

# Customer preferences Fuzzy Clustering based dynamic multivariate process quality control method

Jingping Yang<sup>1, a</sup>, Wanlei Wang<sup>1, b</sup>, Shikuan Zhou<sup>1, c</sup> and Bin Wu<sup>1, d</sup>

<sup>1</sup>Department of electromechanical engineering, Dalian University of Nationalities, Dalian, China

<sup>a</sup>yjp@dlnu.edu.cn, <sup>b</sup>wwl@dlnu.edu.cn, <sup>c</sup>zsk@dlnu.edu.cn, <sup>d</sup>wb@dlnu.edu.cn

**Keywords:** customer preference. fuzzy c-means. fuzzy clustering. dynamic process quality control

**Abstract.** Faced with fierce competition and various customer demands in marketplaces, manufacturers need to determine the appropriate settings of statistical quality control criterion for dynamic batches composed of similar products with different customer preferences so that the best customer satisfaction and minimized production cost can be obtained. To achieve this, functional models of dynamic batch generating and customer preferences related to quality control criterion need to be developed. In this paper, a dynamic statistical quality control model considering customer preferences is suggested to solve such statistical quality control problems. Production batches with appropriate quality control criterion are formed regarding various customer preferences, and the small sample capacity for applying SPC is considered. Our application results in special steel product process quality designs which involves various customer preferences is presented to demonstrate the effectiveness of the proposed method.

## Introduction

Although industrial processes are totally different in appearance, the problems they are facing is similar: how to improve product quality and product yield in a brief period of time to face the keener competitive environment, as while as to meet the various and changing customer requirements. Statistical quality control methods are shown advantages in industrial processes monitoring and control due to their abilities of help engineers to remove the factors that cause abnormal variations as early as possible, and they are also beneficial to the improvements of the product quality by giving suggestions of setting appropriate quality control criterion. Conventional statistical process control charts such as Shewhart control charts, cumulative sum (CUSUM) control charts, and exponentially weighted moving average (EWMA) control charts are well established based on statistics and large amount of data for monitoring univariate processes.

Customer oriented industrial processes are, however, multi-variable systems consisting of a large number of mutually correlated variables, and with a relatively small amount of data. Unfortunately, conventional univariate SPC charts do not function well for such multi-variable processes with limited samples because they can neither cope with correlation among variables, nor reveal statistical pattern with small volume of data. To monitor multi-variable processes, MSPC has been developed<sup>[1]</sup>. To apply SPC method in small batch production, the concept of 'families of parts' based on group technology<sup>[2]</sup> is widely used. In this kind of methods, the analogical processes are based on manufacturing system itself, however, the different customer's preferences which actually result in the varieties of products or processes are seldom considered. As information regarding the relationships between customer preferences and functional specifications or process parameters are generally subjective and blurred in nature, development of the customer preference models is a difficult work. In this paper, a fuzzy clustering based dynamic multivariate process quality control method is proposed to develop the statistical process control method with small batches composed by products with similar customer preferences.

## Dynamic Multi-variate Process Quality Control Model

Customer preference models are widely studied in recent years due to the awareness of its importance for both businesses and academic researches. Some researcher paid attention at the mapping of the customer preferences to customer satisfaction<sup>[3]</sup>, and some built customer preference model to evaluate company's performance<sup>[4]</sup>. In this paper, we try to set up a Fuzzy Dynamic Multi-Variate Process Quality Control (FDMPQC) model to present the relationship between customer preferences, process programs, and relevant statistical quality control criterion based on Fuzzy Set theory, QFD, group technology and SQC, as shown in Fig.1.

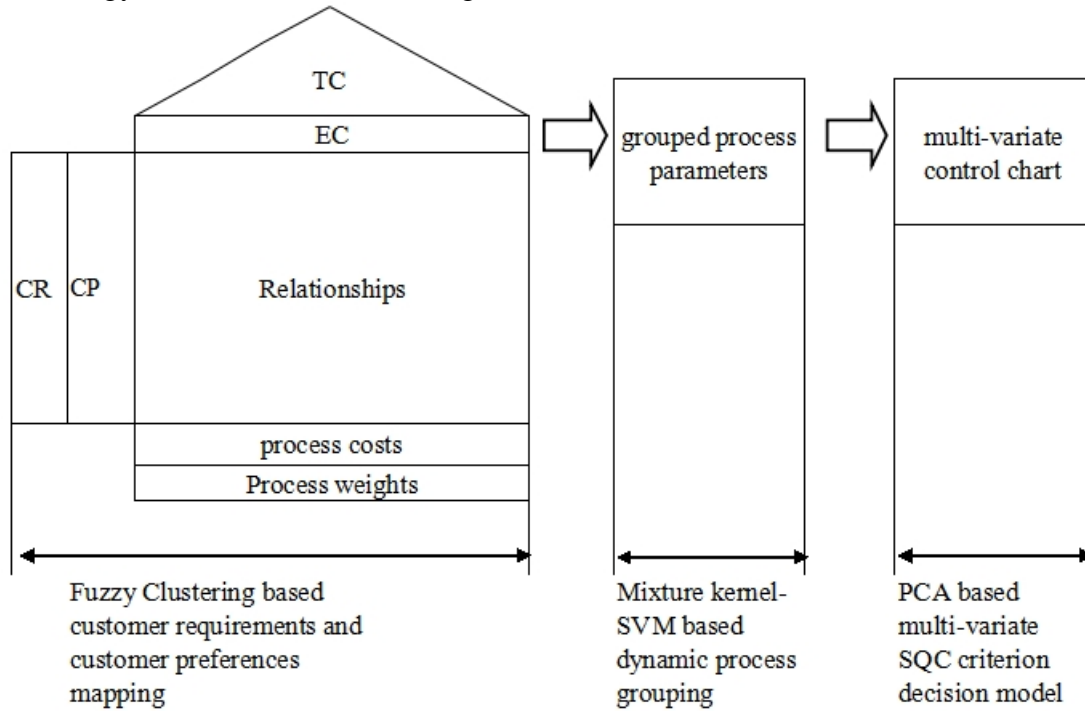


Fig. 1 FDMPQC model based on customer preferences

### Step 1: Fuzzy QFD customer requirements and customer preferences mapping

Quality Function Deployment (QFD)<sup>[5]</sup> is one of the prevailing tools which interprets customer needs into engineering characteristics or process parameters in manufacturing industries. Since the traditional QFD is limited to structured data, a large amount of relevant researches are devoted to recognize customer requirements and preferences that are usually presented as free text data, which are not unified and sometimes even too vague to be recognized by computers.. Fuzzy Set theory<sup>[6]</sup> has a good ability to cope with such problems, in this paper, we apply Fuzzy QFD method to defuzzificate the fuzzy process program translated from fuzzy customers' requirements and preferences.

### Step 2: Mixture kernel-SVM based dynamic process grouping

In order to apply statistical process control method to small batches based on customer preferences, similar process programs should be recognized and grouped into one control batch sample. SVM clustering is a widely known multivariate analysis technique that inserts observations (samples) into classes (clusters) so that observations in the same cluster are as similar as possible, and items in different clusters are as dissimilar as possible<sup>[7]</sup>. The mixtures of kernels provide more optimal performance than any single kernel<sup>[8]</sup>. In the second step of the MQCM model, the mixture kernel-SVM based process programs grouping method is presented to form the statistical quality control batches for the next step of the model.

### Step 3: PCA based multi-variate SQC criterion deciding

The mixture kernel-SVM based process grouping method may successfully supply a group of processes which has similar monitor variables, but still, these variate characters need to be dimension reduction so that we can apply control chart to monitor and control them. Primary Content Analysis(PCA)<sup>[9]</sup> is a multivariate technique capable of compressing data and reducing data dimensionality so that it can be used to monitor possibly correlated variables. In the third step, PCA based SQC criterion decision method is designed to optimize multi-variate online quality control.

### Model Expressions And Algorithms

The model starts by using fuzzy set for the features that the customer requires (internal variables "WHAT"), and then it seeks to establish the relevant process programs with specific technique parameters(external variables "HOW"), whereas customer preferences are treated as the importance of the "WHATs". Fuzzy Set Theory offers a mathematically precise way of modeling vague preferences. The customers' requirements and preferences are identified and analyzed using a linguistic variable in fuzzy numbers. This linguistic set is shown as  $F=\{VL(\text{very low}), L(\text{low}), M(\text{Medium}), H(\text{high}), \text{and } VH(\text{very high})\}$ , and translated into the fuzzy number as  $(2,3,4), (4,5,6), (6,7,8), (8,9,10)\}$ . The fuzzy number of the customer special requirement can be expressed as data  $U = (u_{ij})_{n \times m}$ , where  $u_{ij}$  represent the  $i$ th "WHAT" of the  $j$ th customer, where  $i=1,2,\dots,n$  and  $j=1,2,\dots,m$ .

The importance of the "WHATs" are mainly derived from customer's preferences and can be computed as Eq.(1):

$$w_{ij} = \frac{u_{ij}}{\sum_{i=1}^n u_{ij}} \quad (1)$$

To build the relationship matrix between WHATs and HOWs, Eq.(2) is involed to express the technique measurements to customers' preferences by using linguistic variables to assess the relation matrix between their requirements and the process measurements, whereas fuzzy numbers are also used in this step to translate these variables:

$$r_{ijk} = \frac{1}{l} \sum_{l=1}^q r_{ijk}^l \quad (2)$$

Where  $r_{ijk}^l$  represent the WHAT-HOW matrix of weighted fuzzy numbers for the  $i$ th WHAT of the  $j$ th customer;  $k$  = the number of HOWs;  $l$ = the number of process decision-makers.

Based on the WHAT-HOW matrix ( $r_{ijk}$ ), the weights of HOWs ( $w_{Hjk}$ ) can be formulated as Eq. (3):

$$w_{Hjk} = \sum_{i=1}^n w_{ij} r_{ijk} \quad (3)$$

In order to form a control set with enough samples, similar technique measurement considering customers' preferences with inherent vagueness and subjective judgments should be clustered based on clustering method. The mathematical description for the clustering problem is as follows: The sample are consist of the set of the technique measurements weights  $W_H$ , which can be formulated as Eq.(4):

$$W_H = (w_{Hjk})_{m \times p} \quad (4)$$

The target of the clustering method is to find a division of the sample set  $W_H$  as Eq.(5) that minimize an objective function given by Eq. (6).under constrains (7-9):

$$W_H = \{w_1, w_2, \mathbf{L}, w_c\} \quad (5)$$

$$F(W_H, U) = \sum_{g=1}^c \sum_{k=1}^p (u_{gk})^s d_{gk} \quad (6)$$

$$W_H = \bigcup_{g=1}^c w_g \quad (7)$$

$$w_g \neq \Phi(g=1, 2, \mathbf{L}, c) \quad (8)$$

$$w_{g1} w_{g2} = \Phi(g1, g2=1, 2, \mathbf{L}, c; g1 \neq g2) \quad (9)$$

Where  $s$  is the level of cluster fuzziness between 0 and 1, and  $d_{gk}$  is the squared Euclidean distance that measures the dissimilarity between the feature vectors  $W_{Hk}$  of the technique measurements  $k$  and  $w_g$  of the group  $g$ . The distance is calculated using Eq. (10).

$$d_{gk} = \sum_{j=1}^m (w_{Hjk} - w_{gk})^2 \quad (10)$$

The groups that minimize criterion  $F$  are updated according to Eq. (11).

$$w_g = \frac{\sum_{k=1}^p (u_{gk})^s x_{gk}}{\sum_{k=1}^p (u_{gk})^s} \quad (11)$$

The membership degrees that minimize criterion  $F$  are updated using Eq. (12) under the restriction in Eq. (13).

$$u_{gk} = \left[ \sum_{a=1}^c \left( \frac{d_{gk}}{d_{ak}} \right)^{\frac{1}{s-1}} \right]^{-1} \quad (12)$$

$$\sum_{g=1}^c u_{gk} = 1 \quad (13)$$

Fuzzy c-means (FCM) combined with PSO together is proved to be a effective method to solve clustering problem<sup>[10]</sup>. In FCM-PSO (FPSO) algorithm, particle position and velocity are defined as

$$X_g = \begin{bmatrix} u_{11} & \mathbf{L} & u_{c1} \\ \mathbf{M} & \mathbf{O} & \mathbf{M} \\ u_{1p} & \mathbf{L} & u_{cp} \end{bmatrix}$$

follows: let the position of particle  $i$ , represented by  $X_g$ , be the fuzzy relation between technique measurements and groups, which means that each particle represents one possible solution for the membership degree matrix  $U$ , where  $u_{gk}$  is the membership degree of technique measurement  $k$  to group  $i$  subjected to Eq.(13). Eq.(14) is used for updating the position of particle  $i$  and Eq(15) for updating the velocity.

$$X_g(t+1) = X_g(t) + V_g(t+1) \quad (14)$$

$$V_g(t+1) = wV_g(t) + c_1 r_1 (p_g(t) - X_i(t)) + c_2 r_2 (g_g(t) - X_g(t)) \quad (15)$$

Where  $p_g$  represents the best solution achieved by particle  $g$  and  $g_g$  represents the best solution achieved by all particles.

For technique measurements grouped by weights  $WH$ , there consists an optimized quality monitor and control plan involving multiple quality control characteristics  $Q$ . For these quality characteristics, they may dimensional inconsistency or have an order of magnitude difference. So we have to

diminish these difference by a magnitude standardization to access  $Q = (q_{gk})_{c \times p}$ , where

$q_{gk} = (qx_{gk} - q\bar{x}_k) / s_k$  is each standardization value of  $qx_{gk}$ ,  $q\bar{x}_k$  and  $s_k$  are the average and the standard variation of the elements in sample  $k$ , respectively. After standardization, the characteristics still exist positive or negative correlation, in order to deal with this problem, a relationship matrix  $C_{q \times q} = \text{correcoef}(Z)$  should be worked out by matlab relationship function, and then the Eigenvalue and Eigenvector of matrix  $C$  are solved by the Jacobi Method.

The square root of the load coefficient of each component divided by corresponding Eigenvalue is obtained which make the corresponding coefficient of the principal component expressions, the expressions is like below:

$$F_1 = a_{11}Z_1 + a_{12}Z_2 + \mathbf{L} + a_{1q}Z_q$$

$$F_2 = a_{21}Z_1 + a_{22}Z_2 + \mathbf{L} + a_{2q}Z_q$$

...

$$F_m = a_{m1}Z_1 + a_{m2}Z_2 + \mathbf{L} + a_{mq}Z_q$$

Where  $a_{i1}, a_{i2}, \mathbf{L}, a_{iq}$  ( $i=1, 2, \mathbf{L}, m$ ) is the Eigenvector corresponding to the Eigenvalue of the matrix  $C$  which is the covariance matrix of  $Z$ , and  $Z_1, Z_2, \mathbf{L}, Z_q$  is the weighted and standardized value of the original data. A comprehensive value of each principal component should be worked out by equation (1)~(m). And then Put with these data, we can obtain the multivariate control charts of average  $\bar{x}$  and range  $R$  to obtain the control level of each batch.

## Experiments and Results

In this section, the method proposed in this work is applied to solve the clustering problem based on customer's preferences which presented by customers' requirements of orders as shown in Tab.1.

Tab. 1 Customers' requirements of orders.

No.	standard	Order Amount	Customer special Requirements					
			C	Si	Mn	P	S	Cu
1	GB/T18254-2002	2	VH	M	M	M	H	L
2	GB/T18579-2001	1	M	M	VL	M	H	H
3	YB9-68	5	M	M	L	H	H	H
4	GB/T18254-2002	10	M	M	L	M	L	L
...								
100	GB/T18254-2002	0.5	VH	M	M	M	L	L

The corresponding weight of the process parameters can be translated, and then using FCM-PSO algorithm to cluster the orders under the amount limitatoin of 30t per group. For each group, PCA method is applied to get a synthetic control limitation criterions for each group, eg. As shown in Fig.2. and Fig.3.

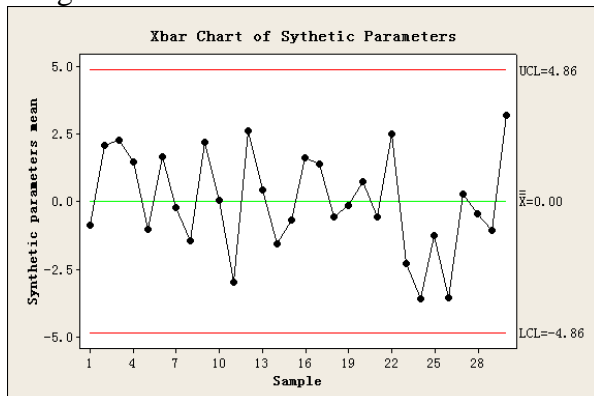


Fig. 2 Xbar chart for monitoring synthetic parameters

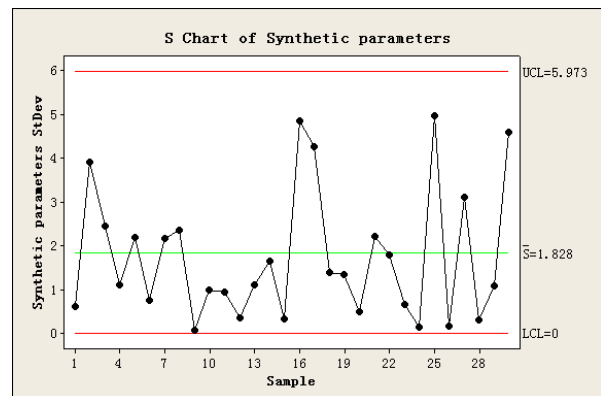


Fig. 3 S chart for monitoring synthetic parameters

## Conclusions

The fuzzy dynamic multi-variate process quality control is helpful in determining the appropriate settings of statistical quality control criterion for dynamic batches composed of similar products with different customer preferences to best satisfy customer's requirements and minimized production cost. Therefore, it is able to diminish possible under-control or over-control problem of the batch process, and has some advantages to the traditional multivariate quality control method.

## Acknowledgements

This work was financially supported by the Liaoning education department Science Foundation (L2013511) and the Fundamental Research Funds for the Central Universities(DC201502010401).

## References

- [1] Manabu Kano, Yoshiaki Nakagawa. Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry. *Computers & Chemical Engineering*, vol. 32, Issues 1–2, January 2008, pp. 12-24.
- [2] Lee, Emma K. A mathematical programming approach to the group technology part family/machine group formation problem. University of Florida, 1988.
- [3] Petric, F. The roles of quality, value and satisfaction in predicting. cruise passengers' behavioral intentions [J]. *Travel Research*, 2004, 42 (4): 397-407.
- [4] Prasun Das, Sandip Mukherjee. Modeling of Customer Preferences on Product Features and Comparing the Competitor's Performances. *Quality Engineering*, 20: 53-62, 2008.
- [5] Jian Jin, Ping Ji, Ying Liu. Translating online customer opinions into engineering characteristics in QFD: A probabilistic language analysis approach. *Engineering Applications of Artificial Intelligence*, Volume 41, May 2015, Pages 115-127.
- [6] Fung, Richard, Y.K., Chen, Yizeng, Tang, Jiafu, 2006. Estimating the functional relationships for quality function deployment under uncertainties. *Fuzzy Sets Syst.* 157 (1), 98–120.
- [7] Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. New Jersey: Wiley Interscience; 2005.
- [8] Sheng Zheng, Jian Liu, JinWen Tian. An efficient star acquisition method based on SVM with mixtures of kernels. *Pattern Recognition Letters*, Volume 26, Issue 2, 15 January 2005, Pages 147-165.
- [9] Jackson, J.E. . A User's Guide to Principal Components. Wiley, 1991.
- [10] Ling Wang, Yi-Kai Juan, Jie Wang, Kai-Meng Li, Colin Ong. Fuzzy-QFD approach based decision support model for licensor selection. *Expert Systems with Applications* 39 (2012) 1484–1491.