Cumulative Sum Quality Control Charts
Design and Applications

by

Galeboe Kesupile

University of Cape Town, South Africa

August 2006

Supervised by Professor Renkuan Guo

Dissertation presented for the Degree of Masters of Science in Quality Control, Reliability and Operation Research

Faculties of Mathematics and Statistics, University of Cape Town, South Africa

Declaration

Copyright:

I hereby grant the University of Cape Town permission to copy and disseminate this work, or any part thereof, for the purposes of study and research.

Plagiarism:

1. Plagiarism is to use another's work and pretend that it is one's own. I know that plagiarism is wrong and therefore my work does not contain any plagiarism.
2. I have not allowed, and will not allow any individual to copy my work with the intention of passing it off as her or his own work.
3. Each significant contribution to, and quotation in this thesis from the work(s) of other people has been attributed, and has been cited and referenced.

NAME	SIGNATURE
Galeboe Kesupile	signature removed
	Signed by candidate

Abstract

Classical Statistical Process Control Charts are essential in Statistical Control exercises and thus constantly obtained attention for quality improvements. However, the establishment of control charts requires large-sample data (say, no less than 1000 data points). On the other hand, we notice that the small-sample based Grey System Theory Approach is well-established and applied in many areas: social, economic, industrial, military and scientific research fields. In this research, the short time trend curve in terms of GM(1,1) model will be merged into Shewhart and CUSUM two-sided version control charts and establish Grey Predictive Shewhart Control chart and Grey Predictive CUSUM control chart. On the other hand the GM(2,1) model is briefly checked its of how accurate it could be as compared to GM(1,1) model in control charts. Industrial process data collected from TBF Packaging Machine Company in Taiwan was analyzed in terms of these new developments as an illustrative example for grey quality control charts.

Acknowledgement

This thesis is a result of two years of research at the department of Statistics of the University of Cape Town. In general, I thank Professor Tim Dunne and Professor Christien Thiart for their support.

A very interesting and appealing journal article on "Predictive Process Control Charts" their application in quality control by Professor Renkuan Guo and Professor Tim Dunne was the motivated reason for my choosing this topic of my dissertation. My first and most earnest acknowledgement must go to Professor Renkuan Guo, for his wisdom, vision, and guidance in my research. He contributed a lot to this thesis, by providing me with many useful suggestions and comments. Thank you, Professor Guo, for being such a guardian of full hearted support with rigor academic stand and parental-style patience. I enjoyed working with you I will always be proud for being "your" student and participant in your field.

I was delighted to work on problems of practical industrial statistics with the data provided by of Mr Liu and Professor C.Y. Yuan of TBF Packing Machines in Taiwan. I am extremely grateful for this Packing Company's generosity.

I thank Miss Danni Guo who gave a helping hand for the final formatting of the thesis.

I would also like to thank the people whom met I when I was still a student. Though they did not all know or understand what I was working on, they always motivated me in my work, especially the ladies by their smiles. However I feel I must apologize to some colleagues, for having not being on my side all the way through. IRE! JAH BLESS! PULA!

Table of Contents

Declaration ii

Abstract iii

Acknowledgements iv

Table of Contents v

List of Figures viii

List of Tables ix

1 Statistical Process Control and Control Charts 1

1.1 Statistical Process Control and Quality 1

1.2 Control Charts 8

1.3 Rational sub-grouping 30

1.4 Runs in random data and an average run length 32

1.5 Literature review 35

1.6 Summary 40

2 Basic Theory on CUSUM Control Charts 41

2.1 One–sided decision scheme for CUSUM Chart 43

2.2 Two–sided decision schemes for CUSUM Chart 44

2.3 Estimating σ for CUSUM charts 49

2.4 Average run lengths of CUSUM charts 50

2.5 The standardized CUSUM 54

2.6 A CUSUM for monitoring process variability 56

2.7 The Fast Initial Response feature in CUSUM charts 57

2.8 Combined Shewhart-CUSUM scheme 59

2.9 Summary 62

3 Exponentially Weighted Moving Average (EWMA) Control Charts 63

3.1 EWMA control chart for monitoring the process mean 65

3.2 EWMA control chart for monitoring the process variance 68

3.3 Average Run Length and design of an EWMA control chart 70

	3.4	EWMA control charts with time-varying control limits	75
	3.5	Fast initial response features for EWMA control chart	78
	3.6	Average run lengths and comparison of charts for control of the process mean level	82
	3.7	Combined Shewhart-EWMA	84
	3.8	Summary	85
4	**Review on Grey System Theory, Grey Differential Equation and Grey Models**		**86**
	4.1	Preview of grey system theory	87
	4.2	Generating of the GM model	92
	4.3	Grey differential equation and grey Models	95
	4.4	The GM(1,1) rolling model	102
	4.5	The Verhulst GM Model	103
	4.6	Non equal-spaced GM(1,1) model	105
	4.7	GM(2,1) model	108
	4.8	The extension of GM model	112
	4.9	GM(h,N) model	116
	4.10	Summary	116
5	**Grey Predictive Control Charts**		**118**
	5.1	A Grey predictive Shewhart control chart	118
	5.2	Transforming data sequence	128
	5.3	Grey predictive CUSUM control chart	130
	5.4	Fuzzy logic and control charts	132
	5.5	Grey- fuzzy predictive control chart	135
	5.6	Summary	137
Chapter 6	**An Industrial Application**		**138**
	6.1	The monitoring process	138
	6.2	Empirical modeling	141
	6.3	Implementation of a GM(1,1) on grey predictive Shewhart	

	control chart	145
6.4	Implementation of a GM(1,1) grey prediction CUSUM control chart	152
6.5	GM(2,1) model based grey predictive Shewhart control chart	159
Chapter 7	**Conclusions**	**162**
References		165
Appendix A	*Notations and Acronyms*	170
Appendix B	*Formulas for Designing Charts*	172
Appendix C	*Factors for Constructing Variables Control Charts*	174
Appendix D	*VBA Codes*	177

List of Figures

1.1	Reducing Variation and Quality Improvement	6
1.2	Statistical Thinking and Quality Improvement	8
1.3	A control chart with control lines	11
1.4	A Control Chart Selection Flow Chart	15
1.5	Control chart with marked "zones"	27
2.1	Tabular CUSUM status chart	46
2.2	ARL curve of Shewhart, CUSUM and Shewhart-CUSUM chart	62
3.1	EWM Control Chart	67
3.2	EWMA Control Charts with Time-varying control limits	78
3.3	EWMA Control Charts with Time-varying control limits and Fast Initial Response	82
3.4	ARL curve of Shewhart, CUSUM and EWMA chart	83
4.1	The Grey system theory concept	88
4.2	The AGO for the jumping sequence	94
4.3	GM(1,1) model concept	98
5.1	Fuzzy set "in-control" defined for characteristic x	134
6.1	Movement of original and predicted GM(1,1) machine failure time	150
6.2	Grey Prediction CUSUM Control Chart	156

List of Tables

1.1	Alarm Rules from Duncan's Quality Control and Engineering Statistics	28
1.2	Nelson's Alarm Rules from the Journal of Quality Technology	29
2.1	ARL Performance of Shewhart control chart and CUSUM	50
2.2	Tabular CUSUM Control Chart Example	55
2.3	ARL values for basic CUSUM and Shewhart- CUSUM	61
3.1	Calculations of Control Limits of EWMA chart	68
3.2	Optimal EWMA Control Schemes	74
3.3	EWMA Example using data from a process initially out-of-control	77
3.4	ARL values for a Shewhart, EWMA and CUSUM chart	82
4.1	The Comparison of Traditional and the Grey Prediction Methods	91
4.2	Summarized results	114
4.3	Comparison of simulation accuracies between three GM models	115
5.1	Critical Values of relative error	123
5.2	Overall model precision ranks	124
6.1	Total number of days on operation for a machine	139
6.2	Quality Measurement Representative Indices value's deviations around the mean	140
6.3	Grey Predictive Shewhart control chart Limits	142
6.4	Linearity Check Results	144
6.5	Results from GM(1,1) on each of output quality measurements representative index	146
6.6	Statistics of Relative accuracy of fitting the GM(1,1) model	148
6.7	Linearity Check for CUSUM Values Prediction Models	152
6.8	Results from GM(1,1) for Grey Predictive CUSUM Control Chart	153
6.9	Mean and Standard Deviation for Original Data Sets	156
6.10	CUSUM Trend Line Identification Rules	156
6.11	Results from GM(2,1) on output quality measurements representative indices	161

Chapter 1 Statistical Process Control and Control Charts

In this chapter, we will discuss the basics of Statistical Process Control (SPC) and Statistical Control Charts. In Section 1.1, we will describe the term "quality" and the phrase "statistical thinking" in the statistical quality control context. We also describe the terms of causes of variation in the process. Section 1.2 describes Shewhart control charts and focus on their settings up, plotting, types, and their applications. In the remaining sections we will review some additional rules for control charts, rational sub-groupings and computation of Average Run Length (ARL) for Shewhart Control Charts. In Section 1.5 we will review the literature on statistical process control charts topics related to this thesis summarizing them in connection with this thesis.

1.1 Statistical Process Control and Quality

Any goods or services produced or provided by any industry or organization must be of good quality. The goods are of good quality if they conform to the requirements and are fit for use. In quality control, *quality* is some attribute or characteristic of a process or product for which a standard is established. Currently the trend of globalization increases the amount of competing products and services worldwide. Globalization also broadens the consumer's choices. There are many factors that influence the choice options among consumers in the environment of highly competing products and services. Quality is one of the leading factors. This makes the manufacturing industries and service organizations pay more attention to the quality of their products or services, so as to satisfy both themselves as providers for business sustainability and the consumers on the receiving end. This can be reached if the manufacturers and organizations comply with the following:

- Design quality goals and manufacturing quality; the quality objectives are set for the process of manufacturing and the manufacturing quality (quality results) measures how well manufactured products will fulfill the design quality goals.

- Quality reliability; the product quality is evaluated on the basis of whether or not the product carries out its intended functions within specified time-duration for the specified function within its specified working environments.
- Consumer-oriented quality; the manufacturing industry must be consumer-oriented on its efforts to create and satisfy demand of customers.

To achieve the above mentioned objectives, the process can be monitored if the information of the production process is collected, then the output is scrutinized with the intent to verify if it reaches the required standards. Output scrutinizing is done by using Statistical process control (SPC). SPC uses statistical monitoring charts, and the core technique in statistical monitoring charts is control charts or Shewhart charts named after Dr. Walter A. Shewhart. There are other techniques like the cumulative sum (CUSUM) control charts and exponentially weighted moving averages (EWMA) control charts, these are discussed in Chapter 2 and 3 respectively.

"Quality" as a word has different contexts of use by people. This makes it hard to give a perfect or precise definition that distinguishes between the bad or high quality of products or services. Different authors have formulated different definitions of the word "quality", which make it clear to realize that the quality of a product or a service is not a single, identifiable characteristic. This makes it necessary to differentiate the different dimensions of quality. To do this we will review a set of eight intuitive quality dimensions proposed by Garving (1987).

1. **Performance** is one of the measures of quality. It refers to the basic functions of a product. Customers usually evaluate a product to determine if it will perform certain specifications and determine how well it performs them;
2. **Reliability** is one of the traditional measures of quality. A reliable product rarely fails. This aspect is sometimes reformulated as 'being free of defectives' is a very important dimension of quality;

3. **Durability** is a measure of product life, either economically or physically. Economically a product is considered to be durable if its expected cost of repair does not exceed its current value. Physically a product is considered durable if it is possible to repair it. A product that lasts longer is usually viewed as being of higher quality;
4. **Serviceability** relates to the time and effort that is needed to repair or routinely maintain a product. The breaking down of a product is usually viewed as an annoyance, but a prompt repair may relieve part of irritation;
5. **Aesthetics** is a subjective dimension of quality. It refers to the look (color, shape and packaging), feel, sound, taste or smell of a product. It is greatly influenced by the preferences of the individual customer. On this dimension of quality it is usually not possible to meet the needs of every customer, but the manufacturer is pleased if the product meets the needs of a majority of regular, trusted and serious consumers;
6. **Features** that are added to the basic functioning of a product attribute to a higher quality and performance. This is why there is always upgrading of products in manufacturing companies;
7. **Perceived quality** is very subjective dimension. When customers do not have full information about a product they may directly base their quality image on past experiences, the reputation of the manufacturer, the quality of other products from the same manufacturer, or the name of the product;
8. **Conformance to Standards** is related to reliability. It refers to the degree to which a product meets pre-established requirements of the designer. This dimension of quality is very important in situations where products are used as components in more complex assembly. Specifications on the individual components are usually expressed as a target and a tolerance. If each of the components is just slightly too big or too small, a tight fit is unlikely, and the final product may not perform as intended by the designer, or may wear out early.

The summary stated above clearly shows that quality is not a one-dimensional characteristic of a product and that quality is determined at various levels during the production process. We should emphasize that the eight dimensions of quality are merely intuitive descriptions not the well-defined quality standards. The eight dimensions are over-lapped. The "durability" dimension is better to be replaced by the term "serviceability (maintainability)". The "feature" dimension is even too vague to cover many other quality dimensions. In industrial quality practices, we must to comply with these international well-accepted quality standards, for example, MIL-STd 721, and DIN 40041 and others, which offer better descriptions and definitions on quality dimensions.

Variability is a critical measure of quality. Montgomery (2001) pointed out *"Quality is inversely proportional to variability."* This description states the relationship between quality and variation. It implies that if the variability in the important characteristics of a product decreases the quality of the product increases. The manufacturer must use statistical process control methods to reduce variability because variation in the product dimensions is a fundamental reality in production. The statistical control methods helps in carefully monitoring of products and reacting to unusual events (variations) by finding and eliminating their roots causes leading to quality improvements in the processes. Identifying the causes for generating product variations is of critically importance in quality improvement processes. There are two types of causes of variations, which will be discussed in the next section.

1.1.1 Causes of Variations

When Dr. Walter A. Shewhart developed the control charts in 1920s he was aiming at quality improvement through reduction of variation. Since then control charts are being used and improved for observing and recording measurements from the process and also detecting changes and the causes of the changes in quality of the products during the process. In statistical quality control theory, the causes of variations are classified into common cause or special cause. This classification guides the management or control chart user to take the appropriate action for quality improvement. These causes of the

variability exist not only in the industrial quality management but also in all other areas, for example, organizational administration. In this thesis we will focus on the area of industrial quality control. Furthermore we note that in the industrial process, quality control can be considered in two ways namely a product orientation or a process perspercive, and they are described as follows.

A product orientation focuses on the parts or units after they are manufactured. A single quality dimension is considered at a time and, the quality of a part is defined based on the target value with the specified limits for that quality dimension. Manufcturing or engineering considerations determine the specifications of the limits specifying the range within which the acceptable quality dimensions should fall in. These will be further discussed in section 1.2.2. and applied in industrial data in Chapter 6.

The process perspective focuses back on the process that is producing items by montoring the system that produces items and trying to identify and fix problems before too many "defective" products are made at the end. In the process perspercive, it is assumed that all the variation in a system arises from one of two causes: the *natural variation* or *non natural variation*. The causes of natural or random variation are called "*common cause*". Common cause of variation is present to some extent in all processes. It is a built-in characteritic of the process which stems from the material variability in inputs to the process and the conditions in which it is operating. When common cause of variation is the only cause of variation present in the process searching for causes of variation will not be necessary because reducing variation due to these causes is normally expensive. Also when common cause of variation is the only cause of variation adjusting the process in response to each deviation from the target value might increase the varaibility.

The causes for non natural or nonrandom variation are called "*special causes*" or "assignable causes". These causes are not always present. They generally occur when something does not go right. These causes of variation be traced and can usually be eliminated relatively easily. When they are eliminated, only natural or random variation

remains. Examples of causes of variation are human error, a problem with raw material or a machine failure. When the process is operating with only natural variation present the process is said to be in a state of statistical control and is referred to as stable process. Figure 1.1 summarizes the reduction of variation so as to improve quality.

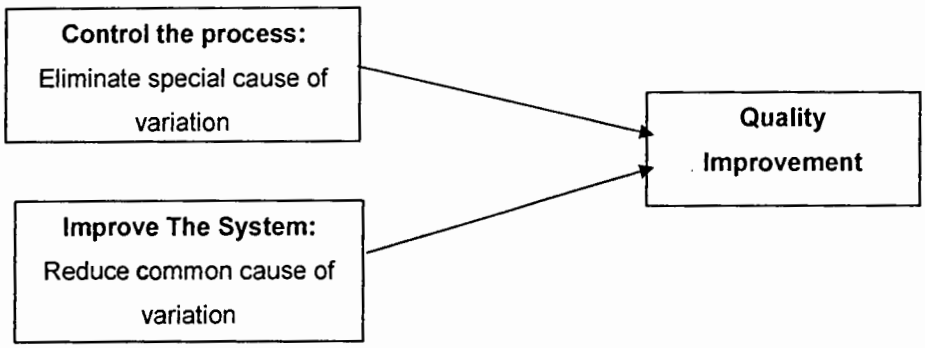


Figure 1.1 Reducing Variation and Quality Improvement

1.1.2 Statistical Thinking

Application of the idea of improving quality by reduction of variability is limited in most cases of manufacturing environments. Snee (1990) discussed Total Quality Management in a broader context in place of Statistical Process Control. He discussed that a successful implementation of idea of improving quality by reduction of variability requires a new way of thinking that he calls "statistical thinking". Statistical thinking is a philosophy of learning and action based on the following three fundamental principles as follows:

- *All work is a system of interconnected process*: In a process one or more activities are connected in which inputs are transformed into outputs for specified purpose. Industries, businesses and other organizations are made up of a collection of processes. The processes interconnect and interact to form a system that provides a product or service for a customer.
- *Variation exists in all processes:* In statistical thinking variation exists in all processes. This provides the opportunities for improvement of the process and hence its product quality. Focusing on variation is a key to improving

performance and achieving consistency, since variation is the enemy of quality. Variation causes the difference between two items/units produced by the same process.

- *Understanding, analyzing, quantifying and reducing variation are key to business success*: In any business activities variation should be identified, characterized, and quantified in so as to understand both the variation and the process generating it. With this knowledge, there can be work to change the process and reduce its variation. Its average or centering and the amount of variation around the average performance influence the performance of a process. Aiming at process improvement, we need to understand the factors that influence the average performance and also those factors that influence variability in the process performance.

Snee described "Statistical thinking" in quality improvement as follows:

> *". . .as thought process, which recognize that variation is all around us and present in everything we do, all work is a series of interconnected process, and identifying, characterizing, quantifying, controlling, and reducing variation provide opportunities for improvement".*

This description is shown schematically in Figure 1.2. The idea integrates the ideas of processes, variation, analysis, developing knowledge, taking action and quality improvement.

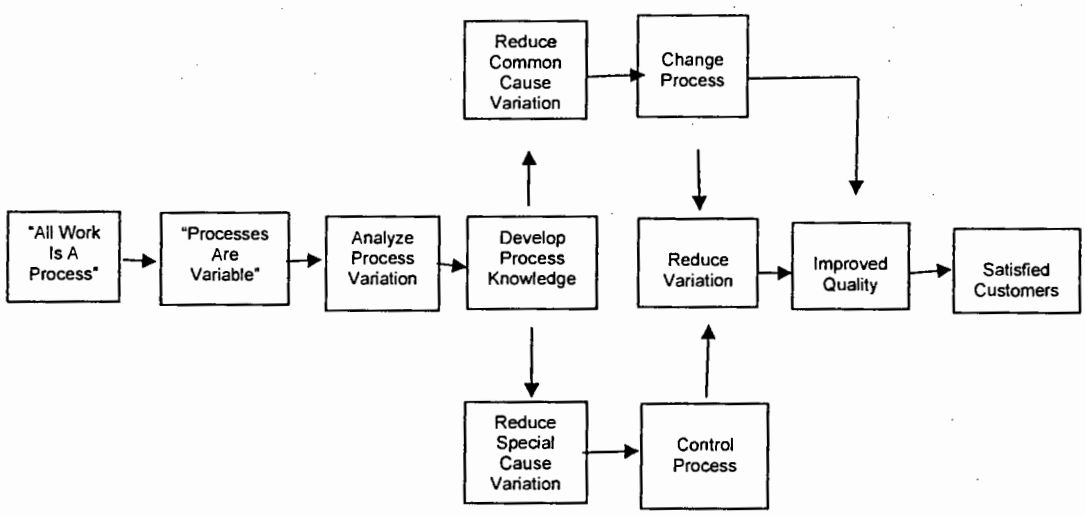


Figure 1.2 Statistical Thinking and Quality Improvement, Snee (1990).

1.2 Control Charts

A control chart is a plot of quality characteristic as a function of time or sample (sequential) number. Control charts are used on-line for monitoring a characteristic of the process for the purpose of showing whether the process is operating within its control limits of expected variation, so they provide the manufacturer or user information for a quick detection of the variation in an important quality charactersitic, and thus allow for an on-time corrective action if the detected variation is undesirable. There are many other benefits associated with the applications of control charts:

- The control charts may also be used to estimate the parameters of a production process, and, through this information, to determine process capability.
- The control charts helps to distinguish between the causes of variation.
- Information obtained from control charts can be used to access the inherent capability of the process after achieving the statistical control.
- Control charts can be used to evaluate the effects of improvement actions, aiming at removing assignable causes of reducing variability associated with common

causes. That is control charts contributes to real record of the qaulity product that provides the information usefull in improving the process.

- Control charts can form a basis of continuous monitoring strategy designed to ensure that the process remains in a reasonable state of statistical control and capable of meeting the requirements over time. That is a "systematic use of a control chart is an excellent way to reduce variability", Montgomery (2001), for keeping greater uniformity of product.
- A successful control chart program will reduce a scrap and rework, which are the primary production-killers in any operation. When scarp and rework is reduced the process will have larger volume of production at no increase in cost.
- Control charts prevent unnecessary process adjustment. The usage of control charts is consistent with the philosophy: "if it is not broken, do not fix it".

1.2.1 Setting up Control Charts

Irrespective of the type of process being monitored or the type of control chart in use, we will have to distinguish between the setting up of the control chart and the on-going plotting of the control chart. There is a so-called "standard procedure" for setting up control charts for sub-grouped data described by Porter and Caulcutt (1992) as follows:

(a) Obtaining sampling data from production process

(b) Putting the data into subgroups

(c) Calculating the mean and range of each subgroup

(d) Calculating the overall mean

(e) Estimating the process standard deviation by using $\overline{R}/d_n$. (d_n is Hartley's constant, which is dependent on the sample size and is tabulated in many statistical process control textbooks), see Table C.1 in Appendix C.

(f) Calculating values for control lines on a mean chart and a range chart by using $(\overline{\overline{x}})$, $(\overline{R})$, or the estimate $\overline{R}/d_n$, and appropriate constants.

(g) Plotting the group means onto the mean chart and plot the group ranges onto the range chart.

(h) If the control charts (plotted) indicate that the process was stable, that is, in-control then these charts, or other charts of a more appropriate type based on the same estimates, can be used to monitor future production.

(i) If the control charts (plotted) indicate that the process was unstable, that is, out-of-control then the assignable causes should be sought and corrective action taken. This is the whole procedure should then be repeated.

Not all the athours of standard textbooks in statistical process control listed the nine steps exactly described above, but all authors did certainly stress the importance of the steps (b) to (e). More discussions for stressing these can be found in Section 3 of Caulcutt R *et al.* (1995).

1.2.2 Plotting a Control Chart

Shewhart control charts compares the observed sample to what is expected from the process on its past peformance, that is, the control of current variables is based on the variables of that past. Control chart process is time series based. The quality characteristic is monitored by repeated sampling of the process. Based on each sample, a test statistic is calculated and plotted on the control chart. This leads to a control chart constructed in way that the horizontal axis of the chart indicates the sample (sequential) number or time at which a quality characteristic is measured and the vertical axis indicates the value of the quality characteristic, see Figure 1.3. There are three basic lines which are superimposed on the plot: the center line and the two control limits. The centre line (abbreviated as CL) represents the specified target level for the quality of the process. The control limits namely the upper control limit (abbreviated as UCL) and the lower control limit (abbreviated as LCL) creates the bandwidth. The general form of control limits is computed by:

$$\begin{aligned} UCL &= E[statistic] + L[\text{standard deviation of } statistic] \\ CL &= E[statistic] \\ LCL &= E[statistic] - L[\text{standard deviation of } statistic] \end{aligned} \qquad (1.0)$$

The upper and lower control limits lines are usually placed at distance of 3-sigma (3σ) of the centre line, that is, usually $L = 3$. It is common to assume that the subgroup means have a normal distribution, the probability that any observed mean will fall within the bandwidth of $\pm 3\sigma$ is approximated by using a normalized variable $z = \frac{x - \mu}{n}$:

$$\Pr(-3 \le z \le 3) = \Pr(z \le 3) - \Pr(z \le -3) = 0.998 - (1 - 0.998) \cong 0.997 .$$

Alternatively, the probability that any observed mean will fall outside the mean is approximately 0.003. Even if the statistic being plotted is not distributed precisely as normal distribution, the control limits are positioned so that the probability that the observed characteristic falls outside the bandwidth is very small, that is, less than one percent when the process is in control, Braverman (1981) The Chebyshev's theorem states that, no matter what the true underlying distriburion is, at least 89% of all observations fall within 3-sigma of the mean or center line.

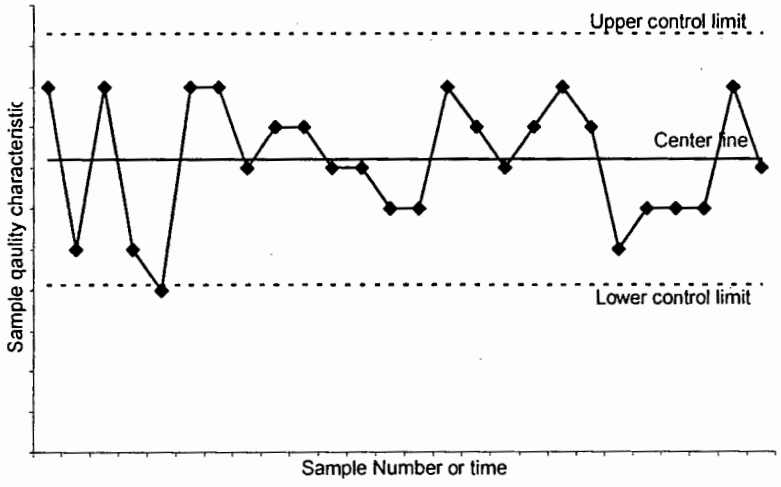


Figure 1.3 A control chart with control lines

The process is said to be under statistical control or stable when the control chart displays a random pattern of variation about the center line within the upper and lower limits, i.e., within the (control) bandwidth. When a process is in control, the control chart provides a method for testing a statistical hypothesis, where the null hypothesis

H_0: "The process is in-control" and the alternative hypothesis H_1: "The process is out-of-control". So long as each sampling statistic value falls within the bandwidth, the null hypothesis can be accepted. If one or more values fall outside the bandwidth, the null hypothesis is rejected in favor of the alternative one. When this occurs, the process should be stopped and an investigation should be conducted to locate the assignable causes. Once the assignable causes are determined and eliminated, the production process will be resuming. It should be noted that the control chart is based on the fact that if the process is in-control, the outcomes are predictable, that is, based on previous observations, it is possible for a given set of limits to determine the probability that future observations fall within the bandwidth.

It should also be noted that the location of control units is closely related to the risk involved in the making an incorrect assessment about the state of control. That is the location of control units might lead to type I or type II error. A type I error occurs when an incorrect conclusion that a special cause is present when in fact it does not exist is reached. This results in the company incurring costs in trying to find a nonexistent problem. A type II error occurs when special causes are present but are not signalled in the control chart because points fall within the control limits by chance. The size of Type I error depends only on the control limits that are used; the wider the limits, the less the chance of falling outside the limits, and consequently there will be a smaller chance of making a Type II error.

1.2.3 Extending Control Charts Applications

There are several key foundational elements that must be used to make the most effective use of the control charts. These foundational elements are stated and briefly and defined below.

> *The control limits will always be set at a distance of three sigma units on either side of the control chart center line.*

This has been justified by Dr Shewhart that the three sigma limits provide control of a practical between the likelihood of committing Type I and Type II errors.

The three sigma control limits must be set using an average measure of dispersion based on the within subgroup variability.

The upper and lower limits are designed to capture the variation in the sample average, range or standard deviation that is due only to common causes. These limits are constructed using an estimate of the standard deviation for the common cause system that applies to the process when it is operating in the state of statistical control. The range or standard deviation are always calculated within each subgroup and then averaged across subgroups to obtain the average range or average standard deviation used to construct the control limits. This will ensure that a practical estimate of the inherent, common cause process variation is obtained that is insensitive to assignable causes of variation occurring across the subgroups, making the control chart very sensitive to changes in variation across the subgroups, see Figure 1.4.

The conceptual basis for the control charts is the notion of rational sampling and rational sub grouping.

The method by which the data are sampled from the process and arranged on the charts will determine what questions can be answered from the data. There is need to plan ahead and consider what questions might be asked of the data because failure to do so results in incorrect sampling schemes, and control charts which produce nonsensical results at best and incorrect results at worst.

The effective use of control charts is limited by the industry's or organization's ability and desire to take action on the knowledge gained from the charts.

Without developing an environment that encourage quick reaction to the knowledge gained from the charts, the analysis activity will only be an academic exercise and not

practical to the industry. When an environment that encourages quick reaction to the information about the performance of the process results, demonstrated by a control chart, in a decision either that the process is in-control or out-of-control is developed action would be taken if the process is out-of-control. The action can be of two types; *direct action* or *indirect action*, Ewan (1963).

Direct action results in the physical alteration of some aspect of the process. The direct action control procedures imply a fairly precise knowledge of cause and effect and are applicable whenever one is trying a process at certain level and knows how to alter the process when it has changed.

Indirect action results in attempts to find out why there has been a change in the process. An indirect action control generally implies to no very specific knowledge about the cause of the change in the process and how immediately to put the process to the right procedure or in some cases, how to maintain the new better process level.

1.2.4 Types of Control Charts

In statistical research field there are two types of data, numerical data and attribute data. As part of statistical methods control charts use numerical data and attribute data, this makes the control charts to be classified according to the type of the product output quality characteristic that they are supposed to monitor. These are quality control charts for variables and quality control charts for attributes, which control chart to use is summarized in Figure 1.4 below.

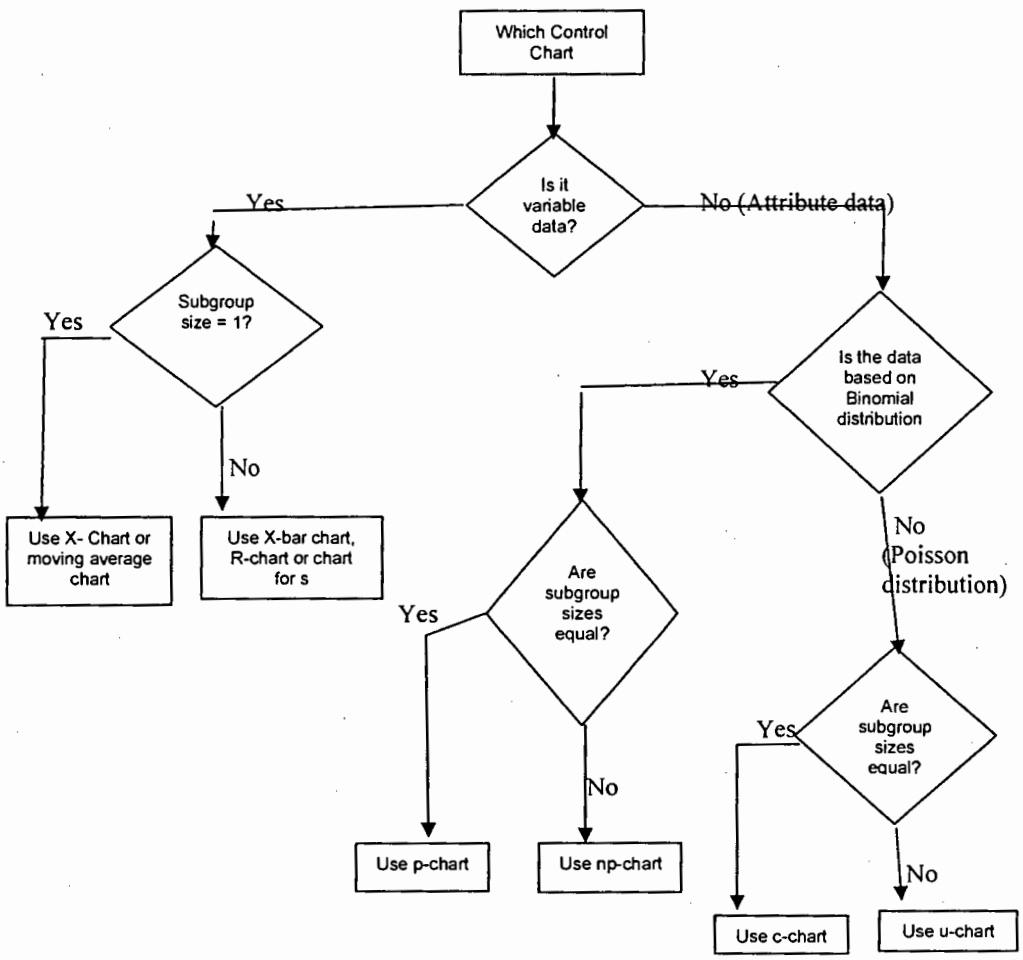


Figure 1.5 A control chart selection flow chart

1.2.4.1 Control Charts for Variables

In control charts a quality characteristic that is measured on a numerical scale is referred to as a variable. The most commonly used charts to chart variable measurements are the X-bar chart and the R chart. The X-bar chart monitors between-sample variability, that is, it monitors the average quality level in the process over time. It is used to control the mean or central tendency of the process. The R chart measures within the sample variability that is the instantaneous process variability at a given time. Other control charts for variables are S chart and S^2 chart. In the S or S^2 chart, the sample *standard deviations* are plotted in order to control the variability of a variable.

X-bar and Range charts are used when one can rationally collect measurements in subgroups. A subgroup could be of between two and ten observations $(2 \le n_i \le 10)$. Each subgroup represents the process at a given point in time. The charts' x-axes are time based, so that the charts show a history of the process. Due to this reason, we must have data that is entered in the sequence from which it was generated, that is, time-ordered sequence. If the sequence is not time ordered, the trends or shifts in the process may not be detected, but instead it might be attributed to random (common cause) variation. X-bar chart and R chart are very effective indicators of problems in the process and also indicate when the problems have been cleared. The center line (CL) is set at the mean μ, the average of all the data. An upper control limit (UCL) and a lower control limit (LCL) are usually based on the mean plus and minus three standard errors of the mean respectively. That is, using equation (1.0) we have

$$\mu \pm 3\sigma_{\bar{x}} = \mu \pm 3\sigma/\sqrt{n} \qquad (1.1)$$

Where μ = the process mean

σ = the process standard deviation

n = the sample size.

The parameters of μ and σ are not known so they have be estimated from the data. The best estimate of μ is $\bar{x}$ where $\bar{\bar{x}}$ is the average of the sample averages. $\bar{x}$ and $\bar{\bar{x}}$ are calculated as

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad \bar{\bar{x}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m}.$$

Where m = the number of subgroups in a sample. The above formulae for computing control limits are used when using X-bar control charts for analyzing past data. Still in a case of analyzing past data a question arises as to the best estimate of the process standard deviation. In this case the process standard deviation σ may be estimated by a function of the average of the sample ranges $\bar{R}$. This is done by dividing the average range by a constant d_2, where d_2 is a function of the sample size. $\hat{\sigma} = \frac{\bar{R}}{d_2}$ where $R = x_{max} - x_{min}$ and $\bar{R} = \frac{\sum_{i=1}^{m} R_i}{m}$. d_2 is calculated as the expectation of relative range, that is, $d_2 = E_w = E R_n/\sigma$ where R_n is a random variable representing the range of a sample size n from a normal distribution of variance σ^2.

The control limits for the X-bar chart may also be calculated from equations which are easier to use as follows:

$$\begin{aligned} UCL_{\bar{x}} &= \bar{\bar{x}} + 3\bar{R}/\left(d_2\sqrt{n}\right) = \bar{\bar{x}} + A_2\bar{R} \\ LCL_{\bar{x}} &= \bar{\bar{x}} - 3\bar{R}/\left(d_2\sqrt{n}\right) = \bar{\bar{x}} - A_2\bar{R} \end{aligned} \tag{1.2}$$

where $A_2 = 3/\left(d_2\sqrt{n}\right)$.

The values of d_2 and A_2 re tabulated in Table C. 1. See Appendix C.

If the X-bar chart is used to control data of the current input, the center line may be a standard value $\bar{x}''$ derived from past data or selected the company's management to

attain certain objectives. The upper and lower control limits may also be based on a standard value for the process standard deviation, like σ'', this implies that a standard value of $\sigma''/\sqrt{n}$ for means samples size n, that has been derived from past data or selected by the management. In this case the formulae for control limits for attaining current control would be

$$\overline{x}'' \pm 3\sigma'' = \mu \pm 3\sigma''/\sqrt{n} \tag{1.3}$$

A range chart shows the variation in the ranges of subgroups in the sample. In the case of range chart the ranges are computed by "artificial" samples of size 2. The center line is set as the mean range estimated by $\overline{R}$. The upper and lower control limits are usually $3\sigma_R$ above and below the center line and these are calculated as $\mu_R \pm 3\sigma_R$ where μ_R or σ_R are the mean and the standard deviation of the distribution of the sample ranges, respectively. The standard deviation of the range is $\sigma_R = d_3\sigma$. This may be estimated by $d_3\overline{R}/d_2$. Therefore the control limits for the range chart would be calculated as

$$\begin{aligned} UCL_R &= \overline{R} + 3d_3\left(\overline{R}/d_2\right) = \left(1 + 3d_3/d_2\right)\overline{R} = D_4\overline{R} \\ LCL_R &= \overline{R} - 3d_3\left(\overline{R}/d_2\right) = \left(1 - 3d_3/d_2\right)\overline{R} = D_3\overline{R} \end{aligned} \tag{1.4}$$

The control chart factors A_2, D_3 *and* D_4 are functions of sample size n used to compute the standard errors or the mean, the range, and the standard deviations have already been computed along with d_2 and d_3 are tabulated in Table C. 1. See Appendix C. The necessary computational formulas for the major types of the control charts discussed above are summarized in Table B.1. See Appendix B.

When constructing X-bar and R-charts from the observed data it is generally advisable to begin with construction of an R-chart because the limits of the X-bar chart will depend upon the estimated standard deviation and the estimated standard deviation should be properly made from the R-chart. It is very important to note that it is incorrect to estimate the process standard deviation from all the data by using the formulae like

$s = \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)}$ and use this for setting up limits for the X-bar chart. The estimate computed this way will be affected by the variation between sample means and so it is not an independent measuring rod for detecting extreme variations in these terms. This is stressed and illustrated by Duncan, A.J. (1986), *Quality Control and Industrial Statistics*, 5th edition, Richard D. Irwin, Inc. with reference to the heights of fragment bomb bases, Table 3.2, p.42. The estimate of the process standard deviation to be used in setting up control limits for the X-bar chart must be computed from the within-sample variation excluding between-sample variation. If the R-chart shows that the variability of the process is in control, such an estimate may be derived from the average of the sample ranges. If the R-chart shows that the variability of the process is out of control, it is better not to set up an X-bar chart until control of the process has been attained. If in any case one or two points are outside the control limits, and assignable causes can be found for these points, one might venture into setting up limits for the X-bar chart on the basis of values $\bar{\bar{x}}$ and $\bar{R}$ of calculated from the remaining points.

Sometimes charts for s, the standard deviation are used with the X-bar chart in place for the sample range charts, in measuring the process dispersion where

$$s_n^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}$$

is the usual estimate of variance. In such cases the standard deviation of each sample would be calculated, the average of the sample standard deviations would be found and the control limits would be calculated from the equations below:

$$UCL_s = \left(1 + \frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{s} = B_4\bar{s}$$
$$LCL_s = \left(1 - \frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{s} = B_3\bar{s} \qquad (1.5)$$

Where $c_4 = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma(n-1/2)} \cong 1 - \frac{1}{4(n-1)}$

If the management provides the standard values of the standard deviation from the start at which it is expects the process can be controlled, the value of the standard deviation is denoted by σ''. The control limits on the s or s^2 chart would be calculated as

$$UCL = \left(c_4 + 3\sqrt{1-c_4^2}\right)\sigma'' = B_5\bar{s}$$
$$LCL = \left(c_4 - 3\sqrt{1-c_4^2}\right)\sigma'' = B_5\bar{s} \qquad (1.6)$$

The values of B_3 , B_4, B_5 and B_6 are tabulated in Table C.1, Appendix C.

At this stage we will now discuss control charts for individuals. Different types of industries produce different products, but these industries use quality control charts. In some industries like those for chemical or other continuous process, such as temperature or pressure it is not practical to take a subgroup from a process. In such cases several samples are taken from the sample at the same time, but only one observation is checked each time from a sample. The reality for such a situation is checking the measuring device and not the manufacturing process itself. Also in some industries data may be available only weekly, fortnightly, monthly, seasonally or even yearly so that we do not get sufficient data quickly enough to be able to divide the data into subgroups. This also forces one to construct a control chart using individual observations or measurements. A control chart based on one observation is called to an *X-chart* or *a chart for individuals.* In an X-chart there is no benefit of the central limit theorem, which tells if the distribution of averages is approximately normal irrespective of the underlying distribution, since we are not using the sample averages. This makes an X-chart more sensitive to lack of normality of the underlying distribution as compared to an X-bar chart.

Using the sample size of n =1 there is no obvious source of an estimate of the standard deviation that can be used to derive the control limits of the X-chart. In this case one can use *moving range* and *standard deviation* methods to derive the control limits. Control limits of the X-chart are calculated in a similar way as for an X-bar chart by setting the control limits at the centerline ± 3 standard deviations. The control would be set at

$$\mu \pm 3\sigma$$

The average of the observations $\bar{x}$ is used to estimate the process mean μ. In the standards-given case, the known standard deviation would be used in the preceding equation. For the non-standards-given case, the standard deviation would be estimated. This estimate can be used either for the moving range or the standard deviation method.

For the moving range method the moving range is calculated as the absolute difference between the two concessive observations.

$$MR_i = |x_i - x_{i-1}|$$

The averages of the moving ranges are then averaged to get an average moving range $\overline{MR}$. The control limits are computed as

$$UCL_x = \bar{x} + 3\overline{MR}/d_2 = \bar{x} + E_2\overline{MR}$$

$$CL_x = \bar{x} \qquad (1.7)$$

$$LCL_x = \bar{x} - 3\overline{MR}/d_2 = \bar{x} + E_2\overline{MR}$$

where $E_2 = 3/d_2$.

$\overline{MR}$ is the estimate of the standard deviation and to make it an unbiased estimator it is divided by the d_2 a function of the sample size n. The values of d_2 are tabulated in Table C.1, Appendix C. The standard deviation σ can be estimated by, s, the standard deviation of all the trial observations. This can be used as an alternative method to the moving range method and the upper and lower control limits are set at $\bar{x} \mp 3s$.

1.2.4.2 Control Charts for Attributes

Quality characteristics of qualitative data cannot be suitably represented numerically, for such cases the items are inspected and classified as either conforming or non-conforming to the specifications on the quality characteristic. The quality characteristics of this type are known as attributes. *P*-chart or "control chart for fraction nonconforming" relates to the fraction of nonconforming or defective products produced by the manufacturing process. This is the most commonly used attributes control charts. The fractions used in

the p-chart are sometimes converted into percentages and refers the chart into the percentage chart. The other control charts for attributes are the *np* chart, *u* chart and the *c* chart. The *np* chart is used to plot number of nonconforming when the sample sizes are all equal. If the samples sizes are not all equal, then a *p*-chart is used instead of the *np* chart.

The p-chart and the np-chart are based on the Binomial distribution with parameters n and p

$$P(x)=\binom{n}{x}p^x(1-p)^{n-x},\quad x=0,1,2,\ldots,n$$

The mean and variance of the binomial distribution are $\mu=np$ and $\sigma^2=np(1-p)$.

Suppose we have a sample of n units and we are to inspect them for possible nonconforming items. We assume that the units are produced in a production process that operating in a stable manner and the successive units produced are independent, such that the probability that any unit will not conform to specifications is p. The fraction of nonconforming is defined as the ratio of the number of nonconforming items in the sample to the total number of items in the sample. If d is the number of nonconforming units in the sample, then d has a binomial distribution with parameters n and p, that is,

$$P(d=x)=\binom{n}{x}p^x(1-p)^{n-x},\quad x=0,1,2,\ldots,n$$

where the mean $E(d)=np$ and $V(d)=np(1-p)$.

The sample fraction nonconforming is defined as the ratio of the number of nonconforming items in the sample to the total number of items in the sample is computed as

$$\hat{p}=\frac{d}{n}$$

where $E(\hat{p})=p$ and $V(\hat{p})=p(1-p)/n$.

Supposing that the true nonconforming p in the production process is known or is standard value specified by management. Then from equation (1.0), the center line and control limits for the p chart are defined as

$$CL = p$$
$$UCL, LCL = p \pm 3\sqrt{\frac{p(1-p)}{n}} \tag{1.8}$$

When the process fraction nonconforming p is not known, then it must be estimated from observed data. The usual procedure is to select m preliminary samples, each of size n. Montgomery (2001) states that the general rule, m should be 20 or 25 Then if d_i nonconforming units in sample i the fraction nonconforming in the i^{th} is computed as

$$\hat{p}_i = \frac{d_i}{n} \qquad i = 1, 2, \ldots, n$$

and the average of these m individual sample fractions nonconforming is

$$\bar{p} = \frac{\sum_{i=1}^{m} d_i}{m} = \frac{\sum_{i=1}^{m} \hat{p}_i}{m}$$

The statistic $\bar{p}$ is an estimate of the unknown nonconforming p. The control limits are computed as

$$CL = \bar{p}$$
$$UCL, LCL = \bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{1.9}$$

For a constant sample size it is also possible to plot on a control chart the number of nonconforming units, rather than the fraction nonconforming. This chart is referred to as np control chart. If the true fraction nonconforming p in the production process is known or is a standard value specified by management, then the control limits are defined as

$$CL = np$$
$$UCL, LCL = np \pm 3\sqrt{np(1-p)} \tag{1.10}$$

where the charting statistic is $n\hat{p}_i$ for each sample.

If the true fraction nonconforming p in the production process is not known, then the average of the m preliminary individual sample fractions nonconforming $\bar{p}$ is used and the control limits are defined as

$$CL = n\bar{p}$$
$$UCL, LCL = n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})} \qquad (1.11)$$

where the charting statistic is $n\hat{p}_i$, for each sample. If we have a signal on a p chart we will have also one in an np chart because of the relation between the two charted statistics. This may lead us to conclude that these charts are equivalent for a constant sample size.

The c-chart is used to plot the number of nonconformities when the sample sizes are equal, and c is the number of nonconformities in each sample not the fraction of nonconforming. The c-chart is used where the possibilities for the nonconforming are infinite and the probability of getting a nonconforming is very small. The u chart is used to plot the nonconformities or defects per unit. The samples need not be equal as in c chart. This type of chart is used for complicated assemblies where the possibilities for the nonconforming are infinite but a constant size is not possible. The u chart is sometimes referred to as a *standardized c chart*. Both the u-chart and c-chart are based on the Poisson distribution.

Suppose that defects or nonconformities occur in an inspection unit according to the Poisson distribution, then

$$P(x) = \frac{e^{-c}c^{x}}{x!} \qquad x = 0,1,2,\ldots$$

where x is the number of nonconformities and $c > 0$ is the parameter of the Poisson distribution. It should be recalled from the probability distributions that both the mean and variance of the Poisson distribution are the parameter c. If the true value of c in the process is known or is a standard value specified by the management, then a control chart for nonconformities with three-sigma limits would be defined as

$$UCL, LCL = c \pm 3\sqrt{c} \tag{1.12}$$

where the charting statistic $\bar{c}_i$ is the number of nonconformities in sample i.

If the true value of c is not known or there is no standard given in the production process, then the control limits are defined as

$$UCL, LCL = \bar{c} \pm 3\sqrt{\bar{c}} \tag{1.13}$$

where $\bar{c}$ is the average number of nonconformities in a preliminary samples examined for lack of control used to estimate the value of c.

If we are to set up a control chart based on the average number of nonconformities per inspection unit, we take x total nonconformities in a sample of n inspections units, then, the average number of nonconformities per inspection is μ calculated as

$$\mu = \frac{x}{n}$$

Since x is distributed as a Poisson random variable the control limits for the average number of nonconformities per unit are computed as

$$CL = u$$

$$UCL, LCL = u \pm 3\sqrt{\frac{u}{n}} \tag{1.14}$$

In this case the true value of μ is known or is specified by the management. If the true value of μ is not known then the control limits are computed as

$$CL = \bar{u}$$

$$UCL, LCL = \bar{u} \pm 3\sqrt{\frac{\bar{u}}{n}} \tag{1.15}$$

where $\bar{u}$ is the average number of nonconformities per inspection unit from the preliminary sample and it is used as an estimate of μ.

1.2.4.3 Control Charts for Variables versus Charts for Attributes

Control charts for variables require actual measurements for the items which are produces. On the other hand control charts for attributes are used in situations where one whishes to count the number of nonconforming items or the number of nonconformities in a sample. Advantages of attributes control charts over variables control charts are:

- Attributes charts can be used to different nonconformities at the same time, whereas a separate chart must be used for each quality characteristic with variables charts.
- For attribute charts inspection required may be much easier as compared to in variable charts. In attribute charts what is merely needed is to know if the inspected item meets the specified requirements, that items are just classified based on various quality criteria. This minimizes time-consuming measurement procedures.
- Attribute charts may be used for visual inspections for some attributes like correct color, required cleanliness etc. This make attribute charts to be more easily understood by managers who are not familiar with quality control procedures; therefore, attributes charts it may provide more persuasive evidence of quality problems to management.
- Attributes charts do not depend on an underlying statistical distribution.

On the other hand attributes charts requires more sample sizes as compared to variable charts. Variable charts uses sample sizes of 4 or 5, whereas attribute charts requires a sample sizes of at least 50. Also variable control charts are more sensitive than attribute control charts (see Montgomery, 2001, p. 329). Therefore, variable control charts may alert us to quality problems before any actual "unacceptable", as detected by the attribute chart, will occur. Montgomery (2001) refers to variable control charts as *leading indicators* of trouble that will sound an alarm before the number of rejects (scrap) increases in the production process.

1.2.5 Sensitivity Rules for Control Charts

As stated previously in section 1.2.2 control charts are the tools for recording observations made on a process and indicating whether the process is in or out of control, that is, the reasons for constructing control charts is to detect variation(s), seeking special causes and tracking common causes. All control charts have three basic components: a centerline, upper and lower statistical control limits. This is done by using the basic criterion of, is there at least one point out side the control limits. There are also supplementary criteria used sometimes to increase the sensitivity of the control charts to small changes in the level or variability of the process. This is done by superimposing a number of limits on the plot to partition the control chart into three zones A, B and C on each side of the center line. Each zone is 1-sigma wide, see Figure 1.5. The results of the product are then plotted onto the charts and fall into four zones, and the zone tests are used. The zone tests can be used for both the variables and the attribute type control charts. This helps in detecting variation and the sooner a process variation is detected by a control chart the sooner the manufacturer can initiate the search for the cause.

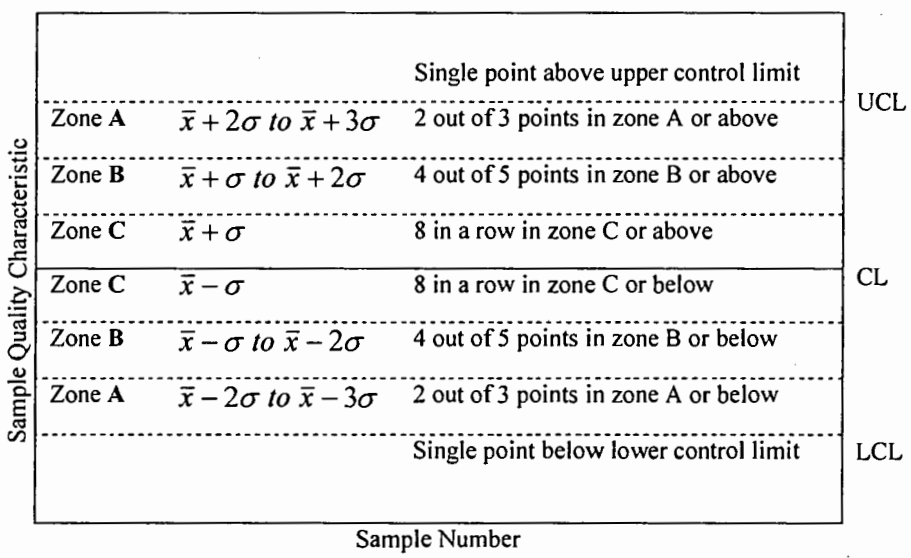


Figure 1.5 Control chart with marked "zones"

In applying the zone tests each half of the chart is considered separately and the decisions are taken using either the Western Electrical Alarm Rules or Alarm Rules from Duncan's *Quality Control and Engineering Statistics* or Nelson's Alarm Rules. The most famous used rules in decision are the Western Electric Alarm Rules and they are summarized as follows:

- A single point outside three-sigma units away from the center line signals that there is a cause of variation.
- Two out of any three consecutive points outside the two-sigma warning limits on one side of the center line but still inside the control limits, in zone A also signals the presence of a cause.
- Four out of any five consecutive points outside the one-sigma limits on one side of the center line in zone B indicates that there is a cause of variation.
- A run of eight consecutive points on the same side of the center line, in zone C indicates the presence of cause of variation.

Two other possible set of rules referred to as warning lines rules, taken from A.J. Duncan's excellent *Quality Control and Engineering Statistics* and the other published by Nelson in the *Journal of Quality Technology* in 1984, are in Table 1.1 and Table 1.2 respectively, are listed below.

Table 1.1 Alarm Rules from Duncan's *Quality Control and Engineering Statistics*

- A single point outside 3-sigma control limits.
- A run of 7 consecutive points up, down or on one side of the center line.
- 2 consecutive points outside the 2-sigma warning limits on one side of the center line but still inside the control limits.
- 4 consecutive points outside the 1-sigma limits on one side of the center line.
- "Obvious" cycles up and down

Table 1.2 Nelson's Alarm Rules from the *Journal of Quality Technology*

- A single point outside 3-sigma control limits.
- 9 consecutive points on one side of center line
- 6 consecutive points increasing or decreasing
- 14 consecutive points alternating up and down
- 2 out of any 3 consecutive points outside the 2-sigma warning limits on one side of the center line but still inside the control limits.
- 4 out of 5 consecutive points outside the 1-sigma limits on one side of the center line
- 15 consecutive points inside 1-sigma limits
- 8 consecutive points with none inside 1-sigma limits

For the range chart to be more sensitive the warning lines could also imposed in the chart. When the data is normally distributed the control limits on the range chart would be asymmetrical about the mean range since the distribution of sample ranges would be a positively skewed distribution. The formulae for setting the action and warning lines on the Range chart would be as follows:

Upper Action Line at: $D^{1}_{0.001}\overline{R}$

Upper Warning Line at: $D^{1}_{0.025}\overline{R}$

Process or Grand mean at: $\overline{R}$

Lower Warning Line at: $D^{1}_{0.975}\overline{R}$

Lower Action Line at: $D^{1}_{0.999}\overline{R}$

The values of $D^{1}_{0.001}$, $D^{1}_{0.025}$, $D^{1}_{0.975}$ and $D^{1}_{0.999}$ are given in Hartley's Constants and Constants used in design of Control Charts for Range Table at Appendix C, Table C.2.

1.2.6 Control Charts for Unequal Sample Sizes

When the samples plotted in the control chart are not of equal size, then the upper and lower control limits around the center line (target specification) cannot be represented by a straight line. There are three ways of dealing with this situation.

- **Average sample size.** If we want to maintain the straight-line control limits, then we can compute the average *n* per sample across all samples, and establish the control limits based on the average sample size. This procedure is not "exact," however, as long as the sample sizes are reasonably similar to each other, this procedure is quite adequate.
- **Variable control limits.** Alternatively, we may compute different control limits for each sample, based on the respective sample sizes. This procedure will lead to *variable* control limits, and result in step-chart like control lines in the plot. This procedure ensures that the correct control limits are computed for each sample. However, one loses the simplicity of straight-line control limits.
- **Stabilized (normalized) chart.** The best of two worlds (straight line control limits that are accurate) can be accomplished by standardizing the quantity to be controlled, the mean, proportion, etc. according to units of *sigma*. The control limits can then be expressed in straight lines, while the location of the sample points in the plot depends not only on the characteristic to be controlled, but also on the respective sample size *n*. The disadvantage of this procedure is that the values on the vertical (*Y*) axis in the control chart are in terms of *sigma* rather than the original units of measurement, and therefore, those numbers cannot be taken at face value.

1.3 Rational Subgrouping

Another issue that needs some general consideration before turning to the details of making specific control charts is the matter of sampling, how one should go about gathering data to be used in control charting. This is referred to as rational subgrouping or rational sampling introduced by Shewhart (1931) and discussed in standard texts such as

Hawkins and Olwell (1998), and Montgomery (2001). A subgroup is a small set of observations on process output taken closely together in time. A rational subgroup is a sample of items or measurements selected in a way that minimizes variation among the items or results in the sample, and maximizes the opportunity for detecting variation between the samples, and all sources of variation contributing to within subgroup variation should also be sources of variation contributing to between subgroup variations. Generally speaking subgroups should be selected in a way that makes each subgroup as homogeneous as possible and gives the maximum opportunity for variation from one subgroup to another. With a rational subgroup, assignable or special causes of causes of variation are not likely to be present, but all of the effects of random or common causes are likely to be shown. In setting up X-bar or R control charts, the first step is based on selection of the samples. When control charts are applied to production processes, the time order of production is a logical basis for rational subgrouping because it allows us to detect assignable causes that occur over time.

The size of the rational sample is governed by the following characteristics:

- Subgroups should be subject to common-cause variation. To achieve this, the sample sample/subgroup size of 4 to 6 is most reasonable for variables charts and samples of 50 to 100 for attributes charts.
- Subgroups should ensure the presence of a normal distribution for the sample means. The construction and interpretation of the $\bar{x}$ chart are based on the assumption that the X values follow the normal distribution
- Subgroups should ensure good sensitivity to the detection of special or assignable causes.
- Subgroups should be enough to be economically appealing from a collection and measurement standpoint.

In some applications, observations may become available individually, one by one, with relatively long time period between each one, so that concentrated sampling is not feasible. This is more practical in many chemical and process industries, observations taken very close together may be highly correlated, so that taking n=4 (use formulae object function) observations at the same time may result in essentially redundant

measurements. In this situation, it would be better to instead to spread out four observations over a stipulated time period to reduce the effect of the high correlation. Reynolds and Stoumbos (2004) stated that the necessity of using dispersed sampling arises when an individual observation is available for every item produced (100% inspection). They further highlights that this situation is becoming more common with sophisticated measuring systems and automated data acquisition technologies used these days.

The use of disperse sampling appears to violate the rational subgroups concepts because it greatly increases the probability that a shift will occur within a sample. The rational subgroups concept was originally developed in the contexts the Shewhart control charts were used. The Shewhart control charts depends only on current data, but the statistics plotted on CUSUM and EWMA charts are functions of current and past data. This means that after a shift occurs, the CUSUM and EWMA statistics will likely contain out-of-control data obtained after the shift, plus in-control data obtained after the shift. Reynolds and Stoumbos (2004) further argued that CUSUM and EWMA charts seem, in some sense to violate the rational subgroup concept and they raise questions about the applicability and generality of the rational subgroups concepts. The CUSUM and EWMA control charts will be discussed in Chapters 2 and 3 respectively.

1.4 Runs in Random Data and an Average Run Length

The theory of runs is very important in analyzing a control chart. A run is a succession of items of the same class. When the output of a process, for example, lets consider series plotted about its mean, a point above the mean may be considered belonging to one class, and a point below the mean may be considered as belonging to the other class. The points exactly equal to the mean are ignored. In such cases we talk of a run above the mean or a run below the mean. There are also "run up" and "run down". A succession of increases in value is a run up and a succession of decrease is a run down. In general a run is defined as a sequence of observations of the same type. Any type of run whether *run up* or *run*

down of eight or more points is a signal of an out-of-control condition. These are explained in great details in the statistical quality control text, Duncan (1986)

The run length provides an important measure of the effectiveness of a chart that is being used to control the quality of the current output. One of these performance measures is measure is average run length (ARL). Average run length can be considered in two parts for the in-control and out-of-control. ARL for the in-control is the average of a run of in-control points that follows immediately after there has been a specified change in the process. ARL for the out-of-control is the average of a run of out-of-control points that follows immediately after there has been a specified change in the process An average run length curve for a chart used to control the quality of current output can be derived if we can calculate that a given chance in a process, say an increase in the process average, will be detected. A commonly used ARL is the in-control ARL and it is mostly used in CUSUM and EWMA control charts.

1.4.1 Computing an Average Run Length

In computing an ARL we consider the general control chart where subgroups of size n are taken from the sample. For each subgroup n_i a statistic T is computed and plotted against the corresponding time. The control limits are computed and inserted in the chart, these are used calculating the probability that the statistic plots outside the control limits for any subgroup, this is denoted by p_d and

$$p_d = P(T > UCL \text{ or } T < LCL)$$

Let Y be the number of subgroups until the event that the statistic plots outside the control limits occurs. The distribution of Y under the assumption that the subgroups are independent of each other and the process is stationary is computed as follows:

$$P(Y=1)=p_d$$
$$P(Y=2)=(1-p_d)p_d$$
$$P(Y=3)=(1-p_d)^2 p_d$$
$$P(Y=j)=(1-p_d)^{j-1} p_d$$

Thus Y has a geometric distribution with parameter p_d. The average run length is defined as E[Y], which is a well-known result that

$$E[Y]=\sum_{j=1}^{\infty} j(Y=j)=\frac{1}{p_d} \tag{1.16}$$

Average run length for a given mean shift and subgroup size n for the X-bar chart can be computed as:

$$ARL=\frac{1}{\Pr\left(Z>L-\frac{\Delta\sqrt{n}}{\sigma}\right)+\Pr\left(Z>L+\frac{\Delta\sqrt{n}}{\sigma}\right)} \tag{1.17}$$

where

Z is N(0,1)

Size of the mean shift $=\Delta=\mu_{\text{'in-control'}}-\mu_{\text{'shifted'}}$

Process variance $=\sigma^2$

'in-control' mean $=\mu$

$L\sigma$ control limits $\mu+L\sigma/\sqrt{n}$

The sample size for a given ARL, and mean shift is computed as:

$$n=\left(\frac{\left(Z_{p_d}-L\right)\sigma}{|\Delta|}\right)^2$$

The ARL for the proportion of nonconforming when the process is in-control for inspected subgroups of size n is computed as:

$$ARL=\frac{1}{p_d}$$

where $p_d = P(\hat{p} > UCL) + P(\hat{p} < LCL)$

$\hat{p} = d/n$ as in subsection 1.2.4.2.

$\hat{p} = d/n$ is the observed proportion nonconforming and d is the number of nonconforming in a particular subgroup. The upper and lower control limits are respectively computed as $UCL \text{ and } LCL = p \pm L\sqrt{\frac{p(1-p)}{n}}$.

During the process the observed proportion nonconforming might change, in this case it is said that the process has shifted to p_s. For this process shift the control limits remain at UCL and LCL, the ARL is calculated as

$$ARL = \frac{1}{p_d}.$$

and $p_d = P(d > nUCL) + P(d < nLCL)$

Where d is binomial with parameters n and p_s

For large n we can use the normal approximation:

$$p_d = P\left(Z > \frac{(p-p_s)\sqrt{n} + L\sqrt{p(1-p)}}{\sqrt{p_s(1-p_s)}}\right) + P\left(Z > \frac{(p_s-p)\sqrt{n} + L\sqrt{p(1-p)}}{\sqrt{p_s(1-p_s)}}\right)$$

where Z is N(0,1)

1.5 Literature Review

As stated before, statistical process control charts are used to monitor for changes in an industrial process, such as manufacturing process. There are different types of control charts used to help in identifying process variations and their causes. Measurements from a process are observed and recorded. Quite often measurements are collected from a subgroup. From these data a control statistic or target value is computed and compared to the control limits. This is done so as to guide the manufacturer to take the appropriate action to rectify a problem or improve the process so as to produce good quality products. This might also enable the management to take action, which could prevent a future

occurrence of a present cause of variation. A dominant type of statistical process control chart called the Shewhart chart commonly known as control chart was proposed in the 1920s which monitors a process to detect the causes and presents the results graphically. From then over years there are further improvements and modifications to the control charts topics and characteristic, paying attention for quality of manufactured products and the fact that an important performance characteristic of the type of control chart is the speed at which it reacts to changes in a process. Many authors contributed to statistical process control theory and methodology, for example, Braverman (1981), Brown and Wetherill (1991), Chiu (1975), DeVor et al (1992), Guenther (1977), Xia (1992), Howell (1949), Wei et al (2002), Johnson (1961), Johnson and Leone (1964), Khan (1981), Klein (2000), Lucas (1985), Levinson (2004), Makrymichalos et al (2005), Mcquater et al (1995), Nelson (2003), North (1982), Pan (2002), Ryan (1989), Sefik (1998), Snee (1990), Steiner (1999), Tien (2005), Vardeman et al (1985), Vander Wiel (1995), Wardel Don et al (1992), Westgard et al (1977), Wetherill and Brown (1991), Wieringa (1999), Woodal (2000), Yeh et al (2004), Yashin (1993), and many others. In our study of the performance, usefulness, improvements and modifications of control charts we are also paying attention to the use of the cumulative sum (CUSUM) control charts and the exponential weighted moving average (EWMA) chart, used as a complementary tool to the X-bar chart, in a way that the small and large changes could be detected. This leads us to review these charts and presenting them in connection to our study. More reviews on the CUSUM control charts and EWMA charts are given in Chapter 2 and 3 respectively.

Shewhart control charts are used to monitor the process to detect causes of variation, they use the 3-sigma this was seen not to be too sensitive to small process changes so more improvements had to be made to make control charts more sensitive. To make the Shewhart chart sensitive to small process changes, quality improvement practitioners usually apply the so-called supplementary run rules see for example Montgomery (2001). In detecting small shifts in a process mean, it was shown by Champ and Woodall hat even with supplementary rules, the Shewhart chart is not as sensitive as the cumulative sum (CUSUM) chart proposed by Page in 1954. Ewan and Kemp (1960) also described the CUSUM control chart procedure for detecting a shift in the process mean as being

more effective in detecting relatively small shifts than the X-bar chart. Ewan (1963) and Geol & Wu (1973) presented a procedure for economical design of the CUSUM chart to control the process average based on a minimum cost criterion. They developed a model for a two-state process, the state being in-control and out-of-control. The model takes into consideration the parameters of the CUSUM chart and the various cost and risk factors associated with the process being controlled. Ewan (1963) discussed "when and how to use CUSUM charts" by outlining the various types of continuous graphical control schemes available and the type of process for which CUSUM control charts are most appropriate. He further explored what decision procedure exactly to be used basing on the information obtained from the process and when to alter the target value.

Another statistical process control chart to detect small shift is the exponential weighted moving average (EWMA) control chart introduced by S.W. Roberts in 1959. Hunter (1986) discussed the differences among the Shewhart, CUSUM and the EWMA control charts. Shewhart control chart depends entirely on the last demarcated point. The CUSUM chart attributes the same weight for the whole sequence, from the oldest to the most recent one. The EWMA chart grants larger weight for the most current information and smaller weight for the most remote ones. Montgomery (2001) presented the exponentially weighted moving average control chart as, a good choice when there is interest in detecting small shifts. Other authors as Crowder (1987), Lucas and Saccucci (1990), and Steiner (1998) also presented the exponential weighted moving average control chart, as a good choice to detect small changes in the process average. Crowder (1987) presented tables for the ARL to the EWMA chart to monitor the average standard deviation, for numeric procedure, using general integral equations, and also using the assumption that the observations are normally distributed. In conclusion Crowder (1987) found that the integral equation approach also extends easily to distributions that are non-normal, an important feature that allows the use of the approach when studying control procedures for process parameters other than a process mean.

Lucas and Saccucci (1990) made an extensive study comparing the properties of the EWMA and the CUSUM control charts. They evaluate the properties of an EWMA

control scheme used to monitor the mean of a normally distributed process that may experience shifts away from the target value. They consider several enhancements to EWMA control scheme these include a fast initial response feature and a combined Shewhart EWMA. The fast initial response feature makes the EWMA control scheme more sensitive to the start–up problem. The Shewhart and EWMA are combined to provide protection against both the large and small shifts in the process. These enhancements work as well for the EWMA control schemes as they do for CUSUM control schemes. The EWMA and CUSUM control schemes present very close results.

Steiner (1998, 1999) proposed a version of exponentially weighted moving average (EWMA) control charts that are applicable to monitoring the grouped data for the process shifts. The run length properties of the new grouped data EWMA chart are compared with similar results previously obtained for EWMA charts for variable data with those for CUSUM scheme based on grouped data. The EWMA charts are shown to be nearly as effective as variables-based EWMA charts and are thus an attractive alternative when the collection of variables data is not feasible. Also it was found that the grouped data EWMA charts are less affected by the discreteness that is inherent in grouped data than are grouped data CUSUM charts.

The form of the CUSUM control chart has been refined over the years to further increase its sensitivity to the fast initial response technique (FIR) by Lucas and Crosier (1982). Using the FIR features they conclude that the fast initial response feature for the CUSUM quality-control schemes permits more rapid response to an initial out-of-control situation than does a standard CUSUM quality-control schemes. Their comparison also shows that if the process starts out in-control, the fast initial response feature has little effect; if the process mean is not at the desired level, an out-of-control signal will be given faster when the FIR feature is used. Steiner (1998, 1999) derived the run length properties for EWMA control charts with varying control limits, the control limits are approximated using non-homogeneous Markov chains. Comparing the average run lengths of the EWMA with time-varying control limits and results previously obtained for asymptotic EWMA charts showed that using time-varying control limits is similar to the fast initial response (FIR)

feature suggested for the CUSUM charts. The ARL of the EWMA scheme with time-varying limits is substantially more sensitive to early process shifts especially when the EWMA weight is small. The variance of the EWMA test statistic is a function of time, time-varying control limits results in improved process shift detection capabilities if the process is initially out-of-control, or if it goes out-of-control quickly. The magnitude of the use of the benefits of using time-varying control limits over traditional asymptotic limits depends on the EWMA constant λ, and the size of the initial process. He also concludes that the proposed approach has the additional benefit of retaining the benefit of the EWMA chart that follows the two-sided detection of problems with a single chart.

Woodall (2000) states that in some cases EWMA and CUSUM charts are very useful, but are not meant to completely replace the Shewhart charts which can be used to detect a wider variety of effects due to assignable causes. He clarifies that it is frequently recommended that the limits of the Shewhart control be used in conjunction with a CUSUM or EWMA control chart. The Shewhart control charts CUMSUM control charts and EWMA control charts are described in great detail in the standard text on statistical process control, for example, Duncan (1986), Montgomery (2001) and Vardeman and Jobe (1999).

Reynolds and Stoumbous (2004) focus on the issue of concentrated versus dispersed sampling and the relationship with the rational subgroups concept. They show that when control chart combinations are based on n> 1 and concentrated sampling is a feasible option, concentrated sampling is better than dispersed sampling for detecting sustained shifts in u and sigma (to be from formulae sheets). This is consistent with the rational subgroup concept in that sampling should be done so that the shift falls between samples. When considering the choice of n, they assumed that the ratio n/d remains constant. Shewhart charts are very sensitive to the choice of n, with individual observations of n=1 best for large sustained shifts. Using an intermediate value of n, such as n=4, provides a reasonable compromise when overall performance is considered. The CUSUM charts are relatively insensitive to the choice of n, but using individual observations of n=1 is the overall best. They conclude that they best overall statistical performance for detecting

small as well as large sustained shifts in u (mean) or sigma is achieved when the samples consist of one observation (n=1), plotting is done after each of these observations and the CUSUM chart is used. In comparison of Shewhart control charts and CUSUM and EWMA charts, they conclude that the Shewhart control chart have inferior overall statistical performance.

1.6 Summary

Statistical process control is a sector of statistics of monitoring specific processes leading to great development and improvement of quality constantly. SPC uses many tools the core one being control charts introduced in the 1920's by Shewhart. These charts give a graphical appearance of the process giving the ability to any manufacture or service provider with or without the knowledge of statistics to immediately understand if the process is under control or not. Control charts helps for recognizing the type and causes of variation in the process. If there are any variations they are reduced or eliminated to bring the process into a state of control about a target value. This brings process to production goods of the required quality.

Research in the area of control charts has been active for over eight decades. There is an increasing interest for this statistical process control tool and it has proved its value in practice, this has been discussed in the literature review section. The improvements on control charts have been of introduction of sensitivity or warning lines, development of the Cumulative Sum charts, Exponential Weighted Moving Averages charts and combinations of these charts.

Chapter 2 Basic Theory on The Cumulative Sum Control Charts

This chapter defines CUSUM control charts and briefly reviews the literature of the improvements made on CUSUM control charts. We focus more on the design and analysis of the types of CUSUM control charts for monitoring the process mean using one-sided and two-sided CUSUM schemes. We compare the effectiveness of using CUSUM charts against that of Shewhart charts. Computations procedures of average run lengths of a CUSUM using Seigmund's (1985) approximation and Markov Chain procedure is covered in section 2.4. The Fast Initial Response in CUSUM charts and combined Shewhart and CUSUM charts are also discussed.

Cumulative sum (CUSUM) control chart was proposed by Page (1954) as an alternative to the Shewhart control chart. CUSUM charts uses the all the available past sequentially accumulated data in order to detect out-of-control conditions in the process not only the last few samples as in Shewhart control charts. Improvement and application of CUSUM control charts were published in several papers of which we will discuss some. Ewan (1963) presented a procedure for the economic design of CUSUM charts to control mean of a process with a normally distributed quality characteristic. He developed the cost model for a two state process of being in-control and out-of-control. The model takes into account the parameters of the CUSUM chart and the various cost and risk factors associated with the process being controlled. Ewan (1963) further reviews and describes cumulative sum charts with special emphasis on when and how to use a CUSUM charts.

Geol and Wu (1973) presented a procedure for the economic design of CUSUM charts to control the process average with a normally distributed quality characteristic. They derived a model which gives the long-run average cost as function of both the design parameters of the chart and the cost and risk factors that are associated to the process. They used the "pattern-search" techniques in determining the sample size, the sampling interval and the decision limit. Further they investigate the effects of changes in the

design parameters. Lucas (1973) proposed a modified control limits for the V-mask to be used in the CUSUM chart. In the modified V-mask scheme he included the parabolic-shaped mask having parameter size constant P which tempt to a V-mask having slope k. This provided better performance when the process goes through a large change in the average from goal conditions. Reynolds (1975) presented an approach of the approximation to the average run length for the CUSUM control charts to signal the out-of-control points. He approximated ARL using an analogy (similarity) among the procedure of the cumulative sum control charts and the sequential probability ratio test for identically normal random variables and independently distributed. In his approximation he also uses a Brownian motion approximation to the CUSUM and this does not require the normality assumption. He then uses the ARL obtained from the approximation to determine the optimal choice of parameters to minimize the ARL at a specified deviation from control, subjected to a fixed ARL value when the process is in control. Johnson and Bagshaw (1975) developed another approximation to the CUSUM which allows to study the run length distribution after a change in level occurred. They determine an approximation to the run length distribution of the one-sided CUSUM test for an increase in the mean by using the Wiener process to the cumulative sums. They concluded that the CUSUM test is not powerful when the observations are not independent. They largely emphasized their conclusion from the results they got of the effects obtained in the distribution of the number of samples necessary to signalize the out-of-control situation when the presence is correlated. Lucas and Crosier (1982) describe the properties of the fast initial response (FIR) for cumulative sum quality-control schemes for controlling a normally distributed process mean. Woodall (1985) presented a method to project quality control charts, based on its statistical performance, using the values from the specified parameters for areas under and out-of-control.

2.1 One–sided Decision Scheme for CUSUM Chart

Like other control charts cumulative sum control charts is primarily used to maintain a current control of the process. This is done by plotting the cumulative sum of the deviations between each data point in the process when using CUSUM charts. For all types of CUSUM charts whether one-sided or two-sided the x-axes of the chart are time based, so that the charts show a history of the process. For this reason, we must have data that entered in the sequence from which it was generated; that is, the data must be time-ordered. If this is not the case, then trends or shifts in the process may not be detected, but instead attributed to random (common cause) variation.

A cumulative sum chart was firstly devised for a protection against a shift in the process mean in a single direction, referred to as one-sided decision scheme of CUSUM chart. This was done by plotting the cumulative sum of "scores" representing the quality measurements. The chart was used to detect an upward shift in the process average. The decision when a change had occurred was made by using the rule of comparing the last point plotted with the previously plotted point. If the difference between the two points exceeded a specified quality H, it was concluded that a change has occurred. The system used in scoring quality was done in a way that the mean sample path on the chart was downward when the quality was satisfactory and upward when unsatisfactory.

When the one-sided scheme of CUSUM chart is used to detect upward shift in the process mean there is no interest in using it for record keeping, this procedure leads to certain difficulties when applied. These difficulties are; if the quality remains good, the trend of the CUSUM chart is negative and the frequent revision of a chart downward is necessary, and, the chart may pick up small shifts in the process mean that may not be of any importance. These reasons had led to modification of the CUSUM procedure for one-sided scheme on the upper side by the adoption of a reference value k. In this modified procedure the qualities plotted on the chart are the cumulative sums of $\bar{X} - k$ where k is a value greater than the value for $\bar{X}$. In this case only positive and the first negative

value of the sums are used. By the time the cumulative sum falls below zero, a new series is started. When the cumulative sum of $\bar{X}-k$ exceeds the decision interval h, it is decided that the process mean has shifted above k the modified procedure brings to attention to the upwards shifts in the process mean that are considered to be of importance only. A one-sided scheme can also be used to detect shifts in the negative direction. When detecting the negative direction k will be less than the target value and the test will be whether $\sum(\bar{X}-k)$ has a negative value that falls below decision interval negative $h\left(^{-}h\right)$.

2.2 Two–Sided Decision Schemes for CUSUM Chart

The basic approach in this form of CUSUM control chart scheme is to detect a shift in the process mean, by directly summing up all the deviations from the nominal or target value μ_0 of the previous information in the sequence of the sample values up to the current one. These deviations are then compared with a target value. This is done by plotting the cumulative sums of the deviations of the sample values from a target value. Using an example, taken form, (Montgomery, 2001) if μ_0 is the target for the process mean, $\bar{x}_j$ is the average of the j^{th} sample, and that the collected samples are of size $n \geq 1$, and then plotting the quality forms the CUSUM control chart:

$$C_i = \sum_{j=1}^{i} (\bar{x}_j - \mu_0) \qquad (2.1)$$

against the sample i. C_i is referred to as the cumulative sum up to the i^{th} sample inclusive, this implies that at any time point i we have i CUSUMs.

If the process remains in control at the target value, the cumulative sum, C_i fluctuate around zero. If the mean shifts upwards to some value $\mu_1 > \mu_o$, implying that the mean of the process μ_1 is larger than the target value μ_o then an upward or positive drift will develop in the cumulative sum statistic, C_i. On the other hand, if the mean shifts

downwards to some value $\mu_1 < \mu_o$, then a downward or negative drift in the CUSUM will develop. Therefore, if a trend develops in the plotted points either upward or downward, we should consider this as evidence that the process mean has shifted. The CUSUM control chart also gives rise to an out-of-control signal whenever its slope becomes too great, in either a positive or negative direction. Tabular or algorithmic CUSUM and CUSUM V-mask are cumulative sum control charts for the process mean that uses a two-sided decision scheme. These will be discussed in the next sub-sections.

2.2.1 Tabular or Algorithmic CUSUM for Monitoring the Process Mean

A standard tabular CUSUM chart for detecting changes in the process mean μ_0 is based on two separate one-sided CUSUM control statistics, the process mean should not be zero $(\mu_0 \neq 0)$. This is done by the tabular CUSUM accumulating deviations from μ_0 that are above target with the statistic C^+ and those that are below target with C^-. C^+ and C^- are called the ***one-sided upper cusum*** and the ***one-sided lower cusum***, respectively. If the nature of the out-of-control is not known, both C^+ and C^- are plotted on the same graph with time or sample number as the horizontal axis, Gan (1996). This is called the ***cusum status chart,*** see Figure 2.1. C^+ and C^- are the accumulation of deviations from the target, μ_0, that are greater than a specified value, k, and both are reset to zero upon coming negative. C^+ and C^- are computed separately, for detecting increases and decreases in the mean respectively as follows:

$$\begin{aligned} C_i^+ &= \max\left[0, x_i - (\mu_0 + K) + C_{i-1}^+\right] \\ C_i^- &= \max\left[0, (\mu_0 - K) - x_i + C_{i-1}^-\right] \end{aligned} \tag{2.2}$$

where the starting values are $C_0^+ = C_0^- = 0$ and both C_0^+ and C_0^- are reset to zero upon becoming negative. If either C^+ or C^- exceeds a value called the *decision interval, H,* the process is considered out-of-control and some corrective action or search for assignable causes of variation is initiated.

When designing a tabular CUSUM the values for the reference value K and decision interval H have to be chosen carefully, because these have a fairly large impact on the performance of the CUSUM. The reference value K in Equation (2.2) is often chosen halfway between the target value μ_0 and an out-of-control value μ_1 that is to be quickly detected. A shift is expressed in standard deviation units as $\mu_1 = \mu_0 + \delta\sigma$, then K is

$$K = \frac{\delta}{2}\sigma = \frac{|\mu_1 - \mu_0|}{2} \tag{2.3}$$

$$H = h\sigma,$$

that is, $K = S/2$ where S is the shift in the mean that we wish to detect. H is often chosen to be five times the process standard deviation σ, $H = 5\sigma$.

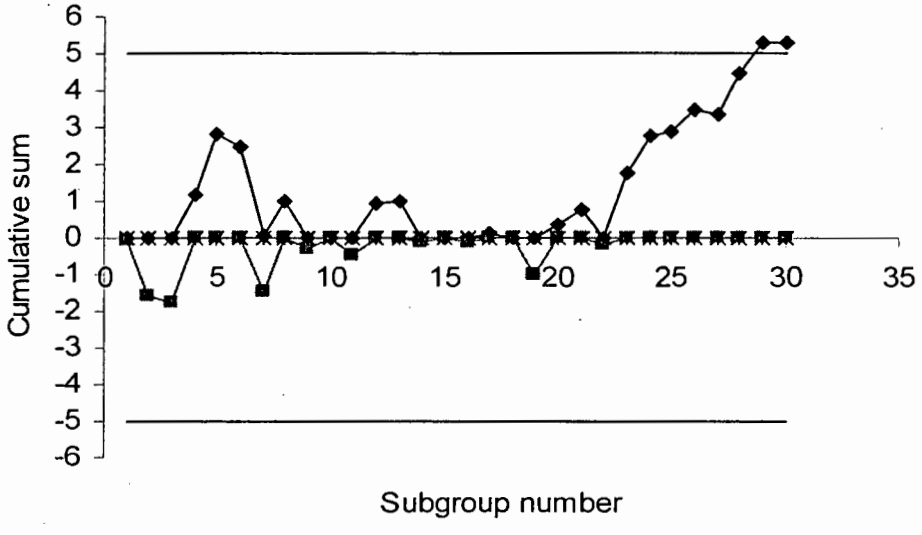


Figure 2.1 Tabular CUSUM status chart

Like other control charts when there is an out-of-control signal on the CUSUM control scheme one should search for the assignable cause and any corrective action required should be taken. At this stage the CUSUM control scheme is then reinitialized to zero. The CUSUM is helpful in determining when the assignable cause has occurred, this is shown by a worked out example in Montgomery (2001), and further explained in Section 2.5 of this thesis. The counters N^+ and N^- are used in the capacity of counting

backward from the out-of-control signal to the time period when the CUSUM lifted above zero to find the first period following the process shift. The counter N^+ records the number of consecutive periods since the upper-side CUSUM C_0^+ rose above the value of zero. The counter N^- records the number of consecutive periods since the lower-side CUSUM C_0^- rose above the value of zero

In some situations the adjustment to some manipulated variable is required in order to bring the process back to the target value μ_0 we may have to have an estimate of the new process mean following the shift. This is computed from:

$$\hat{\mu} = \mu_0 + K + \frac{C_i^+}{N^+} \quad if\ C_i^+ > H$$
$$\hat{\mu} = \mu_o - K - \frac{C_i^-}{N^-} \quad if\ C_i^- > H \tag{2.4}$$

It is usually recommended that the parameters K as the reference value and H as the decision interval be selected to provide good average run length (ARL) performance. $H = h\sigma$ and $K = k\sigma$, where σ is the standard deviation of the sample variable used in the formation of the CUSUSM. Usually $h = 4\ or\ 5$ and $k = \frac{1}{2}$ thesel generally provide a CUSUM that has good ARL properties against a shift of about 1σ in the process.

2.2.2 V-mask procedure

An alternative procedure to the Tabular CUSUM procedure is the V-mask. The V-mask became popular after the article "Control Charts and Stochastic Process" by G.A. Barnard (1959). The V-mask is applied to successive values of statistics CUSUM:

$$C_i = \sum_{j=1}^{i} y_j = y_i + C_{i-1} \tag{2.5}$$

The observation y_i denotes the standardized observation $y_i = \frac{x_i - \mu_0}{\sigma}$.

Like other control charts the CUSUM V-mask charts are interpreted by comparing the plotted points to critical limits. In determining whether the process average has shifted from or is different from the target value μ_0, each point on the CUSUM chart is tested on the time when it is plotted. The critical limits in a CUSUM V-mask chart are neither fixed nor parallel. The critical limits are inserted by superimposing the V-shaped mask on the CUSUM chart, with the vertex pointing horizontally forward set at a distance d ahead of the most recent observation. The angle between the obliques (not at right angles) and the horizontal is denoted by Θ. It should be noted that the angle Θ depends on the scale used in plotting. If all previously plotted points fall inside the angle of the V of the chart, the process is viewed as being in-control. If any of the previously plotted points are covered by the mask, it indicates that a shift has occurred. Points covered by the top of the mask indicate that there is a decrease in the process average, whereas the points covered by the bottom of the mask indicate an increase in the process average. Points are covered by the mask if they cross one of the arms of the V. If any of the previously plotted points are covered by the mask, the process is assumed to be out-of-control and a search for causes of variation should be initiated. In Lucas (1973), it is recommended to scale one unit of the horizontal axis of the graph of the V-mask CUSUM chart equivalent to $2\sigma_x$ on the vertical axis. With this scaling, in the mean gives a 45^0 trend on the CUSUM plot.

There have been modifications of the V-mask scheme that improve the performance of the CUSUM for large shifts include the semi-parabolic CUSUM mask, which replaces the last end of the vertex of the V-mask by a parabolic one, Lucas (1973), and "Snub-nosed V-mask control schemes", by, Rowlands *et al* (1982) which introduces the CUSUM envelope. The CUSUM envelope improved the V-mask's decision interval scheme on detection of both large and small changes in the mean.

Montgomery (2001) strongly advised that the quality engineers should not use the V-mask procedure because it has some disadvantages and problems associated with it. These are as follows:

- The V-mask is a two-sides scheme; it is not useful in not useful in one-sided process monitoring;
- The useful fast initial response initiated by Lucas and Crosier (1982) can not be applied to the V-mask;
- It is not clear and difficult to determine how far backwards the alarms of the V-mask should extend, this complicates interpretation of the V-mask;
- The major problem with the V-mask is when designing the V-mask using Johnson's approach (1961) because of the ambiguity associated with parameters α and β.

2.3 Estimating the Standard Deviation (σ) for CUSUM Charts

If there is sufficient back data to obtain a good estimate of the standard deviation (σ) then we can use it in setting the scales and decision rule in a CUSUM chart. When the observations are sampled in subgroups, the estimate is obtained from the pooled estimate of the within-group variance or from the average of the group ranges. In this case the x_i are presumed to be sample means of the samples of size n and accordingly the variance of x_i is

$$\sigma^2 = \frac{\sigma_I^2}{n}$$

where σ_I^2 is the variance of the individual observations.

If we deal with observations singly, then we use the estimate based on the calculation of successive differences

$$s^2 = \frac{1}{2(N-1)}\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2$$

This could also be computed by

$$\frac{8}{9}\sum_{i=1}^{N-1}|x_{i+1} - xz_i|$$

Because when the sample size n = 2 we have $d_2 = 1.128 \cong 1.125 = 9/8$.

2.4 Average Run Lengths of CUSUM Charts

CUSUM control schemes are usually evaluated by calculating their average run length (ARL). The run length of a control chart is defined as the sample number until a signal is issued by the chart, and the expectation of the run length is commonly known as the average run length. The ARL is the average number of individuals or samples taken before a given change in the process mean is detected or an out-of-control signal is obtained. The ARL is widely used as a performance measure of a control chart. The ARL should be large when the process is in control and small when the process is out-of-control. We would briefly compare the ARL values of the Shewhart chart and the CUSUM chart in Table 2.1.

Table 2.1 ARL Performance of the Shewhart control chart and the CUSUM with $k = 0.5$ *and* $h = 4$ *or* $h = 5$

Shift, in multiples of $\sigma_{\bar{X}}$	Shift - 3	P_a i.e. Probability that $\bar{x} \le UCL$	Shewhart ARL $= 1/(1-P_a)$	CUSUM ARL	
				h =4	h = 5
0	-3.00	0.99865	370.00	168	465
0.25	-2.75	0.99702	281.00	74.2	139
0.50	-2.50	0.99379	155.00	26.6	38
0.75	-2.25	0.98778	81.83	13.3	17
1.00	-2.00	0.97725	43.96	8.38	10.4
1.50	-1.50	0.93319	14.97	4.75	5.75
2.00	-1.00	0.84134	6.30	3.34	4.01
2.50	-0.50	0.69146	3.24	2.62	3.11
3.00	0	0.50000	2.00	2.19	2.57
4.00	1.00	0.15866	1.19	1.71	2.01
5.00	2.00	0.02275	1.02	1.31	1.69

P_a is the probability of a standard normal variable $\ge Shift - 3$ taken form White. J, Yeats. A & Skipworth G (1995), *Tables for Statisticians,* 3rd Edition, Dah Hua Printing Press Co. Ltd. The Shewhart ARL values were calculated using 3-sigma control limits and a constant standard value $\sigma_{\bar{X}}$ for the process standard

deviation. The columns of the CUSUM ARL values were, adapted from, (Montgomery 2001, p. 415). The $K = 0.5\sigma_{\bar{X}}$ and $H = 5\sigma_{\bar{X}}$, values are widely used in practice.

It should be notice that for detecting a $0.75\,\sigma_{\bar{X}}$ shift in the mean, with $n = 1$, the CUSUM ARL is 17, whereas a Shewhart control chart for individual measurements it would require 81.8 samples, on the average, to detect this shift. Critically analyzing the values in Table 2.1 it can be seen that, the Shewhart chart is more likely to signal a warning than a CUSUM chart if the process mean is on target. CUSUM chart improves the efficiency over the standard control chart in detecting changes in the region 0.5σ to 2.5σ. In the region between 0.5σ and 2.5σ the CUSUM can detect sudden and persistent change in the process average more approximately twice more quickly than the Shewhart control chart, this makes a CUSUM chart more affective at less expense. For shifts more than 2.5σ from the target value, Shewhart chart often detects the shifts more quickly than the CUSUM chart. Montgomery (2001) emphasized that it is better to use the Shewhart control chart for shifts larger than 2.5σ.

With further improvements of the CUSUM control charts Lucas in 1976 noted that a V-mask designed to detect a 1σ shift will detect it about four times as fast as a competing Shewhart control chart. A CUSUM chart is very effective visual presentation, it is ease with which changes in the mean level can be detected visually by change in the slope of the chart, and it should be noted that location of the change in the process is remarkably useful in helping to discover its cause. Since a CUSUM chart uses current and past data this provides a useful guide to the nature of the process variation, it implies that a CUSUM chart is likely to be useful tool in the future assessment of the similar data to the past. The CUSUM charts are even furthermore particularly effective with individual observations, that is, samples of size n = 1. This makes the cumulative-sum control chart a good candidate for use in the chemical and process industries where rational subgroups are frequently of size one, and in discrete parts manufacturing with automatic measurement of each part and on-line control using a microcomputer directly at the work center, Montgomery (2001).

A CUSUM does also have a disadvantage, on the fact that with the terms of the relation reversed, the CUSUM is more difficult to use in bringing an out of control process into a state of statistical control because patterns on the CUSUM are not interpretable (because successive CUSUM values are correlated), and interpretation or analysis of patterns is often very useful in identifying assignable cause when control charts are first applied.

There are several techniques that can be used to compute the ARL of a CUSUM depending on the type of variable that the process is dealing with. For continuous variables, direct derivative is used by solving the integral equations that arise in the analytical formulation of the problem or through simulation. For discrete variables a monogram constructed by Kemp (1961) is used for finding the appropriate values of ARLs. A monogram is used only when the variables are normally distributed. Brook and Evans (1972) proposed a method computing the ARL for both discrete and continuous variables based on a Markov Chain. In the case of discrete variables let D be an integer random variable of quality characteristic which is want to be in-control and the reference value k and decision boundary h be positive integers. Let S_n be an integer-valued random variable taking values from set $\{0,1,\cdots,h\}$. The scheme can be regarded as a random walk over the states $E_0, E_1, \ldots, E_h$, where E_h is an absorbing state. The transition probabilities for state E_i, $i = 0,1,\ldots,h$ are determined by the probability distribution of D as follows

$$p_{i0} = \Pr(D \le k-1)$$

$$p_{ij} = \Pr\{E_i \to E_j\} = \Pr(D = k+j-1), \quad 1 \le j \le h-1$$

$$p_{ii} = \Pr\{E_i \to E_h\} = \Pr(D \ge k+h-i)$$

Then the transition probability matrix P has the following form

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0j} & \cdots & p_{0,h-1} & p_{0h} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1j} & \cdots & p_{1,h-1} & p_{1h} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots & p_{i,h-1} & p_{ih} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ p_{h-1,0} & p_{h-1,1} & p_{h-1,2} & \cdots & p_{h-1,j} & \cdots & p_{h-1,h-1} & p_{h-1,h} \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}$$

It is obvious that all the row sums are one. An important relationship is that $(I-R)\mu^{(s)} = sR\mu^{(s-1)}$, $s = 2,3\ldots$ where R is the matrix obtained from the transition probability P by deleting the last row and column, i.e., those relate to the absorbing state E_h, I is an h by h identity matrix and $\mu^{(s)}$ is the vector of the s^{th} factional moments for the random variables $X_0, X_1, \ldots, X_{h-1}$. For $s = 1$ the equation becomes

$$(I-R)\mu = 1$$

where the vector **1** has each of its points h elements equal to unity. The first element of the vector μ gives the average rum length for a CUSUM chart starting from zero and in general the i^{th} element gives the mean of the run length distribution when staring from state E_i, $i = 0,1,\ldots,h$.

The other technique used to approximate the ARL of a CUSUM is Siegmund's approximation proposed by Siegmund in 1985. Siegmund's approximation is recommended by Woodall and Adams (1993) because of its simplicity. For **one-sided cusum**, C_i^+ or C_i^- with parameters *h* and *k*, Siegmund's (1985) approximation is:

$$ARL = \frac{\exp(-2\Delta b) + 2\Delta b - 1}{2\Delta^2} \qquad (2.4)$$

$$ARL = b^2, \text{ if } \Delta = 0$$

where: $\Delta = {}^{-}\ \delta^* - k$ for the upper one-sided CUSUM C_i^+,

$\Delta = \delta^* - k$ for the lower one-sided CUSUM C_i^-

$b = h + 1.166$

$$\delta^* = \frac{(\mu_1 - \mu_0)}{\sigma}$$

If $\Delta = 0$ we, use ARL=b^2. To get the ARL of the two-sided CUSUM from the ARLs of the one-sided statistics, say ARL^+ and ARL^- we use:

$$\frac{1}{ARL} = \frac{1}{ARL^+} + \frac{1}{ARL^-} \tag{2.5}$$

2.5 The Standardized CUSUM

The variable x_i of the CUSUM may be standardized before performing the calculations. The standardized value of x_i is y_i it is computed as:

$$y_i = \frac{x_i - \mu_0}{\sigma} \tag{2.6}$$

where μ_0 is the process goal value, with the assumption that the standard deviation σ is known, and x_i is the i^{th} observation for the CUSUM scheme. The i^{th} observation (x_i) might represent a single reading or the average of a number of observations from a designated routine sampling plan. A CUSUM scheme cumulates deviations more than k (standardized) units from the target or goal mean value. Thus k serves as the reference value of the scheme. This then leads to the standardized two-sided CUSUM defined as:

$$\begin{aligned} C_i^+ &= \max\left[0, y_i - k + C_{i-1}^+\right] \\ C_i^- &= \max\left[0, -k - y_i + C_{i-1}^-\right] \end{aligned} \tag{2.7}$$

where the first equation is for detecting positive mean shifts and the second is for detecting negative mean shifts. Standardizing the CUSUM has two main advantages, these are; firstly many CUSUM charts can now have the same values of k and h, and the choices of k and h are not dependent on σ. The other advantage is that the standardized CUSUM leads to a CUSUM for controlling variability. Table 2.2 provides an example, from Montgomery (2001) page 412, of a CUSUM in Tabular form where values of N^+ and N^- are given using the standardized period variables x_i ,having

$\mu_0 = 10$, $\sigma = 1$, $k = 0.5$ and the subgroup sample size $n = 1$. The value are N_i^+ and N_i^- are expressed as

$$N_i^{\pm} = \begin{cases} N_{i-1}^{\pm} + 1, & C_i^{\pm} > 0 \\ 0 & C_i^{\pm} \le 0 \end{cases} .$$

The standardized period value in Table 2.2 for $i = 1$ is $y_i = -0.55$ therefore, the statistic $C_1^+ = \max\left[0, y_i - 0.5 + C_{1-1}^+\right] = \max\left[0, -0.55 - 0.5 + 0\right] = 0$,and the statistic $C_1^- = \max\left[0, -0.5 - y_i + C_{1-1}^-\right] = \max\left[0, -0.5 + 0.55 + 0\right] = 0.55$ .Since $C_1^+ = 0$, $N_1^+ = 0$. Since $C_1^- > 0$, $N_1^- = N_0^- + 1 = 0 + 1 = 1$. The standardized period value for $i = 1$ is $y_2 = -2.01$ $C_2^+ = \max\left[0, y_i - 2.01 + C_1^+\right] = \max\left[0, -2.01 - 0.5 + 0\right] = 0$, and $C_2^- = \max\left[0, -0.5 - y_i + C_{1-1}^-\right] = \max\left[0, -0.5 + 2.01 + 0\right] = 1.56$.

Since $C_2^+ = 0, N_2^+ = 0$ and since $C_2^- > 0$, $N_2^- = N_1^- + 1 = 1 + 1 = 2$. The standardized period value for $i = 3$ is also less than zero, so $N_3^+ = 0$ and $N_3^- = N_2^- + 1 = 2 + 1 = 3$. The standardized period value for $i = 4$ is greater than zero, this results in $C_4^+ = \max\left[0, y_4 - 0.5 + C_3^+\right] = \max\left[0, +1.66 - 0.5 + 0\right] = 1.66$ and $C_4^- = \max\left[0, -1.66 - 0.5 + 1.77\right] = 0$. Therefore $N_4^- = 0$ and $N_3^- = N_2^- + 1 = 2 + 1 = 3$. The values of $N_i^{\pm}$ for other period i are calculated in the same manner.

Table 2.2 Tabular CUSUM Control Chart Example

Period i	x_i	y_i	Chart for increase			Chart for decrease		
			$y_i - 0.5$	C_i^+	N_i^+	$(-k - y_i)$	C_i^-	N_i^-
1	9.45	-0.55	-1.05	0	0	0.05	0.05	1
2	7.99	-2.01	-2.51	0	0	1.51	1.56	2
3	9.29	-0.71	-1.21	0	0	0.21	1.77	3
4	11.66	1.66	1.16	1.16	1	-2.16	0	0
5	12.16	2.16	1.66	2.82	2	-2.66	0	0
6	10.18	0.18	-0.32	2.5	3	-0.68	0	0
7	8.04	-1.96	-2.46	0.04	4	1.46	1.46	1
8	11.46	1.46	0.96	1	5	-1.96	0	0

9	9.2	-0.8	-1.3	0	0	0.3	0.3	1
10	10.34	0.34	-0.16	0	0	-0.84	0	0
11	9.03	-0.97	-1.47	0	0	0.47	0.47	1
12	11.47	1.47	0.97	0.97	1	-1.97	0	0

2.6 A CUSUM for Monitoring Process Variability

It is also possible to construct CUSUM control chart for monitoring process variability. In this case we assume that the mean is in-control and we let $s_1^2 s_2^2,\ldots$ be successive sample variances observed from a process based on a sample size n. The upper and lower CUSUM are computed separately as

$$C_i^+ = \max\left[0, C_{i-1}^+ + y_i - k\right]$$
$$C_i^- = \max\left[0, -k - y_i - C_{i-1}^-\right]$$

then plotted to on the same graph against i, $i = 1,2,\ldots$, where k is a constant, $y_i = \log\left(s_i^2\right)$, $C_0^+ = u$ for $0 \le u < h^+$ and $C_0^- = v$ for $^-h \le v < 0$. An increase in the variance is detected by the upper CUSUM chart and there is an out-of-control signal at the first i when $C_i^+ > h^+$. A decrease in the variance by the lower CUSUM chart and there is an out-of-control signal at the first i when $C_i^+ < -h^-$. The probability density function of $\log\left(s_i^2\right)$, when the measures of the quality characteristics are independently, independently and normally distributed, is

$$f(y) = \frac{\exp\left[ay - \exp(y)/\beta\right]}{\Gamma(a)\beta^a}, \quad -\infty < y < \infty$$

where $a = (n-1)/2$ and $\beta = 2\sigma^2/(n-1)$.

The variable y_i can also be standardized before performing calculations for setting up the CUSUM control charts for monitoring process variability. This was suggested by Hawkins in 1981 that a new standardized quantity v_i can be created as

$$v_i = \frac{\sqrt{|y_i|} - 0.822}{0.349} \tag{2.8}$$

The statistic v_i is sensitive to the mean and variance changes. Since the in-control distribution of v_i is $N(0,1)$ distributed approximately, two one-sided the standard deviation CUSUMs can then be defined as:

$$\begin{aligned} S_i^+ &= \max\left[0, v_i - k + S_{i-1}^+\right] \\ S_i^- &= \max\left[0, -k - v_i + S_{i-1}^-\right] \end{aligned} \tag{2.9}$$

where $S_0^+ = S_0^- = 0$ when FIR feature is not used, and the values of k and h are selected as in the case of the process mean monitoring. S_i^+ and S_i^- are referred to as *the scale CUSUM*. The scale CUSUM interpretation is similar to that of the CUSUM for the process mean. S_i^+ values will increase and increase eventually exceed h when there is an increase in the process standard deviation. When the standard deviation decreases, S_i^- values will increase and eventually exceed h.

2.7 The Fast Initial Response Feature in CUSUM Charts

Lucas and Crosier (1982) devised the fast initial response (FIR) to improve the sensitivity of CUSUM quality control scheme for controlling a normally distributed process mean. In implementing the FIR feature, the CUSUM is not reset to zero at start-up or after there is an out-of-control signal in the process. Instead the CUSUM is set to initial *headstart* values $S_0^+ = S_0^- = 0$ equal to some nonzero value, typically $H/2$. This is called the fifty percent *headstart*. In using the *headstart* procedure, a fifty percent *headstart* value is $C_0^+ = C_0^- = H/2$. If the process is in-control, this initial headstart has little effect; if the process is out-of-control, a signal will be given faster with nonzero values of C_0^+ *and* C_0^- than with a zero value. The main objective of the fast initial response is to reduce the ARL for the mean shifts, especially for the order of magnitude that one wishes to detect, without significantly reducing the in-control ARL.

probability transformation matrix R is created by letting the vector $\underline{L}$ represent the probability of a run length of length r. From the probability transformation matrix R it is found that

$$\underline{L}_1 = (I - R)1 \text{ and } \underline{L}_r = R\underline{L}_{r-1} = R^{r-1}\underline{L}_1 \tag{2.12}$$

The ARL's obtained using $\underline{u} = [I - R]^{-1}1$ are for the head-start values equal to the midpoints of the states. The ARL for the last state could be used as the ARL at $S_0 = h$.

Lucas and Crosier (1982) calculated the ARL's for five different-size Markov chains (t = 4, 7, 8, 9, 10). Extrapolation to t = ∞ was based on fitting the formula below by least squares:

$$ARL(t) + ASYMPTOTIC\ ARL + B/t^2 + C/t^4 . \tag{2.13}$$

In test cases when large S_0 values and very short ARL's are used, the asymptotic formula gave inconsistent results in that the asymptotic ARL did not increase with the increase of h values.

2.8 Combined Shewhart-CUSUM Scheme

Improving the detection of shifts or variations in the process can be done in several ways like; superimposing warning lines to the Shewhart chart, using the tabular CUSUM or the V-mask chart and adding Shewhart limits to the CUSUM chart. There is also a considerable improvement for detection of causes of variation by using the combination of Shewhart control scheme and CUSUM control scheme, because it is well known that; the Shewhart schemes are more favored when a large shift has occurred, and the CUSUM schemes are good in detection of small and moderates shifts. The combination of Shewhart and two-sided CUSUM schemes was proposed by Lucas (1982). The intension of this combination was to improve CUSUM scheme such that it becomes more competitive with Shewhart scheme for detecting large mean shifts. The result of this

combination is called a combined Shewhart-CUSUM scheme. Numerical results concerning the combined upper one-sided CUSUM–Shewhart schemes for Poisson data can be found in Yashchin (1985).

In a combined Shewhart-CUSUM the X-bar chart limits are used in conjunction with a CUSUM scheme. In developing the combined Shewhart-CUSUM control chart, the method requires that the values be defined for the mean μ_0, and the standard deviation, σ, for a particular control solution when analyzed by the analytical method of interest, and also for the level at which the CUSUM calculations are initiated, k, and the numerical control limit for the CUSUM, h.

To analyze the difference between the Shewhart-CUSUM scheme and the basic CUSUM scheme we will use the average run length values in Table 2.2.

Table 2.3 ARL Values for basic CUSUM and Shewhart-CUSUM schemes with $k = 0.5$ *and* $h = 5$

Mean shift, in multiples of $\sigma_{\bar{x}}$	0	0.25	0.50	0.75	1.00	1.50	2.0	2.50	3.00	4.00	5.00
Basic CUSUM	465	139	38	17	10.4	5.75	4.01	3.11	2.57	2.01	1.69
Shewhart-CUSUM ($y_i = 3.5$)	391	130.9	37.15	16.8	10.21	5.58	3.77	2.77	2.10	1.34	1.07

The ARL values were opted from Ryan, T. P. (1989). *Statistical methods for quality improvement.* New York: Wiley.

In a combined Shewhart-CUSUM an out-of-control signal can only be received not just from C_i^+ *and* C_i^-, but also from the y_i value. Table 2.2 $y_i = 3.5$ is selected as the threshold value so that an out-of-control signal would be received if either the absolute value of y_i exceeds 3.5 C_i^+ or C_i^- either or exceeds h.

Average run length of Shewhart, CUSUM and Shewhart-CUSUM, are shown in Figure 2.2 by plotting them against the deviation from the target in units of the sample standard error. It can be seen that Shewhart-CUSUM has a smaller y-intercept value than a CUSUM. The Shewhart-CUSUM curve becomes more steep than a CUSUM curve when the deviations greater than 2σ.

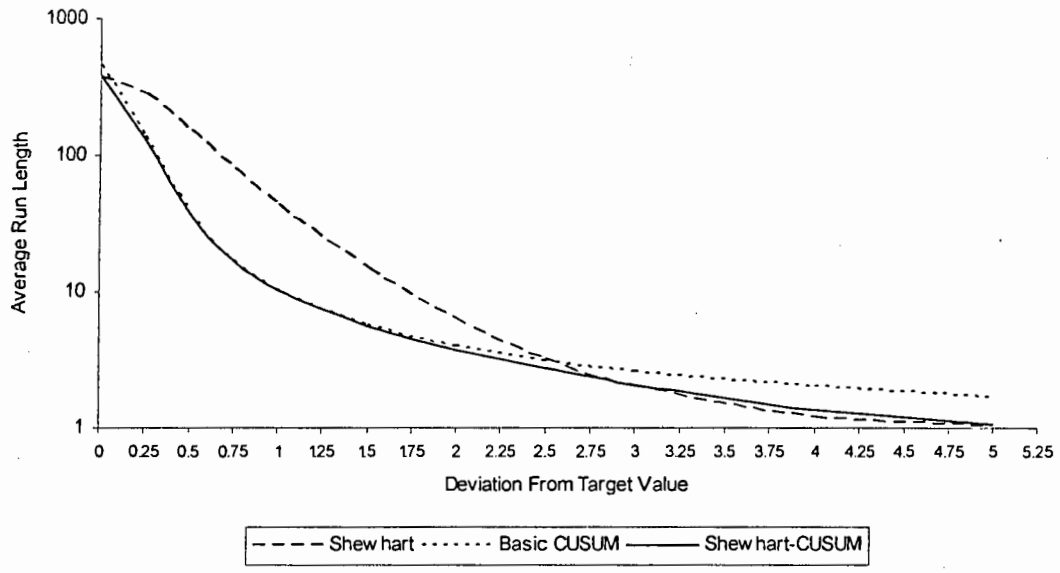


Figure 2.2 ARL curve of Shewhart, CUSUM and Shewhart-CUSUM chart

2.9 Summary

In this chapter, several formulae used for calculating the relevant variables used in cumulative sum control charts were presented. Construction and interpretation of different types CUSUM charts was explained. When comparing CUSUM control charts and a standard Shewhart control charts, CUSUM charts are of typical use in high cost of production or for test in the production process where to shift is common. The improvement of CUSUM chart by adding the limits of a standard control chart was presented. This forms a Shewhart-CUSUM chart that increases the sensitivity of control charts by reducing the average run length required for detecting the mean shift from the target value. The disadvantages of CUSUM charts are complex and hard to explain.

There are no exact formulae for average run lengths for CUSUM charts, they must either be determined empirically or by numerical solution. Comparative numerical figures for ARLs were quoted in Tables 2.1 and 2.2. We noted that CUSUM charts are notably quicker to pick up moderate changes of between 0.5σ and 2.5σ but that they have no advantages for gross changes.

Chapter 3 Exponentially Weighted Moving Average (EWMA) Control Charts

In Chapter 1 and 2, we discussed the Shewhart control charts and CUSUM control charts respectively. We have elaborated that Shewhart control charts only use the last sample to monitor the process. These charts are taken to have no memory because previous observations do not influence the probability of future out-of-control signals. The Shewhart control charts can quickly detect shifts in the mean if the shift is large, but they are not sensitive in detecting small shifts in the mean. To improve the Shewhart control charts and make it sensitive to small shifts in the process the so-called supplementary run rules were introduced. With more improvements to sensitivity of the Shewhart control charts Page (1954) proposed a CUSUM control chart, which made it clear that the so-called supplementary run rules were not so sensitive. The CUSUM control chart uses an unweighted sum of all previous observations. This chart has a long memory, so it is used to detect small level of shifts in the process, that is, it is sensitive to small changes in the process. Roberts (1959) introduced the exponentially weighted moving average (EWMA) control charts as an alternative method to the CUSUM control chart for detecting small levels of shifts in the process. He was more concerned with the control of the process mean. In EWMA control charts the process is monitored using *weighted mean* of all previous observations. This makes the EWMA control charts to also detect much smaller process shifts than a Shewhart control chart would. Like the CUSUM, the EWMA uses the previous observations, but the weight attached to data is exponentially declining, as observations get older and older.

There are other authors who made contributions in the improvement of the EWMA control charts model. Montgomery (2001) presented the EWMA control chart as a good choice when we are interested in detecting small changes in the process. He states that performance of EWMA control chart is, approximately, equivalent to the CUSUM control charts and it is easier to set up and operate the EWMA control chart. Hunter (1986) emphasized the differences among the Shewhart, CUSUM and EWMA control

charts. He called the geometric moving average control chart the EWMA control chart and indicated that the weighted moving average, as a means of dynamic process control, could predict the next process mean. Crowder (1987) presented the integral-equation approach that gives an exact expression for moments of the run length associated with the EWMA scheme. The integral-equation approach also extends to distributions that are non-normal, this feature allows the use of the approach when studying control procedures for process parameters other than a process mean. Lucas and Saccussi (1990) made several enhancements to the EWMA control schemes, one of them is a fast initial response (FIR) feature. A fast the fast initial response exponentially weighted moving average (FIR-EWMA) makes the scheme more sensitive at start-up with control limits narrowed further for the first few observations this promotes the performance of EWMA control charts. Steiner (1998, 1999) stated that time-varying control limits are more sensitive in detecting start-up shifts than fixed control limits. He used FIR to monitor start-up problems or control action after the previous out-of-control signal. Klein (2000) compared the simulation of the composite Shewhart-EWMA control chart with that of the Shewhart control charts with run rules. The results show that the composite Shewhart-EWMA control chart exhibits better ARL characteristics than the Shewhart control charts with run rules.

Form the review of EWMA charts we notice that EWMA has properties in some forecasting and quality monitoring applications. In this chapter we focus on the quality monitoring applications. In Section 3.1, we will define EWMA control charts for monitoring the process mean with the assumption that the sequentially used observations used are independent and give the formulae for calculating its control limits. Section 3.2 states some properties of the EWMA. The formulae used to compute the ARL of EWMA are explained in details in Section 3.3. Usually the properties of EWMA control charts are easer to explained basing on ARL values. In Section 3.4 we discuss the EWMA control charts with time-varying control limits. Fast initial response (FIR) is discussed in Section 3.6 as FIR is one of the enhancements of the EWMA control scheme. The ARL values of the Shewhart, CUSUM and EWMA are compared. The last section presents the formulae for combined Shewhart-EWMA.

3.1 EWMA Control Chart for Monitoring the Process Mean

For a clear definition of quality monitoring process Montgomery (2001) stated that the performance of the EWMA control chart is often equivalent to that of the CUSUM control chart. The EWMA is defined as:

$$\begin{cases} EWMA_0 = target \text{ (or starting value)} \\ EWMA_i = \lambda \bar{x}_i + (1-\lambda) EWMA_{i-1} \text{ for i} = 1,2,3,... \end{cases}$$

Generally $EWMA_i$ is substituted by Z_i, this defines EWMA as:

$$Z_i = \lambda \bar{x}_i + (1-\lambda) Z_{i-1} \tag{3.1}$$

where $\bar{x}_i$ is the mean of the sample of observations in time $i = 1,2,...$, λ is a smoothing parameter or weighting factor for the present observation and it takes the values of between 0 and 1, $(0 < \lambda \leq 1)$. The initial value, requested with the first sample $i = 1$, is the intended for the process, such that $Z_0 = \mu_0$. Z_i is the value of the EWMA after observation i, where the subscript i represents the observation number as well as an index of a point in time. Z_{i-1} is the value of EWMA of the preceding sample number or time x_i can be individually observed values from the process, or the sample averages or sample standard deviation obtained from a sampling plan or any other empirically estimated process parameter. Recursively substituting $\lambda x_{i-j} + (1-\lambda) Z_{i-j-1}$ for z_{i-1}, $j = 1,2,...,i-1$ in Equation (3.1), it can be shown that Z_i is a weighted average of all past current observations as:

$$Z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^i Z_0 \tag{3.2}$$

If the observations x_i are independent random variables with variance σ^2, then the variance of Z_i is:

$$\sigma_{z_i}^2 = \frac{\sigma^2}{n} \left(\frac{\lambda}{2-\lambda} \right) \left[1 - (1-\lambda)^{2i} \right] \tag{3.3}$$

The average of preliminary data is sometimes used as the initial value of EWMA, such that $Z_0 = \bar{x}$. The staring value Z_0 is the target mean, $Z_0 = \mu_0 = CL$ the center line. Finally the upper and lower control limits are defined as:

$$UCL, LCL = CL \pm L\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \tag{3.4}$$

In equation (3.4) the factor $L > 0$, is the extension of the control limits that have to be chosen by the designer of the control chart.

From equation 3.3 it can be seen that $\sigma_{z_i}^2$ increases over time, that is, as observation number or index of a point in time increase the variance of Z_i also increases. Hence also the control limits in equation 3.4 becomes wider. On the other hand, unless λ is very small, $\sigma_{z_i}^2$ converges very quickly to $\lambda/(2-\lambda)\frac{\sigma^2}{n}$ When constant limits are preferred, the computations will be based on the asymptotic standard deviation instead of the exact σ_{z_i}.

From Equation (3.4) we should note that the term $\left[1-(1-\lambda)^{2i}\right]$ approaches unit as i becomes big. Montgomery (2001)[made it clear and more comprehensible that after the control chart has been running for several time periods, the control limits will approach stable-state values, parallel to the center line, see Figure 3.1, given by:

$$CL = \mu_0$$

$$UCL, LCL = \mu_0 \pm L\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{(2-\lambda)}} \tag{3.5}$$

Usually, μ and σ will not be known in practice. Therefore, μ and σ in equation 3.5 will be replaced by the estimators $\hat{\mu}$ and $\hat{\sigma}$.

Figure 3.1 provides an example of the EWMA control chart with $n = 1$, $\lambda = 0.10$, $L = 2.7$, $\mu = 10$, and $\sigma = 1$ from Montgomery (2001) page 429. As with other control charts, EWMA charts are used to monitor processes over time. The charts' x-axes are

time based, so that the charts show a history of the process. For this reason, we must have data that is time-ordered; that is, entered in the sequence from which it was generated. If this is not the case, then trends or shifts in the process may not be detected, but instead attributed to random (common cause) variation. It can be seen that the EWMA provides a fast initial response of the chart to the process variation that occurs early after the process start up due to this a narrow bandwidth at the first observations.

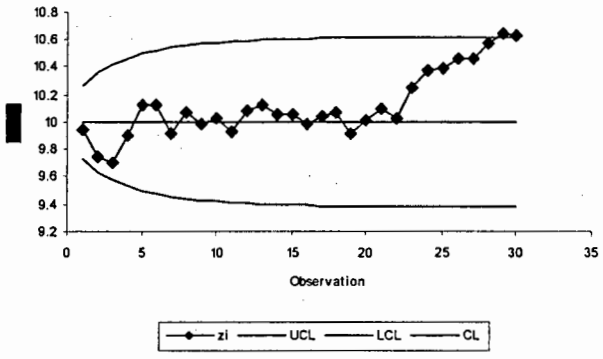


Figure 3.1 EWMA control chart

The choice of λ depends on the shift that one wants to detect so that ARL is as short as possible. When the value of λ is close to 0, the EWMA can detect small to moderate shifts in the process mean, when the value of λ is close to unity the EWMA can detect large shifts in the process mean. When choosing the value of lambda (λ) it is recommended to use small values, such as 0.2, to detect small shifts, and larger values, between 0.2 and 0.4, for larger shifts. When $\lambda = 1$ the EWMA is actually the X-bar chart, since when $\lambda = 1$ the control limits are

$$CL = \mu_0$$

$$UCL, LCL = \mu_0 \pm L\frac{\sigma}{\sqrt{n}}\sqrt{\frac{1}{(2-1)}} = \mu_0 \pm L\frac{\sigma}{\sqrt{n}}.$$

This is shown in Table 3.1 where the control limits are calculated using Equation 3.4 and changing the values of λ only. When $\lambda = 1$, $UCL = 12.7$ and $LCL = 7.3$ which is approximately equal to $UCL = 13$ and $LCL = 7$ the control limits of the X-bar chart of the same observations. The observations used are from Montgomery (2001) page 429.

Table 3.1 Calculations of Control Limits of EWMA chart

Observation i	x_i	$\lambda = 0.1$		$\lambda = 0.8$		$\lambda = 0.9$		$\lambda = 0.95$		$\lambda = 1$	
		UCL	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL	LCL
1	9.45	10.270	9.730	9.560	12.160	12.430	7.570	12.673	7.327	12.70	7.30
2	7.99	10.363	9.637	8.304	12.203	12.442	7.558	12.673	7.327	12.70	7.30
3	9.29	10.424	9.576	9.093	12.204	12.442	7.558	12.673	7.327	12.70	7.30
4	11.66	10.467	9.533	11.147	12.205	12.442	7.558	12.673	7.327	12.70	7.30
5	12.16	10.500	9.500	11.957	12.205	12.442	7.558	12.673	7.327	12.70	7.30
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
26	11.08	10.618	9.382	12.205	7.795	12.442	7.558	12.673	7.327	12.70	7.30
27	10.38	10.618	9.382	12.205	7.795	12.442	7.558	12.673	7.327	12.70	7.30
28	11.62	10.619	9.381	12.205	7.795	12.442	7.558	12.673	7.327	12.70	7.30
29	11.31	10.619	9.381	12.205	7.795	12.442	7.558	12.673	7.327	12.70	7.30
30	10.52	10.619	9.381	12.205	7.795	12.442	7.558	12.673	7.327	12.70	7.30

3.2 EWMA Control Chart for Monitoring the Process Variance

EWMA chart for detecting a shift in standard deviation was proposed by Wortham and Ringer in 1971, this is referred to as exponentially weighted mean square error (EWMS) defined as

$$S_i = \lambda(x_i - \mu_0)^2 + (1-\lambda)S_{i-1} \tag{3.6}$$

and $S_o = \sigma_o^2$ where λ is the smoothing parameter that lies between 0 and 1. S_o is the initial estimated value of the mean squared error. MacGregor and Harris (1993) extended the use of this EWMA-based statistics for monitoring the process variance that under normality the quality S_i/σ^2 is approximately distributed as $\chi^2(v)/v$ where the degrees of freedom of v depend on the parameter λ, the correlation of the x_i'v and the degrees of freedom associated with the initial value. Assuming that the process is on target and the variance is σ_o^2 then the control limits of S_i are the $\alpha/2$ and $1-\alpha/2$ percentiles of $\sigma_0^2\chi^2(v)/v$ distribution. When the observations are independently and normally distributed we could plot $\sqrt{S_i^2}$ and the corresponding control limits as

$$UCL = \sigma_0 \sqrt{\frac{\chi^2_{v,\alpha/2}}{v}}$$
$$LCL = \sigma_0 \sqrt{\frac{\chi^2_{v,1-(\alpha/2)}}{v}} \qquad (3.7)$$

McGregor and Harris (1993) argued that the EWMS statistic can be sensitive to shifts in both the process mean and the standard deviation. They suggest replacing μ_0 in equation (3.7) with an estimate $\hat{\mu}_i$ at each point in time. A usually used estimate of the true population mean at the i^{th} point in time turns to be EWMA statistic for the mean Z_i. They computed control limits for the usually addressed as the Exponentially Weighted Moving Average Variance (EWMV) defined as:

$$S_i^2 = \lambda(x_i - z_i)^2 + (1-\lambda)S_{i-1}^2 \qquad (3.8)$$

Another approach to monitoring the process standard deviation with the EWMA is by Crowder (1987). They proposed a different control chart based on $\ln(\hat{\sigma}_i^2)$. The scheme is

$$C_i = \max\left\{(1-\lambda)C_{i-1} + \lambda y_i \ln(\sigma_0^2)\right\}$$

where $C_0 = \ln(\sigma_0^2)$, λ is the usual constant taking of the values between 0 and 1 and $y_i = \ln(\hat{\sigma}_i^2)$. This statistic can be used to identify only upward shifts in the variance using the UCL in the case of independent observations given by

$$UCL = K\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left\{\frac{2}{n-1} + \frac{2}{(n-1)^2} + \frac{4}{3(n-1)^3} - \frac{16}{15(n-1)^5}\right\}}$$

where K is a constant chosen together with λ so as to achieve the desired ARL. If $L(u)$ is the ARL of this chart with u the starting value then

$$L(u) = 1 + L(0)F\left(\frac{-(1-\lambda)u}{\lambda}\right) + \frac{1}{\lambda}\int_0^{UCL} f\left(\frac{y-(1-\lambda)u}{\lambda}\right)L(y)dy$$

where $F(x)$ is the cumulative distribution function of the log-normal distribution and $f(x)$ is the probability distribution function of the log-normal distribution.

3.3 Average Run Length and Design of an EWMA Control Chart

EWMA control chart is very effective in situations where small changes in the process happen. The EWMA control chart has two parameters, the multiple of sigma in the control limits (L) and the value of λ. Their value is set based on requirements on the average run length (ARL) curve. The ARL is proportional to the amount of production from a process. Hence the design procedures of EWMA control scheme are usually based on the ARL properties of control schemes. "The ARL should be long when the process is operating near its target value and short when the process shifts to an unacceptable level", Lucas and Saccussi (1990).

Several studies were made for the properties of the ARL of EWMA control charts. Usually one of the methods is used to determine the properties of EWMA control charts, the Markov chain or the Integral equation approach. Robinson and Ho (1978) used a numeric procedure to determine the ARL, presenting several combinations for control limits, smoothing constant and level of change in the process mean. Crowder (1987) considered an integral equation approach for evaluating the run length distributions of the two-sided EWMA chart to monitor the average and standard deviation. He represents ARLs as the solution of a Fredholm integral equation of the second kind. Replacing the integral equation with a system of linear algebraic equations and solving it numerically obtain the solutions to the integral equation. This is done by using the successive values describes by

$$Q_i = (1-\lambda)Q_{i-1} + \lambda y_i, \quad 0 < \lambda \le 1, \quad i = 1,2,\ldots$$

And λ is a smoothing constant, Q_i is the value of the EWMA after the observation i, where the i subscript represents the observation number as well as an index of in time. Let the $L(u)$ be the ARL of a two sided EWMA chart for the mean given that the

EWMA starts with $Q_0 = u$. If the first observation y_1 is such that $\left|(1-\lambda u)+\lambda y_1\right|$ is greater than a specified constant h, a signal is given. On the other hand, if the relationship does not hold the run length continues to move from $(1-\lambda)u+\lambda y_1$ with $L\left((1-\lambda)u+\lambda y_1\right)$ representing and additional expected run length. Thus the relationship is computed as

$$L(u) = 1\cdot \Pr\left(\left|(1-\lambda)u+\lambda y_1\right| > h\right) + \int_{\{|(1-\lambda)u+\lambda y|\le h\}} \left(1+L\left((1-\lambda y)u+\lambda y\right)\right) f(y)dy$$

$$= 1+\frac{1}{\lambda}\int_{-h}^{h} L(y) f\left(\frac{y-(1-\lambda)u}{\lambda}\right)dy$$

Where $y_i's$ are assumed to be independent identically distributed observations with probability density function $f(\cdot)$ and $h, -h$ are the upper and lower control limits respectively. Using the above method the desired mean value is taken as $u = 0$ and the known process variance to be $\sigma^2 = 1$. Using the $Q_i's$ to monitor the process mean level, the process is said to be out-of-control if Q_i is too large or small. This implies that the conclusion that the mean has shifted at observation i would be made if $|Q_i| > h$ for h a specified constant.

Lucas and Saccussi (1990) consider the Markov chain approach in studying the run length distributions of the two-sided EWMA chart. They model the EWMA statistic as a continuous-state Markov chain and evaluate its properties by discretising the infinite-sate transition probability matrix. This procedure involves dividing the interval between the upper and lower limits into $t = 2m+1$ subintervals of width 2δ. The control statistic, Z_i, is said to in transient state j at time i if $S_j - \delta < Z_i \le S_j + \delta$ for $j = -m, -m+1, ..., m$, where S_j represents the midpoint of the j^{th} interval. The control statistic is said to be at an absorbing state (a) if Z_i falls out side the control limits. The process is in the absorbing state whenever Z_i is in a transient state and is assumed to be out-of-control whenever Z_i is in the absorbing state. The transient states are often referred to as in-

control states and the absorbing state is often is often referred to as the out-of-control state.

The run-length of an EWMA is determined by its initial probability vector and transition matrix. The initial probability vector is presented by

$$\begin{aligned} P_{\text{int}}^{T} &= \left(p_{-m},\ldots,p_{-1},p_0,p_1,\ldots,p_m\right) \\ &= \left(p^{T} \mid 0\right) \end{aligned}$$

where p_j represents the probability that Z starts in state j. It should be noted that p_a is equal to 0 since the control statistic is assumed to start in control. The transition probability matrix is given by

$$P = \begin{pmatrix} R & \left(I_{2m+1} - R\right)\iota \\ 0^{T} & 1 \end{pmatrix}$$

where I_{2m+1} is the identity matrix of order $2m+1$, ι is a column vector of ones column vector of zeros and R is a $(2m+1)\times(2m+1)$ submatrix containing the probabilities of going from one transient of the other. A typical element of R is denoted by p_{ij} and represents the probability that w_t goes from i state to state j in one step. Lucas and Saccussi (1990) approximate the entries of R by assuming that the EWMA equals S_i whenever it is in the state i. This yields:

$$\begin{aligned} p_{ij} &= \Pr\left(going\ to\ S_j \mid in\ S_i\right) \\ &\approx \Pr\left[\lambda^{-1}\left\{\left(S_j-\delta\right)-\left(1-\lambda\right)S_i\right\} < \lambda y_{it} \le \left\{\left(S_j+\delta\right)-\left(1-\lambda\right)S_i\right\}\right] \quad i,j=-m,-m+1,\ldots,m \end{aligned}$$

The probabilities of going from an in-control state i to the out-of-control state a are found by subtracting the entries in row i from 1. The i^{th} stage transition probability matrix P^i contains the probabilities that EWMA jumps from one state to another in step i steps. That is

$$P^{i} = \begin{pmatrix} R^{i} & \left(I_{2m+1} - R^{i}\right)\iota \\ 0^{T} & 1 \end{pmatrix}$$

it can be obtained by multiplying P i times with itself. From P^i, $\Pr(run\ length \le i)$ can be computed as

$$\Pr(run\ length \le i) = p^T \left(I_{2m+1} - R^i\right)\iota$$

and

$$\Pr(run\ length = i) = p^T \left(R^{i-1} - R^i\right)\iota$$

Using the above expression, ARL values based on $2m+1$ in-control states can be computed as follows:

$$\begin{aligned} ARL(2m+1) &= \sum_{i=1}^{\infty} i \Pr(run\ length = i) \\ &= \sum_{i=1}^{\infty} i p^T \left(R^{i-1} - R^i\right)\iota \\ &= \sum_{i=1}^{\infty} p^T R^{i-1}\iota \\ &= p^T \left(I_{2m+1} - R\right)^{-1}\iota \end{aligned} \tag{3.8}$$

As $(2m+1)$ approaches infinity, the discretized Markov chain approaches a continuous state Markov chain, and the approximation for the ARL of the EWMA gets better.

To design an EWMA control scheme Lucas and Saccusssi (1990) recommends the following procedures:

- To start of by specifying the desired in-control ARL and the shift in the process is to be detected quickly.
- Obtain the EWMA parameters that will result in the minimum ARL for the specified shift in the process using **Table 3.1**.
- The entire ARL profile for the EWMA should be evaluated to determine whether it provides sufficient protection against shifts.

Table 3.2 Optimal EWMA Control Schemes

Shift		In-control average run length					
		100	300	500	1000	2000	5000
0.5	λ	0.07-0.06	0.60-0.05	0.05	0.04	0.04-0.03	0.03
	L	2.015-1.1954	2.462-2.399	2.616	2.817	3.069-2.989	3.299
	ARL_{min}	17.73	24.9	28.7	34.3	40.1	47.7
1.0	λ	0.19-0.16	0.15-0.14	0.15-0.12	0.13-0.10	0.12-0.10	0.09
	L	2.346-2.298	2.723-2.707	2.907-2.858	3.113-3.059	3.317-3.283	3.538
	ARL_{min}	6.97	9.14	10.2	11.7	13.2	15.2
2.0	λ	0.52-0.47	0.42-0.38	0.37-0.36	0.35-0.31	0.32-0.28	0.29-0.26
	L	2.538-2.526	2.538-2.526	2.895-2.885	3.047-3.044	3.253-3.241	3.445-3.433
	ARL_{min}	2.62	3.23	3.51	3.90	4.29	4.81
3.0	λ	0.81-0.77	0.74-0.71	0.70-0.66	0.66-0.59	0.61-0.53	0.53-0.47
	L	2.572-2.569	2.931-2.930	3.086-3.084	3.286-3.283	3.477-3.473	3.714-3.711
	ARL_{min}	1.45	1.72	1.86	2.06	2.26	2.51
4.0	λ	1.00-0.85	0.97-0.84	0.95-0.62	0.91-0.80	0.91-0.75	0.84-0.72
	L	2.576-2.573	2.935-2.934	3.090-3.089	3.290-3.289	3.480-3.480	3.719-3.718
	ARL_{min}	1.08	1.16	1.21	1.29	1.39	1.53

Source: [adapted from Lucas and Saccussi (1990)]

Table 3.2 shows the average run length performance for some EWMA control schemes. The optimal design procedure would consist of specifying the desired in-control and out-of-control ARLs and the magnitude of the anticipated process shift, and the combination of λ and L that provides the desired ARL performance are selected. Montgomery (2001) found that generally the values of λ that lies in the interval between 0.05 and 0.25 work well in practice, with λ =0.05, λ = 0.10 and λ = 0.20 being popular choices. His emphasis is that one should use smaller values of λ to detect smaller shifts and use value of *L* between 2.6 and 2.8 so as to reduce the width of the limits.

Weringa (1999) stated that the ARL of the EWMA control chart depends on the smoothing parameter and the width of the control limits, which is determined by L in equation (3.5) and care must be taken when choosing these parameters. He used the LCL

and UCL, which appears in equation (3.5) to determine the value of the two parameters and L by fixing two points on the ARL curve. He firstly fixed the desired in-control ARL point this is related to the *producer's risk*. If the process is process is stopped accidentally due to a false out-of-control signal the producer will suffer a certain loss. This demand on the ARL curve is usually formulated in terms of a high minimal in-control ARL. The second point fixed on the ARL curve relates to the *consumer's risk*. He further argued that it is desirable for the customer to have a low ARL if the products are of unacceptable quality. The second point that is fixed is usually formulated in terms of a maximal ARL when there is a certain large shift in the mean. Fixing this second point, the chart is designed for detecting a shift in the mean of a certain size as fast as possible. He concludes that the EWMA is more sensitive to small shifts in the mean of a series of observations.

3.4 EWMA control charts with time-varying control limits

Taking EWMA test statistic Z_t as a function of time from equation (3.1) and the EWMA control chart is defined as

$$Z_t = \lambda \bar{x}_t + (1-\lambda) Z_{t-1}, \qquad o < \lambda \leq 1$$

in time *t*, where the starting value Z_0 is set equal to and estimate of the process mean $\left(\bar{\bar{x}}\right)$ calculated from previous data. $\bar{x}_t$ is the sample mean from time period *t*. Assuming the $\bar{x}_i s$ are independent random variables with mean μ_x and variance σ_x^2/n, where n is the sample size used at each time interval, $\bar{x}_i$ is calculated to get the mean value of Z_t as $\mu_{z_t} = \mu_x$ and the variance of Z_t as

$$\sigma_{z_t}^2 = \frac{\sigma_x^2}{n}\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2t}\right]$$

The control limits for the EWMA control charts are also derived based on $\pm L$ sigma limits, where L is equal to three sigma (3σ) as in the design of Shewhart control chart limits. The time-varying upper EWMA control limit ($UCL_{(t)}$) and the time-varying lower EWMA control limit ($LCL_{(t)}$) are given by

$$UCL_{(t)}, LCL_{(t)} = \mu_x \pm L\sigma_x \sqrt{\frac{\lambda\left[1-(1-\lambda)^{2t}\right]}{(2-\lambda)n}} \qquad (3.9)$$

where the sample mean μ_x and sample standard deviation σ_x as estimated from the preliminary data. As t gets bigger the control limits $UCL_{(t)}$ and $LCL_{(t)}$ converge to the asymptotic control limits, denoted as UCL and LCL, given by

$$UCL, LCL = \mu_x \pm L\sigma_x \sqrt{\frac{\lambda}{(2-\lambda)n}} \qquad (3.10)$$

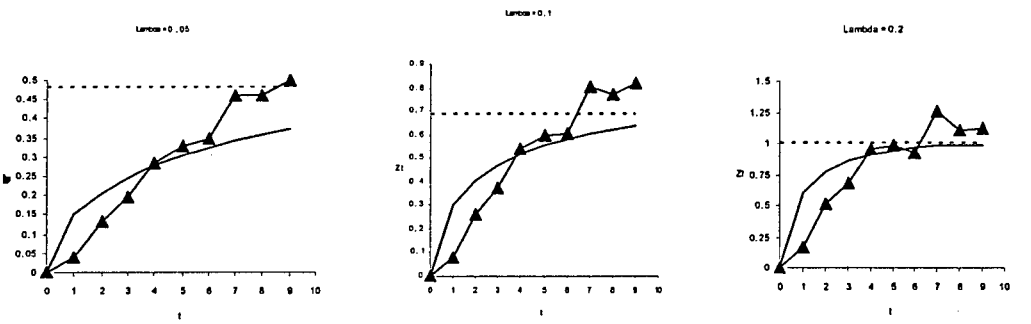
The convergence rate of this asymptotic values depends critically on λ with the convergence being much lower for small λ.

To illustrate the effect of time-varying limits web would use a set of simulated observations taken from Lucas and Crosier (1990). We would assume $\mu_x = 0, \quad \sigma_x = 1 \quad and \quad L = 3$. The raw data is given by $\bar{x}_t$ in Table 3.2 and represents an initial out-of-control situation. The table also gives the EWMA values, Z_t, derived from equation (3.8) and time-varying control limits derived from equation (3.10) with λ = 0.5, λ = 0.25, λ = 0.2, λ = 0.1, and λ = 0.05.

Table 3.3 EWMA Example using data from a process initially out-of-control [adapted from Lucas and Crosier (1990)]

Sample number	$\bar{x}_t$	λ = 0.5 CL = 1.7302 Z_t	$UCL_{(t)}$	λ = 0.25 CL = 1.1334 Z_t	$UCL_{(t)}$	λ = 0.2 CL = 1 Z_t	$UCL_{(t)}$	λ = 0.1 CL = 0.6882 Z_t	$UCL_{(t)}$	λ = 0.05 CL = 0.4804 Z_t	$UCL_{(t)}$
0	-	0	0	0	0	0	0	0	0	0	0
1	0.8	0.4	1.5	0.2	0.75	0.16	0.6	0.08	0.3	0.04	0.15
2	1.9	1.15	1.677	0.625	0.938	0.508	0.768	0.262	0.404	0.133	0.207
3	1.4	1.275	1.718	0.819	1.028	0.686	0.859	0.376	0.471	0.196	0.247
4	2.0	1.638	1.729	1.114	1.076	0.949	0.912	0.538	0.519	0.287	0.279
5	1.1	1.369	1.731	1.111	1.102	0.979	0.945	0.594	0.555	0.327	0.304
6	0.7	1.034	1.732	1.008	1.116	0.923	0.965	0.605	0.583	0.346	0.326
7	2.6	1.817	1.732	1.406	1.124	1.259	0.978	0.804	0.604	0.459	0.344
8	0.5	1.159	1.732	1.179	1.128	1.107	0.986	0.774	0.621	0.461	0.359
9	1.2	1.179	1.732	1.185	1.131	1.126	0.991	0.817	0.634	0.498	0.373

In quality control it is often useful to graphically display control schemes. Figure 3.1 shows the resulting EWMA charts for different values of λ. The solid line shows time-varying upper control limit $UCL_{(t)}$, the dashed line shows the asymptotic control limit UCL and the solid line with triangles represents EWMA values.



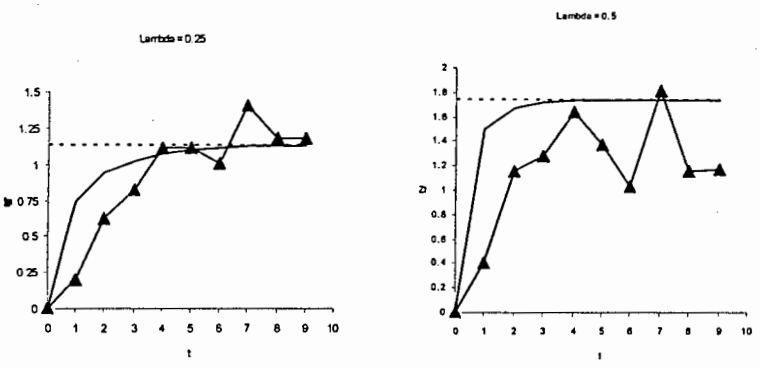


Figure 3.2: Plot of EWMA Control Charts with Time-varying control limits

Figure 3.2 shows the number of observations needed to generate an out-of-control signal depends on both the λ and whether the time-varying limits are used. When λ value is 0.05, 0.1, 0.2 or 0.25 EWMA chart with time-varying control limits gives the signals after four observations whereas using the asymptotic limits (UCL) a signal will not be generated until seven observations for λ equals 0.1, 0.2 and 0.25 or observation nine for λ = 0.05. The time-varying control limit converges to the asymptotic value quickly after λ = 0.5 so it has little effect. When $\lambda = 0.5$ a signal occurs after seven observations using either $UCL_{(t)}$ or UCL as control limit. Figure 3.1 also shows that using asymptotic control limits rather than the time-varying limits makes the EWMA chart less sensitive to process shifts in the first few observations.

3.5 Fast initial response features for EWMA control chart

Lucas and Crosier (1982), [?] describe the properties and enhancements of the EWMA control schemes, the enhancements include the so-called FIR feature, that they had also showed that the feature is useful for the CUSUM control schemes. The FIR feature is useful for EWMA control schemes designed with small values of λ. In the situation when λ is small the variance of the control statistic converges slowly to its asymptotic value so that control schemes based on the asymptotic standard deviation tend to be insensitive at the start-up. The FIR feature requires the simultaneous implementation of two one-sided EWMA control schemes with different starting values than zero, head start (HS). One

EWMA is started above the target, and the other is started below the target value. If the EWMA started on the high side falls outside the upper control limit or if the EWMA started on the low side falls outside the lower control limit an out-of-control signal will be given.

The method used by Lucas and Crosier (1982) to evaluate the properties of the FIR feature for the EWMA is the transition probability matrix. The transition probability matrix requires t^2 in-control states, with $p_{jkj'k'}$ representing the probability that the EWMA with a head start on the high side moves from j to state k and the EWMA with a head start on the low side from state j' to state k'. An upper bound for the FIR ARL's can be obtained using t in-control state only. There are two vectors required for the initial probability, these are $p_L^T = (0,\ldots,1,\ldots 0,0,0,\ldots,0,\ldots 0)$ and $p_U^T = (0,\ldots,0,\ldots,0,0,0,\ldots,1,\ldots,0)$. p_L^T and p_U^T represent the initial probability vectors for the EWMA with a head start on the low side and initial probability vectors for the EWMA with a head start on the high side respectively. When the desired head start does not correspond to the midpoint of any discrete state, the ARL can be approximated using quadratic interpolation of the three closest states.

The transition probability matrix requires two out-of- control states

$$P = \begin{pmatrix} 1 & 0^T & 0 \\ 1_1 & R & h_1 \\ 0 & 0_T & 1 \end{pmatrix}$$

Where the vector 1_1 contains the probabilities that the EWMA of the low-sided scheme goes out-of-control on the low side and h_1 contains the probabilities that the EWMA of the high-sided scheme goes out-of-control on the high side. The upper bound for the FIR ARL's is computed as follows:

$$ARL_{FIR}(t) = \sum_{i=1}^{\infty} i \Pr(RL = i)$$

$$= \sum_{i=1}^{\infty} i \Pr(RL_L = i \text{ or } RL_H = i)$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{\infty} i\{\Pr(RL_L = i) + \Pr(RL_H = i)\} \\
&= \sum_{i=1}^{\infty} i\left\{p_H^T(h_i - h_{i-1}) + p_L^T(1_i - 1_{i-1})\right\} \qquad (3.11) \\
&= \sum_{i=1}^{\infty} i\left(p_H^T R^{i-1} h_1 + p_L^T R^{i-1}\right) \\
&= p_H^T\left(\sum_{i=1}^{\infty} iR^{i-1}\right)h + p_L^T\left(\sum_{i=1}^{\infty} iR^{i-1}\right)1_1 \\
&= p_H^T(I-R)^{-2}h_1 + p_L^T(I-R)^{-2}1_1
\end{aligned}
$$

Where $1_i = 1_{i-1} + R^{i-1}1_1$ and $h_i = h_{i-1} + R^{i-1}h_1$

Steiner (1994) suggested a different from Lucas and Crosier's approach in improvement in the FIR performance. He suggested that approach retains the simplicity of a single control chart. He improved the FIR by further narrowing the control limits for the first few sample points to give EWMA charts with time-varying control limits a FIR feature. Steiner takes the time-varying control limits to be exponentially approaching the asymptotic limits and uses an exponentially decreasing adjustment to further narrow the limits. He designed Steiner's method as follows:

$$FIR_{adj} = 1 - (1-f)^{1+a(t-1)} \qquad (3.12)$$

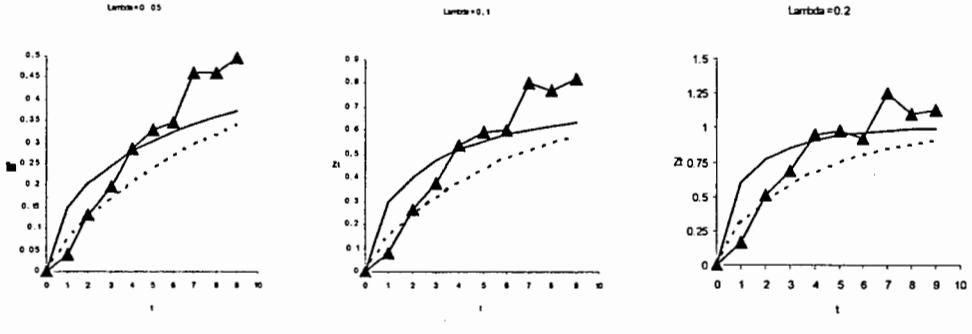
Where FIR_{adj} is FIR adjustment and the set up of the FIR_{adj} makes the control limits for the first sample point, $t = 1$, a proportion f of the original distance from the starting value. The effect of the FIR adjustment decreases with time to ensure that the long term run length properties of the EWMA will be virtually unchanged. The adjustment parameter a is set so that the FIR adjustment has very little effect after observation n_j *and* $n_j < n$, say that the adjustment FIR_{adj} at observation n_j is 0.99. This should be sufficient to allow the detection of quality problems in the startup. This idea implies that we should set $a = \dfrac{\left(-2/\log(f) - 1\right)}{n_j - 1}$

This gives the FIR-EWMA control limits:

$$\mu_x = \pm L\left(1-(1-f)^{1+a(t-1)}\right)\sqrt{\frac{\lambda\left[1-(1-\lambda)^{2t}\right]}{(2-\lambda)n}} \tag{3.13}$$

It can be proved that for a reasonable setup n_j should be equal to 20 ($n_j = 20$) since FIR adjustment has very little effect after observation 20. The control limits given by Equation (3.13) are time-varying; this implies that the run lengths properties of the FIR-EWMA can be also determined using the non-homogeneous Markov chain methodology.

Example for FIR-EWMA control limits is shown in Figure 3.2. Using $f = 0.5$ and $n_j = 20$ it yields $a = \frac{\left(-2/\log(f)-1\right)}{n_j - 1} = \frac{\left(-2/\log(0.5)-1\right)}{20-1} = 0.3$ using equation (3.13) on the previously discussed values Section 3.5 and previously illustrated in Figure 3.1. Dashed lines show FIR time varying control limits computed using equation (3.13), solid lines show the time-varying control limits computed using $UCL_{(t)}$, and the solid lines with triangles shows EWMA values. The values of EWMA control limits used to plot graphs in Figure 3.3 are at Appendix C, Table C.3. $f = 0.5$ is selected so as to use 50% head start. In Figure 3.3 for all the different values of λ the FIR-EWMA signals is substantial over run lengths obtained with only the time-varying control limits. This is evident when they of λ is large.



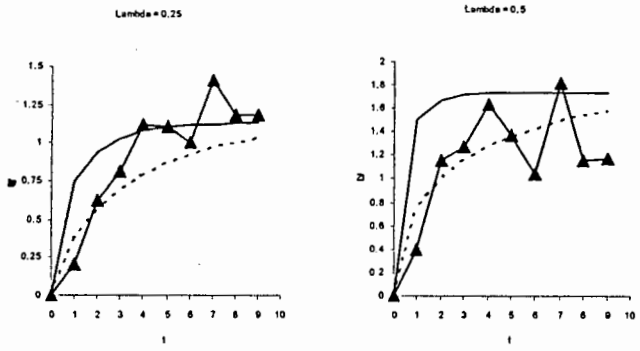


Figure 3.3: Plot of EWMA Control Charts with Time-varying control limits and Fast Initial Response.

3.6 Average run lengths and comparison of charts for control of the process mean level

For comparison of Shewhart, CUSUM and EWMA charting techniques we would use the average run lengths (ARL). The ARL gives how many observations or number of runs, on average, are required to detect the various process mean shifts. As explained in previous chapters control charts are used to detect and monitor changes in the process, EWMA control charts are generally used for detecting small shifts in the process mean. We will use the values in Table 3.3 which shows, ARL values for each column, related to the Shewhart control chart with control limits $L = 3\sigma$, the EWMA control chart when $\lambda = 0.2$ and $L = 2.962\sigma$, and the CUSUM control chart with $k = 0.5$ and $h = 5$, taking into consideration the shift in the mean. This table is adapted from Montgomery (2001) and Lucas and Saccussi (1990). The parameters of the EWMA's were chosen so that the initial control ARL's would match those of the CUSUM control scheme. The matching is done for the zero-state.

Table 3.4 ARL values for a Shewhart, EWMA and CUSUM chart [Adapted from Montgomery (2001) and Lucas and Saccussi(1990)]

Shift in mean (Multiple of σ)	Shewhart chart for individuals	EWMA chart for individuals	CUSUM chart for individuals
0	370	500	465
0.25	281	150	139

0.50	155	41.8	38.0
0.75	81.2	18.2	17.0
1.00	43.9	10.5	10.4
1.50	15.0	5.50	5.75
2.00	6.30	3.74	4.01
2.50	3.24	2.88	3.11
3.00	2.00	2.38	2.57
4.00	1.19	1.86	2.01

For clear comparison the values in ARL curve of the Shewhart, CUSUM and EWMA scheme are given in Figure 3.2

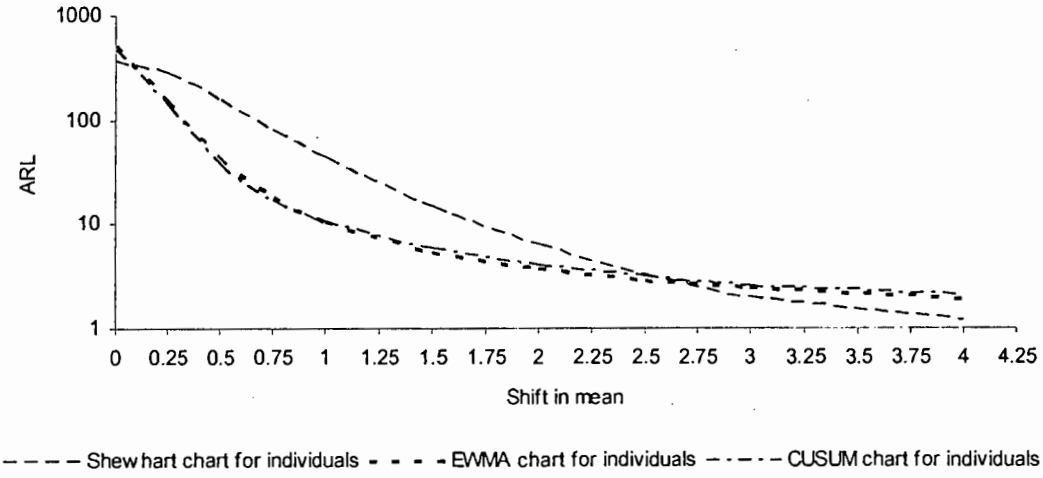


Figure 3.4 ARL curve of Shewhart, CUSUM and EWMA chart

From figure 3.4 on target the Shewhart chart is more likely to signal a warning than the CUSUM and EWMA chart. Between about the mean shift of 0.25σ and 2.5σ, from the target, the CUSUM and the EWMA charts are much more efficient in detecting the shift much more quickly than the Shewhart chart. More than 3σ from the target, the Shewhart chart detects the shift more quickly than the CUSUM and EWMA charts. Further more Figure 3.2 show that the properties of the CUSUM's are very close to those of the EWMA schemes and they both outperforms the Shewhart technique. This implies that the benefits of the CUSUM and EWMA charts are quite attractive for improving process monitoring and detecting small persistent process shifts. Lucas and Saccussi (1990) also emphasis that there is very little difference between CUSUM and EWMA procedures in

terms of the ARL for detecting persistent process mean shifts. The EWMA is better than the CUSUM charting technique for detecting larger mean shifts above 2 as seen in Figure 3.2. EWMA control charts may also be preferred when the subgroups are of size n = 1. In this case, an alternative chart might be the individual X- bar chart, in which case we would need to estimate the distribution of the process in order to define its expected boundaries with control limits. Like the CUSUM control charts the EWMA control charts have an advantage that each plotted point includes several observations, so the central limit theorem can be used to say that the average of the points or the moving average in this case is normally distributed and the control limits are clearly defined. The EWMA also has some attractive properties in particular:

- Unlike X-bar, R and individual charts (without the Western Electric rules which aim to increase sensitivity, see **Appendix B, Table B.3**), all of the data collected over time may be used to determine the control status of a process.
- The EWMA is often superior to the CUSUM charting technique for detecting "larger" shifts.
- EWMA schemes may be applied for monitoring standard deviations in addition to the process mean.
- There exists the ability to use EWMA schemes to forecast values of a process mean.
- The EWMA methodology is not sensitive to normality assumptions.

3.7 Combined Shewhart-EWMA

A combined Shewhart EWMA was developed to give improved properties when both large and small shifts are to be detected. This is done by adding Shewhart limits to an EWMA control scheme so that the out-of-control signal is given if the EWMA statistic is outside the control limits or if the current observation is outside the Shewhart limits.

The properties of the combined Shewhart EWMA can be obtained by modifying the transition probability matrix for an EWMA control scheme discussed in Section 3.4. the modified one-step transaction probabilities are given by:

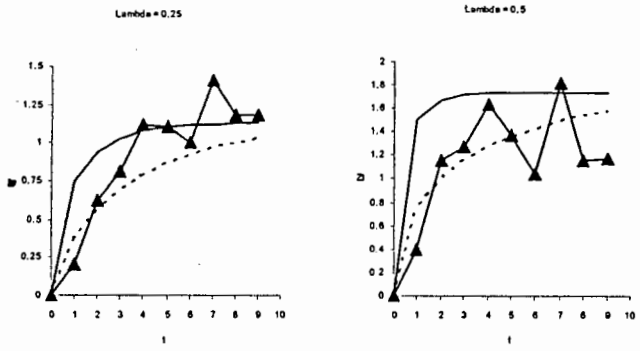


Figure 3.3: Plot of EWMA Control Charts with Time-varying control limits and Fast Initial Response.

3.6 Average run lengths and comparison of charts for control of the process mean level

For comparison of Shewhart, CUSUM and EWMA charting techniques we would use the average run lengths (ARL). The ARL gives how many observations or number of runs, on average, are required to detect the various process mean shifts. As explained in previous chapters control charts are used to detect and monitor changes in the process, EWMA control charts are generally used for detecting small shifts in the process mean. We will use the values in Table 3.3 which shows, ARL values for each column, related to the Shewhart control chart with control limits $L = 3\sigma$, the EWMA control chart when $\lambda = 0.2$ and $L = 2.962\sigma$, and the CUSUM control chart with $k = 0.5$ and $h = 5$, taking into consideration the shift in the mean. This table is adapted from Montgomery (2001) and Lucas and Saccussi (1990). The parameters of the EWMA's were chosen so that the initial control ARL's would match those of the CUSUM control scheme. The matching is done for the zero-state.

Table 3.4 ARL values for a Shewhart, EWMA and CUSUM chart [Adapted from Montgomery (2001) and Lucas and Saccussi(1990)]

Shift in mean (Multiple of σ)	Shewhart chart for individuals	EWMA chart for individuals	CUSUM chart for individuals
0	370	500	465
0.25	281	150	139

0.50	155	41.8	38.0
0.75	81.2	18.2	17.0
1.00	43.9	10.5	10.4
1.50	15.0	5.50	5.75
2.00	6.30	3.74	4.01
2.50	3.24	2.88	3.11
3.00	2.00	2.38	2.57
4.00	1.19	1.86	2.01

For clear comparison the values in ARL curve of the Shewhart, CUSUM and EWMA scheme are given in Figure 3.2

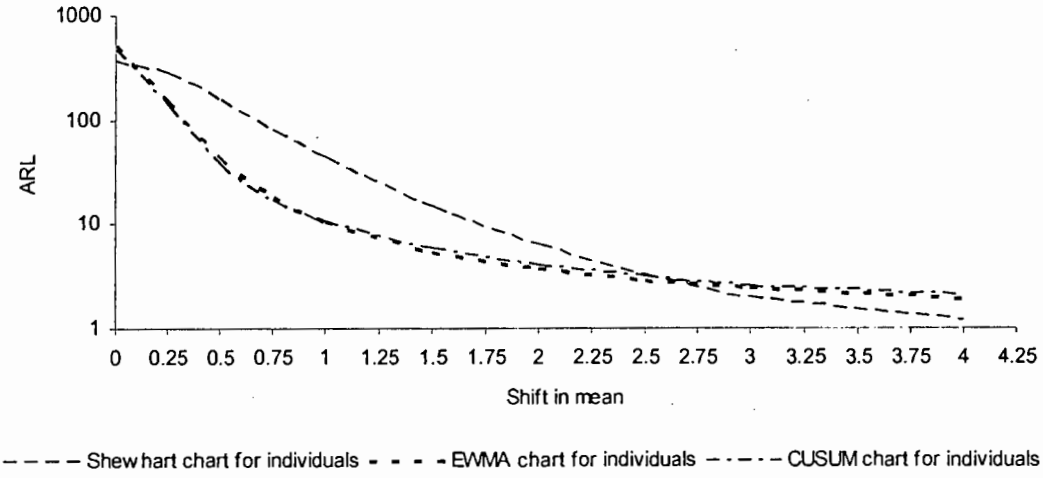


Figure 3.4 ARL curve of Shewhart, CUSUM and EWMA chart

From figure 3.4 on target the Shewhart chart is more likely to signal a warning than the CUSUM and EWMA chart. Between about the mean shift of 0.25σ and 2.5σ, from the target, the CUSUM and the EWMA charts are much more efficient in detecting the shift much more quickly than the Shewhart chart. More than 3σ from the target, the Shewhart chart detects the shift more quickly than the CUSUM and EWMA charts. Further more Figure 3.2 show that the properties of the CUSUM's are very close to those of the EWMA schemes and they both outperforms the Shewhart technique. This implies that the benefits of the CUSUM and EWMA charts are quite attractive for improving process monitoring and detecting small persistent process shifts. Lucas and Saccussi (1990) also emphasis that there is very little difference between CUSUM and EWMA procedures in

terms of the ARL for detecting persistent process mean shifts. The EWMA is better than the CUSUM charting technique for detecting larger mean shifts above 2 as seen in Figure 3.2. EWMA control charts may also be preferred when the subgroups are of size n = 1. In this case, an alternative chart might be the individual X- bar chart, in which case we would need to estimate the distribution of the process in order to define its expected boundaries with control limits. Like the CUSUM control charts the EWMA control charts have an advantage that each plotted point includes several observations, so the central limit theorem can be used to say that the average of the points or the moving average in this case is normally distributed and the control limits are clearly defined. The EWMA also has some attractive properties in particular:

- Unlike X-bar, R and individual charts (without the Western Electric rules which aim to increase sensitivity, see **Appendix B, Table B.3**), all of the data collected over time may be used to determine the control status of a process.
- The EWMA is often superior to the CUSUM charting technique for detecting "larger" shifts.
- EWMA schemes may be applied for monitoring standard deviations in addition to the process mean.
- There exists the ability to use EWMA schemes to forecast values of a process mean.
- The EWMA methodology is not sensitive to normality assumptions.

3.7 Combined Shewhart-EWMA

A combined Shewhart EWMA was developed to give improved properties when both large and small shifts are to be detected. This is done by adding Shewhart limits to an EWMA control scheme so that the out-of-control signal is given if the EWMA statistic is outside the control limits or if the current observation is outside the Shewhart limits.

The properties of the combined Shewhart EWMA can be obtained by modifying the transition probability matrix for an EWMA control scheme discussed in Section 3.4. the modified one-step transaction probabilities are given by:

$$p_{jk} \approx \Pr[\min\{SCL_U, \max(SCL_L, Y_L)\} < Y_i \le \max\{SCL_L, \min(SCL_U, Y_U)\}]$$

$$Y_L = \lambda^{-1}\{(S_k - \delta) - (1-\lambda)S_j\}$$

and $Y_U = \lambda^{-1}\{(S_k + \delta) - (1-\lambda)S_j\} \quad j = -m, -m+1, \cdots, m,$

where SCL_U is the upper Shewhart control limit and SCL_L is the lower Shewhart control limit.

3.8 Summary

The EWMA is almost as good as the CUSUM control charts. They are both have a greater ability to detect small shifts in the process mean between 0.5σ to 2σ faster than Shewhart charts. The control limits of EWMA approaches stable-state values as the process has been running for a some time. It is recommended to use a small value of the weighting factor λ so that the shift that is to be detected gives a the ARL as short as possible, usually $\lambda = 0.2$, to detect small shifts, and larger values, and $0.2 < \lambda \le 0.4$, for larger shifts. We have derived the run length properties for EWMA control charts with time varying control limits. This has shown that the time varying control limits have improved process shift detection capabilities. Adjusting the EWMA control limits to FIR and then exponentially approach to create FIR-EWMA is shown to perform better than EWMA.

Chapter 4 Review on Grey System Theory, Grey Differential Equation and Grey Models

The statistical process control has been applied successfully in industries in relation to massive quality production. In massive production the process is said to be stable or unstable using control charts based on large-samples or more numbers of individuals for 'control chart for individuals'. These statistical process control methodologies applicable to massive production may not be useful today because of rapid changes in business environments. Pan (2002) timely pointed out that "the life cycle of products has decreased rapidly and customized short-run manufacturing process become quite common for achieving custom satisfactions".
Due to this high competitions business environments have become "unstable". This has lead to sparse information availability from "unstable" short-run businesses. Sparse information makes it hard to use large-sample based techniques in process control. Therefore the development of a small-sample based process control theory and methodology is an urgent task in quality control and management communities.

In probabilistic developments, small-sample asymptotic theory has its own developments since 1950's and it is still active research area today. However, we notice the difficulty and complexity in small sample asymptotic theory. The demands of business globalization environments force us to consider alternative theory and methodologies for small-batch based quality control. The grey theory initiated by Deng (1982) offered us part of the solution.

In this chapter we will review the grey theory and grey differential equation models. Particularly, the most predictive model, the first order one variable grey differential equation model, abbreviated as GM(1,1) model, will be reviewed and examined. Except Deng's creative works (1985, 1989, 1993, 2002), there are many researchers contributing to grey theory developments with enthusiasm, for example, Dang et al (2004), Lin et al (2004), Liu and Liu (2006), Ng et al (1995), Tien (2005), Tseng et al (2001), Wang et al (2004), Wen (2004) Wen et al (2005), Xia (1997) and others. We have to mention that most journal articles in grey theory were published in Chinese. Even these listed works in English were not widely accepted partially due to the poor

English readability. In terms of the review of grey differential equation models, it is expected to understand why a grey differential equation model can be used to monitor whether a product conforms to the required specifications. The reviewed model will be carried forward to Chapter 5 for the purpose of using the model in quality control. In subsequent sections in this chapter experimental data are used as examples in some of the formulae of formation of a grey system model.

4.1 Preview of grey system theory

The grey system theory was initiated by Professor Julong Deng, in 1982. In Deng's original opinion (Deng 1985), a grey system is the system with incomplete information. It seems that using "incomplete information" for defining grey system is too vague. Strictly speaking, a system with "incomplete" information can not be a scientific definition of grey system because there are many other systems being also having incomplete information, for example, Hidden Markov Processes in stochastic process theory. The key focus for identifying a grey system is motivated by the fact that this system only provides very limited data available for scientific study, see in Figure 4.1. Therefore, we give a working definition for grey system. A grey system is the one with sparse data information extracted from it for investigation purpose.

The goal of the grey system is to bridge the gap between the social science and natural science initially, that is, the grey system is interdisciplinary, cutting across a variety of specialized fields. Intuitively in modern control theory, black system means we have no information on it, the white system means we have full information on it, i.e., the information in a white system is clear and "completely" known. Systems lie between black systems and white systems are grey. However, the above descriptions of three systems are merely providing some background picture from the stand of information amount. We emphasize again: in this thesis, a grey system is the one with *small sample* information extracted from the system for investigation purpose. The grey number is used to describe the system state in grey system research.

Input (Grey system) Output

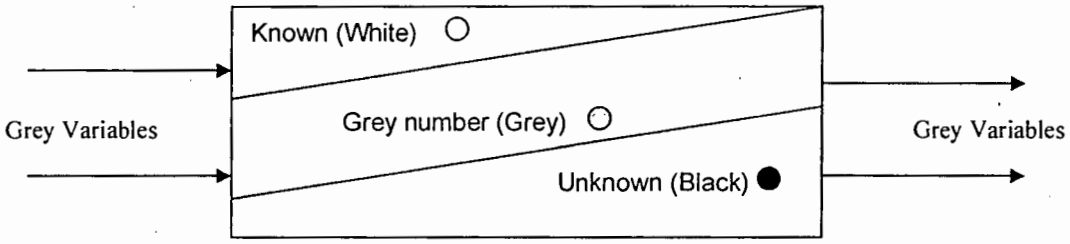


Figure 4.1 The Grey system theory concept

The concepts and models of grey theory has been tested and applied over the past two decades in China, Deng (1989) to a variety of projects of various fields and it has proven success. Some of these projects are: Economy: regional economic planning for several provinces, an analysis of agricultural economy, estimate the economic effect; Predictions of Agriculture: to forecast yield for some provinces; Meteorology: to forecast weather mould rains season; Medical Science Research; to build models available for biological protection, to build a diagnosis model available for medicine; and Others: to make satisfactory planning of irrigation, to control water level for boilers by grey prediction control, researches in industry.

4.1.1 The contents of grey theory

For over 20 years in its development of the grey system has been summarized in six parts namely; grey generating, grey rational analysis, grey model, grey prediction model, grey decision making and grey control.

4.1.2 Grey generating

The idea of grey generating is that a raw data may not reveal any pattern at the first glance, but after grey generating (operations) it may reveal a strong trend with minor fluctuations, which paves the way for further modeling efforts. It is aimed to process the raw and fluctuating data in

order to gain a clear rule, which is whitening of a sequence of numbers, that is, this is data processing to supplement information. The grey generating including four parts:

1. Grey relational generating operation (GRGO): The intention of the GRGO is to add new information for the system needs, that is, it is based on the processed data to find the rule of data.
2. Accumulated generating operation (AGO): The intentions of the AGO is using the data accumulated operation to reduce the randomness of original data.
3. Inverse accumulative generating operation (IAGO): Inverse accumulated generating operation is the anti operation of accumulated generating operation.
4. Localization generating: The main purpose of localized generating is the sequence on non-equigap, or missing data, and then we use this method to reconstruct the lost data and make the sequence become equal-gap.

4.1.3 Grey relational analysis (GRA)

This model is an impact measurement model, which takes the measurements of relations that change in two systems or between two elements into the system in time. In other words GRA quantifies all influences of various factors and their relation, which is called the whitening of factor relation.

4.1.4 Grey model (GM)

This method translates difference equations using dummy concepts. The model can be constructed in different types, some of which are:

- GM(1,1) model: This means first differential and one variable, the function of this GM(1,1) model is being used in the prediction field and useful for sequences that do almost satisfy the exponential law.
- GM(2,1) model: This means second differential and one variable, the function of this GM(2,1) can be applied to describe non-monotonic process of change.

- GM(1,N) model: In this model we first find the differentials of N variables, the function of the GM(1,N) model is being used in analysis where there are more than one variables in the function i.e. in multi variable analysis field.
- GM(0,N) model: The GM(0,N) model means no differential and N variables, the GM(0,N) model is analyzed in the same step as GM(1,N) model. GM(0,N) model is also the special type of GM(1,N) model.
- GM(h,N) model: This means h differentials and N variable. "In the GM(h,N) model, $h \geq 2$ does not exist", Kun-Li Wen (2004).

4.1.5 Grey prediction model

Deng Julong (1989) proposed the Grey system theory to construct a Grey model for forecasting. The subjects of grey forecasting include: series forecasting; calamities forecasting; season calamities forecasting; topological forecasting; and systematic forecasting. All these predictions are made based on the GM(1,1) model. The Grey prediction model (GM) is the core of Grey system theory. The intension in using these grey forecasting is to make forecasting useful for decision and policy makers who need future predictions.

Wang Tien-Chin et al (2004) applied grey prediction to construct a forecasting model on forecasting the output values of exchange rate between the Taiwan Dollar (TWD) and the U.S. Dollar (USD). Their results show that the average residual error of the grey prediction model is lower than 0.09696%. They further show that the Grey prediction model exhibits highest prediction of accuracy. This proves the ability of Grey system theory to effectively deal with incomplete and uncertain information.

The Grey system theory treats all variables as a Grey quantity within a certain range. Grey prediction model then collects available data to obtain the internal regularity. The model examines the nature of internal regularity in managing the disorganized primitive data. Some prediction methods are listed in Table 4.1 below, these includes the traditional and grey method.

Table 4.1 The Comparison of Traditional and the Grey Prediction Methods

Mathematics Model	Minimum data for prediction	Data type	Interval
Simple Exponential	5 to 10	Equal gap	Short
Holt's	10 to 15	Same trend and regular	Short and middle
Winter's	At least 5	Same trend and regular	Short and middle
Regression	At least 10 or 20	Same trend and regular	Short and middle
Time series method	At least 2 pick value	Mixture type and regular	Short and middle
Grey method	Only 4	Equal gap and not equal gap	Short, middle and long

4.1.6 Grey decision making

The other methods are combined with the GM(1,1) model so as to solve decision making problems. A decision is made under imperfect countermeasure and unclear situation, which is called the whitening of status. This is done in three types of model construction, these are:

1. Grey rules of the situation: This deals with strategy making based on multi-objects that are contradictory in the ordinary way.
2 Grey decision-making group: This field includes grey relational space, grey statistics, grey clustering and grey prediction.
3 Grey programming: This is based on the traditional programming, to incorporate the grey prediction model in order to make dynamic programming and to regard the coefficients as grey number with provision for adapting the environment.

4.1.7 Grey control

The grey control uses the data, including general control systems with grey parameters and the controls built on the analysis, modeling, prediction and decision making. This is based on the fact that, there exist errors in recognition and measuring, the structure and related parameters of the system containing a degree of greyness. Yin Lin, et al (2004) stressed that the selection of input

and output data values in different time intervals might have an impact on the determination of the parameters.

4.2 Generating of the GM model

We would start by briefly reviewing the two basic operations of grey generating mathematical models used to build the GM model. These basic operations are: accumulated generation and inverse accumulated generation.

4.2.1 Accumulated generating operation (AGO)

Mathematically speaking, Accumulated Generating Operation (abbreviated as AGO or from now on), is merely a partial sum of data sequence given. However, in grey theory, AGO has been regarded as the most creative approach for the revelation of the pattern underneath of the raw data. We should emphasize here, the applicability of AGO is the discrete data sequence must be *sequentially ordered and strictly positive.* Without such consciousness of the restrictions of AGO, it will be inevitably misleading. Using the AGO technique efficiently reduces noise by converting ambiguous strictly positive time-series data to a monotonically increased series. The AGO technique is capable of reducing the fluctuations in the series and increasing the smoothness of the and help us to identify the systematic regularity quickly and easily. The Grey system must apply the order of accumulated generating operation (AGO) is the times to perform AGO. One time AGO is called as 1-AGO, another AGO on the AGOed sequence is the 2-AGO to the raw data sequence, and so on.

Assume $x^{(0)}$ is the original discrete data sequence, denoted by $x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n)\right) = \left(x^{(0)}(k);\ \ k = 1,2,3,\cdots,n\right)$. n must always be equal to or larger than 4. Computing the AGO we let the first element of the original series be the first element of the new series; let the sum of the first and second elements of the original series be the second element of the new series; let the sum of the first, second and third elements of the original series be the third element of the new series, and so on. The derived new non-negative

series is called the one time accumulated generating series of the original series and is denoted by $x^{(1)}$. The accumulated generating operation (AGO) series is mathematically written as:

$$\begin{aligned}
&1\text{-AGO}\left\{x^{(0)}(k),\ k=1,2,\cdots,n\right\}\\
&=\left\{x^{(1)}(k),\ k=1,2,\cdots,n\right\}\\
&=\left\{\sum_{m=1}^{k}x^{(0)}(m),,\ k=1,2,\cdots,n\right\}\\
&=\left\{x^{(0)}(1),\sum_{k=1}^{2}x^{(0)}(k),\cdots,\cdots,\cdots,\sum_{k=1}^{n}x^{(0)}(k)\right\}
\end{aligned}\qquad (4.1)$$

If there is not yet evident regularity in the series of the first AGO, the AGO can be continued for the second-time, third-time, ..., and n^{th}-time. The superscript (1) in $x^{(1)}(k)$ represents the data is at the first AGO level, which is usually denoted as 1-AGO. If the superscript is (r), it represents the r times AGO and is often denoted as r-AGO. The elements of the r-AGO series are got by rearranging equation (4.1) in standard form then we get

$$\left(x^{(r)}(k);r=1,2,3,\cdots,n\right)=\sum_{m=1}^{k}x^{(r-1)}(m)\qquad (4.2)$$

For an intuitive illustration AGO, we generate a series of random numbers between 0 and 1 which we treat as the original data (with sample size 4).

$$x^{(0)}=\{0.964875,0.245855,0.756121,0.335144\},$$

by using the sequence of first order accumulating generation, equation (4.1), we get

$$1\text{-AGO}\left\{x^{(0)}\right\}=x^{(1)}=\{0.964875,1.21073,1.966851,2.301995\},$$

this result is graphically presented in Figure 4.2 it can be seen that the original series after accumulated generating operation has changed into monotonic increasing type, it is quite like a linear function so it can be simulated with linear regression function. We check by calculating values for the linearity of $x(k)$ with relation to k using R-squared. $R^2\left(x^{(0)}\right)=0.0346$ and $R^2\left(x^{(1)}\right)=0.964$ so there is linearity between $x(k)$ and k of AGO series. Sefin Liu (1998) states

that characteristic function of relationship $x(k)$ between and k should be checked at first order AGO so that the sequence would be simulated using a correct function.

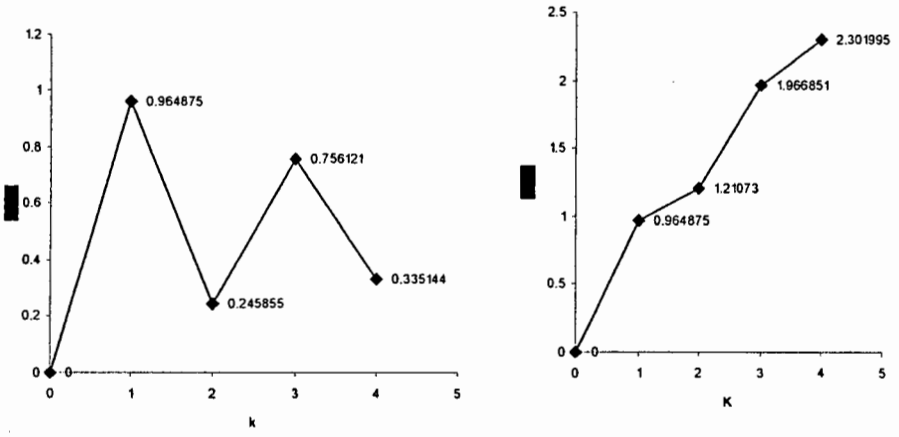


Figure 4.2 The AGO for the jumping sequence

4.2.2 Inverse-accumulated generating operation (IAGO)

The inverse accumulated generating operation (abbreviated as IAGO) is the inverse operation of the accumulated generating operation, which is nothing but the difference operation in mathematical literature. Let the first element of the series to be operated on be the first element of the new series; let the difference between the third element and the second element of the series to be operated on be the third element of the new series; and so on. The derived new series is called the one-time IAGO of the series to be operated on. The basic relationship of inverse accumulated generating operation is defined as:

$$1\text{-IAGO}\left\{x^{(0)}\right\}=\left\{x^{(-1)}(k)=x^{(0)}(k)-x^{(0)}(k-1)\right\} \tag{4.3}$$

When $r=1$, equation (4.3) can be reduced into a simple form, to give the mathematics model of 1-IAGO as:

$$x^{(1)}(k)=x^{(1)}(k-1)+x^{(0)}(k) \tag{4.4}$$

Then the acquired sequence one-order inverse-accumulated generating operation (IAGO) is acquired and the sequence must be reduced as in equation (4.5) in order to obtain statistics confirm the efficiency of the proposed prediction model.

$$\begin{aligned} & x^{(1)}(k)=x^{(1)}(k-1)+x^{(0)}(k) \\ \Rightarrow\ & \hat{x}^{(0)}(k)=x^{(1)}(k)-x^{(1)}(k-1) \end{aligned} \tag{4.5}$$

A final remark on the AGO and IAGO is that they are not the inventions of Deng (1985) but the roles were critically idendified by Deng (1985). IAGO and AGO correspond to the derivative and integral operators for smooth functions in calculus although they are merely approximations to the derivative and integral of a smooth function respectively. We should be fully aware that grey modeling is in nature a special class of *approximation theory* to help the building the approximate dynamics of the system.

4.3 Grey differential equation and grey models

Due to short term production these days there is always a small sample data series available. Small sample sizes do not clearly exhibit whether the relation of product and time in the series is of linear, exponential or of any other mathematical characteristic. This can lead to incorrect monitoring or prediction when using some statistical methods as they require large sample sizes. This makes a grey system model to be a good method that can be used in short term production process, as it uses small sample data series as small as of sample size of at least four elements in a data set. A Grey Model (GM) is contracted based on the Grey Differential Equation. The model is constructed on a series by accumulated generating operation sequence (AGO). AGO can clearly show which relation characteristic is in the series, and hence leads to use of a correct method to use.

Liu and Lin (2006) stated that the qualitative analysis of a study must be done first, and the relevant quantitative analysis must be closer to the qualitative analysis, when studying an abstract system or establishing a mathematical model for the system is the qualitative studying of the system with respect to its overall functions, synergic functions, incidence relations among its factors, casual relations, and dynamical relations, etc. Therefore the establishment of a system's

model must go through the five steps of modeling; development of thoughts, factor analysis, quantification, dynamicalization, and optimization. These steps are as follows:

Step 1: Develop thoughts and form thoughts. The research directions, goals, paths, and how to implement all the details must be clarified through the qualitative analysis and research. Language modeling must be carried out by verbally and precisely describing the results.

Step 2: Examine the factors and the relationship between the factors contained in the language model to find the causes and effects affecting the development of the system under consideration. This can be done by constructing a diagram to depict the casual relations.

Step 3: Analyze quantitative the casual relation of each link and initially derive conceptual and conceptual and quantified relations at some low level. This is referred to as quantification modeling.

Step 4: Collect input and out put data values of each link to establish a dynamic GM model. This is referred to as dynamic modeling.

Step 5: The dynamic model obtained in Step 4 must be systematically studied. This is done trough adjustments on organizations, mechanisms, and parameters, reorganize the system in order to optimize allocations so that the goal of improving the system's dynamics can be reached. In this way one constructs a model referred to as the optimal modeling.

To make the entire model more mature we have to keep on constantly feedback the results obtained in later steps to earlier steps. The above five-step-modeling establishes a bridge connecting the social and natural sciences, which mathematically, computerizes the research in social sciences.

4.3.1 GM(1,1) models

Original grey model GM(1,1) is a time series forecasting model. The GM(1,1) has three basic operations: accumulated generation, inverse accumulated generation and grey modeling. A

GM(1,1) uses these operations to construct differential equations. The GM(1,1) a single variable first order grey model defined as:

$$x^{(0)}(k)+az^{(0)}(k)=b,\ k=2,3,\cdots,n \tag{4.6}$$

where

a is the development coefficient

b is grey input

$x^{(0)}$ is the original discrete data sequence

Accumulated generating operating (AGO):

$$AGO\left\{x^{(0)}(k)\right\}=x^{(1)}(k)=\left(\sum_{k=1}^{1}x^{(0)}(k),\sum_{k=1}^{2}x^{(0)}(k),\cdots,\cdots,\cdots,\sum_{k=1}^{n}x^{(0)}(k)\right)$$

Inverse accumulated generating operating (AGO) is defined as:

$$\begin{aligned}x^{(1)}(k)&=\sum_{m=1}^{k-1}x^{(0)}(m)+x^{(0)}(k)\\&=x^{(1)}(k-1)+x^{(0)}(k)\Rightarrow x^{(0)}(k)\\&=x^{(1)}(k)+x^{(1)}(k-1)\end{aligned}$$

And the mean operation for $x^{(1)}(k)$ and $x^{(1)}(k-1)$, i.e.,

$$z^{(1)}(k)=\frac{1}{2}\left(x^{(1)}(k)+x^{(1)}(k-1)\right),\ k=2,3,\cdots,n$$

The source model of GM(1,1) model in Equation (4.6) is $dx^{(1)}/dt+ax^{(1)}=b$, which is also called the whitening (or shadow) equation to GM(1,1) model. The differential equation $dx^{(1)}/dt+ax^{(1)}=b$ is actually the model restriction (imposed or identified) to the data-assimilation process. Tt uses the original sequence $x^{(0)}$ to build the approximation differential equation in the steps shown below. The reason why the GM(1,1) used the first differential equation type to approach is that from Figure 4.1, we know the main purpose of the AGO translated the jumping type sequence into monotonic increasing type sequence, and it has also been found that the solution of the first differential equation is the simplest type for this type. Then the GM(1,1) model as defined in equation (4.6) is used to fix the existed data using mathematical analysis steps shown below:

1. We perform $\dfrac{dx^{(1)}}{dt}$ into the difference of the forward and backward, means

$$\frac{dx^{(1)}}{dt} \rightarrow x^{(1)}(k+1) - x^{(1)}(k) \tag{4.7}$$

2. Based on IAGO, $x^{(1)}(k+1) - x^{(1)}(k) = x^{(0)}(k+1)$

3. Also according to definition of $x^{(1)}(t)$, we define the background value as

$$x^{(1)}(t) \rightarrow 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) = z^{(1)}(k) \tag{4.8}$$

Clustering the mentioned above, we can get the grey difference equation of GM(1,1) as

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{4.9}$$

the GM(1,1) model means have first differential and only one variable, and equation (4.7) usually is called the source model of GM(1,1) model, and can use the sequence to build the approximation differential equation.

Although equation (4.7) is an approach to the differential equation, it not a true differential equation, therefore, we use a normally differential equation $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$ to replace the source model $x^{(0)}(k) + az^{(1)}(k) = b$, this kind of method does not have mathematics to prove, it is called the "Whiteness Processing", and $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$ is called the "shadow equation" of $x^{(0)}(k) + az^{(1)}(k) = b$.

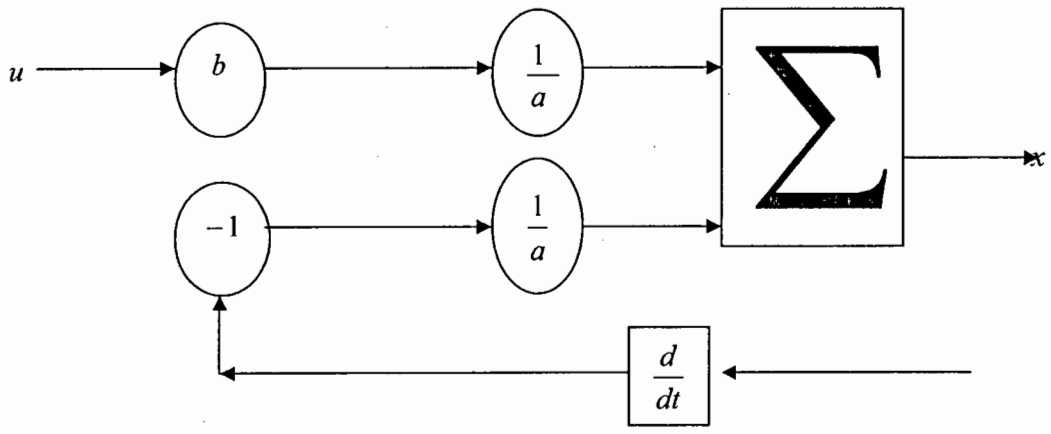


Figure 4.3 GM(1,1) model concept.

According to the solution of the first differential equation, the initial value is $x^{(0)}(1)$, and $x^{(0)}(1)=x^{(1)}(1)$, from the generally method, we can find the discrete response of GM(1,1) model $\frac{dx^{(1)}}{dt}+ax^{(1)}=b$ is

$$\hat{x}(k+1)=\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-ak}+\frac{b}{a} \tag{4.10}$$

where $x^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),x^{(0)}(3),\cdots,x^{(0)}(k)\right)$

$x^{(1)}=\left(x^{(1)}(1),x^{(1)}(2),x^{(1)}(3),\cdots,x^{(1)}(k)\right)$

$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)$

reducing equation (4.10) into equation (4.11), and equation (4.12) then

$$\hat{x}^{(1)}(k+1)=\left(x^{(0)}(1)\right)e^{-ak}+\frac{b}{a}\left(1-e^{-ak}\right) \tag{4.11}$$

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k) \tag{4.12}$$

we have the final prediction result formula as

$$\hat{x}^{(0)}(k+1)=\left(1-e^{a}\right)\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-ak} \tag{4.13}$$

4.3.2 Finding the parameters of GM(1,1) model

From equation (4.11), the first thing for GM(1,1) model, is to calculate the values of the parameters a and b. This has two kinds, called the least square method and parameter method. From equation $x^{(0)}(k)+az^{(1)}(k)=b$, substitute all values

$$x^{(0)}(2)=-az^{(1)}(2)+b$$

$$x^{(0)}(3)=-az^{(1)}(3)+b$$

$$x^{(0)}(4)=-az^{(1)}(4)+b$$

.....

$$x^{(0)}(n)=-az^{(1)}(n)+b \tag{4.14}$$

transferring equation (4.14) into matrix form $Y=B\hat{a}$, we have

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \cdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ \cdots & 1 \\ -z^{(1)}(n) & 1 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix} \tag{4.15}$$

where: $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \cdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ \cdots & 1 \\ -z^{(1)}(n) & 1 \end{bmatrix}, \hat{a} = \begin{bmatrix} a \\ b \end{bmatrix}.$

We use the inverse and matrix method to get the values of $\hat{a}$, that is, we use $\hat{a} = (B^T B)^{-1} B^T Y$ formula which gives the solution of the values a and b.

The values a and b can also be computed using the parameter method as follows:

$$a = \frac{CD - (n-1)E}{(n-1)F - C^2}, \; b = \frac{DF - CE}{(n-1)F - C^2} \tag{4.16}$$

where $C = \sum_{k=2}^{n} z^{(1)}(k)$, $D = \sum_{k=2}^{n} x^{(0)}(k)$, $E = \sum_{k=2}^{n} z^{(1)}(k) x^{(0)}(k)$, $F = \sum_{k=2}^{n} z^{(1)}(k)^2$, and n: is the number of data.

4.3.3 The error analysis of GM(1,1) model

This part deals with the residual error, or accuracy checking, so as to adopt constructed GM(1,1) model applicants to obtain error of the predicted value and the true (actual) value.

1 Defining the error

The traditional error of GM(1,1) model is defined as

$$e(k) = \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\%, \quad k \geq 2 \tag{4.17}$$

where $x^{(0)}(k)$: True value;

$\hat{x}^{(0)}(k)$: Predicted value.

2 The influence factor of error.

From section 4.2.1 equation (4.8), we know that the definition of $z^{(1)}(k)$ in GM(1,1) model is

$$z^{(1)}(k)=\alpha x^{(1)}(k)+(1-\alpha)x^{(1)}(k-1)=\alpha x^{(0)}(k)+x^{(1)}(k-1),\quad k\geq 2 \tag{4.18}$$

and the value of α is equal to 0.5. A GM(1,1) model is influenced by the $z^{(1)}(k)$ error, hence we focus on the adaptive of $z^{(1)}(k)$.

The analysis steps are:

1 Building original sequence as

$$x^{(0)}=\{x^{(0)}(1),x^{(0)}(2),x^{(0)}(3),x^{(0)}(4),\cdots,x^{(0)}(k)\},\qquad k\in N \tag{4.19}$$

2 Building AGO sequence

$$x^{(0)}=\{x^{(0)}(1),x^{(0)}(1)+x^{(0)}(2),\cdots,x^{(0)}(1)+x^{(0)}(2)+x^{(0)}(3)+\cdots+x^{(0)}(k)\},\qquad k\in N$$

3 According to equation (4.18) we have

$$z^{(1)}(k)=(\alpha x^{(0)}(2)+x^{(1)}(1),\alpha x^{(0)}(3)+x^{(1)}(2)),\cdots,\alpha x^{(0)}(k)+x^{(1)}(k-1)\quad k\geq 2 \tag{4.20}$$

4 Using the least square method to find the values of a and b, and substituting them into equation (4.9) to solve the response GM(1,1) model.

5 Based on the error formula to calculate the error

$$\hat{x}^{(0)}(k)-x^{(0)}(k)=(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a(k-1)}-x^{(0)}(k)\qquad k\geq 2 \tag{4.21}$$

the errors are

$$\hat{x}^{(0)}(2)-x^{(0)}(2)=(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a\times 1}-x^{(0)}(2),$$

$$\hat{x}^{(0)}(3)-x^{(0)}(3)=(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a\times 2}-x^{(0)}(3),$$

$$\hat{x}^{(0)}(4)-x^{(0)}(4)=(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a\times 3}-x^{(0)}(4),$$

$$\cdots,$$

$$\hat{x}^{(0)}(n)-x^{(0)}(n)=(1-e^{a})\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a\times(n-1)}-x^{(0)}(n)$$

6 Calculating the total error

$$\bar{e}(\alpha)=\sum_{k=2}^{n}|e(k)|=|e(2)|+|e(3)|+|e(4)|+\cdots+|e(n)| \tag{4.22}$$

From equation (4.20), the value of α, $\alpha \in (0,1]$, could be changed at each step based on the minimum error to decide the optimal α value and use the known α value to calculate the next point in GM(1,1) model. That is a change in α value of is done so as to decrease the error in the GM(1,1) model.

4.4 The GM(1,1) rolling model

The grey system GM(1,1) rolling model is a technique based on the forward data sequence, it updates the input data by discarding old data for each cycling and predicting the value of the next point. This provides a means of guarantee input data are always the most recent values. The procedure is repeated till the end of the sequence. The use of a new information method of the same order can effectively correct the error caused by long-term prediction and eliminate interference. The analysis step for the GM(1,1) rolling model is:

Assume the original sequence is

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)\right), \quad n \geq 4 \tag{4.23}$$

Take the partial of original sequence, referred to as sub sequence

$$x^{(0)}(i;k) = \left(x^{(0)}(i), x^{(0)}(i+1), x^{(0)}(i+2), \ldots, x^{(0)}(k)\right) \tag{4.24}$$

When $i = 1$ means $x^{(0)}(1;k) = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(k)\right)$.

Equation (4.24) is called normal sequence, and basing on this sequence we build the GM(1,1) model, which gives the prediction value $\hat{x}^{(1)}(k+1)$.

In this section, the symbols are defined as:

$$G\hat{M}_p \bullet AGO : x^{(0)}(i;k) \rightarrow (a,b) \tag{4.25}$$

$$G\hat{M}_p \bullet AGO : x^{(0)}(i;k) \rightarrow x^{(0)}(k) + az^{(1)}(k) = b \tag{4.26}$$

$$IAGO \bullet G\hat{M}_p \bullet AGO : x^{(0)}(i;k) \rightarrow \hat{x}^{(0)}(k+1) \tag{4.27}$$

when $i = 1, k = 4,5,6,\ldots,n-1$, the error is defined as

$$e(k=1) = \left| \frac{x^{(0)}(k=1) - x^{(0)}(k+1)}{x^{(0)}(k=1)} \right| \times 100\%, \quad k+1 \leq n \tag{4.28}$$

and $e(k+1)$ is called the prediction error of GM(1,1) model at $(k+1)$ the rolling error of GM(1,1) is also defined as:

$$e = \frac{1}{n-4}\sum_{k=4}^{n-1} e(k+1)\% \tag{4.29}$$

then we have the precision equals $\varepsilon = (1+e)\%$

4.5 The Verhulst GM model

The Verhulst model was proposed by a German biologist Verhulst in 1837 as a modification of the Malthusian model to make the model more suitable to real situations. The formula for Malthusian model is $p(t) = p(t_0)e^{a(t-t_0)}$, where t_0, is the beginning of time and $p(t_0)$ is the initial value. The formula of Verhulst model is as follows:

$$p(t) = \frac{ap_0}{bp_0 + (a - bp_0)^{e-a(t-t_0)}}$$

where t_0 is the beginning time and $p(t_0)$ is the initial value.

The Verhulst model is mainly used in many fields to describe and study process states, like market prediction, population prediction, biology multiple prediction etc. It is also used in the grey system model for predicting so as to limit the prediction value.

Assume that $x^{(0)}$ and $z^{(1)}$ is a sequence of original data, $x^{(0)}$ is the AGO sequence of $x^{(0)}$ and $z^{(1)}$ is the mean generated of consecutive neighbors of $x^{(0)}$, then

$$x^{(0)}(k) + az^{(1)}(k) = b\left(z^{(1)}(k)\right)^{\alpha} \tag{4.30}$$

is called the GM(1,1) power model, and the equation

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b\left(z^{(1)}(k)\right)^2 \tag{4.31}$$

Is the whitenization equation of the GM(1,1) power model.

When $\alpha = 2$,

$$x^{(0)}(k) + az^{(1)}(k) = b\left(z^{(1)}(k)\right)^2 \tag{4.32}$$

Is called the grey Verhulst model and the equation

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b\left(x^{(1)}\right)^2 \tag{4.33}$$

which is defined as the whitenization equation of the grey Verhulst model. Using the mathematics concept, of the grey difference equation of equation (4.18) is

$$x^{(0)}(k) + az^{(1)}(k) = b\left(z^{(1)}(k)\right)^2 \Rightarrow x^{(0)}(k) = -az^{(1)}(k) + b\left(z^{(1)}(k)\right)^2 \tag{4.34}$$

same as the method of GM(1,1) model, to calculate the values of parameters and parameters a and b.

$$x^{(0)}(2) = -az^{(1)}(2) + b\left(z^{(1)}(2)\right)^2$$

$$x^{(0)}(3) = -az^{(1)}(3) + b\left(z^{(1)}(3)\right)^2$$

$$x^{(0)}(4) = -az^{(1)}(4) + b\left(z^{(1)}(4)\right)^2$$

$$\cdots\cdots\cdots\cdots$$

$$x^{(0)}(n) = -az^{(1)}(n) + b\left(z^{(1)}(n)\right)^2 \tag{4.35}$$

transfer equation (4.35) into matrix form $Y = B\hat{a}$, then we have

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \cdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & \left(z^{(1)}(2)\right)^2 \\ -z^{(1)}(3) & \left(z^{(1)}(3)\right)^2 \\ -z^{(1)}(4) & \left(z^{(1)}(4)\right)^2 \\ \cdots & \cdots \\ -z^{(1)}(n) & \left(z^{(1)}(n)\right)^2 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix} \tag{4.36}$$

where: $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \cdots \\ x^{(0)}(n) \end{bmatrix}$, $B = \begin{bmatrix} -z^{(1)}(2) & \left(z^{(1)}(2)\right)^2 \\ -z^{(1)}(3) & \left(z^{(1)}(3)\right)^2 \\ -z^{(1)}(4) & \left(z^{(1)}(4)\right)^2 \\ \cdots & \cdots \\ -z^{(1)}(n) & \left(z^{(1)}(n)\right)^2 \end{bmatrix}$, $\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix}$ and use $\hat{a} = \left(B^T B\right)^{-1} B^T Y$ formula to

solve the values a and b then the values a and b are given by

$$a = \frac{DH - HE}{FG - D^2},\ b = \frac{EF - DE}{FG - D^2} \tag{4.37}$$

where $D = \sum_{k=2}^{n}\left(z^{(1)}(k)\right)^3$, $E = \sum_{k=2}^{n} x^{(1)}(k) \times x^{(0)}(k)$, $F = \sum_{k=2}^{n}\left(z^{(1)}(k)\right)^2$, $G = \sum_{k=2}^{n}\left(z^{(1)}(k)^2\right)^4$,

$H = \sum_{k=2}^{n}\left(z^{(1)}(k)\right)^2 \times x^{(0)}(k)$, n: the number of data.

Substituting the values from D to H into a and b, the solution of Verhulst model is shown in the equation (4.38) below

$$\hat{x}^{(1)}(k)=\frac{\frac{a}{b}}{1+\left(\frac{a}{b}\times\frac{1}{x^{(0)}(1)-1}\right)e^{a(k-1)}} \qquad k\geq 2 \tag{4.38}$$

and use IAGO method, reducing equation (4.38) into

$$\hat{x}^{(1)}(k)=\frac{\left(1-e^{-a}\right)\times\frac{a}{b}\times\left[1+\left(\frac{a}{b}\times\frac{1}{x^{(0)}(1)}\right)e^{a(k-1)}\right]}{\left[1+\left(\frac{a}{b}\times\frac{1}{x^{(0)}(1)}-1\right)e^{a(k-1)}\right]\left[1+\left(\frac{a}{b}\times\frac{1}{x^{(0)}(1)}-1\right)e^{a(k-2)}\right]} \qquad k\geq 2 \tag{4.39}$$

from equation (4.39), we know that if $a<0$ then $\lim_{k\to\infty}\hat{x}^{(0)}(k)\to\frac{a}{b}$, it means that the saturation point in equation (4.39) is $\frac{a}{b}$, this value limits the prediction value, in the other word, it is saturation point of $\hat{x}^{(0)}(k)$.

4.6 Non equal-spaced GM(1,1) model

Assume $x^{(0)}$ an original series to be $x^{(0)}=\left(x^{(0)}(t_1),x^{(0)}(t_2),x^{(0)}(t_3),\cdots,x^{(0)}(t_n)\right)$, where $x^{(0)}(t_i)$ is corresponding to the system output at time t_i. If time taken by the process for a product to be completed differs, time to time between the consecutive products, the sampling interval $\Delta t_k=t_k-t_{k-1}\neq constant\ \ (k=2,3,\ldots,n)$. In this situation $x^{(0)}$ is referred to as a non equal-spaced series. In this section we be considering GM(1,1) model for unequal-spaced data, denoted by $GM_n(1,1)$.

The AGO sequence of $x^{(0)}$ is $x^{(1)}=\left(x^{(1)}(t_1),x^{(1)}(t_2),x^{(1)}(t_3),\cdots,x^{(1)}(t_n)\right)$, that is,

$$AGO\left\{x^{(0)}(t_k)\right\}=x^{(1)}(t_k)=\left(\sum_{k=1}^{1}x^{(0)}(t_k),\sum_{k=1}^{2}x^{(0)}(t_k),\cdots,\cdots,\cdots,\sum_{k=1}^{n}x^{(0)}(t_k)\right) \tag{4.40}$$

which can be summarized as:

$$x^{(1)}(t_k) = \sum_{m=1}^{k} x^{(0)}(t_m) \quad k = 1,2,\ldots,n \tag{4.41}$$

The grey derivative for first order grey differential equation in accordance AGO is defined in the following manner:

$$\frac{dx^{(1)}(t_k)}{dt_k} = \lim_{\Delta t \to 0} \frac{x^{(1)}(t_k + \Delta t_k) - x^{(1)}(t_k)}{\Delta t_k}$$

and $$\frac{dx^{(1)}(t_k)}{dt_k} = \frac{\Delta x^{(1)}(t_k)}{\Delta t_k}$$

Therefore $$\frac{dx^{(1)}(t_k)}{dt_k} + ax^{(1)}(t_k) = b$$

$$\frac{\Delta x^{(1)}(t_k)}{\Delta t_k} + ax^{(1)}(t_k) = b, \qquad k = 2,3,\ldots,n \tag{4.42}$$

Where

$$\Delta x^{(1)}(t_k) = x^{(1)}(t_k) - x^{(1)}(t_{k-1}) = x^{(0)}(t_k)$$

$$\Delta t_k = t_k - t_{k-1}$$

According to definition of $x^{(1)}(t_k)$, we define the background value as:

$$x^{(1)}(t_k) \to 0.5\left[x^{(1)}(t_k) + x^{(1)}(t_{k-1})\right] = z^{(1)}(t_k), \quad k = 2,3,\ldots,n \tag{4.43}$$

Using equation $x^{(0)}(t_k)$ to replace $\Delta x^{(1)}(t_k)$ in equation (4.42) we get:

$$\frac{x^{(0)}(t_k)}{\Delta t_k} + ax^{(1)}(t_k) = b$$

and then using $z^{(1)}(t_k)$, to replace $x^{(1)}(t_k)$ the $GM_n(1,1)$ model can be written as:

$$\frac{x^{(0)}(t_k)}{\Delta t_k} + az^{(1)}(t_k) = b$$

$$x^{(0)}(t_k) + az^{(1)}(t_k)\Delta t_k = b\Delta t_k \quad k = 2,3,\ldots,n$$

The values of the parameters a and b can be estimated using the least square method from equation $x^{(0)}(t_k) + az^{(1)}(t_k)\Delta t_k = b\Delta t_k$ by substituting all values we have:

$$x^{(0)}(t_2) = -az^{(1)}(t_2)\Delta t_2 + b$$

$$x^{(0)}(t_3) = -az^{(1)}(t_3)\Delta t_3 + b$$

.....

$$x^{(0)}(t_n) = -az^{(1)}(t_n)\Delta t_n + b$$

Then we transfer into matrix form $Y = B\hat{a}$, which gives

$$\begin{bmatrix} x^{(0)}(t_2) \\ x^{(0)}(t_3) \\ x^{(0)}(t_4) \\ \cdots \\ x^{(0)}(t_n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(t_2) * \Delta t_2 & \Delta t_2 \\ -z^{(1)}(t_3) * \Delta t_3 & \Delta t_3 \\ -z^{(1)}(t_4) * \Delta t_4 & \Delta t_4 \\ \cdots & \cdots \\ -z^{(1)}(t_n) * \Delta t_n & \Delta t_n \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix}$$

The values of $\hat{a}$, are then got using the inverse and matrix method $\hat{a} = (B^T B)^{-1} B^T Y$ which estimates the values for the parameters a and b. These values for the parameters can also be computed using the parameter method.

$$a = \frac{CD - (n-1)E}{(n-1)F - C^2}, \text{ and } b = \frac{DF - CE}{(n-1)F - C^2} \tag{4.44}$$

where $C = \sum_{k=2}^{n} z^{(1)}(t_k)\Delta t_k$, $D = \sum_{k=2}^{n} x^{(0)}(t_k)$, $E = \sum_{k=2}^{n} z^{(1)}(t_k)x^{(0)}(t_k)$, $F = \sum_{k=2}^{n} [z^{(1)}(t_k)\Delta t_k]^2$, n: the number of data.

The estimated values of $x^{(1)}(t_k)$ can be obtained by

$$\hat{x}^{(1)}(t_k) = \left(x^{(0)}(t_1) - \frac{b}{a}\right)e^{-a(t_k - t_1)} + \frac{b}{a} \tag{4.45}$$

Applying IAGO to the above equation (4.45) it leads to

$$\hat{x}^{(0)}(t_1) = x^{(0)}(t_1)$$

$$\hat{x}^{(0)}(t_k) = \left(x^{(0)}(t_1) - \frac{b}{a}\right)\left(1 - e^{a\Delta t_k}\right)e^{-a(t_k - t_1)}, \qquad k = 2,3,\ldots,n \tag{4.46}$$

Where $\hat{x}^{(0)}$ denotes the fitting series of raw data. Zhang et al [??] pointed out that the condition to contrast this model is that "the ratio of the maximal sampling interval to the minimum sampling interval should be less than 2".

4.7 GM(2,1) model

This means second differential and one variable, the function of this GM(2,1) can be applied to describe non-monotonic process of change.

Assuming that

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n)\right) = \left(x^{(0)}(k);\quad k = 1,2,3,\cdots,n\right)$$

is a sequence of raw data, its first order accumulated generating operation (1-AGO) sequence $x^{(1)}$ is

$$x^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \cdots, x^{(1)}(n)\right)$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i) \qquad k = 1,2,\ldots,n,$$

and the first order inverse accumulative generating operation (1-IAGO) sequence $\alpha^{(1)}x^{(0)}$ of $x^{(0)}$ is

$$\alpha^{(1)}x^{(0)} = \left(\alpha^{(1)}x^{(0)}(1), \alpha^{(1)}x^{(0)}(2), \cdots, \alpha^{(1)}x^{(0)}(n)\right) \tag{4.47}$$

where

$$\alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1)$$

$k = 1,2,\ldots,n$ and the sequence mean generated of consecutive neighbor of $x^{(1)}$ is

$$z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)\right)$$

where

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \qquad k = 1,2,\ldots,n \text{ then}$$

$$\alpha^{(1)}x^{(0)} + a_1 x^{(0)} + a_2 z^{(1)} = b \tag{4.48}$$

is the GM(2,1) grey differential equation.

A whitenization equation of a GM(2,1) grey differential equation is

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2 x^{(1)} = b \tag{4.49}$$

GM(2,1) differential equation $\alpha^{(1)}x^{(0)}(k) + a_1 x^{(0)}(k) + a_2 z^{(1)}(k) = b$ can be written in the form

$$\alpha^{(1)}x^{(0)}(2)+a_1x^{(0)}(2)+a_2z^{(1)}(2)=b$$

$$\alpha^{(1)}x^{(0)}(3)+a_1x^{(0)}(3)+a_2z^{(1)}(3)=b$$

$$\cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots$$

$$\alpha^{(1)}x^{(0)}(n)+a_1x^{(0)}(n)+a_2z^{(1)}(n)=b$$

Which makes it to be transferable to matrix form $Y = B\hat{a}$, then we have

$$\begin{bmatrix}\alpha^{(1)}x^{(0)}(2)\\ \alpha^{(1)}x^{(0)}(3)\\ \alpha^{(1)}x^{(0)}(4)\\ \cdots\\ \alpha^{(1)}x^{(0)}(n)\end{bmatrix}=\begin{bmatrix}-x^{(0)}(2) & -z^{(1)}(2) & 1\\ -x^{(0)}(3) & -z^{(1)}(3) & 1\\ -x^{(0)}(4) & -z^{(1)}(4) & 1\\ \cdots & \cdots & \cdots\\ -x^{(0)}(n) & -z^{(1)}(n) & 1\end{bmatrix}\times\begin{bmatrix}a_1\\ a_2\\ b\end{bmatrix}$$

Where:

$$Y=\begin{bmatrix}\alpha^{(1)}x^{(0)}(2)\\ \alpha^{(1)}x^{(0)}(3)\\ \alpha^{(1)}x^{(0)}(4)\\ \cdots\\ \alpha^{(1)}x^{(0)}(n)\end{bmatrix}=\begin{bmatrix}x^{(0)}(2)-x^{(0)}(1)\\ x^{(0)}(3)-x^{(0)}(2)\\ x^{(0)}(4)-x^{(0)}(3)\\ \cdots\\ x^{(0)}(n)-x^{(0)}(n-1)\end{bmatrix},\ B=\begin{bmatrix}-x^{(0)}(2) & -z^{(1)}(2) & 1\\ -x^{(0)}(3) & -z^{(1)}(3) & 1\\ -x^{(0)}(4) & -z^{(1)}(4) & 1\\ \cdots & \cdots & \cdots\\ -x^{(0)}(n) & -z^{(1)}(n) & 1\end{bmatrix},\ \hat{a}=\begin{bmatrix}a_1\\ a_2\\ b\end{bmatrix}.$$ We the

use and use $\hat{a}=\left(B^TB\right)^{-1}B^TY$ to get the least square estimate of the parameter values of $\hat{a}$.

Before we state the "theorems" for GM(2,1) models, we must pointed out the solution to a ordinary second order one-variable differential equation can be a standard calculus exercise in university calculus courses. However, in grey differential equation modeling these results are critical so that no matter in Deng's publications or Liu and Lin (2006) they were all stated as theorems. We just accepted them with the consciousness that grey theory is a part of approximation theory.

Theorem 4.1. (Sifeng Lui and Yi Lin, 2006)

Having the solved for the values of $\hat{a}$ the solution of the GM(2,1) whitenization equation (4.49) the following hold true:

1. If $x^{(1)^*}$ is a special solution of

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2x^{(1)} = b$$

and $\bar{x}^{(1)}$ is a general solution oh the homogeneous equation

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2x^{(1)} = 0 \qquad (4.50)$$

then $x^{(1)^*} + \bar{x}^{(1)}$ is the general solution of the GM(2,1) whitenization equation;

2. Considering the homogeneous equation homogeneous second-order differential equation (4.50): the roots of its characteristic equation

$$r^2 + a_1r + a_2 = 0 \qquad (4.51)$$

are solved by using a quadratic formula we get two values of r as

$$r_1, r_2 = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2}$$

There are three cases of this general solution as follows:

Case 1: If $a_1^2 - 4a_2 > 0$ and $r_1 \neq r_2$ where r_1 *and* r_2 are real numbers(R), then

$$\bar{x}^{(1)} = C_1e^{r_1t} + C_2e^{r_2t};$$

Case 2: If $a_1^2 - 4a_2 = 0$, the roots of the characteristic equation (4.51) are equal $r_1 = r_2 = 0.5a_1$, it implies that characteristic equation has a real solution of multiplicity 2, $r_1 = r_2 = r$, $r \in R$ then

$$\bar{x}^{(1)} = e^{r_2t}(C_1 + C_2t);$$

Case 3: If $a_1^2 - 4a_2 < 0$, the two roots of the characteristic equation (4.51) has two complex conjugate solutions $r_1 = \alpha + \beta i$, $r_2 = \alpha - \beta i$, $\alpha = -a_1/2$, $\beta = \sqrt{4a_2 - a_1^2}/2$, then

$$\bar{x}^{(1)} = e^{r_2t}(C_1\cos\beta t + C_2\sin\beta t).$$

3. There exist three possibilities for a special solution of a whitenization equation:

(a) When zero is not a solution of the characteristic equation,

$$x^{(1)^*} = C;$$

(b) When zero is a solution of multiplicity 1 of the characteristic equation,

$$x^{(1)^*} = Cx;$$

(c) When zero is a multiplicate solution of the characteristic equation,

$$x^{(1)^*} = Cx^2.$$

4.7.1 Non equal-spaced $GM_u(2,1)$ model

For the original discrete data sequence that is not equal spaced $x^{(0)} = \left(x^{(0)}(t_1), x^{(0)}(t_2), \cdots, x^{(0)}(t_n)\right)$, the second order one variable grey differential equation, abbreviated as $GM_u(2,1)$, possesses the following form

$$a^{(1)}x^{(0)}(k) + a_1 x^{(0)}(k)\Delta t_k + a_2 z^{(1)}(k)(\Delta t_k)^2 = b(\Delta t_k)^2,\ k = 2,3,\cdots,n \qquad (4.52)$$

and its least-square estimator for parameter vector $P=(a_1, a_2, b)^T$ is given by equation $P = \left(B^T B\right)^{-1} B^T Y$ where the matrices **B** and **Y** are in the following manner

$$B = \begin{bmatrix} -x^{(0)}(2)\Delta t_2 & -z^{(1)}(2)(\Delta t_2)^2 & (\Delta t_2)^2 \\ -x^{(0)}(3)\Delta t_3 & -z^{(1)}(3)(\Delta t_3)^2 & (\Delta t_3)^2 \\ -x^{(0)}(4)\Delta t_4 & -z^{(1)}(4)(\Delta t_4)^2 & (\Delta t_4)^2 \\ \vdots & \vdots & \vdots \\ -x^{(0)}(n)\Delta t_n & -z^{(1)}(n)(\Delta t_n)^2 & (\Delta t_n)^2 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} a^{(1)}x^{(0)}(2) \\ a^{(1)}x^{(0)}(3) \\ a^{(1)}x^{(0)}(4) \\ \vdots \\ a^{(1)}x^{(0)}(n) \end{bmatrix}$$

Once the parameter vector $P=(a_1, a_2, b)^T$ is determined, the solution to the $GM_u(2,1)$ model (4.52) can be determined.

There is a special case of GM(2,1) model, named DGM(2,1) by Liu et al (2004). DGM(2,1) model possesses the form

$$a^{(1)}x^{(0)}(k)+ax^{(0)}(k)=b,\ k=2,3,\cdots,n \tag{4.53}$$

with the corresponding whitenization equation

$$\frac{d^2x^{(1)}}{dt^2}+a\frac{dx^{(1)}}{dt}=b \tag{4.54}$$

The parameter vector $P=(a,b)^T$ can be estimated in terms of least-square approach and then the response sequence of DGM(2,1) is

$$\hat{x}^{(1)}(k+1)=\left(\frac{b}{a^2}-\frac{x^{(0)}(1)}{a}\right)e^{-ak}+\frac{b}{a}(k+1)+\frac{1+a}{a}\left(x^{(0)}(1)-\frac{b}{a}\right) \tag{4.55}$$

And the filtered value sequence is

$$\hat{x}^{(0)}(k+1)=a^{(1)}\hat{x}^{(1)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k),\ k=2,3,\cdots,n \tag{4.56}$$

The unequal-spaced DGM(2,1) model, denoted as $DGM_u(2,1)$, can be defined and treated similar to the $GM_u(2,1)$ model.

4.8 Extensions of GM Model

Theorem 4.2 (Liu and Lin, 2006)

The model GM(1,1)

$$x^{(0)}(k)+az^{(1)}(k)=b$$

can be transformed into

$$x^{(0)}(k)=\beta-\alpha x^{(1)}(k-1)$$

where

$$\beta=\frac{b}{1+0.5a},\quad \alpha=\frac{a}{1+0.5a}$$

Theorem 4.3 (Liu and Lin, 2006)

Assume that $\beta = \frac{b}{1+0.5a}, \quad \alpha = \frac{a}{1+0.5a}$

and $\hat{X}^{(1)} = \left(\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \ldots, \hat{x}^{(1)}(n)\right)$

is the time response sequence of GM(1,1) model, where

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)} + \frac{b}{a}, \qquad k = 1,2,\ldots,n,$$

then $x^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right]e^{-a(k-2)}$

Example Assume that

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), x^{(0)}(5)\right)$$
$$= (3.34986, 3.75386, 3.81286, 3.86586, 4.15486)$$

is a given sequence of raw data. We will be using the following GM models to simulate to simulate $x^{(0)}$ and compare their simulation accuracies:

- $x^{(0)}(k) + az^{(1)}(k) = b$
- $x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1)$
- $x^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right]e^{-a(k-2)}$

The solution to this is done via many steps which we have listed below and their results are summarized in Table 4.2.

1. We apply the first order AGO on $x^{(0)}$ to get $x^{(1)}$
2. We perform a quasi-smoothness check on $x^{(0)}$ using the equation

   $\rho(k) = \frac{x^{(0)}(k)}{x^{(1)}(k-1)}$. From Table 4.1 $\rho(4) < 0.5$ *and* $\rho < (0.5)$ this implies that for the case $k > 3$, the condition of being quasi-smooth is satisfied.

3. Checking to see whether or not $x^{(1)}$ satisfies the law of quasi-exponentiality we use $\sigma^{(1)}(k)=\frac{x^{(0)}(k)}{x^{(1)}(k-1)}$. For $k>3$, $\sigma^{(1)}(k)\in[1,1.5]$, $\delta=0.5$. This satisfies the law of quasi-exponentiality. Therefore we can establish GM(1,1) model for $x^{(1)}$.
4. Applying a consecutive neighbor generation to $x^{(1)}$ we let $x^{(1)}(t)\to 0.5x^{(1)}(k)+0.5x^{(1)}(k-1)=z^{(1)}(k)$
5. Use a least square estimate for a parametric sequence $\hat{a}=[a,b]^T$. We obtain that $\hat{a}=(B^TB)^{-1}B^TY$.
6. Determining the model we have a = -0.034070644 and b = 3.511275535 and the time response sequence is $\hat{x}^{(1)}(k+1)=\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a(k-1)}+\frac{b}{a}$, $k=1,2,\ldots,n$,

Table 4.2 Summarized results

(k)	$x^{(0)}$	$x^{(1)}$	$\rho(k)$	$\sigma^{(1)}(k)$	$Z^{(1)}$	$\hat{x}^{(1)}(k+1)$
1	3.34986	3.34986			3.34986	3.34986
2	3.75386	7.10372			5.22679	7.037734823
3	3.81286	10.91658	0.536741313	1.536741	9.01015	10.85342289
4	3.86586	14.78244	0.354127392	1.354127	12.84951	14.80135392
5	4.15486	18.9373	0.281067266	1.281067	16.85987	18.88611115

7. Solve for the simulation values of $x^{(1)}$
8. Restore to find the simulation value of $x^{(0)}$. From $\hat{x}^{(0)}(k)=\alpha^{(1)}\hat{x}^{(1)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$
9. We evaluate the error, these are given in Table 4.2

Table 4.3 Comparison of simulation accuracies between three GM models

	Real data	Simulated data $\hat{x}^{(0)}(k)$			Errors $\varepsilon(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$			Relative error $\Delta_k=\frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$		
		Model			Model			Model		
No	$x^{(0)}(k)$	1	2	3	1	2	3	1	2	3
1	3.34986									
2	3.75386	3.68787	3.68824	3.68824	0.06599	0.06562	0.06562	1.758%	1.748%	1.748%
3	3.81286	3.81569	3.81835	3.81832	-0.00283	-0.00549	-0.00546	0.074%	0.144%	0.143%
4	3.86586	3.94793	3.95051	3.95298	-0.08207	-0.08465	-0.08712	2.123%	2.190%	2.254%
5	4.15486	4.08476	4.08450	4.09240	0.07010	0.07036	0.06246	1.687%	1.693%	1.503%
						Average relative errors		1.411%	1.444%	1.412%

Key: Model 1 $\hat{x}^{(1)}(k+1)=\left(x^{(0)}(1)-\frac{b}{a}\right)e^{-a(k-1)}+\frac{b}{a}$

Model 2 $x^{(0)}(k)=\beta-\alpha x^{(1)}(k-1)$

Model 3 $x^{(0)}(k)=\left[\beta-\alpha x^{(0)}(1)\right]e^{-a(k-2)}$

square sum of errors(s) and average relative error () are computed using the following formulae respectively $s=\varepsilon'\varepsilon$ and $\Delta=\frac{1}{4}\sum_{k=2}^{5}\Delta k$.

We have computed a and b

So, $\beta=\frac{b}{1+0.5a}$ and $\alpha=\frac{a}{1+0.5a}$

and it follows that

$\hat{x}^{(0)}(k)=\beta-\alpha x^{(1)}(k-1)$ and $\hat{x}^{(0)}=\left(\hat{x}^{(0)}(1),\hat{x}^{(0)}(2),\hat{x}^{(0)}(3),\hat{x}^{(0)}(4),\hat{x}^{(0)}(5)\right)$ are computed and the errors are evaluated. Further the square sum of errors and average relative error are and the values of α *and* β are computed and so we have $x^{(0)}(k)=\left[\beta-\alpha x^{(0)}(1)\right]e^{-a(k-2)}$.

From the square sums of errors and average relative errors of the three models used, it can be seen that the exponential models 1 and 3 have relatively high accuracy, than the accuracy of the difference model 2, that accuracy is relatively low.

Theorem 4.4 (Liu and Lin, 2006)

If is a quasi-smooth sequence, then the development coefficient a of the GM(1,1) model of its first order AGO sequence X(1) can be written as

$$\alpha = \frac{\dfrac{b}{x^{(1)}(k-1)} - \rho(k)}{1+0.5\rho(k)}$$

where $\rho(k) = \dfrac{x^{(0)}(k)}{x^{(1)}(k-1)}$

4.9 GM(h,N) model

The GM(h,N) model can be written in the form

$$\sum_{i=0}^{h} a_i \frac{d^{(i)}x_1^{(i)}}{dt^{(i)}} = \sum_{j=2}^{N} b_j x_j^{(1)}(k) \qquad (4.57)$$

where a_1 and b_j are determined coefficients.

$x_1^{(1)}(k)$ is the major sequence

$x_j^{(1)}(k)$ is the influencing sequence

$$AGO\ x^{(0)} = x^{(1)} = \left[\sum_{k=1}^{1} x^{(0)}(k), \sum_{k=1}^{2} x^{(0)}(k), \ldots, \sum_{k=1}^{n1} x^{(0)}(k)\right]$$

The main function of GM(h,N) model of the grey system is to carry out the calculation of measurement among the discrete sequence and compensate the shortcoming in the traditional methodology.

4.10 Summary

We have discussed that the sparse data information availability is the main characteristic of a grey system. One of the grey models, GM(1,1) model has been applied, to obtain the required accuracy, for predicting a reduced number of measurement points required when monitoring the process. This approach helps to reduce the time incurred when using statistical quality control measurements and estimating values when the information is incomplete. It has also been shown

that a grey system has practical significance, since it has minor error due to the fact that it uses the rolling model technique which uses current values in the process for prediction.

Chapter 5 Grey predictive Control Charts

We have reviewed quality control charts and grey system theory separately in previous chapters. In this chapter we will are aiming at applying a grey system theory in quality control charts and reach the balance between the smallest number of data required in process monitoring and grey system prediction accuracy. We also want to prove that is it possible to avoid too many number of or huge sample size, which increases the monitoring time, while achieving the required output quality accuracy. Through out the process the accuracy of the prediction model is also taken into consideration. The grey predictive model will be used in Shewhart and CUSUM quality control charts. Fuzzy logic's membership function will also be used for classifying and grading the output in a Grey-Fuzzy Predictive Control Chart. The application of the proposed grey predictive models will be in Chapter 6.

5.1 A Grey Predictive Shewhart Control Chart

One of the aspects of statistics is dealing with variability. This aspect of statistics is used to predict and control the performance of a system based on measurements of the output from the system in a manufacturing industry. Measurements of output form the system in any company will always exhibits variability which can be described in statistical terms. This characteristic enables the manufacturers to use statistical methods to monitor and control processes. This is done by taking samples from the output of a process, measuring them, plotting them in a quality control chart and interpreting the resulting quality control chart. This helps the manufacturer to see whether the process is under control. If it process is under control it implies that the output is of good quality and it must be maintained. Tracking or monitoring of the process is usually done using a process control chart Shewhart control charts.

Shewhart control charts require a large number of data as discussed in Chapter 1, but currently many production industries are operating in short-run manufacturing process.

This has lead to the numbers of data available to be small and also incomplete. In these situations we feel that the grey system would be the most relevant method since the grey system only needs a minimum of four data. Having small sample size data and monitoring a process to ensure that the output is of good quality we will combine a grey predictive model GM(1,1) and a Shewhart control charts to come up with a Grey Predictive Shewhart control chart. Grey Predictive Shewhart control chart would be useful in forecasting the change in the trend of the data sequences and to decide the variation regulations. Historical data from a company is used as experimental data to verify the output values for grey prediction values by setting the control limits. When a predicted value exceeds the control limits, the manufacturing process is examined immediately. The procedures of constructing a control chart of this type are defined in the following section.

5.1.1 Procedures of construction a Grey Predictive Shewhart control chart

Assume that the original raw data series in the process is denoted by $x^{(0)}(k), k = 1,2,\cdots,n$ where k is the observational number or time of sampling or the batch number of sampling and n is the total observed observations on the quality index Q. The sample size should be between four and six, $4 \le n \le 6$. Recall that in a Shewhart control chart the process is said to be 'in-control' if the values fall within the upper and the lower control limits, the bandwidth. In a Grey Predictive Shewhart control chart the process would also be also monitored basing on control limits. This procedure would be broken down into stages that includes control limits and establishment of a grey model combined with properties analyses and conformation of the grey system model with properties and quantification analyses.

Stage 1: The first stage deals with computing the control limits this would be done in the same way as it is done in Shewhart's control limits for individual observations, which we will give us the grey interval.

1 We compute the center line by averaging the first five observations and denote it by C_s

$$C_s = \overline{x^{(0)}(k)} = \frac{1}{n}\sum_{k=1}^{n} x^{(0)}(k) \qquad k = 1, 2, \ldots, n$$

2 For the non-standards-given case, we estimate the standard deviation by using the moving range or the standard deviation method, since a grey system deals with discrete individual observations. The concept of standard deviation is as follows:

(a) Moving range method: This is calculated as the absolute difference between the two concessive observations

$$MR = \left| x^{(0)}(k) - x^{(0)}(k-1) \right|$$

The differences are then averaged to get an average moving range

$$\overline{MR} = \frac{1}{n-1}\sum_{i=2}^{m} MR_i$$

This is then divided by d_2 to estimate the standard deviation, $d_2 = 1.128$, from Table C.1 in Appendix C, since $n = 2$ in moving range method.

(b) Standard deviation method: This is calculated as

$$s = \sigma = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(x^{(0)}(k) - \overline{x^{(0)}}\right)^2}$$

3. Compute the upper limit and the lower limit, denoted by L_s and U_s respectively, so the bandwidth is $[L_s, U_s]$.

$$U_s = C_s + 3\overline{MR}/d_2 = C_s + E_2\overline{MR}$$
$$L_s = C_s - 3\overline{MR}/d_2 = C_s - E_2\overline{MR}$$

where $E_2 = 3/d_2 = 3/1.128 \approx 2.659574468$

or by $U_s = C_s + 3s$
$L_s = C_s - 3s$

When the control limits which we refer to as action limits are developed and the process is on operation we will also look for signals on the chart, so that if any occurs, we will

attempt to determine the cause behind it. The limits for these warning lines are computed as:

$C_s \pm 2s$
$C_s \pm 1s$

Stage 2: Constructing a grey prediction model:

1 We get the first five values from the raw data series and denote them by $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ *and* Δ_5.

2 Let $x^{(0)} = X_1 = (\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5)$

3 Derive the one-order accumulated generating operation (1-AGO) for the original series 1 $AGO\left\{x^{(0)}(k)\right\} = x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m), \quad k = 1,2,,3,4,5$

4 Plot the $x^{(1)}(k)$ values against their k values, to check if there is a monotonic process of change. If there is monotonic process of change proceed on to the next step, where we will be using the GM(1,1) model. If there is no monotonic process of change used the GM(2,1) model described in subsection 5.1.2.

5 Derive the one-time inverse-accumulated generating operation (IAGO) $x^{(1)}(k) = \sum_{m=1}^{k-1} x^{(0)}(m) + x^{(0)}(k) = x^{(1)}(k-1) + x^{(0)}(k)$

6 Derive the average value $z^{(1)}(k) = x^{(1)}(t) \rightarrow 0.5\left[x^{(1)}(k) + x^{(1)}(k-1)\right]$

7 Using the least-squares method:

(a) Obtain the matrix B and data vector Y

(b) Use the inverse and matrix method $\hat{a} = \left(B^T B\right)^{-1} B^T Y$ to obtain the values of $\hat{a} = \left[a \quad b\right]^T$

8 Ascertain the model: $\hat{x}_j^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a}$ this gives us the simulation value of the model when $k \leq n$, our process $n = 5$.

Stage 3 Feasibility and qualification test:

After generation and modeling there is always a need to test a model for feasibility and qualification before it is used for prediction. This is important because there is bound to be error between the predicted and real values. Liu and Lin (2006) stated that; "only the models passing all the checks of different criteria can be used as prediction models". The error inspection methods that we will use for testing the model are the mean relative error and the post-sample error. For all the original raw data series $x^{(0)} = \left(x^{(0)}(k),\ k = 1,2,\ldots,n\right)$ we get the corresponding model simulated sequence solution series $\hat{x}^{(0)} = \left(x^{(0)}(k),\ldots k = 1,2,\ldots,n\right)$.

For the mean relative error method we will use the following steps:

1. Calculate the sequence of errors residuals $e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$
2. Calculate the sequence of relative errors at each point $\left|\frac{e(k)}{x^{(0)}(k)}\right|,\quad k = 1,2,\ldots,n$

   these

   are usually referred to as relative simulated error at the point k
3. Then calculate the mean relative simulation error $MRSE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{e(k)}{x^{(0)}(k)}\right|$

For any given α, when MRSE $<\alpha$ and $\left|\frac{e(n)}{x^{(0)}(n)}\right| < \alpha$ hold true, the model is said to be error satisfactory. For group of chosen values of α each level of simulation accuracy is determined. The commonly used levels of accuracy are given in Table 5.1 below and these will be used as a reference in this thesis.

Table 5.1 Critical Values of relative error

Critical values level	Relative error α
Level 1	0.01
Level 2	0.05
Level 3	0.10
Level 4	0.20

The post-sample error specific error is derived from the probability prediction method and its procedure is as follows:

1. Compute the standard error of elements the original series

$$S_1 = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(x^{(0)}(k) - \overline{x^{(0)}}\right)^2}$$

where $\overline{x^{(0)}} = \frac{1}{n}\sum_{k=1}^{n} x^{(0)}(k)$ is the average value of elements in original series.

2. Compute the standard error of elements of the residual series

$$S_2 = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(e(k) - \bar{e}\right)^2}$$

where $\bar{e} = \frac{1}{n}\sum_{k=1}^{n} e(k)$, is the average value of the residual series.

3. Compute the post-sample error specific value c as the ratio of the standard errors

$$c = \frac{S_2}{S_1}$$

4. Compute the specific value of the small error p

$$p = P\left\{\left|e(k) - \bar{e}\right| < 0.6745S_1\right\}$$

p is the calculated value as follows: If $k = 1$ and the equation $\left|e(k) - \bar{e}\right| < 0.6745S_1$ holds, then it is determined that there is a p value that pass the specific value of small error. After the completion of all calculations of k, we check all the specific values of the specific error and the passing rate of all p values. If p values pass the specific values, then p is denoted as $p = 1$ etc. The precision of a grey model is decided by c and p. Using model precision assessment (M.P), then

$$M.P = Max\{Rank\ of\ c, Rank\ of\ p\}$$

Table 5.2 shows the precision ranks. The smaller the rank number, the better the grey model is.

Table 5.2 Overall model precision ranks

Model precision	c	p
Rank 1 good	$c \le 0.35$	$0.95 \le p$
Rank 2 qualified	$0.35 < c \le 0.50$	$0.80 \le p < 0.95$
Rank 3 just	$0.50 < c \le 0.65$	$0.70 \le p < 0.80$
Rank 4 unqualified	$0.65 < c$	$p < 0.70$

Stage 4 Conducting the grey prediction operation:

1 Conduct the Inverse AGO for the series $\hat{x}(k+1)\ \ k=1,2,3,4,5$, $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$ and set $\hat{x}^{(0)}(6)=\hat{\Delta}_6$ as the predicted value, since in this case $k > n$, that is, 6 *is greater than where the n is* 5.

2 Solve for the predicted value of the sixth time or sample number $\hat{\Delta}_6$ where $\hat{\Delta}_6=\hat{x}^{(0)}(6)$. We get the 6th predicted value by $\hat{\Delta}_6=\hat{x}^{(0)}(6)=\hat{x}^{(1)}(6)-\hat{x}^{(1)}(5)$.

3 Conduct post sample checking to derive that the predicted value falls within the bandwidth of the control chart. If the predicted value falls within the in-control region then we go to the next step. If it is in the out-of-control region, then stop the manufacturing process to inspect the manufacturing process.

Stage 5 Eliminate the original value Δ_1 and add the value Δ_6 to a new series 2, $X_2=(\Delta_2,\Delta_3,\Delta_4,\Delta_5,\Delta_6)$. Continue to predict the value $\hat{\Delta}_7$ by using the series 2 data. Conduct the grey prediction operation as shown in step 2 for series 2 data. At this time, it also has five values in the data series. Δ_2 becomes the first value of series data and Δ_6

becomes the fifth value of the series data. The feasibility and qualification test of the model is conducted and then, the value $\hat{\Delta}_7$ can be predicted and.

Eliminate the value of Δ_2 in series 2 and add the value of Δ_7 to form a new series 3, $X_3 = (\Delta_3, \Delta_4, \Delta_5, \Delta_6, \Delta_7)$. Using the same procedure discussed above, continue to predict the value $\hat{\Delta}_8$ by using the series 3 data. Conduct the grey prediction operation and the feasibility and qualification tests series 3 data. Derive the predicted value of $\hat{\Delta}_8$. If the value of $\hat{\Delta}_8$ exceeds the control limits of the control chart, stop the manufacturing process and inspect it. We continue to carry out the same procedure as in the above-mentioned steps for the values in monitoring the process. If *yes*, then we stop the procedure, see Flow Chart of the Program at Appendix D.

In some production industries the time taken by the process for a product to be completed differs, to time between the consecutive products, the sampling interval in this process is given an unequal-gapped data series. In this case this first order one variable grey differential equation is abbreviated as $GM_n(1,1)$.The procedures of construction a grey predictive Shewhart control charts using unequal-gapped data series is the same as the one for using equal-gapped data series. For the $GM_n(1,1)$ model we ascertain the model

$$\hat{x}^{(0)}(t_k) = \left(x^{(0)}(t_1) - \frac{b}{a} \right)\left(1 - e^{a\Delta t_k}\right)e^{-a(t_k - t_1)}, \qquad k = 2,3,\ldots,n$$

At *Stage* 2 when constructing a grey prediction model we can still ascertain the model proposed by Dang et al (2002);

$$\hat{x}^{(1)}(n) = ce^{-an} + \frac{b}{a}$$

where n is the sample size and c is a constant. The constant c is obtained from the time response function of the grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b$$

given by

$$\hat{x}^{(1)}(k)=\left(x^{(0)}(n)-\frac{b}{a}\right)e^{-a(k-n)}+\frac{b}{a}$$

since

$$\hat{x}^{(1)}(t)=\underbrace{\left(x^{(0)}(n)-\frac{b}{a}\right)e^{-an}}_{c}e^{-at}+\frac{b}{a}$$

$$=ce^{-at}+\frac{b}{a}$$

is the general solution of the equation

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b$$

when substituting

$$x^{(1)}(t)\big|_{t=n}=x^{(1)}(n)$$

it is obtain that

$$\hat{x}^{(1)}(n)=ce^{-an}+\frac{b}{a}$$

making c subject to the formula it gives

$$c=\left(x^{(1)}(n)-\frac{b}{a}\right)e^{-an}$$

5.1.2 Grey Predictive Shewhart control chart using GM(2,1) model

When one-order accumulated generating operation (1-AGO) for the original series, that is $x^{(1)}(k)$ values are plotted against their k values, and the plot shows a non-monotonic process of change we used the GM(1,1) model. The procedures of construction a grey predictive Shewhart control charts using GM(2,1) model are the same as in using GM(1,1) model.

1 We get the first five values from the raw data series and denote them by $\Delta_1, \Delta_2, \Delta_3, \Delta_4 \text{ and } \Delta_5$.

2 Let $x^{(0)} = X_1 = (\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5)$

3 Derive the one-order accumulated generating operation (1-AGO) for the original series 1 $AGO\left\{x^{(0)}(k)\right\} = x^{(1)}(k) = \sum_{m=1}^{k} x^{(0)}(m), \qquad k = 1,2,,3,4,5$

4 Derive the one-time inverse-accumulated generating operation (1-IAGO) $\alpha^{(1)}x^{(0)}(k) = \left(\alpha^{(1)}\Delta_1, \alpha^{(1)}\Delta_2, \ldots \alpha^{(1)}\Delta_5\right)$

where $\alpha^{(1)}x^{(0)}(k) = \alpha^{(1)}\Delta_k - \alpha^{(1)}\Delta_{k-1}, \quad k = 1,2,\cdots,5$

5 Derive the average value $z^{(1)}(k) = x^{(1)}(t) \rightarrow 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$

6 Using the least-squares method:

a. Obtain the matrices B and data Y

b. Use the inverse and matrix method $\hat{a} = \left(B^T B\right)^{-1} B^T Y$ to obtain the values of $\hat{a} = \left[a_1 \quad a_2 \quad b\right]^T$

7 Having the solved for the values of $\hat{a}$ the solution of the GM(2,1) whitenization equation the following hold true. We consider a general solution of the homogeneous equation

$$\frac{d^2x^{(1)}}{dt} + a_1\frac{dx^{(1)}}{dt} + a_2x^{(1)} = 0$$

The roots of its characteristic equation

$$r^2 + a_1r + a_2 = 0$$

are solved by using a quadratic formula we get two values of r as

$$r_1, r_2 = -a_1 \pm \frac{1}{2}\sqrt{a_1^2 - 4a_2}$$

8 Ascertain the model, depending on the three cases of this general solution as follows:

(a) If $r_1 \neq r_2$ then $\bar{x}^{(1)} = C_1e^{r_1t} + C_2e^{r_2t}$

(b) If $a_1^2 - 4a_2 = 0$ then $\bar{x}^{(1)} = e^{r_2t}\left(C_1 + C_2t\right)$

(c) If $a_1^2 - 4a_2 < 0$ then $\bar{x}^{(1)} = e^{r_2t}\left(C_1\cos\beta t + C_2\sin\beta t\right)$.

We continue on with the last three stages: *Stage* 3 Feasibility and qualification test; *Stage* 4 Conducting the grey prediction operation, the predicted value is

$\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$; and *Stage* 5 Eliminate the original value Δ_1 and add the value Δ_6 to a new series 2, $X_2=(\Delta_2,\Delta_3,\Delta_4,\Delta_5,\Delta_6)$. These stages are carried on in a similar way when using a GM(1,1) model.

5.2 Transforming data sequence

The bandwidth of a grey predictive Shewhart control chart is rectangular and it is divided into two equal subsets by the center line C_s. This means that the $x^{(0)}=\{x^{(0)}(i)\}$ values that are in-control can belong to any of the two subsets of which some are between the center line and the upper control limit and some between the center line and the lower control limit. This implies that the data sequence can be transformed by subtracting the value of the center line from each of the values in the in-control bandwidth. This data sequence is transformed as $x^{(0)}=\{x^{(0)}(i),\ x^{(0)}(i)=X^{(0)}(i)-C_s,\ i=1,2,\cdots,n\}$. This decomposes the sequence $x^{(0)}=\{x^{(0)}(i)\}$ into the positive component and negative component. The positive and negative component is summarized as follows:

$$x^{(0)+}(i)=\begin{cases}x^{(0)}(i) & \text{if } 0<x^{(0)}(i)<U_s-C_s\\ 0 & \text{otherwise}\end{cases} \tag{5.1}$$

and

$$x^{(0)-}(i)=\begin{cases}-x^{(0)}(i) & \text{if } L_s-C_s<x^{(0)}(i)<0\\ 0 & \text{otherwise}\end{cases} \tag{5.2}$$

Denoting the upper under-control sequence by $X^{(0)+}=\{x^{(0)+}(i_k)>0,\ \ i_k\in\{1,2,\cdots,n\}\}$ and the lower under-control sequence by $X^{(0)-}=\{x^{(0)-}(j_k)>0,\ \ j_k\in\{1,2,\cdots,n\}\}$, two predictive curves can be established based on $X^{(0)+}$ and $X^{(0)-}$ respectively. These are as follows:

$$x^{(1)+}(k+1)=\left[x^{(0)+}(i_1)-\frac{b^+}{a^+}\right]e^{(-a^+k)}+\frac{b^+}{a^+} \tag{5.4}$$

and

$$x^{(1)-}(k+1)=\left[x^{(0)-}(j_1)-\frac{b^-}{a^-}\right]e^{(-a^-k)}+\frac{b^-}{a^-} \tag{5.5}$$

respectively. The matrix B and data vector Y are obtained separately from each in-control subset, B^+, Y^+ and B^-, Y^- for upper and lower in-control respectively. Then using the inverse and matrix method $\hat{a}^{\pm}=\left[\left(B^{\pm}\right)^T B^{\pm}\right]^{-1}\left(B^{\pm}\right)^T Y^{\pm}$ to obtain the values for $\hat{a}^+=\begin{bmatrix}a^+\\ b^+\end{bmatrix}$ and $\hat{a}^-=\begin{bmatrix}a^-\\ b^-\end{bmatrix}$. Then the decision rule for the grey predictive Shewhart control chart will be that the process is under-control if and only if $x^{(1)+}(n+1)-x^{(1)+}\left(i_{\max(k)}\right)<U_s-C_s$ or $x^{(1)-}(n+1)-x^{(1)-}\left(j_{\max(k)}\right)<C_s-L_s$.

The grey system *GM(1,1)* model can also be used to perform the out-of-control prediction if the sampled data sequence contains more than 4 out-control data. Denote the out-control sequence by $X^{(0)>}=\left\{x^{(0)>}(q(1)),x^{(0)>}(q(2)),\cdots,x^{(0)>}(q(K)),K\geq 4\right\}$ and correspondingly, the time sequence of out-control will be $Q^{(0)}=\{q(1),q(2),\cdots,q(K)\}$. Then the out-control times will be

$$\begin{cases}\hat{q}^{(1)}(k+1)=\left[q^{(0)}(1)-\dfrac{b_q}{a_q}\right]e^{(-a_qk)}+\dfrac{b_q}{a_q}\\ \hat{q}^{(0)}(k+1)=\hat{q}^{(1)}(k+1)-\hat{q}^{(1)}(k)\end{cases} \tag{5.5}$$

It is a well-known theoretical foundation that ARL is used for establishing the optimal design of Shewhart control charts as discussed in Chapter 1 Section 1.4. However, as a matter of facts, in small batch processing, ARL is not meaningful. Therefore, in terms of Equation (5.5) to describe the potentially next out-of control time point and take preventive action.

Finally, we have to point out that the partition of the sampled observation into the positive sequence that lies above the central line $X^{(0)+}=\left\{x^{(0)+}(i_k)>0,\ \ i_k\in\{1,2,\cdots,n\}\right\}$ and negative sequence that lies above the central line $X^{(0)-}=\left\{x^{(0)-}(j_k)>0,\ \ j_k\in\{1,2,\cdots,n\}\right\}$ is actually imposing information or sample size requirements and also keep unnecessary trace of Shewhart control charts. As we emphasized in Section 5.1, only two feasible tasks of a grey control chart can be achievable: pattern identification (in monotone or fluctuation) and the in-control (or out-of-control) grey prediction. Therefore, it is better to utilize all available the sampled process indices, of a minimum of 5 points, to establish grey differential equation models to facilitate the two tasks.

5.3 Grey predictive CUSUM control chart

CUSUM charts are mostly used and becoming more popular to among practitioners for monitoring quality control as compared to Shewhart control charts especially in manufacturing industries since they can detect small disturbances more rapidly on averages. In most writings on the application of CUSUM control the analysis of CUSUM charts properties is based on the implicit assumption that the process monitored will be operating continuously and indefinitely. These days most production processes are periodically set up to produce a specific quality over a specified time period like an of 8-hour shift. In cases like these, the limited duration of the production run has to be taken into account, in the designing the process control scheme so as to make the scheme to be maximally effective. It is not always that easy to effectively use CUSUM schemes in shorts runs, since there are limited number of samples or individual observations that can be taken to fit well with the accumulative character of the scheme. This has led us to propose a quality control chart of grey predictive CUMSUM control charts, since a grey system deals with very few sample data.

In constructing a grey predictive CUSUM control chart we will be using a similar manner as in developing a grey predictive Shewhart control chart. We will be using two-sided Tabular CUSUM control chart where we standardize the values raw data series. A standardized CUSUM control chart has been discussed in details in Chapter 2, Section

2.5. The procedure for constructing a grey predictive CUSUM control chart is defined below.

1. We calculate the average and the standard deviation of all the original raw data series in the process $x^{(0)}(k), k = 1,2,\cdots,n$.

$$\overline{x^{(0)}} = \mu_0 = \frac{1}{n}, \sum_{k=1}^{n} x^{(0)}(k) \qquad k = 1,2,\ldots,n$$

$$s = \sigma = \sqrt{\frac{1}{2(n-1)} \sum_{k=1}^{n-1} \left(x^{(0)}(k+1) - x^{(0)}(k) \right)^2}$$

2. Compute the reference value k and the control limit h

$$k = \frac{\delta}{2}\sigma = \frac{1}{2}|\mu_1 - \mu_0|$$

where $\delta = |\mu_1 - \mu_0|/\sigma$ and $\mu_1 = \mu_0 + \delta\sigma$

$h = 5\sigma$

3. Standardize all the value $x^{(0)}$ to $y^{(0)}$.

$$y^{(0)} = \frac{x^{(0)}(k) - \overline{x^{(0)}}}{s}$$

4. Compute the values of C_i^+ and C_i^- for detecting increases and decreases in the mean respectively as follows:

$$C_i^+ = \max\left[0, y_i^{(0)} - k + C_{i-1}^+\right]$$
$$C_i^- = \max\left[0, -k - y_i^{(0)} + C_{i-1}^-\right]$$

If the values of C_i^+ or C_i^- is greater than h then the process is said to be out-of-control and for the value of C_i^+ or C_i^- less than h the process is in-control. For in-control series the upper-sided and the lower-sided sequence are denoted as $C_u^{(0)} = \left\{ c_u^{(0)}\left(q(i)\right) = C_{q(i)}^+ < h, \;\; i = 1,2,\cdots,I > 3 \right\}$ and $C_l^{(0)} = \left\{ c_l^{(0)}\left(p(j)\right) = C_{p(j)}^- < h, \;\; j = 1,2,\cdots,J > 3 \right\}$ respectively. The upper and lower under-control sequences is denoted by $Q^{(0)} = \{q(1), q(2), \cdots, q(I)\}$ and $P^{(0)} = \{p(1), p(2), \cdots, p(J)\}$ respectively. Then in terms of

$$B=\begin{bmatrix} -x^{(0)}(2) & -z^{(1)}(2) & 1 \\ -x^{(0)}(3) & -z^{(1)}(3) & 1 \\ -x^{(0)}(4) & -z^{(1)}(4) & 1 \\ \vdots & \vdots & \vdots \\ -x^{(0)}(n) & -z^{(1)}(n) & 1 \end{bmatrix} \quad Y=\begin{bmatrix} a^{(1)}x^{(0)}(2) \\ a^{(1)}x^{(0)}(3) \\ a^{(1)}x^{(0)}(4) \\ \vdots \\ a^{(1)}x^{(0)}(n) \end{bmatrix}$$

Then the least-square estimator for the parameter vector $P=(a_1,a_2,b)^T$ is $P=\left(B^TB\right)^{-1}B^TY$

where $X^{(0)}$, $Z^{(1)}$, $X^{(1)}$, and $a^{(1)}X^{(0)}$ are defined as

$$X^{(0)}=\left\{x^{(0)}(i),\ i=1,2,\cdots,n\right\}$$

$$z^{(1)}(k)=\frac{1}{2}\left[x^{(1)}(k)+x^{(1)}(k-1)\right]$$

$$x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i)$$

$$k=2,\cdots,n$$

$$a^{(1)}x^{(0)}(k)=x^{(1)}(k)-x^{(1)}(k-1),\ k=2,3,\cdots,n$$

With these then the two under-control time predictive equations can be established $\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{\beta}{\alpha}\right]e^{-bk}+\frac{\beta}{\alpha}$. It is then a routine to check if the process is under-control via a CUSUM chart.

5.4 Fuzzy logic and control charts

There are a number of techniques that can be used to improve the reliability, sensitivity and grading of output of the process via control charts, like warning lines and fuzzy logic. In this section we will briefly discuss fuzzy logic and its application to control charts especially on "A Grey Fuzzy Predictive Control Chart".

Fuzzy logic is founded on the alteration of set theory by Professor Lofti Zedah in 1965, called fuzzy set theory, which enables an individual value to have a level of membership over a continuum of values. Fuzzy set theory can manage identify if a value belongs to a set and to what degree is it classified to in the set. We will consider a fuzzy set denoted by A and an element of set A is denoted by x. x is characterized by a membership function $f_A(x)$ which associates x as an element of A or not and x will be mapped onto the membership range between 0 and 1, $0 < f_A(x) \leq 1$. When x does not belong to A $f_A(x) = 0$.

Quality itself is a fuzzy phenomenon. What is a "good" quality? What is a "poor quality"? "good" or "poor" is a concept without clear boundary. The boundary is very vague. Therefore, fuzzy membership is used for quantification of vague boundary concept or event or perception. Pedrycz (1993) offered a details on fuzzy nature of quality concept. We will now give a brief discussion on application of fuzzy logic on control charts. Control charts are the most common technical tools used in statistical process control for quality control and improvement. Control charts plot the measurements of a quality characteristic taken from the process versus time or the sample number. Shewhart, CUSUM and EWMA control charts as discussed in previous Chapters 1, 2 and 3 detects if the process is in-control or out-of-control and also detects if there is any shift in the mean or variance. Their control limits provides a crisp description between good and bad quality of the process monitored. As long as the values of x falls within the bandwidth defined by the control limits, the quality of the process can be considered good. Otherwise the quality is characterized as bad. However, the transition from good to bad quality is often continuous and gradual. In order to be able to represent a gradual change in the quality of a process we use a fuzzy set "in-control". In control charts we have the three basic values of $L_s, C_s, U_s \in R$ with $L_s \leq C_s \leq U_s$ which forms an in-control set. The triangular fuzzy number defined by L_s, C_s *and* U_s, denoted by $f_{(L_s, C_s, U_s)}(x)$ would be transformed to a fuzzy subset of denoted by R and it bounded by the upper and lower control limits $L_s \leq R \leq U_s$, and its membership function over the

domain of the monitored characteristic of x is denoted by $f_R(x)$. When an output value during the process is not within the control limits it would be given the value 0 and when it is within the control limit it will be given the value 1.

For every output value that falls within the control limit we calculate its degree of membership using a triangular membership function:

$$f_R(x) = \begin{cases} 0, & if \quad x < L_s \\ \dfrac{x - L_s}{0.5(U_s - L_s)}, & if \quad L_s \le x \le C_s \\ 1, & if \quad x = C_s \\ \dfrac{x - L_s}{0.5(U_s - L_s)}, & if \quad C_s \prec x \le U_s \\ 0, & if \quad x > U_s \end{cases} \tag{5.6}$$

Figure 5.1 below shows that the further the actual value of x deviates from the centerline, the smaller its degree of membership to the fuzzy set "in-control".

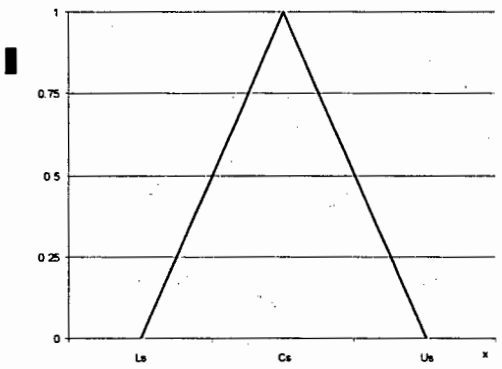


Figure 5.1 Fuzzy set "in-control" defined for characteristic x.

Fuzzy control charts are some of the improvements towards statistical quality control charts due to their ability to recognize gradual changes in a system's state. Also in fuzzy

control charts the degree of membership of the in-control quality of x makes it easy in ranking or classifying products, this is discussed in the coming section.

5.5 Grey- fuzzy predictive control chart

In this section we will discuss the application of the fuzzy logic in a grey predictive control chart, especially using a reasonable membership function for grading in-control quality products, referred to as membership grades.

During the production process the variability of the output occurs. This variability is usually measured in terms of the process standard deviation as one of the parameters of the process. This leads to capability studies which are conducted for the purpose of estimating the process parameters and the information got from these estimations being used to determine the ability of the process to meet standards and specifications. These standards can be used as a basis for constructing control charts. A standard is a basis for comparison of a reference point against which measurement of the quality can be evaluated. Standards can also be established in a variety of ways; but they should be based upon the capability of the process itself in order for them to be effective in the control of process output. Standards of the process can also be based upon the requirements of a customer, like specified purchase order or a contract. With the demand of goods of good quality the manufacturer's products should conform to the specifications and control limits.

In using a grey predictive control charts the values should fall within the bandwidth, and they should also be of high conformance. The values could be treated not just as individual values but in set-valued functions, in particular functions whose values are intervals. These values of X are coverage of the specification-tolerance interval. The specification limits of a product are the lower and upper limits, $[L,U]$. Then designing the specification-tolerance requirements, a membership function of high-conformance $\tilde{Q}$ is given by

$$\mu_{\tilde{Q}}(x)=\begin{cases} l(x) & if \quad \dfrac{U-L}{6}\le x\le \dfrac{U-L}{2} \\ r(x) & if \quad \dfrac{U-L}{2}\le x\le \dfrac{5(U-L)}{6} \\ 0 & otherwise \end{cases} \tag{5.7}$$

where $0\le l(x)\le 1$ is a monotone-increasing function and $0\le r(x)\le 1$ a monotone-decreasing function with $l\left(\dfrac{U-L}{2}\right)=r\left(\dfrac{U-L}{2}\right)=1$. It is should be noticed that the measurement x here is the shifted one, that is, $x=X-L$ where X is the actual observation value.

Let a processed item have the following measurement sequence by equal time sampling $X^{(0)}=\left\{X^{(0)}(i),\ i=1,2,\cdots,n\right\}$, where $n\ge 4$. Then according to the membership function $\mu_{\tilde{Q}}$ of high-conformance, the measurement sample can be converted into membership grade sequence $\mu_{\tilde{Q}}\left(x^{(0)}\right)=\left\{\mu_{\tilde{Q}}\left(x^{(0)}(i)\right),x^{(0)}(i)=X^{(0)}(i)-L,\ i=1,2,\cdots,n\right\}$.

Taking into consideration the estimates parameters (α,β) and the differential equation theory in the prediction equation discussed in Chapter 4 sub-section 4.3.3, the predictive equation could be established via a membership grade sequence:

$$\hat{\mu}_{\tilde{Q}}\left(x^{(0)}(i)\right)=\hat{\alpha}+\hat{\gamma}\exp\left(-\hat{\beta}\ i\right),\ i=1,2,\cdots,n \tag{5.8}$$

Then according to the grey predictive $GM(1,1)$ model, the predictive value of $\hat{\mu}_{\tilde{Q}}\left(x^{(0)}(n+1)\right)$ at time $n+1$ can be calculated, and the quality can be classified using a grey-fuzzy predictive control scheme below proposed by Guo and Dunne (2005).

$$\begin{cases} \text{perfect} & \text{if} \quad 0.75<\hat{\mu}_{\tilde{Q}}\left(x^{(0)}(n+1)\right)\le 1 \\ \text{under control} & \text{if} \quad 0.50<\hat{\mu}_{\tilde{Q}}\left(x^{(0)}(n+1)\right)\le .75 \\ \text{questionable} & \text{if} \quad 0.25<\hat{\mu}_{\tilde{Q}}\left(x^{(0)}(n+1)\right)\le 0.50 \\ \text{out of control} & \text{if} \quad \hat{\mu}_{\tilde{Q}}\left(x^{(0)}(n+1)\right)\le 0.25 \end{cases} \tag{5.9}$$

They proposed this scheme basing on the following facts:

- Quality of conformance itself is a fuzzy event and therefore using a membership to describe its behavior may be very appropriate;
- GM(1,1) model possesses high predictive power, according to the working experiences, the predictive values will not have relative error larger than 5% as long as the predictive time range is less than $n/3$ and therefore, one-step prediction is in general very safe to tell the quality conformance status of the real process
- *GM(1,1)* modeling on membership grade sequence $\mu_{\tilde{Q}}(x^{(0)}) = \{\mu_{\tilde{Q}}(x^{(0)}(i)), x^{(0)}(i) = X^{(0)}(i) - L,\ i = 1,2,\cdots,n\}$ avoids the shortage of data and increases the model efficiency because the grade values are all positive.

5.6 Summary

In a grey predictive Shewhart control chart or grey predictive CUSUM control chart the predicted values should fall within a uniform linear band of upper and lower control limits. A prediction model used is always tested for feasibility and accuracy this minimizes the error of prediction. A fuzzy logic can still be used in grey predictive control charts. This improves the reliability, sensitivity and grading of the output products other than just using the warning lines or sensitivity zones.

Chapter 6 Industrial Application

In this chapter we implement the grey models on industrial data from TBF Packaging Machine in Taiwan. We take into consideration that the company is on operation seven days in a week. We have ten machines given the k values of 1 to 10. The recorded values are the failure times of a machine in operations. The commonly accepted practices for quality control methods in production companies is to use control limits and take action by the time it is found that the product is just about to exceed or has exceeded the control limits. In packaging industries using this practice is not very appropriate, because by the time the packing machine fails a number of products may have already been produced. In such case, the quality control goal of real-time monitoring may not be really satisfied especially if the goods are produced are perishable. This has led us in this thesis used the grey predictive model so as to predict the next failure time of the machine based on the previous failure times of the machine. This predicted failure time will be relayed to management department in time for necessary actions to be taken. This will reduce the packing failure time and increase the production quality.

6.1 The monitoring process

In many application areas of process monitoring for quality control it has become increasingly important to monitor the behavior of the systems based on multiple measurements. In this study we are monitoring the packaging process detecting gradual and abstract failures in the machines, where failure represents the typical state of the machines which leads to the quality of a product.

In order to detect changes in the failure pattern for the machines we categorized the failure times in five ways of output quality measurements representative index. The observed values of these indices are presented in Table C.4, see Appendix C. These output quality measurements representative indices are used to measure the quality of the products in the manner defined below:

1. *Mean time between failures (MTBF)*: This is the average time between machine failures. The more the machine's mean time between failures increases the poor the quality of the output and vice versa.
2. *Maximal functioning times*: This is the maximum or longest time the machine operates well before it fails or stops to operate. This can also be related to reliability, that is, a measure that indicates time between failures occurring in a system. When the functioning time of approaches the maximal functioning time the higher the quality of the product, since the machine is reliable.
3. *Signal-noise ratio*: computed as

$$\text{Signal-noise ratio} = \frac{\text{Mean time between failures}}{\text{Standard Deviation}} = \frac{\bar{x}_i}{s_i} \qquad i = 1, 2, \ldots, 10$$

4. *Number of failures*: This is the number of failures the machine had since it started operating up to the last recording date. The more the number of failures the poor the quality of the product and vice versa.
5. *Average failure rate*: This is the average rate at which the machine has failed in its operation since it started operating up to the last recording date. This is can be referred to as in-availability, that is, a measure of the percentage of time a system or component does not performs its specified function. The higher the machine's failure rate increases the poorer the quality of the output. This is computed as

$$\text{Average failure rate} = \frac{\text{No of failures}}{\text{total number of days on operation}}.$$

Total number of days on operation for each machine is given in Table 6.1

Table 6.1 Total number of days on operation for a machine

Machine No.	Date and time the failure occurred		Number of days in operation
	First	Last	
1	8/3/2005 10:30	4/4/2006 8:05	243.90
2	8/3/2005 10:20	4/27/2006 9:00	266.94

3	8/3/2005 10:10	2/25/2006 15:40	206.23
4	8/3/2005 10:10	4/29/2006 8:30	268.93
5	8/4/2005 8:00	5/4/2006 13:20	273.22
6	8/3/2005 9:50	5/18/2006 10:50	288.04
7	8/4/2005 11:25	4/16/2006 9:00	254.90
8	8/5/2005 15:40	4/16/2006 9:30	253.74
9	8/3/2005 8:30	5/4/2006 15:00	274.27
10	8/3/2005 10:30	5/4/2006 15:10	274.19

After we categorized the data we check the spread of the data for each quality measurements representative index. This is done by calculating the average squared deviation around the mean, see Table 6.2.

Table 6.2 Quality Measurement Representative Indices value's deviations around the mean

Standard deviation	Quality Measurement Representative Index				
	Mean time between failures	Maximal functioning times	Signal-noise ratio	Number of failures	Average failure rate
σ	87.897*	723.301*	0.185	10.286*	0.038

The standard deviations of for the "mean time between failures", "maximal functioning times" and "number of failures" data sets are too high this means that there is a very large variability between larger numbers and smaller numbers within these indices data sets. This variability has to be minimized so as to get good predicted values and make correct conclusions after processing the data. To reduces this variability of a number while still maintaining the original order of the data we transformed these data sets by using a natural logarithm, that is, logarithm to the base e, $(e = 2.7183)$. For the transformed data's standard deviation of mean time between failures", "maximal functioning times" and "number of failures" are of 0.347, 0.529 and 0.37, see Table 6.3.

6.2 Empirical modeling

Each quality measurement representative index is treated as a different data set and the quality control limits for control charts are computed for each data set. Then the first five elements in each data set are used to build a grey system model at each step. We check the linearity of $x(k)$ with relation to k by using R-square for each data set. If on the original data set $x^{(0)}$ there is no linearity we will check linearity in on $x^{(1)}$. If there is linearity in on $x^{(1)}$ we build the model on $x^{(1)}$. If there is no linearity in $x^{(1)}$ we will stop trying higher order AGO and seek other data processing methods.

The grey model built is checked for prediction. This is done by computing the mean relative simulation error and post-sample error specific error for each model of each data set. Checking the model is emphasized by Kocijan, J et al [29] as they say that "the quality of control depends on the quality of the model". We also use the rolling method to make two steps prediction and also check the prediction. The predicted values will be checked if they fall within the control limits.

The program for the above statements is written in VBA and ran in MS Excel, see Appendix D.

6.2.1 Control limits for a grey predictive Shewhart control chart

The control limits of a grey predictive Shewhart control chart model for each data set are computed using the steps in Subsection 5.1.1 in Chapter 5. The standard deviation is estimated using the standard deviation method. The results are in Table 6.3. These control limits results would be used to interpret the results of the predicted values of the grey prediction model in whether the value is of the recommended quality or not, and even when the predicted value is of the recommended quality we would check if there any step or steps that should be taken for the process. This would be tested by using zone rules.

Table 6.3 Grey Predictive Shewhart control chart Limits

Control Limits		Mean time between failures	Maximal functioning times	Signal-noise ratio	Number of failures	Average failure rate
C_s		5.486	7.112	0.8192	3.224	0.1058
U_s		6.527	8.700	1.376	4.334	0.220
L_s		4.445	5.524	0.263	2.115	-0.009
Warning Lines	$C_s+2\sigma$	6.180	8.171	1.190	3.964	0.182
	$C_s-2\sigma$	4.792	6.054	0.448	2.485	0.029
	$C_s+\sigma$	5.833	7.642	1.005	3.594	0.144
	$C_s-\sigma$	5.139	6.583	0.634	2.855	0.068
σ		0.347	0.529	0.185	0.370	0.038

6.2.2 Test of regression

The linearity of $x(k)$ with relation to k is checked using regression analysis. In this process we use the coefficient of multiple determination, commonly referred to as R-square $\left(R^2\right)$. The larger value of R-square, the more useful is the regression model for predicting new observations, but Colton J.A at al [10] advises researchers not select a model based solely on the high value of R^2, because this is true when many terms are included in a model to fit a relatively small number of observations. For the case when no

repeat runs are used, R^2 can reach the value of 1, that is 100%, when we include the same number of terms as the number of dependent observations. "In this situation, the practitioner will be modeling the error, in addition to any deterministic relationships that may exist" Colton J.A at al [10].This made us to also use the F-test in the analysis of variance (ANOVA) at 5% level of significance. These tests is done for each data set at three steps, using the first five pairs of elements, see Table 6.4. Steps 1, 2 and 3 are used for predicting the sixth, seventh and eighth value. These are all predicted using a rolling method.

Table 6.4 Linearity Check Results

Step 1	$x^{(0)}$			$x^{(1)}$			$x^{(2)}$		
	R^2	F	P	R^2	F	P	R^2	F	P
Mean time between failures	0.1002	0.3344	0.6036	0.9995	5927.82	0.000005	0.9724	105.84	0.002
Maximal functioning times	0.0618	0.198	0.687	0.9992	3912.23	0.000009	0.9725	106.16	0.002
Signal-noise ratio	0.0850	0.2786	0.6342	0.9935	462.03	0.000220	0.9726	106.34	0.002
Number of failures	0.0561	0.1784	0.7012	0.9984	1918.43	0.000026	0.9713	103.02	0.0021
Average failure rate	0.0814	0.2660	0.6417	0.9814	187.22	0.000845	0.9705	98.734	0.0022
Step 2	$x^{(0)}$			$x^{(1)}$			$x^{(2)}$		
	R^2	F	P	R^2	F	P	R^2	F	P
Mean time between failures	0.002	0.0005	0.9830	0.9995	6024.21	0.0000047	0.9718	103.44	0.0021
Maximal functioning times	0.0136	0.0414	0.8519	0.9992	3660.91	0.0000099	0.9713	101.60	0.0021
Signal-noise ratio	0.0704	0.2272	0.6662	0.9925	399.06	0.000274	0.9757	120.63	0.0017
Number of failures	0.0697	0.2246	0.6679	0.9988	2594.16	0.000017	0.9707	99.32	0.0022
Average failure rate	0.0038	0.0114	0.9218	0.9860	211.95	0.000703	0.9715	102.32	0.0021
Step 3	$x^{(0)}$			$x^{(1)}$			$x^{(2)}$		
	R^2	F	P	R^2	F	P	R^2	F	P
Mean time between failures	0.1613	0.5771	0.5027	0.9996	7251.61	0.000004	0.9704	98.188	0.0022
Maximal functioning times	0.0258	0.0794	0.7964	0.9993	4058.12	0.000009	0.9705	98.75	0.0022
Signal-noise ratio	0.0018	0.0054	0.9461	0.9924	389.42	0.000284	0.9753	118.39	0.0017
Number of failures	0.0182	0.0556	0.8288	0.9936	2172.44	0.000022	0.9739	112.06	0.0018
Average failure rate	0.1722	0.6242	0.4872	0.9814	158.607	0.001079	0.9820	163.5351	0.00103

The factors used in to check the linearity of $x(k)$ with relation to k are mathematically calculated as:

$$R^2 = \frac{SST - SSE}{SST}$$

where SST is the total sum of squares in the response about the mean, and SSE is the sum of squares in the response about the regression line;

$$F = \frac{MSS}{MSSE}$$

where MSS is the mean sum of squares in the response about the mean, and MSSE is the mean sum of squares error.

At 1-AGO the computed values of R-square are greater than 0.90 $\left(R^2 > 0.90\right)$, which indicates that in the model we will be using for prediction over 90% of the variability in the predicted values is accounted for by the model. Also at 1-AGO the computed P-values in the F-tests are less than 0.05, it implies that the linearity has improved considerably at 1-AGO and worse at 2-AGO for all the quality measurement representative indices so the AGO should be stopped and we can proceed on building grey predictive models.

6.3 Implementation of a GM(1,1) on grey predictive Shewhart control chart

We will take five consecutive values from the raw data and build a GM(1,1) on each quality measurement representative index data set. We will use these models to predict values and take two steps rolling method prediction. Thus we build three models for each data set. The solution for the GM(1,1) for each data set at each step would be derived from the IAGO formula

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a}\right)e^{-ak} + \frac{u}{a}$$

and using IAGO formula to get the predicted value $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) \quad k \geq 2$.

After predicting the values we would calculate their residual errors so as to get their degree of accuracy.

6.3.1 GM(1,1) Modeling on the output quality measurements representative indices

Having checked the smoothness and we now process the raw data series for each index and results are in Table 6.5.

Table 6.5 Results from GM(1,1) on each of output quality measurements representative index

	Machine No.	Original data	a	u	Simulated data	Errors	Relative error	MRSE
Mean time between failures	1	5.239	-0.0028931	5.500796899				0.0520789 4
	2	5.747			5.523940666	-0.223	-3.881%	
	3	5.444			5.539945111	0.096	1.762%	
	4	5.082			5.555995933	0.474	9.327%	
	5	5.919			5.572093266	-0.347	-5.861%	
	6	5.496			5.588237222	0.092	1.678%	
	7	5.622	-0.018056977	5.185623816	5.737217178	0.115	2.049%	0.0448189 2
	8	5.193	-0.021099487	5.184691625	5.827508311	0.635	12.219%	0.0402879 5
Maximal functioning	1	7.006	-0.01657	6.78834				0.060493
	2	7.202			6.961994	0.24001	3.33%	
	3	7.150			7.078345	0.07166	1.00%	
	4	6.360			7.196641	-0.8366	-13.15%	
	5	7.843			7.316914	0.52609	6.71%	
	6	7.050			7.439197	-0.3892	-5.52%	

	7	7.072	-0.01665	6.746825	7.401116	-0.3291	-4.65%	0.064572
	8	6.616	-0.0182	6.696319	7.409362	-0.7934	-11.99%	0.057719
Signal-noise ratio	1	0.833	0.068292476	0.989306791				0.180594
	2	0.867			0.901293	0.034	3.955%	
	3	0.764			0.841796	0.078	10.183%	
	4	1.073			0.786227	-0.287	-26.726%	
	5	0.559			0.734326	0.175	31.364%	
	6	0.814			0.685851	-0.128	-15.743%	
	7	0.919	0.044569667	0.914704137	0.716914377	-0.202	-21.990%	0.209945
	8	0.901	0.028619032	0.912006979	0.782606122	-0.118	-13.140%	0.212526
Number of failures	1	3.526	-0.01206	3.031169				0.083317
	2	2.996			3.092279	-0.0963	-3.21%	
	3	2.996			3.129784	-0.1338	-4.47%	
	4	3.714			3.167743	0.54626	14.71%	
	5	2.890			3.206163	-0.3162	-10.94%	
	6	3.332			3.245049	0.08695	2.61%	
	7	3.045	-0.00559	3.180201	3.278458	-0.2335	-7.67%	0.086344
	8	3.638	0.049854	3.727944	2.859728	0.77827	21.39%	0.068459
Average failure rate	1	0.139	-0.02317	0.089844				0.285592
	2	0.075			0.09415	-0.0192	-25.53%	
	3	0.097			0.096357	0.00064	0.66%	
	4	0.152			0.098615	0.05339	35.12%	
	5	0.066			0.100926	-0.0349	-52.92%	
	6	0.097			0.103291	-0.0063	-6.49%	
	7	0.082	0.079068	0.126068	0.084192	-0.0022	-2.67%	0.260869
	8	0.150	0.208439	0.165508	0.056972	0.09303	62.02%	0.250869

The mean relative simulation error (MRSE) for each measurement representative index data set is satisfactory at each step for prediction to be made. Having fitted the coefficients a and u in the function predicted machine failure time values we would now verify the accuracy of fitting the values by the GM(1,1) model. This done by calculating the filtering accuracy of each predicted value in each data set and also calculating the average relative accuracy of a data set. These are calculated as:

Filtering accuracy is $1-\left|\frac{e(k)}{x^{(0)}(k)}\right|$ (the k^{th} predicted value) where $e(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$

And the average relative is $1-\frac{1}{n}\left[\sum_{k=1}^{n}\left|\frac{e(k)}{x^{(0)}(k)}\right|\right]$. The results are summarized in table form below.

Table 6.6 Statistics of Relative accuracy of fitting the GM(1,1) model

Quality Measurement Representative Index	Model fitting error			Accuracy
	Minimum	Maximum	Average	
Mean time between failures	1.68%	12.22%	5.25%	94.75%
Maximal functioning times	1.00%	13.15%	6.62%	93.38%
Signal-noise ratio	3.96%	31.36%	17.59%	82.41%
No of failures	2.61%	21.39%	9.29%	90.71%
Average failure rate	0.66%	62.02%	26.49%	73.51%

Quality measurement representative indices have prediction with precision accuracy that is accurate even though some have the precision less than 90% but they are all above 70% correct in predicting the failure times. The indices with average relative error less than 10% are the ones with more prediction accuracy since their predicted values will be at least 90% correct.

6.3.2 Are the predicted values in control

In the first three chapters of this thesis we discussed that the randomness of the variations can be best illustrated by presenting the results or out-put of a process in the form of control charts. The packaging machines also have randomness in the output values even when the machine is in operation good condition, because there are always random or common causes during any production process. There might also be time when the process is out of control due to special causes of variation, these leads to a process to be stopped for these causes to be eliminated. If the process is in control there is always need to be ready to fix it when it gets out of control.

Packaging machine failure leads to loss in production and profit to the company. This has made it a point for us to asses the machine failure time using control charts. With control charts we set the limits to detect whether the machine is in good conditions or not, even when it is control we calculated warning lines of all the five data sets, used for warning signs.

For the management to minimized the failure time and rate of a machines they should have a rough idea of when the machine might fail and produce out of control products. This makes us to used a grey predictive control chart, where we use the past data to set the control limits for the values to be predict and asses the state of their control. Figure 6.1 present the grey predictive Shewhart control charts.

In interpreting Figure 6.1 grey predictive Shewhart control charts we will focus much on the predicted values in each data set especially the last three values. For all the quality measurement representative indices, all the original and predicted values fall within the control limits, see Figure 6.1 (a) to (e). This compiles us to make our analysis basing on zone tests, that is, stating zone rules violated for each figure. We would use Western Electric Alarm Rules and Alarm Rules due to the fact that we used five individual values to make three predictions per data set so we have a total of eight pairs of predicted values, this is less than nine in number so we would not use the Nelson's Alarm Rule.

Figure 6.1 Movement of original and predicted GM(1,1) machine failure time

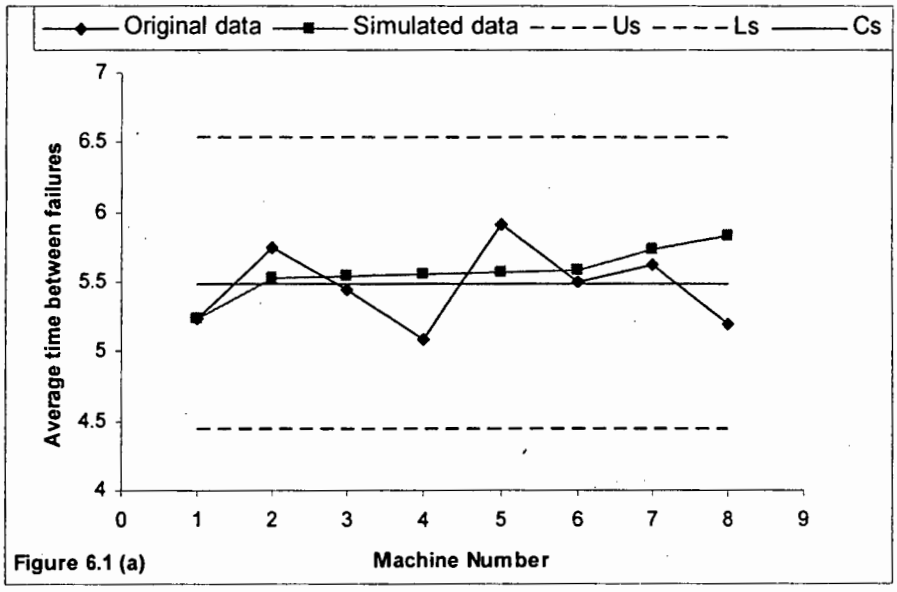


Figure 6.1 (a)

There is a long run of predicted seven consecutive points on the upper side of the center line. Even though six of these points are in Zone C and only one is in Zone B this provides an evidence that the process mean or variability has shifted from the centerline. This is also shown by the last three consecutive predicted values are in increasing order.

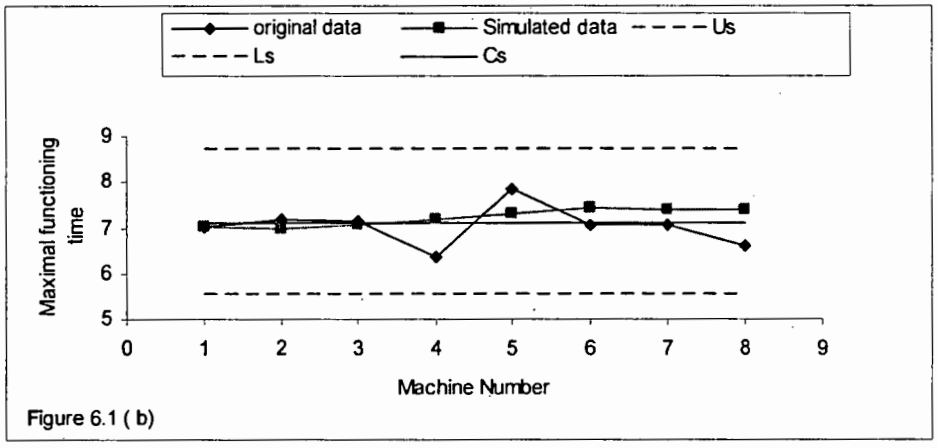


Figure 6.1 (b)

All the predicted points are in Zone C but they are too close to the center line. There is no much problem of Zone Violation.

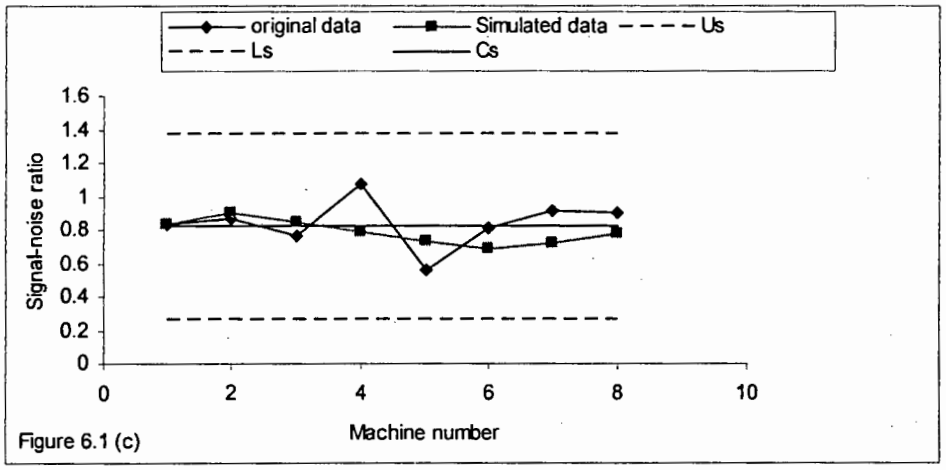


Figure 6.1 (c)

All predicted values are in Zone C. No zone violations.

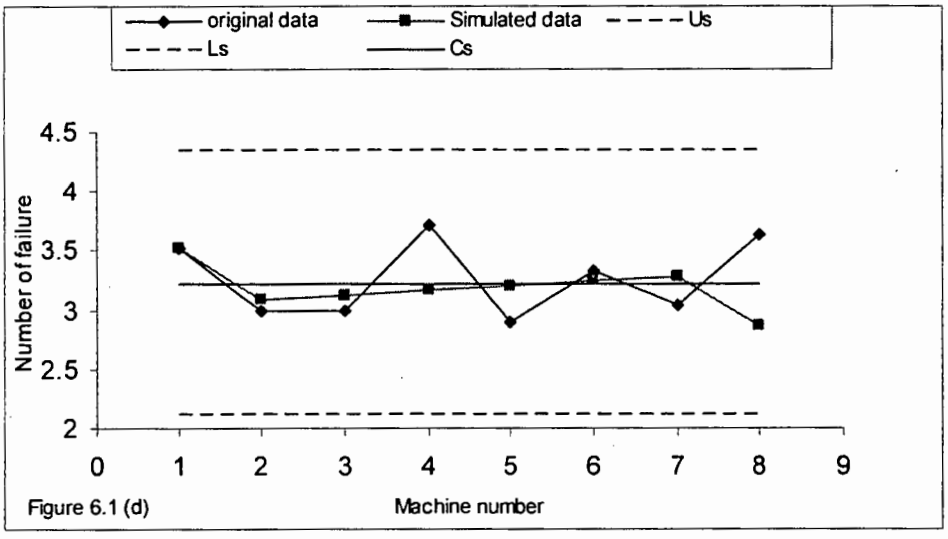


Figure 6.1 (d)

The eighth predicted value show a large swing even thought it is generally centered.

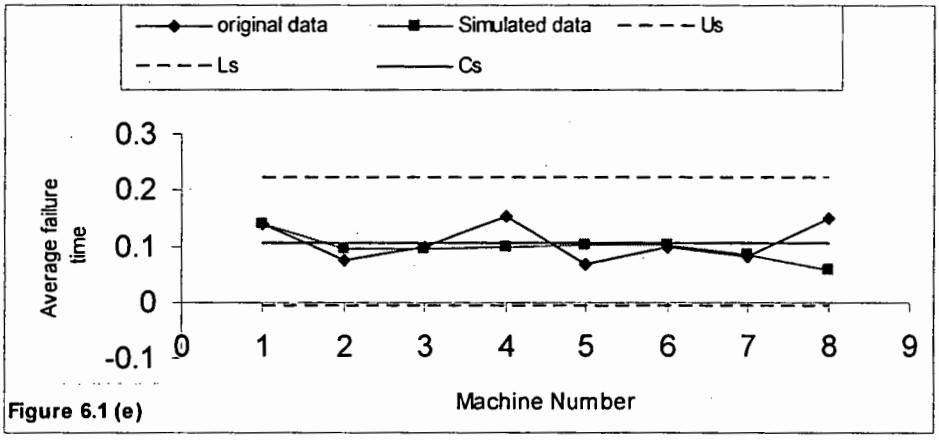


Figure 6.1 (e)

There is a run of predicted seven consecutive points on the lower side of the center line.

6.4 Implementation of a GM(1,1) grey prediction CUSUM control chart

We believe that we used the proper sampling technique so there is no biasness in the results we got in the previous section even though almost all the predicted values fall within the same zone. We will now use the CUSUM trend line identification rules to detect if there was any shift in the mean or variability in the predicted values. This will be done using a grey predictive CUSUM charts for each data set. However it should be noted that the use of CUSUM charts does not require other control as a prerequisite, so we will not be necessarily comparing the grey predictive CUSUM control charts with the grey predictive Shewhart control charts.

Having a sample of 10 machines we will use the first nine original data values to predict the tenth value in construction the GM(1,1) model of CUSUM charts. We have to test the linearity of $x(k)$ with relation to k using regression analysis. The results are in Table 6.7 below:

Table 6.7 Linearity Check for CUSUM Values Prediction Models

	$x^{(0)}$			$x^{(1)}$			$x^{(2)}$		
	R^2	F	P	R^2	F	P	R^2	F	P
Mean time	0.09013	0.0573	0.8176	0.99994	59591.4	5E-15	0.9792	162.992	4.19E-06

between failures									
Maximal functioning times	0.286062	0.6239	0.4555	0.99986	25073.02	1.06E-13	0.9796	166.1718	3.93E-06
Signal-noise ratio	0.3832	1.2049	0.3087	0.9983	2029.824	6.93E-10	0.9789	160.7763	4.39E-06
Number of failures	0.0144	0.0014	0.9707	0.9998	16698.81	4.38E-13	0.9794	164.3113	4.08E-06
Average failure rate	0.13001	0.1203	0.7389	0.9976	1446.147	2.26E-09	0.9809	177.8144	3.12E-06

Using R-square and F-test at $\alpha = 0.05$ criterion we conclude that there is good linearity has improved better at 1-AGO than at 2-AGO for all the quality measurement representative indices. At 1-AGO, that is, at $x^{(1)}$ $R^2 > 0.98$ which shows that the models new would be using to predict will predict almost 100% of the variability in the predicted values accounted for by the model, and the P (probability) values for $F_{0.05,1,7}$ are all less than 0.05, so we can proceed on building grey GM(1,1) predictive models.

Table 6.8 Results from GM(1,1) for Grey Predictive CUSUM Control Chart

Quality Measurement Representative Index	Original data	a	u	Simulated data	Relative error	MRSE	Accuracy
Mean time between failures	5.787	0.002401581	5.57820304	5.453161	0.057688	0.040952	95.90%
Maximal functioning times	6.815	0.008821	7.303146	6.718273	0.985807	0.144157	85.58%
Signal-noise ratio	1.181	-0.03681	0.727153	1.036264	0.122554	0.151409	84.86%
No of failures	2.944	-0.00842	3.076006	3.336256	0.133239	0.082843	91.72%
Average failure rate	0.069	-0.01378	0.09256	0.106214	0.539328	0.285779	71.42%

All quality measurement representative indices have prediction with precision of good accuracy. Average failure rate still has an accuracy of less than 80%, this would be looked at critically when analyzing the mean shift from the results of CUSUM control charts.

6.4.1 Plotting a Grey Prediction CUSUM Control Chart

We have a record of 10 predicted values for each of the data sets we are dealing with. We would the CUSUM trend line identification rules in analyzing the shifts because our data is small and the reference chart of these rules goes up to number points on the line of $n = 8$. Any trend line with nine or more points in a row could still be considered a signal even when the slope may be extremely small.

We would use the following procedure in plotting grey predictive CUSUM control charts:

- Let a series of the predicted values be $x_1, x_2, \ldots, x_{10}$ with an unknown mean μ_x and standard deviation σ_x, results in Table 6.9.
- We use the average of the original data sets of nine the first 9 values as the desirable or target value μ_0 as m, which may differs from μ_0 by δm.
- At each prediction time x_t is a predicted value of a process with mean that is different from the reference value m by $(\delta m + \delta\mu_0)$.
- A CUSUM C_t is calculated as $C_t = \sum_{i=1}^{t}(x_i - m)$, its expectations depends on $(\delta m + \delta\mu_0)$.
- If $(\delta m + \delta\mu_0) = 0$ for all predicted values through t, the expected value of C_t will be zero. Under this condition the value of t CUSUMs will fluctuate randomly about a zero line on a CUSUM plot without any significant trend lines.
- If $(\delta m + \delta\mu_0) \neq 0$ at least some of the predicted values, the CUSUM data will not vary about zero line and may show some trend line.

- We standardize C_t and the quantity $(\delta m + \delta\mu_0)$, this is denoted by $\delta = \frac{\delta m + \delta\mu_0}{\sigma}$, where σ is the standard deviation of the original data series.
- We use the rules in Table 6.10 for the trend line identification and make analyze the grey predictive CUSUM control chart results.

Table 6.9 Mean and Standard Deviation for Original Data Sets

Statistical Measure	Quality Measurement Representative Index				
	Mean time between failures	Maximal functioning times	Signal-noise ratio	Number of failures	Average failure rate
μ_0	5.482111111	6.993333	0.876444	3.247556	0.104222
σ	0.273456781	0.428254	0.173065	0.311172	0.033704

Table 6.10 CUSUM Trend Line Identification Rules

• Number of points on the line, n n	• Shift Is Greater than • $\delta \geq$	• Size of type I Error • α
• 8	• 0.25	• <0.0017
• 7	• 0.40	• <0.0017
• 6	• 0.60	• <0.0016
• 5	• 0.90	• <0.0012
• 4	• 1.25	• <0.0012
• 3	• 1.75	• <0.0016
• 2	• 3.00	• <0.0014

Opted from DeVor R.E et al (1992, p 410)

Figure 6.2 Grey Prediction CUSUM Control Chart

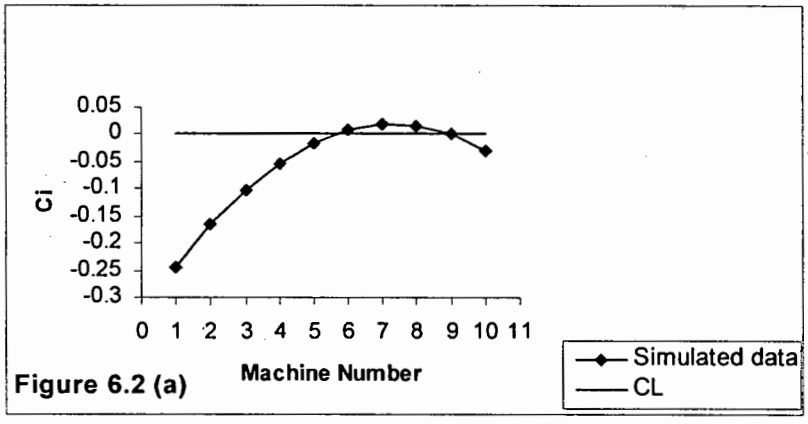


Figure 6.2 (a)

Figure 6.2 (a) indicates that the mean time between failures the signal that there was a shift in the mean of the predicted values by at least 1.25 standard deviations towards the target value m, $m = 5.482$. This brought the process to the be at a more steady state from the forth value to the eighth.

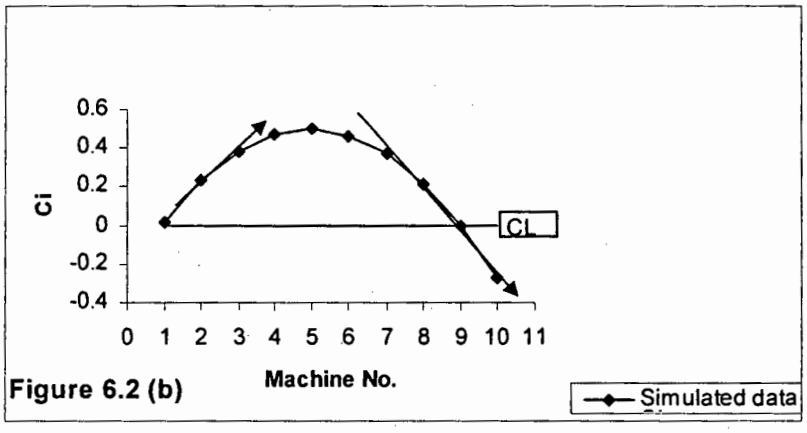


Figure 6.2 (b)

Figure 6.2 (b) is for maximal functioning times. From the second to the third predicted values of machines numbered 2 and 3 respectively there is an indication that there was at least a 3 standard deviations shifts from the reference value 6.993 with a risk of 0.14%. This shows a very steep slope. From the 7th to the 10th predicted values there was a 1.75 standard deviations shifts towards the reference value with a risk of 0.12%. The shift made the 10th predicted value to deviate with decrement from the target value.

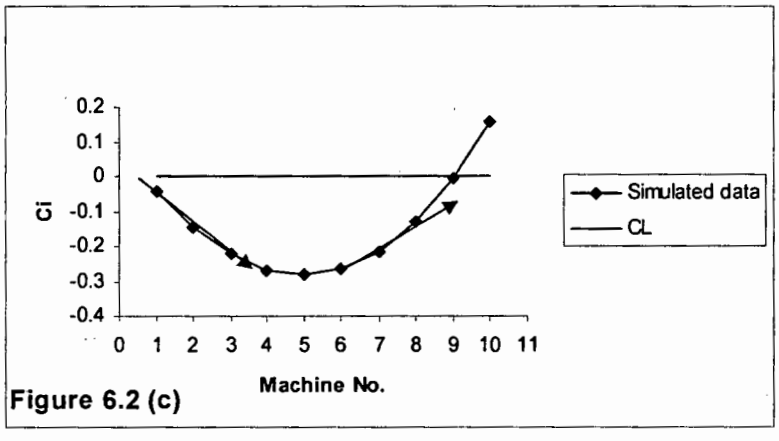


Figure 6.2 (c)

Signal-noise ratio CUSUM signals a slope of 1.75 standard deviation decrement from the reference value 0.876444 and becomes a bit steady from the 4th to the 6th predicted values. From the 7th value there are varying upward slopes to the reference value.

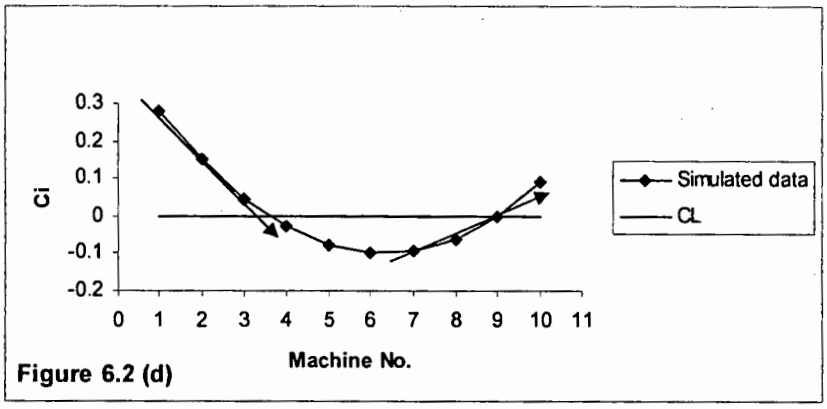


Figure 6.2 (d)

Number of failures simulated values signals that there is a shift in the mean of at least 1.75 standard deviation towards the mean value 3.247556 for the first three predicted values. From the predicted 7th value on wards the mean shifts upwards to the target value.

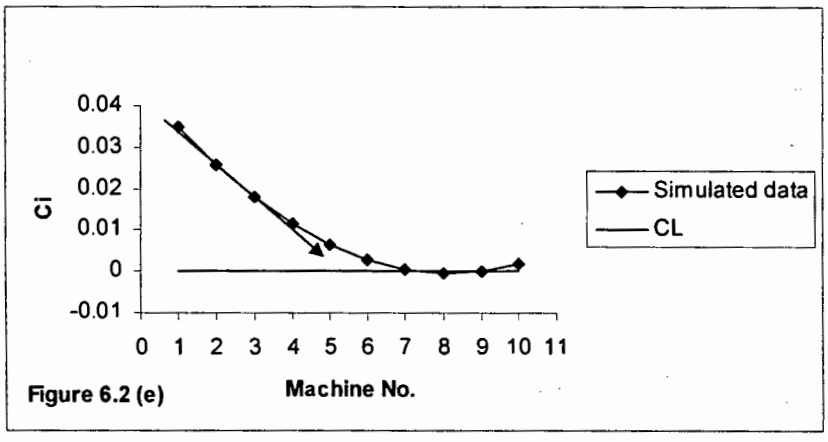


Figure 6.2 (e)

The trend line shows that there is a long but not much steep slope to the horizontal line. This trend line signals a shift in the mean of the average failure rate by at least 1.25 to the reference value 0.104222, with a risk of 0.12%. The predicted values initially were above the reference value this might have been contributed by the prediction model which we used since its accuracy was 71.42% as we mentioned on the model fitting section.

6.5 GM(2,1) model based grey predictive Shewhart control chart

We have implemented the GM(1,1) prediction model on Shewhart and CUSUM control charts for which we would now at this stage implement GM(2,1) predictive model on these charts. Our objective is to get the results and compared them with that of GM(1,1) predictive model and conclude which model could be more appropriate in predictions for machine failure time.

The original data used for computing the GM(2,1) model is the same one as used in computing GM(1,1) we already have the linearity of the accuracy of the variables. So use the least squares estimate to calculate the GM(2,1) parameter sequence given by

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y = \left[a_1 \;\; a_2 \;\; u\right]^T$$

All the parameters calculated were used in the homogeneous equation and had two distinct solutions for λ_1 and λ_2.

Having computed the parameters, λ_1 and λ_2 .we compute the response function

$$\hat{x}^{(1)}(k+1) = C_1 e^{\lambda_1 k} + C_2 e^{\lambda_2 k} + \frac{\mu}{a_2}, \quad k = 0,1,2,\ldots \tag{6.1}$$

where $\lambda_{1,2} = \dfrac{-a_1 \pm \sqrt{a_1^2 4a_2}}{2}$

we have to solve for C_1 and C_1 using the original values.

For $k = 0$ $x^{(1)}(1) = x^{(0)}(1)$ and

$$C_1 + C_2 + \frac{\mu}{a_2} = x^{(0)}(1) \tag{6.2}$$

We differentiate equation (6.2) on the left and right hand sides, we have

$$\frac{dx^{(1)}}{dk} = \lambda_1 C_1 e^{\lambda_1 k} + \lambda_2 C_2 e^{\lambda_2 k} \tag{6.3}$$

We estimate the L.H.S of equation (6.3) by

$$\frac{dx^{(1)}}{dk} = x^{(1)}(k+1) - x^{(1)}(k) \tag{6.4}$$

For $k=1$

$$\frac{dx^{(1)}}{dk} = x^{(1)}(2) - x^{(1)}(1) = x^{(0)}(2) \tag{6.5}$$

Substituting equation (6.5) into equation ((6.2) we get

$$\lambda_1 e^{\lambda_1} C_1 + \lambda_2 e^{\lambda_2} C_2 = x^{(0)}(2) \tag{6.7}$$

We combine equations (6.2) and (6.7) we get

$$\begin{cases} C_1 + C_2 + \dfrac{\mu}{a_2} = x^{(0)}(1) \\ \lambda_1 e^{\lambda_1} C_1 + \lambda_2 e^{\lambda_2} C_2 = x^{(0)}(2) \end{cases}$$

At this stage we can now solve the equations simultaneously to get the values of C_1 and C_1.

We compute

$$\hat{x}^{(1)}(k+1) = C_1 e^{\lambda_1 k} + C_2 e^{\lambda_2 k} + \frac{u}{a_2}$$

and from the IAGO restoration, the predicted value is:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$$

Table 6.5 Results from GM(2,1) on output quality measurements representative indices

	Machine No.	Original data	a_1	a_2	u	Simulated data s	Relative error	MRSE
Mean time between failures	1	5.239	-1.49668	-0.01252	-8.33784	5.24	0.00%	289.61%
	2	5.747				5.63	2.09%	
	3	5.444				6.12	12.36%	
	4	5.082				8.48	66.94%	
	5	5.919				19.30	226.09%	
	6	5.496	-8.9525			68.18	1140.57%	
	7	5.622		0.014004	-8.4669	345.67	6656.40%	1672.61%
	8	5.193	-1.72553	0.03331	-1.5745	166.24	3201.24%	772.55%
Maximal functioning times	1	7.006	-1.43258	-0.02132	-10.4683	7.01	0.00%	108.31%
	2	7.202				7.17	0.43%	
	3	7.150				7.37	3.08%	
	4	6.360				8.55	34.51%	
	5	7.843				13.92	77.54%	
	6	7.050				37.08	426.00%	
	7	7.072	-1.66722	0.028109	-11.2791	2541.91	35843.33%	8799.70%
	8	6.616	-1.69299	0.029619	-11.3817	-262.82	4072.49%	990.50%

From the above results we find that the GM(2,1) predictive model does not fit well in predicting the values for all the quality measurements indices, all the predicted values are out-of-control. The even though the minimum relative error is 0.43% the maximum is 35843.33% this is just too high. The maximum MRSE is also very high 8799.70% and the minimum is 108.31%. Some of the predicted values are even negative

We conclude that currently we should use the GM(1,1) predictive model instead of the GM(2,1) predictive model in control charts. More research has to be done on computing the parameters and the constants of the response function before we could be confident that the GM(2,1) model could be used in quality control charts. We were briefly testing the application of the GM(2,1) model in control charts in this subsection of the thesis.

Chapter 7 Conclusions

The main aim of this thesis is to review the Statistical Process Control techniques used in quality control and contribute to the improvement of these techniques by using the Grey System Theory Approach. Several improvements have been made in quality control charts as we stated in the literature review section. Shewhart quality control charts were improved by introduction of sensitivity or warning lines. The other techniques development that we reviewed in details are the Cumulative Sum charts (CUSUM) and Exponential Weighted Moving Averages charts (EWMA) which generally applied as alternatives to the Shewhart control charts.

In chapter 4 we reviewed the grey system theory from which we used its prediction technique for formulating grey predictive control charts in chapter 5. These formulated control charts were applied on industrial data of packaging machines failure time. Characteristics of the parking machine has some technology that causes difficulties for it full strength in operation, this leads to that there should always be checking for its error of failure of operation. The failure of a machine adds negatively to the quality of the products. As a result, machine's process fault-tolerant design in terms of control limits is always to be implemented so as to be always alert on the machine operation conditions. We used the GM(1,1) model to compute and control limits of a grey prediction Shewhart control chart. The warning lines were also calculated as it is necessary to determine whether an error is caused by a permanent physical failure requiring repair or a transient failure that can be recovered without physical parts replacement, and these can be shown when having warning lines.

From the results we got from the process, the failure time detection has been shown that it can be done easily and accurately using quality control limits. The implementation of grey predictive Shewhart control charts in quality control can be done for high-performance, failure time detection during the process and also prediction of the process operations. When all the points in the process fall within the control limits or even within

one zone especially when the sample size is small, as small less than eight we should use a grey predictive GM(1,1) model. This model can detect the variations no matter how small they are. We have proved this in using the grey predictive CUSUM chart with a reference chart. This is very good way to monitor the progress of the machine failure times for the purpose of evaluating the trend or impact of certain modifications or initiatives that were introduced in the system. This would help in keeping a plot of predicted accidents of failures in the packaging process so as to identify their major causes. The number machine of failures accidents in a given time period is not a continuous variable it is discrete. This is it is either the machine fails or it does not fail at each particular instant.

7.1 Recommendations for future research

Grey predictive control charts could be very useful when applied as pre-control charts, because they have an accurate prediction of, more than 90% for the quality of the products that are to be produced. From the results we got in Chapter 6 we suggest that there should be signal given to the predicted values and with the signals the following setup rules could be introduced in packaging machines:

- A green light signal is to be produced if a predicted quality characteristic of a unit would fall within $\pm 2\sigma$ of the control limits of a Grey Predictive Shewhart Control Chart, that is, within Zones B and C.
- A yellow light is to be produced if a predicted quality characteristic of a unit would fall above $\pm 2\sigma$ but below $\pm 3\sigma$ region of the control limits for a Grey Predictive Shewhart Control Chart, that is, it will fall within Zone A.
- A red light signal is to be produced if a predicted quality characteristic of a unit would fall outside $\pm 3\sigma$ control limits of a Grey Predictive Shewhart Control Chart, that is, out-of-control region.

Setting up this would make it easy for the management to be alert about what is expected of the operations of the machines. We would emphasis that one should always predict at least two values and apply the following rules, these are similar to the ones by Steiner (1997):

- If either of the predicted values is of red signal, stop the process and adjust or fix the machine.
- If both the predicted values are of a green signal, continue operation.
- If either or both of the predicted values are yellow, predict up to the third value. Continue on with the machine operation if the combined predicted values are of the green signal, and stop the process if three yellow signals are for the predicted consecutive values or a single red signal is observed.

7.2 Cautious warning on grey control chart research

We should emphasize that that grey theory is not a rigorous mathematical branch yet and it is in a developing stage. The reason why we appreciate it in some degree and try to explore the potential applications in quality control charts is because within the current development range in grey theory there are many good ideas and models there, particularly these dealing with sparse data with high efficiency. Our work does not mean that we fully accept all the existing contents and statements in grey theory but just pick these useful and correct ones. In our applications, we have realized many potential improving spaces in grey differential equation models and we will explore them and apply them to quality charts more and more.

References

1. Braverman, J.D. (1981), *Fundamentals of Statistical Quality Control*, Reston Publishing Company, Inc.
2. Brown D.W. and Wetherill G.B. (1991) *Statistical Process Control; Theory and practice*, T.J Press (Padstow) Ltd, Padstow
3. Caulcutt, R., (1995) The Rights and the Wrongs of Control Charts, *Applied Statistics*, Vol. 44, No. 3, pp. 279-288.
4. Chang S. I., and Samuel T. R (1998) A Control point methodology for CUSUM Control Charts, *Computers Industrial Engineering*, Vol. 34, No. 3, pp 565 –572.
5. Chiu W. K. (1975), Economic Design of Attribute Control Charts, *Technometrics*, Vol. 17, No. 1, pp 81-87.
6. Crowder, S.V., (1987), A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts, Technometrics, *Vol. 29*, pp 401-407.
7. Yaoguo Dang, Liu Sifeng and Chen Kejia (2004), The GM models that x(n) be taken as initial value. Vol. 33, No. 2, pp 247-254 www.emeraldingsight.com.com/0368-492X.htm
8. Deng, J. L. (1982) Control Problems of Grey Systems, Systems and Control Letters, 1 (6), March
9. Deng, J. L., (1989) Introduction to Grey System Theory, *The Journal of Grey System*, Vol. 1, No. 1, pp 1-24
10. Deng, J. L. (1985) Grey Systems (Social · Economical), The Publishing House of Defense Industry, Beijing.
11. Deng, J. L. (1993) Grey Control System, Huazhong University of Technology Press, Wuhan City, China.
12. Deng, J. L. (2002[a]) The Foundation of Grey Theory, Huazhong University of Technology Press, Wuhan City, China.
13. Deng, J. L. (2002[b]) Grey Prediction and Decision, Huazhong University of Technology Press, Wuhan City, China.
14. DeVor, E. R.; Chang T.; and Sutherland, J. W. (1992) *Statistical quality Design and Control: Contemporary Concepts and Methods*, Macmillan, Inc.
15. Duncan, A.J. (1986), *Quality Control and Industrial Statistics*, 5th edition, Richard D. Irwin, Inc.
16. Ewan, W. D. (1963). When and how to use Cu-sum charts. *Technometrics*, Vol. *5*, pp 1-32.
17. Ewan, W. D. and kemp, K. W. (1960). Sampling Inspection of Continuous Processes with No Autocorrelation Between Successive Results. *Biometrika*, 47, pp363-380.
18. Fang-Mei Tseng, Hsiao-Cheng, and Gwo-Hsiung Tzeng (2001), Applied Hybrid Grey Model to Forecast seasonal Time Series, *Technology Forecasting and Social Change*, Vol. 67, pp 291-302
19. Gan, F. F. (1996), Average Run Length For Cumulative Sum Control Charts Under Linear Trend , Applied Statistics, Vol. 45 No.5, pp 505-512

20. Garving, D. A.. (1987) Competing in the eight dimensions of quality, *Harvard Business* Review, Sept.-Oct, 1987.
21. Goel, A.L., and Wu, S.M. (1973) Economically Optimum Design of Cusum Charts, *Management Science*, Vol. 19, No. 11: pp 1271–1282.
22. Guenther, W. C. (1977*), Sampling inspection in statistical quality control*, Charles Griffin & Company Limited.
23. Guo, R. and Dunne, T. (2005). Grey Predictive Process Control Charts. *Proceedings of the 4th International Conference on Quality and Reliability (ICQR 2005)*, Beijing, China, pp. 55-62.
24. Guo, R. and Dunne, T. (2006) Grey Predictive Control Charts. *Communications in Statistics – Theory and Methods*, Vol. 35, No. 10, pp. 1857-1868, 2006.
25. Howell J. M. (1949) Control Chart for Largest and Smallest Values. *The Annals of Mathematical Statistics,* Vol. 20, No. 2: pp. 305-309.
26. Hunter, S. J. (1986). The exponentially weighted moving average", *Journal of Quality Technology*, Vol. 18 No. 4, pp203-210.
27. Johnson N.L (1961), A simple Theoretical Approach to Cumulative Sum Control Charts, *Journal of the American Statistical Association*, Vol. 56, No. 296, pp 835-840.
28. Johnson N. L. and Leone F. C., (1964) *Statistics and Experimental Design in Engineering and the Physical Sciences*, Vol. 1, John Wiley & Sons Inc,
29. Johnson R. A. and Bagshaw M. (1975), The effect of serial Correlation on the Performance of CUSUM Test II, *Technometrics*, Vol. 17, No. 1, pp 73-75.
30. Kemp, K. W. (1961). The Average Run Length of the Cumulative Sum Chart When a V-mask is used. *J. R. Statist. Soc. B.* Vol. 23, pp149-153.
31. Khan R. A. (1981), A Note on Page's Two-Sided Cumulative Sum Procedure. *Biometrika*, Vol. 68, N0. 3, pp 717-719.
32. Klein, M. (2000) Two Alternatives to the Shewhart $\overline{X}$ Control Chart, *Journal of Quality Technology*, Vol. 32, No. 4, pp 427- 431.
33. Levinson, W.A (2004) Control Charts Control Multiple Attributes, http://www.qualitymag.com/CDA/Archives/118f74ce57c38010VgnVCM100000f932a8c0
34. Liu, S. F. and Lin Y. (2006) Grey Information –Theory and Practical Applications. Springer, London.
35. Lucas, J.M. (1985), Counted Data CUSUM's, *Technometrics*, Vol. 27 No 2. pp 129-144.
36. Lucas, J. M. and Crosier, R. B. (1982), "Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM a Head Start," *Technometrics*, Vol. 24, pp 199 –205.
37. Lucas, J.M. and Saccucci, M.S. (1990), Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics*, Vol. 32. pp 1-12.
38. MacGregor, J. F. and Harris, T. J. (1993). The exponentially weighted moving variance. *Journal of Quality Technology*, Vol. 25, No. 2, pp106-118.
39. MacGregor, J.F., Jenkins, G.M. and Box, G.E.P. (1974) Some Recent Advances in Forecasting and Control. *Applied Statistics*, Vol. 23, No. 2. pp 158-179.

40. Makrymichalos. M., Antony, J., Antony, F. and Kumar M. (2005), Statistical thinking and its role for industrial engineers and managers in the 21st century. *Managerial Auditing Journal*, Vol. 20, No. 4. pp 354-363.
41. Mcquater, R.E., Scurr, C.H., Dale, B.G. and Hill, P.G. (1995), using quality tools and techniques successfully. *The TQM Magazine*, Vol. 7 No. 6 pp 37-42
42. Montgomery, D.C. (2001), *Introduction to Statistical Quality Control*, 4th edition, John Wiley & Sons, Inc.
43. Nelson, L. S. (2003), When Should the Limits on a Shewhart Control Chart be other Than a Center line –Sigma? *Journal of Quality Technology*, Vol. 35 No. 4 pp 424-425
44. David K. W. Ng and Ulon Eng (1995), Constructing Grey System Theory to Probability and Fuzzy, *ACM SIGICE Bulletin*, Vol. 20 No. 3. pp 3-9.
45. North W. R. S (1982) The Quangle-A Modification of the CUSUM Chart. *The Statistician,* Vol. 31, No. 2: pp. 155-158
46. Page, E.S. (1954). Continuous inspection schemes. *Biometrika,* Vol. 41, pp100-115.
47. Page, E.S. (1962). A Modified Control Chart with Warning Lines, *Biometrika,* Vol. 49, No. ½, pp171 – 176.
48. Page, E.S. (1963), Controlling the Standard Deviation by Cusums and Warning Lines, *Technometrics*, Vol. 5, No. 3, pp 307 – 315
49. Pan, R. (2002). Statistical Process Adjustment Methods for Quality Control in Short-Run Manufacturing. (PhD Thesis, the Pennsylvania State University, Department of Industrial and manufacturing Engineering). WEBADRESS???????
50. Pedrycz W. (1993), *Fuzzy Control and Fuzzy Systems*, Second, extended, edition, Research Studies Press Ltd.
51. Reynolds M. R. Jr. (1975), Approximations to the Average Run Length in Cumulative Sum Control Charts, *Technometrics*, Vol. 17, No. 1, pp 65-71
52. Reynolds, Jr. M. R. and Stoumbos Z. G. (2004). Should Observations Be Grouped for Effective Process Monitoring? *Journal of Quality Technology,* Vol. 36, No. 4, October 2004, pp 343-366.
53. Roberts, S. W. (1959), Control Chart Tests Based on Geometric Moving Averages, *Technometrics*, Vol.1, No. 3, pp 239 -250.
54. Robinson, P. B. and Ho, T.Y. (1978). Average run lengths of geometric moving average charts by numerical methods", *Technometrics*, Vol. 20 No. 1, pp85-93.
55. Rowlands R. J., Nix A. B. J. and Abdollahian M. A. (1982) Snub-nosed V-mask Control Schemes, *The Statistician,* Vol. 31, No. 2: pp. 133-142.
56. Ryan, T. P. (1989). *Statistical methods for quality improvement.* New York: Wiley.
57. Sefik M. (1998), Importance of the Rational Subgroups in designing Control Charts, *Computers Industrial Engineering*, Vol. 35, Nos. 1-2, pp 205 – 208.
58. Siegmund, D. Sequential Analysis: Tests and Confidence Intervals. Springer-Verlag, New York, 1985.
59. Snee, R. D. (1990), Statistical Thinking and its Contribution to Total Quality, *The American Statistician*, Vol. 44, No. 2, pp 116-121

60. Steiner S. H. (1998), Grouped data exponentially weighted moving average control charts, *Royal Statistics Society, Applied Statistics,* Vol. 47, Part 2, pp. 203-216.
61. Steiner, S.H. (1999), Exponentially Weighted Moving Average Control Charts with Time-Varying Control Limits and Fast Initial Response, *Journal of Quality Technology*, Vol. 31, pp 75-86.
62. Tzu-Li Tien (2005), The indirect measurement of tensile strength of material by the grey prediction model GMC(1,n), *Measurement Science and Technology*, Vol 16 pp 1322-1328.
63. Vardeman S. and Ray Di-ou (1985), Average Run lengths For CUSUM Schemes When Observations Are Exponentially Distributed, *Technometrics*, Vol. 27, No. 2, pp 145-150.
64. Vardeman, S. B., and Jobe, M. J. (1999) *Statistical Quality Assurance Methods for Engineers*, John Wiley & Sons, Inc.
65. Vander Wiel S. A. (1995), Monitoring Process that Wander Using Integrated Moving Average models, AT&T Bell Laboratories, Murray Hill, NJ 07974
66. Tien-Chin, Wang, Lui Mei-Chu and Hung Hsiu-Huang (2004), Application of Grey Theory on Forecasting The Exchange rate Between TWD and USD, http://atisr.org/conference/BAI2005CD/Paper/2005bai120.doc
67. Wardell Don G., Moskowitz H. and Plante R. D. (1992) Control Charts in the Presence of Data Correlation. *Management Science,* Vol. 38, No. 8: pp. 1084-1105.
68. Jian Wei, Wu Huaiqing, Tsung Fugee Nair. N. Vijayan and Tsui Kwok-Leung (2002), Proportional Integral Derivative Charts for Process Monitoring. *Technometrics*, Vol. 44, No. 3. pp 1-10.
69. Kun-Li Wen, (2004) Grey Systems: Modeling and Prediction, *Yang's Scientific research Institute*, USA.
70. Kun-Li Wen and Ting-Cheng Chang (2005) The research and development of Completed GM(1,1) Model Toolbox Using Matlab, *International Journal of Computational Cognition*, Vol. 3, No. 3 pp 41-47.
71. Westgard, J.O., Groth, T., Aronsson T. and Carl-Hendric de Verdier (1977) Combined Shewhart-Cusum Control Chart for Improved Quality Control in Clinical Chemistry, *Clinical Chemistry.* Vol. 23, No. 10 pp 1881-1887
72. Wetherill, G. B. and Brown, D. W. (1991), Statistical Process Control: Theory and Practice, London: Chapman and Hall.
73. Wetherill, G. B. and Glazebrook, K. D. (1986), Sequential Methods in Statistics, 3rd edition. London: Chapman and Hall
74. Wieringa, J. E. (1999) *Statistical Process Control For Serial Correlated Data*, PhD thesis,
75. Woodall, W.H. (1985), The Statistical Design of Quality Control Charts, *The Statistician,* Vol. 34, No. 2: pp. 155-160
76. Woodall, W.H. (2000), Controversies and Contradictions in Statistical Process Control, *Journal of Quality Technology*, Vol. 32, No. 4
77. Yeh A. B., Lin D. K. J. and Venkataramani Chandaramouliswaran, (2004) Unified CUSUM Charts for Monitoring Process Mean and Variability, *Quality Technology Quantitative Management*, Vol. 1, No. 1, pp. 65-86.

78. Yi Lin, Mian-yun Chen and Sifeng Liu, (2004) Theory of grey systems: capturing uncertainties of grey information, *Kybernets*: *The International Journal of Systems and Cybernetics.* Vol. 33, No. 2, pp 196-218
79. Hongjian Xia, (1992), *A Cumulative Score Control Scheme*, Applied Statistics, Vol. 41, No. 1. pp 47-54.
80. Jun, Xia and Baoqing, Hu. (1997), A Grey System Model for Predicting Trend Change of Urban Waste Water Load, *Journal of Enviromental Hydrology*, Vol. 5. Paper 2. pp 1-10.
81. Yashchin, E. (1993) Performance of CUSUM Control Schemes for Serially Correlated Observations. Technometrics, Vol. 35, No 1 37–52, 1993.

Appendix A: Notations and Acronyms

This appendix summarizes notations and acronyms used in this thesis.

CL Center Line

UCL Upper Control Limit

LCL Lower Control Limit

n Sample Size

m Sample Number

$\bar{x}$ Average Measurements

$\bar{\bar{x}}$ Average of Averages

R Range

$\bar{R}$ Average of Ranges

σ Progress Standard Deviation (sigma)

D_i Nonconforming units in sample i.

$\hat{p}$ The fraction nonconforming.

$\bar{p}$ The average of the individual sample fractions nonconforming.

A A coefficient vector

B An accumulated matrix

$e(k)$ The k^{th} error

m A constant based on $x^{(1)}$

$x^{(0)}$ The original sequence

$x^{(0)}(i)$ The time series data at time i in $x^{(0)}$

$\hat{x}^{(0)}(k+1)$ The predicted value $x^{(0)}(k+1)$ of at time $(k+1)$

$x^{(1)}$ A new time sequence based on $x^{(0)}$

$x^{(1)}(i)$ The time series data at time i in $x^{(1)}$

$\hat{x}^{(1)}(k+1)$ The predicted value $x^{(1)}(k+1)$ of at time $(k+1)$

Y A constant vector

$z^{(1)}(k)$ The k^{th} background value

Appendix B: Formulas for Designing Charts

Formulas for designing Shewhart Charts, CUSUM charts and EWMA Charts

Table B.1. Control limits for various variable control charts

$\bar{x}$ *Chart*	CL	Use Factor	LCL	UCL
σ assumed known	$\bar{\bar{x}}$	A	$\bar{\bar{x}} - A\sigma$	$\bar{\bar{x}} + A\sigma$
$\bar{R}$ *for estimating* σ	$\bar{\bar{x}}$	A_2	$\bar{\bar{x}} - A_2\bar{R}$	$\bar{\bar{x}} + A_2\bar{R}$
$\hat{\sigma}$ *for estimating* σ	$\bar{\bar{x}}$	A_1	$\bar{\bar{x}} - A_1\bar{\sigma}$	$\bar{\bar{x}} + A_1\bar{\sigma}$

R Chart	CL	Use Factor(s)	LCL	UCL
σ assumed known	$\bar{R} = d_2\sigma$	D_1 *and* D_2	$D_1\sigma$	$D_2\sigma$
$\bar{R}$ *for estimating* σ	$\bar{R}$	D_3 *and* D_4	$D_3\bar{R}$	$D_4\bar{R}$

The factors $A, A_1, A_2, D_1, D_2, D_3$ *and* D_4 and are listed in the table of factors in Appendix C TableC.1

Definitions:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{\bar{x}} = \frac{\sum_{i=1}^{m} \bar{x}_i}{m} \qquad \hat{\sigma} = \frac{\bar{R}}{d_2} \qquad R = x_{\max} - x_{\min} \qquad \bar{R} = \frac{\sum_{i=1}^{m} R_i}{m}$$

$$A = \frac{3}{\sqrt{n}} \qquad A_2 = \frac{3}{d_2\sqrt{n}} \qquad A_3 = \frac{3}{c_4\sqrt{n}}. \qquad B_3 = 1 - \frac{k}{c_4} \qquad B_4 = 1 + \frac{k}{c_4}$$

$$B_5 = c_3 - 3\sqrt{1 - c_4^2} \qquad B_6 = c_3 + 3\sqrt{1 - c_4^2} \qquad D_1 = d_2 - 3d_3 \quad D_2 = d_2 + 3d_3 \quad D_3 = 1 - 3\frac{d_3}{d_2}$$

$$D_4 = 1 + 3\frac{d_3}{d_2} \qquad c_4 = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma(n-1/2)} \cong 1 - \frac{1}{4(n-1)} \qquad k = 3\sqrt{1 - c_4^2}$$

Table B.2. Control Charts formulas for attribute data

	p (fraction)	np (number of nonconforming)	c (count of nonconformances)	u (count of nonconformances/unit)
CL	$\bar{p}$	$n\bar{p}$	$\bar{c}$	$\bar{u}$
UCL	$\bar{p}+3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$	$n\bar{p}+3\sqrt{n\bar{p}(1-\bar{p})}$	$\bar{c}+3\sqrt{\bar{c}}$	$\bar{u}+3\sqrt{\frac{\bar{u}}{n}}$
LCL	$\bar{p}-3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$	$n\bar{p}-3\sqrt{n\bar{p}(1-\bar{p})}$	$\bar{c}-3\sqrt{\bar{c}}$	$\bar{u}-3\sqrt{\frac{\bar{u}}{n}}$
Notes	If n varies use $\bar{n}$ or individual n_i	n must be a constant	n must be a constant	If n varies use $\bar{n}$ or individual n_i

Definitions:

$$\hat{p}=\frac{D_i}{n} \qquad i=1,2,\ldots,m \qquad \bar{p}=\frac{\sum_{i=1}^{m} D_i}{mn}=\frac{\sum_{i=1}^{m}\hat{p}_i}{m} \qquad u=\frac{c}{n} \quad \bar{u}=\frac{\sum_{i=1}^{n} u_i}{n}$$

Appendix C: Factors for Constructing Variables Control Charts

Table C.1. Factors for Constructing Variables Control Charts

	Charts for Averages					Charts for Standard Deviations						Charts for Ranges				
	Factors for Control limits			Factors for Center Line		Factors for Control limits				Factors for Center Line		Factors for Control limits				
n	A	A_2	A_3	c_4	$1/c_4$	B_3	B_4	B_5	B_6	D_2	$1/d_2$	d_3	D_1	D_2	D_3	D_4
2	2.12132034	1.87996982	2.658695974	0.7979	1.2533213	0	3.266567573	0	2.60632893	1.128	0.88622627	0.853	0	3.68588	0	3.266524
3	1.73205081	1.02332595	1.954403267	0.8862	1.1283753	0	2.568144252	0	2.27596648	1.693	0.59081751	0.888	0	4.35768	0	2.574594
4	1.50000000	0.72859745	1.628098815	0.9213	1.0853992	0	2.266026462	0	2.0877355	2.059	0.48573163	0.880	0	4.69818	0	2.282055
5	1.34164079	0.57681907	1.427292616	0.9400	1.0638411	0	2.088954115	0	1.96359598	2.326	0.42993555	0.864	0	4.91817	0	2.114496
6	1.22474487	0.48324654	1.287132168	0.9515	1.050939	0.0303324	1.969667624	0.0288622	1.87419783	2.534	0.39456915	0.848	0	5.07853	0	2.003831
7	1.13389342	0.41928346	1.18191461	0.9594	1.0423507	0.117699	1.882300975	0.1129169	1.80582309	2.704	0.36977326	0.833	0.20476	5.20396	0.0757148	1.924285
8	1.06066017	0.37252746	1.099095543	0.9650	1.0362372	0.185084	1.814916008	0.1786116	1.75144839	2.847	0.35122225	0.820	0.38771	5.30669	0.1361724	1.863828
9	1.00000000	0.33669694	1.031661697	0.9693	1.0316617	0.2391237	1.760876286	0.231785	1.70683499	2.970	0.33669694	0.808	0.54654	5.39352	0.1840183	1.815982
10	0.94868330	0.30826327	0.975349349	0.9727	1.0281085	0.2837155	1.716284532	0.2759587	1.66936131	3.078	0.32493802	0.797	0.68636	5.46866	0.2230245	1.776976
11	0.90453403	0.28508386	0.927394303	0.9754	1.025273	0.321279	1.678720953	0.3133595	1.63734048	3.173	0.31517207	0.787	0.81094	5.5348	0.2555856	1.744414
12	0.8660254	0.26577752	0.885905115	0.9776	1.0229551	0.3535215	1.64647851	0.3455885	1.60953153	3.258	0.30689344	0.778	0.92305	5.59387	0.283278	1.716722
13	0.83205029	0.24941705	0.849542372	0.9794	1.0210229	0.3816238	1.618376211	0.3737662	1.58505385	3.336	0.29976199	0.770	1.02475	5.64721	0.3071811	1.692819
14	0.80178373	0.23535081	0.817337661	0.9810	1.0193992	0.4062223	1.593777685	0.3984919	1.5634481	3.407	0.29353403	0.763	1.1177	5.69582	0.328083	1.671917
15	0.77459667	0.22310904	0.788538021	0.9823	1.0179982	0.428263	1.571736976	0.4206913	1.54394867	3.472	0.28803254	0.756	1.20323	5.74043	0.3465694	1.653431
16	0.75000000	0.21234548	0.762598121	0.9835	1.0167975	0.4478276	1.552172351	0.4404295	1.52653046	3.532	0.28312731	0.750	1.28228	5.78168	0.3630485	1.636952
17	0.72760688	0.20279577	0.739054834	0.9845	1.0157337	0.465739	1.534261006	0.4585247	1.5104953	3.588	0.27871612	0.744	1.35576	5.82	0.3778722	1.622128
18	0.70710678	0.1942569	0.717576218	0.9854	1.014806	0.4818482	1.518151832	0.474818	1.496002	3.640	0.27472075	0.739	1.42432	5.8558	0.3912903	1.60871
19	0.68824720	0.18656944	0.697870841	0.9862	1.0139828	0.4965613	1.503438666	0.4897138	1.48270625	3.689	0.27107911	0.733	1.48855	5.88937	0.4035148	1.596485
20	0.67082039	0.17960626	0.679704126	0.9869	1.0132431	0.510149	1.48985098	0.5034814	1.47037863	3.735	0.2677412	0.729	1.54891	5.92099	0.414707	1.585293
21	0.65465367	0.17326489	0.662886724	0.9876	1.0125762	0.5227213	1.477278725	0.5162291	1.45893092	3.778	0.26466649	0.724	1.60586	5.95082	0.4250173	1.574983
22	0.63960215	0.1674623	0.647259226	0.9882	1.0119716	0.5344046	1.465595442	0.5280826	1.44825745	3.819	0.2618226	0.720	1.65968	5.97908	0.4345417	1.565458
23	0.62554324	0.1621284	0.632692669	0.9887	1.0114291	0.545137	1.454862966	0.538977	1.43842301	3.858	0.25918016	0.716	1.71068	6.00596	0.4433743	1.556626
24	0.61237244	0.15720601	0.619064523	0.9892	1.0109281	0.555274	1.444725998	0.5492715	1.42910851	3.895	0.25671634	0.712	1.75917	6.03153	0.4516077	1.548392
25	0.6	0.15264729	0.606281072	0.9896	1.0104685	0.5647777	1.435222304	0.5589266	1.4203534	3.931	0.25441214	0.708	1.80534	6.05592	0.4593004	1.5407

Table C.2. Hartley's Constants and Constants used in design of Control Charts for Range

Sample size n	d_n	$D'_{0.999}$	$D'_{0.001}$	$D'_{0.975}$	$D'_{0.025}$
2	1.128	0.00	4.12	0.04	2.81
3	1.693	0.04	2.98	0.18	2.17
4	2.059	0.10	2.57	0.29	1.93
5	2.326	0.16	2.34	0.37	1.81
6	2.534	0.21	2.21	0.42	1.72
7	2.704	0.26	2.11	0.46	1.66
8	2.847	0.29	2.04	0.50	1.62
9	2.97	0.32	1.99	0.52	1.58
10	3.078	0.35	1.93	0.54	1.56
11	3.173	0.38	1.91	0.56	1.53
12	3.258	0.40	0.187	0.58	1.51

Table C.3 FIR-EWMA Example using data from a process initially out-of-control

Sample number	$\bar{x}_t$	FIR-EWMA Control limits				
		$\lambda = 0.5$	$\lambda = 0.25$	$\lambda = 0.2$	$\lambda = 0.1$	$\lambda = 0.05$
0		0	0	0	0	0
1	0.8	0.75	0.375	0.3	0.15	0.075
2	1.9	0.996	0.557	0.456	0.24	0.123
3	1.4	1.152	0.689	0.576	0.316	0.166
4	2	1.265	0.787	0.668	0.38	0.204
5	1.1	1.354	0.862	0.739	0.435	0.238
6	0.7	1.426	0.919	0.794	0.48	0.268
7	2.6	1.483	0.962	0.837	0.518	0.294
8	0.5	1.53	0.997	0.871	0.549	0.318
9	1.2	1.568	1.024	0.897	0.574	0.338

Table C.4 Machine Failure Time Quality Measurement Representative Indices

Machine No	Quality Measurement Representative Index				
	Mean time between failures	Maximal functioning times	Signal-noise ratio	Number of failures	Average failure rate
1	188.436	1103.750	0.833	34	0.139
2	313.196	1342.000	0.867	20	0.075
3	231.408	1273.500	0.764	20	0.097
4	161.091	578.083	1.073	41	0.152
5	372.157	2546.667	0.559	18	0.066
6	243.648	1153.417	0.814	28	0.097
7	276.345	1178.417	0.919	21	0.082
8	179.923	747.250	0.901	38	0.150
9	269.655	766.167	1.158	22	0.080
10	325.8860	911.667	1.181	19	0.069

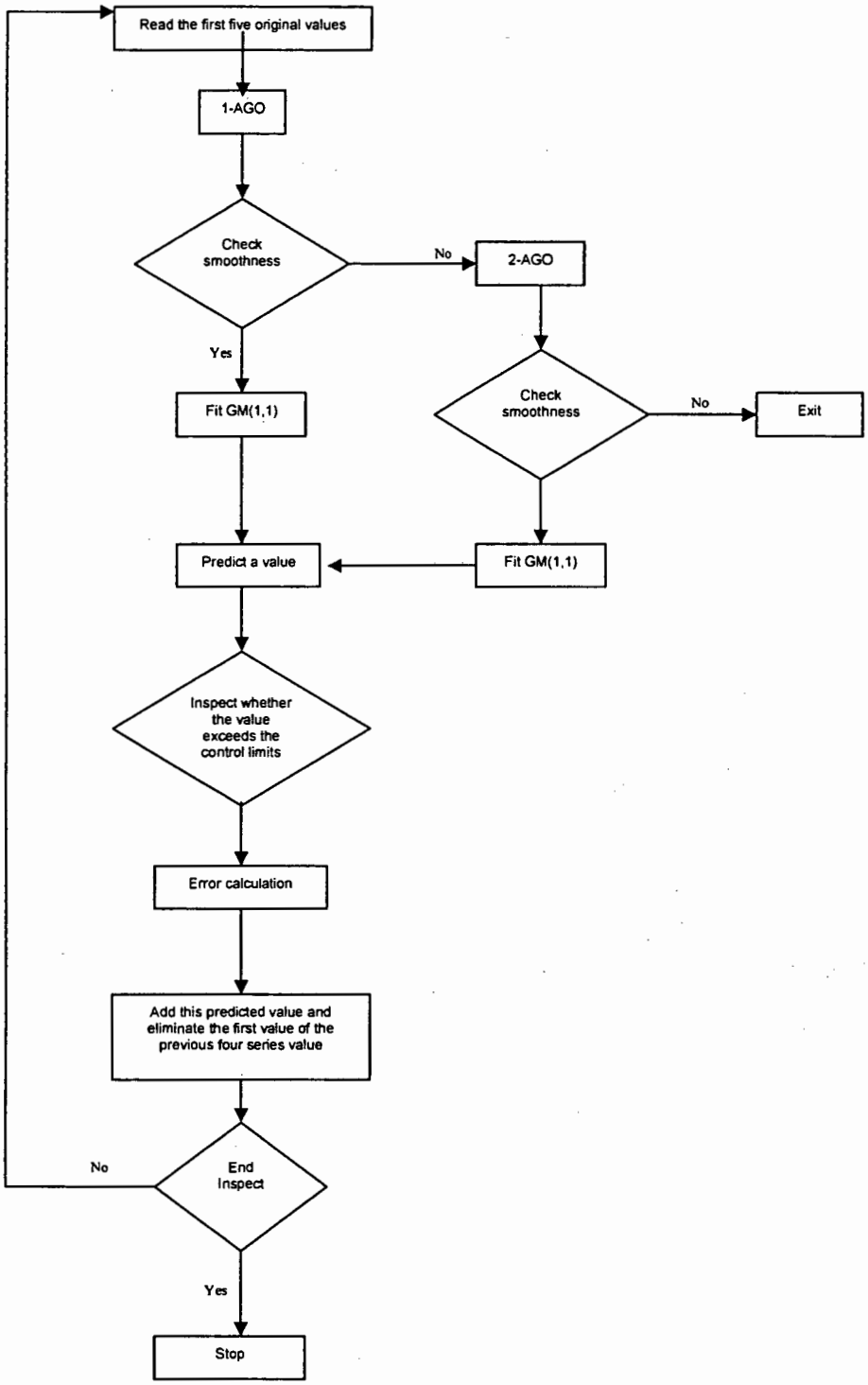
Source: TBF Packaging

Table C.5 Data Transformed to logarithm to base e ($e = 2.7183$)

Machine No	Quality Measurement Representative Index				
	Mean time between failures	Maximal functioning times	Signal-noise ratio	Number of failures	Average failure rate
1	5.239	7.006	-0.183	3.526	-1.973
2	5.747	7.202	-0.143	2.996	-2.590
3	5.444	7.150	-0.269	2.996	-2.333
4	5.082	6.360	0.070	3.714	-1.884
5	5.919	7.843	-0.582	2.890	-2.718
6	5.496	7.050	-0.206	3.332	-2.333
7	5.622	7.072	-0.084	3.045	-2.501
8	5.193	6.616	-0.104	3.638	-1.897
9	5.597	6.641	0.147	3.091	-2.526
10	5.787	6.815	0.166	2.944	-2.674

Appendix D: VBA Codes

1 Flow Chart of the Program



2 Codes for GM(1,1)

```
Option Base 1
Sub GM_1_1()
'machine numbers
Range("A1").Value = "k"
'original data series, ie,machine failure
Range("B1").Value = "X(0)"
'1-AGO series
Range("C1").Value = "X(1)"
'background value z(k)
Range("D1").Value = "z(1)"
Range("B2").Copy Range("C2")
Range("C3").Select
ActiveCell.FormulaR1C1 = "=R[-1]C+RC[-1]"
Range("C3").Copy Range("C4:C6")
Range("D3").Formula = "=SUM(C2:C3)*0.5"
Range("D3").Copy Range("D4:D6")
Range("E1").Value = "X_hat(i+1)"

'other columns headings
Range("F1").Value = "Predicted data"
Range("G1").Value = "Residuals"
Range("H1").Value = "Relative error"

'Matrices setting, B and Y are matrix as defined in Grey Theory
Range("A16").Value = "Y ="
Dim Y As Range
Set Y = Range("B15:B18")
Range("B3:B6").Copy Range("B15:B18")
Range("C16").Value = "B ="
Range("E15:E18").Formula = "=1"
Range("D15").Formula = "=Product(D3*-1)"
Range("D15").Copy Range("D16:D18")
Dim B As Range
Set B = Range("D15:E18")

'calculating for the values of a and u
Range("A20").Value = "Sol ="
Dim sol As Range
Set sol = Range("B20:B21")
sol =
Application.MMult(Application.MMult(Application.MInverse(Application.MMult(Appli
cation.Transpose(B), B)), Application.Transpose(B)), Y)
Range("C20").Value = "a"
```

```
Range("C21").Value = "u"
Dim a, u, d
Set a = Range("B20")
Set u = Range("B21")
'Range("D20").Value = "d="
'Set d = Range("E20")
'Range("E20") = u / a

'get X1hat, fitting values of X1, from the response function

Dim k1, k2, k3, k4, k5, k6 As Integer
k1 = Range("A2")
k2 = Range("A3")
k3 = Range("A4")
k4 = Range("A5")
k5 = Range("A6")
k6 = Range("A7")
Range("E2") = (Range("B2") - u / a) * Exp(-a * (k1 - 1)) + u / a
Range("E3") = (Range("B2") - u / a) * Exp(-a * (k2 - 1)) + u / a
Range("E4") = (Range("B2") - u / a) * Exp(-a * (k3 - 1)) + u / a
Range("E5") = (Range("B2") - u / a) * Exp(-a * (k4 - 1)) + u / a
Range("E6") = (Range("B2") - u / a) * Exp(-a * (k5 - 1)) + u / a
Range("E7") = (Range("B2") - u / a) * Exp(-a * (k6 - 1)) + u / a

Dim Pred As Range
Set Pred = Range("F2:F6")
Range("F3") = Range("E3") - Range("E2")
Range("F4") = Range("E4") - Range("E3")
Range("F5") = Range("E5") - Range("E4")
Range("F6") = Range("E6") - Range("E5")
Range("F7") = Range("E7") - Range("E6")
Range("F3:F7").Font.Bold = True

'calculating the sequence errors residuals
Range("G3") = Range("B3") - Range("F3")
Range("G4") = Range("B4") - Range("F4")
Range("G5") = Range("B5") - Range("F5")
Range("G6") = Range("B6") - Range("F6")

'calculating sequence relative error
Range("H3") = Abs((Range("G3") / Range("B3")))
Range("H4") = Abs((Range("G4") / Range("B4")))
Range("H5") = Abs((Range("G5") / Range("B5")))
Range("H6") = Abs((Range("G6") / Range("B6")))

'Mean Realtive Simulation Error MRSE
```

```
Range("G12").Value = "MRSE ="
Range("H12").Formula = "=(AVERAGE(H3:H6))"

Range("A13").Value = "S1 ="
Range("B13").Formula = "=StDev(B2:B6)"
Range("F13").Value = "S2 ="
Range("G13").Formula = "=StDev(G3:G6)"
Range("A12:H13").Font.Bold = True

Dim C
C = Range("E20")
Range("D20") = "C  ="
Range("E20") = Range("G13") / Range("B13")
Range("F20") = "PreciC ="
Range("G14") = "=Average(G3:G6)"
Range("F14") = "AvrRes ="


If C <= 0.35 Then
PreciC = 1
ElseIf C > 0.35 And C <= 0.5 Then
PreciC = 2
ElseIf C > 0.5 And C <= 0.65 Then
PreciC = 3
ElseIf C > 0.65 Then
PreciC = 4
End If
If P <= 0.95 Then
PreciP = 1
ElseIf P >= 0.8 And P < 0.95 Then
PreciP = 2
ElseIf P >= 0.7 And P < 0.8 Then
PreciP = 3
ElseIf P < 0.7 Then
PreciP = 4
End If
Precision = Application.Max(PreciC, PreciP)

'determining precision
Dim g
Range("G" & j + 20) = Precision
If Precision = 1 Then
Range("G" & j + 21) = "Good"
ElseIf Precision = 2 Then
Range("G" & j + 21) = "Acceptable"
ElseIf Precision = 3 Then
```

```
Range("G" & j + 21) = "Unacceptable"
Range("G" & j + 21).Font.Color = vbRed
Range("F21") = "Precision is"

End If

Range("A20:H21").Font.Bold = True

End Sub
```

3 Codes For Shewhart Control Limits

```
Option Base 1
Sub ShewhartControlLimits()
Dim dataRange As Range
Set dataRange = Selection
Range("A12").Value = "Mean ="
Range("A13").Value = "StdDev ="
Range("A14").Value = "Us ="
Range("A15").Value = "Ls ="
Range("A16").Value = "Cs+2StdDev ="
Range("A17").Value = "Cs-2StdDev ="
Range("A18").Value = "Cs+StdDev ="
Range("A19").Value = "Cs-StdDev ="
Dim theRange As Range
Set theRange = Range("B2:B6")
Range("B12").Formula = "=AVERAGE(B2:B6)"
Range("B13").Formula = "=StDev(B2:B6)"
Range("B14").Formula = "=AVERAGE(B2:B6)+3*Stdev(B2:B6)"
Range("B15").Formula = "=AVERAGE(B2:B6)-3*Stdev(B2:B6)"
Range("B16").Formula = "=AVERAGE(B2:B6)+2*Stdev(B2:B6)"
Range("B17").Formula = "=AVERAGE(B2:B6)-2*Stdev(B2:B6)"
Range("B18").Formula = "=AVERAGE(B2:B6)+Stdev(B2:B6)"
Range("B19").Formula = "=AVERAGE(B2:B6)-Stdev(B2:B6)"
Range("B12:B19").Copy Range("C12:G19")
Range("B12:G19").Font.Bold = True
End Sub

Sub GM_2_1()
'machine numbers
Range("A1").Value = "k"
'original data series, ie,machine failure
Range("B1").Value = "X(0)"
Range("C1") = "alphaX(0)"
```

4 Codes For GM(2,1)

```
'alphaX(0)
Range("C2") = 0
Range("C3").Select
   'Application.CutCopyMode = False
   ActiveCell.FormulaR1C1 = "=RC[-1]-R[-1]C[-1]"
   Range("C3").Copy Range("C4:C6")
'1-AGO series
Range("D1").Value = "X(1)"
Range("D2").Select
   ActiveCell.FormulaR1C1 = "=RC[-2]"
   Range("D3").Select
   ActiveCell.FormulaR1C1 = "=R[-1]C+RC[-2]"
   Range("D3").Copy Range("D4:D6")
Range("E1") = "Z(1)"
   Range("E3").Select
   ActiveCell.FormulaR1C1 = "=0.5*(R[-1]C[-1]+RC[-1])"
   Range("E3").Copy Range("E3:E6")

'Matrices setting, B and Y are matrix as defined in Grey Theory
Range("A16").Value = "Y ="
Dim Y As Range
Set Y = Range("B15:B18")
Range("B15") = Range("C3")
Range("B16") = Range("C4")
Range("B17") = Range("C5")
Range("B18") = Range("C6")
Range("C16").Value = "B ="
Range("F15:F18").Formula = "=1"
Range("D15").Formula = "=Product(B3*-1)"
Range("D15").Copy Range("D16:D18")
Range("E15").Formula = "=Product(E3*-1)"
Range("E15").Copy Range("E16:E18")
Dim b As Range
Set b = Range("D15:F18")

'calculating for the values of a and u
Range("A21").Value = "Sol ="
Dim sol As Range
Set sol = Range("B21:B23")
sol =
Application.MMult(Application.MMult(Application.MInverse(Application.MMult(Appli
cation.Transpose(b), b)), Application.Transpose(b)), Y)
Range("C21").Value = "a1"
Range("C22").Value = "a2"
Range("C23").Value = "u"
```

```
'the fitting coefficient vector of a and u
End Sub
```