Improved Dynamic Latent Variable Modeling for Process Monitoring and Diagnosis

Haitian Zhang
PhD Candidate in the Department of Chemical Engineering
University of Waterloo
Waterloo, Canada
haitian.zhang@uwaterloo.ca

Qinqin Zhu
Assistant Professor in the Department of Chemical Engineering
University of Waterloo
Waterloo, Canada

Ali Elkamel
Full Professor in the Department of Chemical Engineering
University of Waterloo
Waterloo, Canada
aelkamel@uwaterloo.ca

Abstract

Due to the advancement of modern industrial processes, a considerable number of measured variables enhance the complexity of systems, progressively leading to the development of multivariate statistical analysis (MSA) methods to exploit valuable information from the collected data for predictive modeling, fault detection and diagnosis, such as partial least squares (PLS), canonical correlation analysis (CCA) and their extensions. However, these methods suffer from some issues, involving the irrelevant information extracted by PLS, and CCA's inability to exploit quality information. Latent variable regression (LVR) was designed to address these issues, but it has not been fully and systematically studied.

A concurrent kernel LVR (CKLVR) with a regularization term is designed for collinear and nonlinear data to construct a full decomposition of the original nonlinear data space, and to provide comprehensive information of the systems. Further, dynamics are inevitable in practical industrial processes, and thus a dynamic auto-regressive LVR (DALVR) is also proposed based on regularized LVR to capture dynamic variations in both process and quality data. The comprehensive monitoring framework and fault diagnosis and causal analysis scheme based on DALVR are developed. Their superiority can be demonstrated with case studies, involving the Tennessee Eastman process, Dow's refining process and three-phase flow facility process.

Keywords
Multivariate statistical analysis, Nonlinearity, Dynamics, Monitoring and Fault diagnosis.

1. Introduction
To assure the safe operation of industrial processes and the quality of their products, multivariate statistical analysis (MVA) has been extensively used for the purpose of predictive modeling (Gillo and Shelly 1976), fault detection and diagnosis (Qin 2003), and causal analysis (Peerally et al. 2017), which are critical for the enhancement of safety, reliability, and maintainability of industrial processes. When the collected data contains both process and quality data, supervised algorithms are preferred to fully exploit the information in the data, such as partial least squares (PLS) (Geladi and Kowalski 1986), canonical correlation analysis (CCA) (Hardoon et al. 2004), and latent variable regression (LVR) (Zhu 2020). For instance, PLS maximizes covariances between input and output, but its extracts

space may contain extraneous variances of process data, leading to ineffective quality prediction. Alternatively, CCA addresses this issue by maximizing their correlations, but it fails to make full use of the quality information (Liu et al. 2018). LVR was proposed to maximize the prediction projection of quality data on the latent space, and it has shown superiority over PLS and CCA in the terms of quality-relevant modeling and monitoring (Zhu 2020).

The aforementioned approaches have been established to deal with the complexity of the real-world industrial applications, involving nonlinearity, collinearity, and dynamics. However, several limitations are still involved in these methods, such as inadequate exploitation of process and quality data, incomplete decomposition of monitoring spaces, and prospects for further improvements on fault diagnosis. The main contributions of this work are outlined as follows:

- A kernel LVR (KLVR) method is designed to handle nonlinearity by incorporating Gaussian kernel function with rLVR, and a concurrent modeling and monitoring framework is developed to obtain comprehensive monitoring performance.
- A dynamic auto-regressive LVR (DALVR) method is constructed to fully utilize information in both process and quality data via an ARX model, and the corresponding concurrent modeling and monitoring scheme is also proposed.
- A DALVR-based fault diagnosis and root cause identification framework is established to locate the root causes of the detected fault and to enhance the diagnosability and interpretability.
- The proposed methods are tested for both simulation and industrial data.

The remaining part of this thesis is organized as follows. Section 2 includes the literature review on nonlinear and dynamic MVA methods, as well as fault diagnosis. Section 3 presents the proposed CKLVR algorithm and its corresponding quality and process-related monitoring framework together with case studies to show the superiority of CKLVR. The details of DALVR with the corresponding concurrent monitoring scheme is elaborated in Section 4, and the DALVE-based fault diagnosis and root cause identification framework is also demonstrated in Section 4, together with the case studies to demonstrate the effectiveness of the proposed methods. Finally, conclusions are drawn in Section 5.

2. Literature Review
2.1 Nonlinear MVA Algorithms
The typical MVA methods including PLS, CCA, and LVR are based on a linearity assumption, and their performance is usually poor in practical industrial processes, since nonlinearity is inevitable in these processes. To address it, researchers have explored their nonlinear extensions, such as kernel functions (Lai and Fyfe 2000), neural networks (NNs) (Alauddin et al. 2020), and support vector machine (SVM) (Liu et al. 2021).

Kernel MVA approaches construct the kernel matrices with some selected kernels, such as linear kernel (Fomin and Thilikos 2004), polynomial kernel (Shashua 2009), and sigmoid kernel (Camps-Valls et al. 2004). They transform the nonlinear relations in the original space into the linear relations in the kernel space. A kernel PLS (KPLS) was proposed by Rosipal and Trejo (2001) via incorporating the reproducing kernel Hilbert space with PLS to project the original data into a higher-dimensional feature space. The multi-scale KPLS algorithm (Zhang and Hu 2011) was designed to investigate the multi-scale nature of nonlinear data by combining KPLS with the wavelet analysis. Zhou et al. (2020) put forward a novel fault detection and identification method for KPLS-based monitoring to handle the issue of strong nonlinearity and few faulty samples. Nonlinear counterparts have also been developed for CCA (Chen and Wang 2021).

Qin and McAvoy (1992) investigated a neural net PLS algorithm by embedding multilayer NNs into PLS. A NN-based soft sensor was designed by Shang et al. (2014) to model the massive nonlinear data. Alauddin et al. (2020) proposed a hybrid NN model for the fault detection and diagnosis of complex process systems. Compared with NNs, SVM eliminates the local minimum problems, thus giving rise to better generalization capacity. For instance, an SVM-based algorithm designed for fault detection in high-speed trains was studied by Liu et al. (2021) with cost-sensitive strategy handling imbalanced data and weighted-feature strategy differentiating features. Other nonlinear methods such as k-nearest neighbor algorithm (Patrick and Fisher III 1970) and denoising autoencoder (Yu and Zhao 2019) are also applied for process monitoring. However, compared with kernel variants of MVA methods, these algorithms have high computation cost and low interpretability due to their complicated model structures.

2.2 Dynamic MVA Algorithms

Dynamics is one of the typical characteristics of real-world processes, and two kinds of dynamics exist, namely the temporal dependence between adjacent samples, and the dynamic cross-correlations between various variables. It is of key importance to capture temporal relations, and various dynamic models were proposed, such as dynamic PLS (Dong and Qin 2015), dynamic CCA (Dong and Qin 2018), and dynamic rLVR (Zhu et al. 2020). These methods are effective to capture cross-correlations between the input and output data.

2.3 Fault Diagnosis

Once an anomaly is detected with the monitoring statistics, it is necessary to determine its assignable causes, and this task is called fault diagnosis. Several approaches are available for fault diagnosis, such as contribution plots (Ramaker et al. 2006), subspace extraction methods (Hu et al. 2021), and reconstruction-based contribution (RBC) (Alcala and Qin 2009).

As one of the early proposed fault diagnosis methods, contribution plots diagnose the causes by identifying the corresponding contribution of each variable to the monitoring indices, with the assumption that faulty variables contribute more to the monitoring statistics. A novel contribution plots method with a new indicator was developed by Wang et al. (2018) for quality relevant fault diagnosis. Bounoua and Bakdi (2021) improved the contribution plots to analyze the cause for abnormal process conditions, indicating more accurate information. Additionally, an enhanced comprehensive contribution plots approach was designed for Bayesian fusion based distributed multivariate statistical process monitoring, which improved its fault isolation efficiency (Mou et al. 2021).

Subspace extraction methods are also developed for root cause analysis. For instance, a PCA-based subspace reconstruction approach was proposed by Dunia and Qin (1998) for multidimensional fault identification and reconstruction. Subsequently, Gertler et al. (1999) put forward an isolation enhanced PCA method. Based on a robust reconstruction error calculation, a novel fault identification method (Choi et al. 2005) was formulated to handle nonlinear relations by kernel PCA. Besides, to deal with inaccurate fault subspace extraction and unidentified false alarms, a modified PLS algorithm (Hu et al. 2021) was designed to gain a precise subspace through orthogonal decomposition and extract purer quality-related and quality-unrelated fault subspaces.

To address the issues involving non-rigorous diagnosability analysis and inability to diagnose faults with unknown directions in the previous methods, Alcala and Qin (2009) proposed an RBC approach, which denotes the amount of reconstruction in the direction of each variable to minimize the fault detection statistics. Later, several variants of RBC have also been proposed. Specifically, a generalized RBC approach with a total PLS model was created by Li et al. (2010), showing its superiority to contribution plots-based diagnosis methods. To identify the variables that are the most responsible for the fault, He et al. (2012) established a generic reconstruction based multivariate contribution analysis framework for fault diagnosis with a branch and bound algorithm for the efficient solution to the combinatorial optimization problem. Two modified RBC approaches including a generalized RBC method and RBC ratio method were proposed by Mnassri and Ouladsine (2015) to tackle complex faults and remedy the defective of the traditional RBC method.

Since RBC retains satisfactory fault diagnosis performance, it is used as the basis of our work, and its idea is reviewed briefly in the following. For a faulty sample $\mathbf{x}$, a reconstructed vector along its fault direction ξ_i with a fault magnitude of f_i is expressed as

$$\mathbf{z}_i = \mathbf{x} - \xi_i f_i \tag{2.1}$$

where $\mathbf{z}_i$ represents a fault-free sample. Then, the general form of a reconstructed fault detection index is constructed as

$$\text{Index}(\mathbf{z}_i) = \mathbf{z}_i^\top \mathbf{M} \mathbf{z}_i = \|\mathbf{z}_i\|_\mathbf{M}^2 = \|\mathbf{x} - \xi_i f_i\|_\mathbf{M}^2 \tag{2.2}$$

where $\mathbf{M}$ is determined by the corresponding monitoring index. For instance, $\mathbf{M}$ for T^2 and Q indices are $\mathbf{R}\mathbf{\Lambda}^{-1}\mathbf{R}^\top$ and $\mathbf{I} - \mathbf{P}\mathbf{R}^\top$, respectively.

RBC identifies the fault detection by minimizing $\text{Index}(\mathbf{z}_i)$ for different variable directions, and it assumes that the direction with the largest fault magnitude is the faulty direction (Alcala and Qin 2009). Taking the first derivative of $\text{Index}(\mathbf{z}_i)$ with regard to f_i and making it equal to zero yield

$$f_i = (\xi_i^\top \mathbf{M} \xi_i)^{-1} \xi_i^\top \mathbf{M} \mathbf{x}. \tag{2.3}$$

Then the amount of reconstruction along the fault direction ξ_i, denoted as $\text{RBC}_i^{\text{Index}}$, is

$$\text{RBC}_i^{\text{Index}} = \|\xi_i f_i\|_{\mathbf{M}}^2 = \mathbf{x}^\top \mathbf{M} \xi_i (\xi_i^\top \mathbf{M} \xi_i)^{-1} \xi_i^\top \mathbf{M} \mathbf{x}. \tag{2.4}$$

2.4 Root Cause Identification and Causal Analysis

Root cause identification for quality-related problems is a key step to identify the main reason of the process faults or failures. To locate the causes of faults, causal analysis, which is also known as root cause analysis, is developed to explore causal relations between components of time series, and it has attracted plenty of attractions in the related fields including public health, policy, finance, physics, medicine, biology, environmental science, and public health (Liu et al. 2021). Many efficient methods have been designed for causal analysis, such as dynamic causal model (Friston et al. 2003), Granger causality (Granger 1969), transfer entropy (Vicente et al. 2011), randomization test (La Fond and Neville 2010), and phase slope index (Nolte et al. 2008). Among them, Granger causality and transfer entropy are the most commonly used and reliable approaches, but Granger causality is preferred in complex processes due to its computation efficiency (Lindner et al. 2019).

Originated from econometrics, on the basis of linear vector-autoregressive (VAR) modeling, Granger causality (Granger 1969) was proposed to handle stochastic time-series datasets. Conceptually, as a measure of the predictive ability from a signal to another, if the predictive ability of the signal is statistically different from that in the opposite direction, there exists Granger causal relationship between these two signals. Due to its simplicity, usability, robustness, and extendibility (Brovelli et al. 2004), Granger causality has been attached great importance to by many researchers (Liu et al. 2021). At the early stage, focusing on linear relations, Geweke (1982) proposed the definition of causality for multiple time series, which was seen as the supplement and extension of Granger's previous work. Later, to address early clarification of statistical issues, Granger causality was used for local-field potentials (LFPs) (Goebel et al. 2003).

Since Granger causality can determine whether a signal Granger-causes another one, it is feasible to evaluate whether a set of variables includes valuable information that can be used to enhance the predictive performance of another set of variables for root cause analysis, which is meaningful for the causal analysis for dynamic systems (Yuan and Qin 2014). Goebel et al. (Goebel et al. 2003) combined an adaptive multivariate VAR modeling technique with Granger causal analysis to capture fast-changing cortical dynamics. Incorporating sparse regression approaches with Granger causality, a modified Granger causality framework (Valdés-Sosa et al. 2005) was provided to estimate high-dimensional dynamical models. In consideration of hemodynamics responses between regions in the brain, a measure of Granger causality was validated by David et al. (2008) to determine the connectivity from time series. Integrating explicit observation expressions for hemodynamics responses, Ryali et al. (2011) analyzed a state-space framework combined with Granger causality. Integrating the Geweke's spectral approach with the time-varying autoregressive with exogenous input (ARX) modeling method, a novel parametric scheme for conditional Granger causality was established by Li et al. (2019) to capture dynamic interaction relations in oscillatory neocortical sensorimotor networks. To reveal dynamic interaction patterns of cortical regions in surgical training on effective brain connectivity, dynamic spectral Granger causality was applied by Kamat et al. (2021) via the short-time Fourier transform approach. Considering the existence of nonlinear relations in datasets, Shen et al. (2021) proposed a root cause analysis framework which combines a recurrent NN with Granger causality test.

In the following two subsections, two most frequently used Granger causality methods, namely time-domain Granger causality (TDGC) (Yuan and Qin 2014) and conditional spectral Granger causality (CSGC) (Claessen et al. 2019), are reviewed in details and will be adapted in the proposed work.

Time-Domain Granger Causality

Given two time series $\mathbf{x}_1(k)$ and $\mathbf{x}_2(k)$ in stationary stochastic processes, a bivariate auto-regressive (AR) model is constructed as (Yuan and Qin 2014)

$$
\begin{aligned}
\mathbf{x}_1(k) &= \sum_{l=1}^{r} a_{11} \mathbf{x}_1(k-l) + \sum_{l=1}^{r} a_{12} \mathbf{x}_2(k-l) + \boldsymbol{e}_1(k) \\
\mathbf{x}_2(k) &= \sum_{l=1}^{r} a_{21} \mathbf{x}_1(k-l) + \sum_{l=1}^{r} a_{22} \mathbf{x}_2(k-l) + \boldsymbol{e}_2(k)
\end{aligned}
\tag{2.5}
$$

where r defines the AR model order, $a_{i,l}$ ($l = 1,2, \dots, r$) can represent the AR coefficients, and $\boldsymbol{e}_i(k)$ is the residual or prediction error.

To perform the AR model on each time series, one variable, $\mathbf{x}_1$ or $\mathbf{x}_2$, can be excluded. Therefore, Eq. (2.5) can be simplified into

$$
\begin{aligned}
\mathbf{x}_1(k) &= \sum_{l=1}^{r} b_{1,l}\mathbf{x}_1(k-l) + \boldsymbol{e}_{1(2)}(k) \\
\mathbf{x}_2(k) &= \sum_{l=1}^{r} b_{2,l}\mathbf{x}_2(k-l) + \boldsymbol{e}_{2(1)}(k)
\end{aligned}
\tag{2.6}
$$

where $\boldsymbol{e}_{i(j)}(k)$ expresses the prediction error that excludes the jth variable to predict the ith variable, which can be used to determine whether a causal-and-effect relationship exists.

It is defined that $\mathbf{x}_j$ Granger-causes $\mathbf{x}_i$ if the exclusion of $\mathbf{x}_j$ reduces the ability to predict $\mathbf{x}_i$ when all the other variables are involved in the regression model. The TDGC from Variable j to Variable i is measured by the index

$$
F_{\mathbf{x}_j \to \mathbf{x}_i} = \ln \frac{\mathrm{var}(\boldsymbol{e}_{i(j)})}{\mathrm{var}(\boldsymbol{e}_i)}
\tag{2.7}
$$

where $\mathrm{var}(\cdot)$ denotes the variance. Thereafter, the statistical significance is assessed by an F test.

- **Conditional Spectral Granger Causality**

As the theoretical foundation of CSGC, it is necessary to review spectral Granger causality (SGC) first (Dhamala et al. 2008). Due to oscillating variables in real-world industrial processes, it is feasible and appropriate to use SGC to investigate cause-and-effect relationship among them. SGC is the spectral decomposition of TDGC for multiple time series. Compared with TDGC, SGC provides more information on the granger causality at a specific frequency range, so it can be used to combine with TDGC for Granger causality analysis. Theoretically, SGC is defined as

$$
\mathrm{SGC}_{\mathbf{x}_j \to \mathbf{x}_i}(f) = \ln \frac{S_{ii}(f)}{S_{ii}(f) - \left(\boldsymbol{\Gamma}_{jj} - \dfrac{\boldsymbol{\Gamma}_{ij}^2}{\boldsymbol{\Gamma}_{ii}}\right)\left|\mathbf{H}_{ij}(f)\right|^2}
\tag{2.8}
$$

where $S_{ii}(f)$ defines the spectral density of the target variable $\mathbf{x}_i$ at frequency f, which represents components in spectral density matrix $\mathbf{S}$. Dhamala et al. (2008) provided the detailed calculation of $S_{ii}(f)$.

In addition to the mutual relations between $\mathbf{x}_i$ and $\mathbf{x}_j$, other variables $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_v$ may also affect them, which should be included to analyze the causality. However, SGC fails to consider this effect. Alternatively, conditional SGC (CSGC) was developed (Claessen et al. 2019). With the spectral transformation and Wilson factorization (Wilson 1972), a multivariate system with $v + 2$ variables $(\mathbf{x}_j, \mathbf{x}_i, \mathbf{z}_1, \dots, \mathbf{z}_v)$ is formulated as

$$
\begin{aligned}
\mathbf{S}(\mathbf{x}_j, \mathbf{x}_i, \mathbf{z}_1, \dots, \mathbf{z}_v, f) &= \mathbf{H}(f)\boldsymbol{\Sigma}\mathbf{H}^*(f) \\
\mathbf{U}(\mathbf{x}_j, \mathbf{z}_1, \dots, \mathbf{z}_v, f) &= \mathbf{G}(f)\boldsymbol{\Gamma}\mathbf{G}^*(f)
\end{aligned}
\tag{2.9}
$$

where $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_v$ are variables that may influence $\mathbf{x}_i$ and $\mathbf{x}_j$. $\mathbf{S}$ and $\mathbf{U}$ define the spectral matrices of complete system and system with Variable $\mathbf{x}_i$ excluded for the causality test. $\mathbf{H}$ and $\mathbf{G}$ represent the spectral transfer function matrices, while $\boldsymbol{\Sigma}$ and $\boldsymbol{\Gamma}$ are the error covariance matrix of complete and incomplete systems, respectively. Thus, the corresponding CSGC index can be computed as

$$
\mathrm{CSGC}_{\mathbf{x}_j \to \mathbf{x}_i \mid \mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_v}(f) = \ln \frac{\boldsymbol{\Gamma}_{ii}}{\left|\mathbf{Q}_{ii}(f)\boldsymbol{\Gamma}_{jj}\mathbf{Q}_{ii}^*(f)\right|}
\tag{2.10}
$$

where

$$\mathbf{Q} = \begin{bmatrix} \tilde{G}_{ii} & 0 & \tilde{G}_{i\mathbf{z}_1} & \cdots & \tilde{G}_{i\mathbf{z}_v} \\ 0 & 1 & 0 & \cdots & 0 \\ \tilde{G}_{\mathbf{z}_1 i} & 0 & \tilde{G}_{\mathbf{z}_1 \mathbf{z}_1} & \cdots & \tilde{G}_{\mathbf{z}_1 \mathbf{z}_v} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{G}_{\mathbf{z}_v i} & 0 & \tilde{G}_{\mathbf{z}_v \mathbf{z}_1} & \cdots & \tilde{G}_{\mathbf{z}_v \mathbf{z}_v} \end{bmatrix}^{-1} \times \begin{bmatrix} \tilde{H}_{ii} & \tilde{H}_{ij} & \cdots & \cdots & \tilde{H}_{i\mathbf{z}_v} \\ \tilde{H}_{ji} & \cdots & \cdots & \cdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{\mathbf{z}_v i} & 0 & \tilde{H}_{\mathbf{z}_v j} & \cdots & \tilde{H}_{\mathbf{z}_v \mathbf{z}_v} \end{bmatrix}.$$

The calculation of $\tilde{G}_{ii}$ can refer to literature (Physical Review Letters).

If $\mathbf{x}_j$ does not Granger-cause $\mathbf{x}_i$, but other variables among $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_v$ affect both $\mathbf{x}_j$ and $\mathbf{x}_i$, $\mathbf{\Gamma}_{ii}$ will be equal to $\left| \mathbf{Q}_{ii}(f)\mathbf{\Gamma}_{jj}\mathbf{Q}_{ii}^*(f) \right|$, thus leading to $\text{CSGC}_{\mathbf{x}_j \to \mathbf{x}_i | \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_v}(f) = 0$; if there is Granger causality between $\mathbf{x}_j$ and $\mathbf{x}_i$, $\text{CSGC}_{\mathbf{x}_j \to \mathbf{x}_i | \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_v}(f) > 0$. Moreover, to determine whether $\mathbf{x}_j$ Granger-causes $\mathbf{x}_i$, the robustness of CSGC is supposed to be tested against the null hypothesis.

3. Nonlinear Latent Variable Regression for Process Monitoring

With the advancement of modern science and technologies, the industrial processes tend to be increasingly complex, leading to collinearity and nonlinearity. Collinearity in collected adjacent samples can be solved by introducing a regularization term in matrix inversions, while nonlinearity can be handled via nonlinear variants of MVA approaches. However, the existing nonlinear MVA methods such as KPLS and KCCA suffer from the same issues as described in Section 2. Considering the advantage of LVR over PLS and CCA [14], a nonlinear extension of LVR is proposed in this chapter, and a concurrent decomposition with subsequent PCA operation is conducted to obtain a comprehensive modeling and monitoring performance.

3.1 Kernel Latent Variable Regression and Its Monitoring Scheme

Define ϕ as a nonlinear projection indicator to map process variables from the original space into the feature space F, which is a higher-dimensional feature space. The original process matrix $\mathbf{X}$ is transformed into $\mathbf{\Phi} = [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \ldots, \phi(\mathbf{x}_n)]^\top \in \mathbb{R}^{n \times m_f}$ in the feature space. It is assumed that in the feature space, variables are linearly related with each other. The kernel matrix is constructed as $\mathbf{K} = \mathbf{\Phi}\mathbf{\Phi}^\top \in \mathbb{R}^{n \times n}$, where $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ $(i, j = 1, 2, \ldots, n)$, and $\langle \mathbf{a}, \mathbf{b} \rangle$ denotes the dot product between vectors $\mathbf{a}$ and $\mathbf{b}$. For simplicity, a Gaussian kernel function is chosen in this work as:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{c} \right) \tag{3.1}$$

where c represents the width of a Gaussian function, which determines the robustness of kernel mapping. With the kernel function in Eq. (3.1), the optimization objective of KLVR is designed as

$$\max_{\boldsymbol{\alpha}_{\mathbf{\Phi}}, \mathbf{q}} J = \mathbf{q}^\top \mathbf{Y}^\top \mathbf{\Phi} \boldsymbol{\alpha}_{\mathbf{\Phi}} - \frac{\gamma}{2} \|\boldsymbol{\alpha}_{\mathbf{\Phi}}\|^2$$
$$s.t. \ \|\mathbf{\Phi}\boldsymbol{\alpha}_{\mathbf{\Phi}}\| = 1, \|\mathbf{q}\| = 1 \tag{3.2}$$

where $\boldsymbol{\alpha}_{\mathbf{\Phi}} = \mathbf{\Phi}^\top \boldsymbol{\alpha}$ and $\mathbf{q}$ are projection vectors for $\mathbf{\Phi}$ and $\mathbf{Y}$, respectively. Replacing the calculation of $\mathbf{\Phi}$ with kernel matrix $\mathbf{K}$ simplifies Eq. (3.2), leading to

$$\max_{\boldsymbol{\alpha}, \mathbf{q}} J = \mathbf{q}^\top \mathbf{Y}^\top \mathbf{K} \boldsymbol{\alpha} - \frac{\gamma}{2} \boldsymbol{\alpha}^\top \mathbf{K} \boldsymbol{\alpha}$$
$$s.t. \ \|\mathbf{K}\boldsymbol{\alpha}\| = 1, \|\mathbf{q}\| = 1 \tag{3.3}$$

where $\boldsymbol{\alpha}$ and $\mathbf{q}$ are weighting vectors for $\mathbf{K}$ and $\mathbf{Y}$, respectively. In KLVR, the process score vector can be rewritten as $\mathbf{t} = \mathbf{\Phi}\boldsymbol{\alpha}_{\mathbf{\Phi}} = \mathbf{K}\boldsymbol{\alpha}$.

Lagrange multipliers are used to solve Eq. (3.3):

$$\mathcal{L} = \mathbf{q}^\top \mathbf{Y}^\top \mathbf{K} \boldsymbol{\alpha} - \frac{\gamma}{2} \boldsymbol{\alpha}^\top \mathbf{K} \boldsymbol{\alpha} + \frac{\lambda_\alpha}{2}(1 - \boldsymbol{\alpha}^\top \mathbf{K}^2 \boldsymbol{\alpha}) + \frac{\lambda_q}{2}(1 - \mathbf{q}^\top \mathbf{q}). \tag{3.4}$$

After taking derivatives with respect to $\boldsymbol{\alpha}$ and $\mathbf{q}$ and setting them equal to zero, the following expressions are retained from Eq. (3.4).

$$\mathbf{KYq} = \lambda_\alpha(\mathbf{K} + \kappa\mathbf{I})\mathbf{K\alpha}$$
$$\mathbf{Y}^\top\mathbf{K\alpha} = \lambda_q\mathbf{q} \tag{3.5}$$

where $\kappa = \gamma/\lambda_\alpha$.
Arrange terms in Eq. (3.5), obtaining

$$(\mathbf{K} + \kappa\mathbf{I})^{-1}\mathbf{YY}^\top\mathbf{K\alpha} = \lambda_\alpha\lambda_q\mathbf{\alpha}$$
$$\mathbf{Y}^\top(\mathbf{K} + \kappa\mathbf{I})^{-1}\mathbf{KYq} = \lambda_q\lambda_\alpha\mathbf{q}. \tag{3.6}$$

Eq. (3.6) implies that λ_α and λ_q are eigenvalues of $(\mathbf{K} + \kappa\mathbf{I})^{-1}\mathbf{YY}^\top\mathbf{K}$ and $\mathbf{Y}^\top(\mathbf{K} + \kappa\mathbf{I})^{-1}\mathbf{KY}$, respectively, while $\mathbf{\alpha}$ and $\mathbf{q}$ are their corresponding eigenvectors.

Accordingly, the outer structure of KLVR can be obtained by iterating the following relations until convergence.
1. $\mathbf{\alpha} = (\mathbf{K} + \kappa\mathbf{I})^{-1}\mathbf{u}$, $\mathbf{\alpha} = \mathbf{\alpha}/\|\mathbf{K\alpha}\|$;
2. $\mathbf{t} = \mathbf{K\alpha}$;
3. $\mathbf{q} = \mathbf{Y}^\top\mathbf{t}$, $\mathbf{q} = \mathbf{q}/\|\mathbf{q}\|$;
4. $\mathbf{u} = \mathbf{Yq}$.

Afterwards, perform deflation on $\mathbf{\Phi}$ and $\mathbf{Y}$ to remove the effects of the extracted latent variables.
$$\mathbf{\Phi} := (\mathbf{I} - \mathbf{tt}^\top)\mathbf{\Phi}$$
$$\mathbf{Y} := (\mathbf{I} - \mathbf{tt}^\top)\mathbf{Y}.$$

Given $\mathbf{K} = \mathbf{\Phi\Phi}^\top$, $\mathbf{K}$ is deflated as

$$\mathbf{K} := (\mathbf{I} - \mathbf{tt}^\top)\mathbf{K}(\mathbf{I} - \mathbf{tt}^\top).$$

Due to the consistent inner and outer objectives, similar to LVR, inner model is not needed in KLVR. The detailed KLVR algorithm is summarized as follows:
1. Pre-process process and quality matrices to obtain centered $\mathbf{K}$ and scaled $\mathbf{Y}$.
2. The iteration number is set as $i = 1$, and initialize $\mathbf{K}$ and $\mathbf{Y}$ as $\mathbf{K}_1 = \mathbf{K}$, $\mathbf{Y}_1 = \mathbf{Y}$, and $\mathbf{u}_i$ as the first column of $\mathbf{Y}$.
3. Perform the following (1)-(4) iteratively until convergence.
(1) $\mathbf{\alpha}_i = (\mathbf{K}_i + \kappa\mathbf{I})^{-1}\mathbf{u}_i$, $\mathbf{\alpha}_i = \mathbf{\alpha}_i/\|\mathbf{K}_i\mathbf{\alpha}_i\|$;
(2) $\mathbf{t}_i = \mathbf{K}_i\mathbf{\alpha}_i$;
(3) $\mathbf{q}_i = \mathbf{Y}_i^\top\mathbf{t}_i$, $\mathbf{q}_i = \mathbf{q}_i/\|\mathbf{q}_i\|$;
(4) $\mathbf{u}_i = \mathbf{Y}_i\mathbf{q}_i$.

4. Perform deflation on $\mathbf{K}$ and $\mathbf{Y}$ as
$$\mathbf{K}_{i+1} = (\mathbf{I} - \mathbf{t}_i\mathbf{t}_i^\top)\mathbf{K}_i(\mathbf{I} - \mathbf{t}_i\mathbf{t}_i^\top)$$
$$\mathbf{Y}_{i+1} = (\mathbf{I} - \mathbf{t}_i\mathbf{t}_i^\top)\mathbf{Y}_i.$$
5. Set $i := i + 1$, and return to Step 3 until l latent variables are extracted.
With KLVR, $\mathbf{\Phi}$ and $\mathbf{Y}$ are decomposed by KLVR as
$$\mathbf{\Phi} = \mathbf{TP}^\top + \mathbf{\Phi}_r$$
$$\mathbf{Y} = \mathbf{TC}^\top + \mathbf{Y}_r$$

where $\mathbf{P} = \mathbf{\Phi}^\top\mathbf{T}$ and $\mathbf{C} = \mathbf{Y}^\top\mathbf{T}$ represent the loading matrices for $\mathbf{\Phi}$ and $\mathbf{Y}$, while $\mathbf{\Phi}_r$ and $\mathbf{Y}_r$ are residuals of $\mathbf{\Phi}$ and $\mathbf{Y}$, respectively.

According to the score matrix $\mathbf{T} = \mathbf{\Phi R}$,
$$\mathbf{R} = \mathbf{W}(\mathbf{P}^\top\mathbf{W})^{-1} = \mathbf{\Phi}^\top\mathbf{A}(\mathbf{T}^\top\mathbf{KA})^{-1} \tag{3.7}$$
where $\mathbf{A} = [\mathbf{\alpha}_1, \mathbf{\alpha}_2, ..., \mathbf{\alpha}_n]^\top \in R^{n\times l}$, and $\mathbf{W} = \mathbf{\Phi}^\top\mathbf{A}$.

For test samples $\mathbf{X}_t$, its kernel matrix $\mathbf{K}_t$ should be first centered as
$$\mathbf{K}_t := (\mathbf{K}_t - \frac{1}{n}\mathbf{1}_k\mathbf{1}_n^\top\mathbf{K})(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)$$
where $\mathbf{K}$ is the kernel matrix of the training data $\mathbf{X}$.
For a single test sample $\mathbf{x}_{new}$, its kernel is designed as $k(\mathbf{x}_i, \mathbf{x}_{new}) = \langle\phi(\mathbf{x}_i), \phi(\mathbf{x}_{new})\rangle$, where $\mathbf{x}_i$ represents the ith sample of the training data. The score vector $\mathbf{t}_{new}$ is calculated as

$$\mathbf{t}_{\text{new}} = \mathbf{R}^{\top}\phi(\mathbf{x}_{\text{new}}) = \mathbf{\Phi}^{\top}\mathbf{A}(\mathbf{T}^{\top}\mathbf{K}\mathbf{A})^{-1}\mathbf{k}_t \tag{3.8}$$

where $\mathbf{k}_t = \mathbf{\Phi}\phi(\mathbf{x}_{\text{new}})$.

Two statistics, T^2 and Q, are employed to detect the variations in the principal component subspace and residual subspace, which are decomposed with KLVR. T^2 and Q are defined as

$$T^2 = \mathbf{t}_{\text{new}}^{\top}\mathbf{\Lambda}^{-1}\mathbf{t}_{\text{new}} \tag{3.9}$$
$$Q = \|\phi_r(\mathbf{x}_{\text{new}})\|^2 \tag{3.10}$$

where $\mathbf{\Lambda} = \frac{1}{n-1}\mathbf{T}^{\top}\mathbf{T}$ includes the variances of principal components, and $\phi_r(\mathbf{x}_{\text{new}}) = \phi(\mathbf{x}_{\text{new}}) - \mathbf{P}\mathbf{t}_{\text{new}}$ represents the residual of $\phi(\mathbf{x}_{\text{new}})$. With the model parameters obtained from KLVR model, the Q index is denoted as

$$Q = \phi(\mathbf{x}_{\text{new}})^{\top}\phi(\mathbf{x}_{\text{new}}) - 2\mathbf{k}_t^{\top}\mathbf{T}\mathbf{t}_{\text{new}} + \mathbf{t}_{\text{new}}^{\top}\mathbf{T}^{\top}\mathbf{K}\mathbf{T}\mathbf{t}_{\text{new}} \tag{3.11}$$

where $\phi(\mathbf{x}_{\text{new}})^{\top}\phi(\mathbf{x}_{\text{new}}) = 1 - \frac{2}{n}\sum_{i=1}^{n}k(\mathbf{x}_i, \mathbf{x}_{\text{new}}) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}k(\mathbf{x}_i, \mathbf{x}_j)$ $(i, j = 1, 2, \ldots, n)$.

The control limits of T^2 and Q indices are designed as [3]

$$T_{\text{cl}}^2 = \frac{l(n^2 - 1)}{n(n - l)}F_{l,n-l,\alpha} \tag{3.12}$$
$$Q_{\text{cl}} = g\chi_{h,\alpha}^2$$

where α is the confidence interval, and the confidence level is represented by $(1 - \alpha) \times 100\%$. $F_{l,n-l}$ denotes an F-distribution with l and $n - l$ degrees of freedom. $\chi_{h,\alpha}^2$ defines a χ^2-distribution with h degrees of freedom, and g is a scaling factor [19]. The calculation of g and h are calculated according to Literature [17]. When the sample number n is large enough, T_{cl}^2 can also be approximated by $T_{\text{cl}}^2 = \chi_{l,\alpha}^2$ [18].

3.2 Concurrent KLVR Based Monitoring

KLVR-based monitoring only focuses on the monitoring of quality anomalies, failing to provide a comprehensive monitoring for variations in both process and quality variables. Thus, a concurrent modeling and fault detection framework is developed based on KLVR, which constructs a multilayer monitoring structure to capture and monitor variations in both quality and process spaces. The proposed method is named as concurrent KLVR (CKLVR).

CKLVR decomposes the original data space into five subspaces, including covariation subspace (CVS), input-principal subspace (IPS), input-residual subspace (IRS), output-principal subspace (OPS) and output-residual subspace (ORS). The details of CKLVR algorithm are outlined as follows.

1. Scale $\mathbf{X}$ and $\mathbf{Y}$ to zero mean and unit variance, and $\mathbf{X}$ is processed to obtain the mean-centered $\mathbf{K}$.
2. Perform KLVR on the mean-centered $\mathbf{K}$ and $\mathbf{Y}$ with l latent variables to obtain $\mathbf{T}$, $\mathbf{C}$, and $\mathbf{R}$.
3. Construct the predictable output matrix in CVS as $\hat{\mathbf{Y}} = \mathbf{T}\mathbf{C}^{\top}$, and then perform singular value decomposition (SVD) on $\hat{\mathbf{Y}}$ to capture the predictable variations with l_c principal components as

$$\hat{\mathbf{Y}} = \mathbf{U}_c\mathbf{D}_c\mathbf{V}_c^{\top} \equiv \mathbf{U}_c\mathbf{Q}_c^{\top} \tag{3.13}$$

where $\mathbf{U}_c$ reflects the covariations in $\mathbf{\Phi}$ which are predictable from $\hat{\mathbf{Y}}$. $\mathbf{Q}_c = \mathbf{V}_c\mathbf{D}_c$ includes all l_c nonzero singular values in descending order and their corresponding right singular vectors. Since $\mathbf{V}_c$ is orthonormal,

$$\mathbf{U}_c = \hat{\mathbf{Y}}\mathbf{V}_c\mathbf{D}_c^{-1} = \mathbf{T}\mathbf{C}^{\top}\mathbf{V}_c\mathbf{D}_c^{-1} = \mathbf{\Phi}\mathbf{R}\mathbf{C}^{\top}\mathbf{V}_c\mathbf{D}_c^{-1} \equiv \mathbf{\Phi}\mathbf{R}_c \tag{3.14}$$

where $\mathbf{R}_c = \mathbf{R}\mathbf{C}^{\top}\mathbf{V}_c\mathbf{D}_c^{-1} = \mathbf{\Phi}^{\top}\mathbf{A}(\mathbf{T}^{\top}\mathbf{K}\mathbf{A})^{-1}\mathbf{C}^{\top}\mathbf{V}_c\mathbf{D}_c^{-1}$.

3. The unpredictable output matrix is denoted as $\tilde{\mathbf{Y}}_c = \mathbf{Y} - \mathbf{U}_c\mathbf{Q}_c^{\top}$, and perform PCA on $\tilde{\mathbf{Y}}_c$ with l_y principal components

$$\tilde{\mathbf{Y}}_c = \mathbf{T}_y\mathbf{P}_y^{\top} + \tilde{\mathbf{Y}} \tag{3.15}$$

where $\mathbf{T}_y = \tilde{\mathbf{Y}}_c\mathbf{P}_y$ is the output-principal score matrix, $\mathbf{P}_y$ denotes the output-principal loading matrix, and $\tilde{\mathbf{Y}}$ represents the output residual matrix.

4. Variations that are quality-irrelevant but process-relevant can be extracted by $\tilde{\mathbf{\Phi}}_c = \mathbf{\Phi} - \mathbf{U}_c \mathbf{R}_c^{\dagger}$, where $\mathbf{R}_c^{\dagger} = (\mathbf{R}_c^{\mathsf{T}} \mathbf{R}_c)^{-1} \mathbf{R}_c^{\mathsf{T}}$. Theoretically, the next step is to perform PCA on $\tilde{\mathbf{\Phi}}_c$ with l_x principal components

$$\tilde{\mathbf{\Phi}}_c = \mathbf{T}_x \mathbf{P}_x^{\mathsf{T}} + \tilde{\mathbf{\Phi}} \tag{3.16}$$

where $\mathbf{T}_x = \tilde{\mathbf{\Phi}}_c \mathbf{P}_x$ denotes the input-principal score matrix, $\mathbf{P}_x$ is the input-principal loading matrix, and $\tilde{\mathbf{\Phi}}$ represents the input residuals.

The explicit nonlinear calculation can be circumvented by performing PCA on $\mathbf{K}_c = \tilde{\mathbf{\Phi}}_c \tilde{\mathbf{\Phi}}_c^{\mathsf{T}}$, and the input-principal score matrix $\mathbf{T}_x$ can be expressed as

$$\mathbf{T}_x = \tilde{\mathbf{\Phi}}_c \mathbf{P}_x = \mathbf{K}_c \mathbf{W}_x \tag{3.17}$$

where $\mathbf{W}_x$ contains scaled eigenvectors of $\frac{1}{n} \tilde{\mathbf{\Phi}}_c \tilde{\mathbf{\Phi}}_c^{\mathsf{T}}$ corresponding to the l_x largest eigenvalues. Besides, the input-principal loading matrix $\mathbf{P}_x$ can be represented as $\mathbf{P}_x = \tilde{\mathbf{\Phi}}_c^{\mathsf{T}} \mathbf{W}_x$.

Consequently, CKLVR decomposes the process and quality data as

$$\begin{cases} \mathbf{\Phi} = \mathbf{U}_c \mathbf{R}_c^{\dagger} + \mathbf{T}_x \mathbf{P}_x^{\mathsf{T}} + \tilde{\mathbf{\Phi}} \\ \mathbf{Y} = \mathbf{U}_c \mathbf{Q}_c^{\mathsf{T}} + \mathbf{T}_y \mathbf{P}_y^{\mathsf{T}} + \tilde{\mathbf{Y}} \end{cases} \tag{3.18}$$

where $\mathbf{U}_c \mathbf{R}_c^{\dagger}$ represents the predictable quality-relevant variations in process variables, showing the relations between process and quality variables, while $\mathbf{U}_c \mathbf{Q}_c^{\mathsf{T}}$ includes predictable quality variations, and they build up CVS together. $\mathbf{T}_x \mathbf{P}_x^{\mathsf{T}}$ in IPS captures major process-relevant but quality-irrelevant variations, while $\tilde{\mathbf{\Phi}}$ in IRS represents the process residuals. $\mathbf{T}_y \mathbf{P}_y^{\mathsf{T}}$ in OPS contains major quality-relevant variations that are unobservable for the process data, and ORS is constructed by the residuals $\tilde{\mathbf{Y}}$.

For a new data sample pair $(\mathbf{x}_{\text{new}}, \mathbf{y}_{\text{new}})$ can be decomposed by CKLVR into

$$\begin{cases} \phi(\mathbf{x}_{\text{new}}) = \mathbf{R}_c^{\dagger\mathsf{T}} \mathbf{u}_{c,\text{new}} + \mathbf{P}_x \mathbf{t}_{x,\text{new}} + \tilde{\phi}(\mathbf{x}_{\text{new}}) \\ \mathbf{y}_{\text{new}} = \mathbf{Q}_c \mathbf{u}_{c,\text{new}} + \mathbf{P}_y \mathbf{t}_{y,\text{new}} + \tilde{\mathbf{y}}_{\text{new}} \end{cases} \tag{3.19}$$

where score vectors $\mathbf{u}_{c,\text{new}} = \mathbf{R}_c^{\mathsf{T}} \tilde{\phi}(\mathbf{x}_{\text{new}})$, $\mathbf{t}_{x,\text{new}} = \mathbf{P}_x^{\mathsf{T}} \tilde{\phi}_c(\mathbf{x}_{\text{new}})$, and $\mathbf{t}_{y,\text{new}} = \mathbf{P}_y^{\mathsf{T}} \tilde{\mathbf{y}}_{c,\text{new}}$ are calculated directly through the kernel form $\mathbf{k}_t = \mathbf{\Phi}\phi(\mathbf{x}_{\text{new}})$, which are specifically presented as follows:

$$\mathbf{u}_{c,\text{new}} = \mathbf{D}_c^{-\mathsf{T}} \mathbf{V}_c^{\mathsf{T}} \mathbf{C}(\mathbf{T}^{\mathsf{T}} \mathbf{KA})^{-\mathsf{T}} \mathbf{A}^{\mathsf{T}} \mathbf{k}_t \tag{3.20}$$

$$\mathbf{t}_{x,\text{new}} = \mathbf{W}_x^{\mathsf{T}} \left(\mathbf{k}_t - \mathbf{U}_c \mathbf{\Gamma} \mathbf{k}_t - \mathbf{K} \mathbf{\Gamma}^{\mathsf{T}} \mathbf{u}_{c,\text{new}} + \mathbf{U}_c \mathbf{\Gamma} \mathbf{K} \mathbf{\Gamma}^{\mathsf{T}} \mathbf{u}_{c,\text{new}} \right) \tag{3.21}$$

$$\mathbf{t}_{y,\text{new}} = \mathbf{P}_y^{\mathsf{T}} \mathbf{Y} - \mathbf{P}_y^{\mathsf{T}} \mathbf{Q}_c \mathbf{D}_c^{-\mathsf{T}} \mathbf{V}_c^{\mathsf{T}} \mathbf{C}(\mathbf{T}^{\mathsf{T}} \mathbf{KA})^{-\mathsf{T}} \mathbf{A}^{\mathsf{T}} \mathbf{k}_t \tag{3.22}$$

where $\mathbf{\Gamma} = (\mathbf{D}_c^{-\mathsf{T}} \mathbf{V}_c^{\mathsf{T}} \mathbf{C}(\mathbf{T}^{\mathsf{T}} \mathbf{KA})^{-\mathsf{T}} \mathbf{AKA}(\mathbf{T}^{\mathsf{T}} \mathbf{KA})^{-1} \mathbf{C}^{\mathsf{T}} \mathbf{V}_c \mathbf{D}_c^{-1})^{-1} \mathbf{D}_c^{-\mathsf{T}} \mathbf{V}_c^{\mathsf{T}} \mathbf{C}(\mathbf{T}^{\mathsf{T}} \mathbf{KA})^{-\mathsf{T}} \mathbf{A}$ is defined for ease of presentation. In addition, $\tilde{\phi}(\mathbf{x}_{\text{new}})$ and $\tilde{\mathbf{y}}_{\text{new}}$ can be attained by

$$\begin{cases} \tilde{\phi}(\mathbf{x}_{\text{new}}) = \tilde{\phi}_c(\mathbf{x}_{\text{new}}) - \mathbf{P}_x \mathbf{t}_{x,\text{new}} \\ \tilde{\mathbf{y}}_{\text{new}} = (\mathbf{I} - \mathbf{P}_y \mathbf{P}_y^{\mathsf{T}}) \tilde{\mathbf{y}}_{c,\text{new}}. \end{cases} \tag{3.23}$$

Then, the monitoring statistic of each subspace is designed to indicate process and quality anomalies with different alarming levels.

Since the predictable quality-relevant score vector $\mathbf{u}_{c,\text{new}}$, process-relevant score vector $\mathbf{t}_{x,\text{new}}$, and unpredictable quality-relevant score vector $\mathbf{t}_{y,\text{new}}$ involves variations of the whole system, for a new testing sample pair $(\mathbf{x}_{\text{new}}, \mathbf{y}_{\text{new}})$, the Hotelling's T^2, T_c^2, T_x^2, and T_y^2 are applied for monitoring:

$$T_c^2 = (n-1) \mathbf{u}_{c,\text{new}}^{\mathsf{T}} \mathbf{u}_{c,\text{new}} \tag{3.24}$$

$$T_x^2 = \mathbf{t}_{x,\text{new}}^{\mathsf{T}} \mathbf{\Lambda}_x^{-1} \mathbf{t}_{x,\text{new}} \tag{3.25}$$

$$T_y^2 = \mathbf{t}_{y,\text{new}}^{\mathsf{T}} \mathbf{\Lambda}_y^{-1} \mathbf{t}_{y,\text{new}} \tag{3.26}$$

where $\mathbf{\Lambda}_x = \frac{1}{n-1} \mathbf{T}_x^{\mathsf{T}} \mathbf{T}_x$ and $\mathbf{\Lambda}_y = \frac{1}{n-1} \mathbf{T}_y^{\mathsf{T}} \mathbf{T}_y$ represent the variances of process and quality variables with respect to principal components l_x and l_y, respectively.

The Q_x and Q_y indices are designed for residuals of process variations, where Q_x index is

$$Q_x = \left\| \tilde{\phi}(\mathbf{x}_{\text{new}}) \right\|^2 = \left\| \tilde{\phi}_c(\mathbf{x}_{\text{new}}) - \mathbf{P}_x \mathbf{t}_{x,\text{new}} \right\|^2$$
$$= \tilde{\phi}_c^{\top}(\mathbf{x}_{\text{new}}) \tilde{\phi}_c(\mathbf{x}_{\text{new}}) - 2\tilde{\phi}_x^{\top}(\mathbf{x}_{\text{new}}) \mathbf{P}_x \mathbf{t}_{x,\text{new}} + \mathbf{t}_{x,\text{new}}^{\top} \mathbf{P}_x^{\top} \mathbf{P}_x \mathbf{t}_{x,\text{new}} \qquad (3.27)$$

where $\tilde{\phi}_c(\mathbf{x}_{\text{new}}) = \tilde{\phi}(\mathbf{x}_{\text{new}}) - \mathbf{R}_c^{\dagger\top} \mathbf{u}_{c,\text{new}} = \phi(\mathbf{x}_{\text{new}}) - \boldsymbol{\Phi}^{\top} \boldsymbol{\Gamma}^{\top} \mathbf{u}_{c,\text{new}}$, and every term in Eq. (2.26) can be computed as

$$\tilde{\phi}_c^{\top}(\mathbf{x}_{\text{new}}) \tilde{\phi}_c(\mathbf{x}_{\text{new}}) = 1 - 2\mathbf{u}_{c,\text{new}}^{\top} \boldsymbol{\Gamma} \mathbf{k}_t + \mathbf{u}_{c,\text{new}}^{\top} \boldsymbol{\Gamma} \mathbf{K} \boldsymbol{\Gamma}^{\top} \mathbf{u}_{c,\text{new}}$$
$$\tilde{\phi}_x^{\top}(\mathbf{x}_{\text{new}}) \mathbf{P}_x \mathbf{t}_{x,\text{new}} = \left(\mathbf{k}_t^{\top} - \mathbf{u}_{c,\text{new}}^{\top} \boldsymbol{\Gamma} \mathbf{K} \right)(\mathbf{I} - \mathbf{U}_c \boldsymbol{\Gamma})^{\top} \mathbf{W}_x \mathbf{t}_{x,\text{new}}$$
$$\mathbf{t}_{x,\text{new}}^{\top} \mathbf{P}_x^{\top} \mathbf{P}_x \mathbf{t}_{x,\text{new}} = \mathbf{t}_{x,\text{new}}^{\top} \mathbf{W}_x^{\top}(\mathbf{I} - \mathbf{U}_c \boldsymbol{\Gamma}) \mathbf{K}(\mathbf{I} - \boldsymbol{\Gamma}^{\top} \mathbf{U}_c^{\top}) \mathbf{W}_x \mathbf{t}_{x,\text{new}}$$

where $\tilde{\boldsymbol{\Phi}}_c = \boldsymbol{\Phi} - \mathbf{U}_c \mathbf{R}_c^{\dagger} = \boldsymbol{\Phi} - \mathbf{U}_c \boldsymbol{\Gamma} \boldsymbol{\Phi}$, $\mathbf{P}_x = \tilde{\boldsymbol{\Phi}}_c^{\top} \mathbf{W}_x$, and $\tilde{\boldsymbol{\Phi}}_c \tilde{\boldsymbol{\Phi}}_c^{\top} = (\mathbf{I} - \mathbf{U}_c \boldsymbol{\Gamma}) \mathbf{K}(\mathbf{I} - \boldsymbol{\Gamma}^{\top} \mathbf{U}_c^{\top})$.
In addition, the Q_y statistic can be calculated as

$$Q_y = \| \tilde{\mathbf{y}}_{\text{new}} \|^2 = \left\| \tilde{\mathbf{y}}_{c,\text{new}} - \mathbf{P}_y \mathbf{t}_{y,\text{new}} \right\|^2 = \tilde{\mathbf{y}}_{c,\text{new}}^{\top} \left(\mathbf{I} - \mathbf{P}_y \mathbf{P}_y^{\top} \right) \tilde{\mathbf{y}}_{c,\text{new}}$$
$$= \left(\mathbf{y}_{\text{new}} - \mathbf{Q}_c \mathbf{u}_{c,\text{new}} \right)^{\top} \left(\mathbf{I} - \mathbf{P}_y \mathbf{P}_y^{\top} \right) \left(\mathbf{y}_{\text{new}} - \mathbf{Q}_c \mathbf{u}_{c,\text{new}} \right). \qquad (3.28)$$

Table 1. Concurrent monitoring statistics and corresponding control limits

Statistics	Control limits	Alarming level
T_c^2	$T_{c,\text{cl}}^2 = \chi_{\alpha,l_c}^2$	L_1
T_x^2	$T_{x,\text{cl}}^2 = \chi_{\alpha,l_x}^2$	L_3
Q_x	$Q_{x,\text{cl}} = g_x \chi_{\alpha,h_x}^2$	L_2
T_y^2	$T_{y,\text{cl}}^2 = \chi_{\alpha,l_y}^2$	L_2
Q_y	$Q_{y,\text{cl}} = g_y \chi_{\alpha,h_y}^2$	L_3

The corresponding control limits of T_c^2, T_x^2, T_y^2, Q_x, and Q_y indices are summarized in Table 1, where g_x and g_y represent scaling factors of control limits, while h_x and h_y denote the degrees of freedom of χ^2-distribution used in the computation of $Q_{x,\text{cl}}$ and $Q_{y,\text{cl}}$, respectively [99].
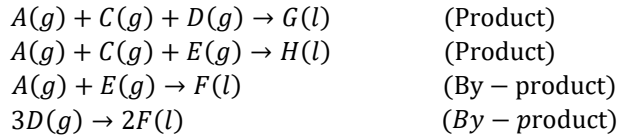
The CKLVR-based multilayer monitoring framework decomposes the original process and quality data into five subspaces, namely CVS, OPS, ORS, IPS, and IRS. Different fault alarms align with faults detected in different subspaces, which is founded on whether anomalies observed by statistics are quality-relevant or not. More concretely, L_1, L_2, and L_3 are attached for faults with descending importance. The monitoring procedure for a new sample is summarized.

1. If $T_c^2 > T_{c,\text{cl}}^2$, the L_1 fault alarm is fired in CVS with $(1 - \alpha) \times 100\%$ confidence, indicating that an output-relevant fault is detected.
2. If $T_y^2 > T_{y,\text{cl}}^2$, a quality-relevant fault is detected with $(1 - \alpha) \times 100\%$ confidence, and it is attached with a L_2 alarming level. This fault is in OPS, and it is unpredictable from process variables.
3. If $Q_y > Q_{y,\text{cl}}$, a quality-irrelevant fault with a L_3 alarming level is detected with $(1 - \alpha) \times 100\%$ confidence.
4. If $T_x^2 > T_{x,\text{cl}}^2$, an output-relevant but input-irrelevant fault is detectable for IPS with $(1 - \alpha) \times 100\%$ confidence, showing that the fault is attached to the L_3 alarming level.
5. If $Q_x > Q_{x,\text{cl}}$, a potentially output-relevant fault can be observed in IRS with $(1 - \alpha) \times 100\%$ confidence, which implies the fault pertains to the L_2 alarm.

3.3 Tennessee Eastman Process Case Study
In this section, the Tennessee Eastman Process (TEP) proposed by Downs and Vogel (1993) is used for case study to further investigate the effectiveness of the CKLVR-based multilayer monitoring framework. TEP is a benchmark industrial process to evaluate different methods of process monitoring. The whole process contains five major unit operation plants, involving a chemical reactor, product condenser, recycle compressor, vapor-liquid separator and

product stripper. Eight components (A-H) are included in TEP: four gaseous reactants (A, C, D, E), the inert material (B), and two liquid products (G,H) along with a by-product (F), which can be shown as follows:

$$A(g) + C(g) + D(g) \rightarrow G(l) \qquad (Product)$$
$$A(g) + C(g) + E(g) \rightarrow H(l) \qquad (Product)$$
$$A(g) + E(g) \rightarrow F(l) \qquad (By-product)$$
$$3D(g) \rightarrow 2F(l) \qquad (By-product)$$

Reactions involved in TEP are exothermic, irreversible and approximately 1-order with respect to concentrations of reactants. In addition, TEP is under closed-loop control.

TEP is composed of two blocks of variables with a total number of 53, including 12 manipulated variables (XMV (1-12)) and 41 measured variables (XMEAS (1-41)). Here, XMEAS (1-22) and XMV (1-11) are chosen as process variables to construct **X**, while XMEAS (35-36) are selected as quality variables to form **Y**. 500 samples are simulated in TEP for training and 960 samples for testing. Process variables XMEAS (1-22) and XMV (1-11) are sampled at an interval of 3 minutes, while quality variables XMEAS (35-36) are sampled every 6 minutes, which gives rise to the irregular sampling intervals for process and quality variables. To balance the sampling discrepancy, data sampling is necessary, and only 1/2 samples without duplicate samples are employed (Zhu et al. 2017). Accordingly, only half of the data (250 normal samples and 480 faulty samples) are used for modeling and testing.

By reference to the criteria established by Zhou et al. (2010), 15 disturbances (IDV (1-15)) can be classified into two categories, quality-relevant and quality-irrelevant disturbances. IDV (1-2, 5-8, 10, 12-13) are identified as quality-relevant disturbances, while IDV (3-4, 9, 11, 14-15) are quality-irrelevant disturbances. For the purpose of better illustration, IDV (5, 11) are chosen as examples of two kinds of disturbances to show the monitoring performance of CKLVR in detail.

By 10-fold cross validation, the numbers of principal components for CKLVR are determined as $l = 1, l_c = 2, l_x = 14$, and $l_y = 2$, while for KLVR, $l = 1$. The width of Gaussian kernel function is $c = 5000$. For CKLVR-based monitoring, Q_y is null, since $l_y = 2$ while $p = 2$. The control limit is selected to be 99 %.

Fault detection rates (FDRs) and false alarm rates (FARs) of quality-relevant and quality-irrelevant disturbances (Zhu et al. 2017) are summarized in Tables 2 and 3, respectively. The FDR and FAR are defined on the premise of KPCA-based monitoring which is considered as ground truth and taken as the criterion to n classify normal and faulty samples, where FDR is the ratio of *number of actual faults that are detectable for the proposed algorithm* and *number of total faulty samples*, while FAR is the ratio of *number of normal samples that are incorrectly categorized as faults* and *number of total normal samples*.

To investigate quality-relevant monitoring performance, T^2 of CKLVR is defined as a combination of T_c^2 and T_y^2, which illustrates the monitoring results of predictable and unpredictable quality-relevant disturbances.

Table 2. FDRs and FARs for quality-relevant disturbances in TEP

| Disturbance | FDR | | | | | FAR | | | | |
| | KLVR | CKLVR | | | | KLVR | CKLVR | | | |
	T^2	T^2	T_c^2	T_y^2	T_x^2	T^2	T^2	T_c^2	T_y^2	T_x^2
IDV (1)	0.8750	**1**	0.7639	0.9028	**1**	0.1397	0.0956	0.0931	**0.0025**	0.0122
IDV (2)	0.9536	0.9969	0.8824	0.9907	**1**	0.1847	0.1083	0.1019	**0.0064**	0.0227
IDV (5)	0.7778	**1**	0.7556	0.9778	0.9429	0.0920	0.0598	0.0552	**0.0046**	0.0187
IDV (6)	0.0731	**1**	0.0705	**1**	0.8333	0.0825	0.0722	0.0722	**0**	**0**
IDV (7)	0.8769	**1**	0.8462	0.9231	**1**	0.1687	0.1446	0.1398	**0.0048**	0.0125
IDV (8)	0.8811	**1**	0.8238	0.9736	0.9974	0.4466	0.4229	0.4190	**0.0079**	0.0217
IDV (10)	0.7308	**1**	0.6538	0.9615	0.8721	0.1740	0.1410	0.1410	**0**	0.0309
IDV (12)	0.8371	**1**	0.7841	0.9773	0.9949	0.4213	0.3889	0.3843	0.0139	**0**
IDV (13)	0.8697	**1**	0.8275	0.9718	0.9894	0.3469	0.3163	0.3163	0.0051	**0**

Table 3. FARs for quality-irrelevant disturbances in TEP

Disturbance	KLVR	CKLVR			
	T^2	T^2	T_c^2	T_y^2	T_x^2
IDV (3)	0.0972	0.0346	0.0302	**0.0043**	0.0128
IDV (4)	0.0395	0.0148	0.0148	**0**	0.2578
IDV (9)	0.1017	0.0297	0.0275	**0.0021**	0.0411
IDV (11)	0.1199	0.0407	0.0321	**0.0086**	0.1446
IDV (14)	0.0169	0.0021	**0**	0.0021	0.0361
IDV (15)	0.1064	0.0489	0.0468	**0.0021**	0.0176

As shown in Table 2, for quality-relevant disturbances in TEP, CKLVR achieves higher quality-relevant monitoring FDRs than KLVR, since it monitors quality-relevant anomalies that are both predictable and unpredictable from process variable, while KLVR only has the process information that are related with quality variables. In terms of FARs of quality-relevant disturbances, except IDV (13), the Hotelling's T^2 statistics of CKLVR and subspaces are lower than T^2 of KLVR, which indicates improved monitoring performance of CKLVR. For IDV (13), the difference of T^2 between KLVR and CKLVR is in a reasonable range, which will not affect the advantage of CKLVR over KLVR.

For quality-irrelevant disturbances, FARs are of the top priority to focus on, since the total number of faulty samples in FDR is close to zero and it is meaningless to calculate FDRs. As shown in Table 3, CKLVR gains better monitoring results than KLVR due to its smaller FARs for all cases. Furthermore, process-relevant FDRs and FARs are also available in CKLVR with T_x^2, which yet attracts lower attention than T_c^2 and T_y^2. Consequently, it is concluded that CKLVR is superior to KLVR because of its comprehensive monitoring ability and fault detection performance.

3.4 Dow's Refining Process Case Study
The Dow's refining process was presented by Braun et al. (2020) as a data science challenge problem, and the variables are anonymized due to privacy issues. The whole process consists of three distillation columns, including Primary Column, Feed Column and Secondary Column, and the Primary Column unit is controlled based on the reflux to feed ratio. 45 variables are collected from the process, including 44 process variables ($x1$-$x44$) and a quality variable (y). Among the process variables, Variables $x1$-$x21$ and $x41$-$x44$ are from the Primary Column, Variables $x22$-$x35$ are from the Secondary Column, and $x36$-$x40$ are from the Feed Column. The quality variable y is the impurity level sampled from the Primary Column. The sampling interval of the dataset is one per hour.

In this case study, the dataset is first preprocessed to remove outliers and missing values. Then 3000 samples are selected as training data, and 1753 samples are chosen as test data, among which the first 1153 samples are normal data, while the latter 600 are faulty samples. KLVR, KPLS, rLVR and PLS models are developed with the training data, and their parameters are determined by cross-validation, which are listed in Table 4. The mean squared errors (MSEs) (Braun et al. 2020) of KLVR, KPLS, rLVR and PLS for normal test samples are also summarized in Table 4. For KLVR and rLVR, $\kappa = 0.005$. As shown in the table, KLVR and KPLS obtain lower MSEs than rLVR and PLS, implying the importance to employ kernel techniques for nonlinear processes. Further, KLVR has the lowest MSE, which indicates the superiority of KLVR in terms of prediction performance over KPLS.

Table 4. Parameters and MSEs of KLVR, KPLS, rLVR and PLS in Dow's refining process

Algorithm	KLVR	KPLS	rLVR	PLS
Principal component number l	4	5	2	1
Kernel width c	900	2000	-	-
MSEs	0.5647	0.6496	1.1427	4.2561

Table 5 summarizes the FDRs and FARs of CKLVR, CKPLS, KLVR, KPLS, rLVR and PLS for faulty test samples. For both CKLVR and CKPLS, $l_c = 1$, $l_x = 5$, and $l_y = 1$. As shown in Table 6, the concurrent decomposition divides the input and output data into different subspaces including CVS, OPS, IPS and IRS, and of T_c^2, T_y^2, T_x^2, and Q_x

indices present their monitoring performance, respectively. It is noted that the ORS subspace does not exist since the output only has one variable in this case study. Due to the subsequent decomposition, concurrent kernel algorithms retain more comprehensive monitoring results than others. In addition, as shown in Table 5, CKLVR performs better than CKPLS with lower FARs, while its FDRs are comparable to those of CKPLS.

Table 5. FDRs and FARs of algorithms for faulty case in Dow's refining process

Algorithm	FDR					FAR				
	T^2	T_c^2	T_y^2	T_x^2	Q_x	T^2	T_c^2	T_y^2	T_x^2	Q_x
CKLVR	**0.9921**	0.9868	0.9553	0.9868	0	0.1545	**0**	0.1545	**0**	**0**
CKPLS	0.9921	0.9842	0.9263	**0.9973**	0.0297	0.2773	**0**	0.2773	**0**	0.0076
KLVR	0.9921	-	-	-	-	0	-	-	-	-
KPLS	0.9816	-	-	-	-	0.0182	-	-	-	-
rLVR	0.9829	-	-	-	-	0	-	-	-	-
PLS	0.9816	-	-	-	-	0	-	-	-	-

4. Dynamic Latent Variable Modeling for Temporal Modeling and Monitoring

Due to the advancement of industrial technologies, another typical characteristic involved in the process and quality data is the complicated dynamics. The dynamic MVA methods reviewed in Section 2 focus only on capturing the dynamic cross-correlations between process and quality variables, leaving the dynamics in remaining spaces unexploited. However, both dynamic cross-correlations and auto-correlations are important for accurate dynamic quality modeling. Thus, inspired by the work of Xu and Zhu (2021), a dynamic auto-regressive model is developed for LVR, namely dynamic auto-regressive LVR (DALVR), which realizes a comprehensive exploitation of dynamics in quality data.

4.1 Dynamic Auto-Regressive Latent Variable Regression

Different from the dynamic algorithms reviewed in Section 2, DALVR considers both the dynamic cross-correlations between process variations and quality variations, and the dynamic relations among quality variables. With the same denotations as in DrLVR, the score vectors t_k and u_k for $(\mathbf{x}_k, \mathbf{y}_k)$ are

$$t_k = \mathbf{x}_k^\top \mathbf{w}$$
$$u_k = \mathbf{y}_k^\top \mathbf{q}. \tag{4.1}$$

DALVR assumes that the current output scores u_k are dependent on the past output scores and input scores.

$$u_k = \beta_0 t_k + \beta_1 t_{k-1} + \cdots + \beta_s t_{k-s} + \delta_1 u_{k-1} + \cdots + \delta_d u_{k-d} + \varepsilon_k \tag{4.2}$$

where s and d represent the dynamic orders for input and output variables, respectively. β_i $(i = 0,1,\ldots,s)$ and δ_i $(i = 0,1,\ldots,d)$ are weighting parameters for input and output scores. ε_k is the regression error. The prediction of u_k can be presented as

$$\hat{u}_k = \sum_{j=1}^s \beta_j \, t_{k-j} + \sum_{j=1}^d \delta_j \, u_{k-j} = [\mathbf{x}_k^\top, \mathbf{x}_{k-1}^\top, \ldots, \mathbf{x}_{k-s}^\top](\boldsymbol{\beta} \otimes \mathbf{w}) + [\mathbf{y}_{k-1}^\top, \ldots, \mathbf{y}_{k-d}^\top](\boldsymbol{\delta} \otimes \mathbf{q}) \tag{4.3}$$

where $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_s]^\top$ is the weighting coefficient vector for $\mathbf{w}$ and $\boldsymbol{\delta} = [\delta_1, \delta_2, \ldots, \delta_d]^\top$ is the weighting coefficient vector for $\mathbf{q}$. Eq. (4.3) can be transformed into a matrix form for the whole training data $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{g+N+1}]^\top$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{g+N+1}]^\top$, which is

$$\hat{\mathbf{u}}_g = \mathbf{Z}_x(\boldsymbol{\beta} \otimes \mathbf{w}) + \mathbf{Z}_y(\boldsymbol{\delta} \otimes \mathbf{q}) \tag{4.4}$$

where $g = \max(s, d)$, and

$$\mathbf{Z}_x = [\mathbf{X}_g, \mathbf{X}_{g-1}, \ldots, \mathbf{X}_{g-s}] \in \mathbb{R}^{N \times (s+1)m}$$
$$\mathbf{Z}_y = [\mathbf{Y}_{g-1}, \mathbf{Y}_{g-2}, \ldots, \mathbf{Y}_{g-d}] \in \mathbb{R}^{N \times dp}$$
$$\mathbf{X}_i = [\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_{i+N}]^\top \in \mathbb{R}^{N \times m} \tag{4.5}$$
$$\mathbf{Y}_i = [\mathbf{y}_i, \mathbf{y}_{i+1}, \ldots, \mathbf{y}_{i+N}]^\top \in \mathbb{R}^{N \times p}$$

$i = 0,1,2, \dots, g$. It is also noted that the score vectors $\mathbf{t}_i$ and $\mathbf{u}_i$ can be calculated for each $\mathbf{X}_i$ and $\mathbf{Y}_i$.

$$\mathbf{t}_i = \mathbf{X}_i \mathbf{w}$$
$$\mathbf{u}_i = \mathbf{Y}_i \mathbf{q}. \qquad (4.6)$$

DALVR is designed to maximize the covariance between $\mathbf{u}_g$ and $\hat{\mathbf{u}}_g$, and its objective is

$$\max_{\mathbf{w},\mathbf{q},\boldsymbol{\beta},\boldsymbol{\delta}} \quad \mathbf{q}^\top \mathbf{Y}_g^\top \left[\mathbf{Z}_x(\boldsymbol{\beta} \otimes \mathbf{w}) + \mathbf{Z}_y(\boldsymbol{\delta} \otimes \mathbf{q}) \right] - \frac{\gamma_w}{2}\|\mathbf{w}\|^2 - \frac{\gamma_\beta}{2}\|\boldsymbol{\beta}\|^2 - \frac{\gamma_\delta}{2}\|\boldsymbol{\delta}\|^2$$
$$s.t. \ \|\boldsymbol{\beta} \otimes \mathbf{w}\| = 1, \|\boldsymbol{\delta} \otimes \mathbf{q}\| = 1, \|\mathbf{q}\| = 1 \qquad (4.7)$$

where γ_w, γ_β and γ_δ are regularized parameters to handle collinearity issues. For ease of presentation, define $\mathbf{X}_\beta$, $\mathbf{T}_s$, $\mathbf{Y}_\delta$, and $\mathbf{U}_d$ as

$$\mathbf{X}_\beta = \mathbf{Z}_x(\boldsymbol{\beta} \otimes \mathbf{I}_m) = \sum_{i=0}^{s} \beta_i \mathbf{X}_{g-i}$$
$$\mathbf{T}_s = \mathbf{Z}_x(\mathbf{I}_{s+1} \otimes \mathbf{w}) = \left[\mathbf{t}_g, \mathbf{t}_{g-1}, \dots, \mathbf{t}_{g-s} \right]$$
$$\mathbf{Y}_\delta = \mathbf{Z}_y(\boldsymbol{\delta} \otimes \mathbf{I}_p) = \sum_{i=1}^{d} \delta_i \mathbf{Y}_{g-i} \qquad (4.8)$$
$$\mathbf{U}_d = \mathbf{Z}_y(\mathbf{I}_d \otimes \mathbf{q}) = \left[\mathbf{u}_{g-1}, \mathbf{u}_{g-2}, \dots, \mathbf{u}_{g-d} \right].$$

Then the objective in Eq. (4.7) can be rewritten as

$$\max_{\mathbf{w},\mathbf{q},\boldsymbol{\beta},\boldsymbol{\delta}} \quad \mathbf{q}^\top \mathbf{Y}_g^\top (\mathbf{T}_s \boldsymbol{\beta} + \mathbf{U}_d \boldsymbol{\delta}) - \frac{\gamma_w}{2}\|\mathbf{w}\|^2 - \frac{\gamma_\beta}{2}\|\boldsymbol{\beta}\|^2 - \frac{\gamma_\delta}{2}\|\boldsymbol{\delta}\|^2$$
$$s.t. \ \|\mathbf{T}_s \boldsymbol{\beta}\| = 1, \|\mathbf{U}_d \boldsymbol{\delta}\| = 1, \|\mathbf{q}\| = 1 \qquad (4.9)$$

which defines the outer structure of DALVR.

The Lagrange multipliers $\lambda_{q\delta}$, $\lambda_{w\beta}$, and λ_q are employed to solve the objective in Eq. (4.7):

$$\mathcal{L} = \mathbf{q}^\top \mathbf{Y}_g^\top (\mathbf{T}_s \boldsymbol{\beta} + \mathbf{U}_d \boldsymbol{\delta}) - \frac{\gamma_w}{2}\|\mathbf{w}\|^2 - \frac{\gamma_\beta}{2}\|\boldsymbol{\beta}\|^2 - \frac{\gamma_\delta}{2}\|\boldsymbol{\delta}\|^2 + \frac{\lambda_{w\beta}}{2}(1 - \boldsymbol{\beta}^\top \mathbf{T}_s^\top \mathbf{T}_s \boldsymbol{\beta})$$
$$+ \frac{\lambda_{q\delta}}{2}(1 - \boldsymbol{\delta}^\top \mathbf{U}_d^\top \mathbf{U}_d \boldsymbol{\delta}) + \frac{\lambda_q}{2}(1 - \mathbf{q}^\top \mathbf{q}). \qquad (4.10)$$

Take derivatives with regard to $\mathbf{w}$, $\mathbf{q}$, $\boldsymbol{\beta}$, and $\boldsymbol{\delta}$ and set these equations to zero, leading to

$$\left(\mathbf{X}_\beta^\top \mathbf{X}_\beta + \kappa_w \mathbf{I} \right)^{-1} \mathbf{X}_\beta^\top \mathbf{u}_g = \lambda_{w\beta} \mathbf{w} \qquad (4.11)$$
$$\left(\mathbf{Y}_\delta^\top \mathbf{Y}_\delta + \kappa_q \mathbf{I} \right)^{-1} \left(\mathbf{Y}_g^\top \mathbf{T}_s \boldsymbol{\beta} + \mathbf{Y}_g^\top \mathbf{U}_d \boldsymbol{\delta} + \mathbf{Y}_\delta^\top \mathbf{u}_g \right) = \lambda_{q\delta} \mathbf{q} \qquad (4.12)$$
$$\left(\mathbf{T}_s^\top \mathbf{T}_s + \kappa_\beta \mathbf{I} \right)^{-1} \mathbf{T}_s^\top \mathbf{u}_g = \lambda_{w\beta} \boldsymbol{\beta} \qquad (4.13)$$
$$(\mathbf{U}_d^\top \mathbf{U}_d + \kappa_\delta \mathbf{I})^{-1} \mathbf{U}_d^\top \mathbf{u}_g = \lambda_{q\delta} \boldsymbol{\delta} \qquad (4.14)$$

where $\kappa_w = \gamma_w/\lambda_{w\beta}$, $\kappa_\beta = \gamma_\beta/\lambda_{w\beta}$, $\kappa_\delta = \gamma_\delta/\lambda_{q\delta}$, and $\kappa_q = \lambda_{q\delta}/\lambda_q$. κ_w, κ_β, and κ_δ can be processed by the iteration calculation of Eqs. (4.11) - (4.14).

As shown in Eqs. (4.11) - (4.14), the compact form of $\mathbf{w}$, $\mathbf{q}$, $\boldsymbol{\beta}$, and $\boldsymbol{\delta}$ cannot be obtained. Thus, the model parameters are calculated iteratively as follows.

1. Scale $\mathbf{X}$ and $\mathbf{Y}$ to zero mean and unit variance.
2. Initialize $\mathbf{u}$ as some column of $\mathbf{Y}$, $\mathbf{u}_g$ as some column of $\mathbf{Y}_g$, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ as unit random vectors.
3. Repeat the following relations iteratively until convergence.

(1) Construct $\mathbf{X}_\beta = \sum_{i=0}^{s} \beta_i \mathbf{X}_{g-i}$, $\mathbf{Y}_\delta = \sum_{i=1}^{d} \delta_i \mathbf{Y}_{g-i}$, and $\mathbf{U}_d = \left[\mathbf{u}_{g-1}, \mathbf{u}_{g-2}, \dots, \mathbf{u}_{g-d} \right]$;
(2) $\mathbf{w} = \left(\mathbf{X}_\beta^\top \mathbf{X}_\beta + \kappa_w \mathbf{I} \right)^{-1} \mathbf{X}_\beta^\top \mathbf{u}_g$;
(3) $\mathbf{t} = \mathbf{X}\mathbf{w}$, and form $\mathbf{T}_s = \left[\mathbf{t}_g, \mathbf{t}_{g-1}, \dots, \mathbf{t}_{g-s} \right]$;
(4) $\mathbf{q} = \left(\mathbf{Y}_\delta^\top \mathbf{Y}_\delta + \kappa_q \mathbf{I} \right)^{-1} \left(\mathbf{Y}_g^\top \mathbf{T}_s \boldsymbol{\beta} + \mathbf{Y}_g^\top \mathbf{U}_d \boldsymbol{\delta} + \mathbf{Y}_\delta^\top \mathbf{u}_g \right)$, and $\mathbf{q} = \mathbf{q}/\|\mathbf{q}\|$;
(5) $\mathbf{u}_g = \mathbf{Y}_g \mathbf{q}$;
(6) $\boldsymbol{\beta} = \left(\mathbf{T}_s^\top \mathbf{T}_s + \kappa_\beta \mathbf{I} \right)^{-1} \mathbf{T}_s^\top \mathbf{u}_g$, and $\boldsymbol{\beta} = \boldsymbol{\beta}/\|\mathbf{T}_s \boldsymbol{\beta}\|$;

(7) $\boldsymbol{\delta} = (\mathbf{U}_d^\top \mathbf{U}_d + \kappa_\delta \mathbf{I})^{-1} \mathbf{U}_d^\top \mathbf{u}_g$, and $\boldsymbol{\delta} = \boldsymbol{\delta}/\|\mathbf{U}_d \boldsymbol{\delta}\|$.

To align with the modeling of dynamic outer structure, a dynamic inner structure is built based on the latent variable modeling structure to describe the relations among $\mathbf{u}_g$, $\mathbf{T}_s$ and $\mathbf{U}_d$:

$$\mathbf{u}_g = \mathbf{T}_s \boldsymbol{\zeta} + \mathbf{U}_d \boldsymbol{\psi} + \boldsymbol{\varepsilon}_g \tag{4.15}$$

where $\boldsymbol{\zeta} = [\zeta_0, \zeta_1, \ldots, \zeta_s]^\top$ and $\boldsymbol{\psi} = [\psi_1, \psi_2, \ldots, \psi_d]^\top$ are weighting vectors of input and output matrices respectively, and $\boldsymbol{\varepsilon}_g$ is the regression error. With the obtained $\boldsymbol{\zeta}$ and $\boldsymbol{\psi}$, $\mathbf{u}_g$ can be predicted by

$$\hat{\mathbf{u}}_g = \mathbf{T}_s \boldsymbol{\zeta} + \mathbf{U}_d \boldsymbol{\psi}. \tag{4.16}$$

After the outer and inner models are obtained, deflation should be performed on $\mathbf{X}$ and $\mathbf{Y}$ to eliminate the effects of the extracted latent variables. The input matrix $\mathbf{X}$ can be deflated by

$$\mathbf{X} := \mathbf{X} - \mathbf{t}\mathbf{p}^\top \tag{4.17}$$

where $\mathbf{p} = \mathbf{X}^\top \mathbf{t} / \mathbf{t}^\top \mathbf{t}$.

The deflation of $\mathbf{Y}$ needs to be partitioned into two parts: the static part $\mathbf{Y}_{0:g-1} \equiv \{\mathbf{y}_i\}_{i=0}^{g-1}$ and the dynamic part $\mathbf{Y}_g \equiv \{\mathbf{y}_i\}_{i=g}^{n}$. $\mathbf{Y}_{0:g-1}$ is deflated by

$$\mathbf{Y}_{0:g-1} := \mathbf{Y}_{0:g-1} - \mathbf{t}\mathbf{c}_{st}^\top \tag{4.18}$$

where $\mathbf{c}_{st} = \mathbf{Y}_{0:g-1}^\top \mathbf{t} / \mathbf{t}^\top \mathbf{t}$; while $\mathbf{Y}_g$ is deflated by

$$\mathbf{Y}_g := \mathbf{Y}_g - \hat{\mathbf{u}}_g \mathbf{c}^\top \tag{4.19}$$

where $\mathbf{c} = \mathbf{Y}_g^\top \hat{\mathbf{u}}_g / \hat{\mathbf{u}}_g^\top \hat{\mathbf{u}}_g$.
After modeling with DALVR, $\mathbf{X}$ and $\mathbf{Y}_g$ are decomposed as

$$\begin{cases} \mathbf{X} = \sum_{i=1}^{l} \mathbf{t}_i \mathbf{p}_i^\top + \tilde{\mathbf{X}} = \mathbf{T}\mathbf{P}^\top + \tilde{\mathbf{X}} \\[2mm] \mathbf{Y}_g = \sum_{i=1}^{l} \hat{\mathbf{u}}_{g,i} \mathbf{c}_i^\top + \tilde{\mathbf{Y}}_g = \hat{\mathbf{U}}_g \mathbf{C}^\top + \tilde{\mathbf{Y}}_g \end{cases} \tag{4.20}$$

where $\hat{\mathbf{U}}_g = [\hat{\mathbf{u}}_{g,1}, \hat{\mathbf{u}}_{g,2}, \ldots, \hat{\mathbf{u}}_{g,l}] \in \mathbb{R}^{N \times l}$, and $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}_g$ are residuals of $\mathbf{X}$ and $\mathbf{Y}_g$. $\mathbf{t}_i$, $\mathbf{p}_i$, $\hat{\mathbf{u}}_{g,i}$, and $\mathbf{c}_i$ correspond to the ith latent variable ($i = 1, 2, \ldots, l$), respectively.

4.2 Concurrent Dynamic Decomposition
To realize a comprehensive modeling and monitoring of the process and quality spaces as for KLVR in Section 3, the modified concurrent dynamic modeling framework is also developed for DALVR, which is summarized as follows.

1. DALVR is performed on scaled $\mathbf{X}$ and $\mathbf{Y}$ for the modeling parameters $\mathbf{R}$, $\mathbf{Q}$, $\mathbf{C}$, $\boldsymbol{\zeta}$ and $\boldsymbol{\psi}$.
2. Divide the dynamic predictable output $\hat{\mathbf{Y}}_g$ into the *dynamic* cross-correlation output part $\hat{\mathbf{Y}}_{xg}$ and *dynamic* auto-correlation output part $\hat{\mathbf{Y}}_{yg}$:

$$\hat{\mathbf{Y}}_g = \hat{\mathbf{Y}}_{xg} + \hat{\mathbf{Y}}_{yg} = [\mathbf{Z}_x (\boldsymbol{\zeta} \otimes \mathbf{R})] \mathbf{C}^\top + [\mathbf{Z}_y (\boldsymbol{\psi} \otimes \mathbf{Q})] \mathbf{C}^\top \tag{4.21}$$

where $\hat{\mathbf{Y}}_{xg}$ are the quality variations that can be predicted from input variables, and $\hat{\mathbf{Y}}_{yg}$ are auto-regressive quality variations.
(1) SVD is performed on $\hat{\mathbf{Y}}_{xg}$ to capture the *dynamic* cross-correlations that are observable from input variables:

$$\hat{\mathbf{Y}}_{xg} = \mathbf{U}_c \mathbf{D}_c \mathbf{V}_c \equiv \mathbf{U}_c \mathbf{Q}_c^\top \tag{4.22}$$

where $\mathbf{U}_c$ represents the covariations in $\mathbf{X}$ that are relevant to $\hat{\mathbf{Y}}_{xg}$; $\mathbf{Q}_c = \mathbf{V}_c \mathbf{D}_c$ contains all non-zero singular values in a descending order and the corresponding right singular vectors. Due to the orthogonality of $\mathbf{V}_c$, $\mathbf{U}_c$ can be rewritten as

$$\mathbf{U}_c = \hat{\mathbf{Y}}_{xg} \mathbf{V}_c \mathbf{D}_c^{-1} \equiv \mathbf{X}_\zeta \mathbf{R}_c \tag{4.23}$$

where $\mathbf{X}_\zeta = [\mathbf{Z}_x (\boldsymbol{\zeta} \otimes \mathbf{I})] = \sum_{i=0}^{s} \zeta_i \mathbf{X}_{s-i}$, and $\mathbf{R}_c = \mathbf{R}\mathbf{C}^\top \mathbf{V}_c \mathbf{D}_c^{-1}$.
(2) Define $\hat{\mathbf{U}}_\psi = [\mathbf{Z}_y (\boldsymbol{\psi} \otimes \mathbf{Q})]$, and the auto-regressive part $\hat{\mathbf{Y}}_{yg}$ can be written as

$$\hat{\mathbf{Y}}_{yg} = [\mathbf{Z}_y (\boldsymbol{\psi} \otimes \mathbf{Q})] \mathbf{C}^\top \equiv \hat{\mathbf{U}}_\psi \mathbf{C}^\top. \tag{4.24}$$

3. The quality residual $\mathbf{Y}_c = \mathbf{Y} - \widehat{\mathbf{Y}}_{xg} - \widehat{\mathbf{Y}}_{yg}$ only contain *static* variations, and thus perform static PCA on $\mathbf{Y}_c$ to extract *static* principal variations:

$$\mathbf{Y}_c = \mathbf{T}_{sy}\mathbf{P}_{sy}^\top + \widetilde{\mathbf{Y}} \tag{4.25}$$

where $\mathbf{T}_{sy}$ and $\mathbf{P}_{sy}$ are the score and loading matrices of static output principal variations, respectively. $\widetilde{\mathbf{Y}}$ is the static output residual.

4. Dynamic inner PCA (DiPCA) (Dong and Qin 2018) is performed on $\mathbf{X}$ to capture both dynamic and static input variations simultaneously:

$$\mathbf{X} = \widehat{\mathbf{T}}_{dx}\mathbf{P}_{dx}^\top + \mathbf{T}_{sx}\mathbf{P}_{sx}^\top + \widetilde{\mathbf{X}}_s \tag{4.26}$$

where $\widehat{\mathbf{T}}_{dx}$ is the score matrix of dynamic input variations and $\mathbf{P}_{dx}$ is the loading matrix of dynamic input variations, which are predictable for the past input scores. $\mathbf{T}_{sx}$ is the score matrix of static input variations and $\mathbf{P}_{sx}$ is the loading matrix of static input variations, and $\widetilde{\mathbf{X}}_s$ represents the static input residuals.

For better understanding, the concurrent DALVR (CDALVR) based modeling framework is summarized in Figure 1, where the subspaces names are also defined.
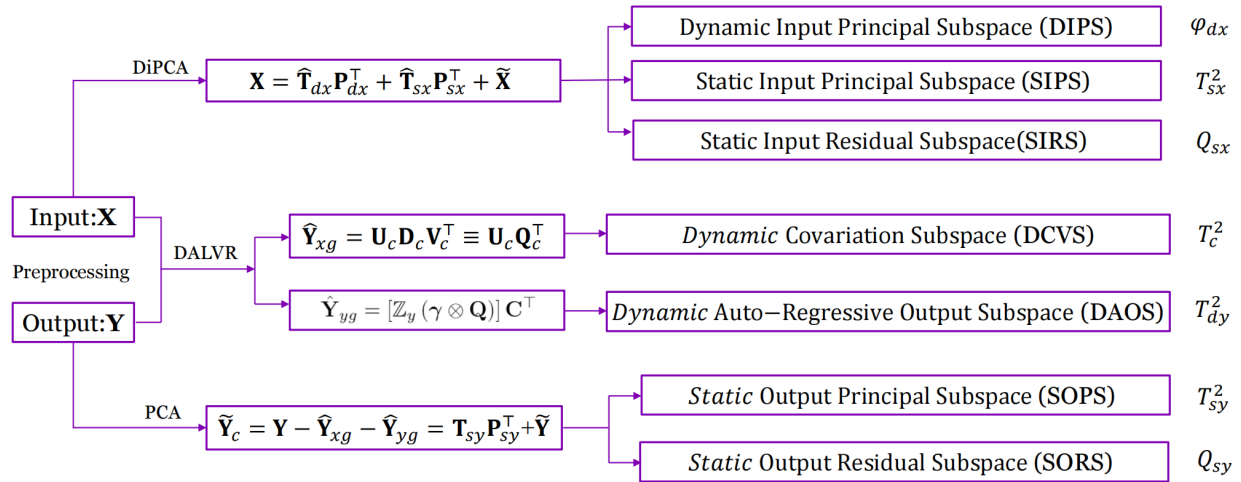
Figure 1. CDALVR-based monitoring framework

With CDALVR, $\mathbf{X}$ and $\mathbf{Y}$ are decomposed as

$$\begin{cases} \mathbf{X} = \widehat{\mathbf{T}}_{dx}\mathbf{P}_{dx}^\top + \mathbf{T}_{sx}\mathbf{P}_{sx}^\top + \widetilde{\mathbf{X}}_s \\ \mathbf{Y}_g = \mathbf{U}_c\mathbf{Q}_c^\top + \widehat{\mathbf{U}}_\psi\mathbf{C}^\top + \mathbf{T}_{sy}\mathbf{P}_{sy}^\top + \widetilde{\mathbf{Y}}. \end{cases} \tag{4.27}$$

After the DALVR model is obtained with training data, it can be employed for online modeling, and it decomposes a new test sample $(\mathbf{x}_k, \mathbf{y}_k)$ into

$$\begin{cases} \mathbf{x}_k = \mathbf{P}_{dx}\widehat{\mathbf{t}}_{dx,k} + \mathbf{P}_{sx}\mathbf{t}_{sx,k} + \widetilde{\mathbf{x}}_k \\ \mathbf{y}_k = \mathbf{Q}_c\mathbf{u}_{c,k} + \mathbf{C}\widehat{\mathbf{u}}_{\psi,k} + \mathbf{P}_{sy}\mathbf{t}_{sy,k} + \widetilde{\mathbf{y}}_k \end{cases} \tag{4.28}$$

where for $\mathbf{x}_k$, $\widehat{\mathbf{t}}_{dx,k}$ is the score vector in DIPS, $\mathbf{t}_{sx,k}$ is the score vector in SIPS, and $\widetilde{\mathbf{x}}_k$ is the residual vector in SIRS, which are obtained by

$$\widehat{\mathbf{t}}_{dx,k} = \sum_{i=1}^{s} \boldsymbol{\Theta}_{x,i}\mathbf{t}_{dx,k-i} \tag{4.29}$$

$$\mathbf{t}_{sx,k} = \mathbf{P}_{sx}\widetilde{\mathbf{x}}_c \tag{4.30}$$

$$\widetilde{\mathbf{x}}_k = (\mathbf{I} - \mathbf{P}_{sx}\mathbf{P}_{sx}^\top)\widetilde{\mathbf{x}}_c \tag{4.31}$$

where $\mathbf{t}_{dx,k} = \mathbf{R}_x\mathbf{x}_k$, and $\widetilde{\mathbf{x}}_c = \mathbf{x}_k - \mathbf{P}_{dx}\widehat{\mathbf{t}}_{dx,k}$. $\mathbf{P}_{dx}$, $\mathbf{P}_{sx}$, $\boldsymbol{\Theta}_x$, and $\mathbf{R}_x$ are parameters determined by DiPCA (Dong and Qin 2018).

For the decomposition of $\mathbf{y}_k$ in Eq. (4.28), $\mathbf{u}_{c,k}$ is the score vector in DCVS, $\widehat{\mathbf{u}}_{\psi,k}$ is the score vector in DAOS, $\mathbf{t}_{sy,k}$ is the score vector in SOPS, and $\widetilde{\mathbf{y}}_k$ is residual vector in SORS, which are

$$\mathbf{u}_{c,k} = \mathbf{R}_c^\top\mathbf{x}_{\zeta,k} \tag{4.32}$$

$$\hat{\mathbf{u}}_{\psi,k} = \sum_{i=1}^{s} \psi_i \mathbf{y}_{k-i} \mathbf{Q} \tag{4.33}$$

$$\mathbf{t}_{sy,k} = \mathbf{P}_{sy}^{\top} \tilde{\mathbf{y}}_c \tag{4.34}$$

$$\tilde{\mathbf{y}}_k = (\mathbf{I} - \mathbf{P}_{sy} \mathbf{P}_{sy}^{\top}) \tilde{\mathbf{y}}_c \tag{4.35}$$

where $\mathbf{x}_{\zeta,k} = \sum_{i=0}^{s} \zeta_i \mathbf{x}_{k-i}$, and $\tilde{\mathbf{y}}_c = \mathbf{y}_k - \mathbf{Q}_c \mathbf{u}_{c,k} - \mathbf{C}\hat{\mathbf{u}}_{\psi,k}$.

4.3 Comprehensive Dynamic Monitoring Scheme

To develop a comprehensive dynamic monitoring scheme for the extracted subspaces, their monitoring statistics are designed.

For a test sample pair $(\mathbf{x}_k, \mathbf{y}_k)$, the dynamic output-relevant score vector $\mathbf{u}_{c,k}$ is monitored with the T_c^2 index in DCVS.

$$T_c^2 = \mathbf{u}_{c,k}^{\top} \mathbf{\Lambda}_c^{-1} \mathbf{u}_{c,k} = \mathbf{x}_{\zeta,k}^{\top} \mathbf{R}_c \mathbf{\Lambda}_c^{-1} \mathbf{R}_c^{\top} \mathbf{x}_{\zeta,k} \tag{4.36}$$

where $\mathbf{\Lambda}_c$ is the variance matrix of $\mathbf{U}_c$. Note that the variations in this subspace are predictable from input variables.
For the auto-regressive output score vector $\hat{\mathbf{u}}_{\psi,k}$, which is not predictable from input, its corresponding statistic T_{dy}^2 is

$$T_{dy}^2 = \hat{\mathbf{u}}_{\psi,k}^{\top} \mathbf{\Lambda}_{dy}^{-1} \hat{\mathbf{u}}_{\psi,k} \tag{4.37}$$

where $\mathbf{\Lambda}_{dy}$ is the variance matrix of $\hat{\mathbf{U}}_{\psi}$.

The static output score vector $\mathbf{t}_{sy,k}$ and the static output residual vector $\tilde{\mathbf{y}}_k$ are monitored by

$$T_{sy}^2 = \mathbf{t}_{sy,k}^{\top} \mathbf{\Lambda}_{sy}^{-1} \mathbf{t}_{sy,k} \tag{4.38}$$

$$Q_{sy} = \|\tilde{\mathbf{y}}_k\|^2 = \tilde{\mathbf{y}}_c^{\top} (\mathbf{I} - \mathbf{P}_{sy} \mathbf{P}_{sy}^{\top}) \tilde{\mathbf{y}}_c \tag{4.39}$$

where $\mathbf{\Lambda}_{sy}$ is the variance matrix of $\mathbf{T}_{sy}$.

For the decomposed input subspaces, variations in DIPS, SIPS and SIRS are captured by φ_{dx}, T_{sx}^2 and Q_{sx} respectively, which are defined as

$$\varphi_{dx} = \hat{\mathbf{t}}_{dx,k}^{\top} \mathbf{\Phi}_{dx} \hat{\mathbf{t}}_{dx,k} = \frac{T_{dx}^2}{T_{dx,\text{cl}}^2} + \frac{Q_{dx}}{Q_{dx,\text{cl}}} \tag{4.40}$$

$$T_{sx}^2 = \mathbf{t}_{sx,k}^{\top} \mathbf{\Lambda}_{sx}^{-1} \mathbf{t}_{sx,k} \tag{4.41}$$

$$Q_{sx} = \|\tilde{\mathbf{x}}_k\|^2 = \tilde{\mathbf{x}}_c^{\top} (\mathbf{I} - \mathbf{P}_{sx} \mathbf{P}_{sx}^{\top}) \tilde{\mathbf{x}}_c \tag{4.42}$$

where φ_{dx} is a combined index integrating Hotelling's T^2 with Q statistics (Qin 2003), and $\mathbf{\Phi}_{dx}$ is

$$\mathbf{\Phi}_{dx} = \frac{\mathbf{P}_{dx} \mathbf{\Lambda}_{dx}^{-1} \mathbf{P}_{dx}^{\top}}{T_{dx,\text{cl}}^2} + \frac{\mathbf{I} - \mathbf{P}_{dx} \mathbf{P}_{dx}^{\top}}{Q_{dx,\text{cl}}} \tag{4.43}$$

where $\mathbf{\Lambda}_{dx}$ is the variance matrix of $\hat{\mathbf{T}}_{dx}$, $\mathbf{P}_{dx}$ is the loading score matrix of $\hat{\mathbf{T}}_{dx}$, and $T_{dx,\text{cl}}^2$ and $Q_{dx,\text{cl}}$ are the corresponding control limits (Zhu et al.. 2017). $\mathbf{\Lambda}_{sx}$ in Eq. (4.41) represents the variance matrix of $\mathbf{T}_{sx}$.
The details of the control limits of the aforementioned statistics are summarized in Table 7, where l, l_{dy}, l_{sy}, and l_{sx} denote the number of latent variables for $\mathbf{u}_{c,k}$, $\hat{\mathbf{u}}_{\psi,k}$, $\tilde{\mathbf{y}}_c$ and $\tilde{\mathbf{x}}_c$, respectively.

To differentiate the importance of detected anomalies, as shown in Table 6, different alarming levels L_1, L_2, and L_3 are attached to faults or disturbances observed in subspaces decomposed by CDALVR, where L_1 and L_3 represent the highest and lowest alarming level, respectively. More concretely, the faults with L_1 alarming level will affect output variables, the L_2 alarming faults have potential influence on the output, and the L_3 alarms are raised for the input-relevant and output-irrelevant faults.

Table 6. Monitoring statistics and control limits for CDALVR

Subspace		Characteristics	Statistics	Control limit	Alarming level
Output	DCVS	Dynamic; predictable from input	T_c^2	$T_{c,\text{cl}}^2 = \chi_{\alpha,l_c}^2$	L_1
	DAOS	Auto-regressive; unpredictable from input	T_{dy}^2	$T_{dy,\text{cl}}^2 = \chi_{\alpha,l_{dy}}^2$	L_1

	SOPS	Static; unpredictable from input	T_{sy}^2	$T_{sy,\text{cl}}^2 = \chi_{\alpha,l_{sy}}^2$	L_1
	SORS	Static; unpredictable from input	Q_{sy}	$Q_{sy,\text{cl}} = g_y\chi_{\alpha,h_{sy}}^2$	L_3
Input	DIPS	Dynamic	φ_{dx}	$\varphi_{dx,\text{cl}} = g_{dx}\chi_{\alpha,h_{dx}}^2$	L_3
	SIPS	Static	T_{sx}^2	$T_{sx,\text{cl}}^2 = \chi_{\alpha,l_{sx}}^2$	L_3
	SIRS	Static	Q_{sx}	$Q_{sx,\text{cl}} = g_y\chi_{\alpha,h_{sy}}^2$	L_2

4.4 Additional Tennessee Eastman Process Case Study

In this section, the additional simulation data of TEP (Rieth, et al. 2017) is utilized to investigate the modeling and monitoring performance of DALVR and CDALVR. The dataset contains 50,000 training samples and 96,000 testing samples for each disturbance. 500 simulation runs are included in the dataset, and each run contains 41 measured variables (XMEAS (1-41)) and 11 manipulated variables (XMV (1-11)). To illustrate the dynamic modeling performance of DALVR, XMEAS (1-22) and XMV (1-11) are selected as input variables **X**, while XMEAS (38) is chosen as the output variable **Y**. Note that the variables in **X** are sampled with an interval of 3 minutes, while the output variable **Y** is sampled every 15 minutes. Thus, to address the irregular sampling frequency, the filtering approach described by Zhu et al. (Zhu et al. 2017) is employed to balance the sampling discrepancy.

Table 7. MSEs for DrLVR and DALVR in additional simulation of TEP

Variable	DrLVR	DALVR
XMEAS (38)	0.3570	0.3308

In terms of modeling performance, determined by the cross-validation method, the modeling parameters for DALVR are $l = 2$, $s = 2$, and $d = 2$; for DrLVR, $l = 5$ and $s = 1$. The regularized parameters for DALVR are $\gamma_w = 0.005$, $\gamma_\beta = 0.005$, and $\gamma_\delta = 0.005$, while for DrLVR, $\kappa = 0.005$. The prediction performance is measured by MSEs, which are summarized in Table 7, and the output variations predicted by DALVR and DrLVR are presented in Figure 2. Compared with DrLVR, DALVR achieves a more comprehensive exploitation on the quality space, and thus as shown in Table 7, its MSE is lower than that of DrLVR, which is in line with the predictive performance shown in Figure 2.
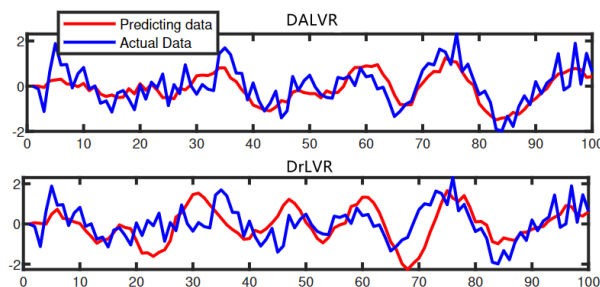

Figure 2. Predicted output variations for normal case in additional simulation of TEP

For concurrent monitoring performance, the TEP simulates 20 disturbances, and the first 15 known disturbances (IDV (1-15)) are adopted for analysis in this case study. With the control limit selected as 99%, the FDRs and FARs [100] results of output-relevant and output-irrelevant disturbances are summarized in Tables 8 and 9, respectively. Note that the T^2 statistic of CDALVR in these two tables is the combination of the T_c^2, T_{dy}^2, and T_{sy}^2 statistics, which is for output-relevant anomaly detection, including both predictable and unpredictable, dynamic and static faults. The T^2 statistic realizes a comprehensive monitoring for output relevant data.

As shown in Table 8, CDALVR achieves comparable FDRs with DrLVR, with better performance for IDV (7) and IDV (10), and its false alarms are raised much less than DrLVR. Furthermore, CDALVR-based monitoring provides detailed information for each subspace, monitoring dynamic and static variations in both input and output data. For output-irrelevant disturbances, only FARs are listed in Table 9, since there are few faulty samples and FDRs are noisy. As shown in the table, in most cases, CDALVR achieves better monitoring results than DrLVR with smaller FARs. The input-relevant FARs are also available in CDALVR with T_{sx}^2, which yet should receive lower attention than

output-related monitoring indices. As observed, CDALVR shows superiority to DrLVR due to its comprehensive monitoring ability and improved anomaly detection performance.

Table 8. FDRs and FARs for output-relevant disturbances in additional simulation of TEP

Disturbance	FDR							FAR						
	DrLVR	CDALVR						DrLVR	CDALVR					
	T^2	T^2	T_c^2	T_{dy}^2	T_{sy}^2	φ_{dx}	T_{sx}^2	T^2	T^2	T_c^2	T_{dy}^2	T_{sy}^2	φ_{dx}	T_{sx}^2
IDV (1)	**1**	0.9600	0.8429	0.4411	0.3711	0.8229	0.9998	0.9282	**0.3723**	0.1956	0	0.2501	0	0.1672
IDV (2)	**1**	0.8696	0.7446	0.0365	0.1394	0.7912	0.9998	0.9034	**0.3556**	0.1267	0	0.2899	0	0.1707
IDV (5)	**1**	0.9986	0.5994	0.2470	0.9951	0.6667	0.6783	0.9455	**0.9428**	0.9126	0	0.9380	0.2143	0.0261
IDV (6)	**1**	1	0.9689	0.8585	0.9468	0.9583	0.9934	0.6970	**0.4457**	0.3783	0.5556	0.3065	0	0.2888
IDV (7)	**0.9497**	0.9749	0.9421	0.5014	0.2703	0.8958	1	0.4395	**0.1703**	0.1241	0	0.0793	0.3846	0.0202
IDV (8)	**0.9940**	0.9741	0.9374	0.5892	0.2602	0.8077	0.9831	0.7742	**0.2325**	0.1755	0.3000	0.1267	0.3636	0.1383
IDV (10)	0.9001	**0.9565**	0.8195	0.1435	0.7580	0.3585	0.2824	0.7418	**0.7081**	0.5625	0.1667	0.6691	0.1489	0.0311
IDV (12)	**0.9950**	0.9876	0.8670	0.6576	0.6759	0.7952	0.9746	0.8092	**0.6175**	0.4327	0.3333	0.5613	0.2353	0.1812
IDV (13)	**0.9993**	0.9946	0.8040	0.6721	0.7802	0.9451	0.9732	0.6194	**0.4907**	0.3565	0.3846	0.4546	0	0.1134

Table 9. FARs for output-irrelevant disturbances in additional simulation of TEP

Disturbance	DrLVR	CDALVR					
	T^2	T^2	T_c^2	T_{dy}^2	T_{sy}^2	φ_{dx}	T_{sx}^2
IDV (3)	0.0123	0.0161	0.0100	0	0.0072	0	0.0067
IDV (4)	0.1907	0.0178	0.0106	0	0.0077	0	0.2253
IDV (9)	0.0124	0.0196	0.0118	0	0.0077	0	0.2253
IDV (11)	0.3765	0.0208	0.0133	0	0.0080	0	0.0768
IDV (14)	0.9595	0.0814	0.0260	0	0.0590	0	0.0752
IDV (15)	0.0145	0.0211	0.0144	0	0.0070	0	0.0096
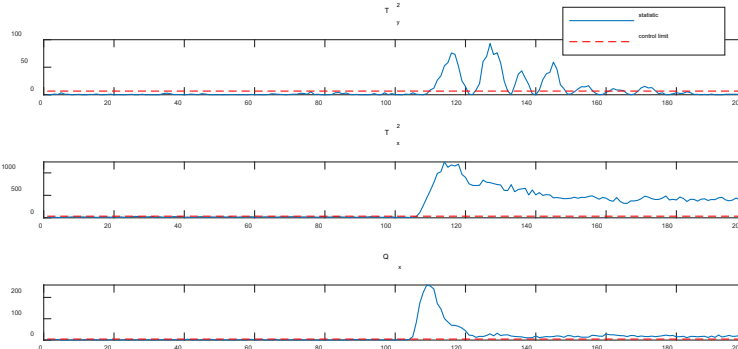
Figure 3. PCA-based process and quality monitoring results for IDV (1)

For better illustration, IDV (1) (an output-relevant disturbance) and IDV (14) (an output-irrelevant disturbance) are chosen to visualize the monitoring performance of CDALVR and DrLVR. The first 100 samples are training data, while the 101st-200th samples are the first round of testing data.
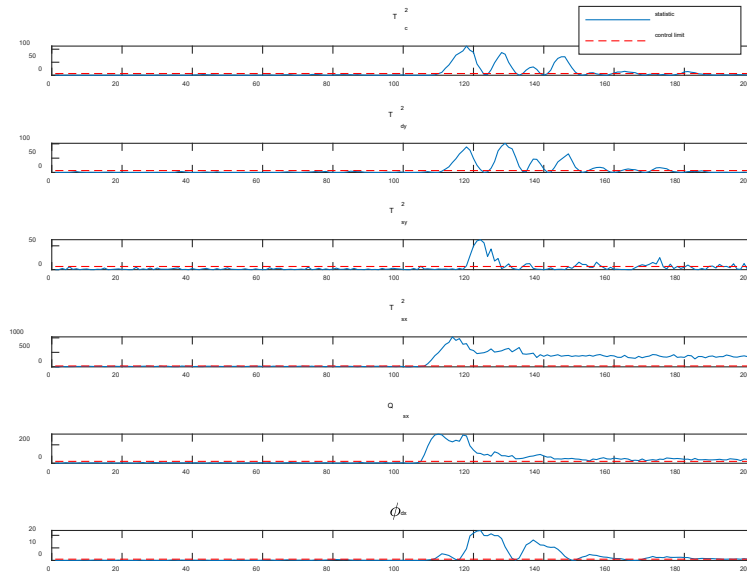
Figure 4. CDALVR-based monitoring results for IDV (1)

IDV (1) introduces a step change of A/C feed ratio in Stream 4 (Downs and Vogel 1993), and as shown in the PCA-based monitoring results in Figure 3, it affects the output variable transiently. Specifically, output variable is abnormal in the interval [110th, 174th], while for input variables, T_x^2 and Q_x detect the faults from 106th and 105th sample respectively.

The monitoring results of CDALVR and DrLVR are presented in Figures 4 and 5, respectively. Figure 5 shows that all quality relevant monitoring indices of CDALVR detect the anomalies, and they return to normal conditions later, which is consistent with the PCA-based monitoring results in Figure 4. T_c^2 is constructed by the predictable dynamic output-relevant variations, and in it captures the anomalies in DCVS for the 111st-150th faulty samples. T_{dy}^2 detects dynamic output-relevant faulty samples that are unpredicted from the input in the range of [112nd, 175th]. T_{sy}^2 also successfully detects additional static output-relevant anomalies in the interval of [118th, 173rd].
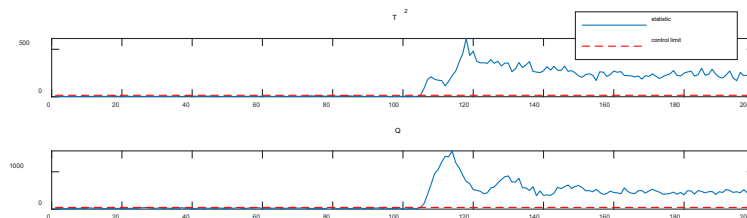
Figure 5. DrLVR-based monitoring results for IDV (1)

The monitoring results of output-relevant statistics in CDALVR are in line with the T_y^2 of PCA-monitoring in Figure 4. In addition to quality-relevant anomalies, CDALVR also detects anomalies in the input data with φ_{dx} from the 109th sample, T_{sx}^2 from the 107th sample and Q_x from the 106th sample. It is observed that IDV (1) affects input data longer than quality variables. In contrast, both monitoring indices of DrLVR keep alarming from the 105th sample, and it cannot provide a detailed analysis for the sources of the detected faults.
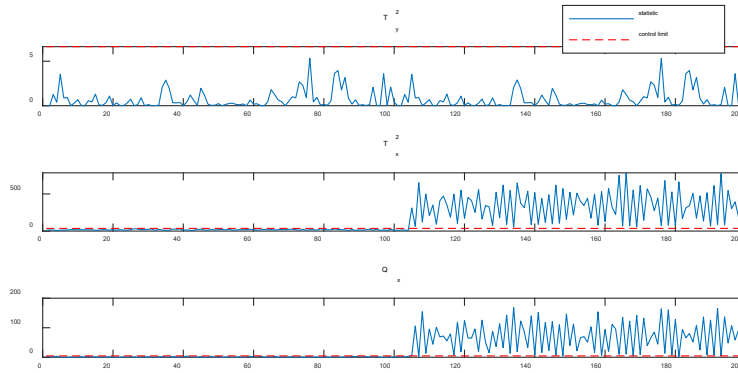
Figure 6. PCA-based process and quality monitoring results for IDV (14)
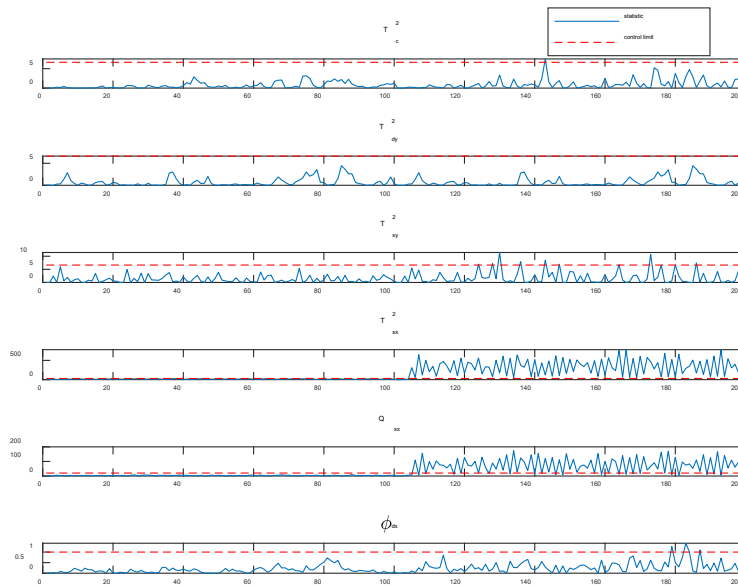

Figure 7. CDALVR-based monitoring results for IDV (14)

The IDV (14) disturbance is caused by the sticking issue of the reactor cooling water valve (Downs and Vogel 1993). The PCA-based monitoring in Figure 6 indicates that it affects T_x^2 and Q_x statistics from the 105th sample, while its T_y^2 index remains unaffected, showing that this disturbance is input-relevant and output-irrelevant. Figures 7 and 8 summarize the CDALVR-based and DrLVR-based monitoring results, respectively. CDALVR 's monitoring results generally align with Figure 6:
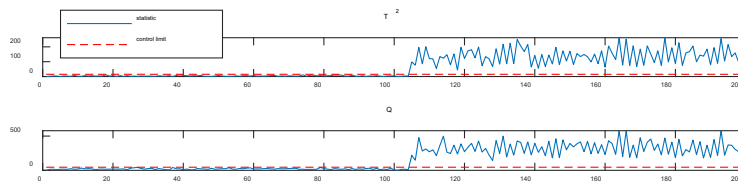
Figure 8. Dr LVR-based monitoring results for IDV (14)

only the input-relevant statistics T_{sx}^2 and Q_x raise the L_3-level alarms from the 105th and 106th faulty samples, respectively, and quality-relevant indices T_c^2, T_{dy}^2, T_{sy}^2 and dynamic input-relevant index φ_{dx} are within their corresponding control limits. DrLVR has both monitoring indices exceeding their control limit, and it is hard to collect any further information from Figure 8. Thus, DrLVR based monitoring method is less informative than CDALVR.

5. Dynamic Latent Variable Modeling for Further Fault Diagnosis

5.1 Root Cause Identification and Causal Analysis Framework
The monitoring scheme based on DALVR includes T^2 and Q statistics, which can be obtained in a similar way. To ensure the monitoring and diagnosis efficiency, a combined monitoring index φ is first designed by integrating T^2 and Q indices as

$$\varphi = \frac{T^2}{T_{cl}^2} + \frac{Q}{Q_{cl}} = \mathbf{x}^\top \mathbf{\Phi x} \tag{5.1}$$

where

$$\mathbf{\Phi} = \frac{\mathbf{R\Lambda^{-1}R^\top}}{T_{cl}^2} + \frac{\mathbf{I - PR^\top}}{Q_{cl}}.$$

φ follows a χ^2-distribution approximately, and its control limit is obtained by

$$\varphi_{cl} = g^\varphi \chi_{h^\varphi,\alpha}^2 \tag{5.2}$$

where

$$g^\varphi = \frac{tr(\mathbf{S}_x\mathbf{\Phi})^2}{tr(\mathbf{S}_x\mathbf{\Phi})}, h^\varphi = \frac{[tr(\mathbf{S}_x\mathbf{\Phi})]^2}{tr(\mathbf{S}_x\mathbf{\Phi})}.$$

and $\mathbf{S}_x$ is the covariance matrix of input $\mathbf{X}$.

Then, RBC is adopted and modified as the first step of root cause analysis. As discussed in Section 2, RBC is an effective approach for fault diagnosis, but it is not reliable, since the RBC values of normal samples are different for various variables, and thus it is difficult to interpret the RBC results for abnormal samples. To alleviate these effects that potentially weaken the performance of RBC, a relative RBC index is designed and employed to identify variables with largest fault amplitudes as contributing variables in the proposed framework. The relative RBC index $rRBC_i^{Index}$ is defined as the ratio of RBC_i^{Index} to the average values calculated with the normal data RBC_{avg}^{Index}; that is

$$rRBC_i^{Index} = \frac{RBC_i^{Index}}{RBC_{avg}^{Index}} = \frac{\mathbf{x}^\top \mathbf{M}\xi_i(\xi_i^\top \mathbf{M}\xi_i)^{-1}\xi_i^\top \mathbf{Mx}}{RBC_{avg}^{Index}} \tag{5.3}$$

where other symbols have the same meaning as those in Eq. (2.4). $rRBC_i^{Index}$ can be used to select contributing variables to the faults, and it successfully removes the interference brought by the correlations among various variables. The variables that have larger $rRBC_i^{Index}$ values are regarded as the main contributing variables, which are applied for root cause identification.
Accordingly, $rRBC_i^\varphi$ is formulated as

$$rRBC_i^\varphi = \frac{RBC_i^\varphi}{RBC_{avg,i}^\varphi} = \frac{\mathbf{x}^\top \mathbf{\Phi}\xi_i(\xi_i^\top \mathbf{\Phi}\xi_i)^{-1}\xi_i^\top \mathbf{\Phi x}}{RBC_{avg,i}^\varphi} \tag{5.4}$$

$rRBC_i^\varphi$ can remove the interference brought by the correlations among various variables. The variables that have larger $rRBC_i^\varphi$ values are regarded as the main contributing variables, which are used for further root cause identification. To further attribute the root causes of the detected fault, TDGC (Yuan and Qin 2014) and CSGC (Claessen et al. 2019) described in Section 2 are employed to identify the causal relations of the contributing variables identified by RBC.

TDGC is easy to implement and understand, so it is extensively used in root cause analysis. However, sometimes it fails to provide enough information to locate the root causes due to the existence of other potential causes. Thus, it is necessary to enhance reliability. In this work, we integrate another causality approach with TDGC to address this issue. CGSC has good causal inference performance since it can identify the causal relationship in the frequency domain, providing strong support for further root cause analysis. However, its computational cost is relatively large.

If the range of variables is narrowed down, it will be more efficient. To combine their advantages and improve the performance of causal analysis, the integration of TDGC and CSGC will lead to the improvement of reliability and effectiveness.

In the integrated framework, DALVR is first performed on historical normal data to obtain model parameters, and fault detection and diagnosis statistics are developed with DALVR. For online monitoring, new samples are processed and monitored with the established statistics. Once a faulty sample is identified, relative RBC is conducted to obtain contributing variables, which will then be employed for further root cause identification with TDGC and CSGC.

5.2 Tennessee Eastman Process Case Study

The Tennessee Eastman Process (TEP) (Downs and Vogel 1993) is analyzed to show the efficacy of the proposed framework. In this work, XMEAS (1-22) and XMV (1-11) are selected as input variables (Variables 1-33), while XMEAS (38) serves as the output variable.

The modelling performance of DrLVR and DALVR are first compared, which are trained with 500 normal samples. With the cross-validation method, the parameters are determined as $l = 1, s = 1, d = 1, \gamma_\omega = 0.005, \gamma_\beta = 0.005$, and $\gamma_\delta = 0.005$ for DALVR; and $l = 1, s = 3$, and $\kappa = 0.005$ for DrLVR. Apart from DrLVR, DAPLS is also used to reflect the effectiveness of DALVR on prediction. For DAPLS, $l = 3$, $s = 3$, and $d = 1$. The prediction performance is evaluated with MSEs, which are summarized in Table 10.

Table 10. MSEs for DrLVR, DAPLS, and DALVR in TEP

Variable	DrLVR	DAPLS	DALVR
XMEAS (38)	0.3872	0.1726	**0.1464**

Compared with DrLVR, DALVR achieves a much lower MSE than DrLVR, since it exploits more information with the aid of historical quality data. The MSEs of DALVR, its cross-correlation part, and auto-correlation part are shown in Table 11. As shown in Table 11, both cross-correlation part and auto-correlation part have higher MSEs than the overall MSE value of DALVR, which implies that the modelling performance with past process samples only or past quality samples only is not satisfactory.

Table 11. MSEs for different parts of DALVR in TEP

Variable	DALVR	Cross-correlation part	Auto-correlation part
XMEAS (38)	**0.1464**	0.6799	0.1920

15 disturbances (IDV (1–15)) are simulated in TEP (Downs and Vogel 1993), and based on the criterion proposed in the work of Zhou et al. (2010), they can be classified into two categories: quality-relevant and quality-irrelevant disturbances. IDV (1, 2, 5–8, 10, 12, and 13) are identified as quality-relevant disturbances, while IDV (3, 4, 9, 11, 14, and 15) are quality-irrelevant disturbances.

Table 12. Monitoring metrics of T^2 for quality-relevant disturbances in TEP

Faulty type	FDR			FAR			AR			PR		
	DrLVR	DAPLS	DALVR	DrLVR	DAPLS	DALVR	DrLVR	DAPLS	DALVR	DrLVR	DAPLS	DALVR
IDV (1)	0.7826	1	0.8261	0.1575	0.7808	**0.1370**	0.8281	0.4063	**0.8542**	0.6102	0.2875	**0.6552**
IDV (2)	0.9167	1	1	0.7667	0.7889	**0.7389**	0.2760	0.2604	0.2969	0.0738	**0.0779**	0.0764
IDV (5)	1	0.8414	1	0.8025	**0.0896**	0.7963	0.3187	**0.5000**	0.3240	0.1815	**0.3256**	0.1826
IDV (6)	0.9905	1	1	0.2818	**0.2364**	0.2773	0.9281	0.9458	**0.9533**	0.9220	0.9343	**0.9436**
IDV (7)	0.8182	0.9273	**0.8364**	0.3431	0.2336	**0.1387**	0.7031	0.8125	**0.8542**	0.4891	0.6145	**0.6970**
IDV (8)	0.8860	**0.9485**	0.9051	0.3263	0.5158	**0.1789**	0.7708	0.7188	**0.8125**	0.7304	0.6525	**0.8298**
IDV (10)	0.5818	0.7091	**0.7818**	0.3577	**0.2774**	0.2818	0.6250	**0.7188**	0.6938	0.3951	**0.5065**	0.4538
IDV (12)	0.7964	**0.9273**	0.8291	0.5244	0.4878	**0.4268**	0.7010	**0.7500**	0.7281	0.7145	**0.7183**	0.6796
IDV (13)	0.8145	0.9274	**0.8306**	0.2353	0.3676	**0.1018**	0.7969	0.8229	**0.8333**	0.8632	0.8214	**0.9035**

With the control limit selected as 99%, the fault detection rates (FDRs) and false alarm rates (FARs) of these two categories are summarized in Tables 12 and 13, respectively. FDR is the ratio of the number of actual faults that are

detected with the selected algorithm and the number of total faulty samples, while FAR is the ratio of the number of normal samples that are incorrectly categorized as faults and the number of total normal samples.

FDR and FAR are used to measure the monitoring performance of algorithms. The larger the FDR is, the better the monitoring performance is, while the smaller the FAR is, the better the monitoring performance is. Compared with DrLVR, DALVR can capture more variance of the data, leading to better modelling performance. In this case, theoretically, the performance of DALVR on fault detection and alarm will be better than DrLVR.

Apart from FDR and FAR, accuracy rate (AR) and precision rate (PR) can be also used as measures of observational error to compare the monitoring performance. AR is the percentage of correct predictions for a given dataset, which evaluates how close or far off the samples are to their true values. PR measures that fraction of examples classified as faulty samples that are truly faulty, which measures how close or dispersed the samples are to each other (Davis and Goadrich 2006).

Table 13. FARs of T^2 for quality-irrelevant disturbances in TEP

Fault Type	DrLVR	DAPLS	DALVR
IDV (3)	0.0957	0.0213	0.0319
IDV (4)	0.2116	0.0212	0.0847
IDV (9)	0.0851	0.0319	0.0266
IDV (11)	0.1341	0.0391	0.0391
IDV (14)	0.4063	0.0885	0.0521
IDV (15)	0.0737	0.0105	0.0211

As shown in Table 12, for output-relevant disturbances, for IDV (1, 5, 8, 12, and 13), DALVR achieves the smallest FARs and the largest ARs and PRs, while its FDRs are comparable to those of DrLVR and DAPLS. For IDV (2, 6, 7, and 10), FDRs of DALVR are largest among these algorithms, while their ARs and PRs are either the largest or relatively large, and FARs are either the smallest or relatively small. Thus, in general, DALVR has better fault detection ability, accuracy, and precision performance in quality-relevant monitoring, since compared with other static and dynamic algorithms, DALVR employs both past process and quality information for predictive modelling, and it can capture the data variations more accurately.

For output-irrelevant disturbances, only FARs are listed in Table 13, since there are few faulty samples and FDRs, ARs, and PRs are noisy. As shown in Table 13, DALVR obtains comparable or smaller FARs compared with other algorithms. FARs of DALVR are comparable or smaller compared to other algorithms, which means that there are only a few normal samples classified as faulty samples erroneously. Thus, as observed, DALVR shows superiority due to its comprehensive monitoring ability and improved anomaly detection performance.
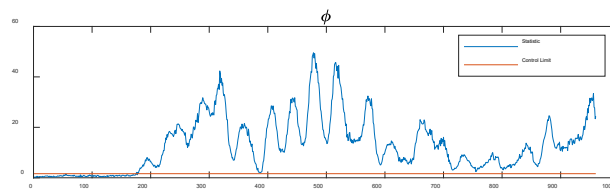

Figure 9. DALVR-based monitoring result for IDV (8)

One of the quality-relevant faults, IDV (8), is selected as an instance to illustrate the proposed fault diagnosis and root cause identification framework. IDV (8) contains a disturbance caused by a random variation of A, B, C Feed composition in Stream 4 with standard derivation of A, B feed as 3% and 0.3%, which will directly affect Stream 4 and the stripper (Liu et al. 2021). Besides, other variables are influenced by the flows or feedback control.
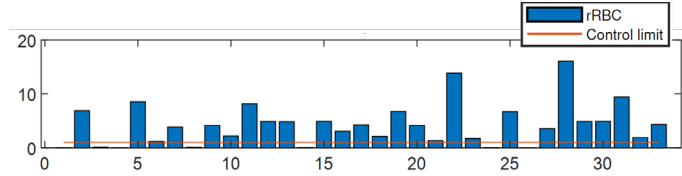
Figure 10. Relative reconstruction-based contribution results for IDV (8)

The DALVR-based monitoring result with the combined index φ is illustrated in Figure 9. It is shown that, from the 176th sample, the disturbance is detected by the combined index φ. As is shown in Figure 10, the corresponding relative RBC identifies many faulty variables due to the interrelations among variables, including Variables 3, 5-7, 8-13, 15-23, 25, 27-33. To further accurately locate the root causes, these variables are selected as the potential contributing candidates in the subsequent Granger causal analysis.
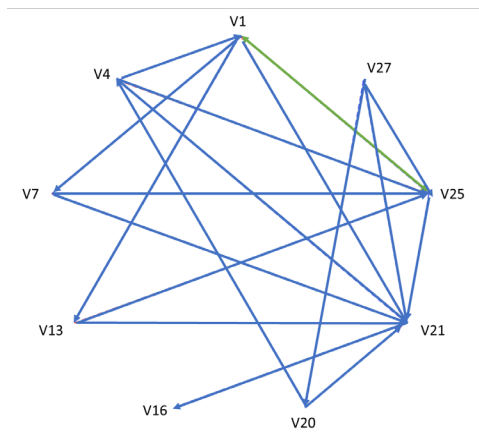
Figure 11. Time-domain Granger causality map for IDV (8)

Figure 11 shows the TDGC map of the 176th sample in IDV (8), where V1, V4, V7, V13, V16, V20-21, V25, and V27 represent Variables 1, 4, 7, 13, 16, 20-21, 25, and 27, respectively. In Figure 11, the green arrows connecting two nodes represent the causality between these two variables, while the blue lines denote the mutual effects between those two variables, failing to provide evidence of their causal relations. Moreover, by reference to the definition of causal flow in a node as the number of outgoing flows minus that of incoming flows,[31] the causal flows of each variable are shown in Figure 12.
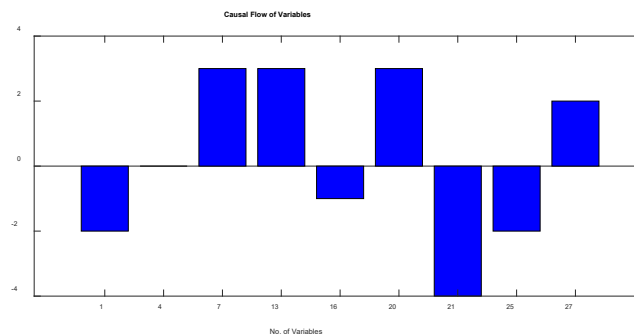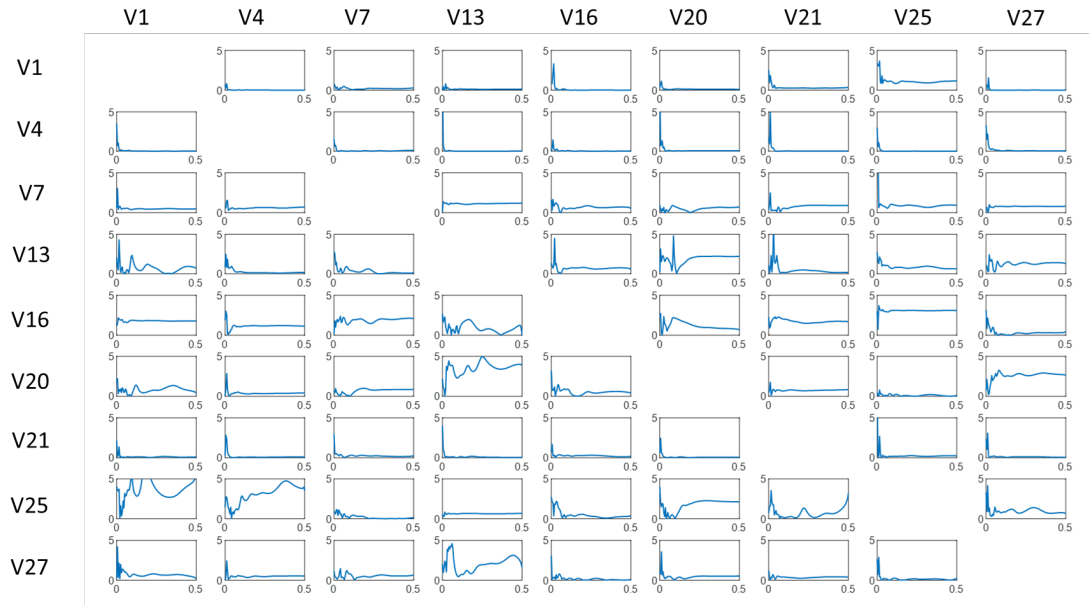

Figure 12. Causal flow results for IDV (8)

Figure 13. Conditional spectral Granger causality results for IDV (8)

It can be seen that Variables 7, 13, 20, and 27 have a positive causal flow, so they are potentially supposed to be the root causes, while Variables 1, 16, 21 and 25 have a negative causal flow, indicating that they are less likely to be the faulty variables. However, the TDGC map fails to provide further information on their causal relations. Thus, CSGC is further applied on these contributing variables. Figure 13 depicts the CSGC graph for these variables of the 176th sample. As illustrated in Figure 13, the subplots under Column V7 on Row V13, V20, and V27 do not have strong peak, and so do the subplots under Column V27 on Row V7, V13, and V20. Thus, V7 and V27 are not considered as the root cause. The subplot under Column V13 on Row V20 has strong variations, while the fluctuations of the subplot under Column V20 on Row V13 are smaller, which indicates that V13 has larger causal effect than V20. Therefore, it is concluded that Variable 13 is the root cause of IDV (8).

6. Conclusions

In this work, CKLVR is proposed by incorporating the concurrent modeling of process and quality variables with KLVR algorithm, which improves the nonlinear modeling performance and enhances the comprehensiveness of data exploitation. After performing CKLVR, the process and quality data are concurrently projected into five subspaces, and their corresponding indices designed for comprehensive process monitoring. In addition, DALVR method is developed by modeling both dynamic cross-correlations and dynamic auto-correlations simultaneously, leading to the enhancement of ability to handle dynamic relations. On this basis, a dynamic comprehensive modeling and monitoring framework based on DALVR is proposed with subsequent decompositions, while a DALVR-based root cause identification and causal analysis scheme is put forward as well for better fault diagnosis. The superiority of the aforementioned proposed methods is shown by case studies on simulation and industrial datasets.

References

Alauddin, M., Khan, F., Imtiaz, S. and Ahmed, S., A variable mosquito flying optimization-based hybrid artificial neural network model for the alarm tuning of process fault detection systems, *Process Safety Progress*, 2020 vol. 39, pp. e12122, 2020.

Alcala, C. F. and Qin, S. J., Reconstruction-based contribution for process monitoring, *Automatica*, vol. 45, no. 7, pp. 1593-1600, 2009.

Bounoua, W. and Bakdi, A., Fault detection and diagnosis of nonlinear dynamical processes through correlation dimension and fractal analysis based dynamic kernel PCA, *Chemical Engineering Science*, vol. 229, pp. 116099, 2021.

Braun, B., Castillo, I., Joswiak, M., Peng, Y., Rendall, R., Schmidt, A., Wang, Z., Chiang L. and Colegrove, B., Data science challenges in chemical manufacturing, *Proceedings of 21st IFAC World Congress*, Berlin, Germany, July 12-17, 2020, pp. 11496-11499.

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R. and Bressler, S. L., Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality, *Proceedings of the National Academy of Sciences*, vol. 101, no. 26, pp. 9849-9854, 2004.

Camps-Valls, G., Martín-Guerrero, J. D., Rojo-Alvarez, J. L. and Soria-Olivas, E., Fuzzy sigmoid kernel for support vector classifiers, *Neurocomputing*, vol. 62, pp. 501-506, 2004.

Chen, Q. and Wang, Y., Key-performance-indicator-related state monitoring based on kernel canonical correlation analysis, *Control Engineering Practice*, vol. 107, pp. 104692, 2021.

Choi, S. W., Lee, C., Lee, J. M., Park, J. H. and Lee, I. B, Fault detection and identification of nonlinear processes based on kernel PCA, *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 1, pp. 55-67, 2005.

Claessen, J., Molini, A., Martens, B., Detto, M., Demuzere, M. and Miralles, D. G., Global biosphere–climate interaction: a causal appraisal of observations and models over multiple temporal scales, *Biogeosciences*, vol. 16, no. 24, pp. 4851-4874, 2019.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C. and Depaulis, A., Identifying neural drivers with functional MRI: an electrophysiological validation, *PLoS Biology*, vol. 6, no. 12, pp. e315, 2008.

Davis J. and Goadrich M., The relationship between Precision-Recall and ROC curves, *Proceedings of 23rd International Conference on Machine Learning*, Pittsburgh, Pennsylvania, USA, Jun 25-29, 2006, pp. 233-240.

Dhamala, M., Rangarajan, G. and Ding, M., Estimating Granger causality from Fourier and wavelet transforms of time series data, *Physical Review Letters*, vol. 100, no. 1, pp. 018701, 2008.

Dong, Y. and Qin, S. J., Dynamic-inner partial least squares for dynamic data modeling, *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 117-122, 2015.

Dong, Y. and Qin, S. J., A novel dynamic PCA algorithm for dynamic data modeling and process monitoring, *Journal of Process Control*, vol. 67, pp. 1-11, 2018.

Dong, Y. and Qin, S. J., Dynamic-inner canonical correlation and causality analysis for high dimensional time series data, *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 476-481, 2018.

Downs, J. J. and Vogel, E. F., A plant-wide industrial process control problem, *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245-255, 1993.

Dunia, R. and Qin, S. J., Subspace approach to multidimensional fault identification and reconstruction, *AICHE Journal*, 44(8), 1813-1831, 1998.

Fomin, F. V. and Thilikos, D. M., Fast parameterized algorithms for graphs on surfaces: Linear kernel and exponential speed-up, *Proceedings of 31st International Colloquium on Automata, Languages, and Programming*, Turku, Finland, July 12-16, 2004, pp. 581-592.

Friston, K. J., Harrison, L. and Penny, W., Dynamic causal modelling, *Neuroimage*, vol. 19, no. 4, pp. 1273-1302, 2003.

Geladi, P. and Kowalski B.R., Partial least-squares regression: a tutorial, *Analytica Chimica Acta*, vol. 185, pp. 1-7, 1986.

Gertler, J., Li, W., Huang, Y. and McAvoy, T., Isolation enhanced principal component analysis, *AIChE Journal*, vol. 45, no. 2, pp. 323-334, 1999.

Geweke, J., Measurement of linear dependence and feedback between multiple time series, *Journal of the American statistical association*, vol. 77, no. 378, pp. 304-313, 1982.

Gillo, M.W., Shelly, M.W., Predictive modeling of multivariable and multivariate data, *Journal of the American Statistical Association*, vol. 69, no. 347, pp. 646-653, 1974.

Goebel, R., Roebroeck, A., Kim, D. S. and Formisano, E., Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping, *Magnetic Resonance Imaging*, vol. 21, no. 10, pp. 1251-1261, 2003.

Granger, C. W., Investigating causal relations by econometric models and cross-spectral methods, *Econometrica: Journal of the Econometric Society*, pp. 424-438, 1969.

Hardoon, D. R., Szedmak, S. and Shawe-Taylor, J., Canonical correlation analysis: An overview with application to learning methods, *Neural computation*, 16(12), 2639-2664, 2004.

He, B., Yang, X., Chen, T. and Zhang, J., Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach, *Journal of Process Control*, vol. 22, no. 7, pp. 1228-1236, 2012.

Hu, C., Luo, J., Kong, X. and Feng, X., Novel fault subspace extraction methods for the reconstruction-based fault diagnosis, *Journal of Process Control*, vol. 105, pp. 129-140, 2021.

Kamat, A., Makled, B., Norfleet, J., Intes, X., Dutta, A. and De, S., November. Brain network effects related to physical and virtual surgical training revealed by Granger causality, *Proceedings of 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, November 1-5, 2021, pp. 1014-1017.

La Fond, T. and Neville, J., Randomization tests for distinguishing social influence and homophily effects, Proceedings of the 19th International Conference on World Wide Web, Raleigh, USA, April 26-30, 2010, pp. 601-610.

Lai, P. L. and Fyfe, C., Kernel and nonlinear canonical correlation analysis, *International Journal of Neural Systems*, vol. 10, no. 5, pp. 365-377, 2000.

Li, G., Joe Qin, S. and Zhou, D., Output relevant fault reconstruction and fault subspace extraction in total projection to latent structures model, *Industrial & Engineering Chemistry Research*, vol. 49, no. 19, pp. 9175-9183, 2010.

Li, Y., Lei, M., Cui, W., Guo, Y. and Wei, H. L., A parametric time-frequency conditional Granger causality method using ultra-regularized orthogonal least squares and multiwavelets for dynamic connectivity analysis in EEGs, *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 12, pp. 3509-3525, 2019.

Lindner, B., Auret, L., Bauer, M. and Groenewald, J. W., Comparative analysis of Granger causality and transfer entropy to present a decision flow for the application of oscillation diagnosis, *Journal of Process Control*, vol. 79, pp. 72-84, 2019.

Liu, J., Hu, Y. and Yang, S., A SVM-based framework for fault detection in high-speed trains, *Measurement*, vol. 172, pp. 108779, 2021.

Liu, Q., Zhu, Q., Qin, S. J. and Chai, T., Dynamic concurrent kernel CCA for strip-thickness relevant fault diagnosis of continuous annealing processes, *Journal of Process Control*, vol. 67, pp. 12-22, 2018.

Liu, Y., Taniguchi, M. and Ombao, H., Statistical Inference for Local Granger Causality, *arXiv preprint arXiv:2103.00209*, 2021.

Mnassri, B. and Ouladsine, M., Reconstruction-based contribution approaches for improved fault diagnosis using principal component analysis, *Journal of Process Control*, vol. 33, pp. 60-76, 2015.

Mou, T., Li, S. and Zou, Y., Enhancing comprehensive contribution plot for fault isolation of distributed systems, *Proceedings of the 2021 40th Chinese Control Conference (CCC)*, Shanghai, China, July 26-28, 2021, pp. 4479-4484.

Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., Brismar, T. and Müller, K. R., Robustly estimating the flow direction of information in complex physical systems*, Physical Review Letters*, vol. 100, no. 23, 234101, 2008.

Patrick, E. A. and Fischer III, F. P., A generalized k-nearest neighbor rule, *Information and Control*, vol. 16, no. 2, pp. 128-152, 1970.

Peerally, M.F., Carr, S., Waring, J. and Dixon-Woods, M., The problem with root cause analysis, *BMJ Quality & Safety*, vol. 26, no. 5, pp. 417-422, 2017.

Qin, S. J., Statistical process monitoring: basics and beyond, *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 17, no. 8-9, pp. 480-502, 2003.

Qin, S. J. and McAvoy, T. J., Nonlinear PLS modeling using neural networks, *Computers & Chemical Engineering*, vol. 16, no. 4, pp. 379-391, 1992.

Ramaker, H. J., Van Sprang, E. N., Westerhuis, J. A., Gurden, S. P., Smilde, A. K. and Van Der Meulen, F. H., Performance assessment and improvement of control charts for statistical batch process monitoring, *Statistica Neerlandica*, vol. 60, no. 3, pp. 339-360, 2006.

Rieth, C. A., Amsel, B. D., Tran, R. and Cook, M. B., Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation, Harvard Dataverse, V1, 2017.

Rosipal, R. and Trejo, L. J., Kernel partial least squares regression in reproducing kernel hilbert space, *Journal of Machine Learning Research*, vol. 2, pp. 97-123, 2001.

Ryali, S., Supekar, K., Chen, T. and Menon, V., Multivariate dynamical systems models for estimating causal interactions in fMRI, *Neuroimage*, vol. 54, no. 2, pp. 807-823, 2011.

Shang, C., Yang, F., Huang, D. and Lyu, W., Data-driven soft sensor development based on deep learning technique, *Journal of Process Control*, vol. 24, no. 3, pp. 223-233, 2014.

Shashua, A., Introduction to machine learning: Class notes 67577, *arXiv preprint arXiv:0904.3664*, 2009.

Shen, G., Wang, P., Hu, K. and Ye, Q., Fault Root Cause Diagnosis Method Based on Recurrent Neural Network and Granger Causality, *Proceedings of 2021 CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes (SAFEPROCESS)*, Chengdu, China, December 17-18, 2021, pp. 1-6.

Valdés-Sosa, P. A., Sánchez-Bornot, J. M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L. and Canales-Rodríguez, E., Estimating brain functional connectivity with sparse multivariate

autoregression, *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 969-981, 2005.

Vicente, R., Wibral, M., Lindner, M. and Pipa, G., Transfer entropy — a model-free measure of effective connectivity for the neurosciences, *Journal of Computational Neuroscience*, vol. 30, no. 1, pp. 45-67, 2011.

Wang, G., Li, J., Sun, C. and Jiao, J., Least squares and contribution plot based approach for quality-related process monitoring, *IEEE Access*, vol. 6, pp. 54158-54166, 2018.

Wilson, G. T., The factorization of matricial spectral densities, *SIAM Journal on Applied Mathematics*, vol. 23, no. 4, pp. 420-426, 1972.

Xu, B. and Zhu, Q., Concurrent auto-regressive latent variable model for dynamic anomaly detection, *Journal of Process Control*, vol. 108, pp. 1-11, 2021.

Yu, W. and Zhao, C., Robust monitoring and fault isolation of nonlinear industrial processes using denoising autoencoder and elastic net, *IEEE Transactions on Control Systems Technology*, vol. 28, no. 3, pp. 1083-1091, 2019.

Yuan, T. and Qin, S. J., Root cause diagnosis of plant-wide oscillations using Granger causality, *Journal of Process Control*, vol. 24, no. 2, pp. 450-459, 2014.

Zhang, Y. and Hu, Z, Multivariate process monitoring and analysis based on multi-scale KPLS, *Chemical Engineering Research and Design*, vol. 89, no. 12, pp. 2667-2678, 2011.

Zhou, D., Li, G., and Qin, S. J., Total projection to latent structures for process monitoring, *AIChE Journal*, vol. 56, no. 1, pp. 168-178, 2010.

Zhou, P., Zhang, R., Liang, M., Fu, J., Wang, H. and Chai, T., Fault identification for quality monitoring of molten iron in blast furnace ironmaking based on KPLS with improved contribution rate, *Control Engineering Practice*, vol. 97, pp. 104354, 2020.

Zhu, Q., Latent variable regression for supervised modeling and monitoring, *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 3, pp. 800-811, 2020.

Zhu, Q., Liu, Q., and Qin, S. J., Concurrent quality and process monitoring with canonical correlation analysis, *Journal of Process Control*, vol. 60, pp. 95-103, 2017.

Zhu, Q., Qin, S. J. and Dong, Y., Dynamic latent variable regression for inferential sensor modeling and monitoring. *Computers & Chemical Engineering*, vol. 137, pp. 106809, 2020.

Biographies

Haitian Zhang is a Ph.D. Candidate in Chemical Engineering (University of Waterloo, CA). Her research focuses on process monitoring and fault diagnosis based on machine learning algorithms in the field of process systems engineering. Prior to her doctoral study, she secured her Bachelor's and Master's degrees in Chemical Engineering from Xi'an Jiaotong University. She had two PRES conference papers and three journal papers in her masteral program. During the period of her doctoral study, she also published two papers related to multivariate statistical analysis.

Qinqin Zhu is an Assistant Professor in the department of Chemical Engineering (University of Waterloo, CA). She is also a faculty member at the Waterloo Artificial Intelligence Institute (Waterloo.AI), Waterloo Institute for Sustainable Energy (WISE) and Waterloo Institute for Nanotechnology (WIN). She received her PhD degree from the Chemical Engineering department at the University of Southern California, advised by Prof. Joe Qin. Prior to UW, she worked as a senior research scientist at Facebook Inc. in the United States. Her research mainly focuses on developing advanced statistical machine learning methods, process data analytics techniques and optimization algorithms in the era of big data with applications to statistical process monitoring and fault diagnosis. Her research addresses theoretical challenges and problems of practical importance in the area of process systems engineering. By leveraging the power of mathematical modeling and optimization, her group strives to develop advanced multivariate statistical analysis algorithms that enhance decision making in complex engineering systems.

Ali Elkamel is a Full Professor of Chemical Engineering (University of Waterloo, CA). He is also cross appointed in Systems Design Engineering. He holds a BSc in Chemical Engineering and BSc in Mathematics from Colorado School of Mines, MSc in Chemical Engineering from the University of Colorado, and PhD in Chemical Engineering from Purdue University. His specific research interests are in computer-aided modeling, optimization, and simulation with applications to energy planning, sustainable operations, and product design. His activities include teaching graduate and undergraduate courses, supervising post doctorate and research associates, and participation in both university and professional societal activities. He is also engaged in initiating and leading academic and industrial teams, establishing international and regional research collaboration programs with industrial partners, national laboratories,

and international research institutes. He supervised over 120 graduate students (of which 47 are PhDs) and more than 45 post-doctoral fellows/research associates. He has been funded for several research projects from government and industry. Among his accomplishments are the Research Excellence Award, the Excellence in Graduate Supervision Award, the Outstanding Faculty Award, and IEOM Awards. He has more than 425 journal articles, 175 proceedings, 50 book chapters, and has been an invited speaker on numerous occasions at academic institutions throughout the world. He is also a co-author of six books.