A new neural network-based control scheme for fault detection and fault diagnosis in fuzzy multivariate multinomial data

Mohammad Reza Maleki,
Seyed Meysam Mousavi* and
Amirhossein Amiri

Industrial Engineering Department,
Faculty of Engineering,
Shahed University,
Tehran, Iran
Email: reza.maleki@shahed.ac.ir
Email: sm.mousavi@shahed.ac.ir
Email: amiri@shahed.ac.ir
*Corresponding author

Abstract: In some multivariate statistical control applications, the data of the process cannot be precise and defined linguistically in practice. Using multivariate control charts in such situations with non-precise data leads to misleading results. In this paper, a new neural network-based monitoring scheme is presented by considering fuzzy multivariate multinomial data. The proposed approach is also able to identify the attribute(s) that cause an out-of-control signal. An application example is provided to evaluate the performance of the proposed approach in detecting different shifts as well as diagnosing the out-of-control attribute quality characteristic(s). The results of applying the proposed approach in both fault detection and the fault diagnosis are satisfactory.

Keywords: linguistic form; fuzzy sets theory; multi-attribute process; multilayer perceptron neural network; detection and diagnosis.

Reference to this paper should be made as follows: Maleki, M.R., Mousavi, S.M. and Amiri, A. (2015) 'A new neural network-based control scheme for fault detection and fault diagnosis in fuzzy multivariate multinomial data', *Int. J. Applied Decision Sciences*, Vol. 8, No. 2, pp.127–144.

Biographical notes: Mohammad Reza Maleki is a PhD student of Industrial Engineering at Shahed University in Iran. He received his BS and MS in Industrial Engineering from Isfahan University of Technology and Shahed University, respectively. His research interests are statistical quality control, artificial neural network and fuzzy theory.

Seyed Meysam Mousavi is an Assistant Professor at the Department of Industrial Engineering, Faculty of Engineering, Shahed University in Iran. He received his PhD degree from the Department of Industrial Engineering at the University of Tehran, Iran, and is currently a member of Iran's National Elite Foundation. His main research interests include: cross-docking systems planning, quantitative methods in project management, engineering optimization under uncertainty, facilities planning and design, multiple criteria decision making under uncertainty, soft computing, applied artificial intelligence and meta-heuristics algorithms. He has published many papers and book chapters in reputable journals and international conference proceedings.

Amirhossein Amiri is an Assistant Professor at Shahed University in Iran. He received his BS, MS and PhD in Industrial Engineering from Khajeh Nasir University of Technology, Iran University of Science and Technology, and Tarbiat Modares University, respectively. He is now the Head of the Industrial Engineering Department at Shahed University in Iran. He has published many papers in the area of statistical process control in high quality international journals and a book with John Wiley and Sons in 2011 entitled *Statistical Analysis of Profile Monitoring*. His research interests are statistical quality control, profile monitoring and change point estimation.

1 Introduction

Nowadays incorporating fuzzy concepts in various areas in industrial engineering is noticed by researchers. Some recent papers are presented by Oke et al. (2009) in maintenance scheduling, Khalili-Damghani et al. (2013) and Kazemi et al. (2014b) in supplier selection, Shah and Soni (2011) and Sharifi et al. (2015) in inventory control.

Incorporating fuzzy theory in statistical process control (SPC) is inevitable because in many real production environments, the data corresponding to the quality of the product or process is expressed in the presence of uncertainty. The presence of uncertainty in the process data is due to two concepts, including randomness and non-preciseness. In the second case, fuzzy sets theory and fuzzy logic can be combined to quality control tools for monitoring the process mean as well as the process variability. In recent years, some efforts have been done by quality engineers to fulfil this purpose. Raz and Wang (1990) introduced two approaches for constructing control charts in the case that the process observations are expressed in linguistic form. Taleb and Limam (2002) suggested different procedures of constructing control charts for linguistic data based on fuzzy and probability theories. Gülbay et al. (2004) developed some α-cut control charts for attributes; they also determined the tightness of the inspection by selecting a suitable α-level. Gülbay and Kahraman (2006) proposed fuzzy control charts for monitoring the attributes under vague data; further, they defined unnatural patterns for the proposed control charts using probabilities. Gülbay and Kahraman (2007) focused on fuzzy control charts based on fuzzy transformation methods by the use of α-cut in order to provide the ability of determining the tightness of the inspection. Zarandi et al. (2008) presented a hybrid methodology based on the combination of fuzzified sensitivity criteria and fuzzy adaptive sampling rules. They pointed out that using this methodology makes the control charts more sensitive and proactive while keeping false alarms rate acceptably low.

Erginel (2008) constructed fuzzy control limits for individual (X) and moving range (MR) control charts using α-cuts. Engin et al. (2008) proposed a fuzzy approach for attribute control charts in multi-stage processes and then solved it by genetic algorithm. Senturk and Erginel (2009) investigated the construction of the fuzzy $\overline{X}/R$ and $\overline{X}/S$ control charts using α-cuts. They found that the flexibility of traditional control limits increases when fuzzy criterion is used. Alaeddini et al. (2009) proposed a hybrid approach based on the combination of fuzzy clustering and statistical methods in order to estimate the change point in processes with both fixed and variable sample sizes.

Ertuğrul and Aytaç (2009) constructed control charts based on fuzzy sets theory by considering the quality in terms of grades of conformance as opposed to absolute conformance and non-conformance. Then, they presented the control chart based on probability theory. They also compared the results of control charts based on two different approaches and found that the fuzzy sets theory performs better than probability theory in monitoring the product quality. Amirzadeh et al. (2009) proposed a fuzzy *p*-chart based on the mean degree of non-conformity. To do that, they used the degree of non-conformity based on fuzzy concepts instead of considering an item to be either conforming or non-conforming. Shu and Wu (2010) constructed a *p* control chart for monitoring the fraction of non-conforming items in the manufacturing processes with fuzzy sample data.

Faraz and Shapiro (2010) constructed a fuzzy statistical control chart in order to explain fuzziness in data with considering the essential variability between observations. Zarandi and Alaeddini (2010) then developed a general fuzzy statistical clustering (FSC) approach in order to estimate the time of change in different control charts with either fixed or variable sampling strategy. They also provided a new objective function and examined its relation with maximum likelihood estimator. Alizadeh et al. (2010) incorporated the multivariate control charts in a fuzzy environment. To do that, they assumed that each observation of samples is a canonical fuzzy number. Demirli and Vijayakumar (2010) developed a rule based fuzzy inference system for $\bar{X}$ control chart to prioritise the assignable causes based on the accumulated evidence. M'Hallah and Melloy (2010) introduced an adaptive economic control scheme for a fuzzy quality characteristic. Alizadeh and Ghomi (2011) developed mean and range control charts in fuzzy environment using different transformation methods. They assumed that the observations of each sample are fuzzy random variables with triangular membership functions.

Shu and Wu (2011) proposed the fuzzy $\bar{X}$ and *R* control charts whose fuzzy control limits are obtained based on the resolution identity, a well-known theory in the fuzzy set field. Kaya and Kahraman (2011) incorporated the fuzzy sets theory to process capability analysis. They also derived fuzzy control charts for fuzzy measurements of the related quality characteristic. Shu and Wu (2012) proposed a constructive methodology for obtaining the fuzzy estimate of loss-based capability index (C_{pm}) based on resolution identity in fuzzy sets theory. They proposed four decision rules to judge the process state by simultaneous introducing randomness and fuzziness. M'Hallah and Melloy (2012) investigated the impact of prior information on the total cost of adaptive control charts with fuzzy quality characteristic using a dynamic quality control scheme.

Tong and Wang (2012) designed fuzzy sampling plans for quality inspection of geospatial mineral products with ambiguity in process quality characteristics. Kazemi et al. (2014a) proposed a FSC method in order to estimate the drift time in different processes. Shu et al. (2014a) proposed a demerit-fuzzy rating system and monitoring scheme in manufacturing processes. Shu et al. (2014b) developed the maximum generally weighted moving average (MaxGWMA) control chart in a fuzzy environment. Wang et al. (2014) employed the weighted possibilistic mean (WPM) and weighted interval valued possibilistic mean (WIVPM) in environments with fuzzy attribute data and constructed fuzzy control charts with WPM and WIVPM. Pastuizaca Fernández et al.

(2014) first reviewed the studies in the literature on the development of fuzzy multivariate control charts. Then, they proposed a method to control the fuzzy multiple attribute quality characteristics using hotelling T^2 control chart.

Nowadays, the application of neural networks (NNs) is noticed as an effective alternative of control charts due to their satisfactory performance; however, we can conclude from the related literature that only few efforts have been made regarding the combination of NNs and fuzzy sets theory. These few papers in the scope of using NNs in monitoring the processes whose data are expressed in the presence of non-preciseness, such as linguistic and human dependent data are reviewed. Chang and Aw (1996) proposed a neural fuzzy control chart for identifying process mean shifts. They designed a supervised multi-layer back-propagation NN to detect various mean shifts in a production process. In identifying mean shifts in real-time usage, they classified the NNs' outputs into various decision regions using a fuzzy set scheme. Yang and Yang (2002) presented a supervised competitive learning network approach, called a fuzzy-soft learning vector quantisation, for the control chart pattern recognition. Wang and Chen (2002) proposed a two-module method for detecting mean shifts in multivariate processes as well as classifying the magnitude of these shifts. In the first module, they used an artificial NN to detect different shifts in the multivariate process mean. Then, in the second module, in order to identify the magnitude of shifts, they classified the outputs of NN into different decision intervals by fuzzy classifier and an additional two-point-in-an-interval decision rule.

On the other hand, only few methods are available in the related literature about fuzzy multi-attribute procedures. Taleb et al. (2006) suggested two control charts for monitoring multi-attribute processes. They developed two statistics including T_f^2 and W^2 based on fuzzy and probability theories. Alipour and Noorossana (2010) developed a fuzzy multi-attribute exponentially weighted moving average control chart.

This paper combines artificial NN and fuzzy concepts in order to monitor multivariate multinomial attributes which are presented linguistically. For example in a textile process, there can be three linguistic attribute quality characteristics including appearance, colour as well as material type of the produced units which should be monitored simultaneously. In order to monitor the linguistic multivariate multinomial quality characteristics, firstly, we estimate parameters of the multinomial attributes. Then, in phase 2, an artificial NN-based approach is proposed based on the estimated parameters of linguistic attributes. The proposed neural-fuzzy approach not only can detect various shifts in the process, but also is able to diagnose the attribute or the group of attributes responsible for out-of-control signals. Notably that we refer the task of detecting different shifts and identifying attribute(s) that cause out-of-control signals as fault detection and fault diagnosis, respectively.

This paper is organised as follows: in Section 2, a multivariate multinomial process whose data are expressed linguistically is defined. In Section 3, the proposed approach for detection and diagnosis purposes is illustrated. An application example based on simulation is given in Section 4 in order to evaluate the performance of the proposed approach. Finally, conclusion and recommendations for future studies are discussed in Section 5.

2 Problem definition

Consider a process in which p linguistic attribute quality characteristics including x_1, $x_2,\ldots,x_p$ should be monitored simultaneously. These attribute quality characteristics are expressed as words or phrases rather than crisp numerical values. In such situations, human dependent terms, such as standard, acceptable and non-acceptable, are used to characterise the quality of products. It is supposed that for any attribute quality characteristic x_j, $j = 1,\ldots,p$, there are q_j linguistic terms that are expressed thorough fuzzy subsets where each fuzzy subset is related to a membership function. The term set $T(x_j)$ is used in fuzzy sets theory for representing linguistic attribute x_j. The membership function corresponding to the linguistic attribute x_j in linguistic term h, is denoted by $\mu_{jh}(u)$, where u is the standardised value of quality level. The values of standardised quality levels belong to the range of [0, 1], which 0 and 1 represent the best and the worst quality levels, respectively.

Let a taken sample with n observations is denoted as follows:

$$S = \left\{ \left\{ \left(F_{11}, n_{11}\right), \ldots, \left(F_{1q_1}, n_{1q_1}\right) \right\}; \ldots; \left\{ \left(F_{p1}, n_{p1}\right), \ldots, \left(F_{pq_p}, n_{pq_p}\right) \right\} \right\}, \tag{1}$$

where n_{jh} is the number of observations related to attribute x_j and h^{th} corresponding category and F_{jh} is the related fuzzy set. Then, for each attribute quality characteristic x_j, $j = 1,\ldots,p$ the fuzzy subset $\{(F_{j1}, n_{j1}),\ldots,(F_{jq_j}, n_{jq_j})\}$ is associated with one fuzzy subset using the following equation:

$$F_j = \frac{1}{n} \sum_{h=1}^{q_j} n_{jh} F_{jh}. \tag{2}$$

It is assumed that fuzzy variables F_{jh} are triangular in this paper. Hence, it is obvious that F_j remains a triangular fuzzy number after using equation (2). In order to transform the fuzzy subset F_j, $j = 1,\ldots,p$ and calculate its corresponding representative value, the median transformation approach is used according to equation (3):

$$R_j = \begin{cases} a_{3j} - \sqrt{\dfrac{\left(a_{3j} - a_{1j}\right)\left(a_{3j} - a_{2j}\right)}{2}}, & \text{for} \quad a_{2j} < \dfrac{a_{3j} + a_{1j}}{2} \\ a_{1j} + \sqrt{\dfrac{\left(a_{3j} - a_{1j}\right)\left(a_{2j} - a_{1j}\right)}{2}}, & \text{for} \quad a_{2j} > \dfrac{a_{3j} + a_{1j}}{2} \end{cases}. \tag{3}$$

The equation (1) can be replaced with representative column vector of $\mathbf{R_S}$ as follows:

$$\mathbf{R_S} = \left(R_{S1}, R_{S2}, \ldots, R_{Sp}\right)' \tag{4}$$

The representative matrix **R** which is computed based on m samples with sample size of n is expressed as follows:

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1p} \\ R_{21} & R_{22} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ R_{m1} & \cdots & \cdots & R_{mp} \end{pmatrix}, \tag{5}$$

where R_{ij} is the representative value corresponding to attribute x_j in the i^{th} sample taken.

3 Proposed approach

In this section, the proposed NN-based approach for monitoring the linguistic multivariate multinomial quality characteristics is discussed. In Sub-section 3.1, the proposed NN structure is defined. Then in Sub-section 3.2, the proposed training algorithm as well as the proposed method for determining the threshold values of the designed NN is described.

3.1 NN structure

One of the most commonly architectures that is used in different scopes of statistical quality control is the multilayer perceptron neural network (MLP), first proposed by Rumelhart et al. (1986). It is pointed out in the literature that in many situations multilayer NNs outperform the performance of the statistical methods and consequently they are considered as the effective alternatives of control charts. In this Sub-section, the proposed neural-fuzzy approach that is able to detect different shifts in the multivariate multinomial fuzzy processes as well as diagnose attribute(s) responsible for the out-of-control signals is illustrated. For this purpose, the MLP is suggested with the following structure:

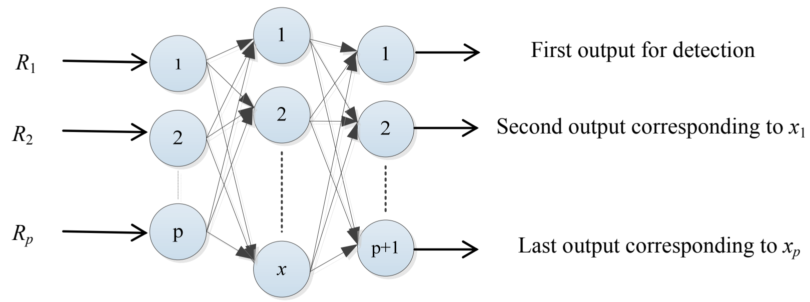
- Input layer: the number of nodes in input layer of the proposed NN is equal to the number of process attributes. Hence, in a process whose quality is represented through p fuzzy attribute quality characteristics; there will be p nodes in its input layer. The input vector of the proposed NN is the column vector of $(R_{i1},\ldots, R_{ij},\ldots, R_{ip})'$ where R_{ij} is the representative value corresponding to attribute x_j in the i^{th} sample.
- Output layer: the number of nodes in the output layer is equal to the number of attributes plus one. Hence in a p-attribute process, $p + 1$ nodes will exist in the output layer of the designed NN. The first output node is used to decide whether the process is in-control or not, whereas other output nodes are used to diagnose the attribute or the group of attributes responsible for out-of-control signals.
- Hidden layers: it is pointed out in the literature that there are no systematic guidelines to determine the number of hidden layers as well as the number of nodes in each hidden layer. Hence, a trial and error approach is usually applied to determine these values.
- Activation function: the most widely used activation functions are sigmoid and hyperbolic tangent functions, respectively. In the proposed NN, a sigmoid function

whose output values are in the range of 0 to 1 is used according to the following equation:

$$f(n) = \frac{1}{1 + e^{-cn}} c > 0. \tag{6}$$

- Training algorithm: the multi-layer perceptron NNs use supervised training rule named feed-forward back-propagation algorithm.
- Evaluation criterion: the mean squared error (MSE) criterion is used to evaluate the designed NN in training step as the performance function. Figure 1 represents the architecture of the proposed NN.

Figure 1 Proposed MLP architecture (see online version for colours)



3.2 Training method

In this section, the proposed procedure of training the NN is illustrated as follows:

Consider a process in which the linguistic data come from a p-attribute distribution and the sample size of n is used for monitoring the process. In the in-control situations, each attribute quality characteristic x_j, $j = 1,\ldots,p$ follows a multinomial distribution with parameters n and $\mathbf{p_j}$, where $\mathbf{p_j}$ (i.e., the vector of probabilities) is defined as follows:

$$\mathbf{p_j} = \left(p_{j1}, \ldots, p_{jh}, \ldots, p_{jq_j}\right). \tag{7}$$

In equation (7), p_{jh} is the probability of classifying in-control attribute x_j into h^{th} category. For out-of-control situations, the vector of probabilities corresponding to out-of-control attribute(s) x_j, $j = 1,\ldots,p$ is characterised according to equation (8):

$$\mathbf{p'_j} = \left(p'_{j1}, \ldots, p'_{jh}, \ldots, p'_{jq_j}\right), \tag{8}$$

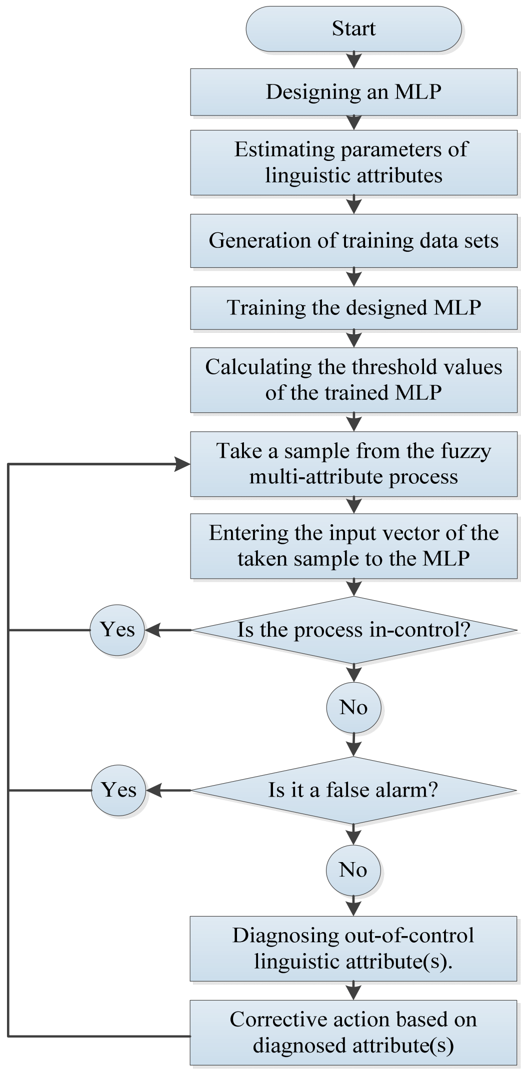
where p'_{jh} is the probability of classifying out-of-control linguistic attribute x_j into h^{th} category. In order to train the desired MLP, the input data as well as their corresponding target values in both in-control and out-of-control situations should be entered to the NN. Note that in a process in which the linguistic data come from a p-attribute distribution, there are totally $2^p - 1$ out-of-control situations. In the proposed training procedure, firstly for each out-of-control situation, we generate 100 dataset of sample size n. After that the in-control datasets are prepared as equal to the total out-of-control datasets. Hence, in a p-attribute process, the number of in-control and out-of-control generated random samples are both equal to 100 × (2^p – 1). The matrix of probabilities corresponding to each in-control and out-of-control samples with size of n are denoted according to equations (9) and (10), respectively:

$$\mathbf{p} = \left(\mathbf{p}_1\ \mathbf{p}_2 \ldots \mathbf{p}_p\right)' \tag{9}$$

$$\mathbf{p}' = \left(\mathbf{p}_1'\ \mathbf{p}_2' \ldots \mathbf{p}_p'\right)' \tag{10}$$

After generating all training datasets, the column vector of $(R_1, R_2,\ldots,R_p)'$ for each sample is calculated as the input vectors of the designed MLP. The target values of the NN are the column vectors each of size $p + 1$ whose elements are zero or one.

Figure 2 Flowchart of the proposed approach (see online version for colours)



The elements of target vector corresponding to in-control input vectors are all considered equal to zero. For out-of-control input vectors, the first element of the target vector as well as the elements corresponding to the out-of-control linguistic attributes is considered equal to one. In out-of-control input vectors of the NN, the elements corresponding to the in-control attributes are regarded equal to zero. After generating the training datasets including the input vectors as well as the corresponding target vectors of input vectors in both in-control and out-of-control situations, the MLP is trained using back-propagation algorithm. The flowchart of the proposed approach is depicted in Figure 2.

After designing and training steps, the values of output neuron thresholds should be determined. The first output neuron threshold is used for detecting the process state and outers are used for diagnosing which attribute(s) is responsible for the out-of-control signals. In order to calculate the first output threshold, firstly, we generate 10,000 in-control random samples through simulation and enter them to the trained MLP. Then, we sort the observed values of first output neuron in an ascending order in a vector like $\mathbf{a_1}$. Finally, we consider the element of $\mathbf{a_1}$ as the first threshold such that the value of in-control average run length (ARL_0) is obtained equal to a predetermined value.

For diagnosis purposes, in order to calculate the value of output threshold i, $i = 2, 3\ldots, p + 1$ the following steps are proposed:

1 Generating 10,000 out-of-control random samples in which the linguistic attribute corresponding to the i^{th} output neuron ($i - 1^{\text{th}}$ attribute) is out-of-control. Note that, in this state the other linguistic attributes are in-control.

2 Entering the prepared datasets to the designed MLP.

3 Sorting the observed values of i^{th} output neuron in an ascending order in a vector like $\mathbf{a_i}$.

4 Finally, the i^{th} threshold is the element of vector $\mathbf{a_i}$ that 90% of data exceed it.

4 Simulation study

In this section, the performance of the proposed neuro-fuzzy approach in both fault detection as well as fault diagnosis purposes is evaluated. To do that, we use the dataset of a food industry process that is first presented by Taleb et al. (2006). In the presented datasets, three attribute quality characteristics including appearance (x_1), colour (x_2) and taste (x_3) of the produced units are important and should be monitored simultaneously. The linguistic terms as well as the membership functions of each linguistic attribute quality characteristic are summarised in Table 1.

The datasets of $m = 20$ samples each of size $n = 220$ are summarised in Table A1 in Appendix.

Table 1 Fuzzy set relative to linguistic attributes

Attribute	*Term set*	*Linguistic term*	*Membership function*
Appearance	x_{11}	Good	[0, 0, 0.25]
	x_{12}	Medium	[0, 0.25, 0.75]
	x_{13}	Poor	[0.25, 1, 1]
Colour	x_{21}	Standard	[0, 0, 0.5]
	x_{22}	Acceptable	[0, 0.5, 0.75]
	x_{23}	Rejected	[0.5, 1, 1]
Taste	x_{31}	Perfect	[0, 0, 0.25]
	x_{32}	Good	[0, 0.25, 0.75]
	x_{33}	Medium	[0.25, 0.75, 1]
	x_{34}	Poor	[0.75, 1, 1]

Source: Taleb et al. (2006)

Table 2 Representative values of the samples

Sample	R_{j1}	R_{j2}	R_{j3}	T_f^2
1	0.09	0.172	0.141	2.158
2	0.089	0.171	0.124	3.241
3	0.098	0.179	0.138	0.263
4	0.091	0.171	0.146	3.965
5	0.093	0.192	0.122	8.247
6	0.092	0.172	0.127	1.945
7	0.096	0.178	0.134	0.12
8	0.099	0.178	0.139	0.314
9	0.102	0.185	0.134	0.314
10	0.128	0.185	0.133	13.74
11	0.121	0.196	0.154	9.287
12	0.101	0.183	0.145	2.182
13	0.107	0.193	0.131	2.995
14	0.110	0.193	0.133	2.659
15	0.095	0.169	0.136	2.274
16	0.098	0.183	0.132	0.534
17	0.097	0.180	0.131	0.279
18	0.092	0.174	0.134	0.639
19	0.096	0.174	0.131	1.173
20	0.092	0.175	0.136	0.673
Average	0.099	0.18	0.135	

As noted, in the in-control state each linguistic attribute quality characteristic x_j, j = 1,2,3 follows a multinomial distribution with parameters n = 220 and $\mathbf{p_j}$. In order to estimate $\mathbf{p_j}$ first using equation (2), each sample m, m = 1,2,…,20 is converted to the fuzzy subset of

$F = (F_1, F_2, F_3)$ which in turn is comprised of three fuzzy subsets of F_1, F_2 and F_3. Then, using equation (3), each fuzzy subset F_j of any sample is transformed into its corresponding representative value. After that, the value of statistic T_f^2 that is proposed by Taleb et al. (2006) corresponding to each sample is computed. The transformed values of each samples as well as their corresponding statistic T_f^2 are summarised in Table 2.

Taleb et al. (2006) calculated the value of upper control limit of T_f^2 control chart equal to 8.121 based on the false alarm rate of 0.05. Hence, the samples 5, 10 and 11 fall outside the upper control limit. After eliminating these out-of-control samples, the in-control vectors $\mathbf{P_j}$, $j = 1, 2, 3$ are estimated equal to $\mathbf{p}_1 = (0.942, 0.035, 0.023)$, $\mathbf{p}_2 = (0.925, 0.045, 0.030)$ and $\mathbf{p}_3 = (0.774, 0.206, 0.004)$, respectively. After estimating the linguistic attributes parameters, the artificial NN should be designed for detection and diagnosis purposes in phase 2. To do that, we use an MLP NN with 3 (number of linguistic attributes) and 4 (number of linguistic attributes plus one) nodes in the input and output layers, respectively. The designed MLP also has one hidden layer whose the number of its nodes is determined through the trial and error equal to 20.

In order to train the MLP, first the in-control and out-of-control datasets should be prepared. As noted, for each out-of-control situation, 100 random datasets with sample size of 220 is generated. In this example, there are seven out-of-control states. Hence, totally 700 in-control random samples each of size n=220 is generated. After that, 700 in-control datasets each of size $n = 220$ are also generated. Finally, the designed MLP is trained with all 1,400 datasets as well as their corresponding target vectors using back-propagation algorithm. The MSE of training the MLP is obtained equal 0.00171. In order to generate out-of-control datasets during training step, the vectors of $\mathbf{p}'_1 = (0.842, 0.105, 0.053)$, $\mathbf{p}'_2 = (0.825, 0.115, 0.060)$ and $\mathbf{p}'_3 = (0.674, 0.276, 0.036, 0.014)$ are considered as the out-of-control vectors. Table 3 summarises the in-control and out-of-control training datasets, the corresponding target values of each state as well the number of dataset required for each category.

Table 3 Datasets of training the MLP

Shifted QC(s)	*Parameter of x_1*	*Parameter of x_2*	*Parameter of x_3*	*Target vector*	*Number of data*	*Process state*
-	$\mathbf{p}_1$	$\mathbf{p}_2$	$\mathbf{p}_3$	$[0, 0, 0, 0]^T$	700	In-control
x_1	$\mathbf{p}'_1$	$\mathbf{p}_2$	$\mathbf{p}_3$	$[1, 1, 0, 0]^T$	100	Out-of-control
x_2	$\mathbf{p}_1$	$\mathbf{p}'_2$	$\mathbf{p}_3$	$[1, 0, 1, 0]^T$	100	Out-of-control
x_3	$\mathbf{p}_1$	$\mathbf{p}_2$	$\mathbf{p}'_3$	$[1, 0, 0, 1]^T$	100	Out-of-control
x_1, x_2	$\mathbf{p}'_1$	$\mathbf{p}'_2$	$\mathbf{p}_3$	$[1, 1, 1, 0]^T$	100	Out-of-control
x_1, x_3	$\mathbf{p}'_1$	$\mathbf{p}_2$	$\mathbf{p}'_3$	$[1, 1, 0, 1]^T$	100	Out-of-control
x_2, x_3	$\mathbf{p}_1$	$\mathbf{p}'_2$	$\mathbf{p}'_3$	$[1, 0, 1, 1]^T$	100	Out-of-control
x_1, x_2, x_3	$\mathbf{p}'_1$	$\mathbf{p}'_2$	$\mathbf{p}'_3$	$[1, 1, 1, 1]^T$	100	Out-of-control

In order to apply the designed MLP for monitoring the multi-attribute linguistic quality characteristics, the threshold values of output neurons of the proposed MLP are determined as discussed in Sub-section 3.2. For detection purpose, the first output threshold of the designed MLP is determined such that the value of the ARL_0 becomes roughly equal to 200. To do that, the value of the first output threshold is determined equal to 0.82. As the result, the value of ARL_0 of the designed MLP will be equal to 194.4394.

The value of the second (corresponding to x_1), third (corresponding to x_2) and forth (corresponding to x_3) output thresholds that are used for diagnostic purpose are determined equal to 0.7071, 0.5442 and 0.7403, respectively. After calculating the threshold values of output neurons, the proposed MLP can be applied for detecting faults in the process as well as diagnosing the linguistic attribute(s) that causes the out-of-control signals. In order to determine the process state, we focus on the value of the first output neuron and compare it with the first threshold. If the observed value of the first output neuron is equal or less than 0.82, the process will be considered in an in-control state. In such situations, we do not notice the value of other output neurons of the designed MLP. Otherwise, if the observed value of the first output neuron is more than 0.82, the designed MLP will signal an out-of-control alarm. If an out-of-control signal occurs, the corrective action based on the linguistic attribute(s) that is responsible for out-of-control alarms should be done in order to restore the process to the in-control state. For diagnosing purposes the following rules exist:

1 If the observed value of second output neuron is more than 0.7071 and the observed value of third and fourth output neurons are less than 0.5442 and 0.7403, respectively, then x_1 is the source of out-of-control signal.

2 If the observed value of third output neuron is more than 0.5442 and the observed value of second and fourth output neurons are less than 0.7071 and 0.7403, respectively, then x_2 is the source of out-of-control signal.

3 If the observed value of fourth output neuron is more than 0.7403 and the observed value of second and third output neurons are less than 0.7071 and 0.5442, respectively, then x_3 is the source of out-of-control signal.

4 If the observed values of second and third output neurons are more than 0.7071 and 0.5442, respectively and the observed value of fourth output neuron is less than 0.7403, then both x_1 and x_2 are the sources of out-of-control signal.

5 If the observed values of second and fourth output neurons are more than 0.7071 and 0.7403, respectively and the observed value of third output neuron is less than 0.5442, then both x_1 and x_3 are the sources of out-of-control signal.

6 If the observed values of third and fourth output neurons are more than 0.5442 and 0.7403, respectively and the observed value of second output neuron is less than 0.7071, then both x_2 and x_2 are the sources of out-of-control signal.

7 If the observed values of all second, third and fourth output neurons are more than 0.7071, 0.5442 and 0.7403, respectively, then all the linguistic attributes are introduced as the sources of signals.

8 In the situations that the observed value of the first output neuron is more than 0.82, but the observed values of other output neurons are less than their corresponding thresholds, the attribute whose corresponding output neuron has the maximum value is considered as the source of out-of-control signal.

The magnitude of shifts in the in-control vectors $\mathbf{p_j}$, j = 1,2,3 are summarised in Table 4. In order to simulate out-of-control datasets, for each linguistic attribute we consider three shift magnitudes. We denote $\boldsymbol{\delta}_{\mathbf{ij}}$; i, j = 1,2,3 as the j^{th} shift magnitude in the i^{th} linguistic attribute. Note that when a shift equal to $\boldsymbol{\delta}_{\mathbf{ij}}$ in the i^{th} attribute occurs, the out-of-control vector of $\mathbf{p}'_{\mathbf{ij}}$ will be resulted. For example, vector $\boldsymbol{\delta}_{\mathbf{11}}$ = (–0.100, 0,070, 0.030) denotes the first shift magnitude in the in-control vector of the first attribute ($\mathbf{p_1}$). Consequently, the out-of-control vector $\mathbf{p}'_{\mathbf{11}} = (0.842, 0.105, 0.053)$ will be resulted after this step shift.

Table 4 Magnitude of different step shifts

Shift magnitude	*Out-of-control QC*	*Out-of-control probability vector*
$\boldsymbol{\delta}_{11}$ = (–0.100, 0.070, 0.030)	x_1	$\mathbf{p}'_{11}$ = (0.842, 0.105, 0.053)
$\boldsymbol{\delta}_{12}$ = (–0.50, 0.035, 0.015)		$\mathbf{p}'_{12}$ = (0.892, 0.070, 0.038)
$\boldsymbol{\delta}_{13}$ = (–0.150, 0.105, 0.045)		$\mathbf{p}'_{13}$ = (0.792, 0.140, 0.068)
$\boldsymbol{\delta}_{21}$ = (–0.100, 0.070, 0.030)	x_2	$\mathbf{p}'_{21}$ = (0.825, 0.115, 0.060)
$\boldsymbol{\delta}_{22}$ = (–0.50, 0.035, 0.015)		$\mathbf{p}'_{22}$ = (0.875, 0.080, 0.045)
$\boldsymbol{\delta}_{23}$ = (–0.150, 0.105, 0.045)		$\mathbf{p}'_{23}$ = (0.775, 0.150, 0.075)
$\boldsymbol{\delta}_{31}$ =(–0.100, 0.070, 0.020, 0.010)	x_3	$\mathbf{p}'_{31}$ = (0.674, 0.276, 0.036, 0.014)
$\boldsymbol{\delta}_{32}$ = (–0.50, 0.035, 0.010, 0.005)		$\mathbf{p}'_{32}$ = (0.724, 0.241, 0.026, 0.009)
$\boldsymbol{\delta}_{33}$ = (–0.150, 0.105, 0.030, 0.015)		$\mathbf{p}'_{33}$ = (0.624, 0.311, 0.046, 0.019)

Table 5 Results of proposed NN-based approach in single out-of-control attribute

Shift / *Shifted QC*	δ_{11}	δ_{12}	δ_{13}	δ_{21}	δ_{22}	δ_{23}	δ_{31}	δ_{32}	δ_{33}
x_1	9,676	9,510	9,420	1	30	0	3	55	1
x_2	9	72	1	9,785	9,650	9,741	3	48	0
x_3	0	38	0	2	52	0	9,485	9,488	9,272
x_1, x_2	179	183	373	22	63	13	0	0	0
x_1, x_3	128	187	188	0	1	0	252	206	343
x_2, x_3	0	3	0	187	203	246	254	199	375
x_1, x_2, x_3	0	7	18	3	1	0	3	4	9
Correct diagnosis	96.76	95.10	94.20	97.85	96.50	97.41	94.85	94.88	92.72
ARL_1	1.254	5.065	1.409	1.451	5.444	1.236	1.288	4.829	1.309

Table 6 Results of proposed NN-based approach under two out-of-control attributes

Shift / *Shifted QC*	$\delta_{11}, \delta_{21}, 0$	$\delta_{11}, 0, \delta_{31}$	$0, \delta_{21}, \delta_{31}$	$\delta_{12}, \delta_{22}, 0$	$\delta_{12}, 0, \delta_{32}$	$0, \delta_{22}, \delta_{32}$	$\delta_{13}, \delta_{23}, 0$	$\delta_{13}, 0, \delta_{33}$	$0, \delta_{23}, \delta_{33}$
x_1	352	607	0	2,710	3,103	15	3	8	0
x_2	890	0	511	3,624	11	2,844	36	0	8
x_3	0	186	362	11	2,233	2,812	0	1	2
x_1, x_2	8,590	11	3	3,488	86	32	9,612	2	0
x_1, x_3	2	8,639	1	45	4,360	30	0	8,016	0
x_2, x_3	5	9	9,079	21	63	4,227	0	0	9,960
x_1, x_2, x_3	161	548	44	101	144	40	349	1,973	30
Correct diagnosis	85.90	86.39	90.79	34.88	43.60	42.27	96.12	80.16	99.60
ARL_1	1.068	1.039	1.097	2.837	2.528	2.777	1.274	1.386	1.482

Table 7 Results of proposed NN-based approach under three out-of-control attributes

Shift / *Shifted QC*	$\delta_{11}, \delta_{21}, \delta_{31}$	$\delta_{12}, \delta_{22}, \delta_{32}$	$\delta_{13}, \delta_{23}, \delta_{33}$
x_1	21	1,492	0
x_2	33	648	0
x_3	5	795	0
x_1, x_2	510	2,063	4
x_1, x_3	152	1,678	0
x_2, x_3	560	1,108	283
x_1, x_2, x_3	8,719	2,216	9,713
Correct diagnosis	87.19	22.16	97.13
ARL_1	1.087	2.445	1.155

The results of simulation experiments based on 10,000 runs under different out-of-control situations are summarised in Tables 5, 6 and 7. The last row of Tables 5–7 represents the performance of the proposed neuro-fuzzy method in detecting different step shifts (fault detection) in terms of out-of-control average run length (ARL_1) criterion. The simulation results in rows 1–7 of Tables 5–7 represent the number of simulation runs in which the proposed neuro-fuzzy method diagnose the corresponding attribute(s) as the source of out-of-control signal. For example, the 4th row represents the situations where x_1 and x_2 are jointly diagnosed as the source of variation under each considered step shifts. Row 8 of Tables 5–7 also shows the percentage of correct diagnostic performance of the proposed method in identifying the contributed attribute(s) after receiving an out-of-control signal. Table 5 tabulates the results of the proposed MLP in fault detection and fault diagnosis under different step shifts where only one linguistic attribute causes an out-of-control signal. The results of Table 5 confirm that in situations where only one linguistic attribute is the source of variation in the process, the ARL_1 values obtained from the proposed neuro-fuzzy approach are small. Therefore, we can conclude that in such situations, the proposed neuro-fuzzy approach can detect different out-of-control states quickly. The results of Table 5 from diagnosis point of view show that the performance of the proposed NN-based approach in diagnosing the single out-of-control linguistic attribute under different step shifts is satisfactory. Table 5 also show that the diagnostic performance of the proposed neuro-fuzzy approach in situations that linguistic attribute x_2 is out-of-control is better than the other situations.

Table 6 tabulates the performance of the proposed neuro-fuzzy method in fault detection and fault diagnosis in situations where joint shifts in two linguistic attributes are occurred. The results of Table 6 show that the proposed method performs adequately in detecting considered shifts as well as in diagnosing the corresponding two contributed attributes for the out-of-control signals. Table 6 also shows that the proposed method performs better under the presence of two out-of-control attributes rather than one out-of-control linguistic attribute.

The results of detecting and diagnosing performance of the proposed neuro-fuzzy method under shifts where all linguistic attributes cause the out-of-control signals is also summarised in Table 7. According to Table 7, we conclude that in situations which all the linguistic attributes are the sources of out-of-control signals, the ARL_1 values of the proposed MLP approach are small. Table 7 also shows that in situations which all attributes are out-of-control, as the magnitude of shifts increases, the correct diagnosis percent obtained from the designed MLP will improve.

5 Conclusions

In this paper a new NN-based monitoring scheme was suggested for monitoring fuzzy multivariate multinomial data in phase 2. Before applying the proposed neuro-fuzzy approach for monitoring multivariate data, the parameters of the multivariate multinomial linguistic attributes were estimated. In addition, the proposed NN-based approach not only is able to detect different shifts in the linguistic multivariate multinomial processes, but also can diagnose the quality characteristic(s) responsible for the out-of-control signals. An application example based on Taleb et al. (2006) data was presented to investigate the performance of the proposed approach in both fault detection as well as the fault diagnosis purposes. The results of application example confirmed the satisfactory performance of the proposed method in both fault detection as well as fault diagnosis.

References

Alaeddini, A., Ghazanfari, M. and 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*, Vol. 179, No. 11, pp.1769–1784.

Alipour, H. and Noorossana, R. (2010) 'Fuzzy multivariate exponentially weighted moving average control chart', *The International Journal of Advanced Manufacturing Technology*, Vol. 48, Nos. 9–12, pp.1001–1007.

Alizadeh, H.M. and Ghomi, S.F. (2011) 'Fuzzy development of mean and range control charts using statistical properties of different representative values', *Journal of Intelligent and Fuzzy Systems*, Vol. 22, No. 5, pp.253–265.

Alizadeh, H.M., Khamseh, A.A. and Ghomi, S.F. (2010) 'Fuzzy development of multivariate variable control charts using the fuzzy likelihood ratio test', *Scientia Iranica Transaction E-Industrial Engineering*, Vol. 17, No. 2, pp.150–163.

Amirzadeh, V., Mashinchi, M. and Parchami, A. (2009) 'Construction of p-charts using degree of nonconformity', *Information Sciences*, Vol. 179, No. 1, pp.150–160.

Chang, S.I. and Aw, C.A. (1996) 'A neural fuzzy control chart for detecting and classifying process mean shifts', *International Journal of Production Research*, Vol. 34, No. 8, pp.2265–2278.

Demirli, K. and Vijayakumar, S. (2010) 'Fuzzy logic based assignable cause diagnosis using control chart patterns', *Information Sciences*, Vol. 180, No. 17, pp.3258–3272.

Engin, O., Çelik, A. and Kaya, İ. (2008) 'A fuzzy approach to define sample size for attributes control chart in multistage processes: an application in engine valve manufacturing process', *Applied Soft Computing*, Vol. 8, No. 4, pp.1654–1663.

Erginel, N. (2008) 'Fuzzy individual and moving range control charts with α-cuts', *Journal of Intelligent and Fuzzy Systems*, Vol. 19, No. 4, pp.373–383.

Ertuğrul, İ. and Aytaç, E. (2009) 'Construction of quality control charts by using probability and fuzzy approaches and an application in a textile company', *Journal of Intelligent Manufacturing*, Vol. 20, No. 2, pp.139–149.

Faraz, A. and Shapiro, A.F. (2010) 'An application of fuzzy random variables to control charts', *Fuzzy Sets and Systems*, Vol. 161, No. 20, pp.2684–2694.

Gülbay, M. and Kahraman, C. (2006) 'Development of fuzzy process control charts and fuzzy unnatural pattern analyses', *Computational Statistics & Data Analysis*, Vol. 51, No. 1, pp.434–451.

Gülbay, M. and Kahraman, C. (2007) 'An alternative approach to fuzzy control charts: direct fuzzy approach', *Information Sciences*, Vol. 177, No. 6, pp.1463–1480.

Gülbay, M., Kahraman, C. and Ruan, D. (2004) 'α-cut fuzzy control charts for linguistic data', *International Journal of Intelligent Systems*, Vol. 19, No. 12, pp.1173–1195.

Kaya, İ. and Kahraman, C. (2011) 'Process capability analyses based on fuzzy measurements and fuzzy control charts', *Expert Systems with Applications*, Vol. 38, No. 4, pp.3172–3184.

Kazemi, M.S., Bazargan, H. and Yaghoobi, M.A. (2014a) 'Estimating the drift time for processes subject to linear trend disturbance using fuzzy statistical clustering', *International Journal of Production Research*, Vol. 52, No. 11, pp.3317–3330.

Kazemi, N., Ehsani, E. and Glock, C.H. (2014b) 'Multi-objective supplier selection and order allocation under quantity discounts with fuzzy goals and fuzzy constraints', *International Journal of Applied Decision Sciences*, Vol. 7, No. 1, pp.66–96.

Khalili-Damghani, K., Dideh-Khani, H. and Sadi-Nezhad, S. (2013) 'A two-stage approach based on ANFIS and fuzzy goal programming for supplier selection', *International Journal of Applied Decision Sciences*, Vol. 6, No. 1, pp.1–14.

M'Hallah, R. and Melloy, B.J. (2010) 'An adaptive control procedure for a fuzzy quality characteristic', *International Journal of Quality Engineering and Technology*, Vol. 1, No. 3, pp.276–300.

M'Hallah, R. and Melloy, B.J. (2012) 'Impact of prior information on the total cost of an adaptive quality control procedure for a fuzzy quality characteristic', *International Journal of Quality Engineering and Technology*, Vol. 3, No. 1, pp.31–49.

Oke, S.A., Adeoye, S.A., Oluwo, A., Oyekeye, M.O., Alozie, S.I. and Johnson, A.O. (2009) 'Application of neurofuzzy framework to maintenance scheduling activity monitoring', *International Journal of Applied Decision Sciences*, Vol. 2, No. 3, pp.299–326.

Pastuizaca Fernández, M.N., Carrión García, A. and Ruiz Barzola, O. (2014) 'Multivariate multinomial T2 control chart using fuzzy approach', *International Journal of Production Research*, Vol. 53, No. 7, pp.2225–2238.

Raz, T. and Wang, J.H. (1990) 'Probabilistic and membership approaches in the construction of control charts for linguistic data', *Production Planning & Control*, Vol. 1, No. 3, pp.147–157.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) 'Learning internal representations by error propagation', a chapter in book entitled *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, pp.318–362, MIT Press, Cambridge, MA.

Senturk, S. and Erginel, N. (2009) 'Development of fuzzy and control charts using α-cuts', *Information Sciences*, Vol. 179, No. 10, pp.1542–1551.

Shah, N.H. and Soni, H. (2011) 'Periodic review inventory model with service level constraint in fuzzy environment', *International Journal of Applied Decision Sciences*, Vol. 4, No. 2, pp.129–147.

Sharifi, E., Shabani, S., Sobhanallahi, M.A. and Mirzazadeh, A. (2015) 'A fuzzy economic order quantity model for items with imperfect quality and partial backordered shortage under screening errors', *International Journal of Applied Decision Sciences*, Vol. 8, No. 1, pp.109–126.

Shu, M.H. and Wu, H.C. (2010) 'Monitoring imprecise fraction of nonconforming items using p control charts', *Journal of Applied Statistics*, Vol. 37, No. 8, pp.1283–1297.

Shu, M.H. and Wu, H.C. (2011) 'Fuzzy $\overline{X}$ and R control charts: fuzzy dominance approach', *Computers & Industrial Engineering*, Vol. 61, No. 3, pp.676–685.

Shu, M.H. and Wu, H.C. (2012) 'Manufacturing process performance evaluation for fuzzy data based on loss-based capability index', *Soft Computing*, Vol. 16, No. 1, pp.89–99.

Shu, M.H., Chiu, C.C., Nguyen, T.L. and Hsu, B.M. (2014a) 'A demerit-fuzzy rating system, monitoring scheme and classification for manufacturing processes', *Expert Systems with Applications*, Vol. 41, No. 17, pp.7878–7888.

Shu, M.H., Nguyen, T.L. and Hsu, B.M. (2014b) 'Fuzzy MaxGWMA chart for identifying abnormal variations of on-line manufacturing processes with imprecise information', *Expert Systems with Applications*, Vol. 41, No. 4, pp.1342–1356.

Taleb, H. and Limam, M. (2002) 'On fuzzy and probabilistic control charts', *International Journal of Production Research*, Vol. 40, No. 12, pp.2849–2863.

Taleb, H., Limam, M. and Hirota, K. (2006) 'Multivariate fuzzy multinomial control charts', *QTQM*, Vol. 3, No. 4, pp.437–453.

Tong, X. and Wang, Z. (2012) 'Fuzzy acceptance sampling plans for inspection of geospatial data with ambiguity in quality characteristics', *Computers & Geosciences*, Vol. 48, pp.256–266.

Wang, D., Li, P. and Yasuda, M. (2014) 'Construction of fuzzy control charts based on weighted possibilistic mean', *Communications in Statistics-Theory and Methods*, Vol. 43, No. 15, pp.3186–3207.

Wang, T.Y. and Chen, L.H. (2002) 'Mean shifts detection and classification in multivariate process: a neural-fuzzy approach', *Journal of Intelligent Manufacturing*, Vol. 13, No. 3, pp.211–221.

Yang, M.S. and Yang, J.H. (2002) 'A fuzzy-soft learning vector quantization for control chart pattern recognition', *International Journal of Production Research*, Vol. 40, No. 12, pp.2721–2731.

Zarandi, M.H., Alaeddini, A. and Turksen, I.B. (2008) 'A hybrid fuzzy adaptive sampling-run rules for Shewhart control charts', *Information Sciences*, Vol. 178, No. 4, pp.1152–1170.

Zarandi, M.H.F. and Alaeddini, A. (2010) 'A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts', *Information Sciences*, Vol. 180, No. 16, pp.3033–3044.

Appendix

Table A1 Datasets of the process

m	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}	x_{31}	x_{32}	x_{33}	x_{34}
1	210	7	3	206	9	5	167	48	3	2
2	211	6	3	207	8	5	176	42	2	0
3	206	9	5	202	12	6	163	55	2	0
4	211	5	4	207	8	5	163	51	5	1
5	203	16	1	194	18	8	175	45	0	0
6	210	6	4	206	9	5	174	44	1	1
7	208	7	5	204	9	7	174	40	5	1
8	207	7	6	204	9	7	169	46	3	2
9	206	7	7	202	9	9	169	48	2	1
10	186	25	9	200	12	8	169	48	3	0
11	196	13	11	196	13	11	163	46	10	1
12	203	12	5	200	13	7	167	44	9	0
13	203	9	8	198	11	11	174	42	3	1
14	202	9	9	198	11	11	174	40	6	0
15	209	6	5	207	9	4	172	42	5	1
16	210	3	7	205	5	10	172	44	4	0
17	205	11	4	201	13	6	172	45	2	1
18	210	6	4	206	8	6	169	48	2	1
19	206	10	4	203	13	4	172	46	0	2
020	206	12	2	202	14	4	169	46	5	0

Source: Taleb et al. (2006)