

# Fault Detection and Diagnosis using Reconstruction-based DiGLPP: Application to Industrial Distillation System

Husnain Ali\*, Furong Gao\*<sup>†</sup>

*\*Department of Chemical and Biological Engineering, Hong Kong University of Science and Technology, Hong Kong*

*<sup>†</sup>Guangzhou HKUST Fok Ying Tung Research Institute, Guangzhou 511458, China  
(e-mail: \*haliaa@connect.ust.hk, \*\*kefgao@ust.hk).*

---

**Abstract:** The rapid advancement of Industry 4.0, artificial intelligence, and big data sensor technologies has made industrial systems highly complex and dynamic. Classical fault detection and diagnosis (FDD) techniques depend on insufficient information and variables with equivalent uncertainty. This paper introduced an advanced dynamic inner reconstruction-based contribution with global-local preservation projection (DiGLPP-RBC) for fault detection and diagnosis. Firstly, inner data statistics are extracted to develop an augmented matrix, which is used to characterize the dynamic latent variable using the DiGLPP framework. Secondly, reconstruction-based contribution (RBC) is used to determine fault contribution. The proposed method employs Hotelling's  $T^2$  and squared prediction error ( $SPE$ ) to detect and diagnose variable contributions and kernel density of faults in the ethanol-water industrial distillation system. The proposed framework's robustness is compared with traditional baseline frameworks such as dynamic inner principal component analysis (DiPCA) and bi-directional long short-term memory-autoencoder (BiLSTM-AE). The results indicate that the DiGLPP-RBC technique detects, identifies, and diagnoses irregularities and faults more effectively and reliably than traditional approaches.

**Keywords:** Industry 4.0, artificial intelligence, fault detection and diagnosis, variable contributions, DiGLPP-RBC, industrial distillation system.

---

## 1. INTRODUCTION

Industrial process plants are intrinsically large-scale and highly interrelated systems with multiple process parameters. Early fault detection and detailed variability diagnosis constitute the first phase of risk assessment, allowing operators to prevent a malfunction from escalating into a significant hazard (Ali, Maulud, Zabiri, Nawaz, Suleman, et al., 2022). Multiple catastrophic events in the past have caused human suffering, environmental degradation, and mortality, increasing the focus on process safety (Ali, Safdar, Zhou, et al., 2024). The fault detection and diagnosis (FDD) method could be enhanced by smart automation, which reduces human intervention and risk, facilitating the transition to Industry 4.0. The use of artificial intelligence (AI) based automation is also important in fault detection and diagnosis (FDD) (Silva et al., 2020). Consequently, fault detection and diagnosis (FDD) are crucial for developing the safety and risk estimation system (Ali et al., 2025).

Fault detection and diagnosis (FDD) methodologies are primarily categorized into two basic types: knowledge and data-driven models (Ali, 2022; Ali, Safdar, Rasool, et al., 2024). Multiple statistical models have been developed, devised, and widely used in system fault detection and diagnosis, including principal component analysis (PCA) (Ali, Zhang, & Gao, 2023), fisher discriminant analysis (FDA), canonical correlation analysis (CCA) (Ali & Gao, 2023), and independent component analysis (ICA) (Ali, Zhang, et al., 2024). In the past few decades, several preservation techniques have achieved significant results in detection and diagnosis.

Nonetheless, their practical implementations face several challenges (Ali, Maulud, Zabiri, Nawaz, & Ismail, 2022). These approaches often need substantial quantities of data with labels for development (Ding, Ali, Gao, Zhang, & Gao, 2025). Acquiring extensive labeled information within the fault diagnosis domain is still challenging. Zhang introduced a fault-detection approach for global-local structural evaluation by formulating a dual-objective function that integrates the attributes of PCA and local preserving projection (LPP) (Zhang, Ge, Song, & Fu, 2011). Yu introduced local and global PCA (Yu, 2012). A dual-objective variable was formulated using LPP and PCA, employing the ratio of variables for concurrently identifying global and local characteristics. Luo introduced an algorithm named global-local preserving projection (GLPP) (Luo, 2014). This approach effectively integrates global and local elements inside a single framework. Luo introduced a combined structure-preserving projection technique to identify abnormalities (Luo, Bao, Mao, & Tang, 2016). These approaches could reveal latent intrinsic traits from highly dimensional information while preserving their local structural characteristics. The reconstruction approach ensures accurate variable diagnosis when the fault direction is identified and included in the possible locations (Dunia & Joe Qin, 1998). However, it cannot provide diagnostic findings for failures with undetermined directions. The contribution graphs may contain faults that might result in misleading and wrong diagnoses. (Yoon & MacGregor, 2001) demonstrate that conventional contribution charts do not yield accurate diagnostic outcomes.

This paper introduced an advanced dynamic inner reconstruction-based contribution with global-local preservation projection (DiGLPP-RBC) for fault detection and diagnosis. The proposed method employs Hotelling's  $T^2$  and squared prediction error ( $SPE$ ) to detect and diagnose variable contributions and kernel density of faults in the ethanol-water industrial distillation benchmark. The proposed framework's robustness is compared with traditional baselines such as dynamic inner principal component analysis (DiPCA) and bi-directional long short-term memory-autoencoder (BiLSTM-AE).

The following portions of this paper are ordered as outlined below. Section 2 outlines the dynamic inner global-local preservation projection (DiGLPP), reconstruction-based contribution (RBC), and the proposed methodological framework (DiGLPP-RBC). Section 3 presents a performance benchmark of the ethanol-water industrial distillation system. Section 4 analyzes, compares, and evaluates the effectiveness of the intended and existing baselines. The conclusion is elaborated in Section 5.

## 2. METHODOLOGY

### 2.1 Dynamic inner global-local preservation projection

The DiGLPP approach is a dynamic approximation of the Laplacian characteristic map used for feature extraction from data while maintaining both local and global manifold knowledge associated with the dataset. For the original dataset  $X$ , the transform matrix  $W$  is computed and projected onto a space with a low dimension to get the data set  $Y$  after reducing its dimensionality (Luo, 2014).

$$X_k(l) = [X(k) X(k-1) \cdots X(k-l)] \quad (1)$$

where,  $X(k-1)$  denotes the  $m$ -dimensional variables at period  $k-l$ , while  $l$  is the duration interval.

The objective functionality of DiGLPP is outlined below:

$$J_{DiGLPP}(\bar{a}) = \min_{\bar{a}} \frac{1}{2} \left\{ \begin{array}{l} \underbrace{\eta \sum_{ij} (y_i - y_j)^2 W_{ij}}_{\text{Local structure preservation}} \\ -(1-\eta) \underbrace{\sum_{ij} (y_i - y_j)^2 \bar{W}_{ij}}_{\text{Global structure preservation}} \end{array} \right\} \quad (2)$$

The converse  $y_i = \bar{a}^T \bar{x}_i \in \mathfrak{R}^m$  of a transformation information matrix in  $A$  is indicated  $y_i = a^T \bar{x}_i$ , while the  $\eta \in [0, 1]$  array coefficient regulates the balance between global and local structure preservation.  $W_{ij}$  and  $\bar{W}_{ij}$  signified the weighted array coefficient reflecting the adjacent and non-adjacent relationships to the relevant elements  $\bar{x}_i$  and  $\bar{x}_j$ , respectively.

$$W_{ij} = \begin{cases} e^{-\frac{\|\bar{x}_i - \bar{x}_j\|^2}{\sigma^2}} & \text{if } \bar{x}_j \in \Omega_k(\bar{x}_i) \ \& \ \bar{x}_i \in \Omega_k(\bar{x}_j) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$\bar{W}_{ij} = \begin{cases} e^{-\frac{\|\bar{x}_i - \bar{x}_j\|^2}{\sigma^2}} & \text{if } \bar{x}_j \notin \Omega_k(\bar{x}_i) \ \& \ \bar{x}_i \notin \Omega_k(\bar{x}_j) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The subset  $\Omega_k(\bar{x})$  is defined by  $k$ -adjacent that involves  $k(x)$ , where  $\sigma_1$  and  $\sigma_2$  are the variables factors. Equation 2 may be calculated as follows (Luo, 2014):

$$J_{DiGLPP}(\bar{a}) = \min_{\bar{a}} \frac{1}{2} \left\{ \begin{array}{l} \underbrace{\eta \sum_{ij} (y_i - y_j)^2 W_{ij}}_{\text{Local structure preservation}} \\ -(1-\eta) \underbrace{\sum_{ij} (y_i - y_j)^2 \bar{W}_{ij}}_{\text{Global structure preservation}} \end{array} \right\} \\ = \min_{\bar{a}} \frac{1}{2} \sum_{ij} (y_i - y_j)^2 R_{ij} \quad (5)$$

$$= \min_{\bar{a}} \left\{ \sum_i y_i H_{ii} y_i^T - \sum_{ij} y_i H_{ij} y_j^T \right\}$$

$$= \min_{\bar{a}} \left\{ \sum_i \bar{a}^T \bar{x}_i H_{ii} \bar{x}_i^T \bar{a} - \sum_{ij} \bar{a}^T \bar{x}_i H_{ij} \bar{x}_j^T \bar{a} \right\}$$

$$= \min_{\bar{a}} \bar{a}^T X(H - R)X^T \bar{a}$$

$$= \min_{\bar{a}} \bar{a}^T XMX^T \bar{a}$$

where,  $H$  indicates the diagonal information

$$H_{ii} \sum_j R_{ij}, R_{ij} = \eta W_{ij} - (1-\eta) \bar{W}_{ij}, \text{ and } M = H - R \text{ is the}$$

Laplacian information array (Luo, 2014).

$$\bar{a}^T [\eta XHX^T + (1-\eta)I_m] \bar{a} = \bar{a}^T N \bar{a} = 1 \quad (6)$$

$$\eta = \frac{\rho(\hat{L})}{\rho(L) + \rho(\hat{L})} \quad (7)$$

where  $\eta \in [0, 1]$  is ascertained using the spectral radii  $\rho(L)$  and  $\rho(\hat{L})$  of the vectors  $L$  and  $\hat{L}$  that correspond to the global and local manifold structures, respectively.

The DiGLPP model is trained in solving the specified optimization challenge.

$$\min_{\bar{a}} \bar{a}^T XMX^T \bar{a} \quad (8)$$

$$s.t. \quad \bar{a}^T N \bar{a} = 1$$

$$XMX^T \bar{a} = \lambda N \bar{a} \quad (9)$$

After the development and implementation of the DiGLPP algorithm, the two statistical control charts, Hotelling's  $T^2$  and  $SPE$ , are used to assess the framework's efficacy and the variation of the residual matrix in dimension, respectively.

The  $T^2$ -monitored control score will be calculated appropriately.

$$T^2 = x^T \Lambda^{-1} x < T_{lim} \quad (10)$$

The control limit for the  $T^2$  statistic may be computed as follows:

$$T_{lim} = \frac{\alpha(m+1)(m-1)}{m(m-\alpha)} F_{\alpha}(\alpha, m-\alpha) \quad (11)$$

The Fisher distribution is  $F_{\alpha}(\alpha, m-\alpha)$  where the degrees of freedom are indicated by  $\alpha, (m+1)(m-1)$  and  $(\alpha)$  denotes the significance scale.

The  $SPE$  is calculated subsequently.

$$SPE = \|x - Ay\|^2 < SPE_{lim} \quad (12)$$

The control limit for the  $SPE$  statistic may be computed as follows:

$$SPE_{lim} = \varphi_1 \left[ \frac{h_0 c_\alpha \sqrt{2\varphi_2}}{\varphi_1} + 1 + \frac{\varphi_2 h_0 (h_0 - 1)}{\varphi_1^2} \right]^{h_0} \quad (13)$$

where,  $\varphi_i = \sum_{j=1}^n \lambda_j^i$  ( $i = 1, 2, 3, \dots, n$ ),  $h_0 = 1 - \frac{\varphi_1 \varphi_i}{3\varphi_1^2}$  and the

value  $c_\alpha$  derived from the substantial universal groupings.

A fault will be identified when  $T^2$  and (SPE) statistical control limits are above the threshold.

## 2.2 Reconstruction-based contribution

Reconstructing a diagnosis index after detecting a variable mitigates the influence of that particular parameter on the index. Reconstruction describes the degree to which a parameter contributes to the overall reconstruction of the fault index. Consequently, this specific degree of reconstruction will be designated as the reconstruction-based contribution (RBC) of this descriptive parameter to fault diagnosis.

Let us assess an operating system with  $n$  sensors. In the occurrence of an anomaly in sensor  $x_i$ , the misleading signal is indicated as  $x \in \square^n$ . Furthermore, the position of the abnormalities is denoted by  $\xi_i$  (Alcala & Qin, 2009).

$$Z_i = x - \xi_i f_i \quad (14)$$

The anomaly diagnosis index of the reconstructed value can be articulated in a typical way.

$$\text{Index}(Z_i) = Z_i^T M Z_i = \|Z_i\|_M^2 = \|x - \xi_i f_i\|_M^2 \quad (15)$$

Here,  $f_i$  denotes the anticipated magnitude of the abnormalities. The anomaly diagnosis index  $Z_i$  of the reconstructed data is computed.

The aim of reconstruction is to choose a threshold  $f_i$  that optimizes the  $\text{Index}(Z_i)$  ratio. The reduction method entails computing the initial derivatives  $\text{Index}(Z_i)$  of a variable and  $f_i$  equating them to zero.

$$\frac{d(\text{Index}(Z_i))}{df_i} = -2(x - \xi_i f_i)^T M \xi_i \quad (16)$$

Consequently, the RBC of the parameterized variable  $x_i$  about the anomaly diagnostic index is articulated as (Alcala & Qin, 2009):

$$RBC_i^{Index} = \|\xi_i f_i\|_M^2 \quad (17)$$

$$RBC_i^{Index} = \|\xi_i (\xi_i^T M \xi_i)^{-1} \xi_i^T M x\| \quad (18)$$

$$RBC_i^{Index} = x^T M \xi_i (\xi_i^T M \xi_i)^{-1} \xi_i^T M x \quad (19)$$

The reconstructed index, denoted as  $\text{Index}(Z_i)$ , is obtained by substituting an integer  $f_i$  into Equation (17) and is calculated to be.

$$\begin{aligned} \text{Index}(Z_i) &= x^T M [1 - \xi_i (\xi_i^T M \xi_i)^{-1} \xi_i^T M] x \\ &= x^T M x - x^T M \xi_i (\xi_i^T M \xi_i)^{-1} \xi_i^T M x \\ &= \text{Index}(x) - RBC_i^{Index} \end{aligned} \quad (20)$$

The formulation of the RBC-aided DiGLPP control limits entails the calculation of statistical heatmap plots  $RBC_i^{DiGLPP}$ , respectively.

$$RBC_i^{DiGLPP} = x^T M \xi_i (\xi_i^T M \xi_i)^{-1} \xi_i^T M x = \frac{(\xi_i^T M x)^2}{(\xi_i^T M \xi_i)} \quad (21)$$

## 2.3 Reconstruction-based DiGLPP methodological framework

This study presents an advanced machine learning (ML) framework for fault detection and diagnosis to tackle contemporary difficulties and improve industrial systems. The proposed algorithm integrates machine learning techniques with reconstruction-aided dynamic inner global-local preservation projection (DiGLPP-RBC). The scenario dataset has **1,000 observations** across **10 variables**, resulting in an array of **1000 X 10**. This dataset encompasses standard operations and abnormal circumstances, offering a compact, precise sample for system training and validation. The proposed framework's robustness is compared with traditional baselines such as dynamic inner principal component analysis (DiPCA) and bi-directional long short-term memory-autoencoder (BiLSTM-AE). Figure 1 and Table 1 depict the proposed advanced machine learning system and the steps for building models.

Table 1. Methodological Development Steps

Steps	DiGLPP-RBC Methodological Framework
(Standardization)	<p><b>Datasets Structuring</b></p> <ul style="list-style-type: none"> <li>The proposed approach acquires data for training and testing by implementing an industrial distillation system using AspenTech and MATLAB.</li> <li>The assessment metrics in the gathered data sets accurately reflect standard process conditions without any anomalies.</li> <li>Assessment metrics consist of aggregating data that shows unusual conditions.</li> <li>The training and testing information is normalized to possess unit variance and a zero mean.</li> </ul>
	<p><b>Dynamic Inner GLPP Model Development</b></p> <ul style="list-style-type: none"> <li>Assess training and testing information.</li> <li>Compute the time-delay of each parameter.</li> <li>Executing the sliding window to obtain a dynamic augmented vector.</li> <li>Compute dimensional reduction of augmented vector to get DiGLPP model.</li> <li>Calculate threshold control limits: <math>T^2</math> and <math>SPE</math>.</li> <li><math>DiGLPP-T^2</math> and <math>DiGLPP-SPE &lt; Threshold\ limit</math>.</li> </ul>
	<p><b>Reconstruction-based DiGLPP Model Development</b></p> <ul style="list-style-type: none"> <li>Compute the augmented vector.</li> <li>Compute the contribution index using DiGLPP variables.</li> <li>Apply kernel density estimation to training and testing information.</li> <li>Execute the variable density of each parameter based on the DiGLPP-RBC model.</li> </ul>
	<p>(Fault Diagnosis)</p>

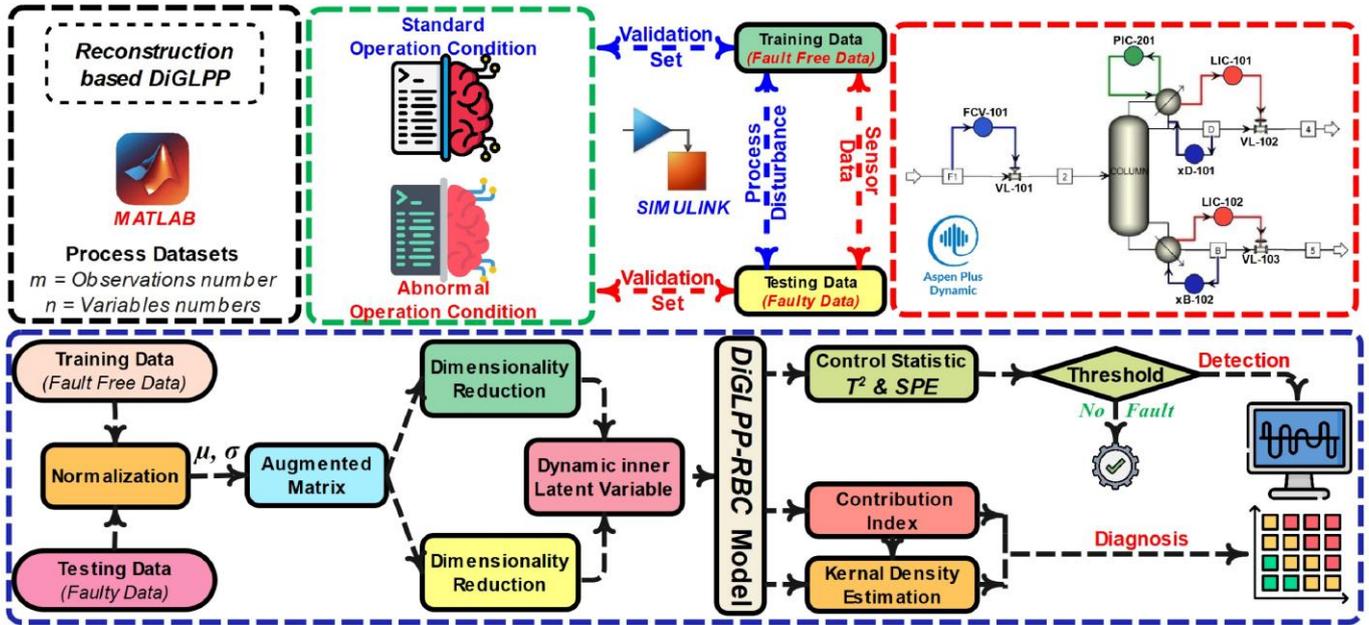


Figure 1. Proposed methodological framework

### 3. INDUSTRIAL DISTILLATION SYSTEM: CASE SCENARIO

This study employs a distillation system (DS) inside an equilibrium ethanol-water combination as a reference for imitating a particular scenario. Figure 2 presents a systematic representation of the DS ethanol-water reaction system. The simulation findings are loaded into MATLAB and Simulink to provide data on standard and abnormal circumstances. Tables 2-4 thoroughly illustrate the model parameters of anticipated attributes of fault categories, including Process Disruption and Sensor failure.

Table 2. Distillation system variables

Industrial Distillation System		
No.	Parameters	Description
1	$F_i$	Feed flow rate
2	$X_F$	Feed composition
3	$R_f$	Reflux flow rate
4	$T_i$	Feed temperature
5	$T_c$	Condenser exhaust temperature
6	$T_r$	Reboiler exhaust temperature
7	$P_i$	Feed pressure
8	$P_d$	Reflux column pressure
9	$P_c$	Condenser exhaust pressure
10	$R$	Reflux ratio

Table 3. Fault case scenarios

No.	Fault Type	Fault Description	Fault Nature
1	Ramp	Discrepancy in condenser out temperature	Process Disruption
2	Step	Instability in reflux pressure	Sensor failure

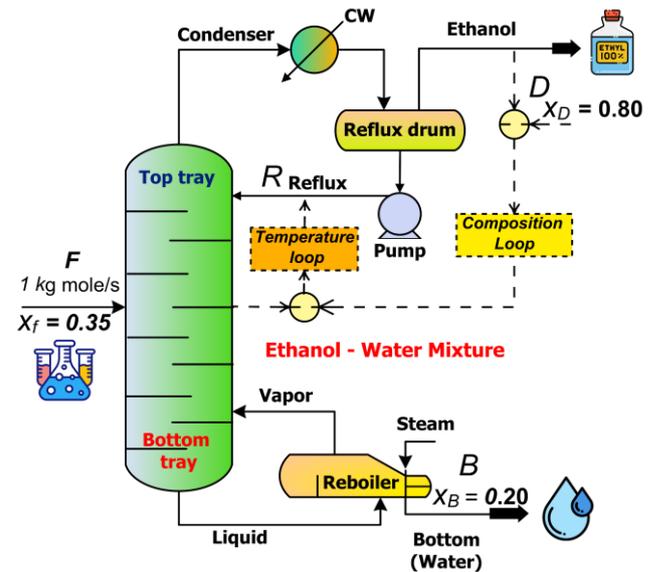


Figure 2. A systematic portrayal of the industrial distillation system

## 4. RESULTS AND DISCUSSION

### 4.1 Fault detection outcomes

The discrepancy in condenser outlet temperature of the industrial distillation system has been influenced by ramp fault resulting from process disruption. The fault was caused at a labeled point of collection at 600. Figure 3 illustrates the outcomes of the dynamic inner principal component analysis (DiPCA), bi-directional long short-term memory-autoencoder (BiLSTM-AE), and the proposed Reconstruction-based dynamic inner global-local preservation projection (DiGLPP-RBC) techniques. Figures 3 (a) and (b) illustrate the detection outcomes of DiPCA. The DiPCA can identify the problem at (821 and 686) with a substantial false alarm rate and inadequate detection effectiveness. Figures 3 (c) and (d) similarly illustrate the detection findings of BiLSTM-AE. The

Table 4. Model development parameters

Feed	Value	Top	Value	Bottom	Value	Parameters	Value
F	12 L/min	D	4.8 L/min	B	7.2 L/min	Reflux flowrate rate	0.45 L/min
T	60 °C	T	83 °C	T	92 °C	Reflux ratio	2.4
P	1.5 atm	P	2 atm	P	3.5 atm	Diameter / Height	2.1 m / 8 m
X <sub>F</sub>	0.35	X <sub>D</sub>	0.80	X <sub>B</sub>	0.20	No. of trays / Type	19/bubble cap tray
		Q <sub>Cond</sub>	-8.2KW	Q <sub>reb</sub>	10.3KW	Tray spacing	0.35 m

BiLSTM-AE can identify the problem at (662 and 638) with modest detection accuracy. The monitored statistics datasets stay under the threshold limit throughout the fault's length. Figures 3 (e) and (f) show the detection outcomes of the developed DiGLPP-RBC framework. The findings indicate that the proposed strategy may identify irregularities and failures promptly upon their emergence in the overall system. The DiGLPP-RBC can identify faults at (608 and 603) with a high detection rate with robust efficiency. The proposed strategy establishes the dynamic attributes of a system through developing a link between historical and current data.

#### 4.2 Fault diagnosis outcomes

Upon detecting irregularities and failures in the distillation system, the subsequent challenge is ascertaining the root cause variable and diagnosis density. The fault density diagnostic outcomes for this fault situation are shown in Figures 4. The findings shown in Figure 4 (a) demonstrate that the RBC index

of the DiPCA reveals the presence of confusing factors that are incorrectly identified. The results mark the targeted regions in the heat map using variables (1, 2, 8, and 10), yielding an impartial result. Comparably, The findings presented in Figure 4 (b) suggest that the RBC index of the BiLSTM-AE reveals the presence of influencing factors that cannot be precisely identified. The projected impact readily distinguishes the targeted regions in the heat map by variables (2, 8, and 10), yielding a consistent result.

Figure 4 (c) illustrates the fault density diagnostic findings derived from the proposed DiGLPP-RBC. Figure 6 demonstrates that the RBC index of the DiGLPP-RBC reveals the presence of complicating variables (2 and 10). The DiGLPP-RBC approach effectively identifies and elucidates genuine fault interpretations. The results of DiGLPP-RBC elucidate the underlying fault components and density interpretation that contribute to the misleading variable.

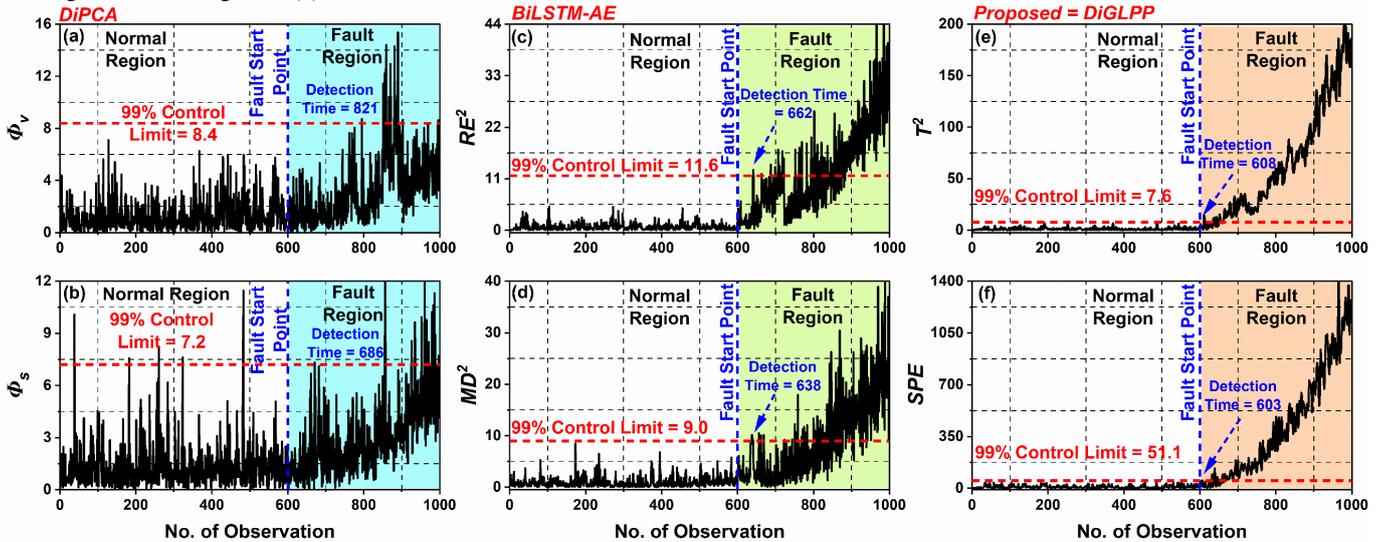


Figure 3. Fault detection results of industrial distillation system (a) DiPCA- $\Phi_v$  (b) DiPCA- $\Phi_s$  (c) BiLSTM- $RE^2$  (d) BiLSTM- $MD^2$  (e) DiGLPP- $T^2$  (f) DiGLPP- $SPE$

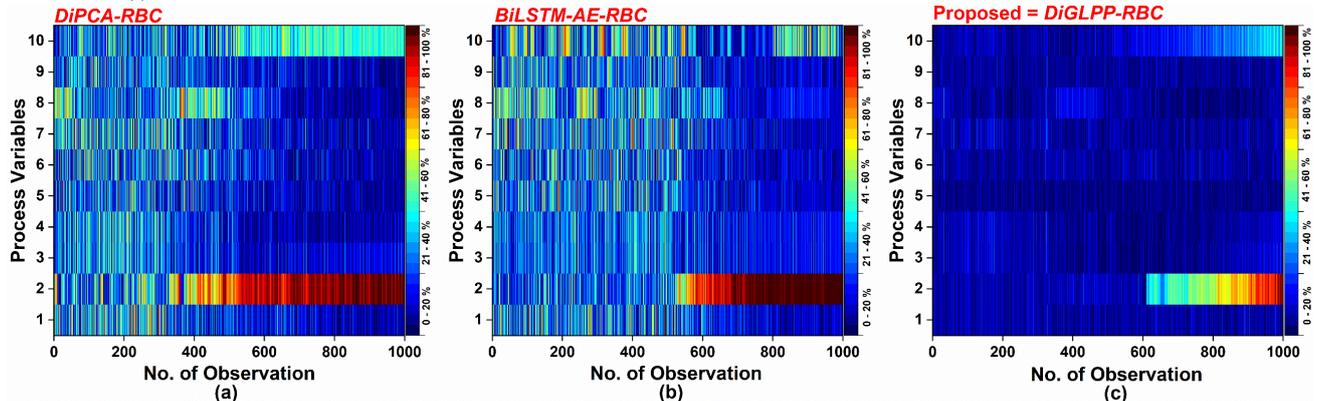


Figure 4. Reconstruction-based fault variable contribution (a) DiPCA-RBC (b) AE-BiLSTM-RBC (c) DiGLPP-RBC

## 5. CONCLUSIONS

This paper introduced an advanced dynamic inner reconstruction-based contribution with global-local preservation projection (DiGLPP-RBC) for fault detection and diagnosis. The proposed framework's robustness is compared with traditional baseline frameworks such as dynamic inner principal component analysis (DiPCA) and bi-directional long short-term memory-autoencoder (BiLSTM-AE). The results indicate that DiPCA and BiLSTM-AE can detect the abnormalities in the process at (821 and 686) and (662 and 638) with a substantial false alarm rate and inadequate detection effectiveness. The DiGLPP-RBC can detect abnormalities at (608 and 603) with a high detection rate with robust efficiency. The proposed DiGLPP-RBC technique can detect, identify, and diagnose irregularities and faults more effectively and reliably than traditional approaches. It can establish the dynamic attributes of a system through developing a link between historical and current data.

## ACKNOWLEDGEMENTS

This study is funded by the Hong Kong Research Grant Council (RGC) grant number 16203322, National Natural Science Foundation of China, grant number 62333010, and the RGC-NSFC joint project number N\_HKUST628/22.

## REFERENCES

- Alcala, C. F., & Qin, S. J. (2009). Reconstruction-based contribution for process monitoring. *Automatica*, 45(7), 1593-1600. doi:<https://doi.org/10.1016/j.automatica.2009.02.027>
- Ali, H. (2022). *Multiscale Process Monitoring In A CSTR System Using Principal Component Analysis And Signed Directed Graph*. Universiti Teknologi PETRONAS,
- Ali, H., & Gao, F. (2023). Multiscale Detection of Chemical Process using Improved Distributed CCA-Wavelet Approach. *IFAC-PapersOnLine*, 56(2), 11675-11680. doi:<https://doi.org/10.1016/j.ifacol.2023.10.517>
- Ali, H., Maulud, A. S., Zabiri, H., Nawaz, M., & Ismail, L. (2022). Fault diagnosis by using multi-scale signed directed graph. *AIP Conference Proceedings*, 2472(1), 040002. doi:10.1063/5.0093249
- Ali, H., Maulud, A. S., Zabiri, H., Nawaz, M., Suleman, H., & Taqvi, S. A. A. (2022). Multiscale Principal Component Analysis-Signed Directed Graph Based Process Monitoring and Fault Diagnosis. *ACS Omega*, 7(11), 9496-9512. doi:10.1021/acsomega.1c06839
- Ali, H., Safdar, R., Rasool, M. H., Anjum, H., Zhou, Y., Yao, Y., . . . Gao, F. (2024). Advance Industrial Monitoring of Physio-Chemical Processes using Novel Integrated Machine Learning Approach. *Journal of Industrial Information Integration*, 100709. doi:<https://doi.org/10.1016/j.jii.2024.100709>
- Ali, H., Safdar, R., Zhou, Y., Yao, Y., Yao, L., Zhang, Z., . . . Gao, F. (2025). A novel dynamic machine learning-based explainable fusion monitoring: application to industrial and chemical processes. *Machine Learning: Science and Technology*, 6(1), 015005. doi:10.1088/2632-2153/ada088
- Ali, H., Safdar, R., Zhou, Y., Yao, Y., Yao, L., Zhang, Z., . . . Gao, F. (2024). Robust statistical industrial fault monitoring: A machine learning-based distributed CCA and low frequency control charts. *Chemical Engineering Science*, 299, 120460. doi:<https://doi.org/10.1016/j.ces.2024.120460>
- Ali, H., Zhang, Z., & Gao, F. (2023). Multiscale monitoring of industrial chemical process using wavelet-entropy aided machine learning approach. *Process Safety and Environmental Protection*, 180, 1053-1075. doi:<https://doi.org/10.1016/j.psep.2023.10.066>
- Ali, H., Zhang, Z., Safdar, R., Rasool, M. H., Yao, Y., Yao, L., & Gao, F. (2024). Fault detection using machine learning based dynamic ICA-distributed CCA: Application to industrial chemical process. *Digital Chemical Engineering*, 11, 100156. doi:<https://doi.org/10.1016/j.dche.2024.100156>
- Ding, W., Ali, H., Gao, K., Zhang, Z., & Gao, F. (2025). Novel deep learning based soft sensor feature extraction for part weight prediction in injection molding processes. *Journal of Manufacturing Systems*, 78, 58-68. doi:<https://doi.org/10.1016/j.jmsy.2024.11.011>
- Dunia, R., & Joe Qin, S. (1998). Subspace approach to multidimensional fault identification and reconstruction. *AIChE Journal*, 44(8), 1813-1831. doi:<https://doi.org/10.1002/aic.690440812>
- Luo, L. (2014). Process Monitoring with Global-Local Preserving Projections. *Industrial & Engineering Chemistry Research*, 53(18), 7696-7705. doi:10.1021/ie4039345
- Luo, L., Bao, S., Mao, J., & Tang, D. (2016). Nonlocal and local structure preserving projection and its application to fault detection. *Chemometrics and Intelligent Laboratory Systems*, 157, 177-188. doi:<https://doi.org/10.1016/j.chemolab.2016.07.014>
- Silva, D. D., Sierla, S., Alahakoon, D., Osipov, E., Yu, X., & Vyatkin, V. (2020). Toward Intelligent Industrial Informatics: A Review of Current Developments and Future Directions of Artificial Intelligence in Industrial Applications. *IEEE Industrial Electronics Magazine*, 14(2), 57-72. doi:10.1109/MIE.2019.2952165
- Yoon, S., & MacGregor, J. F. (2001). Fault diagnosis with multivariate statistical models part I: using steady state fault signatures. *Journal of Process Control*, 11(4), 387-400. doi:[https://doi.org/10.1016/S0959-1524\(00\)00008-1](https://doi.org/10.1016/S0959-1524(00)00008-1)
- Yu, J. (2012). Local and global principal component analysis for process monitoring. *Journal of Process Control*, 22(7), 1358-1373. doi:<https://doi.org/10.1016/j.jprocont.2012.06.008>
- Zhang, M., Ge, Z., Song, Z., & Fu, R. (2011). Global-Local Structure Analysis Model and Its Application for Fault Detection and Identification. *Industrial & Engineering Chemistry Research*, 50(11), 6837-6848. doi:10.1021/ie102564d