



# A normal operating zone model for condition monitoring of reaction-separation-recycle processes\*

Mengyao Wei<sup>a</sup>, Jiandong Wang<sup>a,\*</sup>, Zhishan Zhang<sup>b</sup>

<sup>a</sup> College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, Shandong Province, China

<sup>b</sup> College of Chemical and Biological Engineering, Shandong University of Science and Technology, Qingdao, Shandong Province, China



## ARTICLE INFO

### Keywords:

Normal operating zone model  
Uncertainty measurements  
Condition monitoring  
Reaction-separation-recycle processes

## ABSTRACT

Normal operation of reaction-separation-recycle processes is important for the safety of chemical plants. This study presents a normal operating zone (NOZ) model for condition monitoring of reaction-separation-recycle processes. The NOZ model is a geometric space in higher dimensions that is developed based on physical principles, historical data, and operational requirements. A main technical issue is the uncertainties of NOZ models from historical data with noise disturbance. Bayesian estimation theory is utilized to address this issue. The NOZ model represents an allowable variation region of multiple related process variables, and compensates certain limitations of existing methods. In particular, the NOZ model can account for normal conditions that are not present in historical data compared to data-based methods; the NOZ model can identify abnormal conditions that do not change physical models in contrast to model-based methods. The effectiveness of the proposed method is illustrated via Aspen-Hysys-based examples with comparison to existing methods.

## 1. Introduction

Many processes in different chemical industries, such as ethylene oxide hydration, hydrotreating, and esterification, are reaction-separation-recycle processes (Roostaei and Eslamloueyan, 2024). The reaction-separation-recycle process has been widely used in the design and control of chemical processes (Pattison et al., 2017). With the development of renewable energy, the electrification of chemical processes is universal (Barton, 2020). In order to match renewable energy, operation conditions of chemical processes change frequently, which leads to an increasing probability of abnormal conditions. Therefore, it is necessary to monitor the reaction-separation-recycle processes for the safety and economy of chemical plants.

Condition monitoring of reaction-separation-recycle processes can mainly be classified into model-based and data-based methods. Model-based methods identify abnormalities by comparing the measured process outputs with the counterparts from mathematical models of chemical processes. Bhagwat et al. (2003) presented a fault detection scheme based on nonlinear process models during process transitions. Yu et al. (2020) proposed a deep neural network model for fault detection in chemical processes. Taqvi et al. (2020) developed a nonlinear autoregressive with exogenous input network-based fault detection method

for distillation columns. Du et al. (2024) derived discrete-time models of chemical reactor systems and constructed specific residual generators for coolant inlet temperature and feed concentration faults. Fast Fourier Transform filtering was employed for fault diagnosis. Data-based methods detect abnormalities based on relationships among multiple variables via multivariate statistical analysis using historical data. Principal component analysis (PCA) and its improved techniques were used to address chemical process monitoring (Mansouri et al., 2016; Nawaz et al., 2020; You and Chen, 2021). Slow feature analysis methods were presented to monitor process dynamic abnormalities (Shang et al., 2015; Xu et al., 2022; Zhong et al., 2020). Botre et al. (2017) provided a fault detection method for chemical processes using a multiscale partial least squares-based generalized likelihood ratio test. Madakyaru et al. (2019) adopted data-driven abnormal detection schemes based on a nonlinear latent variable model and statistical monitoring algorithms. Ali et al. (2024) designed a machine learning-based distributed canonical correlation analysis for fault detection. Villagomez et al. (2024) used steady-state simulation data to create a latent space model of how chemical processes normally work.

This paper builds a normal operating zone (NOZ) model for condition monitoring in reaction-separation-recycle processes. The NOZ model is a geometric space developed based on physical principles, historical data,

\* This research was partially supported by the National Natural Science Foundation of China under Grant no. 62273215.

\* Corresponding author.

E-mail addresses: [mywei9701@163.com](mailto:mywei9701@163.com) (M. Wei), [jiandong@sdust.edu.cn](mailto:jiandong@sdust.edu.cn) (J. Wang), [tjizza@163.com](mailto:tjizza@163.com) (Z. Zhang).

## Nomenclature

$\alpha$	recycle ratio
$\rho_A$	density of component A
$\rho_B$	density of component B
$\rho_E$	density of component E
$C_{0,A}$	concentration of A in the make-up feed flow
$C_{b,E}$	concentration of E in DC bottom product flow
$C_{d,A}$	concentration of A in DC top product flow
$C_{in,A}$	concentration of A in CSTR feed flow
$C_{r,A}$	concentration of A in reactor product flow
$C_{r,E}$	concentration of E reactor product flow
$E_a$	activation energy
$F_0$	make-up feed flow rate
$F_b$	product flow rate of DC bottom
$F_d$	product flow rate of DC top
$F_{in}$	feed flow rate of CSTR
$F_r$	reactor product flow rate
$k_0$	pre-exponential factors
$k_1$	reaction rate constant
$m_A$	molar mass of component A
$m_B$	molar mass of component B
$m_E$	molar mass of component E
$R$	universal gas constant
$s$	volume change per kilomole in CSTR
$T_r$	reactor temperature
$V_r$	reaction volume in CSTR
max	maximums of process variables in normal conditions
min	minimums of process variables in normal conditions

and operational requirements. First, a physical model is established to describe the relationships among multiple process variables, and unknown model parameters are identified based on historical data. Second, according to the physical model and operational requirements, the geometric space consisting of all normal conditions is obtained, and is described by a convex hull as the NOZ model. A main technical issue is the uncertainties of NOZ models from historical data with noise disturbance. This technical issue is addressed by utilizing Bayesian estimation theory to yield convergent distributions of model parameters.

The NOZ model represents an allowable variation region of multiple related process variables for reaction-separation-recycle processes. Such a region is analogous to a map containing all normal conditions and showing where the current operating point of processes is located in the map. Boundaries of the NOZ model (see e.g., the black solid lines

in Fig. 10 given later in Section 4) are like cliff edges in the map. If the distance of an operating point away from the boundaries is larger, then the location is safer. Therefore, the NOZ model has potential applications for condition monitoring of reaction-separation-recycle processes. In addition, the NOZ model compensates for certain limitations of existing methods. In particular, it can account for normal conditions that are not present in historical data compared to data-based methods. In contrast to model-based methods, it can identify abnormal conditions that do not change physical models.

The rest of the paper is organized as follows. Section 2 presents the problem description. Section 3 gives detailed steps of the proposed method. Application examples are provided in Section 4. Section 5 makes some concluding remarks.

## 2. Problem description

A reaction-separation-recycle process is illustrated in Fig. 1. The process produces E via the irreversible exothermic addition of A to B as



The goal of this process is to convert the feed components A and B to product E. The make-up feed flow (components A and B) mixes with the recycle flow (also primarily components A and B) and is fed to the isothermal continuous stirred-tank reactor (CSTR). A first-order reaction occurs; the reaction is incomplete, so the reactor product flow (which is fed to the distillation column) contains both components A, B, and E. The distillation column (DC) separates this feed flow into two relatively pure product flows. The distillate product, containing primarily A and B, is recycled; the bottoms product, which contains primarily component E, is sent to storage.

The process involves many related variables, as shown in Fig. 1. Among them, the key process variables are the make-up feed flow rate  $F_0$ , concentration of A in the make-up feed flow  $C_{0,A}$ , reactor temperature  $T_r$ , reactor product flow rate  $F_r$  (which is the feed flow rate of DC), product flow rate  $F_b$  of DC bottom, and concentration of E in the bottom product flow  $C_{b,E}$ . When the process is in normal conditions, these process variables are within certain variation ranges. In particular, the reactor temperature  $T_r$  needs to meet a range requirement to ensure that the reaction proceeds normally; the feed flow rate  $F_r$  needs to meet a load requirement that the DC can withstand; the flow rate  $F_b$  needs to be greater than the minimum load requirement; and the concentration  $C_{b,E}$  needs to meet the product specifications. These operational requirements are formulated as follows:

$$T_{r,\min} \leq T_r \leq T_{r,\max}, \quad (2)$$

$$F_{r,\min} \leq F_r \leq F_{r,\max}, \quad (3)$$

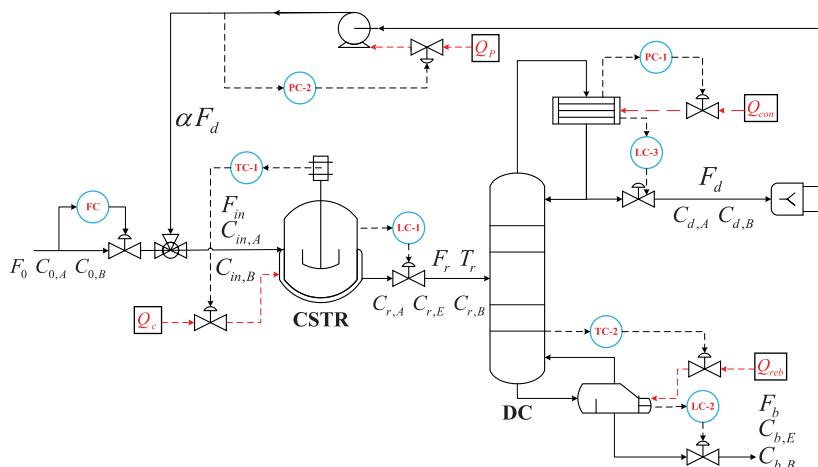
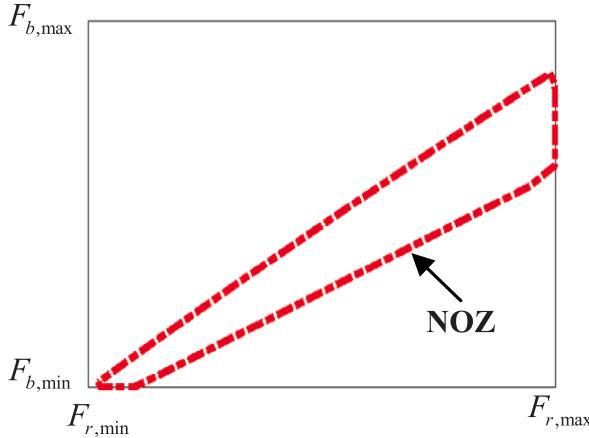


Fig. 1. A schematic diagram of a reaction-separation-recycle process.



**Fig. 2.** Schematic diagram of a NOZ for two related process variables.

$$F_{b,\min} \leq F_b \leq F_{b,\max}, \quad (4)$$

$$C_{b,E,\min} \leq C_{b,E} \leq C_{b,E,\max}. \quad (5)$$

These process variables are closely related in physics, so that the requirements in (2)–(5) should not be isolated. For instance, the allowable variation region of process variables  $F_r$  in (3) and  $F_b$  in (4) is a rectangular area without considering their relationship, as illustrated in Fig. 2. Considering the physical relationships of process variables  $F_r$  and  $F_b$ , the allowable variation region of  $F_r$  and  $F_b$  is an irregular area, depicted as the area enclosed by the red dashed line in Fig. 2, which is called as a normal operating zone (NOZ). For the reaction-separation-recycle process, its NOZ model is an irregular geometric space in six dimensions, which takes into account relationships among the key process variables  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$ . In other words, the process is in normal conditions when operating points are located inside the NOZ model.

Given historical data of the reaction-separation-recycle process, the objective of this paper is to establish a NOZ model that is a geometric space consisting of  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$  for the reaction-separation-recycle process. However, there are uncertainties in the NOZ model because unknown model parameters are identified based on historical data with noise disturbance. In order to assess the credibility of the established NOZ model in a statistical sense, the uncertainty measurements of NOZ models need to be addressed.

### 3. Proposed method

This section first establishes physical models of reaction-separation-recycle processes. Second, the NOZ model is developed based on the physical model. Next, the uncertainty of the NOZ model is measured through Bayesian estimation theory.

#### 3.1. Physical models

The reaction-separation-recycle process, as shown in Fig. 1, consists of the reactor, distillation column, and recirculation. The reactor is where the reaction of components A and B provides the required conditions, such as temperature and pressure. It is assumed that the reactor is well mixed and only the primary reaction in (1) occurs. The model of CSTR is developed based on the material balances of the reactants A and B and the desired product E (Du et al., 2024),

$$F_{in}C_{in,A} = F_rC_{r,A} - r(C_{r,A}, C_{r,B}, T_r)V_r, \quad (6)$$

$$F_{in}C_{in,B} = F_rC_{r,B} - r(C_{r,A}, C_{r,B}, T_r)V_r, \quad (7)$$

$$F_rC_{r,E} = r(C_{r,E}, C_{r,B}, T_r)V_r, \quad (8)$$

$$F_{in} = F_r + r(C_{r,A}, C_{r,B}, T_r)V_r s. \quad (9)$$

Here  $r(C_{r,A}, C_{r,B}, T_r)$  denotes the reaction rate depending on  $C_{r,A}$ ,  $C_{r,B}$ , and  $T_r$ ,

$$r(C_{r,A}, C_{r,B}, T_r) = k_1(T_r)C_{r,A}C_{r,B}, \quad (10)$$

where  $k_1(T_r)$  is the reaction rate constant. It can be obtained from the Arrhenius equation,

$$k_1(T_r) = k_0 \exp\left(-\frac{E_a}{RT_r}\right), \quad (11)$$

where  $k_0$  is the pre-exponential factors and  $E_a$  is the activation energy. The values of  $k_0$  and  $E_a$  vary under different catalytic conditions. They are generally estimated based on empirical or experimental data (Altintokka and Akyalcin, 2009).

Distillation columns are used to separate A, B and E in the reactor output mixture. It is assumed that the bottom product of the DC does not contain component A, while the top product does not contain component E. The material balances of the reactant A and the desired product E are respectively (Taqvi et al., 2020),

$$F_rC_{r,A} = F_dC_{d,A}, \quad (12)$$

$$F_rC_{r,B} = F_bC_{b,B} + F_dC_{d,B}, \quad (13)$$

$$F_rC_{r,E} = F_bC_{b,E}. \quad (14)$$

The total material balance of the DC is

$$F_r = F_b + F_d. \quad (15)$$

The recirculation is mainly the refluxing of the solution containing components A and B from the distillate product to the CSTR. Therefore, the total feed flow of the CSTR is the sum of the make-up feed flow and the recycled flow.

$$F_{in} = F_0 + \alpha F_d. \quad (16)$$

Here  $\alpha$  is a recycle ratio of the circulation to the DC top product. The volume flow of A and B into the CSTR are

$$F_{in}C_{in,A} = C_{0,A}F_0 + \alpha C_{d,A}F_d, \quad (17)$$

and

$$F_{in}C_{in,B} = C_{0,B}F_0 + \alpha C_{d,B}F_d. \quad (18)$$

In addition, the concentration of B in each flow stream satisfies the following relationship with the concentrations of A and E (Kumar and Kaistha, 2018),

$$C_{in,B} = \frac{\rho_B}{m_B} \left( 1 - \frac{m_A C_{in,A}}{\rho_A} \right), \quad (19)$$

$$C_{b,B} = \frac{\rho_B}{m_B} \left( 1 - \frac{m_E C_{b,E}}{\rho_E} \right), \quad (20)$$

$$C_{d,B} = \frac{\rho_B}{m_B} \left( 1 - \frac{m_A C_{d,A}}{\rho_A} \right), \quad (21)$$

$$C_{r,B} = \frac{\rho_B}{m_B} \left( 1 - \frac{m_A C_{r,A}}{\rho_A} - \frac{m_E C_{r,E}}{\rho_E} \right). \quad (22)$$

The physical model of the reaction-separation-recycle process is described by (6)–(22). Here  $\rho_E$ ,  $\rho_A$ ,  $\rho_B$ ,  $m_E$ ,  $m_A$ ,  $m_B$ ,  $V_r$ ,  $s$ , and  $\alpha$  are known parameters.

For unknown parameters  $k_0$  and  $E_a$ , their relationship with process variables is derived by (6), (10), and (11),

$$\frac{F_{in}C_{in,A} - F_rC_{r,A}}{V_rC_{r,A}C_{r,B}} = k_0 \exp\left(-\frac{E_a}{RT_r}\right). \quad (23)$$

The model parameters  $k_0$  and  $E_a$  are estimated from historical data samples  $\{F_{in}(t), C_{in,A}(t), F_r(t), C_{r,A}(t), C_{r,B}(t), T_r(t)\}_{t=1}^T$ . By using a steady-state data extraction technique (Wang et al., 2022),  $M$  steady-state values where process variables are simultaneously in steady-state conditions are found as  $\{T_r[m], F_{in}[m], C_{in,A}[m], F_r[m], C_{r,A}[m], C_{r,B}[m]\}_{m=1}^M$ . By defining  $\theta := [k_0, E_a]$  and  $y := \frac{F_{in}C_{in,A}-F_rC_{r,A}}{V_rC_{r,A}C_{r,B}}$ , Eq. (23) can be rewritten as

$$y = f(T_r; \theta) = k_0 \exp\left(-\frac{E_a}{RT_r}\right). \quad (24)$$

The model parameter  $\theta$  is estimated by minimizing the sum of squared error between  $y[m]$  and the estimated output  $\hat{y}(T_r[m]; \theta)$ ,

$$\hat{\theta} = \arg \min_{\theta} \left( \sum_{m=1}^M (y[m] - \hat{y}(T_r[m]; \theta))^2 \right), \quad (25)$$

where  $y[m] := \frac{F_{in}[m]C_{in,A}[m]-F_r[m]C_{r,A}[m]}{V_rC_{r,A}[m]C_{r,B}[m]}$  and  $\hat{y}(T_r[m]; \theta) := f(T_r[m]; \theta)$ . The non-convex optimization problem in Eq. (25) can be solved by exploiting some global optimization algorithms, such as the genetic algorithm (Aryanezhad and Hemati, 2008).

### 3.2. Normal operating zone models

This subsection formulates NOZ models by convex hulls for reaction-separation-recycle processes.

First, all normal operating points of the reaction-separation-recycle process are obtained. Given the step sizes  $\Delta F_0$ ,  $\Delta C_{0,A}$ , and  $\Delta T_r$ , the input variables  $F_0$ ,  $C_{0,A}$ , and  $T_r$  take values from their own minimums to maximums as,

$$\begin{aligned} F_0[n_F] &= F_{0,\min} + n \cdot \Delta F_0, n_F = 1, \dots, N_{F_0}; \\ C_{0,A}[n_C] &= C_{0,A,\min} + n \cdot \Delta C_{0,A}, n_C = 1, \dots, N_{C_{0,A}}; \\ T_r[n_T] &= T_{r,\min} + n \cdot \Delta T_r, n_T = 1, \dots, N_{T_r}. \end{aligned} \quad (26)$$

Here  $N_{F_0} = \left\lfloor \frac{F_{0,\max}-F_{0,\min}}{\Delta F_0} \right\rfloor$ ,  $N_{C_{0,A}} = \left\lfloor \frac{C_{0,A,\max}-C_{0,A,\min}}{\Delta C_{0,A}} \right\rfloor$ , and  $N_{T_r} = \left\lfloor \frac{T_{r,\max}-T_{r,\min}}{\Delta T_r} \right\rfloor$  are the number of possible values for  $F_0$ ,  $C_{0,A}$ , and  $T_r$ , respectively. The symbol  $\lfloor \cdot \rfloor$  is the largest integer less than or equal to the operand. Considering all possible combinations of the input variables in (26), a model-input set of the physical models in (6)–(22) is formed as  $\{F_0[n], C_{0,A}[n], T_r[n]\}_{n=1}^N$  with  $N = N_{F_0}N_{C_{0,A}}N_{T_r}$ . Correspondingly, the model outputs  $F_r$ ,  $F_b$ , and  $C_{b,E}$  are generated through the physical models in (6)–(22), where the unknown parameters  $k_0$  and  $E_a$  in (11) are replaced by their estimated value in (25). As a result,  $N$  steady-state operating points consisting of six process variables are obtained as

$$\begin{aligned} X &= \{x[n; \theta]\}_{n=1}^N \\ &= \{F_0[n], C_{0,A}[n], T_r[n], F_r[n; \theta], F_b[n; \theta], C_{b,E}[n; \theta]\}_{n=1}^N. \end{aligned} \quad (27)$$

The operating points in the set  $X$  that satisfy in Eqs. (3)–(5) are selected to form a set  $X_n$  of normal operating points,

$$X_n = \{x[n; \theta] | F_{r,\min} \leq F_r[n; \theta] \leq F_{r,\max}, F_{b,\min} \leq F_b[n; \theta] \leq F_{b,\max}, C_{b,E,\min} \leq C_{b,E}[n; \theta] \leq C_{b,E,\max}, \forall x \in X\} \quad (28)$$

Second, a NOZ model is formulated via convex hulls to describe the geometric space composed by six-dimensional normal operating points  $X_n$  in (28). These normal operating points are normalized to eliminate differences in the value range and units of process variables. For a normal operating point  $x[n; \theta]$  in  $X_n$ , it is normalized as

$$\bar{x}[n; \theta] = \{\bar{F}_0[n], \bar{C}_{0,A}[n], \bar{T}_r[n], \bar{F}_r[n; \theta], \bar{F}_b[n; \theta], \bar{C}_{b,E}[n; \theta]\}. \quad (29)$$

Here  $\bar{F}_0 = \frac{F_0-\mu_{F_0}}{\sigma_{F_0}}$ , where  $\mu_{F_0}$  and  $\sigma_{F_0}$  denote the sample mean and standard deviation of  $F_0$ , respectively.  $\bar{C}_{0,A}$ ,  $\bar{T}_r$ ,  $\bar{F}_r$ ,  $\bar{C}_{b,E}$ , and  $\bar{F}_b$  are defined similarly to  $\bar{F}_0$ . Thus, a set  $\bar{X}_N = \{\bar{x}[n; \theta]\}_{n=1}^{N_n}$  is derived based on  $X_N$  in (28) by (29). Here  $N_n$  denotes the number of normal operating points. A

convex hull is defined as the smallest convex polytope that contains all the operating points in  $\bar{X}_N$ , which is formulated as Yu and Wang (2019)  $A_a \bar{x}^T[n; \theta] - B_b \leq 0$ , (30)

where

$$\begin{cases} A_a = [a_1, \dots, a_i, \dots, a_I]^T \\ B_b = [b_1, \dots, b_i, \dots, b_J]^T \\ \mathbf{0} = [0, \dots, 0, \dots, 0]^T. \end{cases} \quad (31)$$

Here  $a_i = [a_{i,1}, \dots, a_{i,j}, \dots, a_{i,J}]$  is the unit normal vector of the  $i$ th hyperplane  $p(i)$ , and  $b_i$  is the distances of  $p(i)$  away from the origin point. The hyperplane  $p(i)$  can be expressed as

$$a_i \bar{x}_{p(i)}^T - b_i = 0, \quad (32)$$

where  $\bar{x}_{p(i)}$  is an operating point on the  $i$ th hyperplane  $p(i)$ .  $p(i)$  passes through  $J$  vertices, namely,  $\bar{x}[H_{i,1}]$ , ...,  $\bar{x}[H_{i,j}]$ , and  $\bar{x}[H_{i,J}]$ . These vertices are generated based on a  $I \times J$  matrix  $H$  returned by the quick-hull algorithm (Barber et al., 1996), where  $I$  represents the number of hyperplanes in the convex hull. The operating point  $\bar{x}[H_{i,j}]$  is connected to the element  $H_{i,j}$  on the  $i$ th row and  $j$ th column of  $H$ , which is the  $j$ th vertex of the  $i$ th hyperplane. These  $J$  vertices are substituted into (32) to get

$$a_i \bar{x}^T[H_{i,j}] - b_i = 0, j = 1, 2, \dots, J, \quad (33)$$

where  $\bar{x} = [\bar{F}_0, \bar{C}_{0,A}, \bar{T}_r, \bar{F}_r, \bar{F}_b, \bar{C}_{b,E}]$ . Since  $a_i$  is a unit vector, that is

$$a_{i,1} \cdot a_{i,1} + \dots + a_{i,j} \cdot a_{i,j} + \dots + a_{i,J} \cdot a_{i,J} = 1, \quad (34)$$

$\hat{a}_i$  and  $\hat{b}_i$  are estimated by solving (33) and (34), so that the convex hull model in (30) is founded based on  $\hat{A}$  and  $\hat{B}$ .

### 3.3. Uncertainties measurements

This subsection presents uncertainty measurements for the NOZ model through Bayesian estimation theory.

Data samples of process variables are usually disturbed by noise for reaction-separation-recycle processes. This results in uncertainties in the physical models in (6)–(22) identified based on historical data, which in turn lead to uncertainties in the NOZ model in (30). Therefore, it is essential to provide the uncertainty measurements of the NOZ model for reaction-separation-recycle processes.

First, the uncertainty of the physical model is described by confidence intervals of the model parameter  $\theta$ ,

$$\iint_{\Theta_u} \hat{p}(\theta) d\theta = 1 - \beta. \quad (35)$$

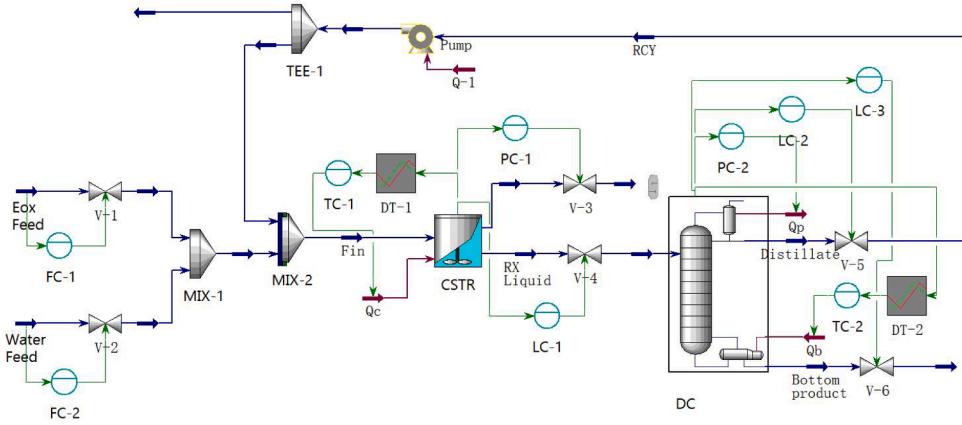
Here  $\beta$  is the significance level,  $\Theta_u$  denotes the bounded closed region of the  $k_0 - E_a$  plane to measure the uncertainty of  $\theta := [k_0, E_a]$ , and  $\hat{p}(\theta)$  denotes the probability density function (PDF) of  $\theta$ .

Second, the PDF of  $\theta$  is obtained from samples  $\{\hat{\theta}(q)\}_{q=1}^Q$  through Bayesian estimation theory (Walpole et al., 1993). The samples  $\{\hat{\theta}(q)\}_{q=1}^Q$  are identified from multiple groups of historical data by (25). In order to count the distribution of  $\theta$ , the sample space of  $\theta$  is separated into  $N_H$  tiny spaces equally. The frequency  $f_h$  of  $\theta$  in the  $h$ -th space is calculated by the ratio of the number  $q_h$  of samples in  $\{\hat{\theta}(q)\}_{q=1}^Q$  belonging to the  $h$ -th space to the total number  $Q$  of samples  $\{\hat{\theta}(q)\}_{q=1}^Q$ , i.e.,

$$f_h = \frac{q_h}{Q}. \quad (36)$$

$f_h$  as a realization of a random variable  $F_h$  denotes the probability that  $\theta = \theta_d$ . Any value within the  $h$ -th space has the same probability of being the true value of  $f_h$ . Without loss of generality, the prior PDF of  $F_h$  is assumed as

$$p_{F_h}(f_h) = \begin{cases} 1, & \text{if } 0 < f_h < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (37)$$



**Fig. 3.** Simulation model of the hydration of ethylene oxide process in Aspen Hysys.

Let us assume that  $q_h$  represents a realization of a random variable  $Q_h$ . Given  $f_h$ , the conditional probability of  $Q_h$  is

$$p_{Q_h|F_h}(q_h|f_h) = \frac{Q!}{q_h!(Q-q_h)!} \cdot (f_h)^{q_h} \cdot (1-f_h)^{Q-q_h}. \quad (38)$$

The posterior PDF of  $f_h$  is derived through the Bayesian formula, i.e.,

$$p_{F_h|Q_h}(f_h|q_h) = \frac{p_{Q_h|F_h}(q_h|f_h) \cdot p_{F_h}(f_h)}{\int p_{Q_h|F_h}(q_h|f_h) \cdot p_{F_h}(f_h) df_h}. \quad (39)$$

The Bayesian estimate of  $f_h$  is

$$\hat{f}_h = \int f_h \cdot p_{F_h|Q_h}(f_h|q_h) df_h. \quad (40)$$

Thus, the Bayesian estimates of  $f_h, h = 1, 2, \dots, N_H$  are calculated by (36)–(40). When the sample size is  $Q$ , the PDF of  $\theta$  is estimated as

$$\hat{p}_Q(\theta_h) = \hat{f}_h, h = 1, \dots, N_H. \quad (41)$$

Then, a convergent PDF of  $\theta$  and the optimal model parameter  $\hat{\theta}_{opt}$  are determined. The posterior PDF in (41) is growing narrower as the sample size  $Q$  increases. Hypothesis testing is used to judge whether  $\hat{p}_Q(\theta_g)$  is converging or not. A null hypothesis is

$$H_0 : \hat{p}_Q(\theta) = \hat{p}_{Q-1}(\theta). \quad (42)$$

Here  $\hat{p}_{Q-1}(\theta)$  and  $\hat{p}_Q(\theta)$  are the estimated PDFs of  $\theta$  based on samples  $\{\hat{\theta}(q)\}_{q=1}^{Q-1}$  and  $\{\hat{\theta}(q)\}_{q=1}^Q$ , respectively. A chi-square statistic  $\chi_Q^2$  is used to quantify the difference between  $\hat{p}_Q(\theta)$  and  $\hat{p}_{Q-1}(\theta)$ ,

$$\chi_Q^2 = \sum_{d=1}^D \frac{(q_d - (Q-1) \cdot \hat{p}_{Q-1}(\theta_h))^2}{(Q-1) \cdot \hat{p}_{Q-1}(\theta_h)}. \quad (43)$$

When  $\chi_Q^2$  is less than the chi-square critical value subject to the desired confidence level, the null hypothesis is accepted, and  $\hat{p}_Q(\theta)$  estimated based on the samples  $\{\hat{\theta}(q)\}_{q=1}^Q$  is convergent. The optimal model parameters  $\hat{\theta}_{opt} = [\hat{k}_{0,opt}, \hat{E}_{a,opt}]$  are yielded as

$$\hat{\theta}_{opt} = \sum_{d=1}^D \hat{f}_h \theta_h. \quad (44)$$

Finally, the uncertainty of the NOZ model is estimated based on the convergent PDF  $\hat{p}_Q(\theta)$ . Surrogate data of  $\theta$ , denoted as  $\theta_m = [k_{0,m}, E_{a,m}]$  for  $m = 1, \dots, M$ , are randomly generated according to the convergent PDF  $\hat{p}_Q(\theta)$ . Thus,  $M$  sets of normal operating points are calculated from (28),

$$X_{n,m} = \{x[n; \theta_m]\}_{n=1}^{N_{n,m}}, m = 1, \dots, M, \quad (45)$$

and their corresponding the NOZ models are formulated in the same way in (30) as

$$A_{a,m} \bar{x}^T[n; \theta_m] - B_{a,m} = 0, m = 1, \dots, M. \quad (46)$$

The model set consisting of  $A_{a,m}$  and  $B_{a,m}$  for  $m = 1, \dots, M$  represents the uncertainty of the NOZ model.

**Remark 1.** NOZ models are built based on all steady-state values of processes in normal conditions. It is not assumed here that all unsteady-state values fall within NOZs. For instance, input variables in dynamic states may temporarily go outside of NOZs in order to change output variables quickly. However, data samples in dynamic states are not be used in NOZ modeling.

#### 4. Examples

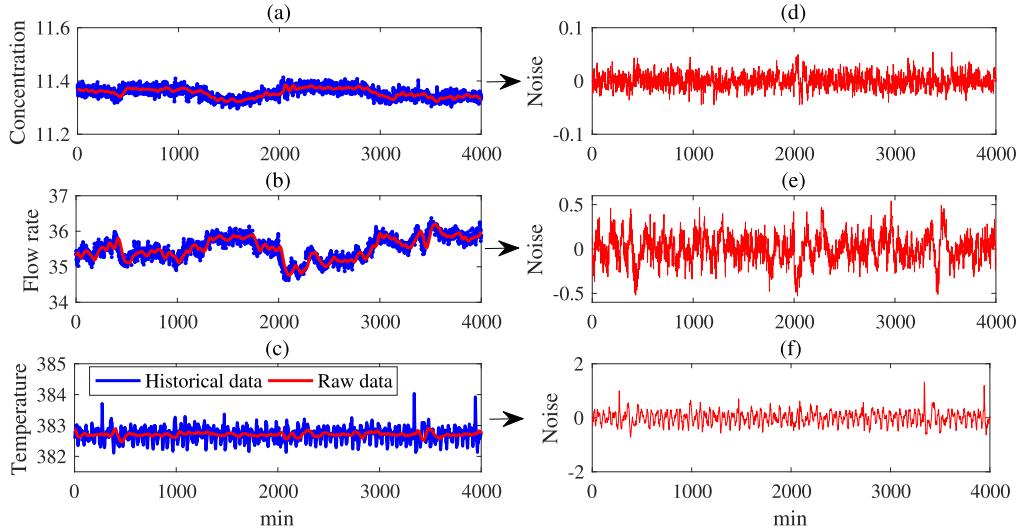
This section establishes a NOZ model for the reaction-separation-recycle process using the hydration of ethylene oxide to ethylene glycol process as an example. Two examples are provided to demonstrate the potential application of the NOZ model for condition monitoring.

##### 4.1. Establishment of normal operating zone models

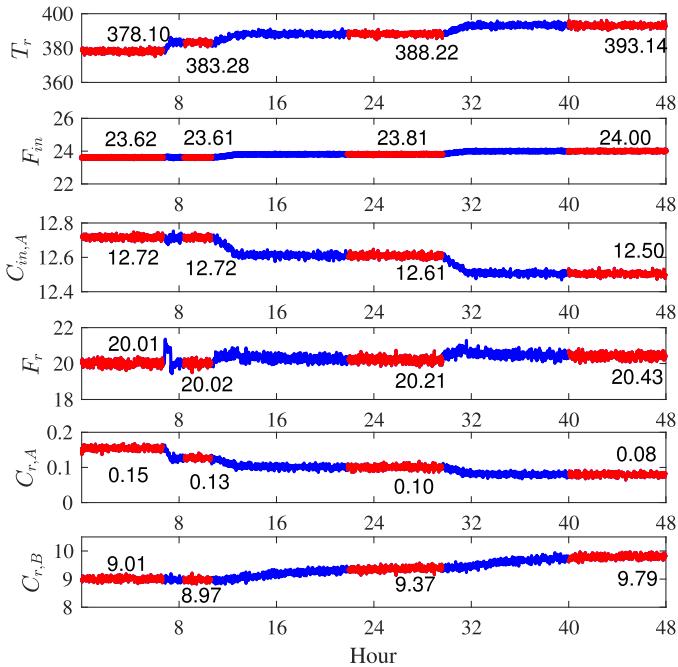
Take the hydration of ethylene oxide to ethylene glycol process as an example, a simulation model is established based on the nonrandom two-liquid model by Aspen Hysys V11 (Hayday, 2019), as depicted in Fig. 3. The ethylene oxide is component A, water is component B, and ethylene glycol is component E in (1). Table 1 lists the reaction kinetics and reaction heat, as well as the equipment parameters of CSTR and DC in Aspen Hysys. In order to achieve dynamic simulation of the process, several controllers are designed, including flow controllers, level controllers, temperature controllers, and pressure controllers. These con-

**Table 1**  
Parameters of the simulation model in Aspen Hysys.

Equipment/controller	Parameter
CSTR	Reaction kinetics: $4123 \cdot \exp\left\{-\frac{50,000}{RT}\right\}$ (kmol m <sup>-3</sup> s <sup>-1</sup> ) Reaction heat: $-9.5e + 4$ (kJ kmol <sup>-1</sup> ) Reactor vessel pressure: 333 (kPa) Volume: 120 (m <sup>3</sup> )
DC	Number of theoretical plates: 10 Feed plate: 5 Plate spacing: 0.55 (m) Diameter: 1.0 (m) Reflux ratio: 3 Top pressure: 200 (kPa) Bottom pressure: 208 (kPa)
FC-1	$K_c = 0.127, T_i = 0.07$
FC-2	$K_c = 0.172, T_i = 0.07$
TC-1	$K_c = 3.04, T_i = 10.7$
TC-2	$K_c = 0.14, T_i = 9.55$
PC-1	$K_c = 7.82, T_i = 0.11$
PG-1	$K_c = 1.72, T_i = 0.33$
LC-1	$K_c = 8.49, T_i = 14.2$
LC-2	$K_c = 2.13, T_i = 204$
LC-3	$K_c = 18.3, T_i = 1.74$



**Fig. 4.** Industrial data of (a) concentration, (b) flow rate, and (c) temperature variables, as well as noise components from (d) concentration, (e) flow rate, and (f) temperature data.



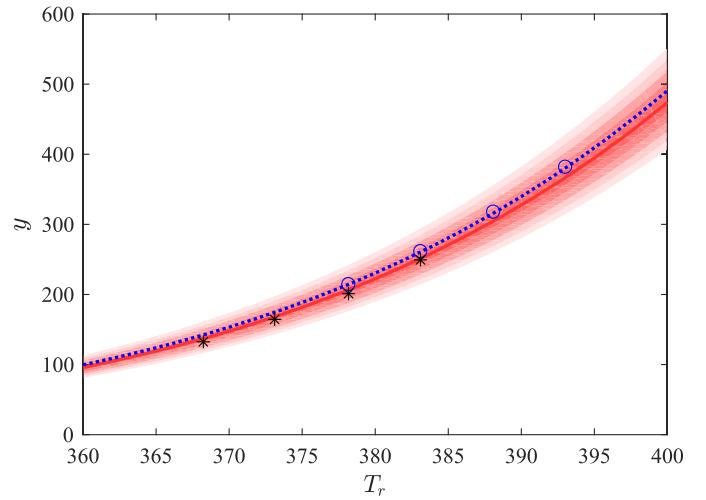
**Fig. 5.** Historical data and steady-state values of process variables.

trollers take the standard proportional-integral controller form as

$$u_c(t) = K_c \left( e_c(t) + \frac{1}{T_i} \int_0^t e_c(t) dt \right).$$

Here  $u_c(t)$  is the controller output and  $e_c(t)$  is the control error. The controller parameters  $K_c$  and  $T_i$  are listed in Table 1.

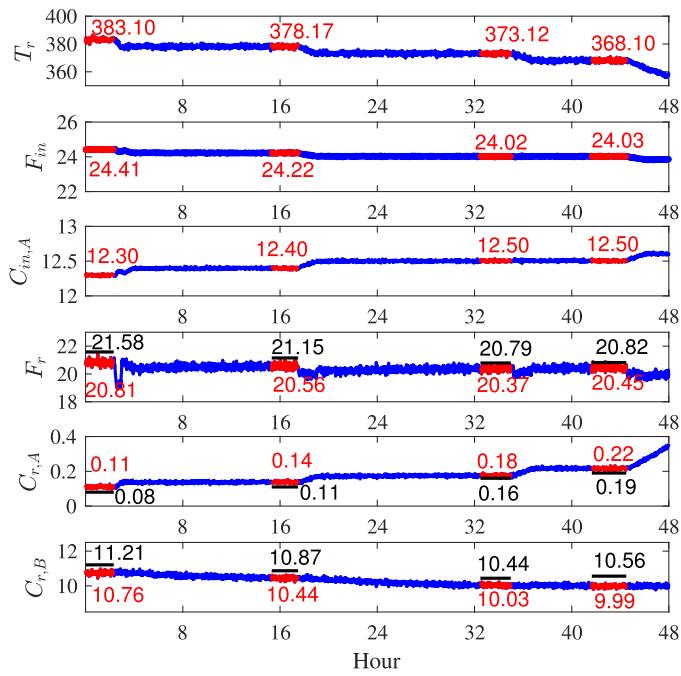
The operating data collected from Aspen Hysys is noise-free. In order to mimic actual chemical processes, noise components of industrial data are introduced as follows. Industrial data of process variables are collected from a reaction-separation-recycle process of a large-scale chemical plant at Qingdao. For instance, three variables of the concentration, flow rate, and temperature are shown in Fig. 4(a)–(c). Using the empirical mode decomposition method (Kopsinis and McLaughlin, 2009), noise components are extracted from industrial data of these process variables, as shown in Fig. 4(d)–(f). These noise components are added



**Fig. 6.** The measured output (blue circle) from historical data, the estimated output (blue dashed line), the new measured output (black star) from testing data, the optimal model (red solid), and its uncertainty range (red shade).

to the corresponding noise-free data from Aspen Hysys to yield historical data with noise.

The physical model is built in (6)–(22) to describe the relationships among process variables  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$ . Some parameters are given as  $\rho_E = 1111 \text{ kg m}^{-3}$ ,  $\rho_A = 882 \text{ kg m}^{-3}$ ,  $\rho_B = 998 \text{ kg m}^{-3}$ ,  $m_E = 62.069 \text{ kg kmol}^{-1}$ ,  $m_A = 44.054 \text{ kg kmol}^{-1}$ ,  $m_B = 18.015 \text{ kg kmol}^{-1}$ ,  $s = \frac{m_A}{\rho_A} + \frac{m_B}{\rho_B} - \frac{m_E}{\rho_E} = 0.0121 \text{ m}^3 \text{ kmol}^{-1}$ ,  $V_r = 100 \text{ m}^3$ , and  $\alpha = 0.1$ . The historical data  $\{F_{in}(t), C_{in,A}(t), T_r(t), F_r(t), C_{r,A}(t), C_{r,B}(t)\}_{t=1}^{48 \times 60}$  for two days with the sampling period 1 min is collected from Aspen Hysys, as shown in Fig. 5. Four groups steady-state values are extracted from the historical data as  $\{F_{in}[m], C_{in,A}[m], T_r[m], F_r[m], C_{r,A}[m], C_{r,B}[m]\}_{m=1}^4$ , as given in Fig. 5. The model parameter  $\hat{\theta} = [\hat{k}_0, \hat{E}_a] = [4023, 52, 155]$  is estimated by (25). Fig. 6 illustrates the fitting results for the measured output  $y[m]$  (blue circle) and the estimated output  $\hat{y}(T_r([m]); \theta)$  (blue dashed line) in (24). In order to verify the accuracy of the estimated parameter  $\hat{\theta}$ , another historical data set is generated as testing data shown in Fig. 7. As depicted in Fig. 6, the new measured output (black star) extracted from the testing data is closed to the estimated output (blue dashed line) with the model parameter  $\hat{\theta}$ . Their corresponding detailed val-



**Fig. 7.** Testing data, steady-state values (red), and the estimated values (black) based on the model parameter  $\hat{\theta}$  of process variables.

**Table 2**  
The true and estimated model parameters.

$\theta$	$\hat{\theta}$	$\hat{\theta}_{opt}$
$k_0 = 4123$	$\hat{k}_0 = 4023$	$\hat{k}_{0,opt} = 4077$
$E_a = 50,000$	$\hat{E}_a = 52,155$	$\hat{E}_{a,opt} = 50,750$

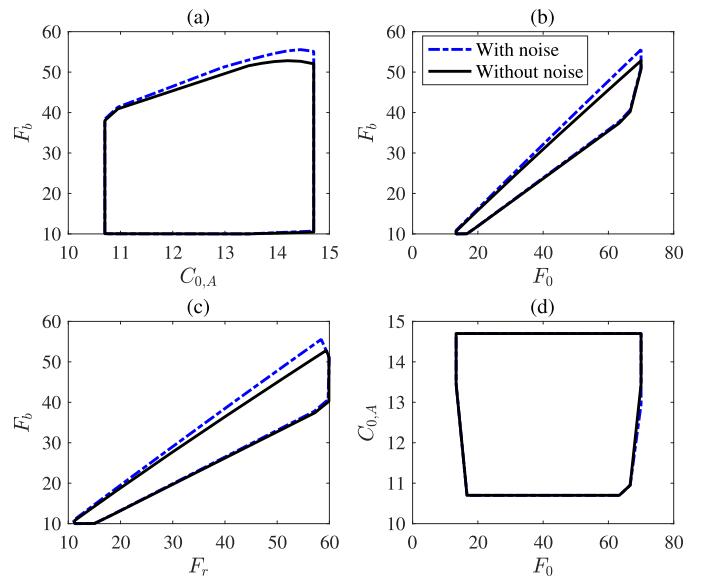
ues are presented in Fig. 7. Therefore, the established physical model is accurate.

The NOZ model is developed for the ethylene glycol production process. According to the process knowledge, the minimums and maximums in (2)–(5) and (26) of the variation ranges for six process variables are given as follows:

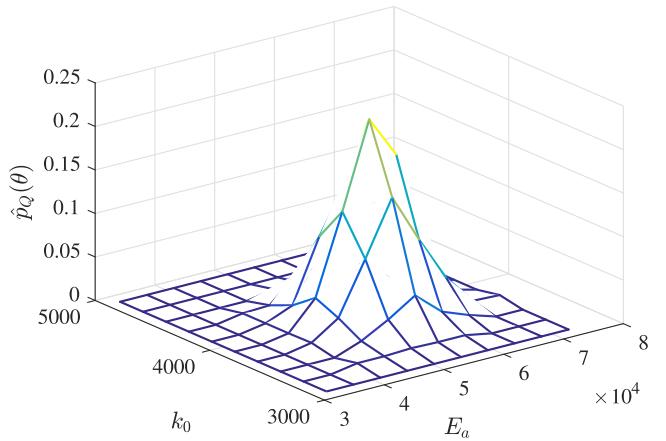
$$\begin{aligned}
F_{0,min} &= 10 \text{ m}^3 \text{ h}^{-1}, F_{0,max} = 100 \text{ m}^3 \text{ h}^{-1}; \\
C_{0,A,min} &= 10.7 \text{ kmol m}^{-3}, C_{0,A,max} = 14.7 \text{ kmol m}^{-3}; \\
T_{r,min} &= 260 \text{ K}, T_{r,max} = 400 \text{ K}; \\
F_{r,min} &= 10 \text{ m}^3 \text{ h}^{-1}, F_{r,max} = 60 \text{ m}^3 \text{ h}^{-1}; \\
F_{b,min} &= 10 \text{ m}^3 \text{ h}^{-1}, F_{b,max} = 80 \text{ m}^3 \text{ h}^{-1}; \\
C_{b,E,min} &= 17.59 \text{ kmol m}^{-3}, C_{b,E,max} = 17.89 \text{ kmol m}^{-3}.
\end{aligned} \tag{47}$$

Given the step sizes  $\Delta F_0 = 3 \text{ m}^3 \text{ h}^{-1}$ ,  $\Delta C_{0,A} = 0.2 \text{ kmol m}^{-3}$ , and  $\Delta T_r = 5 \text{ K}$  in (26), a model-input set  $\{F_0[n], C_{0,A}[n], T_r[n]\}_{n=1}^{17400}$  is obtained. All operating points in (27) are collected through the physical model with parameter  $\hat{\theta}$ . The set  $X_n$  of all normal operating points is acquired by (28). The NOZ model in (30) is described through a convex hull. Fig. 8 displays the projections of the NOZ on 2-D planes, because the NOZ model in 6-D space cannot be visualized in the Cartesian coordinate. As a comparison, a NOZ model without noise is built based on the correct kinetic parameters in Table 1, as shown in Fig. 8 by the solid black line. Obviously, noise leads to minor differences in some boundaries of the two NOZ models.

The uncertainty of the NOZ model is measured based on Bayesian theory. Multiple model parameters are estimated from multiple group modeling data. As the sample size of the estimated model parameters increases, the PDF  $\hat{p}_0(\theta)$  of  $\theta$  is determined by (36)–(43), as shown in Fig. 9. The optimal parameter  $\hat{\theta}_{opt} = [\hat{k}_{0,opt}, \hat{E}_{a,opt}] = [4077, 50, 750]$  in (44) is obtained via the Bayesian theory based on the PDF  $\hat{p}_0(\theta)$ . Table 2



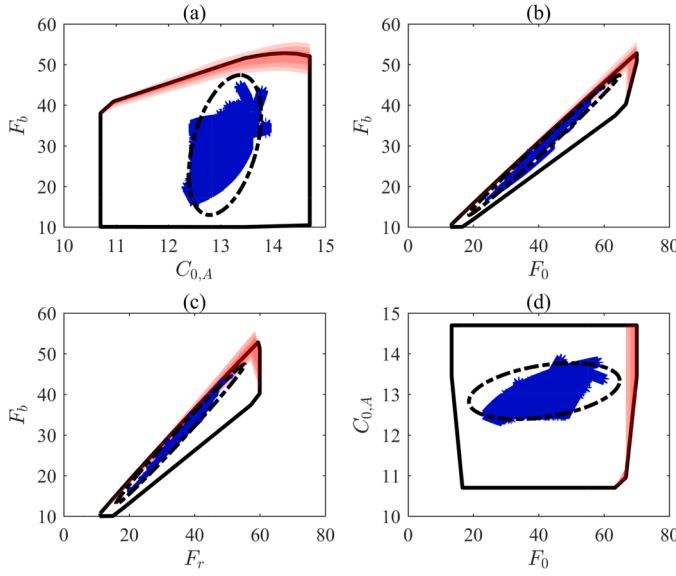
**Fig. 8.** Projections of the NOZ model built based on historical data (blue dashed) and the NOZ model without noise (black solid) on 2-D planes.



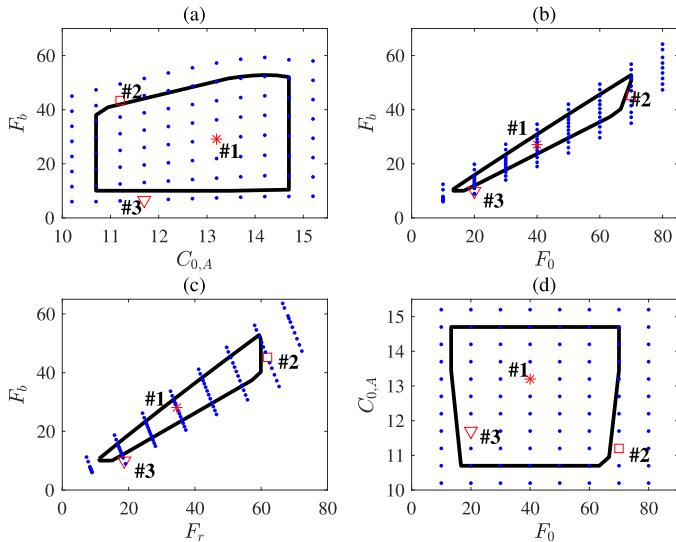
**Fig. 9.** The PDF  $\hat{p}_Q(\theta)$  of the model parameter  $\theta$ .

compares the true parameter  $\theta = [k_0, E_a] = [4123; 50,000]$ , the optimal parameter  $\hat{\theta}_{opt}$  obtained using the Bayesian theory, and the estimated parameter  $\hat{\theta} = [\hat{k}_0, \hat{E}_a] = [4023, 52, 155]$  in (25) based on one set of modeling data. Clearly, the optimal parameter  $\hat{\theta}_{opt}$  based on Bayesian theory is closer to the true value  $\theta$ . The absolute percentage errors for parameters  $k_0$  and  $E_a$  are 1.1% and 1.5%, respectively. Fig. 6 presents the optimal parameter (red solid) and its uncertainty range (red shade) for the physical model. Surrogate data of the model parameter  $\theta$  are randomly generated based on  $\hat{p}_Q(\theta)$ . The corresponding NOZ models are described by (46), as illustrated by the red shaded area in Fig. 10, depicting the uncertainty of the NOZ model. The red shaded area denoting the uncertainty of the NOZ model is tight. Thus, the PDF  $\hat{p}_Q(\theta)$  of the model parameter  $\theta$  is convergent.

The accuracy of the established NOZ model is validated as follows. Based on the Aspen Hysys simulation, 88 operating points inside and outside the NOZ model are validated, as shown in Fig. 11. Due to limited space, three operating points are provided for illustration. Figs. 12–14 present the historical data for operating points #1, #2, and #3, respectively. Among them, the data sample of all process variables in Fig. 12 is within the high and low alarm thresholds, so that operating point #1 is in a normal condition. As shown in Fig. 13, the make-up feed flow rate  $F_0$  increases and product concentration  $C_{hF}$  decreases, which causes the



**Fig. 10.** Projections of the NOZ model (black solid), model uncertainty (red shade), hyperellipsoid (black dashed) and historical data (blue dot) on 2-D planes.



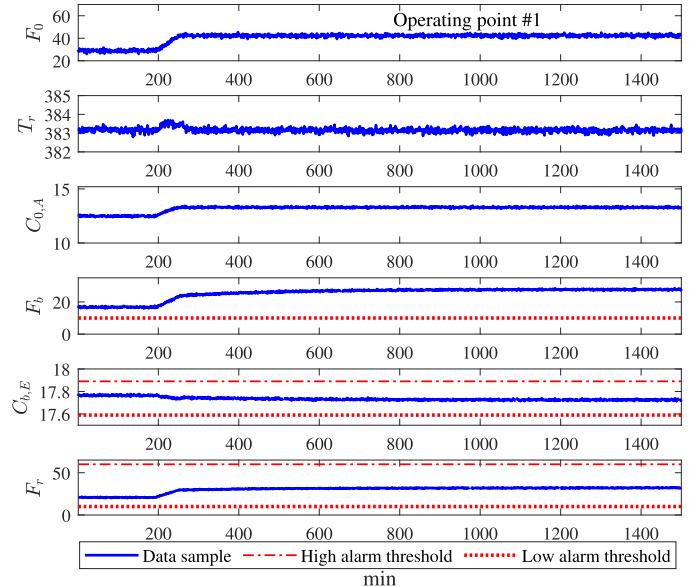
**Fig. 11.** NOZ models, 88 operating points (blue dot) for validation, and three special operating points #1, #2, and #3.

feed flow rate  $F_r$  to the DC to increase and exceed the high alarm threshold. Thus, operating point #2 is in an abnormal condition. The outlet valve at the DC bottom is clogged when the make-up feed flow rate  $F_0$  decreases for operating point #3 in Fig. 14, causing the product flow rate  $F_b$  to decrease and fall below its low alarm threshold. As a result, operating point #3 is also in an abnormal condition. These results are the same as the actual operation conditions. Therefore, the established NOZ model is accurate.

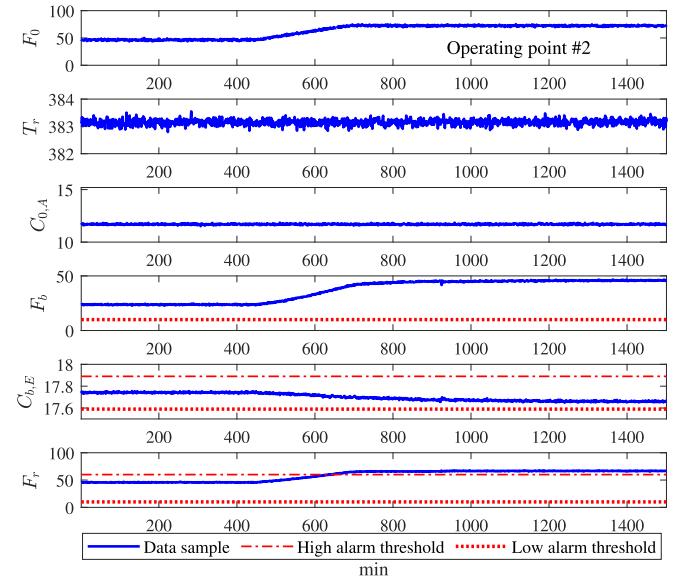
#### 4.2. Potential application of normal operating zone models

The subsection presents two examples to demonstrate the potential application of NOZ models for condition monitoring in the reaction-separation-recycle process of ethylene oxide hydration.

In the first example, the NOZ model and the PCA method as a representative of data-based methods are applied to monitor the operation condition of the process. Time series of the process variables  $F_0$ ,  $C_{0,A}$ ,



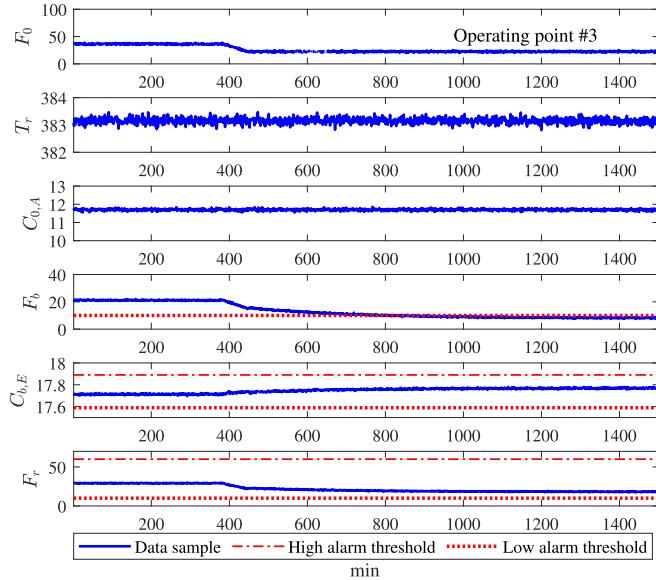
**Fig. 12.** Historical data of process variables for operating point #1.



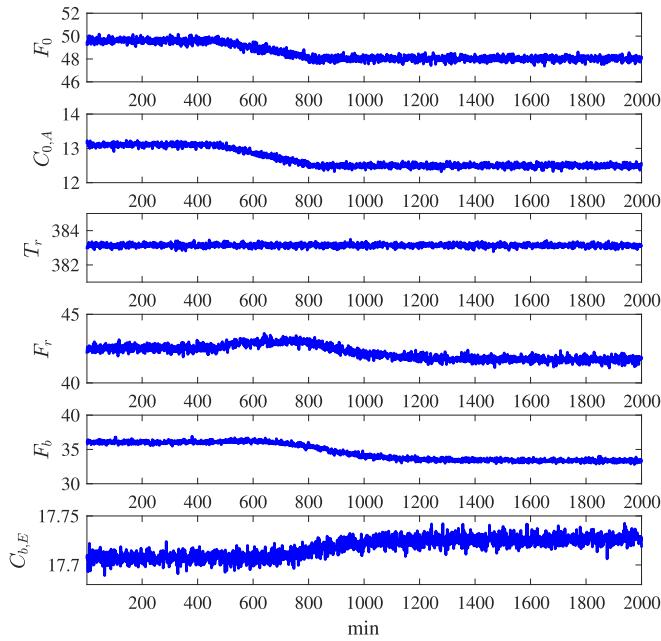
**Fig. 13.** Historical data of process variables for operating point #2.

$F_r$ ,  $F_b$ ,  $F_r$ , and  $C_{b,E}$  are shown in Fig. 15. The input variables  $F_0$  and  $C_{0,A}$  began to change at  $t = 462$  min, decreasing from  $49.6 \text{ m}^3 \text{ h}^{-1}$  to  $48.0 \text{ m}^3 \text{ h}^{-1}$  and  $13.1 \text{ kmol m}^{-3}$  to  $12.5 \text{ kmol m}^{-3}$ . The corresponding output variables  $F_b$  and  $F_r$  decrease, while  $C_{b,E}$  increases. No abnormal conditions appear.

The NOZ model is used for condition monitoring of data samples in Fig. 15. Fig. 16 depicts the projections of the data samples and the NOZ model. The data points are all located inside the NOZ model. Thus, the process is in normal conditions. For comparison, the PCA method is applied to the same data in Fig. 15. The abnormality detection logic of the PCA method is equivalent to judging whether a new data point locates inside a hyperellipsoid corresponding to the normal condition of a multivariate process. The hyperellipsoid is established through the historical data for building the NOZ. The projections of the hyperellipsoid on 2-D planes are presented in Fig. 17. As shown in Fig. 17(c), the data points are located outside the hyperellipsoid after  $t = 774$  min. As a result, the PCA method produces false detection of normal conditions. There are two reasons that may cause false detection results. The first



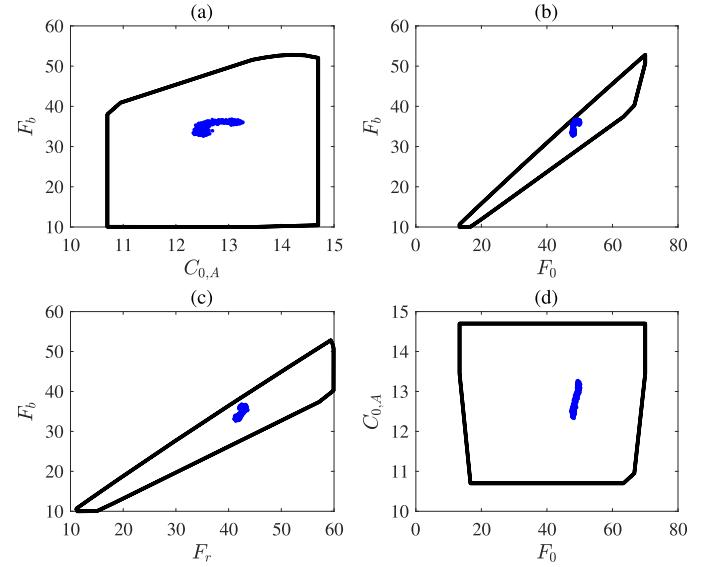
**Fig. 14.** Historical data of process variables for operating point #3.



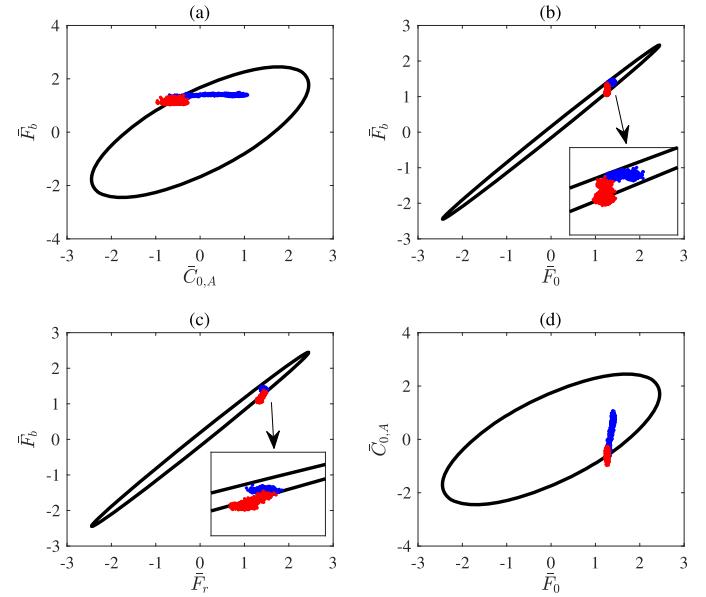
**Fig. 15.** Time series of  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$  in normal conditions.

one is because the PCA method is data-based and does not have extrapolation properties. As shown in Fig. 10, the historical data (blue dot) cannot cover all normal conditions. Those blue dots only occupy a part of the NOZ formed by all normal conditions. Therefore, the PCA method can not identify normal conditions that do not appear in historical data. The second reason is that the PCA method requires the assumption that data points follow the Gaussian distribution. In the case of non-Gaussian distributed data points, the hyperellipsoid does not accurately describe all normal conditions, which yields false detection results. As shown in Fig. 10, there are some discrepancies between the blue dots and the black dash ellipsoid from the PCA method.

In the second example, the NOZ model and the model-based method are employed for condition monitoring of the ethylene glycol production process. Time series of the process variables  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$  as illustrated in Fig. 18. Before  $t = 560$  min, the process is in normal conditions with  $F_0 = 28.81 \text{ m}^3 \text{ h}^{-1}$ ,  $C_{0,A} = 13.30 \text{ m}^3 \text{ h}^{-1}$ ,  $T_r = 383 \text{ K}$ ,  $F_r =$



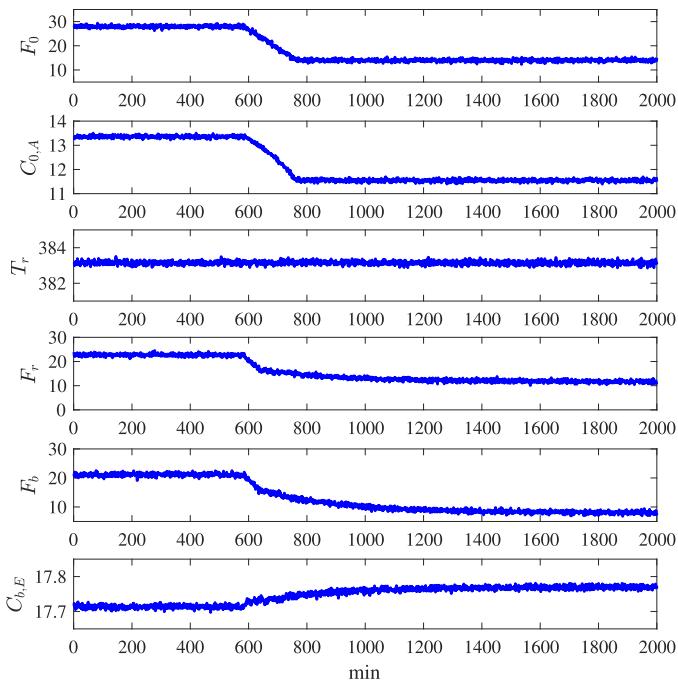
**Fig. 16.** Application result of the NOZ model for normal conditions.



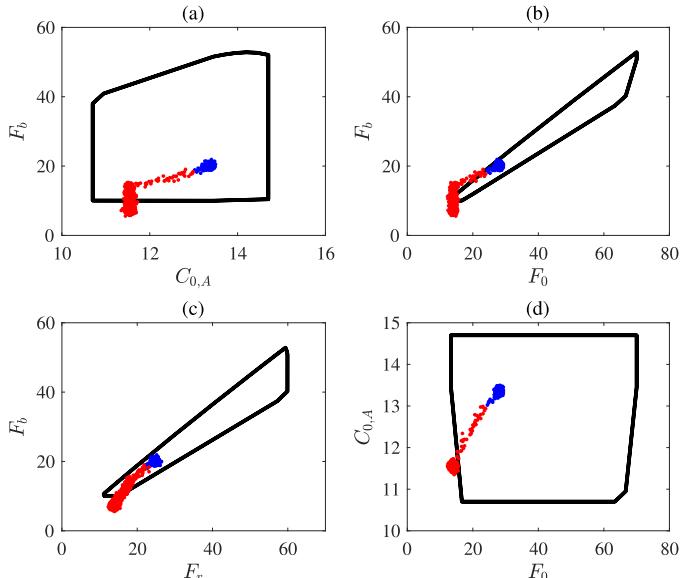
**Fig. 17.** Application result of the PCA method for normal conditions (blue dot: 1–773 min; red dot: 774–2000 min).

$23.02 \text{ m}^3 \text{ h}^{-1}$ ,  $F_b = 21.04 \text{ m}^3 \text{ h}^{-1}$ , and  $C_{b,E} = 17.74 \text{ m}^3 \text{ h}^{-1}$ .  $F_0$  and  $C_{0,A}$  begin to decrease at  $t = 560$  min. Meanwhile, the feed valve of the ethylene oxide is clogged, which reduces the feed flow rate  $F_0$  and the concentration  $C_{0,A}$ , ultimately resulting in a product flow rate  $F_b$  that is too low below its minimum value  $F_{b,min} = 10 \text{ m}^3 \text{ h}^{-1}$  in (3). Thus, the process is in abnormal conditions.

The NOZ model is utilized to monitor the data samples in Fig. 18. As depicted in Fig. 19(b), the data points are outside the NOZ model caused by the  $F_0$  decrease at  $t = 601$  min. After  $t = 1060$  min, the data points in Fig. 19(a), (c), and (d) are outside the NOZ model caused by the  $F_b$  decrease, so that the process is in abnormal conditions. As a contrast, the model-based method is employed to detect the same data samples shown in Fig. 18. Fig. 20(a) and (c) present data samples and model-based predictions for the output variables  $F_r$  and  $F_b$ . Their corresponding absolute residuals are given in Fig. 20(b) and (d), respectively. The process is in normal condition based on the given residual threshold. As a result, the model-based method cannot detect the above abnormal conditions. The



**Fig. 18.** Time series of  $F_0$ ,  $C_{0,A}$ ,  $T_r$ ,  $F_r$ ,  $F_b$ , and  $C_{b,E}$  in abnormal conditions.

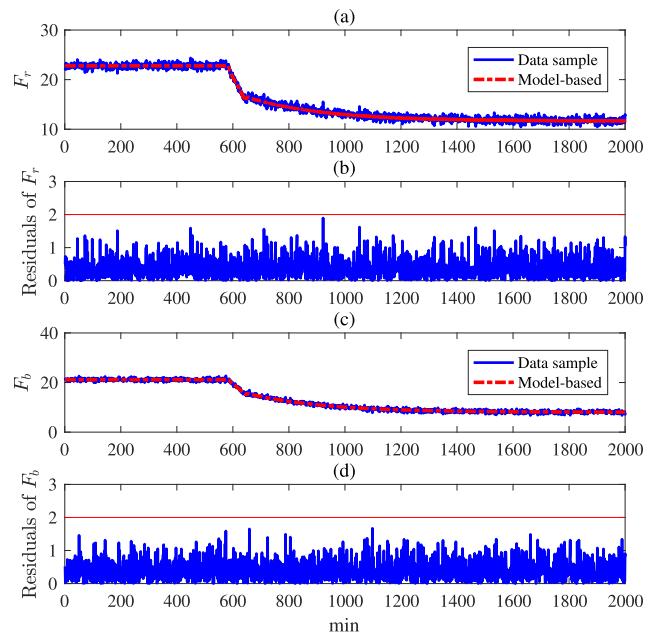


**Fig. 19.** Application result of the NOZ model for abnormal conditions (blue dot: 1–600 min; red dot: 601–2000 min).

NOZ model is able to detect abnormalities that do not change physical models compared to the model-based method.

## 5. Conclusion

This paper built a NOZ model for condition monitoring in reaction-separation-recycle processes. The NOZ model was developed based on physical principles, historical data and operational requirements. Bayesian estimation theory was utilized to address a main technical issue for measuring the uncertainties of the NOZ model. By representing an allowable region of related process variables, the established NOZ model had good potential for condition monitoring of the reaction-separation-recycle processes, compensating certain limitations of existing methods. In particular, the NOZ model explained normal conditions that are not



**Fig. 20.** Application result of the model-based method for abnormal conditions.

present in the historical data compared to data-based methods. In contrast to model-based methods, the NOZ model identified abnormal conditions that do not change physical models.

The NOZ model in this paper has potential applications for condition monitoring of reaction-separation-recycle processes. However, there are quite a few problems to be resolved in the future work. For instance, the uncertainty in the NOZ caused by noise generates false alarms and missed alarms for condition monitoring. Multivariate alarm systems need to be carefully designed to achieve an optimal balance between the false alarms and missed alarms. Another future work is to validate the NOZ models in industrial applications. Some promising results have been obtained for an air separation device, a syngas compressor and a steam generation boiler. It would take more efforts to build the NOZ models for a number of devices and equipments in a large-scale chemical process.

## CRediT authorship contribution statement

**Mengyao Wei:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation; **Jiandong Wang:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization; **Zhishan Zhang:** Writing – review & editing, Visualization, Validation, Investigation.

## Data availability

Data will be made available on request.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jiandong Wang reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Ali, H., Safdar, R., Zhou, Y., Yao, Y., Yao, L., Zhang, Z., Rasool, M.H., Gao, F., 2024. Robust statistical industrial fault monitoring: a machine learning-based distributed cca and low frequency control charts. *Chem. Eng. Sci.* 299, 120460.
- Altiokka, M.R., Akyalcin, S., 2009. Kinetics of the hydration of ethylene oxide in the presence of heterogeneous catalyst. *Ind. Eng. Chem. Res.* 48 (24), 10840–10844.
- Aryanezhad, M.B., Hemati, M., 2008. A new genetic algorithm for solving nonconvex nonlinear programming problems. *Appl. Math. Comput.* 199 (1), 186–194.
- Barber, C.B., Dobkin, D.P., Huhdanpaa, H., 1996. The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 22 (4), 469–483.
- Barton, J.L., 2020. Electrification of the chemical industry. *Science* 368 (6496), 1181–1182.
- Bhagwat, A., Srinivasan, R., Krishnaswamy, P.R., 2003. Fault detection during process transitions: a model-based approach. *Chem. Eng. Sci.* 58 (2), 309–325.
- Botre, C., Mansouri, M., Karim, M.N., Nounou, H., Nounou, M., 2017. Multiscale pls-based girt for fault detection of chemical processes. *J. Loss Prev. Process. Ind.* 46, 143–153.
- Du, P., Wilhite, B., Kravaris, C., 2024. Model-based fault diagnosis for safety-critical chemical reactors: an experimental study. *AIChE J.* 70 (12), 18565.
- Haydary, J., 2019. Chemical Process Design and Simulation: Aspen Plus and Aspen Hysys Applications. John Wiley & Sons.
- Kopsinis, Y., McLaughlin, S., 2009. Development of emd-based denoising methods inspired by wavelet thresholding. *IEEE Trans. Signal Process.* 57 (4), 1351–1362.
- Kumar, V., Kaistha, N., 2018. Real time optimization of a reactor-separator-recycle process i: steady state modeling. *Ind. Eng. Chem. Res.* 57 (37), 12429–12443.
- Madakyaru, M., Harrou, F., Sun, Y., 2019. Monitoring distillation column systems using improved nonlinear partial least squares-based strategies. *IEEE Sens. J.* 19 (23), 11697–11705.
- Mansouri, M., Sheriff, M.Z., Baklouti, R., Nounou, M., Nounou, H., Hamida, A.B., Karim, N., 2016. Statistical fault detection of chemical process-comparative studies. *J. Chem. Eng. Process Technol.* 7 (1), 282–291.
- Nawaz, M., Maulud, A.S., Zabiri, H., Suleman, H., Tufa, L.D., 2020. Multiscale framework for real-time process monitoring of nonlinear chemical process systems. *Ind. Eng. Chem. Res.* 59 (41), 18595–18606.
- Pattison, R.C., Tsay, C., Baldea, M., 2017. Pseudo-transient models for multiscale, multiresolution simulation and optimization of intensified reaction/separation/recycle processes: framework and a dimethyl ether production case study. *Comput. Chem. Eng.* 105, 161–172.
- Roostaei, M., Eslamloueyan, R., 2024. Design and optimization of reaction-separation-recycle systems using a pseudo-transient continuation model. *Chem. Eng. Res. Des.* 211, 235–252.
- Shang, C., Yang, F., Gao, X., Huang, X., Suykens, J.A., Huang, D., 2015. Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis. *AIChE J.* 61 (11), 3666–3682.
- Taqvi, S.A., Tufa, L.D., Zabiri, H., Maulud, A.S., Uddin, F., 2020. Fault detection in distillation column using narx neural network. *Neural Comput. Appl.* 32, 3503–3519.
- Villagomez, E.L., Mahyar, H., Mahalec, V., 2024. Fault detection and analysis via latent space differences between the plant and the model representing normal operation. *Comput. Aided Chem. Eng.* 53, 2983–2988.
- Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., 1993. Probability and Statistics for Engineers and Scientists. Macmillan, New York.
- Wang, J., Wei, M., Xing, X., 2022. Static gain estimation for nonlinear dynamic systems from steady-state values hidden in historical data. *ISA Trans.* 120, 78–88.
- Xu, Y., Jia, M., Mao, Z., 2022. A novel auto-regressive dynamic slow feature analysis method for dynamic chemical process monitoring. *Chem. Eng. Sci.* 248, 117236.
- You, L., Chen, J., 2021. A variable relevant multi-local pca modeling scheme to monitor a nonlinear chemical process. *Chem. Eng. Sci.* 246, 116851.
- Yu, J., Liu, X., Ye, L., 2020. Convolutional long short-term memory autoencoder-based feature learning for fault detection in industrial processes. *IEEE Trans. Instrum. Meas.* 70, 1–15.
- Yu, Y., Wang, J., 2019. Alarm monitoring for multivariate processes based on a convex-hull normal operating zone. *IEEE Trans. Control Syst. Technol.* 28 (6), 2649–2656.
- Zhong, K., Ma, D., Han, M., 2020. Distributed dynamic process monitoring based on dynamic slow feature analysis with minimal redundancy maximal relevance. *Control Eng. Pract.* 104, 104627.