12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Slow-Varying Dynamics Supervised Data Decomposition for Batch Process Quality Prediction

Yan Qin * **Xiaodong Xu** **,*** **Xunyuan Yin** ****

* *Singapore University of Technology and Design, 8 Somapah Road, 487372, Singapore (e-mail: yan_qin@sutd.edu.sg)*
** *School of Automation, Central South University, Changsha 410083, China*
*** *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, Canada (e-mail: xx1@ualberta.ca)*
**** *School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, 62 Nanyang Dr, 637459, Singapore (e-mail: xunyuan.yin@ntu.edu.sg)*

Abstract: Ensuring high prediction accuracy is essential for maintaining high quality of products in batch processes, given their inherent multi-phase characteristics and dynamic variations in real-world applications. Identifying quality-relevant process variations is crucial to address these challenges and produce interpretable and accurate predictions. This work aims to uncover critical quality-relevant process variables from raw measurements collected from batch processes. The proposed method consists of two key components. First, a dynamic subspace is designed for batch processes to extract the slow-varying features that are relevant to the quality index. Second, the quality-relevant features have been employed to achieve the reliable prediction of the performance index. Through the simulated experiment on a concentration batch production process, the proposed method is illustrated.

Keywords: Quality prediction, data decomposition, slow feature analysis, batch processes.

1. INTRODUCTION

The increasing demands for customized products in modern manufacturing industries have greatly boosted the success of batch manufacturing [1-3]. Batch manufacturing produces low-volume and high-value-added products with minor to major differences in features and/or specifications to meet the dynamically changing demands of the market. While batch manufacturing provides great flexibility in customizing the products, it also faces critical issues that have limited wider adoption of this type of production method. In particular, there can be frequent changes in the operating conditions due to variations in the specifications of the items being produced, which pose great challenges to the maintenance of product quality and safe operation. Reliable prediction of product indices through quality-relevant data decomposition is a promising way to confront process failures and ensure operating safety and production consistency.

With the rapid progress and advancement in data sensing and storage through Industrial Internet of Things Technologies [4,5], extensive process data has become accessible. This provides an opportunity for modern batch manufacturing industries to adopt data-driven methods for product quality assessment. The use of data can reduce the dependency on first-principles knowledge for process monitoring and fault detection, bypassing physical-based process monitoring [6,7]. This is advantageous in practical applications, as acquiring precise first-principles knowledge for complex processes can be challenging [8]. Data-driven batch product prediction can be traced back to the 1990s. Due to its interpretation ability, partial least square (PLS) [9] demonstrates advantages and benefits in regressing raw measurements on the performance indices. Nomikos et al. [10] applied PLS to extract the low-dimension primary features from the high-dimension raw data to the batch process for the first time. Specifically, in [10], a multiway partial least squares (MPLS) was proposed to handle the three-dimensional structure of batch data. This approach treats each batch data as a sample with two dimensions. However, the premise of MPLS assumes that only one operation mode exists during the entire batch manufacturing process. This assumption may not hold for complex batch processes with multiple physical stages, in which multiple data evolution behaviors are observed. In [11], Lu et al. introduced the concept of multi-phase partitioning for batch processes, wherein a relevant model was developed in each segmented phase. Consequently, a phase division strategy, which recorded the measurements of variables over all batches at a specific time within the batch, was developed. However, the adjacent time-slice matrices will likely be classified into different phases, which

⋆ This research is supported by Ministry of Education, Singapore, under its Academic Research Fund Tier 1 (RS15/21).

[1] Corresponding author: Xunyuan Yin. Tel: (+65)83413571. Email: xunyuan.yin@ntu.edu.sg

leads to discontinuous temporal characteristics within the same phase.

To address this challenge, quality-relevant phase partitioning and product quality prediction have received much research attention. The core issue is how to establish the linkage between the quality performance and the process variations to enable reasonable phase partition. Lu et al. [12] obtained phase partition results by clustering the regression relationship in each time-slice data matrix. However, this method suffers from the challenges associated with the clustering-based division algorithm, such as rigid partition and uneven allocation of the sampling instants. To overcome this issue, Zhao et al. [13] proposed a sequential phase partition method to ensure that similar regression correlations are assigned to the same phase. Through iteratively comparing the historical and present regression correlations, sequential partitions considering quality interpretation were achieved when the change of regression correlation was detected. In [13], the transition zone among two adjacent phases is further distinguished and handled separately in the subsequent modeling process. The approaches mentioned above focused on obtaining more reasonable phase partition results by finding a static subspace from process variations. However, the primary issue lies in that the derived static subspace may fail to reflect the temporal dynamics over time in each phase. Specifically, the slow-varying dynamics exist in each phase of the operation of the physical asset. For instance, the part weight of plastic products in the injection molding process slowly accumulates with time. In the penicillin fermentation process, the density of penicillin slowly increases over time within a single batch. It is logical to deduce that the gradual change in the performance index results from the inherent characteristics of the process operation, which exhibits a slow-varying nature. Therefore, the analysis of slow-varying quality relevant subspace has the potential to enhance the quality prediction accuracy.

Based on the above observations, this work proposes a feature decomposition guided by slow-varying batch process dynamics to find the most quality-relevant process variations from raw measurements. The proposed approach consists of two key components. First, a dynamic subspace is exploited to extract the slow-varying features that are closely related to the quality index of the batch process. Second, both the static and the dynamic information of the extracted features are employed to achieve the reliable prediction of the performance index.

The remainder of the paper is organized as follows. Section 2 describes the structure of the batch process data. Section 3 formulates the proposed quality prediction model based on data decomposition. Section 4 presents the simulation results. Section 5 concludes this work.

2. BATCH DATA DESCRIPTION AND DATA PREPARATION

In this section, we describe the characteristics of batch manufacturing processes and the process data for both single-batch and multiple-batch case scenarios. The basic data unit with the time-slice matrix is also illustrated.

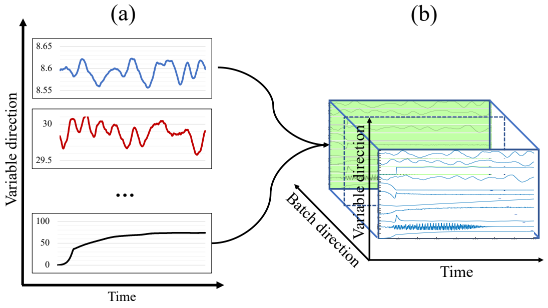


Fig. 1. Data description of the batch process data with (a) a single batch and (b) a three-dimensional presentation with dozens of batches.

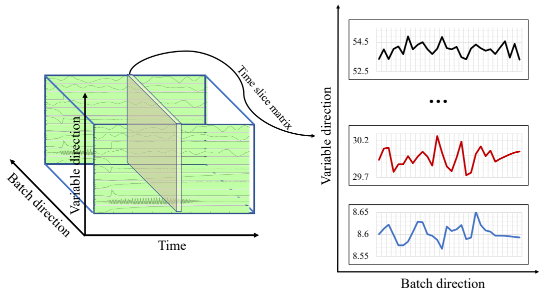


Fig. 2. Data unfolding of the three-dimensional data matrix to a series of two-dimensional time-slice data matrices.

Batch manufacturing can produce a batch of identical products with the same quality over a finite period [14]. In the subsequent batches, more products of the same or different specifications can be produced in a repetitive manner. The typical data structure for batch manufacturing is described as follows:

(1) Data structure of a single batch: Denote by J the number of variables that are measured, a vector $\mathbf{x}_k \in \mathbb{R}^{1\times J}$ is obtained at sampling instant $k \in [1, K]$ where K is the total number of sampling instants. Within a batch, the samples over all the K sampling instants form a two-dimensional data matrix with dimensions $K \times J$, as shown in Fig. 1 (a).
(2) Data structure of multiple batches: Consider the case when data for a series of production batches are collected. Denote the two-dimensional data for the ith batch by $\mathbf{X}_i \in \mathbb{R}^{K\times J}$. Stack of each batch data $\mathbf{X}_i$ along the batch direction creates a three-dimensional data matrix, which is denoted by $\mathbf{X} \in \mathbb{R}^{I\times K\times J}$. The order of the variables and the chronological order of the samples for each batch remain unchanged. A visual illustration of the three-dimensional data matrix is presented in Fig. 1 (b).
(3) Time-slice data matrix: This concept is introduced to elucidate changes in the correlation of variables across various time instants. Derived from the two-dimensional matrix from $\mathbf{X}$, the time-slice matrix $\mathbf{T}_k$ is defined as the measurements of all variables over I batches collected at sampling instant $k \in [1, K]$. This matrix contains information about the distribution of the data over different batches. The relationship between the time-slice data matrix for time instant k and the three-dimensional data matrix, and the

information contained in each time-slice data matrix are delineated in Fig. 2.

In addition to the raw measurements collected from the production process, the quality-relevant performance indices can be obtained at the end of each batch, for example, the part weight in injection molding. Typical batch processes exhibit multi-phase characteristics, which signifies the variations in variable correlations from one time segment to another.

The influence of the variable scale on variable correlation is confronted through data normalization. Performing Z-score normalization on $\mathbf{T}_k$ with zero mean and unit variation gets the normalized time-slices matrix, which is still denoted as $\mathbf{T}_k$ for clarity. Correspondingly, the normalization information with mean and standard variation at each sampling time are stored as $\mathbf{u}_k$ and δ_k, respectively.

To achieve quality-relevant data decomposition and regression, it is relevant to consider the overall evolution of the variable correlations along the batch direction. First, the data matrix in a batch $\mathbf{X}_i$ is normalized based on the information $\mathbf{u}_k$ and δ_k at each time, and the normalized matrix is denoted by $\tilde{\mathbf{X}}_i$. Then, by aggregating all the normalized batches along the variable direction, we obtain a data matrix $\tilde{\mathbf{X}} \in \mathbb{R}^{IK \times J}$. The corresponding quality measurements are recorded as $\mathbf{Y} \in \mathbb{R}^{I \times 1}$.

Additionally, the presence of slow-varying dynamics necessitates an exploration of temporal variations between different time instants. Consequently, $\Delta\mathbf{X}_i(k) = \mathbf{X}_i(k) - \mathbf{X}_i(k-1)$ for $k \in [2, K]$ is computed to represent the differences in $\mathbf{X}_i$, where $\mathbf{X}_i(k)$ denotes the kth row of $\mathbf{X}_i$. The aggregation of each $\Delta\mathbf{X}_i$ along the variable direction creates a comprehensive matrix $\Delta\mathbf{X} \in \mathbb{R}^{I(K-1) \times J}$.

3. BATCH PROCESS QUALITY PREDICTION SUPERVISED BY SLOW VARYING DYNAMICS

The proposed quality prediction method is introduced in this section. This method consists of regression-based data decomposition and quality prediction based on the quality-relevant subspaces.

3.1 Quality-relevant data decomposition considering slow-varying dynamics

This part details the prediction model with the quality-relevant slow feature analysis for batch process in each phase and outputs the regression relationship for online application when incoming data are available.

The essential is to find the quality-relevant subspace $\boldsymbol{\omega}$, covering the process variations that are strong prediction interpretability, low slowness, and large variation. Prediction interpretability requires that the features are highly relevant to the performance indices, i.e., *max* $L_1(\boldsymbol{\omega}) = \frac{(\tilde{\mathbf{X}}\boldsymbol{\omega})^T\mathbf{Y}\mathbf{Y}^T(\tilde{\mathbf{X}}\boldsymbol{\omega})}{\sqrt{(\tilde{\mathbf{X}}\boldsymbol{\omega})^T(\tilde{\mathbf{X}}\boldsymbol{\omega})}}$. Slowness focuses on the varying speed of the features, which is assumed to be as slow as possible, i.e., *min* $L_2(\boldsymbol{\omega}) = (\Delta\mathbf{X}\boldsymbol{\omega})^T(\Delta\mathbf{X}\boldsymbol{\omega})$. The process interpretability calls for more process variations, i.e., *max* $L_3(\boldsymbol{\omega}) = \sqrt{(\tilde{\mathbf{X}}\boldsymbol{\omega})^T(\tilde{\mathbf{X}}\boldsymbol{\omega})}$.

The desired subspace $\boldsymbol{\omega}$ is ensured through the joint optimization as follows,

$$\max L(\boldsymbol{\omega}) = \frac{L_1 \cdot L_3}{L_2} = \frac{(\tilde{\mathbf{X}}\boldsymbol{\omega})^T\mathbf{Y}\mathbf{Y}^T(\tilde{\mathbf{X}}\boldsymbol{\omega})}{(\Delta\mathbf{X}\boldsymbol{\omega})^T(\Delta\mathbf{X}\boldsymbol{\omega})} \tag{1}$$

Given the constraint as $(\Delta\mathbf{X}\boldsymbol{\omega})^T(\Delta\mathbf{X}\boldsymbol{\omega}) = 1$, the Lagrange multiplier method is adopted to solve Eq. 1, which is arranged below,

$$\max L(\boldsymbol{\omega}) = \boldsymbol{\omega}^T\tilde{\mathbf{X}}^T\mathbf{Y}\mathbf{Y}^T\tilde{\mathbf{X}}\boldsymbol{\omega} - \lambda\left(\boldsymbol{\omega}^T\Delta\mathbf{X}^T\Delta\mathbf{X}\boldsymbol{\omega} - 1\right) \tag{2}$$

where λ is the Lagrange multiplier.

3.2 Solving Procedure

Through performing derivation on Eq. 2 with respect to $\boldsymbol{\omega}$, we have the following equation,

$$\frac{\partial L(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} = 2\tilde{\mathbf{X}}^\top\mathbf{Y}\mathbf{Y}^\top\tilde{\mathbf{X}}\boldsymbol{\omega} - 2\lambda\Delta\mathbf{X}^T\Delta\mathbf{X}\boldsymbol{\omega} \tag{3}$$

Setting Eq. 3 to zero yields,

$$\tilde{\mathbf{X}}^T\mathbf{Y}\mathbf{Y}^T\tilde{\mathbf{X}}\boldsymbol{\omega} = \lambda\Delta\mathbf{X}^T\Delta\mathbf{X}\boldsymbol{\omega} \tag{4}$$

Multiplying both sides of Eq. 3 by $\boldsymbol{\omega}^T$ contributes to derive the following equation,

$$\boldsymbol{\omega}^T\tilde{\mathbf{X}}^T\mathbf{Y}\mathbf{Y}^T\tilde{\mathbf{X}}\boldsymbol{\omega} = \lambda\boldsymbol{\omega}^T(\Delta\mathbf{X}^T\Delta\mathbf{X})\boldsymbol{\omega} = \lambda \tag{5}$$

Before solving Eq. 4, a whitening transformation is employed to eliminate the cross-correlation among the first-order different matrix. Assuming that the whitening data matrix is $\mathbf{H}$, data matrix $\tilde{\mathbf{X}}$ will be transformed to $\mathbf{z} = \mathbf{H}^{\mathrm{T}}\Delta\mathbf{X}^{\mathrm{T}}$. To ensure that $\boldsymbol{z}^T = \mathbf{H}^T\Delta\mathbf{X}^T\Delta\mathbf{X}\mathbf{H} = \mathbf{1}$, the $\mathbf{H}$ is defined as $\boldsymbol{\Lambda}^{-1/2}\mathbf{U}^T$, where $\Delta\mathbf{X}^T\Delta\mathbf{X} = \mathbf{U}\mathbf{A}\mathbf{U}^T$.

Left multiply the both sides of Eq. 5 by $\mathbf{H}$ and then right multiply the both sides of Eq. 5 by $\mathbf{H}^T$, we have the following equation,

$$\mathbf{H}^T\tilde{\mathbf{X}}^T\mathbf{Y}\mathbf{Y}^T\tilde{\mathbf{X}}\mathbf{H}\boldsymbol{\omega} = \lambda\mathbf{H}^T\Delta\mathbf{X}^T\Delta\mathbf{X}\mathbf{H} = \lambda \tag{6}$$

Eq. 6 is a typical eigenvalue decomposition, yielding the largest eigenvalue with the first eigenvector $\boldsymbol{\omega}$. Correspondingly, the loading coefficients $\mathbf{p}_i$ and q_i for process variables and quality index can be obtained, respectively,

$$\mathbf{p}_i^T = \frac{(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)^T\tilde{\mathbf{X}}_i}{(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)^T(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)} \tag{7}$$

$$q_i = \frac{(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)^T\mathbf{Y}}{(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)^T(\tilde{\mathbf{X}}_i\boldsymbol{\omega}_i)} \tag{8}$$

To avoid introducing redundancy into the extracted latent variables, data deflation is conducted to ensure that the latent variables are orthogonal to each other. To achieve this, the covariance matrix $(\tilde{\mathbf{X}}^T\mathbf{Y})_i$ and $(\Delta\mathbf{X}^T\Delta\mathbf{X})_i$ are updated at each iteration below,

$$\begin{gathered}\Delta\mathbf{X}_{i+1}^T\Delta\mathbf{X}_{i+1} = \left(\mathbf{I} - \boldsymbol{\omega}_i^T\mathbf{p}_i\right)^T\left(\Delta\mathbf{X}^T\Delta\mathbf{X}\right)_i\left(\mathbf{I} - \boldsymbol{\omega}_i^T\mathbf{p}_i\right) \\ (\tilde{\mathbf{X}}^T\mathbf{Y})_{i+1} = (\mathbf{I} - \boldsymbol{\omega}_i\mathbf{p}_i^T)(\tilde{\mathbf{X}}^T\mathbf{Y})_i\end{gathered} \tag{9}$$

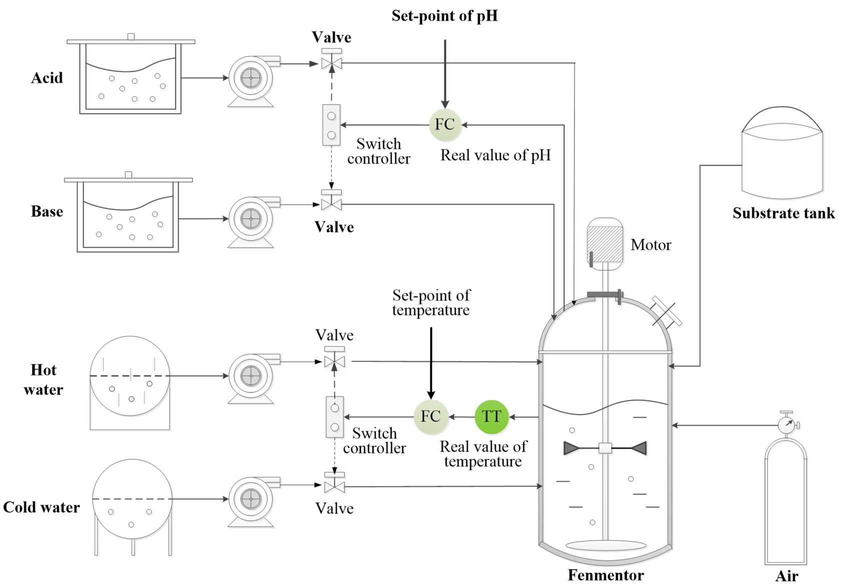


Fig. 3. A schematic of the penicillin fermentation process [15].

Table 1. Variable descriptions of the penicillin fermentation process.

No.	Variable Name	Unit	No.	Variable Name	Unit
1	Aeration rate	L/R	7	Culture volume	L
2	Agitator power	W	8	Carbon dioxide concentration	1
3	Substrate feed rate	L/h	9	pH	1
4	Substrate feed temperature	K	10	Fermentor temperature	K
5	Substrate concentration	g/L	11	Generated heat	kcal
6	Dissolved oxygen concentration	g/L			

Table 2. Variable descriptions of the penicillin fermentation process.

Phase Name	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Duration	1 → 45	46 → 104	105 → 151	152 → 248	249 → 401

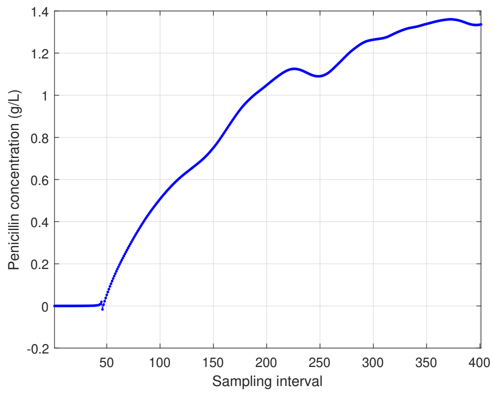


Fig. 4. The trajectory of the penicillin concentration in a batch during the four-stage fermentation process.

Updating $(\tilde{\mathbf{X}}^T\mathbf{Y})_{i+1}$ and $(\Delta\mathbf{X}^T\Delta\mathbf{X})_{i+1}$ into Eq. 6 and recursively repeat Steps to calculate the following terms, including weighting vector $\tilde{\mathbf{X}}_i$ and loading coefficients $\mathbf{p}_i^T$ and q_i until all features are extracted from $\tilde{\mathbf{X}}$.

Through iteratively executing Steps 2 through 5, the weighting matrix $\mathbf{W} = [\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_{J_x}]$ and the loading matrix $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_J]$ are calculated.

Table 3. Performance comparison between the proposed method and its counterparts regarding the index RMSE.

Phase No.	Algorithm Name		
	sub-PLS [12]	LSTM [17]	QSFR (Proposed)
Phase 2	0.0114	0.0031	**0.0029**
Phase 3	0.0657	0.0062	**0.0041**
Phase 4	0.0536	0.0122	**0.0129**
Phase 5	0.0405	0.0185	**0.0162**

3.3 Establishment of the quality prediction model

With data matrices $\mathbf{P}$ and $\mathbf{W}$, the quality-relevant slow features can be expressed as,

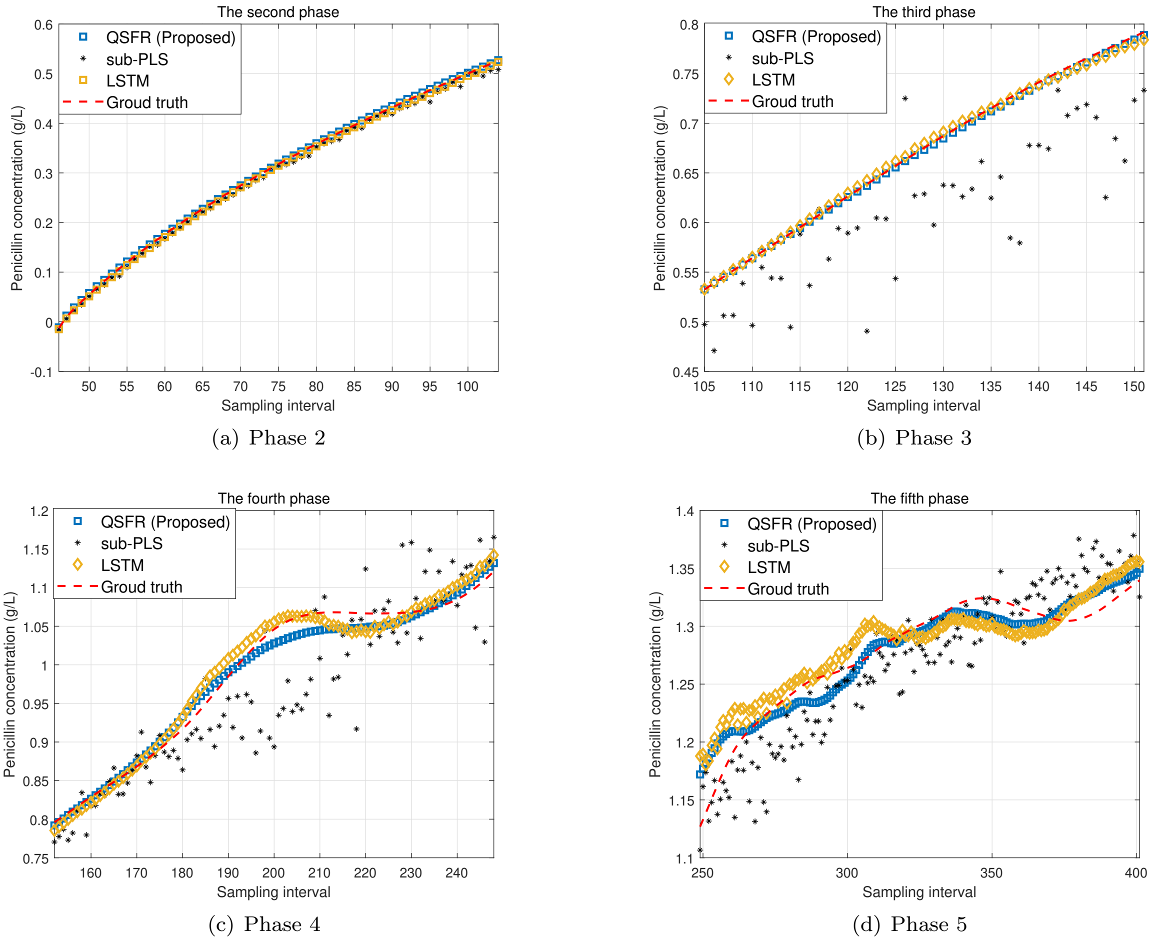
$$\mathbf{S} = \mathbf{X}\mathbf{R}_M \tag{10}$$

where $\mathbf{R} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}$ and $\mathbf{R}_M$ is the first M weighting vectors kept in $\mathbf{R}$.

Finally, the quality-relevant slow feature regression (QSFR) model is completed, and the predicted quality index is computed as,

$$\hat{\mathbf{Y}} = \mathbf{S}\mathbf{q}_M \tag{11}$$

where $\mathbf{q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_J]$ are given according to Eq. 5.



(a) Phase 2

(b) Phase 3

(c) Phase 4

(d) Phase 5

Fig. 5. Performance comparison between the proposed method and its counterparts in (a) Phase 2, (b) Phase 3, (c) Phase 3, and (d) Phase 4.

4. RESULTS AND DISCUSSION

In this section, we utilize the penicillin fermentation process, which is a multi-phase batch process, to illustrate the efficacy of the proposed method. This penicillin fermentation process is a typical multiple-phase batch process.

4.1 Description of penicillin fermentation process and process data

Penicillin is produced via fermentation in a fed-batch manner. Enzymatic hydrolysis is added to enhance the productivity. A schematic of the penicillin fermentation process is presented in Fig. 3. As shown in Fig. 3, this process mainly consists of four operation stages, each of which serves a distinct purpose. In the first stage, the preculture stage promotes the cell densities by introducing microorganisms in batch mode. During this stage, no penicillin is produced. Second, in the exponential growth stage of cells, penicillin starts to be produced concurrently with the continuous feeding of glucose into the process, which aims to sustain an optimal rate of cell growth. The third stage is the stationary stage, in which penicillin is produced at a high rate. In the cell death stage which is the final stage, the cells initiate self-dissolve, which leads to a rapid decline in the production rate. Fig. 4 shows a typical evolution of the penicillin concentration during the four stages, for an illustrating purpose. It is observed that the density of penicillin concentration is almost zero before the 51^{st} sample. A rapid increase in the penicillin concentration starts from the 51^{st} sample and ends around the 300^{th} sample. From the 301^{st} sample, the penicillin concentration increases very slowly, and a slight decline is observed towards the end of the procedure.

The research group at the Illinois Institute of Technology developed a physics-informed simulator (Pensim v1.0) through detailed mechanistic modeling [15]. This simulator is accessible from the website `http://simulator.iit.edu/web/pendownload.html`. The model development involved measuring eleven variables, as listed in Table 1. All batch data are simulated under initial working conditions and set points. With a sampling interval of 1 hour, each batch spans 401 hours, resulting in a data matrix with dimensions 11 × 401. A total of 55 normal batches are generated for analysis: the first 30 batches are used for training, and the the subsequent 15 batches are used for validation, and the remaining ten batches are used for testing.

Based on the phase partition information presented in [16], five sequential phases characterize the penicillin fermentation process, as outlined in Table 2. By integrating with the physical operation stage, the phase division results are accurately derived using data, and the results indicate the starting and ending points of each phase. Notably, these results align with the physical operation but also offer insightful findings, and they provide insights into the subdivision of the third physical stage into Phase 3 and Phase 4. As penicillin production is absent in the first phase, quality prediction and process monitoring are conducted throughout the second phase to the fifth phase.

4.2 Quality performance prediction and comparison

This subsection illustrates the quality prediction performance of the proposed method. By applying the proposed method, four specific regression models, each corresponding to one of the four stages, are developed according to the details presented in Section 3.2. Cross-validation data are used to determine appropriate values for the hyperparameter R, which represents the number of quality-relevant components. Subsequently, the prediction performance is assessed using the testing dataset. Additionally, two benchmark methods, including the sub-PLS [12] and the long short-term memory (LSTM) [17], are employed for comparative analysis.

Both sub-PLS and LSTM are deployed from phase to phase with the same training data, validation data, and testing data of the proposed method for fair comparison. The basic idea of sub-PLS and LSTM are briefly described and their configurations in each phase are specified.

Both sub-PLS and LSTM are applied across phases using the same training, validation, and testing data to ensure a fair comparison. A concise overview of the fundamental concepts of sub-PLS and LSTM is provided, and the specifications of their configurations in each phase are discussed.

(1) sub-PLS: The PLS algorithm is used for regression analysis in each phase. The retained principal components from Phase 2 to Phase 4 are 9, 9, 8, and 8, respectively.
(2) LSTM: Due to the advantage of temporal learning ability, LSTM has been widely adopted in temporal regression issues. Here, a two-layer LSTM network is constructed, and each layer encompasses 100 neurons. The Adam algorithm is adopted for network optimization.

Based on the root mean square error (RMSE), Table 3 summarizes the results in each phase. The proposed method outperforms sub-PLS and LSTM in Phases 2, 3, and 5. In Phase 4, the proposed exhibits comparable performance. Taking the first testing batch as an example, the prediction results of the proposed method and sub-PLS in each phase are illustrated in Fig. 5. In Phases 2 and 3, the proposed method yields estimations that are closely aligned with the ground truth. While estimation errors are observed in Phases 4 and 5, the proposed method outperforms the sub-PLS method.

5. CONCLUSION

To achieve high prediction accuracy, we proposed a feature decomposition guided by slow-varying dynamics. This approach can be used to identify the most quality-relevant process variations from raw measurements. The extraction of a dynamic subspace renders these process variations interpretable and crucial for prediction results. The effectiveness of the proposed method was validated using a penicillin production process. Furthermore, our method demonstrated superiority compared to existing methods, including sub-PLS and long short-term memory networks.

6. REFERENCE

1 Y. Qin, C. Zhao, X. Wang, F. Gao. Subspace decomposition and critical phase selection based cumulative quality analysis for multiphase batch processes. *Chemical Engineering Science*, 166:130–143, 2017.

2 J. Chen, K.-C. Liu. On-line batch process monitoring using dynamic PCA and dynamic PLS models. *Chemical Engineering Science*, 57(1):63-75, 2002.

3 K. Peng, K. Zhang, B. You, J. Dong, Z. Wang. A quality-based nonlinear fault diagnosis framework focusing on industrial multimode batch processes. *IEEE Transactions on Industrial Electronics*, 63(4): 2615–2624, 2016.

4 Q. Jiang, X. Yan, B. Huang. Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes. *Industrial & Engineering Chemistry Research*, 58(29):12899–12912, 2019.

5 A. Arunan, Y. Qin, X. Li, C. Yuen. A federated learning-based industrial health prognostics for heterogeneous edge devices using matched feature extraction, *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2023.3274648.

6 X. Yin, and J. Liu, Distributed output-feedback fault detection and isolation of cascade process networks. *AIChE Journal*, 63(10):4329-4342, 2017.

7 X. Yin, B. Decardi-Nelson, J. Liu. Distributed monitoring of the absorption column of a post-combustion CO_2 capture plant. *Int. J. Adapt. Control Signal Process*, 34:757–776, 2020.

8 X. Yin, Y. Qin, H. Chen, W. Du, J. Liu, B. Huang. Community detection based process decomposition and distributed monitoring for large-scale processes. *AIChE Journal*, 68(11):e17826, 2022.

9 S. Wold, M. Sjöström, L. Eriksson. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2):109–130, 2001.

10 P. Nomikos, J. F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):1361–1375, 1994.

11 N. Lu, F. Gao, F. Wang. Sub-PCA modeling and on-line monitoring strategy for batch processes *AIChE Journal*, 50(1):255–259, 2004.

12 N. Y. Lu, F. Gao. Stage-based process analysis and quality prediction for batch processes. *Industrial & Engineering Chemistry Research*, 44(10):3547—3555, 2005.

13 C. Zhao. A quality-relevant sequential phase partition approach for regression modeling and quality prediction analysis in manufacturing processes. *IEEE Transactions on Automation Science and Engineering*, 11(4):983–991, 2014.

14 Y. Qin, X. Yin. Start-up monitoring for intermittent manufacturing based on hierarchical stationarity analysis. *Chemical Engineering Research and Design*, 185:26–36, 2022.

15 G. Birol, C. Undey, S. J. Parulekar, A. Cinar. A morphologically structured model for penicillin production. *Biotechnology and Bioengineering*, 77(5):538–552, 2002.

16 Y. Qin, C. Zhao, F. Gao. An iterative two-step sequential phase partition (ITSPP) method for batch process modeling and online monitoring. *AIChE Journal*, 62(7):2358–2373, 2016.

17 S. Hochreiter, J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.