IJE TRANSACTIONS B: Applications Vol. 28, No. 11, (November 2015) 1634-1643

International Journal of Engineering

Journal Homepage: www.ije.ir

Online Monitoring and Fault Diagnosis of Multivariate-attribute Process Mean Using Neural Networks and Discriminant Analysis Technique

M. R. Maleki*, R. Sahraeian

Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran

PAPER INFO

Paper history:
Received 01 August 2015
Received in revised form 14 November 2015
Accepted 19 November 2015

Keywords:
Discriminant Analysis
Multivariate-attribute
NORTA Inverse
Neural Network
Fault Detection
Fault Diagnosis

ABSTRACT

In some statistical process control applications, the process data are not Normally distributed and characterized by the combination of both variable and attributes quality characteristics. Despite different methods which are proposed separately for monitoring multivariate and multi-attribute processes, only few methods are available in the literature for monitoring multivariate-attribute processes. In this paper, we develop discriminant analysis technique for monitoring the mean vector of correlated multivariate-attribute quality characteristics in the first module. Then in the second module, a novelty approach based on the combination of artificial neural network (ANN) and discriminant analysis is proposed for detecting different mean shifts. The proposed approach is also able to diagnose quality characteristic(s) responsible for out-of-control signals after detecting different step mean shifts. A numerical example based on simulation is given to evaluate the performance of the proposed methods for detection and diagnosis purposes. The detecting performance of the second module is also compared with the extended T^2 control chart and with the extension of an ANN in the literature. The results confirm that the proposed method outperforms both methods.

***doi**: 10.5829/idosi.ije.2015.28.11b.11*

1. INTRODUCTION

In most statistical process monitoring applications, the quality of the product is characterized by one or more variable or attribute quality characteristics (QCs). Monitoring such processes is well documented in literature such as in Keramatpour et al. [1] and Owlia and Fallahnezhad [2].

Sometimes, the combination of both variable and attribute QCs should be monitored simultaneously. For instance, in semiconductor manufacturing process, the impurity of particle counts, the average of oxide thickness, and the range of the thickness are correlated QCs and follow Poisson, Normal and Gamma distributions, respectively [3]. In such situations when the QCs tend to be correlated, monitoring each QC by a separate control chart leads to misleading results. In order to overcome this problem, two main multivariate procedures are proposed. The first one is multivariate-attribute control chart that takes into account the correlation structure between the combination of variable and attribute QCs while the second approach is applying artificial neural networks (ANNs). The multivariate-attribute control schemes are addressed as follows:

Kang and Brenneman [4] introduced a bootstrap method to construct a $(1-\alpha)\times 100$ upper confidence bound for the overall defect rate of a product whose quality assessment involves multiple pass/fail Binary data and multiple continuous data based on independent assumption. Ganjali [3] proposed a model for mixed continuous and discrete responses with possibility of missing responses. Ning and Tsung [5] mentioned the difficulties in monitoring a process with mixed-type and high-dimensional data and discussed potential solutions. They also presented some practical examples in processes whose quality are characterized thorough high-dimensional data. Doroudyan and Amiri [6]

*Corresponding Author's Email: *Reza.maleki@shahed.ac.ir* (M. R. Maleki)

Please cite this article as: M. R. Maleki, R. Sahraeian, Online Monitoring and Fault Diagnosis of Multivariate-attribute Process Mean Using Neural Networks and Discriminant Analysis Technique, International Journal of Engineering (IJE), TRANSACTIONS B: Applications Vol. 28, No. 11, (November 2015) 1634-1643

extended T^2 Hoteling control chart for detecting different mean shifts in multivariate-attribute processes using root transformation method. Maleki, Amiri and Doroudyan [7] suggested an ANN for monitoring the multivariate-attribute quality characteristics as well as diagnosing quality characteristic(s) that cause out-of-control signals. Amiri et al. [8] suggested a new approach based on multivariate process capability index and NORTA inverse transformation in a multi response optimization problem with mixed continuous-discrete responses. Maleki, Amiri and Rasouli [9] developed two exponentially weighted moving average (EWMA)-based control charts for monitoring the variability of multivariate-attribute processes. Doroudyan and Amiri [10] used four transformation methods to monitor the multivariate-attribute processes. Doroudyan and Amiri [11] utilized NORTA inverse technique for monitoring multivariate-attribute processes. First, they used NORTA inverse method and transformed the original data to a multivariate Normal distribution. Then, they used multivariate charts such as T^2 and MEWMA for transformed data. Amiri et al. [12] proposed an ANN approach for monitoring multivariate-attibute process variability. Maleki and Amiri [13] proposed an ANN to identify the magnitude of shifts in the out-of-control QCs. Maleki and Amiri [14] studied the simultaneous monitoring of multivariate-attribute process mean and variability. Maleki et al. [15] proposed a modular ANN-based method for estimating the time of step changes in multivariate-attribute process variability. They also extended maximum likelihood estimation (MLE) and then provided a comparison study between ANN and MLE approaches.

In this paper, artificial neural network and discriminant analysis (DA) technique are combined to develop a new methodology for on-line monitoring the mean vector of correlated multivariate-attribute QCs. The proposed procedure can be applied in fault detection as well as fault diagnosis, simultaneously. In fault detection, the model investigates whether an out-of-control state has occurred or not; whereas fault diagnosis is faced with identifying the QC(s) that cause out-of-control signals. The rest of the paper is organized as follows: In section 2 the proposed method is described. In section 3, the performance of the proposed methodology is assessed thorough two simulation studies. In section 4, the concluding remarks are provided.

2. PROPOSED METHODOLOGY

2. 1. First Module In this subsection, we develop discriminant analysis technique and apply it to detect different step shifts in the mean vector of a multivariate-attribute process. Discriminant analysis is an effective classification technique that categorizes the data into some predetermined classes via maximizing the separation between classes while minimizing the separation within each class. We focus on deriving the discriminant function that makes the best separation between two classes in a multivariate-attribute process mean including in-control and out-of-control classes. In order to estimate the discriminant function, first for each out-of-control state we generate 500 random samples of size n. Then, equal to the number of out-of-control samples, the in-control samples are simulated. In a multivariate-attribute process where the quality of a product is expressed by the combination of p variable and q attribute, the t^{th} ; $t=1,2,...$ sample is denoted as:

$$\mathbf{X_t} = (\mathbf{X_{t1}}, \mathbf{X_{t2}}, ..., \mathbf{X_{tp}}, \mathbf{X_{t,p+1}}, ..., \mathbf{X_{t,p+q}}), \quad (1)$$

where $\mathbf{X}_{tj} = (x_{tj1}, x_{tj2}, ..., x_{tjn})'$ is a vector of size n which corresponds to j^{th}; $j=1,...,p+q$ QC in $t^{th}; t=1,2,...$ sample. Note that, the first p vectors are related to the variables while the last q ones are related to attributes.

One of the preliminary assumptions in discriminant analysis technique is that the data are Normally-distributed. However, the combination of correlated multivariate-attribute QCs violates the Normality assumption. Hence, after generating multivariate-attribute QCs, the distribution of original data is transformed to the multivariate Normal distribution. Note that in this paper NORTA inverse technique is used for transforming the original data to multivariate Normal distribution.

After applying NORTA inverse technique, the vector corresponding to t^{th} sample is denoted by $\mathbf{Y_t} = (\mathbf{Y_{t1}}, \mathbf{Y_{t2}}, ..., \mathbf{Y_{tp}}, \mathbf{Y_{t,p+1}}, ..., \mathbf{Y_{t,p+q}})$ in which the marginal probability distribution of each transformed QC is approximately Normal. ($\mathbf{Y_{tj}} = (y_{tj1}, y_{tj2}, ..., y_{tjn})$ is corresponding to j^{th}; $j=1,2,...,p+q$ transformed QC). The mean vector of $t^{th}; t=1,2,...$ transformed sample is calculated as follows:

$$\overline{\mathbf{Y_t}} = (\bar{y}_{t1}, \bar{y}_{t2}, ..., \bar{y}_{t,p+q}), \quad (2)$$

where $\overline{y_{tj}} = \sum_{k=1}^{n} y_{tjk} \big/ n, \; j=1,2,...,p+q$. After determining the vector $\overline{\mathbf{Y_t}}$ for all samples, we define matrix $\mathbf{T}$ as the total SSCP matrix in a ($p+q$)-dimensional space as follows:

$$\mathbf{T} = \mathbf{B} + \mathbf{W} \quad (3)$$

where $\mathbf{B}$ and $\mathbf{W}$ are, between groups and within-group SSCP matrices for the p+q QCs. Then:

$$|\mathbf{W}^{-1}\mathbf{B} - \lambda \mathbf{I}| = 0 \quad (4)$$

That is, the problem reduces to finding eigenvalues of the non-symmetric matrix, $\mathbf{W}^{-1}\mathbf{B}$, with the eigenvectors giving the weight matrix for forming the discriminant function. For more information refer to Subhash [16]. In a special case with one variable and one attribute, the following equation should be maximized to calculate the discriminant function:

$$\lambda = {SS_B}/{SS_W}. \tag{5}$$

where SS_B and SS_W are the between-group sum of squares and within-group sum of squares, respectively. The discriminant function that best separates between the in-control and out-of-control data via maximizing the Equation (5) is obtained as:

$$DA = \overline{y_1}\cos\theta^* + \overline{y_2}\sin\theta^*. \tag{6}$$

After estimating the discriminant function (optimal angel θ^*), the future observations of multivariate-attribute process are classified into in-control or out-of-control states. To do that, the cutoff value of discriminant score is determined as follows:

$$\text{cutoff value} = C_0\overline{DA_0} + C_1\overline{DA_1}, \tag{7}$$

where $\overline{DA_0}$ is the mean value of discriminant function for all in-control samples, while $\overline{DA_1}$ is the mean value of discriminant function for all out-of-control random samples. Note that C_0 and C_1 are cost parameters and are determined such that the misclassification cost be minimized.

2. 2. Second Module

In this subsection, the second module that is based on the combination of ANN and discriminant analysis method (called ANN-DA method) is described. Recall that the proposed method not only is able to detect different step shifts in the multivariate-attribute process mean, but also can identify the QC(s) contributed to out-of-control signals. We refer the task of detecting mean shifts as fault detection and diagnosing QC(s) that cause signals as fault diagnosis. For this purpose, we develop a multilayer perceptron (MLP) ANN which uses back-propagation training algorithm with following structure:

According to Figure 1, the input layer of the proposed ANN includes p+q+1 nodes. The input vector of the proposed ANN corresponding to a given sample is a column vector of $(\overline{y}_1, \overline{y}_2, ..., \overline{y}_{p+q}, DA)$. As noted, $\overline{y}_j$ is the mean of the j[th]; $j = 1, 2, ..., p+q$ transformed QC and DA is the discriminant score of the sample taken. The number of nodes in the output layer is equal to 2^{p+q}. It is pointed out in the literature that there is not a direct guideline for determining the optimal number of hidden layers as well as the number of nodes in each one. However, many theoretical and simulative researches represent that one or two hidden layer in multilayer perceptron neural networks may be sufficient. In this research, such values are determined by trial and error.

Choosing a proper activation function has a crucial role in training the ANNs. The sigmoid function whose output values are in the range of [0,1] is used in the suggested neural network. In order to train the designed ANN the following steps are suggested:

- Generating 1000 random samples of size n for each out-of-control state (totally $1000 \times (2^{p+q} - 1)$ data sets).
- Generating in-control training data sets as equal to the number of out-of-control data.
- Transforming the generated input vectors to multivariate Normal distribution using NORTA inverse technique.
- Calculating the mean value of each transformed QCs in the generated samples as the first $p+q$ elements of input vector.
- Calculating the discriminant score corresponding to each sample as the last element of input vector.
- Entering the generated input data as well as their target vectors into the designed ANN.
- Training the designed ANN using back-propagation algorithm with considering mean squared error (MSE) criterion.

The target vectors for the training input data is a column vector of $\mathbf{T} = \left[t_1, t_2, ..., t_{(2^{p+q})}\right]^T$ with 2^{p+q} elements in which the first one is used for fault detection while the others are used for fault diagnosis. The value of t_1 for in-control and out-of-control samples are considered equal to zero and one, respectively. The value of other elements in target vector of in-control samples are also considered equal to zero. For an out-of-control state, the values of all elements in target vector are zero expect t_1 and $t_i; i = 2, ..., 2^{p+q}$ that are equal to one. Note that $t_1 = 1$ shows that the process mean is out-of-control and $t_i = 1$ means that the $(i-1)^{\text{th}}$ out-of-control situation has occurred. The target vectors in a p-variate /q-attribute process are given in Table 1. The columns of Table 1 are explained as follows (See also Figure 1 for more clarification.):

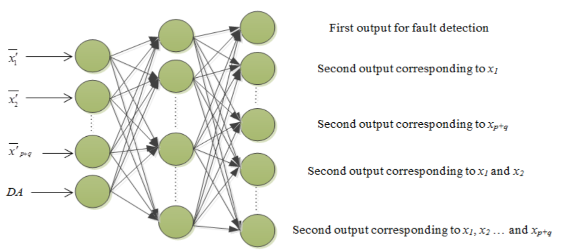

Figure 1. Structure of the proposed neural network

Column 1- number of QCs This column shows the number of possible QC(s) that cause an out-of-control state. Obviously, in a p-variate/q-attribute process, $i; i=1,...,p+q$ QC(s) may cause a signal.

Column 2- number of out-of-control states There are totally $\binom{p+q}{i}; i=1,...,p+q$ different out-of-control states in a p-variate /q-attribute process, when i QC(s) cause an out-of-control state. For example in a process with 2 variables (x_1 and x_2) and one attribute (x_3), there are totally $\binom{2+1}{2}=3$ out-of-control states where two QCs are responsible for the signal. These out-of-control states are:

- The out-of-control state when x_1 and x_2 are responsible for the signal.
- The out-of-control state when x_1 and x_3 are responsible for the signal.
- The out-of-control state when x_2 and x_3 are responsible for the signal.

Column 3 and 4 Beginning and Finishing Elements It shows the output nodes of the designed ANN corresponding to the state in which $i; i=1,...,p+q$ QCs are responsible for the out-of-control state. For example, in a process with 2 variables (x_1 and x_2) and one attribute (x_3), there are totally $2^{2+1}=8$ output nodes in the ANN in which the first output neuron is used for fault detection and the others are for fault diagnosis. The second, third and fourth output nodes are related to the out-of-control states in which x_1, x_2 and x_3, respectively are responsible for the out-of-control signal. Hence, the beginning output neuron corresponding to the out-of-control situation in which one QC is responsible for the signal is 2 and the finishing is 4. Similarly, the fifth, sixth and seventh output nodes are assigned to the out of-control states in which two QCs are responsible for the out-of-control signal. Hence, the beginning output node is 5 and the finishing is 7. Finally, the eighth output neuron is assigned to the out-of-control state in which all three QCs are responsible for the signal. Hence the beginning and finishing output neuron is 8.

In order to apply the proposed ANN-DA method in fault detection, we set a threshold for the first output node and then compare it with the observed value of the first output node. For this purpose the following steps are suggested:

1. Generating 10000 in-control samples of size n and compute vector $(\overline{y}_1, \overline{y}_2, ..., \overline{y}_{p+q}, DA)$ for each simulated sample.
2. Entering generated input vectors to the ANN.
3. Sorting the values of the first output node and saving them in an ascending order.
4. Determining the element of the sorted vector such that the in-control average run length of ANN becomes a predetermined value.

Note that the run length criterion is defined as the number of consecutive samples taken until the first sample statistic falls outside the control limits interval. In in-control and out-of-control situations, the expected value of run lengths is called in-control ARL (ARL_0) and out-of-control ARL (ARL_1), respectively. When an input vector is entered to the trained ANN, we only focus on the value of first observed output. If the value of the first output node is less than or equal to the first threshold, the process is in-control and we do not notice the other output nodes. Otherwise, if the first observed value is more than the first threshold, the process is out-of-control. In out-of-control situations, the QC(s) whose corresponding output neuron has the maximum value amongst the output values except the first node is (are) considered as the out-of-control QC(s).

TABLE 1. Target values of the proposed neural network

Number of qCs	Number of out-of-control states	Beginning element	Finishing element	QCs responsible for signal
1	$\binom{p+q}{1}$	2	$p+q+1$	x_1 x_2 ⋮ x_{p+q}
2	$\binom{p+q}{2}$	$p+q+2$	$p+q+\binom{p+q}{2}+1$	x_1 and x_2 x_1 and x_3 ⋮ x_{p+q-1} and x_{p+q}
⋮	⋮	⋮	⋮	⋮
p+q	1	2^{p+q}	2^{p+q}	$x_1, x_2, ..., x_{p+q}$

TABLE 2. The information of generating data sets for estimating the discriminant function

Signal source	Number of data	Mean of x_1	Mean of x_2	Process state
-	1500	$\lambda = 4$	$\mu = 3$	in-control
x_1	500	$\lambda = 8$	$\mu = 3$	out-of-control
x_2	500	$\lambda = 4$	$\mu = 7$	out-of-control
x_1 and x_2	500	$\lambda = 8$	$\mu = 7$	out-of-control

TABLE 3. Classifying performance of the proposed discriminant analysis-based control scheme

Shift	$(0,0)$	$(\sigma_1,0)$	$(0,\sigma_2)$	$(1.5\sigma_1,0)$	$(0,1.5\sigma_2)$
CCP (%)	90.12	64.42	78.46	86.33	96.37
Shift	$(2\sigma_1,0)$	$(0,2\sigma_2)$	(σ_1,σ_2)	$(1.5\sigma_1,1.5\sigma_2)$	$(2\sigma_1,2\sigma_2)$
CCP (%)	96.03	99.85	98.85	100	100

3. PERFORMANCE EVALUATION

In this section, the performance of both methods is investigated and compared with the extension of T^2 control chart and an ANN approach in the literature through two numerical examples. he ANN approach in the literature is proposed by Amiri et al. [12] for monitoring multivariate-attribute process variability. In order to provide a comparison study, we extend the ANN approach of Amiri et al. [12] and utilize it for monitoring multivariate-attribute process mean. Note that, all methods are simulated in MATLAB computer software. In this paper, we use the normal to anything (NORTA) method which is a simulation-based method to generate the multivariate-attribute QC(s). See Appendix for more information about NORTA method.

3. 1. Simulation Study 1

Consider a 1-variable/1-attribute process where $x_1 \sim N(\mu=3,\sigma^2=4), x_2 \sim Poisson(\lambda=4)$, the correlation coefficient between QCs is $\rho = 0.357$ and the samples of size 10 is used. The information of generating data sets for estimating the discriminant function are summarized in Table 2. As seen in Table 2, we generate 1500 out-of-control and 1500 in-control samples.

According to Section 2, the discriminant function is obtained as $DA = \overline{y_1}\cos 48^{''} + \overline{y_2}\sin 48^{''}$. The values of C_0 and C_1 are determined with following assumptions:

The cost of assigning an out-of-control sample to the in-control state is equal to the cost of a false alarm multiplied by 3. Hence, the cutoff value by selecting $C_0 = 0.75$ and $C_1 = 0.25$ is obtained as equal to 0.6343. The performance of DA method in detecting step mean shifts in terms of correct classification percentage (CCP) criterion is summarized in Table 3. Table 3 shows that the DA method can accurately classify the data into in-control and out-of-control states. Table 3 also represents that as the magnitude of mean shifts increases, the performance of the extended DA method improves. We can also conclude from Table 3 that when a mean shift occurs in x_1, the extended DA method performs better in comparison with situations when the change is occurred in x_2.

To combine the ANN with DA method, a three layer Perceptron ANN whose input and output layers has three and four nodes, respectively is suggested. The presented ANN through trial and error has one hidden layer with 12 nodes. The information on training the ANN is summarized in Table 4. Note that for generating out-of-control data for estimating the discriminant function and training the ANN, the mean shifts with the magnitude of 2σ are used.

After training the designed ANN with 6000 samples, the mean squared error (MSE) of training step is obtained to be equal to 0.00203. For comparing the detecting performance of the proposed ANN-DA method with the extended T^2 chart and the extended ANN proposed by Amiri et al. [12], we consider $ARL_0 = 200$ for all methods. For this purpose, the threshold of the first output node in ANN-DA method is set equal to 0.26 and consequently $ARL_0 = 199.98$. The upper control limit of the extended T^2 chart is also selected equal to 10.6 and then $ARL_0 = 200.93$.

The ARLs of ANN-DA method is compared with two other methods and the results are summarized in Table 5.Table 5 shows that the proposed ANN-DA method can effectively detect different mean shifts. Table 5 also shows that the proposed ANN-DA method outperforms T^2 and ANN methods in detecting almost all mean shifts. In fault diagnosis, if we denote the observed value of i^{th} node in output layer by $Y_i,\ i = 2,3,4$ the out-of-control QC(s) is identified as follows:

If $\max\{Y_2,Y_3,Y_4\} = Y_2$, then x_1 is considered as responsible QC for signal. If $\max\{Y_2,Y_3,Y_4\} = Y_3$, then x_2 is considered as responsible QC for signal. If

$\max\{Y_2, Y_3, Y_4\} = Y_4$, then both x_1 and x_2 are considered as responsible QC(s).

The performance of the ANN-DA method in fault diagnosis based on 10000 replicates is evaluated in Table 6. The first, second and third rows of Table 6 represent the number of runs in which the proposed method diagnoses x_1, x_2 and both x_1 and x_2 as the source(s) of signal, respectively. The results of Table 6 show that the performance of ANN-DA method in fault diagnosis is quite satisfactory. Table 6 also shows that the percentage of correct diagnosis increases as the magnitude of mean shifts increases.

3. 2. Simulation Study 2 Consider a 2-variable/1-attribute process in which $X_1 \sim N(\mu_1 = 4, \sigma_1^2 = 3^2), X_2 \sim N(\mu_1 = 5, \sigma_1^2 = 3^2)$, and $X_3 \sim Binomial(n = 12, p = 0.2)$. Suppose that $\rho_{ij} = 0.2;\ i, j = 1,2,3$ and $\forall i \neq j$ and the samples with size of 10 is used for monitoring the process. The information of generating data for estimating the discriminant function are summarized in Table 7. The discriminant function that best separates in-control and out-of-control data is obtained as: $DA = 0.6034\overline{y_1} + 0.5979\overline{y_2} + 0.5277\overline{y_3}$. The cutoff value considering $C_0 = 0.8$ and $C_1 = 0.2$ is equal to 0.6123. The classifying performance of the DA method in terms of CCP criterion is summarized in Table 8. Table 8 shows that the DA method accurately classifies the data into the in-control and out-of-control states.

In module 2, a three layer Perceptron ANN with four and eight nodes in input and output layers is designed. The presented ANN has one hidden layer with 18 nodes. The information of training the ANN is summarized in Table 9. The MSE of training the designed ANN is equal to 0.000512.

The performance of the proposed ANN-DA control scheme is compared with the T^2 chart and the extension of Amiri et al. [12] ANN and the results are given in Table 10. Table 10 shows that the proposed method outperforms both extended T^2 and ANN approaches under all step mean shifts. The capability of the ANN-DA approach in fault diagnosis is given in Table 11. The results confirm the satisfactory diagnosis capability of the proposed method under all step mean shifts.

TABLE 4. The information of training the MLP

Signal source	Number of data	Mean of x_1	Mean of x_2	Process state	Target vector
-	3000	$\lambda = 4$	$\mu = 3$	in-control	$[0,0,0,0]^T$
x_1	1000	$\lambda = 8$	$\mu = 3$	out-of-control	$[1,1,0,0]^T$
x_2	1000	$\lambda = 4$	$\mu = 7$	out-of-control	$[1,0,1,0]^T$
x_1 and x_2	1000	$\lambda = 8$	$\mu = 7$	out-of-control	$[1,0,0,1]^T$

TABLE 5. ARL values of detecting different step mean shifts

Shift	$(0.5\sigma_1, 0)$	$(0, 0.5\sigma_2)$	$(0.5\sigma_1, 0.5\sigma_2)$	$(0.75\sigma_1, 0)$	$(0, 0.75\sigma_2)$	$(0.75\sigma_1, 0.75\sigma_2)$
ANN-DA	7.1110	10.3170	6.7940	2.7820	3.1110	2.6209
T^2	10.9105	13.4450	8.2055	3.7340	3.8470	2.6905
ANN	8.1820	17.1423	8.1051	3.1255	5.4159	3.0848
Shift	$(\sigma_1, 0)$	$(0, \sigma_2)$	(σ_1, σ_2)	$(1.5\sigma_1, 0)$	$(0, 1.5\sigma_2)$	$(1.5\sigma_1, 1.5\sigma_2)$
ANN-DA	1.5886	1.5730	1.4460	1.0566	1.0249	1.0194
T^2	1.8610	1.7735	1.4050	1.0870	1.0375	1.0140
ANN	1.8777	2.4200	1.8711	1.1267	1.1151	1.0616

TABLE 6. Performance of the proposed ANN-DA method in fault diagnosis

Shift	$(0.5\sigma_1, 0)$	$(0, 0.5\sigma_2)$	$(0.5\sigma_1, 0.5\sigma_2)$	$(0.75\sigma_1, 0)$	$(0, 0.75\sigma_2)$	$(0.75\sigma_1, 0.75\sigma_2)$
x_1	9920	84	4971	9965	11	3760
x_2	17	9653	2847	1	9843	2275
x_1 and x_1	63	263	2182	34	146	3965
Correct diagnosis (%)	99.20	96.53	21.82	99.65	98.43	39.65
Shift	$(\sigma_1, 0)$	$(0, \sigma_2)$	(σ_1, σ_2)	$(1.5\sigma_1, 0)$	$(0, 1.5\sigma_2)$	$(1.5\sigma_1, 1.5\sigma_2)$
x_1	9985	0	2407	9997	0	450
x_2	0	9909	1596	0	9955	425
x_1 and x_1	15	91	5997	3	45	9125
Correct diagnosis (%)	99.85	99.09	59.97	99.97	99.55	91.25

TABLE 7. The information of generating data sets for estimating the discriminant function

Signal source	Number of data	Mean of x_1	Mean of x_2	Mean of x_3	Process state
-	3500	$\mu=4$	$\mu=5$	np	in-control
x_1	500	$\mu=10$	$\mu=5$	np	out-of-control
x_2	500	$\mu=4$	$\mu=11$	np	out-of-control
x_3	500	$\mu=4$	$\mu=5$	$np+2\sqrt{p(1-p)}$	out-of-control
x_1 and x_2	500	$\mu=10$	$\mu=11$	np	out-of-control
x_1 and x_3	500	$\mu=10$	$\mu=5$	$np+2\sqrt{p(1-p)}$	out-of-control
x_2 and x_3	500	$\mu=4$	$\mu=11$	$np+2\sqrt{p(1-p)}$	out-of-control
x_1, x_2 and x_3	500	$\mu=10$	$\mu=11$	$np+2\sqrt{p(1-p)}$	out-of-control

TABLE 8. Classifying performance of the proposed discriminant analysis-based control scheme

Shift	$(0,0)$	$(\sigma_1,0,0)$	$(0,\sigma_2,0)$	$(0,0,\sigma_3)$	$(1.5\sigma_1,0,0)$
CCP(%)	87.35	71.50	69.60	76.55	89.95
Shift	$(0,1.5\sigma_2,0)$	$(0,0,1.5\sigma_3)$	$(0.75\sigma_1,0.75\sigma_2,0)$	$(0.75\sigma_1,0,0.75\sigma_3)$	$(0,0.75\sigma_2,0.75\sigma_3)$
CCP(%)	90.20	92.55	89.70	92.50	92.25
Shift	$(0.75\sigma_1,0.75\sigma_2,0.75\sigma_3)$	$(\sigma_1,\sigma_2,0)$	$(\sigma_1,0,\sigma_3)$	$(0,\sigma_2,\sigma_3)$	$(\sigma_1,\sigma_2,\sigma_3)$
CCP(%)	99.15	98.80	99.15	98.50	100

TABLE 9. The information of training the MLP

Signal source	Number of data	Mean of x_1	Mean of x_2	Mean of x_3	Process state	Target vector
-	7000	$\mu=4$	$\mu=5$	np	in-control	$[0,0,0,0,0,0,0,0]^T$
x_1	1000	$\mu=10$	$\mu=5$	np	out-of-control	$[1,1,0,0,0,0,0,0]^T$
x_2	1000	$\mu=4$	$\mu=11$	np	out-of-control	$[1,0,1,0,0,0,0,0]^T$
x_3	1000	$\mu=4$	$\mu=5$	$np+2\sqrt{p(1-p)}$	out-of-control	$[1,0,0,1,0,0,0,0]^T$
x_1 and x_2	1000	$\mu=10$	$\mu=11$	np	out-of-control	$[1,0,0,0,1,0,0,0]^T$
x_1 and x_3	1000	$\mu=10$	$\mu=5$	$np+2\sqrt{p(1-p)}$	out-of-control	$[1,0,0,0,0,1,0,0]^T$
x_2 and x_3	1000	$\mu=4$	$\mu=11$	$np+2\sqrt{p(1-p)}$	out-of-control	$[1,0,0,0,0,0,1,0]^T$
x_1, x_2 and x_3	1000	$\mu=10$	$\mu=11$	$np+2\sqrt{p(1-p)}$	out-of-control	$[1,0,0,0,0,0,0,1]^T$

TABLE 10. ARL values of detecting different step mean shifts

Shift	$(0.5\sigma_1,0,0)$	$(0,0.5\sigma_2,0)$	$(0,0,0.5\sigma_3)$	$(0.75\sigma_1,0,0)$	$(0,0.75\sigma_2,0)$
ANN-DA	9.8877	8.3625	6.8100	3.0970	2.8509
T^2	15.2897	16.0421	9.8768	5.2774	4.4781
ANN	11.4960	14.9188	9.3404	4.0673	4.3141
Shift	$(0,0,0.75\sigma_3)$	$(0.5\sigma_1,0.5\sigma_2,0)$	$(0.5\sigma_1,0,0.5\sigma_3)$	$(0,0.5\sigma_2,0.5\sigma_3)$	$(0.5\sigma_1,0.5\sigma_2,0.5\sigma_3)$
ANN-DA	2.1689	6.1602	5.7344	6.2341	4.3120
T^2	3.0071	8.7665	8.9923	6.8760	5.4141
ANN	2.9042	8.0840	8.2427	5.4346	5.3901

TABLE 11. Performance of the proposed ANN-DA method in fault diagnosis

Shift	$(0.5\sigma_1, 0, 0)$	$(0, 0.5\sigma_2, 0)$	$(0, 0, 0.5\sigma_3)$	$(0.75\sigma_1, 0, 0)$	$(0, 0.75\sigma_2, 0)$
x_1	9252	91	201	9551	26
x_2	153	9769	483	48	9904
x_3	587	134	9314	395	66
x_1 and x_2	6	5	0	5	3
x_1 and x_3	2	0	1	1	0
x_2 and x_3	0	1	1	0	1
x_1, x_2 and x_3	0	0	0	0	0
Correct diagnosis (%)	92.52	97.53	93.14	95.51	99.04
Shift	$(0, 0, 0.75\sigma_3)$	$(0.5\sigma_1, 0.5\sigma_2, 0)$	$(0.5\sigma_1, 0, 0.5\sigma_3)$	$(0, 0.5\sigma_2, 0.5\sigma_3)$	$(0.5\sigma_1, 0.5\sigma_2, 0.5\sigma_3)$
x_1	75	1639	1569	76	1040
x_2	185	1885	78	1988	1234
x_3	9738	23	2222	2642	1343
x_1 and x_2	0	6452	6	3	126
x_1 and x_3	0	0	6122	6	132
x_2 and x_3	2	0	3	5282	99
x_1, x_2 and x_3	0	1	0	3	6026
Correct diagnosis (%)	97.38	64.52	61.22	52.82	60.26

4. CONCLUDING REMARKS

In this study, a two-module method for fault detection and fault diagnosis in multivariate-attribute process mean in Phase II was suggested. In module 1, the DA method was developed for monitoring the multivariate-attribute process mean. In the second module, a control scheme based on the combination of DA method and ANN was proposed. The performance of both methods in monitoring the multivariate-attribute process mean was evaluated through two simulation studies. The results indicated the high performance of proposed methods in fault detection and fault diagnosis. Moreover, the proposed ANN-DA approach in module 2 outperforms the extended T^2 chart and the extension of an ANN in the literature.

5. ACKNOWLEDGEMENT

The authors appreciate the precious and constructive comments by respectful referees which led to significant improvement in the paper.

6. REFERENCES

1. Keramatpour, M., Niaki, S., Khedmati, M. and Soleymanian, M., "Monitoring and change point estimation of AR (1) autocorrelated polynomial profiles", ***International Journal of Engineering-Transactions C: Aspects***, Vol. 26, No. 9, (2013), 933-942.
2. Owlia, M. and Fallahnezhad, M., "A new statistical approach for recognizing and classifying patterns of control charts (research note)", ***International Journal of Engineering-Transactions A: Basics***, Vol. 28, No. 7, (2015), 1040-1048.
3. Ganjali, M., "A model for mixed continuous and discrete responses with possibility of missing responses", ***Journal of Sciences Islamic Republic of Iran***, Vol. 14, No. 1, (2003), 53-60.
4. Kang, L. and Brenneman, W.A., "Product defect rate confidence bound with attribute and variable data", ***Quality and Reliability Engineering International***, Vol. 27, No. 3, (2011), 353-368.
5. Ning, X. and Tsung, F., "Monitoring a process with mixed-type and high-dimensional data", in Industrial Engineering and Engineering Management (IEEM), 2010 IEEE International Conference on, IEEE., (2010), 1430-1432.
6. Doroudyan, M.H. and Amiri, A., "Root transformation method for monitoring correlated variable and attribute quality characteristics", ***Islamic Countries Society of Statistical Sciences***, (2011), 195-214.
7. Maleki, M., Amiri, A. and Doroudyan, M., "Monitoring multivariate-attribute processes using artificial neural network", in 42th conference on Computers and Industrial Engineering, Cape Town, South Africa.(CIE42), (2012).
8. Amiri, A., Bashiri, M., Mogouie, H. and Doroudyan, M.H., "Non-normal multi-response optimization by multivariate process capability index", ***Scientia Iranica***, Vol. 19, No. 6, (2012), 1894-1905.
9. Maleki, M.R., Amiri, A., Rasouli, M. and in 3rd International Conference on Production and Industrial Engineering, J., India, "Monitoring variability of multivariate-attribute processes using ewma control charts based on norta inverse technique", (2013).
10. Doroudyan, M.H. and Amiri, A., "Monitoring multivariate–attribute processes based on transformation techniques", ***The International Journal of Advanced Manufacturing Technology***, Vol. 69, No. 9-12, (2013), 2161-2172.
11. Doroudyan, M.H. and Amiri, A., "Monitoring correlated variable and attribute quality characteristics based on norta

inverse technique", ***International Journal of Productivity and Quality Management***, Vol. 14, No. 2, (2014), 247-262.

12. Amiri, A. H., Maleki, M.R. and Doroudyan, M.H., "Monitoring variability of multivariate-attribute processes using artificial neural network", (2014).
13. Maleki, M.R. and Amiri, A., "Simultaneous monitoring of multivariate-attribute process mean and variability using artificial neural networks", ***Journal of Quality Engineering and Production Optimization***, (2015), 44-54.
14. Maleki, M.R. and Amiri, A., "Designing an artificial neural network for simultaneous detecting, diagnosing and quantifying the magnitude of mean shift(s) in multivariate-attribute processes", ***Journal of Industrial Engineering Research in Production Systems***, Vol. 2, No. 3, (2014), 21-31.
15. Maleki, M.R., Amiri, A. and Mousavi, S.M., "Step change point estimation in the multivariate-attribute process variability using artificial neural networks and maximum likelihood estimation", ***Journal of Industrial Engineering International***, Vol. 11, No. 4, (2015), 505-515.
16. Subhash, S., "Applied multivariate techniques", John Wily & Sons Inc., Canada, (1996).
17. Cherubini, U., Luciano, E. and Vecchiato, W., "Copula methods in finance", John Wiley & Sons, (2004).

APPENDIX

There are several methods for generating multivariate random numbers in the literature. The proposed methods in which the QC(s) are correlated and follow a known marginal distribution are categorized into three main types including analytic, numeric and simulation based methods. In the analytic and numeric approaches, it is assumed that the joint distribution of QC(s) is known. However, in most situations this assumption is violated. Moreover, these methods are mainly applicable in bivariate situations. In the simulation based methods, the random vectors can be generated only by having the marginal distributions of the QC(s) and their correlation matrix and knowing the joint distribution of QC(s) is not required. NORTA method (a simulation-based method) is based on the transformation of Normal random vectors into our desirable random vectors. For generating vector **X** in a multivariate-attribute process where F_i denotes the marginal distribution of ith; $i=1,...,p+q$ QC, the following steps should be considered. Suppose the correlation matrix between the original QCs is:

$$\mathbf{R_X}=\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1,(p+q)} \\ \rho_{12} & 1 & \cdots & \rho_{2,(p+q)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,(p+q)} & \rho_{2,(p+q)} & \cdots & 1 \end{pmatrix}. \qquad \text{(A-1)}$$

1. Generate multivariate standardized Normal distribution of $\mathbf{Z}=(z_1,...,z_{p+q})$ with the following correlation matrix:

$$\mathbf{R_Z}=\begin{pmatrix} 1 & \rho'_{12} & \cdots & \rho'_{1,(p+q)} \\ \rho'_{12} & 1 & \cdots & \rho'_{2,(p+q)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho'_{1,(p+q)} & \rho'_{2,(p+q)} & \cdots & 1 \end{pmatrix}. \qquad \text{(A-2)}$$

2. The vector **X** is calculated as follow:

$$\mathbf{X}=\begin{pmatrix} F_{x_1}^{-1}(\Phi(z_1)) \\ \vdots \\ F_{x_{p+q}}^{-1}(\Phi(z_{p+q})) \end{pmatrix}, \qquad \text{(A-3)}$$

where $\Phi(.)$ is the cumulative distribution function of standard Normal distribution. The correlation matrix of $\mathbf{R_Z}$ in step 1 depends on the original correlation matrix of $\mathbf{R_X}$. Finding the elements of the $\mathbf{R_Z}$ to achieve the original correlation matrix are usually done through simulation or Newton methods which are time consuming. This issue is important especially when the number of QCs increases. In this paper, the Gaussian copula function is used to find the best values of $\Phi(z_i)$ which leads to obtaining matrix $\mathbf{R_X}$. This issue facilitates using NORTA method in generating multivariate-attribute QC(s). For more information about the Gaussian copula method, refer to Cherubini et al. [17].

Online Monitoring and Fault Diagnosis of Multivariate-attribute Process Mean Using Neural Networks and Discriminant Analysis Technique

M. R. Maleki, R. Sahraeian

Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran

PAPER INFO

Paper history:
Received 01 August 2015
Received in revised form 14 November 2015
Accepted 19 November 2015

Keywords:
Discriminant Analysis
Multivariate-attribute
NORTA Inverse
Neural Network
Fault Detection
Fault Diagnosis

چکیده

در برخی از کاربردهای کنترل کیفیت آماری، داده‌های فرآیند از توزیع نرمال پیروی نمی‌کنند و به وسیله ترکیبی از مشخصه‌های کیفی متغیر و وصفی بیان می‌شوند. علی‌رغم روش‌های متعددی که برای پایش فرآیندهای چند متغیره و چند مشخصه وصفی به طور جداگانه ارائه شده است، تنها روش‌های معدودی برای پایش فرآیندهایی با مشخصه‌های کیفی متغیر و وصفی آمیخته در ادبیات موضوع موجود است. در این مقاله، ابتدا در گام اول روشی مبتنی بر تحلیل تمایز برای تعیین وضعیت بردار میانگین فرآیندهایی با مشخصه‌های کیفی متغیر و وصفی آمیخته تدوین می‌شود. سپس، در گام دوم یک رویکرد جدید ترکیبی مبتنی بر روش تحلیل تمایز و شبکه عصبی مصنوعی به منظور پایش بردار میانگین فرآیند و هم‌چنین تشخیص مشخصه (های) کیفی عامل ایجاد وضعیت خارج از کنترل ارائه می‌شود. هم‌چنین یک مثال عددی مبتنی بر شبیه‌سازی به منظور ارزیابی روش‌های ارائه شده در کشف شیفت و تشخیص عامل خطا آورده می‌شود. در پایان، عمل‌کرد روش ارائه شده در گام دوم به منظور کشف شیفت با نمودار کنترل چند متغیره مربع تی و همچنین توسعه‌ای از یک شبکه عصبی موجود در ادبیات موضوع مقایسه می‌شود

doi: *10.5829/idosi.ije.2014.28.11b.11*