ICIC Express Letters
Part B: Applications
Volume **7**, Number **2**, February **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 415–419

CLASSIFICATIONS OF PROCESS VARIANCE FAULTS USING ARTIFICIAL NEURAL NETWORKS AND MULTIVARIATE ADAPTIVE REGRESSION SPLINES

Yuehjen E. Shao, Chia-Ding Hou*, Yu-Jing Chou and Shih-Chieh Lin

Department of Statistics and Information Science
Fu Jen Catholic University
No. 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 24205, Taiwan
*Corresponding author: stat0002@mail.fju.edu.tw

Received July 2015; accepted September 2015

Abstract. *Correct identification of the source of process disturbance for a multivariate process is an important research issue and has excited considerable interest in recent years. There have been many methods developed for detection of the source of process shifts, either by statistical-based approaches or machine learning techniques. In this study, we employ two computational intelligence approaches, artificial neural networks (ANN) and multivariate adaptive regression splines (MARS), to classify the sources of variance shifts in a multivariate normal process. The five quality variables are considered in the multivariate process. In addition, the simulation experiments are conducted to evaluate the performance of the ANN and MARS approaches.*

Keywords: Multivariate normal process, Variance shift, Artificial neural networks, Multivariate adaptive regression splines

1. **Introduction.** With the recent development of production and sensing techniques, the monitoring and diagnosis of multivariate process data have attracted considerable attention in industrial research. The multivariate process control charts have played an important role in industry, and are a useful tool and widely used for detecting multivariate process disturbance because quality characteristics are usually highly correlated. The multivariate control charts would be able to trigger a signal when disturbance occurred in the multivariate process. However, it is difficult to determine which quality characteristics are responsible for this signal. Consequently, the recognition of the source of process shifts becomes a very important research issue in industry.

There have been many studies that investigated the identification of the source of process shifts. Most of these studies have devoted to the determination of the source of either mean shifts or variance shifts for a multivariate normal process. These methods can be divided into two lines. On the first line basically the methods were developed using the statistical-based approaches [1-4]; on the second line the methods were developed on the basis of machine learning techniques [5-10]. This study focuses on the use of computational intelligence approaches to determine the possible sets of quality variables that are responsible for process variance shifts. Due to the fact that they possess the excellent forecasting and/or classification capability, this study considers the artificial neural networks (ANN) and multivariate adaptive regression splines (MARS) as the proposed approaches [11-16].

The structure of this study is organized as follows. Section 2 addresses the structure of a multivariate normal process and the variance shifts. Section 3 describes the proposed approaches for determining the source of variance shifts for a multivariate normal process. Section 4 states the simulation results. The final section provides the research findings and presents a conclusion to complete this study.

2. The Structure of the Process and Variance Shifts. Assume that a multivariate normal process is monitored by an $|\mathbf{S}|$ control chart on p quality characteristics. Let

$$\underset{\sim}{X}_i = [X_{i1}, X_{i2}, \ldots, X_{ip}]', \quad i = 1, 2, \ldots, n \tag{1}$$

be a $p \times 1$ vector which denotes p characteristics on the i^{th} observation with multivariate normal distribution. The corresponding sample covariance matrix is

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^{n} \left(\underset{\sim}{X}_i - \overline{\underset{\sim}{X}} \right) \left(\underset{\sim}{X}_i - \overline{\underset{\sim}{X}} \right)' \tag{2}$$

where $\overline{\underset{\sim}{X}} = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{X}_i$. Let Σ_0 be the in-control covariance matrix which is defined as follows:

$$\Sigma_0 = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,j} & \cdots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \vdots & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \sigma_{i,1} & \cdots & \vdots & \sigma_{i,j} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p,1} & \sigma_{p,2} & \cdots & \sigma_{p,j} & \cdots & \sigma_{p,p} \end{bmatrix}_{p \times p} \tag{3}$$

To monitor a multivariate process variance shift, we can apply the sample generalized variance $|\mathbf{S}|$, and the following control limits proposed by [17]:

$$\begin{aligned} UCL &= |\Sigma_0| \left(b_1 + 3\sqrt{b_2} \right) \\ LCL &= \max \left(0, |\Sigma_0| \left(b_1 - 3\sqrt{b_2} \right) \right) \end{aligned} \tag{4}$$

where

$$\begin{aligned} b_1 &= \frac{1}{(n-1)^p} \prod_{i=1}^{p} (n-i) \\ b_2 &= \frac{1}{(n-1)^{2p}} \prod_{i=1}^{p} (n-i) \left(\prod_{i=1}^{p} (n-i+2) - \prod_{i=1}^{p} (n-i) \right) \end{aligned} \tag{5}$$

When an $|\mathbf{S}|$ control chart generates the out-of-control signals, the problem accompanied is how to determine variables that are assignable as responsible for these signals. Apparently, if we monitor p quality variables simultaneously, there are $2^p - 1$ possible types of variance shifts. Let Σ_1 be the out-of-control covariance matrix. This study adopts the suggestion of [6] and considers the following variance shift as the process fault:

$$\Sigma_1 = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \theta\sigma_{1,j} & \sigma_{1,j+1} & \cdots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \theta\sigma_{2,j} & \sigma_{2,j+1} & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \theta\sigma_{i,1} & \theta\sigma_{i,2} & \cdots & \theta^2\sigma_{i,j} & \theta\sigma_{i,j+1} & \cdots & \theta\sigma_{i,p} \\ \sigma_{i+1,1} & \sigma_{i+1,2} & \cdots & \theta\sigma_{i+1,j} & \sigma_{i+1,j+1} & \cdots & \sigma_{i+1,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p,1} & \sigma_{p,2} & \cdots & \theta\sigma_{p,j} & \sigma_{p,j+1} & \cdots & \sigma_{p,p} \end{bmatrix}_{p \times p} \tag{6}$$

where θ is the inflated ratio.

3. The Methodologies. In this section, this study presents the modeling concept of ANN and MARS.

3.1. **Artificial neural networks.** ANN is a parallel system comprised of highly interconnected processing elements that are based on neurobiological models. ANN processes information through the interactions of a large number of simple processing elements, the neurons. ANN modeling can be described briefly as follows. The relationship between output (y) and inputs ($x_1, x_2, \ldots, x_a$) in an ANN model can be formed as:

$$y = \alpha_0 + \sum_{j=1}^{b} \alpha_j g\left(\delta_{0_j} + \sum_{i=1}^{a} \delta_{ij} x_i\right) + \varepsilon \tag{7}$$

where α_j ($j = 0, 1, 2, \ldots, b$) and δ_{ij} ($i = 0, 1, 2, \ldots, a$; $j = 0, 1, 2, \ldots, b$) are model connection weights; a is the number of input nodes; b is the number of hidden nodes, and ε is the error term. The transfer function in the hidden layer is often represented by a logistic function,

$$g(z) = \frac{1}{1 + \exp(-z)} \tag{8}$$

Accordingly, the ANN model in Equation (7) accomplishes a nonlinear functional mapping from the inputs ($x_1, x_2, \ldots, x_a$) to the output y,

$$y = f(x_1, x_2, \ldots, x_a, w) + \varepsilon \tag{9}$$

where w is a vector of all model parameters, and f is a function determined by the ANN structure and connection weights.

3.2. **Multivariate adaptive regression splines.** The general MARS function can be described as follows.

$$\hat{f}(x) = b_0 + \sum_{m=1}^{M} b_m \prod_{k=1}^{K_m} [S_{km}(x_{v(k,m)} - t_{km})] \tag{10}$$

where b_0 and b_m are the parameters, M is the number of basis functions (BF), K_m is the number of knots, S_{km} takes on values of either 1 or -1 and indicates the right or left sense of the associated step function, $v(k, m)$ is the label of the independent variable, and t_{km} is the knot location. The optimal MARS model is chosen in a two-step procedure. Firstly, construct a large number of basis functions to fit the data initially. Secondly, basis functions are deleted in order of least contribution using the generalized cross-validation (GCV) criterion. To measure the importance of a variable, we can observe the decrease in the calculated GCV values when a variable is removed from the model. The GCV is defined as follows:

$$\text{LOF}\left(\hat{f}_M\right) = \text{GCV}(M) = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - \hat{f}_M(x_i)\right]^2 \Big/ \left[1 - \frac{C(M)}{n}\right]^2 \tag{11}$$

where n is the observations and $C(M)$ is the cost penalty measures of a model containing M basis function.

4. **Simulation Results.** The computer experiments were performed in order to show the performance for two computational intelligence approaches. This study considers 5 quality characteristics for a multivariate normal process, and consequently, we have $2^5 - 1$ possible types of variance shifts. They are represented by $(1, 0, 0, 0, 0)$, $(0, 1, 0, 0, 0)$, $\ldots$, and $(1, 1, 1, 1, 1)$, where 1 denotes a quality characteristic that is at fault and 0 denotes a quality characteristic that is not at fault. For an abnormal variance vector structure, this study considers five types of variance shifts for demonstration, and they include $(1, 0, 0, 0, 0)$, $(1, 1, 0, 0, 0)$, $(1, 1, 1, 0, 0)$, $(1, 1, 1, 1, 0)$ and $(1, 1, 1, 1, 1)$. Also, this study considers the case of $\theta = 0.8$ and the sample size $n = 10$.

For the ANN and MARS modeling approaches, we have 5 input variables. They are the averaged sample for the five quality variables in a process, and they are denoted by $\overline{X_1}$,

$\overline{X_2}$, $\overline{X_3}$, $\overline{X_4}$ and $\overline{X_5}$, respectively. There is only one output node (Y) for all two models. This output node indicates the classification results of the types of process variance shifts, where a value of 0 implies that the process is the actual type of the underlying process variance fault, a value of 1 implies that the actual type of the process variance fault is not recognized. In this study, we assume that the underlying process variance fault is the type of $(1, 1, 1, 0, 0)$.

In this study, the training data sets include 1000 data vectors. Whereas the first 500 data vectors are all from the process variance fault of $(1, 1, 1, 0, 0)$, data vectors from 501 to 1000 are from other types of process variance faults. This study employs the testing data sets of 400. The first 200 data vectors are all from the process variance fault of $(1, 1, 1, 0, 0)$, and data vectors from 201 to 400 are from other types of process variance faults.

Table 1 displays the simulation results for the accurate identification rates (AIR) of the two approaches. Observing Table 1, it is apparently seen that the ANN approach outperforms the MARS approach. Also, the overall averaged AIR is 66.25% and 63.44% for ANN and MARS respectively.

Table 1. AIR for four combinations of process variance faults

Combination of variance faults	ANN ($\{n_i$-n_h-$n_o\}$)	MARS
$(1, 1, 1, 0, 0)$ vs. $(1, 0, 0, 0, 0)$	67.75% ({5-9-1})	65.50%
$(1, 1, 1, 0, 0)$ vs. $(1, 1, 0, 0, 0)$	64.25% ({5-11-1})	60.00%
$(1, 1, 1, 0, 0)$ vs. $(1, 1, 1, 1, 0)$	62.50% ({5-9-1})	63.25%
$(1, 1, 1, 0, 0)$ vs. $(1, 1, 1, 1, 1)$	70.50% ({5-9-1})	65.00%

5. **Conclusions.** In this paper, the ANN and MARS approaches are proposed to recognize the quality variables at fault when variance shifts have occurred in a multivariate normal process. The proposed ANN approach is superior to the MARS approach.

The proposed ANN and MARS models in this study are effective in recognizing the types of process variance faults. However, some other computational intelligence techniques, such as support vector machine, rough set or genetic algorithms, can be applied to further refine the structure of the classifiers. Extensions of the proposed procedures to data-driven design or real-time implementation of fault tolerant control system are also possible.

Acknowledgment. This research was supported in part by the Ministry of Science and Technology of Taiwan Grant Nos. MOST 103-2221-E-030-021 and MOST 103-2118-M-030-001.

REFERENCES

[1] R. L. Mason, N. D. Tracy and J. C. Young, Decomposition of T^2 for multivariate control chart interpretation, *Journal of Quality Technology*, vol.27, pp.99-105, 1995.

[2] G. C. Runger, F. B. Alt and D. C. Montgomery, Contributors to a multivariate statistical process control signal, *Communications in Statistics: Theory & Methods*, vol.25, pp.2203-2213, 1996.

[3] F. Aparisi, G. Avendaño and J. Sanz, Techniques to interpret T^2 control chart signals, *IIE Transactions*, vol.38, pp.647-657, 2006.

[4] Y. E. Shao, C.-D. Hou, C.-H. Chao and Y.-J. Chen, A decomposition approach for identifying the sources of variance shifts in a multivariate process, *ICIC Express Letters*, vol.5, no.4(A), pp.971-975, 2011.

[5] L. H. Chen and T. Y. Wang, Artificial neural networks to classify mean shifts from multivariate chart signals, *Computers and Industrial Engineering*, vol.47, pp.195-205, 2004.

[6] C. S. Cheng and H. P. Cheng, Identifying the source of variance shifts in the multivariate process using neural network and support vector machines, *Expert Systems with Applications*, vol.35, pp.198-206, 2008.

[7] Y. E. Shao and B.-S. Hsu, Determining the contributors for a multivariate SPC chart signal using artificial neural networks and support vector machine, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(B), pp.4899-4906, 2009.
[8] Y. E. Shao, H.-Y. Huang and Y.-J. Chen, Determining the sources of variance shifts in a multivariate process using flexible discriminant analysis, *ICIC Express Letters*, vol.4, no.5(A), pp.1573-1578, 2010.
[9] H.-Y. Huang, Y. E. Shao, C.-D. Hou and M.-D. Hsieh, Identifying the contributors of the multivariate variability control chart using hierarchical support vector machines, *ICIC Express Letters*, vol.5, no.9(B), pp.3543-3547, 2011.
[10] C. S. Cheng and H. P. Cheng, Using neural network to detect the bivariate process, *Computers & Industrial Engineering*, vol.60, pp.269-278, 2011.
[11] M. Y. Cheng and M. T. Cao, Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines, *Applied Soft Computing*, vol.22, pp.178-188, 2014.
[12] D. L. Martinez, D. T. Shih, V. C. P. Chen and S. B. Kim, A convex version of multivariate adaptive regression splines, *Computational Statistics & Data Analysis*, vol.81, pp.89-106, 2015.
[13] P. J. García Nieto, J. M. Torres, F. J. de Cos Juez and F. Sánchez Lasheras, Using multivariate adaptive regression splines and multilayer perceptron networks to evaluate paper manufactured using Eucalyptus globulus, *Applied Mathematics and Computation*, vol.219, pp.755-763, 2012.
[14] T. A. Choudhury, C. C. Berndt and Z. Man, Modular implementation of artificial neural network in predicting in-flight particle characteristics of an atmospheric plasma spray process, *Engineering Applications of Artificial Intelligence*, vol.45, pp.57-70, 2015.
[15] O. Makarynskyy, D. Makarynska, M. Rayson and S. Langtry, Combining deterministic modelling with artificial neural networks for suspended sediment estimates, *Applied Soft Computing*, vol.35, pp.247-256, 2015.
[16] J. Szoplik, Forecasting of natural gas consumption with artificial neural networks, *Energy*, vol.85, pp.208-220, 2015.
[17] F. B. Alt, Multivariate quality control, in *Encyclopedia of Statistical Sciences*, S. Kotz and N. L. Johnson (eds.), New York, John Wiley & Sons, Inc., 1985.