

Article

# Correlation Degree and Clustering Analysis-Based Alarm Threshold Optimization

Guixin Zhang and Zhenlei Wang \*

Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology (ECUST), Shanghai 200237, China; hw910706@163.com

\* Correspondence: wangzhen\_l@ecust.edu.cn

**Abstract:** In industrial practice, excessive alarms and high alarm rates are mostly generated from unreasonable settings to variable alarm thresholds, which have become the significant causes of impact on operation stability and plant safety. A correlation degree and clustering analysis-based approach was presented to optimize the variable alarm thresholds in this paper. The correlation degrees of variables are first obtained by analyzing correlation relationships among them. Second, the variables are grouped according to the gray correlation coefficients and clustering analysis, given the weight for fault alarm rate (FAR) in each group. An objective function about the FAR, missed alarm rate (MAR), and the maximum acceptable FAR and MAR is then established with variable weight. Eventually, based on an optimization algorithm, the objective function can be optimized for obtaining the optimal alarm threshold. Cases study of the Tennessee Eastman (TE) industrial simulation process and an actual industrial ethylene production process, in comparison to the initial situation, show that the method can effectively reduce FAR according to correlation degrees among variables in the system, and decrease the number of alarms with reduction rates of 40.5% and 35.3%, respectively.

**Keywords:** alarm threshold; correlation degree; clustering analysis; FAR; MAR



**Citation:** Zhang, G.; Wang, Z.

Correlation Degree and Clustering Analysis-Based Alarm Threshold Optimization. *Processes* **2022**, *10*, 224. <https://doi.org/10.3390/pr10020224>

Academic Editor: Krzysztof Talaśka

Received: 27 November 2021

Accepted: 11 January 2022

Published: 25 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the complexity and refinement of the process of industrial production, the production process is more and more inseparable from real-time monitoring of the system. An alarm management system, as an indispensable part in the safety operation of industrial production, has been paid more and more attention by all walks of life. In industrial practice, cases with more false alarms, a higher false alarm rate (FAR) and a missed alarm rate (MAR) always arise in processes [1], which are mainly caused through the unreasonable threshold settings for variables and ineffective management for alarm systems. Based on the studies from EEMUA, the range of alarm numbers that an operator could effectively handle for one alarm is from every 5 min to 10 min [2].

Regarding the methods of alarm optimization, academia has given many methods, each of which plays a certain role in its corresponding system to a greater or lesser extent. There are many kinds of alarm optimization methods, and classification methods are inconclusive. Generally speaking, they can be divided into univariate methods and multivariate methods, threshold optimization methods and algorithm optimization methods, off-line methods and dynamic methods, etc.

For determining the process variable threshold, academia has studied many optimization methods. For instance, in terms of FAR and MAR, an approach for estimating the threshold was proposed on account of an adaptive fuzzy-neural network and genetic learning algorithm [3]. As the threshold can be determined by the deadband, a method with the objective function about FAR and MAR, and the relation between the optimal threshold and deadband to estimate the threshold, was presented [4]. Combining FAR and

MAR with a correlation coefficient, an off-line method for optimizing thresholds was given to reduce alarms for multi-variables based on time delay [5]. For improving the robustness classification performance in the system with regard to separation threshold selection, different intelligent pattern classifiers were used to mine industrial batch dryer data to determine thresholds [6]. In addition, there also have many methods in setting thresholds in early warning and damage systems [7,8], alarms reduction [9], systems monitoring [10,11], and performance optimization [6,12].

Over the past five years, based on intelligent algorithms, some similar methods have been improved. For instance, to remove chattering alarms, a univariate method was presented for addressing the reduction of alarms with median filters [13]. Taking the missed alarms and false alarms into account, an off-line univariate approach for determining alarm threshold in debris flow forecasting was presented, with the lowest missed-alarm and false-alarm probabilities [14]. By optimizing positioning accuracy, the pulse-width multiplexing  $\Phi$ -OTDR and multisensor information fusion algorithm were utilized to reduce the nuisance alarm rate [15]. For target tracking in a chaotic environment, a multivariate approach was proposed to optimize the joint threshold and power allocation strategy with a two-variable nonconvex optimization problem for the cognitive radar network, containing the detection stage and transmitting stage [16]. Based on the test observations, an off-line simple and robust approach was proposed to determine the detection thresholds for detecting defluidization in the early stage [17]. Some approaches about optimal alarm identification [18], design and evaluation analysis for an alarm system [19–23], management framework [24], alarm threshold [25], and an overview of industrial alarm systems [26] also have appeared. Variables in most of the above approaches have not been clustered with optimized thresholds, which could be suitable for analyzing interlinks among similar variables. Therefore, Zhang et al. presented an off-line multivariate method based on ROC curve and sensitivity, considering the sensitivity relationship and clustering analysis among variables, to optimize the alarm threshold [27]. A multivariate alarm clustering method was proposed that takes advantage of the information contained in the alarm logs themselves, of which the clustering analysis for process alarms was achieved through word embedding [28]. Analyzing alarm data, Lucke et al. presented an on-line method that conducted a practical application for alarm flood classification based on a set of historical alarm floods [29]. In the process, for high-dimension variables, the number of alarms needing addressing increases significantly when the number of measurable variables increases. False alarms caused by redundant disturbances will disturb operators, leading to alarms having more significance on the system being missed as a consequence. Thus, clustering variables into groups is necessary for alarm optimization.

Most of the above alarm optimization methods are based on the off-line system optimization, and the results obtained in the corresponding systems are also obvious, large or small, effectively optimizing the production process and reducing losses. Thus, to promptly detect the chattering alarms and effectively reduce the number of chattering alarms, an on-line method was given to detect alarms in a timely manner [30]. As for the HVAC systems, Chakraborty et al. put forward a novel dynamic threshold method with a data-driven model using extreme gradient boosting (XGBoost), which mainly utilized early fault detection [31]. The static and dynamical performance analyses were used to update evidence in designing the industrial alarm system to reduce unnecessary alarms [32]. There are also some corresponding approaches for optimization, such as alarm management strategy [33], alarming mechanism [34], and threshold setting [35], which have promoted the development of dynamic methods to a certain extent.

However, as the current industrial production processes change irregularly, the preceding production process and the following process cannot be consistent all the time, such as the changes caused by different conditions or an abnormal process. In view of the problem, a new alarm threshold optimization method is proposed, which uses the correlation degrees among the variables and clustering analysis. Herein, this paper mainly has four significant contributions. (1) It considers the gray correlation

degree analysis. Variables with similar influence on the system can be found out through correlation analysis. (2) It could carry on the group sorting according to the intrinsic clustering analysis. (3) It reduces FAR and also has a significant inhibitory effect on MAR (significantly reduced invalid alarms). (4) It can be used as a reference for real-time online optimization. When connecting the current programs to the computer interface in on-line systems, it could meet the requirements of the fast-changing production processes through setting an update period and data, which would consider the alarm rates. In addition, this method could help operators reduce operation load, make more efficient repair measures in a timely manner, and reduce the losses.

## 2. Optimum Design Outline

### 2.1. Alarm Efficiency Index

At present, an alarm system is important for safety, which generally utilizes FAR and MAR as efficiency indices to measure the accuracy of detecting operation conditions [36]. Based on the operation conditions, industrial processes usually contain normal and abnormal situations, which generally use the FAR and MAR to represent the probability directly for a variable when its measured values go beyond the threshold in normal operations, and within the threshold in abnormal operations in an alarm system [37].

The FAR and MAR can be obtained as follows:

Initially, for a variable  $x$ , within a period of time, two groups of data under normal and abnormal situations are obtained. Where a group of data are collected as the normal data when the process runs normally and steadily, another set of data are collected as the abnormal data when the process deviates from normal operation state obviously, which contains added disturbance or failure.

Later, for a variable  $x$ , the probability density functions  $f(x)$  and  $g(x)$  under the two situations, respectively, are obtained by fitting the corresponding data, which can be shown in Figure 1, where  $x_T$  denotes as the alarm threshold. For a certain parameter of the system,  $x_T$  indicates that the parameter has a well running state under the current threshold. When it exceeds or falls below the current threshold, the process may generate redundant false alarms or missed alarms. Here, false alarms will be activated when normal process variable values (the blue line) falls below  $x_T$ , and missed alarms will be activated when the abnormal process variable values (red dotted line) exceed  $x_T$ .

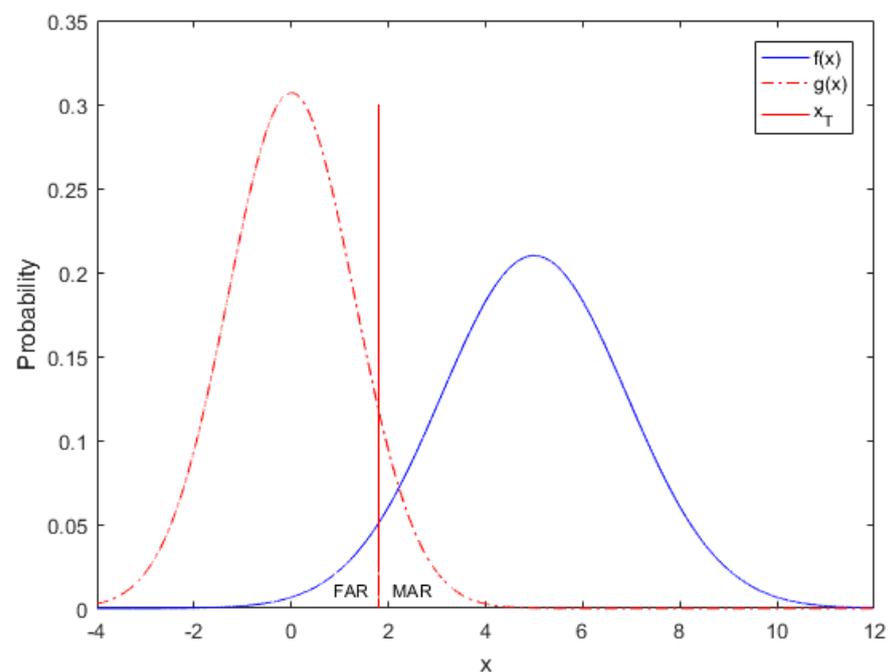


Figure 1. Process probability density curves of  $x$ .

Finally, given the  $x_T$ , based on Figure 1 and the Equations (1) and (2) [5,36], the FAR and MAR can be obtained.

$$FAR = \int_{-\infty}^{x_T} f(x)dx \quad (1)$$

$$MAR = \int_{x_T}^{+\infty} g(x)dx \quad (2)$$

The following work in this paper can be conducted when the functions ( $f(x)$  and  $g(x)$ ) for a variable can be fitted which was irrelevant to the distribution.

## 2.2. Alarm Clustering Analysis

Based on the alarm clustering algorithm, variables can be clustered into groups.

### 1. Correlation degree analysis

A measure of the degree of correlation between two factors in a system that varies from time to time or from object to object is called the correlation degree [38]. In a system process, if the trend of change of the two factors is consistent, that is, the degree of synchronous change is high, then the degree of correlation is high. Conversely, it is lower. Thus, the gray correlation analysis method is a method to measure the correlation degree among factors according to the degree of similarity or difference of development trend among factors, that is, "gray correlation degree".

Specific calculation steps for correlation analysis:

- (1) Determine the reference sequence and comparison sequence. The data sequence that reflects the behavior characteristics of a system is called a reference sequence and the data sequence composed of factors that affect the behavior of a system is called a comparison sequence;
- (2) Conduct dimensionless treatment for the reference sequence and comparison sequence.

Due to the different physical meanings of each factor in the system, the dimensionality of the data may not be the same, which is not convenient for comparison, or it is difficult to get the correct conclusion when comparing. Therefore, in the analysis of gray relational degree, dimensionless data processing should be generally required.

- (3) Determine the reference sequence and comparison sequence of the gray correlation coefficient  $\xi_i(\mathbf{X}_i)$ .

The correlation degree is essentially the difference in geometry among curves. So, the difference among curves can be used as a measure of the correlation degree. For a reference sequence  $\mathbf{X}_0 = \{x_0(1), x_0(2), \dots, x_0(n)\}$ , there are several comparison sequences  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ , the correlation coefficient  $\xi_i(k)$  of each reference sequence and comparison sequence each time is deduced by the following formula:

$$\xi_i(k) = \frac{\min_j \min_i |x_0(l) - x_j(l)| + P \max_j \max_i |x_0(l) - x_j(l)|}{|x_0(k) - x_i(k)| + P \max_j \max_i |x_0(l) - x_j(l)|} \quad (3)$$

where,  $P$  is the distinguish coefficient, the value range of which is generally between 0–1, with 0.5 as the common value;  $x_0(k) - x_i(k)$  represents the absolute difference between the sequences  $\mathbf{X}_i$  and  $\mathbf{X}_0$  at point  $k$ ;  $l = 1, 2, \dots, n$ ,  $\min_i |x_0(l) - x_j(l)|$  is the minimum difference of the first level, which represents the minimum difference between sequences  $\mathbf{X}_j(l)$  and  $\mathbf{X}_0(l)$  at each point;  $\min_j \min_i |x_0(l) - x_j(l)|$  is the minimum difference of the second level, which represents the minimum difference in all sequences based on the minimum difference found in each sequence;  $\max_i |x_0(l) - x_j(l)|$  is the maximum difference of the first level, which represents the maximum difference between sequences  $\mathbf{X}_j(l)$  and  $\mathbf{X}_0(l)$  at each point;  $\max_j \max_i |x_0(l) - x_j(l)|$  is the maximum difference of the second level, which

represents the maximum difference in all sequences based on the maximum difference found in each sequence.

In order to avoid the resulting deviation caused by variable units and other factors, it is necessary to conduct standardized processing on variable data.

(4) Calculate the correlation degree

As the correlation coefficient denotes the value of correlation degree between the comparison sequence and the reference sequence at each time, it has more than one value, which could lead the information to be too scattered to facilitate the overall comparison. Therefore, it is necessary to concentrate the correlation coefficient of each moment into a value, that is, to find its average value, as the value expression of the correlation degree between the comparison sequence and the reference sequence.

Correlation degree  $r_i$  represents the gray correlation degree of comparison sequence  $X_i$  to reference sequence  $X_0$ , also called sequence correlation degree, average correlation degree, and line correlation degree, the formula of which is shown as follows:

$$r_i = \frac{1}{n} \sum_{k=1}^n \zeta_i(k) \quad (4)$$

The closer the value of  $r_i$  is to 1, the better the correlation is.

2. Clustering analysis

Specific clustering steps:

- (1) Calculate the gray correlation coefficients between every two variables, then sum the distances;
- (2) Calculate the correlation degree standard deviations of the above sums, utilizing  $w_d$  to denote the deviation result;
- (3) Based on the relationship between  $w_d$  (the value obtained by 0–1 normalization for the summation of the correlation coefficients of one variable to all other variables) and  $C_g$  (global correlation degree level), and the relation of Pearson correlation coefficients and correlation levels [39], variables are clustered into groups, listed in Table 1. Then, the variable weight of a variable in one group can be calculated through the data of variables in the group.

**Table 1.** Relationship between the values of  $w_d$  and  $C_g$ .

$w_d$	0–0.2	0.2–0.5	0.5–0.7	0.7–1
$C_g$	High	Well	Low	Weak

3. Variable weight calculation

The variable weight of a variable in one group can be determined through the mean square error method with specific steps, as below:

(1) Data normalization

$$Y_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad (5)$$

where,  $x_{ij}$  ( $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ ) denotes the initial data of the  $j$ th variable in group  $i$ .

(2) Mean value

$$\bar{Y}_{ij} = \frac{1}{m} \sum_{j=1}^m Y_{ij} \quad (6)$$

(3) Mean square error

$$\sigma_{ij} = \sqrt{\sum_{j=1}^m (Y_{ij} - \bar{Y}_{ij})^2} \quad (7)$$

(4) Variable weight

$$w_{ij} = \sigma_{ij} / \sum_{j=1}^m \sigma_{ij} \quad (8)$$

Herein, two efficiency indices are introduced totally, FAR and MAR. Compared with MAR, the correlation degree mainly reflects on FAR, which has a significant effect on the system. Therefore, the weight  $w_{ij}$  is given for FAR. Meanwhile,  $MAR/R_{MAR}$  is used in case of overlarge MAR, where  $R_{MAR}$  denotes the maximum acceptable MAR, values of which generally less than the engineering required error (0.05) with 0.01, recommended by [2].

### 2.3. Threshold Optimization

The optimization objective function, shown as Equation (9), is established according to the alarm information under normal and abnormal situations, which is solved by the numerical optimization method from the point of view of minimizing.

$$F(x) = \text{Min}\left(\frac{FAR}{R_{FAR}/(1+w_{ij})} + \frac{MAR}{R_{MAR}}\right) \quad (9)$$

where,  $R_{FAR}$  denotes the maximum acceptable FAR, the value of which generally less than the engineering required error (0.05) with 0.01, recommended by [2].

Figure 2 depicts the flow chart of a quadratic interpolation optimization algorithm with the basic thought shown as: for  $F(x) = \text{Min } \varphi(x) (x \in R^1)$ , the  $\varphi(x)$  can be fitted by  $y(x)$ , which consists of some dots. Then, the extreme point  $\mu$  of  $y(x)$  is an estimate value of  $x^*$ .

A threshold optimization algorithm is implemented as follows:

- (1) Give the initial interval  $[x_1, x_3]$ , three points  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$ , and convergence precision  $\varepsilon$ , where,  $x_1 < x_2 < x_3$ ,  $\varepsilon > 0$ ;
- (2) Calculate  $c_1, c_2$  (where,  $c_1 = (y_3 - y_1)/(x_3 - x_1)$ ,  $c_2 = [(y_2 - y_1)/(x_2 - x_1) - c_1]/(x_2 - x_3)$ ), and  $x_p = 0.5(x_1 + x_3 - c_1/c_2)$ ,  $y_p = f(x_p)$ ;
- (3) If  $|y_2 - y_p| \geq \varepsilon$ , then go step (4), otherwise, go step (9);
- (4) If  $x_p > x_2$ , then go step (5), otherwise, go step (7);
- (5) If  $y_2 \geq y_p$ , then  $x_3 = x_p$ ,  $y_3 = y_p$ , return to step (2) otherwise, go step (6);
- (6) Let  $x_1 = x_2$ ,  $y_1 = y_2$ ,  $x_2 = x_p$ ,  $y_2 = y_p$ , return to step (2);
- (7) If  $y_2 < y_p$ , then  $x_1 = x_p$ ,  $y_1 = y_p$ , return to step (2) otherwise, go step (8);
- (8) Let  $x_3 = x_2$ ,  $y_3 = y_2$ ,  $x_2 = x_p$ ,  $y_2 = y_p$ , return to step (2);
- (9) If  $y_2 < y_p$ , then  $x^* = x_2$ ,  $y^* = y_2$ , otherwise, go step (10);
- (10)  $x^* = x_p$ ,  $y^* = y_p$ ;
- (11) Output  $x^* = x_p, f^* = f(x^*)$ .

### 2.4. Optimization Process Description

Figure 3 gives the optimization algorithm, the specific explanations of which are shown as follows:

To begin with, the correlation degrees of variables are obtained by analyzing correlation relationships among them. Subsequently, the variables are grouped according to the gray correlation coefficients and clustering analysis, given the weight  $\omega_i$  for FAR in each group. An objective function about the FAR, MAR,  $R_{FAR}$ ,  $R_{MAR}$ , and  $\omega_i$  is then established with variable weight. Eventually, based on the optimization algorithm, the objective function is optimized for obtaining the optimal alarm threshold.

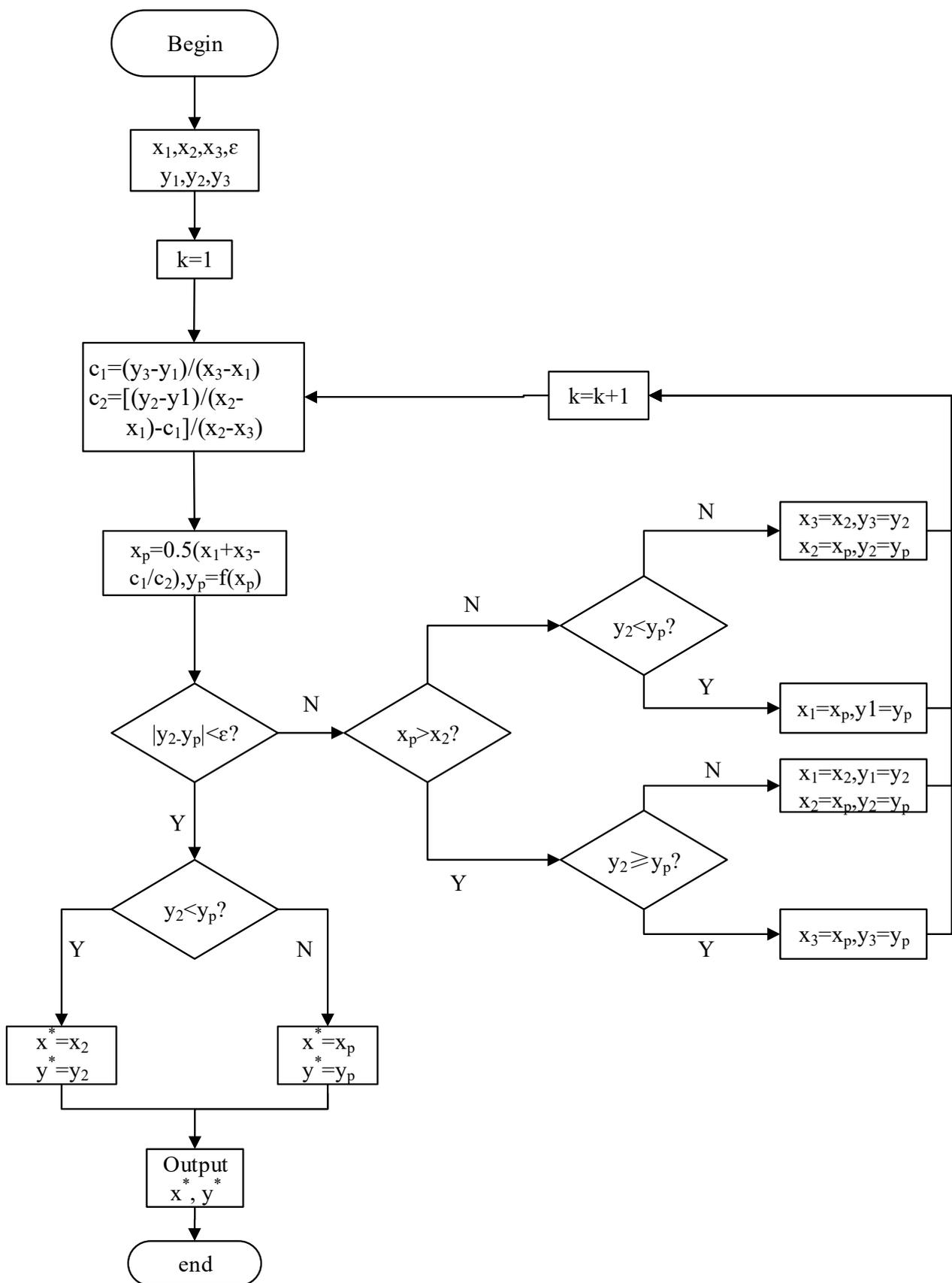
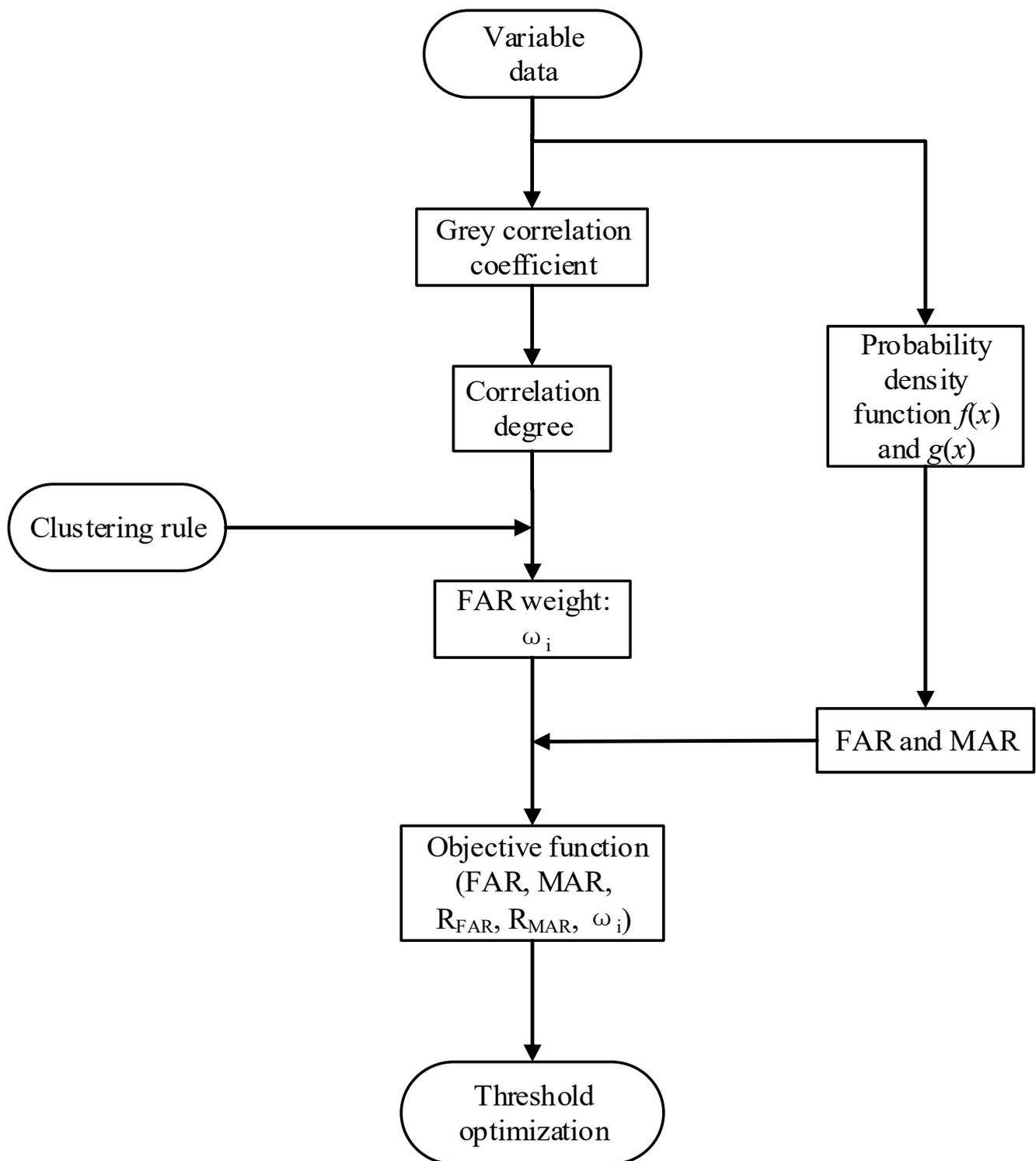


Figure 2. Flow diagram of quadratic interpolation optimization algorithm.



**Figure 3.** Flow chart of alarm threshold optimization process.

### 3. Theory Study—Tennessee Eastman (TE) Simulation Process

#### 3.1. Process Description

TE process was put forward by J. J. Downs and E. F. Vogel. It can be used as a data source, which is commonly utilized for comparing various methods, such as control optimization. Therefore, this work uses the TE simulation process as a case.

Figure 4 shows the flow diagram for a TE process, which contains five major operating units: reactor, condenser, compressor, separator, and stripper. The TE process consists of 15 known failures and 5 unknown failures. Meanwhile, it also consists of 12 operational variables and 41 measured variables. Table 2 lists the selected 10 measured variables for researching the applicability analysis of the method. To verify the accuracy of the results, the same sampling environment (the faults were after 8 simulation hours introduced) should be necessary, sampling interval of which is  $\Delta T = 3$  min by considering the time constants of the process in a closed loop [40,41], as the process under the sampling time can be considered reaches a relatively steady running state, which could reflect the running state with a long period of time, to some extent. Meanwhile, 960 groups of data under a normal condition and 500 groups of data under an abnormal condition with failure 6 are collected.

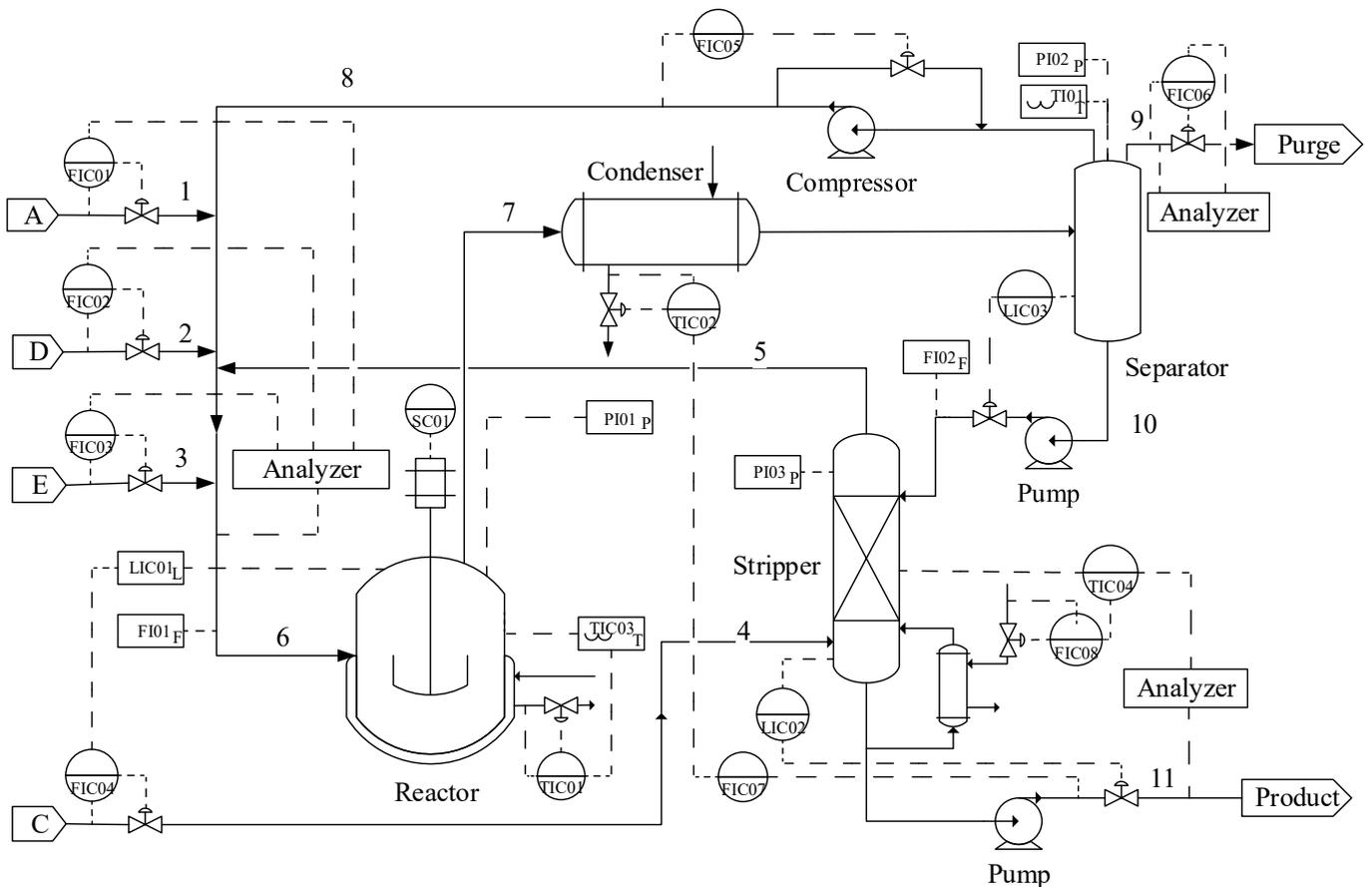


Figure 4. Flow chart of TE process.

Table 2. Ten measured variables in TE process.

Variable	Variable Name	Reference Value	Unit
V1	A feed flow	0.25052	kscmh
V2	D feed flow	3664	kg·h <sup>-1</sup>
V3	E feed flow	4509.3	kg·h <sup>-1</sup>
V4	Feed flow of A and C	9.3477	kscmh
V5	Cycle flow	26.902	kscmh
V6	Reactor feed flow	42.339	kscmh
V7	Reactor pressure	2705	kPa
V8	Reactor level	75	%
V9	Reactor temperature	120.4	°C
V10	Blow-down flow	0.3372	kscmh

### 3.2. Cluster Variables and Calculate Weights

#### (1) Variable clustering

The selected ten variables can be regarded as 10 vectors with 960 dimensions, containing 960 observations of normal data. Table 3 lists the correlation degree of these ten variables.

**Table 3.** Correlation degree of ten variables.

<b>d</b>	<b>V1</b>	<b>V2</b>	<b>V3</b>	<b>V4</b>	<b>V5</b>	<b>V6</b>	<b>V7</b>	<b>V8</b>	<b>V9</b>	<b>V10</b>
V1	0.00	0.10	0.19	0.53	0.16	0.15	0.55	0.32	0.55	0.90
V2		0.0	0.67	0.19	0.37	0.15	0.43	0.67	0.24	0.52
V3			0.00	0.19	0.22	0.24	0.23	0.18	0.22	0.24
V4				0.00	0.56	0.18	0.70	0.90	0.66	0.87
V5					0.00	0.78	0.97	0.88	0.10	0.49
V6						0.00	0.72	0.66	0.84	0.61
V7							0.00	0.10	0.88	0.65
V8								0.00	0.66	0.12
V9									0.00	0.65
V10										0.00

Calculating the sum ( $d_T$ ) of the correlation degree between one variable and all other variables is necessary for treating ten variables as a whole. Table 4 lists the sums and the normalized result ( $w_d$ ).

**Table 4.** Sums of the correlation degree and normalized result of ten variables.

<b>d</b>	<b>V1</b>	<b>V2</b>	<b>V3</b>	<b>V4</b>	<b>V5</b>	<b>V6</b>	<b>V7</b>	<b>V8</b>	<b>V9</b>	<b>V10</b>
$d_T$	3.45	3.34	2.38	4.78	4.53	4.33	5.23	4.49	4.8	5.05
$w_d$	0.38	0.34	0.00	0.84	0.75	0.68	1.00	0.74	0.85	0.94

Based on the normalization result and criterions given in Table 1, the original ten variables are clustered into four groups. Variables V1 and V2 constitute the first group, variable V3 constitutes the second group, variable V6 constitutes the third group and the rest belong to the last group.

#### (2) Variable weight

Table 5 lists the weights for variables in four groups.

**Table 5.** Variable weights.

<b>Group Number</b>	<b>Variable</b>	<b>Weight</b>
1	V1	0.521
	V2	0.479
2	V3	1
3	V4	0.164
	V5	0.147
	V7	0.195
	V8	0.145
	V9	0.166
4	V10	0.183
	V6	1

For the FAR and MAR, the impact on the system caused by the correlation degree of variables commonly reflect on FAR more directly than MAR, therefore, giving the correlation weight to FAR.

### 3.3. Optimization Solution

Steps of the optimization solution are listed as below, using the first variable V1 as an example.

First, the probability density function Equations (10) and (11) of V1 in the normal and abnormal cases are fitted with 960 observations of normal data and 500 observations of abnormal data, respectively. Figure 5 shows the corresponding probability density curves.

$$f(x) = 13.973 \cdot e^{-613.36 \cdot (x-0.25114)^2} \quad (10)$$

$$g(x) = 4.2301 \cdot e^{-56.214 \cdot (x-0.041936)^2} \quad (11)$$

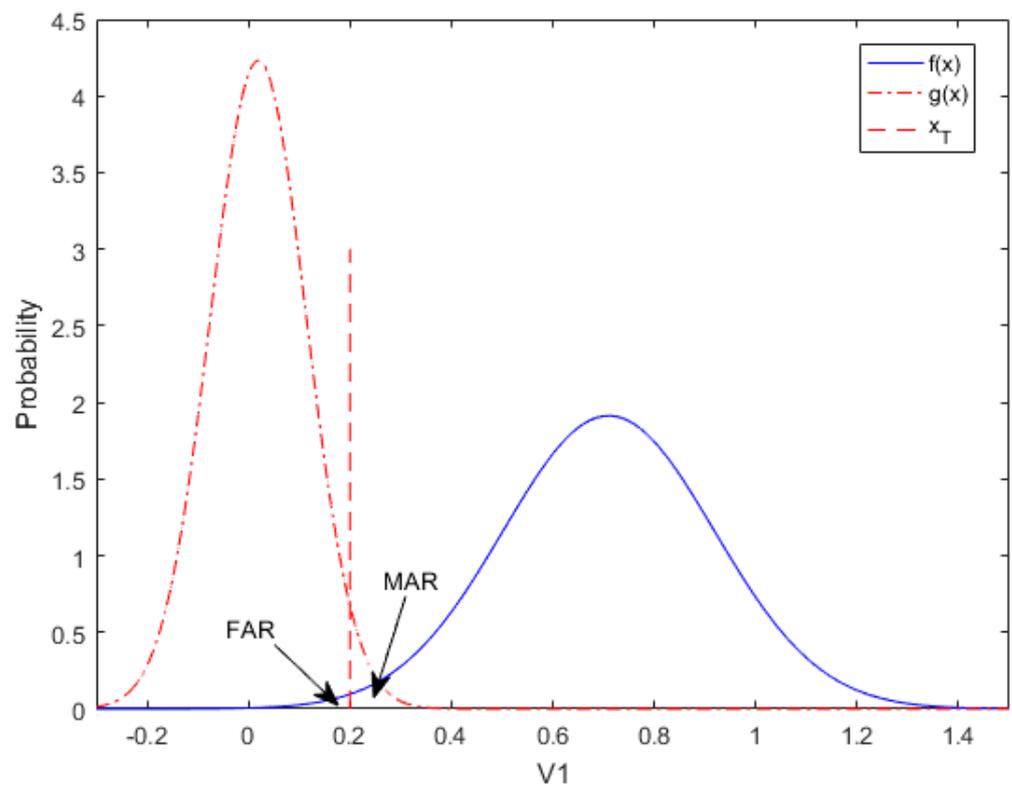


Figure 5. Process probability density curves of V1.

Second, the objective function is obtained as Equation (12), the parameters of which, including weight  $w_1 = 0.521$ , Equations (1) and (2),  $R_{FAR} = R_{MAR} = 0.01$ , are input into function Equation (9).

$$\begin{aligned} F(x) &= \text{Min} \left( \frac{FAR}{R_{FAR}/(1+w_{ij})} + \frac{MAR}{R_{MAR}} \right) \\ &= \text{Min} \left( \frac{\int_{-\infty}^{x_T} 13.973 \cdot e^{-613.36 \cdot (x-0.25114)^2} dx}{R_{FAR}/(1+w_{ij})} \right. \\ &\quad \left. + \frac{\int_{x_T}^{+\infty} 4.2301 \cdot e^{-56.214 \cdot (x-0.041936)^2} dx}{R_{MAR}} \right) \end{aligned} \quad (12)$$

Eventually, the objective function is optimized as Figure 6, with optimum  $x_T = 0.22$ ,  $F(x_T) = 6.5098$ , FAR = 0.00654, MAR = 0.06710. The thresholds of other variables are optimized similarly.

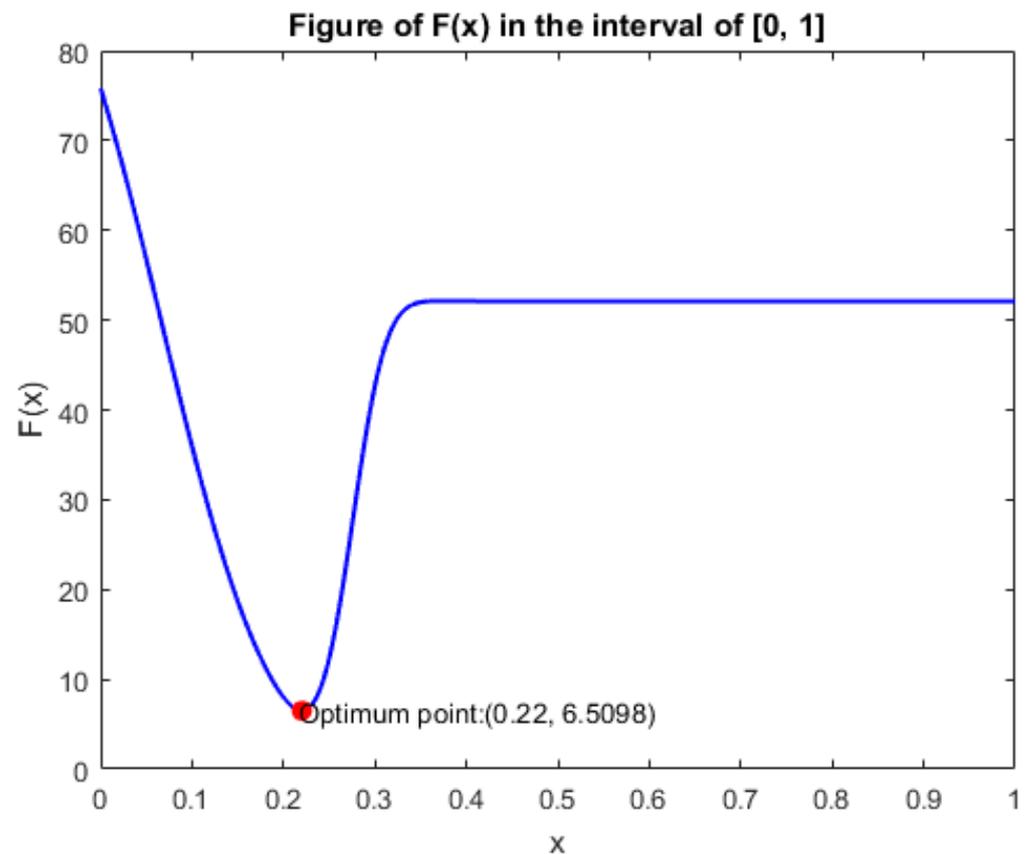
Additionally, to verify the effectiveness of this method, some other methods should be utilized for comparison, such as the deadband [42], alarm delay [36], and moving average filter (MAF) with original reference value [27]. The summarized results listed in Table 6.

To verify the effectiveness of the method, an abnormal data set of fault 8 is added. Table 6 lists the results under the two failures (failure 6: f6, failure 8: f8).

### 3.4. Results and Analysis

Shown in Figure 7, the FAR calculated by the proposed method has an effective reduction in cases with initial thresholds, and the MAR of which under control simultaneously.

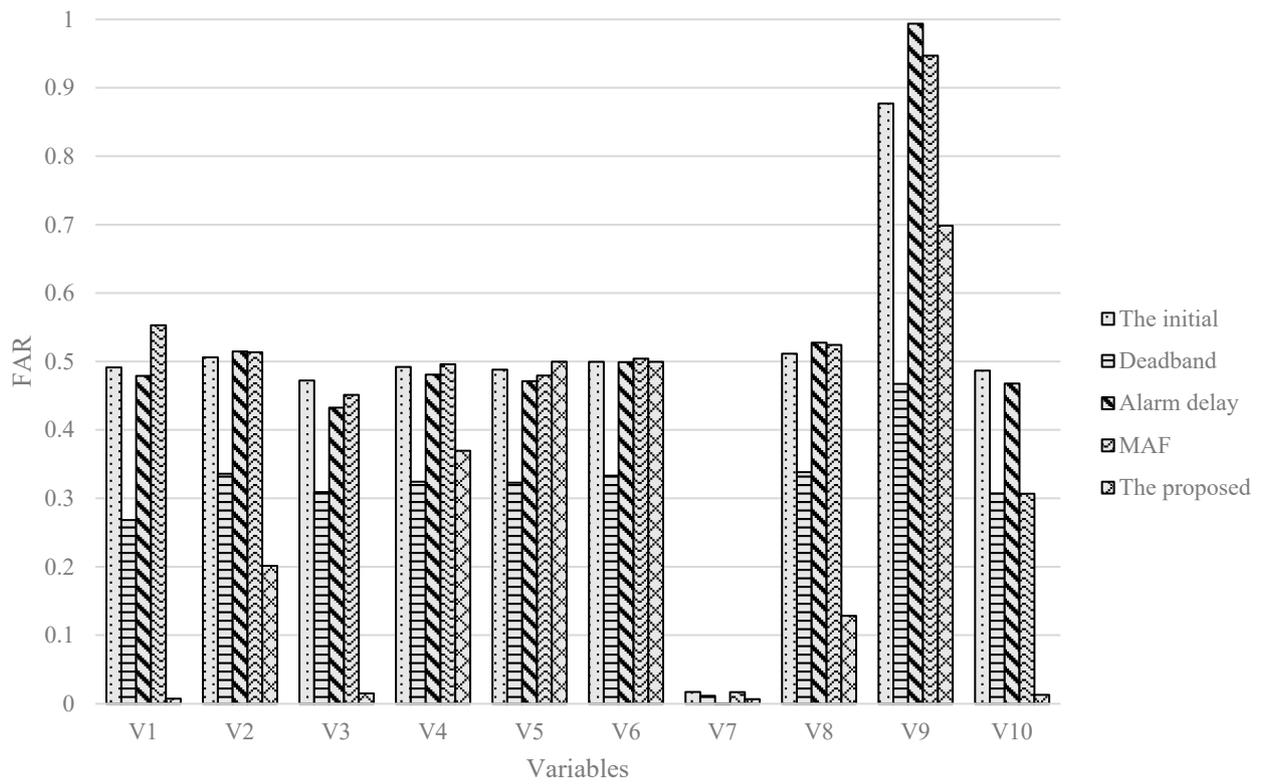
In Figure 8, the numbers of variable alarms in total under 5 cases under the case with failure 6 (8) are 3532, 3414, 2916, 2928, 1733 (3532, 3417, 2968, 2992, 1744), respectively. Compared with the other four methods, the alarm reduction rates calculated by the presented method are 50.9%, 49.2%, 40.5%, and 40.8% (50.6%, 48.9%, 41.2%, and 41.7%), respectively. From which, it reflects the method could have some impact on the alarm threshold optimization, bring a lower FAR and fewer alarms.



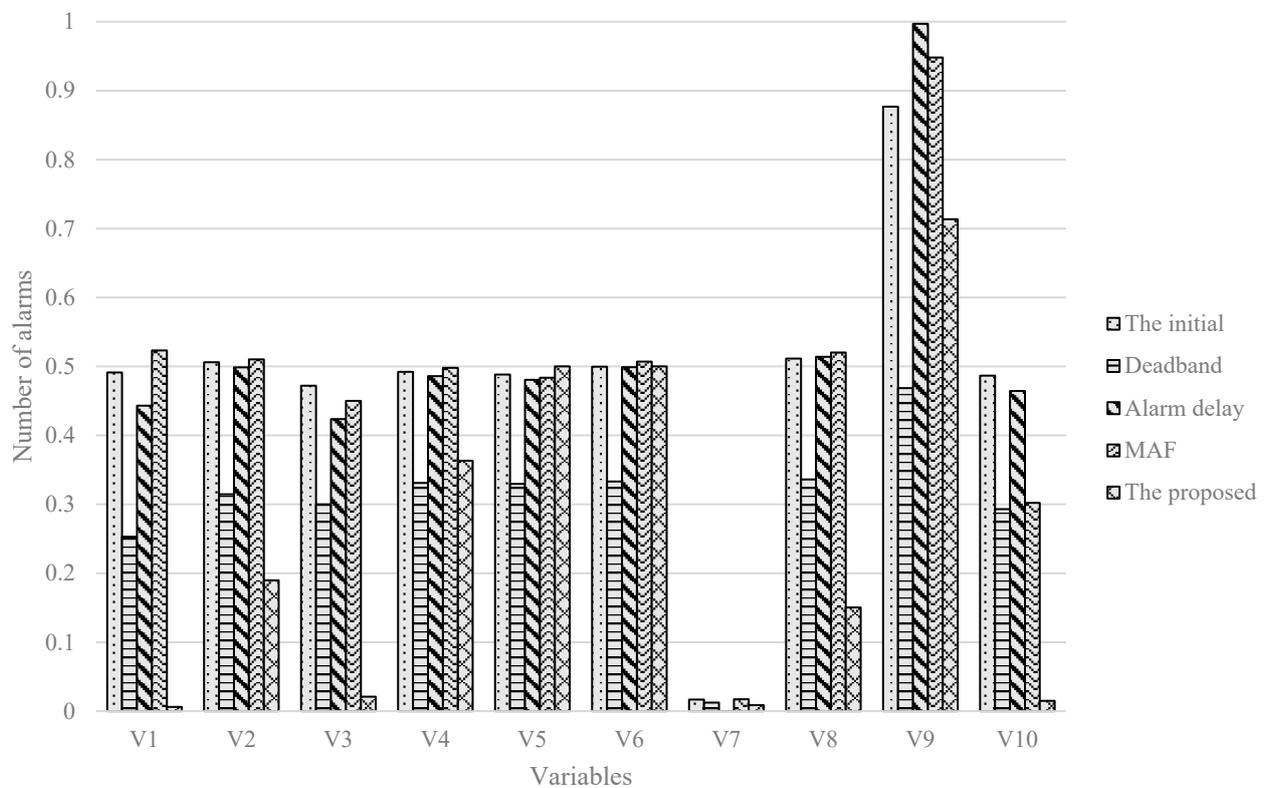
**Figure 6.** Optimization result of V1.

Table 6. Optimized results for ten variables.

Variable		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
Threshold	The initial	0.25052	3664.00	4509.30	9.3477	26.9020	42.3390	2705.00	75.0000	120.400	0.33720
	Deadband	0.25052	3664.00	4509.30	9.3477	26.9020	42.3390	2705.00	75.0000	120.400	0.33720
	MAF	0.25052	3664.00	4509.30	9.3477	26.9020	42.3390	2705.00	75.0000	120.400	0.33720
	Alarm delay	0.25052	3664.00	4509.30	9.3477	26.9020	42.3390	2705.00	75.0000	120.400	0.33720
	The proposed (f6)	0.22000	3637.89	4452.98	9.4501	26.9080	42.3390	2723.21	74.4153	120.409	0.32122
	The proposed (f8)	0.23153	3654.51	4478.39	9.3801	26.9053	42.3380	2721.32	74.4567	120.407	0.33457
FAR	The initial	0.49130	0.50604	0.47213	0.49210	0.48818	0.49959	0.01698	0.51133	0.87689	0.48667
	Deadband (f6)	0.26829	0.33600	0.30906	0.32424	0.32298	0.33306	0.01155	0.33833	0.46720	0.30738
	Deadband (f8)	0.25312	0.31472	0.30001	0.33114	0.32978	0.33309	0.01251	0.33612	0.46851	0.29315
	MAF (f6)	0.47888	0.51467	0.43264	0.48082	0.47132	0.49900	0.00001	0.52749	0.99361	0.46766
	MAF (f8)	0.44321	0.49876	0.42367	0.48587	0.48075	0.49900	0.00001	0.51387	0.99697	0.46451
	Alarm delay (f6)	0.55287	0.51352	0.45118	0.49591	0.47952	0.50421	0.01682	0.52420	0.94702	0.30689
	Alarm delay (f8)	0.52315	0.51031	0.45003	0.49791	0.48367	0.50697	0.01735	0.52032	0.94821	0.30222
	The proposed (f6)	0.00654	0.19138	0.02361	0.35420	0.49999	0.49879	0.00859	0.15455	0.71132	0.01546
The proposed (f8)	0.00623	0.18975	0.02097	0.36317	0.49999	0.49999	0.00873	0.15036	0.71354	0.01512	
MAR	The initial	0.01271	0.04000	0.03830	0.51772	0.51752	0.31243	0.46972	0.08867	0.49999	0.02224
	Deadband (f6)	0.14641	0.00001	0.03410	0.02498	0.02276	0.00041	0.23371	0.00001	0.00001	0.04727
	Deadband (f8)	0.18735	0.00001	0.03595	0.02149	0.01642	0.00029	0.15491	0.00001	0.00001	0.04859
	MAF (f6)	0.00001	0.00020	0.00018	0.54295	0.54247	0.12568	0.42688	0.00230	0.49998	0.00003
	MAF (f8)	0.00001	0.00021	0.00019	0.52137	0.53126	0.11064	0.41978	0.00233	0.49998	0.00003
	Alarm delay (f6)	0.02965	0.02476	0.03712	0.53185	0.53948	0.22475	0.46808	0.07623	0.50000	0.02593
	Alarm delay (f8)	0.03123	0.02574	0.03821	0.49870	0.53125	0.22379	0.44364	0.07711	0.50000	0.02634
	The proposed (f6)	0.06710	0.21004	0.06756	0.87869	0.44817	0.30321	0.95422	0.28201	0.78332	0.06545
The proposed (f8)	0.06857	0.21174	0.06838	0.85413	0.44105	0.30021	0.92154	0.28713	0.78114	0.06637	
Number of total alarms	The initial	303	301	287	301	687	301	340	333	390	289
	Deadband (f6)	188	301	287	301	687	301	340	333	390	286
	Deadband (f8)	193	311	295	295	673	300.973	321	336	393	300
	MAF (f6)	269	230	234	229	574	243	332	252	321	232
	MAF (f8)	273	241	239	231	596	243	332	259	320	233
	Alarm delay (f6)	295	288	298	321	296	320	215	321	313	261
	Alarm delay (f8)	307	290	299	331	310	323	230	323	314	265
	The proposed (f6)	76	190	115	149	320	298	98	170	212	105
The proposed (f8)	81	195	123	137	313	293	94	181	207	120	

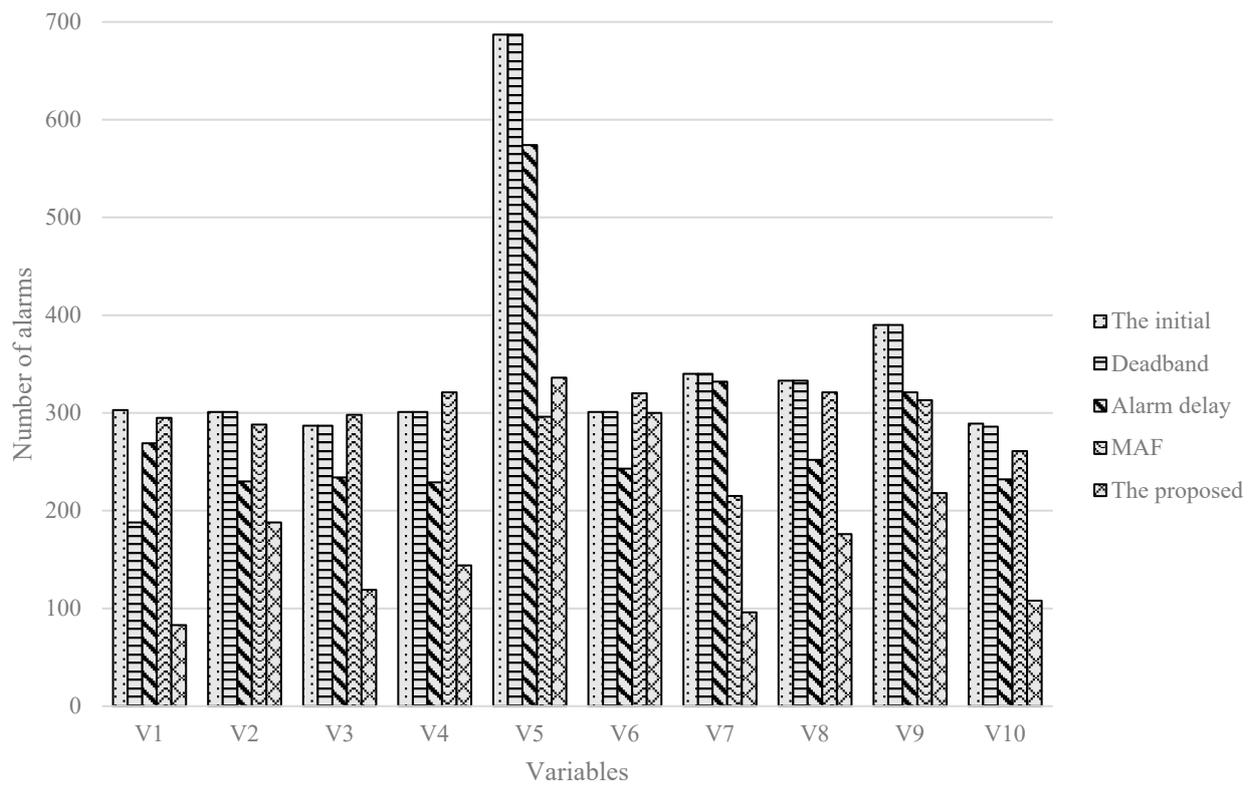


(f6)

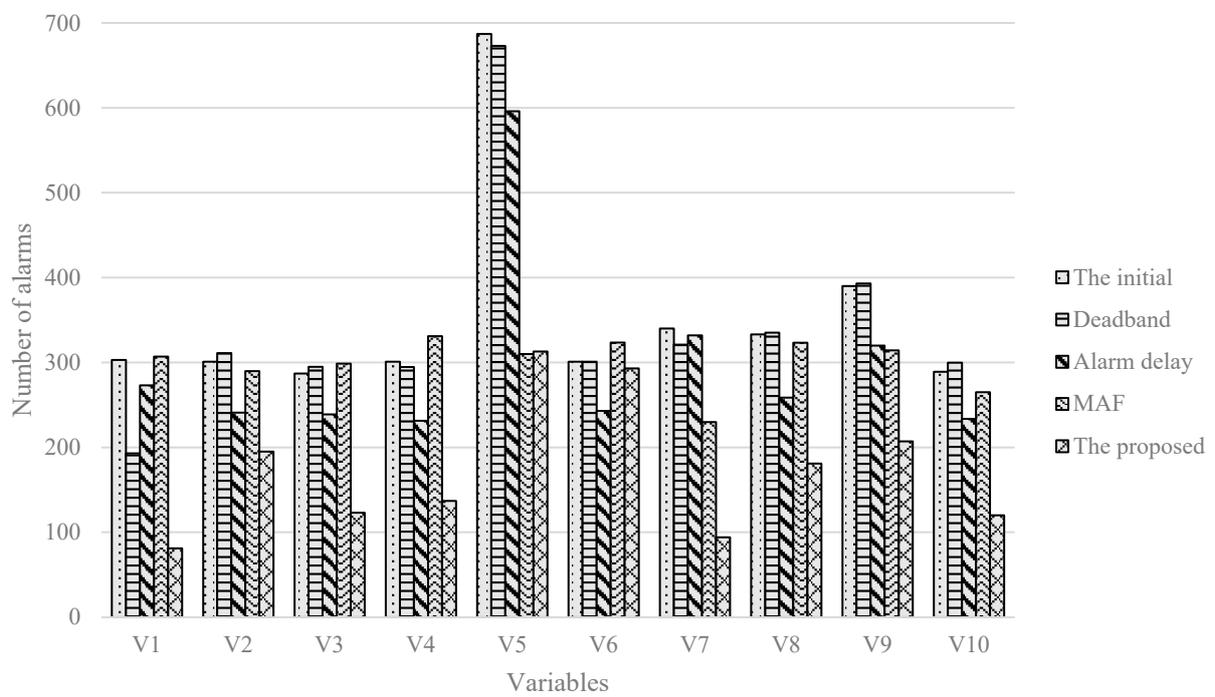


(f8)

Figure 7. Histograms of FARs in five cases with two failures (f6,f8).



(f6)



(f8)

Figure 8. Histograms of number of alarms in five cases with two failures (f6,f8).

## 4. Real Industrial Verification—Industrial Ethylene Production Process

### 4.1. Process Description

To verify the effectiveness of the method in actual industry, an industrial ethylene production process was selected as an actual case, the process of which generally consists of four major processes: cracking, compression, quenching, and separation.

Figure 9 shows the flow chart of the ethylene production process, the separation section of which contains the most of research object. Due to the large number of variables in this process, in order to avoid the influence of blind selection on the results, we try to objectively select variables with correlation relationships among them not easy to judge and which have significant impact on the yield and quality of ethylene through empirical knowledge and analysis. Therefore, ten variables are selected to represent the process, which are listed in Table 7. For this process, the time-lag effect under start-up or shut-down cases is great. Therefore, the data under the steady state are selected as normal data and the data under the state with disturbances are then chosen as abnormal data for study, respectively, are selected for study. In total, 1000 observations of data for 10 variables are extracted with a sampling interval of 1 min, containing 500 observations of normal data and 500 observations of abnormal data with the feed flow of cracking furnace increases by 10%.

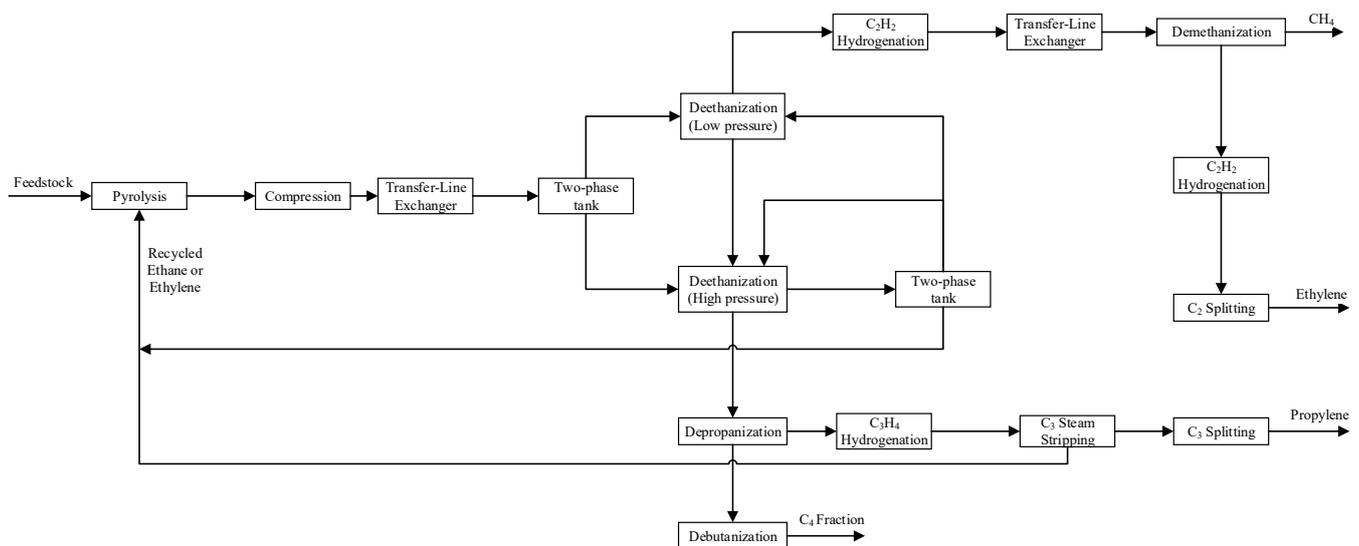


Figure 9. Flow chart of ethylene production process.

Table 7. Ten measured variables in ethylene production process.

Variable	Variable Name	Reference Value	Unit
V1	Vapor fraction of separator to dethanizer	0.71	
V2	Feed flow into C2/C3 column with high pressure	175	t·h <sup>-1</sup>
V3	Feed flow into C2/C3 column with low pressure	88	t·h <sup>-1</sup>
V4	Feed flow into C2 reactor	200	t·h <sup>-1</sup>
V5	Vapor fraction of separator to demethanizer	0.69	
V6	Feed flow into demethanizer	101	t·h <sup>-1</sup>
V7	Bottom flow of ethylene distillation tower	25	t·h <sup>-1</sup>
V8	Feed flow into C3/C4 tower	130	t·h <sup>-1</sup>
V9	Feed flow to C3 reactor	155	t·h <sup>-1</sup>
V10	Circulation flow from C3 stripping tower to the former	1.5	t·h <sup>-1</sup>

#### 4.2. Cluster Variables and Calculate Weights

##### (1) Variable clustering

The selected ten variables can be regarded as 10 vectors with 500 dimensions, containing 500 observations of normal data. Table 8 lists the correlation degree of these 10 variables.

**Table 8.** Correlation degree of 10 variables.

d	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V1	0.00	0.12	0.15	0.43	0.23	0.15	0.43	0.67	0.24	0.52
V2		0.0	0.57	0.43	0.22	0.24	0.23	0.18	0.22	0.24
V3			0.00	0.10	0.56	0.18	0.70	0.90	0.66	0.87
V4				0.00	0.35	0.97	0.78	0.34	0.10	0.49
V5					0.00	0.72	0.66	0.53	0.84	0.61
V6						0.00	0.10	0.22	0.88	0.65
V7							0.00	0.12	0.66	0.12
V8								0.00	0.23	0.18
V9									0.00	0.43
V10										0.00

Calculating the sum ( $d_T$ ) of the correlation degree between one variable and all other variables is necessary for treating ten variables as a whole. Table 9 lists the sums and the normalized result ( $w_d$ ).

**Table 9.** Sums of the correlation degree and normalized result of ten variables.

D	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
$d_T$	2.82	2.33	3.97	3.13	4.27	3.72	3.14	2.52	3.8	3.35
$w_d$	0.25	0.00	0.85	0.41	1.00	0.72	0.42	0.10	0.76	0.53

Based on the normalization result and criterions given in Table 1, the original ten variables are clustered into four groups. Variables V1, V4, and V7 constitute the first group, variables V2 and V8 constitute the second group, variable V10 constitutes the third group and the rest belong to the last group.

##### (2) Variable weight

Table 10 lists the weights for variables in four groups.

**Table 10.** Variable weights.

Group Number	Variable	Weight
1	V1	0.453
	V4	0.241
	V7	0.306
2	V2	0.542
	V8	0.458
3	V3	0.274
	V5	0.176
	V6	0.392
4	V9	0.158
	V10	1

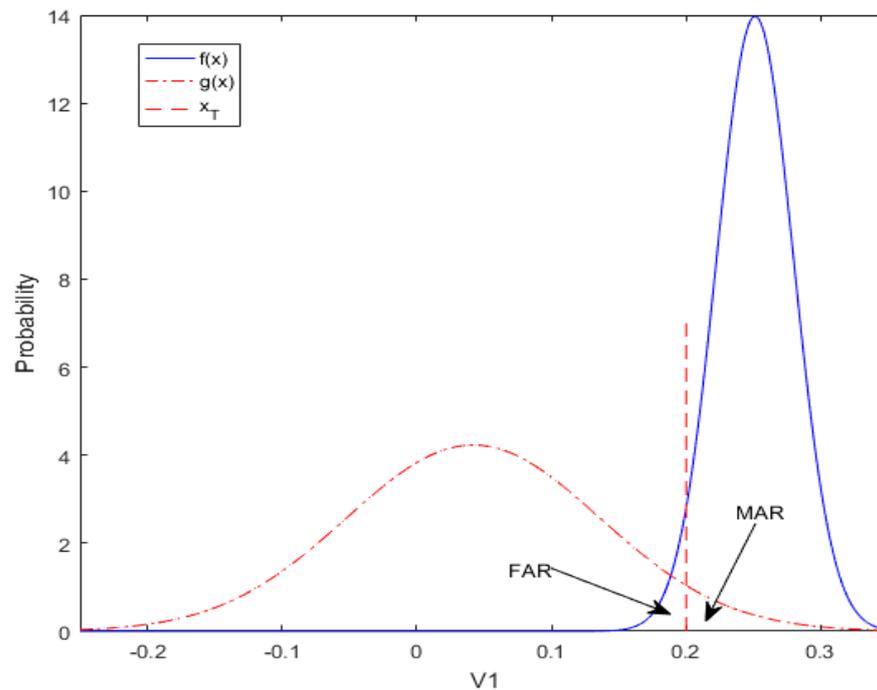
#### 4.3. Optimization Solution

Steps of the optimization solution are listed as below, using the first variable V1 as an example.

First, the probability density functions Equations (13) and (14) of V1 in normal and abnormal cases are fitted with 500 observations of data, respectively. Figure 10 shows the corresponding probability density curves.

$$f(x) = 12.756 \cdot e^{-213.36 \cdot (x-0.71021)^2} \quad (13)$$

$$g(x) = 2.1321 \cdot e^{-34.231 \cdot (x-0.019436)^2} \quad (14)$$



**Figure 10.** Process probability density curves of V1.

Second, the objective function is obtained as Equation (15), parameters of which including weight  $w_1 = 0.453$ , Equations (1) and (2),  $R_{FAR} = R_{MAR} = 0.01$  are inputted into function Equation (9).

$$\begin{aligned} F(x) &= \text{Min} \left( \frac{FAR}{R_{FAR}/(1+w_{ij})} + \frac{MAR}{R_{MAR}} \right) \\ &= \text{Min} \left( \frac{\int_{-\infty}^{x_T} 12.756 \cdot e^{-213.36 \cdot (x-0.71021)^2} dx}{R_{FAR}/(1+w_{ij})} \right. \\ &\quad \left. + \frac{\int_{x_T}^{+\infty} 2.1321 \cdot e^{-34.231 \cdot (x-0.019436)^2} dx}{R_{MAR}} \right) \end{aligned} \quad (15)$$

Eventually, the objective function is optimized as Figure 11, with optimum  $x_T = 0.73$ ,  $F(x_T) = 6.5095$ ,  $FAR = 0.00354$ ,  $MAR = 0.0661$ . The thresholds of other variables are optimized similarly.

To verify the effectiveness of the method, an abnormal data set of the feed flow of cracking furnace decreases by 10% and is added. Table 11 lists the results under the two cases (cracking furnace increases by 10%: c1, cracking furnace decreases by 10%: c2).

#### 4.4. Results and Analysis

Shown as the Figure 12, the FAR calculated by the proposed method has an effective reduction in cases with initial thresholds, and the MAR of which under control simultaneously.

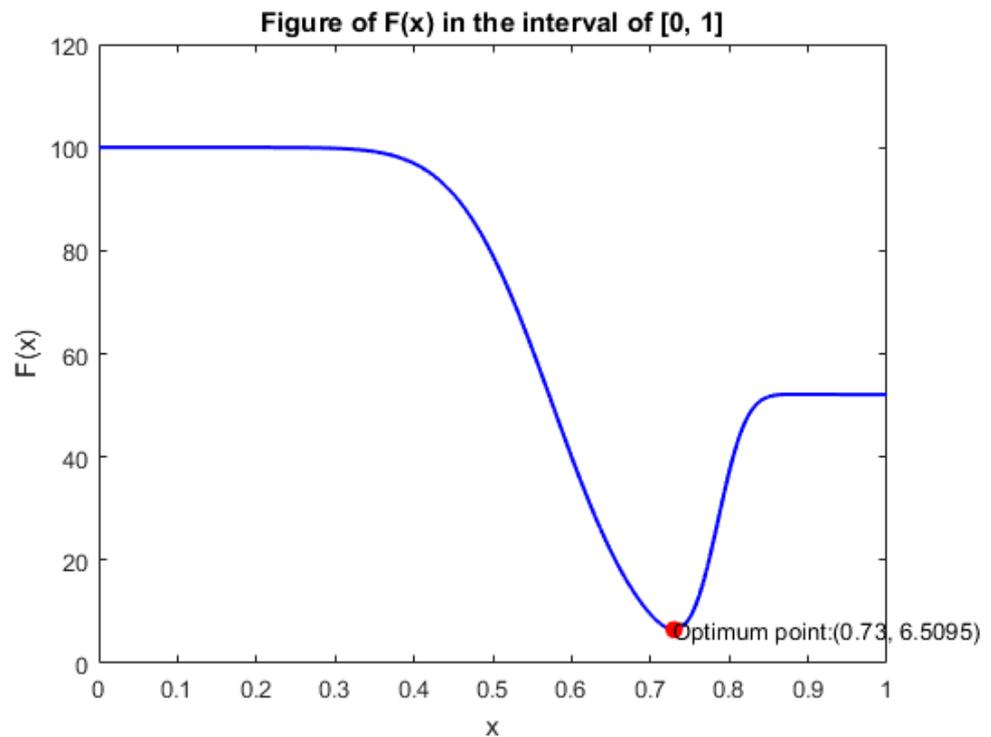
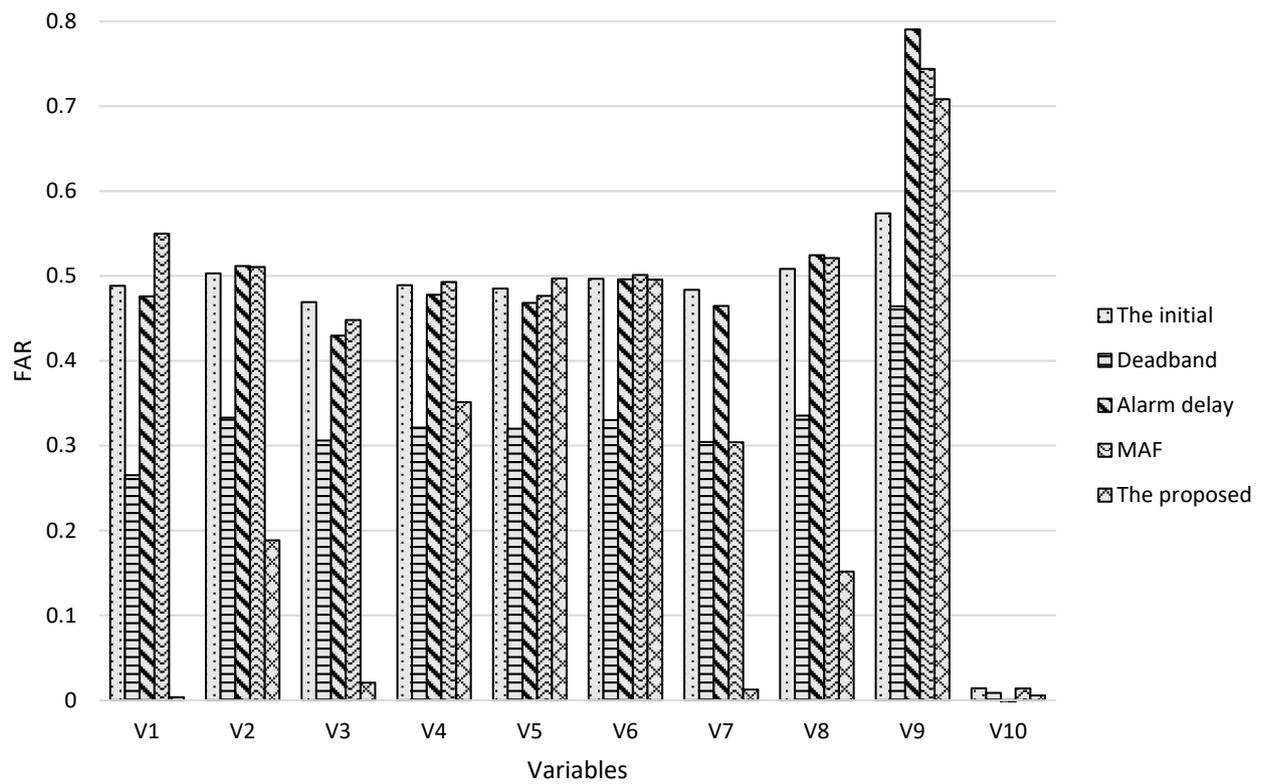


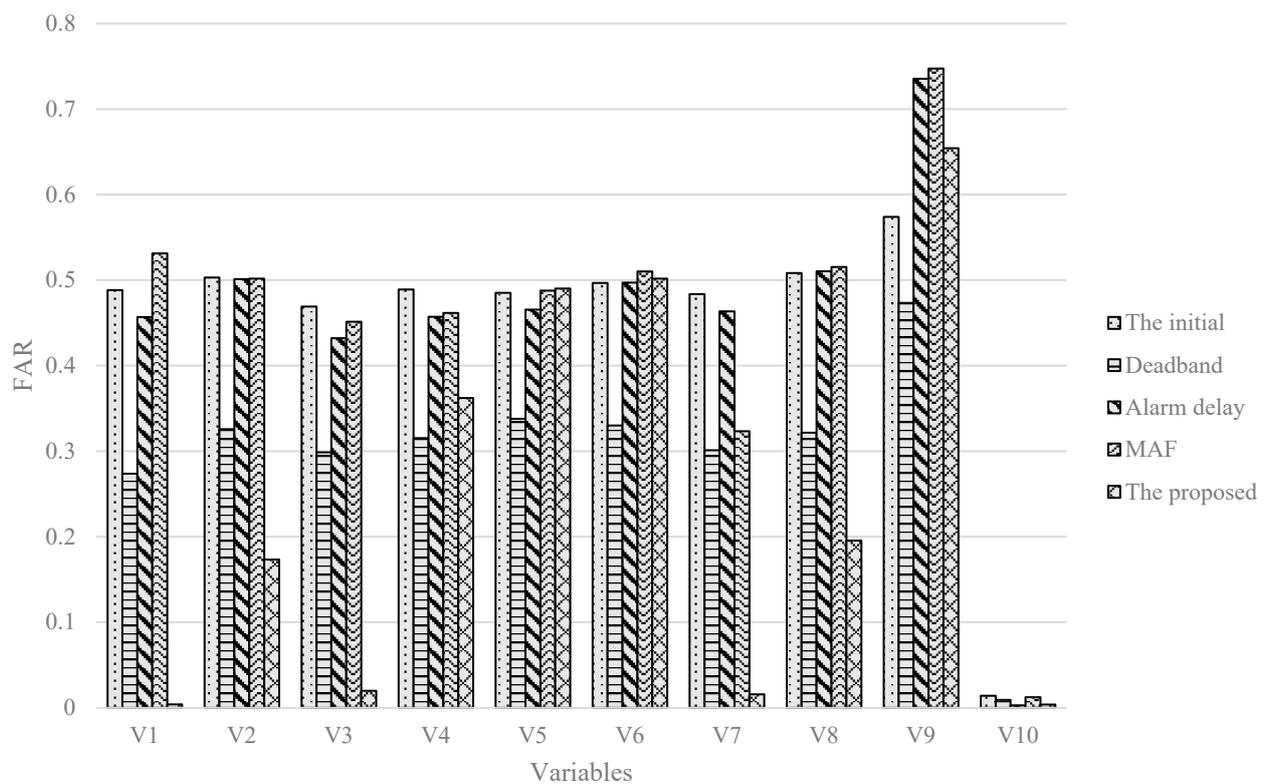
Figure 11. Optimization result of V1.

Table 11. Optimized results for 10 variables.

Variable		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
Threshold	The initial	0.71	175.0	88.0	200.0	0.69	101.0	25.0	130.0	155.0	1.50
	Deadband	0.71	175.0	88.0	200.0	0.69	101.0	25.0	130.0	155.0	1.50
	MAF	0.71	175.0	88.0	200.0	0.69	101.0	25.0	130.0	155.0	1.50
	Alarm delay	0.71	175.0	88.0	200.0	0.69	101.0	25.0	130.0	155.0	1.50
	The proposed (c1)	0.73	174.8	87.7	198.6	0.70	101.7	24.8	131.2	156.1	1.47
	The proposed (c2)	0.72	174.9	87.8	199.3	0.70	101.5	25.0	130.8	155.6	1.49
FAR	The initial	0.48830	0.50304	0.46913	0.48910	0.48518	0.49659	0.48367	0.50833	0.57389	0.01398
	Deadband (c1)	0.26529	0.33300	0.30606	0.32124	0.31998	0.33006	0.30438	0.33533	0.46420	0.00855
	Deadband (c2)	0.27356	0.32570	0.29870	0.31520	0.33780	0.32151	0.30112	0.32154	0.47330	0.00911
	MAF (c1)	0.47588	0.51167	0.42964	0.47782	0.46832	0.49600	0.46466	0.52449	0.79061	0.00299
	MAF (c2)	0.45677	0.50117	0.43213	0.45712	0.46552	0.49710	0.46351	0.51023	0.73542	0.00278
	Alarm delay (c1)	0.54987	0.51052	0.44818	0.49291	0.47652	0.50121	0.30389	0.52120	0.74402	0.01382
	Alarm delay (c2)	0.53123	0.50174	0.45133	0.46151	0.48773	0.51010	0.32331	0.51530	0.74731	0.01231
	The proposed (c1)	0.00354	0.18838	0.02061	0.3512	0.49699	0.49579	0.01246	0.15155	0.70832	0.00559
	The proposed (c2)	0.00401	0.17321	0.01987	0.36215	0.49012	0.50164	0.01574	0.19541	0.65410	0.00379
MAR	The initial	0.01171	0.03900	0.03730	0.51672	0.51652	0.31143	0.46872	0.08767	0.49899	0.02124
	Deadband (c1)	0.14541	0.00099	0.03310	0.02398	0.02176	0.00059	0.23271	0.00099	0.00099	0.04627
	Deadband (c2)	0.14221	0.00103	0.03712	0.02531	0.01827	0.00061	0.24213	0.00101	0.00097	0.04518
	MAF (c1)	0.00099	0.00080	0.00082	0.54195	0.54147	0.12468	0.42588	0.00130	0.49898	0.00097
	MAF (c2)	0.00102	0.00082	0.00081	0.56749	0.54423	0.12440	0.42694	0.00133	0.52643	0.00103
	Alarm delay (c1)	0.02865	0.02376	0.03612	0.53085	0.53848	0.22375	0.46708	0.07523	0.499	0.02493
	Alarm delay (c2)	0.02995	0.02422	0.03623	0.56264	0.53236	0.22202	0.44372	0.07615	0.50217	0.02897
	The proposed (c1)	0.0661	0.20904	0.06656	0.87769	0.44717	0.30221	0.95322	0.28101	0.78232	0.06445
	The proposed (c2)	0.05817	0.22982	0.06913	0.85166	0.45297	0.30167	0.76713	0.23012	0.85964	0.09701
Number of total alarms	The initial	235	233	219	233	319	233	272	265	322	221
	Deadband (c1)	120	233	219	233	319	233	272	265	322	218
	Deadband (c2)	138	229	227	209	303	252	289	247	311	225
	MAF (c1)	201	162	166	161	306	175	264	184	253	164
	MAF (c2)	211	182	181	147	321	187	278	199	248	175
	Alarm delay (c1)	227	220	230	253	228	252	147	253	245	193
	Alarm delay (c2)	234	201	238	248	222	237	152	274	233	200
	The proposed (c1)	18	162	147	81	252	230	43	102	144	137
	The proposed (c2)	21	154	165	75	278	223	31	91	157	151



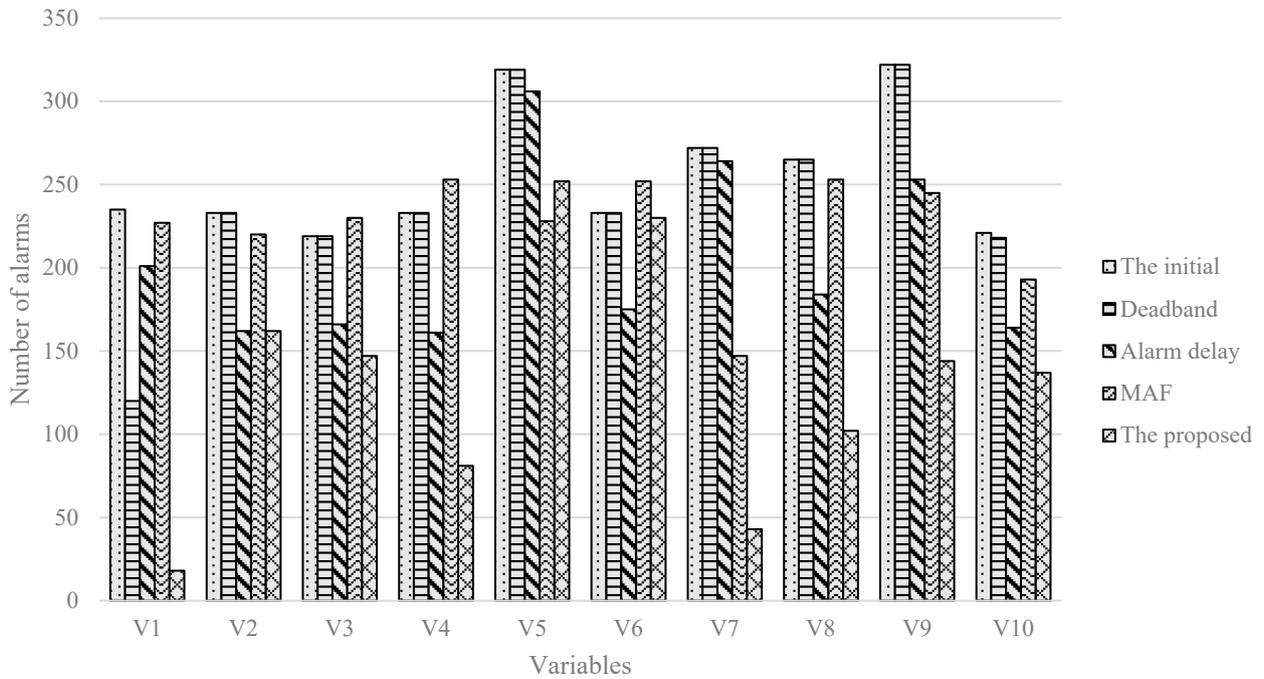
(c1)



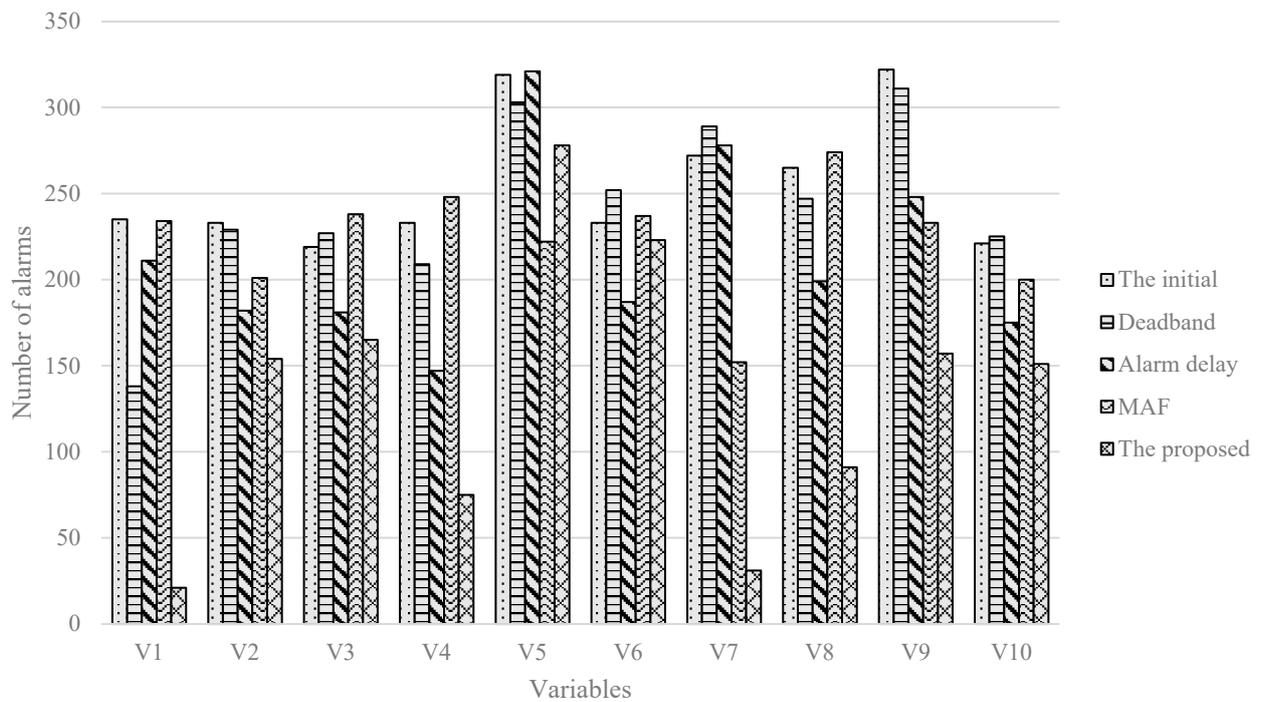
(c2)

Figure 12. Histograms of FARs in five cases with two cases (c1,c2).

In Figure 13, the numbers of variable alarms in total under 5 cases under the case of c1 (c2) are 2552, 2434, 2036, 2248, 1316 (2552, 2430, 2129, 2239, 1346), respectively. Compared with the other four methods, the alarm reduction rates calculated by the method are 48.4%, 45.9%, 35.3%, and 41.4% (47.2%, 44.6%, 36.8%, and 39.9%), respectively. From which, it reflects the method could have some impact on the alarm threshold optimization, bring a lower FAR and fewer alarms.



(c1)



(c2)

Figure 13. Histograms of number of alarms in five cases with two cases (c1,c2).

## 5. Conclusions

In this work, correlation degree and clustering analysis based method is presented to achieve threshold optimization: the gray correlation coefficients of variables are first obtained by analyzing correlation degrees among them; the variables are grouped later according to the correlation degree and clustering analysis, given the weight  $\omega_{ij}$  for FAR in each group; optimization algorithm is finally utilized to optimize objective function about FAR, MAR,  $R_{FAR}$ ,  $R_{MAR}$ , and  $\omega_{ij}$  to complete threshold optimization.

According to the analysis of case theory study with TE simulation process and actual industrial verification for industrial ethylene production process, the results manifest the presented approach can not only reduce FAR, have significant inhibitory effect on MAR, and decrease the number of alarms effectively in total, but could carry on the grouping sorting according to the intrinsic clustering analysis, which could help operators reduce operation load. Meanwhile, it will also leave operators more time to make more efficient repair measures timely and reduce losses through helping them to identify variables that have larger and more rapid impact on system, extend the deteriorative time for abnormality.

**Author Contributions:** Methodology, Z.W. and G.Z.; Writing—original draft preparation, G.Z.; Writing—review and editing, Z.W. and G.Z.; Supervision, Z.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** Support for carrying out this work was provided by the National Key R&D Program of China (2018YFB1701103), International (Regional) Cooperation and Exchange Project (61720106008), and National Science Fund for Distinguished Young Scholars (61725301).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Izadi, I.; Shah, S.L.; Shook, D.S.; Chen, T. An Introduction to Alarm Analysis and Design. *IFAC Proc. Vol.* **2009**, *42*, 645–650. [[CrossRef](#)]
- EEMUA. *Alarm System: A Guide to Design, Management, and Procurement*, 2nd ed.; The Engineering Equipment and Materials Users Association (EEMUA): London, UK, 2007.
- Mezache, A.; Slitani, F. A novel threshold optimization of ML-CFAR detector in Weibull clutter using fuzzy-neural networks. *Signal Process.* **2007**, *87*, 2100–2110. [[CrossRef](#)]
- Naghoosi, E.; Izadi, I.; Chen, T. A study on the relation between alarm deadbands and optimal alarm limits. In Proceedings of the American Control Conference, San Francisco, CA, USA, 29 June–1 July 2011.
- Han, L.; Gao, H.H.; Xu, Y.; Zhu, Q.X. Combining FAP, MAP and correlation analysis for multivariate alarm thresholds optimization in industrial process. *J. Loss Prev. Process Ind.* **2016**, *40*, 471–478. [[CrossRef](#)]
- Levente, L.S.; Konrad, H. Industrial batch dryer data mining using intelligent pattern classifiers: Neural network, neuro-fuzzy and Takagi–Sugeno fuzzy models. *Chem. Eng. J.* **2010**, *157*, 568–578.
- Burgess, L.P.; Herdman, T.H.; Berg, B.W.; Feaster, W.W.; Hebsur, S. Alarm limit settings for early warning systems to identify at-risk patients. *J. Adv. Nurs.* **2009**, *65*, 1844–1852. [[CrossRef](#)]
- Zhou, H.; Ni, Y.; Ko, J. Structural damage alarming using auto-associative neural network technique: Exploration of environment-tolerant capacity and setup of alarming threshold. *Mech. Syst. Signal Process.* **2011**, *25*, 1508–1526. [[CrossRef](#)]
- Chen, T. On reducing false alarms in multivariate statistical process control. *Chem. Eng. Res. Des.* **2010**, *88*, 430–436. [[CrossRef](#)]
- Jablonski, A.; Barszcz, T.; Bielecka, M.; Breuhaus, P. Modeling of probability distribution functions for automatic threshold calculation in condition monitoring systems. *Measurement* **2013**, *46*, 727–738. [[CrossRef](#)]
- Van Wynsberge, S.; Gilbert, A.; Guillemot, N.; Payri, C.; Andréfouët, S. Alert thresholds for monitoring environmental variables: A new approach applied to seagrass beds diversity in New Caledonia. *Mar. Pollut. Bull.* **2013**, *77*, 300–307. [[CrossRef](#)]
- Simşek, B.; İc, Y.T. Multi-response simulation optimization approach for the performance optimization of an Alarm Monitoring Center. *Saf. Sci.* **2014**, *66*, 61–74. [[CrossRef](#)]
- Sun, Y.; Tan, W.; Chen, T. A method to remove chattering alarms using median filters. *ISA Trans.* **2018**, *73*, 201–207. [[CrossRef](#)] [[PubMed](#)]
- Wu, M.-H.; Wang, J.P.; Chen, I.-C. Optimization approach for determining rainfall duration-intensity thresholds for debris flow forecasting. *Bull. Eng. Geol. Environ.* **2018**, *78*, 2495–2501. [[CrossRef](#)]

15. Zhong, X.; Zhao, S.; Deng, H.; Gui, D.; Zhang, J.; Ma, M. Nuisance alarm rate reduction using pulse-width multiplexing  $\Phi$ -OTDR with optimized positioning accuracy. *Opt. Commun.* **2020**, *456*, 124571. [CrossRef]
16. Sun, H.; Li, M.; Zuo, L.; Cao, R. Joint threshold optimization and power allocation of cognitive radar network for target tracking in clutter. *Signal Process.* **2020**, *172*, 107566. [CrossRef]
17. Jaber, S.; Pierre, S.; Jamal, C. A simple and robust approach for early detection of defluidization. *Chem. Eng. J.* **2017**, *313*, 144–156.
18. Venkidasalopathy, J.A.; Mannan, M.S.; Kravaris, C. A quantitative approach for optimal alarm identification. *J. Loss Prev. Process. Ind.* **2018**, *55*, 213–222. [CrossRef]
19. Aslansefat, K.; Gogani, M.B.; Kabir, S.; Shoorehdeli, M.A.; Yari, M. Performance evaluation and design for variable threshold alarm systems through semi-Markov process. *ISA Trans.* **2020**, *97*, 282–295. [CrossRef]
20. Kaced, R.; Kouadri, A.; Baiche, K. Designing alarm system using modified generalized delay-timer. *J. Loss Prev. Process. Ind.* **2019**, *61*, 40–48. [CrossRef]
21. Lucke, M.; Chioua, M.; Grimholt, C.; Hollender, M.; Thornhill, N.F. Integration of alarm design in fault detection and diagnosis through alarm-range normalization. *Control. Eng. Pract.* **2020**, *98*, 104388. [CrossRef]
22. Taheri-Kalani, J.; Latif-Shabgahi, G.; Shoorehdeli, M.A. On the use of penalty approach for design and analysis of univariate alarm systems. *J. Process. Control.* **2018**, *69*, 103–113. [CrossRef]
23. Wang, J.; Li, H.; Huang, J.; Su, C. A data similarity based analysis to consequential alarms of industrial processes. *J. Loss Prev. Process. Ind.* **2015**, *35*, 29–34. [CrossRef]
24. Goel, P.; Pistikopoulos, E.; Mannan, M.; Datta, A. A data-driven alarm and event management framework. *J. Loss Prev. Process. Ind.* **2019**, *62*, 103959–103973. [CrossRef]
25. Spross, J.; Gasch, T. Reliability-based alarm thresholds for structures analysed with the finite element method. *Struct. Saf.* **2018**, *76*, 174–183. [CrossRef]
26. Wang, J.; Yang, F.; Chen, T.; Shah, S.L. An Overview of Industrial Alarm Systems: Main Causes for Alarm Overloading, Research Status, and Open Problems. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 1045–1061. [CrossRef]
27. Zhang, G.; Wang, Z.; Mei, H. Sensitivity clustering and ROC curve based alarm threshold optimization. *Process. Saf. Environ. Prot.* **2020**, *141*, 83–94. [CrossRef]
28. Cai, S.; Zhang, L.; Palazoglu, A.; Hu, J. Clustering analysis of process alarms using word embedding. *J. Process. Control.* **2019**, *83*, 11–19. [CrossRef]
29. Lucke, M.; Chioua, M.; Grimholt, C.; Hollender, M.; Thornhill, N.F. Advances in alarm data analysis with a practical application to online alarm flood classification. *J. Process. Control.* **2019**, *79*, 56–71. [CrossRef]
30. Wang, J.; Chen, T. An online method for detection and reduction of chattering alarms due to oscillation. *Comput. Chem. Eng.* **2013**, *54*, 140–150. [CrossRef]
31. Chakraborty, D.; Elzarka, H. Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold. *Energy Build.* **2019**, *185*, 326–344. [CrossRef]
32. Xu, D.L.; Xu, H.Y.; Hu, Y.Z.; Li, J.N. Evidence updating with static and dynamical performance analyses for industrial alarm system design. *ISA Trans.* **2020**, *99*, 110–122. [CrossRef]
33. Zhu, J.; Shu, Y.; Zhao, J.; Yang, F. A dynamic alarm management strategy for chemical process transitions. *J. Loss Prev. Process Ind.* **2014**, *30*, 207–218. [CrossRef]
34. Qi, X.-G.; Wang, H.; Liu, Y.; Chen, G. Flexible alarming mechanism of a general GDS deployment for explosive accidents caused by gas leakage. *Process. Saf. Environ. Prot.* **2019**, *132*, 265–272. [CrossRef]
35. Yu, C.; Wang, H.; Yao, J.; Zhao, J.; Sun, Q.; Zhu, H. A dynamic alarm threshold setting method for photovoltaic array and its application. *Renew. Energy* **2020**, *158*, 13–22. [CrossRef]
36. Xu, J.; Wang, J.; Izadi, I.; Chen, T. Performance Assessment and Design for Univariate Alarm Systems Based on FAR, MAR, and AAD. *IEEE Trans. Autom. Sci. Eng.* **2012**, *9*, 296–307. [CrossRef]
37. Tian, W.; Zhang, G.; Liang, H. Alarm clustering analysis and ACO based multi-variable alarms thresholds optimization in chemical processes. *Process. Saf. Environ. Prot.* **2018**, *113*, 132–140. [CrossRef]
38. Abdou, L.; Taibaoui, O.; Moumen, A.; Ahmed, A.T. Threshold optimization in distributed OS-CFAR system by using simulated annealing technique. In Proceedings of the 2015 4th International Conference on Systems and Control (ICSC), Sousse, Tunisia, 28–30 April 2015; Volume 4, pp. 295–301.
39. Salleh, F.H.M.; Arif, S.M.; Zainudin, S.; Firdaus-Raih, M. Reconstructing gene regulatory networks from knock-out data using Gaussian Noise Model and Pearson Correlation Coefficient. *Comput. Biol. Chem.* **2015**, *59*, 3–14. [CrossRef]
40. Amin, T.; Khan, F.; Imtiaz, S.A.; Ahmed, S. Robust Process Monitoring Methodology for Detection and Diagnosis of Unobservable Faults. *Ind. Eng. Chem. Res.* **2019**, *58*, 19149–19165. [CrossRef]
41. Hu, C.; Xu, Z.; Kong, X.; Luo, J. Recursive-CPLS-Based Quality-Relevant and Process-Relevant Fault Monitoring With Application to the Tennessee Eastman Process. *IEEE Access* **2019**, *7*, 128746–128757. [CrossRef]
42. Xiao, D.H. Optimization Approached to Multi-Variable Alarm Thresholds with Priorities in Process Productions. Master's Thesis, Beijing University of Chemical Technology, Beijing, China, 2014. Available online: <https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMTEyMDESCFkyNjI4NzI1Gghla3J2a3U2cw%253D%253D> (accessed on 26 November 2021).