

Monitoring fuzzy linear quality profiles: A comparative study

Mohammadreza Nasiri Boroujeni^{a*}, Yaser Samimi^a and Emad Roghanian^a

^aDepartment of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

CHRONICLE

Article history:

Received August 17 2020

Received in Revised Format

October 7 2020

Accepted October 15 2020

Available online

October, 15 2020

Keywords:

Fuzzy linear regression

Multivariate statistical process control

Fuzzy T^2 control chart

Fuzzy profile monitoring

ABSTRACT

Quality profiles representing the quality of a process or product as the functional relationship between one or more dependent variables and one or more explanatory variables, which are nowadays widely recognized in statistical process control (SPC) applications by both researchers and practitioners. Furthermore, in many real-world cases, evaluation of process or product characteristics is carried out with ambiguity or conducted using linguistic values. The theory of fuzzy sets provides an appropriate approach to deal with uncertainty due to ambiguity in human subjective evaluations or vagueness in linguistic variables. The purpose of this study is to introduce two novel methods based on fuzzy regression modeling for monitoring fuzzy linear profiles in phase II of SPC. To accomplish this, the fuzzified Hotelling's T^2 statistic and fuzzy hypothesis testing are used. Moreover, a simulation study is used to compare the performance of the proposed methods with previous methods, based on the average run length (ARL) criterion in order to assess the detectability of charts with regard to the step shifts in profile parameters. Finally, the results of a real-world example in the tile and ceramic industry are presented.

© 2021 by the authors; licensee Growing Science, Canada

1. Introduction

Nowadays, to keep up with the fast pace of technological advancements in both manufacturing and service industries, as well as to take the most advantage of the ever-growing ease of storing huge volume of multifarious multivariate data in highly equipped processes of today necessitate the use of novel scientific approaches and advanced analytical tools for online monitoring of key process or product quality characteristics. Montgomery (2009) postulated that the quality of a process is inversely related to the process variability. Such a definition entails a well-established conceptual framework in quality improvement philosophy that a reduction in the variability of the key quality characteristics will result in an increase in the process or product quality. To reduce the process variability and as a result, elevate the process quality, statistical process control (SPC) provides a set of magnificent effective tools among which control charts are the most widely-used monitoring tool and quite well-known. Although in traditional SPC, process characteristics are often regarded as a single or a collection of correlated random variables, recently, the subject of process monitoring using functional data has received a great deal of attention. Quality profiles are used to represent the functional relationship between one or more response variables and one or more independent variables in the process. Profile monitoring is carried out in two phases. Broadly speaking, the main purpose of the retrospective phase or the so-called phase I is to analyze historical data in order to elucidate an authentic process model and to remove outliers. The in-control process parameters are also estimated in this phase to design process control charts for the prospective stage or phase II. The goal of phase II is to quickly detect the parameter changes with respect to the reference values determined in phase I. Statistical process control charts are well-known effective tools for monitoring purposes in phase II. The performance of the control chart in phase II is usually evaluated based on the average run length (ARL) criterion. The ARL describes the average number of consecutive plotted points on the control chart since the emergence of a persistent change in natural process model until an out of control signal appears.

* Corresponding author Tel.: +989120039488

E-mail: [\(M. Nasiri Boroujeni\)](mailto:mohammadreza.nasiri@alumni.kntu.ac.ir)

So far, numerous control charts have been developed for analyzing and monitoring quality profiles in both phases I and II. Profiles are categorized on the basis of their form as linear, polynomial, generalized linear, nonlinear, nonparametric, along with their multivariate counterparts. More specifically, regarding the subject of our research, Kang and Albin (2000) proposed a multivariate T^2 control chart for monitoring the coefficients of a simple linear profile, and a range control chart for monitoring the variation of points scattered around the linear model. Later on, Kim et al. (2003) used three distinct EWMA control charts for tracking the status of three model parameters including the intercept, slope as well as the variance of the model error term. Employing a group regression approach, Mahmoud et al. (2007) provided a likelihood ratio based control chart for clustering linear profiles in phase I. Additionally, Zou et al. (2006) used a change-point formulation approach for monitoring simple linear profiles and estimating the time of a step shift in process parameters. In order to acquire thorough knowledge about profile monitoring methods such as multivariate control charts, time-weighted control charts, and change point formulation and their application to different forms of profiles in both univariate, as well as multivariate settings, Noorossana et al. (2011), can be referred to as a finely-tuned comprehensive text. On the other hand, in some real-world cases, there are occasions where the precise value of a process or product quality characteristics cannot be measured or determined. Examples can be found in Ghobadi et al. (2014). In some circumstances, quality may be described linguistically on the basis of the expertise of a well-trained or knowledgeable inspector. Material's bending flexibility, surface smoothness/coarseness of materials in the foundry process, and/or transparency/opaqueness of glassware are among frequently-seen instances. Fuzzy sets may be utilized to represent such ambiguous evaluations, and remarkably, the so-called fuzzy random variables are affected by two different uncertainty sources; the first one is related to the inherent random variability of the process or the common-cause variation which may be modeled adopting a probabilistic approach, whereas the next one, the ambiguity in linguistic evaluations or imprecise measurements is formulated from a possibilistic point-of-view using the notion of membership function (Fernandez, 2017; Sabegh et al., 2014). This paper deals with estimating and then monitoring a fuzzy linear function which is used to represent the functional relationship of quality characteristics of a process or product and hence known as the fuzzy linear quality profile. Investigation of the relevant literature reveals a few prominent collections of studies by Moghadam et al. (2015, 2016, 2018), Noghondarian and Ghobadi (2012) and Ghobadi et al. (2014) focused on ceramic/tile industry and tourism industry, respectively. In this study, we propose two other methods for estimation of fuzzy linear profiles and then compare them with the existing approaches in Moghadam et al. (2015) and Ghobadi et al. (2014), in terms of the rate of detection of a change in process parameters. ARL criterion is used and calculated through a simulation study. In this study, we use fuzzy Hotelling's T^2 control chart to monitor a fuzzy quality profile.

The structure of this paper is as follow: in the second section, some important research papers on the estimation of parameters of fuzzy linear regression and monitoring crisp and fuzzy linear profiles are studied. In the third section, fuzzy linear profiles, related assumptions, and four regression models are introduced to define fuzzy linear profiles. In the next section, the fuzzy T^2 statistic besides a method for testing fuzzy hypothesis are discussed. In the fifth section, a case study in tile and ceramic industry is presented. In the sixth section, a comprehensive comparison between the four applied methods is carried out based on average run length (ARL) for various out of control scenarios. Conclusions and possible future research areas are provided in the last part.

2. Literature review

With nearly two decades of precedence, the use of fuzzy sets in SPC have found remarkable applications in order to represent imprecise measurement of process or product quality characteristics or ambiguous subjective human assessments (Raz & Wang, 1990; Kanagawa et al., 1993; Franceschini & Romano, 1999; Gulbay et al., 2004; Taleb & Limam, 2002; Cheng, 2005; Wang & Yasuda, 2014; Senturk et al., 2014; Senturk & Antucheviciene, 2017). Applications involve different types of conventional control charts including Shewhart, as well as EWMA and CUSUM charts in both univariate and multivariate settings encompassing a wide range of industries from food processing and tourism to textile and electronics (Sabegh et al., 2014). Alaeddini et al. (2009) and Zarandi and Alaeddini (2010) proposed a fuzzy clustering approach for identifying the time of a step change of parameters in variable sampling control charts. On the other hand, fuzzy rule-based systems have proved useful to improve the performance of traditional control charts in recognizing unnatural or out-of-control patterns (El-Shal & Morris, 2000; Demirli & Vijayakumar, 2010; Chen & Liang, 2016). Although the idea of combining fuzzy regression and statistical control charts was brought up by researchers nearly one decade ago to deal with problems such as the tool wearing in manufacturing industries, as discussed in the next part, the issue of fuzzy profile monitoring has just been dealt with recently in only a few studies.

2.1. A literature review of fuzzy linear regression

To the best of our knowledge, only a few studies were published on fuzzy linear profile monitoring which studied fuzzy linear profile monitoring in phases I and II. In Noghondarian and Ghobadi (2012) and Ghobadi et al. (2014), both univariate and multivariate approaches were developed for monitoring fuzzy linear profiles in phase I of statistical process control. The multivariate method includes three multivariate fuzzy control charts. Moghadam et al. (2015) presents a new monitoring method for linear profiles with a fuzzy-type response in phase II, which included two multivariate fuzzy control charts known as FT² and FEWMA. In Moghadam et al. (2016), FEWMA control chart is used to monitor fuzzy linear profiles in phase II. When the values of response variable are fuzzy and vague, Moghadam et al. (2018) discusses the phase I of fuzzy profile

monitoring and proposed a maximum likelihood estimator (MLE) for the process change point. Concerning the main issue under discussion, some relevant researches on the estimation of fuzzy linear regression are reported here. In the fuzzy regression analysis, the total prediction error is known as the target for a fuzzy linear programming model. This idea was first introduced by Tanaka et al. (1982). Then, some other fuzzy models with various optimization criteria were introduced. Celmins et al. (1987) and D'Urso (2003) introduced various methods in this area of study. Table 1 summarizes some different fuzzy linear regression estimation methods.

Table 1

Review on fuzzy linear regression

variate type (dependent variate-independent variate)	approach	fuzzy number type	authors
crisp-fuzzy	minimum fuzzy criterion	triangular	Tanaka et al. (1982)
	least squares with maximum compatibility criterion	triangular	Celmins et al. (1987)
	LSSVM ¹ criterion	triangular	Hong and Hwang (2006)
	least square with minimum fuzzy criterion	LR	Coppi et al. (2006)
fuzzy-crisp	least squares	LR	D'Urso (2003)
fuzzy-fuzzy	bootstrap	LR	Akbari et al. (2012)

3. Simple Fuzzy Linear Profile: Parameter Estimation

In this section, four parameter estimation methods for fuzzy linear profiles are introduced. All of the methods consider crisp inputs and fuzzy outputs. In the first three methods, the parameters of model are also considered as fuzzy sets and in the last one, the parameters are supposed to be as crisp. Although, estimation methods are different, all four methods are comparable because the same crisp inputs and fuzzy outputs are presented to the methods. Among these methods, two methods were previously employed by Ghobadi et al. (2014) and Moghadam et al. (2015). We also make use of a third method developed by D'Urso (2003). The last method, which is based on LS-SVM, is a modified version of the basic idea proposed by Hong and Hwang (2006). As it was mentioned before, in all of these methods inputs are crisp and outputs are fuzzy.

3.1. The fuzzy linear regression model with fuzzy coefficients

3.1.1. The Least Squares approach

We first introduce the fuzzy regression method employed by Ghobadi et al. (2014) to tackle the fuzzy linear profile monitoring problem. This parameter estimation method is primarily introduced by Tanaka (1987) and Tanaka et al. (1989). In this method, for the sake of simplicity of calculations, parameters are considered as symmetric triangular fuzzy numbers. In this method, the data are indicated by $\{X_i, \tilde{Y}_i\}_{i=1}^n$ and the fuzzy linear regression model is considered as follows:

$$\tilde{Y}_i = \tilde{B}_0 + \tilde{B}_1 X_i + \tilde{\varepsilon}_i \quad (1)$$

where, $\tilde{Y}_i = (y_i, c_i)$, $X_i = x_0, x_1, \dots, x_n$, $\tilde{B}_0 = (a_0, s_0)$, $\tilde{B}_1 = (a_1, s_1)$, $\tilde{\varepsilon}_i = (e_0, se_0)$ and $i = 0, 1, \dots, n$.

In this method, the least squares criterion was used for optimization as represented by the linear programming model as follows (Gobadi et al. 2014):

$$\min W_1 \sum_{i=1}^n \sum_{r=0}^1 s_r |X_{ir}| + W_2 \sum_{i=1}^n (d_{iu} + d_{il}) \quad (2)$$

$$s.t.: \sum_{r=0}^1 (a_r + (1 - \alpha)s_r)x_{ir} \geq y_i + (1 - \alpha)c_i - d_{iu} \quad (3)$$

$$\sum_{r=0}^1 (a_r - (1 - \alpha)s_r)x_{ir} \leq y_i - (1 - \alpha)c_i + d_{il} \quad (4)$$

$$W_1 + W_2 = 1 \quad (5)$$

$$d_{il}, d_{iu} \geq 0, \quad a_r \in IR, \quad s_r \geq 0, \quad r = 0, 1, \quad i = 1, 2, \dots, n \quad (6)$$

in which, W_i , ($i = 1, 2$), is an appropriate weight factor. The goal, in this model, is first to minimize the width of fuzzy regression coefficients. Next goal aims to achieve the maximum agreement between the α -cut of the observed values and the α -cut of the estimated values. For each objective, there is a weight factor, which is valued based on adjudication. Moreover, in this model:

¹ Least Square Support Vector Machine

$$d_{iu} = \text{Max} \{0, [\text{Max}(supp\tilde{y}_i) - \text{Max}(supp\hat{y}_i)]\} \quad (7)$$

$$d_{il} = \text{Max}\{0, [\text{Min}(supp\tilde{y}_i) - \text{Min}(supp\hat{y}_i)]\} \quad (8)$$

This method can only be used for symmetric triangular fuzzy numbers.

3.1.2. Goal programming approach

Moghadam et al. (2015) adopted a more recent approach for estimating fuzzy linear profiles introduced by Hassanpur et al. (2009). In this method, observations are shown as $\{X_i, \tilde{Y}_i\}_{i=1}^n$ and the fuzzy linear regression model is considered as follows:

$$\tilde{Y}_i = \tilde{B}_0 + \tilde{B}_1 X_i + \tilde{\varepsilon}_i; \quad i = 0, 1, \dots, n \quad (9)$$

in which, $\tilde{Y}_i = (y_i, l_i, r_i)$, $X_i = x_0, x_1, \dots, x_n$, $\tilde{B}_0 = (b_0, \alpha_0, \beta_0)$, $\tilde{B}_1 = (b_1, \alpha_1, \beta_1)$, $\tilde{\varepsilon}_i = (e_0, le_0, re_0)$.

To estimate coefficients in a fuzzy linear regression, three linear programming models are used as shown in Eqs. (10-21).

$$\min z = \sum_{i=1}^n (n_{ia} + p_{ia}) \quad (10)$$

$$\text{s. t.: } \sum_{r=0}^1 (b_r x_{ir}) + n_{ia} - p_{ia} = y_i \quad (11)$$

$$b_r \in R, \quad r = 0, 1 \quad (12)$$

$$n_{ia}, p_{ia} \geq 0, \quad i = 1, 2, \dots, n \quad (13)$$

$$\min z = \sum_{i=1}^n (n_{il} + p_{il}) \quad (14)$$

$$\text{s. t.: } \sum_{r=0}^1 (\alpha_r x_{ir}) + n_{il} - p_{il} = l_i \quad (15)$$

$$\alpha_r \in R, \quad r = 0, 1 \quad (16)$$

$$n_{il}, p_{il} \geq 0, \quad i = 1, 2, \dots, n \quad (17)$$

$$\min z = \sum_{i=1}^n (n_{ir} + p_{ir}) \quad (18)$$

$$\text{s. t.: } \sum_{r=0}^1 (\beta_r x_{ir}) + n_{ir} - p_{ir} = r_i \quad (19)$$

$$\beta_r \in R, \quad r = 0, 1 \quad (20)$$

$$n_{ir}, p_{ir} \geq 0, \quad i = 1, 2, \dots, n \quad (21)$$

where p_{ia} and n_{ia} are respectively positive and negative deviations from the center of estimated value \hat{y}_i and response value \tilde{y}_i . p_{il} and n_{il} are respectively positive and negative deviations from the left spread of estimated \hat{y}_i and response value \tilde{y}_i . And in a similar manner, p_{ir} and n_{ir} are positive and negative deviations from the right spread of estimated \hat{y}_i and \tilde{y}_i , respectively.

3.1.3. LSSVM method

Our third method is based on a modification carried out on the basic model proposed by Hong and Hwang (2006), in such a way that our model allows to accept the crisp inputs and take fuzzy intercept and slope parameters into account. In this method, the data are shown as $\{X_i, \tilde{Y}_i\}_{i=1}^n \subset T(R)^p \times T(R)$ and the fuzzy linear regression model is also described as follows:

$$\tilde{Y}(X) = \langle \tilde{B}_i, X_i \rangle + \tilde{B}_0 = (\langle m_{B_i}, X_i \rangle + m_{B_0}, \langle \alpha_{B_i}, |X_i| \rangle + \alpha_{B_0}, \langle \beta_{B_i}, |X_i| \rangle + \beta_{B_0}) \quad (22)$$

where:

$$X_i = (x_{i1}, \dots, x_{ip}), \quad i = 1, 2, \dots, n$$

$$\tilde{Y}_i = (m_{Y_i}, \alpha_{Y_i}, \beta_{Y_i}), \quad i = 1, 2, \dots, n$$

$T(R) \rightarrow$ Set of triangular fuzzy numbers

$T(R)^p \rightarrow$ Set of a p -dimensional vector of triangular fuzzy numbers

$$\tilde{B}_0 = (m_{B_0}, \alpha_{B_0}, \beta_{B_0})$$

$$\tilde{\mathbf{B}}_i = \left((m_{B_{i1}}, \alpha_{B_{i1}}, \beta_{B_{i1}}), \dots, (m_{B_{ip}}, \alpha_{B_{ip}}, \beta_{B_{ip}}) \right)$$

To estimate fuzzy linear regression coefficients, we resort to the following linear programming model (Hong and Hwang, 2006).

$$\min \frac{1}{2} (\|\mathbf{m}_{\mathbf{B}_i}\|^2 + \|\boldsymbol{\alpha}_{\mathbf{B}_i}\|^2 + \|\boldsymbol{\beta}_{\mathbf{B}_i}\|^2) + \frac{C}{2} \sum_{k=1}^3 \sum_{i=1}^n e_{ki}^2 \quad (23)$$

$$m_{Y_i} - \langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle - m_{B_0} = e_{1i} \quad (24)$$

$$(m_{Y_i} - \alpha_{Y_i}) - (\langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle + m_{B_0} - \langle \boldsymbol{\alpha}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle - \alpha_{B_0}) = e_{2i} \quad (25)$$

$$(m_{Y_i} + \beta_{Y_i}) - (\langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle + m_{B_0} + \langle \boldsymbol{\beta}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle + \beta_{B_0}) = e_{3i} \quad (26)$$

To solve this linear programming by introducing Lagrange multipliers, we construct the Lagrange function as follows:

$$\begin{aligned} L = & \frac{1}{2} (\|\mathbf{m}_{\mathbf{B}_i}\|^2 + \|\boldsymbol{\alpha}_{\mathbf{B}_i}\|^2 + \|\boldsymbol{\beta}_{\mathbf{B}_i}\|^2) + \frac{C}{2} \sum_{k=1}^3 \sum_{i=1}^n e_{ki}^2 - \sum_{i=1}^n \lambda_{1i} (e_{1i} - m_{Y_i} + \langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle + m_{B_0}) \\ & - \sum_{i=1}^n \lambda_{2i} (e_{2i} - (m_{Y_i} - \alpha_{Y_i}) + (\langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle + m_{B_0} - \langle \boldsymbol{\alpha}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle - \alpha_{B_0})) \\ & - \sum_{i=1}^n \lambda_{3i} (e_{3i} - (m_{Y_i} + \beta_{Y_i}) + (\langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle + m_{B_0} + \langle \boldsymbol{\beta}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle + \beta_{B_0})), \end{aligned} \quad (27)$$

where the tuning parameter C is a positive real constant. Then, the following set of equations are utilized to optimize the Eq. (27).

$$\frac{\partial L}{\partial \mathbf{m}_{\mathbf{B}_i}} = 0 \rightarrow \mathbf{m}_{\mathbf{B}_i} = \sum_{k=1}^3 \sum_{i=1}^n \lambda_{ki} \times \mathbf{X}_i \quad (28)$$

$$\frac{\partial L}{\partial \boldsymbol{\alpha}_{\mathbf{B}_i}} = 0 \rightarrow \boldsymbol{\alpha}_{\mathbf{B}_i} = - \sum_{i=1}^n \lambda_{2i} \times |\mathbf{X}_i| \quad (29)$$

$$\frac{\partial L}{\partial \boldsymbol{\beta}_{\mathbf{B}_i}} = 0 \rightarrow \boldsymbol{\beta}_{\mathbf{B}_i} = \sum_{i=1}^n \lambda_{3i} \times |\mathbf{X}_i| \quad (30)$$

$$\frac{\partial L}{\partial m_{B_0}} = 0 \rightarrow \sum_{k=1}^3 \sum_{i=1}^n \lambda_{ki} = 0 \quad (31)$$

$$\frac{\partial L}{\partial \alpha_{B_0}} = 0 \rightarrow \sum_{i=1}^n \lambda_{2i} = 0 \quad (32)$$

$$\frac{\partial L}{\partial \beta_{B_0}} = 0 \rightarrow \sum_{i=1}^n \lambda_{3i} = 0 \quad (33)$$

$$\frac{\partial L}{\partial e_{ki}} = 0 \rightarrow e_{ki} = \frac{\lambda_{ki}}{C}, k = 1, 2, 3 \quad (34)$$

$$\frac{\partial L}{\partial \lambda_{1i}} = 0 \rightarrow m_{Y_i} - \langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle - m_{B_0} - e_{1i} = 0 \quad (35)$$

$$\frac{\partial L}{\partial \lambda_{2i}} = 0 \rightarrow m_{Y_i} - \alpha_{Y_i} - \langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle - m_{B_0} + \langle \boldsymbol{\alpha}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle + \alpha_{B_0} - e_{2i} = 0 \quad (36)$$

$$\frac{\partial L}{\partial \lambda_{3i}} = 0 \rightarrow m_{Y_i} + \beta_{Y_i} - \langle \mathbf{m}_{\mathbf{B}_i}, \mathbf{X}_i \rangle - m_{B_0} - \langle \boldsymbol{\beta}_{\mathbf{B}_i}, |\mathbf{X}_i| \rangle - \beta_{B_0} - e_{3i} = 0 \quad (37)$$

The solution of this set of equations may be found via the matrix representation of the following linear equations:

$$\begin{bmatrix} 0 & 0 & 0 & \mathbf{1}' & \mathbf{1}' & \mathbf{1}' \\ 0 & 0 & 0 & \mathbf{0}' & \mathbf{1}' & \mathbf{0}' \\ 0 & 0 & 0 & \mathbf{0}' & \mathbf{0}' & \mathbf{1}' \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & S_{11} & S_{12} & S_{13} \\ \mathbf{1} & -1 & \mathbf{0} & S'_{11} & S_{22} & S_{23} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & S'_{13} & S'_{23} & S_{33} \end{bmatrix} \begin{bmatrix} m_{B_0} \\ \alpha_{B_0} \\ \beta_{B_0} \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbf{m}_Y \\ \mathbf{m}_Y - \boldsymbol{\alpha}_Y \\ \mathbf{m}_Y + \boldsymbol{\beta}_Y \end{bmatrix} \quad (38)$$

where $\mathbf{1}$, $\mathbf{0}$, λ_1 , λ_2 , λ_3 and S_{11}, \dots, S_{33} are $n \times 1$ and $n \times n$, respectively, and defined by Eqs. (39-44).

$$S_{11} = [\langle \mathbf{X}_i, \mathbf{X}_j \rangle] + I/C \quad (39)$$

$$S_{12} = [\langle \mathbf{X}_i, |\mathbf{X}_j| \rangle] \quad (40)$$

$$\mathbf{S}_{13} = [\langle \mathbf{X}_i, \mathbf{X}_j \rangle] \quad (41)$$

$$\mathbf{S}_{22} = [\langle \mathbf{X}_i, \mathbf{X}_j \rangle] + [|\mathbf{X}_i|, |\mathbf{X}_j|] + I/C \quad (42)$$

$$\mathbf{S}_{23} = [\langle \mathbf{X}_i, \mathbf{X}_j \rangle] \quad (43)$$

$$\mathbf{S}_{33} = [\langle \mathbf{X}_i, \mathbf{X}_j \rangle] + [|\mathbf{X}_i|, |\mathbf{X}_j|] + I/C \quad (44)$$

\mathbf{I} is also an $n \times n$ identity matrix. Now, the estimation of fuzzy linear regression is given as follows:

$$\hat{\tilde{Y}}(\mathbf{X}_q) = \langle \hat{\tilde{\mathbf{B}}}_q, \mathbf{X}_q \rangle + \hat{\tilde{B}}_0 = (\langle \hat{\mathbf{m}}_{B_q}, \mathbf{X}_q \rangle + \hat{m}_{B_0}, \langle \hat{\alpha}_{B_q}, |\mathbf{X}_q| \rangle + \hat{\alpha}_{B_0}, \langle \hat{\beta}_{B_q}, |\mathbf{X}_q| \rangle + \hat{\beta}_{B_0}) \quad (45)$$

3.2. Crisp coefficients

3.2.1. D'Urso's Method

To estimate the parameters of the fuzzy linear regression model, D'Urso (2003) adopted a weighted quadratic error minimization approach. More specifically, in this method, the data are shown as $\{X_i, \tilde{Y}_i\}_{i=1}^n$ and the fuzzy linear regression model is considered as follow:

$$\tilde{Y}_i = B_0 + B_1 X_i; \quad i = 1, 2, \dots, n \quad (46)$$

where, $\tilde{Y}_i = (c_i, p_i, q_i)$ is the estimated fuzzy output and X_i is the independent crisp variable. In a general form, the relationships of the fuzzy linear regression are defined as follows:

$$\mathbf{c} = \mathbf{c}^* + \boldsymbol{\varepsilon}, \quad \mathbf{c}^* = \mathbf{XB}, \quad (47)$$

$$\mathbf{p} = \mathbf{p}^* + \boldsymbol{\lambda}, \quad \mathbf{p}^* = \mathbf{c}^* b + \mathbf{1}d, \quad (48)$$

$$\mathbf{q} = \mathbf{q}^* + \boldsymbol{\rho}, \quad \mathbf{q}^* = \mathbf{c}^* g + \mathbf{1}h, \quad (49)$$

where, \mathbf{X} , an $n \times (k+1)$ matrix, allows including k crisp input variables. \mathbf{c} and \mathbf{c}^* are $n \times 1$ vectors which are respectively assigned to estimated and observed centers. \mathbf{p} and \mathbf{p}^* are $n \times 1$ vectors which are respectively assigned to estimated and observed left spreads. In a similar manner, \mathbf{q} and \mathbf{q}^* are $n \times 1$ vectors which are respectively assigned to estimated and observed right spreads. \mathbf{B} is a $(k+1) \times 1$ vector that includes the coefficients of the linear fuzzy regression model. b, d, g, h are the linear regression parameters for \mathbf{p} and \mathbf{q} . $\mathbf{1}$ is an $n \times 1$ unit vector and $\boldsymbol{\varepsilon}, \boldsymbol{\lambda}, \boldsymbol{\rho}$ are an $n \times 1$ error vectors corresponding to three equations of the regression model. Regarding the aforementioned regression model, a numerical solution is sought with the aim of minimizing an objective function consisting of errors between the observed and estimated values. This model is defined as follows:

$$\min \Delta(\mathbf{B}, b, d, g, h) = (\mathbf{c} - \mathbf{c}^*)'(\mathbf{c} - \mathbf{c}^*)\pi_c + (\mathbf{p} - \mathbf{p}^*)'(\mathbf{p} - \mathbf{p}^*)\pi_p + (\mathbf{q} - \mathbf{q}^*)'(\mathbf{q} - \mathbf{q}^*)\pi_q, \quad (50)$$

where, π_c, π_p, π_q are positive weights indicating the relative priority of the components. In our study, $\hat{\mathbf{B}}, \hat{b}, \hat{d}, \hat{g}, \hat{h}$ values are estimated by solving the above optimization problem using MATLAB (R2017a) software.

4. monitoring fuzzy linear profiles in phase II

After estimating the parameters of the profile and identifying the baseline model of the process, we can evaluate the performance of the process to distinguish process status. In this section, in order to monitor the fuzzy parameters of the linear profile, the fuzzy T^2 statistic is developed on the basis of the fuzzy extension principle. Then, the testing fuzzy hypothesis is introduced to characterize the process status (in-control and out of control) by determining the standardized distance between the sample profile and estimated process means.

4.1. Fuzzy T^2 statistics

Kang and Albin (2000) introduced a method for monitoring linear profiles in phase II using Hotelling's T^2 statistic. Consider the coefficients of a simple linear model as $(A_0, A_1)^T$, and let $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$ represent the error covariance matrix. T^2 Statistic is defined as follows:

$$T_i^2 = (\mathbf{Z}_i - \boldsymbol{\mu})\Sigma^{-1}(\mathbf{Z}_i - \boldsymbol{\mu}) \quad (51)$$

where in the above equation, $\mathbf{Z}_i = (a_{0i}, a_{1i})^T$ is which is obtained from the sample data. Now, suppose that $\tilde{\mathbf{Z}}_i = (\tilde{a}_{0i}, \tilde{a}_{1i})^T$ and $\tilde{\boldsymbol{\mu}} = (\tilde{A}_0, \tilde{A}_1)^T$ represent estimated fuzzy values and the real value of the profile parameters. Moreover, the elements of Σ include variances and covariance of \tilde{a}_{0i} and \tilde{a}_{1i} . It is assumed that these elements are given or are estimated from phase I using in-control process observations. Based on Vieratal (2011) and Korner and Nather (1998), these values can be assumed as crisp numerical values. Then, \widetilde{FT}_i^2 statistic is calculated as follows:

$$\widetilde{FT}_i^2 = \frac{1}{\sigma_{\tilde{a}_{0i}}^2} (\tilde{a}_{0i} - \tilde{A}_0)^2 + \frac{1}{\sigma_{\tilde{a}_{1i}}^2} (\tilde{a}_{1i} - \tilde{A}_1)^2, \quad (52)$$

where $\sigma_{\tilde{a}_{0i}}^2$ and $\sigma_{\tilde{a}_{1i}}^2$ are calculated from below definition.

Definition Based on Viertl (2011) and Korner and Nather (1998), for a random sample of LR fuzzy numbers, $\tilde{X}_1^*, \tilde{X}_2^*, \dots, \tilde{X}_n^*$ where $\tilde{X}_i^* = (x_i, \alpha_i, \beta_i)$, $\sigma_{\tilde{X}^*}^2$ we have,

$$S_{\tilde{X}^*}^2 = S_{xi}^2 + \frac{1}{6} (S_{\lambda i}^2 + S_{\beta i}^2) + \frac{1}{2} (S_{xi, \beta i} - S_{xi, \lambda i}). \quad (53)$$

Moghadam et al. (2015) introduced the $\alpha - cut$ of \widetilde{FT}_i^2 statistics as follows:

$$C_\alpha(\widetilde{FT}_i^2) = [FT_i^{2L}(\alpha), FT_i^{2R}(\alpha)] = \begin{cases} \text{Min}T_i^2(\underline{x}'') & , \text{ st: } \underline{x}'' \in C_\alpha(\underline{x}'') \\ \text{Max}T_i^2(\underline{x}'') & , \text{ st: } \underline{x}'' \in C_\alpha(\underline{x}'') \end{cases} \quad (54)$$

4.2. Fuzzy hypothesis testing for evaluation of process status

A control chart based on the \widetilde{FT}^2 statistic is used to repeatedly carry out the following fuzzy hypothesis testing:

$$\begin{cases} H_0: \tilde{A}_0 = \tilde{a}_0 \\ H_1: \tilde{A}_0 \neq \tilde{a}_0 \end{cases}, \quad \begin{cases} H_0: \tilde{A}_1 = \tilde{a}_1 \\ H_1: \tilde{A}_1 \neq \tilde{a}_1 \end{cases} \quad (55)$$

Taheri and Arefi (2009) introduced a method for testing fuzzy hypothesis consisting of four steps:

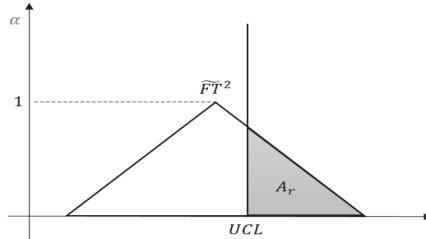


Fig. 1. fuzzy hypothesis testing

- 1) Calculation of the area below the fuzzy test statistic (A_T).
- 2) A_r , the region under the test statistic \widetilde{FT}^2 but to the right of the vertical line through UCL of the \widetilde{FT}^2 control chart.
- 3) The value of $\varphi \in (0,1]$ is chosen as the level of validity, which we consider to be 0.5.
- 4) If $\frac{A_r}{A_T} \geq \varphi$, then the fuzzy zero assumption is rejected. In other words, the point is out of control.

5. A Real-World Example

In this section, a case study including 30 samples of the ceramic and tile industry, presented in Moghadam et al. (2015) is used to show how the proposed methods of monitoring the fuzzy linear profiles can be used and compared in practice. For this study, the specific form of the profile is given as follows:

$$\tilde{y}(t) = \tilde{a}e^{-\tilde{b}t} \quad (56)$$

where $\tilde{y}(t)$ is the dependent variable, \tilde{a} and \tilde{b} are the parameters and t is the independent variable of the profile. To convert the above nonlinear equation into the linear form, the logarithm transformation leads to Eq. (57).

$$\tilde{y}'(t) = \ln(\tilde{y}) = \ln(\tilde{a}) - \tilde{b}t = \tilde{c} - \tilde{b}t \quad (57)$$

We analyzed the data of this case study with the methods under investigation to compare the performance of the methods in practice. This case study cannot be used for the least squares approach of Ghobadi et al. (2014) due to the fact that fuzzy numbers are asymmetric. First, Table 2 shows the values for the goal programming and LSSVM methods. These are specific values of \widetilde{FT}^2 statistics where the area given by the membership function is halved. Finally, the values of T^2 statistic for D'Urso's method are given for each sample.

Table 2
Result of the case study for methods

row	LSSVM	Goal programing	D'Urso	row	LSSVM	Goal programing	D'Urso
1	0.413	0.123	0.126	16	0.914	0.399	2.592
2	0.432	0.207	0.305	17	0.995	0.440	2.683
3	0.451	0.123	0.126	18	0.787	0.321	2.574
4	0.413	0.123	0.345	19	0.787	0.321	2.574
5	0.449	0.111	0.085	20	0.924	0.399	2.783
6	0.623	0.189	0.532	21	0.883	0.373	2.788
7	0.623	0.189	0.532	22	1.301	0.637	4.704
8	0.463	0.207	0.430	23	0.924	0.399	2.783
9	0.389	0.386	0.148	24	1.301	0.637	4.704
10	0.413	0.123	0.126	25	0.792	0.326	2.093
11	0.448	0.144	0.194	26	1.101	0.494	3.961
12	0.434	0.255	0.160	27	0.868	0.351	2.365
13	0.427	0.111	0.001	28	0.726	0.268	2.141
14	0.403	0.207	0.182	29	1.113	0.494	4.139
15	0.380	0.069	0.029	30	1.113	0.494	4.139

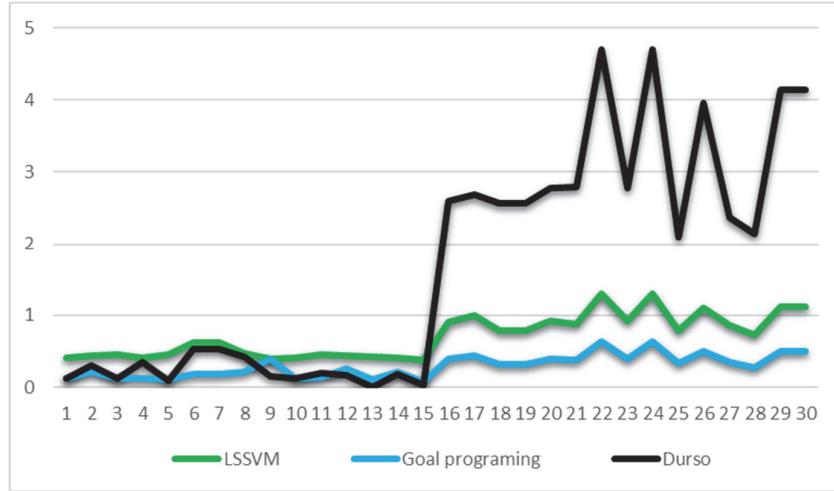


Fig. 2. Performance of the methods

Fig. 2 shows the performance of all methods. Considering a step shift occurred in the process starting from point 16, as shown by Fig. 2, the D'Urso's method indicates superior performance in detecting the persistent process change compared with LSSVM and goal programming approaches.

6. Evaluation and comparison study

In general, the effectiveness of monitoring methods in phase II of SPC is evaluated by the sensitivity of the method with regard to the changes in the process parameters represented usually as a multiple of the process standard deviation. Thus, the performance of the methods is examined in a simulation study by exerting changes in process parameters and calculating detectability of the corresponding control chart. In our numerical example, the given form of the following fuzzy linear regression is considered as follows:

$$\tilde{Y}_i = \tilde{a} + \tilde{b}X_i \quad (58)$$

where, coefficients are known and specified as $\tilde{a} = (20, 0.2)$, $\tilde{b} = (2, 0.2)$ and $X_i = \{6, 7, \dots, 14\}$. We added the random value of the error (\tilde{e}_i) in order to estimate the equation above:

$$\tilde{Y}_i = \tilde{a} + \tilde{b}X_i + \tilde{e}_i \quad (59)$$

in which, $\tilde{e}_i = (e_i, s_{e_i})$. In this section, the performance of the corresponding control charts is compared using the average run length (ARL) criterion which shows the average number of points on the control chart until an out-of-control signal appears.

The method of generating random data needed for our simulation study is as follows. This is the generalized method of the Su et al. (2013) method for this study. As mentioned, the linear profile model under consideration is $\tilde{Y}_i = \tilde{a} + \tilde{b}X_i + \tilde{e}_i$, where fixed values of 6, 7, ..., 14 are used for X . Fuzzy parameters are quantified as $\tilde{a} = (20, 0.2)$, $\tilde{b} = (2, 0.2)$ and finally, $\tilde{e}_i = (e_i, s_{e_i})$ as the random error term is assumed to follow a standard normal distribution with s_{e_i} which is uniformly distributed over the interval [0, 2]. 10000 Simulation runs are considered in order to fix the upper limit of the control chart for the \widetilde{FT}^2 control statistic. Then, for each shift in intercept and slope coefficients, the ARL and standard deviation of run length (SDRL) are calculated for the methods introduced in section 5.

6.1. Shift in model intercept

Table 3 shows the simulation results for out of control conditions produced through step shifts of the intercept parameter. For each method, the amount of ARL and SDRL are given considering various shift sizes. Fig. 3 indicates that D'Urso's approach exhibits a much better performance in detecting of the intercept compared with the rest of the methods.

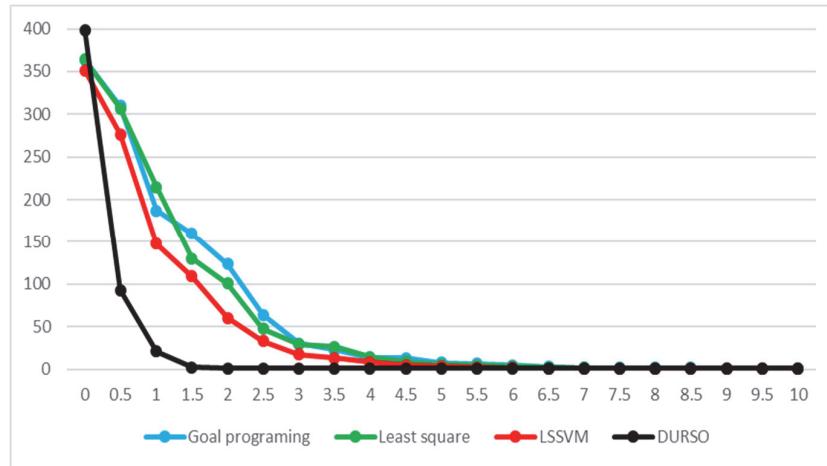


Fig. 3. Performance (ARL) of methods under shifts in intercept

6.2. Shift in model slope

Table 4 shows the simulation results with regard to the various shifts produced by changing the slope coefficient. For each method, the values of ARL and SDRL are given. Fig. 4 shows that the D'Urso's approach has a much better performance in detecting shifts created in the slope than LSSVM, goal programming, and least square methods.

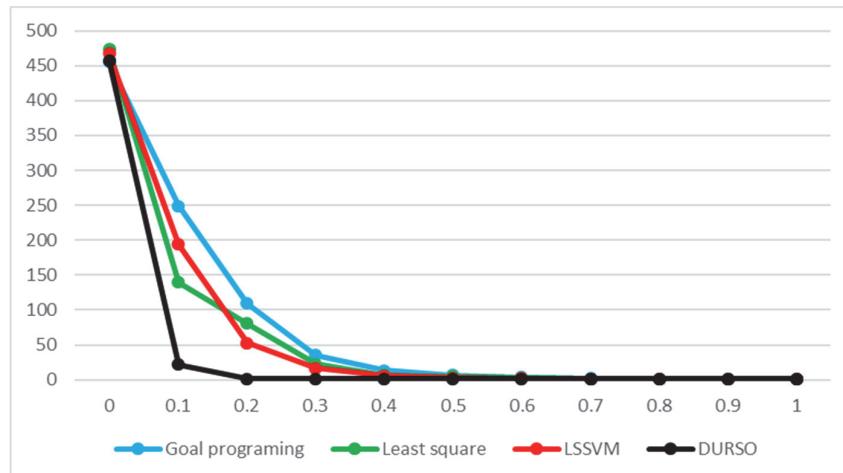


Fig. 4. Performance (ARL) of methods under shifts in slope

Table 3

shift	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10	
Goal	ARL	363	309	187	160	123	63.8	29.8	22.9	13.9	13.2	7.7	6.8	4.6	2.9	1.6	1.8	1.4	1.1	1	1	
programing	SDRL	369	284	177	187	153	52.3	23.6	28.2	11.2	11.1	5.8	5.8	3.4	2.6	0.8	0.9	1.0	0.6	0.3	0	0
Least square	ARL	364	306	214	130	100	47.4	29.2	26.6	14.4	8.8	5.8	4.7	3.6	2.3	1.4	1.1	1	1	1	1	
LSSVM	SDRL	397	364	153	59	93.3	45.6	21.7	18.4	9.1	7.5	5.8	3.4	2.6	1.6	0.7	0.3	0	0	0	0	
D'Urso	ARL	351	276	147	109	60.4	32.8	17.4	13.3	8.2	4.8	3.6	2.6	1.7	1.3	1.2	1	1	1	1	1	
	SDRL	370	254	154	125	65.3	28.1	15.8	13.9	8.1	3.7	3.1	2.0	1.1	0.7	0.5	0	0	0	0	0	

Table 4

7. Conclusions

In this study, the performances of four methods for monitoring fuzzy linear profile in Phase II were discussed and compared. Least square estimation and goal programming are the main two approaches employed in previous research studies in the field of monitoring fuzzy quality linear profiles. In particular, we have developed a new method on the basis of LSSVM for monitoring fuzzy linear profiles and as the last approach, a method introduced by D'Urso has been utilized for the estimation of fuzzy linear profiles. A dataset from a real-world case study in the tile and ceramic industry was investigated to compare the performance of the methods. This comparison has shown that the D'Urso approach possesses a better performance than the LSSVM and the goal programming methods. The simulation study has shown that the D'Urso approach has a much better performance in detecting out-of-control process shifts than other methods and the LSSVM approach is also superior to the least square and goal programming methods.

To extend the present study, all these methods can be investigated in phase I of statistical process control and compared with each other. Moreover, the profile of interest can be chosen among other popular types such as multiple linear, multivariate or nonlinear profiles.

References

- Akbari, M.G., Mohammadalizadeh, R., & Rezaei, M. (2012). Bootstrap statistical inference about the regression coefficients based on fuzzy data. *International Journal of Fuzzy Systems*, 14(4), 549-556.
- Alaeedini, A., Ghazanfari, M., & Nayeri, M. A. (2009). A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts. *Information Sciences*, 179(11), 1769-1784.
- Celmin's, A. (1987). Least-squares model fitting of fuzzy vector data. *Fuzzy Sets and Systems*, 22(3), 245-269.
- Chen, J., & Liang, Y. (2016). Development of fuzzy logic-based statistical process control chart pattern recognition system. *The International Journal of Advanced Manufacturing Technology*, 86(1-4), 1011-1026.
- Cheng, C.B. (2005). Fuzzy process control: construction of control charts with fuzzy numbers. *Fuzzy Sets and Systems*, 154(2), 287-303.
- Coppi, R., D'Urso, P., Giordani, P., & Santoro, A. (2006). Least squares estimation of a linear regression model with LR fuzzy response. *Computational Statistics & Data Analysis*, 51(1), 267-286.
- D'Urso, P. (2003). Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. *Computational Statistics & Data Analysis*, 42(1-2), 47-72.
- Demirli, K., & Vijayakumar, S. (2010). Fuzzy logic based assignable cause diagnosis using control chart patterns. *Information Sciences*, 180(17), 3258-3272.
- El-Shal, S.M., & Morris, A.S. (2000). A fuzzy rule-based algorithm to improve the performance of statistical process control in quality systems. *Journal of Intelligent and Fuzzy Systems*, 9(3-4), 207-223.
- Fernández, M. N. P. (2017, July). Fuzzy theory and quality control charts. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1-6). IEEE.
- Franceschini, F., & Romano, D. (1999). Control chart for linguistic variables: A method based on the use of linguistic quantifiers. *International Journal of Production Research*, 37(16), 3791-3801.
- Ghobadi, S., Noghondarian, K., Noorossana, R., & Mirhosseini, S.M.S. (2014). Developing a multivariate approach to monitor fuzzy quality profiles. *Quality & Quantity*, 48(2), 817-836.
- Gulbuy, M., Kahraman, C., & Ruan, D. (2004). α -Cut fuzzy control charts for linguistic data. *International Journal of Intelligent Systems*, 19(12), 1173-1195.
- Hassanpur, H., Maleki, H.R., & Yaghoobi, M.A. (2009). A goal programming approach for fuzzy linear regression with nonfuzzy input and fuzzy output data. *Asia Pacific Journal of Operational Research*, 26(5), 1-18.
- Hong, D. H., & Hwang, C. (2006, September). Fuzzy nonlinear regression model based on LS-SVM in feature space. In *International Conference on Fuzzy Systems and Knowledge Discovery* (pp. 208-216). Springer, Berlin, Heidelberg., https://doi.org/10.1007/11881599_23
- Kanagawa, A., Tamaki, F., & Ohta, H. (1993). Control charts for process average and the variability based on linguistic data. *International Journal of Production Research*, 31(4), 913-922.
- Kang, L., & Albin, S.L. (2000). On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, 32(4), 418-426.
- Kim, K., Mahmoud, M.A., & Woodall, W.H. (2003). On the monitoring of linear profiles. *Journal of Quality Technology*, 35(3), 317-328.
- Korner, R., & Nather, W. (1998). Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Information Sciences*, 109(1-4), 95-118.
- Mahmoud, M.A., Parker, P.A., Woodall, W.H., & Hawkins, D.M. (2007). A change point method for linear profile data. *Quality and Reliability Engineering International*, 23(2), 247-268.
- Moghadam, G., Raissi Ardali, G.A., & Amirzadeh, V. (2015). Developing new methods to monitor phase II fuzzy linear profiles. *Iranian Journal of Fuzzy Systems*, 12(4), 59-77.
- Moghadam, G., Raissi Ardali, G.A., & Amirzadeh, V. (2016). New fuzzy EWMA control charts for monitoring phase II fuzzy profiles. *Decision Science Letters*, 5(1), 119-128.
- Moghadam, G., Raissi Ardali, G.A., & Amirzadeh, V. (2018). A novel phase I fuzzy profile monitoring approach based on

- fuzzy change point analysis. *Applied Soft Computing*, 71, 488-504.
- Montgomery, D.C. (2009). *Introduction to Statistical Quality Control* (6th ed.). New York: John Wiley & Sons.
- Noghondarian, K., & Ghobadi, S. (2012). Developing a univariate approach to phase I monitoring of fuzzy quality profiles. *International Journal of Industrial Engineering Computations*, 3(5), 829-842.
- Noorossana, R., Saghaei, A., & Amiri, A. (2011). *Statistical analysis of profile monitoring* (Vol. 865). John Wiley & Sons.
- Raz, T., & Wang, J.H. (1990). Probabilistic and membership approaches in the construction of control charts for linguistic data. *Production Planning & Control*, 1(3), 147-157.
- Sabegh, M.H.Z., Mirzazadeh, A., Salehian, S., & Weber, G.W. (2014). A literature review on the fuzzy control chart; classifications & analysis. *International Journal of Supply and Operations Management*, 1(2), 167-189.
- Senturk, S., & Antucheviciene, J. (2017). Interval Type-2 Fuzzy c-Control Charts: An Application in a Food Company. *Informatica*, 28(2), 269-283.
- Senturk, S., Erginol, N., Kaya, I., & Kahraman, C. (2014). Fuzzy exponentially weighted moving average control chart for univariate data with a real case application. *Applied Soft Computing*, 22, 1-10.
- Su, Z., Wang, P., & Song, Z. (2013). Kernel-based nonlinear fuzzy regression model. *Engineering Applications of Artificial Intelligence*, 26(2), 724-738.
- Taheri, S.M., & Arefi, M. (2009). Testing fuzzy hypotheses based on fuzzy test statistic. *Soft Computing*, 13(6), 617-625.
- Taleb, H., & Limam, M. (2002). On fuzzy and probabilistic control charts. *International Journal of Production Research*, 40(12), 2849-2863.
- Tanaka, H., Uejima, S., & Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Transaction of Systems and Man Cybernetics*, 12(6), 903-907.
- Tanaka, H. (1987). Fuzzy data analysis by possibilistic linear models. *Fuzzy Sets and Systems*, 24(3), 363-375.
- Viertl, R. (2011). Statistical methods for fuzzy data (1st ed.). Austria: John Wiley & Sons.
- Wang, D., Li, P., & Yasuda, M. (2014). Construction of Fuzzy Control Charts Based on Weighted Possibilistic Mean. *Communications in Statistics- Theory and Methods*, 43(15), 3186-3207.
- Zarandi, M.H.F., & Alaeddini, M. (2010). A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts. *Information Sciences*, 180(16), 3033-3044.
- Zou, C., Zhang, Y., & Wang, Z. (2006). A control chart based on a change-point model for monitoring linear profiles. *IIE Transactions*, 38(12), 1093-1103.



© 2021 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).