YAŞAR UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

MASTER'S THESIS

UNIVARIATE AND MULTIVARIATE STATISTICAL PROCESS CONTROL CHARTS: AN APPLICATION IN A CHEMICAL INDUSTRY

MERVE ÇAKIR

THESIS ADVISOR: ASST. PROF. DR. ERDİNÇ ÖNER

INDUSTRIAL ENGINEERING

Presentation Date: 31.05.2019

BORNOVA / İZMİR
MAY/2019

We certify that, as the jury, having read this thesis, we found it to be fully adequate in terms of in scope as well as quality which can fulfil the requirement of a thesis for the degree of Master of Science.

Jury Members:	**Signature:**
Asst. Prof. Dr. Erdinç ÖNER Yaşar University	....
Asst. Prof. Dr. Adalet ÖNER Yaşar University	....
Prof. Dr. Selma GÜRLER Dokuz Eylül University	....

Prof. Dr. Cüneyt GÜZELİŞ
Director of the Graduate School

ABSTRACT

UNIVARIATE AND MULTIVARIATE STATISTICAL PROCESS CONTROL CHARTS: AN APPLICATION IN A CHEMICAL INDUSTRY

ÇAKIR, Merve

Msc, Industrial Engineering Programme

Advisor: Asst. Prof. Dr. Erdinç ÖNER

May 2019, 70 pages

One of the tools to control quality of industrial process is the Statistical Process Control (SPC) by improving quality of the process and reducing the variablity of the process from target value for quality to control characteristics. This study aims to develop control charts as univariate and multivariate, which control auto-correlated processes and to compare the univariate and multivariate control charts for the same process. Univariate and multivariate control charts are investigated individually to review the related literature. This thesis investigates how to compare univariate control charts and multivariate control charts.

Key Words: Auto-Correlated Data, Chemical Industry, Multivariate Control Charts, Statistical Process Control, Time-series Models, Univariate Control Charts.

ÖZ

TEK DEĞİŞKENLİ VE ÇOK DEĞİŞKENLİ İSTATİSTİKSEL SÜREÇ KONTROL GRAFİKLERİ: KİMYA ENDÜSTRİSİNDE BİR UYGULAMA

ÇAKIR, Merve
Yüksek Lisans Tezi, Endüstri Mühendisliği
Danışman: Dr. Öğr. Üyesi Erdinç ÖNER
Mayıs 2019, 76 sayfa

İstatistiksel Süreç Kontrolü, sürecin kalitesini arttırarak ve hedef kalite kontrol karakteristiklerinin hedef değerden değişkenliğini düşürerek endüstriyel süreç kontrolünün kalitesini izlemek için kullanılan bir araçtır. Çalışmada, oto korelasyonlu süreçleri izlemek için tek değişkenli ve çok değişkenli kontrol çizelgeleri geliştirmek ve aynı süreç için tek değişkenli ve çok değişkenli kontrol çizelgelerini karşılaştırmak amaçlanmıştır. İlgili literatürde tek değişkenli ve çok değişkenli kontrol tabloları ayrı ayrı incelenmiştir. Bu tezde tek değişkenli kontrol çizelgeleri ile çok değişkenli kontrol çizelgelerinin karşılaştırılması incelenmiştir.

Anahtar Kelimeler: istatistiksel süreç kontrol, tek değişkenli kontrol kartları, çok değişkenli kontrol kartları, oto korelasyonlu veri, zaman çizelgesi modeli, kimya endüstrisi

ACKNOWLEDGEMENTS

I have made attempt to complete this project. However, it would not have been possible to complete this thesis without the sincere support and help of many individuals and organizations. I would like to express my deepest gratitude to those who helped me to complete this thesis.

I would like to thank my thesis advisor Asst. Prof. Dr. Erdinç Öner for his support, guide and patience. I would like to give my deepest gratitude to Dr. Effi Staiou whose contribution to stimulating suggestions and encouragement helped me to write the project.

A special thanks goes to my husband, Ahmet Cumhur Çakır, who supports and encourages me all the time and his sympathy which are make this journey enjoyable.

Finally, yet importantly, many thanks go to my family whose are two important persons in my life; my mother Sevinç Al and my little sister Tuğçe Al for their continuous moral support and endless love, without the support of them I would not succeed this project.

Merve ÇAKIR

İzmir, 2019

TEXT OF OATH

I honestly declare that my study, titled "*Univariate and Multivariate Statistical Process Control Charts: An Application in a Chemical Industry*" and presented as a Master's Thesis, has been written without getting any assistance contradictory with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Merve ÇAKIR

Signature

………………………………..

31.05.2019

TABLE OF CONTENTS

LIST OF FIGURES

.

LIST OF TABLES

.

SYMBOLS AND ABBREVIATIONS

ABBREVIATIONS:

ACF: Auto-correlation Function

ARIMA: Autoregressve Integrated Moving Average

$CaOH_2$: Calcium Hydroxide

CL: Center line

CTQC: Critical to Quality Characteristic

CUSUMCC: Cumulative Sum Control Chart

DOE: Design of Experiment

EWMACC: The Exponentially Weighted Moving Average

I.I.D.: Independent and Identically Distributed

I-MR: Individual Measurements and Moving Range

LCL: Lower Control Limit

MSPC: Mutivariate Statistical Process Control

MV: Multivariate

NaOH: Caustic

R CHART: Range Chart

SCC: Shewhart Control Chart

SPC: Statistical Process Control

TAR: Threshold Autoregressive

TSM: Time Series Model

UCL: Upper Control Limit

CHAPTER 1
INTRODUCTION

One of the biggest progress of the recent years is Statistical Process Control (SPC) (Montgomery, 2005), which is used to control and enhance quality, by reducing variability. There is a set of tools in SPC, which can be used to solve issues easily, underlying statistical principles can be implemented in every process and leave a impact on preventing the quality problems, controlling and improving quality. All individuals can improve the quality continuously using SPC with the support of the top management. Therefore, SPC is an essential part in the proper deployment of a Quality Management Program.

It is also essential to deploy the “magnificent seven” tools to implement SPC properly and understand and better manage a process. One of the most technically sophisticated and commonly used tools of SPC is the traditional Shewhart Control Charts (SCC), which was initially proposed by Walter A. Shewhart in 1920s. SCCs are predominantly used as an online tool, to monitor and control the behavior of a process, especially in phase I implementation of SPC, where processes are influenced by assignable causes and out of control situations occur, because of large shifts in the parameters which are monitored. Typical examples of SCCs are the $\overline{X}$ and R and $\overline{X}$ and S control charts in the cases where rational subgrouping and sampling can take place, or the Individual Measurements and Moving Range Control Charts (I-MRCCs) for the case of individual observations.

Two main assumptions for applying the control charts are; data observations are normally and independently distributed. Shewhart process model indicates that the data, which has been produced with a process in the control state are normally and independently distributed with a mean of μ and a standard deviation of σ (NID (μ, σ)). The sensitivity and the effectiveness of the Shewhart control charts will be affected if these assumptions are violated.

In phase II of SPC implementation, if the processes tend to operate more under an in control state, there are smaller shifts in the process parameters. SCCs are not so sensitive to the smaller shifts in the process parameters and therefore it is harder to detect. In this case, the time-weighted control charts offer an excellent alternative. The process monitoring can be solved using the CUmulative SUM Control Chart (CUSUMCC) and the Exponentially Weighted Moving Average Control Chart (EWMACC). In case of small shifts in the process, the information contained in the entire sequence of observations should be taken into account. This last advantage of the time-weighted control charts is applied when the individual observations control charts are used very commonly in the chemical and process industries, where sampling has no rationale. Moreover, the EWMACC, which can be considered as a weighted average of the old and current observations, is not sensitive to the normality theory and therefore, it will be applied appropriately in the individual views, where violations of the normality assumption are also common. Therefore the EWMACC is considered to be easier to set up, operate, and interpret.

Individual measurement data occur frequently for the chemical and process industries, so the I-MRCC is regarded as the most appropriate control chart. Unfortunately, the performance of these charts is dramatically influenced by even moderate violations of the normality assumption. Also, it is known that in chemical process industries, the data are usually auto-correlated and sometimes not normally distributed.

Especially, if the independence of the observations assumption is violated in the I-MRCC, the number of out of control points exhibiting a situation of an out-of- control process will increase, though this is not true. In reality, these out-of-control points are called false alarm. Due to the nature of the processes and the frequency of the sampling procedure, the autocorrelation is expected to exist in the data in the chemical industry.

However, one of the disadvantages of SCC characteristic is that it can monitor only one critical to quality characteristic (CTQC) at a time.

The complexity of the chemical process industry has increased over time and more than one parameter affects it. Since the chemical industry processes are getting complex, monitoring the CTQCs separately causes to miss the correlation or interaction between these variables. Therefore, by using the traditional control charts, it is not possible to identify these problems.

The use of Multivariate Control Charts (MVCCs) is an alternative approach which can be useful in the chemical industry and can cover more than one CTQC in the same control chart. One of the generally used MVCC is Hotelling T^2 which was first introduced by Harold Hotelling in 1947. More than one variable can be monitored through Hotelling T^2 method at a moment, considering the correlation between the quality characteristics, as the most important advantage of the approach.

In our study, the case of a chemical industry producing pharmaceutical glycerin is examined. In our chemical production process, to control and improve quality of process, final product has four CTQs that we want to monitor. These CTQs are density, ester content, glycerin content, and humidity. Historical data should be used to design Phase I of the control charts. In our study, observations from a three month period are used. 105 data points of daily measurements of the quality characteristics are used to analyze the behavior of the data and to decide upon the appropriate control charts needed. To design the control charts, control limits and center lines are determined, based on past data, and stability of process is evaluated. In case there are out- of - control points, relevant investigation is performed and the control limits and center lines are recalculated, accordingly to decide about the control limits used for Phase II implementation.

In the following chapters, we present the Theoretical Background starting with general information about SPC, the "magnificent seven" tools and the statistical process control techniques applied. Furthermore, the theory of the SCCs and MVCCs, the basic assumptions behind them and interpreting control charts guidelines are addressed. In addition, the Methodology Section gives the details of the design of the appropriate control charts, based on the results of the data analysis. The results that are obtained from the application of the methodology are given in detail and discussed in the Results and Discussion Section. Finally, the Conclusions of the study and proposed Future Works are presented. The concluding Section of this chapter introduces fundamental information of the research.

1.1. Statistical Process Control

Statistical process control (SPC) is a great tool which can be used to enhance the stability of the process and capability by decreasing the variability (Montgomery, 2009). In order to estimate and control the quality of products during the process, the

SPC is used as a methodology which monitors the process behavior. The seven main tools which are also named “the magnificent seven” tools of SPC are as follows (Montgomery, 2009):

1. Histogram or stem-and-leaf plot

2. Check sheet

3. Pareto chart

4. Cause-and-effect diagram

5. Defect concentration diagram

6. Scatter diagram

7. Control chart

For statistical process control; first data is gathered and evaluated to monitor and control process which is an important method to provide continuous improvement. Dr. Walter Shewhart from Bell Laboratories in the 1920's was the first person who developed SPC and Dr. W. Edwards Deming expanded it to introduce SPC to Japanese industry after WWII.

One of the seven major tools of SPC is Control Chart, which is the topic of this thesis and explained below briefly.

Control charts were introduced by Walter A. Shewhart in 1924. A typical control chart is a tool for the graphical representation of the quality aspect versus sample number or time, which points to control and monitor process behavior. Control chart contains the Center Line (CL), Upper Control Limit (UCL), and Lower Control Limit (LCL). While the process is under control, the points will be plotted within the control limits.

A source of variation can be detected with the CC (Montgomery, 2009). The detailed information on the CCs is presented in Section 1.2.

1.2. Shewhart Control Charts

Statistical Process Control (SPC) is used to monitor and control the processes for measuring and controlling the quality. Traditional SCC is used as a basic tool for SPC.

Control charts rely on a fundamental hypothesis, that the observation is individually and identically dispersed (Montgomery, 1991) (Montgomery, 2009). The SCC consists of a CL and UCL and LCL. The UCL and LCL are symmetric according to the center line. The measurements are planned on the control charts versus time or sample number. Each point shows a brief statistic which is calculated from a sample quality characteristic measures. The control limits are typically measured as three sigma limits above and below the CL. A point which is plotted outside of the control limits shows the existence of a particular cause of variation. Furthermore, specific causes test can show an out-of-control situation if there is an observation regarding a statistically unique design of points in the control chart. Even the process is in statistical control, and a point may occur outside the control limits by chance, which results in a false out-of-control signal. Nonetheless, while the existence of a particular cause is correctly signaled by the Shewhart chart, in order to determine and eliminate the nature of the problem, additional action is needed (Mitra, 1998). In the meantime, according to Alwan (Alwan, 1992); control charts are so sensitive towards the presumption of uncorrelated data. The existance of an autocorrelated data in the monitored process results in quite many out of control points, exhibiting a situation of an out of control process, although that is not the case. These out of control points are false alarms in reality. In the case of chemical and process industries, the classical SCCs are not always useful. Because the SCCs are based on two basic assumptions of statistically independent data and normality. However, in the chemical industry, the data are related since there is continuous production, and the sampling is consecutive (Mastrangelo & Montgomery, 1995).

The variability is a part of all theprocesses. Two main causes of variation are special causes and common causes. If the variability is not caused due to the inherent characteristics in the process, it is called as special or assignable cause. However, if the variability is caused due to the inherent characteristics of the process, it is called as common or chance cause (Mitra, 1998).

There are two kinds of control charts based on the type of data available for analysis, which are: Control Charts for Variables and Control Charts for Attributes.

1.2.1. Control Charts for Variables

Most of the quality characteristics are expressed to see numerical measurements. Based on the continuous distribution, variable control charts are applied to data. Variable including dimension, length, temperature, etc. means a quality characteristic that is estimated at a numerical range. The mean and variance of the quality characteristic should be monitored when dealing with the variable type of quality characteristic. (Montgomery, 2009).

1.2.1.1. $\overline{X}$ Control Charts

The control chart is used to control an average of process. The control charts can be used to monitor the standard deviation (s) of the process and the R chart (range chart) can be used to monitor the range of the process. $\overline{X}$ and R charts are used to discover tiny shifts in the process and mostly used with small sample sizes. $\overline{X}$ and s charts are essentially used when the sample sizes are variable or larger. Of the most important techniques used to monitor and control processes are the charts mentioned above (Montgomery, 2009).

The assumption of the normality of quality characteristic is very common in SPC. However, the sample size n is quite important if the quality characteristic is not normally distributed. According to the probability theory, Central Limit Theorem implies that the sum of the n independently distributed random variables is approximately normal, even the distribution of the individual variables are not normally distributed. It is very important that the variables must be identically distributed. The approximation improves as sample size; n goes to infinity (Hogg, McKean and Craig, 2013).

1.2.1.2. Individuals Control Chart

When the sample size (n) equals 1, individuals control chart will be used and it will be impractical to use rational subgroups (NIST/Sematech, 2009). According to Montgomery, individuals control charts can be used when:

- There is no benefit of using subgroup because of automated inspections and measurements.
- There are long intervals between observations because of slow production which cause delays for enough data.
- Chemical processes’s repeating measurements cause dependent data.
- Taking multiple measurements on the same unit of product.
- Measurements may vary very small and produce very small standard deviation to control it in process plants.

Individuals control charts can be used with moving range chart in order to be able to display the difference between two adjacent observations. For individuals control charts, the normal distribution assumption is not required in order to calculate control limits (Wheeler, 2010). The individuals control chart is very efficient to detect large shifts in the mean of the process.

1.2.1.3. The Cumulative Sum Control Chart

The major disadvantage of the SCC is that the chart does not contain the information which come from the entire sequence of observations. It only uses the information of the previous observation. So, this property makes SCCs insensitive to detect small process shifts (Montgomery, 2009).

CUSUMCC is developed by E. S. Page in 1954 and is used to plot observations, which are cumulative sums of observations from target value, versus time (Grigg et al., 2003). CUSUMCCs are very effective to detect small shifts in the process mean; especially when the magnitude of the shift is 1.5 sigma to 2.0 sigma. CUSUMCC can be constructed to monitor the mean of the process both for individual observations and for the averages of the rational subgroups. One of the most important characteristics of the CUSUMCC is that it includes all the information sequence by plotting the cumulative sums of deviations of sample data from the target value (Montgomery, 2009).

1.2.1.4. The Exponentially Weighted Moving Average Control Chart

In order to explore the small changes in the process, the exponentially weighted moving average (EWMA) control chart can be used as an alternative to SCC. The EWMACC is introduced by Roberts (1959) and is similar to CUSUMCC. EWMACC is usually used with individual observations like CUSUMCC and it uses the information of both the past and the present observations (Montgomery, 2009). In addition, EWMACCs performance for non-normal data is good unlike SCC (Montgomery, 2009).

In chemical process industry, generally individual measurements and observations are obtained from the process. Therefore, it is reasonable to implement EWMACCs for individual observations (Mastrangelo & Montgomery, 1995).

1.2.2. Control Charts for Attributes

Most of the critical to quality characteristics are expressed in terms of numerical measurements. However, sometimes, it is impossible to numerically measure or express the quality characteristics; they can be classified as conforming and non-conforming units. Also, there are different quality characteristic classifications as defective and non-defective. These type of quality characteristics called attributes. Attribute charts are very practical to use in service industries and in non-manufacturing quality improvement efforts (Montgomery, 2009). Various types of attribute control charts are explained below.

1.2.2.1. Control Charts for Fraction Nonconforming (P-Chart)

Fraction non-conforming control chart, which is also known as P-chart is used to control the proportion of nonconforming matters in a group to the entire amount of parts in that population. Moreover, P-chart is used to observe the portion of defective items' consistency over time. P-chart follows the binomial distribution and its assumptions (Montgomery, 2009).

1.2.2.2. Number Nonconforming Control Chart (NP-Chart)

The NP-chart is one type of number nonconforming control charts, which is used to monitor number of defectives rather than proportion of defectives' consistency over time. However; for each sample, the subgroup size must be the same. NP-chart is also based on the binomial distribution like P-chart (Montgomery, 2009).

1.2.2.3. Control Charts for Nonconformities

There are two kinds of control charts for nonconformities. The first one is C-chart, which is used to determine the variation in counting type attribute data over time. The sample size must be constant. It is based on Poisson distribution and its assumptions. C-chart is practical to use when large number of products are inspected. The second type of chart for non-conformities is U-chart, which is for sample sizes equal to one inspection unit. However; in this case, the sample size is higher than one, and usually, the average number of nonconformities per unit is used (Montgomery, 2009).

1.2.3. Interpreting Control Charts

Control charts are graphical displays of measured or calculated critical quality characteristics versus observation numbers or time. Control charts contain UCL, LCL and CL. If the process is in control, almost all observations fall between UCL and LCL. However, if the observations fall outside of the control limits, the process is stated as out of control. In out of control situation, the reason for this behavior should be examined and the action should be exercised in order to eliminate the assignable cause. (Montgomery, 2009)

Besides these basic principles to interpret control charts, The Western Electric Rules (1956) with sensitizing rules for Shewhart control charts are used to detect nonrandom patterns. The process is stated as out of control if either one of the following conditions is presented as Western Electric Rules (Montgomery, 2009). While interpreting the Control Charts, these rules are taken as a basis in this study.

According to these zone rules, it can be concluded that the pattern is nonrandom and the process is out of control (Montgomery, 2009). An example of an out of control situation can be seen in Figure 1.1.

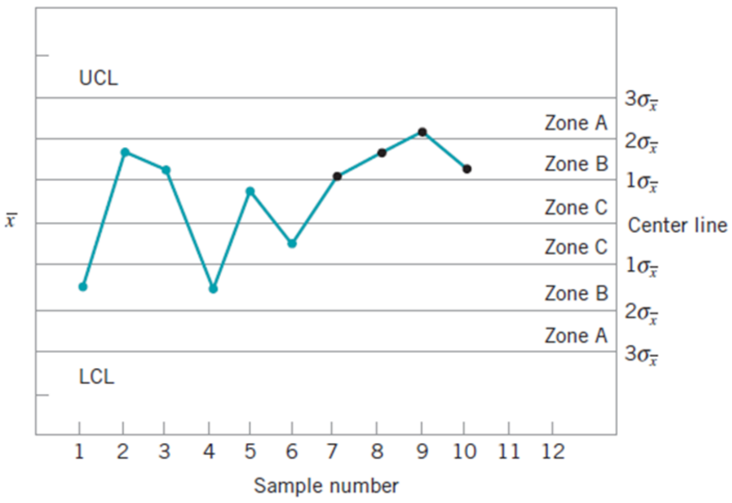

Figure 1.1. Instance of an Out of Control Situation for a Control Chart (Montgomery, 2009)

1.3. Basic Assumptions in Statistical Process Control

In statistical analysis and in constructing SCC, normality, independence, stability, random samples and equal variance are the most common and important data assumptions.

The first and the most important assumption is that the data must be independent and identically distributed (IID). The second assumption is that the data must have normal distribution. In order to understand whether the data is normally distributed or not, statistical tools such as probability plot or histogram can be used. The last assumption is that the data must not be auto-correlated. These assumptions have great importance on constructing appropriate control charts and interpreting them correctly.

Furthermore, stationarity of the process is another important assumption (NITS/Sematech, 2009).

The probability distribution of one of the random variables is not affected by the realization of other means that these random variables are independent. If each random variable have the same distribution and are mutually independent, the random variables can be called as IID (Hogg & Craig, 1978).

When mean, variance and unconditional joint probability distribution of the stochastic process do not change over time, the process is a stationary process. Stationarity assumption become more important for time series analysis. If there is a trend in the mean of the process, the stationarity assumption is violated (Gagniuc, 2017). The process mean level and changes in the amount of variation in the process are the most common changes in process performance over time (Devor, Chang and Sutherland, 2007). In order to be able to understand the behavior of the process or in other words stationarity of the process, time series plot can be used.

1.4. Auto-Correlated Data and the Use of Times Series Models

1.4.1. Time Series Models

Time Series Analysis (TSA) is a statistical technique which deal with TS data measured over successive periods (Hipel and McLeod, 1994). TS models are useful methods for prediction and forecasting. TS modeling is used for working on time based data. Mostly in chemical and process industry, there is continuous production and with continuous production, time becomes an important factor that affects data. Time series can be continuous or discrete. Past observations of the quality characteristic variables are collected and analyzed in order to be able to perform time series forecasting and develop a model which defines the essential relationship (Granger & Anderson, 1978).

It can be said that the time series forecasting is the act of prediction of future by understanding past. An appropriate model fitting is the most important part of time series forecasting (Hipel & McLeod, 1994). There are two main approaches for time series modeling. Moving average, exponential smoothing and autoregressive integrated moving average are traditional statistical models for linear models. Linear models are commonly used because of their simplicity in understanding and implementation. However, in real life, also non-linear problems appear (Granger & Anderson, 1978). Bilinear model, threshold autoregressive (TAR) model and the autoregressive conditional heteroscedastic (ARCH) model are some examples of non-linear models (Tong, 1983; Kandananond 2013; Jensen, Jones-Farmer and Champand, 2000).

Autoregressive Integrated Moving Average (ARIMA) is the most common and popular time series model as a stochastic model as mentioned previously. Linear and having a known statistical distribution are the basic assumptions in order to implement ARIMA model to the data. (Zhang, 2003) The idea of stationarity of a stochastic model are envisioned as style of statistical equilibrium. The statistical parameters like mean and variance of stationary process do not rely on time. Stationarity is a necessary condition in order to construct a time series model for forecasting (Engle, 1982).

1.4.2. Auto-Correlated Data

Individual measurements in a chemical industry arrive consecutively and sampling happens in short time intervals, therefore auto-correlation is something that is expected. Furthermore, distribution of data can be non-normal. However, Shewhart control charts can only be used under two assumptions as normally distributed data and non- auto correlated data (Johnson, 1949).

Control charts are used in chemical industry to monitor the process and its variability. Construction of the control charts are based on IID data. However, assumption of independent observation is violated because of the periodic sampling in a continuous system of chemical industry which causes auto-correlation (AC) (Elevli, Uzgören and Savas, 2008). There are two types of auto-correlation: positive auto correlation which occurs between successive similar observations and negative auto-correlation which occurs between successive non similar observations. Furthermore, if the data have positive autocorrelation, too many false alarms are observed in control charts resulting misleading conclusions (Alwan, 1992).

Although, the level of AC can be measured analytically (Harris & Ross, 1991). One of the analytical methods is measuring autocorrelation over a series of time-oriented samples (Montgomery & Mastrangelo, 1991). An example of an auto-correlation function which shows correlated data can be seen in Figure 1.2.

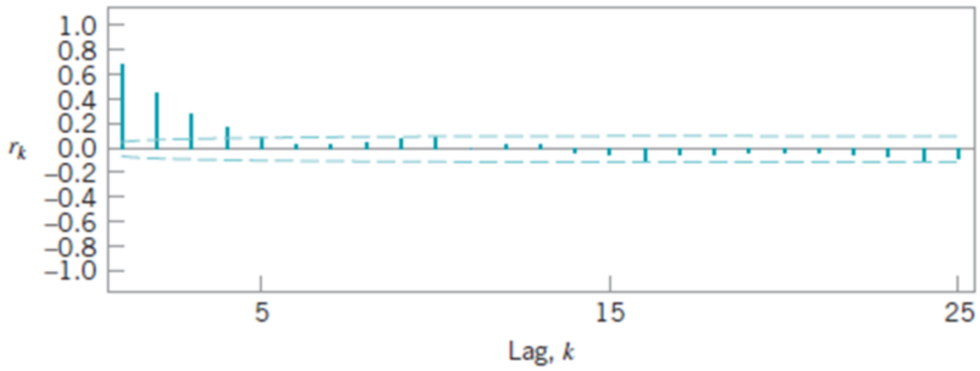

Figure 1.2. Example of Autocorrelation Function (Montgomery, 2009)

Other analytical method in order to detect autocorrelation at first lag in the residuals from a regression analysis is Durbin-Watson statistic (Durbin & Watson, 1971).

According to the literature survey conducted, the proposed approach when data are auto-correlated, is to fit an appropriate time series model (ARIMA) to the original observations and then use the residuals of this ARIMA model to plot the control charts. ARIMA (0,1,1) is used for nonstationary data while ARIMA (1,0,1) is used for stationary data (Alwan & Roberts, 1988). If the residuals are normally distributed and independent, with constant variance, the control charts will present process disturbances caused by any assignable causes as the same way they would be presented in a control chart monitoring the original data (Noskievicova, 2016). An example of auto-correlated process variable can be seen in Figure 1.3.

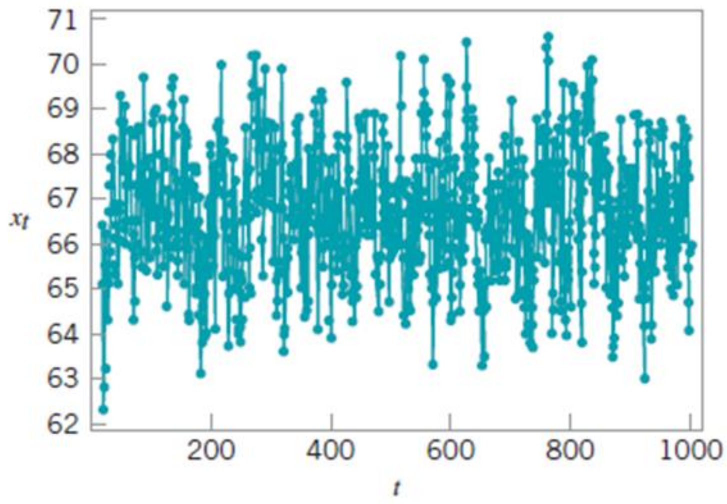

Figure 1.3. A Process Variable with Autocorrelation (Montgomery, 2009)

1.4.3. Control Charts for Non-Normal Data

An important assumption to construct SCC is normality. Individuals control charts are not robust to non-normal data. However, EWMACC is insensitive to non-normal data (Montgomery, 2009). In chemical industry, mostly Individuals control chart is used to monitor the quality characteristics which are not always normally distributed. If data is not normally distributed, the conclusions which are drawn from the control charts on process behavior can be misleading. (Box & Cox, 1964) In order to understand whether the data is normally distributed or not; statistical process control tools like histogram and probability plot can be used.

Johnson Transformation or Box-Cox Transformation can be used, if the data is not normally distributed, to construct Individuals control chart. After transforming the data, Individuals control chart can be constructed. Johnson distribution can be used to normalize the data via transformation when data are not normally distributed. Most of the standard continuous distributions are able to be approximated via Johnson's approach (Bersimis, Psarakis and Panaretos, 2007). Furthermore, Box-Cox transformation for normalization is efficient tool in order to be able to construct robust control charts (Mason & Young, 2002).

1.5. Multivariate Statistical Process Control

One type of the variables control charts is MVariable Control Charts (MVCC). MVCC display correlation or dependence of variables that jointly affects process parameters (Montgomery, 2009).

Previously, process monitoring and control regarding to the univariate perspective are mentioned.It is assumed that there is only one process output variable or quality characteristic of interest. In practice, most of the data are MV. Furthermore, most of the situations in industry for controlling and monitoring more than one quality characteristics is necessary. These quality characteristics might be related. However, monitoring these related quality characteristics independently might be quite misleading. If the data include correlated variables, creating separate control charts for each variable will be misleading because that the process is affected jointly by the variables. And also if univariate control charts are used as separately in MV situation; type I error probability and the probability of a point falling within the control limits are not going to be equal to their expected values. In order to deal with this situation, MV SPC Charts are used (Hogg and Craig, 1978). There are two phases while constructing the MV control charts. Phase I is used to determine the control limits with sample data and phase II is used to monitor the process with future data (Gagniuc, 2017).

In chemical industry, it is very often that the quality characteristics are not independent and they might affect each other. For example while one quality characteristic is increasing, other one might also be increasing, which is positive relation between these two characteristics. Also, while one quality characteristic is increasing, other might decrease which is a negative relation. When this relation is ignored by the analysts, the chance of detecting out-of-control situations will decrease. MV process control can be applied to all kinds of univariate control charts such as MV Shewhart control charts, MV CUSUMCCs and MV EWMACC (Montgomery, 2009).

1.5.1. The Multivariate Process Data

When more than one continuous variable are collected from the same process, the data can be called as MV process data. The multiple variables can be monitored by using one MV control chart if the data are correlated. (Montgomery, 2009)

Creating a correlation matrix of variables are used to decide whether to use univariate or MV control chart. Applying a MV control chart can be considered if the variables are correlated.

The advantages of the MV control charts if the data is MV process data are: (Mitra, 1998)

- The true control region for variables can be presented
- Rate of type I error can be maintained
- All the correlated variables can be monitored on a single chart with a single control limit

1.5.2. The Hotelling's T^2 Control Chart

The Hotelling T^2 control chart is the most common MV process-monitoring and control procedure in order to monitor the mean vector of the process. Hotelling's T^2 combines the dispersion and mean of more than one variable. There are two different application areas of Hotelling T^2: one for sub-grouped data and another for individual observations (Montgomery, 2009; Costa, Corolino and Oliveria, 2014).

The procedure which is used to monitor the process may be represented graphically as in Figure 1.4.

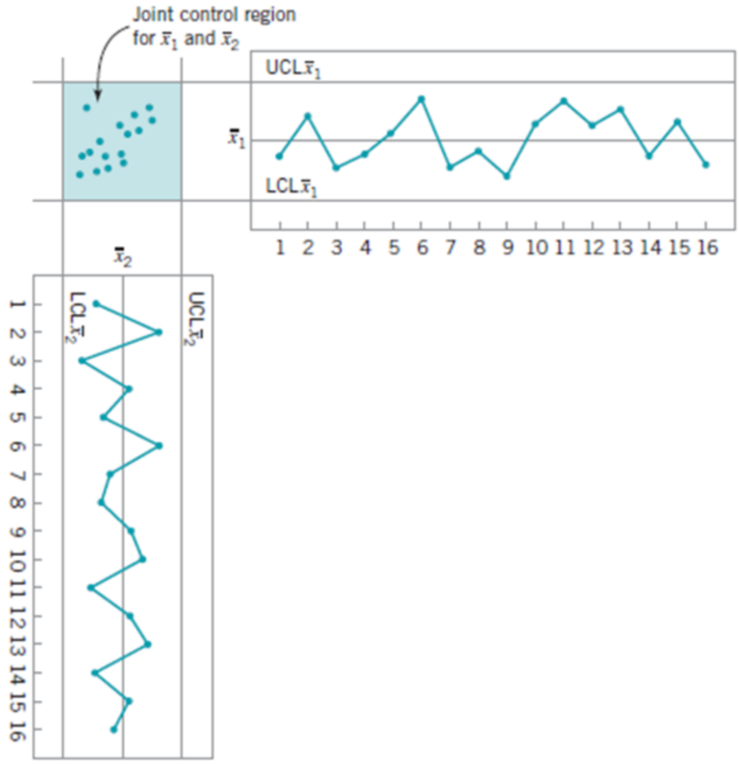

Figure 1.4. Control Region Using Independent Control Limits (Montgomery, 2009)

There are two cases for MV analysis. First case is; there are two arbitrary variables as x_1 and x_2, which are independent; that is σ_{12} is equal to 0. According to the first case, the general equation represents an ellipse centered at (μ1, μ2) with principal axes parallel to $\bar{x}_1$, $\bar{x}_2$ The second case is; there are again two quality characteristics which are dependent, then σ_{12} is not equal to 0. So the related graphical display is different as follows. When two quality characteristics are dependent, the principal axes of the ellipse are no longer parallel to $\bar{x}_1$, $\bar{x}_2$ axes (Montgomery, 2009). However, in real life, mostly μ and σ are not known so $\bar{x}$ and s estimations can be used with n observations.

With the observation of p-variables, as an example $X' = (x_1, x_2,...,x_p)$, is given as $T^2 = (X - \bar{x})'S^{-1} (X - \bar{x})$ where the measure of the process center is represented by sample mean $\bar{x}$. The sample covariance matrix S gives information about individual variables and also shows the relationship within the elements of the observation vector (Mason & Young, 2001).

However, estimate S should not contain redundancies among the process variable. In order to meet this requirement, two quality characteristics must be perfectly correlated.

To deal with redundancy, one of the variables can be deleted from the study. Furthermore, the observations must be independent to be able to use T^2 statistic. However, in industrial applications especially in chemical industry, data is auto-correlated. However, as mentioned previously, time series models can be used to deal with auto-correlation. (Mason & Young, 2001)

The procedure which is used to monitor the process with independent and dependent variables may be represented graphically as in Figure 1.5.

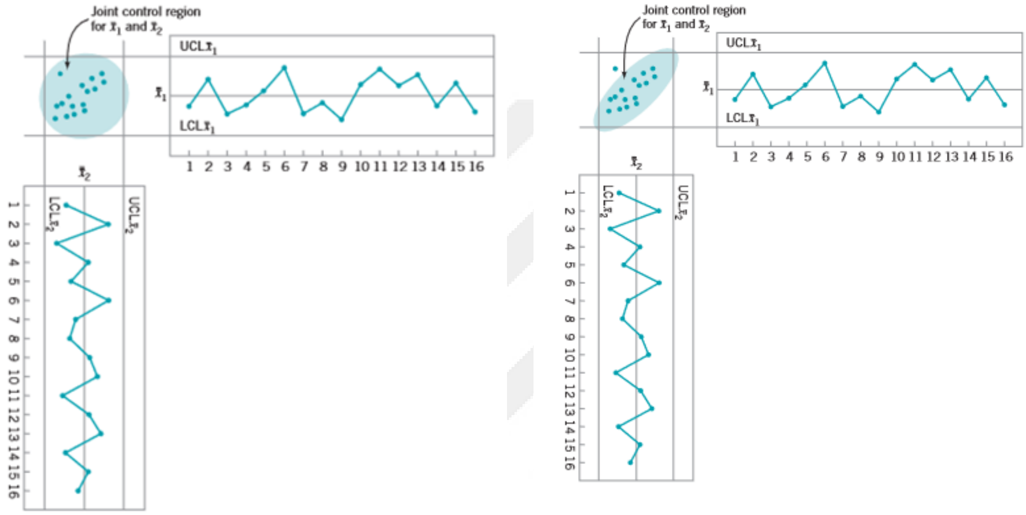

Figure 1.5. Control Ellipse for Two Independent Variables and Dependent Variables (Montgomery, 2009)

The disadvantages of the MV SPC charts are as follows according to Montgomery: (Montgomery, 2009)

- The time order of the plotted points is missed
- With more than two quality characteristics, it is very challenging to build the ellipse
- When more than two variables exist and the out of control situation appears; it is hard to understand which variable cause the out of control situation.
- MV control charts are more complicated to explain than standard SCCs
- The advantages of the multivariate SPC charts are as follows (Mason & Young, 2001):

- The most obvious advantage of the MV control charts is that it minimizes the number of control charts to manage.
- Most of the parameters are related to each other especially in chemical industry (ex. Solubility increases with the temperature increase). Because of these related parameters, considering them together is useful to understand the process correctly.

CHAPTER 2
LITERATURE REVIEW

There are several studies in the literature on various aspects of SPC Charts for Chemical and Process Industry by using univariate and MV control charts.

Elevli et al. (2009), conducted research to determine the effect of auto-correlation to determine process stability of colemanite concentrators. Individuals Control chart and Special Cause Control chart are compared in order to monitor the process while data is auto-correlated. When the data is auto-correlated, control charts which are extremely sensitive to the assumption of independence give a higher false alarm rate. They use ACF and PACF plots in order to monitor auto-correlation in data and then, the control chart is constructed with residuals obtained after fitting the process to an ARIMA model.

Mastrangelo and Montgomery (1995), mentioned in their study that auto-correlated data required alterations or changes to traditional control chart techniques. Violation of the fundamental theory of independent or autocorrelated data results in poor statistical performance and increased number of false alarms.

Alwan and Roberts (1988), introduce and explain statistical modeling and fitting of time series impacts and the purpose of standard control chart procedures to the residuals from the time series model fits when the process is not independent and identically distributed. They observe that ARIMA (0,1,1) model is appropriate for nonstationary processes and ARIMA (1,0,1) model is appropriate for stationary processes.

Rodriguez (1996), conducted a study which shows the application of software for statistical modeling in link with the Shewhart procedure. In the study, SAS is used to analyze data like diagnosing autocorrelation with autocorrelation plot, fitting ARIMA model, and histogram. He suggested that one of the most important assumptions normality is checked with graphical displays.

Bisgaard and Kulahci (2005), demonstrated with an industrial example, which is Temperature Control of a Ceramic Furnace. The study shows how to identify autocorrelation, explained its outcomes for standard control graphs, and described current software packages as MINITAB which makes simple to execute the computations required while dealing AC and using ARIMA time series models.

Lestander et al. (2012), mentioned that when there is more variables to monitor and control, SPC becomes more complex because of overloading information for process operators. However, MSPC can easily defeat the mentioned difficulty. They conducted a study which simulates MSPC, using principal component analysis and partial least squares regression, based on wood pallet production process data; which are used to mark changes in the monitored variables over time; and to forecast pallet drought to determine the potential of using multivariate statistical process control for monitoring and controlling in the wood pallet business.

Rao et al. (2013), propose that classical SPC methods are not optimal to control and monitor multiple variables. Because the impact of a variable may be related to the impacts of another correlated variable. Furthermore, when there is a large number of control charts of each process variable, univariate control charts are hard to control and interpret. They suggested an alternative approach to build a singular MV T2 control chart which helps to minimize the occurrence of false alarms. In their study, they demonstrated a study which shows the use of MSPC charts to control production process and also, T2 diagnosis is applied to analyze the critical to quality process variables.

In the relevant literature, univariate and MV control charts are investigated individually. The difference of this study from previous ones, and in turn its contribution, a comparison of univariate and MV control charts for the same process is introduced.

CHAPTER 3
PROBLEM DEFINITION

Applying SPC correctly is one of the necessary tools to produce high-quality products. An important advantage of the SPC is that it can be used for any process. One of the important assumptions in traditional SCC is independent observations of the method. However, in some industries like chemical processes, food industries and refinery productions, there is correlated data because of consecutive measurements.

The goal of the master thesis is to construct control charts as univariate and multivariate for monitoring auto-correlated processes and to compare the univariate and MV control charts for the same process. In the relevant literature, univariate and MVCCs are investigated individually. However, in this study, univariate control charts and MV control charts are compared.

In the case of chemical and process industries, where individual measurements occur frequently, the most appropriate control chart to be used is the Individual Measurements and Moving Range (I-MR) control chart. Unfortunately, the performance of these charts is dramatically influenced by even moderate violations of the normality assumption.

It is known that in chemical process industries, the data is usually auto-correlated and sometimes not normally distributed.

The real case of glycerin production is studied in this research. Glycerin is produced by hydrolysis treatment of fats and oils. The crude glycerin is used as a raw material which contains some amount of water and impurities. The existing amount of water and impurities are removed by sending it to the reactor tank with water and hydrochloric acid. Calcium Hydroxide ($CaOH_2$) and Caustic ($NaOH$) are added to glycerin. Most of the impurities in the glycerin are separated with a high–efficiency pressure filter. Moreover, after filtration, glycerin water is collected by vacuum concentrator usage in order to eliminate excess water. The removing excess water process includes burning the liquid glycerin below pressure by Evaporating Colons and afterward spraying the glycerin into a vacuum chamber which helps to evaporate the water as steam and by this process glycerin remains as a liquid. Furthermore, the glycerin is clarified by distillation by using fine activated charcoal. Charcoal treatment is repeated in order to be sure about the removal of impurities in the glycerin. Glycerin at this stage is micro-filtered and the glycerin must be perfectly transparent. The filtration is applied to reduce all remaining contaminants that may be started and to remove the smell and color. At the end of this step, the final product, which is pure glycerin, is obtained with excellent color stability upon heating. In the following figure the process flow chart can be seen:

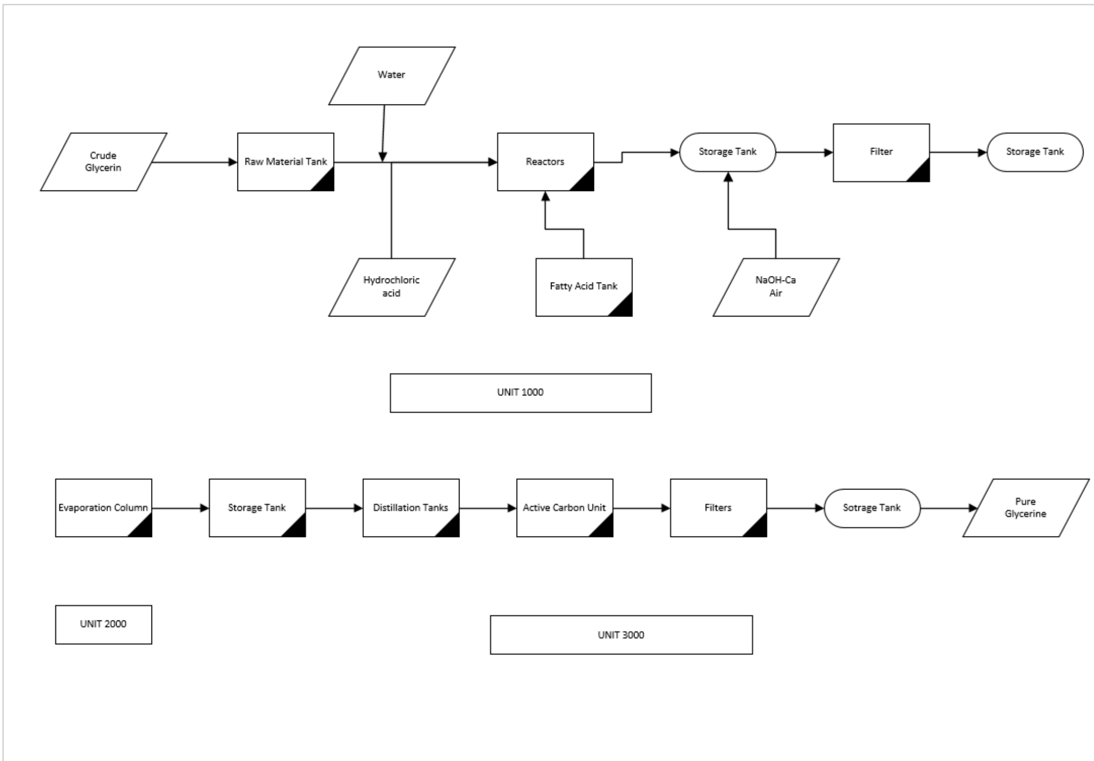

Figure 3.1. Model of Production of Glycerin

CHAPTER 4
METHODOLOGY

In this study, there are four quality characteristics of pharmaceutical glycerin (the final product of our chemical production process), that we want to monitor and control. These critical to quality characteristics are: density, ester content, glycerin content and humidity.

For all these four quality characteristics, the methodological approach followed was the same. An initial data analysis is conducted to understand if the data is auto-correlated and whether they are normally distributed. Our final goal is to prepare appropriate Control Charts, based on the historical data given to us by the company, for the monitoring and control of the production process in the future. It is worth mentioning that Control Charts are an on-line tool for quality control.

The guidelines for univariate control chart selection diagram is shown in Figure 4.1.

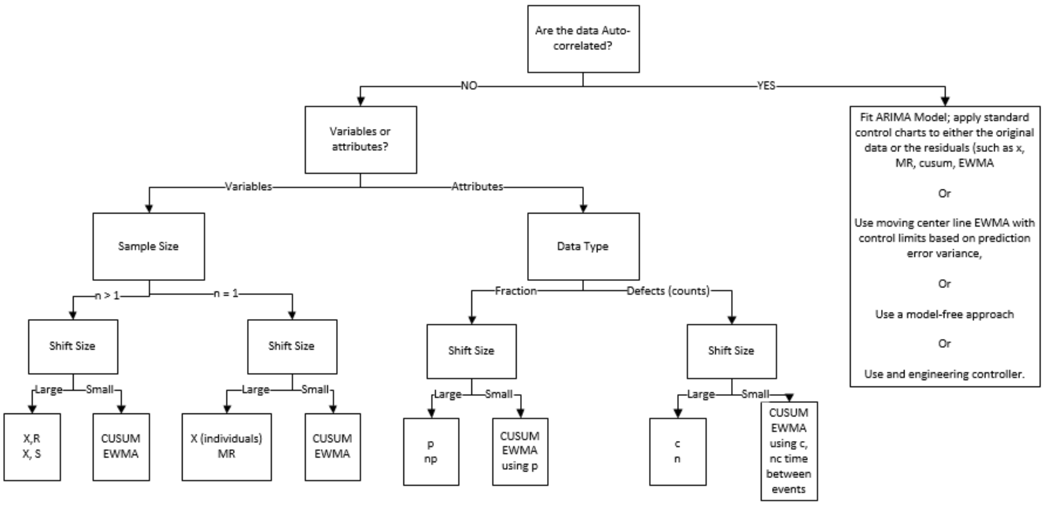

Figure 4.1. Guidelines for Univariate Control Chart Selection (Montgomery, 2009)

In order to be able to apply SPC for our study, there is a path to follow. There are two cases to check for data, auto-correlation and normality of the data.

Firstly, it has to be decided whether the data is auto-correlated or not by using Auto-Correlation Function (ACF) graph. If the data is auto-correlated, then data must be fitted to appropriate TSM (ARIMA). After fitting the data, the residuals obtained from the process should be checked for assumptions, which are independency and normality. After providing all assumptions, I-MR and EWMACCs can be plotted by using residuals.

Secondly, if the data is not auto-correlated, data must be checked whether the data is normally distributed or not. If the data is not auto-correlated and is normally distributed, I-MR and EWMACCs can be plotted directly. If the data is non-normal, the data must be transformed by using the Box-Cox transformation or Johnson transformation. After the transformation of data, I-MR and EWMACCs can be plotted. However, if data is not normally distributed, EWMACC still can be used to monitor critical to quality characteristics (Montgomery, 2009).

For analysis, an amount of at least 100 observations has to be selected. Therefore, from the data available we selected 105 observations for each characteristic that corresponded to the period starting from the 25th of February until the 2nd of August 2016. However, the data are not continuous, there are large gaps between them. For instance, data is missing between 7th March and 4th April and between 30th June and 19th July.

In general, we can summarize the procedure after the initial data analysis as follows:

- As anticipated, after the initial analysis of the data, we observed that the data are auto-correlated for all four-quality characteristics. Since individual measurements in a chemical industry arrive consecutively and sampling is happening in short time intervals auto-correlation was something that was expected.
- One of the quality characteristics (humidity) investigated is also not normally distributed.

According to the literature survey conducted by Alwan and Roberts, 1988, the approach proposed when data are auto-correlated, is to fit an appropriate TSM

(ARIMA) to the original data observations and then use the residuals of this ARIMA model to plot the control charts. If the residuals are normally distributed and independent, with constant variance, the control charts that will be applied will present process disturbances, caused by any assignable causes, the same way they would be presented in a control chart monitoring the original data (Noskievicova, 2016).

In addition, in order to deal with the non-normality of the data, Johnson Transformation is used. After eliminating the non-normality and auto-correlation, I-MR Control Charts and EWMACCs are constructed in order to monitor the process. I-MR Control Charts are used because they are more capable in detecting large shifts in the process mean and EWMACCs are used to capture small shifts.

At first, the data is analyzed by constructing a time series plot, a histogram, a probability plot and the ACF graph in order to see whether the data is normal and non-auto-correlated. After that, if they are auto-correlated, the appropriate time series model (ARIMA) is fitted to the original observations and the normality and independence of the residuals of the model is tested. Finally, the EWMACC and I-MR Charts are developed.

If the data was not auto-correlated and normally distributed, the appropriate EWMACC and I-MR Control Charts would be applied directly, using the original observations.

4.1. Data Analysis of Quality Characteristics

For all these four quality characteristics, data are analyzed for behavior of process, normality of observations and auto-correlation of data:

Glycerin Content

For the initial analysis of the data, time series plot, histogram, probability plot and auto-correlation function graphs are constructed.

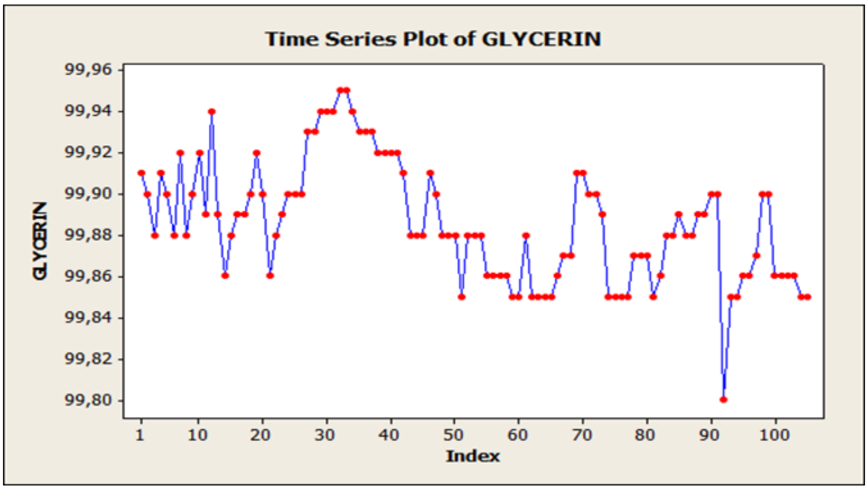

Figure 4.2. Time Series Plot of Glycerin Content

From the time series plot given in Figure 4.2, the behavior of the process can be understood. The behavior of the process looks stationary (data which are quite stable and vary around a target mean) however a shift towards a lower level can be observed after observation 57. Also the variability of the data is higher after that point in time.

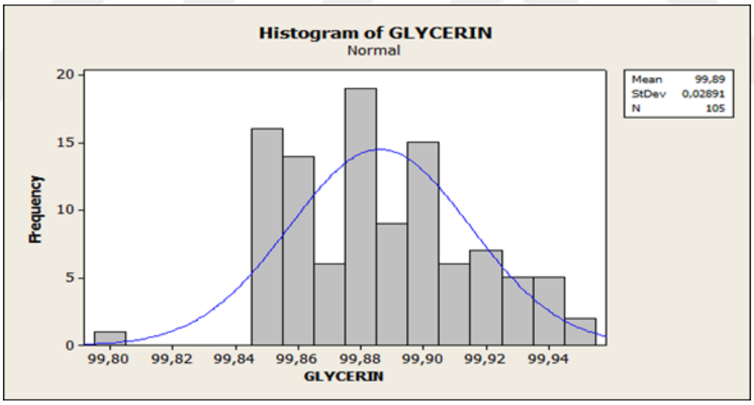

Figure 4.3. Histogram of Glycerin Figure

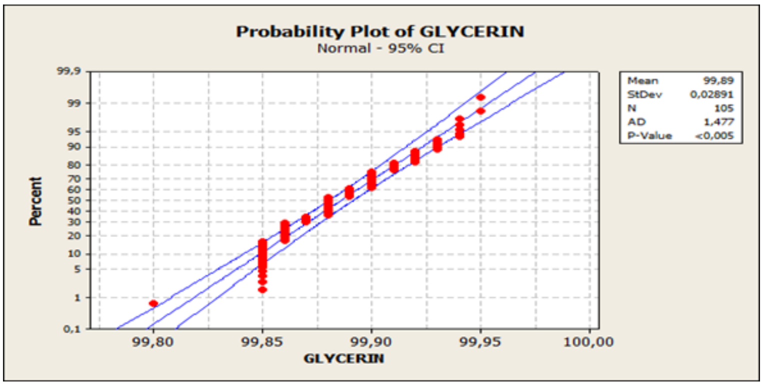

Figure 4.4. Probability Plot of Glycerin

By the help of the probability plot and histogram shown in Figure 4.3 and Figure 4.4 respectively, it can be seen that the data are not normal. From the histogram above, we can see that the data are not symmetrical and appear to be left skewed.

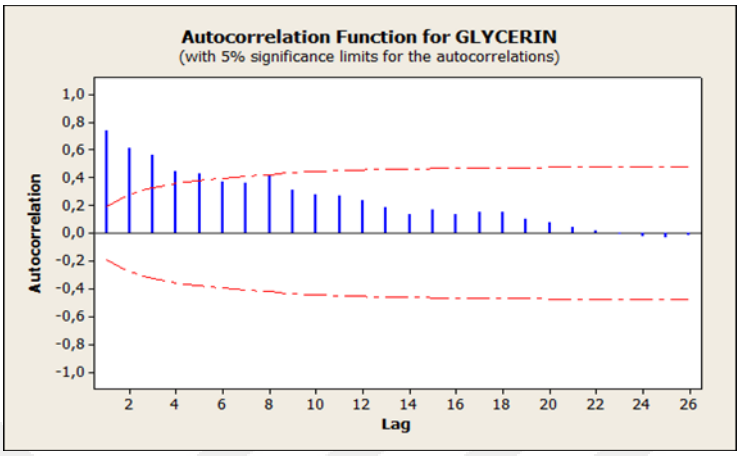

Figure 4.5. Autocorrelation Function for Glycerin Content

Furthermore, according to the auto-correlation function graph given in Figure 4.5, it is seen that the data are highly auto-correlated especially at the first 3 lags.

Ester Content

Time series plot, histogram, probability plot and auto-correlation function graphs are constructed for ester content data.

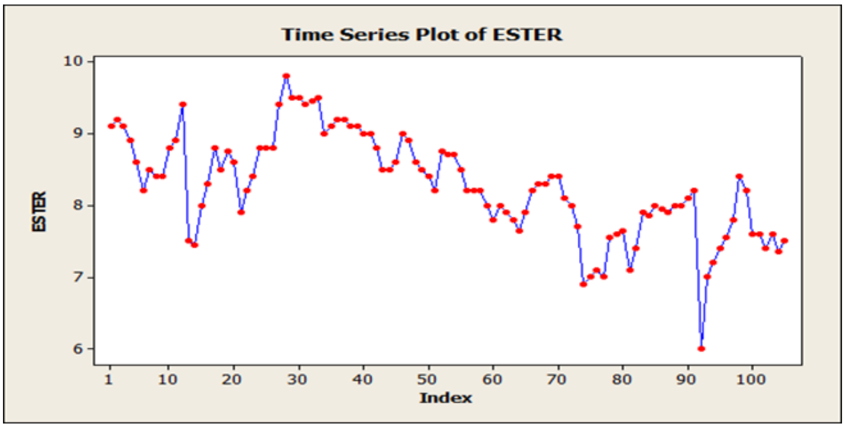

Figure 4.6. Time Series Plot of Ester Content

As can be seen from the time series plot given in Figure 4.6, the behavior of the process is non-stationary. This type of behavior occurs frequently in the chemical and process industries. The process appears to be quite unstable because of the drifts without any sense of a stable or fixed mean.

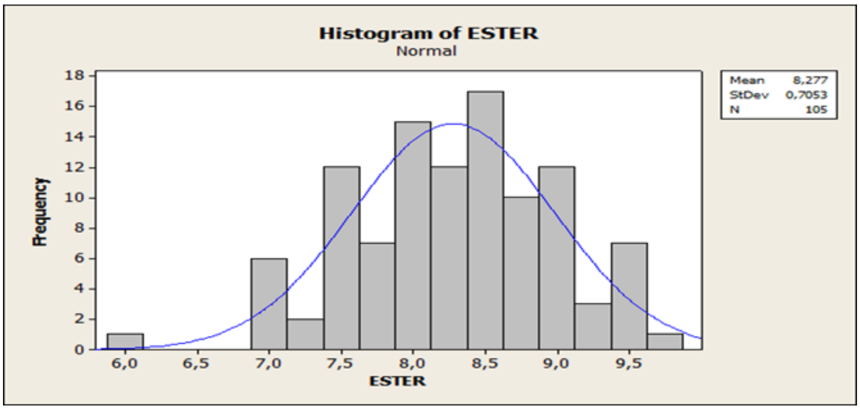

Figure 4.7. Histogram of Ester Content

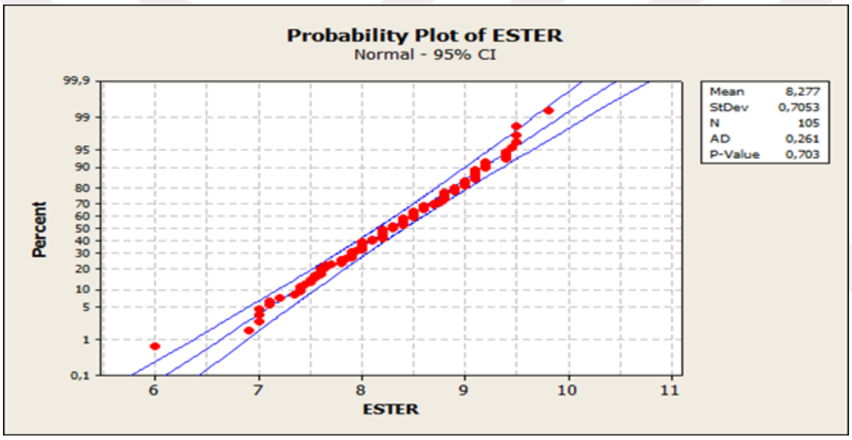

Figure 4.8. Probability Plot of Ester Content

According to the probability plot given in Figure 4.7 and histogram given in Figure 4.8, it can be seen that the data are normally distributed.

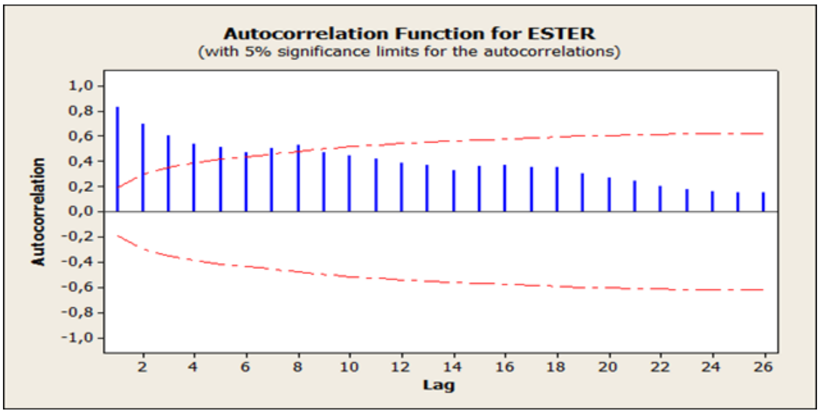

Figure 4.9. Autocorrelation Function of Ester Content

However, as it is seen from the auto-correlation function graph shown in Figure 3.10, that the data are highly auto-correlated.

Density

Time series plot, histogram, probability plot and ACF graphs are constructed for density data.

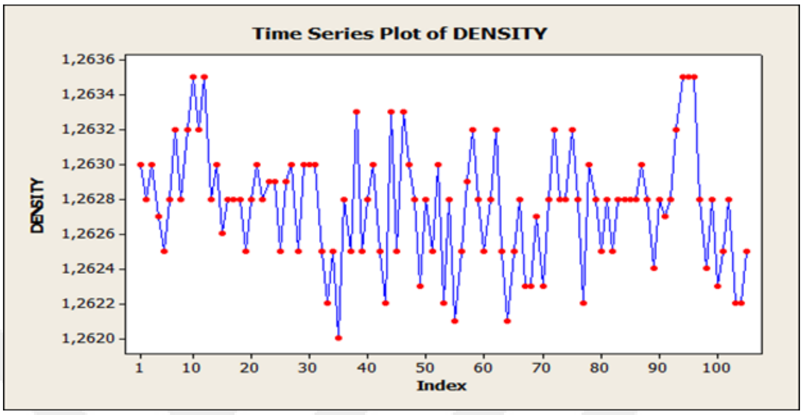

Figure 4.10. Time Series Plot of Density

From the time series plot in Figure 4.10, the behavior of the process is stationary.

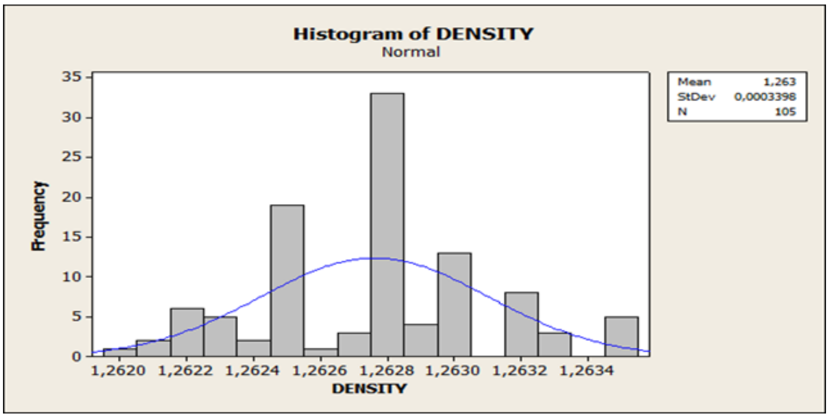

Figure 4.11. Histogram of Density

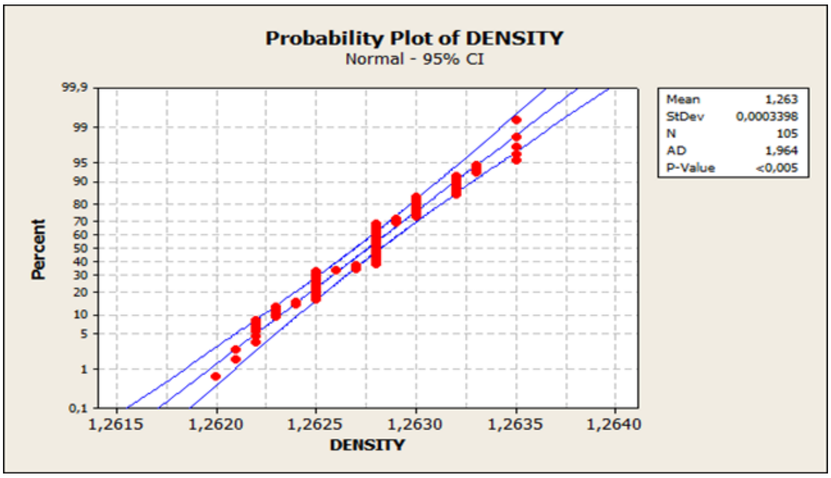

Figure 4.12. Probability Plot of Density

According to the probability plot and histogram given in Figures 4.11 and 4.12 respectively, it can be seen that the data are not normal.

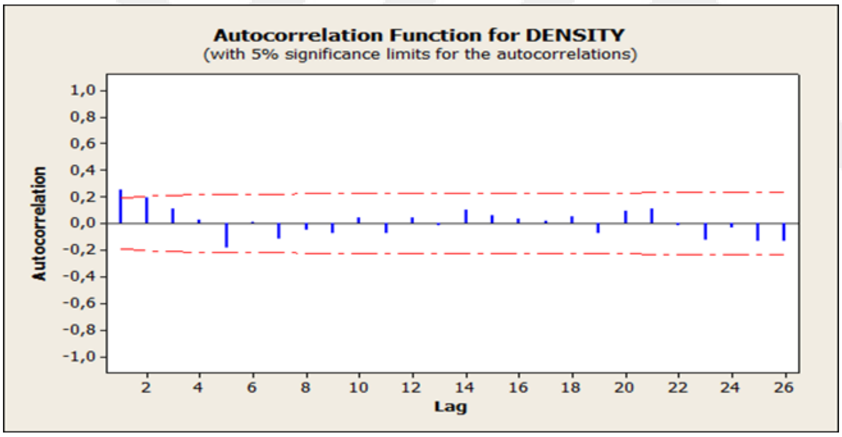

Figure 4.13. Autocorrelation Function of Density

From the ACF graph shown in Figure 4.13, the data are slightly auto-correlated.

Humidity

Time series plot, histogram, probability plot and ACF graphs are constructed for humidity data.

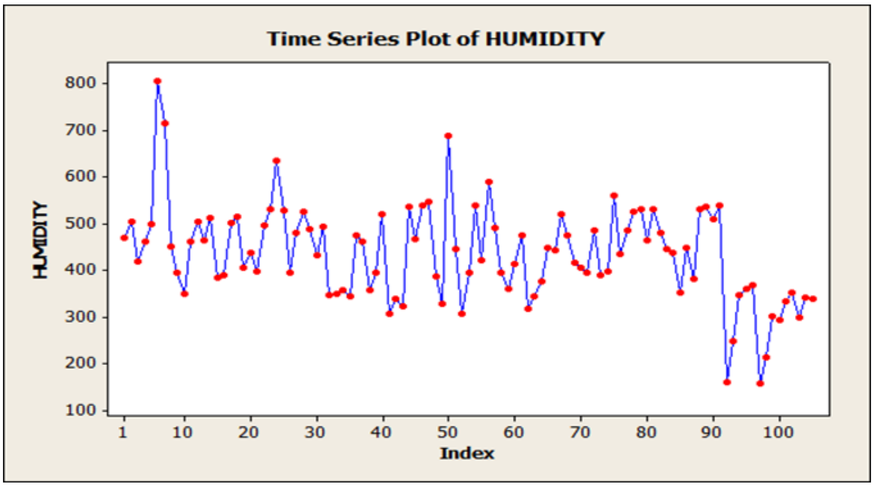

Figure 4.14. Time Series Plot of Humidity

According to the time series plot in Figure 4.14, the behavior of the process is in general stationary, however as it is seen from the graph above, there is a shift towards a lower mean value, starting from the 20/07 as the last part of the data show.

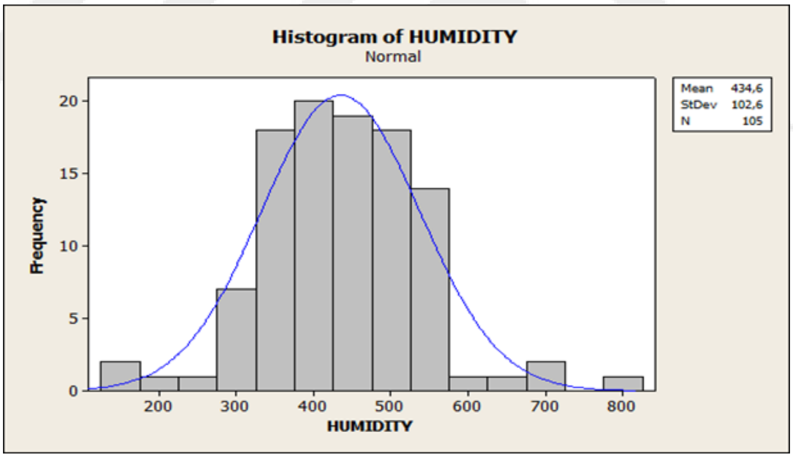

Figure 4.15. Histogram of Humidity

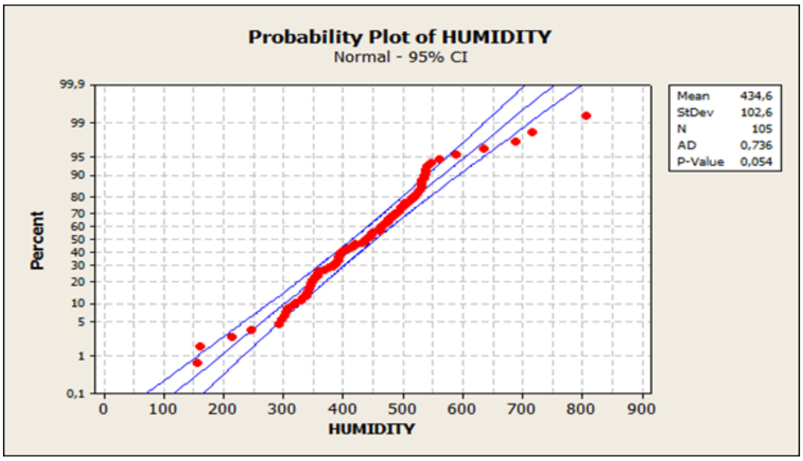

Figure 4.16. Probability Plot of Humidity

According to the probability plot seen in Figure 4.15 and histogram in Figure 4.16, it can be seen that the data are not normal. As it can be seen in the probability plot, there are fluctuations of points around the line and quite a few outliers.

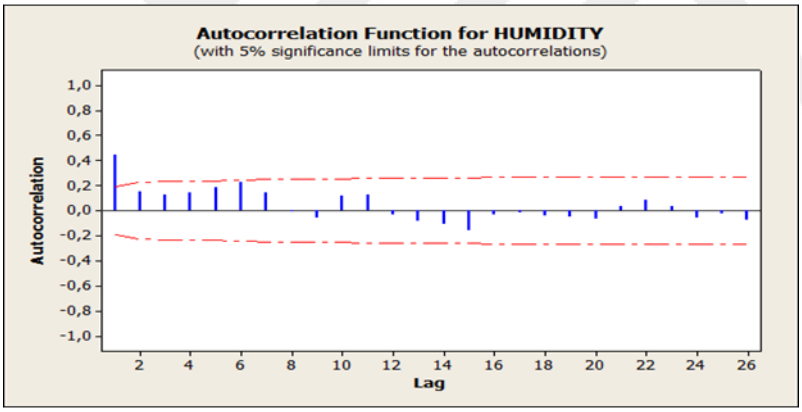

Figure 4.17. Autocorrelation Function of Humidity

From the ACF graph shown in Figure 4.17, the data are auto-correlated at the first lag.

4.2. Application of Control Charts

Case Demonstration A: Univariate Control Charts

According to the data analysis of the quality characteristic shown in Data Analysis section, all four quality characteristics data are correlated. Time series approach is used to deal with the auto-correlated data. TSM helps to remove the auto-correlation from the data. After fitting the data to the appropriate TSM, control chart is constructed by using the residuals.

Modeling the quality characteristic x_t as

$$x_t = \xi + \o x_{t-1} + \varepsilon_t \tag{1}$$

Where ξ and ø (-1<ø<1) are unknown constants and ε_t is normally and independently distributed with mean zero and standard deviation σ.

The residuals are approximately normally and independently distributed with mean zero and constant variance, shown as

$$e_t = x_t - \hat{x}_t \tag{2}$$

The conventional control charts are constructed with the sequence of residuals. It is pointed out that the residual control charts are not sensitive to detect small shifts in the process mean. In order to be able to improve the sensitivity of the residual control charts, it is recommended to use CUSUMCC or EWMACCs on residuals. Furthermore, Montgomery is stated that EWMACC based procedure is effective to control performance and shift detection. (Stoumbus & Reynolds, 200)

CUSUM CONTROL CHART

The Cumulative Sum control chart is used to monitor the process mean which is based on the samples; the samples are taken time basely such as hours, shifts, days, weeks etc. from the related process. The cumulative sum control chart indicates accumulation of current and previous information of data. Because of that, cumulative sum control chart is better for detecting the small shift in the mean of the process.

The tabular CUSUMCC works by using the accumulating derivations from μ_0 which are above the target with one statistic C+ and accumulating derivations from μ_0 which are below target with another statistic C-. The statistic C+ and C- are called as one-sided UC and LC, respectively (Montgomery, 2009).

The Tabular CUSUM:

$$C_i^+=\max[0,\ x_i-(\mu_0+K)+C_{i-1}^+]$$
$$C_i^-=\max[0,\ (\mu_0-K)-x_i+C_{i-1}^-] \tag{3}$$

where the starting values are $C_0^+=C_0^-=0$

K is he reference value and it is usually taken as halfway between the target μ_0 and the out of control value of mean μ_1 for detecting quickly (Montgomery, 2009).

$$K=\frac{\delta}{2}\sigma=\frac{|\mu_1-\mu_0|}{2} \tag{4}$$

The EWMACC is an alternative to the CUSUMCC. The EWMACC has similar properties as CUSUMCC and also which is useful for detecting smaller shifts in the process mean.

EWMA CONTROL CHART

The EWMACC is useful to detect small shifts in the process mean as cumulative sum control chart but it is easier to perform and also it is a good alternative to the Shewhart control chart. The samples are taken based on time such as hours, shifts, days, weeks etc. from the related process. The measurements of the sample data at a given time sequence generates a subgroup. The EWMACC lean on the target value and the standard deviation which can be known or estimated. Because of this reason, it is better to use cumulative sum control chart after establishing process control.

The definition of the EWMA is as follows:

$$z_i=\lambda x_i+(1-\lambda)z_{i-1} \tag{5}$$

where $0 < \lambda \leq 1$ is a constant and the starting value is the process target;

$$z_0 = \mu_0 \tag{6}$$

In some cases as a starting value can be used as the average of the preliminary data $z_0=\bar{x}$.

EWMACC equation can be demonstrated as

$$\begin{aligned} z_i&=\lambda x_i+(1-\lambda)[\lambda x_{i-1}+(1-\lambda)z_{i-2}] \\ &=\lambda x_i+\lambda(1-\lambda)x_{i-1}+(1-\lambda)^2 z_{i-2} \end{aligned} \tag{7}$$

In order to continue substitute repetitively for z_{i-j} = 2,3,….,t, the following equation is obtained

$$z_i=\sum_{j=0}^{i-1}(1-\lambda)^j x_{i-j}+(1-\lambda)^i z_0 \tag{8}$$

So, the weights sum to unity

$$\lambda\sum_{j=0}^{i-1}(1-\lambda)^j=\lambda\left[\frac{1-(1-\lambda)^i}{1-(1-\lambda)}\right]=1-(1-\lambda)^i \tag{9}$$

When λ is equal to 0.2, then 0.2 is assigned to the sample mean as a weight. Using EWMACCs in time series modeling is very common (Mastrangelo & Montgomery, 1995) (Box, Jenkins and Reinsel, 1994).

All past and current information of the process are used in EWMACCs as weighted, because of that EWMACCs are insensitive to the normality assumption. With insensitiveness to the normality assumption, individual observation can be used.

If the observations are independent random variables with the variance σ^2, the variance of z_i is as follows

$$\sigma_{z_i}^2=\sigma^2\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2i}\right] \tag{10}$$

EWMACC is constructed by plotting z_i versus the sample number i (or time). The CL, UCL and LCL are calculated as

The EWMACC:

$$UCL=\mu_0+L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \tag{11}$$

$$\text{Center Line}=\mu_0$$

$$LCL=\mu_0-L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \tag{12}$$

CONSTRUCTING SHEWHART CONTROL CHARTS

Glycerin Content

As mentioned at section 4.1, data analysis of quality characteristics part of glycerin content, if data do not satisfy the assumption of normality and are not independent, the SPC charts would not be reliable. If the control chart is constructed with the individual observations which are not normally distributed and especially not independent, the control chart would be misleading as shown in Figure 4.18:

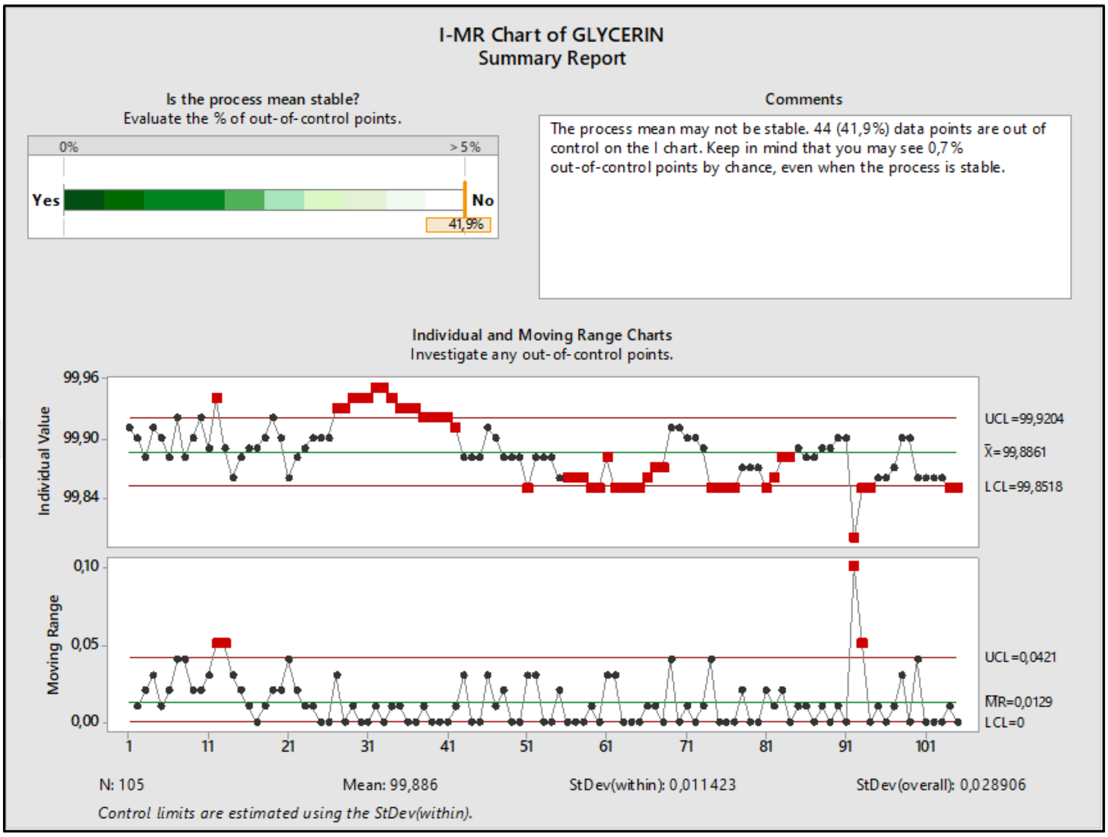

Figure 4.18. Individual and Moving Average Control Chart of Glycerin

According to the I-MR control chart, there are 2 different types of out of control situations:

For individuals control chart; 12, 27-37, 51, 59, 60, 62-65, 74-77, 81, 92-94, 104 points are plotted outside the control limits and 31-42, 56-68, 82-84 points show that there is a shift in the process mean. For moving range control chart, 12, 13, 92, 93 points are plotted outside the control limits. Since the data is auto-correlated, and the independency rule is violated, these out of control points are anticipated as false alarms.

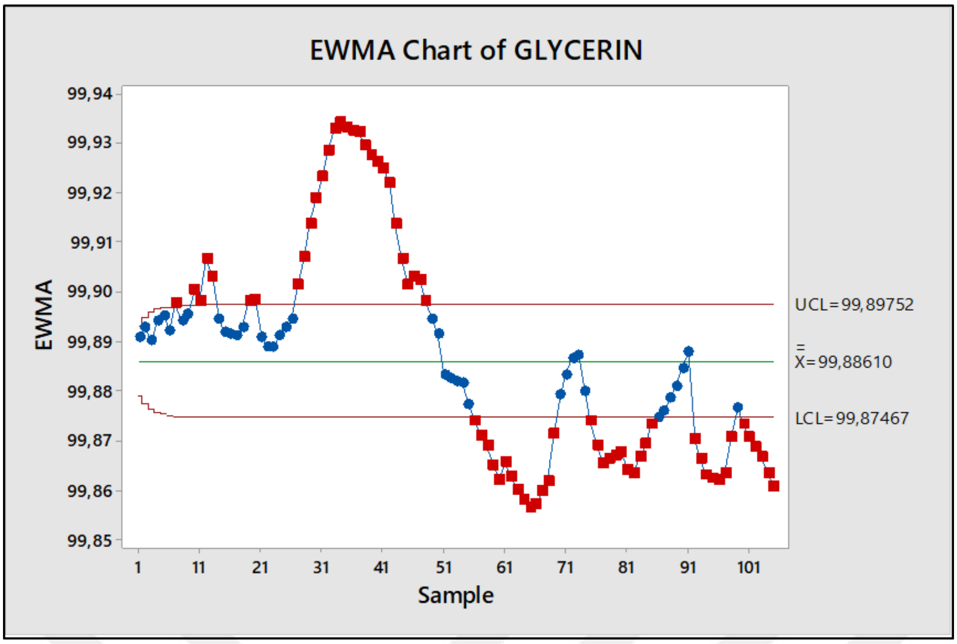

Figure 4.19. EWMACC of Glycerin

As it is seen from the EWMACC, there are too many out of control points.

In order to deal with the auto-correlation, a time series model ARIMA is applied. The most suitable model for the glycerin content data is identified as the ARIMA (1,0,0). After the appropriate TSM is fitted to data, normality and independence of the residuals of the model is tested. According to figure 4.20, the residuals of the model are normal and independent. Furthermore, it is understood from the graphs below that the time series model ARIMA (1,0,0) nicely fits the data.

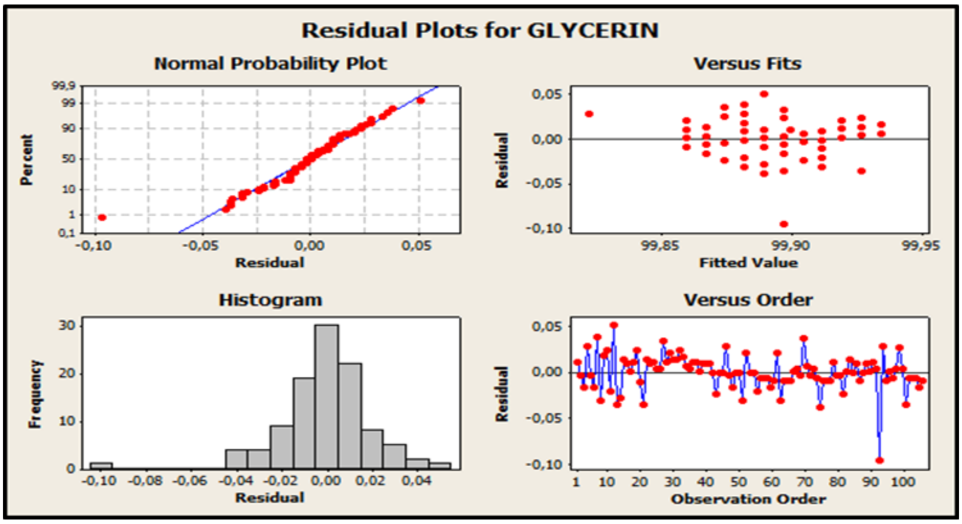

Figure 4.20. Testing Residuals of Time Series Model

Afterwards, the I-MR Control Chart and EWMACC are constructed in order to detect the large and small shifts in the process mean respectively.

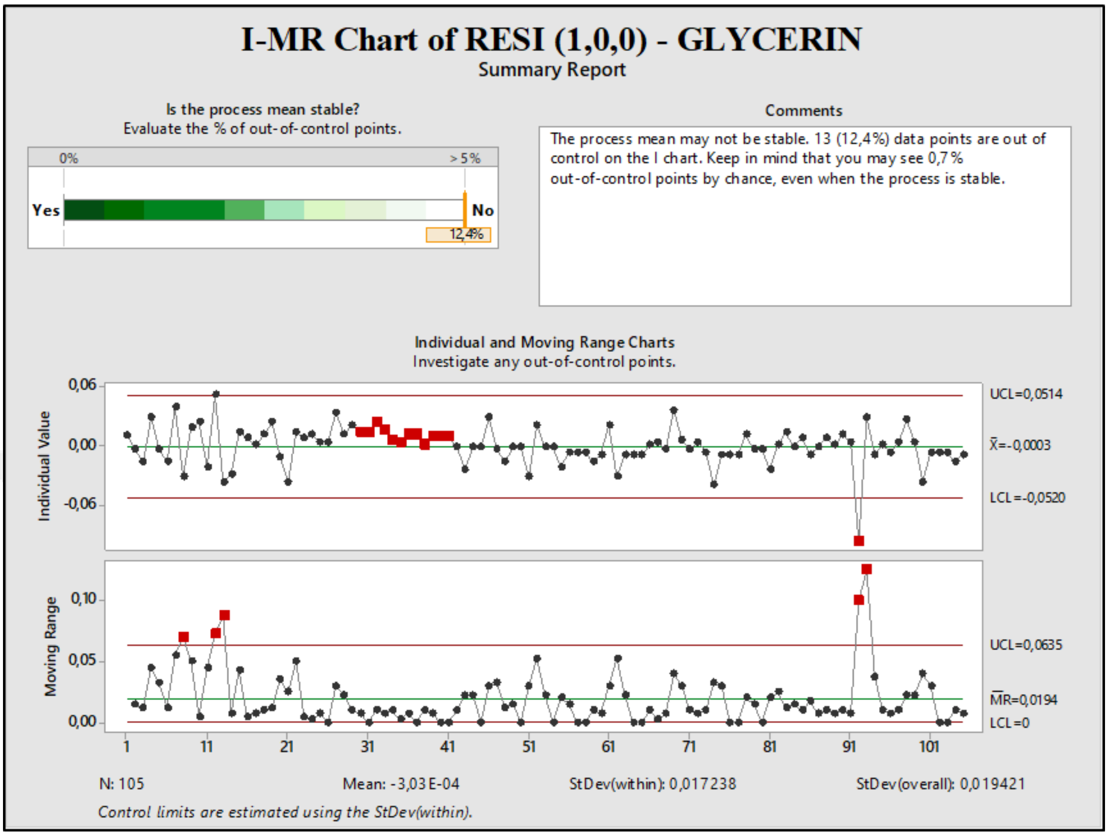

Figure 4.21. I-MR Chart of Residuals of Glycerin Content

Figure 4.21. According to the I-MR Chart, there is an out of control situation because of the consecutive data points very close to the center-line between 28 and 41 and also an out of control situation at 92, which is the first data point after a break between 91 until 92.

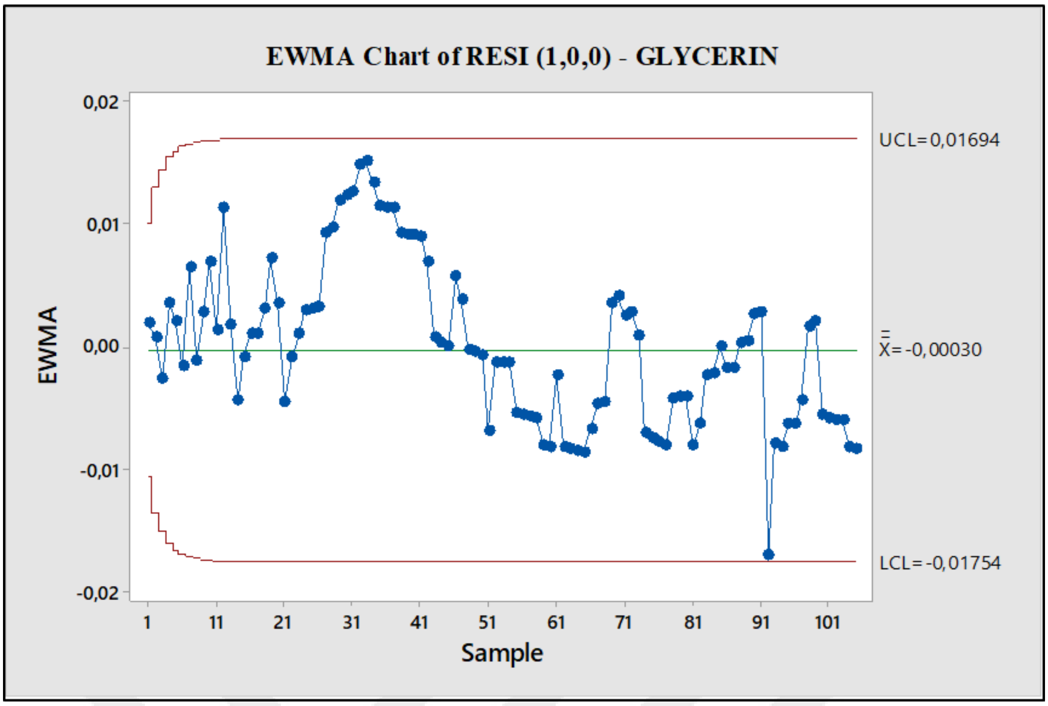

Figure 4.22. EWMACC Residuals of Glycerin Content

According to the EWMACC in Figure 4.22, there are no small shifts detected in the process mean (the process looks in control), but the out of control point at the 92th sample can be also noted.

Ester Content

As mentioned at section 4.1, data analysis of quality characteristics part of ester content, as it is seen from the auto-correlation function graph, that the data are highly auto-correlated.

In order to deal with the auto-correlation, the time series model ARIMA is applied. The most suitable model for ester content data is identified as the ARIMA (0,1,1). After the appropriate TSM is fitted to the data, normality and independence of the residuals of the model is tested. According to figure 4.23, the residuals of the model are normal and independent. Furthermore it is understood that the time series model ARIMA (0,1,1) nicely fits the data.

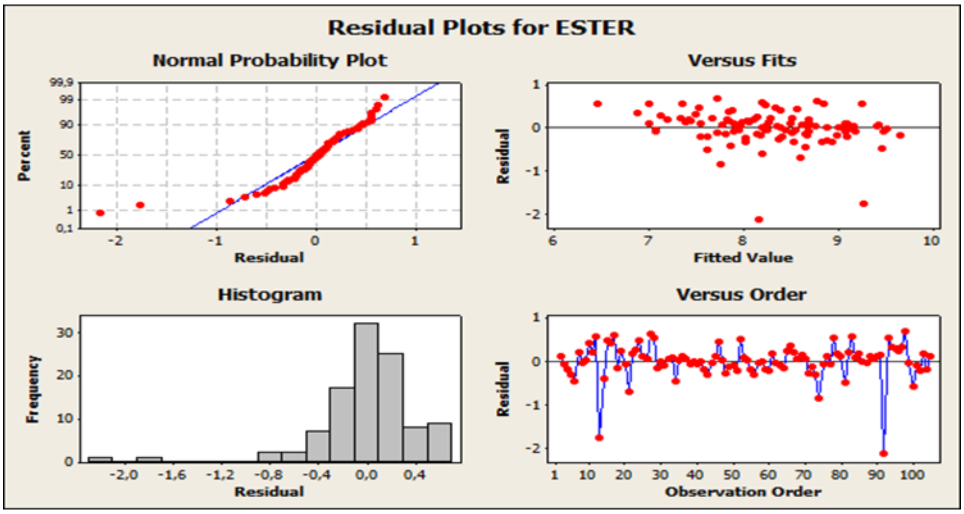

Figure 4.23. Testing Residuals of Time Series Model

Then the I-MR Control Chart is constructed in order to be able to detect large shifts and EWMACC is constructed in order to detect small shifts in the process mean which are shown in Figure 4.23.

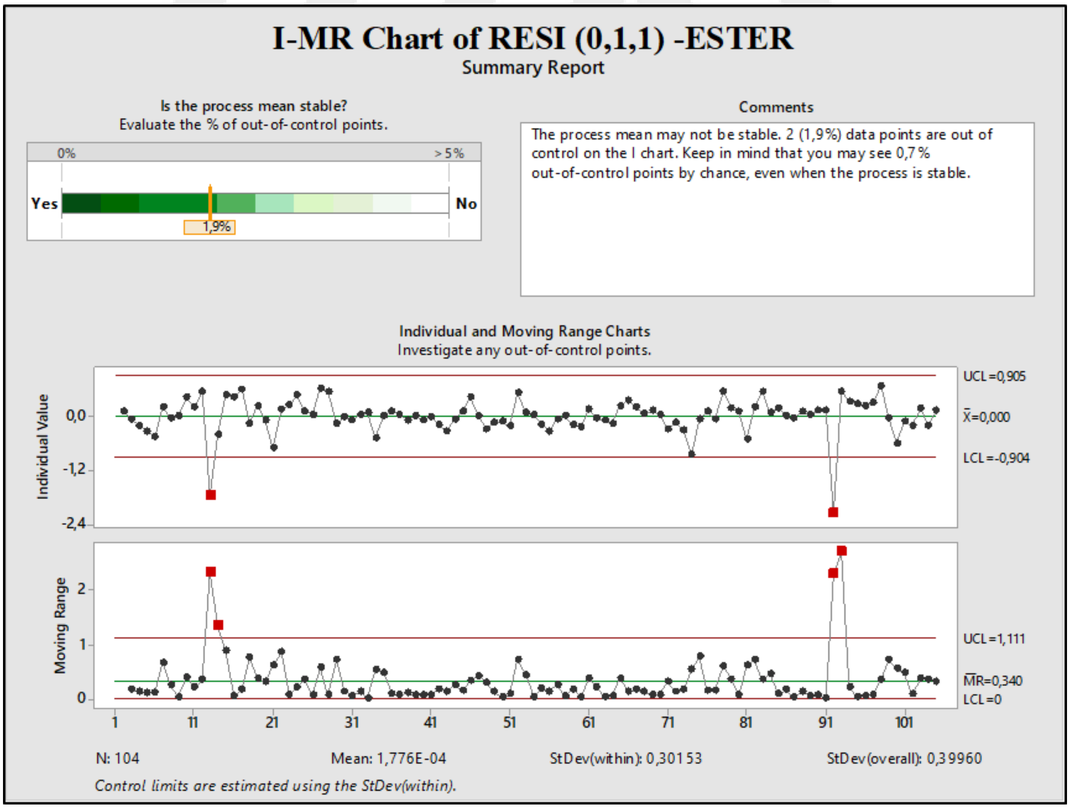

Figure 4.24. I-MR Chart of Residuals of Ester Content

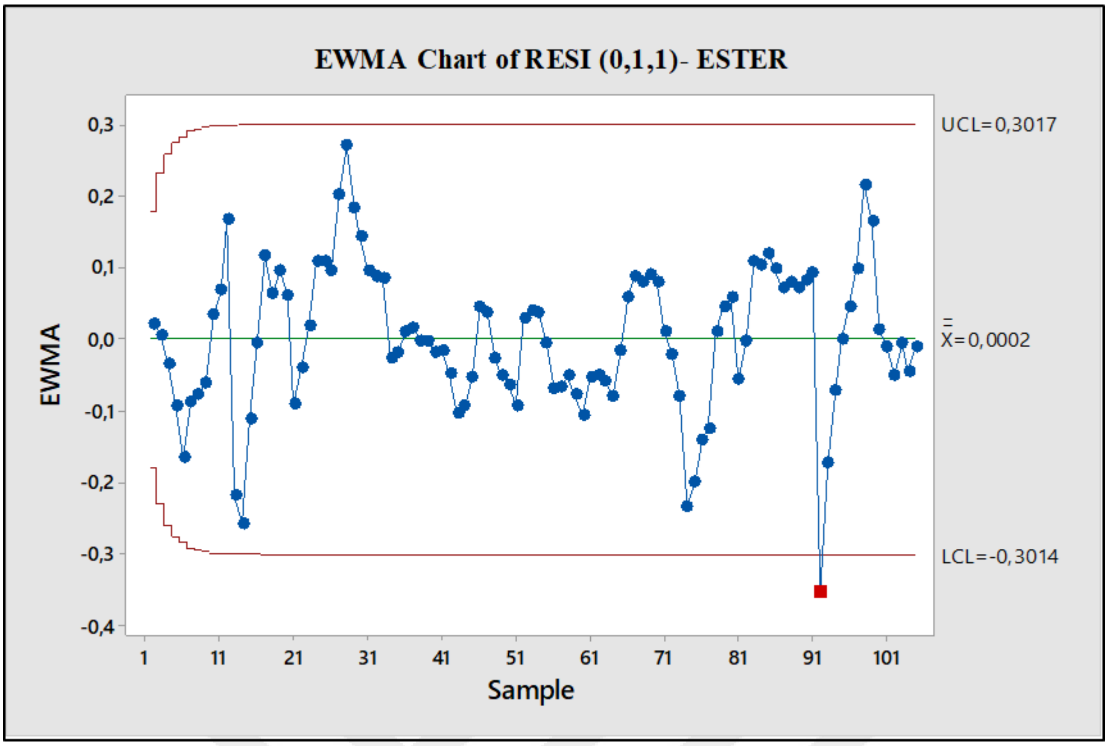

Figure 4.25. EWMACC of Residuals of Ester Content

According to the I-MR in Figure 4.24 and EWMACCs in Figure 4.25, the process looks in control except of two out of control situations, which are on the 13 and the 92 for I-MR and 92 for EWMACC.

Density

As mentioned at section 4.1, data analysis of quality characteristics part of density, the data are slightly auto-correlated.

Although there is a slight auto-correlation in data, in order to deal with the auto-correlation and the non-normality of the data, a time series model ARIMA is applied. The most suitable model for the density data is identified as the ARIMA (1,0,0). After the appropriate TSM is fitted to the data, normality and independence of the residuals of the model is tested. According to figure 4.26, the residuals of the model are normal and independent. Furthermore it is understood that the time series model ARIMA (1,0,0) nicely fits the data.

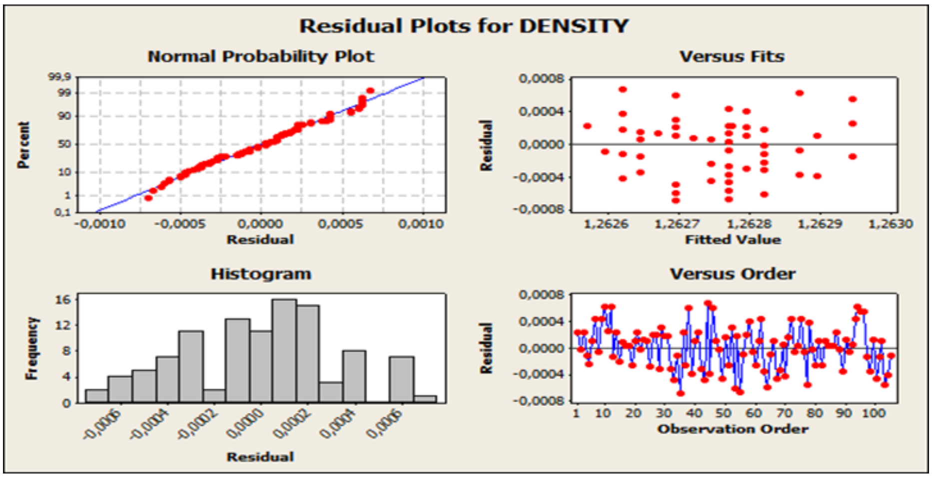

Figure 4.26. Testing Residuals of Time Series Model

Then, the I-MR Control Chart in Figure 4.27 and EWMACC in Figure 4.28 are constructed below in order to detect the large and small shifts in the process mean, respectively.

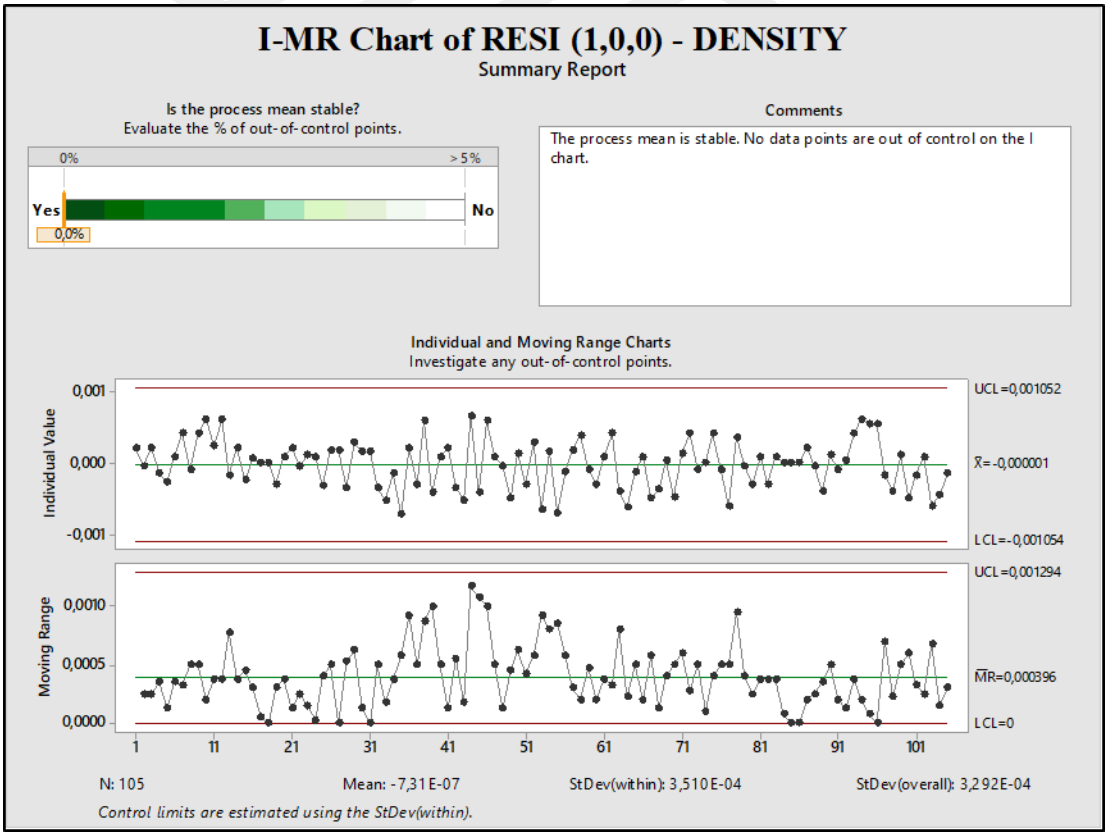

Figure 4.27. I-MR Chart of Residuals of Density

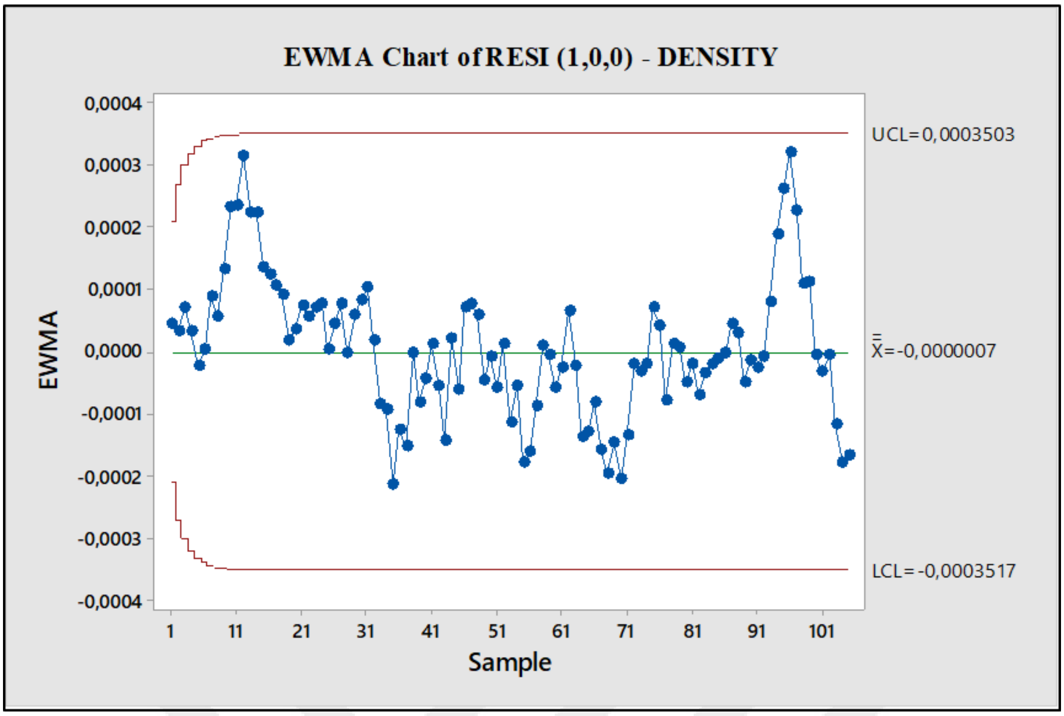

Figure 4.28. EWMACC Chart of Residuals of Density

According to I-MR in Figure 4.27 and EWMACCs in Figure 4.28, the process appears to be in control.

Humidity

As mentioned at section 3.1, data analysis of quality characteristics part of humidity, the data are auto-correlated at the first lag.

In order to deal with the non-normality of the data, Johnson Transformation is applied. After eliminating the non-normality, auto-correlation is eliminated by applying the time series model ARIMA. The most suitable model for glycerin content data is identified as the ARIMA (1,0,0). After the appropriate TSM is fitted to the data, the normality and independence of the residuals of the model is tested. According to figure 4.5, the residuals of the model are normal and independent. Furthermore it is understood that the time series model ARIMA (1,0,0) nicely fits the data.

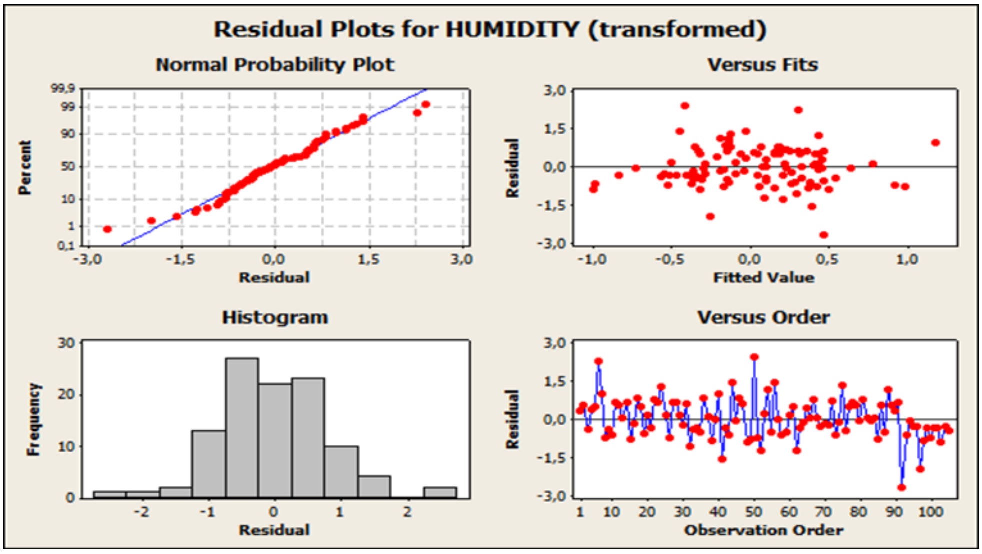

Figure 4.29. Testing Residuals of Time Series Model

Then, I-MR and EWMACCs are constructed in order to detect the large and small shifts in the process mean, respectively.

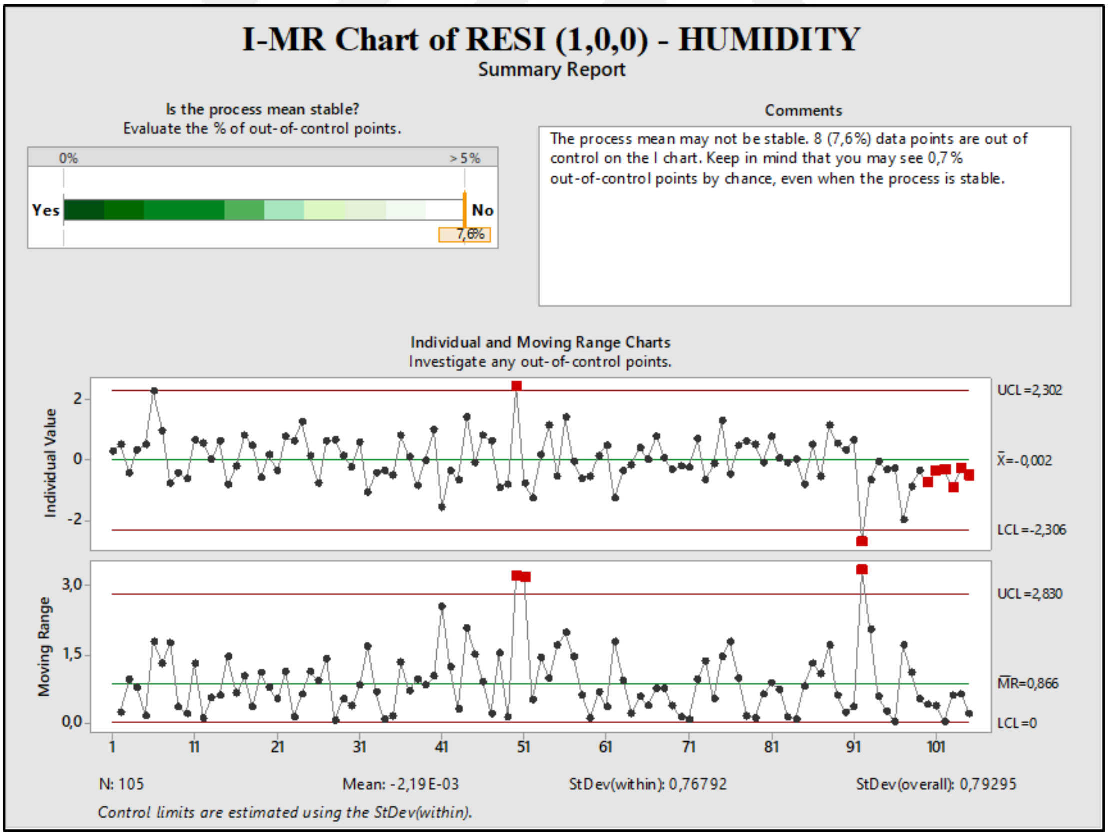

Figure 4.30. I-MR Chart of Residuals of Humidity

According to I-MR Chart in Figure 4.30, there are three out-of-control situations:

- on the 50 because of the out of control point above the upper limit
- on the 91, because of the out of control point below the lower limit and also,
- between the 100 and the 105, where a few out of control points exist because

of their consecutive appearance on the same side and close to the CL which means there is a shift in process mean.

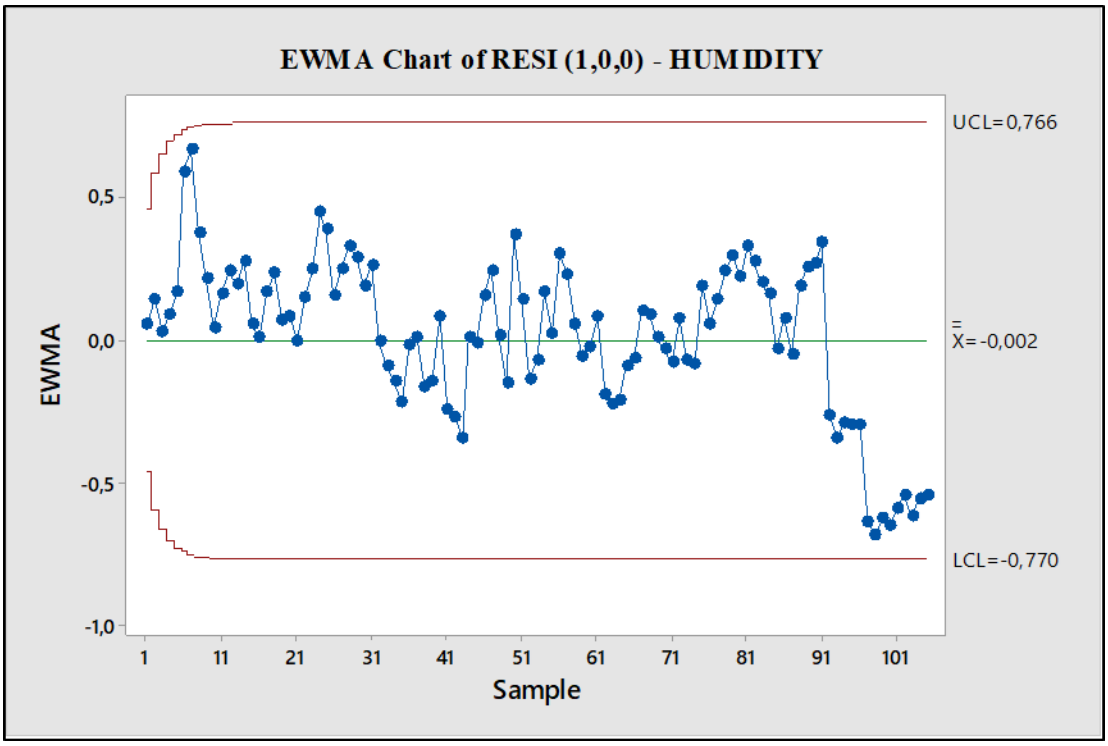

Figure 4.31. EWMACC Chart of Residuals of Humidity

However, according to EWMACC in Figure 4.31, the process appears to be in control (no small shifts in the mean are indicated).

Case Demonstration B: Multivariate Control Charts

Since implementation of the SCC is easy, application of the Shewhart control charts is very common in chemical industry. However, when the data is correlated, monitoring the quality characteristics of the process variable separately does not display the process situation in a correct way. Especially in the chemical process industry, data are mostly correlated such as temperature and pressure. Assuming gas production, there are two critical to quality variables where there is positive correlation between them which are temperature and pressure parameters. In these cases, constructing Shewhart charts becomes insufficient, and also it is hard to manage high number of control charts with all quality characteristic variables.

Most of the time in chemical industry, the process parameters are not independent. There are two possibilities; one is positive relation and another is negative relation. If there is positive relation between the variables, one variable increases while other variable increases. Or if there is negative relation between them, while one of the variables decrease, other increase. However, an out of control situation cannot be detected with correlated data if SCCs are constructed.

Hotelling's T^2 Statistic

Harold Hotelling is the one that realizes constructing MV control charts if correlation exists between the quality characteristic variables. T test with one variable is extended by Hotelling to MV t test with two or more dependent variables. (Hotelling, 1931)

While constructing the Hotelling T^2 statistic, following assumptions must be considered:

- No subpopulations with different population means
- Common variance-covariance matrix
- Independent data
- MV normally distributed data

The Hotelling T^2 MV t test does not have a table of critical t-test values with degrees of freedom values. However, statistical significance can be tested by using an F test. (Hotelling, 1931)

According to the correlation matrix, diagonal values are variances of the terms. The off-diagonal values represent the covariance of only two random terms. The asymptotic variance-covariance matrix displays the variances and covariance of the random terms in the model (Hotelling, 1931).

The correlation of the quality characteristic variables are checked and the results are calculated by using the method as Pearson Correlation via Minitab:

Correlation calculation, which is shown in Table 4.1, between quality characteristics via Minitab: Density, Ester, Glycerin and Humidity

Table 4.1. Correlation Values of Samples

	DENSITY	**ESTER**	**GLYCERIN**
ESTER	0,003	-	-
GLYCERIN	0,067	0,854	-
HUMIDITY	0,075	-0,099	-0,130

If p value is less than the significance level of 0.05, the correlation is significant. According to the Pearson Correlation results, Ester and Glycerin content have large positive relationship with a p value as 0.854.

The Hotelling T^2 statistic follows The Chi-square distribution with two degrees of freedom:

$$\chi_0^2=\frac{1}{s_{11}s_{22}-s_{12}^2}\left[s_{11}(x_2-\bar{x}_2)^2-2(x_2-\bar{x}_2)(x_1-\bar{x}_1)+s_{22}(x_1-\bar{x}_1)^2\right] \qquad (13)$$

The test statistic is plotted on the Chi-square control chart for each sample as follows:

$$\chi_0^2=n(\bar{x}-\mu)'^{\Sigma^{-1}}(\bar{x}-\mu) \qquad (14)$$

Where $\mu' = [\mu_1, \mu_2, \ldots, \mu_p]$ is the vector of in control means for each quality characteristic and Σ is the covariance matrix. Furthermore, the upper control limit of the control chart is

$$UCL=\chi^2_{\alpha,p} \tag{15}$$

It is assumed that the process is in control, data are independent and MV normally distributed with mean μ and covariance matrix Σ. In our case, μ and Σ are unknown. However, $\bar{x}$ and estimated covariance of sample S are used by calculating from data set with n observations when process is assumed to be in control.

Where sample mean is $\bar{x}=(\bar{x}_1,\bar{x}_2,\ldots\bar{x}_p)$ and covariance of sample is

$$S=\begin{bmatrix} s_{11} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{p1} & \cdots & s_{pp} \end{bmatrix}$$

If the μ is replaced with estimated mean and Σ is replaced with estimated covariance matrices in the Chi-square equation, the test statistic becomes

$$T^2=n(\bar{x}-\bar{\bar{x}})'S^{-1}(\bar{x}-\bar{\bar{x}}) \tag{16}$$

The application of the Hotelling T^2 statistic is categorized in two phases which are called as Phase I and Phase II. If the process is assumed to be in control, Phase I is constructed with in control data. With the help of Phase I, control limits are established for Phase II which is used to monitor future observations.

The phase I control limits can be calculated by using the following equation for the T^2 control chart

$$UCL=\frac{p(m-1)(n-1)}{mn-m-p+1}F_{\alpha,p,mn-m-p+1} \tag{17}$$

$$LCL=0 \tag{18}$$

CONSTRUCTING MULTIVARIATE CONTROL CHARTS

The same data set is used as Case Demonstration A: Univariate Control Chart Application. In this case, data set is in control, so only phase I operation is performed and for future work phase II operation will be suggested. The MV control chart is constructed for residual from TSM fitted to the raw data.

As mentioned at section 4.1, data analysis of quality characteristics part of glycerin content, if data do not satisfy the assumption of normality and are not independent, the SPC charts would not be reliable. In order to deal with the auto-correlation, a time series model ARIMA is applied. The most suitable model for the glycerin content data is identified as the ARIMA (1,0,0). After the appropriate TSM is fitted to data, normality and independence of the residuals of the model is tested. According to figure 4.20., the residuals of the model are normal and independent. Furthermore, it is understood from the graphs below that the time series model ARIMA (1,0,0) nicely fit the data.

Data analysis of quality characteristics part of ester content, as it is seen from the auto-correlation function graph, that the data are highly auto-correlated.

In order to deal with the auto-correlation, the time series model ARIMA is applied. The most suitable model for ester content data is identified as the ARIMA (0,1,1). After the appropriate TSM is fitted to the data, normality and independence of the residuals of the model is tested. According to figure 4.23., the residuals of the model are normal and independent. Furthermore, it is understood that the time series model ARIMA (0,1,1) nicely fits the data.

Furthermore, the correlation of the quality characteristic variables is checked and the results are calculated by using the method as Pearson Correlation via Minitab and the results are shown in Table 4.2:

Table 4.2. Pearson Correlation Calculations

	DENSITY	**ESTER**	**GLYCERIN**
ESTER	0,003 0,978		
GLYCERIN	0,067 0,500	0,854 0,000	
HUMIDITY	0,075 0,450	-0,099 0,314	-0,130 0,187

If p value is less than the significance level of 0.05, the correlation is significant. According to the Pearson Correlation results, Ester and Glycerin content have large positive relationship with a p value as 0.854.

According to the correlation results, only Glycerin content and ester are correlated in positive manner. So, only ester and glycerin content quality characteristic data are monitored via constructing MV Hotelling T^2 control chart.

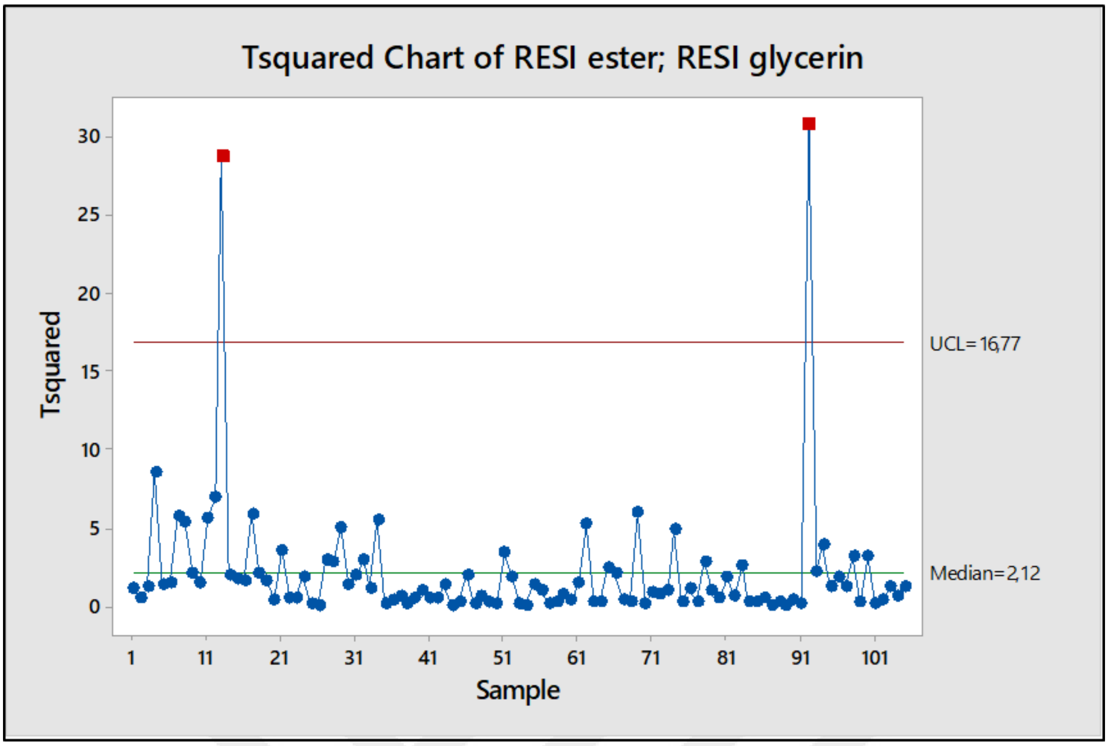

Figure 4.32. Hotelling T^2 Control Chart of Residuals of Ester and Residuals of Glycerin

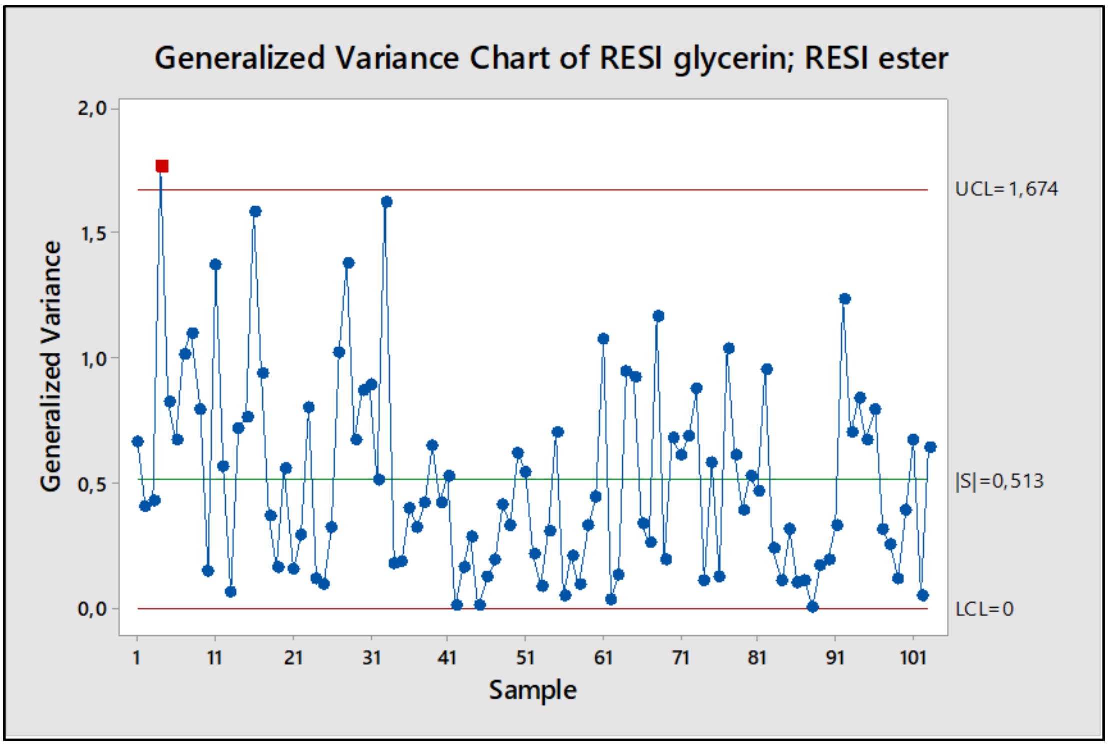

Figure 4.33. Generalized Control Chart of Residuals of Ester and Residuals of Glycerin

According to the Hotelling T^2 control chart in Figure 4.32, the process seems to be in control except observation 13 and observation 92.

Then, MV EWMACC is constructed in order to detect small shifts in the process mean. The ARL is assumed to be 200 and weight is assumed to be 0,1.

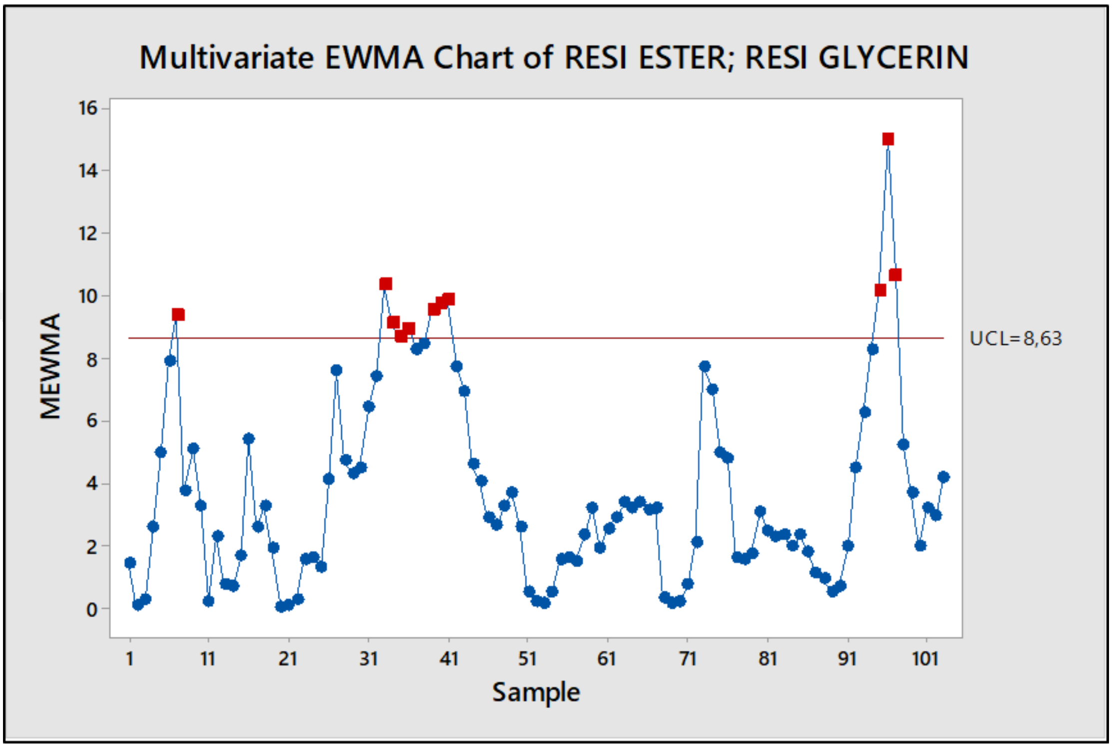

Figure 4.34. Multivariate EWMACC of Residuals of Ester and Residuals of Glycerin

According to MEWMACC in Figure 4.34, there are eleven out-of-control situations:

- On the 7, 33-36, 39-41, 95-97 because of the out of control point above the upper limit.

CHAPTER 5
DISCUSSION OF RESULTS

The SPC is the main topic of this research, various types of SCCs and their applications are mentioned. The univariate statistical process monitoring and control techniques and the MV process monitoring and control techniques are implemented.

The aim of this study is to support chemical process industries and provide them with SPC tools, which help them to control and monitor the processes.

In the case of chemical and process industries, process or product characteristics are often highly correlated because of the consecutive measurements, or automated test and inspection procedures. The most appropriate control chart to use is the Individual Measurements and Moving Range (I-MR) control chart because every quality characteristic is measured on every unit in time order of production. Unfortunately, the performance of these charts is dramatically influenced by even moderate violations of the normality and independence assumptions.

During the study and research, application of SPC on a chemical industry is performed. In our case study, the final product of the chemical production process is pharmaceutical glycerin. There are four quality characteristics which are monitored and controlled. These CTQCs are: glycerin content, ester content, density and humidity. While applying SPC, two different approaches are used. These two approaches are Univariate Control Charts and Multivariate Control Charts in order to monitor the process.

The definition of the research question is as “What kind of control charts are applicable with different kind of data and how to monitor process parameters and interpret them?”

Independence and normality are the basic assumptions behind traditional univariate and MV control charts. However, in many cases especially in chemical and process industry, data are auto-correlated (dependent) because of the process dynamics and consecutive sampling techniques. Furthermore, it is known that the false alarm rate and the shift detection ability of traditional control charts are affected by auto-correlated data. The effect of auto-correlation in the data also influences MV control charts application as well. In this study, the approach followed to overcome auto-correlation is fitting a time series model to the data and then constructing control charts with the residuals of the fitted model.

For all quality characteristics, the methodological approach followed is exactly the same. An initial analysis of the data is conducted in order to understand if the data are auto-correlated and whether they are normally distributed. The final goal is to prepare appropriate Control Charts, based on the historical data given to us by the company, for the monitoring and control of the production process in the future. It is worth mentioning that Control Charts are an on-line tool for quality control.

In the Case Demonstration part A, Univariate control charts are constructed. At first, the data are analyzed by constructing a time series plot, a histogram, a probability plot and the ACF graph in order to see whether the data are normal and independent. After that, if auto-correlation exists, the appropriate time series model (ARIMA) is fitted to the original observations and the normality and independence of the residuals of the model is tested. And finally the EWMACC and I-MR Charts are applied. EWMACC generally detect assignable causes more quickly than the individual moving range chart (Montgomery, 2009).

The revision of the UCL, LCL and CL is required in order to get effective use of control charts. Nevertheless, the same methodology for the analysis of the data can be easily repeated with the current data, in order to compute more appropriate control limits representing the current status of the process.

Starting with the first CTQC, glycerin content, the data analysis shows that data are stationary, they are non-normally distributed and highly auto-correlated. If control charts are constructed with the individual observations which are non-normally distributed and especially not independent, the control chart would be misleading. Therefore, in order to deal with the auto-correlation, a TSM which is used for stationary and auto-correlated data, ARIMA (1,0,0), is applied. The time series model is fitted to the individual observations and then I-MR and EWMACC are constructed using the residuals, in order to be able to detect the large shifts and the small shifts in the process mean, respectively.

As can be observed in the Individual control chart in Figure 4.21, the first out of control situation is starting with observation 28 till observation 41. This can be interpreted as a small shift in the mean. Furthermore, at the EWMACC in Figure 4.22, the same upwards shift can be also observed (points between 28 and 41). As it can be seen above, the ability of detection of the small shift in the mean on the I-MR control chart is weaker than the one of the EWMACC. An unexperienced eye could easily miss the small shift in the mean on the I-MR. However, the same shift on the EWMACC can be easily spotted by the graphical representation of upwards peak. Moreover, there is another out of control point at the individuals control chart which is observation 92 that can be also observed as a point on the lower control limit at the EWMACC.

For the second CTQC, ester content, the data analysis shows that the data are non-stationary; a very common case in the chemical and process industries, they are normally distributed and highly auto-correlated. Therefore, in order to deal with the auto-correlation, a TSM which is used for non-stationary data, ARIMA (0,1,1), is applied. The TSM is fitted to the individual observations and then I-MR and EWMACCs are constructed using the residuals in order to be able to detect the large shift and the small shift in the process mean, respectively.

As can be observed in the Individual control chart in Figure 4.24, the first out of control situation is at observation 13 which can be also observed as a point close to the lower control limit at EWMACC in Figure 4.25. Furthermore, there is another out of control point at the individuals control chart which is observation 92 that is also an out of control point at the EWMACC. In this case, the detection ability of an out of control situation of both control charts could be characterized equal.

For the third of the CTQCs, density, the data analysis shows that the data are stationary, non-normally distributed and slightly auto-correlated. Therefore, in order to deal with the auto-correlation, a time series model that is used for stationary and auto-correlated data, ARIMA (1,0,0), is applied. The time series model is fitted to the individual observations and then I-MR and EWMACCs are constructed using the residuals in order to be able to detect the large shift and the small shift in the process mean, respectively.

According to the individual control chart in Figure 4.27 and EWMACC in Figure 4.28, the process is in control. As can be observed in the EWMACC, there is an upwards shift in the mean which is starting with observation 9 till observation 15. However, if we look carefully at the individual control chart, it can be seen that the small shift in the mean, starting at point 9, can be also spotted by an experienced eye, but not as easily as in the EWMACC. Moreover, there is another upwards shift in the mean at the EWMACC which is starting with observation 93 till observation 99. Same as previously, with a close check, it can be also spotted in the individual chart, with the same observations (93-99). As discussed also above, EWMACCs are more sensitive in detecting smaller shifts in the mean.

For the last CTQC, humidity, the data analysis shows that the data are stationary, non-normally distributed and auto-correlated at the first lag. In order to be able to deal with the auto-correlation, a time series model ARIMA (1,0,0) is applied which is used for stationary and auto-correlated data. The TS model is fitted to the individual observations and then I-MR and EWMACCs are constructed using the residuals in order to be able to detect the large shift and the small shift in the process mean, respectively.

As can be observed in the individual control chart in Figure 4.30, observation 7 is very close to the UCL. There is a second out of control point also, which is observation 50 and, there is a third out of control situation, starting with observation 92 till observation 105. This can be interpreted as a small shift in the process mean. The first and last out of control situations are also observed on the EWMACC in Figure 4.31. Observation 7 can be observed as a point close to the upper control limit, and the downwards shift in the mean can be observed between points 92 and 105. However, the out of control point, observation 50, could not be observed on EWMACC, leading us to believe that it could be a false alarm.

In the Case Demonstration part B, MV control charts are constructed. The same methodology is followed: at first, the data are analyzed by constructing a time series plot, a histogram, a probability plot and the ACF graph in order to see whether the data are normal and independent. In case of auto-correlation, the appropriate time series model (ARIMA) is fitted to the original observations and the normality and independence of the residuals of the model is tested.

In the MV control chart case, another analysis is necessary. The correlation of the quality characteristic variables has to be evaluated. In order to do so, correlation results are calculated by using the Pearson Correlation method via Minitab. According to the Pearson Correlation results, Ester and Glycerin content have large positive correlation with a p value equal to 0.854. The rest of the CTQCs are not correlated to each other. Therefore, only ester and glycerin content quality characteristic data can be monitored via constructing a MV Hotelling T^2 control chart. In order to be able to detect small shifts in the process mean, a multivariate EWMACC is constructed with the residuals of ester and glycerin content.

According to the individual control chart of ester in Figure 4.24, observation 13 and observation 92 are out of control points. Moreover, according to the individual control chart of glycerin content in Figure 4.21, it is seen that observation 92 is an out of control point. Furthermore, it can be observed in the T^2 control chart in Figure 4.32 that the same points can also be observed as out of control points (observation 13 and 92). As expected, the MV T^2 control chart of glycerin and ester, is influenced by the individual charts of the two correlated variables. The assignable causes behind an out of control situation is understood by examining the univariate control charts if the MV control chart detects the cause.

CHAPTER 6
CONCLUSIONS & FURTHER RESEARCH

The objective of this research is to support chemical and process industries and provide them SPC tools, which help them to control and monitor the quality of the processes. Various types of SCCs and its applications are mentioned. The univariate and multivariate statistical process monitoring and control techniques are performed.

In Chapter 2, the thesis starts with a definition of SPC which is very useful in obtaining process stability while reducing the variability. Moreover, the "magnificent seven" tools and techniques are mentioned, as they are used to analyze the process data in methodology part. Then the methodological background of SCC and MV-SPC is addressed. Control charts are graphical displays of measured or derived critical to quality characteristics versus observations or time. Interpreting control charts and the basic assumptions underlying SPC, such as normality and independence, are mentioned. The special case of chemical and process industries, where auto-correlated data exists and the usage of TSMs to handle the violation of the independence assumption, are examined.

In Chapter 3, the methodology and a case study demonstration is presented. The case study is about the monitoring and control of critical to quality characteristics (CTQC) of glycerin production. At first the glycerin production process is explained and four quality characteristics are used to construct control charts for monitoring and controlling the process. The final goal is to design a SPC program, based on historical data, for future monitoring and continuous improvement of the process. The behavior of the process data is checked for normality and independence by using time series plot, histogram, probability plots and ACF graphs respectively.

In Case Demonstration A, Shewhart control charts are elaborated. Analyzed data are examined and if data are correlated, the appropriate TSM is fitted to the data in order to eliminate the auto-correlation. Control charts are designed, starting with the simplest case of I-MR Shewhart control charts mostly appropriate for the chemical and process industries, and which are very useful in phase I implementation of SPC. Then the design of more advanced control charts like EWMACCs is performed since they are more suitable for critical to quality characteristics of processes which are mature and have the need of control charts more sensitive to smaller process shifts.

In the case of the chemical industry, EWMACCs are very useful in implementation of phase II in SPC, since they cover the major disadvantage of a Shewhart control charts, which is containing the information about the last observation in process and ignoring any information given by the entire sequence of points. This feature makes the EWMACC more sensitive to detect the small shifts about 1.5s or less, in the process mean. In cases where the process seems to be in control, the process parameters can be estimated (such as the mean and standard deviation), and assignable causes do not commonly result in large process upsets or disturbances.

In Case Demonstration B, MV control charts are elaborated. In the MV control chart case, the correlation of the quality characteristic variables are calculated by using Pearson Correlation method. As a result of correlation calculation, the quality characteristics of ester and glycerin content which have large positive correlation, are selected to show case the use of MV control charts in our case. At first Hotelling's T^2 control chart is constructed for ester and glycerin content and then, MV EWMACC is constructed in order to detect the small shifts in the process mean.

In Chapter 4, the results of case demonstration A and B are discussed. As a conclusion of Case Demonstration A, it is observed that the ability of detection of the small shift in the mean on the I-MR control chart is weaker than the one of the EWMACC. An unexperienced eye could easily miss the small shifts in the mean on the I-MR control charts. However, the same shifts on the EWMACC can be easily spotted by the graphical representation of upwards peaks. As a conclusion of Case Demonstration B, the MV T^2 control chart of glycerin and ester, is influenced by the individual charts of the two correlated variables. If the MV control chart detects an out of control situation, the assignable causes behind it, can be understood by examining the univariate control charts.

The comparison between univariate and MV control charts indicates that these two different control charts acts as a compatible approach for monitoring the chemical and process industries processes. If the MV control chart detects an out of control situation, then the univariate control charts will be useful to determine the critical to quality characteristic, which caused this out of control situation.

Flow chart of the model is in Figure 6.1:

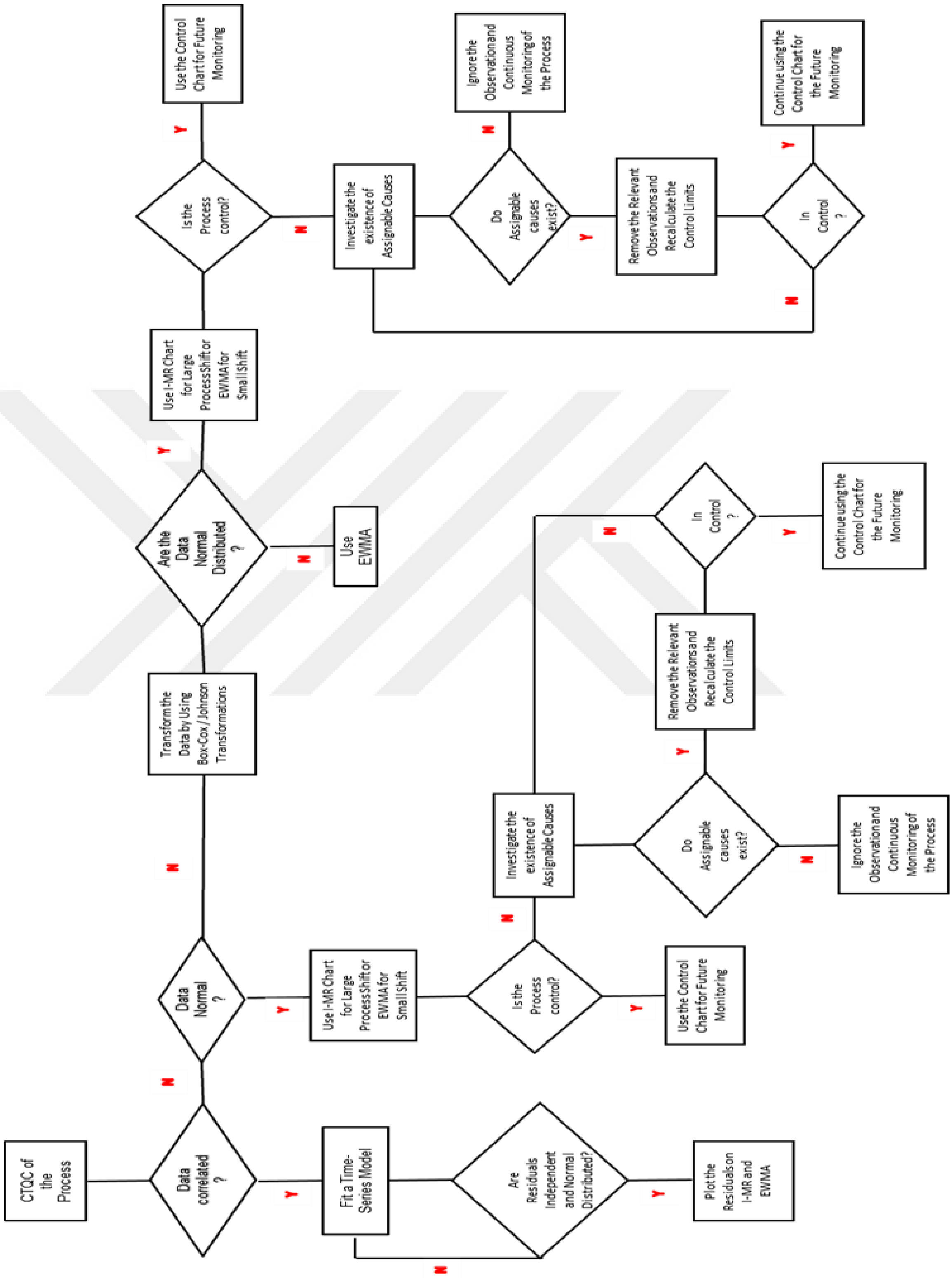

Figure 6.1. Flow Chart of The Model

The following suggestions are made for the future research in this area:

1. Design of experiments (DOE) approach can be used to determine which critical to quality characteristics have the greatest impact on the process quality and at which point during the process they should be monitored. Furthermore, new control charts can be designed for the identified characteristics and the SPC program can be extended throughout the whole process.
2. Revision of the control limits for all designed control charts is necessary, in regular time intervals. Further research could specify the necessary time intervals and provide the methodology for implementing this. For example, in phase II implementation of SPC for T^2 control charts, control limits should be calculated by using the following equation in order to be able to monitor future productions.

$$UCL=\frac{p(m+1)(n-1)}{mn-m-p+1}F_{\alpha,p,mn-m-p+1} \quad (19)$$

$$LCL=0 \quad (20)$$

3. CUSUMCCs application are appearing in the literature review. The performance comparison between EWMACCs and CUSUMCCs could be an extension of this study. Furthermore, MV CUSUMCC and MV-EWMACCs can be compared.

For supporting the everyday work in an industrial environment, a software application could be developed on a later stage. Since in most of the cases, because of the violation of the independence assumption and existence of auto-correlation in data a fitted model is necessary, plotting and maintaining a control chart is not easy. A software application where the quality worker can input the original measured observation and automatically have as a result the appropriate fitted point plotted in the control chart is necessary. Control charts are useful when they can be used as on-line tools in process monitoring. In the case of the chemical industry, this cannot be achieved without a software application.

REFERENCES

Alwan L. C., Roberts H. V. (1988). *Time-Series Modeling for Statistical Process Control*, Journal of Business & Economic Statistics, 6(1), 87-95

Alwan, L. C. (1992). *Effects of autocorrelation on control chart performance. 'Communication in Statistics: Theory and Methods'*, 21, 1025-1049

Alwan, L. C. (1992). *Effects of Autocorrelation on Control Charts.* Communications in Statistics—Theory and Methods, Vol. 21, (4), pp. 1025–1049.

Anderson, T. W. (1984). An Introduction to MV Statistical Analysis, 2nd ed. Wiley, New York.

Bersimis, S., Psarakis, S. & Panaretos, J. (2007). *MV Statistical Process Control Charts: An Overview*, Quality and Reliability Engineering International, vol. 23, p. 517-543

Bisgaard, S. & Kulahci, M. (2005). *Quality Quandaries: The Effect of Autocorrelation on Statistical Process Control Procedures.* Taylor & Francis Inc. 17:481-489

Box, G. E. P & Cox, D. R. (1964). *An Analysis of Transformations.* Journal of Royal Statistical Society B 26(2), 211-252

Box, G. E. P. & Jenkins G., (1970). *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, CA

Box, G. E. P., G. M. Jenkins, & G. C. Reinsel (1994). *Time Series Analysis, Forecasting, and Control*. 3rd ed. Prentice-Hall, Englewood Cliffs, NJ.

Costa, M. T., Corolino, E. & Oliveria, T. (2014). *Application of Hotelling's T2 Charts in Monitoring Quality Parameters in a Drinking Water Supply System*.

Devor, E. R., Chang, T. & Sutherland, J. W. (2007). Statistical Quality Design and Control Contemporary Concepts and Methods (2nd Edition). Pearson Education, Inc. New Jersey

Durbin, J. & Watson, G. S. (1971). Testing for serial correlation in least squares regression III. Biometrika. 58 (1): 1–19. doi:10.2307/2334313.

Elevli S., Uzgören N. & Savas M. (2008). *Control Charts for Autocorrelated Colemanite Data*, Journal of Scientific & Industrial Research, Vol. 68, 11-17

Elevli, S. & Uzgoren, N. & Savas M. (2009). *Control Charts for Auto correlated Colemanite Data*. Journal of Scientific & Industrial Research, Vol. 68, pp. 11-17

Engle, R.F. (1982). *Autoregressive conditional heteroscedasticity with estimates of the variance.* of U.K. in Dation, Econometrica 50 987–1008.

Friendly, M. & Denis, D. (2005). *The early origins and development of the scatterplot.* Journal of the History of the Behavioral Sciences. 41 (2): 103–130

Gagniuc, Paul A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. USA, NJ: John Wiley & Sons. pp. 1–256. ISBN 978-1-119-38755-8.

Granger, C. W. J. & Anderson, A. P. (1978). *An Introduction to Bilinear Time Series Models,* Vandenhoeck and Ruprecht, G Yottingen,

Grigg O. A., Farewell, V. T. & Spiegelhalter, D.J. (2003). The Use of Risk-Adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts. Statistical Methods in Medical Research. 12 (2): 147–170

Harris, T. J. & Ross, W. H. (1991). *Statistical Process Control Procedures for Correlated Observations.* Canadian Journal of Chemical Engineering, Vol. 69(Feb), pp. 48–57.

Hipel, K. W. & McLeod, A.I. (1994). *Time Series Modelling of Water Resources and Environmental System.* Amsterdam, Elsevier

Hogg, R.B. & Craig, A.T. (1978). Introduction to Mathematical Statistics (4th Edition). Macmillan Publishing Co., Inc. New York

Hogg, R.B., McKean, J.W. and Craig, A.T. (2013). Introduction to Mathematical Statistics (7th Edition). Pearson. 307

Hotelling, H. (1931). *The generalization of student's ratio.* The Annals of Mathematical Statistics, v.2, pp.360-378.

Howitt, D. & Cramer, D. (2008) *Statistics in Psychology.* Prentice Hall

Investopedia (2017). What is the ’80-20’ Rule. Retrieved 2017 from http://www.investopedia.com/terms/1/80-20-rule.asp

Ishikawa, K. (1968). *Guide to Quality Control.* Tokyo: JUSE.

Jensen, W. A., Jones-Farmer, L. A., Champand, C. H. & Woodall, W. H. (2000). *Effects of Parameter Estimation on Control Chart Properties: A Literature Review*. J. Quality Technology, 38, 145-187

Johnson, N.L. (1949). *Systems of frequency curves generated by methods of translation*, Biometrika, Vol. 36, 149-176.

Kandananond, K. (2013). *Guidelines for Applying Statistical Quality Control Method to Monitor Autocorrelated Processes*, Elsevier

Lestander, T. A. et al. (2012). *Towards MV Statistical Process Control in the Wood Pellet Industry.* Elsevier. Biomass and Bioenergy 45 pp. 152-158

Mason, R. L. & Young, J. C. (2001). *Implementing MV Statistical Process Control Using Hotelling's T2 Statistics*. Quality Progress. 34(4), 71-73

Mason, R. L. & Young, J. C. (2002). *MV Statistical Process Control with Industrial Applications.* ASA/SIAM: Philadelphia

Mastrangelo, C. M. & Montgomery, D.C. (1995). *SPC with Correlated Observations for the Chemical and Process Industries*, *Quality and Reliability Engineering International*, Vol 11 78-89

Mastrangelo, C. M., & D. C. Montgomery (1995). *SPC with Correlated Observations for the Chemical and Process Industries.* Quality and Reliability Engineering International, Vol. 11(2), pp. 79–89.

Montgomery, D. (2005). *Introduction to Statistical Quality Control.* John Wiley & Sons, Inc.

Montgomery, D. C. & Mastrangelo, C. M. (1991). *Some Statistical Process Control Methods for Auto-correlated Data.* Journal of Quality Technology, Vol. 23(3), pp. 179–204.

Montgomery, D. C. (2009). *Introduction to Statistical Process Control.* (6th Edition), Arizona State University, Wiley.

Montgomery, D. C., Johnson, L. A., & Gardiner, J. S. (1990). *Forecasting and Time Series Analysis*, 2nd ed., McGraw-Hill, New York.

More Steam (2017). Statistical Process Control. Retrieved 30 December 2017, from https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm

NIST/Sematech Engineering Statistics Handbook. (2009). Individuals Control Charts. National Institute of Standards and Technology. Retrieved 10 August 2009

NIST/Sematech Engineering Statistics Handbook. (2010). Proportions Control Charts. National Institute of Standards and Technology. Retrieved 5 January 2010

Noskievicova, D. (2016). *Control Chart Limits Setting when Data are Autocorrelated*, Lund University

Pearson, K. (1895). *Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 186: 343–414

Raicharoen, T., Lursinsap, C. & Sanguanbhoki,P. (2003). *Application of critical support vector machine to time series prediction, Circuits and Systems*, ISCAS '03.Proceedings of the 2003 International Symposium on Volume 5, pages: V-741-V-744.

Rao, O. R. M. et al. (2013). *Application of MV Control Chart for Improvement in Quality of Hot Metal.* International Journal for Quality Research 7(4), pp. 623-640

Rogriguez, R. N. (1996). *Recent Issues in Statistical Process Control: SAS Solutions Using Statistical Modeling Procedures.* SAS Institute Inc. Cary,NC

SAS/QC® (2015). *14.1 User's Guide 2015*, SAS Institute Inc., Cary, NC, USA

Stoumbos, Z. G., & Reynolds, M. R. (2000). *Robustness to non-normality and autocorrelation of individuals control charts*. Journal of Statistical Computation and Simulation, 66(2), 145-187.

Tong, H. (1983). *Threshold Models in Non-Linear Time Series Analysis*. Springer-Verlag, New York

Tracy, N. D., Young, J.C. & Mason, R. L. (1995). *MV Control Charts for Individual observations*. Journal of Quality Technology, 24(2), 88-95

Utts, J. M. (2005). *Seeing Through Statistics.* (3rd Edition, pp 166-167). Thomson Brooks/Cole,

Wheeler, Donald J. (2009). Good Limits from Bad Data, Quality Digest, Retrieved 8 February 2010

Zhang, G. P. (2003). *Time series forecasting using a hybrid ARIMA and neural network model,* Neurocomputing 50 pages: 159–175.