Simultaneous fault diagnosis based on multiple kernel support vector machine in nonlinear dynamic distillation column

Syed Ali Ammar Taqvi1 | Haslinda Zabiri2 | Fahim Uddin1 | Muhammad Naqvi3 | Lemma Dendena Tufa4 | Majida Kazmi5 | Saddaf Rubab6 | Salman Raza Naqvi7 | Abdulhalim Shah Maulud2

1Department of Chemical Engineering, NED University of Engineering & Technology, Karachi, Pakistan
2Chemical Engineering Department, Universiti Teknologi PETRONAS 32610, Seri Iskandar, Malaysia
3Department of Engineering and Chemical Sciences, Karlstad University, Karlstad, Sweden
4School of Chemical and Bioengineering, Addis Ababa Institute of Technology, Addis Ababa, Ethiopia
5Faculty of Electrical and Computer Engineering, NED University of Engineering and Technology, Karachi, Pakistan
6Military College of Signals, National University of Sciences and Technology (NUST), Islamabad, Pakistan
7School of Chemical and Material Engineering, National University of Sciences and Technology (NUST), Islamabad, Pakistan

Abstract
Although numerous works have been done, most of the studies in fault diagnosis are limited to single fault type at a time. Majority of the works reported in the literature do not extend the diagnosis of the root cause of the fault for simultaneous faults specifically in the distillation column. However, an industrial system is susceptible to more than one fault at a time, which may or may not be interrelated. These faults not only reduce the diagnosis performance but also increase the computational complexity of the diagnosis algorithm. In this work, therefore, a multiple kernel support vector machine (MK-SVM) algorithm is proposed to diagnose simultaneous faults in the distillation column. In the developed MK-SVM algorithm, multilabel approach based on various kernel functions has been utilized for the classification of simultaneous faults. Dynamic simulation of a pilot-scale distillation column using Aspen Plus® is used for generating data in normal and faulty operation. Eight different fault types are considered, including valve sticking at reflux and reboiler, tray upsets, loss of feed flow, feed composition, and feed temperature changes. In the classification of simultaneous faults, a combination of two, three, and four faults is introduced for the performance evaluation of the proposed MK-SVM algorithm. The result showed that the proposed MK-SVM has a high fault detection rate (FDR) of 99.51% and a very low misclassification rate (MR) of 0.49%. The MK-SVM–based classification is better with the F1 score of >97% for all combinations of faults. Moreover, it is observed that the proposed MK-SVM shows better fault diagnosis for single, multiple, and simultaneous faults as compared to other established machine-learning algorithms.

KEYWORDS
Aspen Plus dynamics, distillation column, fault diagnosis, machine learning, multiple kernel support vector machines, simultaneous fault classification
1 | INTRODUCTION

In process industries, demands for quality products with a lower rejection rate, minimal downtime, and safer operations are increasing. Human errors cause approximately 70% of industrial accidents. This may affect economics and environmental regulations. Although the advancements in the control system benefited process industries, some disastrous accidents have occurred in history. It has been reported that thousands of people died in the accident of Union Carbide Bhopal, India. Moreover, an estimated loss of 400 million dollars was faced by Mina Al Ahmadi Refinery, Kuwait, in June 2000. It is also reported in the industrial statistics that the minor incidents occur daily, which not only result in injuries but also cost approximately billion dollars per year due to unsafe and abnormal event management, especially in the chemical and process industries. This provides a great interest in the development of fault detection (FD) methods. The FD scheme helps to improve effectiveness, safe operations, and continuous production.

Fault detection and diagnosis have been carried out for various industrial processes such as the Tennessee Eastman process (TEP), reactor system, distillation column, bearing faults, and industrial gas turbine. Distillation is probably the most studied unit since it is energy-intensive and holds a substantial part of the refinery and chemical industry. Any deviation in the column can affect the overall performance of the distillation column.

Efficient monitoring of the distillation column is a demanding task, which depends on various constraints, such as nonlinearity, disturbances, nonstationary behavior, and multivariable coupling. Multivariate statistical process monitoring (MSPM) methods for fault detection and diagnosis in distillation column assume that the system is linear, time-invariant. Due to continuous changes in operating point, and the nonlinearity of the process, the performance of MSPM methods reduced, resulting in a false alarm and missed detection of the faults. The literature presents the significance of the relative efforts to improve the efficiency of monitoring methods for fault detection and diagnosis.

Fault detection and diagnosis can be performed either by first-principle methods, or by data-driven and knowledge-based methods. Due to the nonlinearity and complexity in the modern process industry, there is a demand for data-based methods. Data-driven approaches consider fault detection and diagnosis as classification tasks. This classification can be done by supervised or unsupervised learning methods. These approaches include principal component analysis (PCA), artificial neural network (ANN), Fisher discriminant analysis (FDA), partial least square (PLS), ensemble learning method, Bayesian networks, and support vector machine (SVM). Among all, SVM has been found to be an innovative machine-learning tool and recently used in fault classification of single faults due to its generalization ability.

SVM is based on statistical theory and significantly used as a supervised learning method. SVM has gained much attention due to its generalization ability and becomes a familiar machine-learning and classification tool as compared to the neural network. SVM has been used in various applications including pattern recognition, face detection, text detection, prediction, and classification. In 2004, a fault diagnosis technique using the Fisher discriminant analysis (FDA) and comparison with SVM were done by Chaing. The overall misclassification rate using SVM is found to be three times lower than that using FDA. In this study, only a single fault is considered for fault classification. In the study of Kulkarni et al., SVM is developed to improve the classification performance of the TEP process for binary and multiple faults (three faults). The results have shown substantial improvements in terms of correctly classifying the faults as compared to Chaing. Another technique for single fault classification using fuzzy support vector machine (FSVM) was developed by Yong et al. to diagnose single faults in the TEP process.

For dealing with simultaneous faults in the distillation column, limited analysis can be found in the literature. A multilabel method based on SVM is developed in 2007. Single and simultaneous faults generated from TEP process are tested and diagnosed by the multilabel SVM (ML-SVM) system. The multiclass SVM has also been proposed by Pooyan et al. in dew point process. Nevertheless, only two fault types related to flow and temperature (high/low inlet flow, high/low temperature) have been considered in this study.

Furthermore, fault classification can be considered as monolabel or multilabel approach when one fault may happen at a time or when more than one fault may happen at a time. In our previous study, a fault classification method based on monolabel approach has been proposed to determine the various types of fault in a distillation column.

However, the diagnosis of multiple and simultaneous faults is becoming an important concern in an industry as an industrial system is susceptible to more than one fault at a time, which may or may not be interrelated. Particularly, the interrelated faults may be avoided by timely diagnosis followed by corrective actions. In real processes, there are possible chances of combinations of faults. Hence, there is a need to develop more robust and reliable technique, which can be used to diagnose multiple simultaneous faults in nonlinear systems. However, the current literature does not present any viable solution for simultaneous fault diagnosis. In addition, most studies on fault diagnosis lack comparison with other machine-learning algorithms.

Hence, in this paper, SVM, which has been used so far only for single fault classification in literature using one class technique (binary classification problem), is investigated.
and extended to account for multiple and simultaneous fault diagnosis algorithm through different kernel functions. Hence, the main contributions of this study are as follows:

1. An extended version of MK-SVM for multilabel approach based on various kernel functions has been developed for the classification of two, three, and four simultaneous fault combinations in the nonlinear dynamic distillation column.

2. A rich training dataset of 19,000 observations, which includes 19 single, multiple and simultaneous fault scenarios, is utilized for the training of SVM classifier with a different type of kernels.

3. A full comparison analysis has been performed with other well-known classification algorithms such as ANN and K-nearest neighbors (KNNs).

2 | SIMULTANEOUS FAULT DIAGNOSIS ALGORITHM

2.1 | Linear classification

Using linear separating hyperplane, the classification function is used to separate the training data in two different classes. Figure 1 illustrates the data of two different classes in which one represents a positive class, while other represents a negative class.

The maximum margin is defined as the distance between two imaginary lines. The points that lie on the imaginary lines are known as support vectors. The training data are presented as:

\[(u_1, y_1), (u_2, y_2), \ldots, (u_n, y_n), \quad u \in \mathbb{R}^n, \quad y \in \{-1, +1\}\]

This can be divided by a hyperplane equation as:

\[(a \cdot u) + b = 0\]  \hspace{1cm} (2)

In case of maximum distance between different classes of samples, the hyperplane is known to be optimal hyperplane. The separating hyperplane can be presented in the following form:

\[(a \cdot u) + b = \geq 1, \text{ if } y_i = 1\]  \hspace{1cm} (3)

\[(a \cdot u) + b = \leq 1, \text{ if } y_i = -1\]  \hspace{1cm} (4)

Or equivalently

\[y_i \cdot [(a \cdot u_i) + b] \geq 1; \quad i = 1, 2, \ldots, n\]  \hspace{1cm} (5)

The problem of searching for the optimal hyperplane could be used to obtain the minimum of the function subjected to the following constraint in Equation (6).

\[\delta(a) = \frac{1}{2}||a||^2 = \frac{1}{2}(a \cdot a)\]  \hspace{1cm} (6)

In this way, the optimal hyperplane can be found by solving a simple quadratic programming problem. Lagrange function is constructed in Equation (7).

\[\theta(a, b, \gamma) = \frac{1}{2}a^T \cdot a - \sum_{i=1}^{m} \gamma_i [y_i(a^T \cdot u_i - b) - 1]\]  \hspace{1cm} (7)

\(\gamma_i\) is Lagrange multiplier, which should satisfy the constraint \(\gamma_i \geq 0; \quad i = 1, \ldots, n\). When the condition (6) reaches its extremum, the corresponding points should satisfy Equations (8) and (9).

\[\frac{\partial \theta}{\partial b} = 0 \Rightarrow \sum_{i=1}^{m} \gamma_i y_i = 0\]  \hspace{1cm} (8)

\[\frac{\partial \theta}{\partial u} = 0 \Rightarrow \sum_{i=1}^{m} \gamma_i y_i u_i = a\]  \hspace{1cm} (9)

Equations (8) and (9) are substituted into Lagrange function, and \(a\) and \(b\) are eliminated.

\[a(\gamma) = \sum_{i=1}^{n} \gamma_i - \frac{1}{2} \sum_{i,j}^{n} \gamma_i \gamma_j y_i y_j (u_i, u_j)\]  \hspace{1cm} (10)

**FIGURE 1** (A) Support vector machine hyperplane and (B) support vectors and margin
The above Equation (10) must be under the constraint \( \gamma_i \geq 0, i = 1, ..., n \) and \( \sum_{i=1}^{n} \gamma_i = 0 \).

Those \( u_i \)s with \( \gamma_i > 0 \) are termed support vectors (SVs). The label of a testing data \( u_i \) can then be obtained by

\[
\varphi(u) = \text{sgn} \left( u_i \mathbf{U}_{\text{SVM}} \in \sum_{i} \gamma_i (u \cdot u_i) + b \right)
\]

where

\[
\alpha = x_i \mathbf{U}_{\text{SVM}} \in \sum_{i} \gamma_i u_i
\]

### 2.2 Nonlinear classification

For nonlinear data, the training samples are not linearly separable. Cortes introduced non-negative variables \( \pi_i \) and penalty function \( F(x) = \sum \pi_i \) in order to promote the optimal hyperplane to the general situation. A slack variable is introduced to condition (5)

\[
\gamma_i (a^T \cdot \delta(u_i) + b) \geq 1 - \pi_i
\]

The general classification hyperplane is a minimum value of Equation (13) under the constrain (12).

\[
\min \frac{1}{2}||u||^2 + Z \sum_{i=1}^{n} \pi_i
\]

The formula used in the nonlinear SVM is the same as the one used in the linear SVM, but the constraint \( \gamma_i \) is different, which should satisfy the constrain \( 0 \leq \gamma_i \leq Z \). Therefore, using an appropriate inner product function \( K(u_i, u_j) \) in the optimal classification plane can achieve linear classification after a nonlinear transformation, and the computational complexity is not increased. The objective function is represented below.

\[
A(y) = \sum_{i=1}^{n} \gamma_i - \frac{1}{2} \sum_{i,j=1}^{n} \gamma_i \gamma_j y_i y_j K(u_i, u_j)
\]

The kernel is used in input space because it can map the input samples into feature space. There are various types of kernels, such as linear, polynomial, radial basis, and sigmoid.

### 2.3 SVM model development

SVM is widely used for data classification. The training dataset comprised labeled class and observed variables as labeled class and attribute, respectively, to build a fault diagnosis model. The fault diagnosis methodology using SVM is shown in Figure 2. The trained SVM determines the optimal support hyperplane (OSH) using the training data. The diagnostic system is ready to diagnose the new classes determining OSH.\(^{33}\)

The normal and faulty datasets that have been generated using Aspen Plus dynamic simulation for the fault detection are utilized for the training of the classifier. The dataset is labeled as predictors (ie, the selected observed and manipulated variables) and classes (ie, normal, fault type 1, and fault type 2). The training of the classifier has been done using k-fold cross-validation.

In MATLAB, the function “fitcsvm” trains and cross-validates a support vector machine (SVM) model for one or two classes on a low-dimensional or moderate-dimensional predictor dataset. The function “fitcsvm” supports mapping of the predictor data using kernel functions. The k-fold cross-validation is initially defined before training the model. There are three steps for k-fold cross-validation: (1) randomly partition the data into k sets, (2) for each set, reserve the set as validation data, and train the model using the other k-1 sets; and (3) store the k compact, trained models in the cells of a k-by-1 cell vector in the trained property of the cross-validated model. The values of the kernel parameters are not known beforehand; hence, auto mode is selected in the classifier. The kernel decides the feature space to classify the training data. The kernel function is said to be good if it is able to provide higher bound on the statistical classification dimension. However, it is difficult to accomplish the evaluation of the hypersphere enclosing all data in the nonlinear feature space.

Thus, when it is difficult to separate the nonlinear data, SVM equipped with kernel function is used to project the nonlinear data into higher dimensional space. In this thesis, all four standard SVM kernel functions (ie, linear, quadratic, cubic and Gaussian) are evaluated and the one with the highest training accuracy is selected. The performance of SVM is significantly affected by the kernel scale and penalty parameters. Since there is no defined rule for the kernel parameters for any specific problem, in this study the kernel parameters are selected in the classifier. The kernel decides the feature space to classify the training data. The kernel function is selected in the classifier. The kernel decides the feature space to classify the training data. The kernel function is said to be good if it is able to provide higher bound on the statistical classification dimension. However, it is difficult to accomplish the evaluation of the hypersphere enclosing all data in the nonlinear feature space.

Thus, when it is difficult to separate the nonlinear data, SVM equipped with kernel function is used to project the nonlinear data into higher dimensional space. In this thesis, all four standard SVM kernel functions (ie, linear, quadratic, cubic and Gaussian) are evaluated and the one with the highest training accuracy is selected. The performance of SVM is significantly affected by the kernel scale and penalty parameters. Since there is no defined rule for the kernel parameters for any specific problem, in this study the auto kernel scale has been selected to achieve maximum training accuracy of the classifier. The classifiers are then trained using each of the standard linear, quadratic, cubic, and Gaussian SVM kernels for the comparison of training accuracy.

The classifiers are selected based on the training accuracy (\( A_T \)) with the highest training accuracy (\( A_T \)) being selected and utilized for fault diagnosis of various faults using the test dataset. After a successful diagnosis, the performance evaluation of classifier is done using various performance indices. Table 1 outlines the overall procedure adopted in this study for the development of the proposed fault diagnosis algorithm using SVM.

### 2.4 Performance evaluation

The statistical measures of the classification performance can be obtained by using the several concepts such as true
positive (TP), true negative (TN), false negative (FN), and false positive (FP). The classification performance of the proposed fault diagnosis has been done using various performance indices such as accuracy, recall, precision, and F1 score. These terms are further utilized for the calculation of different performance methods. These terminologies are defined below:

**Accuracy**—It is a ratio of correctly predicted observations to the total observations.

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (16)
\]

**Recall**—the conditional probability of fault diagnosis for predicting fault conditioned to the sample.

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (17)
\]

**Precision (Pre)**—conditioned probability of happening fault.

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (18)
\]

**F1 Score**—F1 performance index is the measure for evaluating the general performance of the fault diagnosis.
system. It is defined as the weighted measure of recall and the precision.

\[ F_1 = \frac{2 \times \text{Prec} \times \text{Rec}}{\text{Prec} + \text{Rec}} \]  

(19)

### TABLE 1  Proposed fault diagnosis algorithm using MK-SVM

| Step | Description |
|------|-------------|
| 1    | Generate normal and fault dataset using Aspen Plus dynamic simulation. |
| 2    | Prepare the dataset based on predictors (Table 4) and classes in which “X” is the matrix of predictor data (where each row is one observation, and each column is one predictor) and “Y” is array of class labels (with each row corresponding to the value of the corresponding row in X). Label the dataset as predictors “X” (observed and manipulated variables) and classes “Y” (ie, normal, fault type 1, and fault type 2). |
| 3    | Normalize the dataset in the range of [0 1]. |
| 4    | Define the cross-validation method; in the case of k-fold cross-validation, the following will be done: |
|      | • Partition the data into k disjoint sets or folds |
|      | For each fold: |
|      | – Train a model using the out-of-fold observations |
|      | – Assess model performance using in-fold data |
|      | • Calculate the average test error overall folds |
| 5    | Observe the scatter plot in order to include or exclude variables |
| 6    | Select kernel function |
| 7    | Set kernel scale to auto to use a heuristic procedure to select the scale value using subsampling. Select manual kernel scale if desired. |
| 8    | Select classification method if the data have 3 or more classes. Train binary SVMs to distinguish the pair of classes (one-vs-one) or each class from all others (one-vs-all). |
| 9    | Train the classifier. |
| 10   | Test the classifier. |
| 11   | Evaluate the performance using performance indices using Equations (16)–(19). |

### 3 | CASE STUDY

#### 3.1 | Dynamic pilot-scale distillation column using Aspen Plus®

The Aspen Plus® dynamic simulation has been used for the generation of normal and faulty datasets of pilot-scale distillation column. Steps 1 and 2 are defined as the development of steady-state and dynamic simulation. In this study, ethanol-water mixture with 25%–75% composition in feed is used in the distillation column as the case study for the generation of normal and faulty datasets. The RADFRAC model is used comprising of a column, condenser, and reboiler. The overall column height is 5.5 m with 0.15 m internal diameter and 0.35 m tray spacing. The top composition of the distillation column contains 85% ethanol. The condenser pressure is maintained at 1.05 bar, which is then increased to 1.24 bar due to the addition of inlet/outlet valves, and the pressure drops across each tray in the dynamic simulation of the distillation column. After running steady-state simulation in Aspen Plus® without errors and warnings, the simulation is then exported to Aspen Plus® dynamics for further control studies and data generation. Feed flow, top and bottom compositions, pressure, and level controller are installed in this distillation column. The Aspen Plus dynamic simulation is presented in Figure 3. The detailed description and development of Aspen Plus® dynamic simulation can be found in our previous publications.45-48

The Aspen RadFrac model can be used in both the steady-state and dynamic simulations. In dynamic mode, RadFrac models the pressure drop across each stage due to liquid and vapor flow resistances. Stage hydraulics can also be modeled by using RadFrac model.
The equipment sizing for the distillation column is presented in Table 2.

The steady-state simulation results have been validated with the published steady-state operating conditions of the distillation column in which the top and bottom temperatures are 81.25°C and 104.0°C, respectively, and the column pressure is 1.05 bar (abs). It should be noted that for the development of the dynamic simulation, the column pressure has to be slightly increased to 1.24 bar for convergence due to the additional inlet/outlet valves and the associated pressure drop across each tray.

In addition, the single-step input does not cover the nonlinear dynamics completely. Therefore, pseudo-random binary sequences (PRBS) have been utilized in this study to excite the plant in normal conditions. The advantages of the PRBS input include ease of implementation and an autocorrelation function similar to white noise. The data generation for normal (no-fault) conditions in process plant can be done using PRBS signals. Zero mean normal distributed noise is added to all measured variables to simulate the sensor noise as mentioned in Table 3. The input variables are reflux, reboiler duty, and condenser duty, whereas the output variables are top and bottom compositions and column pressure. The input and output data are normalized in the range of [0 1]. The generated data from distillation column are divided in 60–20–20 (%) ratio for training, validation, and testing of the network. Furthermore, the data are generated under normal operations neither start-up nor shut down mode.

3.2 | Types of faults

The Aspen Plus dynamic simulation of the distillation column presented in Section 3.1 is used to generate normal and faulty data. The normal operation (no-fault) of the plant has been simulated for 20 h. In this study, eight types of faults are introduced in the developed pilot-scale distillation column dynamic simulation. All the faults considered are introduced at 5 h of normal operation in the distillation column. The detailed descriptions of each fault are as follows:

3.2.1 | Reflux flow valve stiction [fault F1]

Fault (F1) represents the reflux flow control valve suffering from stiction nonlinearity. The reflux flow valve stiction data are generated using the stiction characteristics as described in Ref. [52,53]. The stiction signal is adapted to generate the stiction signal for reflux valve stiction. The magnitude of oscillation was kept at ±35%. The nonlinearity is then introduced in the distillation column to observe the change conditions.

3.2.2 | Reboiler steam valve stiction [fault F2]

Fault (F2) represents the reboiler steam control valve suffering from stiction nonlinearity. The nonlinearity is introduced similar to reflux valve stiction. It is expected that the changes in the reboiler flow may affect the bottom compositions.

3.2.3 | Top tray upset (efficiency reduction) [fault F3]

Fault (F3) represents the top (2nd) tray upset in the distillation column to observe the column behavior with lower tray efficiency. It has been assumed that there is a tray malfunction at the top part of the distillation column. This fault has been simulated by reducing the Murphree efficiency of the second stage to 1%. A significant decrease in the separation efficiency can be due to weeping in the column and will lead to the reduction of the liquid level on the respective tray. In the literature, trays have been bypassed to simulate tray malfunctions, which is equivalent to reducing the tray efficiency to 0%. This paper can be a good example related to tray fault since it reports an FDD study on the pilot-scale distillation column used in this work.

3.2.4 | Bottom tray upset (efficiency reduction) [fault F4]

Fault (F4) represents the bottom tray upset in the distillation column to observe the column behavior with lower tray efficiency. It has been assumed that there is a tray malfunction at the top part of the distillation column. This fault has been simulated by reducing the Murphree efficiency of the second stage to 1%. A significant decrease in the separation efficiency can be due to weeping in the column and will lead to the reduction of the liquid level on the respective tray. In the literature, trays have been bypassed to simulate tray malfunctions, which is equivalent to reducing the tray efficiency to 0%. This paper can be a good example related to tray fault since it reports an FDD study on the pilot-scale distillation column used in this work.
TAQVI ET AL.

3.2.5 Low efficiency in refining section [fault F₅]

Fault (F₅) represents the refining section tray upset in a distillation column to witness the behavior of the column with reduced efficiency. The refining section comprises all trays above the feed tray. Therefore, low tray efficiencies are introduced (1%) in all the trays in the refining section.

3.2.6 Change in feed composition [fault F₆]

A fault related to random variation in feed composition is introduced in the distillation column as fault F₆. A random variation in feed composition as a fault affects the top and bottom composition of the column. A significant impact can also be seen on the reflux flow rate and reboiler duty due to this fault type. It is to note that the steady-state feed composition of ethanol is 25 mole%. For this fault, a random variation in ethanol composition has been introduced in the range of ±16%, that is, 21–29 mol%.

3.2.7 Change in feed temperature [fault F₇]

The change in feed temperature is introduced as a fault (F₇) in distillation column. To maintain the overall heat balance in the column, the feed temperature must be maintained. In this case, a negative step change is introduced in the feed temperature to observe the changed behavior. The steady-state temperature of the column in normal operation is 68°C. However, feed temperature has been reduced to 30°C due to fault. The sudden change in feed temperature has a significant effect on various variables of distillation column such as top and bottom compositions, distillate, and bottom flow rates.

3.2.8 Loss of feed flow [fault F₈]

A fault (F₈) related to feed flow rate has been simulated in the distillation column to observe the column behavior. The feed flow valve has a 50% of valve opening at normal operations. For this fault, an 84% loss (ie, 8% valve opening) in the feed flow rate is simulated. It is assumed that feed lost in the column is due to some malfunction in the feed valve or due to some disruptions from the storage tank. Table 4 shows various faults introduced in the distillation column.

4 RESULTS AND DISCUSSION

In this section, simultaneous faults have been classified using MK-SVM. Combination of two, three, and four simultaneous faults is simulated, and diagnosis results are shown. In the previous section, one fault at a time
Behavior of distillation column at simultaneous faults

This section presents the behavior of the column when faults occur in distillation column simultaneously. Faults F₁ and F₂ are associated with reflux valve and reboiler flow valve malfunctions due to stiction, respectively.

Figure 4 shows the response of the distillation column when F₁ and F₂ (ID1) occur simultaneously. It can be observed that these faults have disturbed the overall behavior of the column after the fault occurs at 5 h. Continuous fluctuation in top and bottom compositions of the column can be seen due to these simultaneous. Both the simultaneous faults have significantly affected the column behavior as clearly exhibited in the process variable responses in Figure 4.

ID2 indicates the simultaneous faults in reflux flow valve stiction (F₁) and top tray (2nd stage) efficiency (F₃), respectively. When these two faults occur simultaneously at time $t = 5.0$ h, most of the observed variables have deviated from the steady-state condition, which depicts the abnormal behavior in the distillation column. It can be seen in Figure 5 that as the anomaly occurs in reflux flow rate and the top stage, a sudden change in the top composition of the column can be observed. As the column is operated in closed-loop operations, the top and bottom compositions are maintained by the reflux flow and reboiler duty. Since there are two simultaneous faults, it is very challenging for the controllers to perform control action. Therefore, it takes around approximately 6 h to retain the top level of the reflux drum back to its original steady state. The top composition started increasing to achieve the desired set-point only after the increase in the top level is observed.

Liquid level, vapor velocity has been significantly affected by the reduction in tray efficiency since these variables tend to maintain the column pressure. The sudden change in differential pressure and weeping in the column may affect the tray efficiency. After fault F₃ occurs, it affects most of the variable. For instance, the reflux flow valve has fully opened to maintain the top composition of the column; however, due to a low level in the reflux drum, it has then decreased. It is observed that the low level in the reflux drum caused the distillate flow rate to decrease to zero flow in the column till $t = 13.0$ h.

The reboiler duty has also decreased due to a direct link with reflux. Reboiler tries to heat the system while reflux tries to cool it. When there is low reflux (ie, low cooling), the heating requirement also goes lower. Similarly, when the reboiler gets limiting heat (ie, when do not have enough steam to heat), it manages the situation by reducing reflux (slight compromise on quality). It has decreased

| Case ID | Description | Type          | Magnitude     |
|---------|-------------|---------------|---------------|
| F₁      | Reflux flow valve stiction | Sticking | ±35% oscillation |
| F₂      | Reboiler steam valve stiction | Sticking | ±35% oscillation |
| F₃      | Top tray upset       | Step         | −99%          |
| F₄      | Bottom tray upset    | Step         | −99%          |
| F₅      | Refining section tray upset | Step | −99%          |
| F₆      | Feed composition     | Random variation | ±16%      |
| F₇      | Feed temperature     | Step         | −45%          |
| F₈      | Loss of feed flow    | Step         | −84%          |

is introduced in the distillation column from F₁ to F₈. Table 5 presents the list of simultaneous faults introduced in the distillation column.

### Table 4 Simulated faults in the distillation column

| No. | Fault ID | Simultaneous fault ID description |
|-----|----------|----------------------------------|
| ID1 | F₁, F₂  | 2 simultaneous faults in reflux flow and reboiler steam flow valves |
| ID2 | F₁, F₃  | 2 simultaneous faults in reflux flow valve and top tray upset |
| ID3 | F₁, F₄  | 2 simultaneous faults in reflux flow and feed temperature |
| ID4 | F₁, F₅  | 2 simultaneous faults in top tray and feed tray |
| ID5 | F₁, F₆  | 2 simultaneous faults in top tray and feed flow |
| ID6 | F₁, F₂, F₃ | 3 simultaneous faults in reflux flow, reboiler, and top tray |
| ID7 | F₁, F₂, F₄ | 3 simultaneous faults in reflux flow, reboiler, and feed composition |
| ID8 | F₁, F₅, F₆ | 3 simultaneous faults in top tray, feed tray, and feed composition |
| ID9 | F₁, F₇, F₈ | 3 simultaneous faults in feed composition, feed temp, and feed flow |
| ID10| F₁, F₂, F₃, F₄ | 4 simultaneous faults in reflux, reboiler, and top and feed tray |
only due to decrease in reflux. Reboiler does two things: (a) heat the feed to column temp and (b) reheat the reflux to column temp.

In ID3, the behavior of the column is observed by introducing the fault in reflux valve stiction ($F_1$) and the decrease in the feed temperature ($F_7$) simultaneously. As
discussed earlier, the valve stiction in the reflux flow rate has a significant impact on the top composition, and the feed temperature affects the overall heat balance of the column. Once the temperature has decreased, the reboiler duty has increased to maintain the temperature profile of the column. These simultaneous faults have also affected top and bottom flow rates, and level of the column along with column pressure, as shown in Figure 6.
Figure 7 shows the behavior of the column when the ID4, that is, the efficiency of top and the bottom tray, has been reduced simultaneously (F2 and F4). These two simultaneous faults show significant effects on the overall performance of the distillation column. It is observed that the reflux flow rate and reboiler duty have decreased to zero due to the low level in the reflux drum and bottom level of the column, respectively. These simultaneous faults have also affected the column pressure significantly. Furthermore, it can be seen that column operation remained abnormal during the whole observation time.

In ID5, a reduction in tray efficiency and loss of feed flow rate are simulated simultaneously to observe the column behavior (F2 and F3). A common effect of both faults can be seen in Figure 8. The loss of feed flow reduces the top and bottom levels of the column that results in the decrement of both the distillate and bottom flow rates. The effect of reduced tray efficiency can also be seen in the sharp increase and decrease in the reflux flow rate between 5.0 and 6.0 h. The abnormal behavior in distillation column due to simultaneous faults in F2 and F3 results in significant changes in the column behavior.

In ID6, a combination of three simultaneous faults, that is, reflux flow valve stiction (F1), reboiler valve stiction (F4), and top tray upset (F3), is simulated, and the responses are shown in Figure 9. It can be observed that the tray upset fault has a significant effect as compared to valve stiction faults. It can be observed that after the fault F3 occurs, it affects all the variables in the distillation column. Similar behavior can be seen in the top composition, which has decreased to 62.0 mol% after the occurrence of faults. The reflux flow rate has suddenly increased to maintain the top composition of the column; however, it has then decreased due to the low level in the reflux drum. The low level in the reflux drum caused the distillate flow rate to decrease to zero flow. The reboiler duty has also decreased due to a direct link with reflux. The presence of faults F1 and F3 results in persistent oscillations in most of the process variables.

For ID7, another scenario based on three simultaneous faults is simulated by introducing reflux flow valve stiction (F1), reboiler valve stiction (F4), and the random variation in feed composition (F6) at time t = 5.0 h. Figure 10 shows the behavior of the column once these faults occur. A noticeable effect can be observed on the product flow rates of the column. These faults significantly affect the top and bottom compositions of the column. It can be observed that these simultaneous faults show the combined effects of all three faults discussed earlier. The valve stiction faults in reflux flow and reboiler duty valves along with the random variation in feed composition result in persistent fluctuations in all the key process variables as can be observed in Figure 10.

In ID8, three simultaneous faults have been introduced in the distillation column in the top and bottom tray efficiencies (F1 and F4), and random variation in the feed composition (F6). It can be witnessed from Figure 11 that these simultaneous faults have a significant impact on all the process variables in the column. Since F1 and F4 are severe faults, the effects of these two faults are significant as compared to random variation in feed composition.

In ID9, faults related to feed are simulated to observe the distillation column behavior. These faults are random variation in feed composition (F6), change in temperature (F7), and loss of feed flow rate (F8). The response of the column after the occurrence of three simultaneous faults at time t = 5.0 h is presented in Figure 12. It can be observed that since these faults contain loss of feed flow rate, its effect is predominant in all the variables.

A combination of four different faults (ID10), that is, reflux valve stiction (F1), reboiler valve stiction (F4), top tray upset (F3), and the bottom tray upset (F4), is simulated to observe the behavior of the distillation column. Figure 13 shows that at t = 5.0 h, faults occur in the system, which has significantly affected the overall behavior of the column. It can be observed that the faults due to tray upsets have substantially affected the column, which results in the low level in the reflux drum. Due to this, the reflux flow rate has been reduced, which affected the top and bottom compositions. Since these faults contain valve stiction, similar trends can also be seen when they were introduced individually. However, among all four faults, the faults related to tray upset have significantly affected the behavior of the column.

All the dataset obtained and discussed in this section (ID1–ID10) from the simultaneous faults is used for the training of MK-SVM for simultaneous fault classification.

4.2 Model selection for simultaneous fault diagnosis

The generated data from all combinations of the faults (ID1–ID10 in Table 4) are used for the classification of simultaneous faults in the distillation column. The training dataset comprises of 1000 observations for each single and simultaneous fault along with the 14 variables as specified in Table 6. The SVM has been trained with 19 different datasets (comprising one normal and 8 single and 10 simultaneous faults). The training dataset comprises of [19,000 * 14] predictor dataset and [19,000 * 1] response dataset (labeled as “class”). Therefore, the SVM has been trained with a total of [19,000 * 15] samples.

Figures 14-16 show the classification projection for all single and simultaneous faults in the distillation column using various kernels (linear, quadratic, and cubic). It can
be seen that all the fault datasets are overlapped. Hence, it is difficult to classify. To classify the overlapped data accurately, the proposed diagnosis method has been trained with different kernel functions. Based on the characteristic of each fault scenario, MK-SVM has an ability to classify most of the faults.

Furthermore, for multiclassification two common approaches are used, which are known as one-Vs-all (OVA) and one-Vs-one (OVO). In OVA, we have trained the M individual binary classifier; each one of them separates one class from the rest. Each classifier has been trained with a full training dataset. Therefore, the training set from

FIGURE 6  Response of distillation column for simultaneous faults F₁ and F₇
each class is required to separate a specific class from the remaining classes. The OVO approach consumes a lot of time as compared to OVA. Though the number of classifiers in OVA is less than OVO the size of the training set for every classifier in OVA is larger than OVO. It has been reported in the literature that OVA is more accurate than OVO. As shown in Table 7, the OVA has less training time as compared to OVO since the size of the training sets is much smaller for every binary classifier as compared to OVA. The training accuracy, time, and prediction speed of various SVM, that is, linear, quadratic, and Gaussian kernels, based on single and simultaneous faults are
presented in Table 7. The model has been trained using both OVA and OVO multiclass methods. It can be seen that OVA has higher training accuracy as compared to OVO. Hence, the OVA approach with quadratic kernel has been selected for performance evaluation of simultaneous faults in the distillation column.

4.3 | Confusion matrix

Table 8 shows the confusion matrix for MK-SVM for all single and simultaneous faults using quadratic kernel function. A clear representation of discriminated faults against each other can be observed from confusion matrix.
The normal and fault cases comprised 1000 observations in which most of the faults are classified in the right category. However, for each single and simultaneous fault, there is some misclassification. It can be observed from the confusion matrix that the higher misclassification rates of 1.5% and 1.8% are found in simultaneous faults F34 (ID4) and F346 (ID8), respectively. The possible reason for misclassification can be due to similar characteristics of both faults. The misclassification rate of 1.0% can also be seen in F12, which is misclassified equally in F1 and F2.
Furthermore, a misclassification rate of 1.8% is observed in F12 (ID1), in which 0.6% of the samples are classified as F1, and 1.2% of the samples are classified as F2. In the case of four simultaneous faults, that is, F1234 (ID10), most of the samples are classified accurately and shows a classification rate of 99.6%. However, F1234 (ID10) has also misclassified with the rate of 0.2%, 0.1% and 0.1%, in faults F12 (ID1) and F34 (ID4), respectively.

From the confusion matrix presented in Table 9, the fault detection rate (FDR) for all single and simultaneous faults is 99.51%, while the misclassification rate (MR) is 0.49%.
4.4 Performance evaluation based on simultaneous faults

Table 9 shows the performance of the proposed algorithm using the magnitude of F1 indices. The simultaneous faults in distillation column are simulated and then successfully diagnosed using SVM. The classification of various faults is done using the combination of two, three, and four simultaneous faults as described in Table 4.6. 500 samples from each simultaneous fault have been used for the testing of the proposed algorithm. It can be observed from the results in Table 4.9 that the proposed MK-SVM has good classification performance at all levels of simultaneous faults. It is witnessed that for two simultaneous faults, the F1 score is...
found to be >99.8%. However, a relatively lower F1 score is obtained for simultaneous faults $F_3$ and $F_4$, which are tray upset at the top and bottom stages, respectively. This is because both faults shared the common characteristics.

Moreover, F1 index for two simultaneous faults $F_1$ and $F_2$ is found to be 100.0%. Similar performance can also be seen when classifying three simultaneous faults in the distillation column. For three simultaneous faults, the F1 index is found to be >99.8% for all cases except for faults $F_3$, $F_4$, and $F_6$, which is found to be 98.50%, 98.90%, and 99.55%, respectively. Four simultaneous faults have been simulated and successfully diagnosed, as shown in Figure 12.

**Figure 12** Responses of distillation column for simultaneous faults $F_6$, $F_7$, and $F_8$
Table 9. For the combination of four simultaneous faults, that is, F_1, F_2, F_3, and F_4, SVM produces promising classification results, and an overall F1 index of >99.0% is achieved. In summary, the proposed MK-SVM approach shows promising diagnosis performance for two, three, and four simultaneous faults in the distillation column.

### 4.5 MK-SVM comparison of simultaneous fault classification

A comparison of the proposed MK-SVM is made with other data-based algorithms for simultaneous fault, as shown in Table 10. Two, three, and four simultaneous faults are considered.
faults have been considered for the comparison of classification performance. Based on the preliminary screening analysis not reported in this paper, comparison results from multiple fault classification KNN and ANN are found to be better as compared to other established classification algorithms. Therefore, for simultaneous fault classification, only these two data-based algorithms have been used to compare with the proposed MK-SVM. It is observed that the classification accuracy of the proposed MK-SVM is better (ie, >98%) for all combinations of faults as compared to the established KNN and ANN.

It can be observed that the classification accuracy of ANN and KNN has decreased with increase in the number of simultaneous faults. It can be observed from Table 9 that both ANN and KNN show poor performance as compared to MK-SVM for three and four simultaneous faults.

5 | CONCLUSION

In this study, simultaneous fault diagnosis in nonlinear dynamic distillation column has been done using multiple kernel support vector machine (MK-SVM) approach. The MK-SVM was used as a multiclass classifier for the classification of simultaneous faults in the distillation column. The study was conducted on two available techniques such as one-vs-one (OVO) and one-vs-all (OVA). It was observed that the OVA has higher accuracy as compared to OVO. The models were trained to find out better training efficiency. Hence, the accuracy of quadratic kernel–based SVM was found to be higher (ie, 99.70%) among all and it was selected to perform fault classification. Combinations of two, three, and four faults were used to evaluate the performance of the proposed fault classifier. From the confusion matrix, it can be observed that the fault detection rate (FDR) for all single and simultaneous faults was 99.51%, while the misclassification rate (MR) was 0.49%. The classification accuracy was measured by F1 score, and it noticed that the F1 score is >98% for all combination of faults. A comparison of the proposed MK-SVM is made with established ANN and KNN for simultaneous fault classification. It is observed that the classification accuracy of the proposed MK-SVM is better (ie, >98%) for all combinations of faults as compared to the established KNN and ANN. Hence, it is concluded that the proposed

Table 6: Manipulated and observed variables of the distillation column

| Variables | Description |
|-----------|-------------|
| OBV (1)   | Top composition |
| OBV (2)   | Bottom composition |
| OBV (3)   | Column pressure |
| OBV (4)   | Feed flow rate |
| OBV (5)   | Feed temperature |
| OBV (6)   | Top level |
| OBV (7)   | Bottom level |
| OBV (8)   | Distillate flow |
| OBV (9)   | Bottom flow |
| OBV (10)  | Top Temp |
| OBV (11)  | Bottom Temp |
| OBV (12)  | Reflux ratio |
| MV (13)   | Reboiler duty |
| MV (14)   | Condenser duty |

Figure 14: Simultaneous fault classification projection using MK-SVM (linear kernel)
FIGURE 15 Simultaneous fault classification projection using MK-SVM (quadratic kernel)

FIGURE 16 Simultaneous fault classification projection using MK-SVM (cubic kernel)

TABLE 7 Performance of various kernel functions using OVA and OVO approach

| MK-SVM model | One-vs-all | One-vs-one |
|--------------|------------|------------|
|              | Accuracy (%) | Training time (s) | Prediction speed obs/s | Accuracy (%) | Training time (s) | Prediction speed (obs/s) |
| Linear       | 86.60       | 4760.30     | 29,000                | 95.10        | 290.3           | 4200                   |
| Quadratic    | 99.70       | 1049.20     | 11,000                | 99.40        | 820.34          | 3500                   |
| Cubic        | 87.00       | 32,027.10   | 18,000                | 96.80        | 987.76          | 3900                   |
| Fine         | 93.50       | 918.08      | 970                   | 97.5         | 653.96          | 250                    |
| Medium       | 96.0        | 990.85      | 1300                  | 96.40        | 180.07          | 1500                   |
| Course       | 95.2        | 1217.5      | 910                   | 93.10        | 215.76          | 1100                   |
TABLE 8
Confusion matrix for single and simultaneous faults

| Single and simultaneous faults | N | F_1 | F_2 | F_3 | F_4 | F_5 | F_6 | F_7 | F_8 | F_12 | F_13 | F_17 | F_34 | F_38 | F_123 | F_126 | F_346 | F_678 | F_1234 |
|--------------------------------|---|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|-------|-------|-------|-------|-------|
| N                             | 100 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.5   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1   |
| F_1                           | 0   | 99.9 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.1    | 0.1    | 0.1    | 0.1    | 0.1   |
| F_2                           | 0   | 0     | 99.5 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.7    | 0.7    | 0.7    | 0.7    | 0.7   |
| F_3                           | 0   | 0.1   | 0.1   | 99.0 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 99.8   | 99.8   | 99.8   | 99.8   | 99.8   |
| F_4                           | 0   | 0.1   | 0.1   | 0.1   | 99.8 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_5                           | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_6                           | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0     | 0     | 0     | 0     | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_7                           | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0     | 0     | 0     | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_8                           | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0     | 0     | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_12                          | 0   | 0.5   | 0.5   | 0.5   | 0.5   | 0.5   | 0.5   | 0.5   | 0.5   | 99.5 | 0     | 0     | 0     | 0     | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_13                          | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0     | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_17                          | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0     | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_34                          | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0.1    | 0.1    | 0.1    | 0.1    | 0.1    |
| F_38                          | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9 | 0     | 0.1    | 0.1    | 0.1    | 0.1    |
| F_123                         | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 99.9   | 99.9   | 99.9   | 99.9   | 99.9   |
| F_126                         | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    | 99.9   | 99.9   | 99.9   | 99.9   |
| F_346                         | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    | 0.1    | 99.9   | 99.9   | 99.9   |
| F_678                         | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    | 0.1    | 0.1    | 99.9   | 99.9   |
| F_1234                        | 0   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    | 0.1    | 0.1    | 0.1    | 99.6   |

Notes: Green shades indicate Classified. Orange shades indicate Misclassified.
MK-SVM classifier was found to produce better results in the case for two, three, and four simultaneous faults.

**ACKNOWLEDGMENT**
The author acknowledges the valuable assistance and support of Chemical Engineering Department, Universiti Teknologi PETRONAS, and NED University of Engineering and Technology.

**CONFLICT OF INTEREST**
All authors declare that they have no conflict of interest.

**ORCID**
Muhammad Naqvi  
https://orcid.org/0000-0002-4359-2232

**REFERENCES**

1. Wakasugi H, Sueyoshi K. Energy-saving Solutions by Advanced Process Control (APC) Technology. Yokogawa Technical Report English Edition; 2010.

2. Kumar M, Ganapathy K, Al Yahyaee N, Ayoub A. Enhance energy efficiency through advanced process control APC. In: Abu Dhabi International Petroleum Exhibition & Conference; 2016.

3. Chetouani Y. Change detection in a distillation column based on the generalized likelihood ratio approach. *J Chem Eng Process Technol.* 2011;2(5):1-6.

4. Behbahani RM, Jazayeri-Rad H, Hajmirzae S. Fault detection and diagnosis in a sour gas absorption column using neural networks. *Chem Eng Technol.* 2009;32:840-845.

5. Economics M-H. Survey of Investment in Employee Safety and Health. McGraw-Hill Publishing Co; 1985.

6. Venkatasubramanian V. A review of process fault detection and diagnosis. Part I: quantitative model-based methods. *Comput Chem Eng.* 2003;27:293-311.

7. Itasca I. The National Safety Council: Injury Facts 1999 Edition. The National Safety Council; 2013:9-12.

8. Lau C, Ghosh K, Hussain M, Hassan CC. Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS. *Chemomet Intellig Lab Syst.* 2013;120:1-14.

9. Tidriri K, Chatti N, Verron S, Tiplica T. A Bond Graph Modeling for Health Monitoring and Diagnosis of the Tennessee Eastman Process. System See (3). Vol. 4; 2017:5.

10. Zhao H, Hu Y, Ai X, Hu Y, Meng Z. Fault detection of Tennessee Eastman process based on topological features and SVM. *IOP Conf Ser: Mater Sci Eng.* 2018;339:012039.

11. Zhou B, Gu X. Multi-block statistics local kernel principal component analysis algorithm and its application in nonlinear process fault detection. *Neurocomputing.* 2020;376:222-231.

12. Zio E, Baraldi P, Popescu IC. A fuzzy decision tree method for fault classification in the steam generator of a pressurized water reactor. *Ann Nucl Energy.* 2009;36:1159-1169.

13. Chetouani Y. Using Artificial Neural networks for the modeling of a distillation column. *IJCSA.* 2007;4:119-133.

14. Chetouani Y. Detecting changes in a distillation column by using a sequential probability ratio test. *Syst Eng Proc.* 2011;1:473-480.

15. Manssouri I, Chetouani Y, Kihel BE. Using neural networks for fault detection in a distillation column. *Int J Comput Appl Technol.* 2008;32:181-186.

16. Rajakarunakaran S, Venkumar P, Devaraj D, Rao KSP. Artificial neural network approach for fault detection in rotary system. *Appl Soft Comput.* 2008;8:740-748.
17. Nozari HA, Shoorehdeli MA, Simani S, Banadaki HD. Model-based robust fault detection and isolation of an industrial gas turbine prototype using soft computing techniques. *Neurocomputing*. 2012;91:29-47.
18. Tian WD, Sun SL, Guo QJ. Fault detection and diagnosis for distillation column using two-tier model. *Can J Chem Eng*. 2013;91:1671-1685.
19. Riggs JB. *Chemical Process Control*. Ferret; 1999.
20. Sari AHA. *Data-Driven Design of Fault Diagnosis Systems: Nonlinear Multimode Processes*. Springer Science & Business; 2014.
21. Geng Z, Zhu Q. Multiscale nonlinear principal component analysis (NLPCA) and its application for chemical process monitoring. *Ind Eng Chem Res*. 2005;44:3585-3593.
22. Jie X, Shousong H. Nonlinear process monitoring and fault diagnosis based on KPCA and MKL-SVM. *Chin J Sci Instrum*. 2010;31:2428-2433.
23. Lee J-M, Yoo C, Choi SW, Vanrolleghem PA, Lee I-B. Nonlinear process monitoring using kernel principal component analysis. *Chem Eng Sci*. 2004;59:223-234.
24. Martínez-Guerra R, Mata-Machuca JL. *Fault Detection and Diagnosis in Nonlinear Systems*. 2012
25. Qin SJ. *Survey on data-driven industrial process monitoring.* *Chem Eng J*. 2004;105:107988.
26. Venkatasubramanian V, Rengaswamy R, Kavuri SN. A re- view of process fault detection and diagnosis: Part II: qualitative models and search strategies. *Comput Chem Eng*. 2003;27:313-326.
27. Tidiriri K, Chatti N, Verron S, Tipica T. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: a review of researches and future challenges. *Annu Rev Control*. 2012;36(2):220-234.
28. Venkatasubramanian V, Rengaswamy R, Kavuri SN. A review of process fault detection and diagnosis: Part II: qualitative models and search strategies. *Comput Chem Eng*. 2003;27:313-326.
29. Taqvi SAA, Tufa LD, Zabiri H, F. Rigorous dynamic modelling and identification of dis- tillation column using Aspen Plus. In: *Ferret Process Loops: State of the Art and Advanced Methods*. Springer Science & Business Media; 2009.
30. Teh WK, Zabiri H, Samyudia Y, et al. An improved diagnostic tool for control valve stiction based on nonlinear principle component analysis. *Ind Eng Chem Res*. 2018;57:11350-11365.
31. di Capaci RB, Scali C. Review and comparison of techniques of analysis of valve stiction: from modeling to smart diagnosis. *Chem Eng Res Des*. 2018;130:230-265.
32. Jelali M, Huang B. *Detection and Diagnosis of Stiction in Control Loops: State of the Art and Advanced Methods*. Springer Science & Business Media; 2009.
54. Belincanta J, Ravagnani T, Pereira J. Hydrodynamic and tray efficiency behavior in parastillation column. *Braz J Chem Eng*. 2006;23:135-146.
55. Luyben WL. *Practical Distillation Control*. Springer Science & Business Media; 2012.
56. Roffel B. Distillation: instrumentation and control systems. In: *Encyclopedia of Separation Science*. Academic Press; 2000.

**How to cite this article:** Taqvi SAA, Zabiri H, Uddin F, et al. Simultaneous fault diagnosis based on multiple kernel support vector machine in nonlinear dynamic distillation column. *Energy Sci Eng*. 2022;10:814–839. doi:10.1002/ese3.1058