STATISTICAL TECHNIQUES TO MINIMIZE PRODUCT LOSS IN HAIR CONDITIONER PROCESS

NATTHANON PHAIBOONSILPA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN CHEMICAL ENGINEERING
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2007

หัวข้อวิทยานิพนธ์	การใช้เทคนิคทางสถิติเพื่อลดความสูญเสียผลิตภัณฑ์ในกระบวนการผลิตครีมนวดผม
นักศึกษา	นายณัฐนนท์ ไพบูลย์ศิลป์
รหัสนักศึกษา	48060102
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมเคมี
พ.ศ.	2550
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.ดร. อัญชลีพร วาริทสวัสดิ์ หล่อทองคำ

บทคัดย่อ

ในการวิเคราะห์มูลค่าการสูญเสียผลิตภัณฑ์เครื่องใช้ส่วนบุคคลชนิดเหลวของบริษัทผลิตสินค้าอุปโภคบริโภคที่เป็นกรณีศึกษาด้วยแผนภูมิพาเรโตพบว่าเกิดการสูญเสียผลิตภัณฑ์ครีมนวดผมมากที่สุดเนื่องจากผลิตภัณฑ์ไม่ได้ตามมาตรฐานและงานรีเวิร์คทั้งนี้เพราะความเบี่ยงเบนของความหนืดของผลิตภัณฑ์ออกจากค่าขอบเขตควบคุมล่าง โครงการนี้ศึกษาการลดความสูญเสียผลิตภัณฑ์ในกระบวนการผสมวัตถุดิบซึ่งเป็นขั้นตอนที่สำคัญที่สุดในการผลิตครีมนวดผม ผลิตภัณฑ์เป้าหมาย คือ ครีมนวดผมชนิดที่ 1 จำนวน 5 สูตร ด้วยเทคนิคทางสถิติตามขั้นตอนต่างๆ ได้แก่ การวิเคราะห์สาเหตุที่เป็นไปได้ต่อความเบี่ยงเบนของความหนืดด้วยแผนภาพสาเหตุและผลและการระดมสมอง การทดสอบความถูกต้องและแม่นยำของระบบการวัดความหนืดของผลิตภัณฑ์ครีมนวดผม การวิเคราะห์การถดถอยและออกแบบการทดลองด้วยโปรแกรมมินิแทป เพื่อระบุตัวแปรที่มีนัยสำคัญที่สุดต่อความเบี่ยงเบนของความหนืดของผลิตภัณฑ์ และกำหนดค่าภาวะการปฏิบัติงานที่เหมาะสมเป็นมาตรฐานการปฏิบัติงานของบริษัทเพื่อลดความเบี่ยงเบนของความหนืดของผลิตภัณฑ์ ได้แก่ ปริมาณสารละลายคลอรีนร้อนที่เติมในถังผสมหลัก 700 กิโลกรัม ระยะเวลาที่ใช้ในการละลายสารลดแรงตึงผิวชนิดที่ 1 ที่ถังผสมหลัก 10 นาที อุณหภูมิที่ใช้ในการละลาย fatty alcohols ที่ถังผสมรอง 75 องศาเซลเซียส และอุณหภูมิที่ถังผสมหลักก่อนเติมน้ำที่นอุณหภูมิห้องเท่ากับ 65 องศาเซลเซียส หลังจากใช้ภาวการณ์ปฏิบัติงานดังกล่าวได้ผลดังนี้ ค่าเฉลี่ยของค่าเบี่ยงเบนมาตรฐานของความหนืดของผลิตภัณฑ์ครีมนวดผมเป้าหมาย 5 สูตร ลดลงจาก 3,256 เป็น 1,718 cP โอกาสของการพบข้อบกพร่องในหนึ่งล้านครั้งของการผลิตครีมนวดผมเท่ากับ 1.57 และดัชนีสมรรถภาพกระบวนการเท่ากับ 1.75 ซึ่งอยู่ในระดับที่ยอมรับได้ คิดเป็นมูลค่าการสูญเสียผลิตภัณฑ์ครีมนวดผมที่ลดลงประมาณ 306 ตัน/ปี

Thesis Title	Statistical Techniques to Minimize Product Loss in Hair Conditioner Process
Student	Mr. Natthanon Phaiboonsilpa
Student ID.	48060102
Degree	Master of Engineering
Program	Chemical Engineering
Year	2007
Thesis Advisor	Assoc. Prof. Dr. Anchaleeporn Waritswat Lothongkum

Abstract

According to loss assessment of personal care liquid products of a case study consumer product industry by the Pareto diagram, it was found that major loss came from the hair conditioners due to the off-spec product and rework because their viscosities deviated from the lower control limit (LSL). This work focused on reduction of product loss in the raw material mixing process which was the most important step in hair conditioner production. Of 8 formulas of hair conditioner type 1, 5 formulas which had similar raw materials and processing steps were selected as the model study. The statistical techniques were used to solve this problem. First, the possible parameters in raw material mixing process that affected the deviation of product viscosity were detected by the cause and effect diagram and brainstorming. The accuracy and precision of hair conditioner viscosity measurement system were assured by the measurement system analysis (MSA). Subsequently, by using the multiple linear regression analysis and the design of experiments (DOEs) in the Minitab program, the strongly significant parameters on the deviation of product viscosity were obtained and the optimum operating conditions were set as the standard operating procedures (SOPs) of the company, i.e., the quantity of hot chlorinated water in the main mixer was 700 kg, time to dissolve cationic surfactant type 1 at the main mixer was 10 mins, the temperature at the side mixer no. 1 to dissolve fatty alcohols was 75 ºC, and the set-point temperature at the main mixer before adding the ambient-temperature water was 65 ºC. From the achieved optimum operating conditions, the average standard deviation of the viscosities of 5 target hair conditioners was reduced from 3,256 to 1,718 cP. The defect per million opportunities (DPMO) was found to be 1.57 corresponding to the potential process capability index (C_p) of 1.75 which was in the acceptance range. Based on the obtained DPMO, the calculated saving was equivalent to hair conditioner of 300 tons/year.

Acknowledgements

The author would like to express his gratitude to Assoc. Prof. Dr. Anchaleeporn W. Lothongkum, Dr. Viwat Krisadhasima, Captain Pichayane Tanprasert R.T.N., Mr. Surapol Patanaphongsanon and Mr. Charean Areeasanadham for their support and fruitful advice to complete this work. The assistance from the supporting staff of the Unilever Thai Holdings Ltd. is also acknowledged.

The partial financial support from the Thailand Research Fund (TRF-MAG), Fiscal Year 2006, and the Unilever Thai Holdings Ltd. is highly appreciated.

Sincere thanks also go to the Technology Promotion Association (Thailand-Japan) for their kindness to allow him to participate the technical training course on six sigma, Faculty of Engineering and the School of Graduate Studies, King Mongkut's Institute of Technology Ladkrabang.

Last but not least, deep gratitude goes to his family for their consistent support and encouragement.

Natthanon Phaiboonsilpa

Contents

Contents (cont.)

Contents (cont.)

List of Tables

List of Tables (cont.)

List of Figures

Figures **Page**

List of Figures (cont.)

List of Figures (cont.)

Technical Terms

Accuracy	The degree of conformity of the measurement or the average value of measurement to the true value.
Bias	The difference between the average value of measurement and the reference value.
Coefficient of determination (R^2)	A static that describes the strength of a linear relationship between two variables. The coefficient of determination can take the values between +1 to 0. A +1 value indicates a perfect correlation, while a zero value indicates no correlation.
Conformance	A condition that the product, process or service conforms to the specifications.
Continuous data	The data that are measured on a continuous scale. Continuous data are presented in the decimal form that can be subdivided into finer and finer increments depending on the resolution of the measurement device.
Continuous distribution	The distribution that describes the probability of the continuous data.
Cycle time	Total time to produce a unit of work/product of interest from the beginning to the end of the process.
Defect	A failure to meet an imposed requirement on a single quality characteristic, or a single instance of nonconformance to the specification.
Defects per million (DPM)	The defect rate based on the one million units of product. Also referred to **the parts-per-million (ppm) defect rate**.
Defects per million opportunities (DPMO)	The number of defects occurred in the process based on one million units of product divided by the number of opportunity to make a defect.
Defects per unit (DPU)	The number of defects occurred in the process per the number of product which are produced, so-called **the defect rate**.

Technical Terms (cont.)

Degree of bias (k)	An index that describes how much the process mean shifts from the target value. The degree of bias takes the values between 0 to any positive number. A zero value indicates there is no off-target process mean.
Degree of freedom	The number of the measurements that are independently available for estimating a population parameter. The number of degree of freedom is equal to the number of observations minus the number of the parameters that are estimated.
Design of experiments (DOEs)	A method of experimentation in which purposeful changes are made to the input variables or factors to observe the corresponding change in output variables or responses in a process
Deviation	The difference between the individuals and the mean value of the observations.
Discrete data	The data that are based on counts. Discrete data are only the finite or integer value.
Discrete distribution	The distribution that describes the probability of the discrete data.
Distribution	The pattern of the measurements or data dispersion which are characterized by the location, spread and shape.
Factorial design experiments	The designed experiments that assess several factors/variables simultaneously in one set of experiments, which all combinations of factor levels are tested.
Failure mode and effects analysis (FMEA)	A systematic approach which examines the possible failures that can occur in the process or product, the effects and causes of the failures. The significant failures are prioritized based on their severity, frequency of occurrence and possibility that the failures are not detected. The process or product are redesigned so that the new model eliminates the possibility of the failures.

Technical Terms (cont.)

Gage repeatability and reproducibility study (Gage R&R study)	The evaluation of a measurement system to determine the variation inherent in the measurement devices, measuring procedures and measurers.
Interaction effect	The effect of one factor on a response variable when the effect is dependent on the level of the other factor.
Level	The values of the factor/independent variable or the number of levels of the independent variable that are used in the experiments to examine the effects of the factor or independent variable on the response variable.
Linearity	The change of bias over a normal measurement range of the measuring device.
Lower control limit (LCL)	The lower limit that the value of process statistic in the control chart, e.g., mean, range, standard deviation, etc. is considered to be out of control.
Lower specification limit (LSL)	The accepted lowest value of a conformance product characteristic or measurement.
Main effect	An estimate of the effect of a factor on a response variable measured independently from other factors.
Mean	The average value of the observations calculated by the sum of the observations divided by the number of observations. The mean describes the central location of the data.
Measurement system	The complete process of measurement including the measurement devices, measuring procedures and measurers.
Measurement system analysis (MSA)	A statistical technique for examining the reliability of the measurement system in terms of the accuracy and precision to obtain the valid data for the process control decision and product acceptance.

Technical Terms (cont.)

Nonconformance	A condition that the product, process or service does not conform to the specifications. Nonconformance is often referred to as a defect or rework.
Normal distribution	A bell-shaped distribution that is often used to describe the continuous distributions. Also referred to as the Gaussian distribution.
Population mean (μ)	The sum of all the observations of the population divided by the population size.
Population standard deviation (σ)	The square root of the population variance. It is denoted by the Greek letter (σ) so-called sigma.
Population variance (σ^2)	A measure of population variation based on the average of the squares of the difference between each observation of the population and the population mean.
Potential process capability index (C_p)	A metric that is used to measure the process capability. It is a ratio of the allowable tolerance spread divided by 6 standard deviations of the process outputs.
Precision	The degree of mutual agreement among the individuals of measurements.
Precision-to-Tolerance (P/T)	The percentage ratio of the variation due to the measurement system in terms of the standard deviation to the tolerance.
Precision-to-Total Variation (P/TV)	The percentage ratio of the variation due to the measurement system in terms of the standard deviation to the total process variations.
Probability value (p-value)	The value of probability that would lead to reject the null hypothesis, in general, the null hypothesis is rejected at the significance level (α) of 0.05 or p-value less than 0.05.
Probability density function (PDF)	Mathematical model that describes the distribution of the data.

Technical Terms (cont.)

Process capability index (C_{pk})	A metric that is used to measure the process capability when the off-target of the data mean is taken into consideration. It is a ratio of distance from the mean values of the process outputs to the nearest specification limit divided by 3 standard deviations.
Productivity	The amount of the conformance process output or product that is produced per one unit input, or the ratio of a number of the conformance process output or product to the total number of the process input. It is often multiplied by 100 and expressed as a percentage value.
Range	The difference between the largest value and the smallest value within the set of observations.
Regression analysis	A statistical technique for determining the linear or polynomial relationship between one response variable and one or more independent variables.
Regression model	The mathematic model that shows the relationships among one response variable and one or more independent variables.
Rejection criterion	The threshold value used to determine whether or not the null hypothesis should be rejected based on the selected significance level or p-value.
Repeatability	The variation in the measurements derived from one measurement device when one operator makes several measurements on one characteristic of the same sample.
Replication	Experiment trials that are made under the identical conditions.
Reproducibility	The variation of the measurements made by the different operators using the same measurement device when measuring one characteristic on the same sample.

Technical Terms (cont.)

Residual	The difference between an observation and the fitted value of the response variable that is predicted by the regression model at the particular value of the regressor or independent variable.
Residual model diagnostics	A check of the reliability of the collected data and adequacy of the regression model in the multiple linear regression analysis. The residuals should be normally and independently distributed with a constant variance.
Resolution	The discrimination of the measurement device explained by the smallest change of the process outputs that the measurement device can detect.
Response variable	The output variable of the product, process or service of interest.
Sample mean ($\bar{x}$)	The sum of the observations of the sample divided by the sample size.
Sample standard deviation (s)	The square root of the sample variance.
Sample variance (s^2)	A measure of sample variation based on the average of the squares of the difference between each observation of the sample and the sample mean.
Sigma quality level or Sigma level	The quality level that describes the capability of a process to meet the process specifications.
Significance	A statistical statement indicating that the different value of a factor causes a difference in a response with a certain degree of risk of being in error.
Significance level (α)	A statistic that indicates the probability or degree of risk to make a wrong decision when doing the hypothesis testing. The significance level has been customarily set at 0.05 and 0.01. The significance levels are sometimes called **alpha level** that is represented by the Greek letter (α).

Technical Terms (cont.)

Standard deviation	The square root of the variance. It is a measure of the consistency of a process.
Stability	The change of bias over time.
Target	The value of the process output characteristic which the process is aimed to reach.
Test statistic (F_0 or t_0)	A numeric value calculated from the sample data that is relevant to the hypothesis. It is used to decide whether or not the null hypothesis should be rejected by comparing its value with the rejection criterion.
Tolerance	The range of upper specification limit and lower specification limit.
Upper control limit (UCL)	The upper limit that the value of process statistic in the control chart, e.g., mean, range, standard deviation, etc. is considered to be out of control.
Upper specification limit (USL)	The accepted highest value of a conformance product characteristic or measurement.
Variation or variability	The dispersion in the measurements described by the difference between the measurements within a population or among population.
Variance	A measure of variation based on the average of the squares of the difference between each observation and the arithmetic **mean** of the data.
Yield	The ratio of a number of the conformance process output or product to the total number of the process output. It is often multiplied by 100 and expressed as a percentage value.

Chapter 1

Introduction

1.1 Rationale

The Unilever Thai Holdings Limited is one of the leading consumer product companies. Three main consumer products are as follows:

1. home care products, e.g., detergents, dish-washing liquid and fabric softener
2. personal care products, e.g., shampoos, hair conditioners, liquid and bar soaps, and lotions
3. foods and ice cream.

Generally, the most important business strategy of any company is to produce products of high quality at the minimum cost to meet the customer satisfaction. Common problems in the consumer product industry are losses in terms of rework, giveaway and long cycle time, resulting in high production cost and low productivity.

This current research has been focused on loss minimization of personal care liquid products, i.e., shampoos, hair conditioners, liquid soaps and lotions. According to the Pareto diagram in Figure 1.1, the main loss of the personal care liquid products came from hair conditioner type 1 of about US$ 46,450 per year. Of 8 formulas of hair conditioner type 1, 5 formulas which have similar raw materials and processing steps were chosen as the target case studies.

The physical properties of the hair conditioner such as viscosity, color and scent are important characteristics to draw the customer attraction. The specifications of hair conditioner in this study are viscosity, pH and % cationic surfactant. The control specifications, i.e., the target value, the upper specification limit (USL) and the lower specification limit (LSL) are written in the standard operating procedures (SOPs).

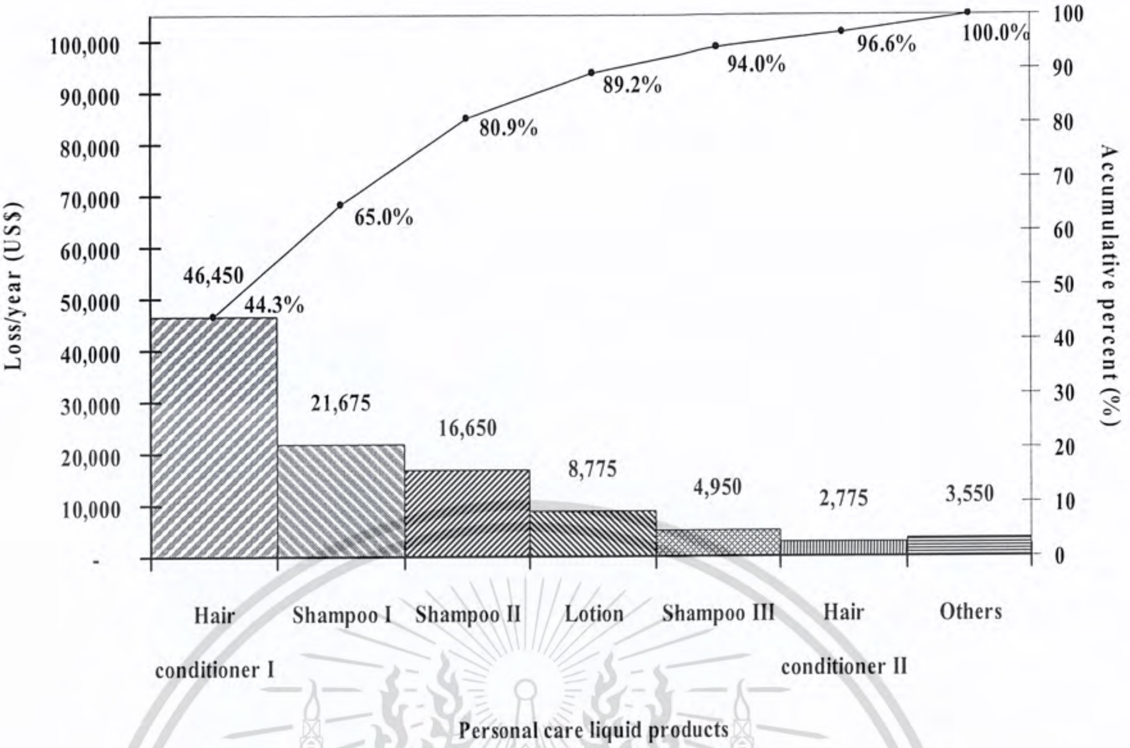


Figure 1.1 Pareto diagram of personal care liquid products.

After collecting the relevant data in raw material mixing process of hair conditioner, it was found that about 90 kg or 3% in a batch lost by cleaning of the mixing equipment and piping systems. The product loss from cleaning became substantial large when hair conditioner viscosity increased. However, for a consecutive mixing batch without cleaning, the viscosity deviated substantially from the target value. Mostly, the viscosity which was obtained from the consecutive batch decreased gradually, and finally deviated from the LSL. For this reason, cleaning of mixing equipment and piping systems must be done frequently, approximately 10 times a month, which on the other hand increased product loss.

The product of its viscosity less than the lower specification limit was reworked. In case the product viscosity was higher than the upper specification limit, the viscosity must be adjusted to be within the control specifications by adding hot chlorinated water and some chemicals, however, a longer production time and higher production cost occurred. The deviation of hair conditioner viscosity played an important role in hair conditioner loss and became an urgent issue that required the corrective action. In this case study, the viscosities of 5 formulas of hair conditioner significantly deviated from the target value of 26,000 cP, mostly the viscosities stayed below the target value as shown in Figure 1.2. Some of the viscosities of 5 formulas of hair conditioner were out of LSL. The viscosities were adjusted to keep the values within the LSL.

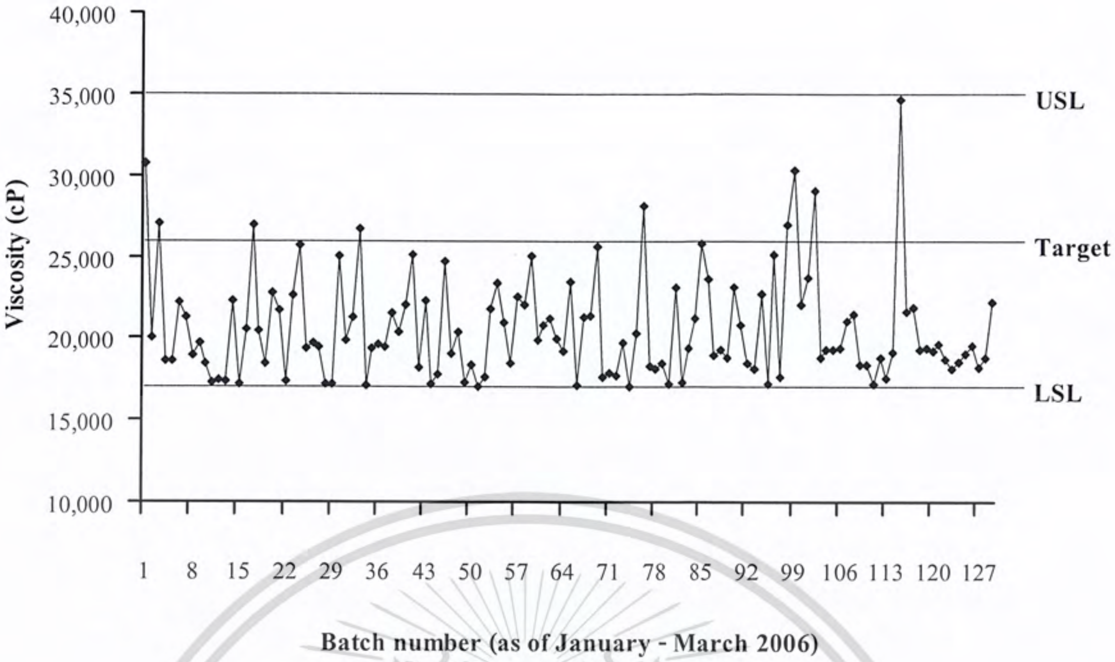


Figure 1.2 The viscosities of 5 formulas of hair conditioner.

Basically, six sigma provides powerful statistical techniques and quality tools to improve the process capability and to reduce the process variations and product defects for a higher profit, high product quality, high employee morale, and high customer satisfaction and loyalty. Six sigma statistically ensures that 99.9997% of all products are of acceptable quality corresponding to 3.4 defects per million opportunities (DPMO) [1].

From the preliminary investigation, it was found that hair conditioner viscosities were apparently deviated from the target viscosity. The standard deviation of the product viscosities was 3,256 cP. To respond high competition in personal care liquid product business, this work aims to minimize the deviation of hair conditioner viscosity during the raw material mixing process using six sigma and statistical techniques to reduce the off-spec product and rework. High productivity, and high product quality and low production cost are expected.

1.2 Objective

To minimize product loss in hair conditioner process by decreasing the deviation of the viscosity using statistical techniques.

1.3 Scopes of this work

1.3.1 To detect the possible parameters in raw material mixing process on the deviation of the viscosity of 5-formula hair conditioner by the cause and effect diagram and brainstorming

1.3.2 To assure the accuracy and precision of the viscosity measurement system by the measurement system analysis (MSA)

1.3.3 To pinpoint the significant parameters in raw material mixing process on the deviation of the viscosity of 5-formula hair conditioner from the possible parameters by using the multiple linear regression analysis in the Minitab program

1.3.4 To determine the strongly significant parameters on the deviation of the viscosity from the significant parameters for the optimum operating conditions by the design of experiments (DOEs) in the Minitab program

1.3.5 To reduce the deviation of the viscosity of hair conditioner by using the obtained optimum operating conditions for product loss minimization

1.4 Research methodology

From a number of parameters in raw material mixing process, the possible parameters that affect the deviation of hair conditioner viscosity were detected by the cause and effect diagram and brainstorming, and the significant parameters were pinpointed by the multiple linear regression analysis in the Minitab program. Before taking the multiple linear regression analysis, the measurement system analysis was employed to assure the accuracy and precision of the measurement devices, inspectors and measuring procedures. Then, the design of experiments was used to determine the strongly significant parameters from the obtained significant parameters to achieve the optimum operating conditions. The potential process capability index (C_p) was evaluated to ensure the in-control process, resulting in product within the control specifications. Finally, the deviation of viscosity and loss in hair conditioner process were minimized.

The research methodology is concluded as follows:

1.4.1 Study the processing steps and viscosities of the target hair conditioners

1.4.2 Review theories and literatures involved, e.g., six sigma, 7 quality tools, the measurement system analysis, the Minitab program, the multiple linear regression analysis, the design of experiments

1.4.3 Identify the possible parameters that affect the deviation of the viscosity by the cause and effect diagram and brainstorming

1.4.4 Check the reliability of the measurement devices, inspectors and measuring procedures of the measurement system for hair conditioner viscosity by the measurement system analysis

1.4.5 Collect the detailed data of the possible parameters. Apply the multiple linear regression analysis in the Minitab program to identify the significant parameters

1.4.6 Determine the strongly significant parameters on the deviation of hair conditioner viscosity from the significant parameters by the design of experiments for the optimum operating conditions

1.4.7 Reduce the deviation of the viscosity of hair conditioner by using the obtained optimum operating conditions for product loss minimization

1.4.8 Conclude the results and write the SOPs

1.5 Outputs of this work

1.5.1 The strongly significant parameters on the deviation of the product viscosity and the optimum conditions

1.5.2 Less deviation of hair conditioner viscosity

1.5.3 Less hair conditioners loss

1.5.4 Higher potential process capability index and less DPMO

1.5.5 The SOPs to produce the target hair conditioners to apply to a similar production

Chapter 2

Theory and Literature Reviews

2.1 Hair conditioners [2-3]

The structure of hair is mainly composed of a protein so-called keratin. Keratin has a high percentage of amino acids which have negative charges. Dirt, natural oils so-called sebum produced by the scalp, and positive charges on hair are removed simultaneously from hair by soap, shampoo, or other cleaners which contain anionic surfactants. Typically, hair conditioners contain positively-charged molecules or cationic surfactants. The cationic surfactants in hair conditioners are attracted to the negative charges on hair and are not rinsed out completely with water. Hair conditioners are hair care products that improve the texture and appearance of human hair. They are coated on dry hair as a thin film, which add weight, make the hair easier to comb, and prevent static electricity from building up and causing the frizzy hair.

There are several types of hair conditioner ingredients, which differ in composition and functionality. The general ingredients are as follows:

1. Cationic surfactants: Hair conditioners usually contain cationic surfactants. The hydrophilic ends of the cationic surfactants strongly bind to negatively-charged amino acids of keratin, and are not rinsed out completely. The hydrophobic ends of the surfactant molecules then act as the new hair surface and help protect the hair from being exposed to the sun light, dry air or harmful environment.

2. Oils (Essential Fatty Acids: EFAs): EFAs are the closest substances to natural oils or sebum. They are added in hair conditioners to make dry or porous hair more soft and pliable.

3. Moisturizers: The moisturizers usually contain high proportions of humectants. They are the main ingredients of hair conditioners and function as a coating layer to hold the moisture in hair.

4. Acids: Acids make the surface of hair scale tighten up. They strengthen the hydrogen bonds between keratins. Theoretically, the acidity of hair conditioners should maintain at the pH of about 2.5–3.5.

5. Detanglers: The detanglers make wet hair slippery and prevent it from entangling, which can damage the hair. Wet hair needs to be treated gently as it is more vulnerable to a

breakage than dry hair. Typical substances used for the detangling purpose are polymers, fatty alcohols, panthenol, dimethicone, etc.

6. Glossers: Glossers are light-reflecting chemicals which can bind to the hair surface. Examples of glossers are polymers and silicones such as dimethicone or cyclomethicone.

7. Perfumes: Perfumes are used to give a pleasant smell.

8. Preservatives: Preservatives are added to retard spoilage and provide a long shelf life of hair conditioners.

2.2 Rheology [4-5]

2.2.1 Viscosity

Viscosity is the measure of the internal friction of a fluid. This friction becomes apparent when a layer of fluid is made to move in relation to another layer. The greater the friction, the greater the amount of force required to cause this movement, which is called shear. Shearing occurs whenever the fluid is physically moved or distributed, such as in pouring, spreading, spraying, mixing, etc. Highly viscous fluids, therefore, require more force to move than less viscous materials.

Isaac Newton defined viscosity by considering the model represented in Figure 2.1.

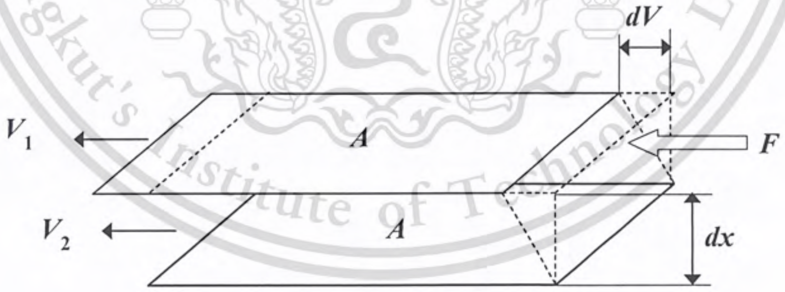


Figure 2.1 Flow model of Newtonian fluids [3].

Two parallel flat areas of fluid of the same size A are separated by a distance dx and are moving in the same direction at different velocities V_1 and V_2. Newton assumed that the force (F) required to maintain this difference in speed was proportional to the difference in speed through the liquid, or the velocity gradient. To express this, Newton wrote:

$$\frac{F}{A} = \eta \frac{dV}{dx} \tag{2.1}$$

where η is a constant for a given material and is called the viscosity.

The velocity gradient ($\frac{dV}{dx}$) is a measure of the change in speed at which the intermediate layers move with respect to each other. It describes the shearing that liquid experiences and is thus called shear rate. It is symbolized as $\dot{\gamma}$ (s^{-1}).

The term F/A indicates the force per unit area required to produce the shearing action. It is referred to as shear stress and symbolized by τ. Its unit of measurement is dyne per square centimeter (dyne/cm^2) or Newton per square meter (N/m^2).

Using these simplified terms, viscosity may be defined mathematically by

$$\eta = \frac{\tau}{\dot{\gamma}} \tag{2.2}$$

or viscosity = shear stress / shear rate.

The fundamental unit of viscosity is poise. A material requiring a shear stress of one dyne per square centimeter to produce a shear rate of one reciprocal second has a viscosity of one poise or 100 centipoises (cP). The viscosity unit can be expressed in pascal-second (Pa•s) or millipascal-second (mPa•s): these units are international system and sometimes used in preference to the metric designations. One pascal-second is equal to ten poises; one millipascal-second is equal to one centipoise.

2.2.2 Newtonian fluids

A Newtonian fluid is represented graphically in Figure 2.2. Graph A shows that the relationship between shear stress (τ) and shear rate ($\dot{\gamma}$) is a straight line. Graph B shows that the fluid viscosity remains constant as the shear rate is varied. Typical Newtonian fluids include water and thin motor oils.

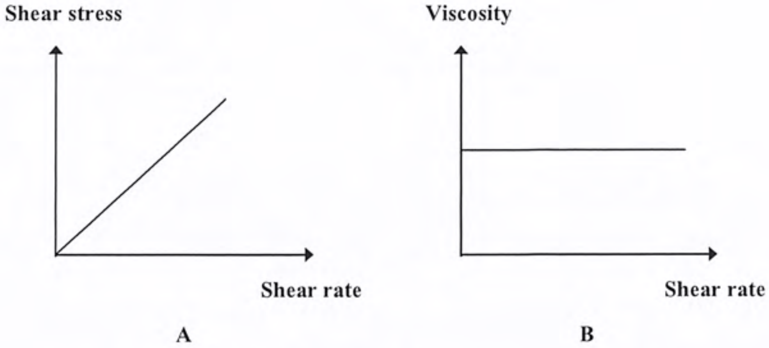


Figure 2.2 Flow behavior of Newtonian fluids.

At a given temperature the viscosity of a Newtonian fluid will remain constant regardless of which viscometer model, spindle or speed are used to measure the viscosity. Newtonian fluids are obviously the easiest fluids to measure and deal with.

2.2.3 Non-Newtonian Fluids

A non-Newtonian fluid is broadly defined as one for which the relationship $\frac{\tau}{\dot{\gamma}}$ is not a constant. In other words, when the shear rate is varied, the shear stress does not vary in the same proportion or even necessarily in the same direction. The viscosity of such fluids will therefore change as the shear rate is varied. Therefore, the measured viscosity of non-Newtonian fluid is called the apparent viscosity.

Non-Newtonian flow can be envisioned by thinking of any fluid as a mixture of molecules with different shapes and sizes. As they pass by each other, as happens during flow, their size, shape, and cohesiveness will determine how much force is required to move them.

There are several types of non-Newtonian flow behaviors, characterized by the way which a fluid viscosity changes in response to variations in shear rate. The most common types of non-Newtonian fluids include:

2.2.3.1 Pseudoplastic

This type of fluid will display a decreasing viscosity with an increasing shear rate, as shown in Figure 2.3.

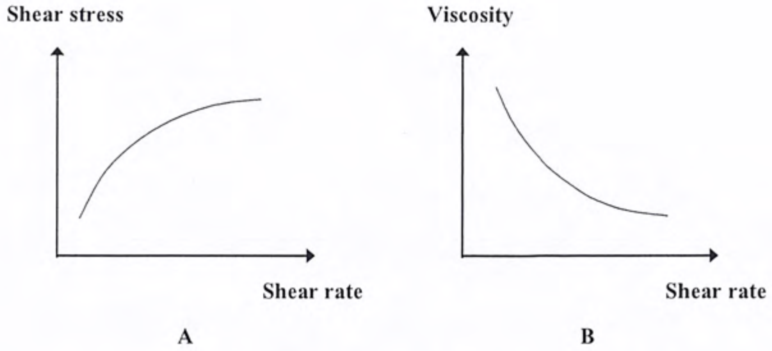


Figure 2.3 Flow behavior of pseudoplastic fluids.

Probably the most common of the non-Newtonian fluids, pseudoplastics include paints and emulsions. This type of flow behavior is sometimes called shear-thinning. An easily understanding model is to imagine that in the moment of tuning the spindle in the sample the structure of molecules of the sample will be destroyed, and the molecule formation will be orientated more parallel to the spindle surface. So the hindering of the spindle rotation will decrease. The faster the rotation will become, the more the structure is destroyed and the less the structure of molecules slide in together, the lower the viscosity will be.

2.2.3.2 Dilatant

Increasing viscosity with an increase in shear rate characterizes the dilatant fluid; see Figure 2.4.

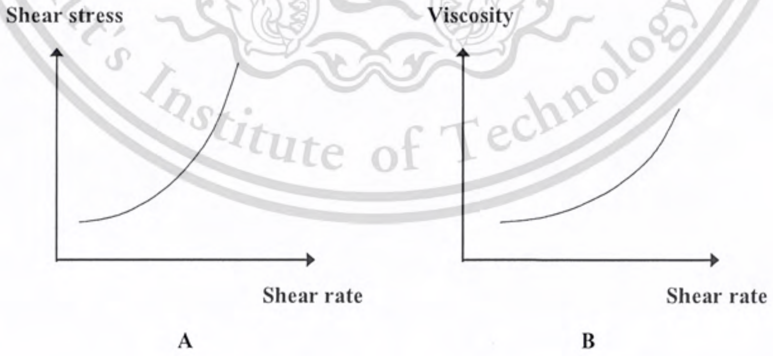


Figure 2.4 Flow behavior of dilatant fluids.

Dilatant fluids is frequently observed in the fluids containing high level of deflocculated solids, such as clay slurries, candy compounds, corn starch in water and sand/water mixtures. This flow behavior is also referred to as shear-thickening.

2.2.3.3 Bingham plastic (Plastic)

This type of fluid will behave as a solid under static conditions. A certain amount of stress must be applied to the fluid before any flow is induced; this stress is called the yield stress. Tomato catsup is a good example of this type fluid. The yield stress will often make the tomato catsup refuse to pour from the bottle until the bottle is shaken or struck, allowing the catsup to gush freely. Once the yield value is exceeded and flow begins, plastic fluids may display Newtonion, pseudoplastic, or dilatant flow characteristics, see Figure 2.5.

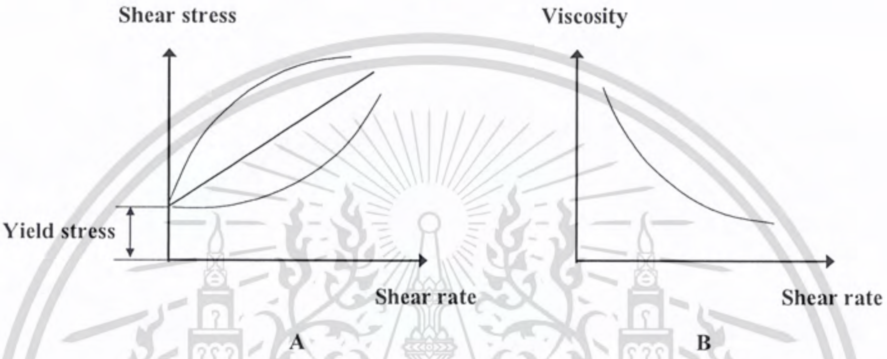


Figure 2.5 Flow behavior of plastic fluids.

So far, the effects of shear rate on non-Newtonian fluids have only been discussed. What happens when the element of time is considered? This question leads us to examination of two more types of non-Newtonian flow which display a change in viscosity with time under conditions of constant shear rate. There are two categories to consider, i.e., thixotropic fluids and rheopectic fluids.

Thixotropic fluids

As shown in Figure 2.6, a thixotropic fluid undergoes a decrease in viscosity with time, while it is subjected to a constant shear rate.

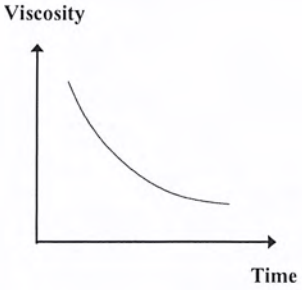


Figure 2.6 Flow behavior of thixotropic fluids.

Rheopectic fluids

This is essentially the opposite of thixotropic behavior, in that the fluid viscosity increases with time as it is sheared at a constant rate, as shown in Figure 2.7.

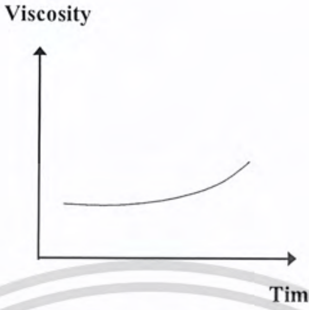


Figure 2.7 Flow behavior of rheopectic fluids.

Both thixotropy and rheopecxy may occur in combination with any of the non-Newtonian flow behaviors, or only at certain shear rates. The time element is extremely variable; under conditions of constant shear, some fluids will reach their final viscosity value in a few seconds, while others may take up to several days. Rheopectic fluids are rarely encountered. Thixotropy, however, is frequently observed in materials such as greases, heavy printing inks and paints.

2.2.5 The power-law model

The power-law model is the fitted model obtained from the experimental results for many fluids over a number of shear rates. It is used extensively to describe the non-Newtonion flow behaviors of fluids in theoretical analyses as well as in practical engineering applications. The power-law model is expressed as:

$$\eta_{app} = m \dot{\gamma}^{n-1} \tag{2.3}$$

where η_{app} is the apparent viscosity (cP), m is the consistency (cP·s^n), $\dot{\gamma}$ is shear rate (s^{-1}), and n (dimensionless) is the degree of deviation from Newtonian behavior or the power-law index. For newtonian fluid, n is equal to 1. The fluid is described as pseudoplastic fluid, when the value of n is less than 1 while the fluid with the value of n greater than unity will behave as dilatant fluid.

2.3 Six sigma [1, 6-7]

Six sigma may be defined in several ways. Tomkins defined six sigma to be “a program aimed at the near-elimination of defects from every product, process and transaction”. Harry defined six sigma to be “a strategic initiative to boost profitability, increase market share and improve customer satisfaction through statistic tools that can lead to breakthrough quantum gains in quality”.

Six sigma was launched by Motorola in 1987. It was the result of a series of changes in the quality area starting in the late 1970s, with ambitious ten-fold improvement drives. The top-level management along with CEO Robert Galvin developed a concept called six sigma. After some internal pilot implementations, Galvin, in 1987, formulated the goal of “achieving Six-sigma capability by 1992” in a memo to all Motorola employees. The results in terms of reduction in process variation were on-track and cost savings totaled US$ 13 billion and improvement in labor productivity achieved 204% increase over the period 1987 – 1997. Six sigma has become very popular throughout the whole world. There are several reasons for this popularity. First, it is regarded as a fresh quality management strategy which can replace TQC, TQM and others. In a sense, we can view the development process of six sigma as shown in Figure 2.8. Many companies, which were not quite successful in implementing previous management strategies such as TQC and TQM, are eager to introduce six sigma.

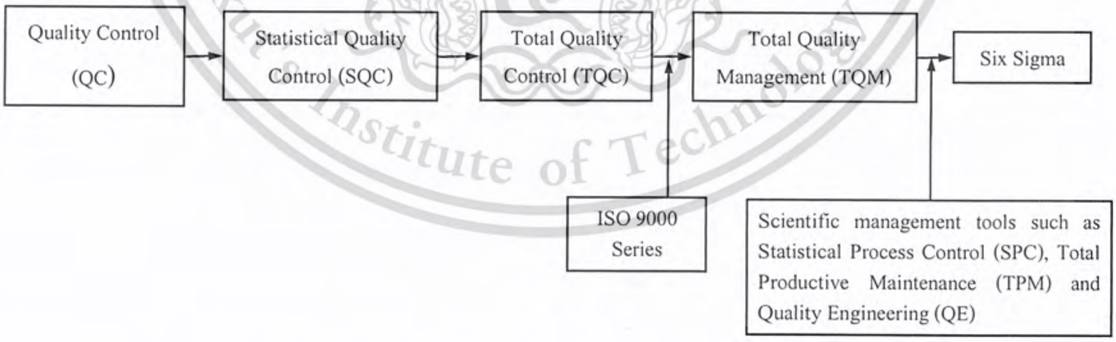


Figure 2.8 Development process of six sigma in quality management.

Six sigma is viewed as a 4S approach (systematic, scientific, statistical and smarter) for management innovation which is quite suitable for use in a knowledge-base information society. Six sigma provides a scientific and statistical basis for quality assessment for all processes through measurement of quality levels. The six sigma method draws comparisons among all processes and tells how good a process is. Through this information, top-level management learns what path to follow to achieve process innovation and customer satisfaction.

2.3.1 Process

A general definition of a process is an activity or series of activities transforming inputs to outputs in a repetitive flow as shown in Figure 2.9.

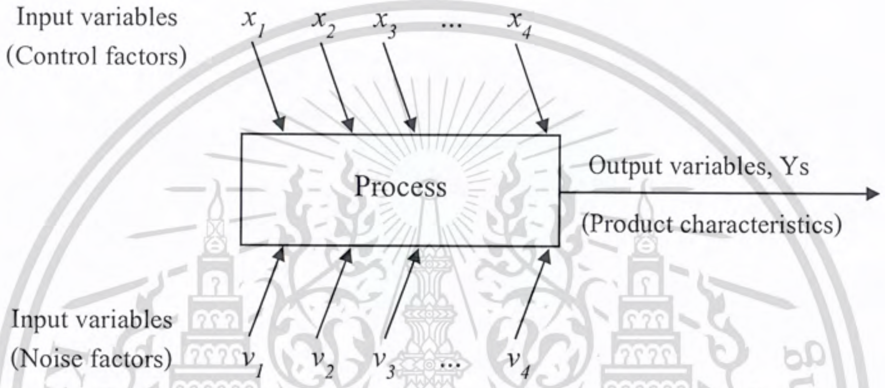


Figure 2.9 Process diagram with input and output variables.

Literally, the inputs can be anything from labor, materials, machines, decisions, information and measurements to temperature, humidity and weight. Inputs are either control factors which can be physically controlled, or noise factors which are considered to be uncontrollable, too costly to control, or not desirable to control.

The model of six sigma in terms of processes and improvement is that y is a function of x and v:

$$y = f(x_1.x_2,....,x_k ; v_1, v_2 ..., v_m) \qquad (2.4)$$

Here, y represents the output or result variable or product characteristic; x represents one or more control factors; and v represents one or more noise factors. The message in the process is to find the optimal level of x variables which give desired values of y as well as being robust to the noise factors *v*. The word "robust" means that the *y* values are not changed much as the level of noise factors are changed.

The characteristic of the collected data from any given process to measure the process performance are classified into continuous and discrete. Continuous characteristics may take any measured value on a continuous scale, providing continuous data, whereas discrete characteristics are based on counts, providing attribute data. Examples of continuous data are thickness, time, speed and temperature. Typical attribute data are the numbers of counts (pass or fail items; acceptable or unacceptable items; or good or bad items).

2.3.2 Variation

The data values for any process or product characteristic always vary. Not any product or characteristic is exactly alike because any process contains many sources of variability. The differences among products may be large or immeasurably small, but they are always present. The variation is the dispersion of the data values. It can be visualized and statistically analyzed by means of the distribution that best fits the observations. This distribution can be characterized by:

- Location using the average value of the data
- Spread using the span or range of the data from the smallest to the largest values
- Shape depending on the pattern of the distribution whatever it is symmetrical, skewed, etc.

Variation is indeed the number one enemy of quality control. It constitutes a major cause of defectives as well as excess costs. Six sigma, through its tracking of process performance and formalizing improvement methodology, focuses on pragmatic solutions for reducing variation. Variation is the key element of the process performance triangle as shown in Figure 2.10. Variation relates to "how close are the measured values to the target value", cycle time relates to "how fast" and yield relates to "how much". Cycle time and yield are the two major elements of productivity.

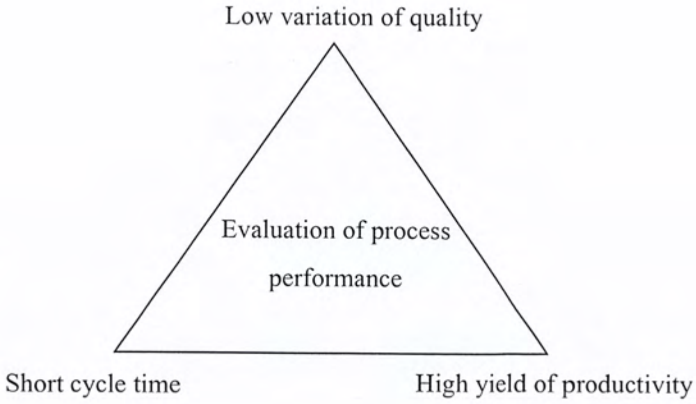
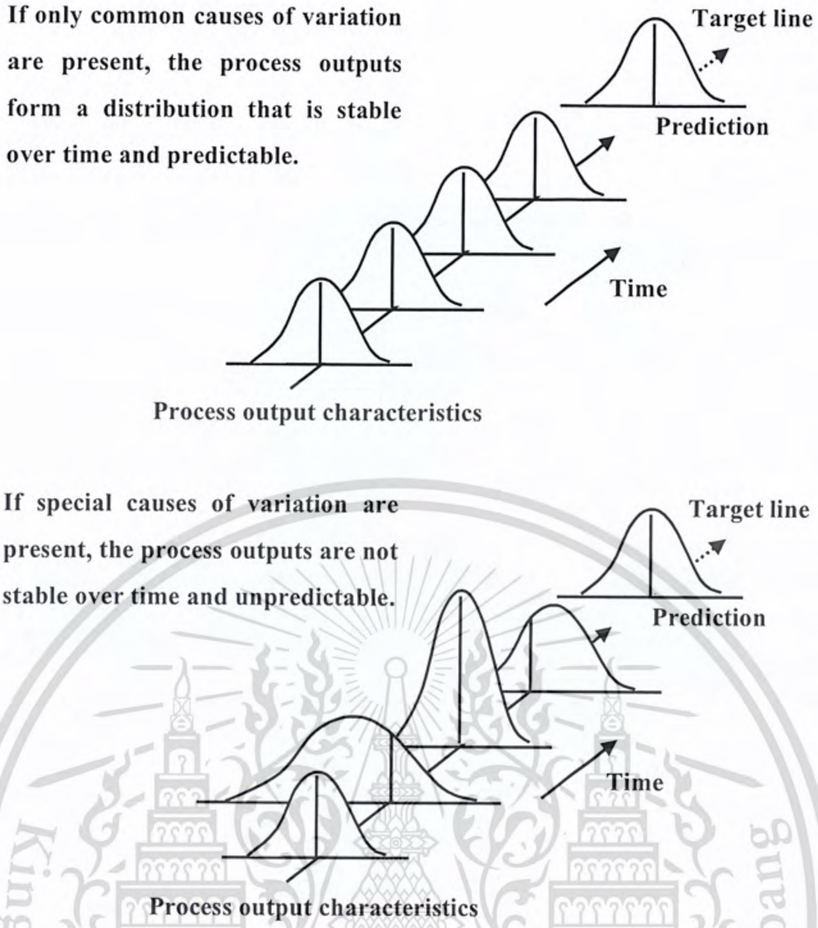


Figure 2.10 Process performance triangle.

It is common to classify sources of variation for process and outputs into two types: common causes and special causes.

Common causes refer to the sources of variation within a process that has a stable and repeatable distribution over time. This is called "in a state of statistical control". The random variation, which is inherent in the process, is not easily removable unless we change the design of the process or product, and is a common cause found everywhere. Common causes behave like a stable system of chance causes. If only common causes of variation are present and do not change, the output of a process is predictable as shown in Figure 2.11.

Special causes, often called assignable causes, refer to any factors causing variation that are usually not present in the process. That is, when the special causes occur, they make a change in the process distribution. Unless all the special causes of variation are identified and acted upon, they will continue to affect the process output in unpredictable ways. If special causes are present, the process output is not stable over time.



If only common causes of variation are present, the process outputs form a distribution that is stable over time and predictable.

If special causes of variation are present, the process outputs are not stable over time and unpredictable.

Figure 2.11 The common and special causes of variation.

2.3.3 Measurement of process performance

Among the dimensions of the process performance triangle in Figure 2.10, variation is the preferred measurement for process performance. Cycle time and yield could have been used, but they can be covered through variation. For example, if a cycle time has been specified in a process, the variation of the cycle time around its target value can be used to indicate the process performance. The same applies to yield.

The distribution of a process output characteristic is usually assumed to be normal or Gaussian distribution for continuous variables, whereas the widely used distributions of the discrete variables are binomial distribution and Poisson distribution. Basically, in both cases of continuous and discrete variables, the two important parameters that describe the distribution profile are mean (μ) and standard deviation (σ). The mean indicates the location of the distributions on the continuous or discrete scales and the standard deviation indicates the dispersion or the spread of the process output characteristic values. In this work, the normal distribution is focused because the hair conditioner viscosity is the continuous variable.

2.3.3.1 Probability distribution and normal distribution

A probability distribution is a mathematical model that relates the value of the population characteristic with the probability of occurrence of that value in the population. There are two types of probability distributions.

1. Continuous distribution

When the variable being measured is expressed on a continuous scale, its probability distribution is called a continuous distribution. The distribution of dimension of a part, the distribution of the piston-ring diameter, and the distribution of hair conditioner viscosity are examples of the continuous distributions.

2. Discrete distribution

When the variable being measured can only take on a certain integer, such as 0, 1, 2, …, the probability distribution is called a discrete distribution, for example, the distribution of the number of nonconformances or defects in printed circuit boards.

If there is a population of size N and all values of a population characteristic are X_1, X_2, …, X_N, then the population parameters, i.e., population mean (μ), population standard deviation (σ) and population variance (σ^2) are calculated as

$$\mu = \frac{X_1 + X_2 + \ldots + X_N}{N} \tag{2.5}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}} \tag{2.6}$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N} \tag{2.7}$$

Generally, the population parameters (μ , σ and σ^2) are usually unknown, and they are estimated by the sample statistics as follows:

$\bar{x}$: sample mean = population mean (μ)

s : sample standard deviation = population standard deviation (σ)

s^2 : sample variance = population variance (σ^2)

If there is a sample of size n and all values of a sample characteristic are x_1, x_2, ..., x_n, then $\bar{x}$, s and s^2 are estimated by

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \tag{2.8}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n-1}} \tag{2.9}$$

$$s^2 = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n-1} \tag{2.10}$$

For the continuous distributions, an important distribution is a normal distribution. Many continuous variables follow the normal distribution. Figure 2.12 illustrates the bell shape of a normal distribution where x is the normal variable, μ is the population mean and σ is the population standard deviation. The probability density function (PDF), *f(x)*, of the normal distribution is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2.11}$$

The normal distribution is usually denoted by $X \sim N(\mu, \sigma^2)$. The cumulative normal distribution is defined as the probability that the normal variable x is less than or equal to any a value, or

$$P\{x \le a\} = F(x = a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \tag{2.12}$$

where *F(x)* is the cumulative distribution function of the normal distribution.

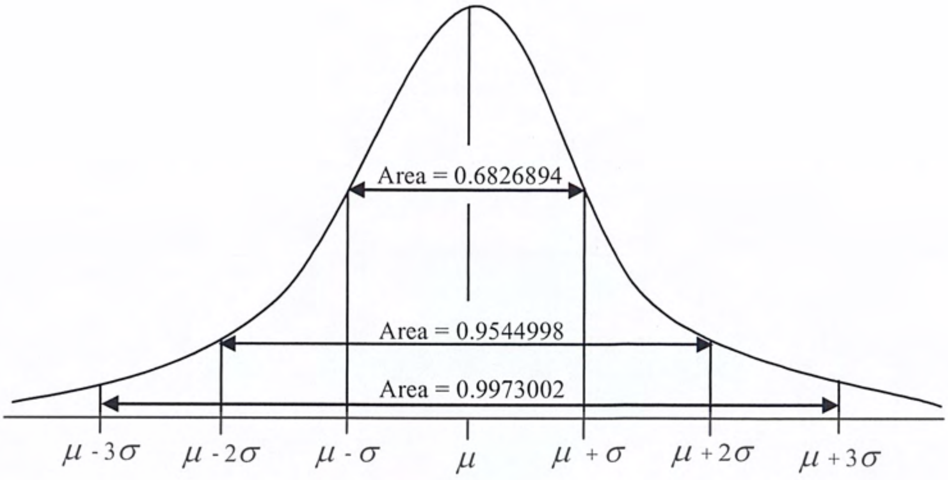


Figure 2.12 Normal distribution: μ = population mean, σ = population standard deviation.

The integral from Equation 2.12 is in an improper form that cannot be evaluated directly. However, $X \sim N(\mu, \sigma^2)$ can be converted into standard normal variable $Z \sim N(0,1)$ by using the relationship of variable transformation,

$$Z = \frac{X - \mu}{\sigma} \tag{2.13}$$

whose probability density function is

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \tag{2.14}$$

The evaluation of the cumulative standard normal distribution can be made independent of μ and σ^2 as follows:

$$P\{x \le a\} \quad = \quad P\left\{z \le \frac{a - \mu}{\sigma}\right\} \quad = \quad \Phi\left(\frac{a - \mu}{\sigma}\right) \tag{2.15}$$

and

$$\Phi\left(\frac{a - \mu}{\sigma}\right) \quad = \quad \Phi(z_a) \quad = \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_a} e^{-\frac{1}{2}u^2} du \tag{2.16}$$

where $\Phi(z_a)$ is the cumulative distribution function of the standard normal distribution (population mean = 0, population standard deviation = 1). The integral values of the cumulative distribution are very useful in working with the normal distribution. However, this integral cannot be integrated by one of the methods of calculus. The cumulative standard normal distribution integral values have been shown in Appendix A-1. The cumulative standard normal distribution is illustrated by the area under the standard normal distribution curve. Total area under the curve represents all numbers of the population and is equal to 1. From Figure 2.12, it is found that 68.26 %, 95.44 % and 99.73 % of the population values fall between the limits defined by $\mu \pm \sigma$, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$, respectively.

2.3.3.2 Defect rate, parts-per-million defect rate and defects per million opportunities

The defect rate, denoted by *p*, is the ratio of the number of defective items which are out of specification to the total number of items which are processed or inspected. Defect rate or fraction of defective items out of one million inspected items is called parts-per-million (ppm) defect rate. Sometimes a ppm defect rate cannot be properly used, particularly, in case of the service work. In this case, defects per million opportunities (DPMO) are often used. DPMO is the number of defective opportunities which do not meet the required specification out of one million possible opportunities.

2.3.3.3 Defects per unit (DPU), Defects per unit opportunity (DPO) and DPMO

Let us suppose that a certain product design may be represented by the area of a rectangle, and each rectangle contains eight equal areas of opportunity for nonconformance or defect to standard. Figure 2.13 illustrates three particular products. The first one has one defect and the third one has two defects.

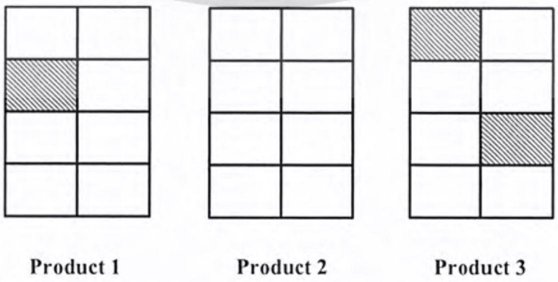


Figure 2.13 Products of eight equal areas of opportunity for nonconformance.

The defects per unit is defined as

$$\text{DPU} = \frac{Total \quad observed \quad number \quad of \quad defects}{Total \quad number \quad of \quad unit \quad products \quad produced} \tag{2.17}$$

From Figure 2.13, the DPU is 3/3 = 1.00 that means, on average, each unit product contains one defect. Of course, this can be assumed that the defect is randomly distributed.

The defects per unit opportunity can be calculated by

$$\text{DPO} = \frac{\text{DPU}}{m} \tag{2.18}$$

where m is the number of independent opportunities for nonconformance per unit. From the illustrated example in Figure 2.13, $m = 8$, therefore,

$$\text{DPO} = \frac{1.00}{8} = 0.125 = 12.5\%$$

Inversely, there is 87.5 percent chance that does not encounter a defect with respect to any given unit area of opportunity. The defects per million opportunities can be calculated by

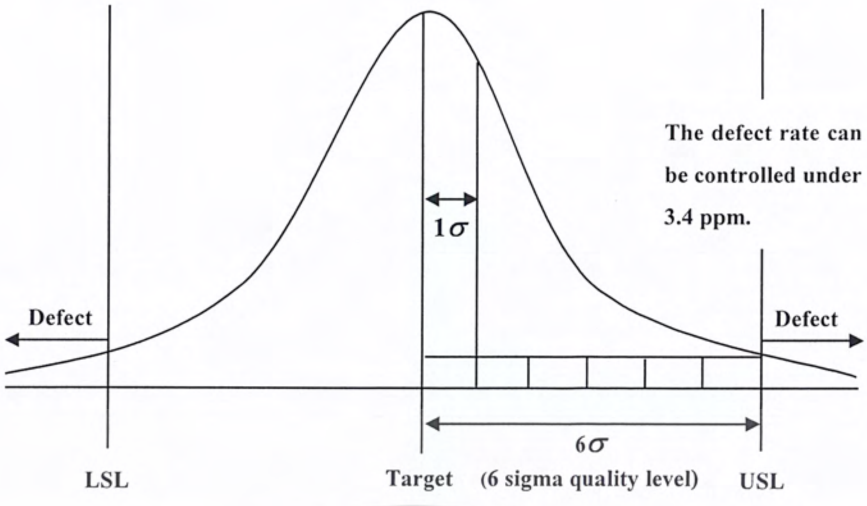
$$\text{DPMO} = \frac{\text{DPU}}{m} \times 1{,}000{,}000 \tag{2.19}$$

From the same example the DPMO becomes

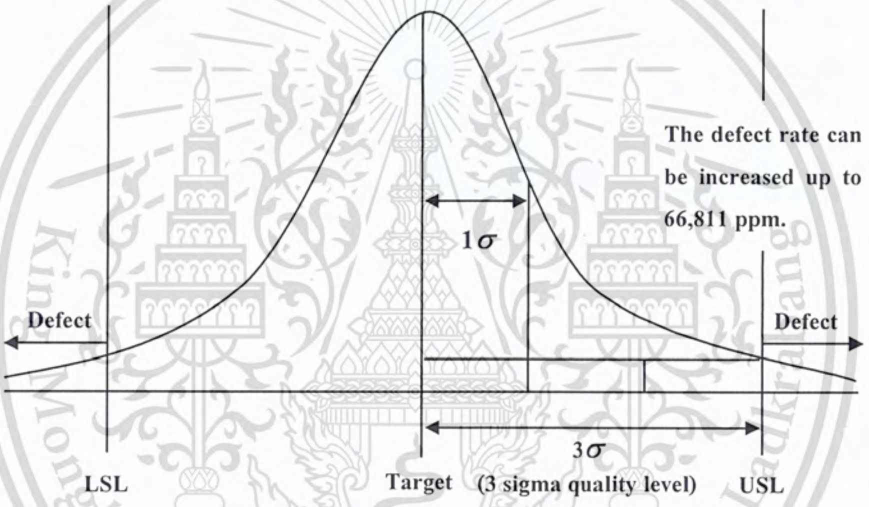
$$\text{DPMO} = \frac{1.00}{8} \times 1{,}000{,}000 = 125{,}000$$

2.3.3.4 Sigma quality level or sigma level

The specification limits are the tolerances or performance ranges that the customers accept the products or processes. Figure 2.14 illustrates the specification limits by the two major vertical lines. The LSL is the lower specification limit, USL is the upper specification limit, and target represents the target value of the process output characteristic. The sigma quality level, in short sigma level, is the distance from the process mean (μ) to the closer specification limit (Figures 2.14 and 2.15).



(a)



(b)

Figure 2.14 Sigma quality levels of 6 σ (a) and 3 σ (b).

In practice, the process mean is set at the target value. However, during one time period the process mean is usually different from another time period for various reasons, such as the different lots of raw materials, different operators, etc. This means the process mean constantly shifts around the target value. To address the typical maximum shift of the process mean, the Motorola added the shift value of ±1.5 σ to the process mean. The calculation of the sigma quality level, in case the shift of the process mean exists, is shown in Figure 2.15.

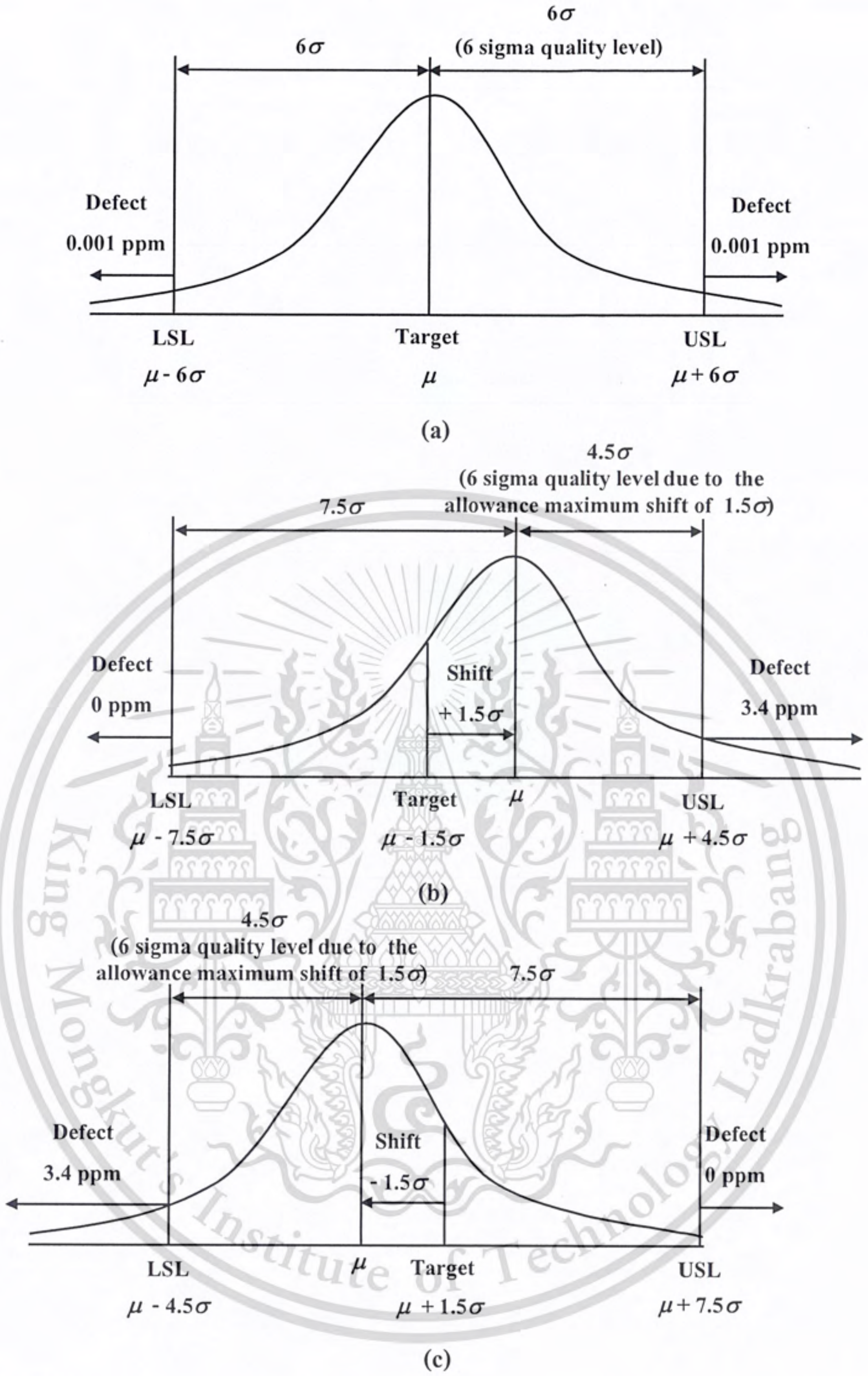


Figure 2.15 Effects of the ±1.5 σ shifts of the process mean from the target on quality level: (a) no shift, (b) +1.5 σ shift and (c) -1.5 σ shift.

From Figure 2.15, the 6 σ quality level corresponds to a 3.4 ppm defect rate. Table 2.1 shows how the sigma quality level relates to the defect rate and the relationship when the process mean is ±1.5 σ shifted.

Table 2.1 Parts-per-million defect rate (ppm) with sigma quality level.

Sigma level	**Process mean without 1.5σ shift**		**Process mean with ±1.5σ shift**	
	Non-defect rate (%)	**Defect rate (ppm)**	**Non-defect rate (%)**	**Defect rate (ppm)**
1σ	68.26894	317,311	30.2328	697,672
2σ	95.44998	45,500	69.123	308,770
3σ	99.73002	2,700	93.3189	66,811
4σ	99.99366	63.4	99.379	6,210
5σ	99.999943	0.57	99.97674	233
6σ	99.9999998	0.002	99.999966	3.4

2.3.3.5 Process capability

Two metrics for measuring the process capability are the potential process capability index (C_p) and the process capability index (C_{pk}).

1. Potential process capability index (C_p)

C_p is defined as the ratio of the specification width over the process spread as shown in Equation 2.20.

$$C_p = \frac{\text{specification width}}{\text{process spread}} = \frac{USL - LSL}{6\sigma} \tag{2.20}$$

The specification width is predefined and fixed. The process spread influences on the C_p. The population standard deviation, σ, is usually estimated from the sample standard deviation (s) in Equation 2.9. Of wide process spread or more variation, the C_p value is small indicating a low process capability. On the other hand, a better process capability is attained when the process spread is narrow or less variation. Table 2.2 presents the recommended minimum values of the C_p and the corresponding DPMO in the processes.

Table 2.2 Recommended minimum values of the potential process capability index (C_p) and the corresponding defects per million opportunities (DPMO).

Type of process	C_p	DPMO
Existing process: general	1.33	6,209.7
New process: general	1.50	1,349.9
Existing process: safety, strength or critical parameter	1.50	1,349.9
New process: safety, strength or critical parameter	1.67	224.1

The C_p does not account for any process shift. The ideal state is assumed when the process is at the desirable target, centered exactly between the two specification limits.

2. Process capability index (C_{pk})

In fact, very few processes are at their desirable target. An off-target process should be "penalized" for shifting from where it should be. The C_{pk} shows this real capability when the off-target penalty is taken into consideration. The penalty or degree of bias (k) is defined as:

$$k = \frac{|\text{target} - \text{process mean}(\mu)|}{\frac{1}{2}(\text{USL} - \text{LSL})} \tag{2.21}$$

and the process capability index is defined as:

$$C_{pk} = C_k\,(1 - k) \tag{2.22}$$

When the process is perfectly on target, $k = 0$ and $C_{pk} = C_p$. Note that the C_{pk} increases when both of the following conditions are satisfied.

- the process mean is very close to the target (small k)
- small process spread or less variation

In general, there are two specification limits, USL and LSL. However, when there is a one-sided specification limit, or when the target is not specified, C_{pk} is calculated by:

$$C_{pk} = \frac{\left|\text{process mean } (\mu) - \text{closer specification limit from } \mu\right|}{3\sigma} \quad (2.23)$$

3. Relationship between the C_p, C_{pk} and sigma quality level

If the process mean is centered, that is μ = target, and the distance from the USL to LSL = 6σ, then the distance from μ to either of the specification limits is 3σ and the $C_p = 1$. The relationship between C_p and the sigma quality level is

$$\text{Sigma quality level} = 3 \times C_p \quad (2.24)$$

In this case the sigma quality level becomes 3σ.

However, in a long run the process mean could shift at most by 1.5σ to the right or left hand side, and the process mean is not centered. Therefore, the relationship between C_{pk} and the sigma quality level is

$$\text{Sigma quality level} = 3 \times C_{pk} + 1.5 \quad (2.25)$$

For example, in long run with the process mean of 1.5σ bias and $C_{pk} = 1$, the sigma quality level becomes $(3 + 1.5)\ \sigma = 4.5\sigma$. Hence, the relationship between C_p and C_{pk} is

$$C_{pk} = C_p - 0.5 \quad (2.26)$$

The relationship between the C_p and C_{pk} and sigma quality level shows in Table 2.3.

Table 2.3 Relationship between the C_p and C_{pk} and sigma quality level.

C_p	C_{pk} ($\pm 1.5\sigma$ shift is allowed)	Sigma level
0.50	0.00	1.5
0.67	0.17	2.0
0.83	0.33	2.5
1.00	0.50	3.0
1.17	0.67	3.5
1.33	0.83	4.0
1.50	1.00	4.5
1.67	1.17	5.0
1.83	1.33	5.5
2.00	1.50	6.0

2.3.4 DMAIC process

The most important methodology in six sigma management is perhaps the formulized improvement characterized by the DMAIC (define-measure-analyze-improve-control) process. This DMAIC process works well as a breakthrough strategy. The methodology works equally well on variation of quality, cycle time, yield of productivity, design, and others. It is divided into five phases as shown in Figure 2.16.

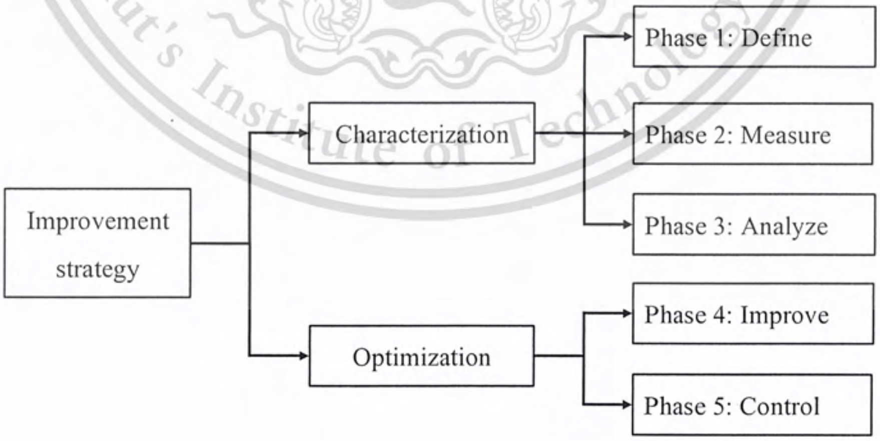


Figure 2.16 Improvement strategy of six sigma.

In each phase the major activities are as follows:

Phase 1: (Define)

This phase concerns the identification of the process or product that needs improvement. It works on benchmarking of the key product and/or process characteristics of other world-class companies.

Phase 2: (Measure)

This phase entails selecting product characteristics; i.e., the dependent variables, mapping the respective processes, making the necessary measurements, recording the results and estimating the short- and long-term process capabilities. The quality function deployment (QFD) and failure modes and effects analysis (FMEA) play a major role in selecting critical product characteristics.

Phase 3: (Analyze)

This phase concerns the analyzing and benchmarking of the key product and/or process performance metrics. Consequently, a gap analysis is often undertaken to identify the common factors of the successful performance; i.e., what factors explain the best-in-class performance. In some cases, it is necessary to redefine the performance goal. In analyzing the product and/or process performance, various statistical and basic quality control (QC) tools are used.

Phase 4: (Improve)

This phase relates to selecting those product and/or process performance characteristics which must be improved to achieve the performance goal. Once this is done, the product and/or process characteristics are diagnosed to reveal the major sources of variations. Next, the key product and/or process characteristics are identified usually by the statistical techniques such as the hypothesis testing, analysis of variance (ANOVA), multiple regression analysis and design of experiments (DOEs). The improved conditions of the key product and/or process characteristics are verified.

Phase 5: (Control)

This last phase is initiated by ensuring that the new product and/or process conditions are documented and monitored via the statistical process control (SPC) methods. After the "settling in" period, the process capability is reassessed. Depending upon the outcomes of such a follow-on analysis, it may become necessary to revisit one or more of the preceding phases.

2.4 Seven quality control (7 QC) tools

The 7 QC tools are graphical and statistical tools which are most often used in QC for continuous improvement. Since the 7 QC tools are so utilized by almost every level of the company, they have been so-called the magnificent seven. The 7 QC tools are applicable to the improvements in all dimensions of the process performance triangle, i.e., the variation of quality, cycle time and yield of productivity.

Each of the 7 QC tools has been used separately before 1960. However, in the early 1960s, they were gathered together by a small group of the Japanese scientists led by Kaoru Ishikawa with the aim of providing the QC circles with effective and easy-to-use tools. They are, in alphabetical order, cause and effect diagram, check sheet, control chart, graph, histogram, Pareto chart and scatter diagram. In six sigma, the 7 QC tools extensively used in all of the DMAIC improvement phases.

2.4.1 Cause and effect diagram

An effective tool as part of a problem solving process is the cause and effect diagram, also known as the Ishikawa diagram or fishbone diagram. This technique is useful to trigger ideas and promote a balanced approach in group brainstorming sessions where individuals list the perceived sources or causes with respect to the outcome or effect. As shown in Figure 2.17, the effect is written in a rectangle on the right-hand side, and the causes are listed on the left-hand side. They are connected with arrows to show the causes and effect relationship.

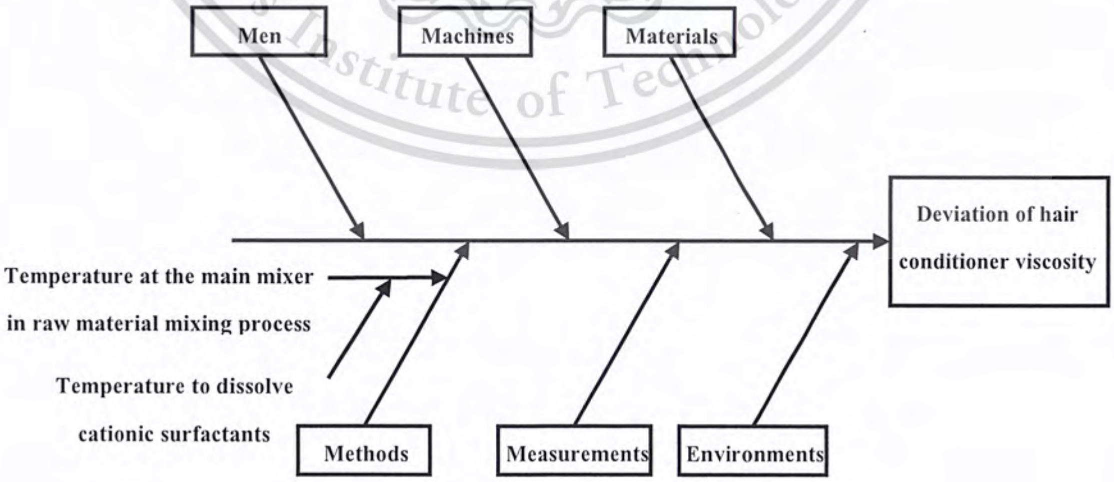


Figure 2.17 An example of the cause and effect diagram.

Constructing of a cause and effect diagram is often appropriate to consider six main causes that can contribute to the effect; so-called 5M1E, i.e., man, machine, material, method, measurement and environment. The first step is to agree on the specific wording of the effect and then to identify the main causes that can possibly produce the identified effect. The main causes can often be identified as any of 5M1E, which helps the users to get started. Using brainstorming techniques, each main cause is analyzed. The aim is to refine the list of causes in more details until the root causes of that particular main cause are established. The same procedure is then followed for each of the other main causes. An example of the cause and effect diagram is given in Figure 2.17. The effect or problem that needs improvements is the deviation of hair conditioner viscosity. For the concern of the main cause in terms of the method, the temperature at the main mixer in raw material mixing process is the subcause, and the main mixer temperature to dissolve cationic surfactants is the root cause of the deviation of hair conditioner viscosity.

2.4.2 Check sheet

The check sheet is used to collect the specific data of any desired characteristics of a process or product that needs improvement. It is frequently used in the measure phase of the six sigma. For practical purposes, the check sheet is commonly formatted as a table. It is important that the check sheet is kept simple. The design should align to the characteristics that are measured. Table 4.2 shows an example of a check sheet for defective items in an assembly process of automobile ratios.

Table 2.4 Example of check sheet for defective items.

Defective item	**Date**					**Sum**
	Aug. 1	**Aug. 2**	**Aug. 3**	**Aug. 4**	**Aug. 5**	
Soldering defect	//	/	///		𝍸	11
Joint defect	//	//		/	///	8
Lamp defect		/	//	//	/	6
Scratch defect	𝍸	𝍸 /	///	𝍸 ///	//	24
Miscellaneous	/	//	///	/	//	9
Sum	10	12	11	12	13	58

2.4.3 Control chart

The control chart is a very important tool in the analyze, improve and control phases of the six sigma. In the analyze phase, control charts are applied to judge if the process and/or product is predictable; in the improve phase, to identify the evidence of special causes of the variation of quality; in the control phase, to verify that the performance of the process and/or product is under control.

The original concept of the control chart was proposed by Walter A. Shewhart in 1924 and the tool has been used extensively in industry since the Second World War, especially in Japan and the USA after about 1980. Control charts offer the study of variations. The control charts can give process monitoring and control, and can also give direction for improvements. They can separate the special causes from the common causes of variations in a process. They can give early identification of the special causes so that there can be timely resolution before many poor quality products are produced.

Shewhart's control charts track product and/or process by plotting data over time in the form shown in Figure 2.18. This chart can track either variables or attribute process parameters. The types of variable charts are process mean ($\bar{x}$), range (R), standard deviation (s), individual value (x) and moving range (Rs). The attribute types are fraction nonconforming (p), number of nonconforming items (np), number of nonconformities (c) and nonconformities per unit (u).

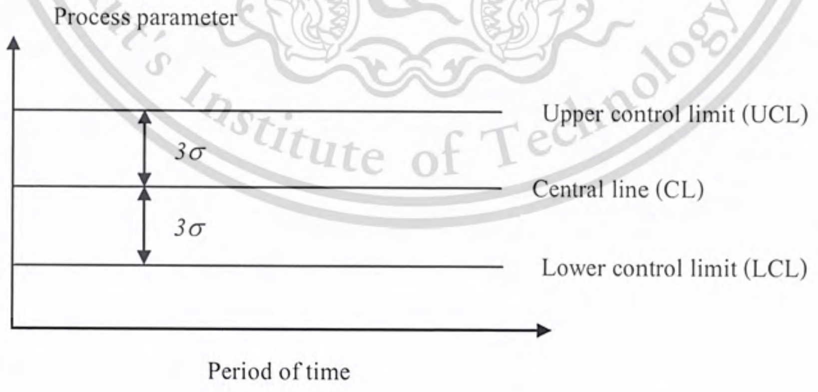


Figure 2.18 Shewchart's control chart format.

The typical control limits are $\pm 3\sigma$, where the standard deviation is calculated by using about 20-30 data points. If a point falls outside these limits, the process is said to be out of control. On the other hand, if a point falls inside these limits, the process is said to be in control.

Types of control charts depend on the nature and quantity of the characteristics of interest. For continuous data, two types of the control charts are simultaneously used in the same way as a single control chart.

The most often used control charts for continuous data are:

$\bar{x}-\mathrm{R}$	average and range chart
$\bar{x}-\mathrm{s}$	average and standard deviation chart
$x-\mathrm{Rs}$	individual observation and moving range chart

The most often used control charts for discrete or attribute data are:

P	fraction of nonconforming items chart
np	number of nonconforming items chart
c	number of defects chart
u	number of defects per unit chart

How to construct the control charts

A detailed generic sequence to construct the control charts are developed.

Step 1: Select the characteristic and type of control chart

First, the decision must be made regarding the characteristic of the process or product that is checked or supervised for predictability of performance. Then, the proper type of control chart is selected.

Step 2: Determine the sample size and sampling interval

Control charts are, in most cases, based on a constant number of samples or observations, n. For continuous data, it is common to use two to six observations. However, there are also charts for subgroup sizes of one, x chart (individual observation chart) and Rs chart (moving range chart). For discrete data, n can be as large as 100 or 200.

Step 3: Calculate the control lines and center line

All control charts have control limits, UCL and LCL. The distance from the center line (CL) to UCL or LCL is 3 standard deviations of the characteristic.

For n individual observations, $x_1, x_2, x_3 \ldots x_n$, of the $\bar{x}$ chart:

$$\text{CL} \quad = \bar{x} \tag{2.27}$$

$$\text{UCL} = \bar{x} + 3 \text{ (standard deviation of } x) \tag{2.28}$$

$$\text{LCL} = \bar{x} - 3 \text{ (standard deviation of } x) \tag{2.29}$$

Step 4: Draw the control chart and check for special causes

The control chart can now be drawn with CL, UCL and LCL. The samples used for calculating the control limits are then plotted on the chart to determine whether the samples used to calculate the control limits embody any special cause variation. Special causes exist if any of the following alarm rules apply:

- A single point falls outside the UCL or LCL.
- 2 out of 3 consecutive points fall outside the ± 2 standard deviations from the CL.
- 7 or more consecutive points fall to one side of the CL.
- A run of 8 or more consecutive points is up for an increasing trend, or down for a decreasing trend.
- At least 10 out of 11 consecutive points are on one side of the CL.
- At least 8 consecutive points make a cycle movement. The cycle movement occurs when a point is on one side of the center line and the next point is on the other side of the center line.

2.4.4 Graphs

Graphs are often used to present information visually. For a complicate information which is difficult to understand and needs an illustration, graphs can help by displaying the points quickly and visually. Line graph, bar graph and pie chart are widely used.

A line graph is used to summarize how two pieces of information are related to each other, for example, the strength of plastic product and the process temperature.

A bar graph consists of a vertical axis, a horizontal axis and a series of labeled horizontal or vertical bars that show different values for each bar.

A pie chart is a circle graph divided into pieces. Pie chart is usually used to display sizes of different parts that make up some whole.

2.4.5 Histogram

It is meaningful to present data in a form that visually illustrates the frequency of occurrence of values. In the analyze phase of the six sigma, the histograms are commonly applied to show the distribution of the data collected in the measure phase, and they are also used to obtain an understanding of the potential improvements.

A histogram is created for the response data on certain discrete values. After a number of data are taken, the tally of the grouping of occurrences is then plotted in the histogram form. Figure 2.19 shows a histogram of 200 rolls of two dices. The sum scores of the two dices are 2 to 12 which are the discrete data. However, for a histogram of the continuous data, the data are grouped into classes or intervals. The height of each bar in the histogram is made proportional to the number of observations within each data value or interval. The histogram shows both the process variation and type of distribution that the collected data entails.

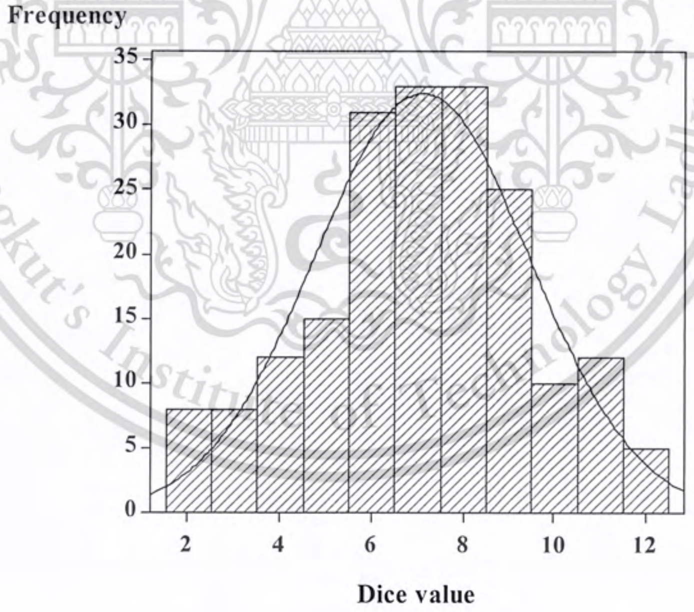


Figure 2.19 Histogram of 200 rolls of two dices.

2.4.6 Pareto chart

The Pareto chart was introduced in the 1940s by Joseph M. Juran, who named it after the Italian economist and statistician Vilfredo Pareto. It is applied to distinguish the vital few from the trivial many causes of the problem. It is closely related to the 80/20 rule; 80% of the problem stems from 20% of the causes, or in the six sigma terms; 80% of the poor values in Y stem from 20% of the Xs.

In the six sigma improvement methodology, the Pareto chart has two primary applications. One is for selecting appropriate improvement projects in the define phase. It offers a very objective basis for selection, based on, for example, frequency of occurrence, cost saving or potential improvement in the process performance.

The other primary application is in the analyze phase for identifying the vital few causes or Xs of the problem that will constitute the greatest improvement in Y if appropriate measurements are taken.

The procedures to construct a Pareto chart are as follows:

1. Define the problem and process characteristics to use in the diagram.
2. Define the period of time for the diagram, for example, weekly, daily, or shift. The quality improvements over time can later be made from the information determined within this step.
3. Obtain the total number of times each characteristic occurred.
4. Rank the characteristics according to the total number from step 3.
5. Plot the number of occurrences of each characteristic in descending order in a bar graph along with a cumulative percentage overlay.
6. Trivial columns can be lumped under one column designation, however, care must be exercised not to omit small but important items.

Table 2.5 shows a summary table in which a total of 50 claims during the first month of 2002 are classified into six different reasons. Figure 2.20 is the Pareto chart of the relevant data in Table 2.5.

Table 2.5 Example of claim reason.

Claim reason	Frequency	Percentage of frequency (%)	Cumulative frequency	Percentage of cumulative frequency (%)
A	23	46	23	46
B	10	20	33	66
C	7	14	40	80
D	3	6	43	86
E	2	4	45	90
Others	5	10	50	100

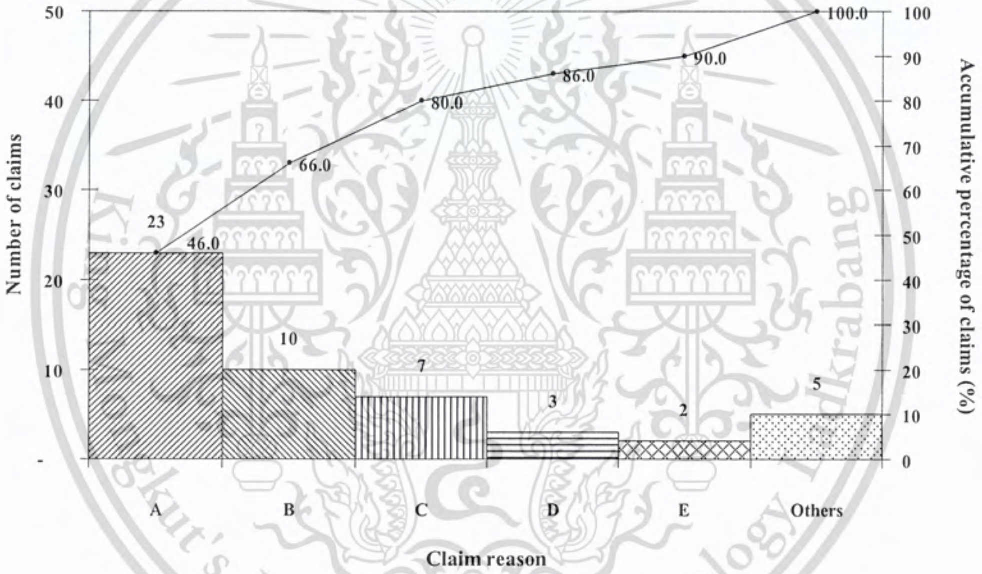


Figure 2.20 Pareto chart of 50 claims in January 2002.

2.4.7 Scatter diagram

The scatter diagram is useful to discover the correlation between two factors, X and Y. An important feature of the scatter diagram is its visualization of the correlation pattern, through which the relationship can be determined. In the improve phase of the six sigma, one often searches the collected data for Xs that have a special influence on Y. Knowing the existence of such relationships, it is possible to identify the input variables that cause special variation of the result variable. It is then able to determine how to set the input variables, if they are controllable, so that the process is improved. Though several Xs may influence the values of Y, one scatter diagram is drawn for each combination of the Xs and Y.

It is common to place the input variable, X, on the X-axis and the response variable, Y, on the Y-axis in constructing the scatter diagram. The two variables can now be plotted against each other and a scatter pattern of plotted points appears. A basic understanding of the relationship between X and Y is given, and a basis for improvement is provided.

Table 2.6 shows a set of data of the relationship between the total batch cycle time (X) and hair conditioner viscosity (Y) in raw material mixing process. Figure 2.21 shows a scatter diagram of the data in Table 2.6.

Table 2.6 Data for the total batch cycle time (X) and hair conditioner viscosity (Y) in raw material mixing process.

X (min)	Y (cP)	X (min)	Y (cP)
131	22,990	129	23,010
135	23,360	135	23,420
136	23,620	134	23,160
130	22,860	126	22,870
132	23,160	133	23,620
133	23,280	134	23,630
132	22,890	130	23,010
131	23,000	131	23,120
128	23,080	136	23,500
134	23,640	133	22,750

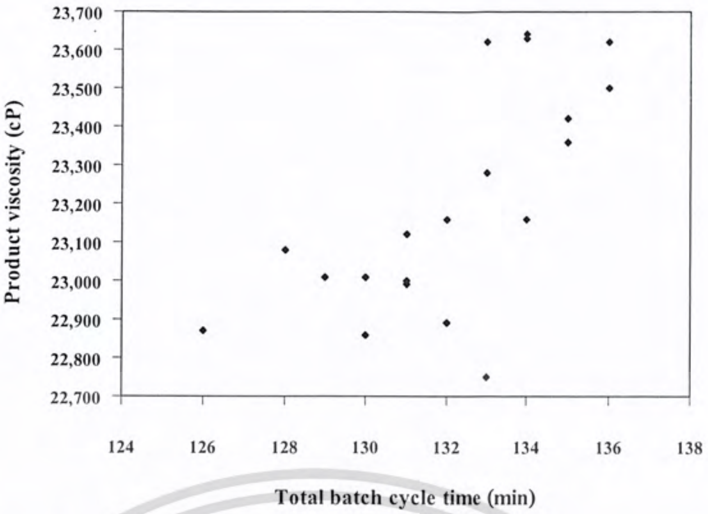


Figure 2.21 Scatter diagram of the total batch cycle time and hair conditioner viscosity.

2.5 Multiple linear regression analysis [8-10]

Engineers frequently use models in problem formulation and solution. Sometimes these models are based on the physical, chemical, or engineering science knowledge of the phenomenon, and in such cases these models are call mechanistic model. Examples of mechanistic models are Ohm's law and Kirchhoff's law. However, there are many situations in which two or more variables of interest are related, and the mechanistic model of these variables is unknown. In these cases it is necessary to build a model relating the variables based on the observed data. This type of model is called an empirical model. The empirical model can be manipulated and analyzed just as the mechanistic model can.

Multiple linear regression models are often used for the empirical model. In a multiple linear regression model, one dependent variable or response (y) is related to k independent variables or regressors ($x_1, x_2, \ldots, x_k$). Therefore the model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \tag{2.30}$$

where the parameters β_j ($j = 0, 1, \ldots, k$) are called regression coefficients, and ε is the error term. The parameter β_j represents the expected change in response y per unit change in x_j when all the remaining regressors x_i $(i \neq j)$ are held constant. In this work, the dependent variable or response is the viscosity of hair conditioner and the independent variables or regressors (x_j) are the parameters that have influential effects on hair conditioner viscosity in raw material mixing process.

The models which are more complex than Equation 2.30 are often analyzed by the multiple linear regression technique. For example, consider the cubic polynomial model in one regressor variable as follows:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon \qquad (2.31)$$

If let $x_1 = x,\ \ x_2 = x^2,\ \ x_3 = x^3$, Equation 2.31 can be written as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \qquad (2.32)$$

which is a multiple linear regression model with three regressors.

Furthermore, the models which include interaction effects are also analyzed by the multiple linear regression method. An interaction between two variables can be represented by a cross-product term in the model, such as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \qquad (2.33)$$

If let $x_3 = x_1 x_2$ and $\beta_3 = \beta_{12}$, Equation 2.33 can be written as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \qquad (2.34)$$

which becomes a linear regression model.

2.5.1 Least-square estimation of the multiple linear regression coefficients

The method of least square can estimate the regression coefficients in the multiple linear regression model. The multiple linear regression model from *n* observations may be written as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \qquad (2.35)$$

Suppose $n > k$ observations are available, where k is the number of regressors in the regression model, and let y_i denotes the i th observed response and x_{ij} denotes the i th level of regressor x_j. The ε_i denotes the error term of the i th observation. Primarily, the ε_i in the model is assumed to be uncorrelated random variables and normally distributed with mean zero and constant variance. The multiple linear regression data from n observations is shown in Table 2.7.

Table 2.7 Data for the multiple linear regression [8].

Observation (i)	**Response** (y_i)	**Regressors** (x_{ij})			
		x_{i1}	x_{i2}	...	x_{ik}
1	y_1	x_{11}	x_{12}	...	x_{1k}
2	y_2	x_{21}	x_{22}	...	x_{2k}
⋮	⋮	⋮	⋮	...	⋮
n	y_n	x_{n1}	x_{n2}	...	x_{nk}

The multiple linear regression model in Equation 2.35 can be written as:

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \qquad i = 1, 2, \ldots, n \tag{2.36}$$

The least-square function is

$$S(\beta_0, \beta_1, \ldots, \beta_k) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2 \tag{2.37}$$

The function S must be minimized with respect to $\beta_0, \beta_1, \ldots, \beta_k$. The least-square estimators of $\beta_0, \beta_1, \ldots, \beta_k$ must satisfy

$$\left. \frac{\partial S}{\partial \beta_0} \right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0 \tag{2.38}$$

and $$\left.\frac{\partial S}{\partial \beta_j}\right|_{\hat{\beta}_0,\hat{\beta}_1,\ldots,\hat{\beta}_k} = -2\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k}\hat{\beta}_j\, x_{ij} \right)x_{ij} = 0,\quad j = 1,\, 2,\, \ldots,\, k \qquad (2.39)$$

Simplifying Equations 2.38 and 2.39, the least-square normal equations are obtained as follows:

$$n\hat{\beta}_0 + \hat{\beta}_1\sum_{i=1}^{n}x_{i1} + \hat{\beta}_2\sum_{i=1}^{n}x_{i2} + \ldots + \hat{\beta}_k\sum_{i=1}^{n}x_{ik} = \sum_{i=1}^{n}y_i$$

$$\hat{\beta}_0\sum_{i=1}^{n}x_{i1} + \hat{\beta}_1\sum_{i=1}^{n}x^2{}_{i1} + \hat{\beta}_2\sum_{i=1}^{n}x_{i1}x_{i2} + \ldots + \hat{\beta}_k\sum_{i=1}^{n}x_{i1}x_{ik} = \sum_{i=1}^{n}x_{i1}y_i$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\hat{\beta}_0\sum_{i=1}^{n}x_{ik} + \hat{\beta}_1\sum_{i=1}^{n}x_{ik}x_{i1} + \hat{\beta}_2\sum_{i=1}^{n}x_{ik}x_{i2} + \ldots + \hat{\beta}_k\sum_{i=1}^{n}x_{ik}^2 = \sum_{i=1}^{n}x_{ik}y_i \qquad (2.40)$$

Note that the number of normal equations is $p = k+1$, one for each unknown regression coefficient. By solving these normal equations simultaneously, the solutions which are the least-square estimators $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ will be obtained.

It is more convenient to deal with multiple regression models if they are expressed in matrix notation. This allows a very compact display of the multiple regression model, data and results. In matrix notation, the model given by Equation 2.36 is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & x_{3k} \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

In general, $\mathbf{y}$ is an $n \times 1$ vector of the observations, $\mathbf{X}$ is an $n \times p$ matrix of the levels of the regressor variables, $\boldsymbol{\beta}$ is a $p \times 1$ vector of the regression coefficients, and $\boldsymbol{\varepsilon}$ is a $n \times 1$ vector of random errors.

The vector of least-square estimators, $\hat{\boldsymbol{\beta}}$, will be obtained by minimizing the following equation.

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Note that $S(\boldsymbol{\beta})$ is expressed as

$$S(\boldsymbol{\beta}) = \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

$$= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

since $\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}$ is a 1 × 1 matrix or a scalar, and its transpose $(\boldsymbol{\beta}'\mathbf{X}'\mathbf{y})' = \mathbf{y}'\mathbf{X}\boldsymbol{\beta}$ is the same scalar. The least-square estimators must satisfy

$$\left.\frac{\partial S}{\partial \boldsymbol{\beta}}\right|_{\hat{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

and simplify to

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y} \qquad (2.41)$$

Matrix equations of Equations 2.41 represent the least-square normal equation, which is the matrix analogue of the scalar normal Equations 2.40.

To solve the normal equations, multiply both sides of Equations 2.41 by the inverse of $\mathbf{X}'\mathbf{X}$. Thus, the least-square estimators of $\boldsymbol{\beta}$, or $\hat{\boldsymbol{\beta}}$ are

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \qquad (2.42)$$

The inverse matrix $(\mathbf{X'X})^{-1}$ will always exist if the regressors are linearly independent. The linear or near-linear dependencies among the regressors are multicollinearity, which can seriously affect the precision of the estimated regression coefficients and will be discussed later in session 2.5.2 (Multicollinearity). From Equation 2.42, the least-square estimators or $\hat{\boldsymbol{\beta}}$, i.e., $\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k$, will be obtained.

The matrix form of the least-square normal equations of Equations 2.41 is identical to the scalar form. Writing out the least-square normal equations of Equations 2.41 in detail, we obtain

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x^2_{i1} & \sum_{i=1}^{n} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{n} x_{i1}x_{ik} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik}x_{i1} & \sum_{i=1}^{n} x_{ik}x_{i2} & \cdots & \sum_{i=1}^{n} x^2_{ik} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik} y_i \end{bmatrix}$$

If the matrix multiplication is performed, the scalar form of the least-square normal equations of Equations 2.40 is obtained.

From the least-square estimation of the model coefficients, the fitted multiple linear regression model is

$$\hat{y} = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j \, x_j \tag{2.43}$$

The difference between the observed value y_i and the corresponding fitted value $\hat{y}_i$ is the residual, which is expressed as

$$e_i = y_i - \hat{y}_i \tag{2.44}$$

2.5.2 Multicollinearity

A serious problem that may dramatically impact the usefulness of a regression model is multicollinearity, or near-linear dependence among the regressors or independent variables. The regressors are the columns of the **X** matrix, therefore, an exact linear dependence would clearly result in the absence of the inverse matrix $(\mathbf{X'X})^{-1}$. The presence of near-linear dependencies can dramatically impact the ability to estimate the regression coefficients.

The main diagonal elements of the inverse matrix $(\mathbf{X'X})^{-1}$, which have the same row (*i*) and column (*j*) subscripts, are called variance inflation factors (VIFs) and are very important multicollinearity diagnostic. The VIF of the *j* regression coefficient can be written as:

$$VIF_j = \frac{1}{1-R_j^2} \tag{2.45}$$

where R_j^2 is the coefficient of the multiple determination obtained from the regressor x_j on the other regressors. Clearly, if x_j is near-linear dependence on the other regressors, then R_j^2 will be near unity and $\mathrm{VIF_j}$ will be large. The VIFs larger than 10 imply serious problems with multicollinearity. Most regression computer programs compute and display the $\mathrm{VIF_j}$.

The regression model is a poor-predictioned equation when strong multicollimearity exists. The values of the regression coefficients are often very sensitive to the particular samples.

2.5.3 Hypothesis testing in multiple linear regression

Once the regressor coefficients in the multiple linear regression model have been estimated, two immediate questions are raised:

1. What is the overall adequacy of the model?
2. Which specific regressors seem important?

Several hypothesis testing procedures prove useful for addressing these questions. Primarily, the hypothesis testing for the multiple linear regression requires some assumptions. Theoretically, the error term (ε_i) of the multiple linear regression model shown in Equation 2.35 is assumed to be uncorrelated random variables and normally distributed with mean zero and constant variance. The residual model diagnostics are employed to verify these assumptions.

In the residual model diagnostics, the residuals from the fitted multiple linear regression model are defined in Equation 2.44, and are checked whether they are normally distributed and constant variance or not. In common, normal residuals are verified by constructing a histogram of the residuals or a normal probability plot of residuals. Since, many computer programs will produce a normal probability plot of residuals and the sample sizes in the regression analysis are often too small to make the histogram meaningful, the normal probability plot is preferable. Similarly, the assumption due to the constant variance is checked from the plot of residuals and the fitted values of the response variable ($\hat{y}_i$) or the observed value of the response variable (y_i).

2.5.3.1 Test for the significance of regression

The significance of regression is tested to check a linear relationship between the response (y) and any of the regressors (x_1, x_2, ..., x_k). This procedure is for an overall or global test of model adequacy. The hypotheses are

$$\begin{aligned} &H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0 \\ &H_1 : \beta_j \neq 0 \quad \text{for at least one } j \end{aligned} \tag{2.46}$$

Rejection of this null hypothesis implies that at least one of the regressors x_1, x_2, ..., x_k contributes significantly to the model.

The analysis of variance is employed in this test. The total variability in the dependent variable or response is described by the total sum of squares (SS_T). The total sum of squares is a sum of squares due to regression (SS_R) and a residual sum of squares (SS_{Res}) as shown:

$$SS_T = SS_R + SS_{\mathrm{Res}} \tag{2.47}$$

The test procedures of the analysis of variance are summarized in Table 2.8.

Table 2.8 The analysis of variance for the significance of regression in multiple linear regression.

Source of variation	**Sum of squares (*SS*)**	**Degree of freedom**	**Mean square (*MS*)**	**Test statistic (F_0)**
Regression	SS_R	k	MS_R	MS_R / MS_{Res}
Residual	SS_{Res}	$n-k-1$	MS_{Res}	
Total	SS_T	$n-1$		

The computational formulas for the sum of squares in the matrix form are as follows:

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \tag{2.48}$$

$$SS_R = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \tag{2.49}$$

$$SS_{\text{Res}} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \tag{2.50}$$

The mean square of regression (MS_R) and the mean square of residual (MS_{Res}) are computed by dividing each sum of squares with its degree of freedom, therefore

$$MS_R = \frac{SS_R}{k} \tag{2.51}$$

and

$$MS_{\text{Res}} = \frac{SS_{\text{Res}}}{n-k-1} \tag{2.52}$$

To draw a conclusion whether at least one of the regressors contributes significantly to the response variable, the test statistic (F_0) would be computed. Rejection of the null hypothesis is made if

$$F_0 \ > \ F_{\alpha,\ k,\ n-k-1} \tag{2.53}$$

where $F_{\alpha,\ k,\ n-k-1}$ is the rejection criterion of the test statistic obtained from the *F* distribution table at the α significance level or (1-α) confident level of the test, *k* regressor or numerator degree of freedom, and *n-k-1* denominator or error degree of freedom. The *F* distribution table is shown in Appendix A-2. The null hypothesis is rejected when the F_0 is greater than the $F_{\alpha,\ k,\ n-k-1}$, indicating that at least one of the regressors contributes significantly to the response variable.

Another test statistic for the significance of regression in multiple linear regression is the p-value test statistic. This method considers the probability value (p-value) and has been adopted widely in practice and reported by computer programs such as the Minitab program.

Probability value (p-value)

Literally, the p-value is the probability that would lead to rejection of the null hypothesis, H_0. The p-value is computed without imposing a preselected level of significance like other test statistics, e.g., *F* test, *t* test. Once after the p-value is calculated, the null hypothesis can be decided to be rejected or not and the significance of the data are attained by the difference between the p-value and the acceptable level of significance. If the p-value is less than the acceptable level of significance, the null hypothesis H_0 will be rejected, indicating that the data are significant. Generally, the acceptable level of significance is selected at 0.05 or 0.10.

2.5.3.2 Check of model adequacy by R^2

In the previous session, the hypothesis testing is employed to conclude that at least one regressor contributes significantly to the model. However, it does not necessarily imply that the model is good enough to predict the response variable. Therefore, the test of model adequacy is ensured this requirement by using the coefficient of determination or R^2, which is computed by

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{\mathrm{Res}}}{SS_T} \tag{2.54}$$

The test of model adequacy indicates how precise the multiple linear regression model can predict the values of the response variable. Large R^2 means that the regression model is adequate and can satisfactorily use to explain and predict the response variable values. The quality of R^2 sometimes reports as 100 R^2. It is referred to the percentage of variation explained by the regression model. Generally, the acceptable value of R^2 depends on how much accuracy of the model that an analyst needs, however, R^2 should not less than 0.6. In this work, R^2 of 0.8 is used.

2.5.3.3 Test for individual regression coefficients

Once at least one of the regressors which is important is determined, a logical question becomes which one is the most important. In general, R^2 never decreases when a regressor is added to the model. However, the addition of a regressor increases the variance of the fitted value ($\hat{y}$). Therefore, in order to include only regressors that are of real value in explaining the response is of concern. Although by adding an unimportant regressor in the regression model may increase the residual mean square, it does not prove to be a good model to predict the response variable values precisely.

The hypotheses in testing the significance of any individual regression coefficient, such as β_j, are

$$\begin{aligned} H_0 &: \beta_j = 0 \\ H_1 &: \beta_j \neq 0 \end{aligned} \tag{2.55}$$

If $H_0 : \beta_j = 0$ is not rejected, then this indicates that the regressor x_j can be deleted from the model. The test statistic (t_0) for this hypothesis is

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \tag{2.56}$$

where C_{jj} is the diagonal element of the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ corresponding to the $\hat{\beta}_j$, and the term $\hat{\sigma}^2$ is given by

$$\hat{\sigma}^2 = MS_{\mathrm{Res}} \tag{2.57}$$

The null hypothesis $H_0 : \beta_j = 0$ is rejected if

$$|\ t_o\ | > t_{\alpha/2, n-k-1} \tag{2.58}$$

where $t_{\alpha/2,n-k-1}$ is the rejection criterion of the test statistic obtained from the *t* distribution table at α significance level or (1-α) confident level of the test, and *n-k-1* degree of freedom. The *t* distribution table is shown in Appendix A-3.

The null hypothesis is rejected when the absolute value of t_0 is greater than the $t_{\alpha/2,n-k-1}$, indicating that the regressor contributes significantly to the response variable. On the other hands, for any individual regression coefficient, the null hypothesis $H_0 : \beta_j = 0$ is not rejected if t_0 is less than the rejection criterion. This indicates that the regressor x_j does not significantly contribute to the response variable and can remove from the model. In some situations, t_0 indicates that more than one regressor is not significant. The corrective approach in this situation is to remove the most insignificant regressor, of which t_0 is the lowest, and refit the model. Then, the hypothesis testing is done again to determine another insignificant regressor in the obtained model, if any. This corrective approach is continued until no insignificant regressor is found in the model. Finally, only the regressors that play important role on the response variable will be determined.

2.6 Design of experiments (DOEs) [6, 9, 11]

Experiments are carried out by researchers or engineers in all fields of study to compare the effects of several conditions or to discover something new. A scientific approach to plan an efficient experiment must be considered. The design of experiments is a process in planning the minimum number of experiments to collect and analyze the data by statistical methods for the reliable results.

The DOEs plays a major role in many engineering activities, for instance, the DOEs is used for

1. Improving the performance of a manufacturing process to determine the optimal values of process variables.

2. The development of new processes. The application of DOE methods in early process development can reduce the production time, variation of target requirements, and enhance process yields.

3. Screening important factors.

4. Engineering design activities such as evaluation of material alternations, comparison of basic design configurations, and selection of design parameters.

5. Building the empirical model to determine the functional relationship between *x* and *y*.

Large manufacturing industries in Japan, Europe and the US have applied the DOEs from the 1970s to achieve cost savings and income growth through reduction of variation and cycle time, and increasing of process yields.

The statistical approach in the experimental design is important to draw meaningful conclusions from the collected data. Two important aspects to any experimental design are the design of experiments and the statistical analysis of the collected data. They are closely related since the statistical analysis method depends on the design employed.

A simple, but very meaningful, model is that $y = f(x)$, where y represents the response variable of importance for the customers, and x represents the input variables so-called factors in the DOEs. The question is which factors have important effects on the response variable and how to determine the levels of the factors.

2.6.1 Classification of DOEs

DOEs are classified as follows according to the allocation of factor combinations and the degree of randomization of experiments.

2.6.1.1 Factorial design

This is a design for investigating all possible treatment combinations which are formed from the factors of interest. The order in which possible treatment combinations are selected is completely random. Examples of the factorial designs are the two-factor and three-factor factorial designs, and k factors at two levels (2^k) and k factors at three levels (3^k) factorial designs.

2.6.1.2 Fractional factorial design

This is a design for investigating a fraction of all possible treatment combinations which are formed from the factors of interest. Designs that use tables of orthogonal arrays, Plackett-Burman design and Latin square design are fractional factorial designs. This fractional factorial design is used for the experiment of high cost and time consuming.

2.6.1.3 Randomized complete block design

All possible treatment combinations are tested in this design but some forms of the restrictions are imposed on randomization. The design in which each block contains all possible treatments, and the only randomization of treatments is within the blocks known as the randomized complete block design.

2.6.1.4 Incomplete block design

The incomplete block design is the randomized complete block design of which every treatment does not exist in each block. This design is used when we may not be able to run all the treatments in each block because of a shortage of experimental apparatus or inadequate facilities.

2.6.1.5 Response surface design

This is a design that is used to explore a regression model and find a functional relationship between the response variable and the factors involved, and to find the optimal conditions of the factors. Central composite designs, rotatable designs, simplex designs, mixture designs and evolutionary operation (EVOP) designs belong to this class.

2.6.1.6 Robust design

In 1986, Taguchi developed the robust design, which is often called the parameter design and tolerance design. The concept of the robust design is to find a set of conditions for design variables which are robust to noise, and to achieve the smallest variation in a product characteristic and a desired target value.

2.6.2 Factorial design

For several variables or factors of interest in an experiment, a factorial design is used. By a factorial experiment, each complete trial or replicate which contains all possible combinations of the levels of the factors are investigated. With two factors A and B of a levels and b levels, each replicate of the experiments contains all ab possible combinations.

The effect of a factor is defined as the change in response produced by a change in the level of the factor. This is called a **main effect** because it refers to the primary factors (A and B) in the study. For example, consider the data that are shown geometrically as the square in Figure 2.22, which is an example of a factorial design of factor A and B. The main effect of factor A is the difference between the average response at the first level of A (A_1) and the average response at the second level of A (A_2).

$$A = \frac{30+40}{2} - \frac{10+20}{2} = 20$$

That is, changing of factor A from level 1 to level 2 causes an increase in average response of 20 units. Similarly, the main effect of B is

$$B=\frac{20+40}{2}-\frac{10+30}{2}=10$$

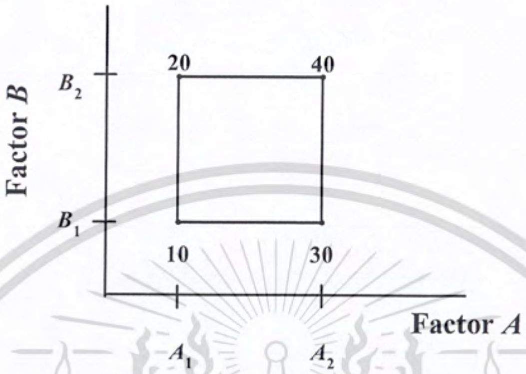


Figure 2.22 A two-factor factorial design and the observed data without interaction.

In some experiments, the difference of responses between the two levels of one factor is not the same at the different levels of the other factor. When this occurs, there is an interaction between the two factors. For example, consider the data that are shown in Figure 2.23.

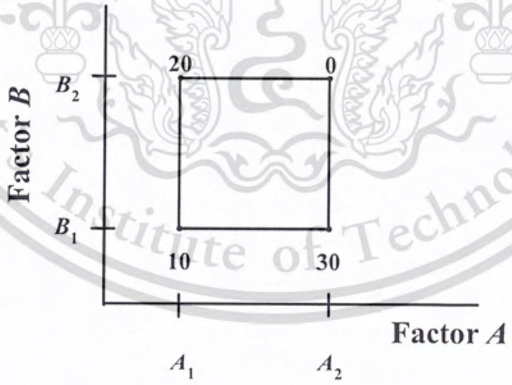


Figure 2.23 A two-factor factorial design and the observed data with interaction.

At level 1 of factor B (B_1), the effect of A is

$$A = 30-10 = 20$$

and at level 2 of factor B (B_2), the effect of A is

$$A = 0 - 20 = -20$$

Because the effect of A depends on the level for factor B, there is an interaction between A and B.

For a large interaction, the corresponding main effects have little meaning. For example, from the data in Figure 2.23, the main effect of A is

$$A = \frac{30+0}{2} - \frac{10+20}{2} = 0$$

showing that there is no effect from factor A. In fact, the effect of factor A exists and depends on the levels of factor B.

It is easy to estimate the interaction effect in the factorial experiment such as those illustrated in Figures 2.22 and 2.23. In this type of experiment, for two factors having two levels, the AB interaction effect is the difference in the diagonal averages. This represents one half of the difference between the A effects at the two levels of B. For example, in Figure 2.22, the AB interaction is

$$AB = \frac{20+30}{2} - \frac{10+40}{2} = 0$$

Thus, there is no interaction between A and B. In Figure 2.23, the AB interaction effect is

$$AB = \frac{20+30}{2} - \frac{10+0}{2} = 20$$

It is mentioned earlier that the interaction effect in this case is very large.

The concept of interaction can be illustrated graphically. Figure 2.24 plots the data of levels B in Figure 2.22 against the levels of A. Lines B_1 and B_2 are roughly parallel, indicating that factors A and B do not interact.

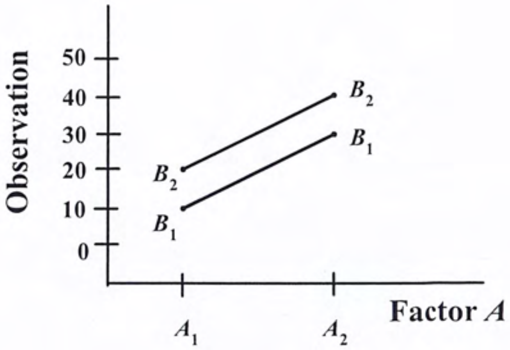


Figure 2.24 An interaction plot of factorial experiment without interaction.

Figure 2.25 plots the data of levels *B* in Figure 2.23 against the level of *A*. Lines B_1 intersects with line B_2, showing the interaction between factors *A* and *B*. Such graphical displays are two-factor interaction or interaction effect plots. They are useful in presenting the results of experiments. The computer programs are used to analyze the data from the designed experiments and construct these graphs automatically.

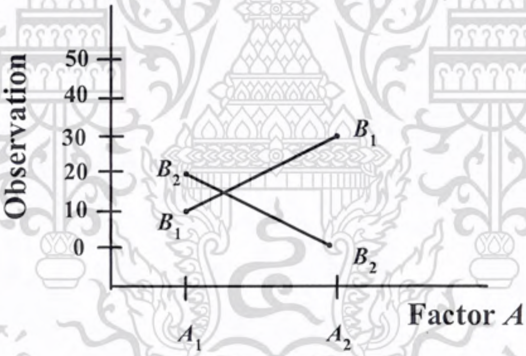


Figure 2.25 An interaction plot of factorial experiment with interaction.

2.6.3 The two-factor factorial design

The simplest types of factorial designs involve with two factors or two sets of treatments. Suppose the *a* levels of factor *A* and *b* levels of factor *B* are arranged in the factorial design, that is, each replicate of the experiment contains all *ab* treatment combinations. In general, there are *n* replicates. Let y_{ijk} represents the observed response when the factor *A* is at *i* th level ($i = 1, 2, \ldots, a$) and factor *B* is at the *j* th level ($j = 1, 2, \ldots, b$) for the *k* th replicate ($k = 1, 2, \ldots, n$).

The two-factor factorial design is shown in Table 2.9. The order in which the *abn* observations are taken is selected randomly so that this design is a completely randomized design.

Table 2.9 General arrangement of a two-factor factorial design.

Factor *B*

	Levels	1	2	…	*b*
Factor *A*	1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
	2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
	a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

The observations in the factorial design are described by the effects model, which is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \in_{ijk} \qquad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases} \tag{2.59}$$

where μ is the overall mean effect, τ_i is the effect of the *i* th level of the row or factor *A*, β_j is the effect of the *j* th level of column or factor *B*, $(\tau\beta)_{ij}$ is the effect of the interaction between τ_i and β_j, and $\in_{ijk}$ is a random error component. Both factors *A* and *B* are fixed, and the treatment effects are defined as the deviation from the overall mean, therefore

$$\sum_{i=1}^{a} \tau_i = 0$$

and
$$\sum_{j=1}^{b} \beta_i = 0 \tag{2.60}$$

Similarly, the interaction effects are fixed and defined as follows:

$$\sum_{i=1}^{a}(\tau\beta)_{ij} = \sum_{j=1}^{b}(\tau\beta)_{ij} = 0 \tag{2.61}$$

Because there are *n* replicates of the experiment, therefore, *abn* total observations exist.

2.6.4 Hypothesis testing of the two-factor factorial design

In the two-factor factorial design, each of the rows and columns of treatments are assumed to have the equal effects on the response variable. The hypothesis testing is based on the equality of the row of treatment effect or factor *A* effect, which is expressed statistically by

$$\begin{aligned} &H_0 : \tau_1 = \tau_2 = \ldots = \tau_a = 0 \\ &H_1 : \text{at least one } \tau_i \neq 0 \end{aligned} \tag{2.62}$$

for the equality of the column of treatment effect or factor *B* effect,

$$\begin{aligned} &H_0 : \beta_1 = \beta_2 = \ldots = \beta_b = 0 \\ &H_1 : \text{at least one } \beta_j \neq 0 \end{aligned} \tag{2.63}$$

To determine whether the row or column of treatment interact or not, the hypotheses are

$$\begin{aligned} &H_0 : (\tau\beta)_{ij} = 0 \\ &H_1 : \text{at least one } (\tau\beta)_{ij} \neq 0 \qquad \text{for all } i, j \end{aligned} \tag{2.64}$$

where H_0 is the null hypothesis and H_1 is the alternative hypothesis. These hypotheses, Equations 2.62, 2.63 and 2.64, are statistically tested by using the two-factor analysis of variance.

Let $y_{i..}$ denotes the total of all observations under the i th level of factor *A*, $y_{.j.}$ denotes the total of all observations under the *j* th level of factor *B*, $y_{.ij}$ denotes the total of all observations in the *ij* th cell, and $y_{...}$ denotes the grand total of all the observations. Define $\bar{y}_{i..}$, $\bar{y}_{.j.}$, $\bar{y}_{ij.}$ and $\bar{y}_{...}$ as the corresponding row, column, cell and grand average, respectively, and are expressed mathematically as follows:

$$y_{i..} = \sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} \quad , \qquad \bar{y}_{i..} = \frac{y_{i..}}{bn} \qquad i = 1, 2, \ldots, a \tag{2.65}$$

$$y_{.j.} = \sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk} \quad , \qquad \bar{y}_{.j.} = \frac{y_{.j.}}{an} \qquad j = 1, 2, \ldots, b \tag{2.66}$$

$$y_{ij.} = \sum_{k=1}^{n} y_{ijk} \quad , \qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n} \qquad i = 1, 2, \ldots, a \text{ and } j = 1, 2, \ldots, b \tag{2.67}$$

$$y_{...} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} \; , \quad \bar{y}_{...} = \frac{y_{...}}{abn} \qquad i = 1, 2, \ldots, a;\; j = 1, 2, \ldots, b \tag{2.68}$$

$$\text{and } k = 1, 2, \ldots, n$$

The total variation in the observed data is described by the total sum of squares (SS_T). The two-factor analysis of variance devides the total sum of squares into four component parts; which are a sum of squares due to row treatment or factor *A* effect (SS_A); a sum of squares due to column treatment or factor *B* effect (SS_B); a sum of squares due to the interaction effect between factors *A* and *B*, (SS_{AB}); and a sum of squares due to error (SS_E). Thus, the total sum of squares is written symbolically as

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \tag{2.69}$$

Many statistical computer programs such as the Minitab program are usually employed to analyze the analysis of variance. However, the manual computing formulas for the sum of squares in Equation 2.69 are obtained easily. The total sum of squares is computed by

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y_{...}^2}{abn} \tag{2.70}$$

The sum of squares of the main effects (SS_A and SS_B) are

$$SS_A = \frac{1}{bn}\sum_{i=1}^{a} y_{i..}^2 - \frac{y_{...}^2}{abn} \tag{2.71}$$

and
$$SS_B = \frac{1}{an}\sum_{j=1}^{a} y_{.j.}^2 - \frac{y_{...}^2}{abn} \tag{2.72}$$

The sum of squares of the interaction effect (SS_{AB}) is convenient to obtain in two stages. First, the sum of squares between the *ab* cell, which is called the sum of squares due to *subtotals*

$$SS_{subtotals} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2 - \frac{y_{...}^2}{abn} \tag{2.73}$$

This sum of squares also contains SS_A and SS_B. Therefore, the second step is to compute SS_{AB}

$$SS_{AB} = SS_{subtotals} - SS_A - SS_B \tag{2.74}$$

Finally, the sum of squares due to error (SS_E) is obtained from Equations 2.69 and 2.74.

$$\begin{aligned} SS_E &= SS_T - SS_A - SS_B - SS_{AB} \\ &= SS_T - SS_{subtotals} \end{aligned} \tag{2.75}$$

The degrees of freedom associated with each sum of squares are shown as follows:

Effect	**Degree of freedom**
A	$a-1$
B	$b-1$
AB interaction	$(a-1)(b-1)$
Error	$ab(n-1)$
Total	$abn-1$

Each sum of squares divided by its degree of freedom is a mean square (*MS*). The expected value of mean squares is

$$E\left(MS_A\right) = E\left(\frac{SS_A}{a-1}\right) = \sigma^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1} \tag{2.76}$$

$$E\left(MS_B\right) = E\left(\frac{SS_B}{b-1}\right) = \sigma^2 + \frac{an\sum_{j=1}^{b}\beta_j^2}{b-1} \tag{2.77}$$

$$E\left(MS_{AB}\right) = E\left(\frac{SS_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\tau\beta)_{ij}^2}{(a-1)(b-1)} \tag{2.78}$$

and

$$E\left(MS_E\right) = E\left(\frac{SS_E}{ab\left(n-1\right)}\right) = \sigma^2 \tag{2.79}$$

If the null hypotheses, no row of treatment or factor *A* effect, no column of treatment or factor *B* effect, and no interaction effect, are true, the expected values of the MS_A, MS_B, MS_{AB} and MS_E are equal to σ^2. However, if there is the row of treatment or factor *A* effect, the MS_A will be larger than MS_E. Similarly, if there is the column of treatment or factor *B* effect, or the interaction effect, the corresponding mean squares will be larger than MS_E. Therefore, to test the significance of both main effects and their interaction, simply divide the corresponding mean squares by the error mean square. Large value of this ratio implies that the data do not correspond to the null hypothesis.

If the error term $\in_{ijk}$ in Equation 2.59 is assumed to be normally and independently distributed with the constant variance (σ^2), therefore, every ratio of mean squares (MS_A / MS_E, MS_B / MS_E, and MS_{AB} / MS_E) is distributed as *F* distribution with the numerator degrees of freedom of $a-1$, $b-1$, and $\left(a-1\right)\left(b-1\right)$, respectively, and the denominator degree of freedom of $ab\left(n-1\right)$. The test procedure is usually summarized in the analysis of variance table, as shown in Table 2.10.

Table 2.10 The analysis of variance table for the two-factor factorial design.

Source of variation	**Sum of squares (*SS*)**	**Degree of freedom**	**Mean square (*MS*)**	**Test statistic (F_0)**
Factor *A*	SS_A	$a-1$	$MS_A = \dfrac{SS_A}{a-1}$	$F_0 = \dfrac{MS_A}{MS_E}$
Factor *B*	SS_B	$b-1$	$MS_B = \dfrac{SS_B}{b-1}$	$F_0 = \dfrac{MS_B}{MS_E}$
AB interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \dfrac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \dfrac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \dfrac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

To make a decision which effect, factor *A*, factor *B* or *AB* interaction, plays a significant role on the response variable, the test statistic (F_0) of each factor is calculated. For the hypothesis testing on the effect of factor A , the null hypothesis is rejected if

$$f_0 \;>\; f_{\alpha,\ a-1,\ ab(n-1)} \tag{2.80}$$

where f_0 is the F_0 value calculated from $\dfrac{MS_A}{MS_E}$, shown in Table 2.10. And $f_{\alpha,\ a-1,\ ab(n-1)}$ is the rejection criterion of test statistic obtained from F distribution table in Appendix A-2 at α significance level or (1-α) confident level of the test, the numerator or the degree of freedom of $a-1$, and the denominator or error degree of freedom $ab(n-1)$.

Similarly, for the hypothesis testing on the effects of factor *B* and *AB* interaction, the null hypotheses are rejected, respectively, if

$$f_0 \;>\; f_{\alpha,\ b-1,\ ab(n-1)} \tag{2.81}$$

and

$$f_0 \;>\; f_{\alpha,\ (a-1)(b-1),\ ab(n-1)} \tag{2.82}$$

2.7 Measurement system analysis (MSA) [9, 12-13]

In many engineering studies, the performance or reliability of the gage or device for a measurement on a system of interest is important. For the problem involving with measurements, some of the observed variations arise from the experimental units, samples that are measured or the measurement error. Two types of errors associated with a gage or measurement device are accuracy and precision. Accuracy refers to the ability to measure the true value of the measured characteristic correctly, and precision reflects the inherent variations in the measurements.

The measurement system analysis is a technique to assure valid data of the measurement values, process control decisions and product acceptance. MSA applies the experimental and statistical methods to determine how much the variation within the measurement process contributes to overall process variation. MSA is essential to ensure the reliability in terms of the accuracy and the precision of the measurement system, i.e., the measurement devices, inspectors or operators and measuring procedures. The accuracy of the measurement system is verified by the resolution, bias, linearity and stability study. Determining accuracy often requires the use of a standard, for which the true value of the measured characteristic is known. The precision of the measurement system is checked by the repeatability and reproducibility study so-called Gage R&R study. The repeatability and reproducibility of the measurement system can be confirmed by using the graphical outputs plotted by the Minitab program, i.e.,

- repeatability: the R-chart by operators
- reproducibility: the X-bar chart by operators, the main effect chart by samples, the main effect chart by operators, and the operator and sample interaction chart

Therefore, the six components in MSA are investigated consecutively, as follows:

1. Resolution: How fine the measurements can be detected?

The resolution or discrimination of a measurement device defines the smallest change or effect that the measurement system can detect. The increments in the measurement system should be one-tenth of the product specifications or process variation.

2. Bias: How much the measurements are shifted from true value?

Bias is a difference between the average of the measurements and the reference value.

3. Linearity: Is the measurement true and/or consistent across the measurement range of the gage?

Linearity is the change in bias over a normal measurement range of the gage.

4. Stability: Does the measurement system remain constant and predictable over time?

Stability is the change in bias over time. It is often tracked on a control chart and corrected with a periodic calibration.

5. Repeatability: How well does an operator get the same result by using one gage?

Repeatability is the variations in the measurements which are derived from one gage or measurement device when one operator makes several measurements on one characteristic of the same sample. The repeatability is the variation caused by the measurement device, may be called the gage repeatability.

6. Reproducibility: How well do the operators get the same result?

Reproducibility is the variation in the measurements made by different operators using the same the measurement device when measuring one characteristic on the same sample. In other words, reproducibility is operator to operator variation.

Gage repeatability and reproducibility (Gage R&R) study

The total process variations are derived from the two components of variations, which are the sum of the process output variation and the variation due to the precision of the measurement system. The precision of the measurement system can be explained by the repeatability and reproducibility. The simple model that is used to explain the measurement variations usually describes the components of variations in terms of variance (σ^2). Therefore, the total process variations may be written symbolically as:

$$\sigma^2_{Total} = \sigma^2_{Process\ outputs} + \sigma^2_{Measurement\ system} \quad (2.83)$$

or

$$\sigma^2_{Total} = \sigma^2_{Process\ outputs} + \sigma^2_{Repeatability} + \sigma^2_{Reproducibility} \quad (2.84)$$

Gage R&R study allows one to estimate the contribution of the variation due to the measurement system to the total process variations. In Gage R&R study, graphical and key numeric outputs are examined. In the Minitab program, six graphical outputs of GR&R study are plotted as shown in Figure 2.26.

1. The components of variation chart

This chart is a graphical representation of the measurement variation of 4 different components or sources of variation,

- Gage R&R
- Repeatability
- Reproducibility
- Part-to-Part

The Part-to-Part variation represents the variation that is due to the different parts or samples. A measurement system of the highest variation components in the Part-to-Part, which describes the variation of different parts or samples, is noted as a good measurement system.

2. R chart by operators

This chart displays the repeatability of each operator. Each plot of the R chart by operators represents the sample range (R) which is the difference between the maximum and minimum measurements taken by each operator on the same sample. If an operator measures consistently, the sample ranges will be small and the points should stay within the control limits (UCL and LCL).

3. X-bar chart by operators

Each point of the X-bar chart by operators represents the average value of the two measurements taken by each operator on the same sample. In the X-bar chart by operators, the interval between the control limits represents the gage or measurement system variation, while the sample range represents the process output variation. The measurement system variation should less than the process output variation to ensure that the measurement system can distinguish the process outputs. Most of the points, 50% or more, are expected to locate outside the control limits.

4. Main effect chart by samples

This chart composes of each single measurement spread or distributions, and the average value by sample. The individual measurement shows the measurement values by every operator on the same sample. The main effect chart by samples with each narrow spread along the different samples shows that there is no effect of different samples on the operators to the measurement system.

5. Main effect chart by operators

This chart plots all measurements taken by each operator. The line connects 4 average values of the measurements by each operator. The horizontal line indicates that the operators have a good reproducibility. In other words, this chart shows that which operator causes the deviation of the measurement system.

6. The operator and sample interaction chart

This chart shows the averages of the measurements taken by each operator on the same sample. Each line connects the averages of the measurements taken by a single operator. Ideally, all of the lines should have the same pattern. In case the patterns are different, it means that the operators are experiencing the difficulties to make the consistent measurements. This chart is used to show if the different samples can affect the measurements taken by each operator or not.

The key numeric outputs are the precision to tolerance (P/T), the precision to total variation (P/TV), and the number of distinct categories.

1. Precision to tolerance (P/T) is the percentage ratio of the variation due to the measurement system in terms of the standard deviation and the range of specification limits so-called the tolerance. The expression is shown by Equation 2.85.

$$\mathrm{P/T} = \frac{6 \times \sigma_{MS}}{\mathrm{Tolerance}} \times 100 \tag{2.85}$$

The P/T values can explain the proportion that the tolerance is consumed by the variation due to the measurement system. The P/T values indicate the ability of the measurement system to discriminate the conformance and nonconformance products precisely at the given specifications. For an acceptable measurement, the P/TV is less than 30%. In the best case, the P/TV is less than 10%.

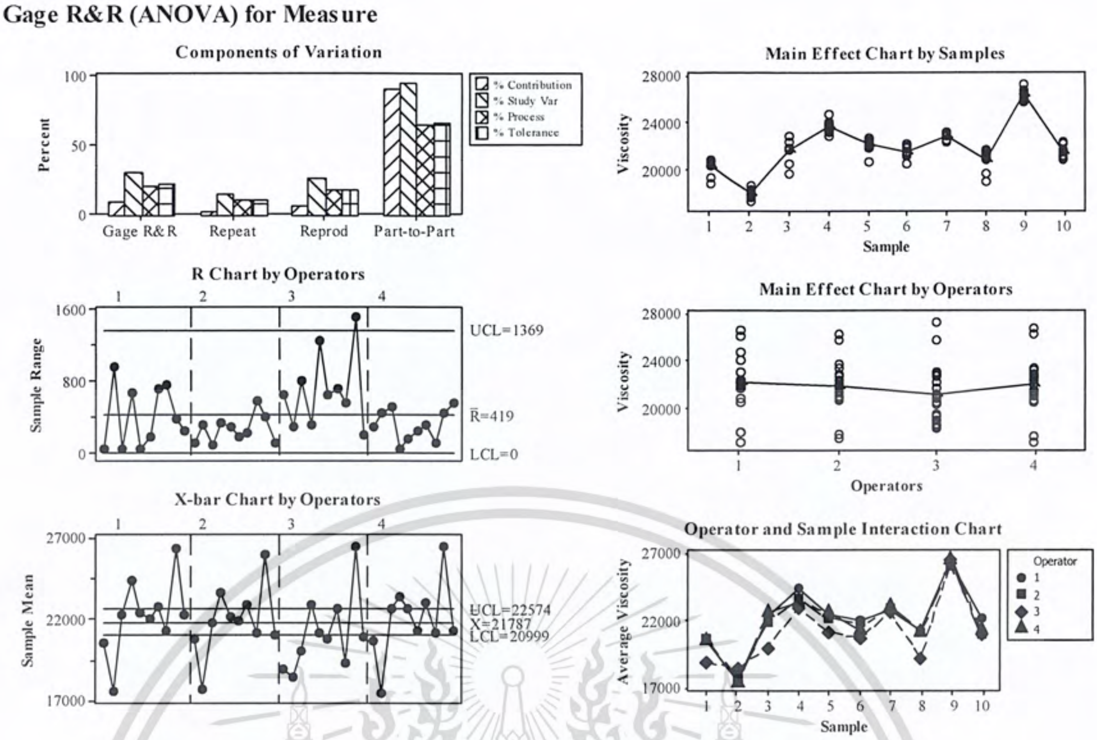


Figure 2.26 Six graphical outputs from GR&R study by the Minitab program.

2. Precision to total variation (P/TV) is the percentage ratio of the variation due to the measurement system to the total process variations, as shown by Equation 2.86.

$$P/TV = \frac{\sigma_{MS}}{\sigma_{Total}} \times 100 \qquad (2.86)$$

The small P/TV value indicates that the variation due to the measurement system is small relative to the total process variations; the measurement system can clearly observe the process output variations; and the profile of the process based on the obtained data can be trust. For an acceptable measurement, the P/TV is less than 30%. In the best case, the P/TV is less than 10%.

3. Number of distinct categories (NDC) is the number of categories within the process output data that are differentiated by the measurement system. The number of distinct categories provides an estimation of numbers of divisions that the measurement system can measure across

the process output variation. The number of distinct categories can be calculated by the following equation:

$$\text{Number of distinct categories} = \left(\frac{\sigma_{Process\ \ outputs}}{\sigma_{MS}} \times \sqrt{2} \right) \tag{2.87}$$

Theoretically, from 10-different sample measurement, the measurement system which can make differentiation of 3-4 numbers of the distinct categories is acceptable. In the best case, the number of distinct categories is more than 4.

2.8 Literature reviews

The correspondence articles are as follows:

Rezayati-Charani et al. [14] applied multiple linear regression analysis by the Minitab program to obtain a correlation between three pulping variables and chemical properties to set up the optimum operating conditions in the pulping industry.

Suwittayaruk [15] applied 5 interconnected-phase of six sigma to minimize seawax chemical PF5060 usage in the Head Stack Assembly (HAS) line manufacturing of the enterprise product. The optimum conditions in the manufacturing process were obtained by the design of experiments by 2-level factorial design using 2 replicates and 5 midpoints or center points.

Teerakitwattana [16] used the Failure Mode and Effect Analysis (FMEA) and the design of experiments by 2-level factorial design to find the substantial causes on weight deviation of ice cream. As a result, the factors that affect weight deviation of ice cream were amount of air in ice cream and the adjustment of filler valves. By using the optimum filling conditions, finally the weight deviation of ice cream was reduced.

Chunkoh [17] improved a mixing process to reduce the deviation of shampoo viscosity by quality management and the statistical techniques. The Failure Mode and Effect Analysis and the design of experiments by 2-level factorial design were used to find the influential causes on the deviation of shampoo viscosity and their optimum values.

Temsawangleart [18-19] applied the multiple linear regression by the Minitab program to specify the specific factors that influenced the deviation of bulk density of non-soap detergent. The important factors were the pressure in the injection tower and the pressure of high pressure pump. By adjusting the corresponding optimum conditions, the deviation of bulk density of the non-soap detergent was reduced.

Chinsutthi [20-21] reduced the rework in a non-soap detergent process by applying the design of experiments using full factorial design, and the statistical process control. The significant parameters on rework and the optimum conditions were observed. After applying the optimum conditions, the rework was reduced.

Thumsavate [22-24] used the design of experiments by 2-level full factorial design in the Minitab program to analyze the significant parameters on the surface flexography printing system on the polypropylene tape to set up a new production line. The significant parameters and their optimum conditions were analyzed and recorded as the standard operating procedures (SOPs).

Tuladhorn [25-27] reduced weight deviation of the consumer liquid products, i.e., dish-washing liquid and liquid soap in the packaging processes using six sigma and the Minitab program. The significant factors that influenced the weight deviation of the case study products and their optimum conditions were determined. By using the obtained process optimum conditions, the standard deviation of product weight, % giveaway and the defect per million opportunities (DPMO) of the dish-washing liquid and liquid soap were reduced.

Im-Orb and Phaiboonsilpa [28] increased the production capacities in a packing line of the personal care products by decreasing non-productive time using the total productive maintenance (TPM). The Pareto diagram, fishbone diagram and why-why analysis were used to figure out the root causes. The root causes were solved by the PDCA cycle, and the standard operating procedures were written.

Chapter 3

Research Methodology

This work aims to minimize the deviation of the viscosity of 5 formulas of hair conditioner during raw material mixing process. From a number of parameters in raw material mixing process, the possible parameters on the deviation of hair conditioner viscosity are detected by the cause and effect diagram and brainstorming. The significant parameters are pinpointed by multiple linear regression analysis. Before taking multiple linear regression analysis, the measurement system analysis was employed to assure the accuracy and precision of the measurement devices, operators and measuring procedures. Finally, the design of experiments is used to determine the strongly significant parameters from the significant parameters to pinpoint the optimum operating conditions. The defects per million opportunities (DPMO) and the potential process capability (C_p) are evaluated to assure the in-control process, resulting in product quality within the specification limits.

3.1 Measurement of hair conditioner viscosity

The apparent viscosities of the 5 formulas of hair conditioner in centipoises (cP) were investigated by using the Brookfield viscometer, RVT model with 6-digit display. The shear rates applied to the hair conditioner were varied by adjusting the T-bar spindle speeds of 0.3, 0.5, 0.6, 1, 1.5, 3, 4, 5, 6, 10, 12, 20, 30, 50, 60 and 100 rpm. The results were analyzed by the power-law model for the power-law index to identify types of fluid. In addition, the effect of time was also studied by applying a constant shear rate on the hair conditioner. The hair conditioner in a 1,000 ml-beaker was stirred by a 5 cm-diameter turbine at a constant speed of 20 rpm. The viscosity was measured at 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 and 60 minutes.

3.2 Determination of the possible and significant parameters on the deviation of hair conditioner viscosity

3.2.1 Possible parameters by the cause and effect diagram and brainstorming

From a number of parameters in hair conditioner raw material mixing process, the possible parameters on the deviation of hair conditioner viscosity were detected by using the cause and effect diagram and brainstorming. The possible parameters in raw material mixing process on the deviation of hair conditioner viscosity were listed.

From many possible parameters, the cause and effect matrix was used. The significance of each possible parameter on the deviation of hair conditioner viscosity was ranked by giving the significance scores. Normally, the score either 1-5 or 1-10 was used. In this work, the range of 1-10 was employed. The parameter with the score of 10 was decided the most significant. For a number of the possible parameters, about 20, the multiple linear regression analysis was employed to pinpoint the significant parameters.

3.2.2 Measurement system analysis

Before taking the multiple linear regression analysis to determine the significant parameters on the deviation of hair conditioner viscosity, the measurement system analysis was employed to assure the reliability (accuracy and precision) of the measurement devices, operators and measuring procedures. The accuracy of the measurement system was verified by the resolution, bias, linearity and stability study. The precision was checked by the repeatability and reproducibility study so-called Gage R&R study using the Minitab program. In Gage R&R study, 10 samples from 5 formulas of hair conditioner of which different viscosities were measured based on the same procedure by four operators using the same viscometer with two replicates. Accordingly, six graphical outputs of Gage R&R study were plotted [13]. The variation of the measurements due to the measurement system comparing with the variation due to the process outputs or samples was examined by the components of variation chart. The repeatability of the measurement system was checked by the R-chart by operators, and the reproducibility of the measurement system was verified by the X-bar chart by operators, the main effect chart by samples, the main effect chart by operators and the operator and sample interaction chart. In addition, the numerical key outputs, i.e., P/T, P/TV and NDC, were calculated. In case a precise measurement system, the values of the P/T and P/TV are less than 30%, and the NDC should be not less than 3. The NDC near 10 is the most preferable. The obtained outputs were analyzed, and the appropriate corrective action(s) and improvement(s) were subsequently implemented.

3.2.3 Significant parameters by the multiple linear regression analysis

The multiple regression analysis in the Minitab program was used to pinpoint the significant parameters on the deviation of hair conditioner viscosity from the possible parameters obtained from the cause and effect diagram and brainstorming based on the variance inflation factors (VIFs) less than 10 and the significance level (α) or p-values of 0.05. The hair conditioner viscosities and the values of the possible parameters were collected and analyzed by the Minitab program. The VIFs and p-values of the possible parameters were calculated and reported. The possible parameters, of which the VIFs were higher than 10 and/or p-values were greater than 0.05, were not significant on the deviation of the viscosity and could be truncated from the multiple linear regression model. After the insignificant parameter was removed from the model one by one, the model was refitted and the new VIFs and p-values were calculated. The other insignificant parameters were truncated from the model in the same manner until the final model of the significant parameters was achieved. The adequacy of the collected data for the final model were examined based on the residual model diagnostics and the coefficient of determination (R^2).

3.3 Determination of the strongly significant parameters on the deviation of hair conditioner viscosity and the optimum operating conditions

From the significant parameters obtained, the design of experiments was finally applied to determine the strongly significant parameters and the optimum operating conditions. Full factorial design with 2 replicates was employed in this work. The factorial design experiments were designed by the Minitab program. The minimum and maximum values of the significant parameters were set in the DOEs by process historical data and the recommendations from the Division of Quality Assurance (QA) of the company. The strongly significant parameters were analyzed by the Pareto plot based on the significance level (α) or p-values of 0.05. The optimum operating conditions of the strongly significant parameters which make the hair conditioner viscosities attain the target viscosity of 26,000 cP were selected by the cube plot in the Minitab program. The potential process capability after the improvement(s) was calculated. The achieved optimum operating conditions were finally used and written in the SOPs.

Chapter 4

Results and Discussion

4.1 Hair conditioner viscosities

The apparent viscosities of 5 formulas of hair conditioner (SCAC, SCBS, SCDR, SCLS and SCSS) were plotted against the shear rate as shown in Figure 4.1. The viscosities decreased with an increase in shear rate, indicating that hair conditioner is the pseudoplastic fluid [3-4]. Figures 4.2 – 4.6 are the plots between log apparent viscosity and log shear rate of hair conditioners. The linear relationships were achieved. The consistencies of 5 formulas of hair conditioner were obtained by taking inverse log of the intercepts. The power-law indexes were calculated from the slopes.

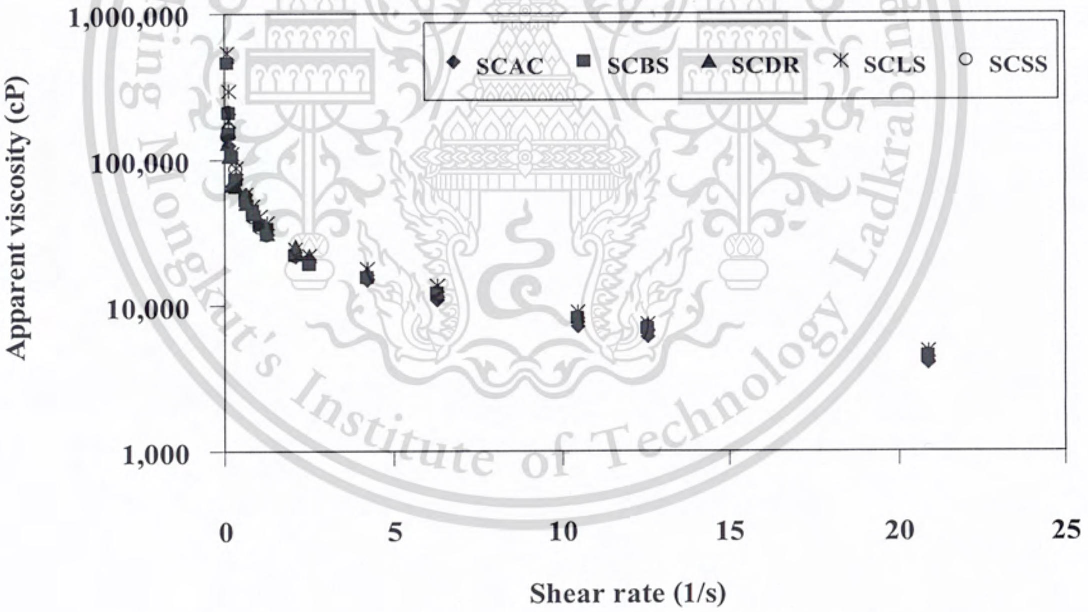


Figure 4.1 The apparent viscosities of 5 formulas of hair conditioner.

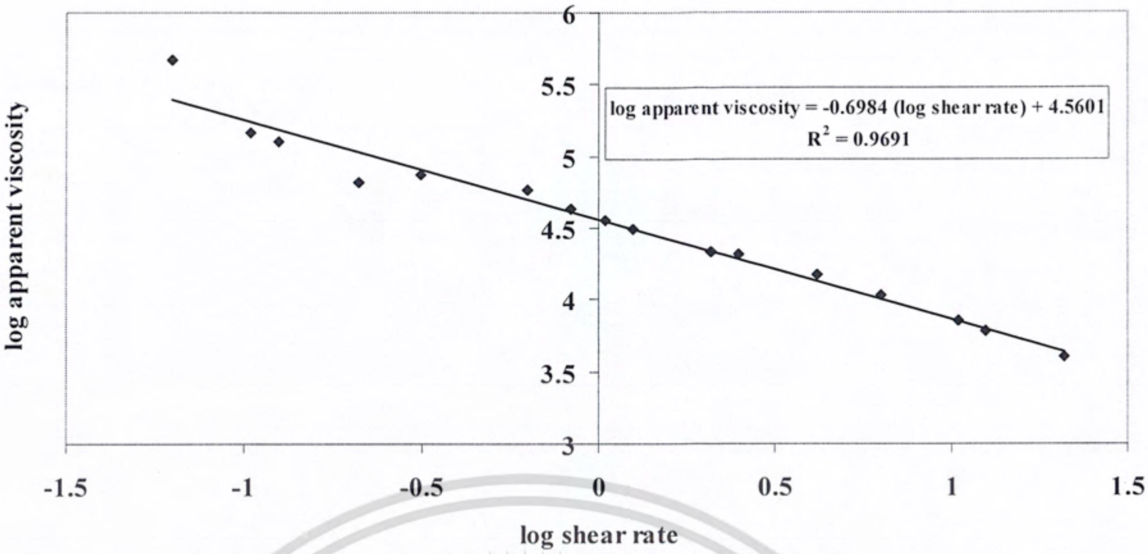


Figure 4.2 Log apparent viscosity and log shear rate of the SCAC hair conditioner.

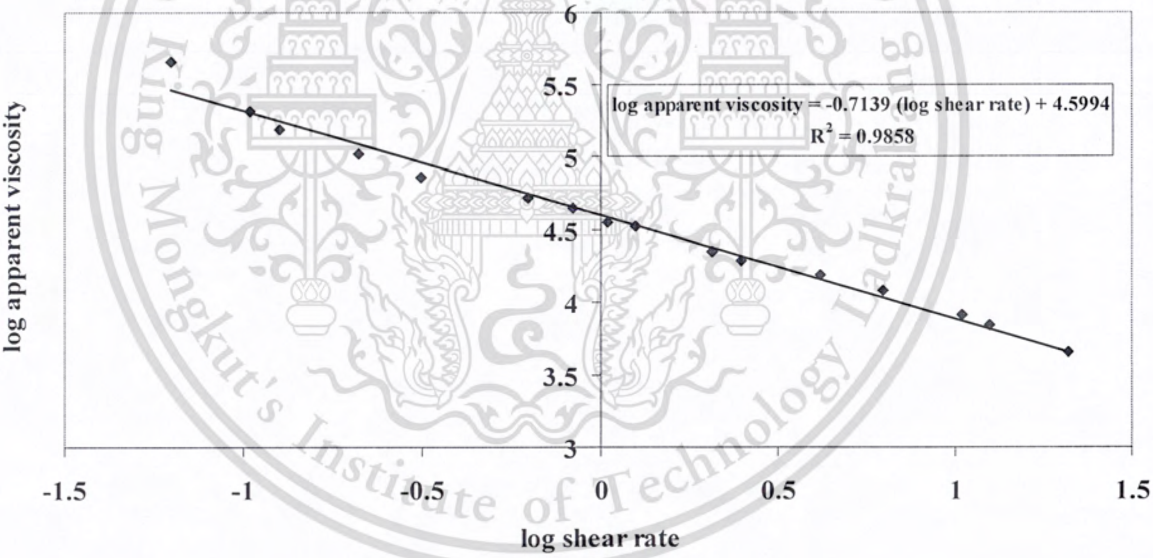


Figure 4.3 Log apparent viscosity and log shear rate of the SCBS hair conditioner.

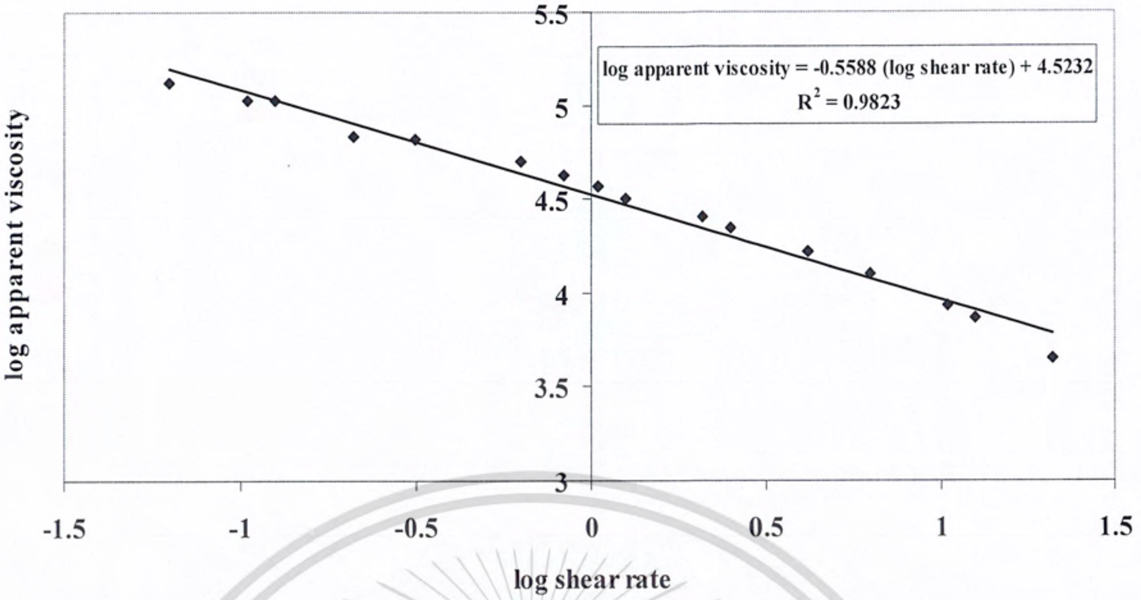


Figure 4.4 Log apparent viscosity and log shear rate of the SCDR hair conditioner.

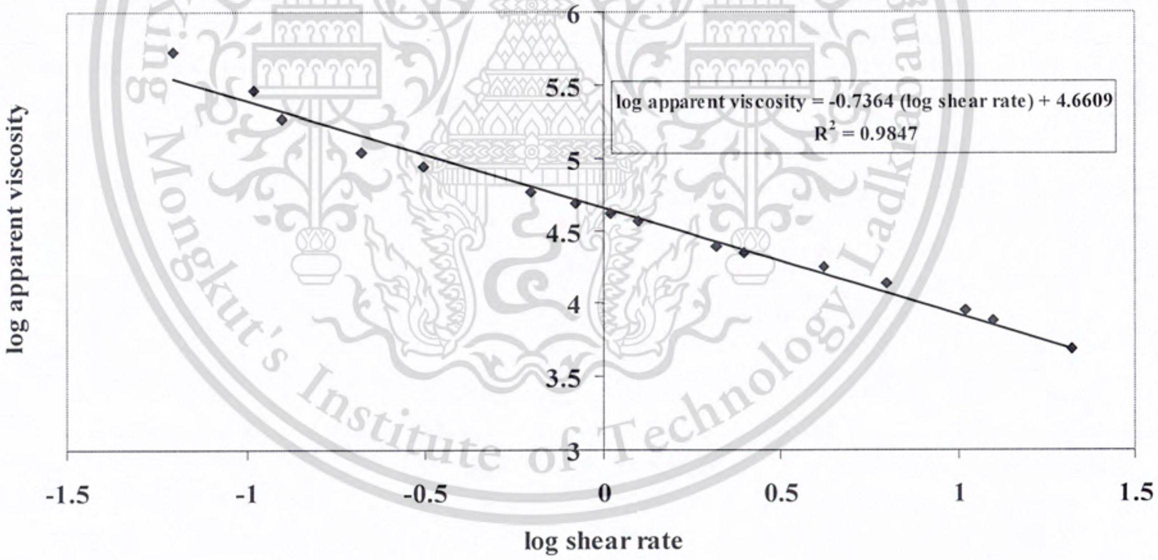


Figure 4.5 Log apparent viscosity and log shear rate of the SCLS hair conditioner.

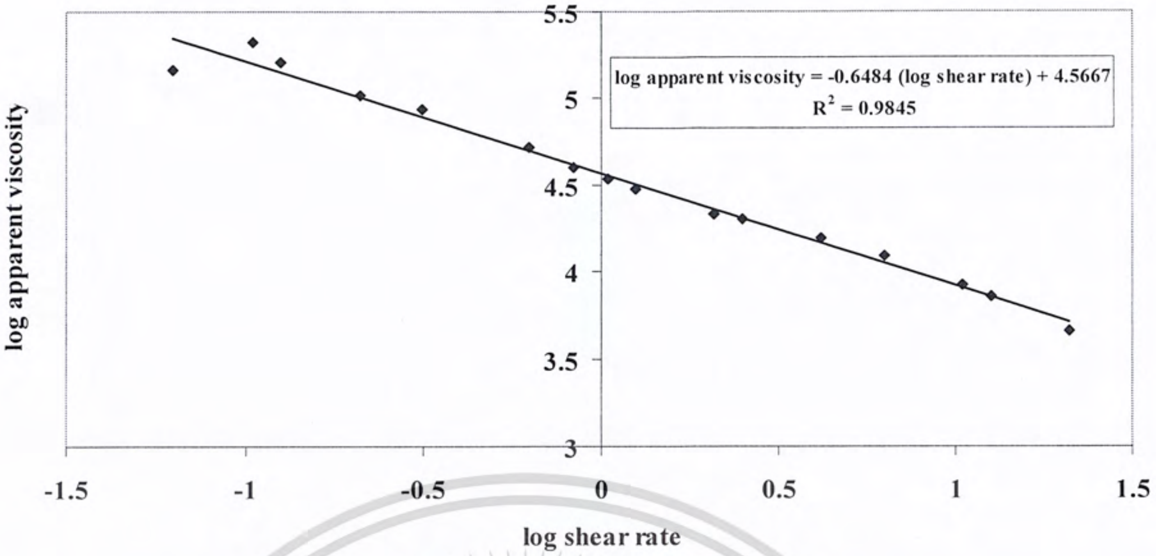


Figure 4.6 Log apparent viscosity and log shear rate of the SCSS hair conditioner.

Table 4.1 shows the consistencies (m) and the power-law indexes (n) of the hair conditioners. The average consistency was 38,421 cP·s^n and the average power-law index was 0.329. The viscosities of the hair conditioners decrease with time at a constant shear rate, as shown in Figure 4.7. Therefore, hair conditioners are time-sensitive to shearing action [4]. In particular, the parameters related to shear rate and time should be considered when the deviation of viscosity is observed.

Table 4.1 The consistencies (m) and the power-law indexes (n) of the hair conditioners.

Hair conditioner	**m (cP s^n)**	**n**
SCAC	36,316	0.302
SCBS	39,756	0.286
SCDR	33,358	0.441
SCLS	45,804	0.264
SCSS	36,872	0.352
Average	**38,421**	**0.329**

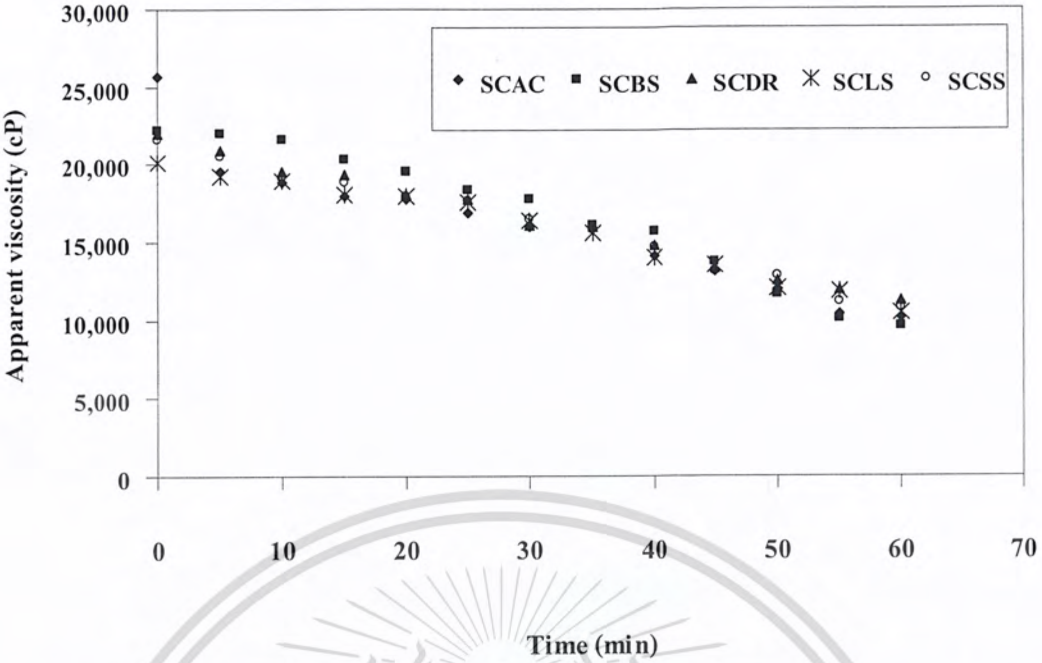
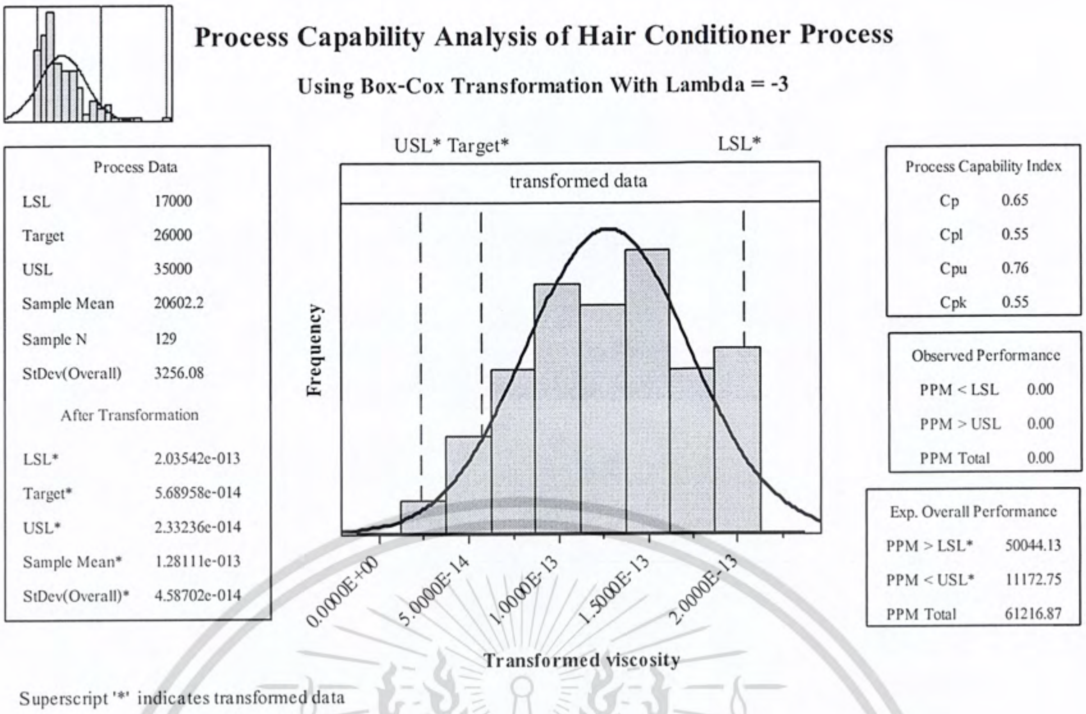


Figure 4.7 The apparent viscosity of 5 formulas of hair conditioner at a constant shear rate of 10 rpm.

4.2 Process study

From the collected data in raw material mixing process of hair conditioner, the deviation of the viscosity of 5 formulas substantially deviated from the target value, mostly the viscosities stayed in the lower specification limit (LSL). The standard deviation (σ) of the product viscosity from 129 numbers of the product viscosities, which was calculated by the Minitab program, was 3,256 cP corresponding to the defect per million opportunities (DPMO) of the process of 61,217 and the potential process capability index (C_p) of 0.65, as shown in Figure 4.8. This means that the number of possible defects occurred in the process based on 1,000,000 total opportunities is 61,217. The DPMO and C_p are commonly used as the indicators to show the performance level of a process. If the process is in control, the DPMO is less than 6,210 and the C_p is higher than 1.33 [6]. The process is recommended for the immediate corrective actions if the DPMO is higher than 6,210 and the C_p is less than 1.33. Therefore, the raw material mixing process of hair conditioner in this case study needs immediate corrective actions. The schematic diagram of raw material mixing process and the simple processing steps of hair conditioners are illustrated in Figures 4.9 and 4.10.



Process Data	
LSL	17000
Target	26000
USL	35000
Sample Mean	20602.2
Sample N	129
StDev(Overall)	3256.08
After Transformation	
LSL*	2.03542e-013
Target*	5.68958e-014
USL*	2.33236e-014
Sample Mean*	1.28111e-013
StDev(Overall)*	4.58702e-014

Process Capability Index	
Cp	0.65
Cpl	0.55
Cpu	0.76
Cpk	0.55

Observed Performance	
PPM < LSL	0.00
PPM > USL	0.00
PPM Total	0.00

Exp. Overall Performance	
PPM > LSL*	50044.13
PPM < USL*	11172.75
PPM Total	61216.87

Figure 4.8 Histogram of hair conditioner viscosities in this case study.

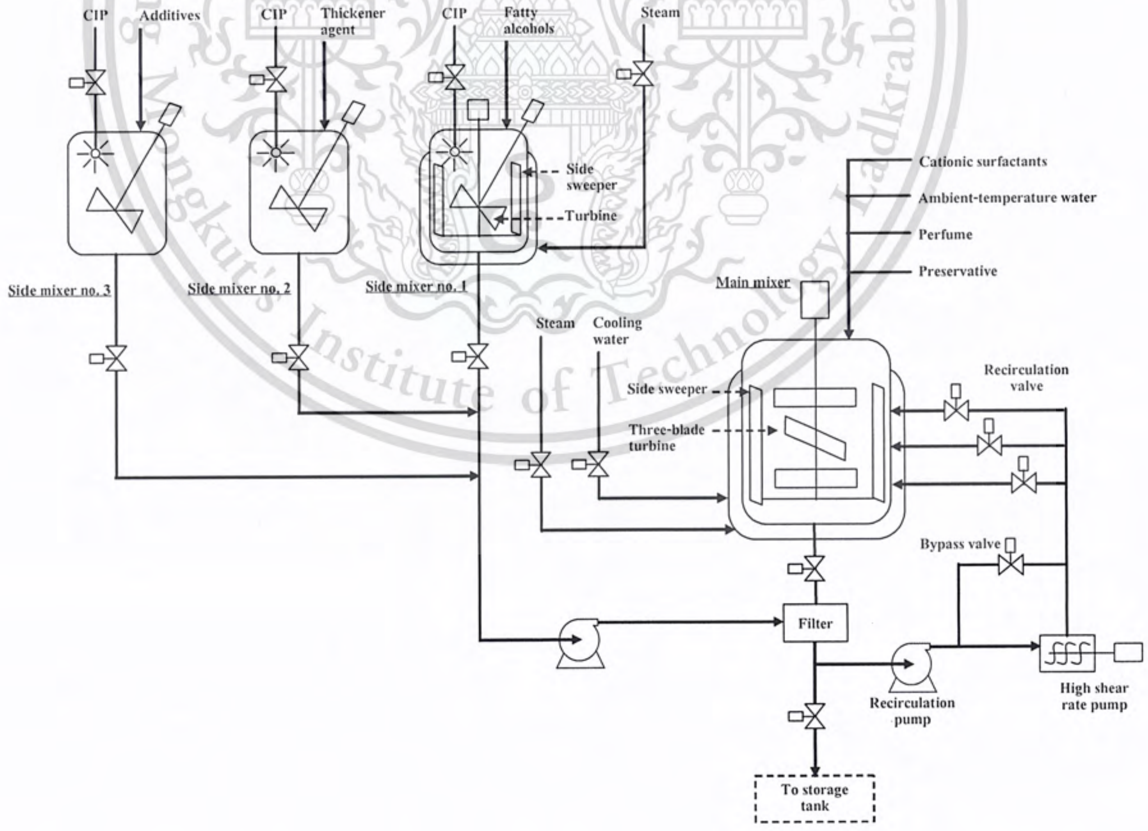


Figure 4.9 Schematic diagram of raw material mixing process of 5 formulas of hair conditioner.

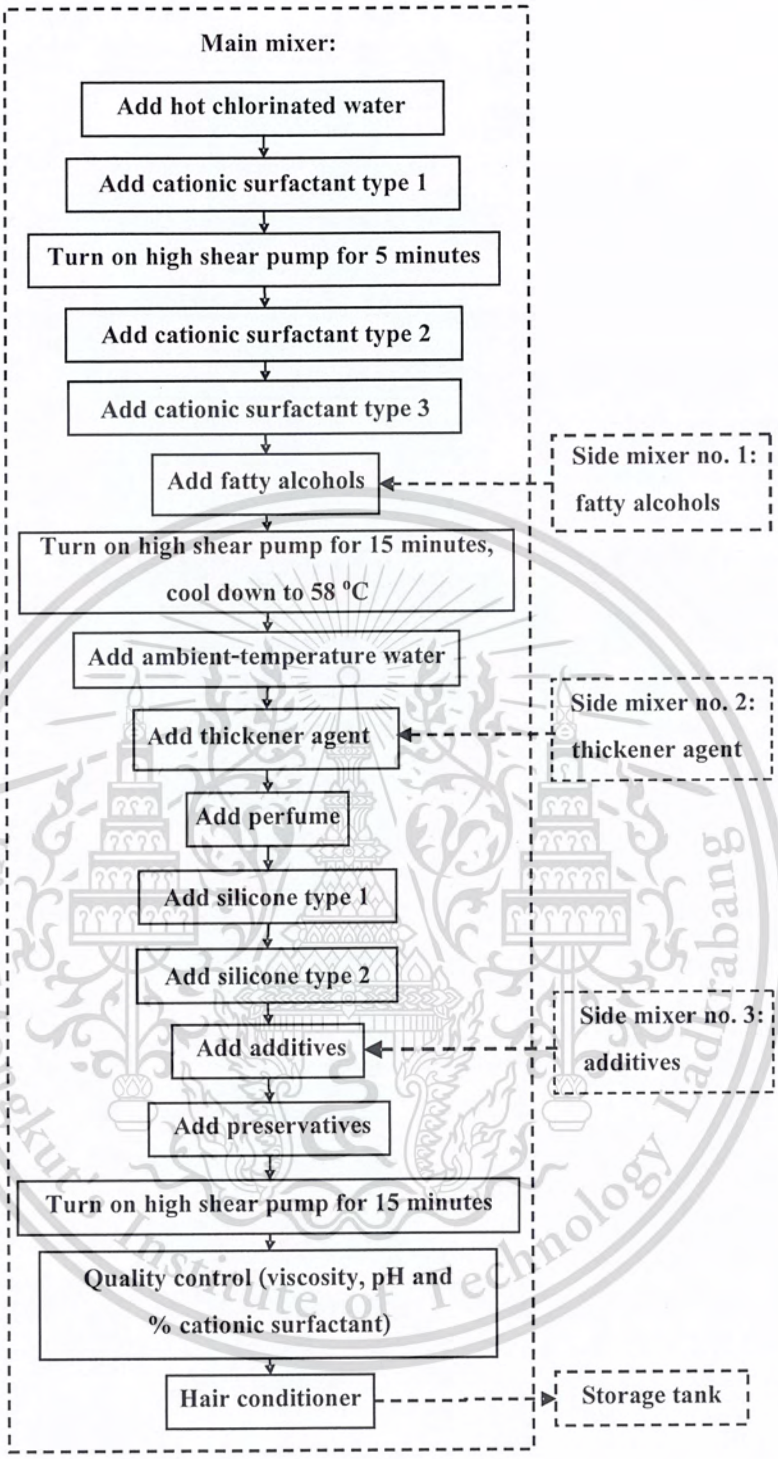


Figure 4.10 Simple processing steps of hair conditioner production in this case study.

4.3 Possible and significant parameters on the deviation of hair conditioner viscosity

4.3.1 Possible parameters by the cause and effect diagram and brainstorming

The possible causes in the cause and effect diagram on the deviation of hair conditioner viscosity were decided by brainstorming with a working group (seven staff from hair conditioner production line, the Division of Quality Assurance (QA) and the Division of Research and Development (R&D)). The results are shown in Figure 4.11.

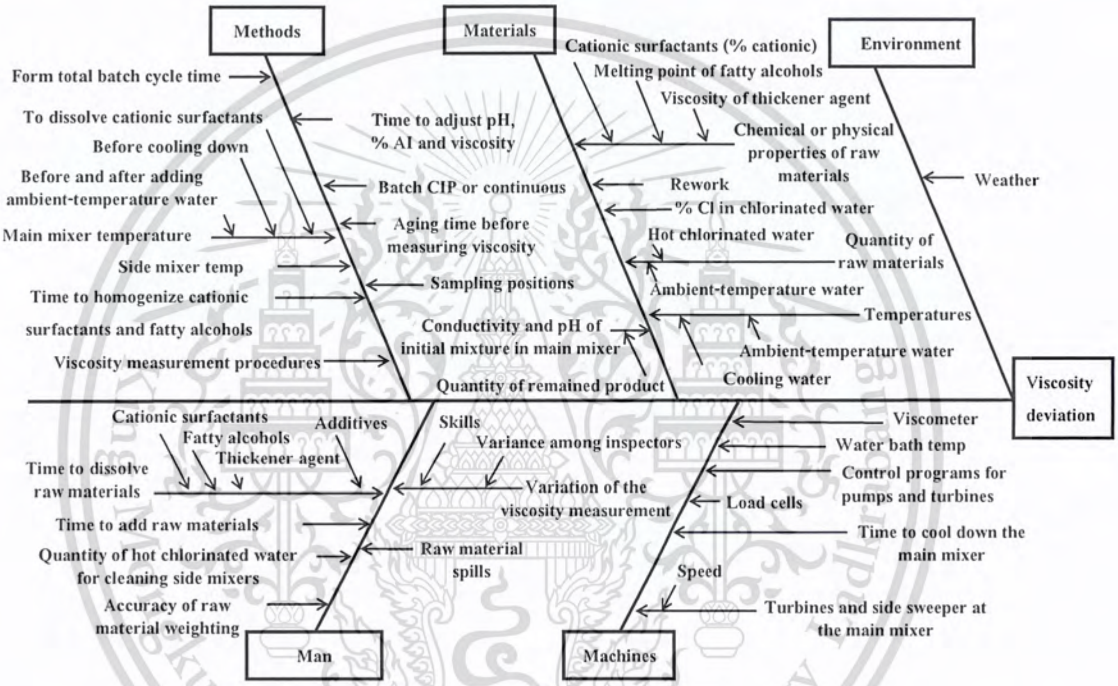


Figure 4.11 Cause and effect diagram of the possible causes on the deviation of hair conditioner viscosity.

One-hundred-and-five possible parameters in the raw material mixing process were ranked in a so-called cause and effect matrix, as shown in Table 4.2. From the cause and effect matrix, the significance score of each possible parameter was evaluated by brainstorming based on a 1-10 points, where 10 indicating the most significant possible parameter on the deviation of the viscosity. Twenty possible parameters of which the significance scores were 10 and 9 were selected and summarized in Table 4.3.

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity.

No	Possible parameter on the deviation of hair conditioner viscosity	Significance score evaluated by brainstorming (based on a 1 -10 points)	Rank of significance
1	Batch CIP or continuous	10	1
2	Quantity of remained product in the main mixer	10	1
3	% Chlorine of hot chlorinated water	0	80
4	Conductivity of hot chlorinated water	1	76
5	pH of hot chlorinated water	2	64
6	Temperature of hot chlorinated water added in the main mixer	0	80
7	Quantity of hot chlorinated water added in the main mixer	10	1
8	Conductivity of the mixtures (hot chlorinated water and remained product) in the main mixer	9	16
9	pH of the mixtures (hot chlorinated water and remained product) in the main mixer	0	80
10	% Cationic surfactant of the mixtures (hot chlorinated water and remained product) in the main mixer	0	80
11	Input steam flow rate at the main mixer	0	80
12	Input steam temperature at the main mixer	0	80
13	Output steam temperature at the main mixer	0	80
14	Alignment of a side sweeper at the main mixer	8	21
15	Turbine speed at the main mixer	8	21
16	Side sweeper speed at the main mixer	8	21
17	Main mixer temperature to dissolve cationic surfactant type 1	10	1
18	Quantity of cationic surfactant type 1 added in the main mixer	8	21

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	Possible parameter on the deviation of hair conditioner viscosity	Significance score evaluated by brainstorming (based on a 1 -10 points)	Rank of significance
19	% Cationic contents of cationic surfactant type 1	9	16
20	Time to add cationic surfactant type 1 in the main mixer	0	80
21	Time to dissolve cationic surfactant type 1 at the main mixer	10	1
22	High shear rate pump speed to homogenize cationic surfactant type 1 at the main mixer	5	48
23	Time for homogenize mixing by high shear rate pump at the main mixer	0	80
24	Quantity of cationic surfactant type 2 added in the main mixer	8	21
25	% Cationic contents of cationic surfactant type 2	7	32
26	Time to add cationic surfactant type 2 to the main mixer	0	80
27	Quantity of cationic surfactant type 3 added in the main mixer	2	64
28	% Cationic contents of cationic surfactant type 3	1	76
29	Time to add cationic surfactant type 3 to the main mixer	0	80
30	Temperature at side mixer no. 1 to dissolve fatty alcohols	10	1
31	Quantity of hot chlorinated water added in side mixer no. 1	5	48
32	Input steam flow rate at side mixer no. 1	3	61
33	Input steam temperature at side mixer no. 1	4	58

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	Possible parameter on the deviation of hair conditioner viscosity	Significance score evaluated by brainstorming (based on a 1 -10 points)	Rank of significance
34	Output steam temperature at side mixer no. 1	4	58
35	Quantity of fatty alcohol type 1 added in side mixer no. 1	8	21
36	Melting point of fatty alcohol type 1	7	32
37	Suppliers of fatty alcohol type 1	6	40
38	Time to add fatty alcohol type 1 to side mixer no. 1	0	80
39	Quantity of preservative type 1 added in side mixer no. 1	2	64
40	Suppliers of preservative type 1	1	76
41	Time to homogenize fatty alcohols at side mixer no. 1	10	1
42	Time to transfer fatty alcohols from side mixer no. 1 to the main mixer	0	80
43	Quantity of hot chlorinated water for cleaning side mixer no. 1	5	48
44	High shear rate pump speed to homogenize cationic surfactants and fatty alcohols at the main mixer	8	21
45	Time to homogenize cationic surfactants and fatty alcohols at the main mixer	9	16
46	Main mixer temperature before cooling down	10	1
47	Temperature of cooling water	10	1
48	Flow rate of cooling water	6	40
49	Time to cool down the main mixer to 58 ºC before adding the ambient-temperature water	10	1

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	**Possible parameter on the deviation of hair conditioner viscosity**	**Significance score evaluated by brainstorming (based on a 1 -10 points)**	**Rank of significance**
50	Main mixer temperature before adding the ambient-temperature water	10	1
51	Quantity of ambient-temperature water added in the main mixer	10	1
52	Temperature of ambient-temperature water	8	21
53	% Chlorine of ambient-temperature water	2	64
54	Conductivity of ambient-temperature water	5	48
55	pH of ambient-temperature water	3	61
56	Main mixer temperature after adding the ambient-temperature water	10	1
57	Time to cool down the main mixer to 48 °C after adding the ambient-temperature water	9	16
58	Quantity of chlorinated water added in side mixer no. 2	7	32
59	Quantity of acid added in side mixer no. 2	6	40
60	Quantity of thickener agent added in side mixer no. 2	8	21
61	Viscosity of thickener agent	10	1
62	Suppliers of thickener agent	8	21
63	Time to add thickener agent in side mixer no. 2	0	80
64	Turbine speed at side mixer no. 2	6	40
65	Time to dissolve thickener agent at side mixer no. 2	9	16
66	Time to add thickener agent in the main mixer	2	64

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	Possible parameter on the deviation of hair conditioner viscosity	Significance score evaluated by brainstorming (based on a 1 -10 points)	Rank of significance
67	Quantity of chlorinated water for cleaning side mixer no. 2	2	64
68	Main mixer temperature after adding thickener agent	5	48
69	Time to cool down the main mixer to 45 °C	7	32
70	Main mixer temperature before adding perfumes	5	48
71	Quantity of perfume	5	48
72	Types of perfumes	7	32
73	Suppliers of perfume	0	80
74	Time to add perfume in the main mixer	0	80
75	Quantity of silicone type 1 added in the main mixer	7	32
76	Viscosity of silicone type 1	6	40
77	Time to add silicone type 1 in the main mixer	0	80
78	Quantity of chlorinated water for cleaning the type 1-silicone tank	2	64
79	Time to homogenize silicone type 1 at the main mixer	5	48
80	Quantity of silicone type 2 added in the main mixer	8	21
81	Viscosity of silicone type 2	7	32
82	Time to add silicone type 2 in the main mixer	0	80
83	Quantity of chlorinated water for cleaning the type 2-silicone tank	2	64
84	Time to homogenize silicone type 2 at the main mixer	6	40
85	Quantity of chlorinated water added in side mixer no. 3	5	48

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	**Possible parameter on the deviation of hair conditioner viscosity**	**Significance score evaluated by brainstorming (based on a 1 -10 points)**	**Rank of significance**
86	Quantity of additive type 1 added in side mixer no. 3	3	61
87	Turbine speed of at side mixer no. 3	2	64
88	Time to homogenize additive type 1 at side mixer no. 3	4	58
89	Quantity of additive type 2 added in side mixer no. 3	2	64
90	Time to homogenize additive type 2 at side mixer no. 3	1	76
91	Time to transfer the additives to the main mixer	2	64
92	Quantity of chlorinated water for cleaning side mixer no. 3	2	64
93	Quantity of preservative type 2 added in the main mixer	0	80
94	High shear rate pump speed to homogenize the final mixture at the main mixer	6	40
95	Final homogenizing time	7	32
96	Waiting time for QC	6	40
97	Reliability of operators in viscosity measurement	5	48
98	Total batch cycle time	10	1
99	Accuracy of load cells in weighting raw materials	0	80
100	Quantity of raw material spills	0	80
101	Quantity of rework added in the main mixer	0	80
102	Sampling positions in the main mixer for the viscosity inspection	0	80

Table 4.2 Cause and effect matrix of possible parameters in raw material mixing process on the deviation of hair conditioner viscosity. (cont.)

No	Possible parameter on the deviation of hair conditioner viscosity	Significance score evaluated by brainstorming (based on a 1 -10 points)	Rank of significance
103	Water bath temperature of for viscosity measurement	0	80
104	Control program for pumps and turbines used to control speed of pumps and turbines	0	80
105	Weather	0	80

Table 4.3 The selected possible parameters in raw material mixing process on the deviation of hair conditioner viscosity.

Parameter (X_i)	**Definition**
X_1	Batch Cleaning in Process (CIP) or continuous
X_2	Quantity of remained product in the main mixer (kg)
X_3	Quantity of hot chlorinated water added in the main mixer (kg)
X_4	Conductivity of initial mixture in the main mixer
X_5	% Cationic contents of cationic surfactant type1
X_6	Main mixer temperature to dissolve cationic surfactant type1 (^{o}C)
X_7	Time to dissolve cationic surfactant type1 (min)
X_8	Temperature at side mixer no. 1 to dissolve fatty alcohols (^{o}C)
X_9	Time to homogenize fatty alcohols (min)
X_{10}	Time to homogenize cationic surfactants and fatty alcohols at the main mixer (min)
X_{11}	Main mixer temperature before cooling down (^{o}C)
X_{12}	Temperature of cooling water (^{o}C)
X_{13}	Time to cool down the main mixer to 58 ^{o}C before adding the ambient-temperature water (min)
X_{14}	Main mixer temperature before adding the ambient-temperature water (^{o}C)
X_{15}	Main mixer temperature after adding the ambient-temperature water (^{o}C)
X_{16}	Quantity of ambient-temperature water added in the main mixer (kg)
X_{17}	Time to cool down the main mixer to 48 ^{o}C after adding the ambient-temperature water
X_{18}	Time to dissolve thickener agent at side mixer no. 2 (min)
X_{19}	Viscosity of thickener agent (cP)
X_{20}	Total batch cycle time (min)

4.3.2 Measurement system analysis

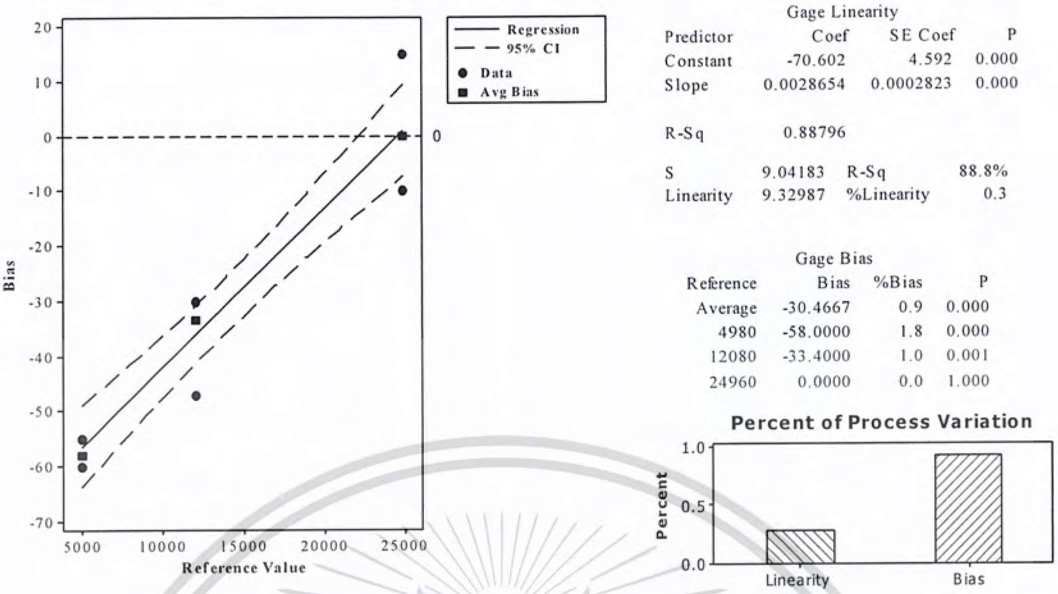
The measurement system analysis was employed to assure the reliability in terms of accuracy and precision of the inspectors and their skills on viscosity measurement, the viscometer and the measuring procedures before taking the multiple regression analysis.

The accuracy of the viscometer is verified by the resolution, bias, linearity and stability. In common, the hair conditioner viscosities were within a range of 10,000–40,000 cP, which can be distinguished clearly by the 6-digit display Brookfield viscometer. Therefore, the resolution of the viscometer is sufficient.

From Table 4.4, the viscosities of three reference fluids, of which the viscosities were known, were measured by the same viscometer with five replicates to analyze the bias and linearity of the viscometer. Figure 4.12 shows the plot of bias and the average value of bias versus the three reference fluid viscosities.

Table 4.4 The viscosities of the reference fluids with known viscosity for bias and linearity analyses.

Viscosity of the reference fluids (cP)	**Measured viscosity (cP)**					**Average**
	1	**2**	**3**	**4**	**5**	
4,980	4,920	4,920	4,925	4,925	4,920	4,922
12,080	12,033	12,050	12,050	12,050	12,050	12,047
24,960	24,950	24,975	24,950	24,975	24,950	24,960



```
              Gage Linearity
Predictor        Coef      SE Coef      P
Constant      -70.602        4.592  0.000
Slope       0.0028654    0.0002823  0.000

R-Sq          0.88796

S            9.04183   R-Sq          88.8%
Linearity    9.32987   %Linearity      0.3

              Gage Bias
Reference        Bias   %Bias      P
  Average    -30.4667     0.9  0.000
     4980    -58.0000     1.8  0.000
    12080    -33.4000     1.0  0.001
    24960      0.0000     0.0  1.000
```

Figure 4.12 Analytical results of bias and linearity of the viscometer from the Minitab program.

From Figure 4.12, the fitted line with a slope of 0.0028654 (close to zero) was observed. The percentages of the average bias and linearity were 0.9% and 0.3% of the total process variations (3,256 cP). Theoretically, the percentages of the average bias and the linearity less than 5% are acceptable [12]. Thus, it is concluded that the viscometer is accurate over the range of the reference fluid viscosities of 4,980 – 24,960 cP which covers the average value of the hair conditioner viscosity of 20,602 cP in this case study.

Due to the periodic calibration program of the company which is done every 4 months or 3 times a year, the stability of the viscometer is inferred to be reliable. According to the resolution, bias, linearity and stability, therefore, it is concluded that the measurement system is accurate.

The precision in terms of the repeatability and reproducibility (Gage R&R) of the measurement system were checked. The different viscosities of 10 samples from 5 formulas of hair conditioner were measured by four operators followed the same procedure using the same viscometer with two replicates. Six graphical outputs of Gage R&R study by the Minitab program are shown in Figure 4.13.

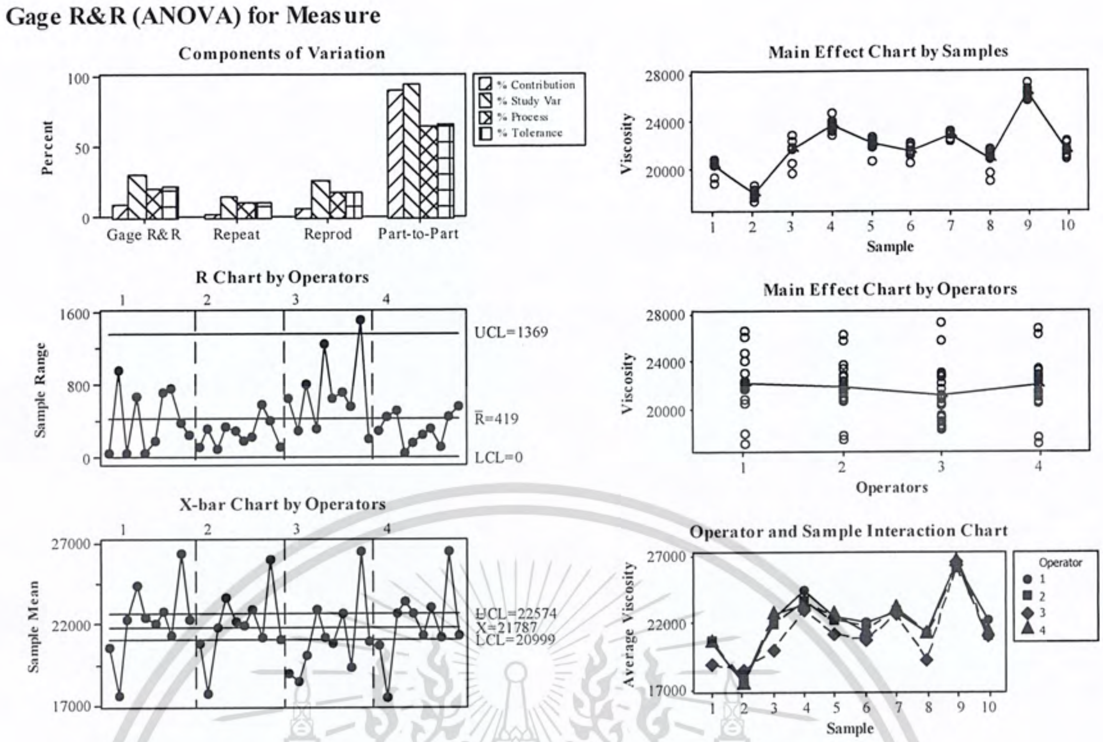


Figure 4.13 The six graphical outputs of GR&R study by the Minitab program.

1. The components of variation chart

As shown in Figure 4.14, 4 separate bars of the Part-to-Part variation, i.e., % contribution, % study variation, % process and % tolerance, are greater than those of other variation components (Gage R&R, Repeatability and Reproducibility). Therefore, the major variation of the measurements in this work came from the Part-to-Part variation (different parts or samples). It means the measurement system is precise.

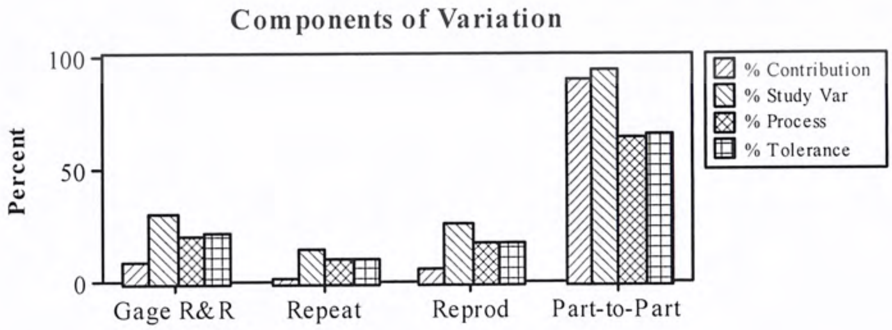


Figure 4.14 The components of variation chart in this work from the Minitab program.

2. R chart by operators

From the sample ranges shown in Figure 4.15, it shows that the first, second and fourth operators have high repeatability, but not the third operator. Because an out-of-control sample range belongs to the third operator and a large difference between two measurement values is found. This means the third operator lacks of the measurement skills. Therefore, the third operator must be trained how to prepare the sample as well as to use the viscometer for viscosity measurement.

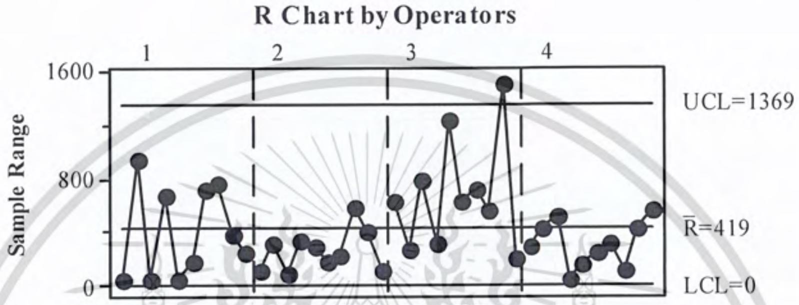


Figure 4.15 R chart by operators in this work from the Minitab program.

3. X-bar chart by operators

In the X-bar chart, the distance between the control limits (UCL and LCL) or measurement system variation represents the variation of the gage or measurement device, while the sample range represents the process output variation. The device variation must be less than the process output variation. Most of the points, at least 50% or more, are expected to locate outside the control limits to meet this requirement. From Figure 4.16, more than 50% of the points locate outside the control limits indicating that the viscometer can be used to distinguish the process output which is hair conditioner viscosity in this work.

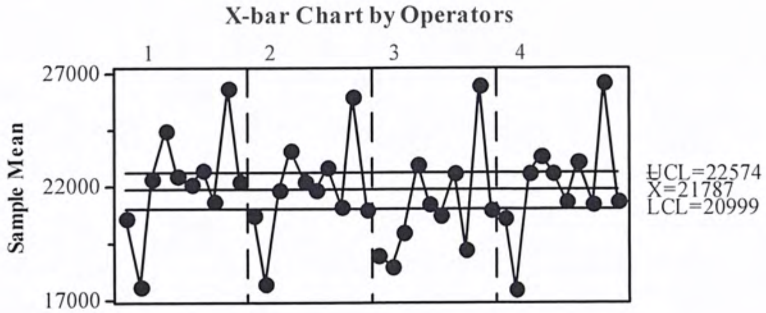


Figure 4.16 X-bar chart by operators in this work from the Minitab program.

4. Main effect chart by samples

From Figure 4.17, the measurements of the same sample by 4 operators are close to each other along 10 different samples. Therefore, the different samples have no effect on the operators to the measurement system.

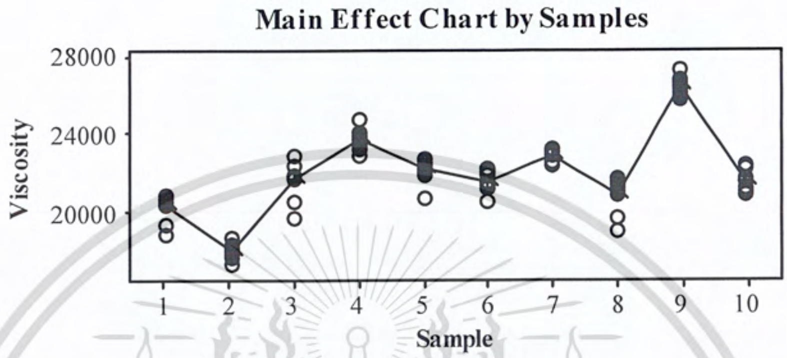


Figure 4.17 The main effect chart by samples in this work from the Minitab program.

5. Main effect chart by operators

From Figure 4.18, the third operator produces lower measurement values. The line connecting the average value of the viscosities approaches horizontal with higher reproducibility for the measurement system in case the average value made by the third operator increases.

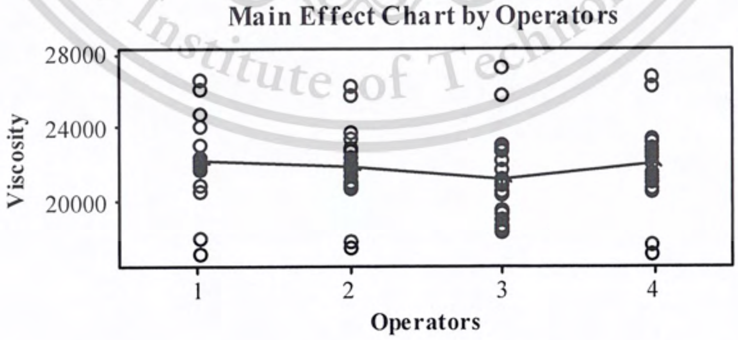


Figure 4.18 The main effect chart by operators in this work from the Minitab program.

6. The operator and sample interaction chart

From Figure 4.19, the viscosity measurements made by the third operator always deviate from others', resulting in the lowest measurement values in samples 1, 3, 5, 6 and 8. According to the results shown in R chart by operators and this chart, the third operator must be trained to improve the viscosity measurement skills on sample preparation, using the viscometer and taking the measurement procedures.

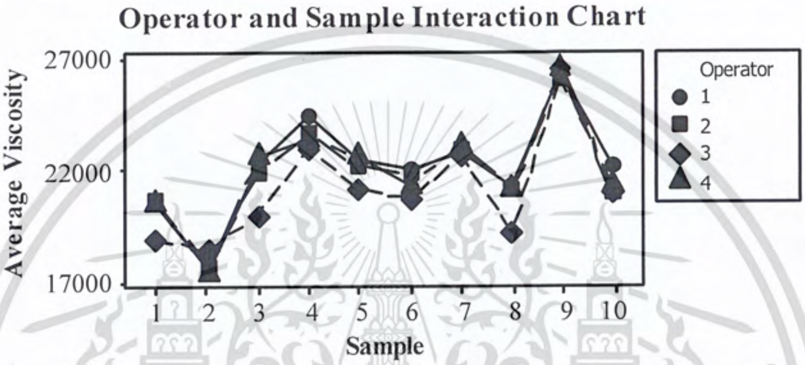


Figure 4.19 The operator and sample interaction chart in this work from the Minitab program.

In addition, the numerical key outputs in this work from GR&R study by the Minitab program were summarized in Table 4.5.

Table 4.5 The numerical key outputs from GR&R study by the Minitab program.

Sources of variation	**Variation (σ^2)**	**% Variation**	**P/T**	**P/TV**
Measurement system or GR&R variation	543,738	10.08	22.12	21.89
Repeatability	140,780	2.61	11.26	11.14
Reproducibility	402,958	7.47	19.04	18.85
Process output or part to part variation	4,851,463	89.92	66.68	65.40
Total variation	5,395,201	100.00	69.68	68.87
Number of distinct categories (NDC) = 4				

From Table 4.5, the main measurement variation of 89.92% came from the process output variation, same as the results in Figure 4.14. Both of the gage precision to tolerance (P/T) and gage precision to total variation (P/TV) values of the measurement system variation are 22.12 and 21.89. The values less than 30% indicate that the measurement system is able to accept a conformance process output or reject an undesired process output precisely at the given specification limits, and the measurement variation is small relative to the total variation. Moreover, the number of distinct categories is in the acceptance range of 3-4.

According to the graphical and key numerical outputs from the Minitab program in Gage R&R study, the precision of the measurement system is confirmed.

From the measurement system analysis, the third operator was trained. Finally, it is concluded that the viscosity measurement system is reliable.

4.3.3 Significant parameters by the multiple linear regression analysis

From 21 numbers of observations among 5 formulas of hair conditioner, the scatter plots between the hair conditioner viscosities (Y) and 20 selected possible parameters (X_1 – X_{20}) are shown in Figure 4.20. The slope of the scatter plot shows how much the relevant parameter, X_i, affects the viscosity. The positive correlation is indicated by a positive slope and vice versa by the negative relation. In case of no relation or no effect, the scatter plot is a horizontal line or does not show any trend.

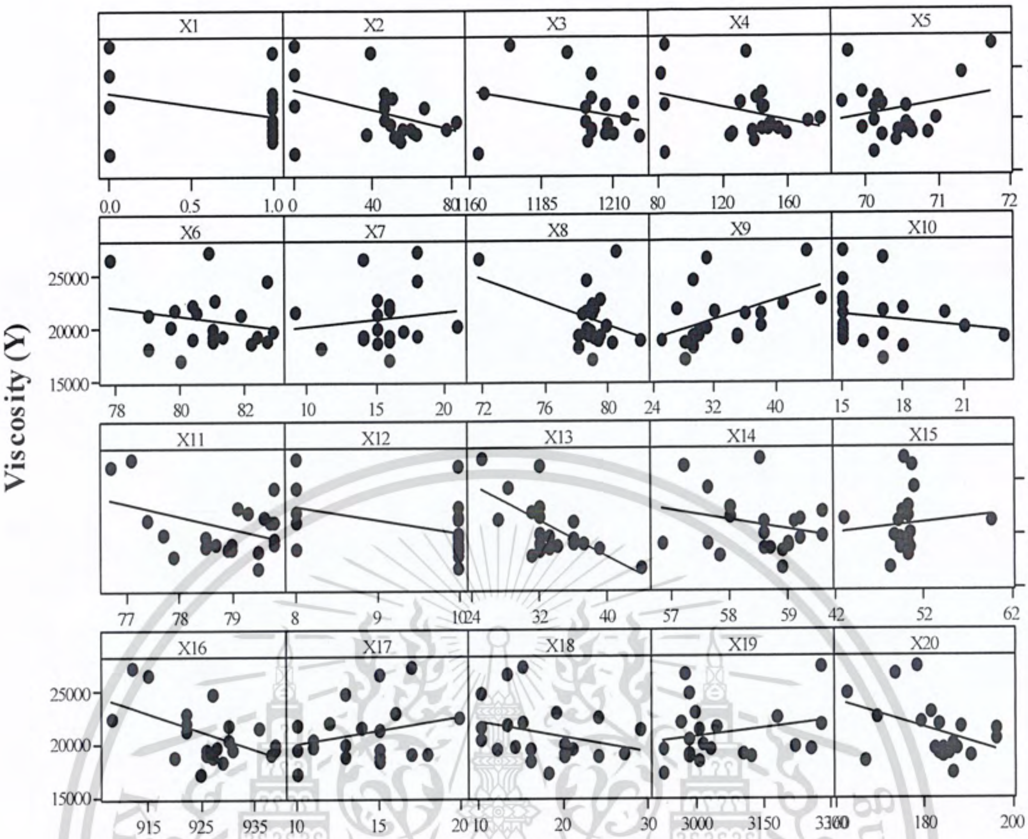


Figure 4.20 The scatter plots between hair conditioner viscosities (Y) and the possible parameters X_1-X_{20}.

From 20 selected possible parameters, the significant parameters were statistically evaluated by the multiple linear regression analysis using the Minitab program at the VIFs less than 10 and the p-value less than 0.05 [8]. From Table 4.6, four significant parameters are the quantity of hot chlorinated water added in the main mixer, time to dissolve cationic surfactant type 1, the temperature at side mixer no. 1 to dissolve fatty alcohols, and time to cool down the main mixer to 58 °C before adding the ambient-temperature water.

Table 4.6 Significant parameters by the multiple linear regression analysis in the Minitab program.

X_i	Parameter	VIF	p-value
X_3	The quantity of hot chlorinated water added in the main mixer (kg)	1.0	0.012
X_7	Time to dissolve cationic surfactant type 1 (min)	1.1	0.018
X_8	The temperature at side mixer no. 1 to dissolve fatty alcohols (°C)	1.1	0.001
X_{13}	Time to cool down the main mixer to 58 °C before adding the ambient-temperature water (min)	1.0	0.000

The correlation between the 4 significant parameters and the viscosity is

$$\text{Viscosity} = 155{,}865 - 57.98\,X_3 + 351.6\,X_7 - 700.3\,X_8 - 469.83\,X_{13} \qquad (4.1)$$

The obtained correlation was confirmed by the coefficient of determination (R^2) of 78.7% and the residual model diagnostics by the Minitab program in Figure 4.21, indicating the linear pattern of the normal probability plot of the residuals. The histogram of the residuals is a normal distribution. As seen in the plot of residuals versus the fitted values, the residuals fluctuate from the zero line without a repeating pattern as the fitted values increase indicating that the residuals has a constant variance. From the plot of residuals versus the observation orders, all of the data are random. According to the results obtained by the residual model diagnostics, it is concluded that the number of the collected data is sufficient and the obtained correlation is reliable [29].

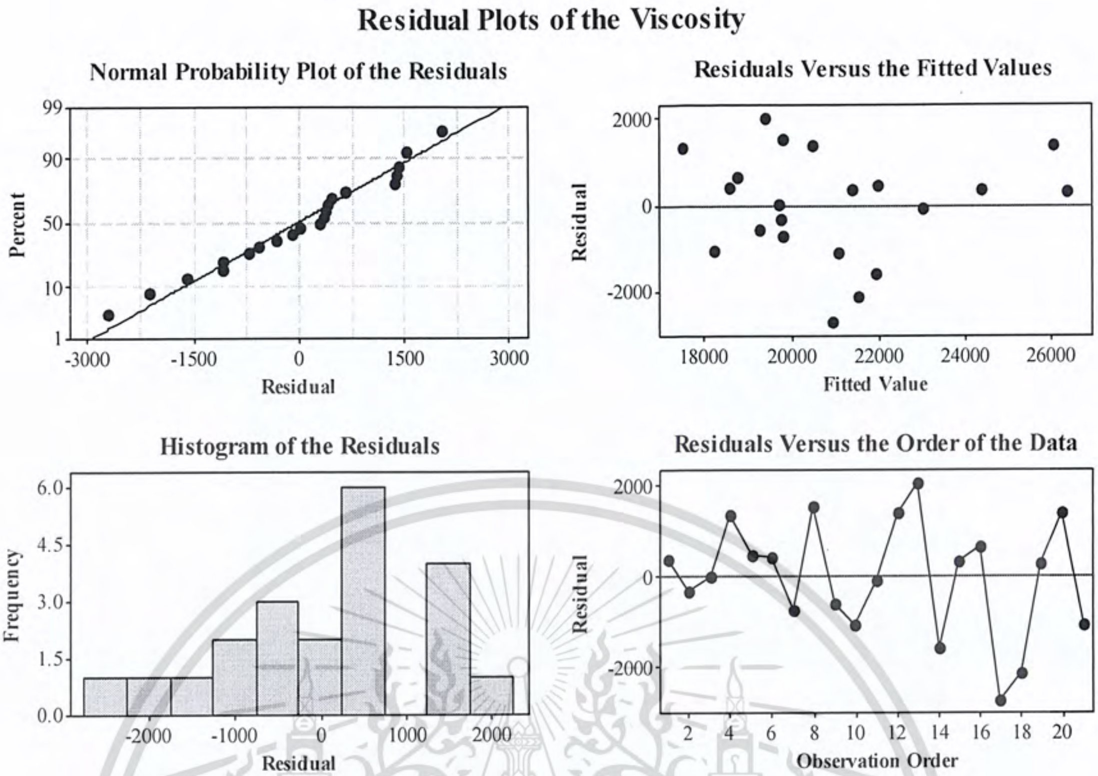


Figure 4.21 Residual model diagnostics for the regression model with R^2 of 78.7 %.

From the correlation in Equation 4.1, the quantity of hot chlorinated water added in the main mixer, the temperature at side mixer no. 1 to dissolve fatty alcohols and time to cool down the main mixer to 58 °C before adding the ambient-temperature water negatively affect the hair conditioner viscosity; but time to dissolve cationic surfactant type 1 shows the positive effect. From the multiple linear regression analysis, the viscosity of hair conditioner increases with a small quantity of hot chlorinated water in the main mixer, low temperature at side mixer no. 1 to dissolve fatty alcohols, and a short time to cool down the main mixer to 58 °C ; but a long time to dissolve cationic surfactant type 1. The correlation between the significant parameters and the hair conditioner viscosity are described as follows:

1. X_3: the quantity of hot chlorinated water added in the main mixer

Basically, the mixture or slurry with water in its composition becomes viscous when the content of water decreases. In hair conditioner production, water is added during raw material mixing process for different purposes, for example,

- hot chlorinated water to dissolve cationic surfactants at the main mixer and to dissolve fatty alcohols at side mixer no. 1

- the ambient-temperature water to cool down the main mixer, to dissolve thickener agent in side mixer no. 2 and to dissolve additives in side mixer no. 3

The largest amount of hot chlorinated water accounts for 40% of the total weight of hair conditioner product per batch of production. From the observed data in the multiple linear regression analysis, the standard deviation of the quality of hot chlorinated water added in the main mixer is 15.42 kg, which is greater than those in other steps. As a result, the quantity of hot chlorinated water has a significant effect on the deviation of hair conditioner viscosity.

2. X_7: time to dissolve cationic surfactant type 1

Two forms of cationic surfactants for hair conditioner production are solid and liquid. The cationic surfactant type 1 is solid, whereas the cationic surfactant types 2 and 3 are liquid. To dissolve all solid of the cationic surfactant type 1 completely in hot chlorinated water at the main mixer, a period of time is required. As shown by Figure 4.10, the mixture of three cationic surfactants and hot chlorinated water is mixed with fatty alcohols. The cationic surfactants act as the emulsifiers to form the emulsion between fatty alcohols and water. The mixture at this stage is a white translucent liquid. It is more viscous when the temperature decreases and becomes a white opaque solid at room temperature. The complete dissolved cationic surfactant type 1 forms a homogeneous and stable mixture so that the hair conditioner viscosity less deviated from the target value.

3. X_8 and X_{13}: the temperature to dissolve fatty alcohols at side mixer no. 1 and time to cool down the main mixer to 58 °C before adding the ambient-temperature water

Fatty alcohols are dissolved by hot chlorinated water at side mixer no. 1 before transferring to the main mixer. From the historical data in hair conditioner production, the temperature at side mixer no. 1 of 75 – 85 °C can dissolve the solid of fatty alcohols. A high temperature at side mixer no. 1 results in high temperature of the mixture of hot chlorinated water, fatty alcohols and the cationic surfactants. The obtained mixture is subsequently cooled down to 58 °C before the ambient-temperature water is added in the main mixer. If the temperature of the mixture is high, a long time is required to cool down the main mixer. The mixture is agitated by a three-blade turbine at constant speed of 150 rpm until the temperature at the main mixer reaches 58 °C. The more time is used to cool down the main mixer, the longer time the mixture is sheared. Because the hair conditioner is pseudoplastic fluid and time-sensitive to shearing, therefore, the viscosity of hair conditioner product decreases if the temperature at side mixer no. 1 increases and when a long cooling down time is used.

4.4 Strongly significant parameters on the deviation of hair conditioner viscosity and the optimum operating conditions

The maximum and minimum values of time to cool down the main mixer to 58 °C, X_{13} in Table 4.6, could not be identified definitely by the DOEs because time varied with the cooling rate and the set-point temperature at the main mixer before adding the ambient-temperature water. The slower cooling rate and/or the lower set-point temperature, the more time was required to cool down the main mixer. However, the set-point temperature was more practically set at the fixed values than the cooling rate. In addition, from the results in Figure 4.22, it was found that the set-point temperature at the main mixer of 60-66 °C had a significant correlation of 88.58% with time to cool down the main mixer to the set-point temperature. Therefore, the set-point temperature at the main mixer was practically able to represent the time to cool down the main mixer and could be checked by the DOEs whether it had a strongly significant effect on the deviation of hair conditioner viscosity or not.

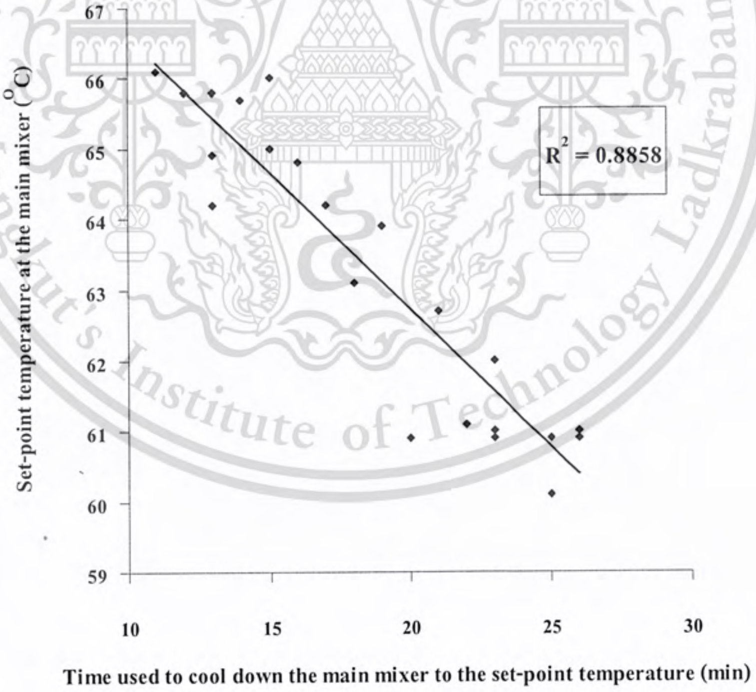


Figure 4.22 Scatter plot of the set-point temperature at the main mixer and time to cool down the main mixer to the set-point temperature.

By using the DOEs in the Minitab program, the 16-run experimental orders with 2 replicates were decided. The viscosities of hair conditioners are shown in Table 4.7. Figure 4.23 is the Pareto plot of the effects of the 4 significant parameters. It was found that all of the 4 significant parameter bars were higher than the alpha line of 0.05 at the significance level of 0.05, indicating that the 4 significant parameters significantly contributed to the response variable which is the observed viscosities. Therefore, it can be concluded that all of the significant parameters had strongly significant effects on the deviations of the viscosities. Furthermore, from Figure 4.23, the 2-way and 3-way interactions among the 4 strongly significant parameters also showed the statistical significance on the viscosity, which were illustrated by the two-parameter and three-parameter interaction bars that exceeded the alpha line.

Table 4.7 Experimental conditions of the relevant strongly significant parameters and hair conditioner viscosities.

Run order	Quantity of hot chlorinated water (kg)	Cationic surfactant dissolving time (min)	Side mixer no. 1 temperature (°C)	Set-point temperature at the main mixer (°C)	Viscosity of hair conditioner (cP) (replicate 1)	Viscosity of hair conditioner (cP) (replicate 2)
1	700	10	75	60	27,777	28,203
2	800	10	75	60	27,600	27,813
3	700	15	75	60	27,032	27,174
4	800	15	75	60	28,558	28,487
5	700	10	80	60	25,186	25,434
6	800	10	80	60	27,759	28,061
7	700	15	80	60	28,948	28,664
8	800	15	80	60	31,184	31,326
9	700	10	75	65	25,736	26,251
10	800	10	75	65	25,505	26,144
11	700	15	75	65	26,535	25,789
12	800	15	75	65	26,641	27,316
13	700	10	80	65	27,032	26,819
14	800	10	80	65	27,742	27,706
15	700	15	80	65	27,138	27,351
16	800	15	80	65	27,245	27,351

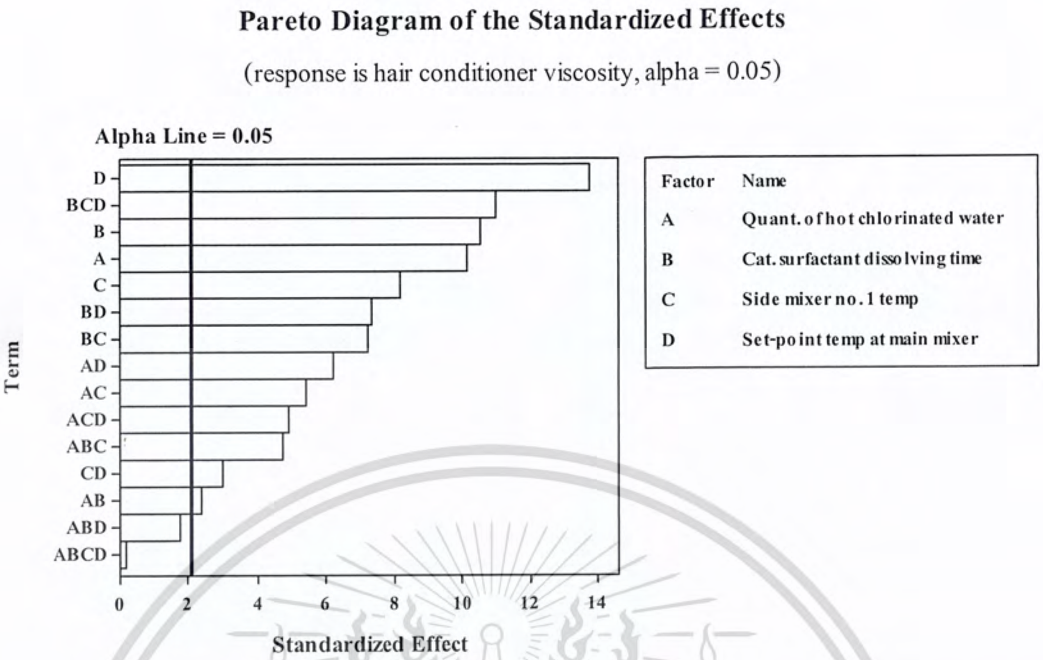


Figure 4.23 The Pareto plot of the effects of the 4 significant parameters on hair conditioner viscosity from the Minitab program.

The main effect plots and interaction effect plots of 4 strongly significant parameters are shown in Figures 4.24 and 4.25. The main effect plots with positive correlation illustrates that the viscosity increases when the quantity of hot chlorinated water added in the main mixer, time to dissolve cationic surfactant type 1, and the temperature at side mixer no. 1 to dissolve fatty alcohols increase. On the other hand, the viscosity decreases when the set-point temperature at the main mixer increases. However, the lower values of the three positive-correlation parameters with the viscosity, and the upper value of the negative-correlation parameter make the viscosity approached the target value of 26,000 cP. From the interaction effect plots, it is found that there is a lack of parallelism of the lines in almost all of the interaction effect plots among 4 strongly significant parameters, which are the interaction effect plots between

1. the quantity of hot chlorinated water and temperature at side mixer no. 1
2. the quantity of hot chlorinated water and the set-point temperature at the main mixer

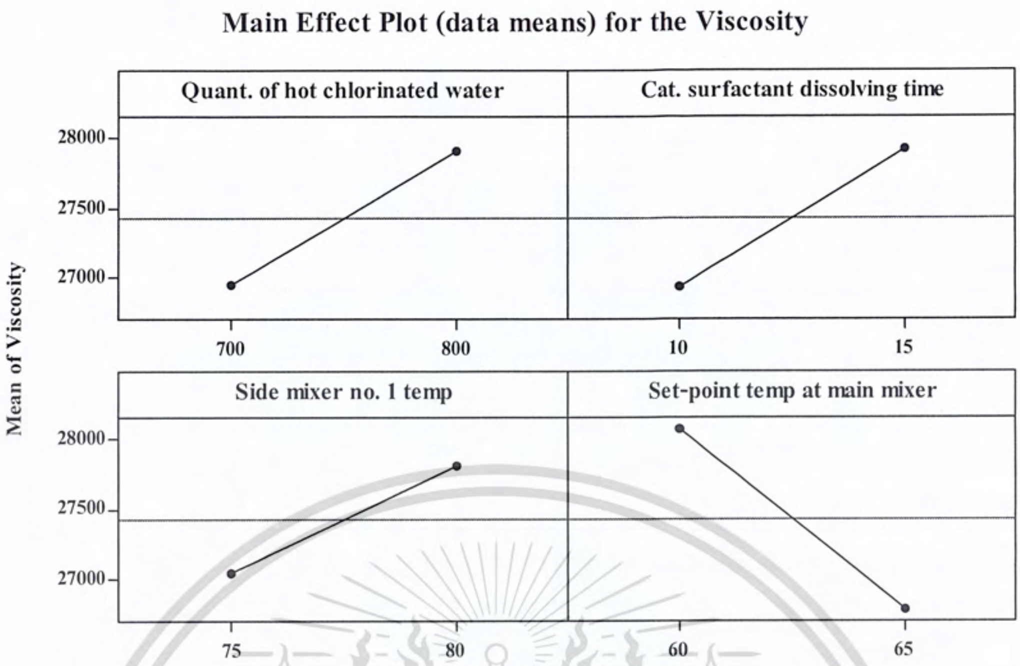


Figure 4.24 Main effect plots of 4 strongly significant parameters on hair conditioner viscosity.

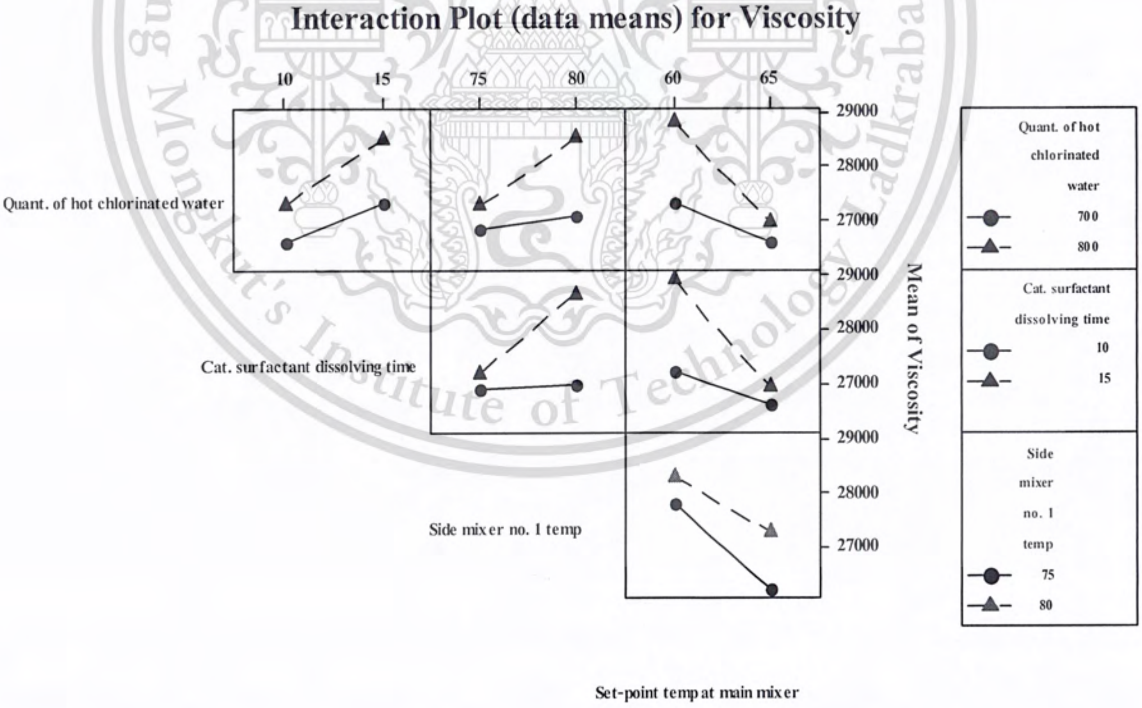


Figure 4.25 Interaction effect plots of 4 strongly significant parameters on hair conditioner viscosity.

3. time to dissolve cationic surfactant type 1 and the temperature at side mixer no. 1
4. time to dissolve cationic surfactant type1 and the set-point temperature at the main mixer
5. the temperature at side mixer no. 1 and the set-point temperature at the main mixer

The optimum operating conditions of 4 strongly significant parameters, which made the hair conditioner viscosities approached the target value of 26,000 cP, were selected by the cube plot in Figure 4.26, i.e.; the quantity of hot chlorinated water of 700 kg, time to dissolve cationic surfactant type 1 of 10 min, the temperature at side mixer 1 to dissolve fatty alcohols of 75 ºC, and the set-point temperature at the main mixer before adding the ambient-temperature water of 65 ºC.

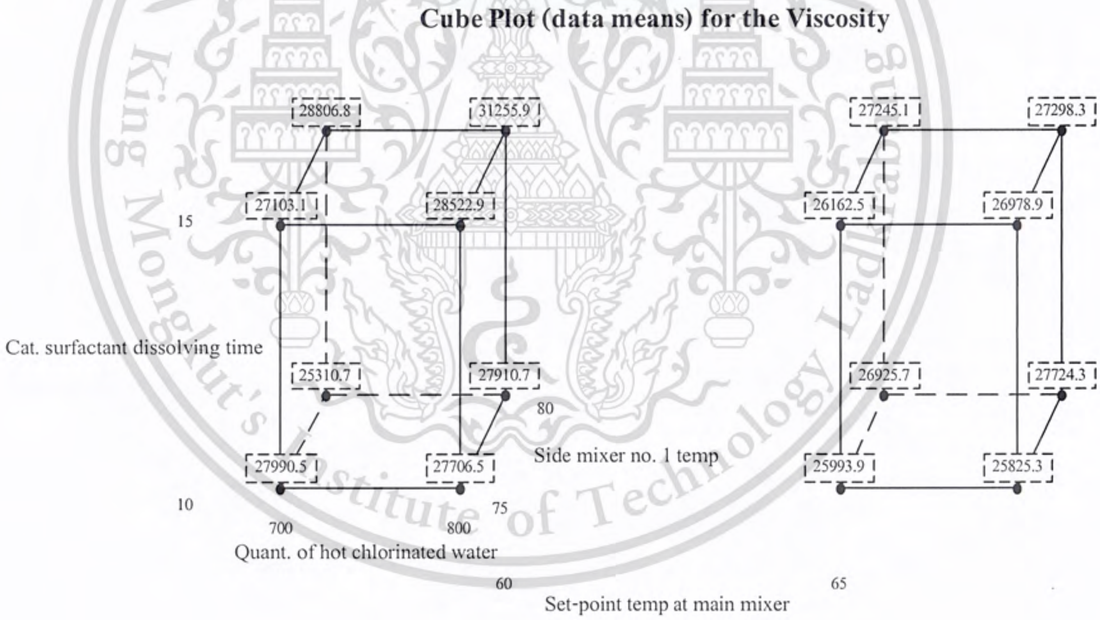


Figure 4.26 The cube plot of hair conditioner viscosities at different conditions of the strongly significant parameters.

The optimum operating conditions were subsequently used in hair conditioner raw material mixing process. The hair conditioner viscosities are shown in Table 4.8.

Table 4.8 Hair conditioner viscosities after using the optimum conditions.

Hair conditioner	Viscosity (cP)	Hair conditioner	Viscosity (cP)	Hair conditioner	Viscosity (cP)
SCDR	24,654	SCSS	27,032	SCAC	26,535
SCBS	29,801	SCSS	27,352	SCAC	27,671
SCBS	28,417	SCSS	27,813	SCDR	24,335
SCAC	27,742	SCDR	24,264	SCBS	25,861
SCSS	25,329	SCSS	28,275	SCLS	27,174
SCDR	25,932	SCAC	27,032	SCLS	27,813
SCDR	28,204	SCAC	27,565	SCLS	26,074
SCAC	28,559	SCBS	28,062	SCSS	25,151
SCLS	28,949	SCBS	26,606	SCDR	24,122
SCLS	31,327	SCLS	27,032	SCBS	24,974

Figure 4.27 shows the histogram of the hair conditioner viscosities, the standard deviation, the DPMO and the potential process capability index of the hair conditioner raw material mixing process after the improvements. It was found that the average standard deviation of hair conditioner viscosities in this case study reduced about 47% from 3,256 to 1,718 cP. The DPMO reduced from 61,217 to 1.57, and the C_p increased from 0.65 to 1.75. In fact, for any existing process the typical value of DPMO must less than 6,210 and the C_p higher than 1.33 [6]. Finally, the achieved optimum operating conditions were used in the process. The standard operating procedures (SOPs) for the raw material mixing process were written, and shown in Appendix B. From this work, hair conditioner loss was reduced.

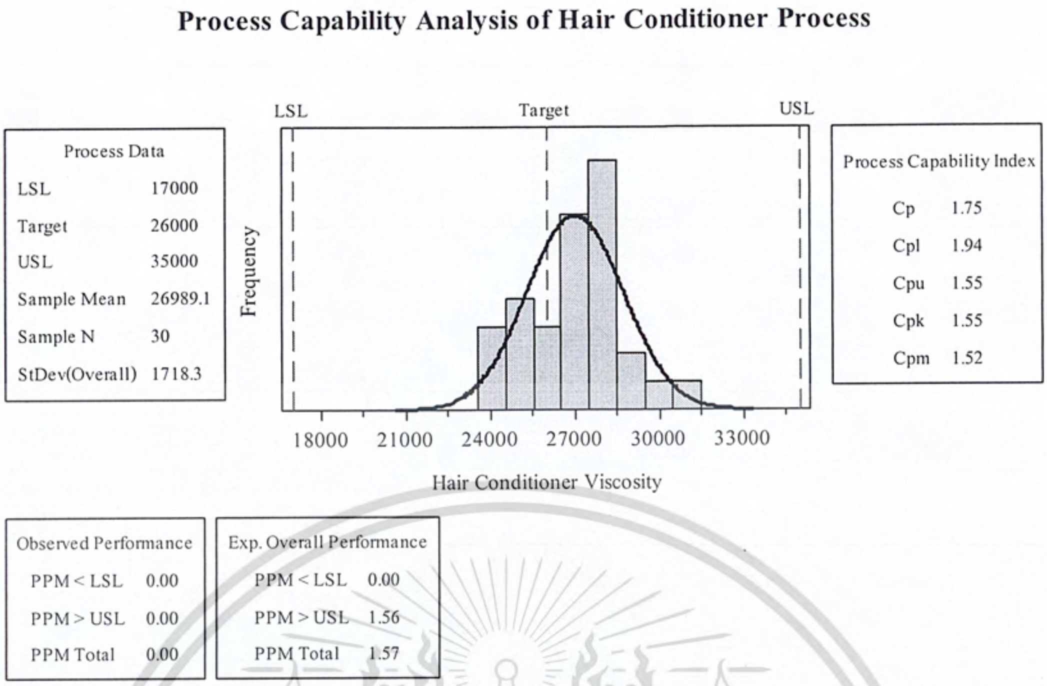


Figure 4.27 Histogram of the hair conditioner viscosities after the improvements.

Chapter 5

Conclusions and Suggestions

5.1 Conclusions

Hair conditioners are pseudoplastic fluids and time-sensitive to shearing. From the high defect per million opportunities (DPMO) and the unpleasant potential process capability index (C_p), the deviation of hair conditioner viscosity in this case study was prioritized. The 4 strongly significant parameters on the deviation of hair conditioner viscosity and their optimum operating conditions are the quantity of hot chlorinated water of 700 kg, time to dissolve cationic surfactant type 1 of 10 min, temperature at side mixer no. 1 to dissolve fatty alcohols of 75°C, and the set-point temperature at the main mixer before adding the ambient-temperature water of 65°C. After using the optimum operating conditions in raw material mixing process, the average standard deviation of hair conditioner viscosities was reduced about 47% from 3,256 to 1,718 cP. The high DPMO was reduced dramatically from 61,217 to 1.57 as well as the unpleasant C_p was increased to the acceptable range of 1.75. Table 5.1 shows the average standard deviation of hair conditioner viscosity, DPMO and C_p in raw material mixing process before and after the improvements. The standard operating procedures (SOPs) for 5 formulas of hair conditioner production were written. Based on the obtained DPMO, the calculated cost savings were equivalent to hair conditioner of 306 tons/year. The details of cost saving calculations are shown in Appendix C.

Table 5.1 The average standard deviation of hair conditioner viscosity, the defect per million opportunities (DPMO) and the potential process capability index (C_p) in raw material mixing process before and after the improvements.

Process performance indicator	Before the improvements	After the improvements
Average standard deviation of hair conditioner viscosity (cP)	3,256	1,718
Defect per million opportunities: DPMO (ppm)	61,217	1.57
Potential process capability index (C_p)	0.65	1.75

5.2 Suggestions

The company should keep monitoring the process performance of raw material mixing process (DPMO and C_p) and controlling the strongly significant parameters in raw material mixing process at the optimum operating conditions.

The methodology to minimize the deviation of hair conditioner viscosity in this work can be applied to a similar production process.

References

[1] Sung H. Park. **Six Sigma for Quality and Productivity Promotion**. Asian Productivity Organization. 2003.

[2] Wikimedia Foundation Inc., "Hair Conditioner", [Online]. Available: http://en.wikipedia.org/wiki/Hair_conditioner/ 2007.

[3] H2G2, "How Hair Conditioner Works?", [Online]. Available: http://www.bbc.co.uk/dna/h2g2/A851627. 2002.

[4] Brookfield Engineering Labs. Inc., **More Solutions to Sticky Problems**. Technical Papers on Viscosity Measurement and Control: Brookfield. 2000.

[5] Barnes, H. A., Hutton, J. F. and F.R.S., K. Walters, **An Introduction to Rheology**. New York: Elsevier Science. 1989.

[6] Douglas C. Montgomery. **Introduction to Statistical Quality Control**. 2nd ed. Canada: John Wiley & Sons. 1990.

[7] Forrest W. Breyfogle, James M. Cupello and Becki Meadows. **Managing Six Sigma**, John Wiley & Sons. 2001.

[8] Douglas C. Montgomery, Elizabeth A. Peck and G. Geoffrey Vining. **Introduction to Linear Regression Analysis**. 4th ed. Canada: John Wiley & Sons. 2006.

[9] Douglas C. Montgomery, George C. Runger and Norma Faris Hubele. **Engineering Statistics**. 3rd ed. Canada: John Wiley & Sons. 2004.

[10] Erwin Kreyszig. **Advanced Engineering Mathematics**. 8th ed. Singapore :John Wiley & Sons. 1999.

[11] Douglas C. Montgomery. **Design and Analysis of Experiments**. 5th ed. John Wiley & Sons. 2001.

[12] กิติศักดิ์ พลอยพานิชเจริญ. **การวิเคราะห์ระบบการวัด**. พิมพ์ครั้งที่ 1. กรุงเทพฯ: สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น). 2542.

[13] The Unilever Thai Holdings Ltd. **Introduction to Measurement System Analysis**; Training Materials for Six sigma: Analyze phase. 2006.

[14] Rezayati-Charati, P., Mohammadi-Rovshandeh, J., Hashemi, S. and Kazemi-Najafi, S. “Influence of Dimethyl Formamide Pulping of Bagasse on Pulp Properties” **Bioresource Technology**, 2005. [Online]. Available: http://www.sciencedirect.com.

[15] Suwittayaruk, P. **“Six Sigma Quality Enhancement of Motors Through Decrease the Seawax Chemical PF 5060 Usage in HSA Line Manufacturing of the Enterprise Product.”** Master Thesis, Department of Production Engineering, Faculty of Engineering, King Mongkut’s Institute of Technology Thonburi. 2001. (in Thai).

[16] Teerakitwattana, L. **“Reduction of Ice-cream Weight Deviation.”** Master Thesis, Department of Production Engineering, Faculty of Engineering, King Mongkut’s University of Technology Thonburi. 2002. (in Thai).

[17] Chunkoh, K. **“Improvement of Shampoo Viscosity in Mixing Process Using Statistical Tecgnique.”** Master Thesis, Department of Production Engineering, Faculty of Engineering, King Mongkut’s University of Technology Thonburi. 2001. (in Thai).

[18] Temsawangleart, P. **“A Reduction of Bulk Density Variation in a Non-Soap Detergent Making Process.”** Master Thesis, Department of Chemical Engineering, Faculty of Engineering, King Mongkut’s Institute of Technology Ladkrabang. 2001. (in Thai).

[19] อัญชลีพร วาริทสวัสดิ์ หล่อทองคำ และพิทักษ์ เต็มแสวงเลิศ “การวิเคราะห์ตัวแปรที่มีผลต่อค่า Bulk Density ในกระบวนการผลิตผงซักฟอกด้วยวิธีการถดถอยแบบหลายเชิง” **วารสารวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่**, Vol. 9, No. 1, March 2001 หน้า84-91.

[20] Chinsutthi, R. **“A Reduction of Rework in a Non-soap Detergent Making Process.”** Master Thesis, Department of Chemical Engineering, Faculty of Engineering, King Mongkut’s Institute of Technology Ladkrabang. (in Thai). 2003.

[21] อัญชลีพร วาริทสวัสดิ์ หล่อทองคำ รดีบูรณ์ ชินสุทธิ และสุรเชษฐ์ จันตรี “การลดปริมาณผงซักฟอกที่ไม่ได้มาตรฐานในกระบวนการผลิตผงซักฟอกชนิดธรรมดา” **วารสารพระจอมเกล้าลาดกระบัง**, ปีที่ 11 ฉบับที่ 1, เมษายน 2546 หน้า 7-17.

[22] Thumsavate, P. **"Analysis of Significant Parameters Affecting Surface Flexography Printing System on Polypropylene Tape by MINITAB Program."** Master Thesis, Department of Chemical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. 2005. (in Thai).

[23] Lothongkum A.W. and Thumsavate P. "Analysis of Significant Variables Affecting Surface Flexography Printing System on Adhesive Laminated Polypropylene Tape by MINITAB Program", **Proceedings of the 7th Triennial AEESEAP Conference on Enhancing Engineering Education and Training (A University-Industry-Government Partnership)**, 8-9 December 2003, Kuala Lumpur, Malaysia, pp.167-173.

[24] Lothongkum W.A., Thumsavate P. and Seangkiatiyuth K. "Significant Parameters Determined by the MINITAB Program to Affect Ink Abrasion Resistance of Polypropylene Tape and their Application to Surface Flexography Printing", **J. Ind. Eng. Chem**., Vol. 12, No.1, 2006. pp. 44-49

[25] Tuladhorn, M. **"Weight- Deviation Reduction of Liquid Products in the Packaging Process for Consumer Products Industry."** Master Thesis, Department of Chemical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. 2006. (in Thai).

[26] Lothongkum A.W., Tuladhorn M., Tanprasert P. and Tanthapanichakoon W. "Reduction of Weight-deviation of Dish-washing Liquid by Minitab Program", **Proceedings of the 1st KMITL International Conference on Integration of Science & Technology for Sustainable Development**, 25-26 August, 2004, Vol.1, KMITL, Bangkok, Thailand, pp. 15-18.

[27] Lothongkum A.W., Tuladhorn M., Tanprasert P. and Tanthapanichakoon W. "Weight-deviation Reduction in Dish-washing Liquid Packaging Process: A Case Study in a Consumer Products Industry", **Proceedings of the 10th Congress of Asian Pacific Confederation of Chemical Engineering** , Session General Area 4I-04, October 17-21, 2004, Kitakyushu, Japan.

[28] Im-Orb K. and Phaiboonsilpa N. **"Increasing of the Production Capacities of Personal Care Products by Total Productive Maintenance."** Senior Project, Department of Chemical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang. 2004. (in Thai).

[29] The Unilever Thai Holdings Ltd. **Multiple Regression**. Training Materials for Six sigma: Analyze phase. 2006.

APPENDICES

Appendix A

Statistical Tables

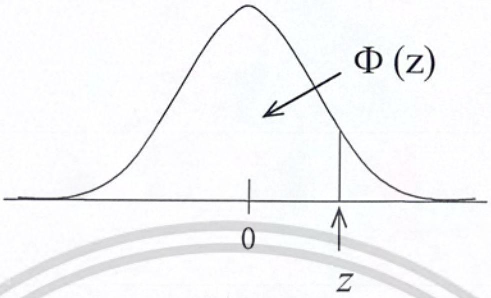
Appendix A-1: Cumulative standard normal distribution table

Appendix A-2: *F* distribution table

Appendix A-3: *t* distribution table

Appendix A-1: Cumulative standard normal distribution table [9]

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}u^2}\, du$$

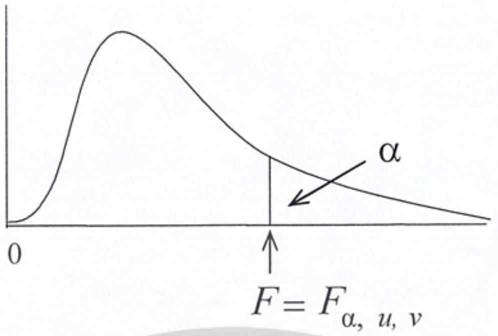


For example, if z = 1.53, the cumulative standard normal distribution table provides the probability $P\,(Z \le 1.53) = 0.936992$, see Table A-1.

Table A-1 Cumulative standard normal distribution

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**	**z**
0.00	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.523922	0.527903	0.531881	0.535856	**0.00**
0.10	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345	**0.10**
0.20	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092	**0.20**
0.30	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732	**0.30**
0.40	0.655422	0.659097	0.662757	0.666402	0.670031	0.676645	0.677242	0.680822	0.684386	0.687933	**0.40**
0.50	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405	**0.50**
0.60	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903	**0.60**
0.70	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236	**0.70**
0.80	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267	**0.80**
0.90	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913	**0.90**
1.00	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143	**1.00**
1.10	0.864664	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977	**1.10**
1.20	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475	**1.20**
1.30	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736	**1.30**
1.40	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888	**1.40**
1.50	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083	**1.50**
1.60	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486	**1.60**
1.70	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.964636	0.962462	0.963273	**1.70**
1.80	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621	**1.80**
1.90	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705	**1.90**
2.00	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691	**2.00**
2.10	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738	**2.10**
2.20	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.955696	0.988989	**2.20**
2.30	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576	**2.30**
2.40	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613	**2.40**
2.50	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201	**2.50**
2.60	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427	**2.60**
2.70	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365	**2.70**
2.80	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074	**2.80**
2.90	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605	**2.90**
3.00	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999	**3.00**
3.10	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289	**3.10**
3.20	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499	**3.20**
3.30	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650	**3.30**
3.40	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758	**3.40**
3.50	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835	**3.50**
3.60	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888	**3.60**
3.70	0.999892	0.999896	0.999900	0.888804	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925	**3.70**
3.80	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950	**3.80**
3.90	0.999952	1.00	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967	**3.90**

Appendix A-2: *F* distribution table [9]



For example, $F_{\alpha,\ u,\ v}$ at the level of significance (α) of 0.05 or confident level 95%, the numerator degree of freedom (u) of 15 and the denominator degree of freedom (v) of 21, the F distribution table (Table A-3) provides the $F_{\alpha,\ u,\ v} = 2.18$.

Table A-2 Percentage points $F_{\alpha,\ u,\ v}$ of the F distribution at the level of significance of 0.10

		Degree of freedom of the numerator (u)															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**30**	**60**	**∞**
Degree of freedom of the denominator (v)	**1**	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.7	61.2	61.7	62.3	62.8	63.3
	2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.46	9.47	9.49
	3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.17	5.15	5.13
	4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.82	3.79	3.76
	5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.17	3.14	3.10
	6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.80	2.76	2.72
	7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.56	2.51	2.47
	8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.38	2.34	2.29
	9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.25	2.21	2.16
	10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.16	2.11	2.06
	11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.08	2.03	1.97
	12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.01	1.96	1.90
	13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.96	1.90	1.85
	14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.91	1.86	1.80
	15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.87	1.82	1.76
	16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.84	1.78	1.72
	17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.81	1.75	1.69
	18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.78	1.72	1.66
	19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.76	1.70	1.63
	20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.74	1.68	1.61
	21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.72	1.66	1.59
	22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.70	1.64	1.57
	23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.69	1.62	1.55
	24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.67	1.61	1.53
	25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.66	1.59	1.52
	26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.65	1.58	1.50
	27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.64	1.57	1.49
	28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.63	1.56	1.48
	29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.62	1.55	1.47
	30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.61	1.54	1.46
	40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.54	1.47	1.38
	100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69	1.66	1.61	1.56	1.49	1.42	1.34	1.21
	∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.34	1.24	1.00

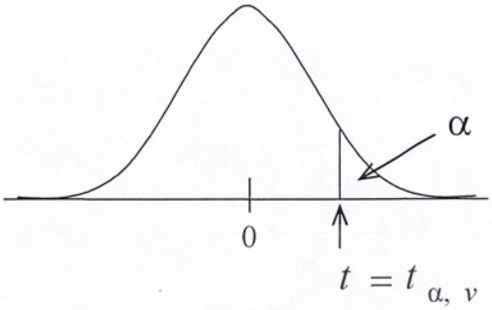
Table A-3 Percentage points $F_{\alpha,\ u,\ v}$ of the F distribution at the level of significance of 0.05

		Degree of freedom of the numerator (*u*)															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**30**	**60**	**∞**
Degree of freedom of the denominator (*v*)	**1**	161	199	215	224	230	233	236	238	240	241	243	245	248	250	252	254
	2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
	3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.53
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.63
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.36
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.67
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.23
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	3.01	2.93
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.71
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.70	2.62	2.54
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.57	2.49	2.40
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.38	2.30
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.38	2.30	2.21
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.31	2.22	2.13
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.07
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.19	2.11	2.01
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.15	2.06	1.96
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.11	2.02	1.92
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.07	1.98	1.88
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.84
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.01	1.92	1.81
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	1.98	1.89	1.78
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	1.96	1.86	1.76
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.94	1.84	1.73
	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.92	1.82	1.71
	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.90	1.80	1.69
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.88	1.79	1.67
	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.87	1.77	1.65
	29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.85	1.75	1.64
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.62
	40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.74	1.64	1.51
	100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.85	1.77	1.68	1.57	1.45	1.28
	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.00

Table A-4 Percentage points $F_{\alpha,\ u,\ v}$ of the F distribution at the level of significance of 0.01

		Degree of freedom of the numerator (u)															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**30**	**60**	**∞**
Degree of freedom of the denominator (v)	**1**	4052	4999	5403	5624	5763	5858	5928	5981	6022	6055	6106	6157	6208	6260	6313	6365
	2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5
	3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.5	26.3	26.1
	4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.8	13.7	13.5
	5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.38	9.20	9.02
	6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.88
	7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	5.99	5.82	5.65
	8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.20	5.03	4.86
	9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.31
	10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	3.91
	11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	3.94	3.78	3.60
	12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.70	3.54	3.36
	13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.51	3.34	3.17
	14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.35	3.18	3.00
	15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.87
	16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.10	2.93	2.75
	17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.00	2.83	2.65
	18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	2.92	2.75	2.57
	19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.84	2.67	2.49
	20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.42
	21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.72	2.55	2.36
	22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.67	2.50	2.31
	23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.62	2.45	2.26
	24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.58	2.40	2.21
	25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.54	2.36	2.17
	26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.50	2.33	2.13
	27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.47	2.29	2.10
	28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.44	2.26	2.06
	29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.41	2.23	2.03
	30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.01
	40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.20	2.02	1.80
	100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.37	2.22	2.07	1.89	1.69	1.43
	∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.00

Appendix A-3: *t* distribution table [9]



For example, $t_{\alpha,\ v}$ at the level of significance (α) of 0.05 or confident level 95% and the degree of freedom (v) of 30, the t distribution table provides the $t_{\alpha,\ v}$ = 1.697, see Table A-5.

Table A-5 Percentage points $t_{\alpha,\ v}$ of the t distribution

		Level of significance (α)					
		0.25	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
Degree of freedom (v)	**1**	1.000	3.078	6.314	12.706	31.821	63.657
	2	0.816	1.886	2.920	4.303	6.965	9.925
	3	0.765	1.638	2.353	3.182	4.541	5.841
	4	0.741	1.533	2.132	2.776	3.747	4.604
	5	0.727	1.476	2.015	2.571	3.365	4.032
	6	0.718	1.440	1.943	2.447	3.143	3.707
	7	0.711	1.415	1.895	2.365	2.998	3.499
	8	0.706	1.397	1.860	2.306	2.896	3.355
	9	0.703	1.383	1.833	2.262	2.821	3.250
	10	0.700	1.372	1.812	2.228	2.764	3.169
	11	0.697	1.363	1.796	2.201	2.718	3.106
	12	0.695	1.356	1.782	2.179	2.681	3.055
	13	0.694	1.350	1.771	2.160	2.650	3.012
	14	0.692	1.345	1.761	2.145	2.624	2.977
	15	0.691	1.341	1.753	2.131	2.602	2.947
	16	0.690	1.337	1.746	2.120	2.583	2.921
	17	0.689	1.333	1.740	2.110	2.567	2.898
	18	0.688	1.330	1.734	2.101	2.552	2.878
	19	0.688	1.328	1.729	2.093	2.539	2.861
	20	0.687	1.325	1.725	2.086	2.528	2.845
	21	0.686	1.323	1.721	2.080	2.518	2.831
	22	0.686	1.321	1.717	2.074	2.508	2.819
	23	0.685	1.319	1.714	2.069	2.500	2.807
	24	0.685	1.318	1.711	2.064	2.492	2.797
	25	0.684	1.316	1.708	2.060	2.485	2.787
	26	0.684	1.315	1.706	2.056	2.479	2.779
	27	0.684	1.314	1.703	2.052	2.473	2.771
	28	0.683	1.313	1.701	2.048	2.467	2.763
	29	0.683	1.311	1.699	2.045	2.462	2.756
	30	0.683	1.310	1.697	2.042	2.457	2.750
	40	0.681	1.303	1.684	2.021	2.423	2.704
	50	0.679	1.299	1.676	2.009	2.403	2.678
	100	0.677	1.290	1.660	1.984	2.364	2.626
	∞	0.674	1.282	1.645	1.960	2.326	2.576

Appendix B

Standard Operating Procedures (SOPs) for Hair Conditioner Productions

Example of standard operating procedures for SCAC hair conditioner production

Appendix B: Example of standard operating procedures for SCAC hair conditioner production

Standard Operating Procedures for SCAC Hair Conditioner Production

1. Objective:	**To explain raw material mixing procedures**
2. Scope:	**Raw material mixing process of SCAC hair conditioner production**
3. Practitioner:	**Mixing operators**

4. Personal protective equipment:

4.1 Safety goggles

4.2 Mask

4.3 Ear plugs

4.4 Safety shoes

5. Raw material mixing procedures:

Step	**Raw material**	**Description**	**Control process parameter**					
			Temperature (°C)	**Turbine speed (rpm)**	**Side sweeper speed rpm)**	**Recirculating pump speed (rpm)**	**High shear rate pump speed (rpm)**	**Raw material weight (kg)**
1	**Hot chlorinated water**	**<u>Side mixer no. 1</u>** **1.1 Add hot chlorinated water** **1.2 Turn on side sweeper** **1.3 Turn on heating system** **1.4 Check temperature at 75 °C**	75	-	××××	-	-	××××
2	**Fatty alcohol type 1**	**<u>Side mixer no. 1</u>** **2.1 Add fatty alcohol type 1**	-	××××	××××	-	-	××××

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix B-1: Example of standard operating procedures for SCAC hair conditioner production (cont.)

Step	Raw material	Description	Control process parameter					
			Temperature (°C)	Turbine speed (rpm)	Side sweeper speed rpm)	Recirculating pump speed (rpm)	High shear rate pump speed (rpm)	Raw material weight (kg)
	Fatty alcohol type 2	**2.2 Add fatty alcohol type 2**						××××
		2.3 Turn on turbine						
	Preservative type 1	**2.4 Add preservative type 1**						××××
		2.5 Homogenize mixing for 15 minutes						
		2.6 Turn off heating system						
		2.7 Drain steam from the side mixer jacket						
		2.8 Adjust turbine and side sweeper speeds	-	××××	××××	-	-	
		2.9 Wait to transfer to main mixer						
3	**Hot chlorinated water**	**<u>Main mixer</u>**	××××	××××	××××	××××	-	
		3.1 Add hot chlorinated water						700.00
		3.2 Turn on side sweeper						
		3.3 Turn on turbine						
		3.4 Turn on recirculating pump						
		3.5 Turn on heating system						
		3.6 Check temperature at ×××× °C						
4	**Cationic surfactant type 1**	**<u>Main mixer</u>**	××××	××××	××××	××××	××××	
		4.1 Add cationic surfactant type 1						××××
		4.2 Dissolve cationic surfactant type 1 for 10 minutes						
		4.3 Turn on high shear rate pump						
		4.4 Homogenize mixing for 5 minutes						
		4.5 Turn off high shear rate pump						

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix B-1: Example of standard operating procedures for SCAC hair conditioner production (cont.)

Step	Raw material	Description	Control process parameter					
			Temperature (°C)	Turbine speed (rpm)	Side sweeper speed rpm)	Recirculating pump speed (rpm)	High shear rate pump speed (rpm)	Raw material weight (kg)
	Cationic surfactant type 2 and 3	4.6 Turn off heating system	-	××××	××××	××××	-	
		4.7 Drain steam from the main mixer jacket						
		4.8 Add cationic surfactant type 2						××××
		4.9 Add cationic surfactant type 3						××××
		4.10 Homogenize mixing for 5 minutes						
5	Hot chlorinated water	<u>Side mixer no. 1</u>	-	××××	××××	-	-	
		5.1 Transfer mixture to the main mixer						
		5.2 Clean side mixer no. 1 with hot chlorinated water						××××
		5.3 Recycle hot chlorinated water to the main mixer						
		5.4 Turn off side sweeper						
		5.5 Turn off turbine						
6		<u>Main mixer</u>	-	××××	××××	××××	××××	
		6.1 Adjust recirculating pump speed and turn on high shear rate pump						
		6.2 Homogenize mixture for 15 minutes						
		6.3 Adjust recirculating pump speed and turn off high shear rate pump	-	××××	××××	××××	-	
7		<u>Main mixer</u>	××××	××××	××××	××××	-	
		7.1 Start cooling down using cooling water						
		7.2 Cool down the main mixer						
		7.3 Check temperature at ×××× °C						

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix B-1: Example of standard operating procedures for SCAC hair conditioner production (cont.)

Step	Raw material	Description	Control process parameter					
			Temperature (°C)	Turbine speed (rpm)	Side sweeper speed rpm)	Recirculating pump speed (rpm)	High shear rate pump speed (rpm)	Raw material weight (kg)
8		<u>Main mixer</u>	××××	××××	××××	××××	-	
	Ambient-temp water	8.1 Add ambient-temperature water						××××
		8.2 Adjust turbine speed	××××	××××	××××	××××	-	
		8.3 Cool down the main mixer						
		8.4 Check temperature at ×××× °C						
9		<u>Side mixer no. 2</u>	-	××××	-	-	-	
	Chlorinated water	9.1 Add chlorinated water						××××
		9.2 Turn on turbine						
	Acid	9.3 Add acid						××××
		9.4 Dissolve acid solution for 5 minutes						
	Thickener agent	9.5 Add thickener agent						××××
		9.6 Homogenize mixing for 20 minutes						
		9.7 Transfer the thickener agent solution to the main mixer						
		9.8 Clean side mixer no. 2 with chlorinated water						
	Chlorinated water	9.9 Recycle hot chlorinated water to the main mixer						××××
10		<u>Main mixer</u>	< ××××	××××	××××	××××	-	
		10.1 Homogenize mixture for 10 minutes						
		10.2 Check temperature less than ×××× °C						
	Perfume	10.3 Add SCAC perfume						××××
		10.4 Homogenize mixing for 10 minutes						

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix B-1: Example of standard operating procedures for SCAC hair conditioner production (cont.)

Step	Raw material	Description	Control process parameter					
			Temperature (°C)	Turbine speed (rpm)	Side sweeper speed rpm)	Recirculating pump speed (rpm)	High shear rate pump speed (rpm)	Raw material weight (kg)
	Silicone type 1	10.5 Add silicone type 1						××××
		10.6 Homogenize mixing for 5 minutes						
		10.7 Adjust turbine and recirculation pump speeds	<××××	××××	××××	××××	-	
	Silicone type 2	10.8 Add silicone type 2						××××
		10.9 Homogenize mixing for 15 minutes						
		10.10 Adjust recirculation pump speed	<××××	××××	××××	××××	-	
11		Side mixer no. 3	-	××××	-	-	-	
	Chlorinated water	11.1 Add chlorinated water						××××
		11.2 Turn on turbine						
	Additive type 1	11.3 Add SCAC additive type 1						××××
		11.4 Homogenize additive type 1 for 5 minutes						
	Additive type 2	11.5 Add SCAC additive type 2						××××
		11.6 Homogenize mixing for 10 minutes						
		11.7 Transfer the additive mixture to the main mixer						
		11.8 Clean the side mixer no. 3 with chlorinated water						
	Chlorinated water	11.9 Recycle hot chlorinated water to the main mixer						××××
12		Main mixer	<××××	××××	××××	××××	-	
		12.1 Homogenize mixing for 5 minutes						
	Preservative type 2	12.2 Add preservative type 2						××××
		12.3 Homogenize mixing for 5 minutes						

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix B-1: Example of standard operating procedures for SCAC hair conditioner production (cont.)

Step	Raw material	Description	Control process parameter					
			Temperature (°C)	Turbine speed (rpm)	Side sweeper speed rpm)	Recirculating pump speed (rpm)	High shear rate pump speed (rpm)	Raw material weight (kg)
		12.4 Homogenize additive type 1 for 5 minutes						
		12.5 % cationic surfactant and pH inspection						
13		Main mixer	-	××××	××××	××××	××××	
		13.1 Turn on high shear rate pump and adjust recirculation pump speed						
		13.2 Homogenize mixing for 15 minutes						
		13.3 Turn off high shear rate pump and adjust recirculation pump speed	-	××××	××××	××××	-	
		13.4 Viscosity inspection						
14		Main mixer	-	××××	××××	××××	××××	
		14.1 Turn off cooling system						
		14.2 Drain cooling water from the main mixer jacket						
		14.3 Turn off recirculation pump						
		14.4 Turn off side sweeper						
		14.5 Turn off turbine	-	××××	××××	××××	-	
15		Main mixer						
		15.1 Record product weight in the main mixer						
		15.2 Check available space of the storage tank						
		15.3 Transfer the SCAC product to the storage tank						

Remarks: ×××× denotes the confidential data. All of the data are completely presented in the report for the company.

Appendix C

Calculation of Cost Saving

Calculated cost saving product loss in hair conditioner process (US$/year)

= [DPMO before the improvement (ppm) - DPMO after the improvement (ppm)] × [Hair conditioner capacity (tons/year)] × [Hair conditioner production cost (US$/ton)] / 1,000,000

DPMO before the improvement = 61,217 ppm
DPMO after the improvement = 1.57 ppm
Hair conditioner capacity = 5,000 tons/year
Hair conditioner production cost = x US$/batch (3 tons)

Therefore,

Cost saving in terms of product saving

= [61,217 - 1.57 ppm] × [5,000 tons/year] × / 1,000,000

= 306 tons/year

[306 tons/year] × [y US$/ton]

Appendix D

Collecting Data Sheets

Appendix D-1: Collecting data sheet for the multiple linear regression analysis

Appendix D-2: Collecting data sheet for the design of experiments

Appendix D-1: Collecting data sheet for the multiple linear regression analysis

Mixing Form (special version for the multiple linear regression analysis) Page 1

Date of production / /
Time to start

Shift M A N

Product SCAC SCBS SCDR SCLS SCSS
Batch code

Batch type CIP / Continuous
Product of the previous batch
Previous batch code
Weight of remained product in the main mixer kg

% Chlorine of hot chlorinated water %
Weight of mixture after adding hot chlorinated water in the main mixer kg
Coductivity of mixture in the main mixer (Take sample before adding cationic surfactant type 1)

Dissolving of fatty alcohols:
Temp at side mixer no. 1 °C
Time to start dissolving **End**
(when adding fatty alcohol type 1 to side mixer no. 1) (when transferring fatty alcohols to the main mixer)

Dissolving of cationic surfactant type 1:
Temp at the main mixer °C
Time to start dissolving **End**
(when adding cationc surfactant type 1 to the main mixer) (when adding cationic surfactant type 2 to the main mixer)

Homogenization of cationic surfactants and fatty alcohols at the main mixer:
Time to start homogenization **End**

Cooling water temp °C

Cooling down before adding the ambient-temp water in the main mixer:
Time to start cooling down **End**
(before adding the ambient-temp water in the main mixer)
Temp at the main mixer when start cooling down °C
Temp before adding the ambient-temp water °C
Temp of ambient-temp water °C

Weight of the mixtures in the main mixer:
before adding the ambient-temp water kg
after adding the ambient water kg

Cooling down after adding the ambient-temp water in the main mixer:
Time to start cooling down **End**
(when adding the ambient-temp water in the main mixer) (before adding the thickener agent in the main mixer)
Temp at the main mixer after adding the ambient-temp water °C
Temp before adding thickener agent °C
(at 10 minutes after adding the ambient-temp water) (before adding the thickener agent in the main mixer)

Thickener agent preparation:
Time to start to dissolve the thickener agent at side mixer no. 2
Time when adding the thickener agent in the main mixer

Remarks: ______________________________

______________________________ **Data collected by** ______________

Appendix D-1: Collecting data sheet for the multiple linear regression analysis (cont.)

Mixing Form (special version for the multiple linear regression analysis) Page 2

Cooling down after adding the thickener agent in the main mixer:

Time when adding the thickener agent

Temp at the main mixer after adding the thickener agent ℃

(at 10 minutes after adding the thickener agent)

Time when adding the perfume

Temp before adding the perfume ℃

(before adding the perfume in the main mixer)

Final homogenization: **Time to start homogenization** **End**

Time to finish mixing

Sampling time **Time to measure the viscosity**

Hair conditioner characteristics:

Viscosity cP **% Cationic surfactants** **pH**

Remarks: ______________________________

______________________________ **Data collected by** ______________________________

Appendix D-2: Collecting data sheet for the design of experiments

Mixing Form (special version for the design of experiments)

Experiment order ☐

Parameter	Definition	Setting value
X_3	The quantity of hot chlorinated water added in the main mixer (kg)	
X_7	Time used to dissolve cationic surfactant type 1 (min)	
X_8	The temperature at side mixer no. 1 to dissolve fatty alcohols (°C)	
X_{13}	the set-point temperature at the main mixer before adding the ambient-temperature water (°C)	

Date of production	/ /				**Shift**	M	A
Product	SCAC	SCBS	SCDR	SCLS	SCSS	**Batch code**	

Weight of remained product ________ kg

Weight of the mixtures after adding hot chlorinated water in the main mixer ________ kg X_3 = ________ kg

Dissolving of fatty alcohols: Temp at side mixer no. 1 ________ °C X_7 = ________ °C

Dissolving of cationic surfactant type 1:

Time to start disssolving ________ End ________ X_8 = ________ min

(when adding cationc surfactant type 1 in the main mixer) (when adding cationc surfactant type 2 in the main mixer)

Cooling down before adding the ambient-temp water in the main mixer:

Time to start cooling down ________ End ________

(before adding the ambient-temp water in the main mixer)

Main mixer temp when starts cooling down ________ °C Temp before adding the ambient-temp water ________ °C

X_{13} = ________ °C

Hair conditioner viscosity:

(Take 3 samples)

Sample 1 ________ cP

Sample 2 ________ cP

Sample 3 ________ cP

Inspector ____________________

Remarks: ____

____ Data collected by ____________________

Appendix E

Data for the Multiple Linear Regression Analysis

Table E-1: Data for the multiple linear regression analysis

Observation	Observed data of the possible parameters on the deviation of hair conditioner viscosity									
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	0	0	1,207	82.5	70.2	80.5	9	78.5	32	17
2	1	55	1,202	143.7	70.61	81.3	16	78.7	29	15
3	1	49	1,208	148.3	70.55	82.9	17	78.8	30	15
4	1	62	1,210	158.3	70.84	82.7	16	79.3	29	15
5	1	50	1,202	140.5	70.15	80.4	16	78.9	41	15
6	1	37	1,207	125	70.62	81	16	82	25	15
7	1	60	1,202	137.2	70.44	80.4	14	79	35	23
8	1	67	1,200	143.2	70.11	81.9	15	78.3	38	20
9	1	51	1,219	123.7	70.2	82.2	15	80.2	28	16
10	1	83	1,200	172.8	70.1	81	15	78.5	31	21
11	1	46	1,164	143.7	69.94	81.1	15	79.4	46	15
12	0	0	1,173	82.2	71.75	80.9	18	80.5	44	15
13	1	46	1,200	145.3	70.55	79	15	78.9	36	15
14	1	46	1,214	180	70.97	79.7	21	79.8	38	15
15	0	0	1,202	80	71.31	82.7	18	78.5	29	15
16	1	78	1,208	154.2	70.53	81	14	78	30	15
17	1	54	1,201	138.1	70.39	79	11	78	29	18
18	1	59	1,208	147.2	69.93	82.4	18	79.4	35	17
19	1	39	1,194	133.4	69.73	77.8	14	71.6	31	17
20	1	46	1,217	130.4	69.64	79.8	16	79.1	27	18
21	0	0	1,162	83.1	70.1	80	16	78.9	28	17

Table E-1: Data for the multiple linear regression analysis (cont.)

Observation	Observed data of the possible parameters on the deviation of hair conditioner viscosity										Hair conditioner viscosity
	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	
1	77.4	10	27	59.2	42.9	930	10	13	3,279	183	21,780
2	79.8	10	36	56.8	49.8	931	15	21	3,259	184	19,440
3	79.8	10	36	59.2	49.1	928	11	14	3,219	186	19,720
4	78.9	8	39	58.6	49.1	920	13	20	2,979	184	18,840
5	79.3	8	31	59.6	49.5	908	20	24	3,179	169	22,440
6	79	10	33	58.7	50.3	926	18	27	2,979	183	19,000
7	78.5	10	34	59	50	938	17	20	3,099	185	19,080
8	79.8	8	36	59.1	48.9	922	15	29	2,999	196	21,340
9	79.5	10	32	58.9	50.1	927	15	24	3,119	190	18,680
10	78.5	10	33	59.6	48.5	939	13	20	3,008	186	20,000
11	79.1	10	32	58	50	922	16	19	2,989	181	22,920
12	77.1	8	25	58.5	49.7	912	17	15	3,279	178	27,460
13	79.7	8	36	58.6	59.6	936	14	10	3,039	188	21,440
14	77.7	10	32	58.6	50.3	930	11	10	2,979	196	20,360
15	79.8	8	28	57.6	50.7	927	13	10	2,979	162	24,720
16	79	10	37	59	49.2	926	15	16	3,029	182	19,400
17	77.9	10	31	57.8	50	929	15	16	2,999	166	18,180
18	78.7	10	32	57.6	50.1	939	10	12	2,916	187	19,400
19	76.7	10	32	57.2	50.6	915	15	13	2,966	173	26,680
20	79.6	10	32	58	50	922	12	15	2,959	179	21,920
21	79.5	10	44	58.9	48.1	925	10	18	2,916	186	17,120

Biography

Name: Mr. Natthanon Phaiboonsilpa

Date of birth: 16 November 1982

Place of birth: Bangkok, Thailand

Education:

2001-2005 B.Eng. (Chemical Engineering), Second Class Honors,
Department of Chemical Engineering, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang

Industrial trainings:

- Thai Polyethylene Co. Ltd., April - May 2004
- The Unilever Thai Holdings Ltd., May, 2004 – May, 2007

Research grants and awards from:

- The Thailand Research Fund (IPUS fiscal year 2004)
- The Association of International Education (AIEJ) and The Association of Japan Student Services Organization (JASSO) collaborated with Tohoku University (24 Oct – 5 Nov 2004)
- The Thailand Research Fund (TRF-MAG fiscal year 2006)
- Bronze Medal of Outstanding Senior Project Award, 2005, "Increasing of the Production Capacities of Personal Care Products by Total Productive Maintenance" from the Thai Institute of Chemical Engineering and Applied Chemistry

Publications:

1. Anchaleeporn W. Lothongkum, Natthanon Phaiboonsilpa, Kanyanee Seangkiatiyuth , Pichayane Tanprasert and Sumet Rojdamrongkarn, "Significant Parameters on the Deviation of Hair Conditioner Viscosity: Part I. Regression Analysis by the Minitab Program", KMITL Science Journal, vol.6, no.2b, 2006, pp.599-606.

Publications (cont.):

2. Anchaleeporn W. Lothongkum, Natthanon Phaiboonsilpa and Issara Choedkiatsakul, “Statistical Problem Solving Techniques for Strongly Significant Parameters on Product Loss in Hair Conditioner Production”, Industrial & Engineering Chemistry Research, 2007. (submitted)
3. Anchaleeporn W. Lothongkum, Natthanon Phaiboonsilpa, Kanyanee Seangkiatiyuth, Pichayane Tanprasert and Sumet Rojdamrongkarn, “Significant Parameters on the Deviation of Hair Conditioner Viscosity: Part I. Regression Analysis by the Minitab Program” Proceedings of the KMITL International Conference on Science and Applied Science (KSAS) 2006, 8-10 March, 2006, Swissotel Le concorde Hotel, Bangkok, Thailand, p. 46.
4. Anchaleeporn W. Lothongkum, Natthanon Phaiboonsilpa, Suraphol Pattanapongsanon and Charean Areeasanadham, “Loss Minimization in Hair Conditioner Process using Six Sigma”, Proceedings of the Association of Engineering Education in South East Asia and the Pacific (AEESEAP) Regional Symposium on Engineering Education 2007, 14 February, 2007, Kompleks Dewan Kuliah, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia, pp. 238-243.