

# Step change point estimation in the multivariate-attribute process variability using artificial neural networks and maximum likelihood estimation

Mohammad Reza Maleki<sup>1</sup> · Amirhossein Amiri<sup>1</sup> · Seyed Meysam Mousavi<sup>1</sup>

Received: 14 January 2015 / Accepted: 20 June 2015 / Published online: 9 July 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** In some statistical process control applications, the combination of both variable and attribute quality characteristics which are correlated represents the quality of the product or the process. In such processes, identification the time of manifesting the out-of-control states can help the quality engineers to eliminate the assignable causes through proper corrective actions. In this paper, first we use an artificial neural network (ANN)-based method in the literature for detecting the variance shifts as well as diagnosing the sources of variation in the multivariate-attribute processes. Then, based on the quality characteristics responsible for the out-of-control state, we propose a modular model based on the ANN for estimating the time of step change in the multivariate-attribute process variability. We also compare the performance of the ANN-based estimator with the estimator based on maximum likelihood method (MLE). A numerical example based on simulation study is used to evaluate the performance of the estimators in terms of the accuracy and precision criteria. The results of the simulation study show that the proposed ANN-based estimator outperforms the MLE estimator under different out-of-control scenarios where different shift magnitudes in the covariance matrix of multivariate-attribute quality characteristics are manifested.

**Keywords** Change point estimation · Covariance matrix · Multilayered perceptron neural network · Multivariate-attribute processes · Phase II

## Introduction

In most statistical process control (SPC) applications, the time when a control scheme triggers an out-of-control signal does not indicate the actual time of change in the process. In such situations, estimating the actual time when the fault is first manifested (called change point) is inevitable, because it can facilitate identifying the assignable causes by searching in a limited time interval. We can conclude from the literature that the most efforts have been focused on change point estimation of univariate processes. As one of the first methods in change point estimation of univariate processes, Samuel et al. (1998) investigated the time of step changes in the X-bar control chart. It is also worth addressing two recent researches in change point estimation of univariate and uni-attribute processes. Assareh et al. (2013) applied Bayesian hierarchical models to estimate change point where there exists a step change, a linear trend and a known multiple number of changes in the Poisson quality characteristic. Amiri et al. (2014) developed a probabilistic neural network (PNN)-based procedure to estimate the variance change point in a univariate process with normal quality characteristic. For more information, refer to the review paper provided by Amiri and Allahyari (2012).

In many manufacturing processes, multivariate or multi-attribute quality characteristics, which are correlated, characterize the quality of the products. For example, El-Midany et al. (2010) proposed modular artificial neural networks (ANNs) to recognize abnormal patterns, to

✉ Amirhossein Amiri  
amirhossein.amiri@gmail.com; amiri@shahed.ac.ir

Mohammad Reza Maleki  
reza.maleki@shahed.ac.ir

Seyed Meysam Mousavi  
sm.mousavi@shahed.ac.ir

<sup>1</sup> Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, Iran

identify the variables responsible for out-of-control signal and classify the abnormal pattern parameters in multivariate processes. On the other hand, Niaki and Nafar (2008) proposed a modular ANN-based method to monitor multi-attribute quality characteristics.

In comparison with processes with a single quality characteristic, fewer methodologies have been developed by researchers for change point estimation of multivariate and multi-attribute processes. Sullivan and Woodall (2000) based on a likelihood ratio statistic proposed a preliminary control chart for change point estimation in the multivariate processes with individual observations when the shift occurs in the mean vector, the covariance matrix or in both. Nedumaran et al. (2000) developed maximum likelihood estimator (MLE) approach to estimate step change point in the multivariate normal process mean. Zamba and Hawkins (2006) presented a model to estimate the time of step change in the mean vector of multivariate processes in situations where the parameters are unknown.

Zarandi and Alaeddini (2010) estimated the time of change in different types of control charts including univariate, uni-attribute as well as the multivariate control charts with either fixed or variable sampling strategy using a general fuzzy-statistical clustering approach. Doğu and Kocakoc (2011) estimated the time of step change in the covariance matrix of multivariate normal processes, in which a multivariate control chart based on sample covariance is used for receiving out-of-control signals. Niaki and Khedmati (2012) proposed a new method to derive the maximum likelihood estimator of the time of a step change in the mean vector of multivariate Poisson processes. To do that, first, they employed two transformations to decrease the inherent skewness involved in multi-attribute processes and made the distribution of quality characteristics almost multivariate normal and diminished correlations between the attributes. After that, they employed a  $T^2$  control chart for detection purposes. Finally, using a maximum likelihood estimator, they found the actual time of a change. Niaki and Khedmati (2013) then proposed a new multi-attribute  $T^2$  control chart based on two transformation methods to monitor the parameter vector of multi-attribute Poisson processes. Then, to estimate the process change point, they developed the maximum likelihood estimators for both linear trend and step change disturbances. Doğu and Kocakoc (2013) estimated the step change point of multivariate normal processes when joint mean vector and covariance matrix shifts occurred.

Besides, some researchers used artificial neural networks for estimating the time of change in multivariate processes. Atashgar and Noorossana (2010) proposed a neural network-based change point estimator to identify the change point in the mean vector of a bivariate normal distribution when the monotonic changes occur. At the

same time, they also diagnosed the variables responsible for the change in the process mean vector. Ahmadzadeh (2009) introduced a neural network to identify the time of change in the multivariate process mean parameters. Noorossana et al. (2011) proposed an integrated supervised learning solution to detect the out-of-control conditions, estimate the change point when the shift occurs in the mean vector, diagnose the variables contributing to the out-of-condition and determine the direction of the shift in the mean of each contributing variable. Allahyari and Amiri (2011) studied step change point estimation problem in multivariate processes using a clustering approach. Atashgar and Noorossana (2011) proposed a supervised learning approach based on artificial neural networks to identify the change point in a bivariate process when the process mean vector changes linearly. They simultaneously performed an analysis to diagnose the variables that contributed to the change in the process mean vector. Ahmadzadeh et al. (2013) developed a multivariate exponentially weighted moving average (MEWMA) control chart by using neural network for identifying the step change point as well diagnosing the variable responsible of the change in the multivariate process mean vector.

In some production environments, the quality of the product is expressed by the combination of both variables and attributes quality characteristics, where a non-zero correlation structure between them exists. Despite some efforts for monitoring multivariate-attribute processes, there is no research in the literature about the change point estimation of such processes. As the main contribution in this paper, we propose a modular methodology based on artificial neural networks for estimating the step changes in the covariance matrix of multivariate-attribute processes.

Considering the literature of using ANNs in different scopes of SPC, we realize that in almost all researches the observations are directly fed to the ANN as the input values. One unique aspect of our research is that we link two ANN modules including (1) module one, which is used for detecting variance shifts (fault detection) as well as diagnosing the source of shifts (fault diagnosis), and (2) module two, which consists of some ANNs used for estimating the time of variance step changes. In the proposed modular methodology, the output values of the neural network designed for detection purpose is used as the input values of the ANN estimators required for change point estimation.

The proposed methodology is presented in Phase II; hence, the parameters of quality characteristics are known based on historical data. The rest of this paper is outlined as follows: in “[Modified maximum likelihood estimator](#)”, the change point estimator based on MLE with some modifications is presented. In “[Proposed model for change point estimation](#)”, the proposed ANN-based methodology consisting of two modules, fault detection/fault diagnosis



module and change point estimation module, is described. Then, in “Performance evaluation”, the performance of the proposed ANN-based method is assessed through the simulation experiments from a given multivariate-attribute process in comparison with the modified MLE estimator. In “Sensitivity analysis”, a sensitivity analysis on the effect of the location of change point on the performance of MLE approach is conducted. In the final section, conclusions as well as some recommendations for future researches are given.

### Modified maximum likelihood estimator

Consider a multivariate-attribute process with  $p$  variables and  $q$  attributes where the quality characteristics are correlated. Let  $\mathbf{X}_{ij} = (X_{ij1}, \dots, X_{ijp}, X_{ij(p+1)}, \dots, X_{ij(p+q)})^T$  be the  $j$ th observations vector ( $j = 1, 2, \dots, n$ ) of the  $i$ th subgroup ( $i = 1, 2, \dots$ ) where the first  $p$  elements represent the variables, whereas the last  $q$  elements represent attribute quality characteristics. In the first step of extending our statistical approach, using Normal to Anything (NORTA) inverse method, the process data are transformed to the multivariate normal distribution. After using NORTA inverse method, the column vector of  $\mathbf{X}'_{ij} = (X'_{ij1}, \dots, X'_{ijp}, X'_{ij(p+1)}, \dots, X'_{ij(p+q)})^T$  with the mean vector of  $\mu_0$  and covariance matrix of  $\Sigma_0$  will be obtained where the marginal distribution of the elements is normal distribution.

In this study, we use a control chart in the literature for monitoring the variability of the transformed quality characteristics. This control chart plots the determinant of the sample covariance matrix ( $|\mathbf{S}|$ ) as a statistic and has the following control limits (Montgomery 2005):

$$\text{UCL} = |\Sigma_0| \left( b_1 + L\sqrt{b_2} \right), \tag{1}$$

$$\text{LCL} = |\Sigma_0| \left( b_1 - L\sqrt{b_2} \right). \tag{2}$$

The value of control limit coefficient ( $L$ ) is set such that a predetermined in-control average run length ( $ARL_0$ ) is obtained. The parameters of  $b_1$  and  $b_2$  depend only on the number of variable and attribute quality characteristics and sample size and can be determined through the following equations, respectively (Montgomery 2005):

$$b_1 = (n - 1)^{-(p+q)} \prod_{i=1}^{p+q} (n - i), \tag{3}$$

$$b_2 = (n - 1)^{-2(p+q)} \times \prod_{i=1}^{p+q} (n - i) \times \left[ \prod_{j=1}^{p+q} (n - j + 2) - \prod_{j=1}^{p+q} (n - j) \right]. \tag{4}$$

The control chart triggers an out-of-control alarm in the  $i$ th subgroup if  $|\mathbf{S}_i|$  exceeds the control limits. The negative value of LCL should be substituted by zero.

Let  $T$  be the time when the extended control chart triggers an out-of-control alarm. Let also after incidence of a given step shift in process variability, the covariance matrix of multivariate-attribute quality characteristics changes from  $\Sigma_0$  to  $\Sigma_1$  at the unknown change point  $\tau$ . After that, the parameters of covariance matrix remain at the new level until the sources of the assignable cause is identified and omitted. Hence, the determinant of sample matrix until the time  $\tau$ , i.e.,  $|\mathbf{S}_1|, |\mathbf{S}_2|, \dots, |\mathbf{S}_\tau|$  come from the in-control state, whereas  $|\mathbf{S}_{\tau+1}|, \dots, |\mathbf{S}_T|$  come from the out-of-control state of the multivariate-attribute process variability. The estimation of the parameter  $\tau$  by the MLE approach ( $\hat{\tau}$ ) is a value that maximizes the likelihood function of observations as follows (Doğu and Kocakoc 2011):

$$\hat{\tau} = \arg \max \{ \text{CP}_t \}, t = 0, 1, \dots, T - 1. \tag{5}$$

The value of  $\text{CP}_t$  is computed according to Eq. (6):

$$\text{CP}_t = \frac{\left( \text{trace} \left( \Sigma_0^{-1} \times \sum_{i=t+1}^T \sum_{j=1}^n (\mathbf{X}'_{ij} - \mu_0)(\mathbf{X}'_{ij} - \mu_0)' \right) \right)}{2} - \frac{n(T-t)}{2} \ln \left( \frac{\sum_{i=t+1}^T \sum_{j=1}^n (\mathbf{X}'_{ij} - \mu_0)(\mathbf{X}'_{ij} - \mu_0)' \times \Sigma_0^{-1}}{n(T-t)} \right) - \frac{n(p+q)(T-t)}{2}, \tag{6}$$

where the trace ( $\mathbf{A}$ ) is the sum of diagonal elements in matrix  $\mathbf{A}$ .

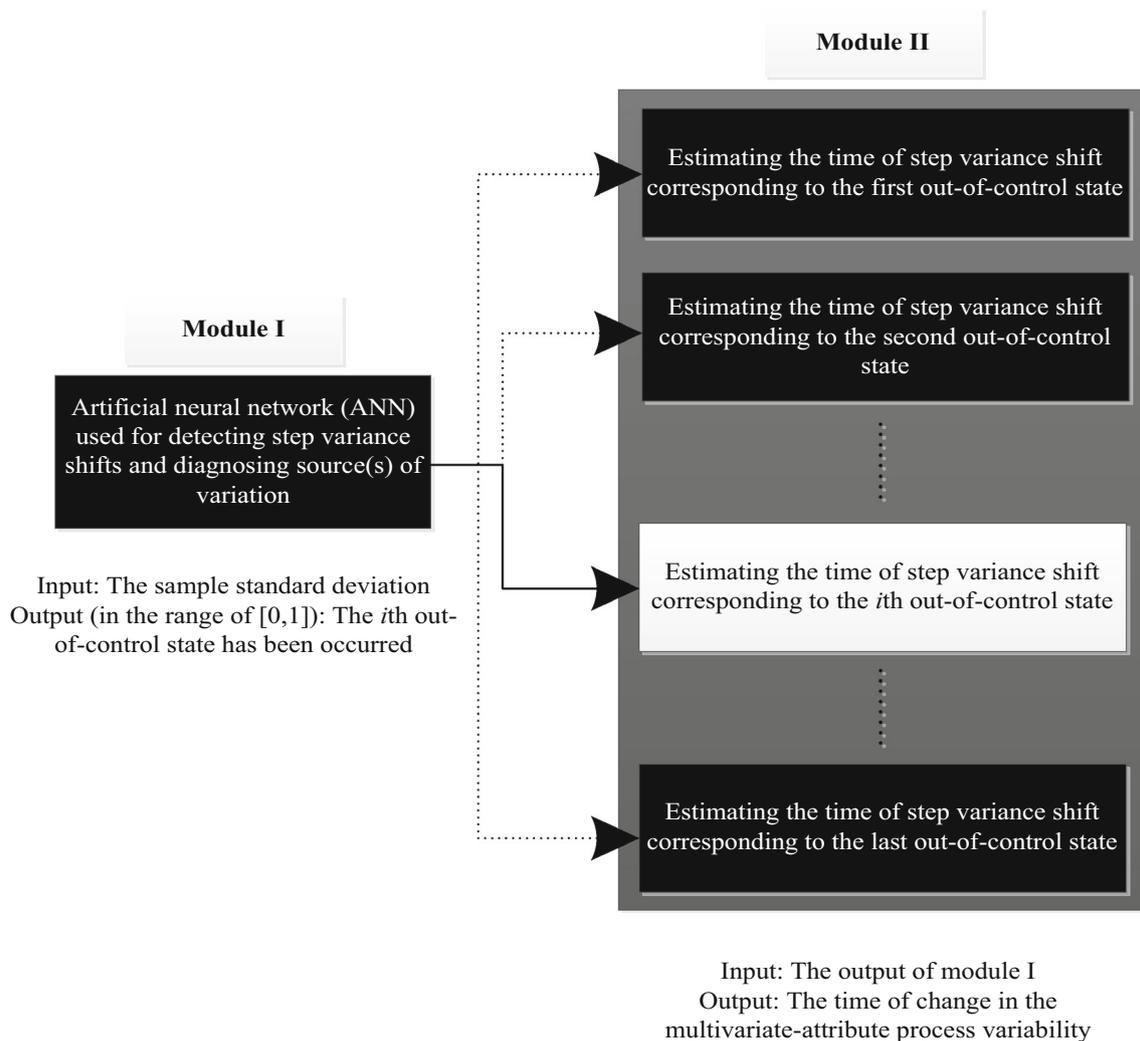
### The proposed model for change point estimation

In this section, the proposed modular model for estimating the change point in the covariance matrix of multivariate-attribute processes is described. In module one, we use a three-layered perceptron neural network that is presented by Amiri et al. (2015) for monitoring the multivariate-attribute process variability. In the first module, we aim to detect the variance shifts and diagnose quality characteristics responsible for out-of-control signals. Note that, diagnosing the out-of-control quality characteristics in the first module is equivalent to identifying which out-of-control state has occurred. Then in the second module, we design an artificial neural network for estimating the change point corresponding to any out-of-control state. In a multivariate-attribute process whose quality is characterized by the combination of  $p$  variables and  $q$  attributes, we

design totally  $\sum_{i=1}^{p+q} \binom{p+q}{i} = 2^{p+q} - 1$  ANN estimators for change estimation in the second module. Then, based on the out-of-control state which is diagnosed in the first module, one of the designed ANN estimators in the second module will be activated for estimating the change point in the covariance matrix of multivariate-attribute quality characteristics. Note that, using only one ANN in change point estimation instead of modular ANNs makes it large and consequently the training process will be complex and time consuming. Figure 1 presents the proposed modular ANN-based methodology (It is supposed that the  $i$ th out-of-control state is diagnosed in the first module).

### Structure of ANN modules

In this subsection, the structure of each ANN estimator required for estimating the time of change in each out-of-control state is illustrated. Due to the successful performance of multilayer perceptron neural networks in various scopes of SPC, this type of neural network is used in designing the ANNs of the change point estimation module. To determine the number of nodes in the input layer of each ANN estimator in the second module, the following procedure is recommended. First in module one, we identify the quality characteristics contributed to the out-of-control signal. In a multivariate-attribute process with  $p$



**Fig. 1** The proposed modular methodology



variables and  $q$  attributes, there are totally  $\binom{p+q}{j}; j, j = 1, 2, \dots, p+q$  states where  $j$  quality characteristics are the sources of variation. Obviously, in the situations where one of the  $p+q$  quality characteristics is responsible for out-of-control signal,  $p+q$  ANNs should be designed for the change point estimation.

The following procedure is used to determine the number of nodes in the input layer of the ANN estimator corresponding to the out-of-control state in which the  $i$ th quality characteristic has been diagnosed as the source of variation: First, we generate random samples in a  $p$ -variate/ $q$ -attribute process, each of size  $n$ . In the simulated random samples  $\delta_i > 1$ , whereas  $\delta_j = 1, j = 1, 2, \dots, p+q, j \neq i$ , where  $\delta_i$  is the standard deviation of the  $i$ th quality characteristic after a given step shift divided by its standard deviation for the in-control state. Then, we enter the generated random samples into the ANN of the first module until the first out-of-control signal is received. This process is repeated in 10,000 replicates and in each replicate we record the run length (RL) values obtained in the first module and save them in a vector like  $c_1$ . Finally, an element in vector  $c_1$  (called  $m_1$ ) such that the value of  $\text{pr}(\text{RL} > m_1)$  obtained in 10,000 simulation replicates is negligible is considered as the number of nodes in the input layer of the neural network designed for change point estimation of the corresponding out-of-control state. To determine the number of input nodes, we should consider two issues: (1) increasing the number of nodes in the input layer of ANN modules leads to the complex ANNs in both training and testing, (2) insufficient number of input nodes is not desired. Because in situations where the run length obtained by module one exceeds the number of input nodes, the ANN module will not be usable. The input vector used in the ANN of the first module is a column vector as follows:

$$S = [s_1, s_2, \dots, s_{p+q}]^T, \tag{7}$$

where  $s_j$  is the sample standard deviation of the  $j$ th quality characteristic in the simulated random sample. Determining the number of input nodes in the ANN modules required for estimating the time of change in the variance of two quality characteristics (totally  $\binom{p+q}{2}$  ANNs) is almost the same as the previous  $p+q$  ANNs. Suppose that the quality characteristics  $u$  and  $v$  are the sources of variation in the covariance matrix of multivariate-attribute quality characteristics. First, we generate random samples of size  $n$  from a  $p$ -variate/ $q$ -attribute process where  $\delta_u, \delta_v > 1$ , whereas  $\delta_j = 1, j = 1, 2, \dots, p+q, j \neq u, v$ . Then, based on 10,000 simulation replicates, we obtain run length values in the first module and save them in a vector like  $c_2$ . Finally, an element like  $m_2$  in the vector  $c_2$  that the value of

$\text{pr}(\text{RL} > m_2)$  obtained from 10,000 simulation replicates is almost equal to zero is considered as the number of nodes in the input layer of these ANNs.

This procedure is performed to determine the number of input nodes in the neural networks that are designed for change point estimations related to other out-of-control states where  $j, j = 3, \dots, p+q$  quality characteristics contribute to the out-of-control situations. The number of nodes in the output layer of each neural network is also considered as equal as the number of its input layer nodes. There is no standard guideline to determine the number of hidden layers as well as the number of nodes in each hidden layer. It is also stipulated in the literature that one or two hidden layers may be sufficient in designing any neural networks. We finalize the number of hidden layers as well as the number of nodes in each one after trial and error experiments. We also use the sigmoid function as the transfer function of all ANNs which send out outputs in the range of  $[0, 1]$ .

### Training procedure of neural networks

The most substantial issue in the training process of each  $2^{p+q} - 1$  ANNs for change point estimation purpose is collecting proper training data sets. The training procedure of the  $k$ th;  $k = 1, 2, \dots, 2^{p+q} - 1$  ANN estimator corresponding to the  $k$ th out-of-control state which contains  $m_k$  nodes in its input layer is presented as follows: The first output value of ANN in the first module is used as the input value of the ANN estimators in the second module.

Assume that the run length value obtained from the first module is equal to  $h, h = 1, 2, \dots, m_k$ . First of all, we generate  $m_k - h$  in-control random samples of size  $n$ . We also simulate  $h$  random samples from a  $p$  variate/ $q$  attribute process where  $k$ th;  $k = 1, 2, \dots, 2^{p+q} - 1$  out-of-control state has occurred. In the next step, for each generated  $m_k$  random samples, we calculate vector  $S$  (the input vector associated with the ANN in the first module) according to Eq. (7). Now, we enter the generated vectors into the ANN in the first module and record the observed values of the first output node. Then, we save these  $m_k$  values in a column vector and consider it as the input vector of the  $k$ th ANN in the second module. Hence, the input vector of the  $k$ th;  $k = 1, 2, \dots, 2^{p+q} - 1$  neural network which contains  $m_k$  nodes in its input layer is a column vector with  $m_k$  elements as follows:

$$O_k = (o_1, \dots, o_{m_k-h}, o_{m_k-h+1}, \dots, o_{m_k})^T. \tag{8}$$

In Eq. (8), the first  $m_k - h$  elements of vector  $O_k$  are related to the in-control samples, whereas the last  $h$  elements are associated with the out-of-control samples. We repeat this process in 100 replicates for each value of RL;  $\text{RL} = 1, \dots, m_k$ . Consequently, for each possible value of

RL, totally 100 column vectors are generated. Consequently,  $100 \times m_k$  column vectors with the size of  $m_k \times 1$  are available as the input data sets required for training the  $k$ th neural network in change point estimation module. Note that to generate out-of-control data in the training step, the magnitude of variance shifts in the out-of-control quality characteristics are determined randomly. Similar to the input vectors, the target vector of  $k$ th;  $k = 1, 2, \dots, 2^p + q - 1$  neural network is a  $m_k \times 1$  vector whose elements are zero, except one element which is equal one. The location of element one in the target vector represents the time when in-control state is terminated, i.e., the first out-of-control sample is taken. After generating the input vectors as well as their corresponding target vectors, using back-propagation algorithm which is the most common supervised training algorithm, the ANN estimators are trained.

### Utilizing the proposed ANN modules for estimating change point

After designing and training all the neural networks required in change point estimation module, we should apply them to estimate the time of actual change in the

multivariate-attribute process variability. As the first step, using the ANN of the first module, the out-of-control state is detected. Then, the quality characteristics which contributed to cause an out-of-control alarm are diagnosed. Recall that the diagnostic process by the ANN in module one is equivalent to determining the out-of-control state. Finally, based on the diagnosed quality characteristics responsible of the change in the multivariate-attribute process variability, only one ANN estimator is activated for change point estimation in the quality characteristics whose variances are changed. The flowchart of the proposed algorithm is depicted in Fig. 2.

Recall that the target values of neural networks in the change point estimation module are zero and one. However, when each ANN is applied in the second module, the observed output values are not exactly equal to zero or one and are in the range of  $[0, 1]$  due to errors. Hence, to overcome this problem, in this paper the maximum observed value of the neural networks is considered as the time when the process goes to the out-of-control state. For example, in an ANN with 50 nodes in its output layer, if the maximum observed value is located in the 35th element of output vector, the estimated time of change by the neural network is obtained equal to  $\hat{\tau} = 34$ .

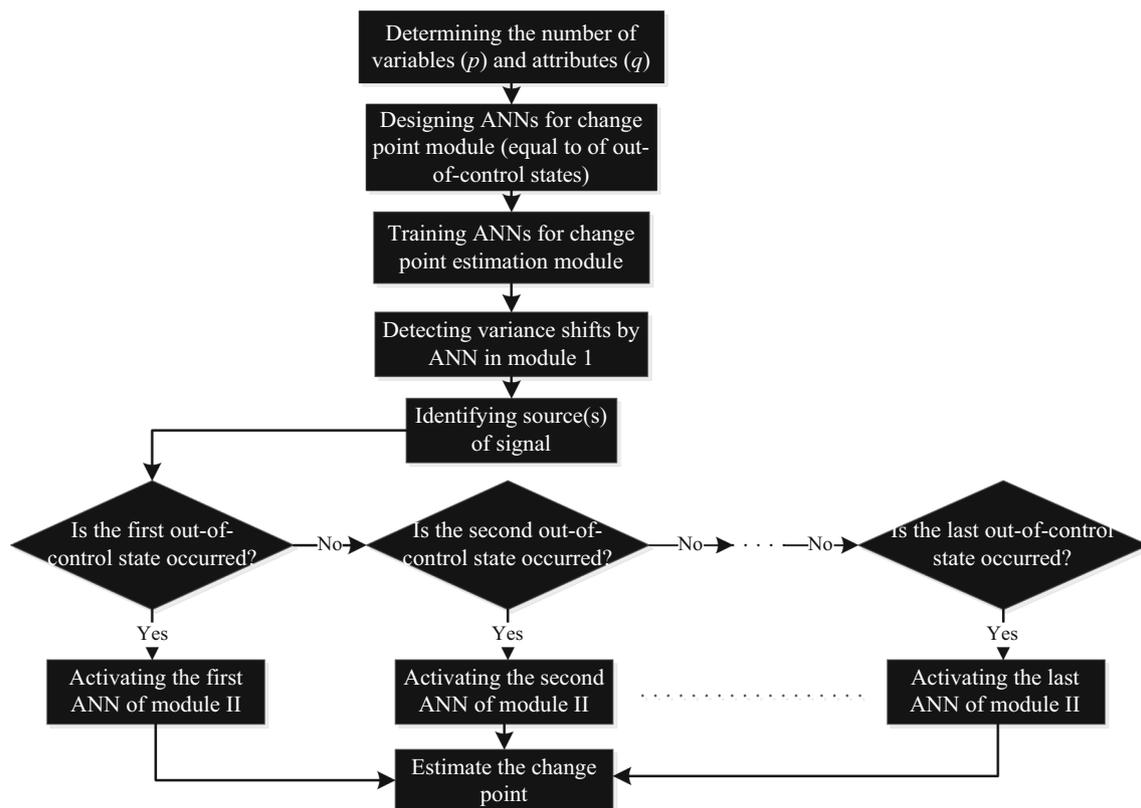


Fig. 2 The proposed ANN-based algorithm



## Performance evaluation

To evaluate the performance of the proposed ANN-based estimator in comparison with the MLE, we present a numerical example based on simulation. In the presented example, we use the random samples of size  $n = 10$  under both methods. Suppose a multivariate-attribute process in which the quality of the products is expressed by one attribute and one variable quality characteristics which are correlated. It is supposed that there is a correlation coefficient equal to  $\rho = 0.357$  between the quality characteristics. The presented numerical example is focused in Phase II. Based on the results of Phase I analysis, we assume that  $x_1$  and  $x_2$  are attribute and variable quality characteristics with the following distributions when the process is in-control:

$$x_1 \sim \text{Poisson}(\lambda = 1), \quad x_2 \sim \text{Normal}(\mu = 3, \sigma^2 = 4).$$

After using NORTA inverse method, the mean vector and the covariance matrix of in-control transformed data are estimated equal to  $\boldsymbol{\mu}_0 = \begin{pmatrix} 0.546 \\ 0 \end{pmatrix}$  and  $\boldsymbol{\Sigma}_0 = \begin{pmatrix} 0.667 & 0.268 \\ 0.268 & 0.999 \end{pmatrix}$ , respectively. To obtain the  $\text{ARL}_0$  roughly equal to 200 by the  $|\text{S}|$  control chart, the control limits are determined equal to  $\text{UCL} = 2.1050$  and  $\text{LCL} = 0$  (the negative value of LCL is substituted by zero). In the MLE method, first we generate 100 in-control data and transform them to obtain bivariate normal distribution data. For the first 100 in-control random samples, it is assumed that no false alarm is received by the extended control chart. Hence, to overcome the probable problem, if the value of  $|\text{S}|$  corresponding to each simulated sample exceeds the UCL, we replace it by another in-control random sample. Starting from the 101th taken sample, the observations are simulated from an out-of-control state until the extended control chart triggers an out-of-control signal. Then, the change point in the process variability is estimated according to Eqs. (5) and (6). For different variance shift magnitudes, we repeat this procedure 10,000 times. Finally, in each replicate we calculate the absolute difference of the actual and estimated change point ( $|\hat{\tau} - \tau|$ ). Then, based on the simulated values of  $|\hat{\tau} - \tau|$ , we investigate the performance of the MLE method.

Obviously in this process ( $p = q = 1$ ), there are three states where the process variability is out of control. The out-of-control states in the process variability include (1) the states in which  $x_1$  contributes to out-of-control signals, (2) states in which  $x_2$  contributes to out-of-control signals and (3) situations in which both  $x_1$  and  $x_2$  contribute to out-of-control signals. Hence, three ANNs are required to be designed for estimating the actual time of change for each out-of-control state of process variability.

## Designing the ANN estimators

Using the ANN of module one for detecting purpose, according to “[Proposed model for change point estimation](#)”, the value of  $\text{Pr}(\text{RL} > 40)$  under a small variance shift of  $\delta_1 = 1.4$  under 10,000 simulation replicates is obtained equal to a negligible value (nearly zero). Hence, it seems reasonable that using 40 nodes in the input layer of the first ANN in module two (network A) is a proper choice. As noted, the number of nodes in the output layer of network A is also considered equal to 40. Designing the other ANNs (networks B and C) that are designed for change point estimation in variable  $x_2$  as well as in joint quality characteristics  $x_1$  and  $x_2$  is almost similar to network A. Using module one, the probability value of  $\text{Pr}(\text{RL} > 40)$  under a variance shift with a magnitude of  $\delta_2 = 1.4$  under 10,000 simulation replicates is obtained equal to 0.002. Hence, we consider 40 nodes in the input and output layers of network B. Similarly, the probability value of  $\text{Pr}(\text{RL} > 30)$  under a joint shift with a magnitude of ( $\delta_1 = 1.4, \delta_2 = 1.4$ ) under 10,000 simulation replicates is obtained equal to zero. Consequently, 30 nodes will exist in the input and output layer of network C. By trial and error experiments, the ANN estimators A, B and C have a single hidden layer with 22, 24 and 22 nodes, respectively.

## Training the ANN estimators

According to “[Training procedure of neural networks](#)”, the input data as well as the corresponding target vectors required for training the networks A and B are both the column vectors with 40 elements. To generate out-of-control data sets in training the neural network A, we use the shifts with the random magnitude of  $\delta_1$  in the standard deviation of  $x_1$  where  $\delta_1$  is uniformly distributed in the range of [1.5, 2.5]. Similarly, to generate out-of-control data sets required for training network B, the shifts with the random magnitude of  $\delta_2$  in the standard deviation of  $x_2$  are used where  $\delta_2$  is also uniformly distributed in the range of [1.5, 2.5]. The input and target vectors that are prepared for training the network C are also the column vectors with 30 elements. Furthermore, we use simultaneous shifts in the standard deviation of both  $x_1$  and  $x_2$  with random magnitude of shifts distributed uniformly in the range of  $\delta_i \in [1.5, 2.5]; i = 1, 2$ . Now, 4000 input data for training network A and B and 3000 ones for training network C are available. Finally, the ANN estimators are trained with the generated input vectors as well as their corresponding target vectors. The value of mean-squared error (MSE) which is an evaluation criterion in training stage for neural networks A, B and C is obtained equal to 0.0166, 0.0138 and 0.0142, respectively. When each trained network is

**Table 1** Performance of the first ANN (Network A) and MLE in estimating change point in the variance of  $x_1$

Shift	(1.4 $\sigma_1$ , 0)		(1.6 $\sigma_1$ , 0)		(1.8 $\sigma_1$ , 0)		(2 $\sigma_1$ , 0)	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	8.74	32.73	2.23	20.48	0.90	14.55	0.53	9.55
Std ( $ \hat{\tau} - \tau $ )	8.71	43.02	3.26	37.94	1.53	33.70	1.04	28.16
$p( \hat{\tau} - \tau  = 0)$	11.76	33.94	30.24	57.20	48.16	72.56	61.74	81.94
$p( \hat{\tau} - \tau  \leq 1)$	24.30	44.12	58.52	66.24	82.88	79.26	92.88	86.56
$p( \hat{\tau} - \tau  \leq 2)$	32.48	48.54	71.58	69.66	92.12	81.04	97.60	87.62
$p( \hat{\tau} - \tau  \leq 3)$	37.60	51.64	78.42	71.32	95.22	81.58	98.54	87.82
$p( \hat{\tau} - \tau  \leq 4)$	42.84	53.52	83.82	72.32	97.04	81.86	99.08	87.92
$p( \hat{\tau} - \tau  \leq 5)$	48.08	55.00	87.70	73.10	98.34	82.00	99.46	87.98
Shift	(2.2 $\sigma_1$ , 0)		(2.4 $\sigma_1$ , 0)		(2.6 $\sigma_1$ , 0)		(2.8 $\sigma_1$ , 0)	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	0.37	5.41	0.30	2.78	0.28	0.93	0.25	0.17
Std ( $ \hat{\tau} - \tau $ )	0.88	21.41	0.73	15.56	0.72	8.90	0.66	2.82
$p( \hat{\tau} - \tau  = 0)$	70.24	88.60	75.16	93.14	76.50	96.66	79.14	98.60
$p( \hat{\tau} - \tau  \leq 1)$	96.58	91.64	97.54	95.52	97.66	98.28	97.54	99.58
$p( \hat{\tau} - \tau  \leq 2)$	99.00	92.08	99.14	95.74	99.36	98.44	99.50	99.64
$p( \hat{\tau} - \tau  \leq 3)$	99.18	92.22	99.34	95.82	99.48	98.44	99.52	99.64
$p( \hat{\tau} - \tau  \leq 4)$	99.50	92.22	99.54	95.84	99.74	98.44	99.68	99.64
$p( \hat{\tau} - \tau  \leq 5)$	99.66	92.24	99.78	95.84	99.82	98.44	99.84	99.64

applied for estimating the time of change in the corresponding quality characteristics, we focus on the maximum observed output and consider it as the estimation of time when the first out-of-control sample is manifested in the process.

To investigate the performance of ANN and MLE methods and provide a comparison study, we use three criteria in change point estimation literature including the mean, the standard deviation as well as the empirical distribution of estimated change point obtained by the methods around the actual time of process change. The performance of neural network A under the step variance shift ( $\delta_1\sigma_1, 0$ ) with various magnitudes of  $\delta_1$  is investigated through 10,000 replicates and compared with MLE. Table 1 shows the performance of the ANN and MLE estimators in estimating the time of change in the variance of  $x_1$ . The results of Table 1 show that under each variance shift magnitudes in  $x_1$ , including small, moderate and large shifts except shift with the magnitude of (1.4 $\sigma_1, 0$ ), the ANN method in all criteria outperforms the MLE method. It is obvious that as the variance shift in the quality characteristic  $x_1$  increases, the preciseness and accuracy of both estimators increase.

Table 2 shows the performance of the neural network B as well as the MLE method under the step variance shift (0,  $\delta_2\sigma_2$ ) with various values of  $\delta_2$  through 10,000 replicates. The computational results in Table 2 indicate that under each out-of-control state considered, the ANN considerably presents more accurate and more precise results in

comparison to MLE. We can see that under both estimators, increasing the magnitude of variance shift in the quality characteristic  $x_2$  leads to better estimates.

In Table 3, we summarize the results of applying neural network C as well as MLE in estimating change point when joint variance shifts in both quality characteristics have occurred. Similar to neural networks A and B, the neural network C performs adequately in estimating the time of change in the process variability when both quality characteristics are the sources of the out-of-control signals. Moreover, the results of Table 3 show the superior performance of the third ANN estimator (Network C) under all change point criteria in comparison with MLE method.

The results of Tables 1 and 3 show that in all out-of-control scenarios, the MLE approach is better than ANN in terms of the  $p(|\hat{\tau} - \tau| = 0)$  criterion. It means that in 10,000 replicates, the number of simulation runs in which the MLE estimates the actual time of change is more than ANN. However, in some simulation experiments, the MLE estimates are imprecise. Consequently, in almost all out-of-control scenarios except (1.4 $\sigma_1, 0$ ) and (1.4 $\sigma_1, 1.4\sigma_2$ ), the ANN outperforms MLE in terms of the  $p(|\hat{\tau} - \tau| \leq i); i = 1, 2, 3, 4, 5$  criterion. In contrast to the results in Tables 1 and 3, it is observed in Table 2 that the ANN outperforms the MLE in terms of  $p(|\hat{\tau} - \tau| = 0)$ . Generally, the ANN estimates are more reliable than MLE, because before applying the ANN modules, they are trained with sufficient data set.

**Table 2** Performance of the second ANN (Network B) and MLE in estimating change point in the variance of  $x_2$

Shift	(0, $1.4\sigma_2$ )		(0, $1.6\sigma_2$ )		(0, $1.8\sigma_2$ )		(0, $2\sigma_2$ )	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	2.81	45.17	1.20	40.48	0.97	35.97	0.73	31.32
Std ( $ \hat{\tau} - \tau $ )	3.79	44.23	1.98	44.73	2.00	44.20	1.86	43.67
$p( \hat{\tau} - \tau  = 0)$	24.86	19.88	43.64	32.54	54.50	43.84	66.06	53.48
$p( \hat{\tau} - \tau  \leq 1)$	49.60	31.86	73.88	44.80	81.40	53.94	86.92	61.02
$p( \hat{\tau} - \tau  \leq 2)$	64.24	37.00	87.06	49.84	90.78	57.08	93.10	62.72
$p( \hat{\tau} - \tau  \leq 3)$	73.66	40.62	92.84	51.10	94.30	58.22	95.68	63.40
$p( \hat{\tau} - \tau  \leq 4)$	80.24	42.72	95.32	52.18	95.30	58.62	96.14	63.72
$p( \hat{\tau} - \tau  \leq 5)$	85.32	44.22	97.32	52.62	97.60	58.70	98.10	63.96
Shift	(0, $2.2\sigma_2$ )		(0, $2.4\sigma_2$ )		(0, $2.6\sigma_2$ )		(0, $2.8\sigma_2$ )	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	0.71	26.54	0.63	22.78	0.59	18.13	0.63	15.52
Std ( $ \hat{\tau} - \tau $ )	2.04	41.33	2.16	39.40	2.06	36.47	2.26	34.55
$p( \hat{\tau} - \tau  = 0)$	73.50	61.62	80.46	68.30	83.40	74.68	84.96	79.06
$p( \hat{\tau} - \tau  \leq 1)$	87.66	67.96	89.30	72.94	89.69	78.82	88.98	82.26
$p( \hat{\tau} - \tau  \leq 2)$	91.08	69.04	91.14	73.64	90.78	79.38	89.98	82.42
$p( \hat{\tau} - \tau  \leq 3)$	94.90	69.40	95.08	73.84	95.36	79.50	94.80	82.58
$p( \hat{\tau} - \tau  \leq 4)$	95.54	69.64	95.44	74.00	95.56	79.56	95.04	82.70
$p( \hat{\tau} - \tau  \leq 5)$	97.72	69.82	97.50	74.20	97.38	79.60	97.04	82.74

**Table 3** Performance of the third ANN (Network C) in estimating change point in the variances of  $x_1$  and  $x_2$

Shift	(1.4 $\sigma_1$ , 1.4 $\sigma_2$ )		(1.5 $\sigma_1$ , 1.5 $\sigma_2$ )		(1.6 $\sigma_1$ , 1.6 $\sigma_2$ )		(1.7 $\sigma_1$ , 1.7 $\sigma_2$ )	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	1.58	32.21	0.76	27.57	0.51	22.38	0.44	15.00
Std ( $ \hat{\tau} - \tau $ )	2.60	42.47	1.33	41.16	1.15	38.59	1.23	33.22
$p( \hat{\tau} - \tau  = 0)$	39.10	47.54	51.78	57.30	61.20	66.32	67.02	76.78
$p( \hat{\tau} - \tau  \leq 1)$	70.18	55.50	86.58	63.84	93.20	71.12	95.04	80.12
$p( \hat{\tau} - \tau  \leq 2)$	82.02	58.12	94.74	65.42	98.62	72.46	99.20	81.04
$p( \hat{\tau} - \tau  \leq 3)$	87.82	59.38	97.50	66.24	99.36	72.80	99.58	81.52
$p( \hat{\tau} - \tau  \leq 4)$	91.34	60.08	98.40	66.76	99.66	73.12	99.68	81.76
$p( \hat{\tau} - \tau  \leq 5)$	93.48	60.74	98.94	67.08	99.70	73.40	99.68	81.84
Shift	(1.8 $\sigma_1$ , 1.8 $\sigma_2$ )		(1.9 $\sigma_1$ , 1.9 $\sigma_2$ )		(2 $\sigma_1$ , 2 $\sigma_2$ )		(2.5 $\sigma_1$ , 2.5 $\sigma_2$ )	
	ANN	MLE	ANN	MLE	ANN	MLE	ANN	MLE
$E( \hat{\tau} - \tau )$	0.37	10.14	0.33	5.39	0.29	2.55	0.22	0.58
Std ( $ \hat{\tau} - \tau $ )	0.89	28.32	0.96	21.25	0.91	14.68	0.91	7.52
$p( \hat{\tau} - \tau  = 0)$	70.94	83.96	74.66	90.78	78.50	94.80	86.22	99.00
$p( \hat{\tau} - \tau  \leq 1)$	95	86.74	95.42	92.88	95.32	96.44	95.46	99.40
$p( \hat{\tau} - \tau  \leq 2)$	99.46	87.42	99.48	93.38	99.56	96.74	99.28	99.42
$p( \hat{\tau} - \tau  \leq 3)$	99.68	87.72	99.68	93.50	99.74	96.82	99.52	99.42
$p( \hat{\tau} - \tau  \leq 4)$	99.82	87.88	99.78	93.68	99.82	96.86	99.74	99.42
$p( \hat{\tau} - \tau  \leq 5)$	99.82	87.96	99.80	93.72	99.82	96.90	99.78	99.42

**Table 4** The effect of parameter  $\tau$  on MLE approach in estimating change point under variance shifts in  $x_1$

Shift	$(1.4\sigma_1, 0)$		$(1.6\sigma_1, 0)$		$(1.8\sigma_1, 0)$		$(2\sigma_1, 0)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	25.57	36.93	15.84	31.62	12.04	28.62	6.43	21.63
100	32.73	43.02	20.48	37.94	14.55	33.70	9.55	28.16
110	38.18	48.56	26.31	45.03	17.53	38.90	12.11	33.34
Shift	$(2.2\sigma_1, 0)$		$(2.4\sigma_1, 0)$		$(2.6\sigma_1, 0)$		$(2.8\sigma_1, 0)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	4.48	18.18	2.55	14.15	0.81	7.46	0.17	2.82
100	5.41	21.41	2.78	15.56	0.94	8.90	0.17	2.82
110	6.49	24.98	3.34	18.04	0.95	9.23	0.27	5.71

**Table 5** The effect of parameter  $\tau$  on MLE approach in estimating change point under variance shifts in  $x_2$

Shift	$(0, 1.4\sigma_2)$		$(0, 1.6\sigma_2)$		$(0, 1.8\sigma_2)$		$(0, 2\sigma_2)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	36.52	39.27	32.79	38.97	30.31	38.95	24.81	38.26
100	45.17	44.23	40.48	44.73	35.97	44.73	31.32	43.67
110	53.41	49.22	47.96	49.89	41.30	49.49	34.65	48.44
Shift	$(0, 2.2\sigma_2)$		$(0, 2.4\sigma_2)$		$(0, 2.6\sigma_2)$		$(0, 2.8\sigma_2)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	21.55	35.80	16.57	32.36	14.15	31.02	11.25	28.26
100	26.54	41.33	22.78	39.40	18.13	36.47	15.52	34.55
110	32.89	47.16	27.50	47.16	21.63	41.47	20.16	40.72

**Table 6** The effect of parameter  $\tau$  on MLE approach in estimating change point under shifts in  $x_1$  and  $x_2$

Shift	$(1.4\sigma_1, 1.4\sigma_2)$		$(1.5\sigma_1, 1.5\sigma_2)$		$(1.6\sigma_1, 1.6\sigma_2)$		$(1.7\sigma_1, 1.7\sigma_2)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	26.53	37.17	21.23	36.11	18.31	33.52	12.22	28.66
100	32.21	42.47	27.57	41.16	22.38	38.59	15.00	32.22
110	42.59	48.95	34.53	47.81	26.92	44.00	20.23	39.82
Shift	$(1.8\sigma_1, 1.8\sigma_2)$		$(1.9\sigma_1, 1.9\sigma_2)$		$(2\sigma_1, 2\sigma_2)$		$(2.5\sigma_1, 2.5\sigma_2)$	
	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )	$E( \hat{\tau} - \tau )$	Std ( $ \hat{\tau} - \tau $ )
90	7.95	24.04	4.22	17.80	1.63	11.07	0.09	2.81
100	10.14	28.32	5.39	21.25	2.55	14.68	0.58	7.52
110	12.49	33.15	8.62	27.89	3.35	18.16	0.55	7.69

**Sensitivity analysis**

In this section, a sensitivity analysis on the location of the change point ( $\tau$ ) in the MLE approach is performed. The results of estimating the time of change considering

separate variance shifts in  $x_1$  and  $x_2$  under different values of  $\tau$  are summarized in Tables 4 and 5, respectively. The results of estimating the change point in both quality characteristics under different values of  $\tau$  are shown in Table 6. It is concluded that in all out-of-control scenarios

including separate and joint shifts under different shift magnitudes, decreasing the location of actual change point  $\tau$  leads to more accurate and precise estimates.

## Conclusions and future researches

In this paper, a modular ANN-based methodology for estimating the time when the out-of-control state is manifested in the multivariate-attribute processes was studied. Before designing the ANNs required for change point estimation in the process variability, a three-layered perceptron ANN was applied for detecting the variance shifts and identifying quality characteristics responsible for the out-of-control states. After that, based on the out-of-control quality characteristics diagnosed by this ANN, a three-layered perceptron neural network was trained to estimate the actual time of change in the responsible quality characteristics. We used back-propagation algorithm to train all ANN modules in the proposed change point estimation methodology. Moreover, we present the MLE approach for estimating the time of variance shifts. The performance of the proposed ANN-based method as well as the modified MLE in estimating the time of change in the contributed quality characteristics was investigated through a numerical example based on simulations. The results show that the proposed ANN-based methodology outperforms the MLE in estimating the time of step change in the process variability with multivariate-attribute quality characteristics. Finally, we investigate the effect of the location of the change point ( $\tau$ ) on the performance of MLE approach through sensitivity analysis.

Developing our proposed methods in the case of multiple change points is highly recommended as a future research. In addition, proposing the ANN-based methodology for estimating drift and isotonic change points can also be fruitful areas for future researches.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Ahmadzadeh F (2009) Change point detection with multivariate control charts by artificial neural network. *Int J Adv Manuf Technol*. doi:10.1007/s00170-009-2193-6
- Ahmadzadeh F, Lundberg J, Strömberg T (2013) Multivariate process parameter change identification by neural network. *Int J Adv Manuf Technol* 69(9–12):2261–2268
- Allahyari S, Amiri A (2011) Clustering approach for change point estimation in multivariate normal processes. In: *Proceedings of the 41st International Conference on Computers & Industrial Engineering*, 38–43
- Amiri A, Allahyari S (2012) Change point estimation methods for control chart postsignal diagnostics: a literature review. *Qual Reliab Eng Int* 28(7):673–685
- Amiri A, Niaki STA, Moghadam AT (2014) A probabilistic artificial neural network-based procedure for variance change point estimation. *Soft Comput* 19:1–10
- Amiri A, Maleki MR, Doroudyan MH (2015) Monitoring variability of multivariate-attribute processes using artificial neural network. *Prod Op Manag* 5(2):21–36
- Assareh H, Noorossana R, Mengersen KL (2013) Bayesian change point estimation in Poisson-based control charts. *J Indus Eng Int* 9(32):1–13
- Atashgar K, Noorossana R (2010) Identifying change point in a bivariate normal process mean vector with monotonic changes. *Int J Ind Eng Prod Manag* 21(1):1–13 (In Farsi)
- Atashgar K, Noorossana R (2011) An integrating approach to root cause analysis of a bivariate mean vector with a linear trend disturbance. *Int J Adv Manuf Technol* 52(1–4):407–420
- Doğu E, Kocakoc ID (2011) Estimation of change point in generalized variance control chart. *Commun Stat Simul Comput* 40(3):345–363
- Doğu E, Kocakoc ID (2013) A multivariate change point detection procedure for monitoring mean and covariance simultaneously. *Commun Stat Simul Comput* 42(6):1235–1255
- El-Midany TT, El-Baz MA, Abd-Elwahed MS (2010) A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks. *Expert Syst Appl* 37(2):1035–1042
- Montgomery DC (2005) *Introduction to statistical quality control*, 5th edn. John Wiley & Sons, USA
- Nedumaran G, Pignatiello JJ Jr, Calvin JA (2000) Identifying the time of a step-change with  $\chi^2$  control charts. *Qual Eng* 13(2):153–159
- Niaki STA, Khedmati M (2012) Detecting and estimating the time of a step-change in multivariate Poisson processes. *Sci Iran* 19(3):862–871
- Niaki STA, Khedmati M (2013) Estimating the change point of the parameter vector of multivariate Poisson processes monitored by a multi-attribute  $T^2$  control chart. *Int J Adv Manuf Technol* 64(9–12):1625–1642
- Niaki STA, Nafar M (2008) An artificial neural network approach to monitor and diagnose multi-attribute quality control processes. *J Indus Eng Int* 4(7):10–24
- Noorossana R, Atashgar K, Saghaei A (2011) An integrated supervised learning solution for monitoring process mean vector. *Int J Adv Manuf Technol* 56(5–8):755–765
- Samuel TR, Pignatiello JJ Jr, Calvin JA (1998) Identifying the time of a step change with  $\bar{X}$  control charts. *Qual Eng* 10(3):521–527
- Sullivan JH, Woodall WH (2000) Change-point detection of mean vector or covariance matrix shifts using multivariate individual observations. *IIE Trans* 32(6):537–549
- Zamba KD, Hawkins DM (2006) A multivariate change-point model for statistical process control. *Technometrics* 48(4):539–549
- Zarandi MHF, Alaeddini A (2010) A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts. *Inf Sci* 180(16):3033–3044

