VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DIAGNOSIS OF BIVARIATE PROCESS VARIATION USING AN INTEGRATED MSPC-ANN SCHEME

Ibrahim Masood, Rasheed Majeed Ali, Nurul Adlihisam Mohd Solihin and Adel Muhsin Elewe
Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia
E-Mail: ibrahim@uthm.edu.my

ABSTRACT

Monitoring and diagnosis of mean shifts in manufacturing processes become more challenging when involving two or more correlated variables. Unfortunately, most of the existing multivariate statistical process control schemes are only effective in rapid detection but suffer high false alarm. This is referred to as imbalanced performance monitoring. The problem becomes more complicated when dealing with small mean shift particularly in identifying the causable variables. In this research, a scheme that integrated the control charting and pattern recognition technique has been investigated toward improving the quality control (QC) performance. Design considerations involved extensive simulation experiments to select input representation based on raw data and statistical features, recognizer design structure based on individual and Statistical Features-ANN models, and monitoring-diagnosis approach based on single stage and two stages techniques. The study focuses on correlated process mean shifts for cross correlation function, ρ = 0.1, 0.5, 0.9, and mean shift, μ = ± 0.75 ~ 3.00 standard deviations. Among the investigated design, an Integrated Multivariate Exponentially Weighted Moving Average with Artificial Neural Network scheme provides superior performance, namely the Average Run Length for grand average ARL1 = 7.55~7.78 (for out-of-control) and ARL0 = 491.03 (small shifts) and 524.80 (large shifts) in control process and the grand average for recognition accuracy (RA) = 96.36 ~ 98.74. This research has provided a new perspective in realizing balanced monitoring and accurate diagnosis of correlated process mean shifts.

Keywords: bivariate process, statistical process control, artificial neural network.

INTRODUCTION

In the manufacturing industry, the process of extraordinary change has become a major source of low quality products (Masood and Hassan, 2010). When the manufacturing process involves two or more correlated variables, the corresponding scheme is necessary to monitor the variables together. In addressing this issue, multivariate statistical process control (MSPC), traditional frameworks such as T^2, the cumulative number of multivariate (CUSUM) and multivariate rapid weighted moving average (MEWMA) known to be effective in detecting mean process changes. However, they lack the ability to identify the sources of the variables responsible for this process mean shifts. In other words, it cannot provide diagnostic information for quality practitioners toward finding the root cause of the error and the solution for corrective action. Therefore, major studies have focused on improving the ability to identify the source variable in reference (Bersimis *et al.*, 2007). In a related study, various artificial neural network (ANN) based framework that SPC-ANN in reference (Chen LH, Wang TY, 2004). In terms of monitoring, SPC developed this framework has been declared a faster switch detection. However, most of them are suffering from the report that one of the high (average run length, ARL0 ≤ 200) compared with the level of de facto framework for univariate SPC (ARL0 ≥ 370). It will be important for the practitioner qualified in carrying solution for a problem that does not need any introduction in process control as out of control. In this study, the condition is called 'monitoring unbalanced and lack of diagnosis'. In terms of diagnosis, on the other hand, they are less accurate in identifying the source (causable) variables, especially when dealing with small changes. It will be more difficult for a practitioner of quality in finding the root cause of the error. To overcome this drawback, an enhanced framework of integrated Statistical Features -Ann been developed to achieve the 'appropriate monitoring, balanced and diagnosis. The proposed framework is intended to allow rapid displacement detection with minimum false alarms and high accuracy in identifying the source of the switch variables. Details of the discussions organized as follows. Section 2 presents a framework enhanced. Section 3 then gives a performance comparison between an integrated MSPC-ANN and Baseline-ANN schemes. Finally, section 4 outlines some conclusions.

M ETHODOLOGY MSPC-ANN SCHEME

An integrated MSPC-ANN scheme was designed and developed based on two stages monitoring and diagnosis technique as shown in Figure-1.

Monitoring process refers to the identification process status whether it is running in a statistical or circumstances beyond the control of the control, while referring to the identification of diagnosis process outsourcing process control variables in the mean change. In the first stage of monitoring, MEWMA control charts used to trigger changes in the multivariate mean by 'one point out of control' procedure. Once the transition is detected, Statistical Features-ANN model is then used to carry out the second stage of monitoring and diagnosis by identifying the data flow pattern contained point (s) out of control. Procedures for the implementation of the proposed framework is summarized in Figure 2.

ARPN Journal of Engineering and Applied Sciences

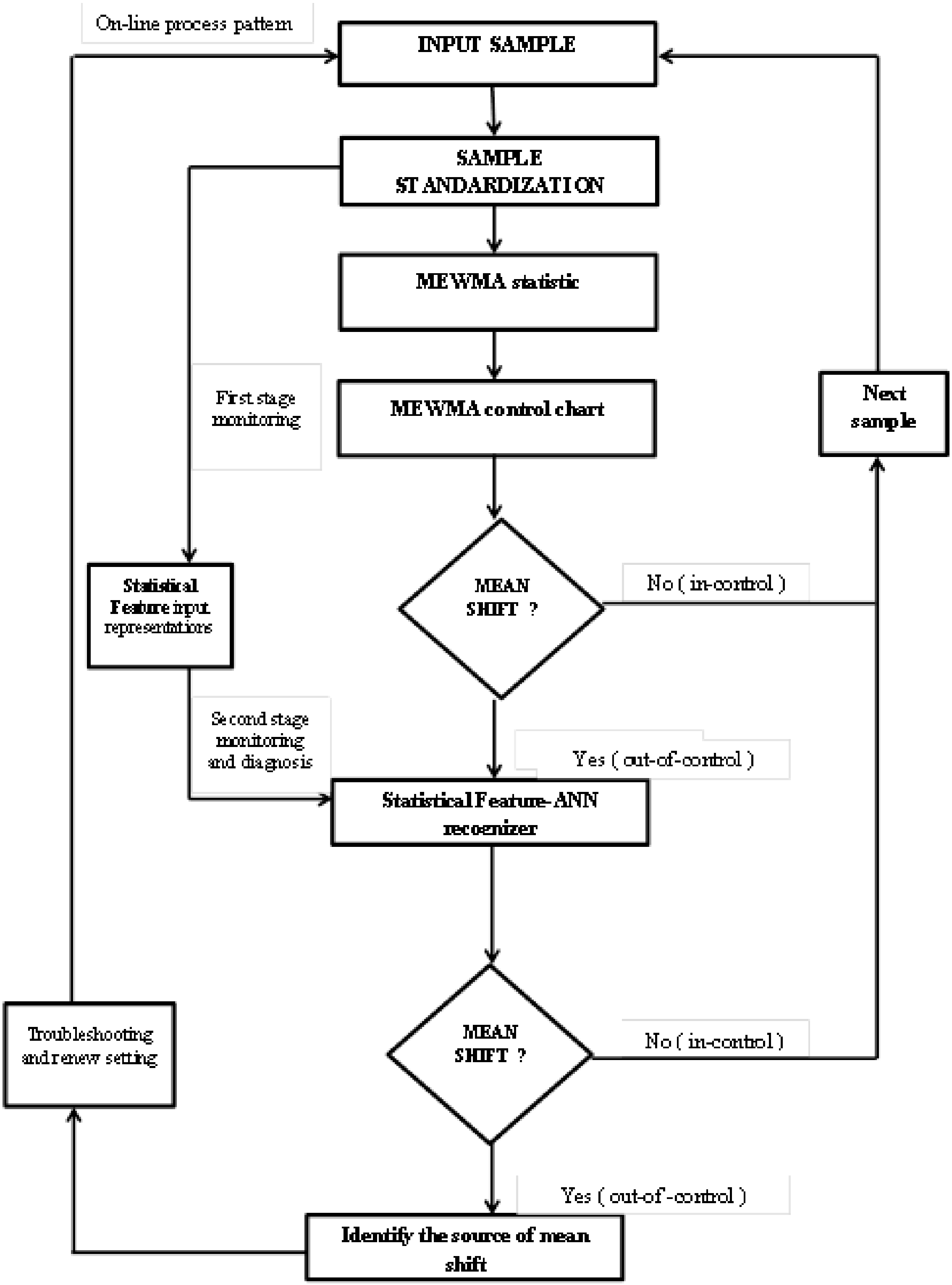

Figure-1. An integrated MSPC-ANN framework.

ARPN Journal of Engineering and Applied Sciences

Step 1: Input samples (bivariate). Window size = 24, starting observation samples are:
$X_{1\text{-}i} = (X_{1\text{-}1},\ldots X_{1\text{-}24})$ and $X_{2\text{-}i} = (X_{2\text{-}1},\ldots, X_{2\text{-}24})$. It is followed by (i + 1), (i + 2) and so on.

Step 2: Samples standardization. Rescale observation samples into a standardized range within [−3, +3]: $Z_1 = (X_1 - \mu_{01})/\sigma_{01}$ and $Z_2 = (X_2 - \mu_{02})/\sigma_{02}$
Input samples (original) and standardized samples can be represented graphically using Shewhart control charts and scatter diagram.

Step 3: Input representation (statistical features):

- Feature extraction: the standardized samples (Z_1, Z_2) are converted into statistical features, namely, the last value of exponentially weighted moving average ($LEWMA_\lambda$) with λ = (0.25, 0.20, 0.15, 0.10), mean (μ), multiplication of mean with standard deviation (MSD), and multiplication of mean with mean square value (MMSV).

- The extracted features are represented as follows: ($LEWMA_{0.25\_P1}$, $LEWMA_{0.20\_P1}$, $LEWMA_{0.15\_P1}$, $LEWMA_{0.10\_P1}$, μ_{P1}, MSD_{P1}, $MMSV_{P1}$, $LEWMA_{0.25\_P2}$, $LEWMA_{0.20\_P2}$, $LEWMA_{0.15\_P2}$, $LEWMA_{0.10\_P2}$, μ_{P2}, MSD_{P2}, $MMSV_{P2}$). Number of samples is denoted by P.

Step 4: Pattern Recognition (statistical features-ANN recognizer):
Decision rule: *If* Maximum output of ANN belongs to N(0,0) pattern:
Process is "in-control"; proceed to the next samples

else
Process is "out-of-control"; identify the sources of mean shifts; perform diagnosis, troubleshooting; renew setting and return to Step 1 (out of scopes of this research).

end

Figure-2. Pseudo code for an integrated MSPC-ANN scheme.

VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Performance of an integrated MSPC-ANN scheme.

STATISTICAL FEATURES-ANN SCHEME

Pattern category	Mean shifts		Average run lengths				Recognition accuracy			
			MEWMA λ = 0.10 H = 8.64							
	X1	X2	ARL_0 for ρ =	0.1	0.5	0.9	RA for ρ	0.1	0.5	0.9
N (0,0)	0.00	0.00		491.03	426.70	524.80		NA	NA	NA
			ARL_1 for ρ =	0.1	0.5	0.9		0.1	0.5	0.9
US (1,0)	0.75	0.00		16.89	17.77	19.50		95.7	93.2	94.6
US (0,1)	0.00	0.75		18.73	19.58	21.43		95	93.9	93.5
US (1,1)	0.75	0.75		15.19	14.56	15.10		81.5	95.5	100
DS (1,0)	-0.75	0.00		16.29	16.74	17.03		97.6	97.2	97.4
DS (0,1)	0.00	-0.75		17.64	18.27	19.88		96.5	95.1	96.4
DS (1,1)	-0.75	-0.75		15.88	14.49	14.41		88.2	98.4	99.7
Average				16.77	16.90	17.89		92.42	95.55	96.93
US (1,0)	1.00	0.00		10.90	11.36	11.84		96.7	96.2	97.2
US (0,1)	0.00	1.00		12.34	12.50	13.35		96.5	95.5	95.7
US (1,1)	1.00	1.00		10.60	10.13	10.34		87.7	96.1	100
DS (1,0)	-1.00	0.00		10.58	10.58	11.00		97.8	98.5	97.5
DS (0,1)	0.00	-1.00		11.59	11.83	12.18		95.4	97	97.7
DS (1,1)	-1.00	-1.00		10.55	9.90	10.23		93.7	99.1	100
Average				11.09	11.05	11.49		94.63	97.07	98.02
US (1,0)	1.50	0.00		6.35	6.40	6.47		98.3	98.1	97.8
US (0,1)	0.00	1.50		7.29	7.44	7.50		97.4	96.3	97.1
US (1,1)	1.50	1.50		6.20	6.03	6.04		90.2	98.2	100
DS (1,0)	-1.50	0.00		6.30	6.28	6.19		98.2	98	98.8
DS (0,1)	0.00	-1.50		6.52	6.70	6.71		97.6	97.7	97.9
DS (1,1)	-1.50	-1.50		6.08	5.95	5.99		96.9	99.1	99.9
Average				6.45	6.47	6.48		96.43	97.90	98.58
US (1,0)	2.00	0.00		4.50	4.38	4.44		99.4	99.8	99.9
US (0,1)	0.00	2.00		4.98	5.02	5.04		99.6	99.2	99.4
US (1,1)	2.00	2.00		4.43	4.35	4.45		92.1	98.3	100
DS (1,0)	-2.00	0.00		4.35	4.28	4.32		99.9	99.7	99.8
DS (0,1)	0.00	-2.00		4.65	4.65	4.49		99.3	99.5	97.9
DS (1,1)	-2.00	-2.00		4.26	4.14	4.22		97.9	99.2	99.9
Average				4.53	4.47	4.49		98.03	99.28	99.48
US (1,0)	2.50	0.00		3.51	3.39	3.35		99.8	99.8	99.9
US (0,1)	0.00	2.50		3.86	3.90	3.82		99.6	99.4	99.4
US (1,1)	2.50	2.50		3.41	3.38	3.53		92.9	98.3	100
DS (1,0)	-2.50	0.00		3.48	3.27	3.29		99.6	99.6	99.6
DS (0,1)	0.00	-2.50		3.56	3.45	3.51		99.4	99.4	99.7
DS (1,1)	-2.50	-2.50		3.31	3.24	3.41		98.1	99.7	100
Average				3.52	3.44	3.49		98.23	99.37	99.77
US (1,0)	3.00	0.00		2.89	2.75	2.73		100	99.8	99.8
US (0,1)	0.00	3.00		3.18	3.12	3.11		99.8	99.3	99.5
US (1,1)	3.00	3.00		2.84	2.79	2.90		93.2	98.5	100
DS (1,0)	-3.00	0.00		2.89	2.75	2.70		99.7	99.8	99.5
DS (0,1)	0.00	-3.00		2.94	2.90	2.89		99.4	99.6	99.3
DS (1,1)	-3.00	-3.00		2.75	2.72	2.83		98.3	99.8	99.9
Average				2.91	2.84	2.86		98.40	99.47	99.67
Grand average ± (0.75 - 3.00)				7.55	7.53	7.78		96.36	98.11	98.74

Note: Design parameters for MEWMA control chart (λ = 0.1, H = 8.64)

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-2. Performance of a baseline-ANN scheme.

BASELINE SCHEME

Pattern Category	Mean Shifts		Average Run Lengths				Recognition Accuracy			
	X1	X2	ARL_0 for ρ =	0.1	0.5	0.9	RA for ρ	0.1	0.5	0.9
N (0,0)	0.00	0.00		163.83	318.20	249.49		NA	NA	NA
			ARL_1 for ρ =	0.1	0.5	0.9		0.1	0.5	0.9
US (1,0)	0.75	0.00		16.32	16.70	17.58		88.4	87.2	87.2
US (0,1)	0.00	0.75		14.04	14.36	14.73		89.1	88.4	89.2
US (1,1)	0.75	0.75		13.42	12.80	12.78		79.9	93.4	99.7
DS (1,0)	-0.75	0.00		13.63	14.44	14.83		89.9	89.1	88.9
DS (0,1)	0.00	-0.75		15.37	16.10	16.64		88.6	88.2	89.7
DS (1,1)	-0.75	-0.75		13.46	13.00	12.82		77	92.3	99.5
Average				14.37	14.57	14.90		85.5	89.8	92.4
US (1,0)	1.00	0.00		10.38	11.04	11.11		92.5	92.1	91.2
US (0,1)	0.00	1.00		9.54	9.71	9.78		91.5	92.9	91.2
US (1,1)	1.00	1.00		9.38	9.23	9.07		85.5	96.1	99.9
DS (1,0)	-1.00	0.00		9.71	9.95	10.07		93.5	92.4	93.4
DS (0,1)	0.00	-1.00		10.59	10.83	10.70		91.9	91.8	91.9
DS (1,1)	-1.00	-1.00		9.66	9.46	9.35		81.9	94.5	99.8
Average				9.88	10.04	10.01		89.5	93.3	94.6
US (1,0)	1.50	0.00		7.17	7.16	7.17		96	95.9	95.3
US (0,1)	0.00	1.50		6.37	6.42	6.44		95.2	95.4	95.1
US (1,1)	1.50	1.50		6.47	6.32	6.30		89.7	97.2	99.9
DS (1,0)	-1.50	0.00		6.58	6.59	6.65		95.8	96.3	96.6
DS (0,1)	0.00	-1.50		7.09	7.09	7.01		94.7	95.4	95.6
DS (1,1)	-1.50	-1.50		6.57	6.49	6.34		87.2	97.1	99.9
Average				6.71	6.68	6.65		93.1	96.2	97.1
US (1,0)	2.00	0.00		5.55	5.51	5.55		96.1	95.9	96.3
US (0,1)	0.00	2.00		5.01	4.99	5.00		95.8	95.7	95.3
US (1,1)	2.00	2.00		5.04	4.97	4.89		90.7	97.1	99.9
DS (1,0)	-2.00	0.00		4.98	5.01	5.04		97.5	97.1	96.5
DS (0,1)	0.00	-2.00		5.33	5.42	5.42		95.9	96.2	95.6
DS (1,1)	-2.00	-2.00		5.18	5.06	5.03		87.9	96.7	99.9
Average				5.18	5.16	5.16		94.0	96.5	97.3
US (1,0)	2.50	0.00		4.52	4.56	4.57		96.8	96.3	97.1
US (0,1)	0.00	2.50		4.15	4.16	4.14		95.7	95.1	95
US (1,1)	2.50	2.50		4.2	4.15	4.12		91.2	97.5	99.9
DS (1,0)	-2.50	0.00		4.04	4.02	4.02		97.8	97.7	96.8
DS (0,1)	0.00	-2.50		4.36	4.38	4.42		96.7	96.4	96.3
DS (1,1)	-2.50	-2.50		4.24	4.22	4.09		88.3	96.3	99.9
Average				4.25	4.25	4.23		94.4	96.6	97.5
US (1,0)	3.00	0.00		3.9	3.91	3.89		97.7	97.5	97.3
US (0,1)	0.00	3.00		3.57	3.59	3.58		95.6	95.5	95.8
US (1,1)	3.00	3.00		3.63	3.60	3.56		91.1	97.7	99.9
DS (1,0)	-3.00	0.00		3.41	3.37	3.42		98	98.3	97.4
DS (0,1)	0.00	-3.00		3.72	3.71	3.74		96.6	96.7	97.2
DS (1,1)	-3.00	-3.00		3.69	3.60	3.56		88.2	96.8	99.8
Average				3.65	3.63	3.63		94.5	97.1	97.9
Grand average ± (0.75 - 3.00)				7.34	7.39	7.43		91.8	94.9	96.1

Table-3. Statistical significant test (small shifts).

Performance measures (PM)	**Result of Paired T-Test mean difference (MSPC-ANN minus Baseline-ANN)**	**Remarks**
ARL_0	N MEAN StDev SE Mean MSPC-ANN 3 480.8 49.8 28.8 Baseline 3 243.8 77.3 44.7 Difference 3 237.0 114.3 66.0 Mean difference of ARL_0: 95% CI: (-46.9, 520.9) T-Test = 0 (vs ≠ 0): T = 3.59, P = 0.070	Increment in ARL0 is Not statistically significant
ARL_1	N MEAN St Dev SE Mean MSPC-ANN 9 11.62 4.66 1.55 Baseline 9 10.42 3.45 1.15 Difference 9 1.198 1.226 0.409 Mean difference of ARL_0: 95% CI: (0.255, 2.140) T-Test = 0 (vs ≠ 0): T = 2.93, P = 0.019	Increment in ARL1 is closely significant
RA	N MEAN StDev SE Mean MSPC-ANN 9 96.39 1.94 0.65 Baseline 9 91.28 4.93 1.64 Difference 9 5.11 3.95 1.32 Mean difference of ARL_0: 95% CI: (2.08, 8.15) T-Test = 0 (vs ≠ 0): T = 3.88, P = 0.005	Increment RA is absolutely significant

Table-4. Statistical significant test (large shifts).

Performance measures (PM)	**Result of Paired T-Test Mean difference (MSPC-ANN minus Baseline-ANN)**	**Remarks**
ARL_0	N MEAN StDev SE Mean MSPC-ANN 3 480.8 49.8 28.8 Baseline 3 243.8 77.3 44.7 Difference 3 237.0 114.3 66.0 Mean difference of ARL_0: 95% CI: (-46.9, 520.9) T-Test = 0 (vs ≠ 0): T = 3.59, P = 0.070	Increment in ARL0 is Not statistically significant
ARL_1	N MEAN St Dev SE Mean MSPC-ANN 9 3.617 0.712 0.237 Baseline 9 4.349 0.667 0.222 Difference 9 -0.7322 0.0540 0.0180 Mean difference of ARL_0: 95% CI: (-0.7732, -0.6970) T-Test = 0 (vs ≠ 0): T = -40.66, P = 0.000	Decrement in ARL1 is absolutely significant
RA	N MEAN StDev SE Mean MSPC-ANN 9 98.544 1.596 0.532 Baseline 9 96.200 1.49 0.497 Difference 9 2.344 2.207 0.736 Mean difference of ARL_0: 95% CI: (0.648, 4.041) T-Test = 0 (vs ≠ 0): T = 3.19, P = 0.013	Increment in RA is closely significant

VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

It should be noted that the following initial setting needs to be performed before it can be put into application:

- Load the trained the raw data-ANN recognizer into the system.
- Set the values of means ($\mu_{01,}$ μ_{02}) and standard deviations ($\sigma_{01,}$ σ_{02}) of multivariate in-control process (for variables X_1 and X_2). These parameters can be obtained based on historical samples.
- Perform in-process quality control inspection until 24 observations to begin the system.

Modeling of data patterns of multivariate process mean shifts, and design and training-testing of the Statistical Features-ANN recognizer can be referred in reference (Masood and Hassan, 2012). The formulation of the MEWMA control chart can be found in reference (Lowry *et al.*,1992) Parameters (λ, H) = (0.10, 8.64) as reported in reference (Prabhu and Runger, 1997) were selected for this scheme.

RESULTS AND DISCUSSIONS

The monitoring and diagnosis performance of the Statistical Features-ANN scheme was compared with the Basic scheme as summarized in Table-1 and Table-2. In order to support the comparison of statistically, the statistical significant test was performed using ' Paired T-Test' as summarized in Table-3 and Table-4. Based on test of statistical significance (Paired T-test) is divided into two part, for small mean shift ($\leq$ 0.75 - 1.50) and large mean shift ($\geq$ 2.00 to 3.00) standard deviation. The test methods have been used since the decision (ARL0, ARL1, and RA) than is dependent upon the magnitude of the displacement mean shift (μ) and correlation of data (ρ). In term of monitoring, the Statistical Features-ANN scheme has caused more ARL0 for low and moderate correlation data (ρ/ ARL0 = 0.1/ 491.03, 0.5/ 426.70), and longer ARL0 for high data correlation (ρ = 0.9) compared to the Basic scheme (ρ/ ARL0 = 0.1/ 163.83, 0.5/ 318.2, 0.9/ 249.49). According to the chart below, shows the average run length (ARL0 and ARL1) and recognition accuracy (RA) for the Baseline scheme and Statistical Features-ANN scheme. For both schemes, showing an increase occurs at ARL0 when an increase in the correlation data.

For small shifts, MSPC-ANN scheme shows a better accuracy in diagnosis the source of variation (root cause error) compared to the Baseline-ANN scheme (MSPC-ANN = 96.39%; BS-ANN = 91.28%). However, this scheme required a slightly large ARL^1 in identifying out-of-control process condition compared to the Baseline scheme (MSPC = 11.62; BS-ANN = 10.42).

For large shifts, MSPC-ANN scheme gave shorter ARL^1 results compared to the Baseline-ANN scheme to identify out-of-control correlation (MSPC-ANN = 3.617; BS = 4.349). Meanwhile, the MSPC-ANN scheme is better than the Baseline-ANN scheme in diagnosing the source of variation (root cause error) (MSPC-ANN = 96.39 %; BS-ANN = 91.28%).

CONCLUSIONS

This research presented that an integrated MSPC-ANN scheme was efficient to achieve the 'balance monitoring and accurate diagnosis' performance in dealing with bivariate process variation in mean shifts. Based on two-stages monitoring and diagnosis approach, the proposed framework and pseudo-code (implementation procedures) would be efficient in terms of rapid detection of process variation and accurate identification of the sources of variation with minimum false alarms (ARL0 $\geq$ 370). The monitoring-diagnosis performances of the scheme are strongly dependent on input representation technique, recognizer design and training, and the procedure for recovering false alarm. In the future work, further investigation will be extended to other causable patterns such as trends and cyclic.

REFERENCES

Bersimis S, Psarakis S, Panaretos J. 2007. Multivariate statistical process control charts: an overview. Quality and Reliability Engineering International. 23: 517-543.

Chen LH, Wang TY. 2004. Artificial neural networks to classify mean shifts from multivariate $\chi 2$ chart signals. Computers and Industrial Engineering. 47: 195-205.

Lowry CA, Woodall WH, Champ CW, Rigdon SE. 1992. A multivariate exponentially weighted moving average control chart. Technometrics. 34(1): 46-53.

Masood I, Hassan A. 2010. Issues in development of artificial neural network-based control chart pattern recognition schemes. European Journal of Scientific Research. 39(3): 336-355.

Masood I, Hassan A. 2012. An Integrated MEWMA-ANN scheme towards balanced monitoring and accurate diagnosis of bivariate process mean shifts. Journal of King Saud University - Computer and Information Sciences. 24: 93-100.

Prabhu SS, Runger GC (1997) Designing a multivariate EWMA control chart. Journal of Quality Technology. 29(1): 8-15.