A Scheme for Balanced Monitoring and Accurate Diagnosis of Bivariate Correlated Process Mean Shifts

Ibrahim Masood
Faculty of Mechanical and Manufacturing Engineering
86400 Parit Raja, Batu Pahat, Johor, Malaysia

Adnan Hassan
Faculty of Mechanical Engineering
81310 UTM Skudai, Johor, Malaysia

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 suffers high false alarm, that is, imbalanced monitoring performance. The problem becomes more complicated when dealing with small mean shift particularly for identifying the causable variables. In this research, a framework to address balanced monitoring and accurate diagnosis was investigated. Design considerations involved extensive simulation experiments to select input representation based on raw data and statistical features, recognizer design structure based on synergistic model, and monitoring-diagnosis approach based on two stages technique. The study focuses on correlated process mean shifts for cross correlation function, ρ = 0.1 ~ 0.9 and mean shift, μ = ± 0.75 ~ 3.00 standard deviations. The proposed design, that is, an Integrated Multivariate Exponentially Weighted Moving Average with Artificial Neural Network scheme gave superior performance, namely, average run length, ARL_1 = 3.18 ~ 16.75 (for out-of-control process), ARL_0 = 452.13 (for in-control process) and recognition accuracy, RA = 89.5 ~ 98.5%. The proposed scheme was validated using an industrial case study from machining process of audio video device component. This research has provided a new perspective in realizing balanced monitoring and accurate diagnosis of correlated process mean shifts.

Keywords
Balance monitoring, bivariate pattern recognition, statistical features, synergistic-ANN, two-stages monitoring and diagnosis

1. Introduction
In manufacturing industries, process variation has become a major source of poor quality. When manufacturing process involves two or more correlated variables, an appropriate procedure/scheme is necessary to monitor these variables jointly. In addressing this issue, the traditional multivariate statistical process control (MSPC) schemes such as T^2, multivariate cumulative sum (MCUSUM) and multivariate exponentially weighted moving average (MEWMA) are known effective in detecting the process mean shifts. Nevertheless, they are lack of capability in identifying the source variable(s) that responsible to the process mean shifts. In other word, it is unable to provide diagnosis information for a quality practitioner towards finding the root cause errors and solution for corrective action. Therefore, major researches have been focused for improving capability in identifying the source variable(s). Further discussions on this issue can be found in Lowry and Montgomery (1995), Kourti and MacGregor JF (1996), Mason *et al.* (1997) and Bersimis *et al.* (2007).

Development in artificial intelligence techniques has motivated researchers to explore the use of artificial neural networks (ANN), among others, for automatically recognizing patterns in relation to monitoring-diagnosis of multivariate process mean shifts. Identification of these patterns coupled with engineering knowledge of the process would lead to more specific diagnosis and troubleshooting. Various types of ANN pattern recognition schemes namely, MSPC-ANN (Chen and Wang, 2004; Niaki and Abbasi, 2005; Cheng and Cheng, 2008; Yu *et al.*, 2009), Novelty Detector ANN (Zorriassatine *et al.*, 2003), Modular-ANN (Guh, 2007), Ensemble-ANN (Yu and Xi, 2009) and Multi-Module-Structure-ANN (El-Midany *et al.*, (2010) have been investigated. The MSPC-ANN schemes consisted of the integration between the traditional MSPC charts and the ANN-based model. The MSPC control chart is utilized for monitoring the existing of mean/variance shifts in multivariate process, whereas the ANN-based model is utilized for diagnosing the source variable(s) that responsible for mean/variance shifts. Besides these integrated schemes, the other schemes (Novelty Detector ANN and others) have been developed for simultaneous monitoring-diagnosis by fully utilization of ANN-based model as pattern recognizer. Further discussion on these schemes can be found in (Masood and Hassan, 2010).

In monitoring aspect, such existing integrated MSPC-ANN and fully ANN-based schemes have provided faster mean shift detection. However, most of them are suffers in high false alarms (average run length, $ARL_0 \leq 200$) in comparison to the *de facto* level for univariate SPC schemes ($ARL_0 \geq 370$). This would be critical for a quality practitioner in conducting unnecessary troubleshooting due to wrong identification of in-control process as out-of-control. In this study, this situation is called ‘imbalance monitoring’. In diagnosis aspect, on the other hand, most of the fully ANN-based schemes are lack in accurately identifying the source variable(s) especially when dealing with small mean shifts. This would be critical for a quality practitioner in searching the root cause errors due to wrong identification of the source variable(s) of mean shifts. In this study, this situation is called ‘lack of diagnosis’. In order to overcome these disadvantages, an improved scheme namely, an integrated MEWMA-ANN was developed towards ‘balanced monitoring and accurate diagnosis’ of bivariate correlated process mean shifts. This proposed scheme aims for enabling faster detection of process mean shifts with minimum false alarm and high accuracy in diagnosing the source variable(s) of process mean shifts. 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 traditional existing schemes. Section 4 finally outlines some conclusions.

2. An Integrated MEWMA-ANN Scheme

An integrated MEWMA-ANN scheme was designed and developed based on two stages monitoring and diagnosis approach as shown in Figure 1. Process monitoring refers to the identification of process status either it is running within a statistically in-control or out-of-control condition, while process diagnosis refers to the identification of the source variable(s) of out-of-control process in mean shifts. In the first stage monitoring, the MEWMA control chart is utilized for triggering the existing of process mean shifts based on ‘one point out-of-control’. Once the mean shift is detected, the Synergistic-ANN model is then utilized for conducting second stage monitoring and diagnosis by recognizing data stream pattern contained point(s) out-of-control.

In general, the MEWMA control chart is known effective for detecting bivariate process mean shifts more rapidly as compared to the T^2 control chart. Unfortunately, based on a quite short ARL_0 (≤ 200), it shows limited capability to minimize false alarms towards achieving the *de facto* level ($ARL_0 \geq 370$). On the other hand, the Synergistic-ANN gave better capability in reducing false alarms (with longer ARL_0). Therefore, it can be concluded that process identification based on ‘recognition of process data stream patterns’ (using the Synergistic-ANN scheme) is more effective compared to ‘detection of one point out-of-control’ (using the MEWMA control chart). Different techniques should have their respective advantages in terms of point/pattern discrimination properties. As such, in order to further improve the monitoring performance towards achieving the *de facto* level ($ARL_0 \geq 370$), it is useful to combine both discrimination properties (MEWMA control chart and Synergistic-ANN recognizer) by approaching two-stages monitoring and diagnosis. In the first stage monitoring, the MEWMA control chart is used for triggering bivariate process mean shifts based on ‘one point out-of-control’ as per usual. Once the shift is triggered, the Synergistic-ANN recognizer will performs second stage monitoring and diagnosis through recognition of data stream patterns that contain ‘out-of-control point(s)’. This approach is suited for ‘recognition only when necessary’ concept as introduced by Hassan (2002). He noted that it is unnecessary to perform recognition while the process lies within a statistically in-control state. Besides, recognition is only necessary for identifying patterns suspected to a statistically out-of-control state. Besides minimum false alarms, this approach will also reduce computational efforts and time consumes for pattern recognition operation.

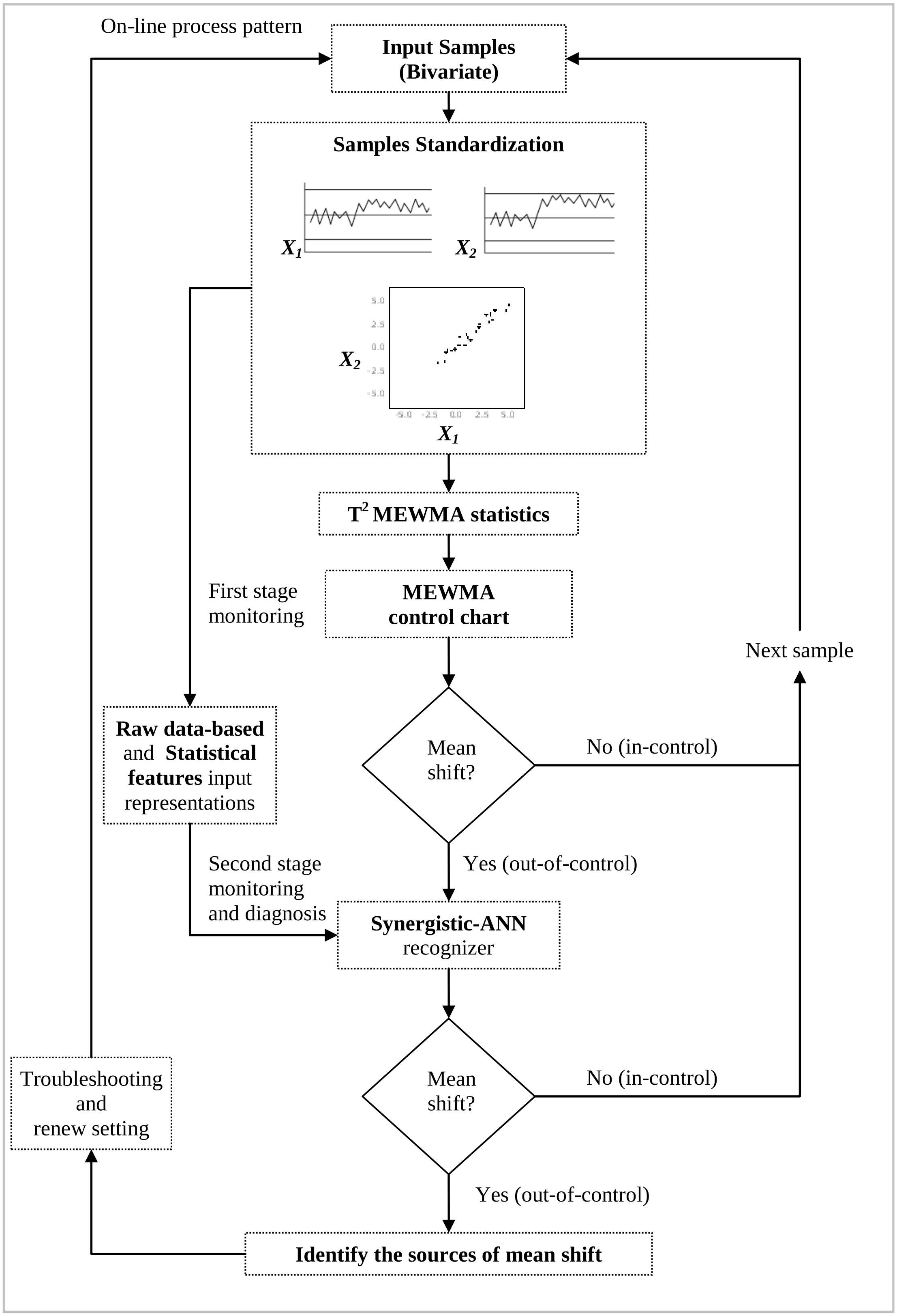

Figure 1: An Integrated MEWMA-ANN scheme

Implementation procedures for an Integrated MEWMA-ANN scheme are summarized in Table 1. It should be noted that an initial setting as follows 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 (μ_{01}, μ_{02}) and standard deviations (σ_{01}, σ_{02}) of bivariate in-control process (for variables X_1 and X_2). These parameters can be obtained based on historical or preliminary samples.
- Perform in-process quality control inspection until 24 observation samples (individual or subgroup) to begin the system.

Recognition window size is set to 24 observation samples (for variables X_1 and X_2) since it provided sufficient training results and statistically acceptable to represent normal distribution. Preliminary experiments suggested that a smaller window size (< 24) gave lower training result due to insufficient pattern properties, while a larger window size (> 24) does not increase the training result but burden the ANN training.

Table 1: Implementation Procedures for the Integrated MEWMA-ANN scheme

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

Step 2: <u>Samples standardization</u>. 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: <u>First stage monitoring (MEWMA control chart):</u>

Standardized samples (Z_1, Z_2) are converted into MEWMA statistics (T^2_{MEWMA}).

Decision rule: *If* T^2_{MEWMA} < Upper control limit (h)

Proceed to the next samples

else

Proceed to the second stage monitoring.

end

Step 4: <u>Input representation.</u>

(i) Raw data-based input representation: $Z_{1-P1}, Z_{2-P1}, \ldots, Z_{1-P24}, Z_{2-P24}$

(ii) Statistical features input representation:

- 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).
- Fourtheen 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 5: <u>Second stage monitoring and diagnosis (synergistic-ANN recognizer):</u>

Decision rule: *If* Maximum output of ANN belongs to N(0, 0):

<u>Process is "in-control"</u>; proceed to the next samples

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

end

2.1 Modeling of Data and Patterns of Bivariate Process Mean Shifts

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

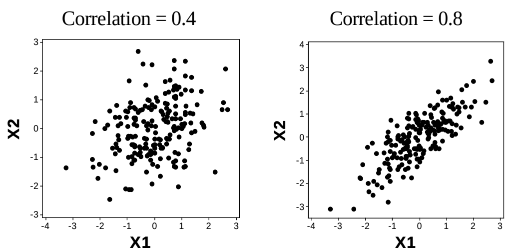

Figure 2: Changes in data correlation

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. Seven possible categories of bivariate patterns as follows were considered in representing the bivariate process variation in mean shifts:

- Normal(0,0)/N(0,0): both variables X_{1i} and X_{2i} remain in-control
- Up-Shift(1,0)/US(1,0): X_{1i} shifted upwards, while X_{2i} remains in-control
- Up-Shift(0,1)/US(0,1): X_{2i} shifted upwards, while X_{1i} remains in-control
- Up-Shift(1,1)/US(1,1): both variables X_{1i} and X_{2i} shifted upwards
- Down-Shift(1,0)/DS(1,0): X_{1i} shifted downwards, while X_{2i} remains in-control
- Down-Shift(0,1)/DS(0,1): X_{2i} shifted downwards, while X_{1i} remains in-control
- Down-Shift(1,1)/DS(1,1): both variables X_{1i} and X_{2i} shifted downwards

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. A large amount of bivariate samples is required to perform training-testing for raw data-ANN recognizer. Ideally, such samples should be tapped from real world. Unfortunately, they are not economically available or too limited. As such, there is a need for modeling of synthetic samples. The synthetic samples of bivariate process were generated based on the following steps (Lehman, 1977):

<u>Step 1</u>: Generate random normal variates for process variable-1 (n_1) and process variable-2 (n_2), which is identically and independently distributed (i.i.d.) within [−3, +3]:

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

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

Parameters (r_1, r_2) and b represent random normal variates (random data) and noise level respectively. Random normal variates are computerized generated data, whereby the noise level is used to rescale its variability. In this research, b = 1/3 is used to ensure that all random data are simulated within ± 3.00 standard deviations (not exceed the control limits of Shewhart control chart).

<u>Step 2</u>: Transform random normal variates for process variable-1 (n_1) into random data series (Y_1):

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

Parameters μ_1 and σ_1 respectively represent the mean and the standard deviation for Y_1.

Step 3: Transform random normal variates for process variable-2 (n_2) into random data series (Y_2) dependent to (Y_1).

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

Parameters μ_2 and σ_2 respectively represent the mean and the standard deviation for Y_2, whereas ρ represents the correlation coefficient between (Y_1, Y_2).

Step 4: Compute mean and standard deviation from (Y_1, Y_2). These values represent in-control process mean (μ_{01}, μ_{02}) and standard deviation (σ_{01}, σ_{02}) for component variables.

Step 5: Transform random data series (Y_1, Y_2) into normal or shift (pattern) data streams to mimic real observation samples (X_1, X_2):

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

$$X_2 = h_2\,(\sigma_{01}/\sigma_2) + Y_2 \tag{6}$$

The magnitudes of mean shifts (h_1, h_2) are expressed in standard deviation of in-control 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). The notations μ_0 and $\Sigma_0 = [(\sigma_1^2\ \sigma_{12})\ (\sigma_{12}\ \sigma_2^2)]$ represent mean vector and covariance matrix for bivariate in-control process with variances (σ_1^2, σ_2^2) and covariance ($\sigma_{12} = \sigma_{21}$).

Step 6: Rescale pattern data streams 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 value and R = [(1 ρ) (ρ 1)]) represent mean vector and general correlation matrix for bivariate in-control process with unity variances ($\sigma_1^2 = \sigma_2^2 = 1$) and covariance equal to cross correlation ($\sigma_{12} = \sigma_{21} = \rho$).

2.2 MEWMA Control Chart

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

2.3 Synergistic-ANN Recognizer

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

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 ($Z_{1\text{ - P1}}$, $Z_{1\text{ - P2}}$, …, $Z_{24\text{ - P1}}$, $Z_{24\text{ - P2}}$). Statistical features input representation consisted of last value of exponentially weighted moving average (LEWMA_λ) 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). Each bivariate pattern was represented by 14 data as follows: $\text{LEWMA}_{0.25\text{-P1}}$, $\text{LEWMA}_{0.20\text{-P1}}$, $\text{LEWMA}_{0.15\text{-P1}}$, $\text{LEWMA}_{0.10\text{-P1}}$, μ_{P1}, MSD_{P1}, MMSV_{P1}, $\text{LEWMA}_{0.25\text{-P2}}$, $\text{LEWMA}_{0.20\text{-P2}}$, $\text{LEWMA}_{0.15\text{-P2}}$, $\text{LEWMA}_{0.10\text{-P2}}$, μ_2, MSD_{P2}, MMSV_{P2}.

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

$$\text{EWMA}_i = \lambda X_i + (1 - \lambda)\,\text{EWMA}_{i-1} \tag{9}$$

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

$$\text{EWMA}_i = \lambda Z_i + (1 - \lambda)\,\text{EWMA}_{i-1} \tag{10}$$

where $0 < \lambda \leq 1$ is a constant parameter and i = [1, 2,…, 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 ($\leq$ 0.75 standard deviations), while the EWMA with large constant parameter (λ = 0.40) were more sensitive in identifying large shifts ($\geq$ 2.00 standard deviations).

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

$$MSD_1 = \mu_1 \text{ x } \sigma_1 \tag{11}$$

$$MSD_2 = \mu_2 \text{ x } \sigma_2 \tag{12}$$

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

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

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

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 Figure 3.

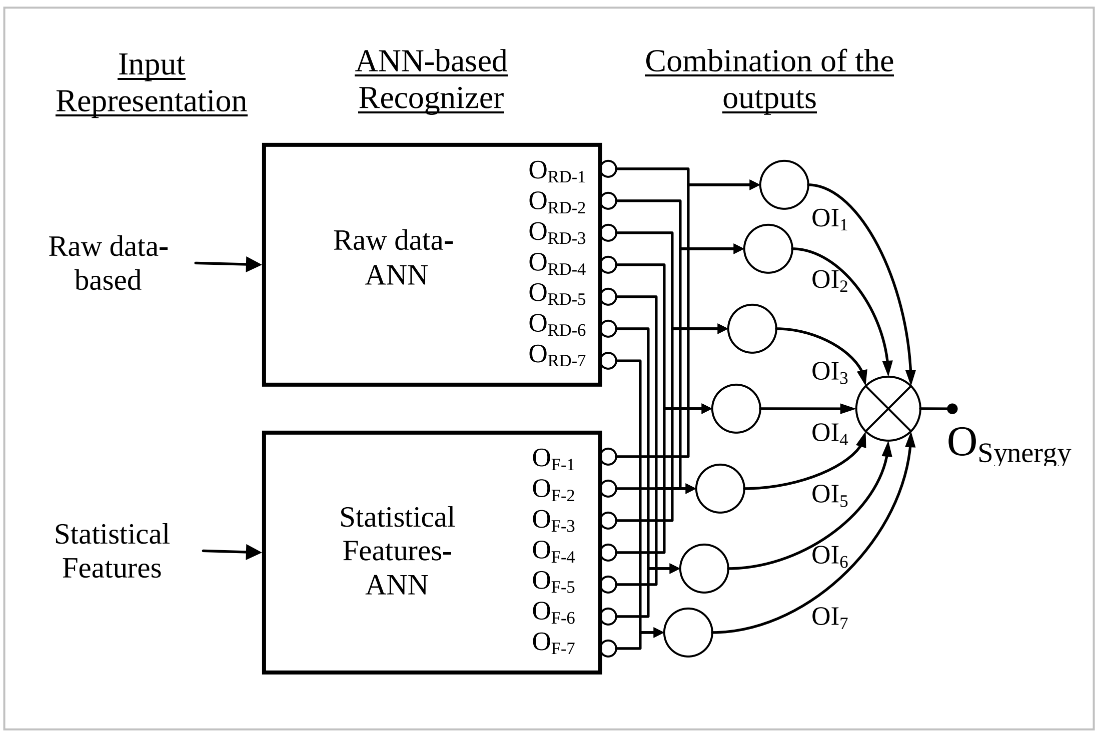

Figure 3: Synergistic-ANN model

Let O_{RD} = (O_{RD-1}, …, O_{RD-7}) and O_F = (O_{F-1}, …, 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,…,7) are the number of outputs. Final decision ($O_{synergy}$) was determined based on the maximum value from the combined outputs:

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

Multilayer perceptron (MLP) model was applied for the individual ANNs as shown in Figure 4. 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. 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.

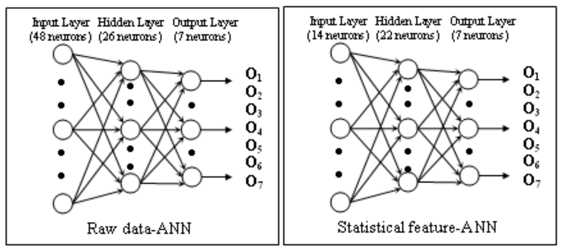

Figure 4: Individual ANN recognizer

<u>Recognizer Training and Testing</u>

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 2. It should be noted that for bivariate process mean shifts, the number of training pattern = [100 x (total combination of shifts) x (total combinations of cross correlation)], while for bivariate normal process, the number of training pattern = [1500 x (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.

Table 2: Parameters for training patterns

Pattern Category	**Mean Shift (σ in Std Dev.)**	**Total Pattern (100 x σ x ρ)**
N (0, 0)	X1: 0.00 X2: 0.00	1,500 x 1 x 5 = 7,500
US (1, 0)	X1: 1.00, 1.25, … , 3.00 X2: 0.00, 0.00, … , 0.00	100 x 9 x 5 = 4,500
US (0, 1)	X2: 0.00, 0.00, … , 0.00 X1: 1.00, 1.25, … , 3.00	100 x 9 x 5 = 4,500
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 x 25 x 5 = 12,500
DS (1, 0)	X1: -1.00, -1.25, … , -3.00 X2: 0.00, 0.00, … , 0.00	100 x 9 x 5 = 4,500
DS (0, 1)	X2: 0.00, 0.00, … , 0.00 X1: -1.00, -1.25, … , -3.00	100 x 9 x 5 = 4,500
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 x 25 x 5 = 12,500

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. Performance 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 3. The performance results involve comparison between an integrated MEWMA-ANN scheme and the existing schemes. The comparison is supported mathematically with statistical significant test as summarized in Table 4.

The monitoring and diagnosis performances of the Integrated MEWMA-ANN scheme as proposed in this research were compared to the existing MPR schemes that are: Modular-ANN (Guh, 2007) and Ensemble-ANN (Yu and Xi, 2009). The related results can be referred to Tables 6.4 and 6.5, and Figures 6.3 and 6.4. ARL_1 and RA results are taken based on average values regardless the differences in pattern category (sources of variation in mean shift). The proposed scheme involves six sources of variation, Modular-ANN scheme involves eight sources of variation, whereas Ensemble-ANN scheme involves three sources of variation. Comparison of ARL_1 and RA results are limited to mean shifts within ± 1.00 ~ 3.00 standard deviations since the results for smaller shift (± 0.75 standard deviations) were not reported in Guh (2007), and Yu and Xi (2009). Comparison of ARL_0 is limited to low data correlation (0.0 ~ 0.1) since the results for moderate and high data correlations (0.5 ~ 0.9) were not reported in the previous works.

Based on monitoring aspect, the proposed scheme provided longer ARL_0 (335.01) as compared to the Modular-ANN scheme (ARL_0 = 198). This result is quite shorter as compared to the Ensemble-ANN scheme (ARL_0 = 364.82). For medium and high data correlations, the proposed scheme provided longer ARL_0 (ARL_0 = 543.93, 477. 45), satisfied the desired *de facto* level ($ARL_0 \geq 370$). This result describes that the proposed scheme is capable to maintain minimum false alarms when dealing with medium and highly correlated processes. Based on ARL_1 results, the proposed scheme showed shorter ARL_1 for overall magnitudes of mean shifts as compared to the ensemble-ANN scheme. This shows that the proposed scheme is more sensitive to detect bivariate process mean shifts, despite it is not statistically significant ($P = 0.117 > 0.05$). The proposed scheme also indicated a similar ARL_1 results when dealing with medium and highly correlated processes. These results, however, are quite longer (less effective) as compared to the Modular-ANN scheme.

Based on diagnosis aspect, the proposed scheme gave higher RA for overall magnitudes of mean shifts for low correlated process as compared to the ensemble ANN scheme. However, statistical significant test shows that overall increment in RA (3.36 %) is not significant ($P = 0.309 > 0.05$). Based on comparison against the Modular-ANN scheme, it can be observed that the proposed scheme is less accurate to classify the sources of variation when involving low and medium correlated processes. In opposite, the modular ANN scheme is less accurate when involving high correlated process. Statistical significant test shows that both scheme has a comparable diagnosis performance with fine difference of RA (0.24 %, $P = 0.879$).

Table 3: Performance comparison between EWMA-ANN and Baseline scheme

Pattern Category	Mean Shift		Average Run Lengths			Recognition Accuracy		
			Modular-ANN (Guh, 2007)	Ensem-ANN (Yu & Xi, 2009)	MEWMA-ANN	Modular-ANN (Guh, 2007)	Ensem-ANN (Yu & Xi, 2009)	MEWMA-ANN
	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	198	364.82	335.01, 543.93, 477.45	NA	NA	NA
			ARL$_1$ for ρ = 0.1, 0.5, 0.9					
US (1, 0)	0.75	0.00	NA	NA	17.60, 18.34, 20.00	89.5, 91.3, 93.0	NA	92.7, 90.4, 89.5
US (0, 1)	0.00	0.75			16.20, 15.99, 16.21			92.9, 89.3, 90.6
US (1, 1)	0.75	0.75			13.64, 13.28, 14.17			82.4, 94.8, 99.9
DS (1, 0)	−0.75	0.00			16.31, 16.43, 17.35			92.3, 89.2, 89.4
DS (0, 1)	0.00	−0.75			16.94, 17.44, 18.75			92.3, 87.8, 88.5
DS (1, 1)	−0.75	−0.75			13.46, 13.37, 14.03			84.1, 96.1, 99.9
Average					15.69, 15.81, 16.75			89.5, 91.3, 93.0
US (1, 0)	1.00	0.00		**10.76**	11.52, 11.57, 11.70	**93.3**, 94.4, 95.6		95.3, 93.1, 94.4
US (0, 1)	0.00	1.00			10.50, 10.22, 10.20			95.8, 93.5, 94.4
US (1, 1)	1.00	1.00			9.16, 9.09, 9.66			90.0, 96.5, 100
DS (1, 0)	−1.00	0.00			10.99, 10.86, 11.06			95.3, 93.2, 92.3
DS (0, 1)	0.00	−1.00			11.08, 11.12, 11.36			93.8, 92.1, 92.6
DS (1, 1)	−1.00	−1.00			9.15, 9.12, 9.63			89.5, 98.0, 100
Average					10.40, 10.33, 10.60			93.3, 94.4, 95.6
US (1, 0)	1.50	0.00	3.28, 3.39, 3.53	**6.83**	7.02, 7.07, 7.03	**95.5**, 97.0, 97.4	99.6, 99.1, 88.8	97.4, 96.5, 97.1
US (0, 1)	0.00	1.50			6.54, 6.33, 6.40			97.1, 96.5, 96.2
US (1, 1)	1.50	1.50			5.82, 5.73, 5.94			91.7, 97.9, 100
DS (1, 0)	−1.50	0.00			6.81, 6.81, 6.92			97.4, 96.3, 95.5
DS (0, 1)	0.00	−1.50			6.82, 6.80, 6.85			96.2, 95.8, 95.6
DS (1, 1)	−1.50	−1.50			5.81, 5.69, 5.98			93.2, 99.0, 100
Average					6.47, 6.41, 6.52			95.5, 97.0, 97.4
US (1, 0)	2.00	0.00	2.15, 2.35, 2.19	**5.14**	5.23, 5.15, 5.19	**95.7**, 97.6, 97.8	99.6, 99.7, 94.0	97.8, 97.1, 97.6
US (0, 1)	0.00	2.00			4.80, 4.72, 4.70			97.7, 97.8, 97.1
US (1, 1)	2.00	2.00			4.36, 4.32, 4.39			91.6, 98.4, 100
DS (1, 0)	−2.00	0.00			5.04, 5.04, 5.02			96.8, 96.7, 96.6
DS (0, 1)	0.00	−2.00			4.97, 5.03, 4.98			96.5, 96.5, 95.6
DS (1, 1)	−2.00	−2.00			4.29, 4.27, 4.33			93.7, 98.9, 100
Average					4.78, 4.76, 4.77			95.7, 97.6, 97.8
US (1, 0)	2.50	0.00		**5.74**	4.10, 4.14, 4.12	**96.2**, 97.8, 98.2		98.0, 98.4, 98.0
US (0, 1)	0.00	2.50			3.83, 3.81, 3.81			97.3, 97.4, 97.0
US (1, 1)	2.50	2.50			3.54, 3.49, 3.53			93.2, 98.4, 100
DS (1, 0)	−2.50	0.00			3.99, 3.96, 3.95			97.3, 97.3, 97.0
DS (0, 1)	0.00	−2.50			3.97, 4.02, 3.98			96.5, 96.6, 97.0
DS (1, 1)	−2.50	−2.50			3.41, 3.40, 3.46			94.9, 98.8, 100
Average					3.81, 3.80, 3.81			96.2, 97.8, 98.2
US (1, 0)	3.00	0.00	2.41, 2.56, 2.47	**7.02**	3.47, 3.46, 3.46	**96.6**, 98.0, 98.5	99.3, 99.3, 92.5	98.6, 98.3, 98.2
US (0, 1)	0.00	3.00			3.20, 3.20, 3.21			97.8, 97.8, 98.0
US (1, 1)	3.00	3.00			2.98, 2.93, 2.98			93.8, 98.4, 100
DS (1, 0)	−3.00	0.00			3.31, 3.30, 3.27			98.0, 97.1, 97.6
DS (0, 1)	0.00	−3.00			3.33, 3.32, 3.32			96.7, 97.1, 97.1
DS (1, 1)	−3.00	−3.00			2.84, 2.85, 2.90			94.6, 99.1, 100
Average					3.19, 3.18, 3.19			96.6, 98.0, 98.5
Grand Average	± (0.75 ~ 3.00)				7.39, 7.38, 7.61			94.5, 96.0, 96.8

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

Bold values represent the comparison results between EWMA-ANN and Ensemble-ANN schemes

Table 4: Statistical significant test of performance results in Table 3

<table>
<tr><th>Performance Measures (PM)</th><th>Result of Paired T-Test
Mean difference of PM</th><th>Remarks</th></tr>
<tr><td rowspan="2">ARL_1</td><td>N Mean StDev SE Mean
MEWMA-ANN 9 4.808 1.421 0.474
Modular-ANN 9 2.703 0.541 0.180
Difference 9 2.104 1.077 0.359

Mean difference of ARL_1
95% CI: (1.277, 2.932)
T-Test = 0 (vs ≠ 0): T = 5.860, P = 0.000</td><td>Increment in ARL_1 is proven to be statistically significant.</td></tr>
<tr><td>N Mean StDev SE Mean
MEWMA-ANN 5 5.730 2.890 1.292
Ensemble-ANN 5 7.098 2.189 0.979
Difference 5 −1.368 1.535 0.686

Mean difference of ARL_1
95% CI: (−3.274, 0.538)
T-Test = 0 (vs ≠ 0): T = −1.990, P = 0.117</td><td>Decrement in ARL_1 is not statistically significant.</td></tr>
<tr><td rowspan="2">RA</td><td>N Mean StDev SE Mean
MEWMA-ANN 9 97.122 1.023 0.341
Modular-ANN 9 96.878 4.064 1.355
Difference 9 0.244 4.659 1.553

Mean difference of RA:
95% CI: (−3.337, 3.826)
T-Test = 0 (vs ≠ 0): T = 0.160, P = 0.879</td><td>Increment in RA is too small (comparable to each other).</td></tr>
<tr><td>N Mean StDev SE Mean
MEWMA-ANN 5 95.460 1.282 0.573
Ensemble-ANN 5 92.100 7.654 3.423
Difference 5 3.360 6.454 2.886

Mean difference of RA:
95% CI: (−4.654, 11.374)
T-Test = 0 (vs ≠ 0): T = 1.160, P = 0.309</td><td>Increment in RA is not statistically significant.</td></tr>
</table>

4. Conclusions

This paper proposed an integrated MEWMA-ANN scheme towards achieving ‘balance 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.

Acknowledgements

The authors would like to thank Universiti Tun Hussein Onn Malaysia (UTHM), Universiti Teknologi Malaysia (UTM), and Ministry of Higher Education (MOHE) of Malaysia who sponsoring this work.

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 χ^2 chart signals. Computers and Industrial Engineering 47: 195–205.

Cheng CS, Cheng HP (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 TT, El-Baz MA, Abd-Elwahed MS (2010) A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks. Expert Systems with Applications 37: 1035–1042.

Guh RS (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.

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

Law AM and Kelton WD, Simulation modeling and Analysis, 3rd ed., Mc Graw-Hill, 2000.

Lehman RS (1977) Computer simulation and modeling: an introduction. Lawrence Erlbaum, London.

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

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

Mason RL, Tracy ND, Young JC (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.

Masood I, Hassan A (2009) Synergistic-ANN recognizers for monitoring and diagnosis of multivariate process shift patterns. International Conference of Soft Computing and Pattern Recognition: 266–271.

Montgomery DC (2005) Introduction to statistical quality control. John Wiley & Sons, Inc., New Jersey.

Niaki STA, Abbasi B (2005) Fault diagnosis in multivariate control charts using artificial neural networks. Quality and Reliability Engineering International 21: 825–840.

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

Yu JB, Xi LF (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 JB, Xi LF, Zhou XJ (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 JDT, O’Brien C (2003) Using novelty detection to identify abnormalities caused by mean shifts in bivariate processes. Computers and Industrial Engineering 44: 385–408.