IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 16, 2019, accepted May 19, 2019, date of publication May 27, 2019, date of current version July 3, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919167

Phase Identification and Online Monitoring for the Uneven Batch Processes

RUNXIA GUO AND YANCHENG JIN

College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China

Corresponding author: Runxia Guo (rxguoblp@163.com)

This work was supported in part by the National Key R&D Project of China under Grant 2016YFB0502405, in part by the National Natural Science Foundation of China under Grant 61603395 and Grant 51707195, and in part by the Special Program of Talents Development for Excellent Youth Scholars in Tianjin.

ABSTRACT In practice, the batch processes are usually uneven and show significantly different variables' characteristics in different sub-phases. Therefore, it is necessary to divide each batch into several sub-phases separately. In this paper, a moving window-based multiway information increment matrix (MWMIIM) algorithm for the uneven batch processes is proposed for single-batch phase identification and online monitoring by combining the moving window technique with an information increment matrix (IIM) algorithm. Similar to the IIM algorithm, the MWMIIM algorithm captures variables' correlation changes by calculating the increment matrix between adjacent covariance matrices, which does not need to extract feature from data or carry out complicated matrix decomposition, thus improving the computational efficiency. Besides, the influence of several vital parameters on the phase identification performance is discussed in detail. After phase identification, the partition points of each sub-phase need not to be strictly aligned. Furthermore, a batch process is divided into three types of regions, namely common region, transition region, and end region. Next, fine modeling and online monitoring strategies are adopted in different regions separately. The comparative experiment is conducted by the window-based stepwise sequential phase partition method for nonlinear uneven batch processes (WNSSPP-U). A practical application on batch processes, namely aircraft steering gear system fault diagnosis experiment, is given to confirm the feasibility and effectiveness of the proposed method.

INDEX TERMS Phase identification, online monitoring, uneven batch processes, moving window-based multiway information increment matrix (MWMIIM), fine modeling.

I. INTRODUCTION

Batch processes are an important mode of production in modern industries, and they are widely applied in fine chemical, pharmaceutical and food industries. Therefore, it is of great significance to realize the real-time fault diagnosis for batch processes by means of fine modeling and precise monitoring, so as to improve the safety and reliability of the operation processes. Batch processes possess many complex process characteristics. Among all the characteristics, multi-phase is the most prominent one, which represents different variables' characteristics in different sub-phases. Therefore, accurate phase partition is a prerequisite for fine modeling and precise monitoring.

Recently, various methods have been proposed for accurate phase partition. Since the prior process knowledge is hard to obtain, multivariate statistical analysis methods that only require to process historical data have attracted considerable attention. Among them, MPCA and MPLS proposed by Nomikos and Macgregor [1], [2] extend the traditional multivariate statistical analysis methods from continuous processes to batch processes. Lu *et al.* [3] developed a phase-based sub-PCA and sub-PLS models based on the improved K-means clustering algorithm. Additionally, Li *et al.* [4] introduced a phase partition method based on the Gaussian mixture model (GMM), in which the posterior probability is used to classify the operation phases. However, the methods state above do not take into account the time sequence of operation phases. Consequently, original phase partition results in [3], [4] may contain some outliers that break the continuity of the operation phases. Mamede *et al.* [5] divided the complete principal component analysis (PCA) model into several independent sub-models to improve the prediction accuracy of phase partition points. Besides,

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

a multiphase (MP) algorithm is developed to monitor batch processes. However, this approach is time-consuming because it requires more than one chronological search across all sampling points to find the appropriate partition points. Considering the effect of phase partition on monitoring performance, Zhao and Sun [6] and Qin *et al.* [7] proposed a stepwise sequential phase partition (SSPP) algorithm which automatically determines segments based on time sequence. In addition, a soft transition multiple PCA (STMPCA) method is proposed by Zhao and Gao [8] and Zhang *et al.* [9] to detect transitions between different sub-phases. Guo *et al.* [10] introduced an innovative algorithm, namely multiway information increment matrix (MIIM), which directly extracts effective information from the covariance matrix. Moreover, the algorithm can separate the process into several different phases based on the accurate capture of variables' correlation changes. However, the aforementioned methods assume that not only do all batches have strictly equal duration, but also the key process events in all batches overlap at the same time interval. Nevertheless, the duration of phase corresponding to different batches is not synchronized in practice because of unavoidable disturbances, fluctuations in initial conditions, as well as the changes of operation conditions. Therefore, it is a key problem in the batch processes to solve the uneven duration problem. That is to say, if there is an effective method for such a problem, the data provided for the modeling and monitoring processes based on multivariate statistical analysis will be more reliable.

Since uneven duration problem commonly arises in batch processes, many studies have been conducted on it. Among all the methods, the most commonly used one is the shortest length method [11], which directly cuts the rest of the batches by the shortest path length. Though this method is easy to implement, it will lead to a large loss of trajectory data and decrease the correlations between variables, thus reducing the reliability of the data. Besides, there are two general solutions that can be utilized to solve uneven-length problem, dynamic time warping (DTW) [12], [13] and correlation optimization warping (COW) [14], [15]. However, DTW and COW suffer from heavy calculation and low efficiency. Moreover, the relationship between process variables may be distorted when the trajectories are stretched or compressed to the reference, resulting in inaccurate results of trajectory synchronization. Additionally, warping technique can be only used for offline modeling owing that online batches cannot be synchronized to the reference. Li *et al.* [15], [16] proposed a method based on time slice sequence alignment. The method captures the changes of process characteristics by continuously calculating the correlation changes of each batch data along the time direction, so as to automatically identify irregular phases. And the generalized time slice is then constructed to build the monitoring model for the irregular phase data. Liu *et al.* [17] and Jiang *et al.* [18] proposed a new Gaussian mixture model (GMM) for offline phase partition. Besides, a multiple hypotheses testing-based operating optimality assessment and nonoptimal cause identification method is extended to handle the online assessment of the batch processes with multiphase and uneven-length characteristics. To improve the monitoring performance of nonlinear batch processes, Liu *et al.* [19] developed a window-based stepwise sequential phase partition method. Moreover, a traversal algorithm is given to determine the optimal choice of the KPCA parameters and the window size for phase partition. Luo *et al.* [20] and Wei *et al.* [21] developed a new phase identification method based on the warped K-means (WKM) clustering algorithm and phase identification combination index (PICI). Zhang *et al.* [22] and Wang and Zong [23] introduced the statistical analysis and online monitoring method for uneven batches based on the multi-phase characteristics. Furthermore, the problem of local modeling is solved according to the detection method of variable moving window-k nearest neighbor (VMW-KNN). In the above studies, the monitoring models are established based on the corresponding all sub-phase data, which makes the scale of data in the monitoring model so huge that the models are not accurate enough and online monitoring is not sensitive enough.

In order to automatically identify each sub-phase according to the time sequence, a phase identification and online monitoring method based on moving window-based MIIM for the uneven batch processes (MWMIIM) is proposed in this paper, which extends the multiway information increment matrix (MIIM) algorithm to the uneven batch processes [10]. Different from the existing MIIM algorithm, the innovative points of this paper are as follows: First, phase identification of each batch is carried out separately by combining moving window technique with MIIM algorithm. To be specific, determine the partition points by the variables' correlation changes in the corresponding moving window. Second, all the sub-phases are subdivided into common region, transition region and end region when offline modeling. Moreover, it is necessary to execute fine modeling in transition region and end region. Third, different monitoring strategies are adopted in different regions when online monitoring. Thereinto, new monitoring strategies are carried out in transition region and end region.

The remainder of this article is organized as follows: In Section II, the phase identification, data consolidation and sub-phase modeling algorithm are introduced in detail. Moreover, parameters N, L and w are discussed theoretically. In Section III, online monitoring is executed and an effective fault location method is proposed. In addition, experimental results are given in Section IV. Finally, conclusion and outlook are stated in Section V.

II. PHASE IDENTIFICATION AND SUB- PHASE MODELING

A. BASIC OPERATION

Assume that the three-dimensional data $X(I \times J \times K_i)$ is collected under normal operating conditions, where I is the number of batches, J is the number of measured variables per sampling point, K_i is the number of measured sampling points in $i_{th}(i \in 1, 2, \cdots, I)$ batch. Suppose that I batches

are considered separately, then each batch corresponds to a set of two-dimensional data $X_i(J \times K_i), i = 1, 2, \cdots, I$ accordingly. Take the i_{th} batch as an example, where K_i sampling points are denoted as $\{x(1), x(2), \cdots, x(k_i)\}$ in the batch. Besides, a moving window data matrix serves as the basic data unit for phase identification, then the working process of the moving window is described as follows. Starting from the first sampling data, the moving window covers the continuous w sampling data to form the first data matrix $X^w_{i,1} = [x(1), x(2), \cdots, x(w)]$, where w is a fixed value representing the length of the moving window, and w will be further discussed in section C. Then $X^w_{i,1}$ is normalized to zero mean and the unit variance, where $X^w_{i,1}$ is the 1_{st} normalized moving window data matrix in the i_{th} batch. Besides, the covariance matrix $C_1^{J\times J}$ of $X^w_{i,1}$ is calculated according to time sequence to characterize the process. Afterwards, update the window data matrix by moving one sampling point backward, namely the first sampling point in the window is removed and a new sampling point is added at the end of the window to ensure that the length of the moving window remains unchanged, and the corresponding data matrix $X^w_{i,2} = [x(2), x(3), \cdots, x(w+1)]$ is obtained. Similarly, $X^w_{i,2}$ is normalized to zero mean and unit variance, and the corresponding covariance matrix $C_2^{J\times J}$ is calculated. In this way, the moving window continues to move one sampling data backward each time until the last sampling point of this batch get covered. In the process of window sliding, $(K_i - w + 1)$ data windows are constructed successively, meanwhile $(K_i - w + 1)$ data matrices are obtained, thus $(K_i - w + 1)$ covariance matrices $C_k^{J\times J}(k \in 1, 2, ..., K_i - w + 1)$ are calculated accordingly. Note that $C_k^{J\times J}$ is also known as the correlation matrix of the process variables, which means that the element in row i and column j of $C_k^{J\times J}$ represents the correlation coefficient between the i_{th} and j_{th} variables. Other batches operate similarly. The whole working process of the moving window is shown in Fig.1.

B. SEQUENTIAL UNEVEN PHASE IDENTIFICATION

As for uneven batch processes, the duration of the corresponding sub-phase and total duration varies with batches. Consequently, executing phase identification in uneven batch processes is more difficult compared with even conditions.

In this work, an improved phase identification algorithm for uneven batch processes is developed by analyzing each batch separately. The detailed calculation process is described following.

Above all, the moving window technique described in part A above is required to obtain a series of correlation matrices $C_k^{J\times J}(k \in 1, 2, ..., K_i - w + 1)$. Furthermore, calculate the increment matrix between two adjacent correlation matrices $C_k^{J\times J}$ and $C_{k+1}^{J\times J}$, then the multidimensional average gain index δ_k corresponding to the increment matrix is derived. The gain index δ_k is defined to capture the dynamic variation

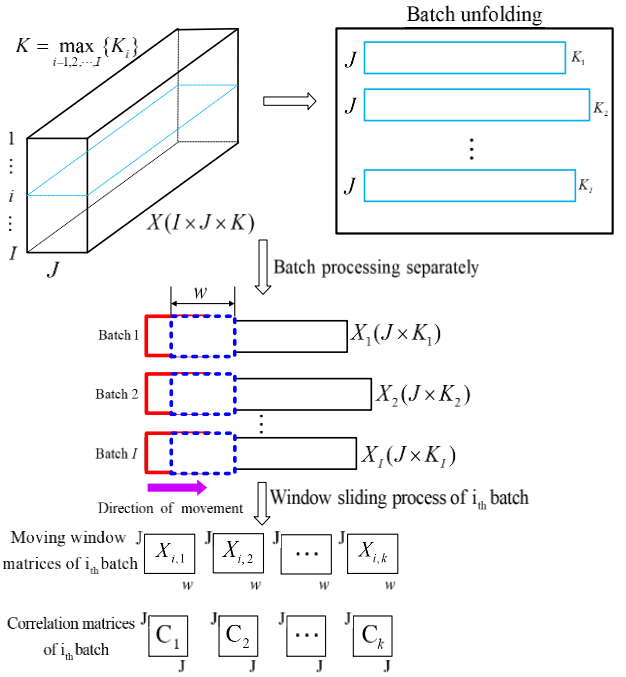


FIGURE 1. The whole working processes of the moving window.

of process characteristics

$$\delta_k = \frac{\sum_{p=1}^{J}\sum_{q=1}^{J}|C_{k+1}(p,q) - C_k(p,q)|}{J^2} \tag{1}$$

The most critical point in the sub-phase identification is to find the 'starting point' of the next sub-phase, namely phase partition point, which requires a benchmark to determine whether a new sub-phase occurs. Specifically, when the moving window moves backward in turn, calculate the increment matrix of two adjacent correlation matrices, and each corresponding gain indexes δ_k is compared with the benchmark to find the phase partition points. The benchmark in this thesis is the switch control limit σ, which is defined as follows

$$\sigma = \frac{1}{L}\sum_{k=1}^{L}\delta_k \tag{2}$$

where L is a fixed value.

As can be seen from (2), σ is associated with the latest L values of δ_k. In order to ensure that there is a switch control limit σ at the initial moment of each sub-phase for comparison, L values of δ_k are required, which in turn requires $L + 1$ correlation matrices, meaning that moving window need to be moved L times. In other words, there are at least $L + w$ sampling points in each sub-phase. Therefore, the basic assumption this algorithm based on is that the first $L + w$ sampling points of each sub-phase must belong to this sub-phase.

Based on the above discussion, let the first $L+w$ sampling points in one batch belong to the first sub-phase, then the

initial switch control limit σ can be obtained. According to the time sequence, once a new sampling point is obtained, the gain index δ_k is calculated and compared with the switch control limit σ afterwards to determine whether a new sub-phase generates. The criteria are as follows: 1) If $\delta_k \geq N\sigma$, then the corresponding sampling point is identified as an outlier that does not belong to this sub-phase, thus it does not contribute to the update of switch control limit. 2) if $\delta_k < N\sigma$, then the current sampling point is regarded as a normal sampling point that will be adopted to update the switch control limit defined in (2), where N is a tunable parameter called tolerance factor that will be further discussed in subsection C. Suppose that continuous L sampling points exceed the switch control limit, then a new sub-phase is generated, and the first sampling point beyond the switch control limit is the starting point of the new sub-phase. Subsequently, the sampling points identified as belonging to the first sub-phase are removed, and the first remaining $L + w$ sampling points are utilized as the initial sampling points for the next sub-phase. Repeat the same steps as in the first sub-phase until the end of the batch. Considering that there are at least $L + w$ sampling points in each sub-phase, if the number of the remaining sampling points of the batch is less than $L + w$, then these sampling points are affiliated into the previous sub-phase herein. By analogy, all subsequent batches take the same operations. In this way, the phase partition points of each batch are determined. The flow chart of the entire phase identification processes for one batch is described in Fig.2

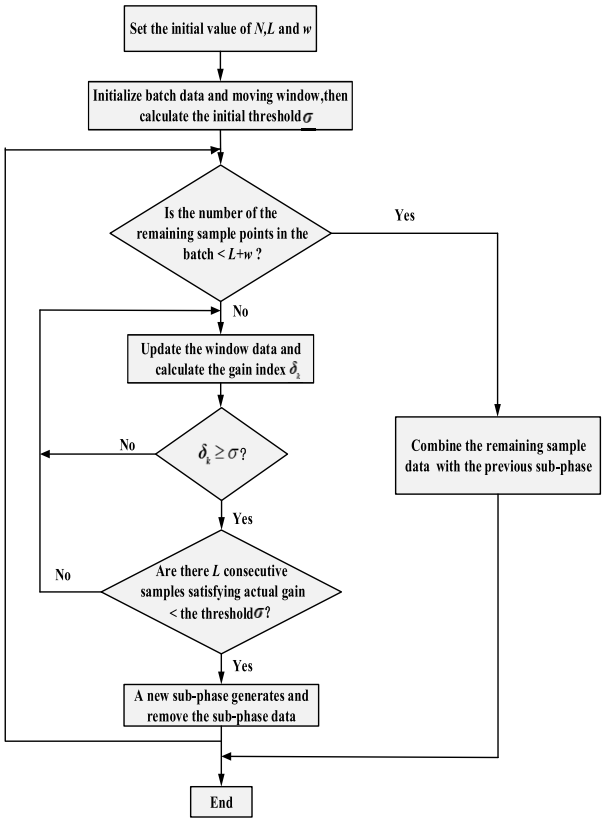


FIGURE 2. Flow chart of the entire phase identification processes for one batch.

C. DISCUSSION ON THE SELECTION OF PARAMETERS

The parameters N, L and w presented in the subsections A and B have a significant influence on the phase identification performance of the proposed algorithm, thus they are further discussed as following.

i) The N defined in $\delta_k \geq N\sigma$ is an adjustable parameter which is called tolerance factor. A larger value of N means that there will be fewer sub-phases in a batch, thus a smaller number of monitoring models need to be established, which significantly reduces the modeling complexity. However, fewer monitoring models also mean degraded monitoring accuracy, because more sampling points are allocated to the same sub-phase and monitored by the same model. While a smaller N can lead to the opposite effect. Under some special circumstances, when N is set to 1, nearly every $L + w$ sampling points are identified as a separate sub-phase. While a sufficiently large value is assigned to N, only one sub-phase is obtained and the whole process is monitored by a single model. Therefore, the choice of N is a trade-off between the model accuracy and the model complexity.

ii) The L in (2) is a fixed value, which represents the number of gain indexes used to calculate the switch control limit σ as well as the number of outliers needed to generate a new sub-phase. It is difficult to form new sub-phases when L is large, thus reducing the number of sub-phases. Similar to N, this results in lower monitoring accuracy. On the contrary, a small L can lead to the opposite effect. Ultimately, the value of L needs to be determined according to the actual situation.

iii) The w mentioned in $X_{i,1}^{w} = [x(1), x(2), \cdots, x(w)]$ refers to the length of the moving window. Take the aircraft steering gear system as an example, data update cycle of aircraft steering gear system is within 30-40 ms. In order to better reflect the process characteristics of the data, the length of the window is set as 3-4 times of the data update cycle. Therefore, the initial window length is set as 10. The larger the value of w is, the more the sampling data contained in the data matrix is, which increases the amount of computation each time and reduces the accuracy of the obtained dynamic switch control limit simultaneously. And the opposite effect occurs when the value of w is smaller. The $L + w$, which is raised to improve the accuracy of the phase identification, represents the minimum length of a new sub-phase. Obviously, the value of $L + w$ is affected not only by L, but also by w.

In summary, adjusting and simultaneously selecting the optimal values of the parameters N, L and w can change the performance of the proposed phase identification algorithm

CHART.I. Summary of Optimal Parameters Selection.

1. Take the aircraft steering gear system as an example, according to a number of previous experience, when N, L and w range from 1 to 5, 2 to 10 and 10 to 15 respectively, the performance is better. Here, initialize $N=1, L=2, w=10$ and $l_1=0.1, l_2=1, l_3=1$, where l_1, l_2 and l_3 is the step length of N, L and w, respectively.

2. Update N via $N=N+l_1$ iteratively, N is used to identify a sampling point. In the phase identification process, confirm whether it belongs to the current sub-phase or not. Besides, ensure whether a sampling point is normal or abnormal when online monitoring.

3. Let $L=L+l_2$ iteratively, then execute phase identification with (2), and confirm the optimal value of L.

4. Let $w=w+l_3$, constantly update w in $X_{i,k}^{w}$, the length of the moving window is enlarged step by step, and the optimal value of w is obtained by means of the final experimental results.

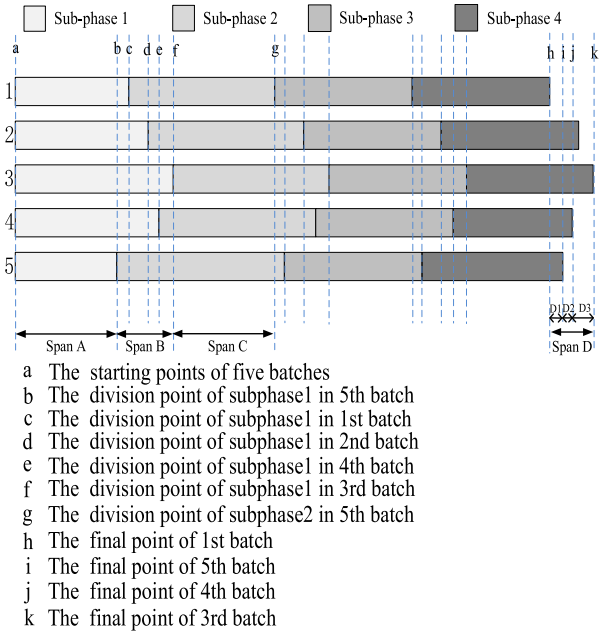


FIGURE 3. The diagram of the batch processes divided into common region, transition region and end region.

to a large extent. The summary of optimal parameters selection is shown in Chart I.

D. SUB-PHASE FINE MODELING

After all the sub-phases of all batches have been identified, the partition points in each batch are available. Subsequently, the corresponding batch data is consolidated accurately and normalized for building monitoring models so that each region can be accurately monitored by an appropriate model. For the ease of understanding, this section takes five uneven batches as an example to introduce the proposed data consolidation and fine modeling method, as shown in Fig.3.

As can be seen from Fig.3 that each batch is divided into sub-phase1, sub-phase2, sub-phase3 and sub-phase4. Obviously, sub-phase partition points of each batch are not aligned. Therefore, it is necessary to further execute fine modeling. Taking into account the disparate process characteristics of disparate regions, all batches are divided into common region, transition region and end region in this paper. Thereinto, common region (such as span A and C) denotes that the sampling points of all batches belong to the same sub-phase, in which only one common monitoring model is required. While transition region (such as span B) indicates that the sampling points in different batches belong to different sub-phases at the same time. Considering that the variables' characteristics between different sub-phases change obviously, the model may be not accurate enough if only one model is utilized to describe the variables' characteristics in the transition region. To improve the accuracy of the model, two monitoring models are established according to the different variables' characteristics of different sub-phases in the transition region. Specifically, all the sampling points belonging to sub-phase1 in the transition region are merged, the corresponding monitoring model is established afterwards, and all the sampling points belonging to sub-phase2 are merged and modeled similarly. For the end region (such as span D as shown in Fig.3), not only is the number of batches gradually decreased, but also the modeling process is restricted by the termination condition. Therefore, span D is subdivided into span D1, span D2 and spanD3. Here is a detailed description of the modeling processes for each of the three regions as shown in Fig.4.

a: COMMON REGION

Take span A as an example, and suppose that there are I batches in the region and there are R sampling points in each batch. Above all, the data segment is unfolded and merged along the batch-wise direction, and the sample data of all batches belonging to the same time interval constitutes a time slice matrix $X_{\mathrm{r}}(I \times J)$, $\mathrm{r} = 1, 2, \cdots, R$. Subsequently, the mean time slice matrix $\overline{X}(I \times J)$ of span A as shown in Fig.5 can be computed via

$$\overline{X}_{I\times J} = \frac{1}{R}\sum_{i=1}^{R} X_r \tag{3}$$

Afterwards, add each row of each time slice matrix X_{r} in the region to the average matrix $\overline{X}(I \times J)$ respectively, and $I * R$ expanded matrices $\overline{X}_{ri}$ can be derived as

$$\overline{X}_{ri} = \begin{bmatrix} \overline{X}_{I\times J} \\ X_r(i,:) \end{bmatrix} \tag{4}$$

where $\mathrm{r} = 1, 2, \cdots, R, i = 1, 2, \cdots, I$. The correlation matrices of $\overline{X}_{ri}$ and $\overline{X}$ can be calculated and express as $\overline{C}_{ri}$ and $\overline{C}$. Then, the gain indexes δ_{ri} between $\overline{C}_{ri}$ and $\overline{C}$ can be

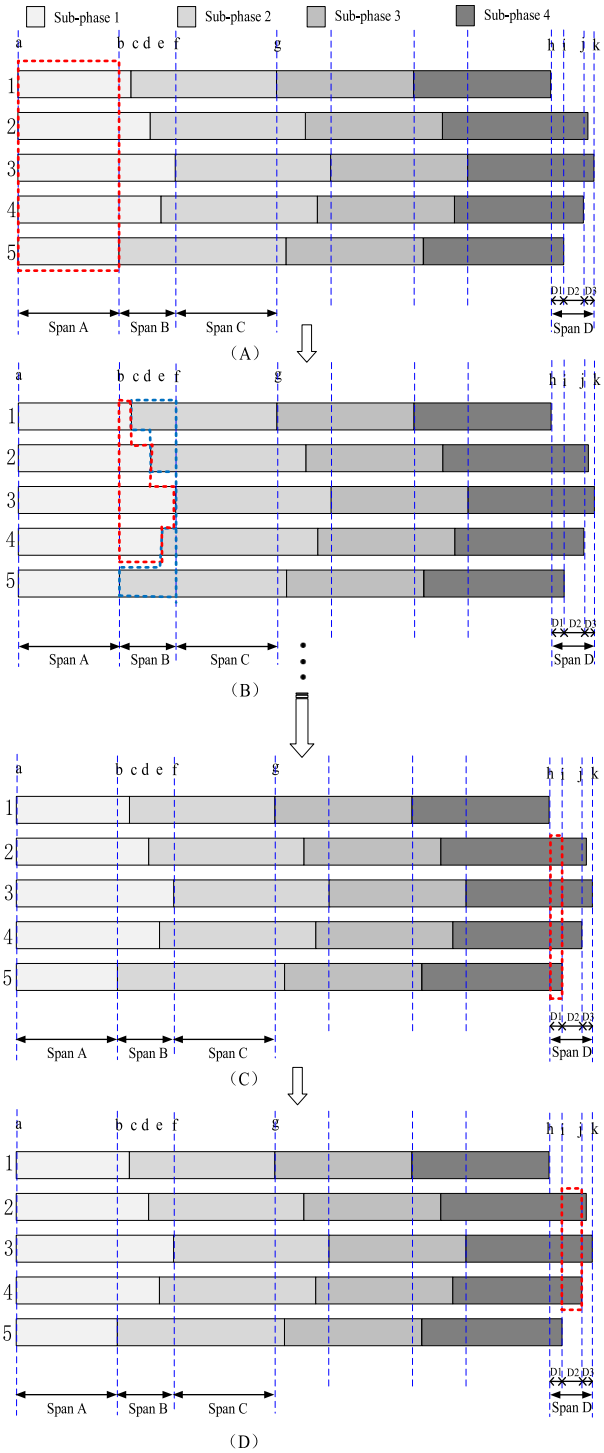


FIGURE 4. The illustration of detailed modeling processes for each region.

obtained by

$$\delta_{ri} = \frac{\sum_{p=1}^{J}\sum_{q=1}^{J}\left|\overline{C}_{ri}(p,q) - \overline{C}(p,q)\right|}{J^2} \tag{5}$$

The statistics δ_{ri} roughly obey the chi-square distribution. And the control limit can be approximated obtained

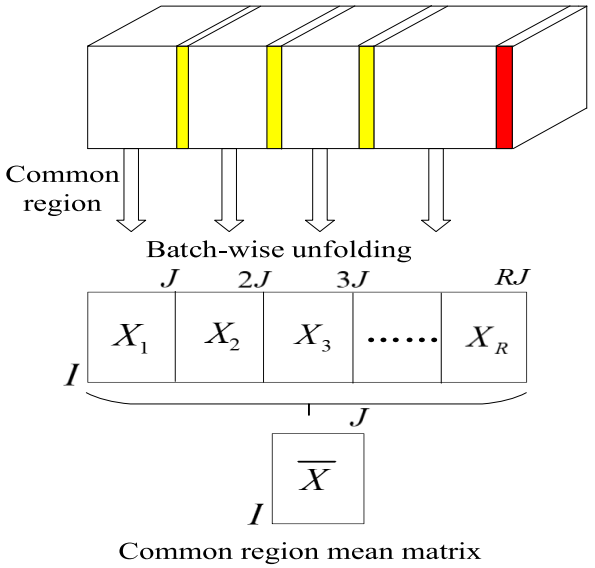


FIGURE 5. The illustration of the mean matrix for common region.

as follows

$$\delta_{limit} \sim g\chi^2_{h,\alpha} \tag{6}$$

where $g = \nu/2m$ and $h = 2m^2/\nu$. Specifically, m is the average of all the δ_{ri} and ν is the variance of δ_{ri}. Other common regions, such as span C, also follow the above procedures to establish the corresponding monitoring model.

b: TRANSITION REGION

The modeling processes in the transition region is similar to that of the common region except for the data used for modeling and the direction to unfold and merge the data segments. Take span B in subgraph B of Fig.4 as an example, the data segments belonging to sub-phase1 and sub-phase2, namely the region in the red dotted line and the blue dotted line, are unfolded and merged along the variable-wise direction separately. What needs illustration is that the length of each batch is differentiated in the red dotted line region, as is the blue dotted line region. Therefore, the average time slice matrix cannot be obtained by batch-wise unfolding and merging similar to that of the common region, while can be just obtained by variable-wise unfolding and merging. Take the red dotted line region as an example, average the data of each batch firstly to obtain its average vector $\overline{x_i}$, which can be obtained by

$$\overline{x}_i = \frac{1}{K_i}\sum_{i=1}^{K_i} x_{i,k} \tag{7}$$

where $i = 1, 2, \cdots (I-1), k = 1, 2, \cdots K_i$. Thereinto, K_i represents the number of vectors of the current sub-phase in i_{th} batch. Then stack the average vectors $\overline{x_i}$ of all batches in line order to construct the average matrix $\overline{X}((I-1)\times J)$ of span B as shown in Fig.6. The rest of the procedures are the same as the common region.

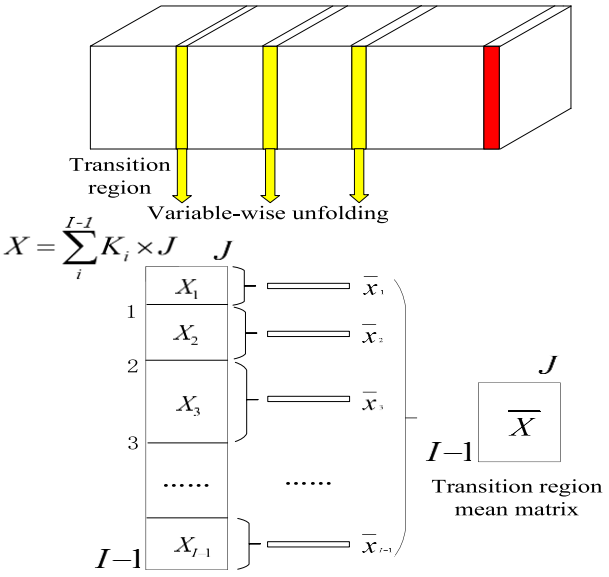


FIGURE 6. The illustration of the mean matrix for transition region.

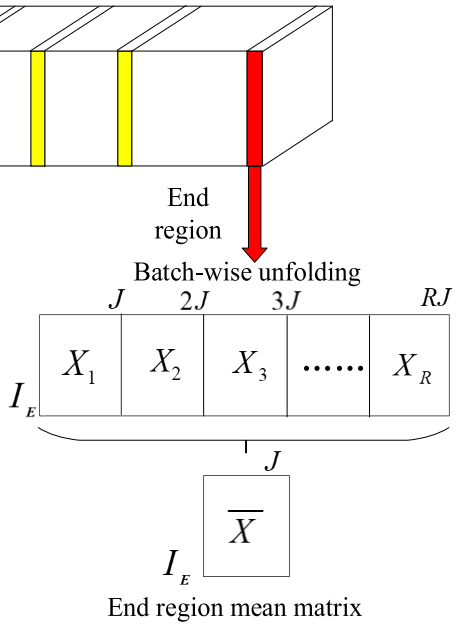


FIGURE 7. The illustration of the mean matrix for end region.

c: END REGION

With respect to end region as shown in span D in Fig.4, the number of remaining batches decreases as moving backwards. According to the numerical relationship between the batches and the variables, which is suggested by Johnson and Wichern [24], the modeling of span D is limited by the termination condition $I_E \geq (2 \sim 3)J$. It is significant to determine whether a separate model is required between two adjacent endpoints of different batches. When the termination condition is satisfied (span D1 and D2 in the Fig.4), the data segments of all batches between two adjacent endpoints are unfolded and merged respectively along the batch-wise direction, which is similar to that of the common region. Therefore, the mean time slice matrices $\overline{X}(I_E \times J)$ of span D as shown in Fig.7 can be computed via

$$\overline{X}_{I_E \times J} = \frac{1}{R}\sum_{i=1}^{R} X_r \tag{8}$$

where I_E is a varying value which represents the number of batches between two adjacent endpoints and satisfies $2J \leq I_E \leq I$. If not satisfy (span D3 in the Fig.4), the remaining sampling points in span D will be supervised by the previous adjacent model (span D2 in the Fig.4).

E. IDENTIFICATION OF VARIABLE CONTRIBUTION RATE TO PHASE TRANSITION

In the batch processes, the correlations between variables remain almost constant in the same sub-phase, while there are significant differences between different sub-phases, and the correlation matrices will also have an obvious change. According to this typical feature, the continuous w sampling points are selected in each sub-phase to form a matrix respectively after phase identification, and the correlation matrix is calculated to represent the data characteristics of the current sub-phase. Afterwards, the difference matrix of two adjacent correlation matrices is calculated, where diagonal and non-diagonal elements of the correlation matrix respectively represent the self-correlation and cross-correlation between variables. Therefore, a column of the correlation matrix represents the correlations of variables corresponding to the column. In the different matrix, the ratio that the sum of the absolute values of each column occupy the sum of the absolute values of the entire matrix represents the contribution of the corresponding variable of this column to the phase transition. In this way, the contribution rate of variables to the phase transition is identified by

$$P(j) = \frac{\left|\sum_{j=1}^{J} C[j,k]\right|}{\left|\sum_{j=1}^{J}\sum_{k=1}^{J} C[j,k]\right|} \tag{9}$$

III. ONLINE MONITORING

A. ONLINE MONITORING

Fine modeling is to better monitor the batch processes. Online monitoring provides real-time monitoring of the batch processes to detect whether it is running properly. For uneven batch monitoring processes, different monitoring strategies should be adopted for different regions.

B. CASE A: COMMON REGION MONITORING

In the common region, when a new sampling point x_{test} of the test batch arrives, there is a definite monitoring model and a corresponding fixed control limit to determine whether it is normal or not according to the time index. Above all, the test sampling point is normalized to zero mean and unit variance. Then the gain index δ can be obtained by (4) and (5). The detailed process is as follows.

Firstly, the test sampling point is added to the average matrix of common region as follows

$$\overline{X}_{\text{test}} = \begin{bmatrix} \overline{X}_{I\times J} \\ x_{test} \end{bmatrix} \tag{10}$$

Then, $\overline{C}_{test}$ and $\overline{C}$ are used to represent the correlation matrices of $\overline{X}_{test}$ and $\overline{X}$ respectively. In addition, the gain index δ_{test} can be obtained by

$$\delta_{test} = \frac{\sum_{p=1}^{J}\sum_{q=1}^{J}\left|\overline{C}_{test}(p,q) - \overline{C}(p,q)\right|}{J^2} \tag{11}$$

By comparing δ_{test} with the control limit $\delta_{\lim it}$ obtained by (6), online monitoring consistently carries out. If $\delta_{test} < \delta_{\lim it}$, the current sample is considered normal; otherwise abnormal. Additionally, the fault location method is proposed to find the root cause of the fault in the next subsection B.

C. CASE B: TRANSITION REGION MONITORING

In the transition region, owing that the phase partition points of each batch are not aligned, the new test sampling point either belong to the current sub-phase or transfer to the next sub-phase, thus it is hard to determine which sub-phase it belongs to. In other words, it is uncertain which monitoring model should be utilized to detect the new test sampling point. Therefore, a flexible monitoring strategy is required. Originally, the current sub-phase model is adopted to detect the new test sampling point. Concretely, the test sampling point is added to the average matrix of transition region of current sub-phase

$$\overline{X}_{\text{test}} = \begin{bmatrix} \overline{X}_{(I-1)\times J} \\ x_{test} \end{bmatrix} \tag{12}$$

Then, the solution procedures of δ_{test} are the same as that of the common region. Provided that the test statistic δ_{test} does not exceed the control limit $\delta_{\lim it}$ of the current sub-phase monitoring model, it is considered to operate normally in the current sub-phase. Otherwise, the current sub-phase model cannot describe the process characteristics of the test sampling point well, and the next sub-phase monitoring model is used to detect it, likewise. Suppose that the test statistics is within the control limit of the model, it is considered to be a normal operation data of the next sub-phase, otherwise it is abnormal, and a fault alarm signal is generated.

D. CASE C: END REGION MONITORING

In the end region, there may be multiple sub-models, and each sub-model corresponds to an average matrix and a control limit. Besides, each control limit and the dimension of each average matrix are different. Different from the above two regions, the test sampling point is firstly added to the average matrix of current sub-model as follows

$$\overline{X}_{\text{test}} = \begin{bmatrix} \overline{X}_{I_E\times J} \\ x_{test} \end{bmatrix} \tag{13}$$

After that, the rest of the monitoring steps are the same as that of the common region. When the sampling point corresponding to the next sub-model arrives, the test sampling point is added to the average matrix of next sub-model, then the monitoring steps of the previous sub-model are repeated until the end of the test batch.

E. FAULT LOCATION

When a fault is detected by the monitoring model, the source of the fault can be found out by calculating the contributions of process variables to the change of the correlation matrix. Specifically, the variable that contributes the most to the change of the correlation matrix is just the source that causes the fault. And the contributions of each variable are computed as follows.

$$V_s = \frac{\sum_{p=1}^{J}\left|\overline{C}_{test}(p,s)\right| + \sum_{q=1}^{J}\left|\overline{C}_{test}(s,q)\right| - \left|\overline{C}_{test}(s,s)\right|}{\sum_{p=1}^{J}\sum_{q=1}^{J}\left|\overline{C}_{test}(p,q)\right|} \tag{14}$$

where $s = 1, 2, \cdots, J$.

In some cases, there may be two or more variables that contribute significantly to the existing fault because process variables may be interrelated and fault can spread quickly between process variables, especially in closed-loop systems. Therefore, the joint effect is further analyzed by

$$V_{st} = \frac{2\overline{C}_{test}(s,t)}{\sum_{p=1}^{J}\left|\overline{C}_{test}(p,s)\right| + \sum_{q=1}^{J}\left|\overline{C}_{test}(s,q)\right| - \left|\overline{C}_{test}(s,s)\right|} \tag{15}$$

where $s = 1, 2, \cdots, J, t = 1, 2, \cdots, J$.

IV. CASE STUDY AND DISCUSSION

A. APPLICATION ON AIRCRAFT STEERING GEAR SYSTEM

The aircraft steering gear is an actuator of automatic flight control system, which changes the civil aircraft's attitude or trajectory by controlling the control surfaces. Fig.8 is a mechanical structure diagram of the aircraft steering gear system. Under working conditions, the servo motor receives the control command from the flight control computer. Then, the control surface is pushed or pulled by the motion of the actuator cylinder to ensure that the aircraft's attitude or trajectory approaches the given command value with certain precision.

During the entire flight of a civil aircraft, the operation processes of aircraft steering gear system are typical batch processes, which show significantly different operation characteristics in disparate flight stages. To verify the proposed phase identification and modeling method, they can be conveniently defined as typical uneven multiphase batch processes for experiments. The six phase-sensitive variables of aircraft steering gear system collected in the experiments are shown in Table 1.

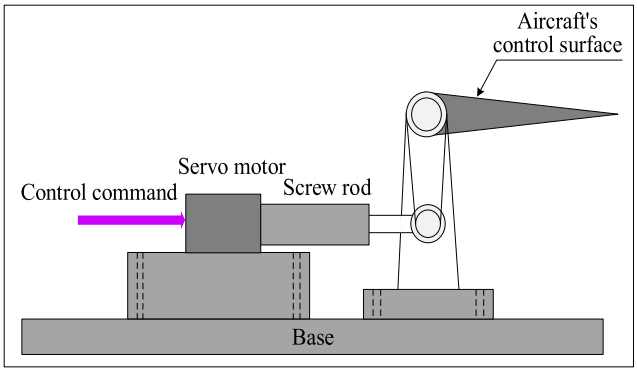


FIGURE 8. The mechanical structure diagram of the aircraft steering gear system.

Table 1. The six variables of aircraft steering gear system.

No.	Description	Units	Value
1	Rudder position	Degree	32
2	Motor speed	r/min	7650
3	Motor current	A	35
4	Winding A phase current	A	20
5	Winding B phase current	A	20
6	Winding C phase current	A	20

The aircraft steering gear system mainly consists of triple closed-loop structures, which are current loop, speed loop and position loop. In this experiment, 50 normal batches are used as training data for modeling, and another 10 batches are used as test data. To approximately simulate the dynamic change of practical operation processes, the frequency of the current loop control is set to 25.0, 26.0, 27.0, 28.0 and 29.0 Hz respectively to generate the training data. Afterwards, 10 batches data are collected for each frequency, and 50 uneven batches are finally obtained, in which batch duration ranges from 400 to 450 samples.

B. PHASE IDENTIFICATION RESULTS

To verify the effectiveness of the proposed phase identification method in addressing the uneven-length problem of the aircraft steering gear system, a comparative experiment is conducted by the window-based stepwise sequential phase partition method for nonlinear uneven batch processes (WNSSPP-U). In this algorithm, the principal component analysis (PCA) algorithm is utilized to extract the principal component of the moving window data matrix, so as to realize phase identification. Actually, the function of α in WNSSPP-U algorithm is basically the same as that of N in the proposed method, both of which are used to adjust the phase identification precision. In this experiment, five different sets of values are tried for N and α. Among them, N is set to 1.3, 1.5, 1.8, 2, and 2.5 respectively, while α is set to 2.5, 2.8, 3.5, 5 and 10 respectively. L and w are set to 3 and 12 here according to experimental results. Take one batch of each frequency as an example, Fig.9 presents the phase identification results regarding different N, and Fig.10 presents the comparison experiment results.

As can be seen from Fig.9, with the increase of the relaxing factor N, the number of sub-phases acquired gradually decreases. And Fig.10 shows a similar pattern, indicating that N and α have similar effects on phase identification. Particularly, when N and α are large enough ($N = 2.5$, $\alpha = 10$), the batch is no longer divided and the whole batch is modeled as a single phase. The possible reason is that with the increase of the relaxation factor N and α, the sampling points identified as abnormal points decrease, thus resulting in a decrease in the number of sub-phases obtained. On the contrary, when the relaxation factor decreases, the sampling points are more likely to be regarded as abnormal points and more sub-phases can be obtained.

Thereinto, when L is set to 3, w is set to 12 and N is set to 1.5, the phase partition result is the closest to the actual situation. It can be seen that the operating status of the steering engine under experimental conditions is divided into four sub-phases according to the time order after the phase identification. Specifically, partition points of the first phase are 95th, 100h, 105th, 120th and 130th respectively; the second phase partition points are 170th, 180th, 185th, 210th and 220th respectively; partition samples of the third phase are: 250th, 260h, 265th, 295th and 315th respectively. At last, the remaining sampling points are the fourth phase data. Subsequently, the corresponding monitoring model is established and the control limit is calculated in the common region and transition region. As for the end region, if the number of batches between two adjacent endpoints is less than 12, no new monitoring model is established, and the remaining sampling points adopt the adjacent model as the monitoring model.

C. ONLINE MONITORING RESULTS

For online monitoring statistics in WNSSPP-U algorithm, T^2 statistics focus on tracking the changes of systematic variations and T^2 monitoring performance may not be significantly affected by the partition results. While the monitoring statistics of squared prediction errors (SPE) are more sensitive to changes of process variable correlations. Therefore, only SPE is used in WNSSPP-U algorithm of this comparative experiment.

Originally, take a normal batch data as test data. The monitoring results of the proposed MWMIIM-U and the WNSSPP-U methods are shown in Fig.11 and Fig.12, respectively. As can be seen from Fig.11, none of the test statistics exceed the control limit of the MWMIIM-U method. While some individual test statistics exceed the control limit of the WNSSPP-U method as shown in Fig.12. It can be concluded that the online monitoring performance of the proposed MWMIIM-U method is superior to the WNSSPP-U method.

To further illustrate the ability of the MWMIIM-U method to monitor fault, five aircraft steering engine system fault types are randomly added to the test batch. In addition, in order to eliminate the influence of random errors, the experiment is repeated 10 times and the average value is taken as

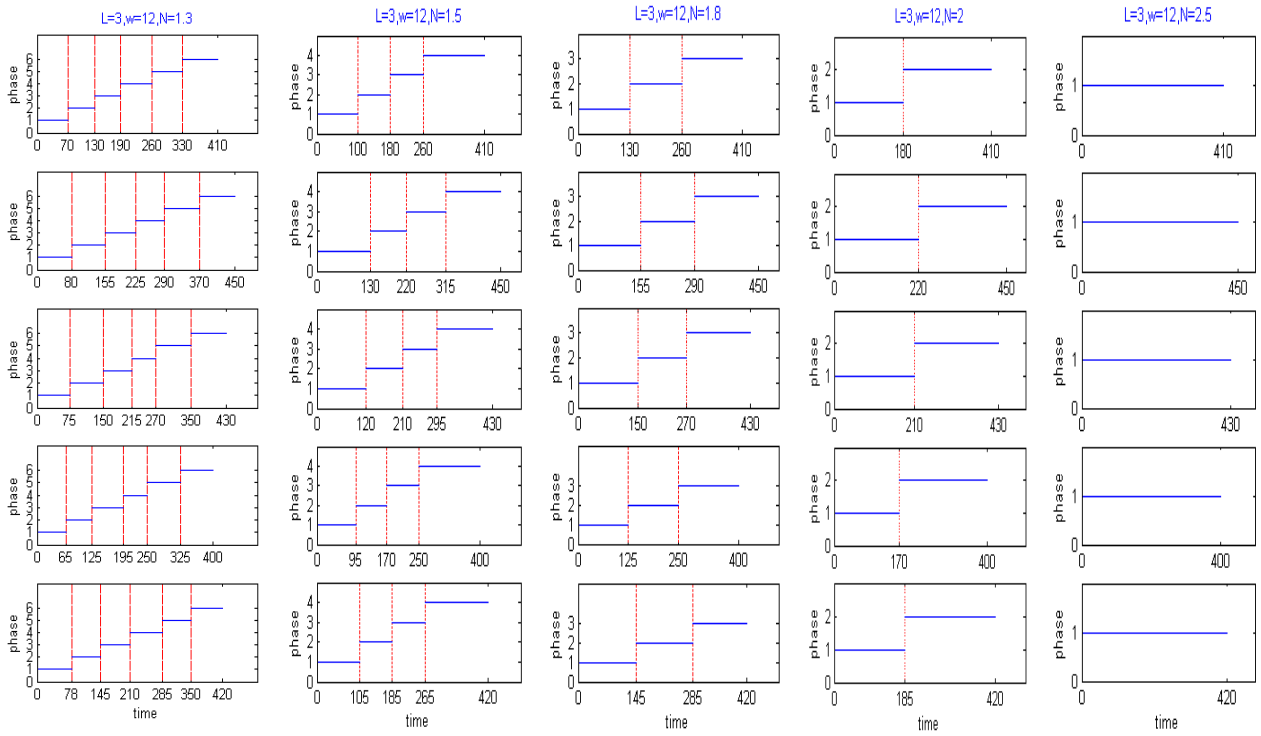


FIGURE 9. The sub-phase identification results regarding different values of N for the proposed MWMIIM-U algorithm.

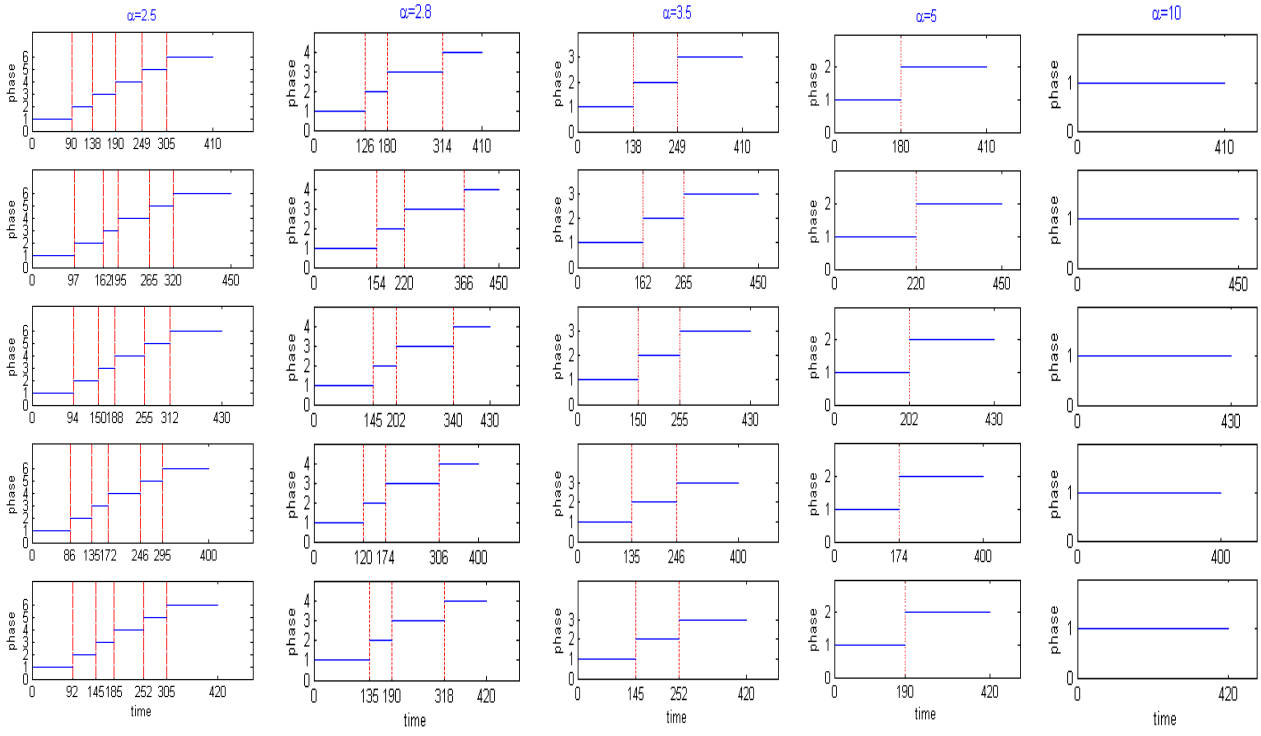


FIGURE 10. The sub-phase identification results regarding different values of α for the WNSSPP-U algorithm.

the final results. Then the monitoring results of the proposed MWMIIM-U method ($L = 3, w = 12, N = 1.5$) and the WNSSPP-U method ($\alpha = 2.8$) for the five faults are given in Fig.13 and Fig.14, respectively. It can be seen that these faults can be detected by both the MWMIIM-U method and the WNSSPP-U method. However, the number and time of

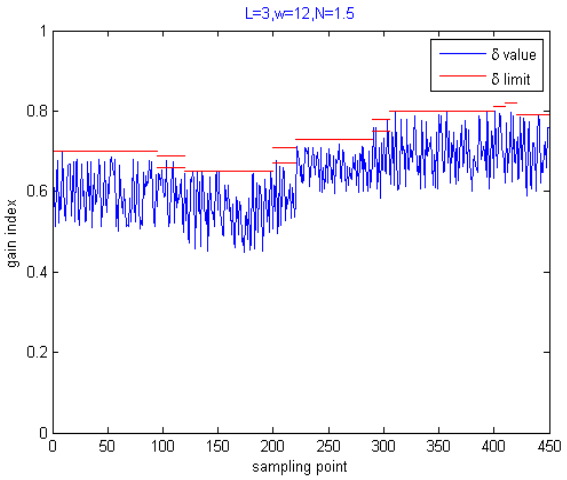


FIGURE 11. A normal batch monitoring results of the proposed MWMIIM-U method.

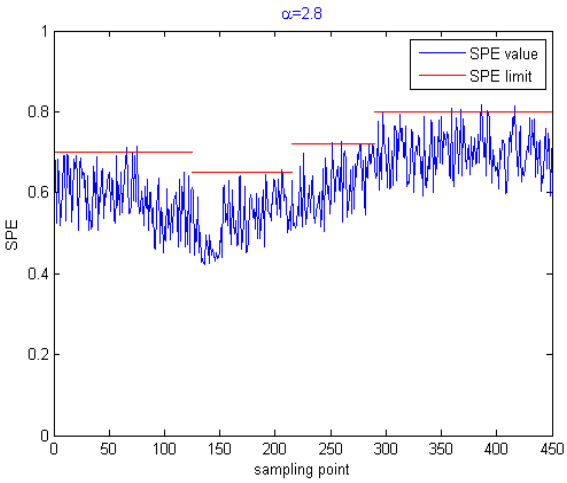


FIGURE 12. A normal batch monitoring results of the WNSSPP-U method.

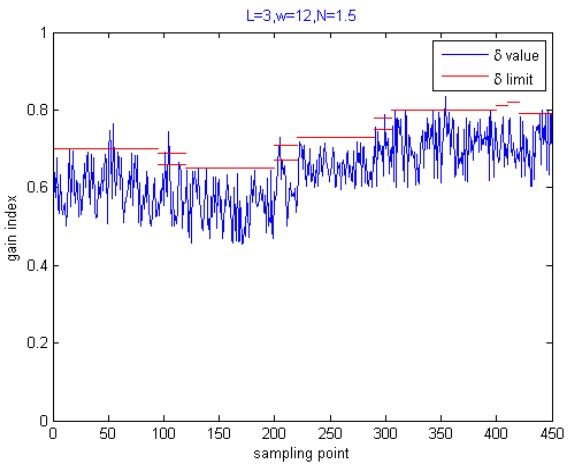


FIGURE 13. Fault monitoring results of the proposed MWMIIM-U method. (blue line, the monitoring statistics; red line, SPE control limits).

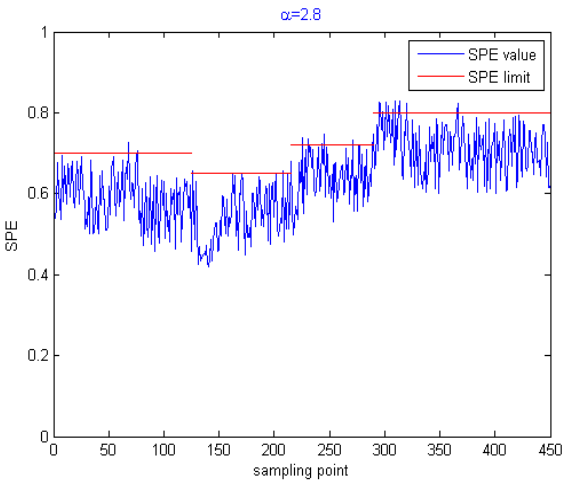


FIGURE 14. Fault monitoring results of the WNSSPP-U method. (blue line, the monitoring statistics; red line, δ control limits).

fault alarm signals are different, resulting in disparate false alarming rate (FAR) and alarming delay time ΔT. Thereinto, $FAR = n/K$, where n is the total number of false alarm signals, and K is the total number of sampling points. $\Delta T = FAT - FOT$, ΔT reveals the fault detection ability (FAT is termed as first alarming time, FOT is called the first occurring time).

The location of 5 fault points in the test batch, and the location information is detected by the monitoring methods of MWMIIM-U and WNSSPP-U are shown in Table 2. Take the third row of data in the table, the fault is introduced at the 200th sampling point, the monitoring statistics of the MWMIIM-U method gives the only alarm signal at the 202th sampling point, as shown in Fig.13, meaning that only 2 delay exists using the MWMIIM-U method. While in Fig.14, the fault is introduced at the 230th sampling point, the monitoring statistics SPE of the WNSSPP-U method gives the first and second alarm signals at the 238th and 240th sampling points, respectively. To put it another way, the delay time of the MWMIIM-U method is much shorter than that of the WNSSPP-U method. Besides, the false alarm rates of the MWMIIM-U method are much lower than that of the WNSSPP-U method. In conclusion, experimental results illustrate that the MWMIIM-U method is superior to WNSSPP-U method in terms of real-time monitoring performance and accuracy.

The contribution rate of each variable identified by the proposed MWMIIM-U method to the transition between different phases is shown in Fig.15. It can be concluded that the contribution rate of the same variable to different phase transition processes is different, and the function of different variables on the same phase transition process is also different. Among them, variable 3, variable 2 and variable 1 contribute the most to the transition from phase 1 to phase 2, from phase 2 to phase 3 and from phase 3 to phase 4, with the contribution rates of 30%, 40% and 35% respectively. In summary, the variables that play a leading role in different phase transitions are not the same.

Table 2. Comparison of two monitoring methods.

Monitoring methods	Fault types	FOT	FAT	ΔT	FAR
MWMIIM-U	1	50	53	3	1.1%
	2	100	104	4	
	3	200	202	2	
	4	290	293	3	
	5	350	354	4	
WNSSPP-U	1	50	58	8	2.4%
	2	200	208, 210, 214	8, 10, 14	
	3	230	238, 240	8, 10	
	4	280	287, 290, 292	7, 10, 12	
	5	350	355, 358	5, 8	

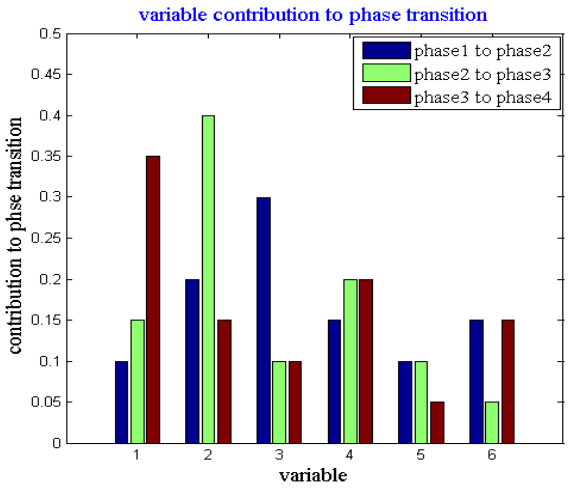


FIGURE 15. Contribution rate of each variable to different phase transition.

Replace a part of the aircraft steering engine system with a part with faults, and collect the corresponding data. Subsequently, apply the proposed MWMIIM-U and WNSSPP-U algorithms for fault identification, and the comparison results of the fault location effects of the two methods are shown in Fig.16. It can be clearly seen that the variables with the highest fault contribution rate obtained by the MWMIIM-U algorithm and WNSSPP-U algorithm are both variable 1. Actually, the part corresponding to variable 1 are exactly the fault part, indicating that both fault location methods can accurately analyze the source of faults. In addition, the contribution rate of variable 1 calculated by the MWMIIM-U method is 33%, which is higher than that of WNSSPP-U method (24%). Moreover, the variable fault contribution rate calculated by the proposed MWMIIM-U method is significantly different from each other, while that of the WNSSPP-U is almost the same. The results illustrate that the MWMIIM-U method can better distinguish the fault

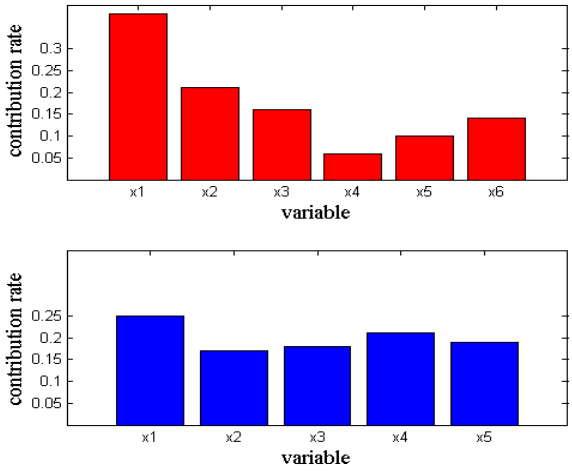


FIGURE 16. Contribution charts for fault 1 with the proposed MWMIIM-U algorithm and the WNSSPP-U algorithm.

parts from the normal parts, so that the obtained results are more credible and reliable.

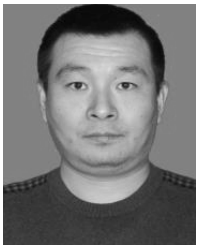
V. CONCLUSIONS

In this paper, a moving window-based multiway information increment matrix method for the uneven batch processes (MWMIIM) is proposed for single batch phase identification. Experimental results show that the algorithm can accurately capture the changes of variables' correlations in the process of phase transition, so as to obtain accurate sub-phase partition results representing different working conditions. Since the corresponding sub-phase partition points of each batch are not aligned, the whole batch processes are divided into three regions. Meanwhile, diverse detailed modeling and online monitoring strategies are adopted in different regions to more accurately reflect the variation trend of the process characteristics. The experiments conducted on aircraft steering gear system show that the proposed strategies provide competitive performance of online monitoring and fault location. Future works will be devoted to the research of online modeling, the improvement of the design of parameters and the universality of the proposed algorithm.

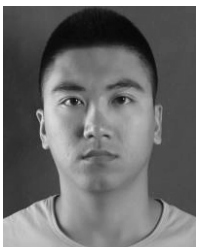
REFERENCES

[1] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE J.*, vol. 40, no. 8, pp. 1361–1375, Aug. 1994.
[2] P. Nomikos and J. F. MacGregor, "Multi-way partial least squares in monitoring batch processes," *Chemometrics Intell. Lab. Syst.*, vol. 30, no. 1, pp. 97–108, Nov. 1995.
[3] N. Lu, F. Gao, and F. Wang, "Sub-PCA modeling and on-line monitoring strategy for batch processes," *AIChE J.*, vol. 50, no. 1, pp. 255–259, Jan. 2010.
[4] Z. Li, Q. Li, Z. Wu, J. Yu, and R. Zheng, "A fault diagnosis method for on load tap changer of aerospace power grid based on the current detection," *IEEE Access*, vol. 6, pp. 24148–24156, 2018.
[5] A. C. F. Mamede, J. R. Camacho, and J. A. Malagoli, "Design and finite element analysis for the single-phase variable reluctance motor," *IEEE Latin Amer. Trans.*, vol. 14, no. 8, pp. 3636–3642, Aug. 2016.
[6] C. Zhao and Y. Sun, "Step-wise sequential phase partition (SSPP) algorithm based statistical modeling and online process monitoring," *Chem. Intell. Lab. Syst.*, vol. 125, pp. 109–120, Jun. 2013.

[7] Y. Qin, C. Zhao, and F. Gao, "An iterative two-step sequential phase partition (ITSPP) method for batch process modeling and online monitoring," *AIChE J.*, vol. 62, no. 7, pp. 2358–2373, Feb. 2016.
[8] C. Zhao and F. Gao, "Critical-to-fault-degradation variable analysis and direction extraction for online fault prognostic," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 3, pp. 842–854, May 2017.
[9] S. Zhang, C. Zhao, and F. Gao, "Incipient fault detection for multiphase batch processes with limited batches," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 1, pp. 103–117, Jan. 2019.
[10] R. Guo, K. Guo, and J. Dong, "Phase partition and online monitoring for batch process based on multiway BEAM," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 4, pp. 1582–1589, Oct. 2017.
[11] L. Wang, X. He, and D. Zhou, "Average dwell time-based optimal iterative learning control for multi-phase batch processes," *J. Process Control*, vol. 40, pp. 1–12, Apr. 2016.
[12] J. M. González-Martínez, O. E. de Noord, and A. Ferrer, "Multisynchro: A novel approach for batch synchronization in scenarios of multiple asynchronisms," *J. Chemometrics*, vol. 28, no. 5, pp. 462–475, May 2014.
[13] T. Sun, H. Liu, H. Yu, and C. L. P. Chen, "Degree-pruning dynamic programming approaches to central time series minimizing dynamic time warping distance," *IEEE Trans. Cybern.*, vol. 47, no. 7, pp. 1719–1729, Jul. 2017.
[14] M. Fransson and S. Folestad, "Real-time alignment of batch process data using COW for on-line process monitoring," *Chemometrics Intell. Lab. Syst.*, vol. 84, nos. 1–2, pp. 56–61, Dec. 2006.
[15] Y. Li, K. Li, and S. Tong, "Finite-time adaptive fuzzy output feedback dynamic surface control for MIMO nonstrict feedback systems," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 1, pp. 96–110, Jan. 2019.
[16] W. Li, C. Zhao, and F. Gao, "Linearity evaluation and variable subset partition based hierarchical process modeling and monitoring," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2683–2692, Mar. 2018.
[17] Y. Liu, F. Wang, Y. Chang, R. Ma, and S. Zhang, "Multiple hypotheses testing-based operating optimality assessment and nonoptimal cause identification for multiphase uneven-length batch processes," *Ind. Eng. Chem. Res.*, vol. 55, no. 21, pp. 6133–6144, Jun. 2016.
[18] Y. Jiang, S. Yin, and O. Kaynak, "Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond," *IEEE Access*, vol. 6, pp. 47374–47384, Aug. 2018.
[19] J. Liu, T. Liu, and J. Zhang, "Window-based stepwise sequential phase partition for nonlinear batch process monitoring," *Ind. Eng. Chem. Res.*, vol. 55, no. 34, pp. 9229–9243, Jul. 2016.
[20] L. Luo, S. Bao, J. Mao, and D. Tang, "Phase partition and phase-based process monitoring methods for multiphase batch processes with uneven durations.," *Ind. Eng. Chem. Res.*, vol. 55, no. 7, pp. 2035–2048, Jan. 2016.
[21] Y. Wei, J. Qiu, P. Shi, and L. Wu, "A piecewise-Markovian Lyapunov approach to reliable output feedback control for fuzzy-affine systems with time-delays and actuator faults," *IEEE Trans. Cybern.*, vol. 48, no. 9, pp. 2723–2735, Sep. 2018.
[22] S. Zhang, C. Zhao, S. Wang, and F. Wang, "Pseudo time-slice construction using a variable moving window k nearest neighbor rule for sequential uneven phase division and batch process monitoring," *Ind. Eng. Chem. Res.*, vol. 56, no. 3, pp. 728–740, Jan. 2017.
[23] R. Wang and M. Zong, "Unsupervised feature selection based on self-representation and subspace learning," *World Wide Web*, vol. 21, no. 6, pp. 1745–1758, 2018.
[24] R. Johnson and D. Wichern, "Applied multivariate statistical analysis," in *Practical Multivariate Statistical Analysis*, 6th ed, Chinese, Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2008.



RUNXIA GUO received the M.S. degree from the Civil Aviation University of China, Tianjin, China, in 2008, and the Ph.D. degree from Tianjin University, Tianjin, in 2012. He is currently a Professor with the School of Electronic Information and Automation, Civil Aviation University of China. His current research interests include fault diagnosis and adaptive control techniques applied to civil aircrafts.



YANCHENG JIN received the B.S. degree from the University of Science and Technology Liaoning, Anshan, China, in 2017. He is currently pursuing the M.S. degree with the School of Electronic Information and Automation, Civil Aviation University of China, Tianjin, China.

His current research interests include fault diagnosis and health management, fault-tolerant control, and signal processing applied to civil aircrafts.