A New Fuzzy Process Capability Estimation Method Based on Kernel Function and FAHP

Zhiqiang Geng, Zun Wang, Chenglong Peng, and Yongming Han

***Abstract*—Because of more and more complexity of an operation environment in today's industrial production process, it is difficult to monitor the process operation quality and to estimate the performance efficiency based on the existing mathematical model and knowledge. This paper proposes a new method to estimate the process capability, and a new criterion for capability and performance assessment. This method is based on kernel function and fuzzy analysis hierarchy process (FAHP), which can improve the adaptation of process capability analysis. The device process capability can be estimated by FAHP with main variables, which are determined by interpretive structure modeling. The estimators of these indices overcome uncertainties caused by data fluctuation in the traditional process capability, and could strongly improve the robustness and adaptability of the process capability estimation and diagnosis. The proposed methods are used in a simulation of the Tennessee Eastman process. The results demonstrate the efficiency and validity of the presented approach. The proposed method can provide more performance decision information of industrial process to help decision makers evaluate and diagnose the state of the production devices, and improve the process operations.**

***Index Terms*—Fuzzy analysis hierarchy process, fuzzy decision, fuzzy process capability, interpretive structure modeling, kernel function.**

I. Introduction

In THE industrial production, the quantification of process variation and location to estimate process capability is important for understanding the quality of process production. For hundreds of process variables in a modern complex device, it is hard to use traditional statistics or mathematical modeling to describe or represent the process production. However, the process capability indices (PCIs), which are unitless, could create a relationship between the process mean and standard deviation with the engineering specifications. These provide a common method for quantifying the process [1]–[3]. The relationship is made by calculating the ratio of the spread between the process specification and the spread of the process value [4]. The most common outputs of process capability analysis are PCIs, which are measured to judge whether it can coincidentally meet customer expectations [5], [6].

The PCIs have become very universal in assessing the capability in lots of fields, and some indices are accessible in different literature works such as [7]–[9]. More and more people have made efforts to study and use them to guide the process production. For example, the PCIs for program planning and growth can be used to enhance product development [10]. The PCIs with C_p and C_{pk} have been used in many automotive industry productions such as Ford Motor Company, etc. [11]. From 1992 to 2000, Kotz and Johson provided elaborate study with interpretations in about 170 publications on PCIs [12]. Ravichandran and Rai had also made the process capability for the software engineering and knowledge management [13].

As the traditional process analysis indicated, the process should be in control and the indices should follow the normal distribution. If the process is out of control, it would fail to display the historical conformity degree well, and the capability of the process would be useless in the future [14]. However, there are conditions that this hypothesis may not be appropriate. For example, sometimes the directional deviations and positional deviations may not be normal, so it is essential to see the aspects of capability indices when the process is nonnormal [15]. The most common techniques used to handle the nonnormal distribution data are mathematical transformation and quantile estimation [16]. The artificial neural network is proposed to estimate PCIs for right skewed distributions without appealing to the probability density function (PDF) of the process [17]. A generalized PCI, defined as the ratio of proportion of specification conformance to proportion of desired conformance, has also been developed [18]. It can be assessed under either unilateral or bilateral specifications. Chang and Lu [19] and Pearn and Chen [20] also made the studies that utilized the percentiles as estimation for the process mean and the standard deviation, which can be used to calculate the PCIs for the distributions of any shape. And these methods are more accurate than the normal PCIs when the quality characteristics possess a specific nonnormal distribution. Sundaraiyer had made studies about the interval estimation of Clements' process capability when the underlying distribution was inverse Gaussian [21]. The root transformation had also been used for the nonnormal process, which did not require to fit the PDF to the data [22], [23]. In the existing methods for dealing with nonnormal processes, the PDF of the process is required. It is difficult to estimate the PDF of the process, and the PCIs estimated by PDF may be far from real values.

Studies of process capability measurements are usually based on the precise process data of manufacture. But sometimes the measurements of product quality are incompletely precise.

Manuscript received May 14, 2015; revised November 16, 2015 and September 9, 2015; accepted December 29, 2015. Date of publication January 29, 2016; date of current version April 15, 2016. Review of this manuscript was arranged by Department Editor B. Jiang. This work was supported in part by the National Natural Science Foundation of China under Grant 61374166 and Grant 61533003, by the Fundamental Research Funds for the Central Universities under Grant YS1404, ZY1502, and by the Doctoral Fund of Ministry of Education of China under Grant 20120010110010.

The authors are with the College of Information Science and Technology, and the Engineering Research Center of Intelligent Process System Engineering, Ministry of Education, Beijing University of Chemical Technology, Beijing 100029, China (e-mail: gengzhiqiang@mail.buct.edu.cn; wz_txdz@163.com; elsonpeng@gmail.com; hanym@mail.buct.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEM.2016.2517337

Uncertainty factors are familiar in the production, for example, observations with coarse scales, measurement errors, sample data estimation, and so on. Intervals can be used to solve the problem when the measured values of equipment are not precise, but only a finite number of decimals are useable. For these reasons, the concept of fuzzy PCIs is proposed, and the fuzzy sets bring benefits to the flexible definition and evaluation to them [24]. Wu [25] presented a set of confidence intervals that produces triangular fuzzy numbers (TFNs) for the estimation of C_{pk} index using Buckley's approach with some modifications. Additionally, a three-decision testing rule and step-by-step procedure are developed to assess process performance based on fuzzy critical values and fuzzy p-values. Kaya and Kahraman [26] proposed the fuzzy PCIs by using fuzzy process mean and fuzzy variance with fuzzy specification limits (SLs), which has provided more information and more sensitiveness on PCIs. The fuzzy values of process mean and variance have been produced by the estimation theory based on confidence intervals. SLs have also been defined as TFNs to increase the flexibility of process capability analysis. The PCIs are obtained by linguistic or approximate values. To check if a statistical control exists, fuzzy control charts are produced, and fuzzy PCIs are derived by Kaya and Kahraman [27]. If the process is not under control, then its parameters are unstable, and the values of these parameters in the future are uncertain. The fuzzy PCIs with asymmetric tolerances also have been studied [28], which have been obtained by using fuzzy SLs, variance, mean, and target values. Parchami and Mashinchi [29] proposed an algorithm based on Buckley's estimation approach and used a family of confidence intervals to estimate PCIs C_p, C_{pk}, and C_{pm}. The thus-obtained estimators of these indices are triangular-shaped fuzzy numbers.

The traditional process capability analysis is generally used for manufacturing production process. When process capability analysis is applied directly to the process industry, it is difficult to accurately evaluate the production process in consideration of the differences of technology, etc. This paper concentrates on the fuzzy process capability with nonnormal process data. The PCIs based on the kernel function are developed under fuzzy environment for the process industry. Meanwhile, a three-layer hierarchy model based on fuzzy analysis hierarchy process (FAHP) is put forward to provide an extensive assessment for devices. The method can provide more information comprehensively based on the PCIs to get more sensitiveness and flexibility analysis. We can not only know the integrated state of the device, but also get the detailed information of process variables.

II. Basics of Process Capability

The process capability analysis is one of the main techniques in statistical process control fields. It is mainly used to describe whether the ability of industrial process meets the expectations of the production. It is a process to measure the consistency and represent manifestation of the smallest fluctuations in steady state. The fluctuation of the process distribution is formed to indicate the standard deviation σ [3]. For the given output of the process, it can assess the ability of the quality characteristics of the process through the measurement and collection of the process data. And the analysis results of the process capability can be used to improve the process operation and decision.

Process capability analysis mainly uses the PCIs to recognize the capability when the production process is in a normal distribution under statistical control. The PCIs are calculated by several ways that include the following major elements [3], [5]:

C_p is defined as the process capability, only for statistical stabilization process [3]. It is calculated only for bilateral tolerance in the following equation:

$$C_p = \frac{\mathrm{USL} - \mathrm{LSL}}{6\sigma} \tag{1}$$

where USL and LSL separately represent the upper and lower SLs of tolerance, and the standard deviation σ is usually estimated by R-bar/d2 or s-bar/c4.

C_{pk} represents the PCI when the process has the shift situation, but the premise is that the process is stable and the data which should be more than 20 groups are normally distributed. Therefore, C_{pk} is called as the potential process capability in the following equation [3]:

$$C_{\mathrm{pk}} = \min\left\{\frac{\mathrm{USL} - \mu}{3\sigma}, \frac{\mu - \mathrm{LSL}}{3\sigma}\right\} \tag{2}$$

where μ is the process mean.

P_p is process performance, describing the actual process capability in the long-running process without considering whether controlled or not [5], which is calculated as

$$P_p = \frac{\mathrm{USL} - \mathrm{LSL}}{6s} \tag{3}$$

where *s* represents the standard deviation of the data *N*.

P_{pk} is the process performance index. It is required in the product manual production part approval procedure when the manufacturing process is a trial and unstable measure process. P_{pk} is always used for its capability assessment. When $P_{\mathrm{pk}} >= 1.67$, the process can reach the mass production stage. Therefore, P_{pk} is also initial capacity index [5]. P_{pk} comes from the minimum of P_{pu} and P_{pl} as

$$P_{\mathrm{pk}} = \min\{P_{\mathrm{pu}}, P_{\mathrm{pl}}\} = \min\left\{\frac{\mathrm{USL} - \mu}{3s}, \frac{\mu - \mathrm{LSL}}{3s}\right\}. \tag{4}$$

The greater value of the PCI, the smaller degree of dispersion relative to the product technology standard tolerances and the larger process capability. Meanwhile, the smaller value of the PCI, the bigger relative dispersion degree and the lower the process capability. Thus, the capability level can be determined by the PCIs. Considering of both economic and quality requirements, the PCI is not always good when the value is large, and it should be taken within the appropriate range.

III. Fuzzy Numbers and Its Operations

A. Triangular Fuzzy Number

For the real world, many descriptions or concepts cannot be clearly defined. According to the analysis, observation, and research of such problems, it is difficult to identify the precise rules. In this aspect, fuzzy set theory has much advantage to solve these problems. It is a very beneficial tool for describing the objective world and such obscure boundary things. It provides a structured formulaic approach that can solve the problem in a fuzzy environment to make a clear decision. And now vague concepts have been widely applied to various disciplines.

For TFN of fuzzy set $\tilde{A} = (a, b, c)$, $a \le b \le c$, and the membership function is obtained as

$$\mu_A = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ \dfrac{c-x}{c-b}, & b \le x \le c \\ 0, & x > c. \end{cases} \tag{5}$$

Let $\tilde{A} = (a_1, b_1, c_1)$ and $\tilde{B} = (a_2, b_2, c_2)$ be any two TFNs. The operations laws are obtained in the literature [30].

B. Satisfaction With Fuzzy Numbers

Definition 1: Suppose that there are two TFNs $\tilde{A} = (a_1, b_1, c_1)$ and $\tilde{B} = (a_2, b_2, c_2)$, and then, the intersection of the membership function of $\tilde{A}$ and $\tilde{B}$ can be calculated as follows:

$$\mu_A = \begin{cases} 1 - \dfrac{b_2 - b_1}{c_1 + a_2}, & 0 \le b_2 - b_1 \le c_1 + a_2 \\ 1 - \dfrac{b_1 - b_2}{a_1 + c_2}, & 0 \le b_1 - b_2 \le a_1 + c_2 \\ 0, & \text{others.} \end{cases} \tag{6}$$

If there is an intersection in two TFNs, (6) can reflect the relationship between two TFNs which have only one point of intersection and it is also the center point of two TFNs. Based on the membership values, the satisfaction formula can be used to compare the fuzzy numbers.

Definition 2: Suppose two TFNs $\tilde{A} = (a_1, b_1, c_1)$, $\tilde{B} = (a_2, b_2, c_2)$, and the satisfaction of $\tilde{A} \le \tilde{B}$ is defined as follows [31]:

$$P_n(\tilde{A} \le \tilde{B}) = \begin{cases} 1 - \alpha_*^n \dfrac{c_1 + a_2}{a_1 + c_1 + a_2 + c_2}, & \tilde{A} \le \tilde{B} \\ \alpha_*^n \dfrac{c_2 + a_1}{a_1 + c_1 + a_2 + c_2}, & \tilde{A} \ge \tilde{B}. \end{cases} \tag{7}$$

For any two TFNs $\tilde{A}$ and $\tilde{B}$, there are the following properties.

1) $0 < P_n(\tilde{A} \le \tilde{B}) \le 1$.
2) When $\tilde{A} = \tilde{B}$, then $P_n(\tilde{A} \le \tilde{B}) = 1 - \frac{c_1 + a_2}{c_1 + a_1 + c_2 + a_2}$.
3) When $b_1 + c_1 \le b_2 - a_2$, $P_n(\tilde{A} \le \tilde{B}) = 1$.
4) $P_n(\tilde{A} \le \tilde{B}) \le 1 - P_n(\tilde{B} \le \tilde{A})$,
   where *n* is a constant and the increase of *n* can more effectively distinguish between TFNs.

It can be generally considered as $n = 3$. The large value of *n* is not the best case, and we should take the real problems into account and get a reasonable amount of computation to select.

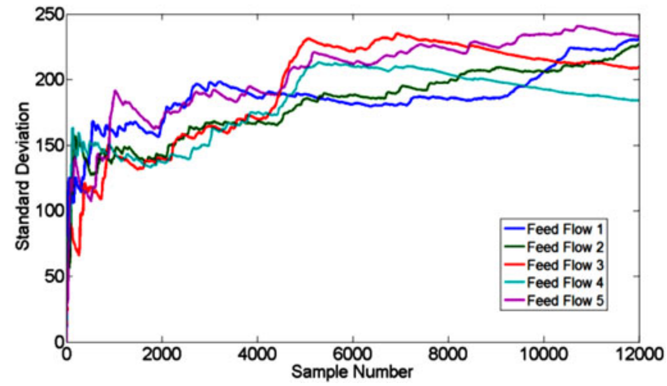

Fig. 1. Standard deviation with the data length incremented trends.

IV. Fuzzy Process Capability Based on Kernel Function

A. Selection of Data Length

For practice, the selection of the data length directly determines the assessment of process capability and process performance, and a reasonable length of data can ensure accurate evaluation for process capability and process performance.

Thornhill gave the selected proposals of the control loop sample length on level, flow, pressure, etc. [32]. Combining with the actual operation of real process variables, we select different lengths of steady-state data repeatedly by comparing with its standard deviation. When the data reach a certain length of samples, the standard deviation will have small fluctuations, and then, we can take the length of the data, which is as shown in Fig. 1. Wherein the abscissa represents the length of the selected data which is incremented from left to right, and the vertical axis represents the standard deviation value.

By integrating experiments with literature suggestions on the sample number, we can choose a suitable sample number and additionally ensure that the process is in a steady state, which cannot have stop or switching status. As shown in Fig. 1, when the data length of flow is less than 1400, the standard deviation has relatively large changes. However, if it is more than 1400, the standard deviation will have small changes. In [32], the suggestion flow data length is 1500. So we make the decision of the experiment as 1500 or so. Similarly, the data length of other data types can be obtained.

B. Definition of Dynamic SLs

In the manufacturing production, the calculation of process capability always provides the upper SL and the lower SL of the product. However, in the process industry, the measurement range of the process variable is often much larger than the upper SL. So if we still use the measurement range of data calculation, most of the results that come from the data are inaccurate. Therefore, the value of the specification needs to be estimated based on the recorded data, and then, we can dynamically get the upper and the lower SLs of the samples.

The selection of the data length should be long enough, stable, and controllable. Here, we take the sample number as *N* and

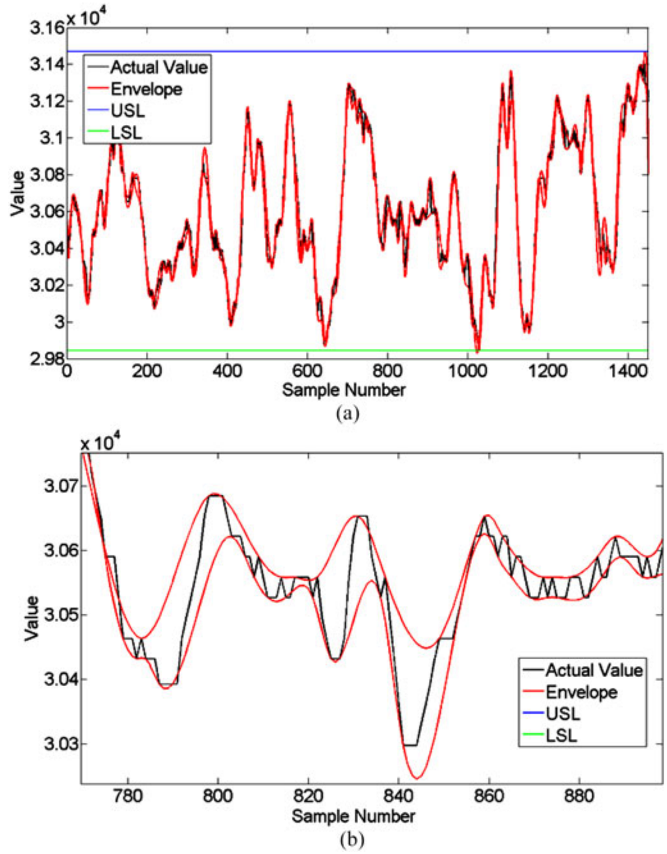

Fig. 2. Schematic cubic spline limits on specifications and its partially enlarged view.

choose continuous M groups (M take 10 to 15). By using the cubic splines [33] on the N data, we can get the upper and the lower envelope interpolated values by the extreme points. We take the maximum of the upper envelope as USL and the minimum of the lower envelope as LSL. K-means clustering algorithm is used to get the clustering centre value u_2 for the obtained M groups of USLs. Then, we get its nearest USL to u_2 as u_1, which is smaller than u_2, and the nearest USL to u_2 as u_3, which is bigger than u_2. Similarly, we can get l_1, l_2, and l_3. As shown in Fig. 2(a), the middle line represents the actual data and there are two envelope lines on its upper and lower sides. The upper line and the lower line represent USL and LSL, respectively. An enlarged partial schematic view of the envelope lines is shown in Fig. 2(b). Then, we can get the fuzzy upper and lower SLs

$$\tilde{\mathrm{USL}} = (u_1, u_2, u_3), \tilde{\mathrm{LSL}} = (l_1, l_2, l_3). \tag{8}$$

C. Fuzzy Process Capability Method Based on Kernel Function and Evaluation Criterion

The kernel method is an effective way to solve nonlinear problems. The kernel function is mainly used in the kernel method, and its structure and parameters play a decisive role on the effect of the kernel method. The main role of the kernel function is to transform the process source data. By the sample data mapped to the high-dimensional feature space, a new feature vector is created, which can achieve the conversion form the nonlinear solving method of original input space to the linear solving method of feature space [34].

The parameters and the structure of kernel function are not restricted by data dimension of feature space. Combined with the kernel function, numerous nonlinear methods have got better improvements of performance and effectiveness for the original nonlinear processing, which brings many innovative algorithms for nonlinear processing problems. The selection of kernel functions takes as a standard whether it is the symmetric function of Mercer condition [35]. The commonly used kernel functions include Gaussian kernel function, Perceptron kernel function, polynomial kernel function, exponential kernel function, and so on.

Considering each group of N data $x_1, x_2, \ldots, x_N$, the kernel function can be established as

$$\hat{f}(x) = \frac{1}{nh_n} \sum_{i=1}^{n} k\left(\frac{x - x_i}{h_n}\right) \tag{9}$$

where $f(x)$ is the kernel density estimation and $k(x)$ is an optional kernel function. Generally, the Gaussian kernel function is chosen for calculation based on the experience and the literature, as

$$k(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \tag{10}$$

where h_n is the bandwidth of process variable, which is generally calculated as $h_n = 1.06\, sN^{-\frac{1}{5}}$, and s represents the standard deviation of the data N.

We can get the estimation of mode m of the given number by the function:

$$\hat{f}(m) = \sup_x \hat{f}_n(x). \tag{11}$$

It is clear that $\hat{f}(m_n) > 0$, and there exists $c \in (0, \hat{f}(m_n))$. There are at least two points x_n^* and y_n^*, making $\hat{f}(x_n^*) = \hat{f}(y_n^*) = c$, where x_n^* is the minimum value of the axis and y_n^* is the maximum value of the axis. Then, we can get (x, y) by the bisection method [23].

For the M sets of the data, we can get $x_1, x_2, \ldots, x_M$ and $y_1, y_2, \ldots, y_M$. They all can be divided into three categories by K-means [36] clustering algorithm, and then, we can take their cluster centers, respectively, as (x_1, x_2, x_3) and (y_1, y_2, y_3).

Then, the fuzzy PCIs based on the kernel function are as follows:

$$\begin{aligned}
\tilde{C}_p &= \frac{\tilde{\mathrm{USL}} - \tilde{\mathrm{LSL}}}{\tilde{d}} = \left(\frac{u_1 - l_3}{y_3 - x_1}, \frac{u_2 - l_2}{y_2 - x_2}, \frac{u_3 - l_1}{y_1 - x_3}\right) \\
\tilde{C}_{\mathrm{pl}} &= \left(\frac{m_1 - l_3}{m_1 - x_1}, \frac{m_2 - l_2}{m_2 - x_2}, \frac{m_3 - l_1}{m_3 - x_3}\right) \\
\tilde{C}_{\mathrm{pu}} &= \left(\frac{u_1 - m_3}{y_3 - m_3}, \frac{u_2 - m_2}{y_2 - m_2}, \frac{u_3 - m_1}{y_1 - m_1}\right) \\
\tilde{C}_{\mathrm{pk}} &= \min\left\{C_{\mathrm{pu}}, C_{\mathrm{pl}}\right\}.
\end{aligned} \tag{12}$$

TABLE I
EVALUATION REFERENCE TABLE OF FUZZY PCI

Range	Diagnosis	Measures
$\widetilde{C}_p > (1.67, 1.67, 1.67)$	Excess process capability	May be appropriate to consider the requirements of the economy
$(1.67, 1.67, 1.67) \geq \widetilde{C}_p >$ $(1.33, 1.33, 1.33,)$	Adequate process capability	For non-primary variables and indicators could relax border operations, reduce costs
$(1.33, 1.33, 1.33,) \geq \widetilde{C}_p >$ $(1.00, 1.00, 1.00)$	Acceptable process capability	In order to detect abnormal fluctuations and take measures to eliminate it, using data envelopment methods to strengthen the supervision of the production process and control it
$(1.00, 1.00, 1.00,) \geq \widetilde{C}_p >$ $(0.67, 0.67, 0.67)$	Inadequate process capability	Find out the real reason for lack of process capability, to develop measures to improve
$(0.67, 0.67, 0.67) \geq \widetilde{C}_p$	Serious shortage of process capability	Should stop production generally, identify the causes, improve the process and increase the $\widetilde{C}_p$ value

TABLE II
RELATIONSHIP BETWEEN $\widetilde{C}_p$ AND $\widetilde{C}_{\mathrm{pk}}$

		$\widetilde{C}_p$	
		Inadequate	adequate
$\widetilde{C}_{\mathrm{pk}}$	inadequate	Lack ability, need to improve and adjust device for getting good results	Adequate process capability, but the average position of the process needs to adapt
	adequate	Impossible situation	Process capability is sufficient and have better adaptability

For the fuzzy process capability calculated by the fuzzy numbers, the precise process capability criterion is made fuzzy [37], [38]. We can get the number of the corresponding fuzzy PCI criterion as shown in Table I. From (12), it can be found that in $\widetilde{C}_p$, the value $(y - x)$ is used to estimate 6σ, and $\widetilde{C}_{\mathrm{pk}}$ is close to the target of the actual value. When the average coincides with the center value, $\widetilde{C}_p = \widetilde{C}_{\mathrm{pk}}$ will appear. In practice, the relationship between the two [38] can be represented in Table II.

D. Fuzzy Process Performance Index and Decisions

Industrial process performance index reflects the degree that the process performance meets the standards and specifications. The instantaneous and real-time process is described as P_p. Therefore, the analysis of the process performance is typically to be calculated by a relatively short period of the data samples.

First, we get the mean value of the overall data as μ_2. Second, we make the number of selected group minimum average value as μ_1 and the maximum average value as μ_3. Then, we can obtain the fuzzy average value as

$$\overset{\cong}{x} = (\mu_1, \mu_2, \mu_3). \tag{13}$$

For the overall fuzzy standard deviation of process performance, the standard deviation for the fuzzy number can be achieved with the same method

$$\tilde{\sigma} = (s_1, s_2, s_3). \tag{14}$$

Then, we can get the process performance index P_p

$$P_p = \frac{\mathrm{U\tilde{S}L} - \mathrm{L\tilde{S}L}}{6\tilde{\sigma}} = \left(\frac{u_1 - l_3}{6s_3}, \frac{u_2 - l_2}{6s_2}, \frac{u_3 - l_1}{6s_1}\right). \tag{15}$$

P_{pk} is used to measure the performance of the process unit, defined as tolerance width divides standard deviation. It is the estimate of the current process performance, which can reflect the total variation of the process. Further, we can carry on the fuzzy process performance index calculation, as

$$\begin{aligned}\tilde{P}_{\mathrm{pu}} &= \left(\frac{u_1 - \mu_3}{3s_3}, \frac{u_2 - \mu_2}{3s_2}, \frac{u_3 - \mu_1}{3s_1}\right)\\ \tilde{P}_{\mathrm{pl}} &= \left(\frac{\mu_1 - l_3}{3s_3}, \frac{\mu_2 - l_2}{3s_2}, \frac{\mu_3 - l_1}{3s_1}\right)\\ \tilde{P}_{\mathrm{pk}} &= \min(\tilde{P}_{\mathrm{pu}}, \tilde{P}_{\mathrm{pl}}).\end{aligned} \tag{16}$$

In order to find out the real cause of insufficient industrial process production capacity, by referring to these papers [5], [37], [39], [40], we can get the decision matrix of the comparison between $\tilde{P}_p$ and $\tilde{P}_p k$ as shown in Table III.

E. ISM Method for Determining the Main Parameters of the System

The ISM method is used to identify and improve the relationship between the different parameters in a system. By creating multiple layers of the relationship between the parameters inherent hierarchical structure of the model, it can be used to analyze the relationship between system elements and distinguish the key factors on the overall system.

First, we need to calculate the correlation coefficient between each two parameters. Taking the *i*th and *j*th parameters as examples, the correlation coefficient between them can be calculated as

$$r_{xy} = \frac{\sum_{i=1}^{M} (x_i - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{i=1}^{M} (x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{M} (y_j - \bar{y})^2}}. \tag{17}$$

TABLE III
DECISION MATRIX OF PERFORMANCE, CAPABILITY, AND QUALITY

Decision Matrix of Performance, capability and quality		The performance and capabilities match? No $\left(\tilde{P}_p < \tilde{P}_{\mathrm{pk}}\right)$	Yes $\left(\tilde{P}_p \geq \tilde{P}_{\mathrm{pk}}\right)$
The ability and quality match?	No $\left(\tilde{P}_{\mathrm{pk}} < (1.0, 1.0, 1.0)\right)$	Improve process equipment, and control loop performance	Improve process equipment
	Yes $\left(\tilde{P}_{\mathrm{pk}} > (1.0, 1.0, 1.0)\right)$	Improve control loop performance	The process equipment and control loop performance come to optimal operation

Then, we can get the correlation coefficient matrix as follows:

$$r = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{\mathrm{mn}} \end{bmatrix}. \tag{18}$$

Inverse correlation matrix is used to obtain partial correlation coefficient between two variables, and then, the partial correlation coefficient of inverse *r* matrix can be calculated as

$$c = \mathrm{inv}(r) = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{\mathrm{mn}} \end{bmatrix}. \tag{19}$$

So, we can obtain the partial correlation coefficient between two variables as

$$R_{\mathrm{ij}} = -\frac{c_{\mathrm{ij}}}{\sqrt{c_{\mathrm{ii}} * c_{\mathrm{jj}}}} \tag{20}$$

among them, $i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$.

After these steps, we can create the following partial correlation coefficient matrix:

$$R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{m1} & R_{m2} & \cdots & R_{\mathrm{mn}} \end{bmatrix}. \tag{21}$$

For each element of this matrix, compare it with the value 0.4. If R_{ij} is positive and bigger than 0.4, then define the adjacency matrix x_i to x_j as 1 and x_j to x_i as 0. If R_{ij} is negative number and its absolute value is bigger than 0.4, then we make the adjacency matrix x_i to x_j as 0 and x_j to x_i as 1. If $i = j$, $R_{\mathrm{ij}} = 1$. Based on these, we can construct the adjacency matrix as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & r_{\mathrm{nn}} \end{bmatrix}. \tag{22}$$

Take the *n*-order identity matrix as

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n*n}. \tag{23}$$

If $A + E = (A + E)^2 = \cdots = (A + E)^{n-1} = (A + E)^n$ exists, then the reachability matrix *R* is as follows:

$$R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{\mathrm{nn}} \end{bmatrix}. \tag{24}$$

For this reachability matrix *R*, in the *i*th row, if $R_{\mathrm{ij}} = 1$ ($j = 1, 2, \ldots, n$), we can add it to the reachability set S_i. In the *j*th column, if $R_{\mathrm{ij}} = 1 (i = 1, 2, \ldots, n)$, we can add it to the first set B_i. Traverse the entire matrix, and make the intersection of the reachability set and first set as $S_i \cap B_j$. Find the highest impact factor of the first stage, and then, remove the row and column from the matrix where the factor is. Similarly, find the rest of the other layers of the factor until identify all of the factors at each level contains. Then, the hierarchical ISM can be established.

F. Fuzzy AHP Method for Assessing the Device Capability

The fuzzy AHP method provides a comprehensive result given by a number of different types of judgment and analysis and reduces to pairwise comparisons of various factors and simple calculations. By analyzing the factors contained in the complex and interrelated issues, we make the problem break down into different elements. And these elements are incorporated into different levels, so as to form a multilevel structure [41].

The industrial process contains lots of parameters. If we just calculate the process capability of each parameter, it may take little effects for the operators to know the state of the device completely. For getting better assessment of the whole device, we establish a three-layer hierarchy model as shown in Fig. 3. The specific steps of the construction of the model are described as follows:

Step 1: Use the ISM model to find the main parameters of the device. Suppose that the parameters are $P_1, P_2, \ldots, P_n$. If the parameters are not a lot, we can consider them all in our analysis.

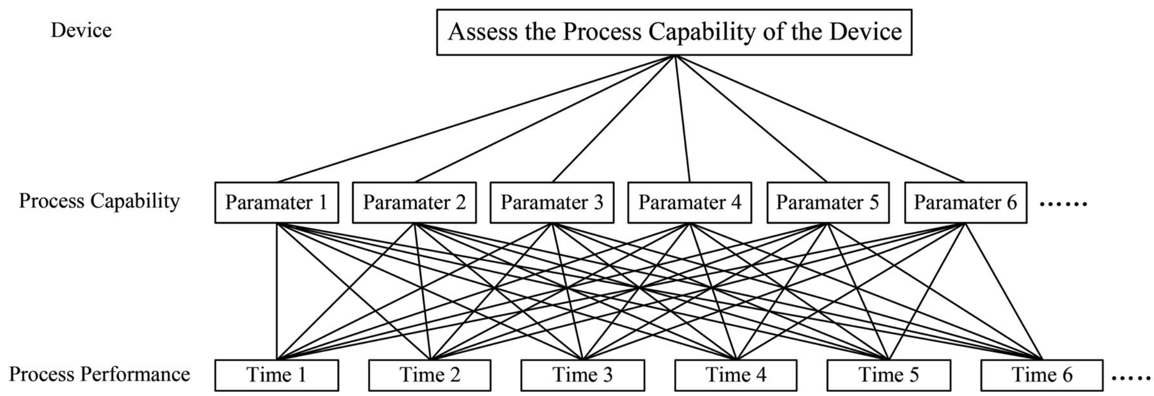

Fig. 3. Three-layer hierarchy model for device.

Step 2: Choose a reasonable length of the parameter by the suggestion of Section IV-A. Then, we can obtain the process performance of each period by the formulation and decisions in Section IV-D. It is called the process performance rank. At this rank, we can get the short-time state of each parameter thoroughly.

Step 3: Calculate the process capability of each parameter in a relatively long-time period by the formulation in Section IV-C. For the *N* parameters, we can get $\tilde{C}_{p1}, \tilde{C}_{p2}, \ldots, \tilde{C}_{\mathrm{pn}}$ and $\tilde{C}_{\mathrm{pk1}}, \tilde{C}_{\mathrm{pk2}}, \ldots, \tilde{C}_{\mathrm{pkn}}$

Then, we can get the matrix of the fuzzy process capability $K_{n\times 1}(C_p)$ and $K_{n\times 1}(C_{\mathrm{pk}})$ as follows:

$$K_{n\times 1}(C_p) = \begin{bmatrix} \tilde{C}_{p1} \\ \vdots \\ \tilde{C}_{\mathrm{pn}} \end{bmatrix} = \begin{bmatrix} \tilde{C}_{p11} & \tilde{C}_{p12} & \tilde{C}_{p13} \\ \vdots & \vdots & \vdots \\ \tilde{C}_{\mathrm{pn1}} & \tilde{C}_{\mathrm{pn2}} & \tilde{C}_{\mathrm{pn3}} \end{bmatrix} \quad (25)$$

$$K_{n\times 1}(C_p) = \begin{bmatrix} \tilde{C}_{p1} \\ \vdots \\ \tilde{C}_{\mathrm{pn}} \end{bmatrix} = \begin{bmatrix} \tilde{C}_{p11} & \tilde{C}_{p12} & \tilde{C}_{p13} \\ \vdots & \vdots & \vdots \\ \tilde{C}_{\mathrm{pn1}} & \tilde{C}_{\mathrm{pn2}} & \tilde{C}_{\mathrm{pn3}} \end{bmatrix}. \quad (26)$$

Step 4: Construct the judgment matrix $A_{n\times n}$ for each fuzzy process matrix. Taking $K_{n\times 1}(C_p)$ as an example, the matrix *R* can be constructed as follows:

$$\begin{aligned} R &= \begin{bmatrix} \sum_{i=1}^{3} C_{p1i} & \sum_{i=1}^{3} C_{p2i} & \cdots & \sum_{i=1}^{3} C_{p3i} \end{bmatrix}^T \\ &= \begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix}. \end{aligned} \quad (27)$$

For each parameter of the matrix $A_{n\times n}$, we can get a_{ij} as follows:

$$a_{\mathrm{ij}} = \begin{cases} \dfrac{r_i - r_j}{r_{\max} - r_{\min}}, & r_i \geq r_j \\ \left(\dfrac{r_j - r_i}{r_{\max} - r_{\min}}(km-1)+1\right)^{-1}, & r_i < r_j \end{cases} \quad (28)$$

where $r_{\max} = \max(r_1, r_2, \ldots, r_n)$, $r_{\min} = \min(r_1, r_2, \ldots, r_n)$, and $km = r_{\max}/r_{\min}$.

Step 5: For the transfer matrix $B_{n\times n}$, $b_{\mathrm{ij}} = \lg a_{\mathrm{ij}}$. The optimal transfer matrix of $B_{n\times n}$ can also be calculated as $C_{n\times n}$, and its each parameter c_{ij} can be obtained as

$$\begin{aligned} c_{\mathrm{ij}} &= \frac{1}{n}\sum_{k=1}^{n}(b_{ik} - b_{jk}) = \frac{1}{n}\sum_{k=1}^{n}(\lg a_{ik} - \lg a_{jk}) \\ &= \lg\left(\frac{\left(\prod_{k=1}^{n} a_{ik}\right)^{1/n}}{\left(\prod_{k=1}^{n} a_{jk}\right)^{1/n}}\right). \end{aligned} \quad (29)$$

We can also get the consistent matrix $D_{n\times n}$, and its parameter

$$d_{\mathrm{ij}} = 10^{c_{\mathrm{ij}}} = \left(\left(\prod_{k=1}^{n} a_{ik}\right)^{1/n}\right) \Big/ \left(\left(\prod_{k=1}^{n} a_{jk}\right)^{1/n}\right). \quad (30)$$

For the *n*-order symmetric matrix $D_{n\times n}$, its eigenvector $W = (w_1, w_2, \ldots, w_n)^T$ is derived through geometric mean method as $w_i = o_i/o$ ($i = 1, 2, \ldots n$), where $o_i = \left(\prod_{j=1}^{n} o_{\mathrm{ij}}\right)^{\frac{1}{n}}$ and $o = \prod_{i=1}^{n} o_i$.

Step 6: As what is shown above, for each matrix $K_{n\times 1}(C_p)$ and $K_{n\times 1}(C_{\mathrm{pk}})$, we can get its own eigenvector $W(C_p)$ and $W(C_{\mathrm{pk}})$.

By making use of the vector to integrate schemes, we can get the integration data $K_{n\times 1}(C_p)$ and $K_{n\times 1}(C_{\mathrm{pk}})$ of the C_p and the C_{pk} values of the process industry device

$$Y(C_p) = \begin{bmatrix} y_1(C_p) \\ y_2(C_p) \\ y_3(C_p) \end{bmatrix}^T = K_{n\times 1}(C_p)^T \times W(C_p) \quad (31)$$

$$Y(C_{\mathrm{pk}}) = \begin{bmatrix} y_1(C_{\mathrm{pk}}) \\ y_2(C_{\mathrm{pk}}) \\ y_3(C_{\mathrm{pk}}) \end{bmatrix}^T = K_{n\times 1}(C_{\mathrm{pk}})^T \times W(C_{\mathrm{pk}}). \quad (32)$$

G. *Process Capability Online Evaluation of the Device*

In practical applications, in order to obtain the evaluation result in time, the online evaluation is required. The proposed method can also be used on line. First of all, we need to prepare the configuration of the proposed algorithm, determine the parameters and the data length *M*, then choose a beginning time point, and let the data length be zero (Len = 0). At last, start recording data as time goes by. If the Len ≥ *M*, then start the

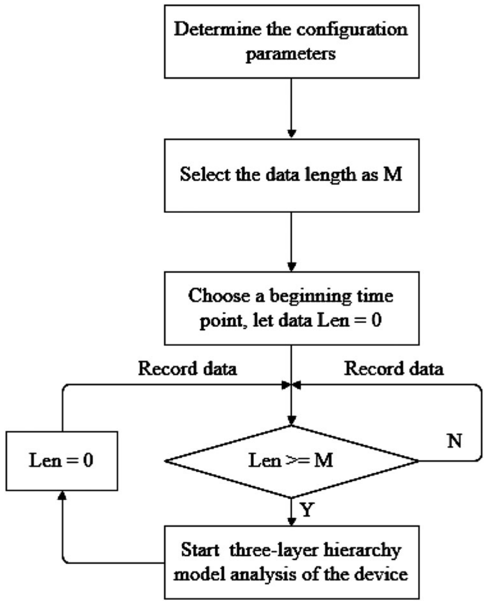

Fig. 4. Flow diagram of online diagnosis.

TABLE IV
FEATURES OF DIFFERENT METHODS INTRODUCED IN THIS PAPER

Method	Characteristics
ISM and FAHP	①Simple and practical decision-making approach ②Subjectivity is too strong, mainly qualitative analysis
Fuzzy PCI	①Relatively objective quantitative analysis, suitable for stable data ②When the data have greater fluctuations, the evaluation results are not accurate
Our Proposed method	①Objective quantitative analysis, obtain the control limits based entirely on data ②Accurate evaluation for fluctuation data

analysis in Section IV-E. Make the Len $= 0$, record data, and go in cycles. The flow diagram is shown in Fig. 4.

From what has been mentioned above, the features and advantages of the proposed method can be fully demonstrated. The method of ISM and FAHP is mainly qualitative analysis to process specific data. Meanwhile, it takes into account the impact of human subjectivity factors, such as the reasonableness of management means and method sets, which can affect the accuracy of analysis evaluation. In Fuzzy PCI [26], [27], for the nonnormality process capability data which have fluctuations, the difference between the upper and lower limits of process control is larger, which leads the evaluation results to be not accurate. The proposed method makes an analysis on the data of objective reality and obtains main parameter variables, which have a decisive impact on the system. Then, the upper and the lower control limits can be obtained by the cubic spline method, so as to ensure that the evaluation of the production process has better adaptability. The specific comparison of methods is shown in Table IV.

V. CASE STUDY

In order to validate the proposed process estimation method, we take the Tennessee Eastman (TE) process system as a research case. The TE process is a simulation software platform of a realistic industrial process, and it is broadly used for process research. The flow diagram is shown in Fig. 5.

In this application, we determine the length of the process variables as 1450, by analysis of the standard deviation trends on the number of different process variables and the data length recommendation by Thornhill *et al.* [32], which are mentioned in Section IV-A.

The TE process has 41 measured variables [42], which are XMEAS1, XMEAS2, ..., XMEAS41. Taking the TE process variable XMEAS10 as an example, the trend is shown in Fig. 6. For Fig. 6(a), (8) is used to get the fuzzy number of the upper and lower SL

$$\mathrm{U\tilde{S}L} = (0.209, 0.215, 0.218), \quad \mathrm{L\tilde{S}L} = (0.182, 0.185, 0.186)$$

Then, by (11), we can get the left and the right margin as $\tilde{x} = (0.185, 0.186, 0.188), \tilde{y} = (0.215, 0.219, 0.219)$ and the fuzzy number of mode *m* as $\tilde{m} = (0.189, 0.190, 0.191)$.

At last, we can get the fuzzy PCIs by (12)

$$\tilde{C}_p = (0.672, 0.914, 1.400), \quad \tilde{C}_{\mathrm{pk}} = (0.640, 0.874, 1.107).$$

Combining with the interval distribution criterion in Table I, choose $\tilde{A} = (1, 1, 1)$, $\tilde{B} = \tilde{C}_p$. By using the TFN comparison method based on the satisfaction function [31], we can get $\alpha_* = 1.051$ and $P_n(\tilde{A} \leq \tilde{B}) = 0.568$.

According to the fuzzy capacity evaluation criterion of both Tables I and II, we can get the following diagnostic conclusions: The process capability of XMEAS10 is acceptable. In order to detect abnormal fluctuations and take measures to eliminate it, data envelopment methods are used to strengthen the supervision of the production process and control it.

In the same way, the process capability of Fig. 6(b) can be calculated as

$$\tilde{C}_p = (0.573, 0.684, 0.875), \quad \tilde{C}_{\mathrm{pk}} = (0.541, 0.624, 0.797).$$

In order to verify the consistency of the diagnosis and the actual condition, we take the period data [see Fig. 6(a)] curve and the period data curve [see Fig. 6(b)] of parameter to make comparison. The capability of the two processes is concluded as follows.

1) Process equipment and control loop performance come to optimal operation.
2) Process equipment constraints exist.

In Fig. 6, the above three dashed lines indicate the fuzzy number of the upper SL, and the following three below dashed lines represent the fuzzy number of the lower SL. It can be found that in the second curve, most of the data values fluctuate too much, less than the SL. It is obvious that the upstream process equipment lacks capabilities, while the first curve data are basically within the SLs and closer to the upper SL, which indicates that the process equipment and the control loop performance are close to the upper limit of the card edge of the operating

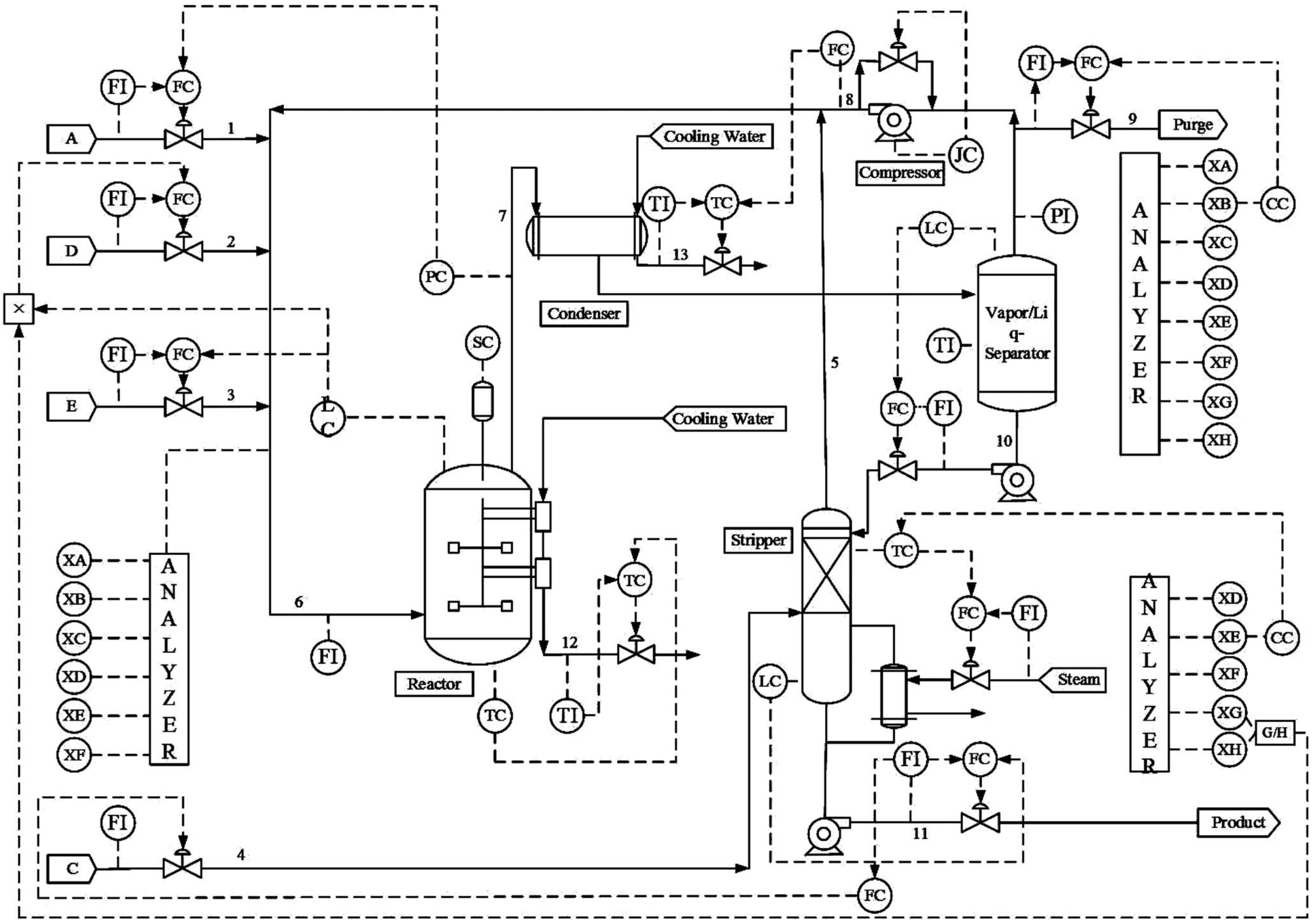

Fig. 5. Flow diagram of the TE process.

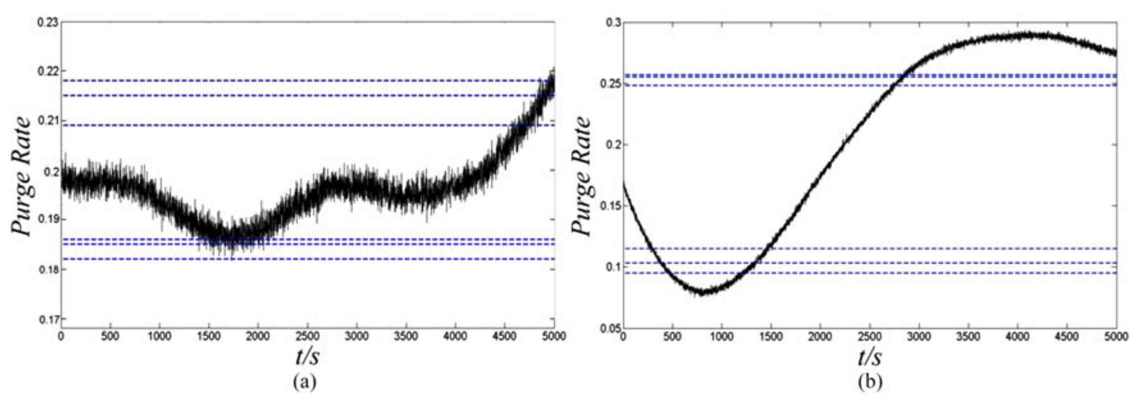

Fig. 6. Different trends of XMEAS10.

state. These validate the diagnostic conclusions consistent with the actual operating conditions.

For the purpose of making comparison with traditional PCIs and fuzzy PCIs [26], [27], XMEAS2 and XMEAS7 are chosen as examples. The data trends are shown in Fig. 7, and then, the process capability analysis can be obtained as shown in Table V.

It can be seen that, when the data trend is smooth, the PCIs of three methods are basically the same, so we can get the same evaluation. However, when the data have fluctuations, the upper and lower control limits obtained by the traditional evaluation method and fuzzy control charts method are far more than the normal range of data, which can lead to greater PCIs and inaccurate evaluation results. The method proposed in this paper has better adaptability for volatility data and ensures the accuracy of the evaluation.

In order to obtain the overall device process capability, we take 22 measured variables for the analysis, including XMEAS1, XMEAS2, ..., XMEAS22. Because XMEAS23, XMEAS24,

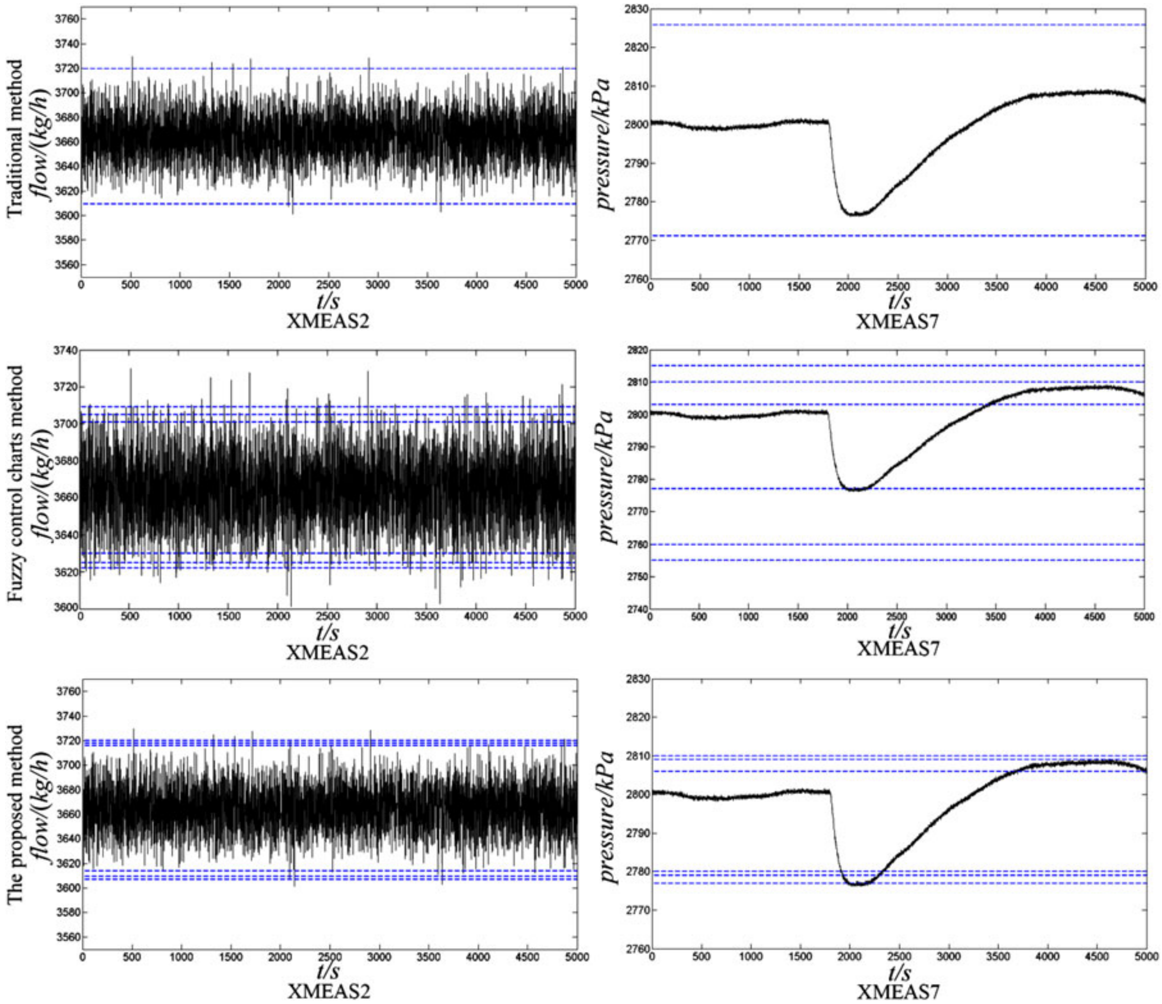

Fig. 7. Data trends of XMEAS2 and XMEAS7.

TABLE V
PCI COMPARISON OF XMEAS2 AND XMEAS7

Variable	Traditional C_{pk}	Fuzzy control charts method	Fuzzy C_{pk} based on kernel funcion
XMEAS2	0.998	(0.968, 0.981, 1.023)	(0.974, 0.988, 1.031)
XMEAS7	1.837	(0.657, 1.117, 1.891)	(0.783, 0.882, 0.959)

TABLE VI
ISM OF TE VARIABLES

Level	Variable Set
L1	XMEAS1, XMEAS2, XMEAS3, XMEAS4, XMEAS5, XMEAS6, XMEAS8, XMEAS9, XMEAS12, XMEAS14, XMEAS15, XMEAS16, XMEAS17, XMEAS19, XMEAS22
L2	XMEAS7, XMEAS10, XMEAS13, XMEAS20, XMEAS21
L3	XMEAS18
L4	XMEAS11

. . . , XMEAS41, are composition variables; we do not take them into consideration. Taking the method as shown in Section IV-E, we can get the reachability matrix *R* as in Fig. 8, and then, we can get its ISMas shown in Table VI.

From Table VI, the variables of Level 1 are the key parameters of this process. We can take them as the presentation of the TE process at this moment. As having got the C_p and C_{pk} of all the Level 1 parameter variables, we can obtain the matrix of fuzzy C_p and C_{pk} as follows:

$$T(C_p) = \begin{bmatrix} 0.668 & 1.008 & 1.493 \\ 0.621 & 0.973 & 1.751 \\ 0.537 & 1.083 & 1.354 \\ 0.476 & 1.021 & 1.279 \\ 0.627 & 1.031 & 1.383 \\ 0.649 & 1.115 & 1.929 \\ 0.549 & 0.874 & 1.344 \\ 0.725 & 1.127 & 1.445 \\ 0.626 & 1.089 & 1.895 \\ 0.704 & 1.162 & 1.515 \\ 0.747 & 1.124 & 1.642 \\ 0.704 & 1.004 & 1.432 \\ 0.464 & 1.356 & 1.398 \\ 0.537 & 1.185 & 1.577 \\ 0.808 & 1.014 & 1.583 \\ 0.351 & 0.977 & 2.013 \end{bmatrix},$$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fig. 8. ISM of TE variables.

$$T(C_{pk}) = \begin{bmatrix} 0.544 & 0.873 & 1.193 \\ 0.371 & 0.774 & 1.251 \\ 0.437 & 0.711 & 0.854 \\ 0.483 & 0.883 & 0.921 \\ 0.588 & 0.782 & 1.183 \\ 0.579 & 1.005 & 1.019 \\ 0.619 & 0.774 & 0.944 \\ 0.525 & 0.927 & 1.113 \\ 0.633 & 0.889 & 1.341 \\ 0.684 & 1.091 & 1.245 \\ 0.683 & 0.994 & 1.132 \\ 0.584 & 0.893 & 1.327 \\ 0.511 & 1.014 & 1.123 \\ 0.339 & 0.965 & 1.061 \\ 0.713 & 0.789 & 1.125 \\ 0.411 & 0.877 & 1.213 \end{bmatrix}.$$

Following the FAHP method, the integrated results are

$$Y(C_p) = \; [T(C_p)]^T \, W(C_p) = [0.615 \quad 1.075 \quad 1.579]$$

$$Y(C_{\mathrm{pk}}) = \; [T(C_{\mathrm{pk}})]^T \, W(C_{\mathrm{pk}}) = \left[0.896 \quad 1.093 \quad 1.275\right].$$

From the decisions in Table I, we can get the diagnosis as acceptable process capability. For further information, we need to check the decision of each parameter variable. In order to detect abnormal fluctuations and take measures to eliminate it, we use data envelopment methods to strengthen the supervision of the production process and control it.

VI. Conclusion

The calculation method of fuzzy PCIs based on kernel function and fuzzy process performance index is proposed in this paper. It can effectively perform in real-time process capability and analyze process performance for the industrial collected data. On the basis of experiments and practical experience, we propose the evaluation method for fuzzy process capability and fuzzy process performance in the process industry. It can provide an objective basis for the evaluation of production process capability and process performance in the continuous process.

The experimental results show that this method can ensure the sensitivity analysis of the production condition results and the consistency with actual working conditions. It is able to effectively ensure that managers can obtain real-time monitoring of industrial processes and make the appropriate decisions based on the results of monitoring analysis. What is more is that the three-layer hierarchy model makes the operators get the state of the device comprehensively.

The proposed method can timely extract the actual plant operating data and effectively convert the data into the operational knowledge. It can be used to guide the device operation by integrating performance evaluation criterion with the decision matrix. It takes effects directly for the realization of the accumulated knowledge from data archiving change and provides an effective way to continuously improve the intelligence level of the operating devices.

References

[1] C. W. Wu, W. L. Pearn, and S. Kotz, "An overview of theory and practice on process capability indices for quality assurance," *Int. J. Prod. Econ.*, vol. 117, no. 2, pp. 338–359, 2009.

[2] K. Rezaie, B. Ostadi, and M. R. Taghizadeh, "Application of process capability and process performance indices," *J. Appl. Sci.*, vol. 6, no. 5, pp. 1186–1191, 2006.

[3] V. E. Kane, "Process capability indices," *J. Qual. Technol.*, vol. 18, no. 1, pp. 41–52, 1986.

[4] D. C. Montgomery, *Introduction to Statistical Quality Control*. Hoboken, NJ, USA: Wiley, 2005.

[5] J. H. Shinde and R. S. Katikar, "Importance of process capability and process performance indices in machine tool," *Int. J. Res. Eng. Appl. Sci.*, vol. 2, no. 2, pp. 1211–1217, 2012.

[6] M. Deleryd, "On the gap between theory and practice of process capability studies," *Int. J. Qual. Rel. Manage.*, vol. 15, no. 2, pp. 178–191, 1998.

[7] S. Basu, P. K. Dan, and A. Thakur, "Experimental design in soap manufacturing for optimization of fuzzified process capability index," *J. Manuf. Syst.*, vol. 33, pp. 323–334, 2014.

[8] S. Kotz and C. R. Lovelace, *Process Capability Indices in Theory and Practice*. London, U.K.: Arnold, 1998.

[9] I. Kaya and C. Kahraman, "Air pollution control using fuzzy process capability indices in the Six-Sigma Approach," *Human Ecol. Risk Assess.*, vol. 15, pp. 689–713, 2009.

[10] W. Z. Huang and Z. Y. Kong, "Process capability sensitivity analysis for design evaluation of multistage assembly processes," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 4, pp. 736–745, Oct. 2010.

[11] V. E. Kane, "Corrigenda," *J. Qual. Technol.*, vol. 18, no. 1, pp. 265–276, 1986.

[12] S. Kotz and N. L. Johnson, "Process capability indices—A review, 1992–2000," *J. Qual. Technol.*, vol. 34, no. 1, pp. 2–53, 2002.

[13] T. Ravichandran and A. Rai, "Structural analysis of the impact of knowledge creation and knowledge embedding on software process capability," *IEEE Trans. Eng. Manage.*, vol. 50, no. 3, pp. 270–284, Aug. 2003.

[14] Z. G. Stoumbos, "Process capability indices: Overview and extensions," *Nonlinear Anal.: Real World Appl.*, vol. 3, pp. 191–210, 2002.

[15] M. Aslam, C. W. Wu, M. Azam, and C. H. Jun, "Variable sampling inspection for resubmitted lots based on process capability index C_{pk} for normally distributed items," *Appl. Math. Model.*, vol. 37, pp. 667–675, 2013.

[16] J. A. Clements, "Process capability calculations for non-normal distributions," *Qual. Progr.*, vol. 22, pp. 95–100, 1989.

[17] A. Abbasi, "A neural network applied to estimate process capability of nonnormal processes," *Expert Syst. Appl.*, vol. 36, pp. 3093–3100, 2009.

[18] S. S. Maiti, M. Saha, and A. K. Nanda, "On generalizing process capability indices," *Qual. Technol. Quant. Manage.*, vol. 7, no. 3, pp. 279–300, 2010.

[19] P. Chang and K. Lu, "PCI calculations with any shape distribution with percentile," *Qual. World Tech. Suppl.*, vol.7, no. 1, pp. 110–114, 1994.

[20] W. L. Pearn and S. Kotz, "Applications of Clements' method for calculating second and third generation process capability indices for non-normal Pearsonian populations," *Qual. Eng.*, vol. 7, no. 1, pp. 139–145, 1994.
[21] V. H. Sundaraiyer, "Estimation of a process capability index for inverse Gaussian distributions," *Commun. Statist.: Theory Methods*, vol. 26, no. 10, pp. 2381–2396, 1996.
[22] S. Z. Hosseinifard, B. Abbasi, S. Ahmad, and M. Abdollahian, "A transformation technique to estimate the process capability index for non-normal process," *Int. J. Adv. Manuf. Technol.*, vol. 40, pp. 512–517, 2009.
[23] T. V. Ramanathan, A.D. Dharmadhikari, and B. Abraham, "Nonparametric capability indices," *Econ. Qual. Control*, vol. 18, no. 1, pp. 31–41, 2003.
[24] I. Kaya and C. Kahraman, "Fuzzy process capability analyses with fuzzy normal distribution," *Expert Syst. Appl.*, vol. 37, pp. 5390–5403, 2010.
[25] C. W. Wu, "Decision-making in testing process performance with fuzzy data," *Eur. J. Oper. Res.*, vol. 193, no. 2, pp. 499–509, 2009.
[26] I. Kaya and C. Kahraman, "Process capability analyses with fuzzy parameters," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11918–11927, 2011.
[27] I. Kaya and C. Kahraman, "Process capability analyses based on fuzzy measurements and fuzzy control charts," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3172–3184, 2011.
[28] I. Kaya and C. Kahraman, "Fuzzy process capability indices with asymmetric tolerances," *Expert Syst. Appl.*, vol. 38, pp. 14882–14890, 2011.
[29] A. Parchami and M. Mashinchi, "Fuzzy estimation for process capability indices," *Inf. Sci.*, vol. 177, pp. 1452–1462, 2007.
[30] Z. Q. Geng, Y. M. Han, X. B. Gu, and Q. X. Zhu, "Energy efficiency estimation based on data fusion strategy: Case study of ethylene product industry," *Ind. Eng. Chem. Res.*, vol. 51, pp. 8526–8534, 2012.
[31] H. Q. Qiu, H. Pan, and J. H. Yang, "Method for comparing triangular fuzzy numbers based on the satisfactory degree function," *J. AnHui Sci. Technol. Univ.*, vol. 24, no. 5, pp. 24–28, 2010.
[32] N. F. Thornhill, M. Oettinger, and P. Fedenczuk, "Refinery-wide control loop performance assessment," *J. Process. Control*, vol. 9, pp. 109–124, 1999.
[33] X. Y. Xu and T. Y. Zhong, "Construction and realization of cubic spline interpolation function," *Autom. Meas. Control*, vol. 25, no. 11, pp. 76–78, 2006.
[34] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[35] B. Üstün, W. J. Melssen, and L. M. C. Buydens, "Facilitating the application of support vector regression by using a universal Pearson VII function based kernel," *Chemometrics Intell. Lab. Syst.*, vol. 281, no. 1, pp. 29–40, 2006.
[36] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient K-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Ana1. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.
[37] S. Steiner, B. Abraham, and J. MacKay. (1997). "Understanding process capability indices," Waterloo, ON, Canada. [Online]. Available: http://www.stats.uwaterloo.ca/~shsteine/papers/cap.pdf
[38] W. L. Pearn and K. S. Chen, "One-sided capability indices C_{pu} and C_{pl}: Decision making with sample information," *Int. J. Qual. Rel. Manage.*, vol. 19, no. 3, pp. 221–245, 2002.
[39] F. Rudisill and L. A. Litteral, "Capability ratios: Comparison and interpretation of short-term and overall indices," *Int. J. Qual. Standards*, vol. 2, pp. 67–86, 2008.
[40] V. Bendic and D. Tilina, "New tendencies in quality management," in *Proc. 7th WSEAS Int. Conf. Eng. Educ.*, Corfu Island, Greece, 2010, pp. 133–136.
[41] S. Zhang, X. Z. Shi, D. S. Gu, and G. H. Huang, "Analysis and evaluation of safety management capability in mine based on ISM and AHP and fuzzy evaluation method," *J. Central South Univ. (Sci. Technol.)*, vol. 42, no. 8, pp. 2406–2416, 2011.
[42] H. H. Chen, P. Tino, and X. Yao, "Cognitive fault diagnosis in Tennessee Eastman process using learning in the model space," *Comput. Chem. Eng.*, vol. 67, pp. 33–42, 2014.

Zhiqiang Geng received the B.Sc. and M.Sc. degrees in process system engineering from Zhengzhou University, China, in 1997 and 2002, respectively, and the Ph.D. degree in control theory and control engineering from the Beijing University of Chemical Technology, Beijing, China, in 2005.

He is currently a Professor with the College of Information Science and Technology, Beijing University of Chemical Technology. His research interests include neural networks, intelligent computing, data mining, knowledge management, and process modeling. He has produced over 80 research papers in his fields.

Zun Wang received the B.Sc. degree in automation and the M.Sc. degree in control engineering from the Beijing University of Chemical Technology, Beijing, China, in 2012 and 2015, respectively, where he is currently working toward the Ph.D. degree.

His research interests include process monitoring and fault diagnosis.

Chenglong Peng received the B.Sc. and M.Sc. degrees in computer science and technology from the Beijing University of Chemical Technology, Beijing, China, in 2012 and 2015, respectively.

His research interests include process capability estimation and data mining.

Yongming Han received the B.Sc. degree in computer science and technology, and the Ph.D. degree in control theory and control engineering from the Beijing University of Chemical Technology, Beijing, China, in 2009 and 2014, respectively.

He is currently a Lecturer with the College of Information Science and Technology, Beijing University of Chemical Technology. His research interests include energy efficiency analysis, neural networks, intelligent computing, data mining and energy saving.