ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Linear Chain Conditional Random Field for Operating Mode Identification and Multimode Process Monitoring

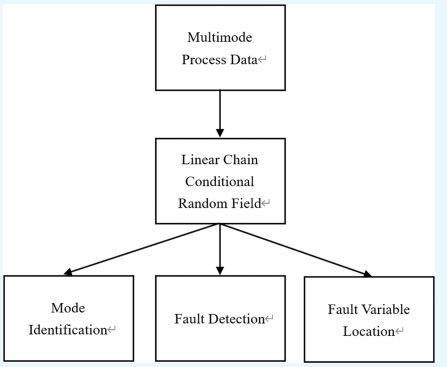
Fan Wang*

Cite This: *ACS Omega* 2022, 7, 29483−29494

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: As a supervised machine learning algorithm, conditional random fields are mainly used for fault classification, which cannot detect new unknown faults. In addition, faulty variable location based on them has not been studied. In this paper, conditional random fields with a linear chain structure are utilized for modeling multimode processes with transitions. A linear chain conditional random field model is trained by normal data with mode label. This model is able to distinguish transitions from stable modes well. After mode identification, the expectation of state feature function is developed for fault detection and faulty variable location. Case studies on the Tennessee Eastman process and continuous stirred tank reactor (CSTR) testify the effectiveness of the proposed approach.



1. INTRODUCTION

Process monitoring in process industry has always been a challenging problem. Massive process data from process control computers and manufacturing execution systems contain fruitful information about the production status. Arunthavanathan et al.[1] analyze process fault diagnosis approaches from safety perspectives. Risk-based methods for fault detection develop risk indication, which creates a new concept.[2−4] Hybrid or integrated approaches exploit multiple algorithms for better performance. Don and Khan use the hidden Markov model for abnormality detection and Bayesian network (BN) for root diagnosis.[5] Amin, Imtiaz, and Khan propose a hybrid approach using PCA and BN.[6] Machine learning, acting as a data analysis tool, has attracted researchers' and engineers' attention in last decades.[7−11] Although successful and promising advances have been made in process data analytics, opportunities still exist due to its three important and unique attributes.[12]

Many efforts based on machine learning have also been made for multimode process monitoring. Clustering and classifying approaches are often used. Hard clustering/classifying divides each sample to a particular cluster/class, such as *k*-means and its variants. Soft clustering/classifying uses probability, such as Gaussian mixture model,[13] mixture of probabilistic principal component analysis,[14] support vector data description,[15] Bayesian decision theory,[16] hidden Markov model (HMM),[17,18] Bayesian network,[19,20] and so on. This probability is often utilized to combine monitoring results. Window-based methods take time series information into account, such as dynamic mutual information similarity,[21] moving window hidden Markov model,[22,23] and so on.

In the meantime, supervised and semisupervised methods are used as classifiers for fault classification and identification. Yu, Khan, and Garaniya use a three-layer nonlinear Gaussian belief network for diagnosis.[24] Yu utilizes a multiway discrete hidden HMM to transform three-dimensional data into two-dimensional for batch processes.[25] Sen et al. combine segmental k-means with the HMM to improve classification accuracy.[26] Yu revises emitting distribution of the HMM to capture high-order statistical information.[27] Sammaknejad et al. develop condition diagnosis based on the transition probability of the HMM.[28] Fang et al. construct a marginalized conditional random field (CRF) for fault diagnosis.[29] After that, Fang et al. design a switching CRF based on multiple linear chain CRFs for multimode fault diagnosis.[30] As a powerful discriminative probabilistic model, a CRF is also used for working condition classification in oil drilling[31] and sucker rod pumps.[32] Similar to other machine learning algorithms, CRFs can also be directly used for modeling. Zhang et al. design a fault detection index based on log conditional probability of the CRF.[33] Arunthavanathan et al.[34] utilize convolution neural network and long short-term memory for fault sequence learning.

However, using CRF as a fault classifier cannot recognize new unknown faults because CRF is a supervised approach.

Received: June 27, 2022
Accepted: August 3, 2022
Published: August 12, 2022

Therefore, establishing a monitoring model based on the CRF may be a better choice. In addition, identifying transitions between stable operating modes has not been studied intensively. Faulty variable location based on the CRF also remains unsolved.

Targeting the above-mentioned problems, a historical normal dataset with mode label is used to train a linear chain conditional random field model (LCCRF). Operating mode identification is achieved by the Viterbi algorithm. During the test phase, this model is able to distinguish stable modes and transitions well. Subsequently, a new fault detection index called expectation of state feature function (ESFF) is proposed. Faulty variable location based on it has a specific and interpretable engineering physical meaning. At last, applications of the Tennessee Eastman process and continuous stirred tank reactor verify effectiveness of the proposed method.

There are two goals to achieve in this paper. The first objective is better normal operating mode identification. The step-by-step procedure is shown as follows. First, construct a historical dataset by labeling each normal sample with its operating mode. Operating modes include stable modes and transitions between them. Then, normalize this dataset using the mean and standard deviation. Subsequently, train an LCCRF using the dataset. When one new online sample comes, its operating mode label can be obtained by the Viterbi algorithm. The second objective is fault detection and fault variable location. The step-by-step procedure is demonstrated below. First, build a historical dataset with normal data. The labels of training data are known. Then, normalize the dataset utilizing the mean and standard deviation. After that, train the LCCRF model via a scaled conjugate gradient algorithm. A new monitoring index named ESFF is calculated for each normal sample. The control limit of every operating mode is separately computed by kernel density estimation afterward. At last, when an online sample comes, calculate its ESFF and compare it with the control limit. If this online sample is regarded as faulty, contribution of every monitored variable can be calculated by ESFF of the variable. Process variables with remarkable contribution are considered fault variables.

The remainder of this manuscript is organized as follows. Section 2 provides a brief introduction to LCCRFs. Section 3 gives the structure of the LCCRF, the construction of ESFF, faulty variable location method, and procedure of methodology. Applications of the Tennessee Eastman process and continuous stirred tank reactor are shown in Section 4. Section 5 presents the conclusion.

2. LINEAR CHAIN CONDITIONAL RANDOM FIELDS

A CRF is just a conditional distribution $p(y|x)$, where output variables y are attributes of entities and input variables x are observed knowledge about the entities.[35] The job of a CRF is to model the conditional distribution using an associated graphical structure. An LCCRF is one important case. Its conditional distribution has the following forms.[36] The probability form is

$$p(y|x) = \frac{1}{Z(x)}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1}, y_i, x, i) + \sum_{i,l}\mu_l s_l(y_i, x, i)\right) \quad (1)$$

$$Z(x) = \sum_{y}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1}, y_i, x, i) + \sum_{i,l}\mu_l s_l(y_i, x, i)\right) \quad (2)$$

where $t_k(\cdot)$ is the feature function of the edge that describes the transition relationship between y_{i-1} and y_i. $s_l(\cdot)$ is the feature function of the node that represents the current state about y_i and x.

Summing the transition and state feature functions at each position i gives the simplified form.

$$f_k(y_{i-1}, y_i, x, i) = \begin{cases} t_k(y_{i-1}, y_i, x, i), & k = 1, 2, ..., K_1 \\ s_l(y_i, x, i), & k = K_1 + l;\ l = 1, 2, ..., K_2 \end{cases} \quad (3)$$

$$f_k(y, x) = \sum_{i=1}^{n} f_k(y_{i-1}, y_i, x, i),\ k = 1, 2, ..., K \quad (4)$$

$$\omega_k = \begin{cases} \lambda_k, & k = 1, 2, ..., K_1 \\ \mu_l, & k = K_1 + l,\ l = 1, 2, ..., K_2 \end{cases} \quad (5)$$

$$p(y|x) = \frac{1}{Z(x)}\exp\left(\sum_{k=1}^{K}\omega_k f_k(y, x)\right) \quad (6)$$

$$Z(x) = \sum_{y}\exp\left(\sum_{k=1}^{K}\omega_k f_k(y, x)\right) \quad (7)$$

The abundant selection of the feature function makes the LCCRF complex, diverse, and powerful. Similar to the HMM, there are three primary problems: calculation of conditional probability, parameter estimation of an LCCRF model, and state inference.

Given training dataset $D = \{x^i, y^i\}_{i=1}^N$, where x^i means inputs and y^i represents corresponding label/state. Parameter estimation is carried out by penalized maximum likelihood. The conditional log likelihood is

$$l(\theta) = \sum_{i=1}^{N}\log p(y^i|x^i) \quad (8)$$

After substituting the LCCRF model (eq 6) into eq 8, the following expression is obtained.

$$l(\theta) = \sum_{i=1}^{N}\sum_{k=1}^{K}\omega_k f_k(y^i, x^i) - \sum_{i=1}^{N}\log Z(x^i) \quad (9)$$

To avoid overfitting, regularization which is a penalty on weight vectors is added. The regularized log likelihood is

$$l(\theta) = \sum_{i=1}^{N}\sum_{k=1}^{K}\omega_k f_k(y^i, x^i) - \sum_{i=1}^{N}\log Z(x^i) - \sum_{k=1}^{K}\frac{\omega_k^2}{2\sigma^2} \quad (10)$$

The partial derivatives of eq 10 are

$$\frac{\partial l}{\partial w_k} = \sum_{i=1}^{N} f_k(y^i, x^i) - \sum_{i=1}^{N}\sum_{y,y'} f_k(y, y', x^i)p(y, y'|x^i) - \sum_{k=1}^{K}\frac{\omega_k}{\sigma^2} \quad (11)$$

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

The first term is expectation of f_k given the empirical distribution.

$$\tilde{p}(y, x) = \frac{1}{N}\sum_{i=1}^{N} 1_{\{y=y^i\}}1_{\{x=x^i\}} \tag{12}$$

The second term is expectation of f_k given the model distribution $p(y|x;\theta)\tilde{p}(x)$. Therefore, these two expectations should be equal when the gradient is zero in an unregularized maximum likelihood method. This important and pleasing interpretation is predictable in exponential family maximum likelihood estimation.

A scaled conjugate gradient algorithm is used for optimizing $l(\theta)$ in this work because quasi-Newton approaches are often numerically unstable. The forward−backward algorithm is utilized to calculate the conditional probability.

For each i, a matrix of order m is defined, where m stands for the number of states y_i.

$$M_i(x) = [M_i(y_{i-1}, y_i|x)] \tag{13}$$

$$M_i(y_{i-1}, y_i|x) = \exp(W_i(y_{i-1}, y_i|x)) \tag{14}$$

$$W_i(y_{i-1}, y_i|x) = \sum_{i=1}^{K} f_k(y_{i-1}, y_i, x, i) \tag{15}$$

Forward variables $\alpha_i(x)$ are defined:

$$\alpha_0(y|x) = \begin{cases} 1, & y = y_0 = \text{start} \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

$$\alpha_i^{\mathrm{T}}(y_i|x) = \alpha_{i-1}^{\mathrm{T}}(y_{i-1}|x)M_i(y_{i-1}, y_i|x),\ i = 1, 2, \ldots, n+1, y_{n+1} = \text{stop} \tag{17}$$

Backward variables $\beta_i(x)$ are defined:

$$\beta_{n+1}(y_{n+1}|x) = \begin{cases} 1, & y = y_{n+1} = \text{stop} \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

$$\beta_i(y_i|x) = M_i(y_i, y_{i+1}|x)\beta_{i-1}(y_{i+1}|x) \tag{19}$$

Normally, the Viterbi algorithm[35−37] is used for state inference as eq 20.

$$y^* = \arg\max P(y|x) \tag{20}$$

3. MODE IDENTIFICATION AND PROCESS MONITORING USING LINEAR CHAIN CONDITIONAL RANDOM FIELD

3.1. Structure of Linear Chain Conditional Random Field. The design of transition and state feature functions leads to different structures of LCCRF. Fang et al.[29] and Zhang et al.[33] have proposed an effective and clear structure. They also analyzed the structural relationship between the HMM and LCCRF, which shows the superiority of the latter. The structure developed by Fang[29] is used in this work for the following reasons. The full connection of states and observations is beneficial to mode identification. A new monitoring index can be constructed using state feature function.

The transition feature function is

$$t_k(y_i, y_{i-1}) = \begin{cases} 1, & \text{if } y_{i-1} = i \text{ and } y_i = j \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

The state feature function is

$$s_l(y_i, x) = \begin{bmatrix} s_{l1}(y_i, x_i) \\ s_{l2}(y_i, x_{i-1}) \\ \ldots \\ s_{lg}(y_i, x_{i-g+1}) \end{bmatrix} \tag{22}$$

$$s_{lg}(y_i, x_{i-g+1}) = \begin{cases} 1, & \text{if } y_i = i \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

3.2. Expectation of State Feature Function. From the point of view of physical meaning, state feature function $s_l(\cdot)$ describes the relationship between observations and states. Observations $x = [x_1, x_2, \ldots, x_q]$ are the values of q monitored process variables. States $y = [y_1, y_2, \ldots, y_m]$ are m possible normal operating modes. In the offline modeling phase, normal data are used for training. The weights of state feature function $\mu_l = [\mu_{l1}, \mu_{l2}, \ldots, \mu_{lm}]$ are calculated. They are parameters of the trained LCCRF model. Therefore, the ESFF for one mode y_i can be constructed.

$$\mathrm{ESFF}(y_i) = \sum_{l} \mu_{li}^{\mathrm{T}} s_l(y_i, x)x \tag{24}$$

where μ_{li} is a vector, which contains the corresponding weights for mode y_i. Its computation is quite easy. The new monitoring index called ESFF is given.

$$\mathrm{ESFF} = (\mathrm{ESFF}(y_i) - \overline{\mathrm{ESFF}(y_i)})^2 \tag{25}$$

In this work, the distribution of ESFF is uncertain, so the control limit is calculated through kernel density estimation. For a fault sample, the contribution of every variable x_j can be calculated for the contribution plot,

$$\mathrm{ESFF}(y_i, x_j) = \sum_{l} \mu_{lij}^{\mathrm{T}} s_l(y_i, x)x_j \tag{26}$$

$$\mathrm{contribution}(x_j) = (\mathrm{ESFF}(y_i, x_j) - \overline{\mathrm{ESFF}(y_i, x_j)})^2 \tag{27}$$

where μ_{lij} is the weight which corresponds to the monitored variable x_j.

3.3. Procedure of Operating Mode Identification. Offline modeling:

(1) Construct a historical dataset. Label each sample with its operating mode (serial number of stable modes and transitions between stable modes).
(2) Normalize dataset using the mean and standard deviation.
(3) Train an LCCRF by a scaled conjugate gradient algorithm. Its structure is defined by eqs 21−23.

Online monitoring:

(1) Normalize online sample utilizing the mean and standard deviation.
(2) Operating mode identification. The mode label of this online sample is obtained by the Viterbi algorithm.

3.4. Procedure of Fault Detection and Fault Variable Location. Offline modeling:

(1) Build a normal historical dataset. Label each normal sample with its known operating mode.

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

(2) Normalize dataset using the mean and standard deviation.
(3) Train an LCCRF by a scaled conjugate gradient algorithm. Its structure is defined by eqs 21−23.
(4) Calculate ESFF of every stable mode separately by eq 25. The control limit for each stable mode is computed by kernel density estimation (KDE).
(5) Compute ESFF of every transition separately via eq 25. The control limit for each transition is calculated via kernel density estimation.

Online monitoring:

(1) Normalize online sample utilizing the mean and standard deviation.
(2) Operating mode identification. Label the online sample by the Viterbi algorithm.
(3) Calculate its ESFF via eq 25. If the label of this online sample is one stable mode, the ESFF should be compared to cognate control limit of stable modes. Otherwise, corresponding transition control limit should be used.
(4) If this online sample is regarded as faulty, contribution of each monitored variable is computed by eqs 26 and 27. The contribution plot points out faulty variables.

Table 1. Numbers and Names of Monitored Variables

number	name	number	name
1	A feed	17	stripper underflow
2	D feed	18	stripper temperature
3	E feed	19	stripper steam flow
4	A and C feed	20	compressor work
5	recycle flow	21	reactor cooling water outlet temperature
6	reactor feed rate	22	separator cooling water temperature
7	reactor pressure	23	D feed flow
8	reactor level	24	E feed flow
9	reactor temperature	25	A feed flow
10	purge rate	26	A and C feed flow
11	product separator temperature	27	purge valve
12	product separator level	28	separator pot flow
13	product separator pressure	29	stripper liquid flow
14	product separator underflow	30	reactor cooling water flow
15	stripper level	31	condenser cooling water flow
16	stripper pressure		

3.5. Methodology. In this section, a simple example is used to explain the steps of the proposed methodology. Consider a basic unit of process industry; there are four monitored variables. They are temperature x_1, pressure x_2, flow x_3, and level x_4. This equipment could run in three modes due to its manufacturing technique. Therefore, three mode labels are $y_i \in [y_1 = 1, y_2 = 2, \text{and } y_3 = 3]$.

The procedure of operating mode identification is as follows. Historical data of four monitored variables are collected to build a training dataset $D = \{x^i, y^i\}_{i=1}^N$. This dataset contains values and mode labels of those variables. After normalization, a LCCRF model is trained by a scaled conjugate gradient algorithm. When an online sample comes, its mode label y_i can be calculated in the trained model by the Viterbi algorithm.

The procedure of fault detection and fault variable location is as follows. Construct a historical dataset with normal data of four monitored variables. The dataset $D = \{x^i, y^i\}_{i=1}^N$ consisted of normal operating data and their mode labels. After normalizing, train an LCCRF model through a scaled conjugate gradient algorithm. Compute *ESFF* of all data with mode y_1 label. The control limit for mode y_1 is calculated by kernel density estimation. The other two modes are the same. There are three control limits for three modes. When a new sample comes, its mode label y_i is obtained by the Viterbi algorithm. Calculate its ESFF and compare it with the corresponding control limit. If the control limit is not exceeded, it is considered as a normal sample. Otherwise, ESFF of every variable x_i is computed to find fault variables.

4. RESULTS AND DISCUSSION

4.1. Tennessee Eastman Process. Revised version of the Tennessee Eastman (TE) process model[38] is used in this work (https://depts.washington.edu/control/LARRY/TE/download.html#Updated%20TE%20Code). The monitored variables contain 22 continuous measurements and nine manipulated variables. Their variable numbers and names are listed in Table 1.

4.1.1. Mode Identification for Stable Modes and Transitions. A total of 3000 samples of mode 3 and mode 1 and the transition between them are produced for training (Figure 1). The change trend of products G and H is shown in Figure 2. The transition starts from the 1001st sample due to the change of several set points. In this experiment, the 1800th sample is considered the end of the transition because all variables are relatively stable after that. The change trend of some typical variables is demonstrated in Figure 3. The number of states is 3. The samples of mode 3 are labeled as 1. Transition samples are marked as 2, 3 is for mode 1. The value of regularization weight σ in eq 10 is selected as 1. Three test datasets are designed for trained LCCRF. Test dataset 1 includes mode 3 and the transition. The transition happens since the 1167th sample. Mode 1 brings out test dataset 2. Test dataset 3 is generated from mode 3.

It can be known from Figures 4−6 that all test samples are labeled correctly. Although in actual production, engineers often judge the beginning and end of transition according to some indicator variables, data-driven approaches have their uses and advantages.

The HMM is used for comparison. Mode identification also relies on the Viterbi algorithm. Training and testing data are the same. The number of hidden states is 3. The Gaussian component of each hidden state is 1. Mode identification results for three test datasets are shown in Figure 7. Although the number of hidden states is set as 3, the trained HMM considers transition and mode 1 as the same class. Three types of labeled data are considered as two types. The transition cannot be successfully identified. Thus, LCCRF has superiority.

4.1.2. Process Monitoring Based on ESFF. First, process disturbance IDV (4) is used. A step change occurs in reactor cooling water inlet temperature. Reactor cooling water flow (monitored variable No. 30) will increase for compensation because of the control loop. As a result, qualities of products G and H are not affected, as well as the other monitored variables. This disturbance happens since 201st sample in the test dataset.

During offline modeling, the number of states is 2 (modes 1 and 3). The samples of mode 1 are labeled as 1. Mode 3 samples are marked as 2. The value of regularization weight σ in eq 10 is

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

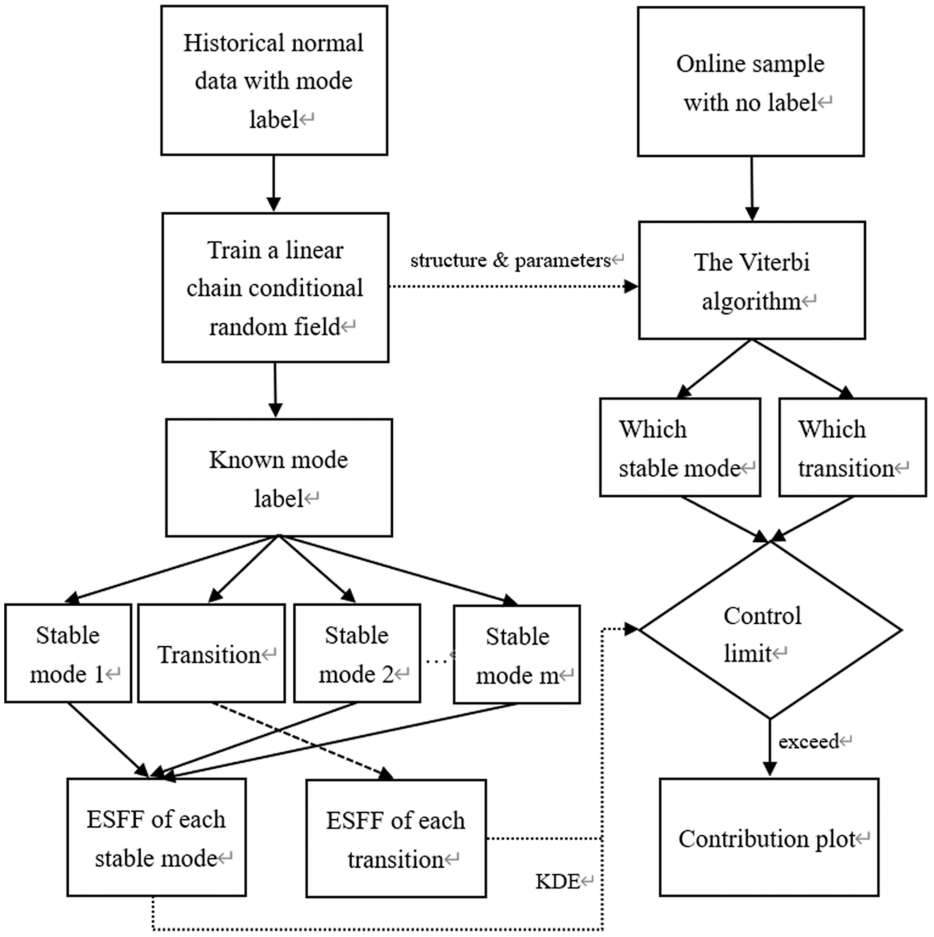


Figure 1. Flowchart of the proposed method.

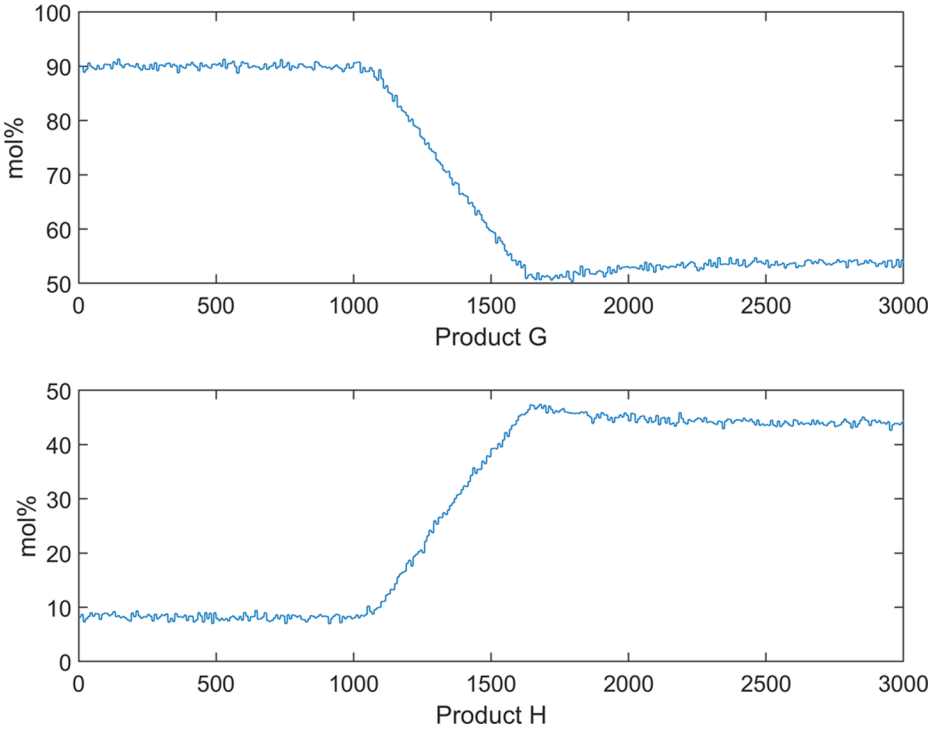


Figure 2. Change trend of products G and H.

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

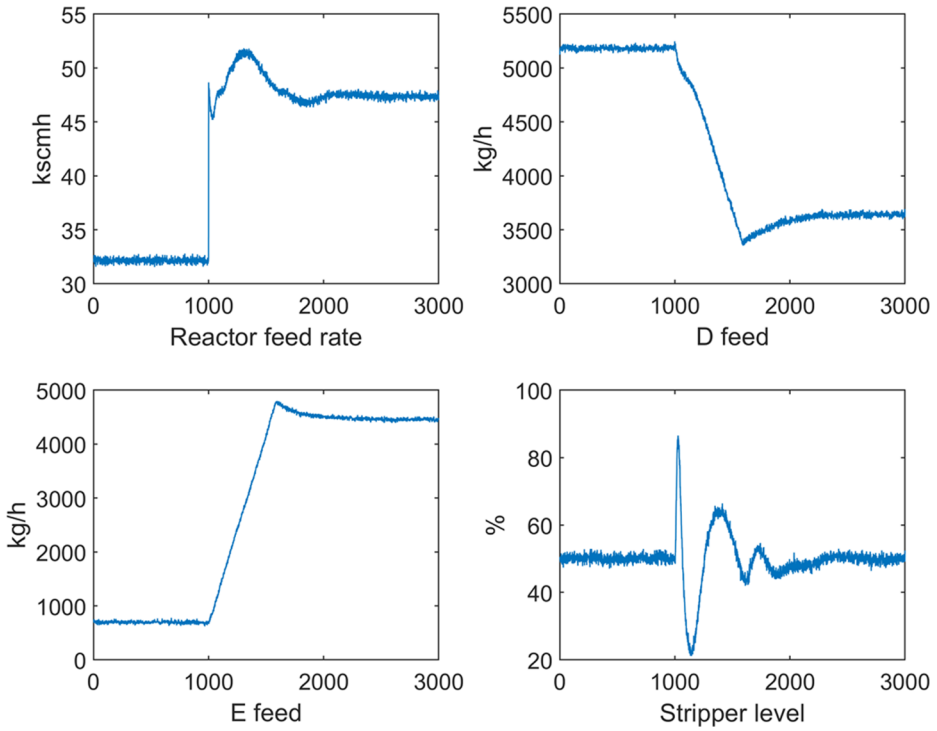


Figure 3. Change trend of some typical variables.

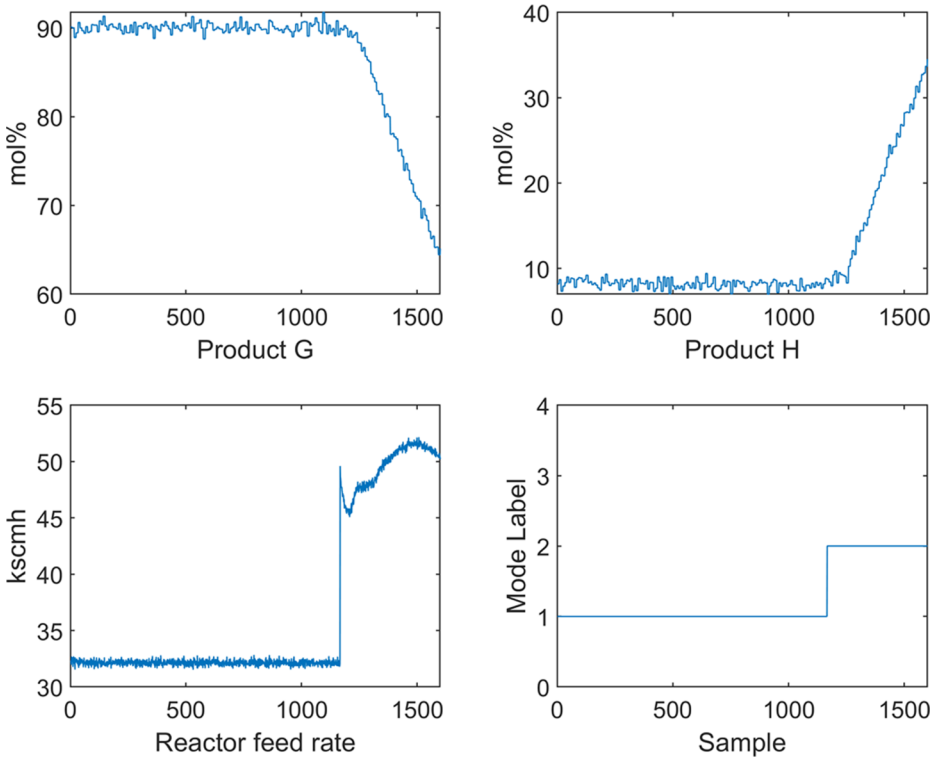


Figure 4. Mode identification result for test dataset 1.

selected as 1. Regularization of weights prevents overfitting by restricting the number of parameters. The value of regularization weight σ determines the severity of penalty, so the ordinary value of this parameter will lead to effective results. The confidence level is 0.99 when calculating both control limits. Figures 8 and 9 describe fault detection results for modes 1 and 3. Figures 10 and 11 present faulty variable location results for modes 1 and 3. These results come from the 600th sample in the test dataset. Figures 12 and 13 provide monitoring results based on the negative log likelihood probability (NLLP) of the HMM for comparison. Training and testing data are the same. The confidence level is also 0.99 when computing the control limit of NLLP.

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

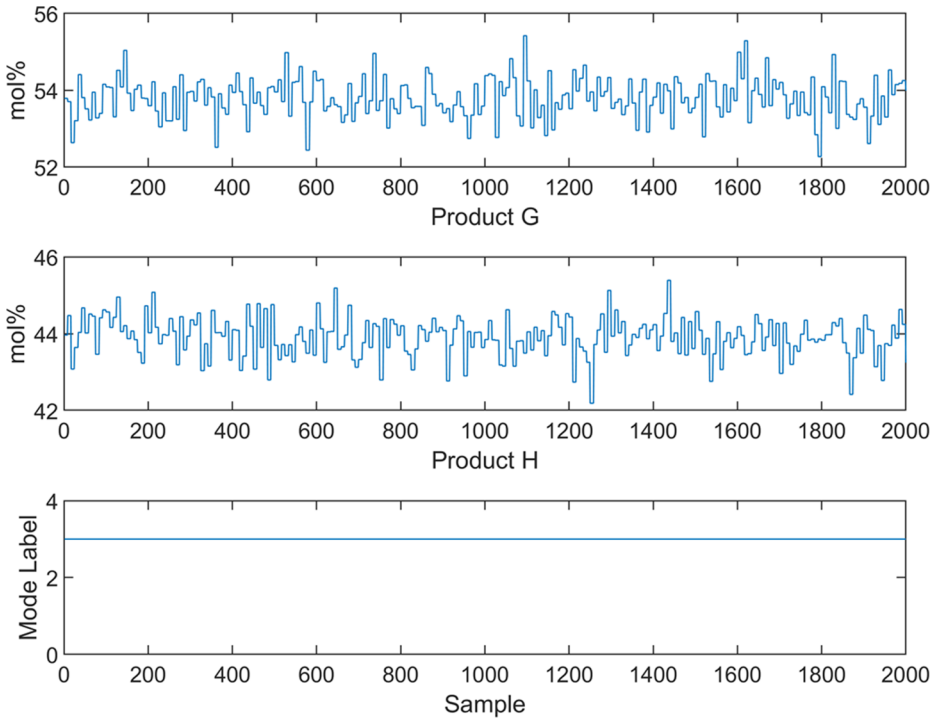


Figure 5. Mode identification result for test dataset 2.

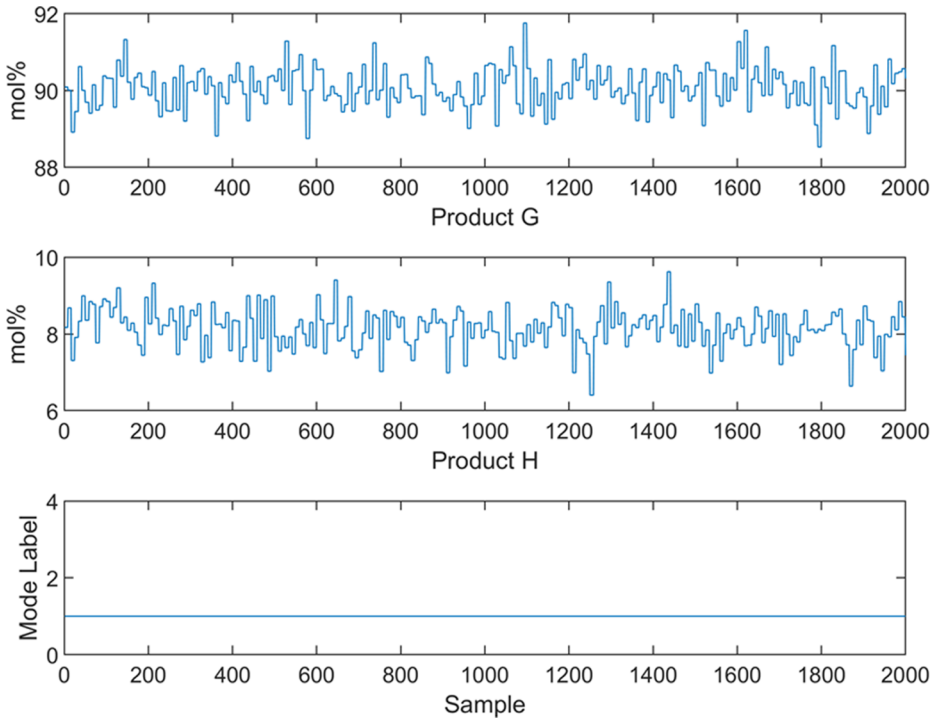


Figure 6. Mode identification result for test dataset 3.

According to those charts, both methods provide satisfactory monitoring performance. All faulty samples are detected. But the negative log likelihood probability lacks the faulty variable location technique. In Figures 10 and 11, the faulty variable is obviously the monitored variable No. 30 due to its significant contribution. This variable is reactor cooling water flow, which is consistent with reality. The effectiveness of the proposed method is verified.

Then, process disturbance IDV (1) is used for modes 1 and 3. This step is about reactants A and C. The change of A/C feed ratio affects several monitored variables at the beginning. Subsequently, the compensatory effect of A feed flow makes qualities and other variables normal. This disturbance also happens since the 201st sample in the test dataset.

During offline modeling, the number of states is 2 (modes 1 and 3). The samples of mode 1 are labeled as 1. Mode 3 samples are marked as 2. The value of regularization weight σ in eq 10 is

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

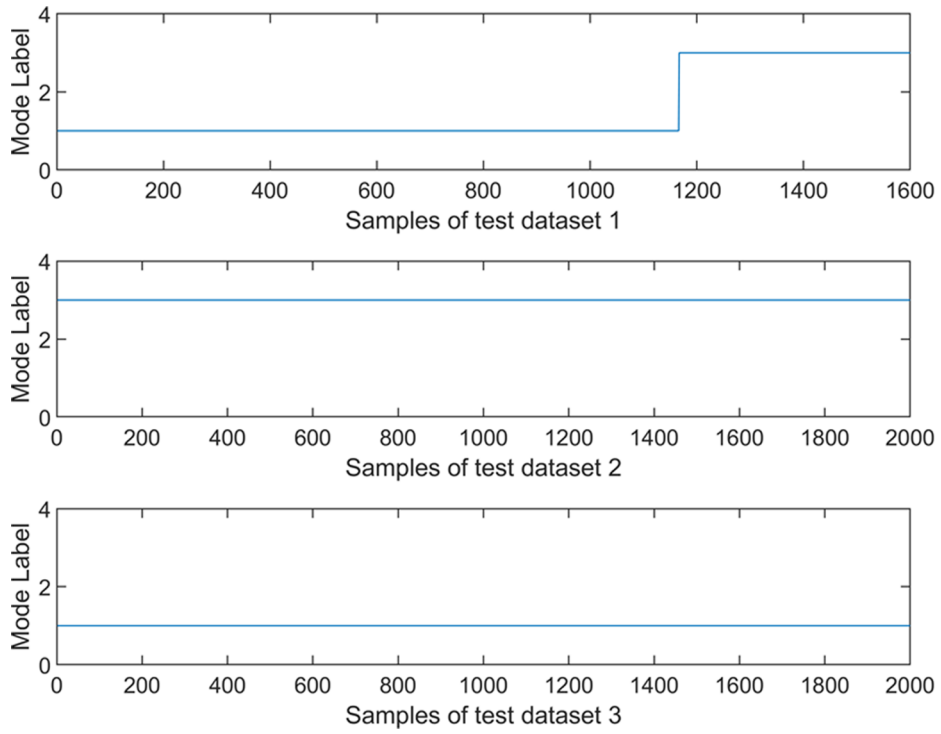


Figure 7. Mode identification result for three test datasets using the hidden Markov model.

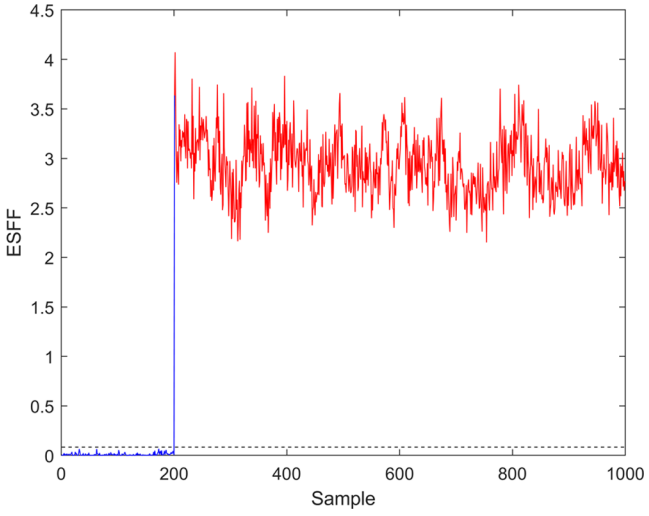


Figure 8. Fault detection result based on ESFF for IDV (4) in mode 1.

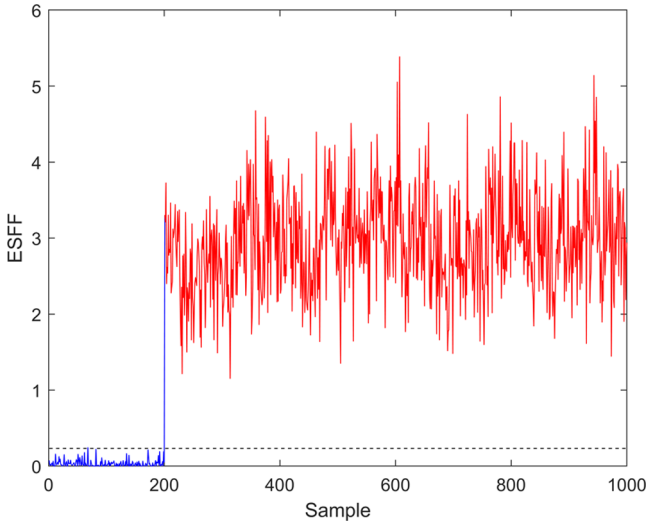


Figure 9. Fault detection result based on ESFF for IDV (4) in mode 3.

selected as 1. The confidence level is 0.99 when calculating both control limits. Detection results of modes 1 and 3 are shown in Figures 14 and 15. Figures 16 and 17 demonstrate the change of faulty variable location with different samples.

Fault detection results are satisfying. At the initial stage of fault, there are many faulty variables. In the end, only two faulty variables remain. They are monitored variable Nos. 1 and 25. Variable No.1 is a continuous process measurement, namely, A feed in stream 1. Variable No. 25 is one manipulated variable called A feed flow in stream 1. This outcome accords with the fact, which testifies the availability of the proposed approach.

4.2. Continuous Stirred Tank Reactor. As a benchmark, the continuous stirred tank reactor (CSTR)[39,40] is used to testify the proposed methodology in this section. This simulation system consisted of feed stream containing reactant stream, product stream, and cooling water stream. Two differential eqs 28 and 29 describe its component material and energy balance. Table 2 lists its primary parameters.

$$\frac{dC_A}{dt} = \frac{q}{V}(C_{Af} - C_A) - k_0 e^{-E/RT} C_A + \nu_1 \quad (28)$$

$$\frac{dT}{dt} = \frac{q}{V}(T_f - T) - \frac{\Delta H}{\rho C_p} k_0 e^{-E/RT} C_A + \frac{UA_H}{\rho V C_p}(T_c - T) + \nu_2 \quad (29)$$

The standard deviations of noises are selected as 0.01 mol/(min L) and 0.02 K/min. Monitored variables are C_A, T, T_c, and q. The standard deviations of those measured noises are 0.001 mol/L, 0.01 K, 0.01 K, and 10 mL/min, respectively. All noises

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

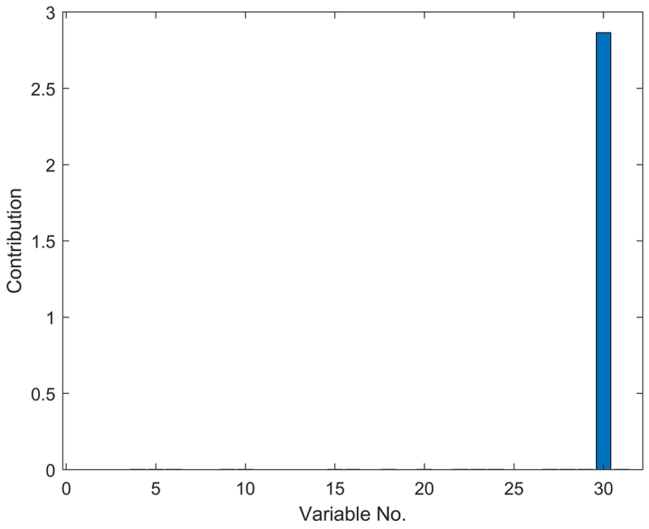


Figure 10. Faulty variable location based on ESFF for IDV (4) in mode 1.

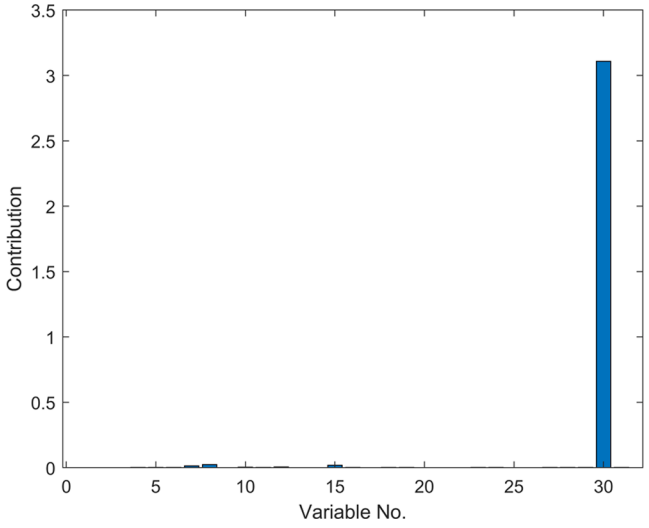


Figure 11. Faulty variable location based on ESFF for IDV (4) in mode 3.

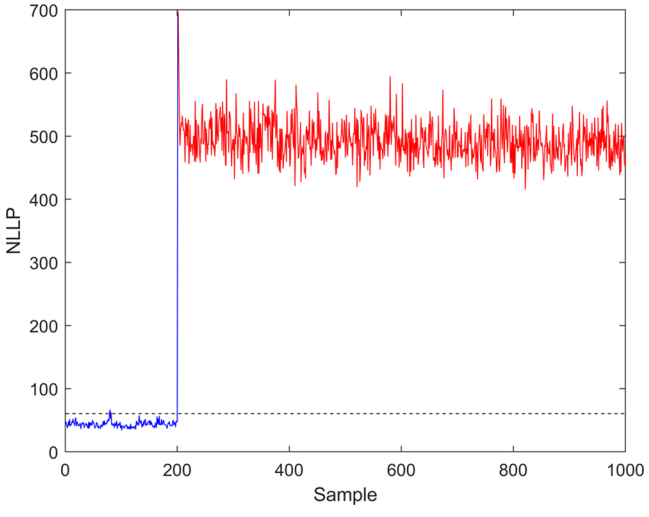


Figure 12. Fault detection result based on NLLP for IDV (4) in mode 1.

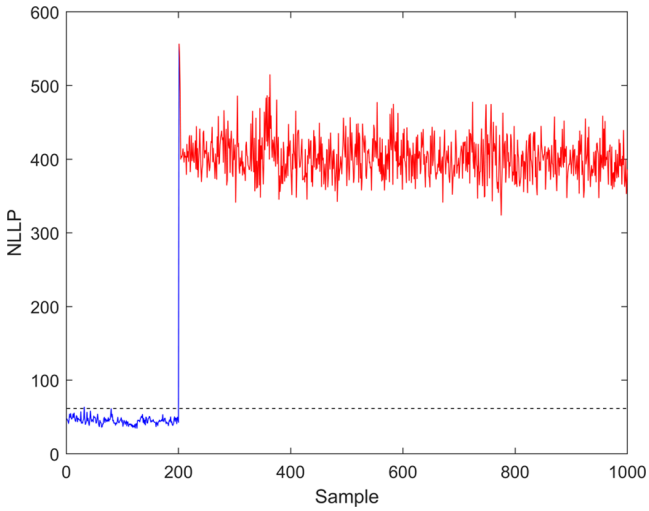


Figure 13. Fault detection result based on NLLP for IDV (4) in mode 3.

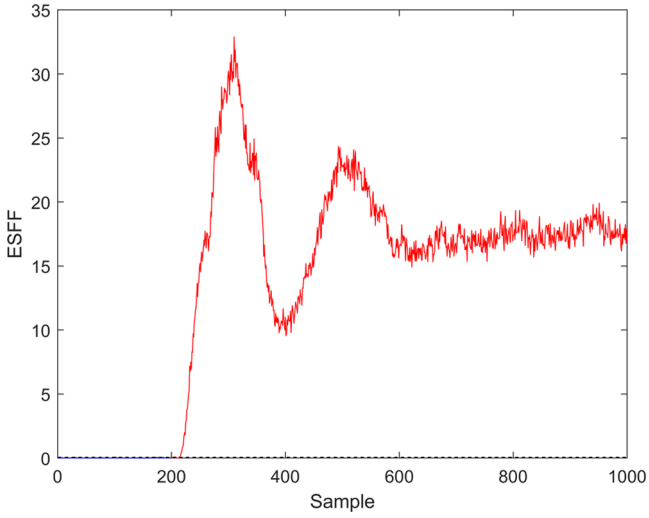


Figure 14. Fault detection result based on ESFF for IDV (1) in mode 1.

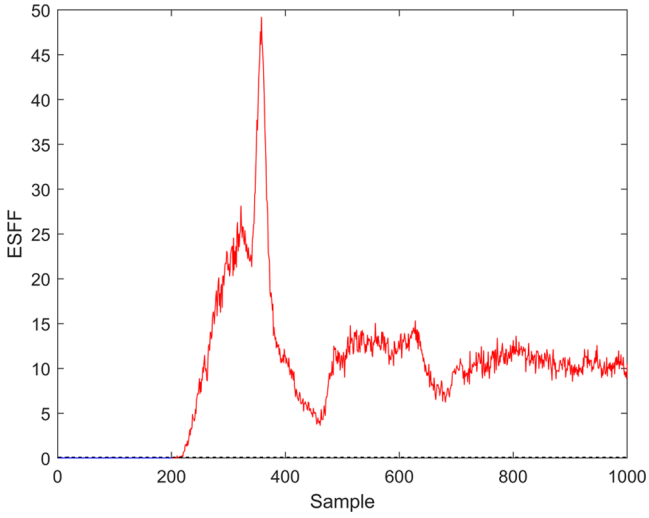


Figure 15. Fault detection result based on ESFF for IDV (1) in mode 3.

are assumed to be independent and Gaussian. The sampling time is 0.5 min.

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

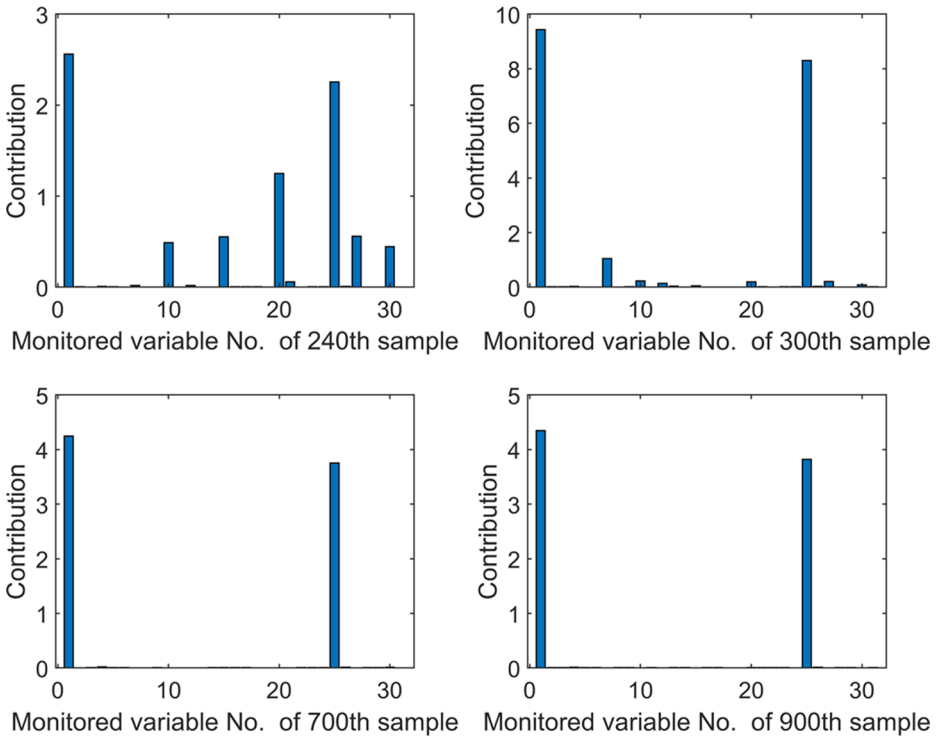


Figure 16. Faulty variable location based on ESFF for IDV (1) in mode 1.

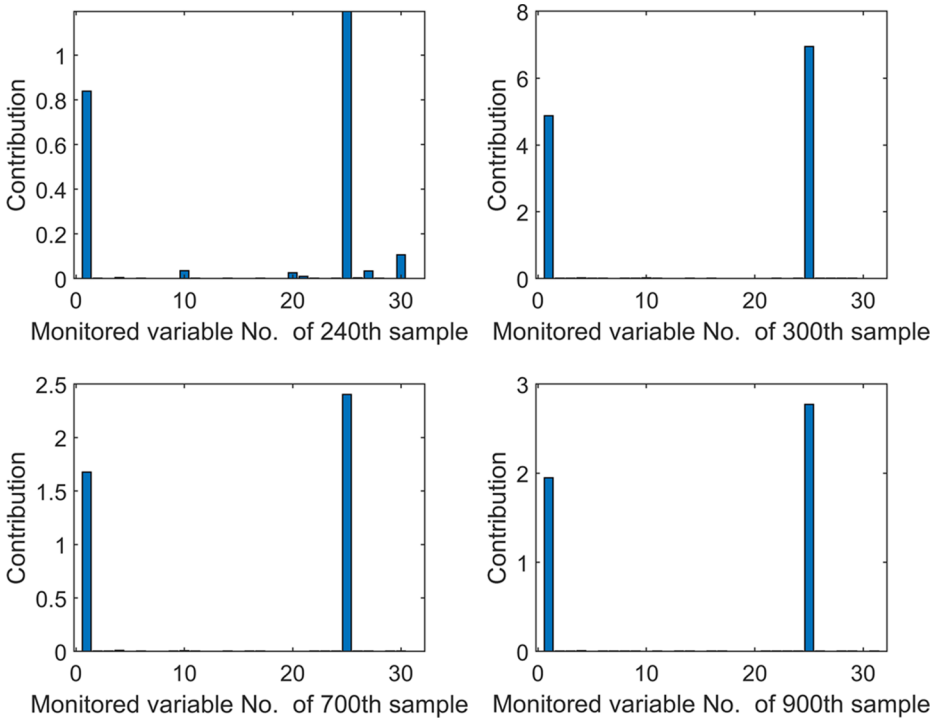


Figure 17. Faulty variable location based on ESFF for IDV (1) in mode 3.

Two operating modes are designed for verification. In mode 1, the value of C_A is 0.816 mol/L. Thousand samples are generated for training. Their labels are marked as 1. C_A equals 1.273 mol/L in mode 2. Thousand samples are also produced for training, and their mode label is 2. For the LCCRF model, the number of states is 2. The value of regularization weight σ in eq 10 is selected as 1. Two test datasets are constructed for the trained model. Test dataset 1 contains 500 samples, which come from mode 1. Similarly, 500 samples are generated from mode 2 to build test dataset 2. It can be known from Figures 18 and 19 that all test samples are labeled correctly.

5. CONCLUSIONS

In this manuscript, linear chain conditional random fields are used for mode identification and process monitoring for multimode processes with transitions. Distinguishing stable

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494

Table 2. Primary Parameters of the CSTR

parameter	description	nominal value
C_A	outlet concentration	0.2 mol/L
C_{Af}	feed concentration	1 mol/L
T	reaction temperature	446 K
T_c	cooling water temperature	419 K
T_f	feed temperature	400 K
q	feed flow rate	100 L/min
ρ	density	1 Kg/L
V	volume	600 L

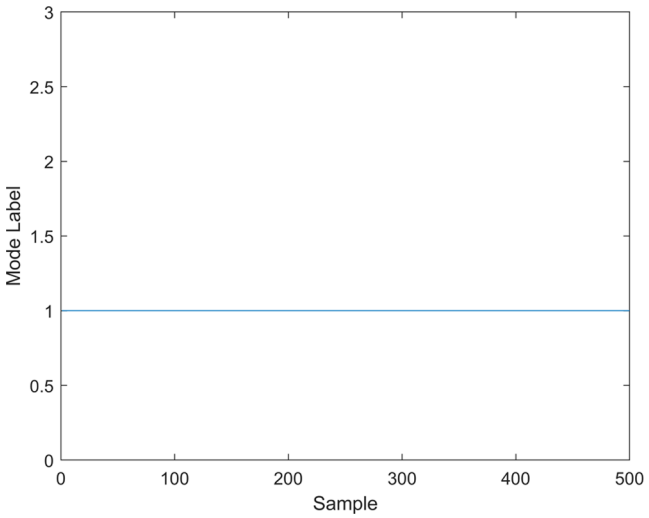


Figure 18. Mode identification result for test dataset 1.

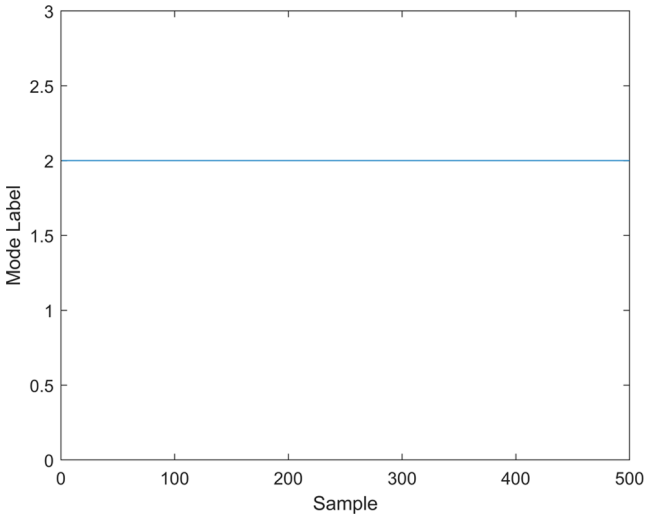


Figure 19. Mode identification result for test dataset 2.

modes from transitions is emphasized. The essence of conditional random field makes it superior to the hidden Markov model. The expectation of state feature function is developed by physical and engineering meaning. Fault variable location based on it is effective. Case studies on the TE process and continuous stirred tank reactor verify the availability of the proposed methodology.

However, there are several problems that need to be further studied. The selection of transition and state feature functions remains an open question. Transition feature functions describe relationship between labels. State feature functions determine connections of observations and labels. How to choose them better according to specific applications is worth exploring. In addition, the structure of a linear chain conditional random field is very important. How to design it can be a future work. Semisupervised conditional random fields can also be a future direction. The training requires correct labels for all samples. It is very useful to have satisfactory performance given only a small amount of labeled data.

AUTHOR INFORMATION

Corresponding Author

Fan Wang − *School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China;* orcid.org/0000-0003-3542-6960; Email: wangfan@tyust.edu.cn

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c04005

Notes

The author declares no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by research start-up funds of Taiyuan University of Science and Technology under Grant 20202035 and 20212051.

REFERENCES

(1) Arunthavanathan, R.; Khan, F.; Ahmed, S.; Imtiaz, S. An analysis of process fault diagnosis methods from safety perspectives. *Comput. Chem. Eng.* **2021**, *145*, No. 107197.

(2) Zadakbar, O.; Imtiaz, S.; Khan, F. Why risk-based multivariate fault detection and diagnosis? *IFAC Proc. Vol.* **2013**, *46*, 672−677.

(3) Yu, H. Y.; Khan, F.; Garaniya, V. Risk-based fault detection using Self-Organizing Map. *Reliab. Eng. Syst. Saf.* **2015**, *139*, 82−96.

(4) Amin, M. T.; Khan, F.; Ahmed, S.; Imtiaz, S. Risk-based fault detection and diagnosis for nonlinear and non-Gaussian process systems using R-vine copula. *Process Saf. Environ. Prot.* **2021**, *150*, 123−136.

(5) Galagedarage Don, M.; Khan, F. Dynamic process fault detection and diagnosis based on a combined approach of hidden Markov and Bayesian network model. *Chem. Eng. Sci.* **2019**, *201*, 82−96.

(6) Amin, M. T.; Imtiaz, S.; Khan, F. Process system fault detection and diagnosis using a hybrid technique. *Chem. Eng. Sci.* **2018**, *189*, 191−211.

(7) Ge, Z. Q.; Song, Z. H.; Ding, S. X.; Huang, B. Data mining and analytics in the process industry: The role of machine learning. *IEEE Access* **2017**, *5*, 20590−20616.

(8) Lee, J. H.; Shin, J.; Realff, M. J. Machine learning: overview of the recent progresses and implications for the process systems engineering field. *Comput. Chem. Eng.* **2018**, *114*, 111−121.

(9) Pilario, K. E.; Shafiee, M.; Cao, Y.; Lao, L. Y.; Yang, S. H. A review of kernel methods for feature extraction in nonlinear process monitoring. *Processes* **2020**, *8*, No. pr8010024.

(10) Wang, F.; Zhang, S.; Yin, Y. X. A new nonlinear process monitoring method based on linear and nonlinear partition. *Ind. Eng. Chem. Res.* **2019**, *58*, 17445−17454.

(11) Zhang, K.; Peng, K. X.; Zhao, S. S.; Wang, F. A novel feature-extraction-based process monitoring method for multimode processes with common features and its applications to a rolling process. *IEEE Trans. Ind. Inf.* **2021**, *17*, 6466−6475.

(12) Qin, S. J.; Chiang, L. H. Advances and opportunities in machine learning for process data analytics. *Comput. Chem. Eng.* **2019**, *126*, 465−473.

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

(14) Ge, Z. Q.; Song, Z. H. Mixture Bayesian regularization method of PPCA for multimode process monitoring. *AIChE J.* **2010**, *56*, 2838−2849.
(15) Li, H.; Wang, H. G.; Fan, W. H. Multimode process fault detection based on local density ratio-weighted support vector data description. *Ind. Eng. Chem. Res.* **2017**, *56*, 2475−2491.
(16) Zhang, S. M.; Zhao, C. H. Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear multimode processes. *Chemom. Intell. Lab. Syst.* **2017**, *168*, 45−61.
(17) Wang, F.; Tan, S.; Yang, Y. W.; Shi, H. B. Hidden Markov model-based fault detection approach for a multimode process. *Ind. Eng. Chem. Res.* **2016**, *55*, 4613−4621.
(18) Wang, F.; Tan, S.; Shi, H. B. Hidden Markov model-based approach for multimode process monitoring. *Chemom. Intell. Lab. Syst.* **2015**, *148*, 51−59.
(19) Amin, M. T.; Khan, F.; Imtiaz, S. Fault detection and pathway analysis using a dynamic Bayesian network. *Chem. Eng. Sci.* **2019**, *195*, 777−790.
(20) Amin, M. T.; Khan, F.; Ahmed, S.; Imtiaz, S. A data-driven Bayesian network learning method for process fault diagnosis. *Process Saf. Environ. Prot.* **2021**, *150*, 110−122.
(21) He, Y. C.; Zhou, L.; Ge, Z. Q.; Song, Z. H. Dynamic mutual information similarity based transient process identification and fault detection. *Can. J. Chem. Eng.* **2018**, *96*, 1541−1558.
(22) Wang, L.; Yang, C. J.; Sun, Y. X. Multi-mode process monitoring approach based on moving window hidden Markov model. *Ind. Eng. Chem. Res.* **2018**, *57*, 292−301.
(23) Wu, D. H.; Zhou, D. H.; Zhang, J. X.; Chen, M. Y. Multimode process monitoring based on fault dependent variable selection and moving window-negative log likelihood probability. *Comput. Chem. Eng.* **2020**, *136*, No. 106787.
(24) Yu, H. Y.; Khan, F.; Garaniya, V. Nonlinear Gaussian Belief Network based fault diagnosis for industrial processes. *J. Process Control* **2015**, *35*, 178−200.
(25) Yu, J. Multiway discrete hidden Markov model-based approach for dynamic batch process monitoring and fault classification. *AIChE J.* **2012**, *58*, 2714−2725.
(26) Sen, D.; Dilshad, R.; Chidambaram, M. Multiway continuous hidden Markov model-based approach for fault detection and diagnosis. *AIChE J.* **2014**, *60*, 2035−2047.
(27) Yu, H. Y. A novel semiparametric hidden Markov model for process failure mode identification. *IEEE Trans. Autom. Sci. Eng.* **2018**, *15*, 506−518.
(28) Sammaknejad, N.; Huang, B.; Xiong, W. L.; Fatehi, A.; Xu, F. W.; Espejo, A. Operating condition diagnosis based on HMM with adaptive transition probabilities in presence of missing observations. *AIChE J.* **2015**, *61*, 477−493.
(29) Fang, M. Q.; Kodamana, H.; Huang, B.; Sammaknejad, N. A novel approach to process operating mode diagnosis using conditional random fields in the presence of missing data. *Comput. Chem. Eng.* **2018**, *111*, 149−163.
(30) Fang, M. Q.; Kodamana, H.; Huang, B. Real-time mode diagnosis for processes with multiple operating conditions using switching conditional random fields. *IEEE Trans. Ind. Electron.* **2020**, *67*, 5060−5070.
(31) Ribeiro, L. C. F.; Afonso, L. C. S.; Colombo, D.; Guilherme, I. R.; Papa, J. P. Evolving neural conditional random fields for drilling report classification. *J. Pet. Sci. Eng.* **2020**, *187*, No. 106846.
(32) Zheng, B. Y.; Gao, X. W.; Pan, R. Sucker rod pump working state diagnosis using motor data and hidden conditional random fields. *IEEE Trans. Ind. Electron.* **2020**, *67*, 7919−7928.
(33) Zhang, H. W.; Shang, J.; Yang, C. J.; Sun, Y. X. Conditional random field for monitoring multimode processes with stochastic perturbations. *J. Franklin Inst.* **2020**, *357*, 8229−8251.
(34) Arunthavanathan, R.; Khan, F.; Ahmed, S.; Imtiaz, S. A deep learning model for process fault prognosis. *Process Saf. Environ. Prot.* **2021**, *154*, 467−479.
(35) Sutton, C.; McCallum, A.An introduction to conditional random fields for relational learning. In *Introduction to Statistical Relational Learning*, MIT Press, 2006; pp 93−128.
(36) Sutton, C.; McCallum, A. An introduction to conditional random fields. *Found. Trends Mach. Learn.* **2012**, *4*, 267−373.
(37) Forney, G. D. The Viterbi algorithm. *Proc. IEEE* **1973**, *61*, 268−278.
(38) Bathelt, A.; Ricker, N. L.; Jelali, M. Revision of the Tennessee Eastman process model. *IFAC-PapersOnLine* **2015**, *48*, 309−314.
(39) Khan, M. A. I.; Imtiaz, S.; Khan, F. Predictive warning system for nonlinear process plants. *J. Process Control* **2021**, *100*, 1−10.
(40) Shang, J.; Chen, M. Y.; Ji, H. Q.; Zhou, D. H. Isolating incipient sensor fault based on recursive transformed component statistical analysis. *J. Process Control* **2018**, *64*, 112−122.

https://doi.org/10.1021/acsomega.2c04005
ACS Omega 2022, 7, 29483−29494