Condition-Driven Soft Transition Modeling and Monitoring Strategy for Complex Nonstationary Process

Xu Chen, Chunhui Zhao*

State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering,
Zhejiang University, Hangzhou, 310027, China
(Corresponding author, email: 3160102260@zju.edu.cn, chhzhao@zju.edu.cn)*

Abstract: Due to the frequent switching of operation condition, the real industrial processes show typical nonstationary characteristics. In some cases, the switching is frequent and may not be instantaneous, revealing typical transition characteristics different from steady operation. In this work, a condition-driven soft transition modeling and monitoring method is proposed to deal with this problem. The condition modes are obtained by an automatic sequential condition-mode division algorithm, then a fine-grained mode recognition strategy is developed to further separate the condition mode into steady and transition submodes. The steady submode model is established by designing a conditional autoencoder network which can more closely describe each steady submode and facilitate evaluation of the relationship between transition submode and each steady submode. Finally, an online monitoring strategy is designed which can capture the nonstationary process changes. A real industrial case illustrates the effectiveness and superiority of the proposed method, which establishes a more accurate model for nonstationary processes by revealing the transition feature.

Keywords: condition-driven; nonstationary process monitoring; soft transition modeling; mode recognition; conditional autoencoder network.

1. INTRODUCTION

During the last few decades, data-driven technologies (Jackson et al., 1991; Cuentas et al., 2017) have been widely focused on to ensure successful operation of industrial processes. Some typical methods, including principal component analysis (PCA) (Jackson et al., 1991), partial least squares (PLS) (Geladi et al., 1986), canonical correlation analysis (CCA) (Lin et al., 2018), construct feature spaces in different ways and analyse the measurement data from respective perspectives. However, these methods are only suited to stationary processes, which is not satisfied in many industrial cases. Under the influence of operating condition switching and production products changes, the real industrial processes often show non-stationary characteristics, which is reflected by a time-variant mean, a time-variant autocovariance, or both (Chen et al., 2021). The characteristics of non-stationary processes change along with time, which cannot be well described by traditional model (Zhang et al., 2020).

Some existing methods may be used to address this problem. Cointegration analysis (CA) is an effective method to model the nonstationary process, which investigates the long-term cointegration relationships among nonstationary variables (Engle et al., 1987). Zhao et al. proposed a full-condition monitoring method which combined CA and slow feature analysis (SFA) to monitor the static and dynamic variations of multifarious operation conditions (Zhao et al., 2018). However, it is difficult to guarantee that nonstationary variables can be integrated in the same order, and at that case, CA will fail to handle the nonstationary processes. Multimode process

modeling strategies divide the data into different modes by clustering methods, and then describe each mode respectively (Yu et al., 2008). Nevertheless, it is difficult to judge which mode the current example belongs to for online application. It may lead to a bad result if the wrong model is applied.

Neither the CA nor the multimode approaches consider the transition characteristics, which are common in nonstationary processes since the switching of operating conditions is not always instantaneous. If these data are strictly partitioned as a certain steady operation, it may lead to false alarm or missing alarm, since the model of steady operation is insufficient to describe the transition characteristics (Zhao et al., 2007; Zhao et al., 2015). In the resent years, the nonstationary processes modeling methods focusing on soft transition have attracted increasing attention. Adaptive strategy, which frequently updates the model as new samples arrives, is able to deal with this problem (Yu et al., 2019; Wu et al., 2020). Just-in-time-learning (JITL) was combined with canonical correlation analysis (CCA) by Chen et al. to monitor the nonlinear nonstationary processes (Chen et al., 2020). However, it is perplexing for adaptive methods to distinguish different types of normal operation switching from real faults. The frequent updating of models also leads to great computational complexity. Qin et al. combined the finite Gaussian mixture model (FGMM) and Bayesian inference strategy to compute the posterior probabilities of each monitored sample belonging to the multiple components (Yu et al., 2008). Nevertheless, the number of Gaussian components is hard to determine in practice. Zhao et al. proposed a soft transition multiple PCA (STMPCA) to handle the hard-partition and misclassification problems of multimodal batch process monitoring (Zhao et al.,

2007). Whereas, this time-driven approach can be incredibly complex. The operation conditions change irregularly and frequently over time, which increases the difficulty in the identification of transition operation

Recently, Zhao et al. proposed a condition-driven data analytics and monitoring strategy for nonstationary and transient processes (Zhao et al., 2020). It reveals that although the characteristics of nonstationary process change with time, the process might follow certain relations within the same condition. An automatic sequential condition-mode division (SCMD) algorithm is developed to divided the data into different condition modes following changes of condition indicator, and the characteristics are similar in the same condition mode and significantly different for different modes. This method fuses the accessible prior knowledge into the model through condition indicator, improving the accuracy of model. But it does not consider the different transition characteristics between different condition modes.

In this paper, a condition-driven soft transition modeling and monitoring method is proposed for nonstationary process. First, based on the SCMD algorithm, the process data are initially divided into different condition modes. Second, a fine-grained mode recognition strategy is proposed to divide the data of each condition mode into steady and transition submodes. Then a conditional autoencoder network is designed for modeling the steady submodes data, which describes each steady submode closely and highlights the the difference across steady submodes synchronously. The relationship between transition submodes and each steady submode are further explored. Finally, an online strategy is developed to monitoring both the steady and transition submodes simultaneously. A real industrial case proves the effectiveness of the algorithm.

The remainder of this article is described as follows. In Section II, automatic sequential condition-mode division (SCMD) algorithm is briefly revisited. The motivation and the proposed method are introduced in Section III. Then we present the application results of this method subsequently. The conclusion is drawn in the last section.

2. REVISIT OF SCMD ALGORITHM

In this section, automatic sequential condition-mode division (SCMD) (Zhao et al., 2020) algorithm is briefly revisited. In industrial processes, some variables are directly related to the working condition and have an important influence on the whole process, which are named condition indicators. The process data with similar condition indicator value have similar characteristics, and should be modelled together, although they are at the different time. Therefore, it is a critical issue to determine which condition-mode data should be put together.

The main steps of automatic SCMD algorithm can be summarized as follows.

Step 1 (Data Preparation). Condition indicator is applied to reorganize the measurement data $\mathbf{x}_t$. The values of condition are reordered to increase monotonously. Then define the

condition interval β to separate the condition values into M conditions, which is named condition slice $\mathbf{X}_m$ ($N_m \times J$) here, where J is the number of variables. These condition slices are standardized, which are denoted as $\bar{\mathbf{X}}_m$. The size of β is adjustable to adapt to cover enough examples.

Step 2 (Condition-Slice-Based Modeling). Perform slow feature analysis (SFA) algorithm on the condition slices $\bar{\mathbf{X}}_m$ to get the condition-slice models for both slow parts $\mathbf{W}_{s,m}$ ($J \times R$) and fast parts $\mathbf{W}_{f,m}$ ($J \times (J-R)$), where R is the number of slow features (SFs). SFA combines the original variables to obtain new latent variables, which are arranged according to the change rate from slow to fast. Slow latent variables ($\mathbf{S}_{s,m}$), which are also named slow features, are generally considered to contain the essential information of the process, while fast features ($\mathbf{S}_{f,m}$) represent noise information. They are expressed as

$$\mathbf{S}_{s,m} = \bar{\mathbf{X}}_m \mathbf{W}_{s,m} \qquad (1)$$

$$\mathbf{S}_{f,m} = \bar{\mathbf{X}}_m \mathbf{W}_{f,m} \qquad (2)$$

Then calculate the monitoring statistics of slow features $\mathbf{S}_{s,m}$ and fast features $\mathbf{S}_{f,m}$, and determine the control limits $\mathrm{Ctr}_{s,m}$ and $\mathrm{Ctr}_{f,m}$ by kernel density estimation (KDE) method (Botev et al., 2010). The statistics of slow and fast features are designed as

$$T_{s,m}^2 = s_{s,m} s_{s,m}^T \qquad (3)$$

$$T_{f,m}^2 = s_{f,m} s_{f,m}^T \qquad (4)$$

Step 3 (Condition-Segment-Based SFA Modeling). The characteristics of adjacent condition slices are similar to each other, which is reflected that the SFA models and control limits are also similar. These similar condition slices should be aggregated into large segment and modelled together, which is named condition segment here. Set the first condition slice (the kth condition slice) that have not been aggregated into segment as a new condition segment $\mathbf{X}_{v,k}(N_k \times J)$. Then add the next condition slice $\mathbf{X}_{v,k+1}$ ($N_{k+1} \times J$) to the segment $\mathbf{X}_{v,k+1}$ $((N_k + N_{k+1}) \times J)$. Perform SFA on the new rearranged data matrix and calculate the monitoring statistic again. After that, the new control limits $\mathrm{Ctr}_{v,s,k+1}$ and $\mathrm{Ctr}_{v,f,k+1}$ are obtained.

Step 4 (Compare Model Accuracy). Compare $\mathrm{Ctr}_{v,s,k+1}$ with $\mathrm{Ctr}_{s,k+1}$ and $\mathrm{Ctr}_{v,f,k+1}$ with $\mathrm{Ctr}_{f,k+1}$ for each condition slice. If their relationship satisfies (5) and (6), it means that the $k+1$th conditional slice is similar to the previous condition segment, which should be incorporated together for modelling.

$$\mathrm{Ctr}_{v,s,k+1} \le \alpha * \mathrm{Ctr}_{s,k+1} \qquad (5)$$

$$\mathrm{Ctr}_{v,f,k+1} \le \alpha * \mathrm{Ctr}_{f,k+1} \qquad (6)$$

Here α is a constant termed relaxing factor. It determines how much the condition-segment SFA model is allowed to be less representative than the condition-slice SFA models. If the condition $k*$ from which three consecutive samples cannot satisfy (5) and (6), we judge that the addition of the current condition slice is not adapted to the condition segment, which should be split off from the current condition segment, and set as the beginning of the next condition segment.

Step 5 (Data Updating and Recursive Implementation). Repeat Steps 3-4 to find the condition segments until to the end. The characteristics within each condition segment are similar, while the characteristics of different condition segments are great difference. Therefore, each condition segment can be treated as a condition mode and analysed respectively.

The illustration of SCMD algorithm is shown in Fig. 1. In automatic sequential condition-mode division algorithm, every condition slice is divided into a certain condition mode, paying no attention to transition processes, which is not consistent with the facts. Besides, SFA is a linear method, which is not able to explore the nonlinear relationship of the process.
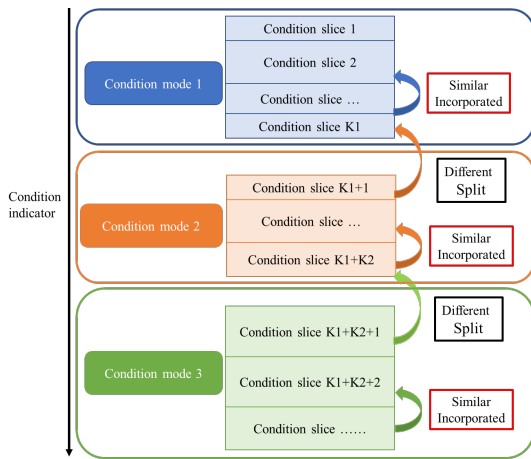

Fig. 1. The illustration of SCMD algorithm. Process data are divided into different condition slices, and the slices with similar characteristics are aggregated into the same condition mode.

3. THE PROPOSED SOFT TRANSITION METHOD

3.1 Motivation

With the change of product requirements or the switching of operation conditions, complex nonstationary industrial processes may operate in different steady operations, which cannot be covered by the traditional monomodal method. It is reasonable to divide data with different characteristics into several modes and model them separately. However, in non-stationary processes, transitional data are worthy of attention. The transition characteristics are generated between different steady mode switching. The characteristics of transition data collected at the beginning of the switch is similar to the previous steady data. With the change of operation condition, the previous model can no longer describe these transition data. Finally, the transition process enters into another steady operation. Therefore, it is inaccurate to divide the transition data into one steady model, which is prone to false alarm or missing alarm due to the insufficient descriptive ability of the original model.

Under the idea of condition-driven, we further divide each condition mode into steady and transition submodes. For example, in Condition mode A, some data obtained when the industrial process operates stably under Condition mode A, while other data is collected only when the process is running temporarily through Condition mode A . These two kinds of data are different, and the latter cannot be described by Model A alone, which is also correlated with other models. There are both steady and transition submode data in the same conditional mode. We should divide these data and model the data separately. The number of condition modes according to the condition indicator is finite, which reduces the complexity of the model.

3.2 Steady and Transition Submode Partition

As mentioned above, steady and transition submode data may be falsely divided into the same condition mode. It is deemed that steady submode data possess similar characteristics, whose distribution is relatively close and compact. In contrast, the quantity of transition submode data is smaller and the characteristics are difference from others, which are more likely to be scattered at the edge of steady distribution. In order to capture the nonlinear relationship, autoencoder (AE) is applied to reduce the dimension of the original data, which is consisted with an encoder and a decoder (Bengio et al., 2013). The encoder reduces the dimension of the data, while the decoder uses low dimension features to reconstruct the original data. If the error of reconstruction is small, it means that these low dimension features have well covered the information of the original data, so as to achieve the purpose of dimension reduction. The mathematical is expressed as

$$\mathbf{F}_m = f_{encoder}(\bar{\mathbf{X}}_m) \tag{7}$$

$$\hat{\bar{\mathbf{X}}}_m = f_{decoder}(\mathbf{F}_m) \tag{8}$$

where $f_{encoder}(\cdot)$ and $f_{decoder}(\cdot)$ are the function of encoder and decoder, $\mathbf{F}_m$ are the low dimension features, and $\hat{\bar{\mathbf{X}}}_m$ is the reconstruction data. The objective of AE is to minimum $\bar{\mathbf{X}}_m$ and $\hat{\bar{\mathbf{X}}}_m$.

Then a density-based spatial clustering method (DBSCAN) (Kumar, 2016) is performed on the low dimension features to divide the data into two submodes. DBSCAN algorithm defines cluster as a region of densely connected points separated by regions of non-dense points. The points are divided into three categories: core points, border points, and noise points. The core radius $r_{neighbor}$ and the minimum adjacent points $N_{neighbor}$ are defined first. If a point has more than

$N_{neighbor}$ points within $r_{neighbor}$, that point will be considered a core point. If a point is within the range of a core point, it is considered a border point. And the rest are noise points. The illustration of three kinds of points is shown in Fig. 2. Here, we consider that both core points and boundary points are closely distributed, which are classified as steady submode data, while the other points (noise points) are recognized as transition submode data.
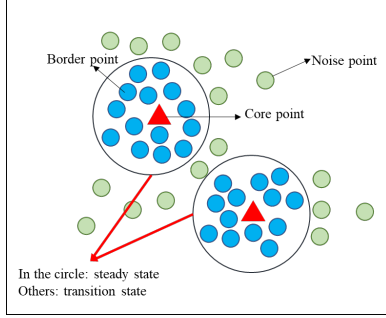

Fig. 2. Illustration: DBSCAN and submode partition. Different colors represent different categories of data.

3.3 Steady Submode Model

In Section 3.2, we divide the data into two submodes, and in this section, we establish the model for steady submode data.

Inspired by the conditional variational autoencoder (CVAE) (Yan et al., 2016), we encode the information of condition mode as a one hot vector, which is spliced with the original data, thus adding the model information. For example, if there are C modes, the information of first mode is noted as the one hot vector $\mathbf{v}_{mode,1} = [0,0,...,0,1]$ ($1 \times C$). Then catenate the first mode data $\mathbf{X}_{mode,1}$ ($N \times J$) with $\mathbf{v}_{mode,1}$ to get the new data $\mathbf{X}_{c\_mode,1}$ ($N \times (J+C)$) with modal information. Put the catenated data $\mathbf{X}_{c\_mode,1}$ into AE for training, and its mathematical is written as

$$\mathbf{F}_{c\_mode,1} = f_{c\_encoder}(\mathbf{X}_{c\_mode,1}) \quad (9)$$

$$\hat{\mathbf{X}}_{mode,1} = f_{c\_decoder}(\mathbf{F}_{c\_mode,1}) \quad (10)$$

where $f_{c\_encoder}(\bullet)$ and $f_{c\_decoder}(\bullet)$ are the function of encoder and decoder, $\mathbf{F}_{c\_mode,1}$ ($N \times R$) are the low dimension features, R is the dimension of features and $\hat{\mathbf{X}}_{mode,1}$ is the reconstruction data. The reconstruction loss is expressed as

$$\text{reconstruction loss} = \left\| \mathbf{X} - \hat{\mathbf{X}} \right\|^2 \quad (11)$$

where $\mathbf{X}$ is the steady submode data, and $\hat{\mathbf{X}}$ is the reconstitution data. We apply autoencoder structure here since neural network method is able to capture the nonlinear information, which cannot be handled by SFA.

For process monitoring, two statistics $A^2_{mode,k}$ and $SPE_{mode,k}$ are designed. The former statistic monitors the principal component subspace, while the latter monitors the residual subspace, which are expressed as

$$A^2_{mode,k} = (\mathbf{f}_{mode,k} - \overline{\mathbf{f}}_{mode,k}) \Sigma^{-1}_{mode,k} (\mathbf{f}_{mode,k} - \overline{\mathbf{f}}_{mode,k})^T \quad (12)$$

$$SPE_{mode,k} = \mathbf{e}_{mode,k} \mathbf{e}^T_{mode,k} \quad (13)$$

where k is the kth mode, $\mathbf{f}_{mode,k}$ is an example of $\mathbf{F}_{c\_mode,k}$, $\overline{\mathbf{f}}_{mode,k}$ is the mean of $\mathbf{F}_{c\_mode,k}$, $\Sigma_{mode,k}$ is a diagonal matrix with the variance of $\mathbf{F}_{c\_mode,k}$, and $\mathbf{e}$ is the residual vector. The control limits are estimated by KDE. The architecture of neural network is shown in Fig. 3.
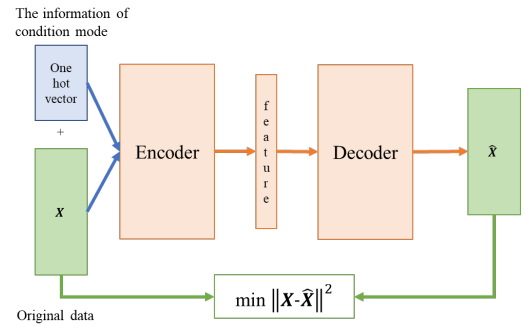

Fig. 3. The architecture of the proposed steady submode model.

3.4 Transition Submode Model

Even within the same condition mode, the transition data are generated from different condition switching, so the characteristics can be quite different. Transition submode data in one condition mode may be related to all steady submode models. To describe the transition data, an attention-based soft transition modeling method is proposed, where all the steady submode models are combined to explore the transition characteristics. The allocation of different model weights is estimated by attention mechanism.

The transition data of all condition modes are noted as $\mathbf{X}_{trans}$. Then the $\mathbf{X}_{trans}$, catenated with the one hot vectors of different modes, are put into the trained encoder ($f_{c\_encoder}(\bullet)$) to obtain the characteristics ($\mathbf{F} = [\mathbf{F}_{c\_mode,1}, \mathbf{F}_{c\_mode,2}, ..., \mathbf{F}_{c\_mode,C}]$) of each mode. For an example of transition submode data $\mathbf{x}_{trans}$, every steady submode feature $\mathbf{f}_{c\_mode,k}$ ($1 \times R, k = 1, ..., C$) may related to it. It is necessary to assign different weights to each steady submode feature, which can be realized by attention mechanism. A new encoder model is trained to obtain the weight assignment vector $\mathbf{v}_{score}$ ($1 \times R$), and calculate the weight of every steady submode.

$$\mathbf{v}_{score} = f_{trans\_encoder}(\mathbf{x}_{trans}) \quad (14)$$

$$\mathbf{w}_k = \mathbf{f}_{c\_mode,k} \mathbf{v}^T_{score} (k = 1, 2, ..., C) \quad (15)$$

Then standardize the weight coefficient.

$$\hat{w}_k = \frac{exp(w_k)}{\sum_{i=1}^{C} exp(w_i)} \quad (16)$$

The new feature is obtained by (18)

$$\mathbf{f}_{trans} = \sum_{k=1}^{C} \hat{w}_k \mathbf{f}_{c\_mode,k} \quad (17)$$

Finally, train a decoder to reconstruct $\mathbf{f}_{trans}$ into original data.

$$\hat{\mathbf{x}}_{trans} = f_{trans\_decoder}(\mathbf{f}_{trans}) \quad (18)$$

Analogously, we develop two statistics A^2_{trans} and SPE_{trans} for process monitoring, which are expressed as

$$A^2_{trans} = (\mathbf{f}_{trans} - \overline{\mathbf{f}}_{trans}) \Sigma^{-1}_{trans} (\mathbf{f}_{trans} - \overline{\mathbf{f}}_{trans})^T \quad (19)$$

$$SPE_{trans} = \mathbf{e}_{trans} \mathbf{e}^T_{trans} \quad (20)$$

where Σ_{trans} is a diagonal matrix with the variance of $\mathbf{F}_{trans} = [\mathbf{f}_{trans,1}, \mathbf{f}_{trans,2}, ..., \mathbf{f}_{trans,N}]^T$ and $\mathbf{e}$ is equal to $\mathbf{x}_{trans} - \hat{\mathbf{x}}_{trans}$.

4. CASE STUDY

In this section, the proposed method is illustrated by a nonstationary process of coal mill, which is one of the most important machines for ultrasupercritical unit of the thermal power plant. This machine works under varying operation conditions, and the coal feed rate is the key input to the coal mill, which is taken as the condition indicator. In this case, 35 measured variables are collected for process monitoring. There are 10000 examples for training, which covers various changes in operation conditions. The value of coal feed rate varies from 35.44 ton/h to 83.47 ton/h. According to previous work (Zhao et al., 2020), the nonstationary process can be partitioned into eight condition modes. The partition point of the indicator variable is 43.36, 44.32, 45.28, 63.04, 67.36, 76.96, and 77.92. As mentioned before, each condition mode contains both steady submode data and transition submode data. Here, we reduce the dimension of each condition mode data through AE, and the low dimensional features are clustered by DBSCAN. Steady and transition submode data are modelled in their respective methods.

A fault case is presented to verify the performance for fault detection. The temperature becomes abnormal from the 750th samples in this case. The performance of proposed method is shown in Fig. 4. T^2 statistic represent the principal component space, which is equivalent to A^2, and SPE represents the information of the residual space. Either of these two statistics exceed the control limit can be considered that the fault has occurred (here we use the logarithm of statistics for more clear comparison). As shown from Fig. 4, the performance of SPE is more accurate than T^2. It demonstrates that when the temperature becomes abnormal, the influence on the residual space is stronger than principal component space, which lead to earlier detection on SPE. It is noted that around the 250th

sample, both T^2 and SPE show disturbances and go beyond the control limits. At that time, the coal feed rate is much less than 43.36 ton/h. There are very few similar samples in the training sample, so it is assumed that an anomaly has occurred.

For comparison, another two methods are applied to explain the necessity of condition mode partition and soft transition modelling. The first method is global AE, which does not divide the data into different modes. The performance of global AE is shown in Fig. 5. According to the result, it can be found that there are a lot of false alarming around the 250th sample and the 700th example. Due to the influence of the other mode data, the modelling accuracy of global AE for the mode with fewer training samples is greatly reduced. The second method is hard partition method, which pays no attention to transition process (Zhao et al., 2020). And the performance of this method is shown in Fig. 6. The results show that it is unreasonable to hardly divide a transition sample into a steady submode. When the steady submode model is insufficient to describe the transition data, the transition data may be treated as an anomaly. Besides, without detailed modelling of the transition submode data, it may be difficult to find outliers in the transition process when it is directly described by a broader steady submode model. Therefore, the outliers around the 1600th sample are not detected. The false alarming rates (FAR) and missing alarming rates (MAR) of three methods are listed in Table 1.
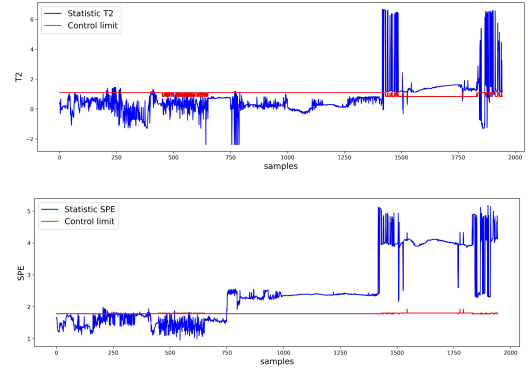

Fig. 4. Monitoring results for two statistics using the proposed method. The blue line denotes the statistics and the red line denotes the control limit.
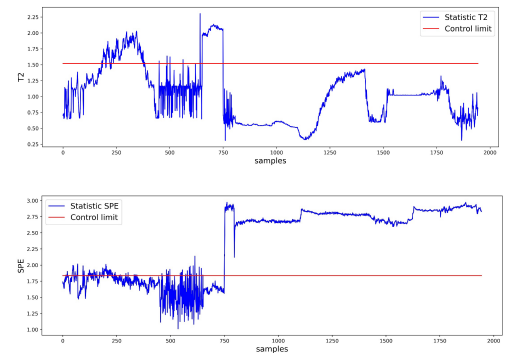

Fig. 5. Monitoring results for two statistics using global AE. The blue line denotes the statistics and the red line denotes the control limit.
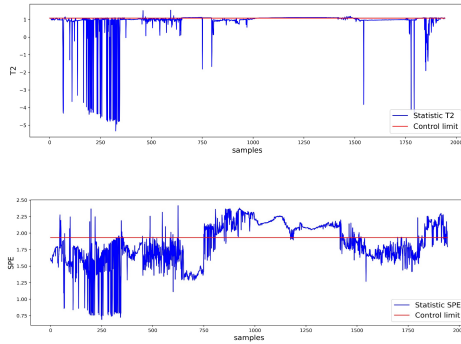
Fig. 6. Monitoring results for two statistics using hard partition method. The blue line denotes the statistics and the red line denotes the control limit.

Table 1. Monitoring comparison evaluated with FAR and MAR for three methods

	The proposed method	Global AE	Hard partition method
FAR	0.132	0.501	0.225
MAR	0.000	0.000	0.310

5. CONCLUSION

In this work, a condition-driven soft transition modeling and monitoring strategy for nonstationary process is proposed. For each condition mode which are available by SCMD algorithm, the data are further divided into steady submode and transition submode, which are then modelled and monitored in a fine-grained way. The changes of nonstationary process can be captured, which can reveal whether the process stays in one steady status or is switching dynamically. Based on the concept of condition-driven, the transition processes are easily identified and modelled, which is a cumbersome task for time-driven methods since the switching along time can be very complex. An industrial case proves the effectiveness and superiority of the proposed algorithm. Compared with global model and hard partition model, the proposed algorithm has lower false alarming rate and missing alarming rate.

ACKNOWLEDGMENT

REFERENCES

Bengio, Y., Courville, A., Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, *35*(8), 1798-1828.

Chen, J., & Zhao, C. (2021). Exponential Stationary Subspace Analysis for Stationary Feature Analytics and Adaptive Nonstationary Process Monitoring. *IEEE Transactions on Industrial Informatics.*

Chen, Z., Ding, S. X., Peng, T., Yang, C., Gui, W. (2017). Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms. *IEEE Transactions on Industrial Electronics*, *65*(2), 1559-1567.

Chen, Z., Liu, C., Ding, S., Peng, T., Yang, C., Gui, W., Shardt, Y. (2020). A Just-In-Time-Learning Aided Canonical Correlation Analysis Method for Multimode Process Monitoring and Fault Detection. *IEEE Transactions on Industrial Electronics*.

Cuentas, S., Peñabaena-Niebles, R., Garcia, E. (2017). Support vector machine in statistical process monitoring: a methodological and analytical review. *The International Journal of Advanced Manufacturing Technology*, *91*(1-4), 485-500.

Geladi, P., Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica chimica acta*, *185*, 1-17.

Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). Kernel density estimation via diffusion. The annals of Statistics, 38(5), 2916-2957.

Jackson, J. E. (2005). *A user's guide to principal components* (Vol. 587). John Wiley & Sons.

Kumar, K. M., Reddy, A. R. M. (2016). A fast DBSCAN clustering algorithm by accelerating neighbor searching using Groups method. *Pattern Recognition*, *58*, 39-48.

Tan, S., Wang, F., Peng, J., Chang, Y., Wang, S. (2012). Multimode process monitoring based on mode identification. *Industrial & Engineering Chemistry Research*, *51*(1), 374-388.

Wu, H., Zhao, J. (2020). Self-adaptive deep learning for multimode process monitoring. *Computers & Chemical Engineering*, *141*, 107024.

Yan, X., Yang, J., Sohn, K., Lee, H. (2016, October). Attribute2image: Conditional image generation from visual attributes. In *European Conference on Computer Vision* (pp. 776-791). Springer, Cham.

Yu, J., Qin, S. J. (2008). Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models. *AIChE Journal*, *54*(7), 1811-1829.

Yu, W., Zhao, C. (2018). Recursive exponential slow feature analysis for fine-scale adaptive processes monitoring with comprehensive operation status identification. *IEEE Transactions on Industrial Informatics*, *15*(6), 3311-3323.

Zhang, H., Shang, J., Yang, C., Sun, Y. (2020). Conditional Random Field for Monitoring Multimode Processes with Stochastic Perturbations. *Journal of the Franklin Institute*.

Zhao, C., Chen, J., Jing, H. (2020). Condition-Driven Data Analytics and Monitoring for Wide-Range Nonstationary and Transient Continuous Processes. *IEEE Transactions on Automation Science and Engineering*.

Zhao, C., Huang, B. (2018). A full-condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis. *AIChE Journal*, *64*(5), 1662-1681.

Zhao, C., Wang, F., Lu, N., Jia, M. (2007). Stage-based soft-transition multiple PCA modeling and on-line monitoring strategy for batch processes. *Journal of Process Control*, *17*(9), 728-741.

Zhao, C., Wang, W., Qin, Y., Gao, F. (2015). Comprehensive subspace decomposition with analysis of between-mode relative changes for multimode process monitoring. *Industrial & Engineering Chemistry Research*, 54(12), 3154-3166.