEP experiments, samples obtained under three typical operations: nonoperation, anode replacement, and metal tapping are used as known-mode data, while samples collected under a nonstandard cell-setting condition are treated as unknown-mode data. To evaluate the proposed method from both cross-condition and cross-equipment perspectives, samples are collected from a 400-kA electrolytic cell under different cell conditions, as well as from 400-kA and 440-kA electrolytic cell located in different plants. Furthermore, to better approximate the

Table 1
Experimental cases in different industrial processes.

Case	Task	Phase	Domain	Mode	Sample number
TEP_{case1}	Cross-condition	Training	Source	Modes 1, 2, and 3	1-1500
			Target	Modes 1, 2, 3, 4, 5, and 6	1501-1680
		Testing	Target	Modes 1, 2, 3, 4, 5, and 6	1681-2280
			Target	Mode 1 with each anomaly	2281-2580
TEP_{case2}	Cross-condition	Training	Source	Modes 1, 2, 3, 4, and 5	1-2500
			Target	Modes 1, 2, 3, 4, 5, and 6	2501-2680
		Testing	Target	Modes 1, 2, 3, 4, 5, and 6	2681-3280
			Target	Mode 1 with each anomaly	3281-3580
CGP_{case1}	Cross-condition	Training	Source	Modes 1 and 2	1-1000
			Target	Modes 1, 2, and 3	1001-1090
		Testing	Target	Modes 1, 2, and 3	1091-1390
			Target	Mode1 with abnormally low liquid level	1391-1690
CGP_{case2}	Cross-condition	Training	Source	Modes 1 and 2	1-1000
			Target	Modes 1, 2, and 3	1001-1090
		Testing	Target	Modes 1, 2, and 3	1091-1390
			Target	Modes 1 with abnormally low flow rate	1391-1690
AEP_{case1}	Cross-condition	Training	Source	Modes 1, 2, and 3	1-1500
			Target	Modes 1, 2, 3, and 4	1501-1620
		Testing	Target	Modes 1, 2, 3, and 4	1621-2020
			Target	Mode 1 with anode effect	2021-2320
AEP_{case2}	Cross-equipment	Training	Source	Modes 1, 2, and 3	1-1500
			Target	Modes 1, 2, 3, and 4	1501-1620
		Testing	Target	Modes 1, 2, 3, and 4	1621-2020
			Target	Mode 1 with anode effect	2021-2320

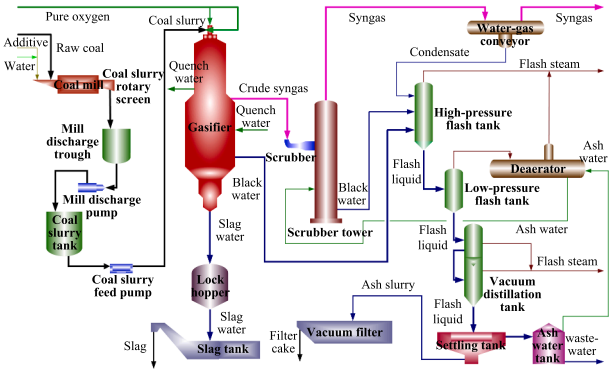


Fig. 3. Schematic diagram of CGP.

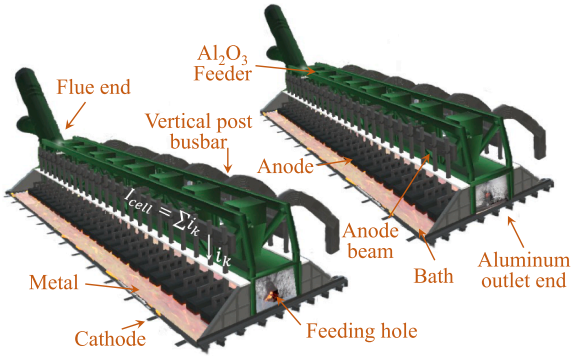


Fig. 4. Virtual simulation model of the electrolytic cell.

practical difficulty of target-domain sampling (Huang et al., 2020a), the numbers of training samples for each mode in the source and target domains are set to 500 and 30, respectively, whereas 100 samples per target mode and 300 abnormal samples are used for testing. A summary of all experimental cases is provided in Table 1.

Table 2
FDRs for 28 anomalies in Case 1 of TEP.

No.	SROSR	TDL	PDR	DSCDA	DFEDTL	**PACDR**
1	1.0000	1.0000	1.0000	0.9400	1.0000	**1.0000**
2	1.0000	1.0000	0.8767	1.0000	1.0000	**0.9200**
3	0.0100	0.0300	0.7200	0.5767	0.0433	**0.7867**
4	0.5400	0.6233	0.9533	0.6600	0.7133	**0.9600**
5	0.0600	0.6600	0.7200	0.7433	0.7400	**0.8367**
6	1.0000	1.0000	1.0000	0.0700	1.0000	**1.0000**
7	0.0133	1.0000	1.0000	1.0000	1.0000	**0.9933**
8	1.0000	1.0000	1.0000	0.3100	1.0000	**1.0000**
9	0.5933	0.4767	0.2167	0.7267	0.1533	**0.7433**
10	0.8267	0.5500	0.7967	0.3467	0.6567	**0.8533**
11	0.9200	0.4567	0.6100	0.2600	0.7900	**0.8867**
12	0.8967	0.7700	0.9033	0.9500	0.6900	**0.8800**
13	1.0000	1.0000	0.4233	0.6533	1.0000	**1.0000**
14	0.6400	0.7833	0.9833	1.0000	0.9033	**0.7167**
15	0.8700	0.8567	0.7533	0.7300	0.8233	**0.8933**
16	0.0100	0.0633	0.1633	0.7567	0.1933	**0.4933**
17	0.2800	0.7533	0.8933	0.8833	0.7133	**0.9700**
18	0.7867	0.6667	0.4467	0.3200	0.1967	**0.9300**
19	0.9500	0.4800	0.8933	0.5600	0.8167	**0.9367**
20	0.6700	0.9967	0.7367	0.2200	0.7400	**1.0000**
21	0.3400	0.0667	0.2133	0.3933	0.2700	**0.4567**
22	0.1900	0.7900	0.9800	0.8600	0.7567	**0.5067**
23	0.0767	0.2700	0.2433	0.6133	0.1967	**0.6267**
24	1.0000	0.9733	0.7033	0.8333	0.6600	**0.9533**
25	0.8267	0.6767	0.8033	0.4733	0.3667	**0.8433**
26	0.1533	0.8067	0.7367	0.9033	0.8267	**0.7833**
27	0.6567	0.8067	0.8167	0.6100	0.7433	**0.8767**
28	0.5000	0.4500	0.7867	0.7800	0.7700	**0.8267**

4.2. Evaluation metrics

In studies on process monitoring methods, the fault detection rate (FDR) and the false alarm rate (FAR) (Chen et al., 2024b) are commonly used to evaluate anomaly detection performance. They are defined as follows:

$$\mathrm{FDR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{40}$$

Table 3
FDRs for 28 anomalies in Case 2 of TEP.

No.	SROSR	TDL	PDR	DSCDA	DFEDTL	**PACDR**
1	1.0000	1.0000	1.0000	1.0000	1.0000	**1.0000**
2	0.9333	0.9500	0.8967	0.7667	0.9933	**1.0000**
3	0.1733	0.6267	0.4867	0.9367	0.0200	**0.9467**
4	0.4267	0.9433	0.2300	0.2033	0.5367	**0.9967**
5	0.4967	0.1100	0.4300	0.2967	0.4767	**0.4700**
6	1.0000	1.0000	1.0000	0.9200	1.0000	**1.0000**
7	0.8067	1.0000	0.0933	1.0000	1.0000	**1.0000**
8	1.0000	0.9333	1.0000	0.7267	1.0000	**1.0000**
9	0.7900	0.9800	0.9233	0.9767	0.5133	**0.9767**
10	0.9233	0.9800	0.7833	0.6367	0.3967	**0.8500**
11	0.1400	0.7467	0.5000	0.5067	0.6167	**0.8333**
12	0.1267	0.6967	0.5633	0.8267	0.5433	**0.8667**
13	1.0000	0.8700	0.4033	1.0000	1.0000	**0.9767**
14	0.2033	0.5333	0.6700	0.5300	0.5367	**0.8167**
15	0.6100	0.7733	0.5967	0.6400	0.5833	**0.7533**
16	0.1667	0.7533	0.9067	0.8933	0.4967	**0.9800**
17	0.8267	0.9133	0.6000	0.9767	0.7767	**0.8600**
18	0.4000	0.6000	0.9833	0.9233	0.8900	**0.9667**
19	0.1800	0.6700	0.9033	0.8067	0.8333	**0.9433**
20	0.9500	1.0000	0.5367	0.9967	0.7567	**1.0000**
21	0.2000	0.7333	0.9067	0.8867	0.9167	**0.9800**
22	0.8267	0.8167	0.8900	0.8433	0.8233	**0.9667**
23	0.9167	0.6400	0.8400	0.7367	0.9133	**0.9233**
24	0.9833	1.0000	0.7767	0.9867	0.6767	**0.9100**
25	0.8833	0.9400	0.8267	0.4333	0.7400	**0.9900**
26	0.3733	0.9967	0.8667	0.7333	0.8700	**0.9867**
27	0.4000	0.7900	0.7933	0.8967	0.8467	**0.9000**
28	0.8700	0.6333	0.8700	0.9200	0.8900	**0.9633**

$$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{41}$$

where TP denotes the number of correctly detected abnormal samples, FN is the number of missed abnormal samples, FP represents the number of normal samples that are incorrectly reported as abnormal, and TN is the number of correctly detected normal samples. To evaluate the performance of open-mode identification, the following metrics from open-set indicators (Fang et al., 2021) are adopted:

$$\mathrm{OM} = \frac{M_k}{N_k} \tag{42}$$

$$\mathrm{ACC} = \frac{M_k + M_u}{N_k + N_u} \tag{43}$$

$$\mathrm{MI} = \frac{M_k^{cond} + M_u^{cond}}{N_k + N_u} \tag{44}$$

where M_k denotes the number of correctly identified known-mode samples, N_k is the total number of known-mode samples, M_u represents the number of samples correctly identified as belonging to an unknown mode, N_u is the total number of unknown-mode samples, and $M_k^{cond} + M_u^{cond}$ denotes the number of samples that are detected as normal and whose modes are correctly identified. In addition, two time-based measures are used to evaluate the performance of the proposed method. The training time reflects the computational complexity of the model, whereas the testing time indicates its real-time capability and detection efficiency in online monitoring.

4.3. Implementation details

To validate the superior performance of the proposed PACDR approach in cross-domain open-mode process monitoring, a range of advanced open-set and cross-domain DL-based methods are employed for comparison, including the sparse representation-based open-set (SROSR) (Zhang & Patel, 2017) algorithm, TDL (Huang et al., 2020a), PDR (Han et al., 2020), dictionary-sharing coding-based DA (DSCDA) (Li et al., 2019), and discriminative Fisher embedding TDL (DFEDTL) (Fan et al., 2023). For closed-set methods, the one-class SVM (OCSVM) (Wang et al., 2025a) is additionally applied to determine unknown modes. For all methods, OSNN is uniformly adopted to initialize target pseudo-labels. In PACDR, following the parameter ranges in the open-set difference (Fang et al., 2021), the parameters α, β, λ, δ, and σ are determined based on the values corresponding to the highest average ACC across ten random sampling trials. To ensure fair comparison, DL-based benchmark methods adopt the same strategy for selecting their optimal parameters. The number of atoms, the maximum number of iterations, and the significance level are set to $m = 60$, $T_{max} = 10$, and $\theta = 0.05$, respectively. All experiments in this paper are conducted on a computer equipped with an Intel Core i7-13650HX CPU.

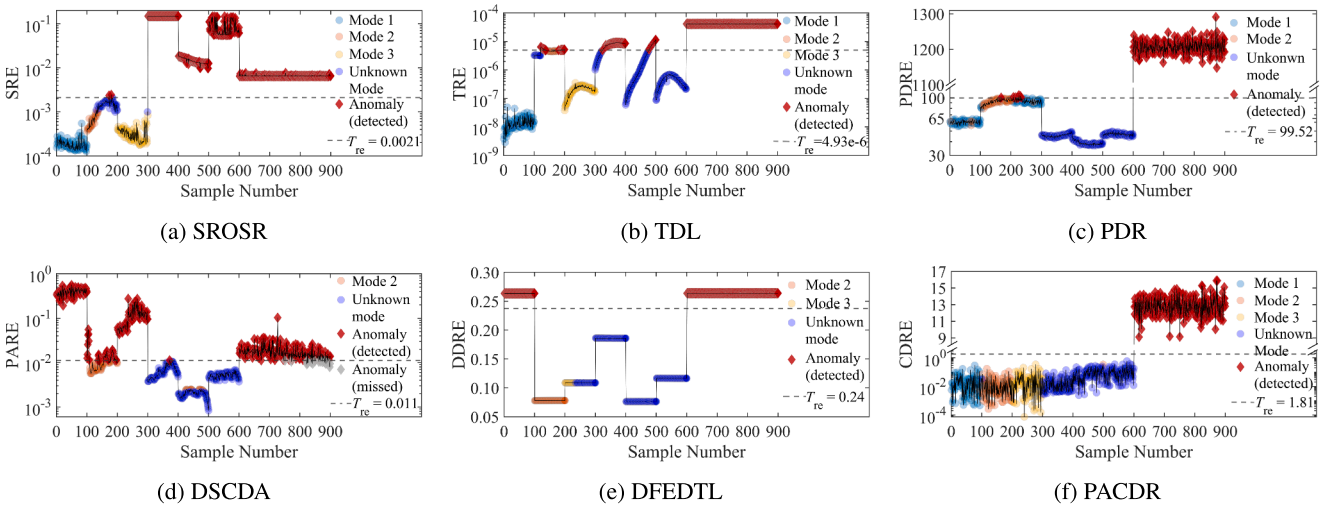
4.4. Experimental results and analyses

The detection results of all methods for the 28 anomalies in Case 1 and Case 2 of TEP are summarized in Tables 2 and 3. Taking the first fault as a representative, Figs. 5 and 6 illustrate the process monitoring outcomes of each method, and Tables 4 and 5 present the corresponding quantitative comparisons. The TEP experimental results demonstrate that, under both the single unknown-mode and multiple unknown-mode scenarios, the proposed method consistently exhibits clear advantages in detecting the 28 anomalies. From the multi-metric evaluations in Tables 4 and 5, it can be observed that all methods achieve a 100% FDR for the first anomaly, with the exception of DSCDA. Overall, PACDR achieves both low FARs and high ACCs. In contrast, while PDR maintains low FARs, it performs poorly in open-set mode identification, evidenced by an ACC of only 45.33% in Case 2. This indicates that cross-domain dual reconstruction can transfer source knowledge and accurately model the underlying structure of normal data in the target domain, enabling the calculation of effective monitoring statistics for testing samples via the learned transfer dictionary pair. However, PDR overlooks the manifold structures of both the source and target domains. Consequently, discrepancies in geometric distributions between domains cause the classifier parameter matrix trained by PDR to underperform in the target open-mode identification. Meanwhile, other comparative methods exhibit significant gaps between their ACCs and MIs, whereas PACDR demonstrates consistently superior performance in both metrics. Considering the definitions of ACC and MI, maintaining a low FAR is a prerequisite and fundamental guarantee for effective open-mode identification. In Case 1, both PDR and DFEDTL show large discrepancies between their OM and ACC results, mainly because they lack a dedicated decision mechanism for distinguishing between known and unknown modes, causing their discovery of unknown modes to rely heavily on OCSVM.

Furthermore, Figs. 7 and 8 illustrate the monitoring results of all methods on the CGP, and the corresponding quantitative comparisons are provided in Tables 6 and 7. The CGP results indicate that TDL and PACDR achieve both high FDR and low FAR, owing to their effective transfer reconstruction mechanisms. In contrast, SROSR is unable to perform domain adaptation, leading to its poor performance in terms of FAR. DFEDTL ignores the learning of data manifold structures, leading to unstable detection performance for complex process anomalies. In Case 1, PDR exhibits a high FAR and low ACC. This is attributed to the lack of a pseudo-label anchoring strategy in PDR to secure high-confidence pseudo-labels, resulting in heavy reliance on the quality of the initial target pseudo-labels. Hence, the PDR model, trained on unreliable unlabeled samples, performs poor reconstruction capability for normal modes. In contrast, our PACDR delivers consistently strong performance in both open-mode identification and anomaly detection. This is mainly because the open-set difference reduces the risk of misclassifying known source modes as unknown modes, while the geometric distribution consistency constraint and the target anchoring strategy sharpen the inter-mode decision boundaries, achieving accurate recognition and effective reconstruction of target modes.

Table 4
Experimental results for the first anomaly in Case 1 of TEP.

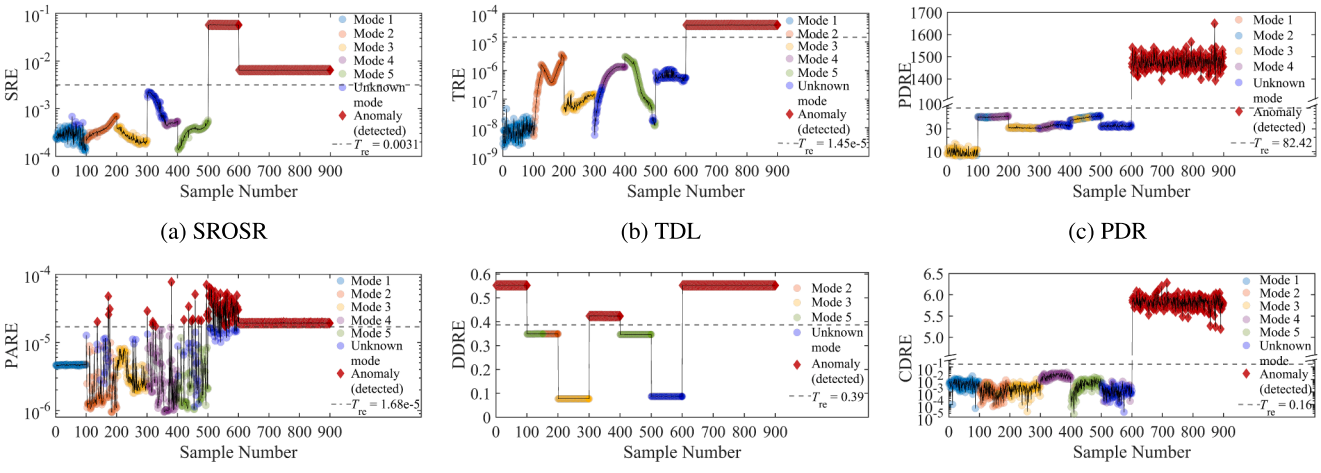
Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	100.00	50.33	39.17	78.33	89.17	4.4742	10.7096
TDL	100.00	21.83	74.67	93.00	96.50	60.2265	0.1199
PDR	100.00	1.50	81.17	62.67	81.33	1.4330	0.0085
DSCDA	94.00	41.17	52.83	95.33	91.67	49.5131	0.1095
DFEDTL	100.00	16.67	73.50	80.33	90.17	3.2486	3.3297
PACDR	**100.00**	**0.00**	**99.17**	**99.67**	**99.17**	**2.7565**	**0.0011**



(a) SROSR (b) TDL (c) PDR

(d) DSCDA (e) DFEDTL (f) PACDR

Fig. 5. Process monitoring for the first anomaly in Case 1 of TEP.



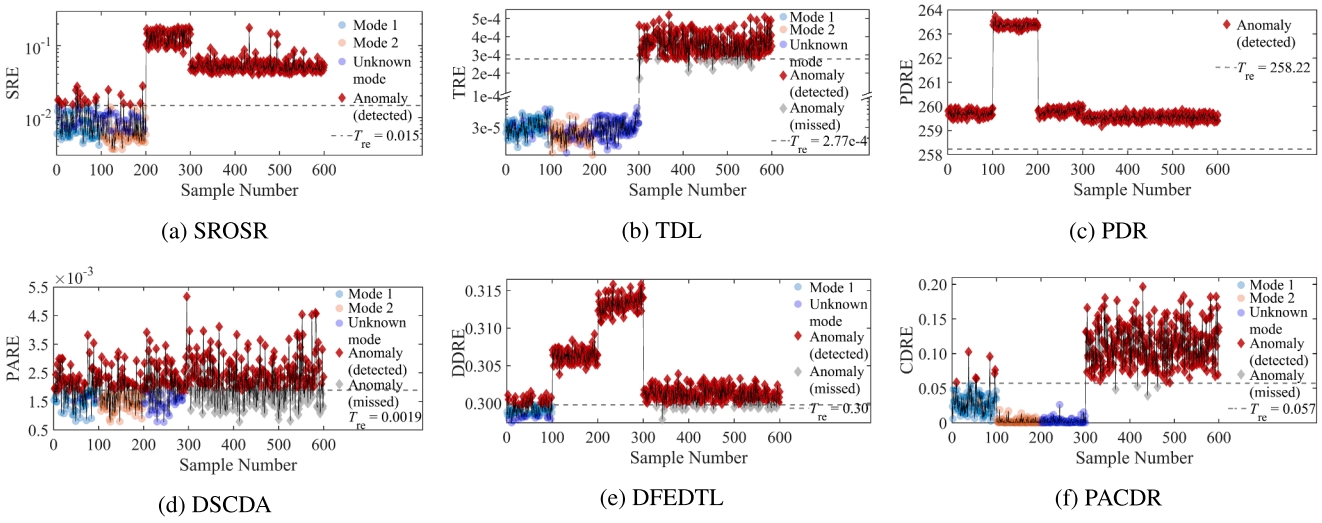
(a) SROSR (b) TDL (c) PDR

(d) DSCDA (e) DFEDTL (f) PACDR

Fig. 6. Process monitoring for the first anomaly in Case 2 of TEP.

Table 5
Experimental results for the first anomaly in Case 2 of TEP.

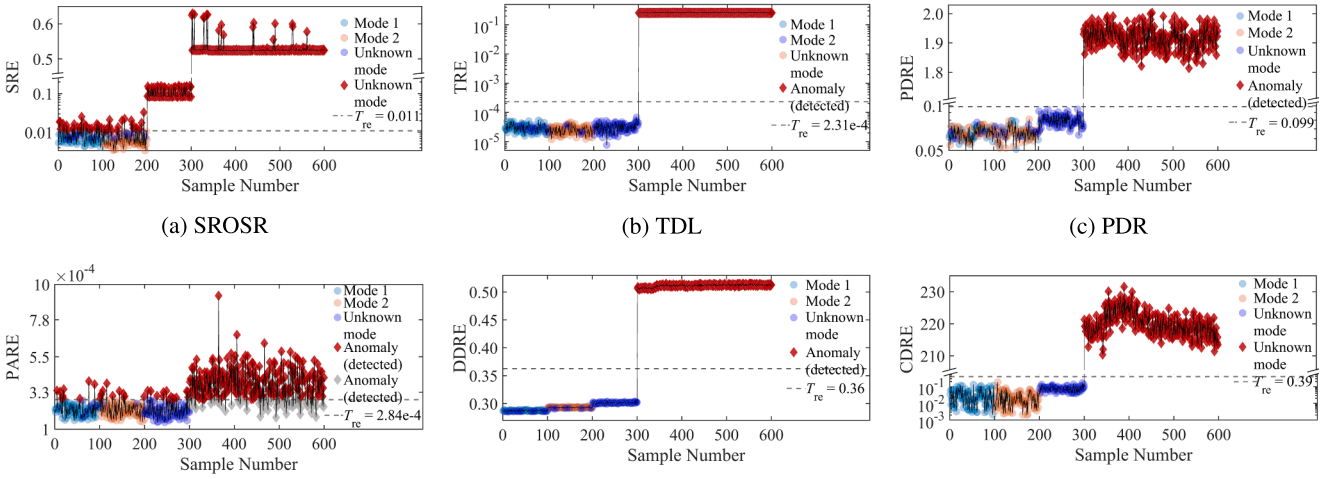
Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	100.00	16.67	73.00	87.60	89.67	5.1335	9.3590
TDL	100.00	0.00	95.00	94.00	95.00	162.6686	0.1838
PDR	100.00	0.00	45.33	34.40	45.33	4.7767	0.0211
DSCDA	100.00	15.50	77.67	89.80	91.50	37.5342	0.1296
DFEDTL	100.00	33.33	58.00	89.60	91.33	9.4414	11.1405
PACDR	**100.00**	**0.00**	**94.17**	**99.00**	**94.17**	**6.2478**	**0.0028**



(a) SROSR (b) TDL (c) PDR

(d) DSCDA (e) DFEDTL (f) PACDR

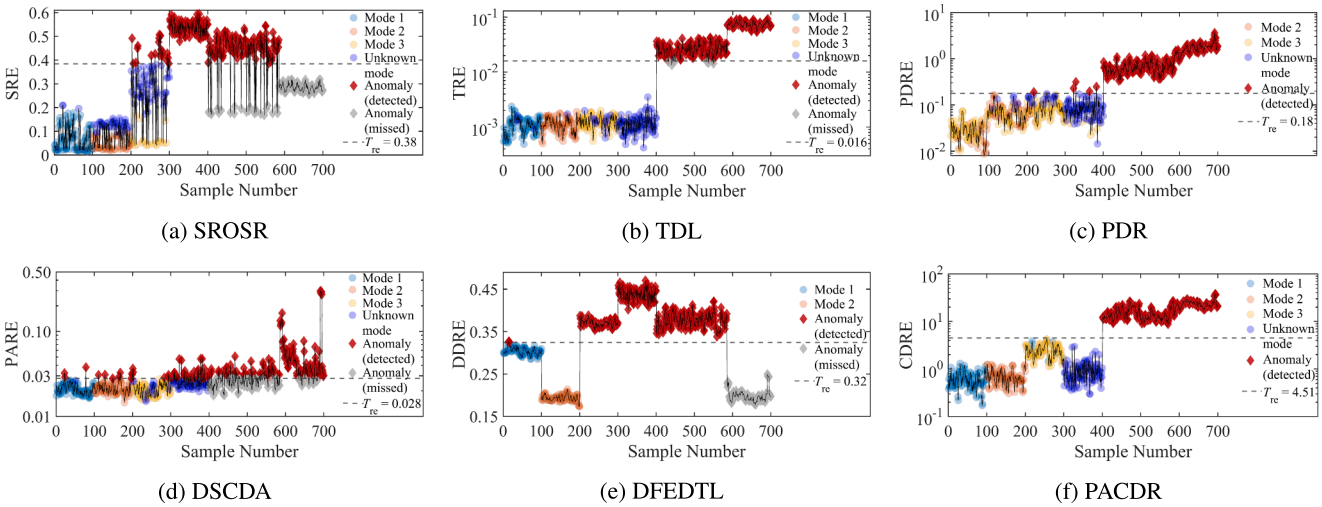
Fig. 7. Process monitoring for Case 1 of CGP.



(a) SROSR (b) TDL (c) PDR

(d) DSCDA (e) DFEDTL (f) PACDR

Fig. 8. Process monitoring for Case 2 of CGP.



(a) SROSR (b) TDL (c) PDR

(d) DSCDA (e) DFEDTL (f) PACDR

Fig. 9. Process monitoring for Case 1 of AEP.

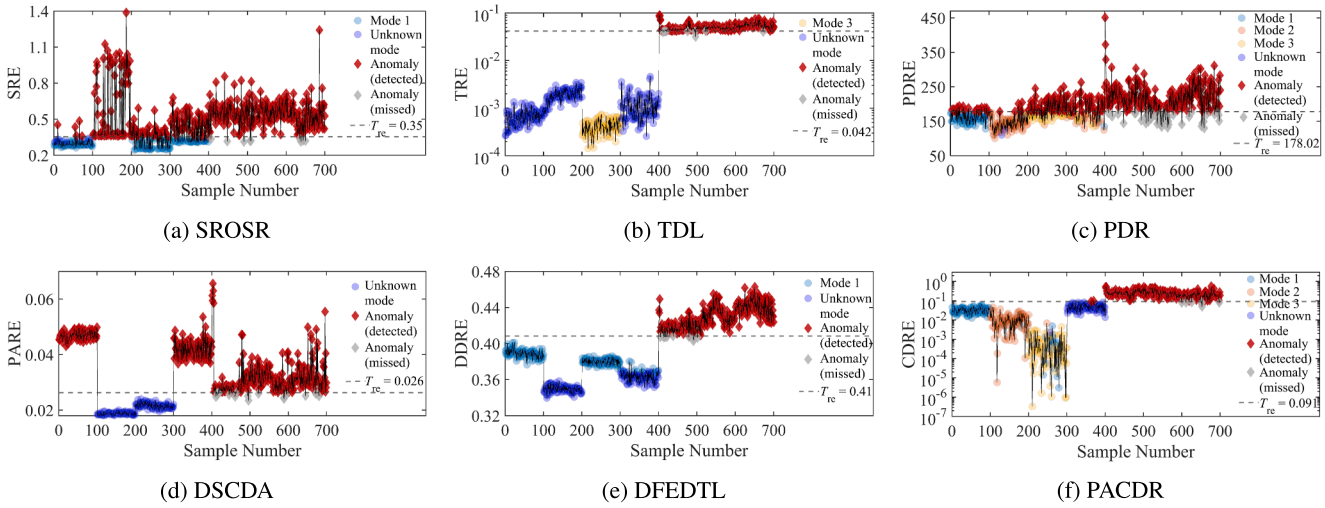


Fig. 10. Process monitoring for Case 2 of AEP.

Table 6
Experimental results for Case 1 of CGP.

Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	100.00	40.00	42.00	67.50	78.33	1.3183	3.9438
TDL	93.67	0.00	91.00	86.50	91.00	79.2263	0.1520
PDR	100.00	100.00	0.00	82.00	54.67	1.0871	0.0025
DSCDA	60.00	46.67	52.00	93.00	95.33	17.1187	0.0436
DFEDTL	91.00	76.67	17.67	86.00	90.67	1.2674	1.0737
PACDR	**98.67**	**2.33**	**97.67**	**100.00**	**100.00**	**1.4290**	**0.0018**

Table 7
Experimental results for Case 2 of CGP.

Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	100.00	49.33	42.00	70.50	80.33	1.8657	4.9630
TDL	100.00	0.00	93.67	90.50	93.67	79.3853	0.5782
PDR	100.00	0.00	68.33	52.50	68.33	0.8789	0.0036
DSCDA	79.00	11.33	85.33	94.00	96.00	16.4632	0.0398
DFEDTL	100.00	0.00	90.33	85.50	90.33	1.3000	1.1645
PACDR	**100.00**	**0.00**	**100.00**	**100.00**	**100.00**	**0.8464**	**0.0015**

In the AEP experiments, the cross-condition and cross-equipment open-mode monitoring results are presented in Figs. 9 and 10, and the corresponding quantitative comparisons are summarized in Tables 8 and 9. Under the cross-condition case, both TDL and PACDR perform well. Because the similarity between the source and target domains is relatively high, TDL can effectively exploit the source-domain information to learn the initial target dictionary. Although DSCDA and DFEDTL achieve high OM values, their anomaly detection performance is poor, indicating that the presence of an unknown mode increases the difficulty of modeling normal samples. In the cross-equipment case, where the domain discrepancy is much larger, the proposed PACDR shows a clear advantage over PDR in discovering unknown modes. This improvement is attributed to the higher reliability of the anchored target samples. According to testing time, processing 700 testing samples by PACDR takes 1.0 to 1.3 ms, i.e., 1.43 to 1.86 μs per sample. Given that the AEP sampling interval is 10 s, the online inference overhead is negligible relative to the sampling period.

From Table 4 through Table 9, it can be observed that TDL and DSCDA incur the longest training times, primarily because their strict sparsity constraints significantly increase the complexity of DL. In addition, the two-stage training strategy of TDL places a substantial computational burden on the model as the number of training samples grows. During the testing phase, SROSR requires the modeling of extreme value distribution to determine unknown modes, and DFEDTL involves additional iterative computation to obtain the coding coefficients. Hence, both methods exhibit longer testing times. In contrast, PDR and the proposed PACDR are DPL-based approaches, which require considerably less computational time in both training and testing. Thus, they demonstrate higher efficiency in both model learning and online monitoring.

4.5. Parameter sensitivity study

In the proposed PACDR method, six parameters are involved: α, β, λ, δ, σ, and m. The first five balancing parameters are selected from the range $\{10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}\}$, while the last dictionary size is chosen from the range $\{36, 60, 84, 108, 132\}$. To analyze the impact of different parameter values on the performance of PACDR, we examine the sensitivity of each parameter by varying one parameter at a time while keeping the others fixed. The performance of open-mode identification is evaluated using ACC, and the corresponding results are shown in Fig. 11. Similarly, the FDR and FAR results for anomaly detection are presented in Fig. 12. As illustrated in Figs. 11 and 12, PACDR is generally insensitive to most choices of the balancing parameters, especially with relatively small variations in FDR. When the open-set parameters α and β fall within $[10^{-3}, 10^{0}]$, ACC remains stable and achieves favorable performance. In addition, the performance of PACDR improves as the

Table 8
Experimental results for Case 1 of AEP.

Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	53.67	30.25	43.00	57.33	68.00	2.5334	7.1908
TDL	97.67	0.00	94.75	93.67	94.75	115.0867	0.0872
PDR	100.00	1.25	53.50	44.00	54.50	4.2425	0.0178
DSCDA	60.33	11.25	84.75	92.33	94.25	37.2489	0.0698
DFEDTL	61.67	50.75	49.25	95.00	71.25	2.5810	2.6170
PACDR	**100.00**	**0.00**	**96.50**	**95.33**	**96.50**	**2.1677**	**0.0013**

Table 9
Experimental results for Case 2 of AEP.

Method	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	Training time (s)	Testing time (s)
SROSR	97.00	50.75	21.25	30.67	23.00	4.7819	10.4600
TDL	97.33	0.00	50.00	33.33	50.00	123.3181	0.0976
PDR	85.00	36.00	49.25	95.67	71.75	4.2195	0.0144
DSCDA	93.67	50.00	0.00	33.33	50.00	34.7297	0.0272
DFEDTL	96.67	0.00	43.50	33.33	43.50	2.9210	3.5366
PACDR	**97.67**	**0.50**	**93.25**	**93.00**	**93.75**	**1.6428**	**0.0010**

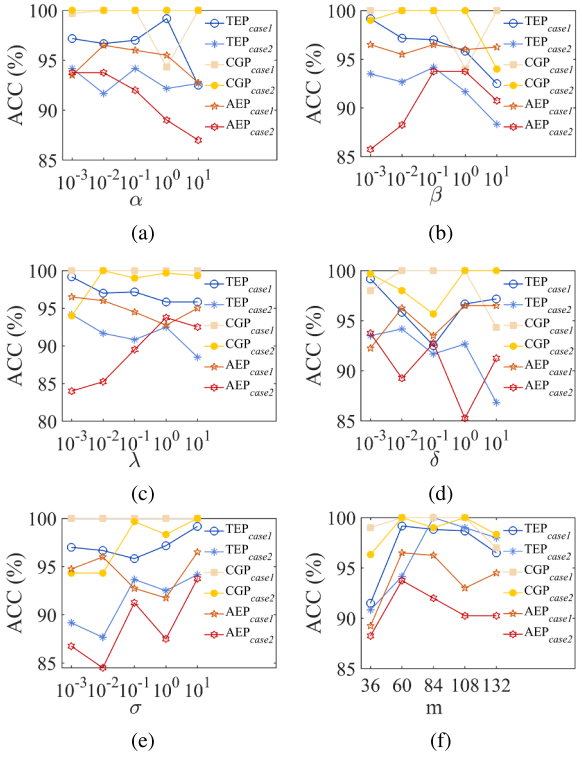


Fig. 11. ACCs (%) with PACDR under different values of (a) α, (b) β, (c) λ, (d) δ, (e) σ, and (f) m.

dictionary size increases. However, when the number of atoms becomes excessively large, the performance gain slows down and may even become limited, which is attributable to the adverse impact of oversized dictionaries on optimization stability.

Furthermore, a parameter-selection scheme that combines a heuristic rule with an unsupervised criterion is proposed to facilitate real-world deployment. Specifically, FAR on the normal target data segments is utilized as a constraint, and the set of feasible parameter values must satisfy the condition FAR$\leq \alpha_f$. Subsequently, the mean reconstruction error RE of the normal target samples is employed as the unsupervised metric. From the candidate set satisfying the FAR constraint, the parameter combination that minimizes RE is selected for final deployment. To validate the performance of this selection strategy, we define $\bar{E} = RE/\sum RE$ as the ratio of RE to better visualize and compare the error magnitudes. $\bar{E}s$ corresponding to different parameter settings is illustrated in Fig. 13. With $\alpha_f = 5\%$, the bars satisfying FAR$\leq \alpha_f$ are shown in solid colors, while the bar corresponding to the minimum $\bar{E}$ is highlighted with a red box. Based on the results in Figs. 11 and 12, the proposed parameter-selection scheme is verified to perform well in process monitoring.

4.6. Ablation discussion

To investigate the influence of the core components of PACDR on monitoring performance, an ablation study is conducted. In this study, the positive term of the open-set difference, the negative term of the open-set difference, the intra- and inter-domain graph regularization, the regularization constraint, and the progressive anchoring strategy are individually removed, corresponding to $\alpha = 0$, $\beta = 0$, $\lambda = 0$, $\delta = 0$, and anchor $= 0$, respectively. The resulting performance is summarized in Table 10. The ablation results show that setting $\alpha = 0$ or $\beta = 0$ leads to a direct decrease in OM and ACC, indicating that incorporating the open-set difference is essential for effectively learning both the target unknown modes and the source known modes. When $\lambda = 0$, the accuracy of open-mode identification drops significantly, because the intra- and inter-domain graph regularization plays a crucial role in enhancing geometric distribution consistency and improving the discriminability of the model. The refined graph structure clarifies the decision boundaries between known and unknown modes, mitigating the side effects introduced by the positive term of the open-set difference on the known target modes. The results obtained with $\delta = 0$ demonstrate that the regularization constraints are necessary for improving the robustness of the proposed method to noise and ensuring optimization stability. Furthermore, when anchor $= 0$, the performance of PACDR deteriorates in both anomaly detection and open-mode identification, which indicates that the progressive anchoring strategy helps guarantee the reliability of target pseudo-labels. Overall, the components of PACDR function synergistically, and each contributes positively to enhancing cross-domain open-mode monitoring performance.

Furthermore, to validate the advantage of the proposed inter-domain graph regularization, we conduct additional experiments on Case 2 of TEP and Case 2 of AEP. Euclidean and cosine weighting schemes are adopted as baselines. In Eq. (15), the inter-domain graph weights are computed as $\sigma_w^{ij} = \exp\left(-\left\|x_i - x_j\right\|_2^2/2\omega^2\right)$ and $\sigma_w^{ij} = \cos\left(x_i, x_j\right)$, respectively. For the Euclidean weighting, ω^2 is set by the median of the

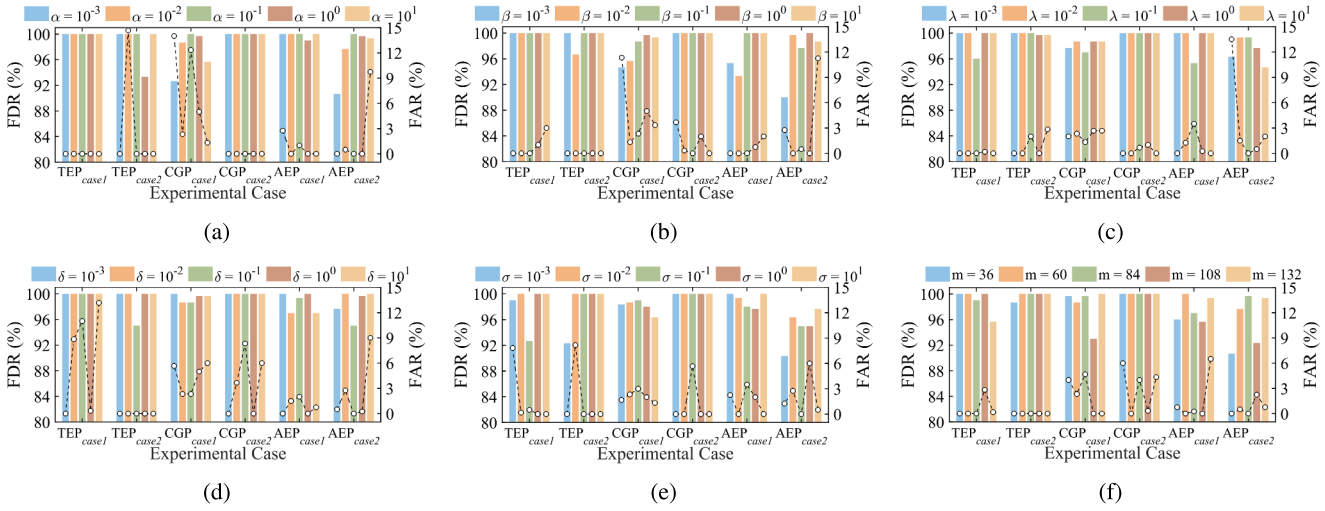


Fig. 12. FDRs (%) and FARs (%) with PACDR under different values of (a) α, (b) β, (c) λ, (d) δ, (e) σ, and (f) m.

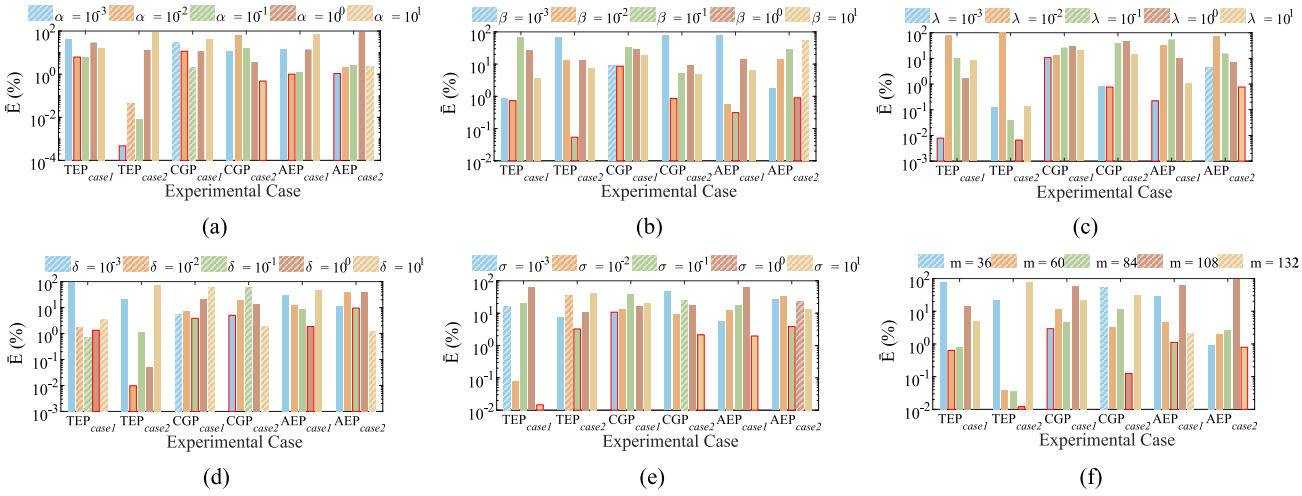


Fig. 13. $\bar{E}s$ (%) with PACDR under different values of (a) α, (b) β, (c) λ, (d) δ, (e) σ, and (f) m.

Table 10
Ablation study results of PACDR.

$\alpha=0$	$\beta=0$	$\lambda=0$	$\delta=0$	anchor = 0	TEP_{case1}					TEP_{case2}				
					FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)
✓	–	–	–	–	87.33	0.00	73.50	91.33	73.50	99.67	11.17	61.00	86.60	72.17
–	✓	–	–	–	100.00	8.33	68.83	54.33	77.17	99.00	0.00	79.00	74.80	79.00
–	–	✓	–	–	100.00	1.00	69.17	40.33	70.17	90.00	7.33	59.50	57.60	64.67
–	–	–	✓	–	100.00	0.00	94.50	91.67	94.50	100.00	13.00	82.83	92.20	93.50
–	–	–	–	✓	92.33	47.17	47.00	87.00	93.50	100.00	2.17	76.17	91.00	76.17
–	–	–	–	–	100.00	0.00	99.17	99.67	99.17	100.00	0.00	94.17	99.00	94.17
$\alpha=0$	$\beta=0$	$\lambda=0$	$\delta=0$	anchor = 0	CGP_{case1}					CGP_{case2}				
					FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)
✓	–	–	–	–	98.33	2.33	71.00	72.00	72.33	100.00	4.33	70.00	60.00	73.33
–	✓	–	–	–	63.00	17.67	58.00	61.00	72.33	100.00	4.67	68.00	61.00	72.33
–	–	✓	–	–	57.67	34.33	40.00	61.50	74.33	100.00	13.00	53.67	50.00	66.67
–	–	–	✓	–	80.00	5.67	74.67	86.00	78.67	100.00	7.00	63.67	88.00	69.33
–	–	–	–	✓	85.33	14.33	50.00	48.50	56.67	91.00	33.33	33.33	50.00	66.67
–	–	–	–	–	98.67	2.33	97.67	100.00	100.00	100.00	0.00	100.00	100.00	100.00
$\alpha=0$	$\beta=0$	$\lambda=0$	$\delta=0$	anchor = 0	AEP_{case1}					AEP_{case2}				
					FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)
✓	–	–	–	–	91.33	1.25	65.50	84.00	66.50	96.00	22.00	45.50	79.00	59.25
–	✓	–	–	–	74.00	29.00	37.50	49.67	62.00	80.33	10.00	56.25	56.33	66.25
–	–	✓	–	–	79.67	9.00	57.25	52.33	63.50	53.33	9.75	64.25	56.33	74.00
–	–	–	✓	–	92.33	5.75	79.50	79.67	84.75	50.67	43.25	45.00	66.67	74.75
–	–	–	–	✓	80.00	25.00	61.25	81.67	86.25	86.00	24.50	63.25	66.33	74.50
–	–	–	–	–	100.00	0.00	96.50	95.33	96.50	97.67	0.50	93.25	93.00	93.75

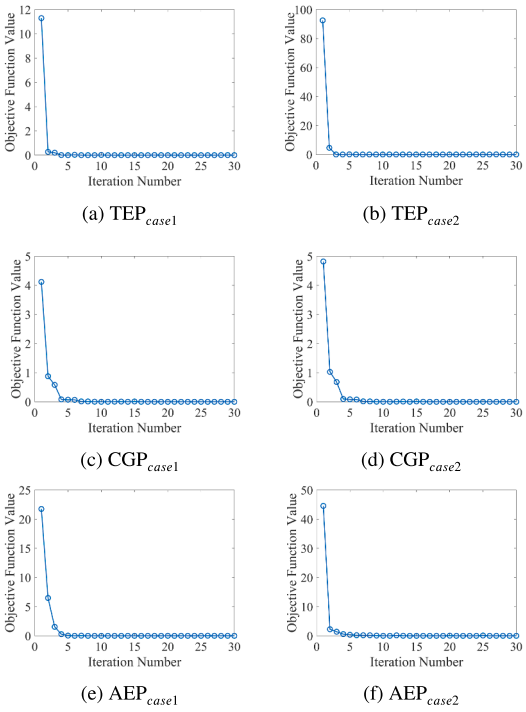


Fig. 14. Convergence curves for the objective function value of PACDR on different industrial processes.

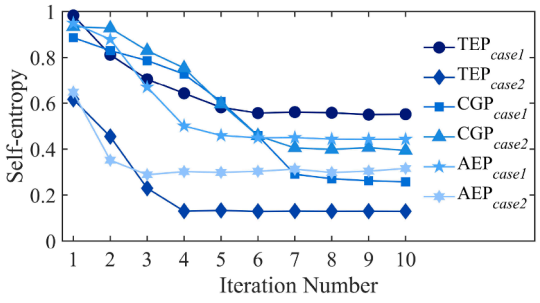


Fig. 15. Adaptive self-entropy threshold versus iteration.

Table 11
Ablation study results on inter-domain edge weighting.

Case	Weighting scheme	FDR (%)	FAR (%)	MI (%)	OM (%)	ACC (%)
TEP$_{case2}$	Euclidean	93.67	0.00	74.83	89.80	74.83
	Cosine	100.00	9.33	76.00	90.80	83.00
	Proposed	100.00	0.00	94.17	99.00	94.17
AEP$_{case2}$	Euclidean	84.67	0.75	79.75	91.67	80.25
	Cosine	92.33	48.75	51.00	66.67	74.50
	Proposed	97.67	0.50	93.25	93.00	93.75

squared pairwise Euclidean distances. The comparative results in Table 11 show that our method achieves better open-mode process monitoring performance. In particular, under Case 2 of AEP with significant domain shift, our ACC improves by 13.5% and 19.25% compared with the Euclidean and cosine weighting, respectively. These results verify that the proposed weighting based on SVM decision-boundary distance is more reliable in promoting the cross-domain consistency of representation coefficients in multi-mode manifold structures.

4.7. Convergence analysis

To verify the convergence behavior of the proposed PACDR method, we conduct experiments on TEP, CGP, and AEP. Following the division of training and testing sets reported in Table 1, we randomly collect samples to investigate how the PACDR objective value evolves with the number of iterations. The convergence curves of PACDR over 30 iterations for all six experiments are shown in Fig. 14. As can be observed, the objective value in the TEP and AEP experiments decreases monotonically with iterations and converges rapidly within five iterations. In the CGP experiments, the objective value drops to a small, stable level within ten iterations.

4.8. Visualization

To analyze the evolution of self-entropy during the iterative process, the adaptive threshold and the self-entropy distributions of both known and unknown modes are visualized in Figs. 15 and 16. The adaptive threshold versus iteration for all cases in the TEP, CGP, and AEP experiments is plotted in Fig. 15. The curves show that the threshold decreases rapidly during the first few iterations and then exhibits only minor fluctuations, stabilizing after several iterations. Moreover, the t-SNE visualization (Deng et al., 2025b) of the cross-device case in AEP is depicted in Fig. 17. While a significant shift exists between the source and target domains, PACDR effectively improves domain alignment. Hence, the threshold behavior suggests that under our domain alignment mechanism, the anchoring criteria rapidly converge to a reliable selection rule. Besides, Fig. 16 compares the self-entropy distributions of known and unknown modes at iteration 1 and iteration 10. As the iterations proceed, the self-entropy distribution of the known modes shifts left and becomes narrower in TEP and AEP, while the unknown-mode self-entropy distribution remains concentrated on the right-hand side. Meanwhile, the adaptive threshold lies near the right tail of the known-mode distribution. The results indicate that, after reliably anchoring known modes, the adaptive threshold can effectively anchor high-confidence unknown-mode instances. Thus, it enables a clear separation between known modes and high-uncertainty unknown modes.

To further examine whether noticeable model drift occurs in early iterations, the evolution of ACC is illustrated in Fig. 18. The results show a rapid increase in ACC within the first three iterations, with only minor transient drops, before stabilizing after iteration 6. The lack of sustained degradation or large oscillations in early ACC confirms that the model avoids noticeable drift even under significant domain shifts, exhibiting excellent self-correction and stable convergence behavior.

5. Conclusions

To address the issues of distribution mismatch between training and testing process data and the emergence of unknown modes, a progressive anchoring-driven consistent dual reconstruction method is proposed for open-mode process monitoring. An open-mode dual reconstruction framework is developed to enable domain-knowledge transfer and information exchange of known modes via the dictionary pair, while simultaneously mitigating the negative transfer caused by unknown modes. To further achieve domain adaptation, an intra- and inter-domain graph regularization is proposed to align geometric distributions and enhance the discriminability of the coding coefficients. In addition, a progressive target anchoring strategy is designed to filter out unreliable pseudo-labels, effectively acquiring target semantic information and updating the classifier parameter matrix. Finally, extensive experiments on both simulated and real industrial processes verify the superior performance of the proposed method. Future research will consider multi-source domain adaptation to exploit the complementary information from multiple source domains, aiming to achieve stronger cross-domain generalization and robustness.

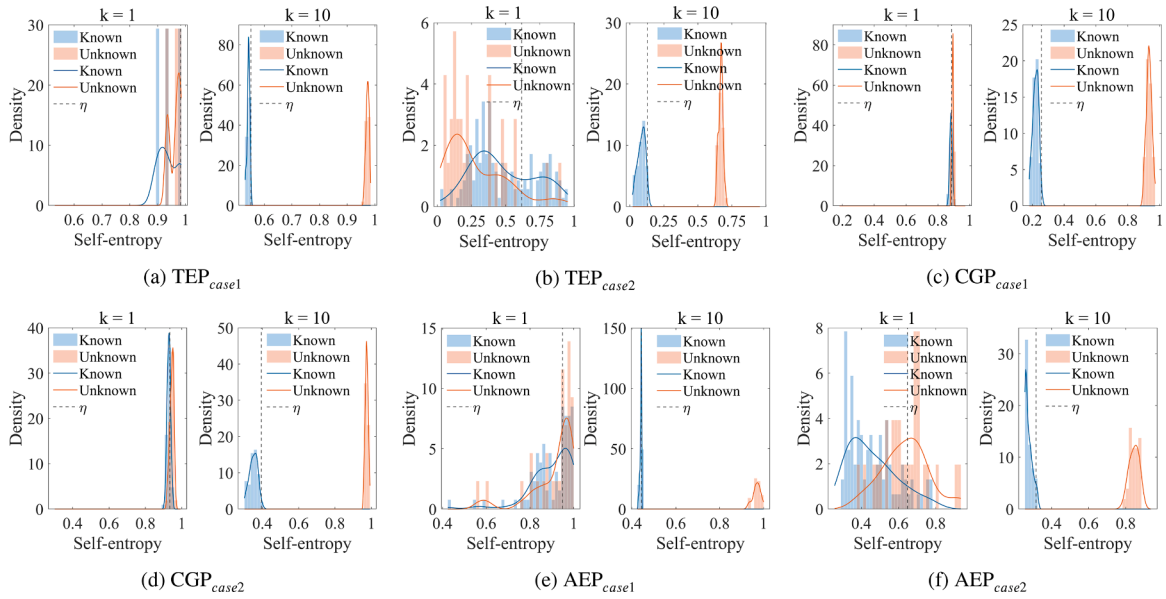


Fig. 16. Self-entropy distributions of known and unknown modes at iterations 1 and 10.

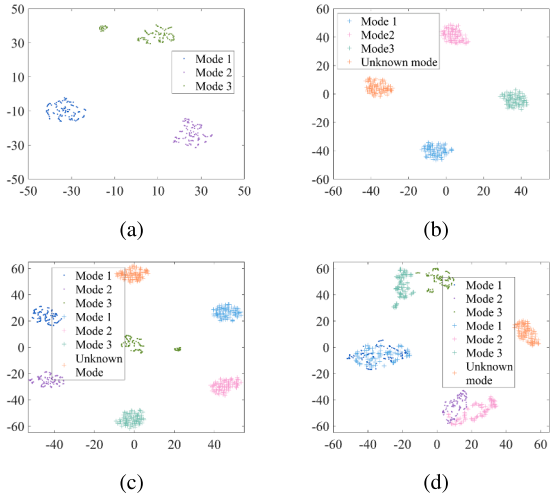


Fig. 17. t-SNE visualization of the cross-device case in AEP. (a) Original source domain. (b) Original target domain. (c) Original source and target domains. (d) Aligned Source and target domains by PACDR.

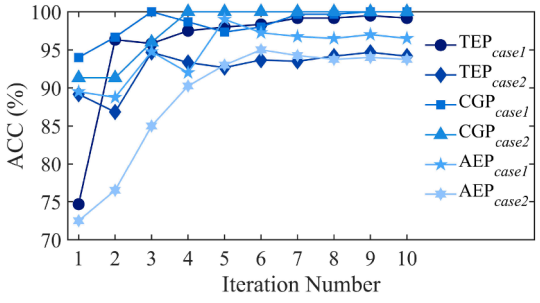


Fig. 18. ACCs (%) versus iteration.

CRediT authorship contribution statement

Ziqing Deng: Data curation, Funding acquisition, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing; **Xiaofang Chen:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing - original draft; **Lipo Wang:** Supervision, Validation, Writing - original draft; **Lihui Cen:** Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision; **Yalin Wang:** Conceptualization, Funding acquisition, Project administration, Supervision; **Shouli Yu:** Data curation, Resources, Validation.

Data availability

The data that has been used is confidential.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported in part by the National Key R&D Program of China (Grant No. 2024YFB4105200), in part by the National Natural Science Foundation of China (Grant Nos. 62473384 and 62133016), in part by the Natural Science Foundation of Hunan Province (Grant No. 2025JJ10007), in part by the Fundamental Research Funds for Central Universities of Central South University (Grant No. 2025ZZTS0614), and in part by the China Scholarship Council (Grant No. 202406370121).

References

Chen, Y., Fang, X., Liu, Y., Zheng, W., Kang, P., Han, N., & Xie, S. (2024a). Two-step strategy for domain adaptation retrieval. *IEEE Transactions on Knowledge and Data Engineering*, *36*(2), 897–912.

Chen, Y., Ma, H., Wang, Y., & Liu, X. (2026). Key performance indicator-related process monitoring for irregular scenarios with incomplete data. *Expert Systems with Applications*, *298*, 129440.

Chen, Z., Huang, K., Wu, D., Yang, C., & Gui, W. (2024b). Global information-based lifelong dictionary learning for multimode process monitoring. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *54*(12), 7182–7194.

Chen, Z., Huang, K., Wu, D., Yang, C., & Gui, W. (2024c). Global information-based lifelong dictionary learning for multimode process monitoring. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *54*(12), 7182–7194.

Deng, Z., Chen, X., Cen, L., Xie, Y., & Zou, Z. (2025a). Zero-shot fault diagnosis in industrial processes using graph-regularized coupled dictionary learning. *IEEE Transactions on Industrial Informatics*, *21*(11), 8562–8573.

Deng, Z., Chen, X., Xie, S., Xie, Y., & Sun, Y. (2021). Distributed process monitoring based on joint mutual information and projective dictionary pair learning. *Journal of Process Control*, *106*, 130–141.

Deng, Z., Chen, X., Xie, S., Xie, Y., & Zhang, H. (2023a). Semi-supervised discriminative projective dictionary pair learning and its application to industrial process. *IEEE Transactions on Industrial Informatics*, *19*(3), 3119–3132.

Deng, Z., Chen, X., Xie, Y., Zhang, H., & Gui, W. (2025b). Generalized cross-domain industrial process monitoring via adaptive discriminative transfer dictionary pair learning with attribute embedding. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(9), 17406–17420.

Deng, Z., Chen, X., Xie, Y., Zou, Z., & Zhang, H. (2023b). Multiple structured latent double dictionary pair learning for cross-domain industrial process monitoring. *Information Sciences*, *648*, 119514.

Dong, J., Li, D., Wei, Y., & Peng, K. (2025). A novel process monitoring and root cause diagnosis strategy based on knowledge-data-integrated causal digraph for complex industrial processes. *IEEE Transactions on Instrumentation and Measurement*, *74*, 1–12.

Du, Z., Liu, D., & Cui, L. (2025). Dynamic model-driven dictionary learning-inspired domain adaptation strategy for cross-domain bearing fault diagnosis. *Reliability Engineering & System Safety*, *258*, 110905.

Fan, Z., Shi, L., Liu, Q., Li, Z., & Zhang, Z. (2023). Discriminative fisher embedding dictionary transfer learning for object recognition. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(1), 64–78.

Fang, X., Jiang, L., Han, N., Sun, W., Xu, Y., & Xie, S. (2022). Cross-domain recognition via projective cross-reconstruction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *52*(12), 7366–7377.

Fang, Z., Lu, J., Liu, F., Xuan, J., & Zhang, G. (2021). Open set domain adaptation: Theoretical bound and algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(10), 4309–4322.

Fu, Y., Ding, J., & Xu, X. (2024a). Low-rank multimanifold embedding learning for multimode process monitoring. *IEEE Transactions on Industrial Informatics*, *20*(3), 3468–3477.

Fu, Y., Luo, C., Xu, X., Song, L., & Xia, C. (2024b). Graph embedding dictionary pair learning for robust process monitoring. *Measurement*, *228*, 114287.

Fu, Y., Xia, F., Ding, J., & Xu, X. (2025). Manifold-aware and sparsity discriminant collaborative representation model and its application to multimode process monitoring. *Knowledge-Based Systems*, *324*, 113869.

Han, N., Wu, J., Fang, X., Teng, S., Zhou, G., Xie, S., & Li, X. (2020). Projective double reconstructions based dictionary learning algorithm for cross-domain recognition. *IEEE Transactions on Image Processing*, *29*, 9220–9233.

Huang, J., Sun, X., Ding, S. X., Yang, X., & Ersoy, O. K. (2025). Variational discriminative stacked auto-encoder: Feature representation using a prelearned discriminator, and its application to industrial process monitoring. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(5), 9383–9394.

Huang, K., Wen, H., Zhou, C., Yang, C., & Gui, W. (2020a). Transfer dictionary learning method for cross-domain multimode process monitoring and fault isolation. *IEEE Transactions on Instrumentation and Measurement*, *69*(11), 8713–8724.

Huang, K., Wu, Y., Wen, H., Liu, Y., Yang, C., & Gui, W. (2020b). Distributed dictionary learning for high-dimensional process monitoring. *Control Engineering Practice*, *98*, 104386.

Li, J., Lu, K., Huang, Z., Zhu, L., & Shen, H. T. (2019). Heterogeneous domain adaptation through progressive alignment. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(5), 1381–1391.

Li, X., Chen, H., Li, S., Wei, D., Zou, X., Si, L., & Shao, H. (2025). Multi-kernel weighted joint domain adaptation network for cross-condition fault diagnosis of rolling bearings. *Reliability Engineering & System Safety*, *261*, 111109.

Liu, J. (2025). Bearing fault diagnosis based on incremental transfer learning and ensemble learning with stochastic configuration network. *Expert Systems with Applications*, *290*, 128398.

Liu, Y., Zeng, J., Jiang, B., Sheng, W., Wang, Z., Xie, L., & Li, L. (2024). Structured collaborative sparse dictionary learning for monitoring of multimode processes. *Information Sciences*, *666*, 120444.

Lyu, Y., Zhou, L., Cong, Y., Zheng, H., & Song, Z. (2024). Multirate mixture probability principal component analysis for process monitoring in multimode processes. *IEEE Transactions on Automation Science and Engineering*, *21*(2), 2027–2038.

Ngo, B. H., Choi, T. J., & Cho, S. I. (2026). Towards enhancing prototypes driven by graph convolutional network for domain adaptation. *Expert Systems with Applications*, *299*, 130010.

Tan, S., Wang, Y., Shi, H., Song, B., & Tao, Y. (2023). A topology model based on common and specific feature separation for multimode process monitoring. *Journal of Process Control*, *130*, 103052.

Tang, W., Panahi, A., Krim, H., & Dai, L. (2019). Analysis dictionary learning based classification: Structure for robustness. *IEEE Transactions on Image Processing*, *28*(12), 6035–6046.

Wang, C., Shu, Z., Yang, J., Zhao, Z., Jie, H., Chang, Y., Jiang, S., & See, K. Y. (2025a). Learning to imbalanced open set generalize: A meta-learning framework for enhanced mechanical diagnosis. *IEEE Transactions on Cybernetics*, *55*(3), 1464–1475.

Wang, Y., Cen, C., Cao, H., Li, Z., & Mu, H. (2025b). Guided representation learning with dictionary-based fuzzy sparse discriminative embedding. *Expert Systems with Applications*, *287*, 128112.

Warke, V., Kumar, S., Bongale, A., & Kotecha, K. (2024). Robust tool wear prediction using multi-sensor fusion and time-domain features for the milling process using instance-based domain adaptation. *Knowledge-Based Systems*, *288*, 111454.

Wei, H., Ma, L., Liu, Y., & Du, Q. (2021). Combining multiple classifiers for domain adaptation of remote sensing image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 1832–1847.

Wu, W., Song, C., Zhao, J., Xu, Z., & Yu, W. (2024). Hybrid-order graph embedded distributed encoder-decoder for multiunit industrial plant-wide process monitoring. *IEEE Transactions on Automation Science and Engineering*, *21*(4), 7298–7311.

Xu, C., Vu, G., Cao, B. T., Liu, Z., Diewald, F., Yuan, Y., & Meschke, G. (2026). Bridging simulation and experiment: A self-supervised domain adaptation framework for concrete damage classification. *Advanced Engineering Informatics*, *69*, 104125.

Xu, X., Ding, J., Liu, Q., & Chai, T. (2021). A novel multimanifold joint projections model for multimode process monitoring. *IEEE Transactions on Industrial Informatics*, *17*(9), 5961–5970.

Yang, C., Liang, H., Huang, K., Li, Y., & Gui, W. (2021). A robust transfer dictionary learning algorithm for industrial process monitoring. *Engineering*, *7*(9), 1262–1273.

Yin, J., Yu, J., Jiang, Q., & Yan, X. (2025). Adaptive temporal diffusion-based reconstruction model for industrial dynamic uncertain process monitoring. *Applied Soft Computing*, *181*, 113407.

Zhang, H., & Patel, V. M. (2017). Sparse representation-based open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(8), 1690–1696.

Zhang, J., Zhou, D., & Chen, M. (2023). Self-learning sparse PCA for multimode process monitoring. *IEEE Transactions on Industrial Informatics*, *19*(1), 29–39.

Zhang, Z., Sun, Y., Wang, Y., Zhang, Z., Zhang, H., Liu, G., & Wang, M. (2021). Twin-incoherent self-expressive locality-adaptive latent dictionary pair learning for classification. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(3), 947–961.

Zhao, Y., Yang, X., Huang, J., Zhou, X., Cui, J., & Li, Q. (2026). A data structure-preserving semi-supervised method for rotating machinery fault diagnosis under low labeled rates. *Advanced Engineering Informatics*, *69*, 103890.

Zheng, J., Liu, X., Yang, Z., & Ge, Z. (2025a). Bagging deep residual principal component analysis for ensemble industrial process monitoring. *Process Safety and Environmental Protection*, *203*, 107995.

Zheng, Z., Teng, S., Teng, L., Zhang, W., & Wu, N. (2025b). Adaptive graph learning with semantic promotability for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *47*(3), 1747–1763.

Zhu, J., Ye, F., Xiao, Q., Guo, P., Zhang, Y., & Yang, Q. (2024). A versatile framework for unsupervised domain adaptation based on instance weighting. *IEEE Transactions on Image Processing*, *33*, 6633–6646.