Statistical Control and Error Measurement of Belt Scales in Bulk Handling Systems

by

Hamed Darban Hosseini

B.A.Sc., Mechanical Engineering, Ryerson University, Ontario, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Mechanical Engineering)

The University of British Columbia

(Vancouver)

April 2012

.

Abstract

Fast-paced integral weighing systems such as conveyor belt scales are used in the bulk material handling industry. Despite their numerous benefits and widespread use, measurement errors caused by mis-calibration, out of calibration, mechanical malfunctioning and maintenance-related errors can be extremely costly; therefore weighing systems need constant monitoring. This research proposes a method to monitor and control the error in conveyor belt scales by the use of Statistical Process Control. The cumulative conveyed weight on a scale network was analyzed to identify process error limits. Further, series of pattern recognition identifiers, along with statistical tools, were used to monitor and locate unnatural patterns within the scale network. A process validation was conducted using four series of scale data that were obtained from industrial scales. The proposed method performed successfully in determining the start points and types of nine defined patterns. It outperformed conventional monitoring techniques that are based on experimental values, and also provided more detailed information on types, start points, exact locations and potential cases of the unnatural patterns than other conventional methods used in manufacturing known as Western Electric Company (WECO) rules.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deepest gratitude to my academic supervisor and professional mentor, Dr. Farrokh Sassani. His encouragement, support, and extensive knowledge in Statistical Process Control provided me the resources to work in this field and complete the presented research in a unique environment. As a mentor, he has genuinely guided me through technical, professional and industrial aspects of the project.

I would also like to extend my appreciation to the resourceful network of faculty and staff at the Department of Mechanical Engineering of University of British Columbia; in particular Dr. Ryozo Nagamune and Dr. Walter Merida for being part of my examining committee and Ms. Yuki Matsumura for her professionalism, effort and positive comments.

I would like to thank my colleagues at the Process Automation and Robotics Laboratory for creating a pleasant work environment and giving me the opportunity to join their team.

I am grateful to my parents and my family for their unconditional support and motivation throughout my studies. Graduate studies would have not been possible for me without their support.

Chapter 1

Introduction

1.1 Preliminary remarks

Global manufacturing and service industries are actively seeking technological solutions that enable them to quickly respond to expanding client demand and improved services. Hou emphasizes that agility has become the key to survival and success [1]. In order to take advantage of intelligent automated machinery and improve the quality and efficiency of services as well as integrate systems, utilizing intelligent information technology is inevitable. New technologies such as dynamic monitoring, fault diagnosis and self-maintaining and prognostic condition monitoring solutions are growing rapidly to meet the industry demand.

Earlier in 20th century, considerable efforts were made to improve the quality of manufacturing as well as reduce the product development cycle. Amongst the biggest contributors to the quality control field were Walter A. Shewhart and W. Edwards Deming, the pioneers of the modern quality control. Chisholm states that "Shewhart, a Bell Labs man, pioneered quality control and was a major inspiration to Deming (who met him at

Bell Labs)" [2]. Deming is well known in his own right for his contributions to such things as the manufacturing quality in post-war Japan and the foundations of Six-Sigma. Today continuous improvement and quality control remain one of the most critical deciding factors in the survival of a business. *Lean* and *Six Sigma* practices are the two general terms that are used to signify the importance of preserving value while reducing costs and resources (*Lean*); while improving quality to ensure the least possible amount of failures or defects (*Six Sigma*) in service or manufacturing processes. These two terms are interrelated in a sense that *Lean* manufacturing leads to shorter production cycles. This agility lead to more frequent results per time unit which can result in improved quality by analysis of causes of variation and process enhancements. (Please see Plate 1)

Plate1: Seven main categories of waste in *Lean* strategies:
- Rework
- Overproduction, *Muda* (Japanese term for "wasteful activity")
- Conveyance (moving of product within the shop)
- Queue time
- Inventory
- Movement (of people or equipment)
- Over processing

Source: BPM Process Excellence Network

The term *Lean* originated from the Toyota Production Systems in the 1990s. It refers to a practice that considers the reduction of wasteful procedures that does not add value to the manufacturing or service process [3]. The core concepts of *Lean* are experimental rather than statistical. The objective of *Lean* is to prevent conditions under which the value of time and the activities in which resources are engaged exceed the value of the product being manufactured within that time and by those resources [4].

Six Sigma has a stronger historical background. The name is derived from statistics and the probability theory that relates to the 3.4 defective parts per a million observations. It largely originated from Motorola in 1986. The focus in Six Sigma is mostly variation, it is said that variation is the evil of the process and therefore it is identified and attacked by some robust statistical analysis and series of techniques. Origins of these techniques for service industries mostly lie in continuous data, which are the quantities that are collected in terms of continuous variables such as time. The main objectives of Six Sigma are to achieve high performance capabilities by eliminating or eradicating variation [4].

Six Sigma practices aim to improve quality by implementing two distinct strategies; towards existing products and towards new products or processes. This research will concentrate on the existing process strategies which consist of five concepts, DMAIC. (Please see Plate 2) Management of quality improvement techniques, Statistical Quality Control, involves implementing a series of tools that enable the making of decisions regarding the status of products or processes to manufacture products. These tools are divided into three main categories:

> Plate 2: Six Sigma Process strategies for existing processes, DMAIC are acronyms for:
> - Define
> - Measure
> - Analyze
> - Improve
> - Control
>
> Source: BPM Process Excellence Network

1.1.1 Descriptive statistics:

Parameters of the process are indicators of the statistical distribution or performance indicators e.g. mean standard deviation, the sample Moving Range, etc.

1.1.2 Statistical process Control (SPC):

SPC is the practice of investigating random samples of the output product or a process variable in order to determine if the process is generating acceptable results.

1.1.3 Acceptance sampling:

Random sampling of a batch and generalizing the decision of pass or fail for a whole population based on the random inspection.

MacCarthy suggests that the applications of SPC have been spread from the primary domain (manufacturing businesses) to other so-called non-conventional fields. There has been growing emphasis on performance measurements, performance monitoring and benchmarking in many organizations [5]. An increase in the availability of analytics has enabled the industry to rely more on process statistics in decision-making and strategies. The capabilities of the tools and methods in process control are large contributors to the quality revolution. (Please see plate 3)

Plate 3: Research performed by the McKinsey Global Institute in June 2011 analyzed how leaps in technology are enabling countless data-mining possibilities. The challenge however, remains for companies to crunch the numbers so they can better service their customers. Business intelligence software is a solution that is gaining more popularity because of its ability to diagnose the trends and key performance indicators. Statistical concepts have enables engineers to decode the big data numbers into comprehendible information.

Source: Technology Review Magazine: Data analytics by the numbers, Tuesday, May 31, 2011

Importance of the service sector in many economies, suggests MacCarthy, is another factor of the rising trend of SPC applications in industry. Sharp surge in world population - 7billion as of 2011, and hence rising commodity and agricultural product prices have made SPC a popular tool to improve efficiency, increase the capabilities of the industry and to reduce the costs of production and transportation [6]. Based on International Monetary Fund analytics, the commodity food and metal prices have gone up by two-fold and five-fold respectively in the past 15 years, Figure 1.
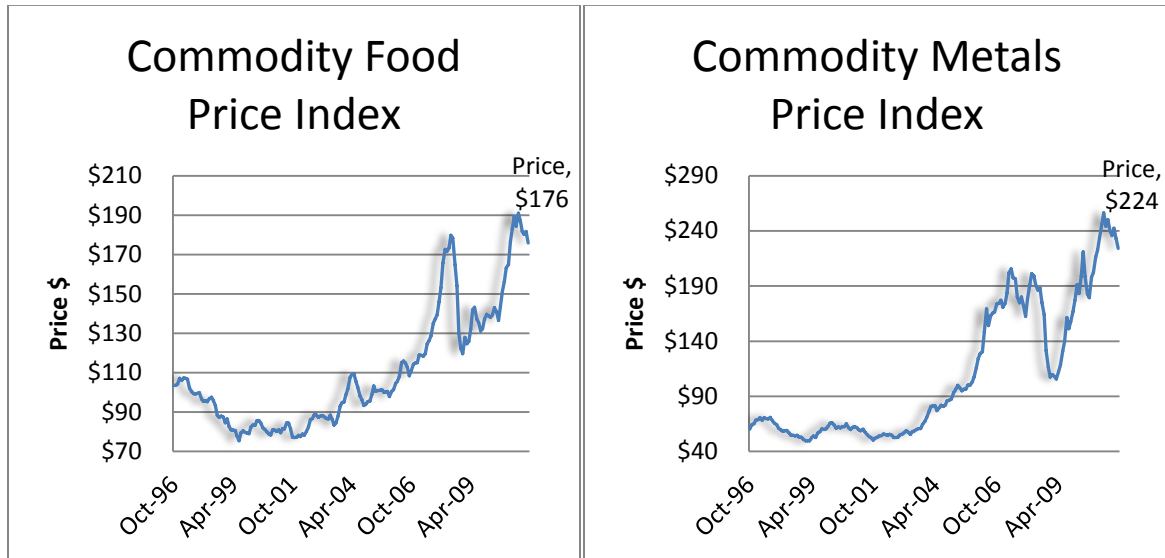
Figure 1: Food and metal commodity index prices [1996-2011]

The motivation of the work reported in this thesis has been the consideration of a process within a complex material handling facility. The process consists of the receiving, weighing, storing, and transporting bulk cargo. This work addresses some of the complexities that arise in the implementation of Statistical Process Control (SPC) in a non-conventional field. It proposes a monitoring system for weight measurement to improve the performance of the weighing process.

1.2 Material handling and challenges within bulk material weighing

Bulk cargo is a commodity that is transported unpackaged in large quantities, Figure 2. It is usually a mass of relatively small solid particles (e.g. ores, coal, oil sand, grains, etc.) that is generally poured or dropped into containers using a shovel bucket. Bulk solids are handled and transported within short distances on bulk material handling systems such as conveyor belts, stackers, re-claimers, bucket elevators, rail car dumpers, ship loaders or hoppers. Conveyors belts are stationary machinery that are typically composed of two or more pulleys, with a continuous loop of material, the belt, that circulates around them. One or more pulleys can be powered and the tension of the belt can be adjusted using the tensioners such as shown in Figure 3.

Figure 2: Potash is transported in bulk form

Raw material in many industries come in the form of bulk solids, therefore exact gravimetric assessment of the bulk material is crucial for product consistency and uniform

high quality services in industries such as plastic processing facilities, batch glass and ceramic processing plants. Trading, weighing, and storing and handling of bulk material are based on wet masses of cargos, where the insurance charges are base in value. Therefore both the buyer and the seller encounter some measure of risk. The magnitude of this risk is determined by the precision and the bias of the measuring equipment and the techniques that have been applied to weigh the wet weight [7]. Uncertainties in the weighing process make it very complex to translate the wet weight into dry weight and the dry weight into monetary value.
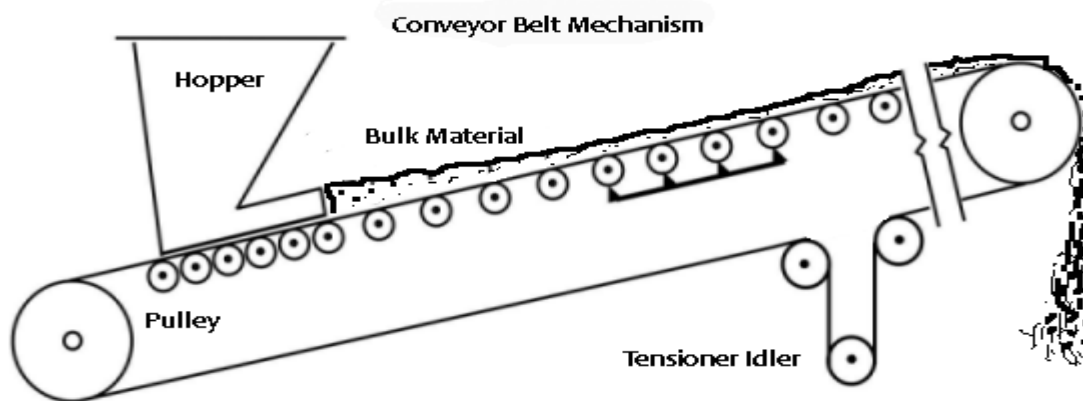
Figure 3: Conventional conveyor belt mechanism (Source: www.NISA.org)

1.3 Literature review

Shewhart suggested that the control may serve to firstly define the goal or standards for a process, secondly as an instrument to attain that goal and finally to review whether or not the goal has been reached [8]. Gibra researched quality control and the implementation of control charts from Second World War until 70's [8] and Vance, Schilling and Nelson

made many contributions on the effects of non-normality and the control charts published in late 70's[8, 9].

The principal domain for Quality Control and specifically Statistical Process Control (SPC) has been manufacturing. Therefore, conventional SPC research has been limited to manufacturing processes. Since the 1990s, more research has been conducted in the application of SPC in other domains. MacCarthy and Wasusri have reviewed and categorized the literature into four main fields [5]: Engineering, industrial and environmental applications, healthcare applications, general service sector, and statistical applications.

Among these four groups about 43 percent of the research has been categorized to belong to engineering and industrial applications. Literature has largely focused on statistics and fault detections criteria rather than the potential complications and challenges in implementation of these methods. One example of implementation complexity that is unique not only to each field but also individual application is the regulation of *Type I* and *Type II* errors (Explained in 1.4).

1.4 Implementation complexities

While working at Western Electric, Shewhart tried on several failed attempts to limit variations in the manufacturing process [10]. It was not until he moved to Bell laboratories that he realized a categorical difference between the types of variations, the common causes and the special causes. The common causes are inherent to the system, whereas the special causes are imposed on the system by operator related issues, mechanical

inaccuracies, etc. and they can be avoided. Application of SPC to non-conventional domains aims to detect and reduce the special causes of variations. A process that is statistically "in control" generates independently and normally distributed outputs [5, 11]. An "out of control" state can be detected when a point is plotted outside of the control limits. Possible "out of control" states are then detectable by the application of so called "run rules" such as the eight consecutive points above and below the process mean [5, 11]. A successful fault detection system will be able to pinpoint all simple (cycles, trends, shifts) or complex special causes. Complex patterns are combinations of two or more simple patterns and are called superimposed patterns. The superimposed patterns can be further divided into four subtypes [12] cycle, mixture, complex and multiple shift). However, there is a tradeoff in over sensitive fault detections systems—false alarm. False alarms are the states that are falsely detected as "out of control" where the process is really "in control". *Type I* error refers to such errors. On the other hand, *Type II* error occurs when the system is not sensitive enough to detect the "out of control" states where there is one. In the literature, several researchers have tried to devise new methods of intelligent fault detections using methods that are either Neural Networks-based pattern recognition or expert system-based pattern recognition. Wang and colleague used Neural Networks to recognize complex patterns and identify process noise. This method cannot easily be incorporated in material handling industry, since it does not provide a mathematical model for fault recognition [12, 13]. Methods based on probability, such as what Shewhart suggested (a series of simple pattern recognition rules that were developed within Western Electric in the 1990s), still remains one of the most practical tools to balance a monitoring system. Nelson used the Shewhart approach to develop "out of control" conditions for a process,

establishing a uniform and scientifically based interpretation for simple patterns. In developing run rules, that are called alarm Modes in this research, I have incorporated some concepts that were initially developed by Shewhart [13].

1.5 Research objectives

As a consequence of the issues outlined in previous sections, the related industry's operational demand has been to develop better control and monitoring means for the bulk handling systems with the focus of applications on conveyor belt s. The proposed research specifically focuses on:

- Analyzing the operation flow and creating a suitable error function model, by assessing the effects of the model on application, performance effectiveness and conformity to integrate on-line detection systems.

- Developing algorithms to identify, locate, and characterize sources of variation and trends in the scale network.

- Using an experimental setup that enables the research to implement the proposed algorithm and verify the results.

- Generating and developing computer software that analyses data and generates error detection signals based on the calculated controls limits.

1.6 Scope of the present work

This research was conducted to introduce a solution for error measurement and control of an industrial conveyor belt network. A non-conventional approach was taken to form the solution. The proposed method utilizes manufacturing Statistical Process Control

methods, and it is driven by industry analytics rather than computer simulated data. Previous studies have been done on conveyor belt error control and control chart pattern methods separately. This research is conducted to develop a method that combines previous research in these fields to address weighing variations in bulk material handling.

1.7 Organization of the thesis

This thesis has been prepared to address the outlined scope. The organization of this thesis is as follows:

Chapter 1 introduces basics of descriptive statistics and quality control methods in bulk material handling. Furthermore, chapter 1 discusses the complexities in monitoring and reducing the errors in metrology of conveyor belt scales as well as introducing the scope and objectives of this research.

Chapter 2 discusses core statistical concepts that are utilized to develop a methodology to identify the statistical limits, as well as natural and unnatural patterns in measurement error. Chapter 2 also introduces a system setup and different ways of bulk cargo measurement.

Chapter 3 reviews previous research in the field, discusses the major issues and ways to address them. As well, Chapter 3 provides a step-by-step explanation of the structure of the proposed method, followed by a detailed example of the process.

Chapter 4 evaluates the proposed method using four data samples "A", "B", "C" and "D" and identifying the faulty patterns. The results are then compared to conventional standards that are used in similar manufacturing industries.

Chapter 5 concludes and remarks on the findings and proposes ideas for future work to improving the implementation and further refining of the technique.

1.8 Notations and symbols

SPC has been used in several major domains; therefore industries in applied statistics have put forth the effort to standardize the terms and symbols of SPC, however different terms can still be shown by various symbols depending on the field of study. Table 1 lists the terms and symbols that are used in this thesis. The definitions for less frequently used variables are given within the context.

Table 1: Notations and sumbols used in this research

Variable	Description
n	Number of measurements (sample size)
k	Total number of samples
X_i	i^{th} data item form a number sequence
μ	Mean of the individual measurements
i	Indexing integer
$σ^2$	variance
σ	Standard deviation
UAL	Upper Action Limit
UWL	Upper Warning Limit
UOSL	Upper One Sigma Limit
CR	Center Line
LOSL	Lower One Sigma Limit
LWL	Lower Warning Limit
LAL	Lower Action Limit

Chapter 2

Descriptive statistics and system setup

2.1 Descriptive statistics

This chapter introduces a set of useful statistical concepts and describes the initial data exploration that has been performed on the data that will be used to verify the methodology. Subsequently, the system component setup and how they function and interact with each other is described in more detail.

In order to understand the applications of quality control it is essential to understand certain statistical theory. In weighing and measurement, statistics are used heavily to develop various tools and techniques. "A single estimate does not provide any information on variability of precision, but the absolute difference between two observations is already an effective measure for precision." [7] By using statistical tools such as mean and the sum of a set of measurements and their variances measures for precision such as the standard

deviation and the coefficient of variation can be used to develop confidence intervals and warning and action limits for future measurements.

2.1.1 The mean

In a process that outputs a desirable quantifiable parameter, the arithmetic average of the population for all the readings is called the distribution mean. In cases where sampling is required, the arithmetic average of the samples is called the sample mean which is an estimate of population mean. Therefore:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Where:

$$\bar{X} = the\ mean$$

$$X_i = observations\ i = 1, 2, 3, \dots, n$$

$$n = number\ of\ observations$$

2.1.2 Variance and types of variation

If one looks at a batch of rings in a manufacturing line, one will notice that the exact diameters of the rings are not the same. One can be 10.93 mm and the other can be 11.08 mm. This inconsistency in outcomes is called variation. In manufacturing processes variations can be caused by many factors e.g. differences in operator skill, differences in raw materials, weather conditions, tools, machines, etc. These causes of the variation are called random/common causes or chance causes of variation. In Quality Control, these variations are part of the process; even though they cannot be reduced to zero, it is

important that the range and the main sources are identified in order to recognize limits of these variations. [14].

Second type of variation that is observed is special causes or assignable cause of variation. Assignable causes are due to events that are not inherent in the system. Shewhart was greatly concerned with detecting and identifying these types of variations. The main goal of Quality Control is to reduce or remove the causes from the system through statistical methods. Chisholm calls this "statistically predictable behavior."[2]

The fundamental measure of variability in a set of data is variance. Variability in metrology is a quality that causes the measurements of the same character to differ in replications. And it is quantified by variance. Variance of a randomly distributed data is a measure of how far a set of data are spread apart from each other and it is calculated:

$$var(X_i) = \frac{\sum(X_i - \bar{X})^2}{n - 1}$$

Where:

$$var(X_i) = Variance\ of\ the\ sample$$

$$\bar{X} = the\ mean$$

$$X_i = observations\ i = 1, 2, 3, \dots, n$$

$$n = number\ of\ observations$$

2.1.3 The *Moving Range* and standard deviation

In the example from the ring manufacturing in the previous section, the amount of natural variation, range, between the smallest diameter (10.93 mm) and the largest diameter(11.08 mm) is 0.15 mm. *Moving range* is the difference between the two consecutive observations. *Moving Range* can carry significance by being an indicator of quality. If, in a process observation, the *Moving Range* surpasses a significant statistical limit, the process can be assumed to be "out of control." This is shown in a separate run chart beside regular control charts (introduced in section 2.1.5).

$$MR_i = (X_i - X_{i-1})$$

Where:

$$MR_i = Moving\ Range\ of\ the\ sample\ i$$

$$X_i = observations\ i = 1, 2, 3, \ldots, n$$

Another measure of variation is standard deviation. Standard deviation is calculated as follows:

$$\sigma = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n - 1}}$$

Where:

$$\sigma = Standard\ deviation\ of\ the\ sample$$

$$\bar{X} = the\ mean$$

$$X_i = observations\ i = 1, 2, 3, \ldots, n$$

$$n = number\ of\ observations$$

$$Var = \sigma^2$$

Small values of the *Moving Range* and the standard deviation are an indication that the population distribution is closely clustered around the mean. And similarly large values mean that the data set is widely spread.

2.1.4 Distribution of data

Distributions can be symmetric or skewed- with large or small variance, Figure 4
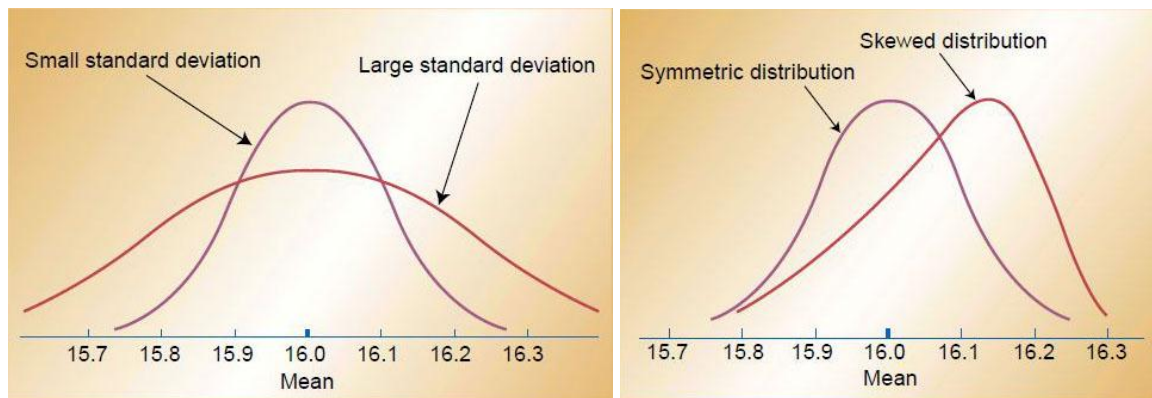
Figure 4: Distributions with different small and large/symmetric or skewed spread

2.1.5 Developing control charts

As discussed earlier, the aim of SPC is to differentiate between the random (common) variations and assignable causes and to identify, investigate and eliminate special variation [14]. This way, the process can be studied to reduce the random variations [15].
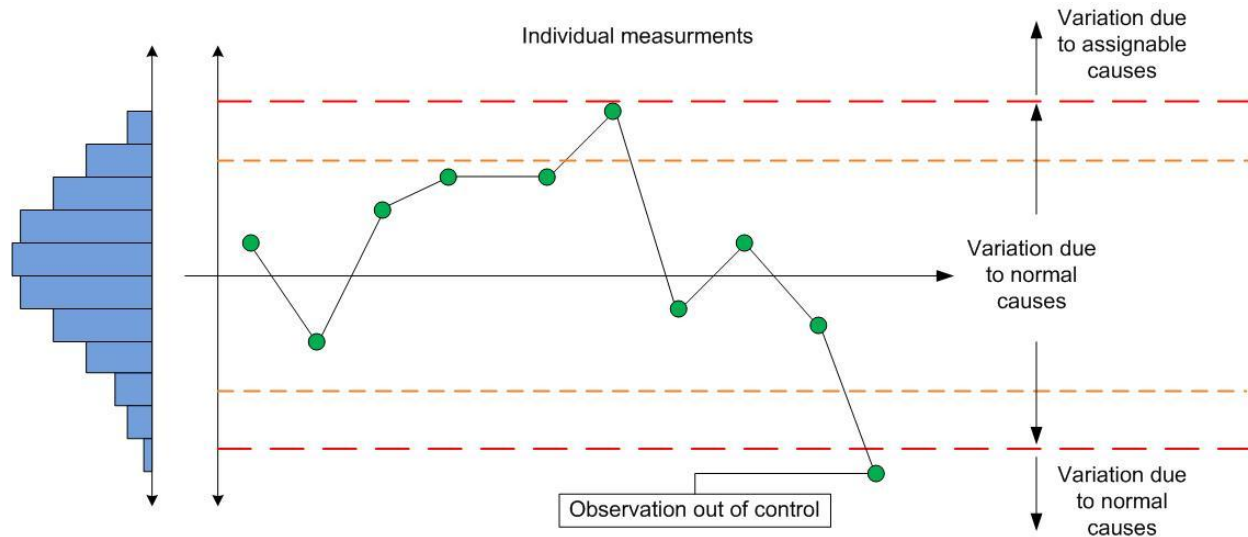
Figure 5: Conventional format of Control Chart proposed by Shewhart

In engineering practices this is usually done by targeting a quantifiable variable that is an indicator of a process characteristic or a desirable outcome. The performance of such a variable is plotted in a run chart as individual readings. An example is shown in Figure 5. The chart area is limited to the control lines and the lines are signified by the previous behavior of the process. It is said that the process is "out of control" if a reading falls outside of the lines. In the control chart in Figure 5 the round dots are the quantifiable readings, the horizontal axis shows the center line and the vertical axis represents the quality characteristics of the variable. The outer dashed lines are the Upper Control Limit (UCL) and the Lower Control Limit (LCL). The UCL is the maximum acceptable variation from the mean and the LCL is the minimum acceptable variation from the mean. The inner dashed line represents a tighter control limit that is usually applied to provide a means of a warning, Upper Warning Limit (UWL) and the Lower Warning Limit (LWL). This chart is

18

usually coupled with another chart called *Moving Range* chart for better performance. The

Moving Range chart records the *Moving Range* between the two consecutive observations similar to X_i control charts. The UCL and LCL are set at 3 standard deviations that capture 99.73% of the distribution assuming a normal distribution. The probability of readings falling within these limits is called the confidence limits. Confidence limits are generally used as 95.45% for 2 standard deviations and 99.73 % for 3 Sigma. (Please see plate 4)

Plate 4: Probabilities in normal distribution

- *About 68.27% of the values of a normally distributed data set lie within 1 standard deviation of the mean.*

- *Similarly, about 95.45% of the values lie within 2 standard deviations of the mean.*

- *Nearly all (99.73%) of the values lie within 3 standard deviations of the mean.*

Source: http://en.wikipedia.org
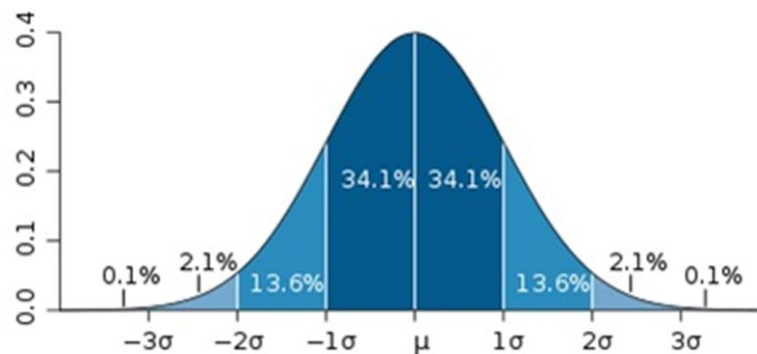
2.2 System setup (Gravimetric measurement for bulk material)

Gravimetric devices generally work in two main ways: differential and integral weighing. Differential weighing devices are most commonly used. They work by taking a

measurement and adding or subtracting another measurement then reporting the difference. All truck scales, rail car scales and most stationary scales work with this method. Integral weighing instruments rely on a simple integrating method to sum the conveyed material and calculate a cumulative result. Conveyor belts are integrating instruments that will be discussed in 2.3.2.

2.3 Weight measurement terminology and measurement concepts

Random variations in measurements will lead to risks of losing or winning by the same degree. By increasing the number of data points the random variable will cancel out to make the buying or the selling partner identically lucky or unlucky. However, the bias may accumulate and grow to show one of the parties has statistically incurred a loss. The random variations can stem from different sources; however, the systematic errors are caused by factors such as malfunctioning machinery or operator.

Uncertainty in weight measurements and weighing devices has been identified by different terms internationally. In some handbooks, the use of the term "error" is restricted to a bias or systematic error, while other sources have referred to "error" as "maximum permissible error", known as "tolerance" as a measure for random variations in a weight measurement technique [7]. In metrology and weighing instrumentations, bias and precision are fundamental concepts and a clear understanding of them facilitates the development of techniques to achieve them.

Accuracy: "A generic term implying closeness of agreement between a measurement and its unknown true value" [7]. Accuracy as a concept cannot be measured; nevertheless, the lack of accuracy can be measured as bias.

Bias: "A statistically significant difference between a single measurement, or a mean of a set, and the most reliable estimate of its unknown true value." Bias is a systematic error in weight measurement.

Precision: "A generic term that relates to the cumulative effect of random variations in a measurements system". Precision is a generic qualifier with no quantitative implications. The variance on weight measurements is a fundamental measure for precision. It also plays a key role in identifying the bias in weight measurements and the sensitivity of a statistical student's t-test.

2.3.1   Static weight measurements

Static scales still have a wide variety of usage in the industry and they are the simplest of all gravimetric equipment. Common industrial scales use strain gauges to measure the deflections. Deflections of the mechanical members then translate to calibrate weight readings using a set of four varying electrical resistances called a "Wheatstone bridge". While the static scales reduce the risk of mis-measurement to the lowest possible degree, they are not feasible in dynamic processing plants since the material has to be stopped for weight measurement.

2.3.2   Conveyor belt scales

This section will focus on how conveyor belt scales work and how the two factors of speed and instantaneous weight are used to determine the total weight conveyed.

Conveyor belt scales are integrating weighing devices. They use a simple integrating method to sum the conveyed material and calculate a cumulative result. A load cell is placed along the belt by replacing a number of the idlers by a series of idlers that sit within a floating frame, Figure 6.
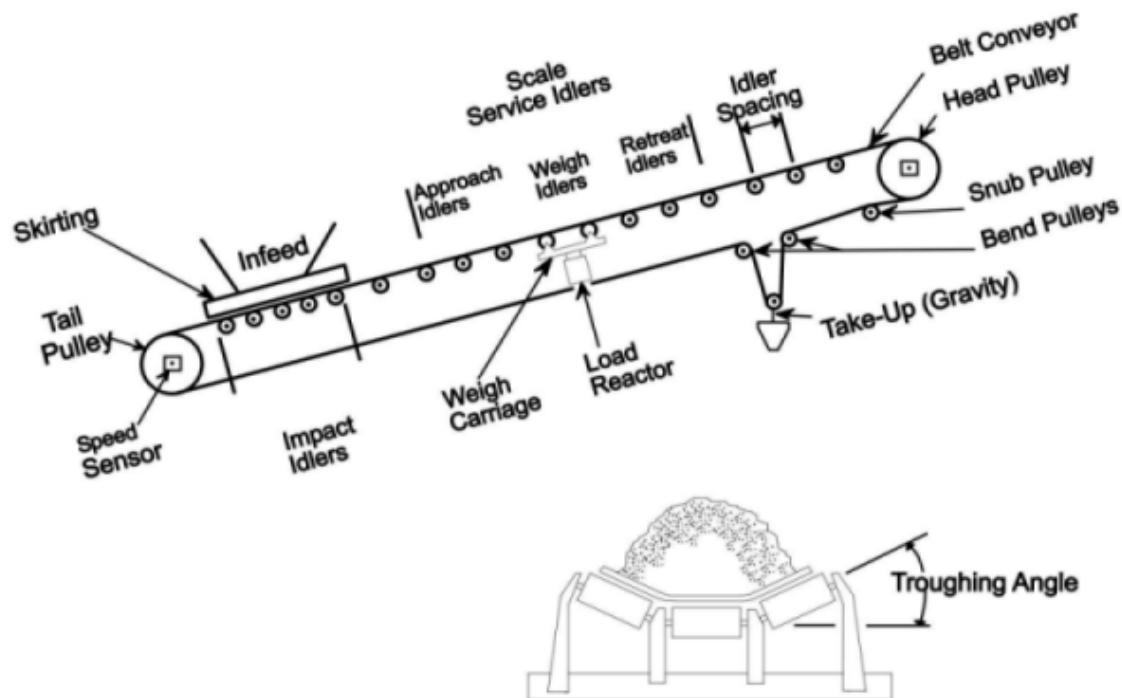
Figure 6: Schematic of a conveyor belt integrated with a scale (source: www.NISA.org)

The principle of operation is very simple. One of belt conveyor idlers' is resting on load reactor (weight scale); therefore signal from its sensor U_q corresponds to weight of bulk material/per idler span.

$$Uq = k_1 \times q$$

Where:

$$q = weight\ of\ bulk\ per\ idler\ span\ [kg/m]$$

$$k_1 = Transfer\ coefficient$$

Rotary speed sensor measures the angular velocity of the idlers that rotate as the belt moves. The velocity sensor produces a signal U_v corresponding to speed of the traveled belt.

$$U_v = k_2 \times V$$

Where:

$$V = liear\ velocity\ [m/s]$$

$$k_2 = Velocity\ coefficient$$

To calculate the instantaneous flow rate:

$$U_Q = k_1 \times k_2 \times q \times V$$

Where:

$$U_Q = instantanious\ flow\ rate\ [kg/s]$$

To calculate the total material passed through the scale where there is continuous flow of material instantaneous flow is integrated with respect to time. This is done by a (totalizer).

$$Q = k \int U_Q dt$$

Where:

$$Q = total\ conveyed\ material\ [kg]$$

$$k = k_1 \times k_2$$

The integrator or the totalizer receives the weight of the floating frame in high frequency. This is coincident with a tachometer pulse generator that is synchronized with the speed of the belt conveyor. The frequency of the weight measurements can be up to 4000 Hz for more accurate readings. This enables high resolutions by integrating the weight functions with regards to the distance traveled. In order to sum the measured weights and generate a total number, the movement of the belt can be considered in very small segments (i.e. 0.01 m). Weight sensors measure the weight of a small portion of the material that passes in a given time. The gross weight that the sensor shows is the sum of the belt weight, conveyor idlers and the material on the belt. The net weight of the material is gross weight minus the weight of the belt and the idlers.

A continuous weight traveling on a conveyor belt crosses through the weighing carriage as seen in Figure 6. The force function that it generates with the variable of distance or time is a trapezoid that starts from zero force and reaches a maximum of less than the full weight of whole continuous material and then back to zero. In other words, at any given position of the belt (or time) the weight of the material on the scale is the sum of all the particles on the weighing area [15]. In a special case where the length of the conveyed weight is the same as the idler span, weight function forms a triangle with a maximum of whole material weight. (Please see plate 5)

Total conveyed material for time t is calculated by:

$$Q = q \times v \times t$$

Where:

$$Q = conveyed\ material\ for\ time\ t\ [kg]$$

As for conveyors we have flow rate measured in tons per hour. Last expression could be written as:

$$Q = 3.6 \times q \times v \times t$$

Where:

$$Q = Conveyed\ material$$

$$for\ time\ t\ [metric\ tones]$$

Conventional scales use one of the following four methods to calculate the sum:

- Left-hand approximation: takes the initial of the two pulses.

- Right-hand approximation: takes the latter of the two pulses.

- Trapezoid approximation: forms a trapezoid between the values of the two pulses.

- Simpson's rule approximation: takes the first, second and the middle values and fits a parabolic curve to approximate the area under the slice of the material. The

Plate 5: Belt scale function for a weight length equal to scale span

Triangular weight function
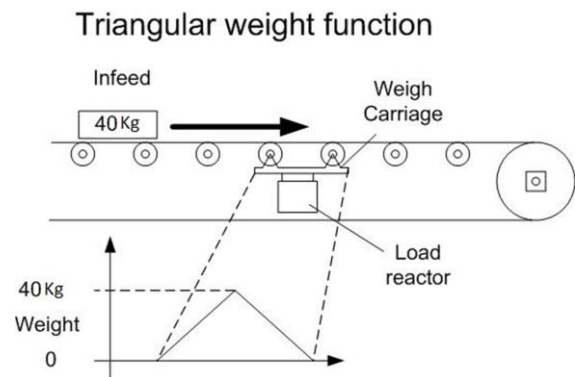
Total weight summation

25
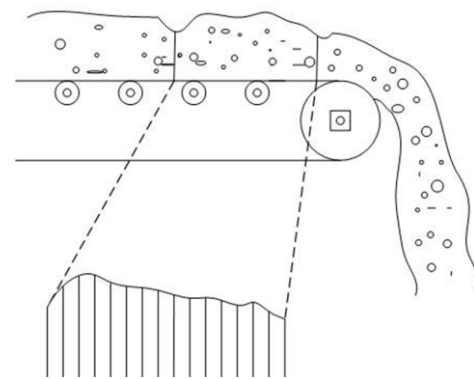
Simpson rule approximation is the most accurate of the four methods which is used in new generations of the scales [15].

2.3.3 Draft surveys (Displacement measurement)

An alternative weighing method of bulk cargos is draft surveys. This method is rooted in Archimedes' Law. The Archimedes law maintains that the mass of a floating vessel is equal to the mass of the fluid it dispenses. (Please see plate 6) Where the means are available (usually when the destination is a cargo vessel) it is possible to use the draft surveying method to verify the total transported weight. This method, due to a large number of variables and the limitations of application, is one of the least practical methods used.

Plate 6: Weight of a floating body
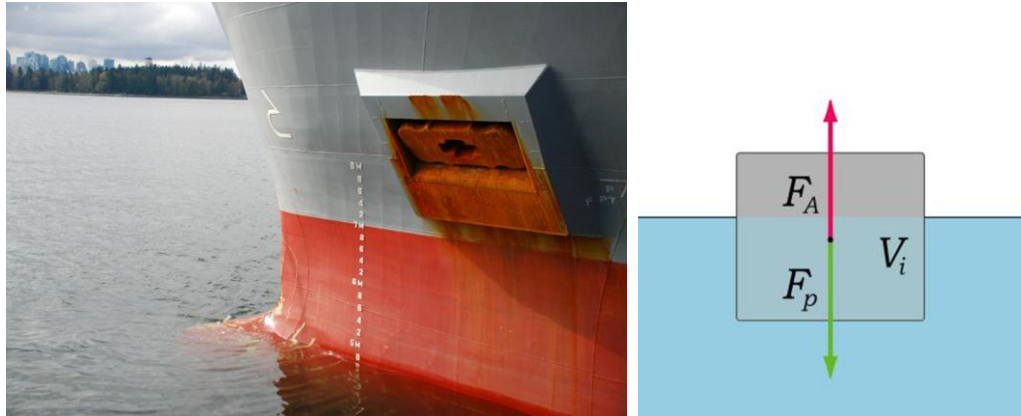
Weight of the displaced volume =

(Ship + bunkers + stores + consumables + cargo)

Draft surveying is used in ports as estimation, since it can lack precision due to a wide range of external conditions and unknown variables. Errors can occur in:

- Difference between draft table (based on drawings) and built vessel
- Observation error due to physical geographical limitations e.g. waves
- Accuracy limitations, draft marks are normally 0.10 m
- Water density estimate
- Complexity in determining bunkers, stores, etc.

Source: www.bulk-online.com

Chapter 3

Application of statistical process control in scale monitoring

As discussed in Chapter 2, the weighing methodology of the studied process revolves around a load sensor with an input and an output stream. The weight imbalance between the two readings at various sections of the network creates a difference between the weight of the material received and the weight of the material processed or shipped. These differences are used to analyze the performance of the individual components of the weighing network using SPC. Figure 7 demonstrates a schematic of the setup that has been used for this research. Each scale unit is shown in upper case letter and the stream of material is shown by arrows.
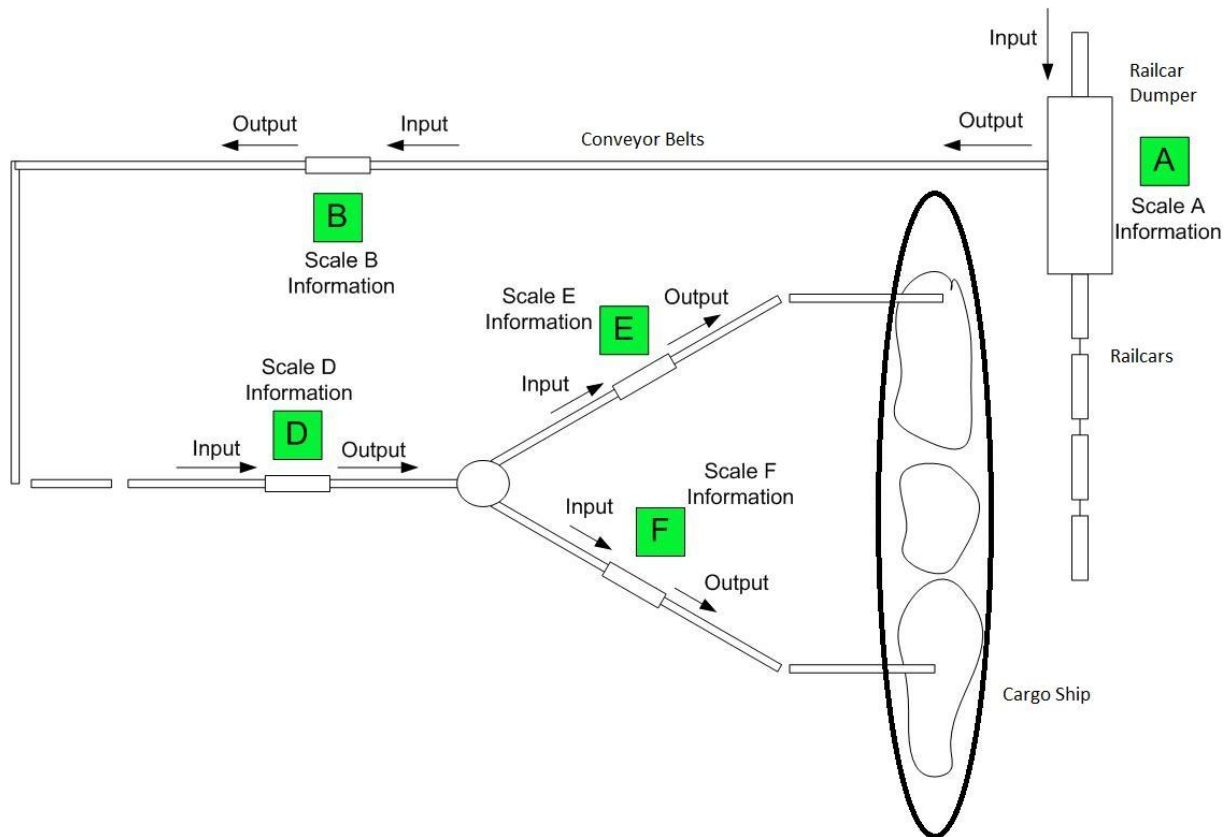
Figure 7: Calculation of weight imbalance between different stages of the process

This chapter concentrates on conditions of previous work on SPC tools, and the methods of pattern recognition in control charts. The following are discussed:

- Selection methodology based on the targeted non-standard application

- Categorical methods of detecting patterns in the literature and the integration of the proposed method

- Step by step structure of the modules in the proposed method

- Introduction to the capabilities and limitations of the proposed method

- Illustration of an example from the process

3.1 Selection methodology

Much of the research on control charts has been categorically based on either *neural networks* or *expert systems* [16, 17]. Both methods have been able to demonstrate advantages and disadvantages for many applications. The *neural networks* approach is successful in detecting highly complex and superimposed patterns within the control charts as well as predicting the starting points of the unnatural patterns; however it is very cumbersome to provide mathematical formulations for complex networks. The *expert systems* approach relies on statistics and probability theory; therefore it is able to predict the faulty behavior through a system by assigning a probability of success vs. failure. The *expert systems* approach works very well with single patterns and has been used widely in the manufacturing industry due to its reliability. In this report the research has been based on *expert systems* that were initially further explored and standardized by Nelson [9].

Here, based on of Nelson's ideas, the *expert system* approach has been chosen for its simplicity. Wang believes that Nelson's work provides a uniform and scientific basis for the considerations of more complex patterns [12]. He divides this category into two subcategories in the sense of pattern recognition. The first approach, which was also adopted by the Western Electric Company, is defined by using descriptive words such as trends, cycles etc. The second approach involves a plot of observation points and limits on a control chart. The control chart is then divided into several zones with a span of 1 standard deviation of distribution. In some research these zones have been termed zones A, B, C, C, B, A as shown in the Figure 8 [18].
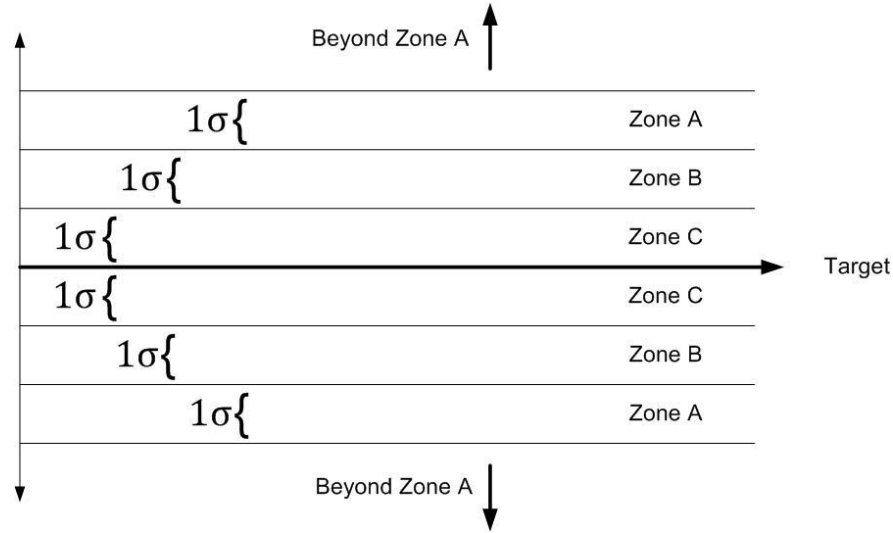
Figure 8: Control chart is devided into 1σ band zones

Anjard suggests an X_i chart along with a *Moving Range* (MR) chart for processes that only utilize one piece of data at a given time for experiments (as compared to a sample) [16]. Since the outcome of a weight imbalance between the input and the output of a conveyor belt scale produces a single percentage error reading per shipment, the X_i-individual chart along with the MR chart have been used for this research. The combination of the Xi-individual chart and MR chart is called an X_i/MR chart for short. In a non-standard application like the one in this research, it would be appropriate to use X_i/MR charts. Stapenhurst supported the work done by Wheeler that the X_i/MR chart could be a fall back option if one is not sure which chart would be more effective[19, 20].

Furthermore, as Anjard reported, the pattern directions have not been explicitly specified in the literature, however this work has includes the direction of the patterns for added clarification in fault detection [16].

MacCarthy has simplified a chart selection methodology in the form of an algorithm. He suggests that assuming lack of autocorrelation and normality are the two key issues in the selection of the X_i-individual chart. In the presence of autocorrelation however, there are ways to successfully detect "out of control" variations Zhang suggests. Using residual charts such as Cumulated Sum Cu Sum chart or Exponentially Weighted Moving Average, EWMA chart is one of the ways of dealing with autocorrelation [21]. Even though Zhang has studied residual control charts and proposed residual Exponentially Weighted Moving Average for Stationary charts (EWMAST), he, as do others, suggest that the detection capabilities of the X_i-individual's chart can exceed that of the residual charts on occasions [21, 22, 23].

3.2 X_i-individual and MR charts

X charts are known as the X_i individual charts. They are used when data sampling is not an option. Stapenhurst argues that they are less sensitive than conventional $\bar{X}/MR$ charts in identifying of "out of control" states. Furthermore, he suggests that although they are generally less robust to non-normality than $\bar{X}/MR$ charts, they have wider applications due to their ease of adjustability to different applications [20].

3.3 Patterns recognition tests

Does and Schriver amongst many others suggested the practicality of a protocol to be used for pattern recognitions in control charts that was initially advocated and widely used by the Western Electric Company, *WECO* for the quality control program. The *WECO* rules have been successful in detecting most common patterns of "out of control" states in

control charts [8] . The scientific basis of these rules lies in statistics and the probability theory of normal distribution. The chances of an independent event occurring between -3σ and +3σ in a normally distributed event is 99.73%. [24].

$$P(\mu - 3\sigma < X_i < \mu + 3\sigma) =$$

$$P(-3\sigma < X_i - \mu < +3\sigma) =$$

$$P(-3 < Z < +3) = 0.9973 = 99.73\%$$

Upper and Lower Control Limits are shown as UCL and LCL which correspond to probabilities of the far ends of normal distribution [8].

$$p = p_{UCL} = 1 - p_{LCL}$$

Therefore, the probability of an event outside the 3σ limit is:

$$(1 - 0.9973) = 0.0027 = 0.27\%$$

3.4 General structure of proposed method - Process modules

This section is a summary of all the steps taken to generate the results. The cycle starts with an inquiry of data over a particular time period for a specific scale or a network of scales.

Module 1

- Data acquisition and filtering from the scales network
- Comparison of the acquired data with reference value and generating percentage error value for each scale

- Generating of control limits and run charts to identify "out-of-control" states

MODULE 2

- SPC decision matrix

- Fault monitoring

- Pareto analysis

- Statistical capabilities

In the proposed method a user interface has been designed to serve two main purposes. First, is to refine the base of the analysis in a specific period of the past performance, and second is to apply alternative assisting tools for that specific period in order to ease the decision making process. The user interface control is responsible for signaling the data acquisition unit to obtain data from the required scales so that it is passed on to modules 1 and 2 for further processing. The generated results are then loaded back into the user interface from the output of module 2. See Figure 9 for the process diagram.

Module 1 explanation

3.4.1 Data acquisition and clean up

In this phase, each shipment is divided into segments that address the amount of material that passes through one scale during a certain time. This includes all the scales that are engaged in a particular shipment. The process time on each scale is defined by T_{zk}.

Where z is the indicator of the scale and k is the indicator of the time period, for $z = A, B, C, \ldots$ and $k = 1, 2, 3, \ldots, n$
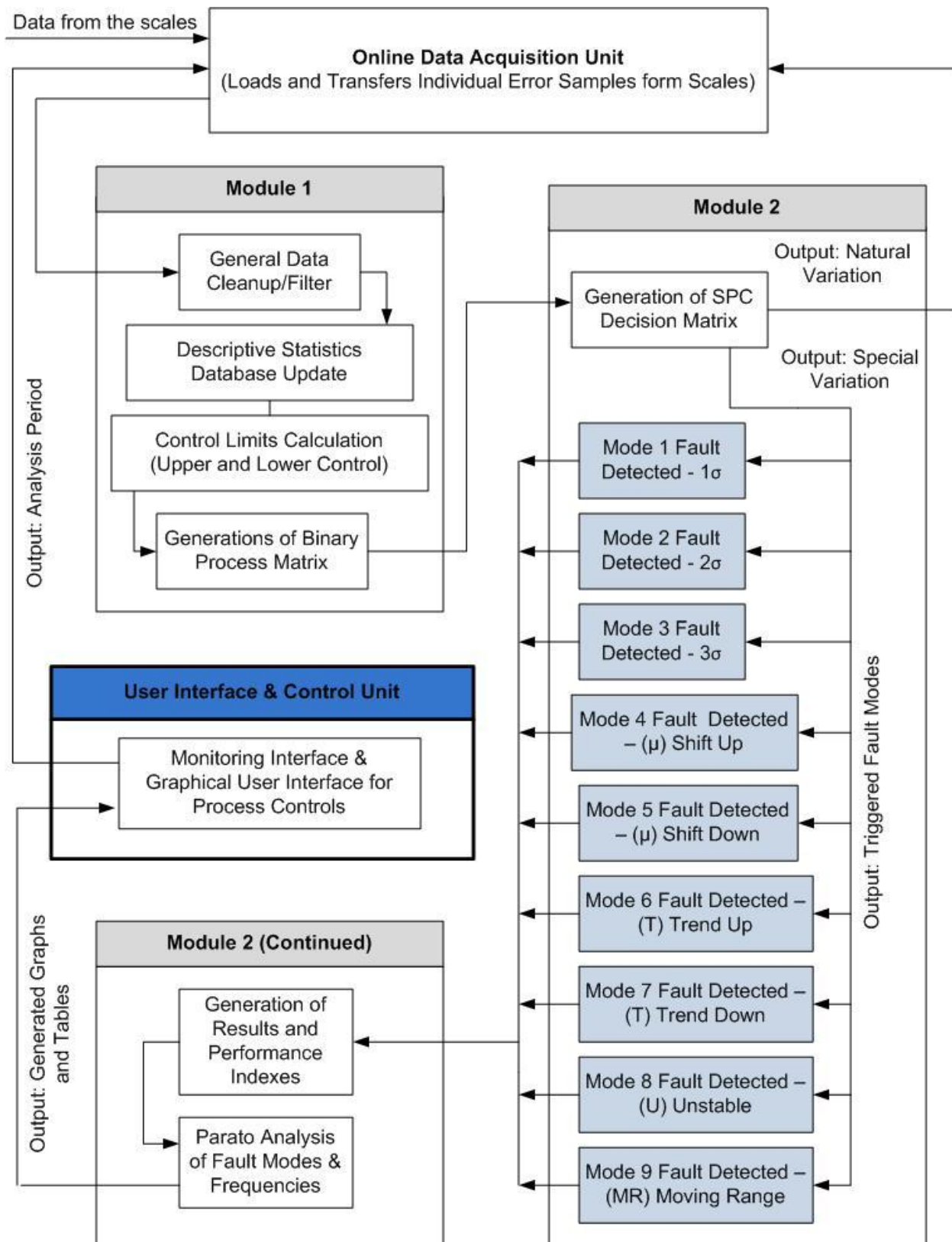
Figure 9: Monitoring flow chart [Starts from the User Interface]

3.4.2 Individual scale error values

As illustrated in Figure 10, the amount of material received from Scale "*A*" in period 1 is compared to the amount of material passed through Scale "*B*" in period 1 and the weight imbalance between the two observations is treated as a percentage error function:

$$X_{A1} = Cumulative\ weight\ from\ scale\ "A"\ during\ period\ 1$$

$$X_{B1} = Cumulative\ weight\ from\ scale\ "B"\ during\ period\ 1$$

$$Percentage\ Error = \frac{X_{A1} - X_{B1}}{X_{A1}} \times 100\%$$
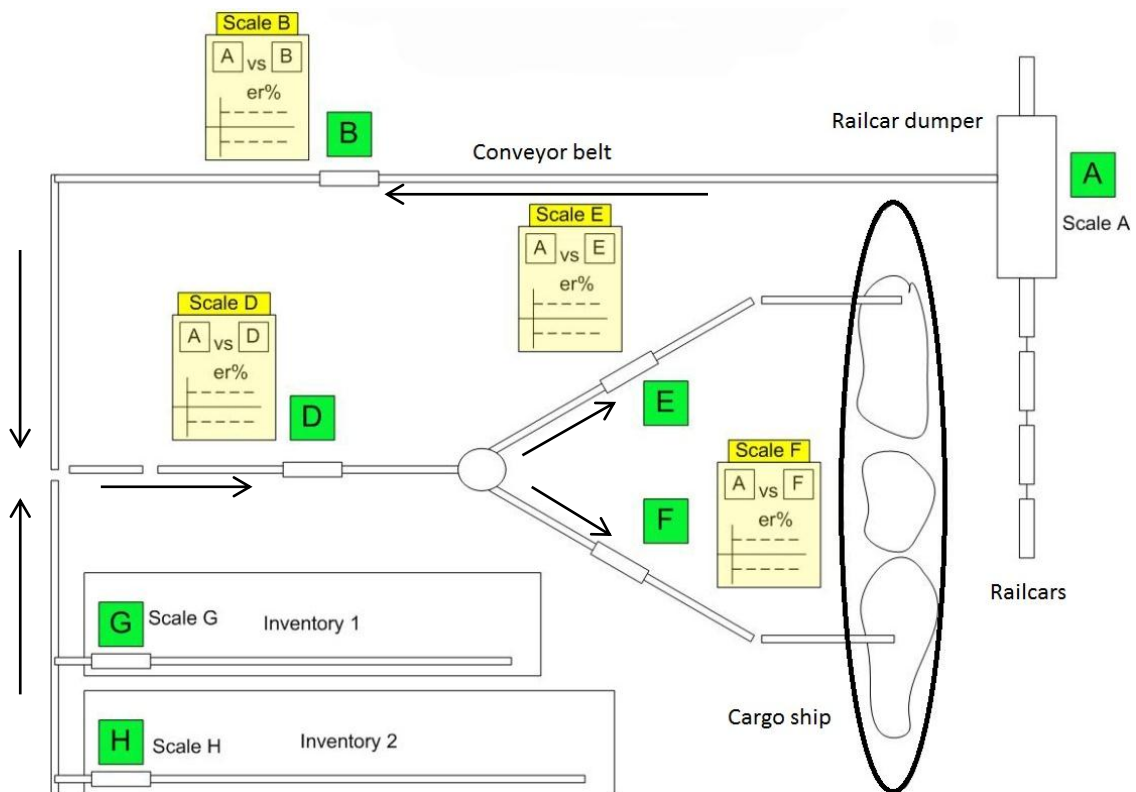
Figure 10: Individual scales generate control charts based on historical performance

Even though the reference values are obtained from very accurate static scales, the absolute accuracy of the reference readings is not the target in this phase. This is because

the absolute accuracy in reference values will affect the readings of all the engaged scales with same intensity; therefore, an individual scale patterns such as wear and tear, mechanical failure, operator related issues over time will be visible regardless of absolute inaccuracy of the reference value in a single reading.

3.4.3 Run charts and identifying "out-of-control" states

The scale data collected from each of these stages includes series of error measurements that correspond to the performance of an individual scale during a period of time in a shipment process.

Let X_i, X_{i+1}, X_{i+2}, ... , X_n be independent normally distributed observations with a mean of μ and variance of σ^2 for i=1, 2, ... , n. In high volume manufacturing processes, observations are sampled or sub-grouped to calculate statistical parameters such as $\bar{X}$ and R.

Where:

$$\bar{X} = \frac{\sum X_i}{n}, \qquad Sample\ mean$$

$R = sample\ range$ (difference between max. value and min. value in one sample)

In cases where sampling is required different types of control charts are used i.e. $\bar{X}/R$. Occasionally observations are taken in such ways that sampling or subgrouping is not possible. In this case the individual measurements are considered as their own sample or subgroup of size 1. In which case short term variability is measured using a parameter, *Moving Range*; *MR* is the absolute difference between the successive observations. These moving ranges are treated like ranges from subgroups of size 2.

$$MR_i = |X_i - X_{i-1}|, \text{ for } i = 1, 2, 3, \ldots, n$$

Where:

$$MR_i = Moving\ Range\ of\ the\ observation\ i$$

$$X_i = observations\ i = 1, 2, 3, \ldots, n$$

The average *Moving Range* (MR) is calculated as:

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

Where:

$$\overline{MR} = Average\ Moving\ Range$$

The estimate of standard deviation is calculated based on the *Moving Range* using Hartley's constant (for sample size 2, d_2 = 1.128) [20].

$$S_{MR} = \frac{\overline{MR}}{1.128}$$

X_i-*Individual* chart:

Plot: X_i for $i = 1, 2, 3, \ldots, n$. X_i-*Individual* control limits can be calculated as follows:

$$Upper\ Action\ Limit, UAL_x = \bar{X} + (3 \times S_{MR})$$

$$Upper\ Warning\ Limit, UWL_x = \bar{X} + (2 \times S_{MR})$$

$$Lower\ Warning\ Limit, LWL_x = \bar{X} - (2 \times S_{MR})$$

$$Lower\ Action\ Limit, LAL_x = \bar{X} - (3 \times S_{MR})$$

***Moving Range, MR* chart:**

Plot: MR_i for $i = 1, 2, 3, \dots, n$. The MR chart there is only one upper limit which is calculated as:

$$Upper\ Action\ Limit, UAL_{\overline{MR}} = 3.267 \times \overline{MR}$$

$$Center\ Line, CL_{\overline{MR}} = \overline{MR}$$

There are differences between British and American conventions for drawing and intercepting control charts, Stapenhurst acknowledges [20]. The British systems suggest that:

Action limits at ± 3.09 standard deviation from the mean

Warning limits at ± 1.96 standard deviations from the mean

These limits correspond to 2 in 1000 probability of an observation outside the action limits by chance, and 2 in 40 probability of an observation falling outside the warning limits by chance given that the distribution is normal. These limits for American systems suggest that:

Action limits at ± 3 standard deviation from the mean

Warning limits at ± 2 standard deviations from the mean

This corresponds to 3 in 1000 probability of "out of control" states by chance. Regardless of these limits, Stapenhurst suggests, the success of control charts have not been only because they follow theoretical models but because they work in practice [20].

3.4.4 Possible causes of control chart signals

As discussed in section 3.1, a set of rules were introduced in the Western Electric Company's statistical Quality Control handbook to detect unnatural patterns in a statistical process [25]. In a system compelling to six sigma standards these rules are based on events whose probability of accruing is 0.27%. In cases where random sampling proves higher probability than the limit, an alarm is triggered. These rules are known as Western Electric Company, *WECO* rules [8, 18].

In processing the X_i values, a matrix including the specifications of each measurement (negative/positive readings, the magnitude of error, and the placement of the reading with respect to the zones A, B, C, C, B, A) is generated, Binary Process Matrix, BPM. Patterns of irregular observations are then detected for each point on the plot. If a faulty behavior is detected up to the current observation, an alarm mode is recorded in a separate matrix, SPC Decision Matrix which will be discussed in module 2. The proposed model is capable of detecting nine different patterns of "out of control" states that are identified as "fault modes".

Module 2 explanation

3.4.5 SPC decision matrix

SPC Decision Matrix consists of 9 rows representing the "out of control" modes and *n* columns representing the number of X_i data points for each scale. Each element of the matrix corresponds to an alarm mode for a specific data point.

For example, in matrix $A_{jxi,}$ (*i*) is observation index number and (*j*) is the "out of control" mode. The matrix A_{9x4} shows that in the observations $X_{i+1,}$ and $X_{i+2,}$ Mode6 alarm is

triggered. In observation X_{i+2} Mode2 alarm is also triggered and in observation $X_{i+3,}$ Mode2, Mode3 and Mode9 alarms are triggered. As shown in Table 2 this is interpreted as:

For X_{i+1}: Six consecutive readings have been trending up. X_{i+2}: the uptrend is still present, but 2 out of three consecutive points are also beyond 2 sigma level. X_{i+3}: observation has fallen out of 3 sigma level as well as the moving range value is out of MR control level.

3.4.6 Fault monitoring rules:

As discussed before, control chart pattern recognition techniques are based on either a neural network approach, which are complex in their learning methods. Or they have been based on statistics and experimental probability theory, which are called *Expert System* methods in literature. The main challenge between the two methods is to use methods that offer a balanced ratio of *Type I* and *Type II* errors [26]. The proposed method defines 9 unnatural patterns signals as seen in Table 2:

Table 2: Decision matrix for four consecutive observations [X_i, X_{i+3}]

		N_i	N_{i+1}	N_{i+2}	N_{i+3}
Mode 1	1 σ Limit Triggered 4 out of 5 Times	0	0	0	0
Mode 2	2 σ Limit Triggered 2 out of 3 Times	0	0	1	1
Mode 3	3 σ Limit Triggered	0	0	0	1
Mode 4	8 Observations Shifted Up	0	0	0	0
Mode 5	8 Observations Shifted Down	0	0	0	0
Mode 6	6 Observations Trended Up	0	1	1	0
Mode 7	6 Observations Trended Down	0	0	0	0
Mode 8	14 Consecutive Points Alternating Around the Mean	0	0	0	0
Mode 9	*Moving Range* Limit Triggered	0	0	0	1

Mode1: In five consecutive observations at least four have fallen beyond 1σ limits (beyond zone C)

Mode2: In three consecutive observations at least two have fallen beyond 2σ limits (beyond zone B)

Mode 3: Current observation has fallen beyond 3σ limits (beyond zone A)

Mode4: Eight consecutive observations have all fallen above the mean (top side of the control chart)

Mode5: Eight consecutive observations have all fallen below the mean (bottom side of the control chart)

Mode6: Six consecutive observations have been trending up

Mode7: Six consecutive observations have been trending down

Mode8: 14 consecutive observations have alternated above and below the mean

Mode9: *Moving Range* between the current reading and the previous have fallen beyond the upper control limit

Several tools are used in statistical process control to facilitate decision making by pointing to suitable corrective direction. In this research Pareto analysis and statistical capability indices have been introduced to assist decision making and proposing methods to reduce causes of variation.

3.4.7 Pareto analysis

Control charts can still offer hints and clues that can direct management into ways of efficiently improving the process. Prato analysis is a tool that assists identifying the "vital few" from the "useful many" causes of variation [20].
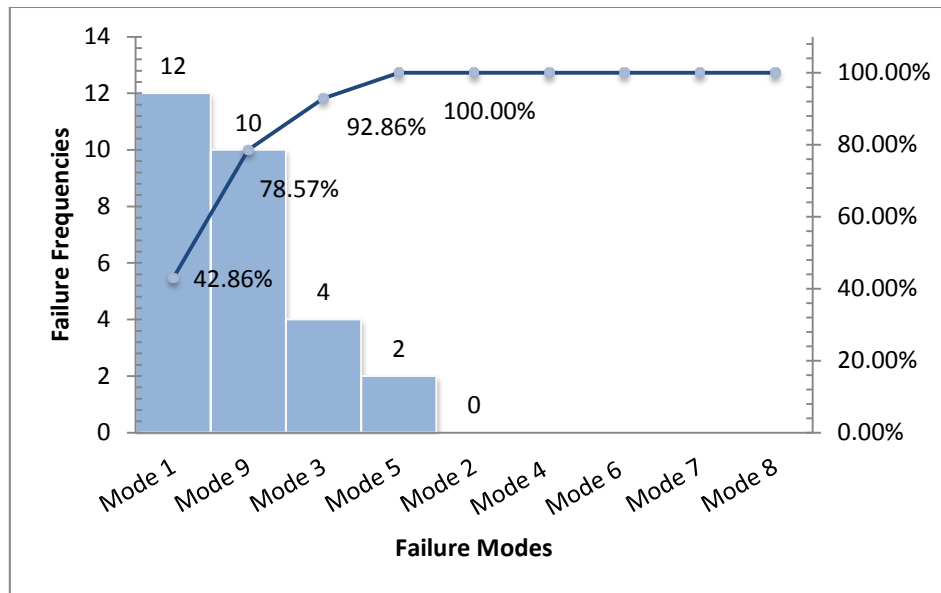
Figure 11: Standard Pareto chart that identifies "vital few" vs. "useful many"

Detecting variations within a process can be challenging, however the next step is to remove the variations causes. As shown in Figure 11 for a random set of data Mode 1 failure has been the most frequent cause of variation followed by Mode 9 and Mode3. Considering this reveals statistical details about a process and assists targeted corrective action for reducing variation. For example: in Figure 11, Mode 1 shows frequent observations that fall outside *zone "A"* (1σ) – refer to 3.1. As well interpreting from Mode 9 alarms, the process shows an unusually frequent number of "out-of-control" *Moving Range* values. This explains some of the Mode 3 alarms that are observations out of *zone "C"* (3σ). These assistive tools indicate a process that is calibrated considerably well due to low

number of Mode 4 and Mode 5 alarms, however random sources of variation are present. These can be caused be structural failures or poor maintenance scheduling.

3.4.8 Statistical capability indices

So far what has been discussed has been how well the process is performing based on descriptive statistics given previous performance. The faulty observations are detected based on the process statistical capability. Sometimes the process requires that certain guidelines or regulations are met. One of the metrics that can be used for this purpose is called process capabilities or *Six Sigma* capabilities when the control limits are selected $\pm 3\sigma$ for the control charts [27]. For an application such as the proposed concept, the ideal mean for the weight imbalance is to get a zero percentage difference between the two sources. This would ensure minimal scale error; however achieving zero error is practically not possible due to series of uncontrollable variables in real processes. Therefore certain levels of accuracy or permitted error are defined within industry to certify measuring devices. In case of conveyor scales the *National Institute of Standards and Technology* provides the error limits for bulk material conveyor scales to be within $\pm 0.5\%$ [28]. Process capability indices enable us to find out whether or not a process is following the instructed guideline. C_p is process capability index which is computed as the ratio of the specification width to the width of the process variability:

$$C_p = \frac{Specification\ width}{Process\ width} = \frac{USL - LSL}{6\sigma}$$

Where:

$$USL = is\ the\ Upper\ Specification\ Limit$$

$$LSL = is\ the\ Lower\ Specification\ Limit$$

Naturally, it is ideal to have a ratio of larger than one, meaning that the process is within the specified limits. C_{pk} is used in cases where the variability of the process is not centered on the specification range. Therefore it measures the capability of each half and takes the smaller of the two.

$$C_{pk} = min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right)$$

In systems where process capabilities are not centered around the mean of the specified capabilities, using C_{pk} ensures that capability index reflects a true measure of process capabilities by using the smallest of the upper and lower indices.

$C_{pk} > 1$, ratio larger than 1: The process variability is tighter than the specified capability.

$C_{pk} < 1$, ratio smaller than 1: The process variability is outside the process specifications.

3.5 Evaluation of the proposed method compared to conventional monitoring systems

Even though the application of the proposed method is specifically tailored to the field being studied, the data sets were later analyzed by conventional *WECO* run rule to compare the results. The *WECO* rules have a generic application. They are considered to be the most suitable generic pattern recognition in related unconventional fields. The conventional run rules are as follows:

Mode A) 8 or more successive observations above or below centerline

Mode B) 8 or more successive points alternating above and below centerline

Mode C) Sets of 5 observations with at least 4 beyond one standard deviation

Mode D) Sets of 3 observations with at least 2 beyond two standard deviations

As mentioned earlier in chapter 3, the proposed method suggests performance and fault diagnosis system using analytics that are obtained from cumulative weights of incoming rail cars. Rail car weights are accurately measured by static scales. The scales are evaluated in the network using reference car weights based on SPC criteria. Degradation patterns become evident once the performance of the scales are studied in long term and therefore improved maintenance and calibration practices are proposed based on data driven system. This ensures the reliability of the instruments that report the conveyed material through the network of hoppers, conveyor belts, surge bins, inventory containers etc. when there is no other redundancy reference is in place. A visual example of the proposed method for a single scale has been given in the following Plate. Chapter four verifies the proposed method using four series of data, where the results using proposed method are compared to the conventional methods of monitoring.

3.6 Illustration of a process example

Plate 7: Shows the steps of the monitoring from data acquisition, process through the two modules until generation of the results.

Phase 1: *Data acquisition from scales B, C and D (specifications are read from user interface)*
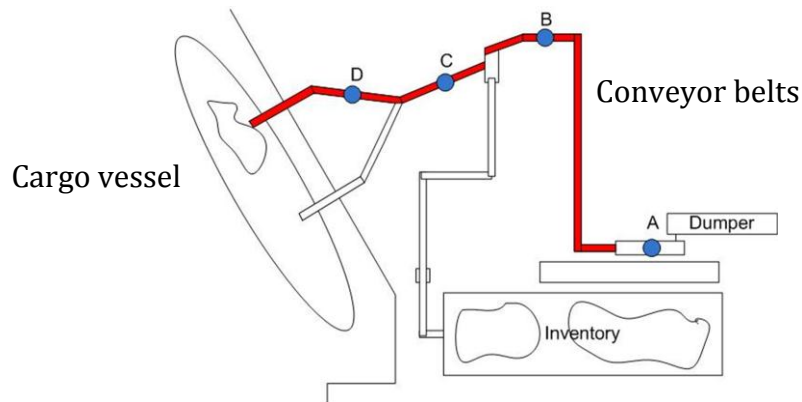
Module 1:

Phase 2: *Comparison of the acquired data with reference value (A) and generation of error values for each scale*

Phase 3: *Calculations of Control limits*

UAL = μ+3σ	0.0048835
UWL = μ+2σ	0.0037694
Mean (μ)	0.0015413
LWL = μ+-2σ	-0.0006869
LAL = μ+-3σ	-0.0018010

Plate 7 continued …

Phase 3: *Generation of run charts and Binary Process Matrix (BPM)*

Shewhart Process Control Chart

UAL= Mean+3σ

UWL= Mean+2σ

LWL=Mean-2σ

LAL=Mean-3σ
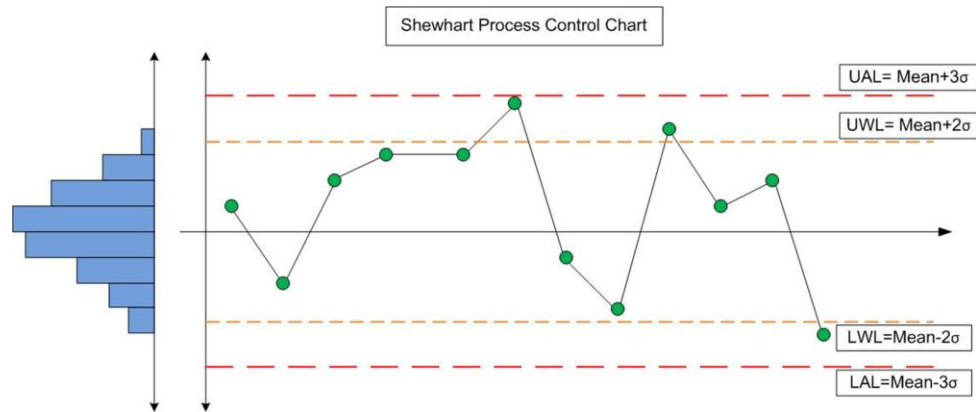
	X_i	X_{i+1}	X_{i+2}	X_{i+3}
MR sign (+~1)	0	1	1	1
Xi Sign (+~1)	1	1	1	1
1 sigma (0/1)	0	0	0	0
2 sigma (0/1)	0	0	1	1
3 sigma (0/1)	0	0	0	1

Module 2:

Phase 4: *Application of pattern recognition rules to detect "out of control" states and generation of SPC Decision Matrix*

		N_i	N_{i+1}	N_{i+2}	N_{i+3}
Mode 1	1 σ Limit Triggered 4 out of 5 Times	0	0	0	0
Mode 2	2 σ Limit Triggered 2 out of 3 Times	0	0	1	1
Mode 3	3 σ Limit Triggered	0	0	0	1
Mode 4	8 Readings Shifted Up	0	0	0	0
Mode 5	8 Readings Shifted Down	0	0	0	0
Mode 6	6 Readings Trend Up	0	1	1	0
Mode 7	6 Reading Trend Down	0	0	0	0
Mode 8	14 Consecutive Points Alternating	0	0	0	0
Mode 9	Moving Range Limit Triggered	0	0	0	1

Plate 7 continued ...

Phase 5*: Generating performance metrics—Pareto analysis & Statistical Capabilities*
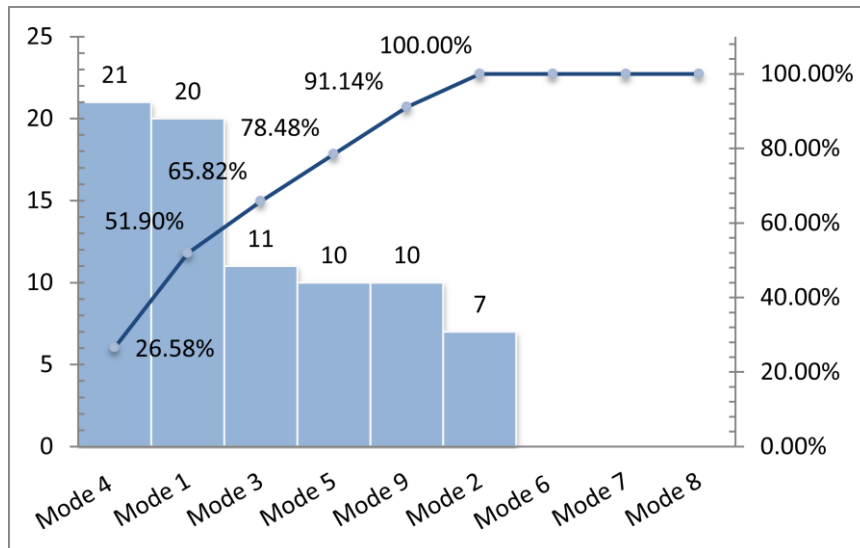
Process Statistical Capability Index	
The Actual Process Capability Cpk~1	0.8072
The Potential Process Capability Cp~1	0.9061
Failures Part Per Million (PPM~2700)	6528 PPM
Upper Specification Limit (USL)	0.005
Lower Specification Limit (LSL)	-0.005
Statistical Standard Deviation	0.0018394
The mean	0.0005455
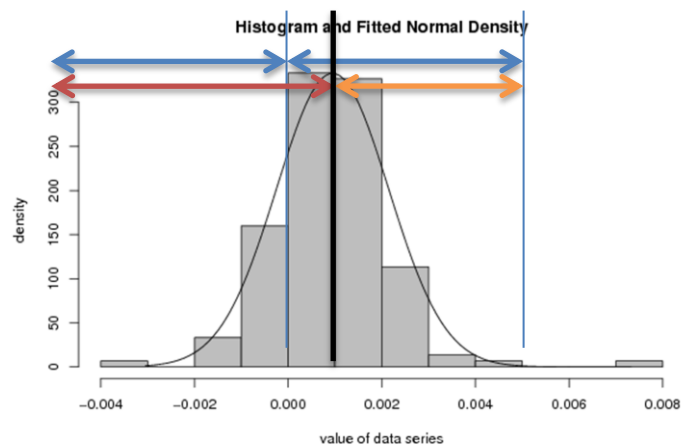
Phase 6: *Report results for individual scales*

Chapter 4

Evaluation of results using four data sets "A", "B", "C", and "D".

4.1 Data set conditioning

In this chapter, four sets of data samples that have been obtained from industrial scale observations are further considered. The data included dates and the error percentage measurements between each scale and the corresponding rail car weight. Initially, the data were analyzed to find out if they were fit to be examined by the proposed model, then the results were generated. To compare the performance of the proposed method the same data set was analyzed by a different model that uses conventional *WECO* rules. Observations are discussed after evaluating the model for each data set and briefly explained at the end of section in 4.8.2, for data set "A", 4.9.2, for data set "B", 4.10.2 for data set "C" and 4.11.2 for data set "D". Conclusive summery is presented in section 4.12.

Correlation between the sets has not been an issue for the purposes of this research since the factors that affect the four scales are common for all in the same way. The proposed method uses an individual observation control chart. Xi-individual/MR control charts require that the data are normally distributed and are not auto-correlated. Therefore, before validating the model the four data sets were tested for assumptions, normality and auto-correlation. Observations obtained from belt scales for each set did not exceed 6000 data points; therefore *Shapiro-Wilk* test for normality was used. (Please see plate 8). P-value for data sets was considerably smaller than the confidence level $\propto= 0.05$ therefore the hypothesis that they were not from a Normal or Gaussian distribution was rejected.

Plate 8: *Shapiro-Wilk* test for normality. The null hypothesis: the data is normally distributed. If p-value > or = α (significance level). It means that there is no evidence to reject null hypothesis. Otherwise the null hypothesis is rejected.

Central limit theorem in statistics and probability theory states that the sum of a large number of independent observations from the same distribution will have a normal distribution. In cases where normality is an issue, sampling can be performed in order to utilize the proposed method [29]. "A key theoretical result, called the central limit theorem, underpins many methods of analysis. It states that the means of random samples from any distribution will themselves have a normal distribution; As a consequence, when we have samples of hundreds of observations we can often ignore the original distribution of the data" [30].

4.2 Data exploration and application of SPC on scale data

In this section the data is analyzed for assumptions of normality and autocorrelation as well as overall distribution specifications such as distribution parameters. Initial data exploration is performed to identify normality using *Shapiro-Wilk* test. The data is also tested for autocorrelation with different lag factors. Normal distribution parameters were calculated using maximum likelihood estimation (MLE) method. Consequently a visual data exploration is performed to identify any visible outliers or unexpected data distribution. This includes four types of visual aids:

- Histogram
- Density plot
- Box plot
- Normal QQ plot

In sections 4.3, 4.4, 4.5, 4.6 initial data exploration for four data sets "A", "B", "C", and "D" have been discussed, respectively.

Table 4, and Figure 13 present the quartiles and visual data exploration of the data set "A" respectively. The distribution is normal with no significant outliers.

Table 7 and Figure 16 present the quartiles and visual data exploration of data set "B", respectively. Data set "B" is also normally distributed with the right skewness.

Table 10 and Figure 19 present the data "C" with one potential outlier which can be neglected and treated as an "out of control" state.

Table 13 and Figure 22 present the quartiles of data set "D". Data set "D" is normally distributed with 3 evident outliers. These will be treated as "out of control" states since

there are three close values that could have been caused by a mechanical "out of control" a lost state or a mis-calibration. It is important to identify outliers and "out of control" states. The outliers are considered to be single significantly large or small numbers, whereas "out of control" states are large error readings that have been caused by an unnatural cause.

4.3 Data exploration of set "A"

Based on the *Shapiro-Wilk* test we cannot reject the idea that readings from data set "A" come from a normal distribution with 95% confidence. The mean is $\mu= 0.000545474$ and the standard deviation is $\sigma= 0.002484827$.
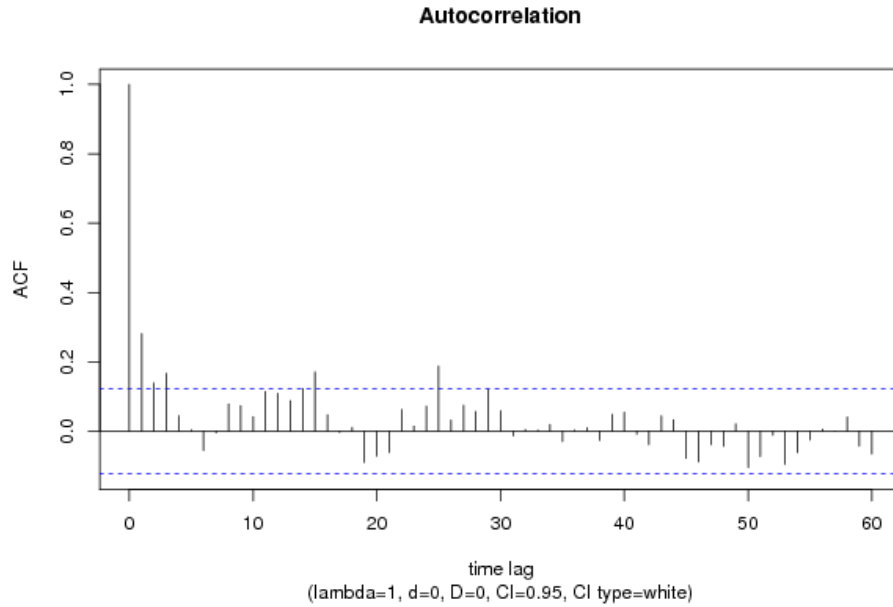
Figure 12: Plot of auto-correlaton for successive observations in set "A"

Table 3: Auto-correlation coefficient for lags of [1-3]

Autocorrelation Function			
Time lag k	ACF(k)	T-STAT	P-value
1	0.281253	4.4736	6e-06
2	0.138939	2.21	0.014002
3	0.16662	2.6503	0.004275

The coefficient of auto-correlation in Figure 12 and Table 3 has been calculated to be at most 0.28; therefore data-set "A" is not significantly auto-correlated.

4.3.1 Application of SPC on data set "A"

Table 4: Exploration of data set "A"

Date	Data point number		*Xi*
30-Dec:	6	Min. :	-0.0108447
12-Jul:	5	1st Qu.:	-0.0008545
28-Mar:	5	Median :	0.0002428
12-Jun:	4	Mean :	0.0005455
29-Jun:	4	3rd Qu.:	0.0015315
1-Jul:	3	Max. :	0.0114670
Data points:	226		
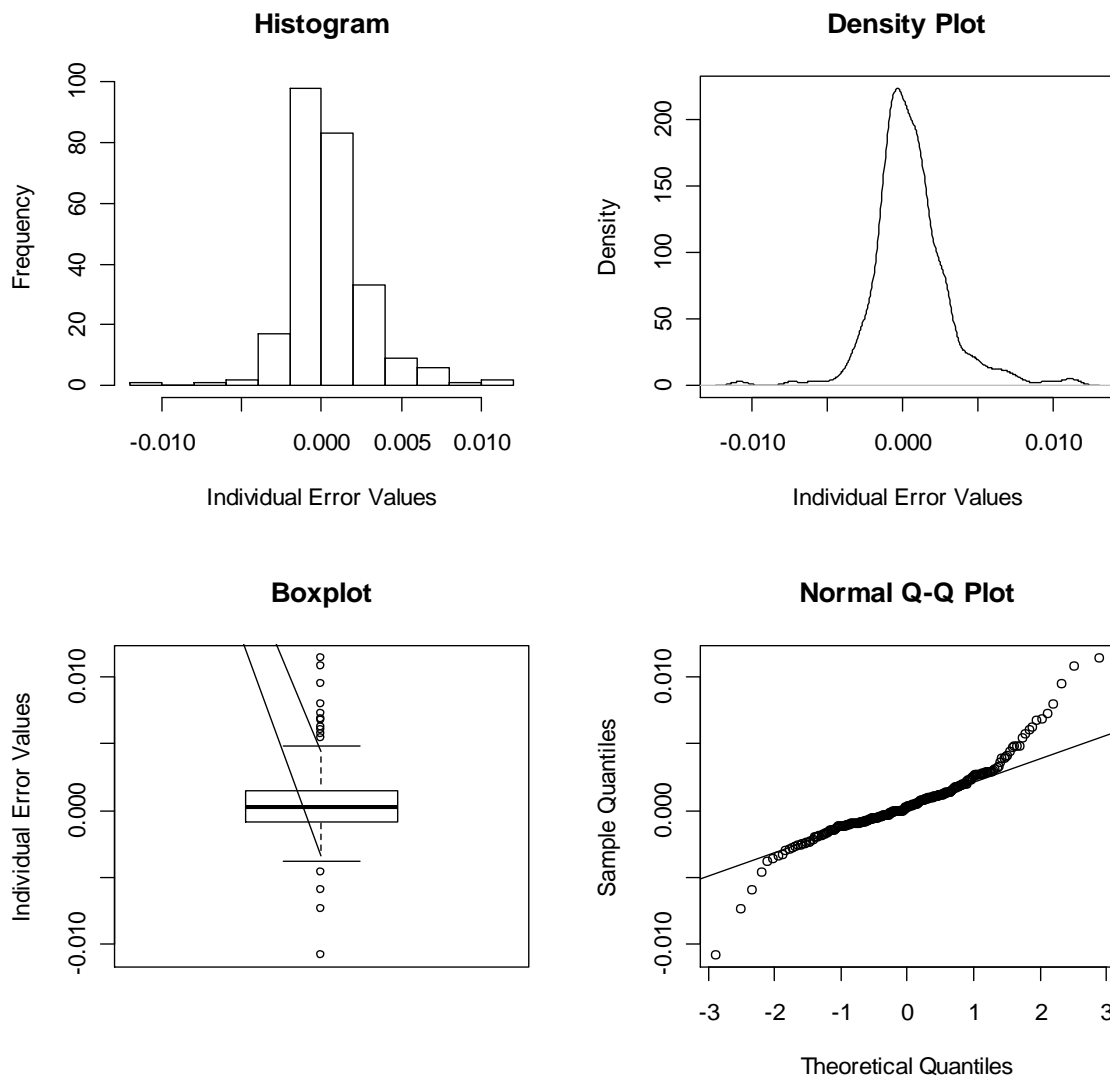
Figure 13: Visual data exploration of data set "A"

The statistical limits of data set "A" are calculated in Table 5 and Figure 14.

Table 5: SPC limits for data set "A"

Control limits (Set A)		
Upper Action Limit	UAL = µ+3σ	0.0060638
Upper Warning Limit	UWL = µ+2σ	0.0042243
Mean Centre Line	Mean	0.0005455
Lower Warning Limit	LWL = µ+-2σ	-0.0031334
Lower Action Limit	LAL = µ+-3σ	-0.0049728
Upper *Moving Range* Limit	UAL=MR × 3.27	0.0067849

Figure 14: Control chart indicating some "out of control" states for data set "A"
Vertical axis values are unit-less percentage error measurments

4.4 Data exploration of set "B"

Based on the *Shapiro-Wilk* test we cannot reject the idea that readings from data set "B" come from a normal distribution with 95% confidence. The mean is μ= -0.00008774 (unit-less error percentage value) and the standard deviation is σ= 0.001678307.

Autocorrelation

time lag
(lambda=1, d=0, D=0, CI=0.95, CI type=white)
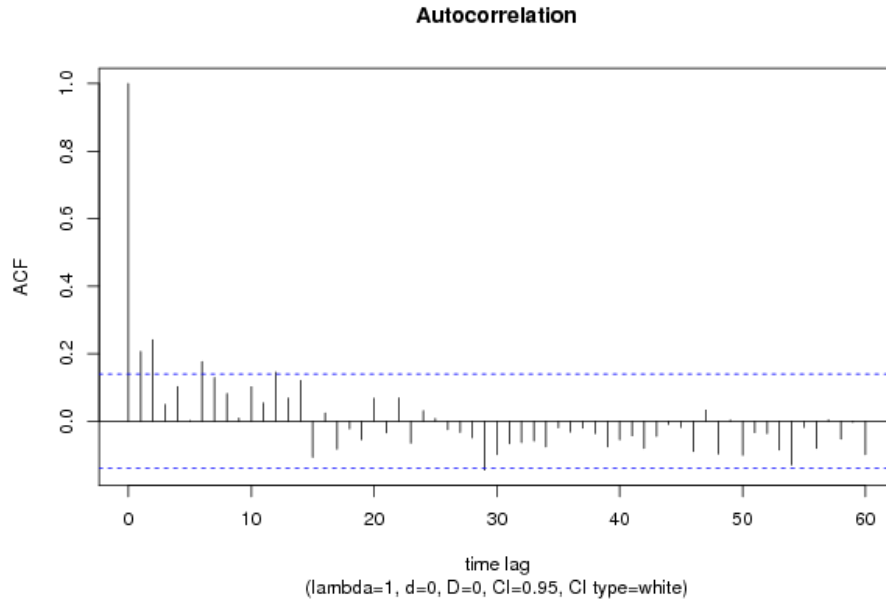
Figure 15: Plot of auto-correlaton for successive observations in set "B"

Table 6: Auto-correlation coefficient for lags of [1-3]

Autocorrelation Function

Time lag k	ACF(k)	T-STAT	P-value
1	0.206735	2.9017	0.002067
2	0.241623	3.3913	0.00042
3	0.049984	0.7016	0.241889

The coefficient of autocorrelation in Figure 15 and Table 6 has been calculated to be at most 0.24; therefore data-set "B" is not significantly auto-correlated either.

4.4.1 Application of SPC on data set "B"

Table 7: Exploration of data set "B"

Date	Data point number	Xi	
27-Nov:	4	Min. :	-0.0056220
30-Aug:	4	1st Qu.:	-0.0008688
31-Dec:	4	Median :	-0.0000614
14-May:	3	Mean :	-0.0000877
15-Jan:	3	3rd Qu.:	0.0007082
21-May:	3	Max. :	0.0099830
Data points :	176		
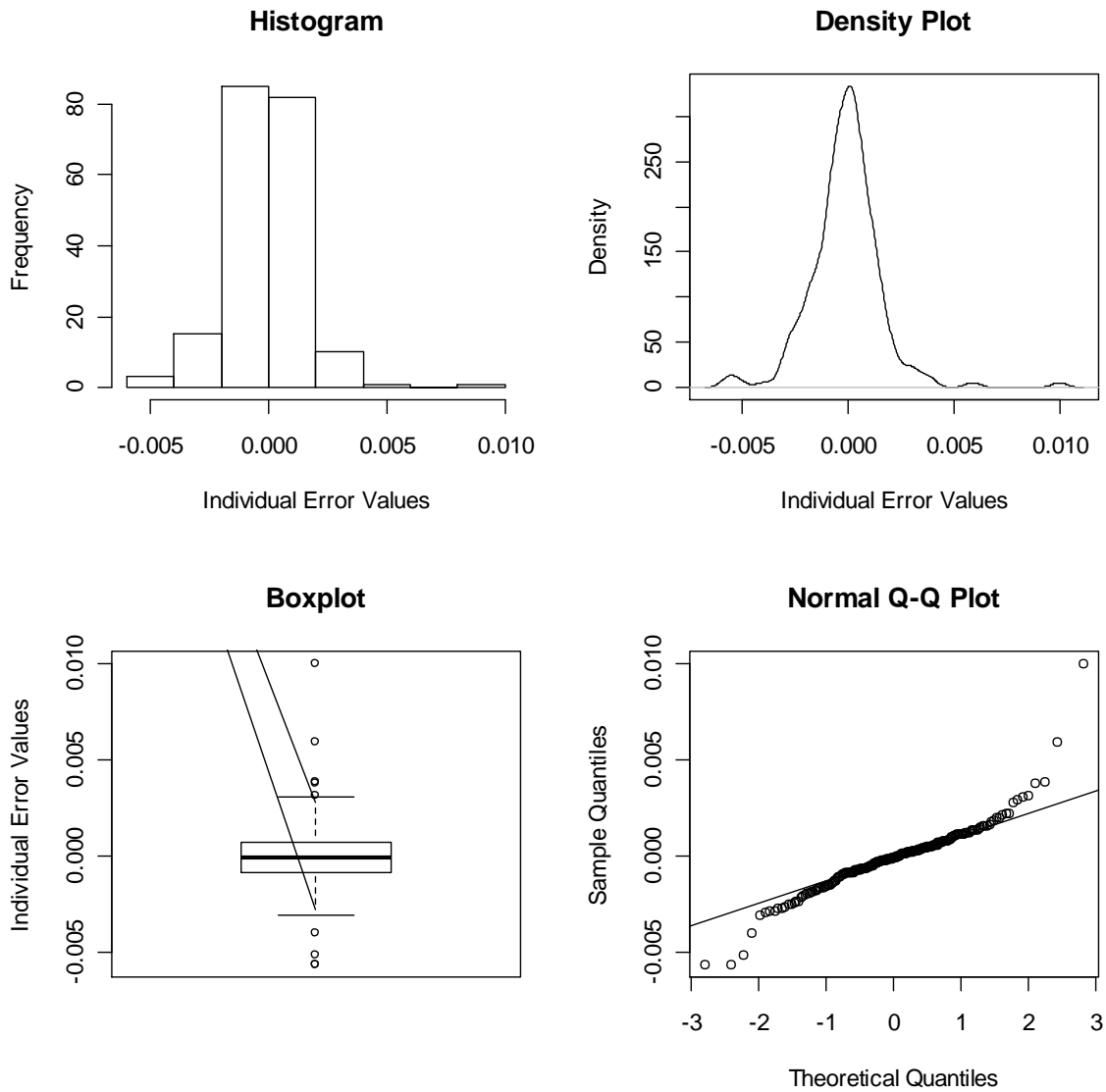
Figure 16: Visual data exploration of data set "B"

The statistical limits of data set "B" are calculated in Table 8 and Figure 17.

58

Table 8: SPC limits for data set "B"

Control limits (Set B)		
Upper Action Limit	UAL = μ+3σ	0.0036236
Upper Warning Limit	UWL = μ+2σ	0.0023429
Mean Centre Line	Mean	-0.0002187
Lower Warning Limit	LWL = μ+-2σ	-0.0027802
Lower Action Limit	LAL = μ+-3σ	-0.0040610
Upper *Moving Range* Limit	UAL=MR × 3.27	0.0047242
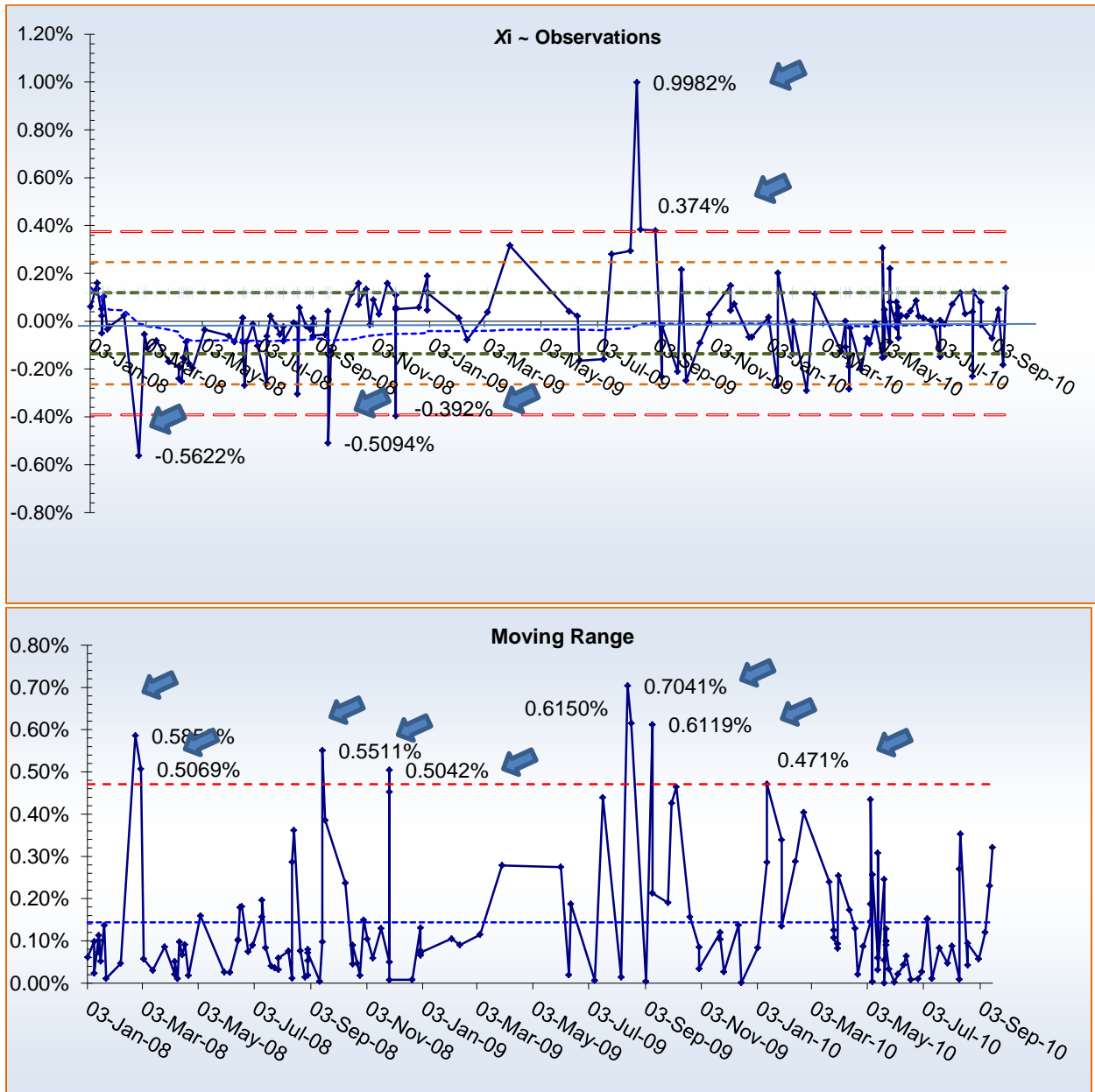
Figure 17: Control chart indicating some "out of control" states for data set "B"
Vertical axis values are unit-less percentage error measurments

4.5 Data exploration of set "C"

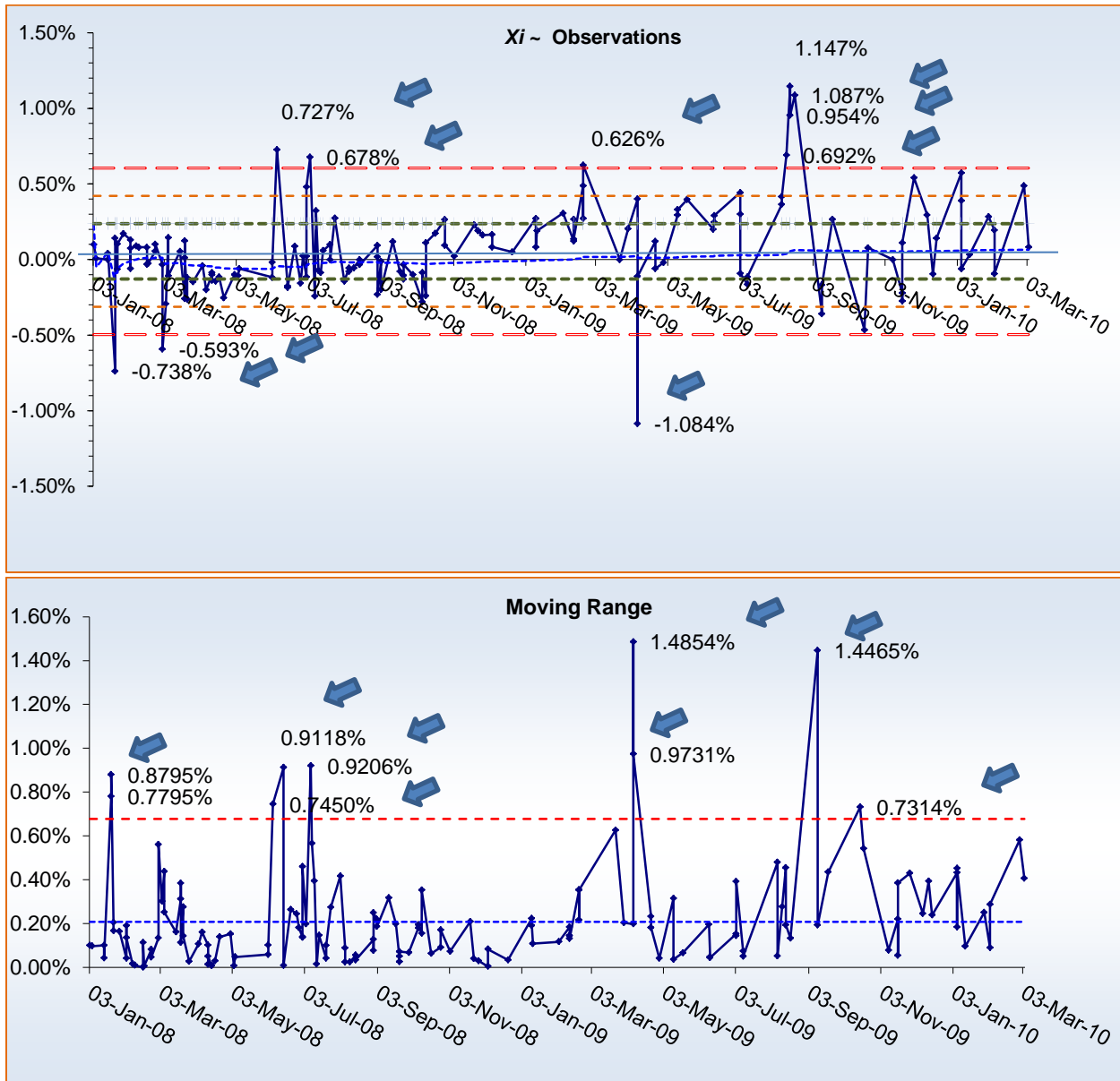Based on the Shapiro-Wilk test we cannot reject the idea that readings from data set "C" come from a normal distribution with 95% confidence. The mean is μ= 0.000948457 (unit-less error percentage value) and the standard deviation is σ= 0.001197873.
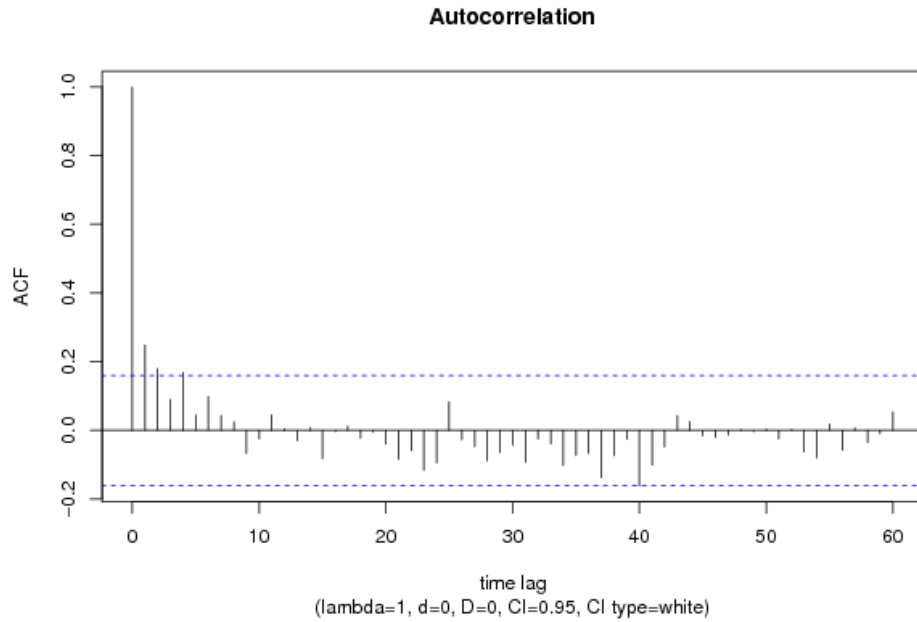
Figure 18: Plot of auto-correlaton for successive observations in set "C"

Table 9: Auto-correlation coefficient for lags of [1-3]

Autocorrelation Function			
Time lag k	ACF(k)	T-STAT	P-value
1	0.247502	3.0313	0.001435
2	0.179014	2.1925	0.014944
3	0.09009	1.1034	0.135816

The coefficient of autocorrelation in Figure 18 and Table 9 has been calculated to be at most 0.25; therefore data-set "C" is not significantly auto-correlated either.

4.5.1 Application of SPC on data set "C"

Table 10: Exploration of data set "C"

Date	Data point number	*Xi*	
8-Feb :	7	Min. :	-0.00305
12-Mar:	6	1st Qu.:	0.000229
30-May:	6	Median :	0.00094
7-Apr :	6	Mean :	0.000949
12-Jul:	4	3rd Qu.:	0.001672
17-Oct:	4	Max. :	0.007323
Data points :	117		
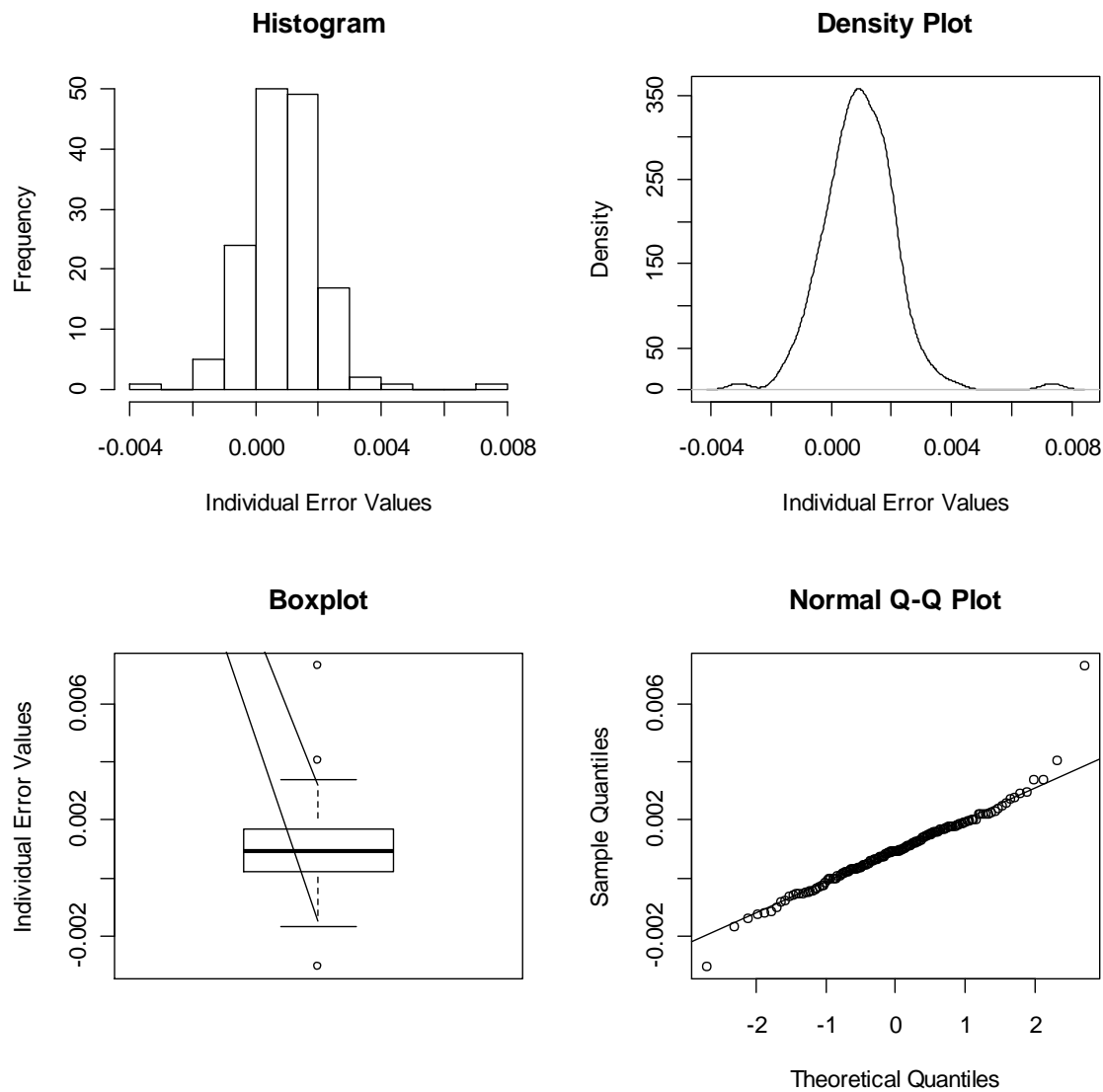
Figure 19: Visual data exploration of data set "C"

The statistical limits of data set "C" are calculated in Table 11, Figure 20.

Table 11: SPC limits for data set "C"

Control limits (Set C)		
Upper Action Limit	UAL = μ+3σ	0.0036424
Upper Warning Limit	UWL = μ+2σ	0.0027444
Mean Centre Line	Mean	0.0009485
Lower Warning Limit	LWL = μ+-2σ	-0.0008475
Lower Action Limit	LAL = μ+-3σ	-0.0017455
Upper *Moving Range* Limit	UAL=MR ×3.27	0.0033122
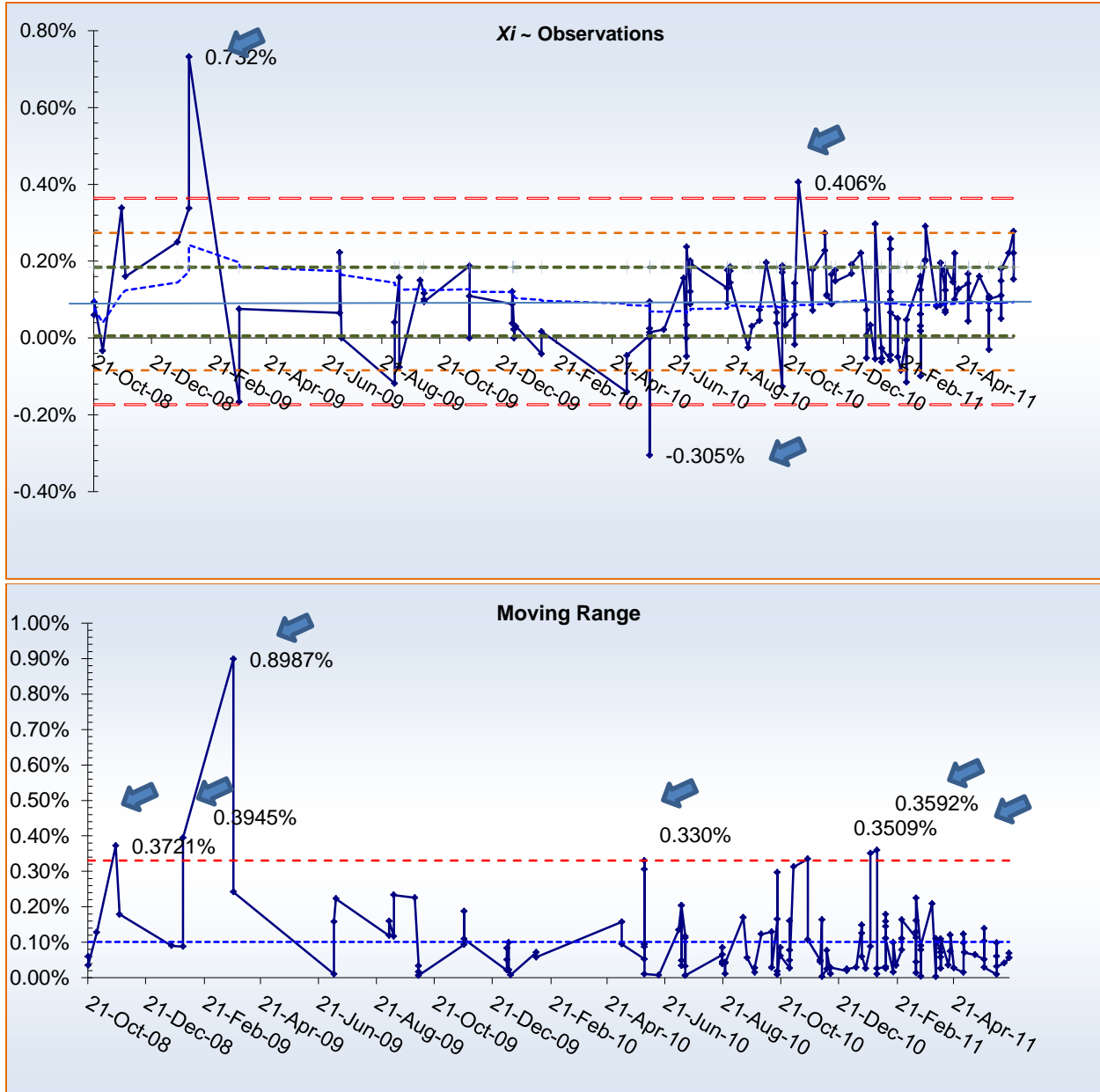
Figure 20: Control chart indicating some "out of control" states for data set "C"
Vertical axis values are unit-less percentage error measurments

4.6 Data exploration of set "D"

Based on the *Shapiro-Wilk* test we cannot reject the idea that readings from data set "D" come from a normal distribution with 95% confidence. The mean is $\mu = 0.001541254$ and the standard deviation is $\sigma = 0.002108014$.
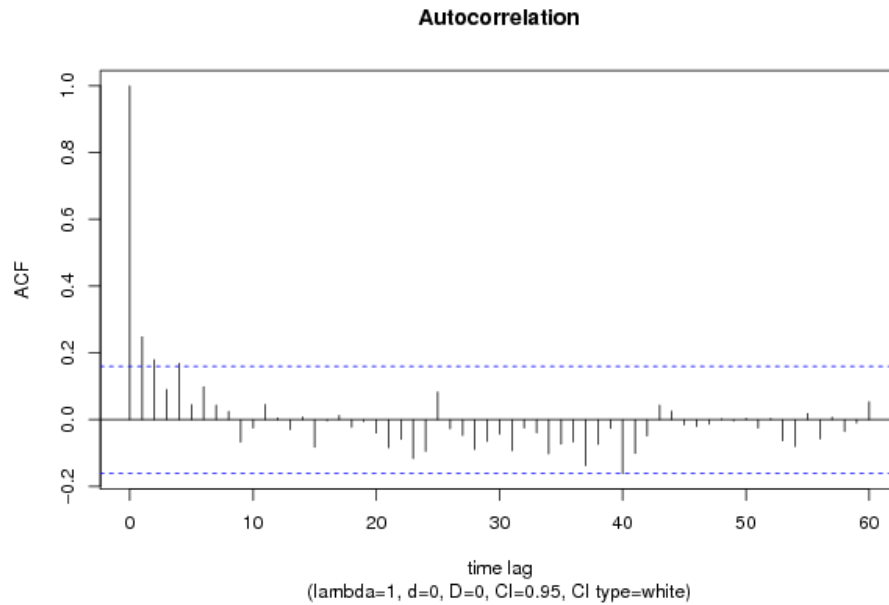
Autocorrelation

Figure 21: Plot of auto-correlaton for successive observations in set "D"

Table 12: Auto-correlation coefficient for lags of [1-3]

Autocorrelation Function			
Time lag k	ACF(k)	T-STAT	P-value
1	0.348184	3.6016	0.000241
2	0.119659	1.2378	0.109257
3	-0.044274	-0.458	0.323952

The coefficient of autocorrelation in Figure 21 and Table 12 has been calculated to be at most 0.34; therefore dataset "D" is not significantly auto-correlated either.

4.6.1 Application of SPC on data set "D"

Table 13: Exploration of data set "D"

Date	Data point number		Xi
12-Jul:	4	Min. :	-0.00209
17-Oct:	4	1st Qu.:	0.000559
20-Aug:	4	Median :	0.001372
30-Oct:	4	Mean :	0.001541
8-Jul:	4	3rd Qu.:	0.001797
23-Aug:	3	Max. :	0.013966
Data points :	84		
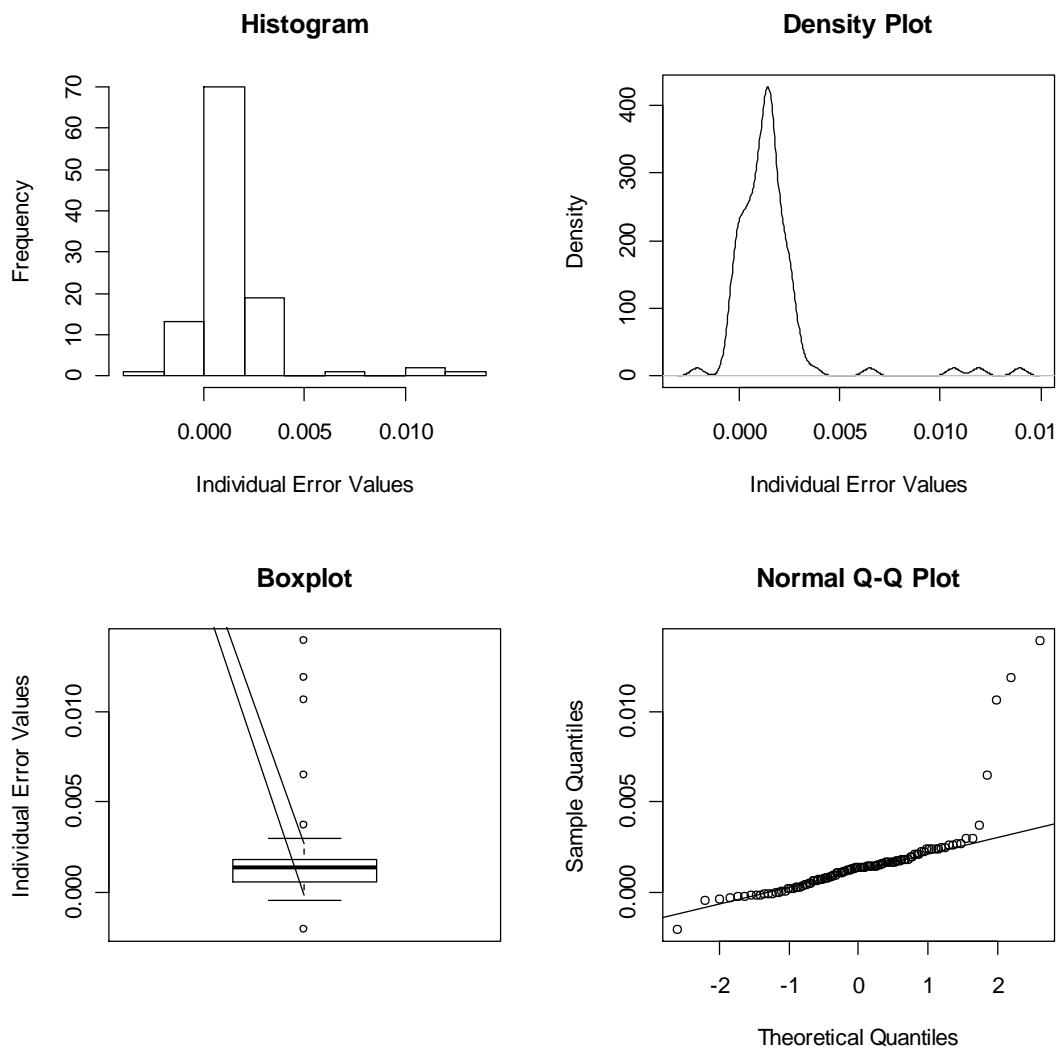
Figure 22: Visual data exploration of data set "D"

The statistical limits of data set "D" are calculated in Table 14 and Figure 23.

64

Table 14: SPC limits for data set "D"

Control limits (Set D)		
Upper Action Limit	UAL = μ+3σ	0.0048835
Upper Warning Limit	UWL = μ+2σ	0.0037694
Mean Centre Line	Mean	0.0015413
Lower Warning Limit	LWL = μ+-2σ	-0.0006869
Lower Action Limit	LAL = μ+-3σ	-0.0018010
Upper *Moving Range* Limit	UAL=MR × 3.27	0.0041093
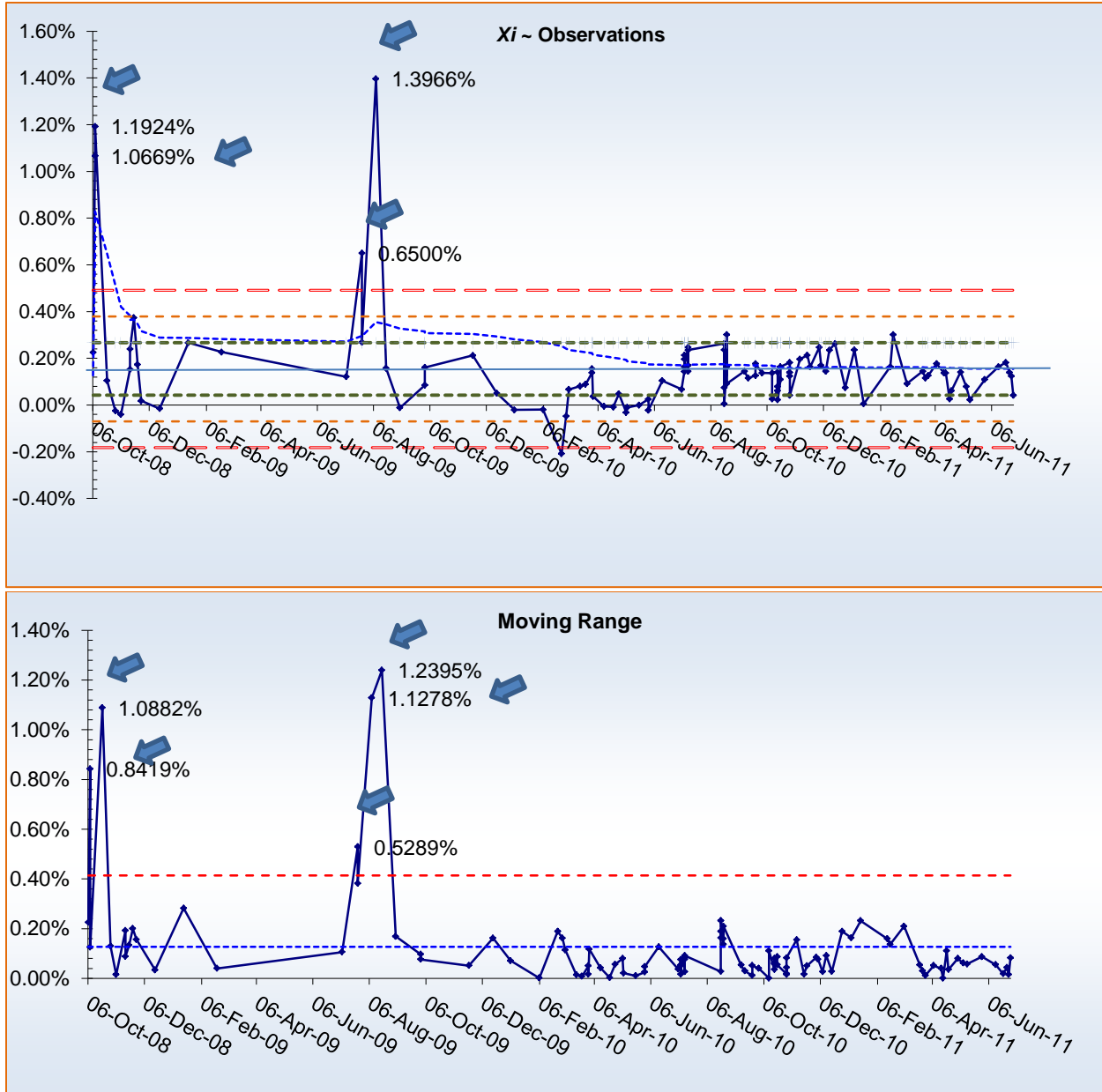
Figure 23: Control chart indicating some "out of control" states for data set "D"
Vertical axis values are unit-less percentage error measurments

4.7 SPC decision matrix

In this section the data generated from the run charts, Process Binary Matrix, were analyzed to generate SPC Decision Matrix. As discussed in section 3.4.5, nine alarm modes have been defined for the proposed model to detect the unnatural patterns within the control charts. A set of significance measure weights have been assigned to each alarm mode which enables the generation of an index called performance index. The performance index is a normalized index that shows the performance of the scale based on the number of the "out of control" alarms and their importance. This information along with other tools such as Pareto analysis and capability indices can be used to facilitate corrective actions to reduce the causes of variation.

Based on the performance of individual scales, Pareto charts have been created for each data set. They signify the importance of the "crucial few" alarms in comparison to the "useful many ".

The capability indices discussed in section 3.4.8 give operators further measures of performance. They represent a measure of how well the process has been within externally imposed limits that could be established as guidelines or process rules. Furthermore they unveil a very quick look at the spread of the data and can estimate the number of failures that could occur by chance in such a data distribution. In cases where the alarms are triggered more frequently than a random normal process, appropriate corrective actions are taken.

4.8 Application of run rules on data set "A":

Decision rules on data set "A" are shown in Table 15 for different modes. Figure 24 shows Pareto analysis of data set "A".

Table 15: Application of decision rules on data set "A"

Failure	Quick Overview	Symbol	Out of Control States
Mode 1	1 σ limit triggered 4 out of 5 times	1 Sigma	20
Mode 2	2 σ limit triggered 2 out of 3 times	2 Sigma	7
Mode 3	3 σ limit triggered	3 Sigma	11
Mode 4	Observations shifted UP	Mean-UP	21
Mode 5	Observations shifted DOWN	Mean-DOWN	10
Mode 6	Observations trending UP	Trend-UP	0
Mode 7	Observations trending DOWN	Trend-DOWN	0
Mode 8	14 consecutive points alternating	Instability	0
Mode 9	*Moving Range* limit triggered	MR	10
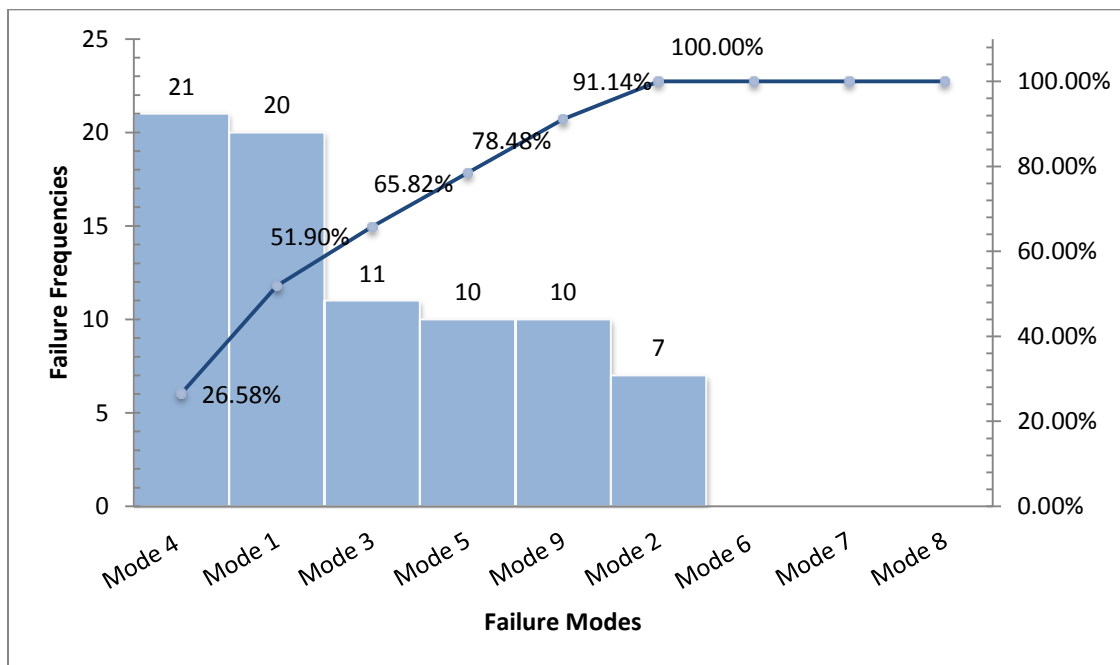
Figure 24: Pareto analysis of data set "A"

Process capabilities of data set "A" are shown in Table 16 and discussed in 4.8.2.

Table 16: Process capabilities of data set "A"

Process Statistical Capability Index	
The Actual Process Capability Cpk~1	0.8072
The Potential Process Capability Cp~1	0.9061
Failures Part Per Million (PPM~2700)	6528 PPM
Upper Specification Limit (USL)	0.005
Lower Specification Limit (LSL)	-0.005
Statistical Standard Deviation	0.0018394
The mean	0.0005455

4.8.1 Conventional SPC run rules

According to *WECO* rules, the out of control states for set "A" are shown in Table 17.

Table 17: Application of *WECO* rules on data set "A"

Failure	Description	Out of control States
Mode A	Runs above or below centerline of length 8 or greater	36
Mode B	Runs up or down of length 8 or greater	0
Mode C	Sets of 5 observations with at least 4 beyond 1.0 sigma	14
Mode D	Sets of 3 observations with at least 2 beyond 2.0 sigma	11
Total		61

4.8.2 Discussion, data set "A"

Set "A" showed a total number of 79 incidents of "out of control" states. This was more than the conventional method's 61 "out of control" states. *Mode 4* followed by *Mode 1* are the main causes of the failure alarms. This means that the variance for the measurements is generally high and they are shifted up. The process capability indices support this. C_p equals 0.906075, which is usually considered to be not good. The mean is located 10.9095% of the way from the center of the specs toward the upper specification limit.

4.9 Application of run rules for data set "B"

Decision rules on data set "B" are shown in Table 18 for different modes. Figure 25 shows Pareto analysis of data set "B".

Table 18: Application of decision rules on data set "B"

| Failure	Quick Overview	Symbol	Out of Control States
Mode 1	1 σ limit triggered 4 out of 5 times	1 Sigma	36
Mode 2	2 σ limit triggered 2 out of 3 times	2 Sigma	10
Mode 3	3 σ limit triggered	3 Sigma	16
Mode 4	Observations shifted UP	Mean-UP	1
Mode 5	Observations shifted DOWN	Mean-DOWN	12
Mode 6	Observations trending UP	Trend-UP	0
Mode 7	Observations trending DOWN	Trend-DOWN	0
Mode 8	14 consecutive points alternating	Instability	0
Mode 9	*Moving Range* limit triggered	MR	22
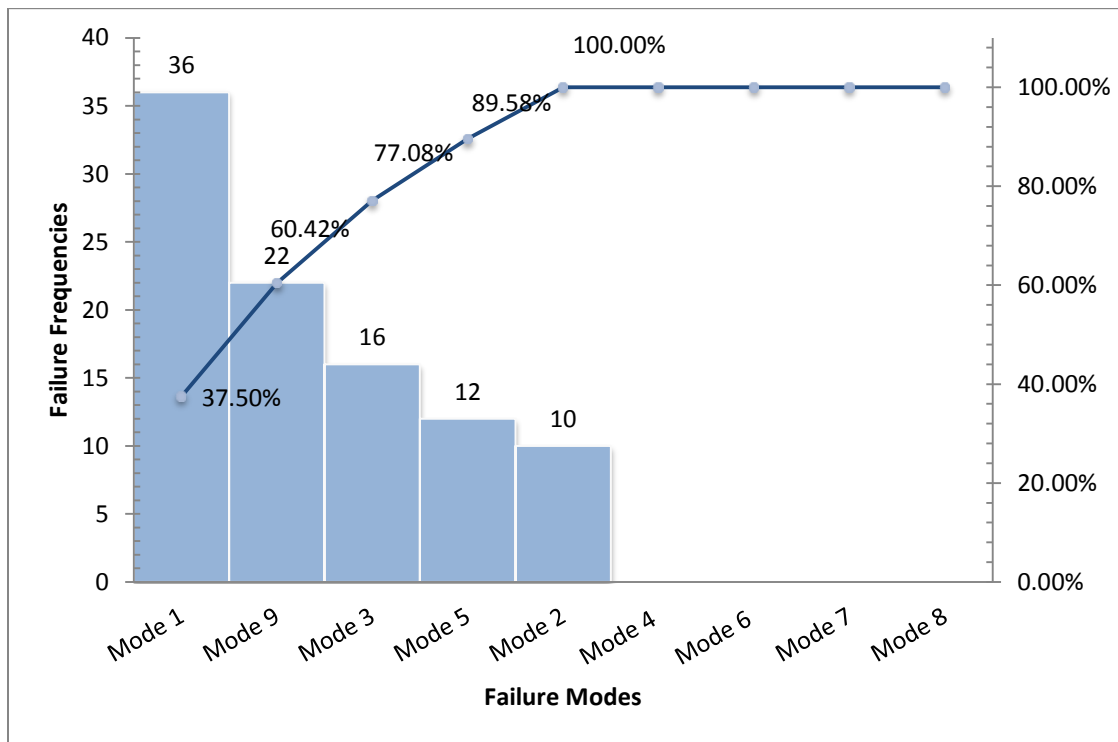
Figure 25: Pareto analysis of data set "B"

Process capabilities of data set "B" are shown in Table 19 and discussed in 4.9.2.

Table 19: Process capabilities of data set "B"

Process Statistical Capability Index	
The Actual Process Capability Cpk~1	1.2785
The Potential Process Capability Cp~1	1.3013
Failures Part Per Million (PPM~2700)	96 PPM
Upper Specification Limit (USL)	0.005
Lower Specification Limit (LSL)	-0.005
Statistical Standard Deviation	0.0012808
The mean	-0.0000877

4.9.1 Conventional SPC run rules

According to *WECO* rules, the out of control states for set "B" are shown in Table 20.

Table 20: Application of *WECO* rules on data set "B"

Failure	Description	Out of control States
Mode A	Runs above or below centerline of length 8 or greater	18
Mode B	Runs up or down of length 8 or greater	0
Mode C	Sets of 5 observations with at least 4 beyond 1.0 sigma	11
Mode D	Sets of 3 observations with at least 2 beyond 2.0 sigma	11
Total		40

4.9.2 Discussion, data set "B"

Set "B" showed a total number of 97 incidents of "out of control" states. This is considerably higher than the conventional method's 40 "out of control" states. *Mode 1* followed by *Mode 9* are the main causes of the failure alarms. This means that the variance for the measurements is spread in zone B more than usual. The *Mode 9* failure shows large jumps between the consecutive measurements that show a degree of randomness. C_p equals 1.30129, which is usually considered to be okay. The mean is located 1.75477% of the way from the center of the specification limits toward the lower specification limit

4.10 Application of run rules to data set "C"

Decision rules on data set "C" are shown in Table 21 for different modes. Figure 26 shows Pareto analysis of data set "C".

Table 21: Application of decision rules on data set "C"

| Failure	Quick Overview	Symbol	Out of Control States
Mode 1	1 σ limit triggered 4 out of 5 times	1 Sigma	17
Mode 2	2 σ limit triggered 2 out of 3 times	2 Sigma	4
Mode 3	3 σ limit triggered	3 Sigma	4
Mode 4	Observations shifted UP	Mean-UP	43
Mode 5	Observations shifted DOWN	Mean-DOWN	0
Mode 6	Observations trending UP	Trend-UP	0
Mode 7	Observations trending DOWN	Trend-DOWN	0
Mode 8	14 consecutive points alternating	Instability	0
Mode 9	*Moving Range* limit triggered	MR	12
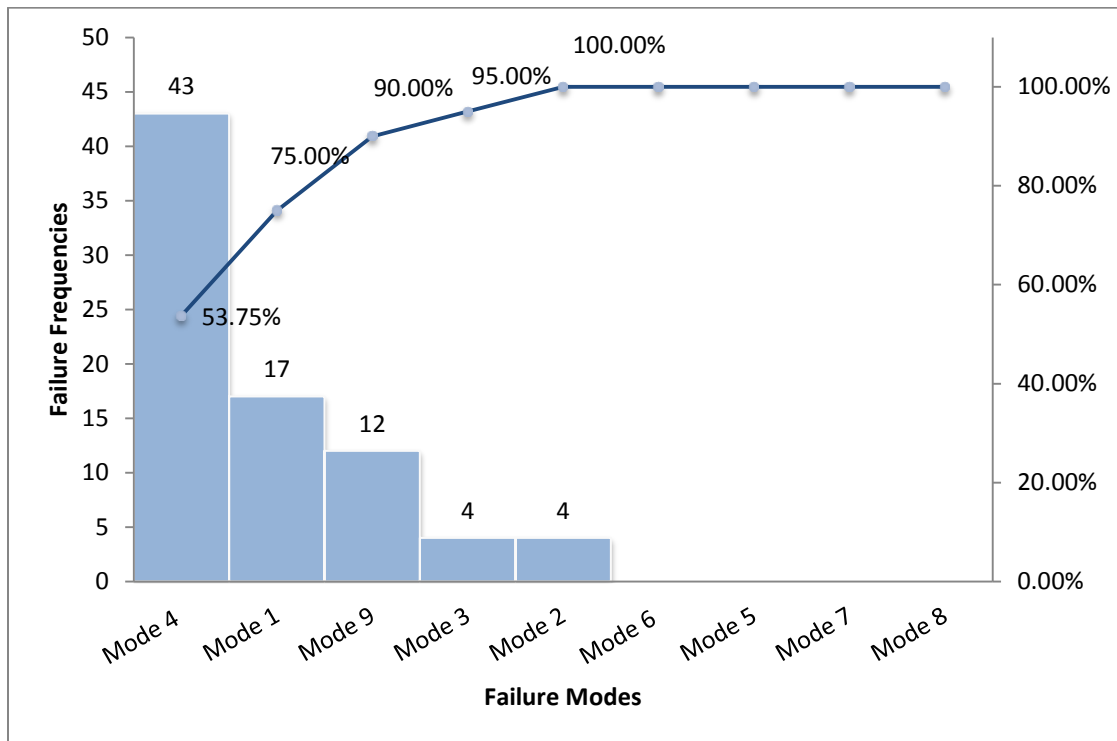
Figure 26: Pareto analysis of data set "C"

Process capabilities of data set "C" are shown in Table 22 and discussed in 4.10.2.

Table 22: Process capabilities of data set "C"

Process Statistical Capability Index	
The Actual Process Capability Cpk~1	1.5039
The Potential Process Capability Cp~1	1.8560
Failures Part Per Million (PPM~2700)	0.03 PPM
Upper Specification Limit (USL)	0.005
Lower Specification Limit (LSL)	-0.005
Statistical Standard Deviation	0.0008980
The mean	0.0009485

4.10.1 Conventional SPC run rules

According to *WECO* rules, the out of control states for set "C" are shown in Table 23.

Table 23: Application of *WECO* rules on data set "C"

Failure	Description	Out of control States
Mode A	Runs above or below centerline of length 8 or greater	15
Mode B	Runs up or down of length 8 or greater	0
Mode C	Sets of 5 observations with at least 4 beyond 1.0 sigma	4
Mode D	Sets of 3 observations with at least 2 beyond 2.0 sigma	5
Total		24

4.10.2 Discussion, set "C"

Set "C" showed a total number of 81 incidents of "out of control" states. This is considerably higher than the conventional method's 24 "out of control" states. *Mode 1* followed by *Mode 9* are the main causes of the failure alarms. This means that the variance for the measurements is spread in zone B more than usual. It is expected that the distribution would be shifted up and would bring down the capability indices. However, it should be noted that the C_p equals 1.85602, which is usually considered to be good. The mean is located on the center of the specs toward the upper specification limit. This shows that there has been an up shift in the readings however the corrective actions have been enough to bring the process to required specifications.

4.11 Application of run rules to data set "D"

Decision rules on data set "D" are shown in Table 24 for different modes. Figure 27 shows Pareto analysis of data set "D".

Table 24: Application of decision rules on data set "D"

| Failure	Quick Overview	Symbol	Out of Control States
Mode 1	1 σ limit triggered 4 out of 5 times	1 Sigma	21
Mode 2	2 σ limit triggered 2 out of 3 times	2 Sigma	6
Mode 3	3 σ limit triggered	3 Sigma	10
Mode 4	Observations shifted UP	Mean-UP	58
Mode 5	Observations shifted DOWN	Mean-DOWN	0
Mode 6	Observations trending UP	Trend-UP	1
Mode 7	Observations trending DOWN	Trend-DOWN	0
Mode 8	14 consecutive points alternating	Instability	0
Mode 9	*Moving Range* limit triggered	MR	10
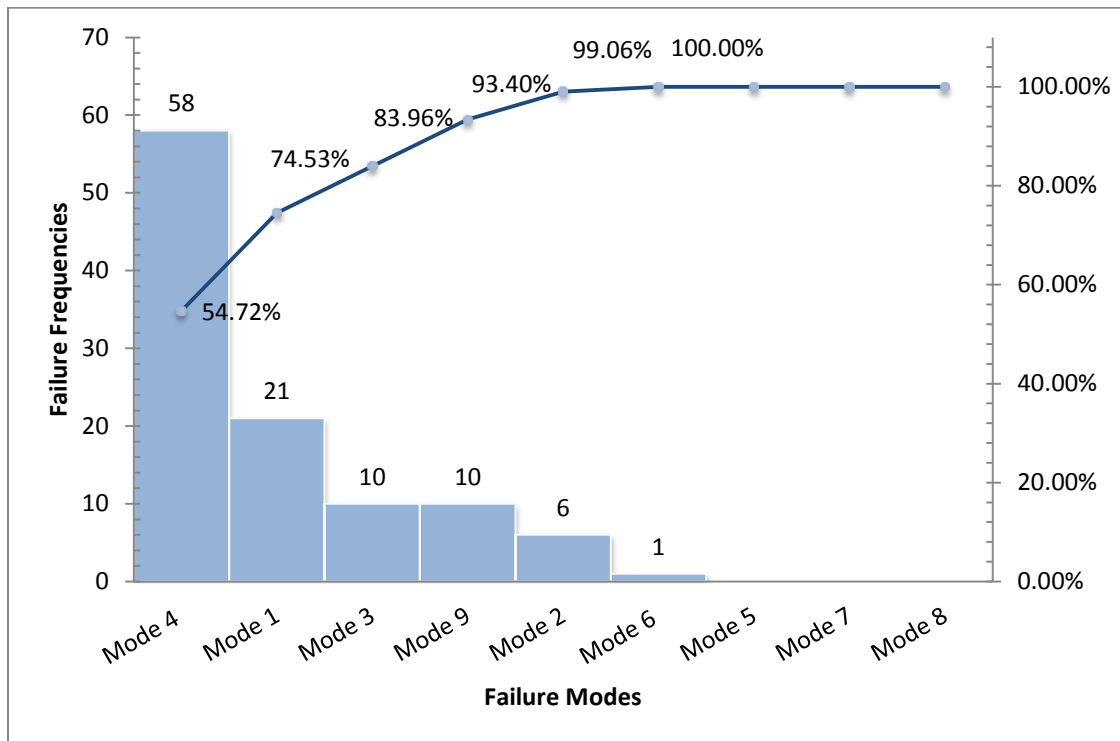
Figure 27: Pareto analysis of data set "D"

Process capabilities of data set "D" are shown in Table 25 and discussed in 4.11.2.

73

Table 25: Process capabilities of data set "D"

Process Statistical Capability Index	
The Actual Process Capability Cpk~1	1.0349
The Potential Process Capability Cp~1	1.4960
Failures Part Per Million (PPM~2700)	7.13 PPM
Upper Specification Limit (USL)	0.005
Lower Specification Limit (LSL)	-0.005
Statistical Standard Deviation	0.0011141
The mean	0.0015413

4.11.1 Conventional SPC run rules, data set "D"

According to *WECO* rules, the out of control states for set "D" are shown in Table 26.

Table 26: Application of *WECO* rules on data set "D"

Failure	Description	Out of control States
Mode A	Runs above or below centerline of length 8 or greater	35
Mode B	Runs up or down of length 8 or greater	0
Mode C	Sets of 5 observations with at least 4 beyond 1.0 sigma	7
Mode D	Sets of 3 observations with at least 2 beyond 2.0 sigma	5
Total		47

4.11.2 Discussion, set "D"

Set "D" showed a total number of 106 incidents of "out of control" states. This is considerably more than the conventional method's 47 "out of control" states. *Mode 4* followed by *Mode 1* are the main causes of the failure alarms. This shows a large shift up in the measurements, which could have been caused by a calibration issue. The spread of the measurements are in zone B more than usual. C_p equals 1.496, which is usually considered to be good which suggests a good process. Therefore, one can conclude that there is definitely a "calibration issue" in the process "D". The mean is located 30.8252% of the way from the center of the specs toward the upper specification limit, which again suggests the up shift of the measurements.

4.12 Summary of the results

Results showed that the proposed method performed better in alarming the shift pattern, direction of the shift and identifying sudden change in process. This is due to *Xi-individual* chart's high sensitivity to unnatural variations. Using $\bar{X}$ charts to sample the data, where possible, reduces the sensitivity of detecting sudden changes in control charts; Standard deviation of the process was calculated based on moving ranges of sample size 2 (Individual observations), therefore the proposed method shows very little leniency to sudden shifts in moving range than the *WECO* rule method. Developed method offers alarm modes indicating shift direction. This is important in monitoring and maintaining scale calibration and zero setting. Similarly, it offers the capability to detect directional trending that is suitable mechanical wear indicators. The data sets did not show any signs of trending in this test. Most frequent alarm modes are identified in Table 27, where importance (weight) is given to number of occurrence in all four data sets.

Table 27: Most to least frequent alarm modes for all sets

	Alarm Mode
Most Frequent	Mode 1
	Mode 4
	Mode 3
	Mode 9
	Mode 2
	Mode 5
	Mode 6
	Mode 7
Least Frequent	Mode 8

Mostly occurred alarm refers to Mode 1, 4 out of 5 points fall in zone "C", 1σ. Next to that is Mode 4, which refers to more than eight consecutive points observed above the mean. Followed by Mode 3, where a single "out of control" state observation falls out of 3σ limit. Mode 9 indicates a sudden shift in moving range caused by mechanical failure.

Chapter 5

Discussions and conclusions

5.1 Summary

This research was conducted to develop a practical method of monitoring bulk material weighing networks. The focus was on moving conveyor belt weighing systems that have the capability to perform high speed gravimetric measurement during material handling process.

In the development of a methodology a real scenario was considered and the solution architecture revolved around that particular application. Referring to the process optimization methodologies, one of the most practical monitoring tools, Statistical Process Control, was used. Since the application of Statistical Process Control has been mainly in manufacturing processes, implementation of the SPC tools such as Control charts for the current research was considered an unconventional use of control charts.

Useful metrics of the process were obtained to form the variables and the structure to be used in the control charts. The following are two challenges that were considered:

a) **Choosing the right type of a control chart**

b) **Utilizing a suitable method to detect the unnatural patterns within the control chart**

a) After reviewing the literature on unconventional applications of SPC in processes outside the manufacturing domain, it was decided to use a combination of *Xi-Individual* chart and *Moving Range* chart. The combination has the ability to detect a variety of faulty measurements by two methods. First, is the absolute performance of the scale within the current process with respect to standards and industry guidelines, and the second, is the relative performance of the current process with respect to the systems' past performance.

b) The second challenge was to use a suitable technique to detect the unnatural variations and patterns within the control chart while forgiving the natural randomness that is inherent in the system. Extensive research has been done on SPC during past decades. They can generally be categorized to have used two approaches: *Neural Networks* approach and the *Expert Systems* approach. Both methods demonstrate advantages and disadvantages for many applications. The *Neural Networks* approach is successful in detecting highly complex and superimposed patterns within the control charts as well as predicting the starting points of the unnatural patterns; however it is not possible to offer solid mathematical formulas leading to results since there is a learning process for the model. The *expert systems* approach relies on statistics and probability theory; therefore it is able to predict the faulty behavior through a system by assigning a probability of success

versus failure. The *expert system* approach works very well with single patterns and has been used widely in the manufacturing industry due to its reliability. In this research, it was decided to construct a series of nine fault modes based on the *expert systems* approach. The nine fault modes suit the application of this study and offer a balance between the *Type I* and *Type II* errors.

The proposed model was tested on four sets of data that were obtained from real conveyor belt scales in a bulk material handling company. In order to make sure that the model would perform as desired, it was required that the data sets were both normally distributed and did not have significant auto-correlation. *Shapiro-Wilk* normality test was satisfied to assure the rejection of the hypothesis that the sets were not coming from normal distributions. Consequently, the data sets were analyzed and met the requirements of the model.

*Pareto analysi*s of alarm modes and *capability indices* were generated for the proposed model as alternative assisting tools of monitoring. The performance of the model satisfied the purpose of the research. The next subsection summarizes the outcomes of the study.

5.2 Concluding remarks

Fast paced integral weighing systems such as conveyor belts are gaining popularity due to their accuracy and instant response compared to other available systems. An important issue frequently faced in bulk material handling practices is to define methods of controlling and minimizing the metrology error and monitor the calibration of the

equipment. Inaccurate measurements will be costly to manufacturers and handlers of the material due to lost material, industry regulations and down time.

The proposed method offers a new way of analysis on belt scales by utilizing available information of the process to produce useful metrics that can be translated to key performance indicators. A monitoring system is developed based on Statistical Process Control techniques. Evaluation results using real data reveal that the proposed method can successfully identify and classify the unnatural measurements based on external guidelines and previous performance. From the results of performance comparisons, the proposed method detects unnatural patterns more efficiently than conventional methods, while also offering the advantage of error classification and process capabilities with respect to expectations.

The proposed method is targeted to a specific scenario for validation. Nevertheless, application of this method to similar cases is possible with minor adjustments in the data acquisition methods. *X_i-individuals* chart was used for this application, since the data was normally distributed and lacked auto-correlation. The *Xi-individuals* chart combined by *Moving Range* chart offered a suitable combination swift response while enabling the use of other tools such as capability indices for a wider range of performance monitoring tools. Other types of charts such as $\bar{X}/R$ charts can be used when data sampling is possible. The proposed method can be utilized in cases where the purpose is to reduce the measurement errors to zero. It is also capable of eliminating the need of use of *EWMA* (*Exponentially Weighted Moving Average*) or *CuSum, (Cumulated Summation)* charts that are used for

detecting the mean shift. Once the process mean and the standard deviation are known, the proposed method can effectively recognize 9 types of unnatural patterns addressed in 3.4

Shewhart charts are very sensitive to trends and shifts in the mean which offers this model its agility in detecting mean shift; however this can cause more than usual number of false alarms. Results showed that the proposed method performed better in alarming the shift patterns in data. Generally, *Xi-individual* charts are more sensitive to unnatural variations; therefore they might generate more *Type I* errors than other charts [31]. (I.e. a Mode 3 alarm is generated, if a measurement point is beyond $\pm 3\sigma$ control limits. In a normally distributed set of observation probability of such an event happening is one in about 371 observations. Using the *WECO* rules increases the frequency of false alarms to about one in every 91.75 points, on the average [32].

In agreement with Does' argument, despite practicality of Statistical Process Control in industry, the common failure of SPC implementation is due to organizational, social and human factors. Lack of management and operational commitment as well as inefficient instructions and operator training would be an impending factor in popularity of applications of SPC. MacCarthy also notes that commitment of the top management in implementation is important especially in non-standard applications of the SPC [5, 8].

5.3 Suggestions to improve capabilities of this work

This research has used control chart pattern method based on expert systems rather than neural networks method due to its successful applications in the industry. A hybrid of control chart pattern method and neural network based approach could be a beneficial

method for applications in non-standard fields such as bulk material handling. Guh has researched unnatural patterns using a hybrid method and has proposed a method for issues in miss-classification of the unnatural patterns [31]. More research can be conducted to predict the optimum starting point of an unnatural pattern and the occurrence of more complex patterns using the agility of the control chart patterns and the extensive ability of neural network based pattern recognition methods to improve the quality of this research.

This research proposes monitoring method that can be used in unconventional fields such as bulk material handling. A crucial factor in quality improvement practice remains to be the corrective actions that are taken towards returning a system to being under control. The latter part requires extensive experimental knowledge and physical involvement with the mechanical systems and the process steps. This work can be offered as a tool to detect and plan the corrective actions and a gauge to direct the experts in right directions based on analytics of the system. There has been extensive consulting with the field experts in this research, however in implementation of the real systems conditions and the limitations can judge the success of a system.

References

[1] Hou, T. H., Liu, W. L., and Lin, L., 2003, "Intelligent Remote Monitoring and Diagnosis of Manufacturing Processes using an Integrated Approach of Neural Networks and Rough Sets," Journal of Intelligent Manufacturing, 14(2) pp. 239-253.

[2] M, C., 2009, "Shewhart, Deming, and Data," 2012(01/19) pp. 3.

[3] Womack, J.P., Jones, D.T., and Roos, D., 1990, "The machine that changed the world: based on the Massachusetts Institute of Technology 5-million dollar 5-year study on the future of the automobile," Scribner, .

[4] Blowers, J., 2009, "Lean Six Sigma: The Next Generation," BPM Process Excellence Network.

[5] MacCarthy, B., and Wasusri, T., 2002, "A Review of Non-Standard Applications of Statistical Process Control (SPC) Charts," International Journal of Quality & Reliability Management, 19(3) pp. 295-320.

[6] Roberts, S., 2011, "U.N. Says 7 Billion Now Share the World ," The New York Times, (November 1, 2011) pp. A4.

[7] Jan Merks, 2001, "Metrology in Mining and Mettalurgy," pp. 209.

[8] Does, R., and Schriever, B., 1992, "Variables Control Chart Limits and Tests for Special Causes," Statistica Neerlandica, 46(4) pp. 229-245.

[9] Nelson, L. S., 1984, "The Shewhart Control Chart-Tests for Special Causes," Journal of Quality Technology, 16(4) pp. 237-239.

[10] Does, R. J. M. M., Roes, K. C. B., and Trip, A., 1999, "Handling Multivariate Problems with Univariate Control Charts," Journal of Chemometrics, 13(3-4) pp. 353-369.

[11] Woodall, W. H., Spitzner, D. J., Montgomery, D. C., 2004, "Using Control Charts to Monitor Process and Product Quality Profiles," Journal of Quality Technology, 36(3) pp. 309-320.

[12] Wang, J., Kochhar, A., and Hannam, R., 1998, "Pattern Recognition for Statistical Process Control Charts," The International Journal of Advanced Manufacturing Technology, 14(2) pp. 99-109.

[13] Wang, J., Kochhar, A., and Hannam, R., 1998, "Noise Tolerance for Pattern Recognition for Statistical Process Control," The International Journal of Advanced Manufacturing Technology, 14(12) pp. 901-909.

[14] Shewhart, W.A., 1986, "Statistical method from the viewpoint of quality control," Dover Publications, .

[15] Richard D. Linville, J., 2000, "Application and Operating Principles Of Conveyor Belt Scales," Anonymous .

[16] Anjard, R. P., 1995, "SPC Chart Selection Process," Microelectronics and Reliability, 35(11) pp. 1445-1447.

[17] Guh, R. S., and Tannock, J., 1999, "A Neural Network Approach to Characterize Pattern Parameters in Process Control Charts," Journal of Intelligent Manufacturing, 10(5) pp. 449-462.

[18] Kahraman, C., Tolga, E., and Ulukan, Z., 1995, "Using triangular fuzzy numbers in the tests of control charts for unnatural patterns," Emerging Technologies and Factory Automation, 1995. ETFA'95, Proceedings., 1995 INRIA/IEEE Symposium on, Anonymous IEEE, 3, pp. 291-298 vol. 3.

[19] Wheeler, D.J., 2003, "Making sense of data: SPC for the service sector," Knoxville, TN: SPC Press, .

[20] Stapenhurst, T., 2005, "Mastering statistical process control," Elsevier, .

[21] Zhang, N. F., 1998, "A Statistical Control Chart for Stationary Process Data," Technometrics, pp. 24-38.

[22] Harris, T. J., and Ross, W. H., 1991, "Statistical Process Control Procedures for Correlated Observations," The Canadian Journal of Chemical Engineering, 69(1) pp. 48-57.

[23] Wardell, D. G., Moskowitz, H., and Plante, R. D., 1992, "Control Charts in the Presence of Data Correlation," Management Science, pp. 1084-1105.

[24] Di Bucchianico, A., 2008, "Lecture Notes 3TU Course Applied Statistics," .

[25] Small, B. B., 1956, "Statistical Quality Control Handbook," Western Electric Company, Mack Printing Company Easton, PA, .

[26] Ghomi, S., Lesany, S., and Koochakzadeh, A., 2011, "Recognition of Unnatural Patterns in Process Control Charts through Combining Two Types of Neural Networks," Applied Soft Computing, .

[27] Reid, R. D., Sanders, N. R., 2006, John Wiley, 2005, pp. 171-218, Chap. 6.

[28] NIST, H., "44(2007) NIST Handbook 44: Specifications," Tolerances, and Other Technical Requirements for Weighing and Measuring Devices, .

[29] Grinstead, C.M., and Snell, J.L., 1997, "Introduction to probability," Amer Mathematical Society, .

[30] Douglas G Altman, and J Martin Bland, 1995, "Statistics Notes: The Normal Distribution," BMJ, 310(6975) pp. 298.

[31] Guh, R. S., 2005, "Real-Time Pattern Recognition in Statistical Process Control: A Hybrid Neural network/decision Tree-Based Approach," Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 219(3) pp. 283.

[32] Champ, C. W., and Woodall, W. H., 1987, "Exact Results for Shewhart Control Charts with Supplementary Runs Rules," Technometrics, pp. 393-399.

Online and software resources

Wessa P. (2010), (Partial) Autocorrelation Function (v1.0.9) in Free Statistics Software (v1.1.23-r7), Office for Research Development and Education, URL

http://www.wessa.net/rwasp_autocorrelation.wasp/

The R code is based on :

Borghers, E. and P. Wessa, Statistics - Econometrics - Forecasting, Office for Research

Development and Education, http://www.xycoon.com/

Appendix

R code for Initial data exploration

Set A

R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

```
>DataA<-read.table(file.choose(),header=T,sep="\t")
>names(DataA) <- c("Date", "Xi")
>summary(DataA)
      Date              Xi
 30-Dec-10:   6      Min.  :       -0.0108447
 12-Jul-10:   5      1st Qu.:      -0.0008545
 28-Mar-11:   5      Median :       0.0002428
 12-Jun-11:   4      Mean  :        0.0005455
 29-Jun-10:   4      3rd Qu.:       0.0015315
 1-Jul-08 :   3      Max.  :        0.0114670
 (Other)  :  226
>str(DataA)
'data.frame':  253 obs. of  2 variables:
 $ Date: Factor w/ 156 levels "1-Apr-10","1-Dec-10",..: 115 42 42 74 74 85 85 133 103 112
...
 $ Xi  :num  0.000999 0 0.000418 -0.007377 0.001418 ...
>attach(DataA)
```

```
>par(mfrow = c(2, 2))
>hist(Xi, xlab="Individual Error Values", main="Histogram")
>plot(density(Xi), xlab="Individual Error Values", main="Density Plot")
>boxplot(Xi, ylab="Individual Error Values", main="Boxplot")
>qqnorm(Xi, main = "Normal Q-Q Plot", xlab = "Theoretical Quantiles", ylab = "Sample
Quantiles")
>qqline(Xi)
>shapiro.test(test1)
```

     Shapiro-Wilk normality test

data: test1
W = 0.9157, p-value = 9.231e-11

Set B

```
>DataB<-read.table(file.choose(),header=T,sep="\t")
>names(DataB) <- c("Date", "Xi")
>summary(DataB)
```
Date		Xi	
27-Nov-08:	4	Min. :	-5.622e-03
30-Aug-08:	4	1st Qu.:	-8.688e-04
31-Dec-08:	4	Median :	-6.140e-05
14-May-10:	3	Mean :	-8.774e-05
15-Jan-08:	3	3rd Qu.:	7.082e-04
21-May-10:	3	Max. :	9.983e-03
(Other) :	176		

```
>str(DataB)
'data.frame':  197 obs. of  2 variables:
 $ Date: Factor w/ 154 levels "1-Apr-10","1-Dec-10",..: 112 12 12 41 41 41 48 68 81 146 ...
 $ Xi  :num  0.000609 0.001594 0.001361 0.000229 -0.000501 ...
>attach(DataB)
>par(mfrow = c(2, 2))
>hist(Xi, xlab="Individual Error Values", main="Histogram")
>plot(density(Xi), xlab="Individual Error Values", main="Density Plot")
>boxplot(Xi, ylab="Individual Error Values", main="Boxplot")
>qqnorm(Xi, main = "Normal Q-Q Plot", xlab = "Theoretical Quantiles", ylab = "Sample
Quantiles")
>qqline(Xi)
>shapiro.test(test2)
```

     Shapiro-Wilk normality test

data: test2
W = 0.9092, p-value = 1.245e-09

Set C

```
>DataC<-read.table(file.choose(),header=T,sep="\t")
>names(DataC) <- c("Date", "Xi")
>summary(DataC)
      Date                  Xi
 8-Feb-11 :    7      Min.  : -0.0030520
 12-Mar-11:    6      1st Qu.:      0.0002287
 30-May-10:    6      Median :      0.0009399
 7-Apr-11 :    6      Mean  :       0.0009485
 12-Jul-10:    4      3rd Qu.:       0.0016724
 17-Oct-10:    4      Max.  : 0.0073228
 (Other) :     117
>str(DataC)
'data.frame':  150 obs. of  2 variables:
 $ Date: Factor w/ 65 levels "1-Dec-10","1-May-11",..: 30 30 46 23 35 16 39 39 33 33 ...
 $ Xi  :num  0.000596 0.000939 -0.000336 0.003385 0.0016 ...
>attach(DataC)
>
>par(mfrow = c(2, 2))
>hist(Xi, xlab="Individual Error Values", main="Histogram")
>plot(density(Xi), xlab="Individual Error Values", main="Density Plot")
>boxplot(Xi, ylab="Individual Error Values", main="Boxplot")
>qqnorm(Xi, main = "Normal Q-Q Plot", xlab = "Theoretical Quantiles", ylab = "Sample
Quantiles")
>qqline(Xi)
>shapiro.test(test3)
```

        Shapiro-Wilk normality test

data: test3
W = 0.9466, p-value = 1.741e-05

Set D

```
>DataD<-read.table(file.choose(),header=T,sep="\t")
>names(DataD) <- c("Date", "Xi")
>summary(DataD)
      Date                  Xi
 12-Jul-10:    4      Min.  : -0.0020868
 17-Oct-10:    4      1st Qu.:      0.0005592
 20-Aug-10:    4      Median :      0.0013723
 30-Oct-10:    4      Mean  :       0.0015413
 8-Jul-10 :    4      3rd Qu.:       0.0017971
 23-Aug-10:    3      Max.  :       0.0139659
 (Other) :     84
>str(DataD)
```

'data.frame':   107 obs. of 2 variables:
 $ Date: Factor w/ 81 levels "1-Dec-10","10-Nov-10",..: 72 80 80 35 60 69 12 12 26 42 ...
 $ Xi :num 0.00225 0.010669 0.011924 0.001042 -0.000249 ...
>attach(DataD)
>
>par(mfrow = c(2, 2))
>hist(Xi, xlab="Individual Error Values", main="Histogram")
>plot(density(Xi), xlab="Individual Error Values", main="Density Plot")
>boxplot(Xi, ylab="Individual Error Values", main="Boxplot")
>qqnorm(Xi, main = "Normal Q-Q Plot", xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
>qqline(Xi)
>shapiro.test(test4)

        Shapiro-Wilk normality test

data: test4
W = 0.6066, p-value = 1.694e-15