MULTIVARIATE QUALITY CONTROL USING AN INTEGRATED ARTIFICIAL NEURAL NETWORK SCHEME: A CASE STUDY IN PLASTIC INJECTION MOLDING INDUSTRY

AFIZ AZRY MOHD KHAIRY

Tesis ini dikemukakan sebagai
memenuhi syarat penganugerahan
Ijazah Sarjana

Fakulti Kejuruteraan Mekanikal dan Pembuatan
Universiti Tun Hussein Onn Malaysia

MAY 2013

ABSTRACT

Manufacturers using traditional process control charts to monitor their processes often encounter out-of-control signals indicating that the process mean has changed. Most manufacturers are unaware how much these changes in the mean inflate the variance in the process output. In actual, many manufacturing processes involve two or more dependent variables and attempting to monitor such variables separately using univariate SPC charting scheme would increase false alarms and leading to wrong decision making The problem becomes more complicated when dealing with small mean shift particularly in identifying the causable variables. In this project study, advances SPC scheme which applied Artificial Neural Networks that were designed to enable balanced monitoring and accurate diagnosis were used. Its performance and effectiveness in actual manufacturing practice were compared with common use traditional process control chart. Thermoplastic Injection Molding was selected in this study since most of the industries applied this method to produce low cost parts. It is important to ensure that good quality parts can be produced according to requirement. The potential benefit from advance SPC schemes was it able to performed rapid detection of process disturbance. However the accuracy of mean shifted diagnosis performance need to improve.

ABSTRAK

Pengeluar yang menggunakan carta kawalan proses tradisional untuk memantau proses mereka sering menghadapi isyarat luar kawalan yang menunjukkan bahawa min proses telah berubah. Kebanyakan pengeluar mungkin tidak menyedari berapa banyak pertambahan perubahan dalam min pembolehubah semasa proses pengeluaran. Dalam persekitaran sebenar, proses pembuatan banyak melibatkan dua atau lebih pembolehubah dan cuba untuk memantau pembolehubah tersebut secara berasingan menggunakan skema kawalan proses statistik satu pembolehubah .Tetapi ini akan meningkatkan penggera palsu dan membawa kepada masalah iaitu membuat keputusan yang salah .Ini akan menjadi lebih rumit apabila berhadapan dengan perubahan min kecil dalam mengenal pasti punca pembolehubah. Dalam kajian projek ini, skema maju kawalan proses statistik yang direka untuk membolehkan pemantauan seimbang dan diagnosis yang tepat telah digunakan. Prestasi dan keberkesanan dalam amalan pembuatan sebenar dibandingkan dengan penggunaan biasa tradisional carta kawalan proses. Acuan suntikan termoplastik telah dipilih dalam kajian ini kerana kebanyakan industri menggunakan kaedah ini untuk menghasilkan komponen kos rendah. Ia adalah penting untuk memastikan bahawa komponen-komponen berkualiti baik boleh dihasilkan mengikut ketetapan. Potensi manfaat daripada skema SPC adalah ia mampu mengesan dengan cepat gangguan proses. Walau bagaimanapun ketepatan purata berubah prestasi diagnosis perlu di tingkatkan.

TABLE OF CONTENTS

LIST OF TABLE

LIST OF FIGURE

LIST OF SYMBOL

α	-	Type I error (α risk)
β	-	Type II error (β risk)
λ	-	Constant parameter for EWMA control chart
ρ	-	Correlation coefficient for bivariate samples
μ	-	Mean
σ	-	Standard deviation
μ_0	-	Mean for in-control samples
σ_0	-	Standard deviation for in-control samples
σ_{ij}	-	Covariance for bivariate samples
Σ	-	Covariance matrix for bivariate samples or basic summation
X_t	-	Original observation samples at time/point t
Z_t	-	Standardized observation samples at time/point t
N	-	Random normal variates
$\bar{x}$	-	Sample mean
X^2	-	Hotelling's generalized distance
S	-	Hotelling's covariance of sample
H_0	-	Null Hypothesis

LIST OF APPENDIX

CHAPTER 1

INTRODUCTION

1.1 Project Background

In Industries, statistical analysis plays an important role as a tool to make in advance judgment for any decision and action that related to business performance. In the manufacturing industries, process variation exists as a main factor that affects the quality of product. Regardless to any kind of process, design specification and tolerance is created to control the product quality. A lot of efforts have been taken to ensure the process variable condition falls within the requirement. Monitoring and controlling quality costs are becoming critical activities of quality improvement programs. Precise and reliable statistical tools are most preferred by the industries for process monitoring. In order to maintain and improve the quality, effort towards minimizing process variation in manufacturing environment has become an important issue in quality control. Statistical quality engineering (SQE) tools have been developed for systematically reducing variability in the key process variables or quality characteristics of the product (Montgomery, 2009). Statistical process control (SPC) charting is one of the SQE tools that useful for monitoring and diagnosing process variation. Recent SQE tools, Integrated Multivariate Exponentially Weighted Moving Average – Artificial Neural Networks (MEWMA-ANN) scheme has been developed for enabling balanced monitoring and accurate diagnosis of multivariate process variation. Since the initial study focused on development of the scheme programme, further verification in actual industrial environment are required to validate its effectiveness in fault diagnosis towards continuous quality improvement.

1.2 Problem Statements

To produce product that meet the specification requirement is the important point that lead them to control and monitor the quality in the process. In manufacturing process, quality monitoring is practice by sampling check on the bulk lots which produced at specific interval of in control process running time according to standard specification. Easy to implement, less investment and applicable to any kind of environment are the reasons why traditional SPC method is common in most of the manufacturing industries. The control charts in traditional SPC are designed to monitor a single product with large production runs. But imbalance in monitoring and diagnosis of multivariate process variation, most of the problems cannot be detected in advance that lead to poor Quality control in manufacturing , rework, scrap, product failures and recalls can severely damage the company business. In addition, process variation from 4Ms (Man, Machine, Material, and Method) contribute to the process quality. The research activity in multivariate control charts has been reported to be at its highest level for the past decade which reflects increased measurement and computing ability (Woodall & Montgomery, 1999). Large and diversified research areas on the application of multivariate control charts in manufacturing areas are extensively discussed in many literatures (Tracy, Mason and young (1992), Lowry and Montgomery (1995), Mason and young (2002) and the references therein).

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 (Masood, 2012). The Integrated MSPC-ANN schemes able to monitor and diagnose jointly, such advances are ultimately aim to minimize human intervention through computerized decision making. The matters that need to be considered is whether this scheme is effective or applicable to actual environment in manufacturing process need to be further investigated.

1.3 Project Objectives

The objectives of this project are:

i. To investigate the effectiveness of the Integrated MEWMA-ANN scheme in monitoring and diagnosing process faults in manufacturing industry
ii. To recognize problems or drawbacks arise from Integrated MEWMA-ANN scheme and discuss the proposal for improvement of scheme.

1.4 Project Scopes

The scopes of this project are:

i. The application of Univariate Statistical Process Control Chart which are Individual Chart and Exponentially Weighted Moving Average (EWMA) Charts
ii. The application of Multivariate Statistical Process Control Chart which are Hotelling T Square Chart and Multivariate Exponentially Weighted Moving Average (EWMA) Charts
iii. The application of Advance Statistical Process Control chart which are Baseline artificial neural network (ANN) pattern recognitions scheme, Statistical Features artificial neural network (ANN) pattern recognitions scheme and Integrated MEWMA-ANN artificial neural network (ANN) pattern recognition scheme
iv. Thermo plastic part fabricated from Plastic Injection Process machine which the dimensions and weight were differed base on the changes in machine parameter and the usage of virgin plastic resin mixed with 20% and 50% of regrind plastic resin.

1.5 Dissertation Outline

This dissertation is organized as follows:

Chapter I

This chapter gives an introduction to Statistical Control Chart (SPC) and its function especially on monitoring process quality and project background including problem statement, project objectives and project scopes.

Chapter II

This chapter is a review of SPC fundamental knowledge describing on the function and application of each type control chart by previous researcher and practitioner.

Chapter III

This chapter elaborates the methodology used in this project.

Chapter IV

This chapter presents the result of performance of recognizing process out of control by Univariate SPC, Multivariate SPC and Advance Multivariate SPC. The diagnosing accuracy of MEWMA-ANN scheme, Statistical Features scheme and Baseline scheme were compared.

Chapter V

Project summary, conclusion and recommendation for future work are described in this chapter.

CHAPTER 2

LITERATURE REVIEW

2.1 Statistical process control (SPC)

Variation in manufacturing process environment causes no parts or products can be produced in exactly the same size and properties. Process variation can be influenced from **chance causes (random error)** and/or **assignable causes (systematic errors)** (Montgomery, 2009). There are many ways to implement process control. Key monitoring and investigating tools include:

i. Histograms
ii. Check Sheets
iii. Pareto Charts
iv. Cause and Effect Diagrams
v. Defect Concentration Diagrams
vi. Scatter Diagrams
vii. Control Charts

According to NIST/SEMATECH e-Handbook of Statistical Methods, Statistical process control (SPC) charting is one of the Statistical quality engineering (SQE) tools that useful for monitoring and diagnosing process variation. SPC charting schemes focused on heuristic, smaller shift detection, process pattern identification and automated pattern recognition. . A control chart is the primary tool of SPC and is basically used to monitor the process characteristics, e.g., the process mean and process variability [Duncan, 1988, Montgomery, 2001]. The most common types of variable control charts for variables include: (1) Average and Range (X bar and R) Charts (2) Average and Standard Deviation ($\overline{X}$ and S) Charts (3) Individual and Moving Range (X and MR) Charts.

The underlying concept of statistical process control is based on a comparison of what is happening today with what happened previously. Stated differently, we use historical data to compute the initial control limits. Then the data are compared against these initial limits. Points that fall outside of the limits are investigated and, perhaps, some will later be discarded. Process capability compares the output of an in-control process to the specification limits by using capability indices. The comparison is made by forming the ratio of the spread between the process specifications (the specification "width") to the spread of the process values, as measured by 6 process standard deviation units (the process width").The definition of the C_p given in Equation (2.1) implicitly assumes that the process is centered at the nominal dimension.

$$C_p = \frac{USL - LSL}{6\sigma} \tag{2.1}$$

If the process is running of center, it actual capability will be less than indicated by the C_p. It is convenient to think of C_p as a measure of potential capability, that is, capability with centered process. If process is not centered, a measure of actual capability is often used. This ratio is called C_{pk} as defined in Equation (2.2).

$$C_{pk} = \min\left[\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right] \tag{2.2}$$

In effect, C_{pk} is a one-sided Process Capability Ratio (PCR) that is calculated relative to the specification limit nearest to the process mean. Montgomery (2006) provides guidelines on appropriate values of the PCR and a relating fallout for a normally distributed process in statistical control to the value of C_p. Many big company use C_p =1.33 as a minimum acceptable target and C_p =1.66 as a minimum target for strength ,safety or critical characteristic, Some company require that internal processes and those at suppliers achieve a C_{pk} =2.0.Figure 2.1 shows various process width.

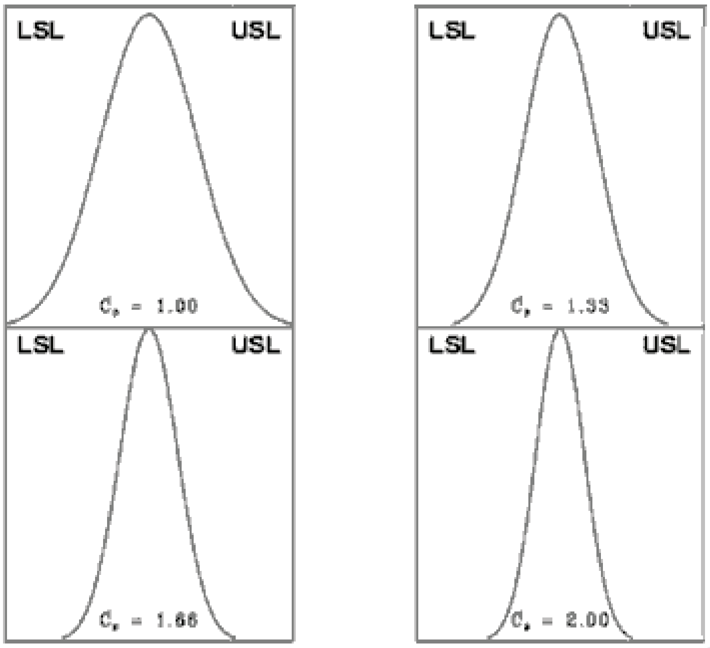

Figure 2.1: Cp condition for varying process widths

The aim of statistical process control (SPC) is to achieve higher quality of final product and lower the production loss due to defect product. Process monitoring with control chart is a basic tool of statistical process control. It monitors the behavior of a production process and signals the operator to take necessary action when abnormal event occurs. A stable production process is the key element of quality improvement. Depending on the number of process characteristics to be monitored, there are two basic types of control charts, Univariate Control Chart and Multivariate Control Chart .

2.2 Univariate Control Chart

It is a graphical display (chart) of one quality characteristic. There are few types of Control Charts that have been developed. Shewhart charts are sensitive to large process shifts. The probability of detecting small mean shifts fast is rather small. The CUSUM (Cumulative Sum) chart is very effective for small shifts. Disadvantages CUSUM is relatively slow to respond to large shifts .Also, special patterns are hard to see and analyze. The Exponentially Weighted Moving Average (EWMA) is a statistic for monitoring the process that averages the data in a way that gives less and less weight to data as they are further removed in time. In contrast, attempting to monitor such variables separately using univariate SPC charting scheme would increase false alarms and leading to wrong decision making. However, monitoring each process variable with separate Shewhart control chart ignores the correlation between variables and does not fully reflect the real process situation. Nowadays, the process industry has become more complex than it was in the past and inevitably that number of process variables need to be monitored has increased dramatically. Thus only monitor a single parameter or output at a time. Therefore they cannot detect changes in the relationship between multiple parameters. Very often, these variables are multivariate in nature and using Shewhart control charts becomes insufficient.

2.2.1 Shewhart Control Chart

The most common use method in current industries is Control Chart or Shewhart Charts. These control charts are constructed by plotting product's quality variable over time in sequence plot as shown in Figure 2.2. A control chart contains a **center line, an upper control limit and a lower control limit.** Points that plots within the control limits indicates the process is in control. In this condition no action is necessary. Points that plots outside the control limits is evidence that the process is out of control. In this condition, investigation and corrective action are required to find and eliminate assignable cause(s) (Mendenhall & Sincich, 2007).

Let w be a sample statistic that measure some quality characteristic of interest and suppose that the mean of w is μ_w and the standard deviation of w is σ_w. Then the center line, upper control limit and lower control limit as shows in equation (2.3).

$$\begin{aligned} UCL &= \mu_W + L\sigma_W \\ Center\ Line &= \mu_W \\ LCL &= \mu_W - L\sigma_W \end{aligned} \qquad (2.3)$$

Where L is "the distance" of the control limits from center line, expressed in standard deviation units.

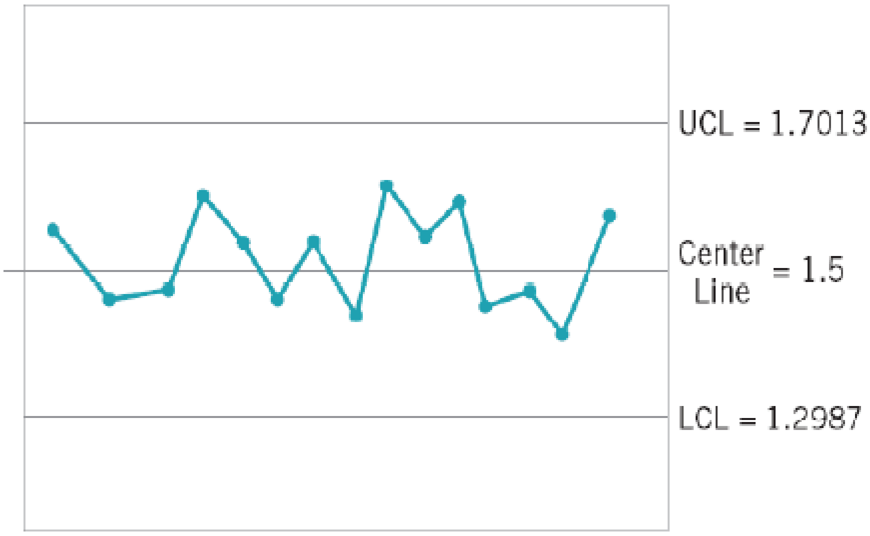

Figure 2.2: How the Control Chart works

2.2.2 Control Limits

A point falling within the control limits means it fails to reject the null hypothesis that the process is statistically in-control, and a point falling outside the control limits means it rejects the null hypothesis that the process is statistically in-control. Therefore, the statistical Type I error α (Rejecting the null hypothesis H_0 when it is true) applied in Shewhart control chart means the process is concluded as out-of-control when it is truly in-control. Same analog, the statistical Type II error β (Failing to reject the null hypothesis when it is false) means the process is concluded as in-control when it is truly false.

2.2.3 Average Run Length (ARL)

The performance of control charts can also be characterized by their average run length. Average run length is the average number of points that must be plotted before a point indicates an out-of-control condition (Montgomery, 1985). We can calculate the average run length for any Shewhart control chart according to,

$$ARL = \frac{1}{p} \tag{2.4}$$

Where p or Type I error is the probability that an out-of-control event occurs. Therefore, a control chart with 3 sigma control limits, the average run length will be

$$ARL = \frac{1}{p} = \frac{1}{0.0027} = 370$$

This means that if the process remains in-control, in average, there will be one false alarm every 370 samples.

2.2.4 Patterns of Process Behavior

Apart from all the measurement should fall with the control limits, the process can be viewed as in-control when there is no systematic pattern shown in the process behavior. Systematic patterns occurring in Shewhart control charts have often been interpreted as indicators of extraneous sources of process variation (Mason, et al. 2003). The process will be improved if the causes of systematic pattern in the process are diagnosed and further eliminated. A statistically in-control process can be indicated by normal pattern, whereas out-of-control process can be indicated by abnormal patterns (upward shift, downward shift, upward trend, downward trend, cyclic, systematic, stratification, and mixture). Figure 2.3 shows in practice, sudden shifts patterns (upward or downward shift) commonly indicate there are changes in material, operator or machine. Trends patterns (upward or downward trend) indicate there are wears and tears in cutting tools. Cyclic patterns indicate there are voltage fluctuations in power supply (Chen *et al.*, 2007).

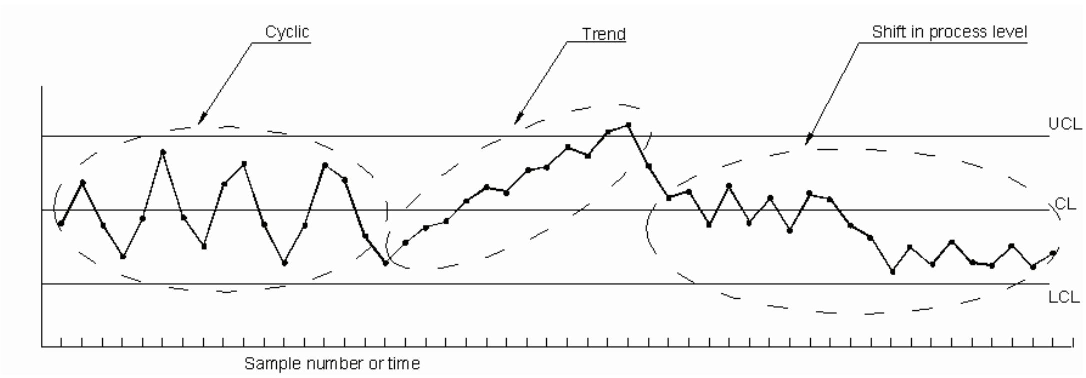

Figure 2.3: Trend of sudden shifts patterns

The Western Electric Handbook (1956) provides a set of guidelines to detect the systematic patterns in the process. A brief summary is shown below. A process is considered as out-of-control if any of the following conditions holds:

i. One point falls outside the 3-sigma control limits (beyond Zone A).
ii. At least two out of three consecutive points fall on the same side of the centerline, and are beyond the 2-sigma control limits (in Zone A or beyond).
iii. At least four out of five consecutive points fall on the same side of the center line and are beyond the 1-sigma limits (in Zone B or beyond).
iv. At least eight successive points fall on the same side of the center line.

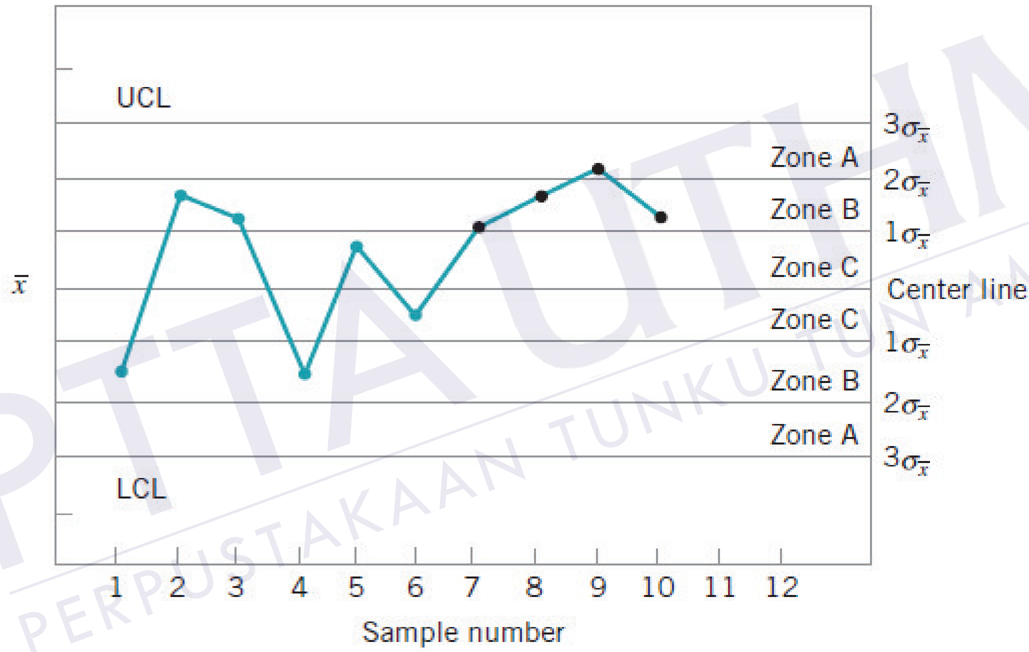

Figure 2.4:The Western Electric run rules.

2.2.5 Control Chart for Mean $\bar{x}$- Chart

The concept of $\bar{x}$- Chart is taken from the Shewhart Chart with Center Line, UCL and LCL. It is applied to monitor a quantitative quality characteristic base on random samples of several units of the product rather than on the characteristic of individual industrial units (Mendenhall & Sincich, 2007). The equation of Center Line and Control Limits are mentioned in equation (2.5)

$$Center\ Line = \overline{\overline{x}} = \sum_{i=1}^{k} \overline{x}_i$$
$$UCL = \overline{\overline{x}} + A_2\overline{R} \qquad (2.5)$$
$$UCL = \overline{\overline{x}} - A_2\overline{R}$$

Where

k = Number of samples, each of size lot

$\bar{x}$i = Sample mean for the ith sample

Ri = Range of ith sample

$$\overline{R} == \frac{\sum_{i=1}^{k} R_i}{k}$$

A2 is given in table of Appendix B

2.2.6 Control Chart for Process Variation R- Chart

In Quality Control variability value of some quality characteristic can be controlled. An increase in the process standard deviation σ means that the quality characteristic variable will vary over a wider range, thereby increasing the probability of producing an inferior product (Mendenhall & Sincich, 2007). The variation in quality characteristic is monitor using a **range chart** or **R-chart**. The location of Center Line and Control Limits for R-chart are mentioned in equation (2.6)

$$Center\ line = \overline{R}$$
$$UCL = D_4\overline{R} \qquad (2.6)$$
$$UCL = D_3\overline{R}$$

Where

k = Number of samples, each of size lot

Ri = Range of ith sample

$$\overline{R} == \frac{\sum_{i=1}^{k} R_i}{k}$$

and D3 and D4 are given in Table in Appendix B for n =2 to n =25

In normal practice, : $\bar{x}$- Chart and R-Chart are shows together as Figure 2.5

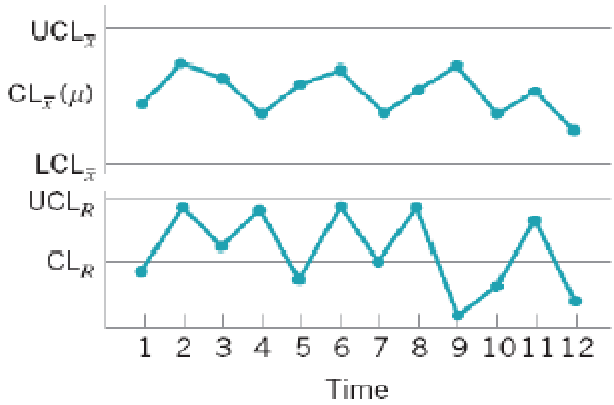

Figure 2.5: $\bar{x}$*-Chart (above) and R-Chart (below)*

2.2.7 Individuals Control Charts

The individuals control chart examines variation in individual sample results over time as shown in Figure 2.6. While rational subgrouping does not apply, thought must be given to when the results will be measured. If the process is in statistical control, the average on the individuals chart is our estimate of the population average. The average range will be used to estimate the population standard deviation. For individual measurements, e.g., the sample size = 1, use the *moving range* of two successive observations to measure the process variability. The moving range is defined as in equation (2.7)

$$MR_i = |x_i - x_{i-1}| \tag{2.7}$$

Which is the absolute value of the first difference (e.g., the difference between two consecutive data points) of the data. Analogous to the Shewhart control chart, one can plot both the data (which are the individuals) and the moving range. For the control chart for individual measurements, the lines plotted are:

$$\begin{aligned} UCL &= \bar{x} + 3\frac{\overline{MR}}{1.128} \\ Center\ Line &= \bar{x} \\ LCL &= \bar{x} - 3\frac{\overline{MR}}{1.128} \end{aligned} \tag{2.8}$$

Where is the average of all the individuals and is the average of all the moving ranges of two observations. Keep in mind that either or both averages may be replaced by a standard or target, if available. (Note that 1.128 is the value of *d*2 for *n* = 2) in Appendix B.

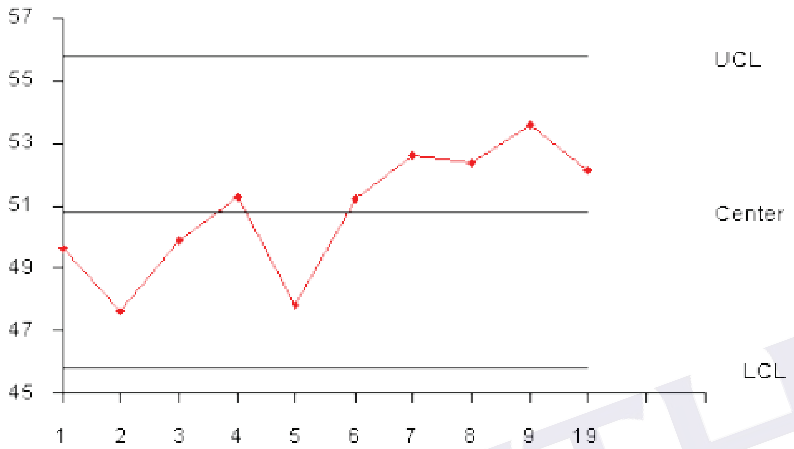

Figure 2.6: Individual Chart shows the process is in control, since none of the plotted points fall outside either the UCL or LCL.

2.3 Exponential Weighted Moving Average (EWMA) Chart

The Exponentially Weighted Moving Average (EWMA) control chart (Figure 2.7) is one of the control charts used to detect the occurrence of a shift in a process mean compared to widely used Shewhart control chart (Montgomery, 2009). The EWMA is used extensively in time series modelling and in forecasting.Upper Control Limit (UCL), Control Limit (CL) and Lower Control Limit (LCL) for the EWMA control chart are given below in the equation form respectively. According to Sharaf El-Din (2006) in [Statistical Process Control Charts Applied to Steelmaking Quality Improvement] values of λ in the interval $0.05 \leq \lambda \leq 0.25$ work well in practice, with $\lambda = 0.05$, $\lambda = 0.10$, and $\lambda = 0.20$ being popular choices. A good rule of thumb is to use smaller values of λ to detect smaller shifts. The width of control limits L = 3 (the usual 3-sigma limits) works reasonably well, particularly with the larger value of λ, although when λ is small, say $\lambda \leq 0.10$, there is an advantage in reducing the width of the limits by using a value of L between about 2.60 and 2.80, (Montgomery ,1997).

$$
\begin{aligned}
&\bar{x}_i(n) = \frac{1}{n}x_i + \frac{1}{n}x_{i-1} + \frac{1}{n}x_{i-2} + \ldots + \frac{1}{n}x_{i-n-1} \qquad n = Moving\ average\\
&Z_i = \lambda\bar{x}_i + (1-\lambda)Z_{i-1},\\
&Z_0 = \mu_0 \qquad (2.9)\\
&Center\ Line = \mu_0\\
&UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda\left[1-(1-\lambda)^{2i}\right]}{(2-\lambda)}}
\end{aligned}
$$

Where

$\lambda = 0.1$

σ = standard deviation of the object xi

L = width of the control limits

Zi = Weighted average of all previous sample means.

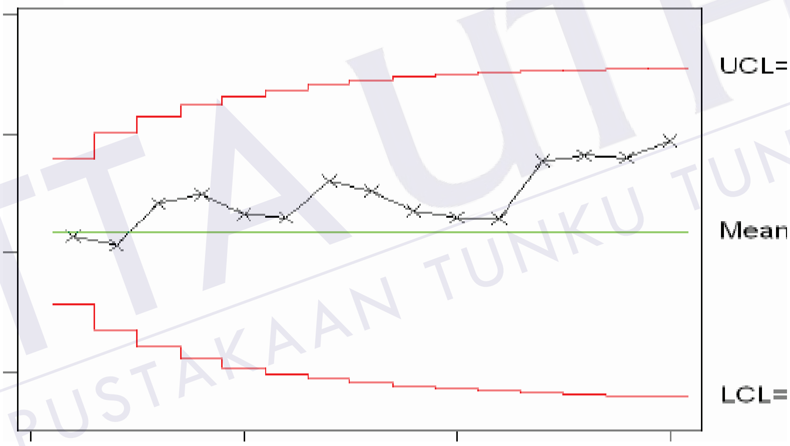

Figure 2.7: EWMA Chart

2.4 Multivariate Control Chart

It is a graphical display of a statistic that summarizes or represents more than one quality characteristic. In the related study, many manufacturing processes involve two or more dependent variables, whereby an appropriate scheme is required to monitor and diagnose such variables jointly. This joint monitoring-diagnosis concept is called multivariate quality control (MQC). The T^2 control chart (Hotelling, 1947) that is developed based on logical extension of univariate SPC chart (Shewhart control chart) was claimed as an original work in MSPC. Nevertheless, it was found to be effective only for detecting mean shift in large magnitudes ($\geq$ 1.5 standard deviations). In order to improve capability for detecting mean shift in smaller magnitudes (< 1.5 standard deviations), the multivariate cumulative sum (MCUSUM) (Crosier, 1988; Pignatiello and Runger, 1990) and the multivariate exponentially weighted moving average (MEWMA) (Lowry et al., 1992; Prabhu and Runger, 1997) control charts were developed based on logical extension of univariate CUSUM and EWMA control charts respectively. These multivariate control charts are commonly known as the traditional MSPC charting schemes. Aparisi and Haro (2001) proposed the T^2 control chart for variable sampling interval (VSI) to improve sensitivity in detecting mean shifts. Khoo and Quah (2003) developed a multivariate control chart for monitoring shifts in the covariance matrix based on individual observations. Alwan and Alwan (1994), Apley and Tsung (2002), and Jiang (2004) investigated the application of T^2 control chart for monitoring mean shifts in univariate autocorrelated processes. Ngai and Zhang (2001) proposed the MCUSUM control chart based on projection pursuit to deal with a specific situation, that is, the process mean is already shifted at the time the control charting begins. The traditional MSPC charting schemes are only effective for monitoring (detecting) mean shifts but they are unable to diagnose (identify) the sources of variation in 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.

Basically, the MSPC charting schemes for variables can be categorized to: (i) statistical design, and (ii) economical design. Both design categories can be further classified to fixed sampling interval (FSI) and variable sampling interval (VSI). The traditional MSPC charting schemes that are T^2, MCUSUM and MEWMA control charts were designed based on statistical consideration for dealing with FSI.

2.4.1 Hotelling's T^2 Chart

Hotelling H. (1931) can be viewed as the originator of multivariate control charts. Hotelling proposed a concept of generalized distance between a new observation to its sample mean. In case of bivariate condition by assuming these x_1 and x_2 are distributed according the bivariate normal distribution. Referring to Figure 2.8, say $\overline{X}_1$ and $\overline{X}_2$ are the mean, s_1 and s_2 are the standard deviation of these two variables respectively.

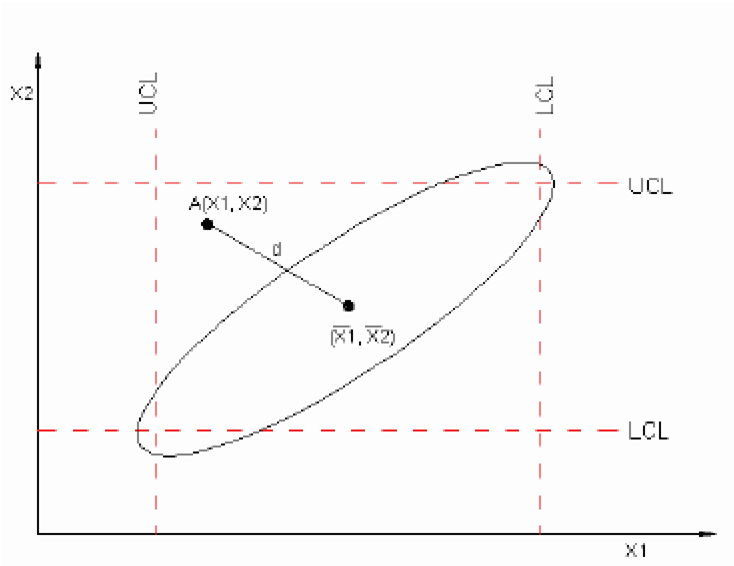

Figure 2.8: A generic bivariate Hotelling's T^2 control region

The covariance s_{12} is used to estimate the dependency between x_1 and x_2. The generalized distance between point A and its mean can be calculated as:

$$X_0^2 = \frac{1}{s_{11}s_{22} - s_{22}^2}\left[s_{11}(x_2 - \bar{x}_2)^2 - 2(x_2 - \bar{x}_2)(x_1 - \bar{x}_1) + s_{22}(x_1 - \bar{x}_1)^2\right] \qquad (2.10)$$

This statistic follows the Chi-square distribution with two degrees of freedom. An ellipse can be graphed with the x_1 and x_2 in this equation. Moreover, all the points lying on the ellipse will generate the same Chi-square statistic. As a consequence, every observation can be determined whether its generalized distance exceeds the ellipse by comparing X_0^2 and $X^2_{2,\alpha}$,where $X^2_{2,\alpha}$ is the upper α percentage point of the Chi-square distribution with 2 degrees of freedom. The observation will be considered as out-of-control if $X_0^2 > X^2_{2,\alpha}$

With the same concept of the generalized distance, it can be extended from bivariate to a multiple *p* variables. Let $X_i = (X_{i1}, X_{i2}, \ldots . X_{ip})$ represent a *p* dimensional vector of measurements made on a process time period i. The value X_{ij} represents an observation on the j^{th} characteristic. Assuming that when the process is in control, the X_i are independent and follow a multivariate normal distribution with mean vector μ and covariance matrix Σ. Normally μ and Σ are unknown, but we can use $\bar{X}$ and S estimated from a historical data set with *n* observations.

Phase I and Phase II

The application of Hotelling's T^2 statistic shall be categorized into two phases. Phase I tests whether the preliminary process was in control and phase II tests whether the future observation remains in-control (Alt, 1985). Phase I operation refers to the construction of in-control data set. Same idea as Shewhart control chart, control limits are estimated from a period of in-control data.

To obtain this in-control data, the raw data set needs to be purged. For instance, the outliers need same idea as Shewhart control chart, control limits are estimated from a period of in-control data. To obtain this in-control data, the raw data set needs to be purged. For instance, the outliers need to be removed and the missing data needs to be substituted with an estimate. During phase I operation, Hotelling's T^2 statistic is calculated for each measurement and compared to the control limit, which will follows Chi-square distribution (according to Richard, A.J. & Dean, W.W., 2002.)

$$T^2 = (X_i - \overline{X})S^{-1}(X_i - \overline{X}) \sim X^2_{\alpha,p}(Chi - square\ distribution) \qquad (2.11)$$

Also other research shows that the control limit follows Beta distribution (Mason, Young & Tracy,1992).

$$T^2 = (X_i - \overline{X})S^{-1}(X_i - \overline{X}) \sim \frac{(n-1)^2}{n} B_{(\alpha,\frac{p}{2},\frac{n-p-1}{2})} \qquad (2.12)$$

n=number of preliminary observations

Both control limits will be approximate when the number of observations is large. The control limit based on Chi-square distribution is established on the assumption that $\overline{X}$ and S are true values μ and Σ, which is just an approximate situation (Mason, Young & Tracy, 1992). Beta distribution is more precise and is a recommendable choice. After purging the raw data with Hotelling's T^2 statistic, the in-control data set is ready for monitoring future observations which is termed as phase II operation. The control limit for determining future observation is different from the one in phase I. It follows an F distribution with *p* and (*n*-*p*) degrees of freedom.

$$T^2 = (X_i - \overline{X})S^{-1}(X_i - \overline{X}) \sim \frac{p(n+1)(n-1)}{n(n-p)} F_{(p,n-p,\alpha)} \qquad (2.13)$$

Where sample mean is and the covariance of sample

$$S = \begin{array}{cccc} S_{11} & S_{12} & S_{13} & S_{p1} \\ & S_{22} & S_{23} & S_{2p} \\ & . & & \\ & & . & S_{pp} \end{array}$$

The idea of using Hotelling's T^2 statistic in phase I and phase II is the same. Each measurement is examined whether it is out-of-control by checking if it deviates extraordinarily from its sample mean. It should be reminded to choose the correct upper control limit on different purposes.

The Hotelling's T^2 statistic can be extended for more than two variables. Instead of a 2-dimensional ellipse control region, the result will be presented in a similar way as Shewhart control chart. The T^2 statistics calculated from all the observation will be plotted in a chart against time or observation serious and compared to the upper control limit. Figure 2.9 is a generic T^2 control chart. It should be noticed that there is no center line and the lower control limit is set to zero, because the meaning of T^2 statistic is a generalized distance between the observation and its sample mean.

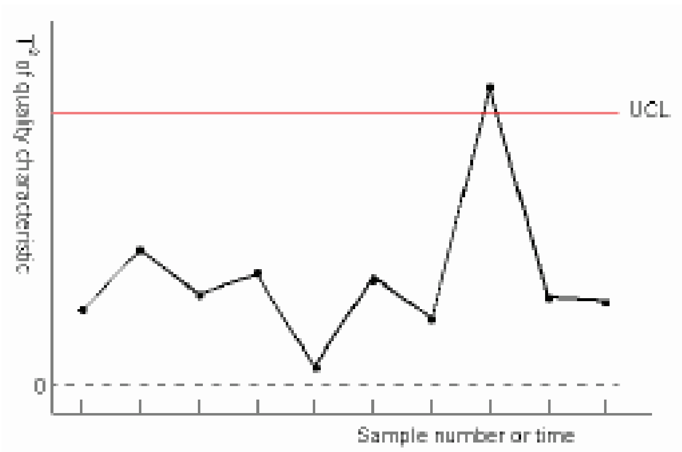

Figure 2.9: A generic T^2 control chart

2.4.2 MEWMA Control Chart

The MEWMA control chart is very sensitive in detecting small shifts ($\leq$ 1.00 standard deviations) as compared to the T^2 control chart. The MEWMA control chart developed by Lowry et al (1992) is a logical extension of the univariate EWMA control chart. In the bivariate case, the MEWMA statistics can be defined as follows:

$$MEWMA_i = [\sigma_2^2(EWMA_{1i} - \mu_1)^2 + \sigma12(EWMA_{2i} - \mu_2)^2 - 2\sigma12^2(EWMA_{1i} - \mu_1)(EWMA_{2i} - \mu_2)]\, n / (\sigma_1^2\sigma_2^2 - \sigma_{12}^2) \quad (2.14)$$

$$EWMA_{1i} = \lambda\, Z_{1i} + (1 - \lambda)\, EWMA1_{i-1} \quad (2.15)$$

$$EWMA_{2i} = \lambda\, Z_{2i} + (1 - \lambda)\, EWMA_{2i-1} \quad (2.16)$$

Covariance matrix of MEWMA:

$$\Sigma MEWMA = (\lambda / (1 - \lambda))\, [\, (\sigma_1^2\ \sigma_{12})\ (\sigma_{12}\ \sigma_2^2)\,] \quad (2.17)$$

The $\sigma_1 = \sigma_2 = 1$; $\sigma_{12} = \rho$. Notations λ and i represent the constant parameter and the number of samples. The starting value of EWMA ($EWMA_0$) was set as zero to represent the process target (μ_0).
The MEWMA statistic samples will be out-of-control if it exceeded the control limit (H). In this study design parameter H=8.64 was used.

2.4 Advances statistical process control

Advances in SPC Pattern Recognition (SPCPR) schemes can be observed has moved from univariate SPC applications towards multivariate SPC applications as shown in Figure 2.10. It can also be observed as influenced by the strong interest to apply soft computing technology such as expert systems (ES), fuzzy inference system (FIS), artificial neural networks (ANN), adaptive neural fuzzy inference system (ANFIS), support vector machine (SVM), and decision tree learning (DT), among others into the basis SPC charting procedures. In Advances statistical process control pattern recognition (SPCPR) scheme, ANN is apply as a recognizer for recognition of unnatural data streams patterns in multivariate quality control (MQC). The existing schemes were categorized into: (i) ANN-based model and (ii) Integrated MSPC-ANN model, based on the external structures. The ANN-based models such as

Novelty Detector-ANN, Modular-ANN, Ensemble-ANN, Multi Modules-ANN scheme etc have been designed to perform process monitoring and diagnosis simultaneously and continuously.

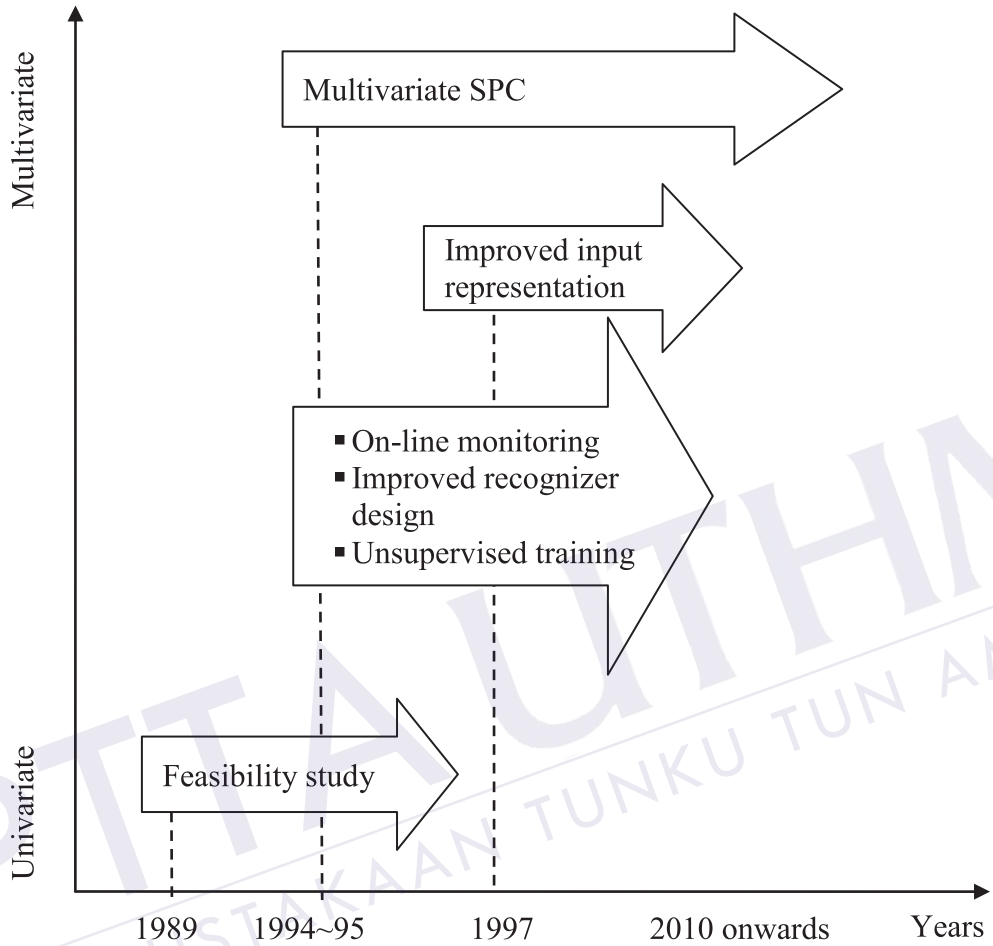

Figure 2.10: Advances in SPCPR schemes

2.4.1 ANN-based model: Novelty Detector-ANN

Zorriassatine *et al.* (2003) applied single ANN recognizer, namely, novelty detector-ANN as shown in Figure 2.11 for recognizing normal pattern and sudden shift patterns (upward shift and downward shift).This scheme was effective to accurately identify the existence of mean shifts when dealing with large magnitudes of mean shifts (≥ 2.0 standard deviations). However, there is difficult to identify the source of variation when involving small magnitudes of mean shifts (1.0 standard deviation).

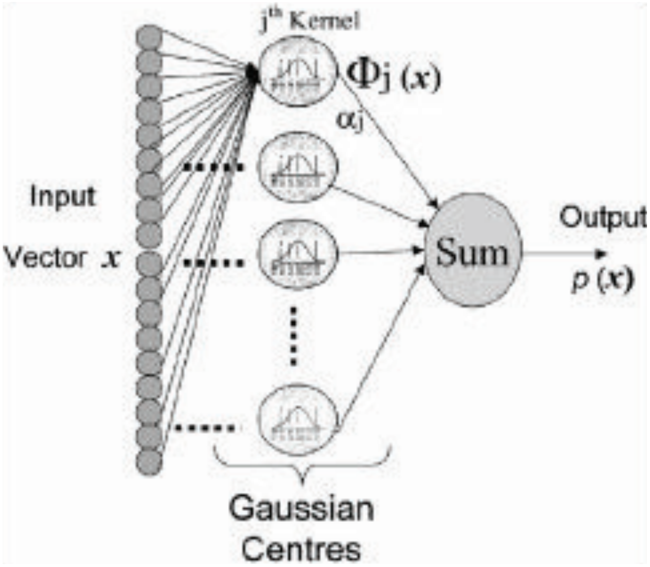

Figure 2.11: Novelty detector-ANN recognizer (Zorriassatine et al., 2003)

2.4.2 ANN-based model: Modular-ANN scheme

Guh (2007) proposed the modular ANN scheme as shown in Figure 2.12 for monitoring and diagnosis of bivariate process mean shifts. The study focused on the overall monitoring-diagnosis performances. In monitoring aspect, this scheme can be observed as so effective to rapidly detect process mean shifts (with short ARL_1) based on limited capability to avoid false alarms ($ARL_0 \approx 200$). This ARL_0 level was determined based on monitoring capabilities of the traditional MSPC charting schemes. In diagnosis aspect, it was also effective to accurately identify the sources of variation (with excellence RA results).

2.4.3 ANN-based model: Ensemble-ANN

Yu and Xi (2009) investigated this scheme for monitoring and diagnosis of bivariate process mean shifts. The sources of variation investigated are limited to three possibilities that are: upward shift (1, 0), upward shift (0, 1) and upward shift (1, 1). The upward shift (1, 0) pattern represents only variable-1 is shifted, upward shift (0, 45 1) pattern represents only variable-2, whereas upward shift (1, 1) pattern represents both variables are shifted. In monitoring aspect, this scheme can be observed as quite slow to detect the moderate and large process mean shifts (≥ 2.00 standard deviations). On the other hand, capability to avoid false alarms has been improved ($ARL_0 \approx 364$) close to the *de facto* level for univariate SPC charting schemes ($ARL_0 \geq 370$). In diagnosis aspect, it was also effective to accurately identify the sources of variation when dealing with moderate and large mean shifts ($\geq$ 2.0 standard deviations). Nevertheless, it has become less accurate when dealing with small mean shift (1.0 standard deviation).

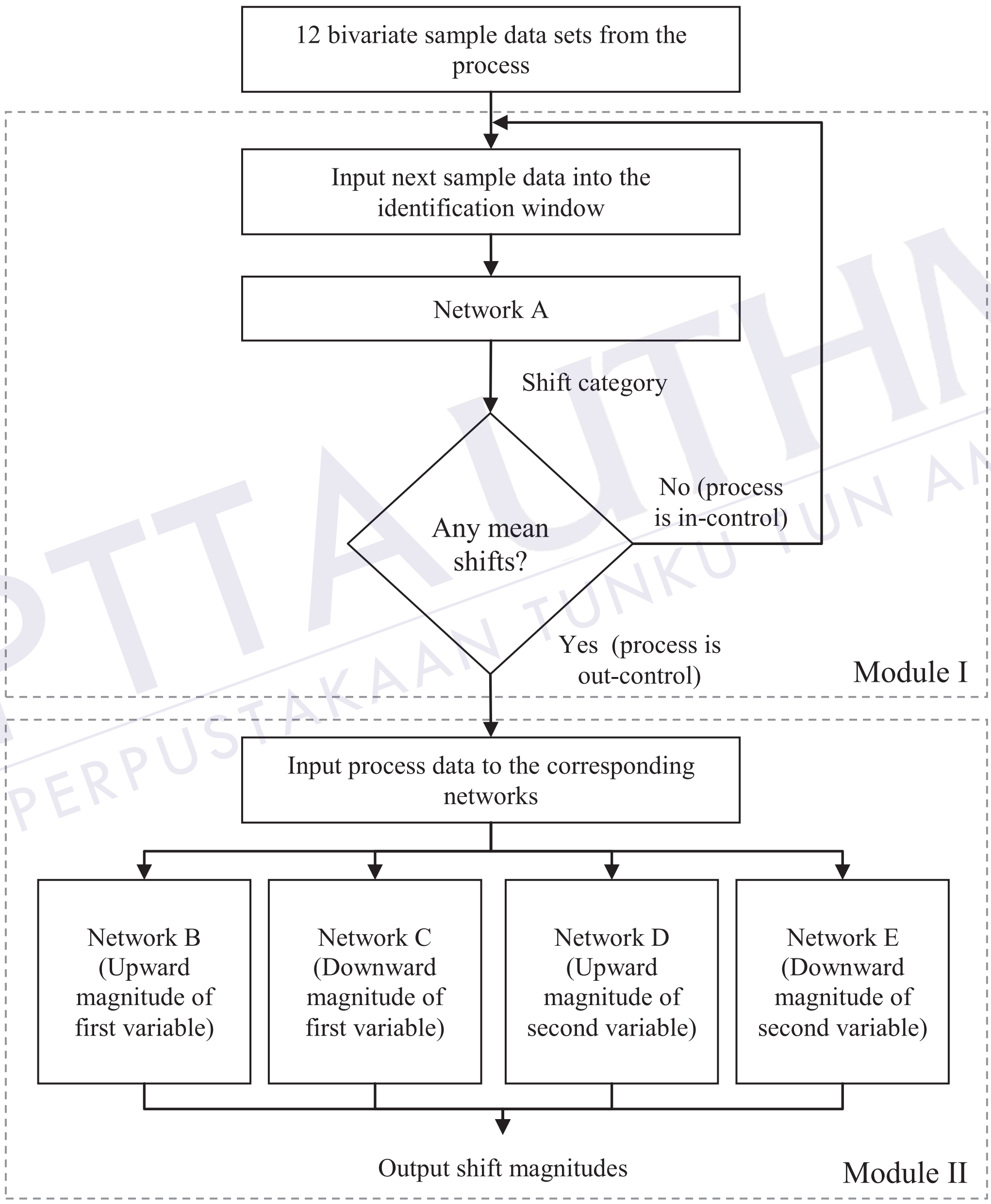

Figure 2.12: Modular-ANN scheme (Guh,2007)

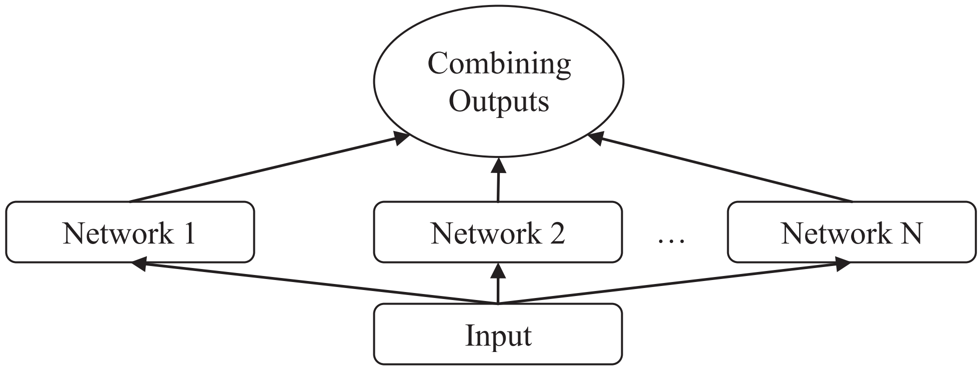

Figure 2.13: Ensemble-ANN recognizer (Yu and Xi, 2009)

2.4.4 ANN-based model: Multi modules-ANN scheme

El-Midany *et al.* (2010) proposed the multi-module-ANN as shown in Figure 2.14 for monitoring and diagnosis of three variates process mean shifts. The χ^2-statistics (56 consecutive χ^2) as shown in Block A were utilized as input representation for all individual ANN-based recognizers. Variation in mean shifts was represented by sudden shift and trend patterns as shown in Block B, which can be recognized using the three-layered MLP neural network recognizer. In Block C, outputs from several specialized-ANN recognizers were combined to determine the sources of variation. There are seven possibility sources of variation that are: (1,0,0), (0,1,0), (0,0,1), (1,1,1), (1,1,0), (1,0,1) and (0,1,1). Notation '1' represents shifted variable, while notation '0' represents normal variable.

2.4.5 Integrated MSPC-ANN

Niaki and Abbasi (2005) and Yu *et al.* (2009) reported the applications of integrated MSPC-ANN model for monitoring and diagnosis of bivariate process mean shifts. Yu *et al.* (2009) provided additional results based on three variables case. Such models as shown in Figure 2.15 were designed to perform sequential 47 process monitoring and diagnosis. Based on "one point out-of-control" charting rules, the traditional MSPC chart (T^2, MCUSUM or MEWMA) was applied to monitor the process mean shifts. Once an out-of-control signal is triggered, the ANN recognizer begins to identify the sources of variation (mean shifts) based on pattern recognition technique.

REFERENCES

Alwan, A. J. and Alwan, L. C. (1994). "Monitoring Autocorrelated Processes Using Multivariate Quality Control Charts." *Proceedings of the Decision Sciences Institute Annual Meeting 3*. pp. 2106 − 2108.

Aparisi, F. and Haro, C. L. (2001). "Hotelling's T2 Control Chart with Variable Sampling Intervals." International Journal of Production Research. Vol. 39. No. 14. pp. 3127 − 3140.

Apley, D. W. and Tsung, F. (2002). "The Autoregressive T^2 Chart for Monitoring Univariate Autocorrelated Processes." *Journal of Quality Technology*. **Vol. 34.** pp. 80 − 96.

Cheng, C. S. and 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*. **Vol. 35** pp. 198 − 206.

Chen, L. H. and Wang, T. Y. (2004). "Artificial Neural Networks to Classify Mean Shifts from Multivariate χ^2 Chart Signals." *Computers and Industrial Engineering*. **Vol. 47.** pp. 195 − 205.

Chen. W, "Multivariate Statistical Process Control in Industrial Plants, Master Thesis, Delft University of Technology, (2005)

Chen, Y. K. (2007). "Economic Design of Variable Sampling Interval T2 Control Charts - A Hybrid Markov Chain Approach using Genetic Algorithms." Expert Systems with Applications. Vol. 33. pp. 683 – 689

Crosier, R. B. (1988). "Multivariate Generalizations of Cumulative Sum Quality Control Schemes." Technometrics. Vol. 30. No. 3. pp. 291 − 303.

El-Midany, T. T., El-Baz, M. A. and Abd-Elwahed, M. S. (2010). "A Proposed Framework for Control Chart Pattern Recognition in Multivariate Process Using Artificial Neural Networks." *Expert Systems with Applications*. **Vol. 37.** pp. 1035 − 1042.

Flores, G. E., Flack, W.W., Avlakeotes, S., and Martin, B. (1995). “Process Control of Stepper Overlay Using Multivariate Techniques.” *OCG Interface* pp. 1 − 17.

Guh. R.S, “Effective Pattern Recognition of Control Charts Using A Dynamically Trained Learning Vector Quantization Network”, Journal of the Chinese Institute of Industrial Engineers, Vol. 25, No. 1, pp. 73-89 (2008)

Huh, Ick, "Multivariate EWMA Control Chart and Application to a Semiconductor Manufacturing Process" (2010).

Ibrahim b Masood , “A Scheme for Balanced Monitoring and Accurate Diagnosis of Bivariate Process Mean Shifts” ,UTHM PhD Thesis , 2012.

Jiang, W. (2004). “Multivariate Control Charts for Monitoring Autocorrelated Processes.” *Journal of Quality Technology*. **Vol. 36.** No. 4. pp. 367 − 379.

Karl D. Majeske & Patrick C. Hammet, “Identifying Sources of Variation in Sheet Metal Stamping” ,Working Paper 01-009, University of Michigan

Khoo, M. B. C. and Quah, S. H. (2003). “Multivariate Control Chart for Process Dispersion Based on Individual Observations.” *Quality Engineering*. **Vol. 15.** No. 4. pp. 639 − 642.

Lowry, C. A. and Montgomery, D. C. (1995). “A Review of Multivariate Control Charts.” IIE Transactions. Vol. 27. No 6. pp. 800 − 810.

Lowry, C. A., Woodall, W. H., Champ, C. W. and Rigdon, S. E. (1992). “A Multivariate Exponentially Weighted Moving Average Control Chart.” Technometrics. Vol. 34. No 1. pp. 46 − 53.

Masood, I. and Hassan, A. (2008). “Application of Full Factorial Experiment in Designing an ANN-based Control Chart Pattern Recognizer”, *International Graduate Conference on Engineering and Science*. pp. 187−193.

Mason, R. L., Chou, Y. M. and Young, J. C. (2001). “Applying Hotelling T^2 Statistic to Batch Processes.” *Journal of Quality Technology*. **Vol. 33** No. 4. pp. 466 − 479.

Mason, R.L., Tracy, N.D. & Young, J.C., (1992), Multivariate control charts for individual observations, Journal of Quality Technology, Vol.24, No.2, pp.88-95.

Montgomery, Douglas C, “Introduction to Statistical Quality Control”, Sixth Edition, John Wiley & Sons, Inc, 2009

Muzalwana,Susila & Shamsuddin, "A Multivariate Control Chart Scheme for Quality Monitoring of Automotive Stamping Process –An Empirical Study of A Malaysian Plant", Advances in Quality Engineering and Management Research (2009)

N. Costa, B. Ribeiro, "A Neural Prediction Model for Monitoring and Fault Diagnosis of a Plastic Injection Moulding Process" Department of Engineering Informatics, Coimbra University

Network Ensemble." *Engineering Applications of Artificial Intelligence*. **Vol. 22.** pp. 141 − 152.

Niaki, S. T. A. and Abbasi, B. (2005). "Fault Diagnosis in Multivariate Control Charts Using Artificial Neural Networks." *Quality and Reliability Engineering International*. **Vol. 21.** pp. 825 − 840.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, date

Owen L. Davies and Peter L. Goldsmith, "Statistical Methods in Research and Production", Fourth Edition, Oliver & Boyd , 1972.

Parra, M. G. L. and Loziza, P. R. (2003 − 2004). "Application of Multivariate T^2 Control Chart and Mason-Tracy Decomposition Procedure to the Study of the Consistency of Impurity Profiles of Drug Substances." *Quality Engineering*. **Vol. 16** No. 1. pp. 127 − 142.

Pignatiello, J. J. and Runger, G. C. (1990). "Comparison of Multivariate CUSUM Charts." Journal of Quality Technology. Vol. 22. No. 3 pp. 173 − 186.

Prabhu, S. S. and Runger, G. C. (1997). "Designing a Multivariate EWMA Control Chart." *Journal of Quality Technology*. **Vol. 29** No. 1. pp. 8 − 15.

Richard, A.J. & Dean, W.W., (2002), Applied multivariate statistical analysis, N.J.: Prentice Hall.

S. Bersimis1 J. Panaretos and S. Psarakis, "Multivariate Statistical Process Control Charts and the Problem of Interpretation: A Short Overview and Some Application in Industry" MPRA Paper No. 6397 (2007)

Singh, R. and Gilbreath, G. (2002). "A Real-Time Information System for Multivariate Statistical Process Control." *International Journal of Production Economics*. **Vol 75.** pp. 161 − 172.

Suzanne L.B.W, Douglas J.C and Blair V.S, "Online Pattern-Based Part Quality Monitoring of Injection Molding Process" Polymer Engineering and Science, 36,11 (1996)

Thyregod, "Modeling and Monitoring in Injection Molding", Ph.D. Degree, Technical University of Denmark, and Novo Nordisk A/S (2001)

Williams, J. D., Woodall, W. H., Birch, J. B., and Sullivan, J. H. (2006). "Distribution of Hotelling's T^2 Statistic Based on the Successive Differences Estimator." *Journal of Quality Technology*. **Vol. 38** No. 3. pp. 217 − 229.

William Mendenhall & Terry Sincich, "Statistics for Engineering and Sciences" Fifth Edition, Pearson Prentice Hall, 2007

Woodall, W. H. and D. C. Montgomery.1999. "Research Issues and Ideas in Statistical Process Control." Journal of Quality Technology 3 I (a): 37 6-86.

Yu, J. B., Xi, L. F. and Zhou, X. J. (2009). "Identifying Source(s) of Out-of-Control Signals in Multivariate Manufacturing Processes Using Selective Neural 166

Zorriassatine, F., Tannock, J. D. T. and O'Brien, C (2003). "Using Novelty Detection to Identify Abnormalities Caused by Mean Shifts in Bivariate Processes." *Computers and Industrial Engineering*. **Vol. 44.** pp. 385 − 408.