ISSN 2278-3091

Volume 8, No.1.3, 2019

International Journal of Advanced Trends in Computer Science and Engineering

Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse1781.32019.pdf

https://doi.org/10.30534/ijatcse/2019/1781.32019

Statistical Features-MLP Neural Network for Recognizing Bivariate SPC Chart Patterns

Ibrahim Masood[1], Mohamad Azrul Azhad Mohd Haizan[2], Ong Pauline[1], Mohd Helmy Abd Wahab[3]

[1]Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia, ibrahim@uthm.edu.my

[2]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia, ibrahim@uthm.edu.my

ABSTRACT

Statistical process control (SPC) charting is an important tool in monitoring, controlling and minimizing unnatural variation in manufacturing process. Nevertheless, its application has become more challenging when involves two correlated quality variables (bivariate) since the conventional multivariate SPC charting scheme is only efficient for triggering unnatural variation but unable to interpret the source of disturbance. Therefore, SPC pattern reStatistical process control (SPC) charting is an important tool in monitoring, controlling and minimizing unnatural variation in manufacturing process. Nevertheless, its application has become more challenging when involves two correlated quality variables (bivariate) since the conventional multivariate SPC charting scheme is only efficient for triggering unnatural variation but unable to interpret the source of disturbance. Therefore, SPC pattern recognition technique has been proposed for solving this issue. In this study, a new SPC pattern recognition scheme is proposed for monitoring and diagnosing nine category of unnatural variation in bivariate cases. Based on single-stage monitoring and diagnosis approach, a generalized multilayer's perceptron model with statistical features input representation was applied as the pattern recognizer. The proposed scheme has provided rapid triggering capability and high accuracy in interpreting the sources of unnatural variation, especially in dealing with moderate and large magnitude of unnatural variation. The selection of the statistical features as the input representation is very important in determining the effectiveness of the recognizer. This study has opened a new perspective in quality controlcognition technique has been proposed for solving this issue. In this study, a new SPC pattern recognition scheme is proposed for monitoring and diagnosing nine category of unnatural variation in bivariate cases. Based on single-stage monitoring and diagnosis approach, a generalized multilayer's perceptron model with statistical features input representation was applied as the pattern recognizer. The proposed scheme has provided rapid triggering capability and high accuracy in interpreting the sources of unnatural variation, especially in dealing with moderate and large magnitude of unnatural variation. The selection of the statistical features as the input representation is very important in determining the effectiveness of the recognizer. This study has opened a new perspective in quality control.

Key words : Bivariate SPC, Statistical Features, Multilayer's Perceptron, SPC Pattern Recognition.

1. INTRODUCTION

Unnatural variation in manufacturing process has become a dominant source of poor quality production. This has motivated the manufacturers to invest more on quality control in order to prevent poor quality products from reaching the market. In spite of the irrefutable accomplishments of production focused quality strategies such as Total Quality Management (TQM), Lean Manufacturing or Six Sigma widely applied since the early 1980s, many manufacturing companies still experience variety of unnatural variation in manufacturing [1].

The statistical process control (SPC) has been used as one of the important tools of TQM, i.e. to monitor, control and minimize unnatural variation. Interpretation of SPC chart pattern of monitored measurement is the most widely technique adopted in identifying the sources of unnatural variation [2]. The specific patterns can be used in deciding whether a manufacturing process is running in its intended procedure or it is not functioning correctly and required adjustment [3].

The cases has become more challenging when involving multi correlated variable (bivariate or multivariate), whereby the conventional multivariate SPC charting scheme was unable to provide information in relation to the source of unnatural variation. In recent years, this subject has been addressed through the evolution of artificial intelligence-based pattern recognition especially using an artificial neural network. The early reported studies focused on limited category of bivariate SPC chart patterns such as normal at both correlated quality variables, shift at 1st quality variable only, shift at 2nd quality variable only, and shifts at both correlated quality variables [4] – [5]. Since the higher possible category of variation is barely reported[6], this paper aims to report a study on nine category of variation that can be identified successfully using pattern recognition technique. Based on sudden shift patterns, this number of category covers overall possible source of variation in the bivariate cases.

2. METDODOLOGY

In this study, a new pattern recognition scheme was proposed based on single stage monitoring-diagnosis approach, a generalized multilayer's perceptron (MLP) neural network model, and statistical features (SF) input representation as shown in Figure 1.

Process monitoring aims to identify the manufacturing process condition either it is running within a statistically in-control or out-of-control. Process diagnosis aims to interpret the source of unnatural variation

2.1 Bivariate SPC chart pattern

In monitoring and diagnosing the source of unnatural variation, there are 9 possible patterns could be considered from the bivariate SPC charting, i.e., normal at both correlated quality variables (N00), upward shift at 1st quality variable only (US10), upward shift at 2nd quality variable only (US01), upward shifts at both correlated quality variables (US11), downward shift at 1st quality variable only (DS10), downward shift at 2nd quality variable only (DS01), downward shifts in both correlated quality variables (DS11), upward shift at 1st quality variable and downward shift at 2nd quality variable (USDS), and lastly downward shift at 1st quality variable and upward shift at 2nd quality variable (DSUS).

Each pattern is represented by 24 standardized observation samples (*Z*) of both correlated quality variables (*V1*, *V2*). The length of observation samples (24) is also called as window size (ws). The 1st quality variable is represented by $Z_{1\_V1}, \dots, Z_{24\_V1}$, whereas 2nd quality variable is represented by $Z_{1\_V2}, \dots, Z_{24\_V2}$.

The unnatural variation is focused on mean shifts within ± 3 standard deviations (σ) based on control limits of *Shewhart* control chart. Each pattern has different correlation (ρ), i.e. US10, US01, DS10 and DS01 have weak correlation (ρ = 0.0 ~ 0.3), US11 and DS11 have strong correlation with positive effect (ρ = 0.7 ~ 0.9), whereas USDS and DSUS have strong correlation with negative effect (ρ = –0.9 ~ –0.7).

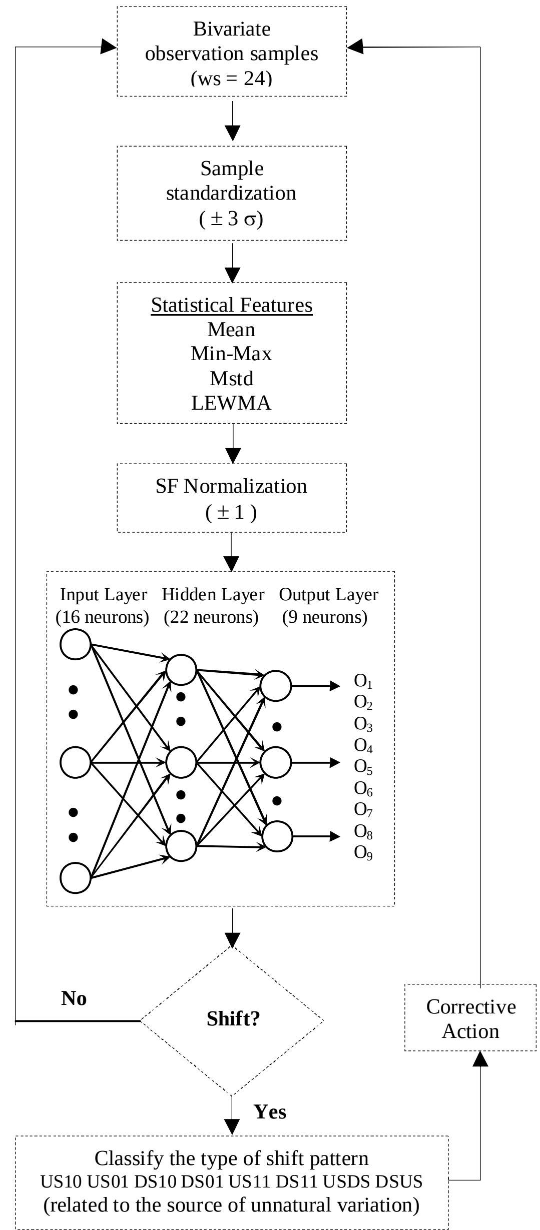

Figure 1: SF-MLP Pattern Recognition Scheme

2.2 Pattern recognizer

The generalized MLP neural network recognizer consists of three layers, i.e. input, hidden and output layers. Each layer can be represented by several numbers of neuron that perform the task of converting input data into desirable output[7]. The number of neuron in its respective layers were determined based on size of input representation, empirical simulation, and number of pattern category.

2.3 Input representation

Input representation is a method to symbolize input signal into the MLP neural network for achieving the impressive result in recognition [7]-[8]. In order to achieve a high recognition performance, selection of an effective input representation is a challenging effort. In this study, the SF input representation was investigated to improve recognition accuracy through reduction in training time, computational efforts, and size of MLP structure.

Initially, 9 SF candidates were considered based on its discrimination capability assumption. In features extraction process, the standardized samples were transformed to the summary statistic, i.e. mean, minimum and maximum (Min-Max), multiplication of mean and standard deviation (Mstd), mean square value (Msv), multiplication of mean and mean square value (Mmsv), slope, skewness, percentile, and last value of exponentially weighted moving average (LEWMA).
Each SF candidate could have possibility either to improve or to deteriorate the recognition performance. In selecting an effective set of SF, the design of experiment (DOE) was used since it can provide a systematic analysis using an optimum number of experiments.

The DOE results in training the SF-MLP recognizer are summarized in Table 1. Out of eight SF candidates, only three SF, i.e. Min-Max, Mstd and LEWMA were selected since it gave the sign of improvement. Noted that mean is an additional priority SF, which was considered without DOE analysis. Based on this selected SF input representation, an impressive recognition performance can be observed in the final stage of recognizer training.

Table 1: Features Selection Based on DOE Analysis

Input Features	Low Level (-1)	High Level (1)	Selected Features
MinMax	81.36	82.31	✓
Mstd	80.69	82.98	✓
Msv	83.80	79.87	✗
Mmsv	85.05	78.62	✗
Slope	85.01	78.66	✗
Skewness	84.43	79.24	✗
Percentile	83.03	80.64	✗
LEWMA	81.95	81.72	✓

Note: Mean is determined as the priority SF features

Table 2 shows the comparison between raw data-MLP and SF-MLP recognizers. The raw data-MLP model with 48 x 26 x 9 structure refers to the size of an input layer with 48 neurons (raw data input representation), a hidden layer with 26 neurons (determined based on empirical simulation), and an output layer with 9 neurons (9 pattern categories). The raw data input representation consists of a series of standardized samples, i.e. Z_{1_V1}, Z_{1_V2}, …, Z_{24_V1}, Z_{24_V2}.

The SF-MLP model with 16 x 22 x 9 structure refers to the smaller size of input layer (SF input representation), a hidden layer with 22 neurons (empirically simulated), and an output layer with 9 neurons. The SF input representation consists of a series of selected features, i.e. Mean$_{_V1}$, Mean$_{_V2}$, Min$_{_V1}$, Min$_{_V2}$, Max$_{_V1}$, Max$_{_V2}$, Mstd$_{_V1}$, Mstd$_{_V2}$, LEWMA$_{_V1}$, LEWMA$_{_V2}$. Smoothing constant 0.1, 0.15, 0.2, and 0.25 are applied to each LEWMA statistic.

Table 2: MLP Structure

Neural Network Layer	**Raw data –MLP**	**SF-MLP**
Input layer	48	16
Hidden layer	26	22
Output layer	9	9

2.4 Training and testing

Back propagation algorithm was applied in training the MLP recognizer. At this stage, static SPC chart pattern were applied. Input representation of each bivariate SPC pattern was normalized to a range between [-1, 1] before it can be processed into the recognizer. Overall patterns were divided to training (60%), validation (20%), and pre-testing (20%) sets. The recognition target was set at 95% accuracy for normal and shifts patterns.

In actual testing, application of dynamic SPC chart pattern can show a deteriorated quality variables (V_1, V_2) as happen in actual case. Each variable begins from in-control condition. Once there is a shift, the variable is deteriorating from partially developed shift to become a fully developed shift. At this condition, the recognition process can be made using dynamic window approach to compute how fast the unnatural variation could be triggered and how accurate the source of unnatural variation could be classified

3. RESULTS AND DISCUSSION

The average run length (ARL_1) and recognition accuracy percentage (RA) results as summarized in Table 3 to Table 5 can be used to evaluate the monitoring and diagnosis performances respectively. Each value of these performance measures can be computed based on 3000 testing patterns at its specific shift and correlation magnitudes.

An impressive performance monitoring can be indicated by a shorter ARL_1, which is fast in detecting an unnatural variation. On the other hand, an impressive diagnosis can be indicated by a higher RA, which is accurate in classifying the source of unnatural variation. This is to ensure that the proposed SF-MLP pattern recognition scheme would be able to identify the unnatural variation rapidly and correctly towards avoiding erroneous decision from SPC tool. Thus, the quality control philosophy 'making it right for the first time' can be realized.
Table 3 shows the performance results in recognizing small magnitude of unnatural variation, i.e. shift = 1.0 standard deviation. It can be observed that the bivariate SPC chart pattern with shift at one quality variable only (US10, US01, DS10, DS01) would be more difficult to be recognized compared to shift at both quality variables (US11, DS11,

USDS, DSUS). This is strongly related to the capability of unnatural variation properties as contributed by SF input representation to distinguish between normal and shifts patterns. The proposed scheme required a longer run (ARL_1 = 8.37 ~ 9.43) and provided lower accuracy (RA = 83.3 ~ 92.2 %) to recognize patterns with shift at one quality variable only compared to shift at both quality variables (ARL_1 = 6.14 ~ 6.45, RA = 93.8 ~ 97.3 %).

Table 3: SF-MLP Performance (Shift = 1σ)

Pattern Category	Shift Magnitude		ARL_1	RA
	X_1	X_2		
ρ = 0.2, –0.2				
US10	1.00	0.00	9.11, 9.43	90.3, 83.5
US01	0.00	1.00	9.30, 9.18	83.3, 87.1
DS10	–1.00	0.00	8.47, 8.37	91.4, 92.2
DS01	0.00	–1.00	8.56, 8.66	87.7, 86.3
ρ = 0.7				
US11	1.00	1.00	6.38	97.3
DS11	–1.00	–1.00	6.14	95.4
ρ = –0.7				
USDS	1.00	–1.00	6.36	96.6
DSUS	–1.00	1.00	6.45	93.8

Table 4 shows the performance results in recognizing moderate magnitude of unnatural variation, i.e. shift = 2.0 standard deviations. The patterns with shift at one quality variable only have become easier to be recognized but still less accurate compared to shift at both quality variables. At this condition, the proposed scheme required a shorter run (ARL_1 = 3.49 ~ 3.96) and gave better accuracy (RA = 90.2 ~ 93.9 %) to recognize patterns with shift at one quality variable only. This trend can also be found when dealing with shift at both quality variables (ARL_1 = 2.80 ~ 2.87, RA = 95.9 ~ 97.4 %). This observation supports the conclusion that the SF input representation shows a stronger discrimination capability in line with increment in magnitude of shift.

Table 4: SF-MLP Performance (shift = 2σ)

Pattern Category	Shift Magnitude		ARL_1	RA
	X_1	X_2		
ρ = 0.2, –0.2				
US10	2.00	0.00	3.96, 3.90	93.7, 93.1
US01	0.00	2.00	3.75, 3.77	90.2, 90.8
DS10	-2.00	0.00	3.56, 3.53	93.7, 93.6
DS01	0.00	-2.00	3.55, 3.49	92.5, 93.9
ρ = 0.7				
US11	2.00	2.00	2.87	97.4
DS11	–2.00	–2.00	2.80	96.1
ρ = –0.7				
USDS	2.00	–2.00	2.80	96.8
DSUS	–2.00	2.00	2.86	95.9

Table 5 shows the performance results in recognizing large magnitude of unnatural variation, i.e. shift = 3.0 standard deviations. Again, it is found that recognition for patterns with shift at one quality variable only have become easier, which is merely comparable with the shift at both quality variables. At this condition, the proposed scheme produced the shortest run (ARL_1 = 2.12 ~ 2.41) and the highest accuracy (RA = 93.5 ~ 95.1 %) to recognize patterns with shift at one quality variable only as compared to the previous condition. This trend can also be found when dealing with shift at both quality variables (ARL_1 = 1.72 ~ 1.86, RA = 96.7 ~ 97.7 %). It supports the assumption that, at large magnitude of unnatural variation, the SF input representation has a sufficient discrimination capability to classify between normal and shift.

Table 5: SF-MLP Performance (shift = 3σ)

Pattern Category	Shift Magnitude		ARL_1	RA
	X_1	X_2		
ρ = 0.2, –0.2				
US10	3.00	0.00	2.41, 2.41	95.1, 94.4
US01	0.00	3.00	2.31, 2.32	93.6, 93.5
DS10	–3.00	0.00	2.16, 2.17	94.7, 95.3
DS01	0.00	–3.00	2.13, 2.12	94.6, 94.9
ρ = 0.7				
US11	3.00	3.00	1.86	97.7
DS11	–3.00	–3.00	1.72	97.3
ρ = -0.7				
USDS	3.00	–3.00	1.80	96.7
DSUS	–3.00	3.00	1.80	97.2

4. CONCLUSION

In quality control of in-progress manufacturing process, proper selection of SPC charting scheme is important in monitoring and diagnosing the source of unnatural variation. When dealing with process disturbance, an industrial practitioner requires a right information to perform corrective action. In this study, a new SPC pattern recognition scheme has been designed to identify nine category of variation in bivariate cases. The generalized MLP neural network with SF input representation was applied as the pattern recognizer.

The overall recognition performance of the proposed SF-MLP scheme is influenced by the discrimination capability of SF properties to distinguish between normal and shift patterns. The strength is the SF properties can be obtained when integrated with a proper design of pattern recognizer model and its training algorithm. In order to achieve the highest recognition performance, the design and selection of SF input representation is challenging. An effective SF set can be obtained using a systematically and an economically analysis, i.e. DOE. It can be concluded that the performance monitoring and diagnosis have become faster and more accurate when the magnitude of unnatural variation increased. At small magnitude of unnatural variation, recognition of patterns with shift at one quality variable only is more difficult compared to shifts at both quality variables.

ACKNOWLEDGEMENT

The authors wish a highly gratitude to Universiti Tun Hussein Onn Malaysia (UTHM) for funding this research (Contract-Industry Grant Scheme, U524).

REFERENCES

1. Eifler T, Murthy BS & Howard TJ (2016), **Toward meaningful manufacturing variation data in design - feature based description of variation in manufacturing processes**. *Procedia CIRP* 43, 190−195. https://doi.org/10.1016/j.procir.2016.02.161
2. Yu J, Xi L & Zhou X (2009), **Identifying source(s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble**. *Engineering Applications of Artificial Intelligence* 22(1), 141−152. Ebrahimzadeh, V. Ranaee. **Control chart pattern recognition using an optimized neural network and efficient features**. *ISA Trans*, 49 (2010) 387−393. https://doi.org/10.1016/j.isatra.2010.03.007
3. Zorriassatine F, Tannock JDT & O'Brien C (2003), **Using novelty detection to identify abnormalities caused by mean shifts in bivariate processes**. *Computers and Industrial Engineering* 44(3), 385−408. https://doi.org/10.1016/S0360-8352(02)00215-2
4. Niaki STA & Abbasi B (2005), **Fault diagnosis in multivariate control charts using artificial neural networks**. *Quality and Reliability Engineering International* 21(8), 825−840. https://doi.org/10.1002/qre.689
5. Guh RS (2007), **On-line identification and quantification of mean shifts in bivariate processes using a neural network-based approach**. *Quality and Reliability Engineering International* 23, 367–386. https://doi.org/10.1002/qre.796
6. Kumar S & Kaur T (2016), **Development of ANN based model for solar potential assessment using various meteorological parameters**. *Energy Procedia* 90, 587−592. https://doi.org/10.1016/j.egypro.2016.11.227
7. Masood I and Hassan A (2014), **Bivariate quality control using two-stage intelligent monitoring scheme**. *Expert Systems with Applications* 41, 7579–7595. https://doi.org/10.1016/j.eswa.2014.05.042