Journal of King Saud University – Computer and Information Sciences (2012) **24**, 93–100

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

An integrated MEWMA-ANN scheme towards balanced monitoring and accurate diagnosis of bivariate process mean shifts

Ibrahim Masood [a,*], Adnan Hassan [b]

[a] *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia*
[b] *Faculty of Mechanical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia*

Received 20 June 2011; accepted 12 October 2011
Available online 7 November 2011

KEYWORDS

Balanced monitoring;
Bivariate pattern recognition;
Statistical features;
Synergistic-ANN;
Two-stages monitoring and diagnosis

Abstract Various artificial neural networks (ANN)-based pattern recognition schemes have been developed for monitoring and diagnosis of bivariate process variation in mean shifts. In comparison with the traditional multivariate statistical process control (MSPC) charts, these advanced schemes generally perform better in identifying process mean shifts and provide more effective information towards diagnosing the root causes. However, it seemly less effective for multivariate quality control (MQC) application due to disadvantages in reference bivariate patterns and imbalanced monitoring performance. To achieve 'balanced monitoring and accurate diagnosis', this study proposes an integrated multivariate exponentially weighted moving average (MEWMA)-ANN scheme for two-stages monitoring and diagnosis of some reference bivariate patterns. Raw data and statistical features input representations were applied into training of the Synergistic-ANN recognizer for improving patterns discrimination capability. The proposed scheme has resulted in better monitoring – diagnosis performances with smaller false alarm, quick mean shift detection and higher diagnosis accuracy compared to the basic scheme.

1. Introduction

In manufacturing industries, process variation has become a major source of poor quality. Manufacturing process may involve two or more correlated variables and an appropriate procedure is required to monitor these variables simultaneously. This issue is sometimes called multivariate quality control (MQC). It has led to extensive research in the field of multivariate statistical process control (MSPC) towards monitoring and diagnosis of multivariate process variation in mean shifts/variances. Further discussions on this issue can be found in Lowry and Montgomery (1995), Kourti and MacGregor (1996), Mason et al. (1997) and Bersimis et al. (2007).

Development in soft computing technology have motivated researchers to explore the use of artificial intelligence techniques such as artificial neural networks (ANN), among others

* Corresponding author.
E-mail address: ibrahim@uthm.edu.my (I. Masood).

Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

1319-1578 © 2011 King Saud University. Production and hosting by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jksuci.2011.10.002

for automatically and intelligently recognizing patterns in relation to MSPC charting. Identification of these patterns coupled with engineering knowledge of the process would lead to more specific diagnosis and troubleshooting. Various ANN-based pattern recognition schemes have been proposed such as MSPC-ANN (Chen and Wang, 2004; Niaki and Abbasi, 2005; Cheng and Cheng, 2008; Yu et al., 2009), novelty detector (Zorriassatine et al., 2003), modular-ANN (Guh, 2007), ensemble-ANN (Yu and Xi, 2009) and multi-module-structure-ANN (El-Midany et al., 2010). The MSPC-ANN schemes combined the MSPC charts (for monitoring any mean/variance shifts in multivariate processes) with ANN recognizer (for diagnosing the source variable(s) that responsible for mean/variance shifts). The other schemes such as novelty detector, modular-ANN, ensemble-ANN and multi-module-structure-ANN were designed to perform continuous monitoring and diagnosis simultaneously. Further discussion on these schemes can be found in Masood and Hassan (2010).

In this study, these ANN-based schemes are referred as bivariate pattern recognition (BPR) since the investigations are mainly focused on two correlated variables. In comparison with the traditional MSPC charts, these schemes have shown faster detection of mean shifts and provided a more detail information of the source variable(s) towards effective diagnosis. Nevertheless, they revealed some disadvantages in terms of:

1.1. Reference bivariate patterns

In MQC, the joint effect (cross correlation) between two dependent variables should be taken into account. Monitoring-diagnostic using Shewhart control chart patterns may provides useful meaning about univariate process mean shifts but it would lead to a higher false alarm than assumed. On the other hand, monitoring-diagnosis using χ^2 control chart patterns would result in lack of diagnosis (unable to identify the source variables). Generally, there are limited works reported on modeling of bivariate correlated process and patterns.

1.2. Imbalanced monitoring

In monitoring aspect, the existing BPR schemes are generally effective to quickly detect mean shifts. Unfortunately, they are mainly limited to a short ARL_0 ($\approx$200), that is inadequate to reduce false alarm towards an original SPC level ($ARL_0 \approx 370$ based on Shewhart control chart). It is critical for a practitioner to conduct unnecessary troubleshooting due to wrong identification of stable process as unstable. In this study, this situation is called 'imbalanced monitoring'.

In order to overcome the above disadvantages, an integrated MEWMA-ANN scheme was developed towards 'balance monitoring and accurate diagnosis' for some reference bivariate patterns. The proposed scheme aims for a reduced false alarm, faster mean shift detection and a more accurate diagnosis. Details discussion is organized as follows. Section 2 presents an integrated MEWMA-ANN scheme. Section 3 then provides performance comparison between an integrated MEWMA-ANN scheme and the Basic scheme. Section 4 finally outlines some conclusions.

2. An integrated MEWMA-ANN scheme

An integrated MEWMA-ANN scheme was developed based on two-stages monitoring and diagnosis approach as shown in Fig. 1. Process monitoring refers to the identification of process status either in a statistically stable or unstable state, whereas process diagnosis refers to the identification of the source variable(s) of an unstable process. In the first stage monitoring, the MEWMA control chart is used for triggering mean shifts based on 'one point out-of-control'. Once the mean shift is triggered, the Synergistic-ANN recognizer is then used to perform second stage monitoring and diagnosis by recognizing data stream pattern contained point(s) out-of-control as truly unstable or not.

2.1. Modeling of bivariate process and patterns

Let $X_{1i} = (X_{1\text{-}1}, \ldots, X_{1\text{-}24})$ and $X_{2i} = (X_{2\text{-}1}, \ldots, X_{2\text{-}24})$ represent bivariate process data streams based on window size = 24. Observation windows for both variables start with samples ith = (1, ..., 24). It is followed by (ith + 1), (ith + 2) and so on.

In a statistically stable state, samples for both variables are identically and independently distributed with zero mean ($\mu_0 = 0$), unity standard deviation ($\sigma_0 = 1$) and zero cross correlation ($\rho = 0$). They yield random patterns when plotted

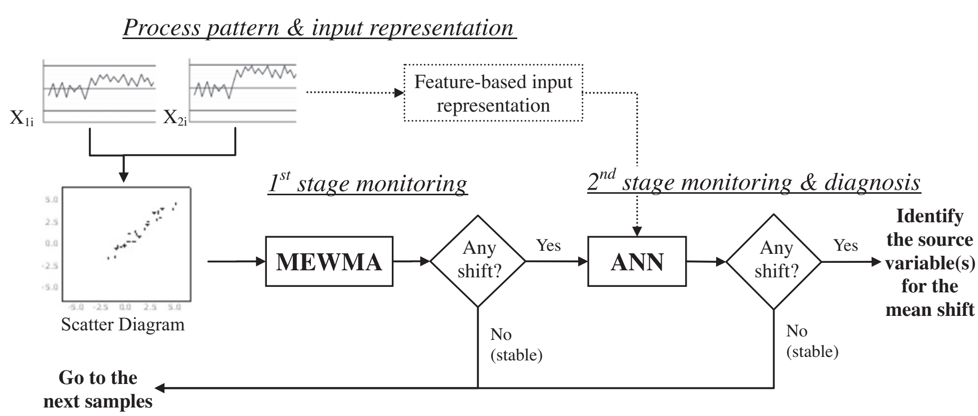

Figure 1 Conceptual diagram of an integrated MEWMA-ANN scheme.

separately on two Shewhart control charts and yield a circle pattern when plotted on a scatter diagram. Scatter diagram yields ellipse patterns when $\rho > 0$ as shown in Fig. 2.

Disturbance from assignable causes may deteriorate data streams into an unstable state. Initially, the pattern structure is in 'partially developed'. Then, it will be more obvious into 'fully developed'. This occurrence could be identified by common causable patterns such as sudden shifts, trends, cyclic, systematic or mixture. However, investigation for this study was focused on sudden shift patterns as follows:

- Normal $(0,0)$: Both X_{1i} and X_{2i} are stable.
- Up-shift $(1,0)$: X_{1i} in up-shift, X_{2i} is stable.
- Up-shift $(0,1)$: X_{2i} in up-shift, X_{1i} is stable.
- Up-shift $(1,1)$: Both X_{1i} and X_{2i} in up-shifts.
- Down-shift $(1,0)$: X_{1i} in down-shift, X_{2i} is stable.
- Down-shift $(0,1)$: X_{2i} in down-shift, X_{1i} is stable.
- Down-shift $(1,1)$: Both X_{1i} and X_{2i} in down-shifts.

Bivariate patterns that attributed to the similar assignable causes would share common structure and properties that are identifiable and recognizable. Changes in mean shift and cross correlation can be identified by center position and ellipse shapes as shown in Fig. 3.

2.1.1. Data generator

Ideally, observation samples for training and testing the scheme should be tapped from real process environment. Since they are not economically available, synthetic data were generated based on the following steps (Lehman, 1977):

- Step 1: Generate random normal variates for process variable 1 (n_1) and process variable 2 (n_2), which are identically and independently distributed within $[-3, +3]$.

$$n_1 = b \cdot r_1 \tag{1}$$

$$n_2 = b \cdot r_2 \tag{2}$$

(r_1, r_2) are random normal variates, whereas $b = 1/3$ is a baseline noise level or random noise.

- Step 2: Transform random normal variates into random data series (Y_1, Y_2):

$$Y_1 = \mu_1 + (n_1\sigma_1) \tag{3}$$

$$Y_2 = \mu_2 + [n_1\rho + n_2(\sqrt{1-\rho^2})]\sigma_2 \tag{4}$$

(μ, σ) are the mean and standard deviation, whereas ρ is the correlation coefficient between (Y_1, Y_2).

- Step 3: Compute mean and standard deviation from (Y_1, Y_2). These values represent in-control (stable) process means (μ_{01}, μ_{02}) and standard deviations $(\sigma_{01}, \sigma_{02})$.

- Step 4: Transform random data series into pattern data series (normal, upward shift and downward shift) to mimic real observation samples (X_1, X_2):

$$X_1 = h_1(\sigma_{01}/\sigma_1) + Y_1 \tag{5}$$

$$X_2 = h_2(\sigma_{02}/\sigma_2) + Y_2 \tag{6}$$

(h_1, h_2) are the magnitudes of mean shift expressed in standard deviation of stable process. A pair observation sample (X_1, X_2) represents a bivariate vector measured at time t (X_t) that follows the random normal bivariate distribution N (μ_0, Σ_0). μ_0 and $\Sigma_0 = [(\sigma_1^2\sigma_{12})(\sigma_{21}\sigma_2^2)]$ are the mean vector and covariance matrix for bivariate stable process with variances (σ_1^2, σ_2^2) and covariance $(\sigma_{12} = \sigma_{21})$.

- Step 5: Rescale pattern data series into a standardize range within $[-3, 3]$.

$$Z_1 = (X_1 - \mu_{01})/\sigma_{01} \tag{7}$$

$$Z_2 = (X_2 - \mu_{02})/\sigma_{02} \tag{8}$$

A pair standardized sample (Z_1, Z_2) represents a standardized bivariate vector measured at time t (Z_t) that follows the standardized normal bivariate distribution N $(0, R)$. Zero and $R = [(1\ \rho)(\rho\ 1)]$ are the mean vector and a general correlation matrix for bivariate stable process with unity variances $(\sigma_1^2 = \sigma_2^2 = 1)$ and covariance equal to correlation $(\sigma_{12} = \sigma_{21} = \rho)$.

2.2. The MEWMA control chart

The formulation of the MEWMA control chart can be found in Lowry et al. (1992). Parameters $(\lambda, H) = (0.10, 8.64)$ as reported in Prabhu and Runger (1997) were selected for this scheme.

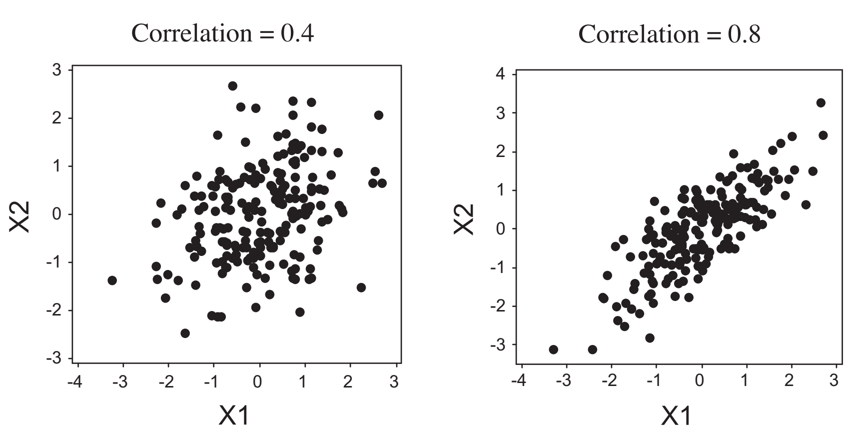

Figure 2 Changes in cross correlation.

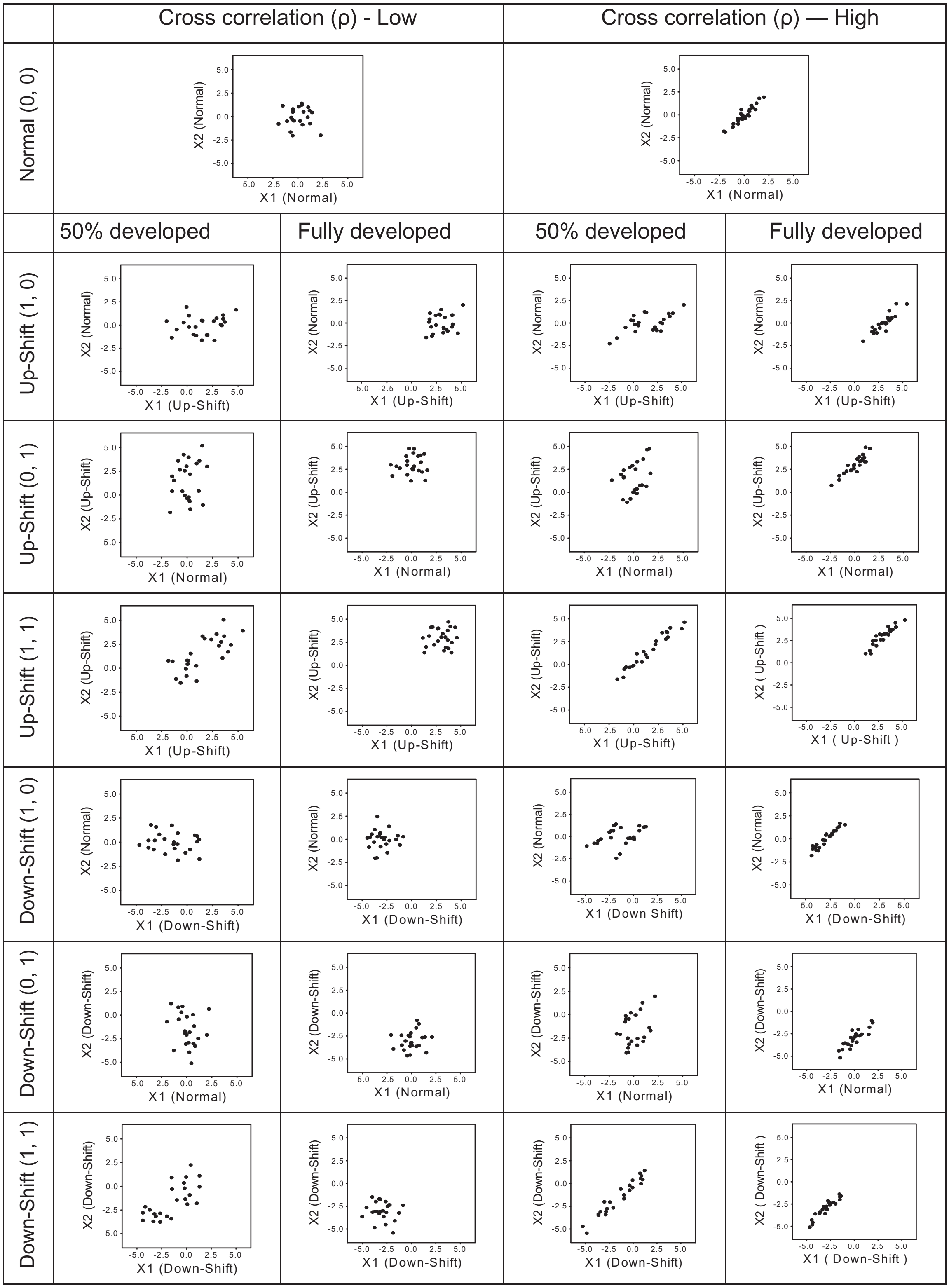

Figure 3 Summary of reference bivariate patterns.

2.3. The Synergistic-ANN recognizer

2.3.1. Input representation

Input representation is a technique to represent input signal into ANN towards achieving effective recognition. In this study, raw data and statistical features input representations were applied into training of Synergistic-ANN recognizer for improving pattern discrimination capability.

Raw data input representation consisted of 48 data, that are: 24 consecutive standardized samples of bivariate process $(Z_{1-P1}, Z_{1-P2}, \ldots, Z_{24-P1}, Z_{24-P2})$.

Statistical features input representation consisted of last value of exponentially weighted moving average ($LEWMA_\lambda$) with $\lambda = [0.25, 0.20, 0.15, 0.10]$, mean ($\mu$), multiplication of mean with standard deviation (MSD), and multiplication of mean with mean square value (MMSV). Each bivariate pattern was represented by 14 data as follows:

$LEWMA_{0.25-P1}$, $LEWMA_{0.20-P1}$, $LEWMA_{0.15-P1}$, $LEWMA_{0.10-P1}$, μ_{P1}, MSD_{P1}, $MMSV_{P1}$, $LEWMA_{0.25-P2}$, $LEWMA_{0.20-P\,2}$, $LEWMA_{0.15-P\,2}$, $LEWMA_{0.10-P\,2}$, μ_2, MSD_{P2}, $MMSV_{P2}$.

$LEWMA_\lambda$ features were taken based on observation window = 24. The EWMA-statistics as derived using Eq. (9) incorporates historical data in a form of weighted average of all past and current observation samples (Lucas and Saccucci, 1990):

$$EWMA_i = \lambda X_i + (1-\lambda)EWMA_{i-1} \tag{9}$$

X_i represents the original samples. In this study, the standardized samples (Z_i) were used instead of X_i so that Eq. (9) becomes:

$$EWMA_i = \lambda Z_i + (1-\lambda)EWMA_{i-1} \tag{10}$$

where $0 < \lambda \leqslant 1$ is a constant parameter and $i = [1, 2, \ldots, 24]$ are the number of samples. The starting value of EWMA ($EWMA_0$) was set as zero to represent the process target (μ_0). Four value of constant parameter (λ = 0.25, 0.20, 0.15, 0.10) were selected based on a range within [0.05, 0.40] recommended by Lucas and Saccucci (1990). Besides resulting longer ARL_0, these parameters could influence the performance of EWMA control chart in detecting process mean shifts. Preliminary experiments suggested that the EWMA with small constant parameter (λ = 0.05) were more sensitive in identifying small shifts ($\leqslant$0.75 standard deviations), while the EWMA with large constant parameter (λ = 0.40) were more sensitive in identifying large shifts ($\geqslant$2.00 standard deviations).

The MSD and MMSV features were used to magnify the magnitude of mean shifts (μ_1, μ_1):

$$MSD_1 = \mu_1 \times \sigma_1 \tag{11}$$

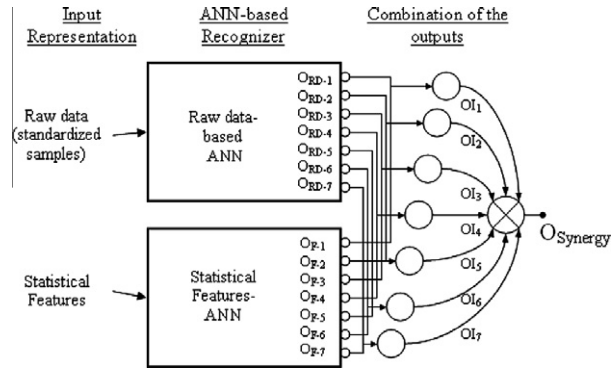

Figure 4 The organization of synergistic-ANN model.

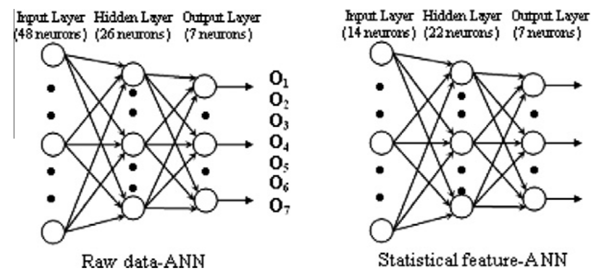

Figure 5 The network structures of the individual ANNs.

$$MSD_2 = \mu_2 \times \sigma_2 \tag{12}$$

$$MMSV_1 = \mu_1 \times (\mu_1)^2 \tag{13}$$

$$MMSV_2 = \mu_2 \times (\mu_2)^2 \tag{14}$$

where (μ_1, μ_2), (σ_1, σ_2) (μ_1^2, μ_1^2) are the means, standard deviations and mean square value, respectively. The mathematical expressions of mean and standard deviation are widely avail-

Table 1 Parameters for the training patterns.

Pattern category	Mean shift (σ in std. dev.)	Total pattern ($100 \times \sigma \times \rho$)
N (0,0)	$X1$: 0.00 $X2$: 0.00	$1.500 \times 1 \times 5 = 7500$
US (1,0)	$X1$: 1.00, 1.25, ..., 3.00 $X2$: 0.00, 0.00, ..., 0.00	$100 \times 9 \times 5 = 4500$
US (0,1)	$X2$: 0.00, 0.00, ..., 0.00 $X1$: 1.00, 1.25, ..., 3.00	$100 \times 9 \times 5 = 4500$
US (1,1)	$X1$: 1.00, 1.25, 1.00, 1.25, ..., 3.00 $X2$: 1.00, 1.00, 1.25, 1.25, ..., 3.00	$100 \times 25 \times 5 = 12{,}500$
DS (1,0)	$X1$: −1.00, −1.25, ..., −3.00 $X2$: 0.00, 0.00, ..., 0.00	$100 \times 9 \times 5 = 4500$
DS (0,1)	$X2$: 0.00, 0.00, ..., 0.00 $X1$: −1.00, −1.25, ..., −3.00	$100 \times 9 \times 5 = 4500$
DS (1,1)	X1: −1.00, −1.25, −1.00, −1.25, ..., −3.00 $X2$: −1.00, −1.00, −1.25, −1.25, ..., −3.00	$100 \times 25 \times 5 = 12{,}500$

able in textbook on SPC. The mean square value feature can be derived as in Hassan et al. (2003).

2.3.2. Recognizer design

A combined ANN model, namely, 'Synergistic-ANN' was developed for pattern recognizer. It is a parallel combination between two individual ANNs that are: (i) raw data-based ANN and (ii) statistical features-ANN as shown in Fig. 4.

Let $O_{RD} = (O_{RD-1}, \ldots, O_{RD-7})$ and $O_F = (O_{F-1}, \ldots, O_{F-7})$ represent seven outputs from raw data-based ANN and statistical features-ANN recognizers respectively. Outputs from the two recognizers can be combined through simple summation: $OI_i = \Sigma(O_{RD-i}, O_{F-i})$, where $i = (1, \ldots, 7)$ are the number of outputs. Final decision ($O_{synergy}$) was determined based on the maximum value from the combined outputs:

$$O_{synergy} = \max(OI_1, \ldots, OI_7) \quad (15)$$

Multilayer perceptron (MLP) model was applied for the individual ANNs as shown in Fig. 5. This model comprises an input layer, one or more hidden layer(s) and an output layer. The size of input representation determines the number of input neurons. Raw data input representation requires 48 neurons, while statistical features input representation requires only 14 neurons. The output layer contains seven neurons, which was determined according to the number of pattern categories.

Table 2 Performance comparison between the EWMA-ANN and the Basic schemes.

Pattern category	Mean shift (σ)		Average run lengths (ARL_0, ARL_1)		Recognition accuracy (RA)	
			Basic scheme	EWMA-ANN	Basic scheme	EWMA-ANN
	$X1$	$X2$	ρ = 0.1, 0.5, 0.9	ρ = 0.1, 0.5, 0.9	ρ = 0.1, 0.5, 0.9	ρ = 0.1, 0.5, 0.9
N (0,0)	0.00	0.00	163.83, 318.2, 249.49	335.01, 543.93, 477.5		
US (1,0)	0.75	0.00	16.32, 16.70, 17.58	17.60, 18.34, 20.00	88.4, 87.2, 87.2	92.7, 90.4, 89.5
US (0,1)	0.00	0.75	14.04, 14.36, 14.73	16.20, 15.99, 16.21	89.1, 88.4, 89.2	92.9, 89.3, 90.6
US (1,1)	0.75	0.75	13.42, 12.80, 12.78	13.64, 13.28, 14.17	79.9, 93.4, 99.7	82.4, 94.8, 99.9
DS (1,0)	−0.75	0.00	13.63, 14.44, 14.83	16.31, 16.43, 17.35	89.9, 89.1, 88.9	92.3, 89.2, 89.4
DS (0,1)	0.00	−0.75	15.37, 16.10, 16.64	16.94, 17.44, 18.75	88.6, 88.2, 89.7	92.3, 87.8, 88.5
DS (1,1)	−0.75	−0.75	13.46, 13.00, 12.82	13.46, 13.37, 14.03	77.0, 92.3, 99.5	84.1, 96.1, 99.9
			14.37, 14.57, 14.90	15.69, 15.81, 16.75	**85.5, 89.8, 92.4**	89.5, 91.3, 93.0
US (1,0)	1.00	0.00	10.83, 11.04, 11.11	11.52, 11.57, 11.70	92.5, 92.1, 91.2	95.3, 93.1, 94.4
US (0,1)	0.00	1.00	9.54, 9.71, 9.78	10.50, 10.22, 10.20	91.5, 92.9, 91.2	95.8, 93.5, 94.4
US (1,1)	1.00	1.00	9.38, 9.23, 9.07	9.16, 9.09, 9.66	85.5, 96.1, 99.9	90.0, 96.5, 100
DS (1,0)	−1.00	0.00	9.71, 9.95, 10.07	10.99, 10.86, 11.06	93.5, 92.4, 93.4	95.3, 93.2, 92.3
DS (0,1)	0.00	−1.00	10.59, 10.83, 10.70	11.08, 11.12, 11.36	91.9, 91.8, 91.9	93.8, 92.1, 92.6
DS (1,1)	−1.00	−1.00	9.66, 9.46, 9.35	9.15, 9.12, 9.63	81.9, 94.5, 99.8	89.5, 98.0, 100
			9.95, 10.04, 10.01	10.40, 10.33, 10.60	**89.5, 93.3, 94.6**	93.3, 94.4, 95.6
US (1,0)	1.50	0.00	7.17, 7.16, 7.17	7.02, 7.07, 7.03	96.0, 95.9, 95.3	97.4, 96.5, 97.1
US (0,1)	0.00	1.50	6.37, 6.42, 6.44	6.54, 6.33, 6.40	95.2, 95.4, 95.1	97.1, 96.5, 96.2
US (1,1)	1.50	1.50	6.47, 6.32, 6.30	5.82, 5.73, 5.94	89.7, 97.2, 99.9	91.7, 97.9, 100
DS (1,0)	−1.50	0.00	6.58, 6.59, 6.65	6.81, 6.81, 6.92	95.8, 96.3, 96.6	97.4, 96.3, 95.5
DS (0,1)	0.00	−1.50	7.09, 7.09, 7.01	6.82, 6.80, 6.85	94.7, 95.4, 95.6	96.2, 95.8, 95.6
DS (1,1)	−1.50	−1.50	6.57, 6.49, 6.34	5.81, 5.69, 5.98	87.2, 97.1, 99.9	93.2, 99.0, 100
			6.71, 6.68, 6.65	6.47, 6.41, 6.52	**93.1, 96.2, 97.1**	95.5, 97.0, 97.4
US (1,0)	2.00	0.00	5.55, 5.51, 5.55	5.23, 5.15, 5.19	96.1, 95.9, 96.3	97.8, 97.1, 97.6
US (0,1)	0.00	2.00	5.01, 4.99, 5.00	4.80, 4.72, 4.70	95.8, 95.7, 95.3	97.7, 97.8, 97.1
US (1,1)	2.00	2.00	5.04, 4.97, 4.89	4.36, 4.32, 4.39	90.7, 97.1, 99.9	91.6, 98.4, 100
DS (1,0)	−2.00	0.00	4.98, 5.01, 5.04	5.04, 5.04, 5.02	97.5, 97.1, 96.5	96.8, 96.7, 96.6
DS (0,1)	0.00	−2.00	5.33, 5.42, 5.42	4.97, 5.03, 4.98	95.9, 96.2, 95.6	96.5, 96.5, 95.6
DS (1,1)	−2.00	−2.00	5.18, 5.06, 5.03	4.29, 4.27, 4.33	87.9, 96.7, 99.9	93.7, 98.9, 100
			5.18, 5.16, 5.16	4.78, 4.76, 4.77	**94.0, 96.5, 97.3**	95.7, 97.6, 97.8
US (1,0)	2.50	0.00	4.52, 4.56, 4.57	4.10, 4.14, 4.12	96.8, 96.3, 97.1	98.0, 98.4, 98.0
US (0,1)	0.00	2.50	4.15, 4.16, 4.14	3.83, 3.81, 3.81	95.7, 95.1, 95.0	97.3, 97.4, 97.0
US (1,1)	2.50	2.50	4.20, 4.15, 4.12	3.54, 3.49, 3.53	91.2, 97.5, 99.9	93.2, 98.4, 100
DS (1,0)	−2.50	0.00	4.04, 4.02, 4.02	3.99, 3.96, 3.95	97.8, 97.7, 96.8	97.3, 97.3, 97.0
DS (0,1)	0.00	−2.50	4.36, 4.38, 4.42	3.97, 4.02, 3.98	96.7, 96.4, 96.3	96.5, 96.6, 97.0
DS (1,1)	−2.50	−2.50	4.24, 4.22, 4.09	3.41, 3.40, 3.46	88.3, 96.3, 99.9	94.9, 98.8, 100
			4.25, 4.25, 4.23	3.81, 3.80, 3.81	**94.4, 96.6, 97.5**	96.2, 97.8, 98.2
US (1,0)	3.00	0.00	3.90, 3.91, 3.89	3.47, 3.46, 3.46	97.7, 97.5, 97.3	98.6, 98.3, 98.2
US (0,1)	0.00	3.00	3.57, 3.59, 3.58	3.20, 3.20, 3.21	95.6, 95.5, 95.8	97.8, 97.8, 98.0
US (1,1)	3.00	3.00	3.63, 3.60, 3.56	2.98, 2.93, 2.98	91.1, 97.7, 99.9	93.8, 98.4, 100
DS (1,0)	−3.00	0.00	3.41, 3.37, 3.42	3.31, 3.30, 3.27	98.0, 98.3, 97.4	98.0, 97.1, 97.6
DS (0,1)	0.00	−3.00	3.72, 3.71, 3.74	3.33, 3.32, 3.32	96.6, 96.7, 97.2	96.7, 97.1, 97.1
DS (1,1)	−3.00	−3.00	3.69, 3.60, 3.56	2.84, 2.85, 2.90	88.2, 96.8, 99.8	94.6, 99.1, 100
			3.65, 3.63, 3.63	3.19, 3.18, 3.19	**94.5, 97.1, 97.9**	96.6, 98.0, 98.5
Overall	±(0.75, 3.00)		7.35, 7.39, 7.43	7.39, 7.38, 7.61	91.8, 94.9, 96.1	94.5, 96.0, 96.8

Based on preliminary experiments, one hidden layer with 26 neurons and 22 neurons were selected for raw data-based ANN and statistical features-ANN. The experiments revealed that initially, the training results improved in-line with the increment in the number of neurons. Once the neurons exceeded the required numbers, further increment of the neurons did not improve the training results but provided poorer results. These excess neurons could burden the network computationally, reduces the network generalization capability and increases the training time.

2.3.3. Recognizer training and testing

Partially developed patterns of bivariate process mean shifts were applied for training the synergistic-ANN recognizer. Detail parameters of the training patterns are summarized in Table 1. It should be noted that for bivariate process mean shifts, the number of training pattern = [100 × (total combination of shifts) × (total combinations of cross correlation)], while for bivariate normal process, the number of training pattern = [1500 × (total combinations of cross correlation)]. On the other hand, dynamic patterns were used for testing the recognizers, which is suited for on-line process monitoring as addressed in Guh (2007).

Input representations were normalized to a compact range between $[-1, 1]$. The maximum and the minimum values for normalization were taken from the overall data of training patterns.

Based on back propagation (BPN) algorithm, 'gradient decent with momentum and adaptive learning rate' (traingdx) was used for training the MLP model. The other training parameters setting were learning rate (0.05), learning rate increment (1.05), maximum number of epochs (1500) and error goal (0.001), whereas the network performance was based on mean square error (MSE). Hyperbolic tangent function was used for hidden layer, while sigmoid function was used for an output layer. The training session was stopped either when the number of training epochs was met or the required MSE has been reached.

3. Results and discussion

The monitoring and diagnosis performances were evaluated based on average run lengths (ARL_0, ARL_1) and recognition accuracy percentage (RA) as summarized in Table 2. The performance results involve comparison between an integrated MEWMA-ANN scheme and Basic scheme. The comparison is supported mathematically with statistical significant test as summarized in Table 3.

The basic scheme as shown in Fig. 6 was developed based on 'single stage monitoring and diagnosis' approach, similar to the existing BPR schemes as discussed in Section 1. Raw data was utilized as input representation, whereas single ANN model (MLP) was applied as the pattern recognizer.

3.1. Monitoring performance

In monitoring aspect, the values of ARL_1 measure how fast it could detect the mean shift, while the values of ARL_0 measure how long it could maintain stable process running without false alarm. These values (ARL_0, ARL_1) were computed based on the correctly classified patterns.

Based on ARL_0 results, an integrated MEWMA-ANN scheme has shown a huge increment (335.01, 543.93, 477.45) compared to the Basic scheme (163.83, 318.2, 249.49). This improvement as proven statistically (mean difference of ARL_0 satisfied 95% confident interval with $P = 0.008$) clearly indicates that the proposed scheme is very efficient

Table 3 Statistical significant test of performance results in this table.

Performance measures (PM)	Result of paired *T*-test mean difference of PM (EWMA-ANN minus Basic)					Remarks
		N	Mean	St. dev.	SE mean	
ARL_0	EWMA-ANN	3	452.130	106.737	61.624	
	Basic	3	243.840	77.340	44.652	
	Difference	3	208.290	32.158	18.566	
	Mean difference of ARL_0:					Increment in ARL_0 is proven as statistically significant
	95% CI:	(128.406, 288.174)				
	T-Test = 0 (vs ≠ 0): $T = 11.22$, $P = 0.008$					
ARL_1	EWMA-ANN	18	7.459	4.665	1.100	
	Basic	18	7.390	3.951	0.931	
	Difference	18	0.069	0.730	0.172	
	Mean difference of ARL_1:					Increment in ARL_1 is not statistically significant
	95% CI:	(−0.293, 0.432)				
	T-Test = 0 (vs ≠ 0): $T = 0.400$, $P = 0.691$					
RA	EWMA-ANN	18	95.744	2.544	0.600	
	Basic	18	94.294	3.331	0.785	
	Difference	18	1.450	1.057	0.249	
	Mean difference of RA:					Increment in RA is proven as statistically significant
	95% CI:	(0.924, 1.976)				
	T-Test = 0 (vs ≠ 0): $T = 5.82$, $P = 0.000$					

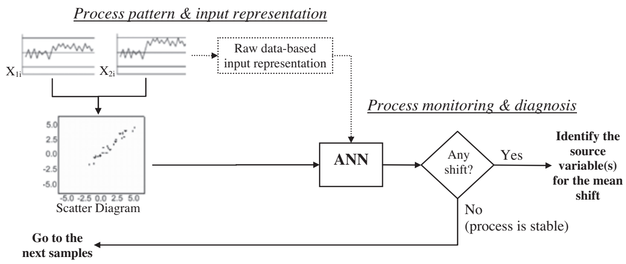

Figure 6 Conceptual diagram of the Basic scheme.

to maintain small false alarm in various sizes of cross correlation samples. On the other hand, based on ARL_1 results, an integrated MEWMA-ANN scheme has shown longer ARL_1 for small mean shifts (0.75–1.00 standard deviations) and shorter ARL_1 for moderate and large mean shifts (1.50–3.00 standard deviations) compared to the Basic scheme. This shows that the proposed scheme is more sensitive to detect moderate and large process variations, while the Basic scheme is more sensitive to detect small process variations. Comparison based on overall magnitudes of mean shifts indicates that both schemes have an equivalent ARL_1 performance as proven statistically (mean difference of $ARL_1 = +0.069$, $P = 0.691$).

3.2. Diagnosis performance

In diagnosis aspect, RA measures how accurate it could identify the source variable(s) towards diagnosing the root cause. The integrated MEWMA-ANN scheme has shown significant increment in RA (94.5%, 96.0%, 96.8%) compared to the Basic scheme (91.8%, 94.9%, 96.1%). This increment as proven statistically (mean difference of RA = +1.45%, $P = 0.000$) is strongly influenced by the application of Synergistic-ANN recognizer towards improving pattern discrimination capability.

4. Conclusions

This paper proposed an integrated MEWMA-ANN scheme towards achieving 'balanced monitoring and accurate diagnosis' performances in dealing with bivariate process mean shifts. Based on two-stages monitoring and diagnosis approach, the proposed scheme has resulted in a smaller false alarm, quick mean shift detection and higher diagnosis accuracy compared to the Basic scheme (based on raw data input representation and single ANN recognizer). 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 517, 517–543.

Chen, L.H., Wang, T.Y., 2004. Artificial neural networks to classify mean shifts from multivariate χ^2 chart signals. Computers and Industrial Engineering 47, 195–205.

Cheng, C.S., Cheng, H.P., 2008. Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines. Expert Systems with Applications 35, 198–206.

El-Midany, T.T., El-Baz, M.A., Abd-Elwahed, M.S., 2010. A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks. Expert Systems with Applications 37, 1035–1042.

Guh, R.S., 2007. On-line identification and quantification of mean shifts in bivariate processes using a NN-based approach. Quality and Reliability Engineering International 23, 367–385.

Hassan, A., Nabi Baksh, M.S., Shaharoun, A.M., Jamaludin, H., 2003. Improved SPC chart pattern recognition using statistical features. International Journal of Production Research 41 (7), 1587–1603.

Kourti, T., MacGregor, J.F., 1996. Multivariate SPC methods for process and product monitoring. Journal of Quality Technology 28 (4), 409–428.

Lehman, R.S., 1977. Computer Simulation and Modeling: An Introduction. Lawrence Erlbaum, London.

Lowry, C.A., Montgomery, D.C., 1995. A review of multivariate control charts. IIE Transactions 27 (6), 800–810.

Lowry, C.A., Woodall, W.H., Champ, C.W., Rigdon, S.E., 1992. A multivariate exponentially weighted moving average control chart. Technometrics 34 (1), 46–53.

Lucas, J.M., Saccucci, M.S., 1990. Exponentially weighted moving average control schemes: properties and enhancements. Technometrics 32 (1), 1–12.

Mason, R.L., Tracy, N.D., Young, J.C., 1997. A practical approach for interpreting multivariate T^2 control chart signals. Journal of Quality Technology 29 (4), 396–406.

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.

Niaki, S.T.A., Abbasi, B., 2005. Fault diagnosis in multivariate control charts using artificial neural networks. Quality and Reliability Engineering International 21, 825–840.

Prabhu, S.S., Runger, G.C., 1997. Designing a multivariate EWMA control chart. Journal of Quality Technology 29 (1), 8–15.

Yu, J.B., Xi, L.F., 2009. A neural network ensemble-based model for on-line monitoring and diagnosis of out-of-control signals in multivariate manufacturing processes. Expert Systems with Applications 36, 909–921.

Yu, J.B., Xi, L.F., Zhou, X.J., 2009. Identifying source(s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble. Engineering Applications of Artificial Intelligence 22, 141–152.

Zorriassatine, F., Tannock, J.D.T., O'Brien, C., 2003. Using novelty detection to identify abnormalities caused by mean shifts in bivariate processes. Computers and Industrial Engineering 44, 385–408.