comparing different instruments for measuring fat content in ground meat

ibrahim al-kaisi

Norwegian University of Life Sciences
Department of Chemistry,biotechnology and food science
Master Thesis 60 credits 2013

DECLARATION OF AUTHORSHIP

I, Ibrahim Al-kaisi, declare that this paper titled “COMPARING DIFFERENT INSTRUMENTS FOR MEASURING FAT CONTENT IN GROUND MEAT” and the work presented in it is my own. I confirm that:

- Where I have consulted the published work of others, this is always clearly attributed.
- Where i have quoted from the work of others, the source is always given.
- I have acknowledged all main sources of halp.
- Where the thesis is based on work done by myself joinly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

Contents

Chapter 3

The sampling (21)

Chapter 4

Theoretical statistic (30)

ACKNOWLEDGEMENTS

First I want to express my gratitude to my supervisor Dr.Trygve Almøy and co supervisor Professor Solve Sæbø for valuable assistance and guidance.

I also want to thank Morten Engen and the worker at AASHEIM KJØTT, Tor Lunder, Steinar schie, Kristian Åsle and Anita Magnusson from STABBURET, Dr.Froydis Bjerke and the workers at ANIMALIA for their kind help.

I want to thank also Dr.Geir Tøgersen from PREDIKTOR, Dr.Geir Stang Hauge from TOMRA, Dr.Jens Petter Wold and Dr.Oliver Tomic from NOFIMA for their kind assistance.

Last but not least, I want to thank my dear wife, Huda and our lovely daughter Tamara for their never ending support, patience and understanding.

ABSTRACT

Knowledge of the fat content in the meat is extremely important subject for the food production industry. Food production's companies like (STABBURET) can save big amount of money by determine some quality attributes. Several fast and nondestructive instrumental methods have been reported, such as use of X-ray, ultraviolet energy, fluorescence, visual light, Raman scatter, infrared energy, radio waves and few moor.

The first part of this thesis shall discuss the multivariate statistic and the calibration methods. I obtained two data sets from NOFIMA. These data sets is taken from the NIR instrument Q-vision 500. In this part I shall try to learn some important principles in modern calibration methods such as PCA, PCR and PLS.

In the second part of this thesis we will emphasis on some methods that have been used by STABBURET, AASHEIM KJØTT, ANIMALIA and EUROFINS. These methods are either x-ray, chemical or Near Infrared Spectroscopic (NIR). One problem with all methods is that we do not get the true value of the fat content, but a prediction depends on some reference methods. This is the reason we have different results from different instruments. The background for our study is that STABBURET has observed some differences in the fat measurements from their instruments (FA DEXA, BUCHI and NIR INRAALYZER) with the fat content value they get from AASHEIM KJØTT. The people in STABBURET confirmed that those differences are always over the measurements from AASHEIM KJØTT. That means, if AASHEIM KJØTT deliver a batch of 800-1000kg grading meat from the category 21% fat content, the people from the laboratory will find that the fat content is 23% or more. In this way STABBURET pay for one fat category, but they get meat with higher fat content.

The most important conclusion in my opinion is that the reference instrument (FOODSCAN) at ANIMALIA is closer to the FA DEXA at STABBURET than Q-monitor at AASHEIM KJØTT. The results from the first and second data collection show that Q-monitor provides results with lower fat content than the other instruments. A new recalibration for Q-monitor according to the results from these two data collection, could improve the performance of this instrument.

CHAPTER 1

INTRODUCTION

The meat industry is one of the hugest industries in the world and in Norway. Most of the meat slaughters prefer to sell the meat in different categories to the food industry. The slaughter workers adjust their cutting to sort the meat into different fat categories depending on their catting scales and experience. This manually process led to the development of automatically measuring instruments, to check the fat content for each batch. There is several companies internationally put emphasis on developing these kinds of instruments. These different instruments supplier lead sometimes to disagreement about the fat content between the meat supplier and food production industry. These disagreements in the fat content can be interpreted to a large amount of money annually. AASHEIM KJØTT and STABBURET have such disagreement about the meat delivered from the first company to the second company. The basic aim of this thesis is to investigate these differences.

Below some introduction for the places which are involved in this thesis and the instruments they have.

1.1 STABBURET AS

STABBURET is a Norwegian food producer founded by (Gunnar Nilsen) in 1943.STABBURET is a part of Orkla Foods, and has ten factories in southern Norway. It sells well-known brands, such Grandiosa, Big One, Nugatti, Fun Light, Idun and Chef.

STABBURET has three kinds of fat measuring instruments. The first and the most important instrument is the FA DEXA x-ray instrument. This instrument analyzes a whole batch (which can be sent from Aasheim). STABBURET has also a laboratory with two instruments or methods to test the fat content, the BÜCHI instrument which is chemical method and the INFRAALYZER instrument which is NIR method.

1.2 AASHEIM KJØTT AS

''AASHEM KJØTT AS was established in 1982 and moved into a new production facility in 2004. AASHEM KJØTT office is in SOLBERGELVA just west of Drammen, and it is a company in constant growth. From being a small family business it has grown to be a professional company with skilled staff with specialist and master craftsman'' (1).

AASHEIM KJØTT gathers slaughtered whole animals both local and imported. AASHEIM KJØTT has production line in which several employees cut these animals to different parts according to the price and the needs of the food production companies like STABBURET.

AASHEM KJØTT AS has a QMonitor NIR instrument which give only the fat content average for the whole batch. AASHEM KJØTT AS sends few samples to the laboratory of EUROFINS to test the accuracy of the QMonitor.

1.3 ANIMALIA

''ANIMALIA is one of the leading academic and development in meat and egg production. ANIMALIA has around 50 highly qualified employees working with professional issues associated with livestock production from the cubicle to the board. ANIMALIA'S users are Norwegian farmers, veterinarians and businesses in meat and egg industry. To these we provide knowledge and skills through e-learning and training, research and development, development and operation of animal control and animal health. Animalia social mission is to provide professional support to Norwegian farmers and Norwegian meat and egg industry to promote the sale of meat and eggs. ANIMALIA does this both through joint initiatives funded by the sales tax and the assignment of individual operators and companies in the industry'' (2).

ANIMALIA has different instruments to test the fat content. In this thesis we will use the FOODSCAN NIR method as a reference method.

1.4 EUROFINS

EUROFINS use The SCHMID-BONDZYNSKI-RATZLAFF (SBR) method calls for acid digestion before liquid-liquid extraction of the sample. According to the EUROFINS web site ''EUROFINS Scientific is an international life sciences company which provides a unique range of analytical testing services to clients across multiple industries. The Group is the world leader in food and pharmaceutical products testing. It is also number one in the world in the field of environmental laboratory services and discovery pharmacology, and one of the global market leaders in agroscience, genomics and central laboratory services. EUROFINS Scientific was founded in 1987 with 10 employees to market the SNIF-NMRâ technology, a patented analytical method used to verify the origin and purity of several types of food and beverages and identify sophisticated fraud not detectable by other methods ''(3).

1.5 DATA SETS

Four data sets have been used in this thesis. The first two of them are for the multivariate calibration part and I borrowed them from Dr. Jens Petter Wold at ANIMALIA. The second two data sets are to use them in the variance analysis.

1. The first data set is 35 samples, for each sample, 15 wavelengths. This data set has been used as a calibration set to create the calibration model. These samples are pork meat from 2% to 80% fat content.

2. The second data set is the test set. This data set is 368 samples, each sample 15 wavelengths. These samples are from pork meat from 5% to 80% fat content.

3. The third data set is the set which is 160 sample units and I selected them by myself from AASHEIM KJØTT. These sample units are from 14% and 21% fat content. The sample units have been analyzed with all the instruments.

4. The fourth data set is 850 sample units and has been selected in cooperation between AASHEIM KJØTT, STABBURET and TOMRA. These sample units from 14% and 21% fat content and have been analyzed with all the instruments except the instrument at EUROFINS.

1.6 THE USED PROGRAMMES

This thesis is written by Microsoft Word and Adobe Reader PDF and $L^{A}T_{E}X$. The calculations are done by R 2.15.2, R commander and Minitab 16 Statistical Software.

1.7 THE MEAN QUESTION

The mean question in this thesis is to compare between Q-monitor instrument at AASHEIM KJØTT and FA DEXA x-ray instrument at STABBURET. The reference instrument would be FOODSCAN instrument at ANIMALIA. We will try to understand the reason of the difference between the measurements from the two first instruments.

CHAPTER 2

INSTRUMENTS AND FAT CONTENT MEASUREMENT METHODS

In this chapter we will emphasis on the instruments and the fat content measuring process for each place involved in this research. We have AASHEIM KJØTT with its own QMonitor (NIR) instrument, STABBURET with three instruments (FA DEXA, BÜCHI and INFRAALYZER), ANIMALIA with its own FOODSCAN (NIR) instrument and EUROFINS with the (SBR) chemical instrument.

2.1 AASHEIM KJØTT AS

AASHEIM KJØTT has a QMonitor instrument which use Near Infrared (NIR) spectroscopy technology. NIR is one of the most important non-destructive analytical techniques in food science. The traditional quality analyses based on wet chemistry like BÜCHI and SBR have several drawbacks, like the time that those methods need (several hours or days), the use of chemicals, destruction of the samples and the physical distance between the process and the analytical instrument. But the biggest difference between the NIR methods and the wet chemical methods that the chemical methods end with testing only few grams from the whole 1000kg batch while the big NIR instruments like QMonitor and Qvision500 can test the surface (as deep as 2cm)of the whole batch. It is much chipper to use NIR instruments than chemical instrument.

''The QMonitor Fat Analyzer is based on a NIR non-contact transflection system, which is an imaging scanner patented and produced by the Norwegian company TOMRA. The QMonitor Fat Analyzer was launched in 2006 and was replaced by the QVision 500 Fat Analyzer in 2010. The QMonitor at AASHEIM KJØTT was installed in 2008.

The QMonitor produces a multispectral image of the meat as it is scanned on a conveyor belt. The spectral data covers a range of wavelengths in the range from 760nm to 1040nm. The principal advantage of this system when compared to other online NIR systems is the fact the light travels through the sample (as deep as 2 cm), increasing the optical path, thus giving more absorbance information when compared to pure reflection. The NIR scanner system makes in-line measurements possible by allowing detection at the illumination meat. Light from the 12 halogen sources is blocked from the detection zone using blackened plates, such that the light reflected from the surface of the meat does not disturb the detection of the illuminated meat where the depth information is present''(4).

2.1.1 THE ANALYSIS PROCESS

The process in AASHEIM KJØTT AS is:

Our sample units' collection starts from AASHEIM KJØTT. The QMonitor instrument is connected to a big meat grinding machine from one side and to a 1000 kg's meat container from the other side. The QMonitor instrument gives us only one measurement of the fat content in the 1000kg's meat container. This instrument provides the estimated average of all the amount of the minced meat

scanned by the instrument. The workers control the fat content in the big meat container according to the wanted fat percentage they want. The workers put in the grinding machine meat with different fat content from the production lines to get the wanted fat content in the big meat container. I will describe all the process in AASHEIM in some points to make this process more understandable.

1. AASHEIM get the slaughtered animals from the slaughterhouse. The parts of those animals are sorted into different categories according to the fat content. They put the sorted meat in 200kg containers (picture 2.1.1) to make ready for the grinding machine.

Picture 2.1.1

2. According to the wanted fat percentage, the workers put the meat container in the grinding machine (picture 2.1.2 and 2.1.3).

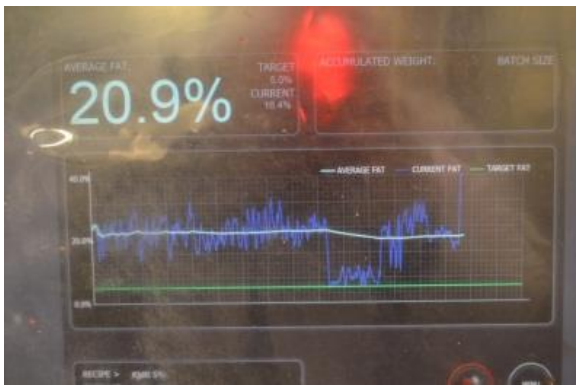
Picture 2.1.2

Picture 2.1.3

3. The minced meats going from the grinding machine throw the QMonitor (picture 2.1.4 and 2.1.5). The QMonitor scan the meat and the control screen shows the cumulative fat content until that moment (picture 2.1.6). In this (figure 6),we can see that the white line in the middle is the average line, while the blue line which move up and down is the measurement from the QMonitor. We can also see that the total average for the whole batch is (20.9%).

Picture 2.1.4

Picture 2.1.5



Picture 2.1.6

4. After the scanning the meat delivers to the 1000kg's container (picture 2.1.7). There the meat be blended and temperature be reduced by inject it with CO_2.

Picture 2.1.7

5. From AASHEIM the meat is sent to STABBURET in two forms (fresh and frozen) and in different fat categories. (In this thesis we will emphasis on 14% and 21% frozen form, only). The fresh meat be sent in 400kg's plastic boxes (picture 2.1.8) while the frozen form be send in 20kg's plastic boxes (picture 2.1.9 and 2.1.10).

Picture 2.1.8

Picture 2.1.9

Picture 2.1.10

2.2 STABBURET

STABBURET has its own fat analyze processes which is depending on chemical (BÜCHI), NIR (IFRAALYZER 2000) and x-ray method (FA DEXA).

2.2.1 FA DEXA

The EAGLE Fat Analysis systems (picture 2.2.1) analyze up to a (28kg) boxes or bulk (120US tons/hour) for chemical lean (CL) value within a 1% margin of error. Unlike traditional sampling methods, the EAGLE Product Inspection DEXA (dual energy x-ray) technology analyzes 100% of the throughput. Cartons can also be tracked by bar code for easy batching and traceability. In addition to the fat analysis, the FA system check weighs and uses the SimulTask software to inspect for contaminants such as: metal, glass, stone bone and wire. The most important specifications for this instrument (6):

1. Inspects 100% of throughput with better than 1% CL measurement accuracy
2. Line speeds up to 30 cartons per minute

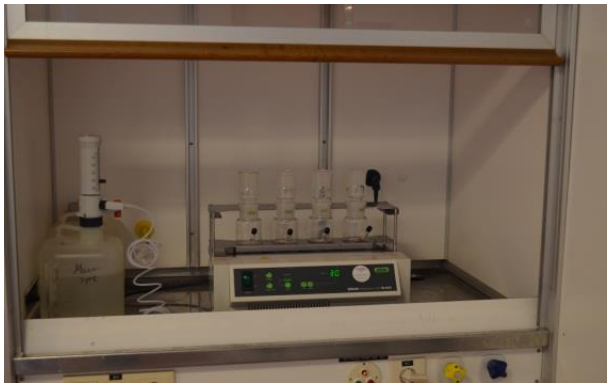
3. FA system reads barcodes from each carton providing the ability to catalog each carton based on CL value
4. Network Capable
5. Multiple inspection capability combined with Fat Analysis provides a fast Return on Investment

Picture 2.2.1

2.2.2 BÜCHI

The second instrument in STABBURET is the BÜCHI B-815. This instrument is wet chemical method. This kind of methods ends with test only few grams from each 1kg sample.

When it comes to the determination of fats in food and feed, the Swiss company BÜCHI plays a substantial role. The fat determination system consists of a sample preparation unit (B-815,Büchi,Switzerland) with four digestion positions with integrated optimized heating and magnetic stirring program (picture 2.2.2). GERSTEL MAESTRO software integrated reporting tool tailored for fat analysis provides detailed sample information (picture 2.2.3).



Picture 2.2.2

Picture 2.2.3

2.2.3 INFRAALYZER

The third fat determination method at STABBURET is the INFRAALYZER (NIR) instrument (picture 2.2.4). ''Both quantitative and qualitative NIR analysis is based on the interaction of Near Infrared light with the molecules of a sample. The sample is irradiated with Near Infrared light of specific wavelengths, selected from up to 19 high-precision interference filters. The penetrating light is partly absorbed by the sample and partly reflected. The wall of the gold-plated integrating sphere contains a lead sulphide detector which measures the reflected portion of the incident light. The resulting spectral information is characteristic for the product and allows a determination of product-relevant properties.

The gold surface for the automated reference sampling also serves as an internal reference, so guaranteeing high drift stability.

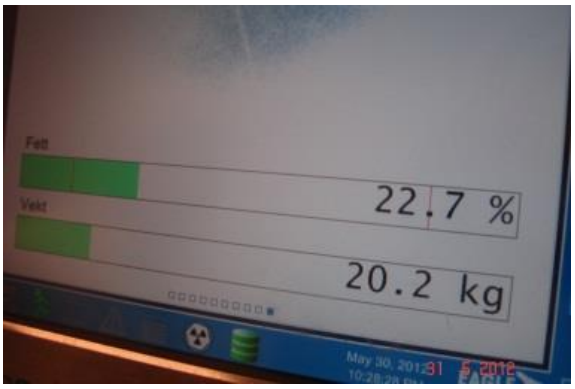
The most important advantages for this instrument are, ''Integrated personal computer and pre-installed, Quanta software allow analysis, calculations and archiving by one system, Electronics module uses digital signal processing for extremely fast analysis (analysis time down to a few seconds), PIOTA precision optics with low noise and automatic reference sampling, Dialog display for easy operation, Only one analytical system required as all properties are determined simultaneously, Environmentally friendly analysis with no chemicals, Direct application at the production site for frequent measuring, Proven calibrations for optimum product quality and Auto-diagnostics for trouble-free system operation'' (8).

Picture 2.2.4

2.2.4 THE FAT MEASURING PROCESS AT STABBURET

1. STABBURET receives the 20kg's frozen boxes from AASHEIM. The next step is to send these boxes to the FA DEXA x-ray to be scanned (picture 2.2.5). For each 20kg frozen box, the instrument give us the fat content and the weight (picture 2.2.6)

Picture 2.2.5



Picture 2.2.6

2. The next step is to select randomly one box from the whole 800-100kg batch. The meat has to be thawed this box and take 4 samples 1kg each (picture 2.2.7). These samples are send to the laboratory for the analyze process.

Picture 2.2.7

3. In the laboratory, the 4 samples have to be homogenized with meat blender (picture 2.2.7).

Picture 2.2.8

4. From those homogenized sample, the laboratory responsible take some grams from each sample according to specific table (picture 2.2.9) and use the BUCHI method to test the fat content in those samples. The BUCHI instrument is chemical method. Four replicates been taken from each sample (picture 2.2.10 and 2.2.11).


VEIETABELL FETT

Forventet Fett %	PRØVE VEKT	INTERN STANDARD
100 %	0,7 gr.	0,2 - 0,25 gr
50 %	1,2 gr.	0,2 - 0,25 gr
40 %	1,5 gr.	0,2 - 0,25 gr
30 %	2,0 gr.	0,2 - 0,25 gr
25 %	2,5 gr.	0,2 - 0,25 gr
20 %	3,0 gr.	0,2 - 0,25 gr
15 %	3,5 - 4 gr.	0,2 - 0,25 gr
10 %	4 - 5 gr.	0,2 - 0,25 gr
<5%	4 - 5 gr.	0,1 - 0,125 gr

Picture 2.2.9

Picture 2.2.10

Picture 2.2.11

5. The Laboratory in STABBURET has a NIR instrument (INFRAALYZER) to analyze the homogenized sample units. This instrument is used on sample units without replicates, so we obtain one value from each sample (picture 2.2.12 and 2.2.13).

Picture 2.2.12

Picture 2.2.13

6. The laboratory has a reference sample. Those reference samples been bought from Sweden and they use them to check the accuracy of the BÜCHI method (picture 2.2.14).

Picture 2.2.14

2.3 ANIMALIA

ANIMALIA use FOODSCAN LAB instrument, this method is some kind of NIR. This instrument uses near-infrared spectral energy to illuminate the sample. By measuring the energy reflected off (or passing through) the sample, chemical information and composition may be determined. This information may be used for quantification of constituents. Both STABBURET and AASHEIM KJØTT were agreed to use an instrument from third place as a neutral reference. This third instrument was the FOODSCAN (NIR) at ANIMALIA.

FOODSCAN Meat Analyzer (picture 2.3.1) is a fast, accurate and easy to use instrument for analyzing all stages of meat production - from checking incoming raw material to final product control. It is pre-calibrated to analyze all key parameters including fat analysis, moisture analysis, protein analysis, salt analysis and collagen analysis delivering results in just 50 seconds.

Picture 2.3.1

2.3.1 THE FAT ANALYSIS PROCESS AT ANIMALIA

The process in ANIMALIA is as the fellow:

1. The sample should be homogenized with the blender until become fine chopped mixture (picture 2.3.2).

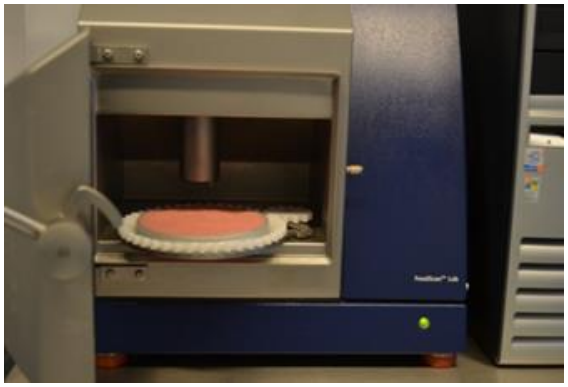
Picture 2.3.2

2. This chopped mixture should be placed in a circle form with 10cm diameter (picture 2.3.3). Empty all the air which is between the form and mixture (picture 2.3.4).

Picture 2.3.3

Picture 2.3.4

3. The last step to analyze the sample unit is to place the form inside the (FOODSCAN LAB) and to wait some second to get the results (picture 2.3.5). This machine gives different values for the fat content, protein, water, and etc. (picture 2.3.6).



Picture 2.3.5

Picture 2.3.6

2.4 EUROFINS

AASHEIM KJØTT AS sends some sample to EUROFINS to test them with the wet chemical (SBR) instrument which is much like the wet chemical instrument from BÜCHI. In this experiment, all the

sample units which have been tested with the other instruments were tested also with the (SBR) instrument.

With the SCHMIID-BONDZYNSKI-RATZLAFF (SBR) instrument, ‘’the sample is boiled with hydrochloric acid to break down triacylglycerols, phospholipids, lipoproteins, glycoproteins and protein, and is then extracted with of diethylether-petroleum ether. The solvents are evaporated and the extracted fat is weighted.

With this method practically all the lipid material in the sample will be extracted, but sometimes non-lipid material is extracted as well. This method therefore tends to give high results for total fat. The method is used widely in Scandinavia and almost now here else ‘’ (9).

CHAPTER 3
THE SAMPLING

The objective of a sample survey is to make an inference about the population from information contained in the sample. The sampling method is the most important process in experiments like the experiment we have in this thesis. One reason is that for some reference methods like the BÜCHI method and the SBR method, the amount of the meat we use is only few grams from the whole batch. Those chemical methods can be the calibration method for the big NIR or X-ray instrument like QVision and FA DEXA.

In this chapter we will discuss sampling theory and the sampling plan for the experiment. The most ideas is taken from (10.Richard L. Scheaffer, William Mendenhall lll, R. Lyman Ott, ELEMENTARY SURVEY SAMPLING, FIFTH EDITION)

3.1 THE SAMPLING THEORY

Needless to say, increasingly accurate and precise analytical methods will not improve on poor sampling technique, nor will it give better estimation for sample quality. This make that the Sampling is the most important technique to make good calibration equation with highly predictable performance. The calibration sample that is selected to make a calibration equation as all the kind of samples should contain representative characteristics of the population for unknown samples that will be predicted by NIR spectroscopy.

3.1.1 GRINDING AND HOMOGENIZING

The selected samples from AASHEIM KJØTT were grinding meat from grinding machine with 10-12mm holes. We get better samples if the holes are smaller, because we reduce the variance between the meat pieces. If we take 1kg sample, the variance in fat content inside the sample will be much smaller.

The homogenizing is very important for all the chemical methods and the NIR spectroscopy hence the experimenter should homogenize the samples as much as possible. But the NIR spectroscopy instruments are sensitive to the temperature. For this reason we have to find the balance between the homogenizing level and the correct temperature e.g. the experimenters in ANIMALIA use the meat blander for 30-45 seconds to get good homogenized samples and with keep the meat in almost the same temperature (0-4C^{o}).

3.1.2 ELEMENTS OF THE SAMPLING PROBLEM

In this part we will discuss some definitions for the sampling problem elements (10. ELEMENTARY SURVEY SAMPLING).

1. The population is a set of measurements, finite or infinite and it is a collection of elements which we want to make inference about. It is important task for the experimenter to define the population carefully and completely before collecting the samples.

2. An observation is an object in the sampling survey. It could be an animal in a farm or a voter in the community.
3. The sampling units are nonoverlapping collections of observations from the population that cover the entire population. In this thesis the entire population is the all the meat in which be produced and send from AASHEIN KJØTT to STABBURET. While the population that cover the entire population are the two fat categories that we researched in this experiment, namely 14% and 21% cattle.
4. A sampling frame is a list of sampling units.
5. A sample is collections of sampling units drown from a sampling frame.

3.1.3 SOURCES OF ERRORS IN SAMPLING SURVEYS

The data from the population will not match up the data observed in a sample, even if we did the sampling with extreme care and accuracy. The difference between what we observed and the entire population is the sampling error. We can divide the sample surveys errors in this thesis in two major groups:

1. The errors of nonobservation: this kind of errors is connected to the sample elements like the error of coverage or/and error of the nonresponse. The error of coverage arise when the sampling frame does not match up perfectly with the target population, like we cannot for one or another reason to collect sample units from every batch we planned to collect sample units from. While the error of non-response arise in one of two ways. The inabilities to collect the sample units or the sample units for one or other reason are damaged.
2. The errors of observation: Once the sample subject is ready to be measured, there are more error sources in the survey. These are the error due to the experimenter and due to the measurement instrument or the method of the data collection.
   Experimenter and the way he or she understand the sampling plan have a big effect on the way of sample collecting. If the experimenter does not fallow the correct process before, during and after the sample units collecting, then this experimenter is an error source.
   The measurement instrument is another source for the error in the survey. In any measurement sample unit, the unit of measurement should be clearly defined, whether it be centimeter or meter, gram or kilogram or different fat content category. The measurement instrument should be correctly calibrate to be as near as the real value of the measured sample unit.
   The amount of the error due to the experimenter and the measurement instrument is depending on the method of data collection

3.1.4 THE DIFFERENT KINDS OF SAMPLE SURVEY DESING

There are three factors determine our inference making procedure, the size of the sample selected from the population, the amount of the variation in the data and the model. The variation in the data can be controlled by the method of selecting the sample. The process to select the sample called the

sample survey design. Since observations cost money, so a precise estimation for the parameters is needed to select the correct number of observations in the sample survey.

There are several kinds of sample survey design. In this thesis we will discuss three of these methods, simple random sampling, stratified random sampling and systematic sampling.

1. SIMPLE RANDOM SAMPLE:
   This kind of sample design is the basic design. If a sample with the size n selected from a population with the size N in a way that each possible sample of the size n has the same chance to be selected, the sampling procedure is called the simple random sampling and the sample called simple random sample. The basic principle in this design is to draw randomly the wanted number of observations without replacement. We might use our judgment to randomize the sample selection. Usually we use a table of random numbers or the computer to select the samples units. For instance If a sample with size 5 wanted to be selected from a population with size 10, we could write 10 papers, each paper represent one element of the 10 elements, and then we have to mix the 10 papers and draw 5 papers one by one without replacement.

2. STRATIFIED RANDOM SAMPLING:
   In this kind of sample design, we separate the population into nonoverlapping groups, called strata, and then select a simple random sample from each stratum. Since the purpose of sample survey design is to maximize the amount of information for given cost. Separating the population in several strata increases the quantity of information for given cost.
   We can put the principal reasons to use the stratified random sample in three main reasons. First, stratification may produce a smaller bound on the error of estimation than the error produced by the simple random sample, especially when the measurements within strata are homogeneous. Second, stratification of the population element into convenient grouping may reduce the cost per observation. Third, we may find the estimates of the parameters for the subgroups of the population. These subgroups should then be identifiable strata.
   To draw a stratified random sample, we should first clearly specify the strata, and then every sampling unit of the population is placed into its appropriate stratum. The second step is to select simple random sample from each strata.

3. SYSTEMATIC SAMPLING:
   Both simple random sample and stratified random sample require very detailed work in the sample selection process. Sampling units on those two methods must be numbered or identified so that randomization device such as a random number table or computer can be used select specific units for the sample. The sample survey design which is widely used primarily because it simplifies the sampling selection process is called systematic sampling. "A sample obtained by randomly selecting one element from the first *k* elements in the frame and every *k*th element thereafter is called a 1-in-k systematic sample, with a random start"(9).
   There are two reasons that systematic sampling provides a useful alternative to simple random sampling. The first reason is that the systematic sampling is easier to practice in the field and hence less error can be done by the field workers than either simple random samples or stratified random samples, especially when we do not have a good frame. The

second reason is the systematic sampling can provide greater information per unit cost than both simple random sample and stratified random sample.

The basic idea with the systematic sampling is that we select the sample units at equal interval. The process to select the systematic sampling is to select randomly one element from the first *k* elements and then select of every *k*th thereafter.

To select a systematic sample of *n* elements from a population of size *N*, *k* must be less than or equal to *N/n* that means (*k≤N/n)*.

3.1.5 SAMPLE SIZE

To decide the size number we need first of all to think about the confidence interval for the sample mean. This confidence interval is defined by:

$$\bar{y} \pm t_{\alpha/2,df} \cdot \sqrt{\frac{\widehat{Var}(\bar{y})}{n}}$$

$$Length = L$$

$$\Rightarrow n = \frac{2.t_{\alpha/2,df}.\widehat{Var}(\bar{y})}{L^2}$$

So to decide the sample size we have to decide the length that we want and the significant level. In this thesis the representatives from AASHEIM KJØTT or STABBURET has to decide the length in fat content before we decide the sample size.

3.2 THE SAMPLE PLAN IN THIS THESIS

3.2.1 THE FIRST ROUND

In the first part of this thesis we discussed the difference in the measurements between Q-monitor at AASHEIM KJØTT and the instruments at STABBURET with using the FOODSCAN instrument at ANIMALIA as a reference instrument. The most correct way to do this comparison is to select sample units and analyze these units with all the instruments. It is an important task to select the most representative sample units especially with the mechanism of the Q-monitor instrument. As I represented in chapter (2.1), the Q-monitor returns only the average for the cumulative fat content for the whole 1000kg batch. The mechanism of the Q-monitor leads to different levels of fat content because the wanted fat content should be controlled by adding meat with different fat content. That means at meat with various fat content levels be sent to the big container in the end of the scanning process and even if the mixture blended, still we get big variation between the sample units. In such situation good sampling method should reduce this variation between the sample units. I visited AASHEIM KJØTT AS three times in the 17.24.31- January-2013. Each visit I selected sample units from two fat categories 14% and 21% and one batch from each fat category. Totally I selected sample units from six batches.

The sample survey design which has been used in this part is a systematic sample design to select the sample units. From each 200kg coming from the big container in the end of the Q-monitor's process,

I toke some sample units. That this I toke one unit from every 30kg-50kg coming out from the big container.

On the 17th of January-2013, I visited AASHEIM KJØT AS for the first time to select the first sample units. We selected two batches which should been sent to STABBURET the next day. The first batch was 14% fat content and the second was 21% fat content. The two measurements from the Q-monitor instrument were 14.1% and 20.9% (it is difficult to obtain the accurate wanted fat category because the fat content has to be adjusted manually by adding meat with lower or higher fat content to obtain the wanted fat level).

After finishing the scanning process for the whole batch, the 1000kg container should be empty gradually. The process is to empty the meat in 200kg metal container and then the container have to be send to the workers to divide the meat amount in 20kg plastic boxes (picture 3.2.1). Those plastic boxes should be frozen and send to STABBURET, where they would be scanned with FA DEXA and I marked the batch with red label with the batch number, the fat category and the date.

From each 200kg container, I took 12 sample units, each unit is 1kg. Four units to be analyzed at ANIMALIA, four units to be analyzed at EUROFINS and the last four to be send to STABBURET. By this way, I had 60x1kg units, 20 units for every place of the three places (ANIMALIA, EUROFINS and STABBURET).

I had the units for both ANIMALIA and EUROFINS in small nylon bags and I marked them with labels which have unit's number, date and batch number (picture 3.2.2) while the units for STABBURET were together in one 20kg box and I had a yellow label with the butch number, the fat category and the date.

Picture 3.2.1

Picture 3.2.2

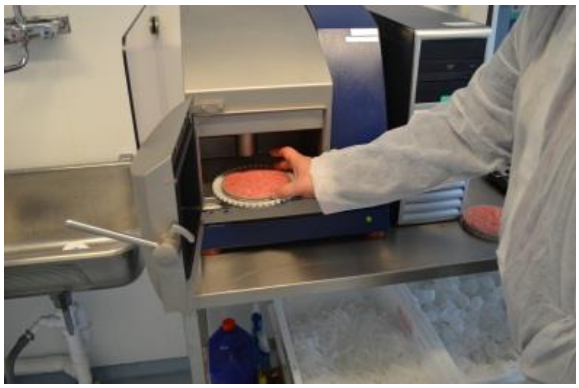
In the end of the sampling process I toke the sample units which are for ANIMALIA with my in two isolated boxes and I left them in the cooling room at ANIMALIA in Oslo to come the next day to analyze the sample units with FOODSCAN instrument. The next day, I analyzed the 40 units (20 from 14% and 20 from 21%) following the process at (2.3.1) (picture 3.2.3). After analyzing all the units, I selected the units in one 20kg box for each fat category and I had a green label (picture 3.2.4). These two boxes have been sent to STABBURET to be analyzed by the FA DEXA and in the laboratory.

Picture 3.2.3

Picture 3.2.4

In this stage, I analyzed the units with take two replicates for every unit with odd number that is for the sample units ($1, 3, \cdots, 19$) I took two 100g samples (picture 3.2.5 and 3.2.6) instead for one to be analyzed with FOODSCAN instrument. But we found that the variation between the samples is almost 15 times more than the variation between the replicates, hence we decided to take 30 sample units instead for 20 sample units from each batch and for each place of the three analysis places.

Picture 3.2.5



Picture 3.2.6

At the second visits in the 24th of January 2013, from every batch of the two batches (one batch 14% and one batch 21%), I selected 90 sample units each one 700g, 30 units for every place of the three places (EUROFINS, ANIMALIA, STABBURET). In this stage the sample units with same number are identical for the three places because I took 30 sample units of 3kg and I divided in three parts for the three places. That is the sample unit with number 1 which has been analyzed at ANIMALIA is identical with the sample unit which has been analyzed at STABBURET and EUROFINS. At the third visit I followed the same process in the second visit.

Totally I selected 6 batches from two fat categories 14% and 21%. From the first 14% batch and the first 21% batch, I selected 20 sample units for every place of the three places (ANIMALIA, STABBURET and EUROFINS). From the second and the third 14% batches and the second and the third 21% batches, I selected 30 sample units from each batch and the units been sent to every place of the three places.

The following part show the process in figures:

1.ÅSHEIM:

This figure is to the process at ÅSHEIM.

And I will make this process for meat

From 14% and 21%.I will repeat this

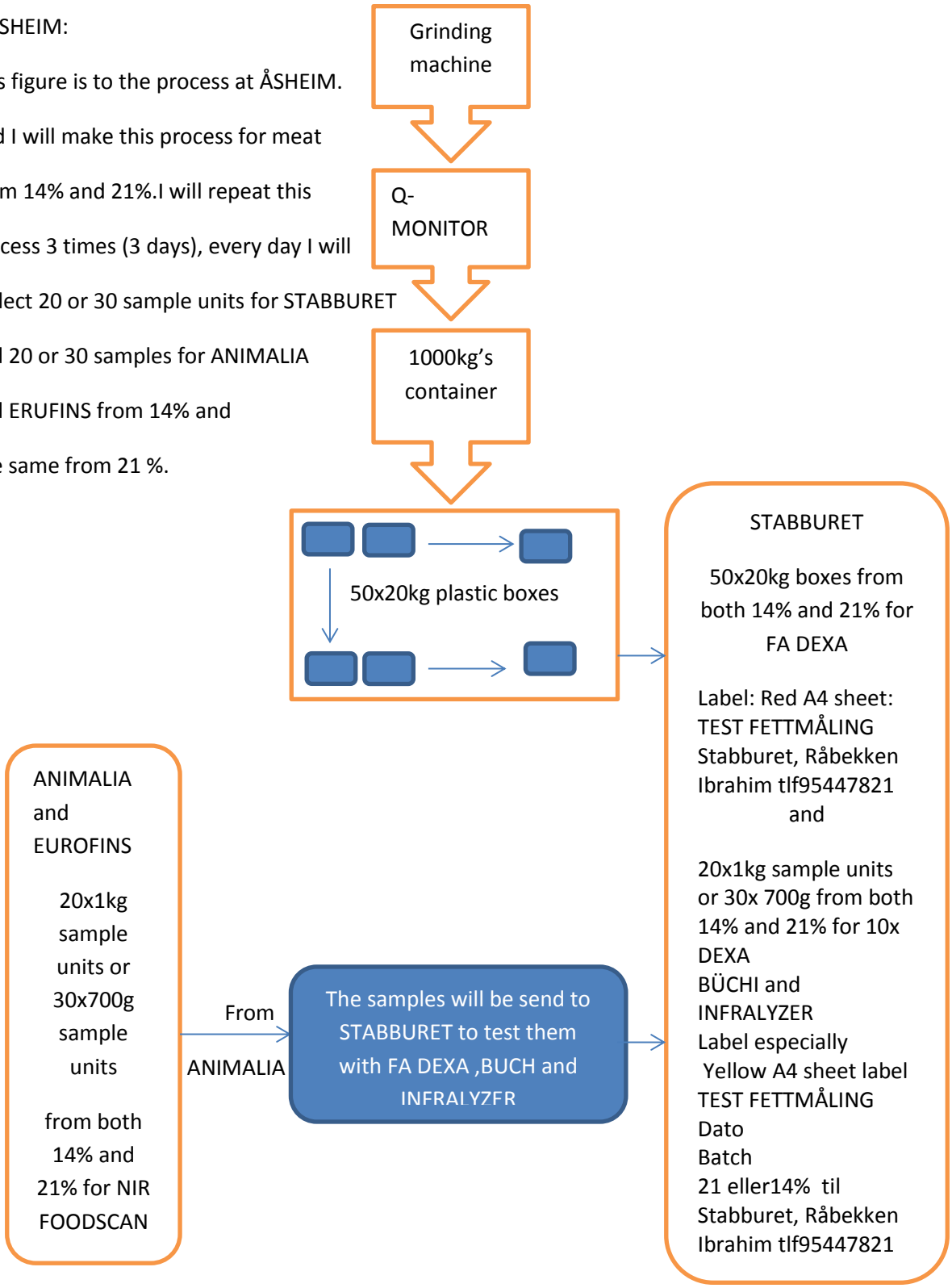
Process 3 times (3 days), every day I will

Collect 20 or 30 sample units for STABBURET

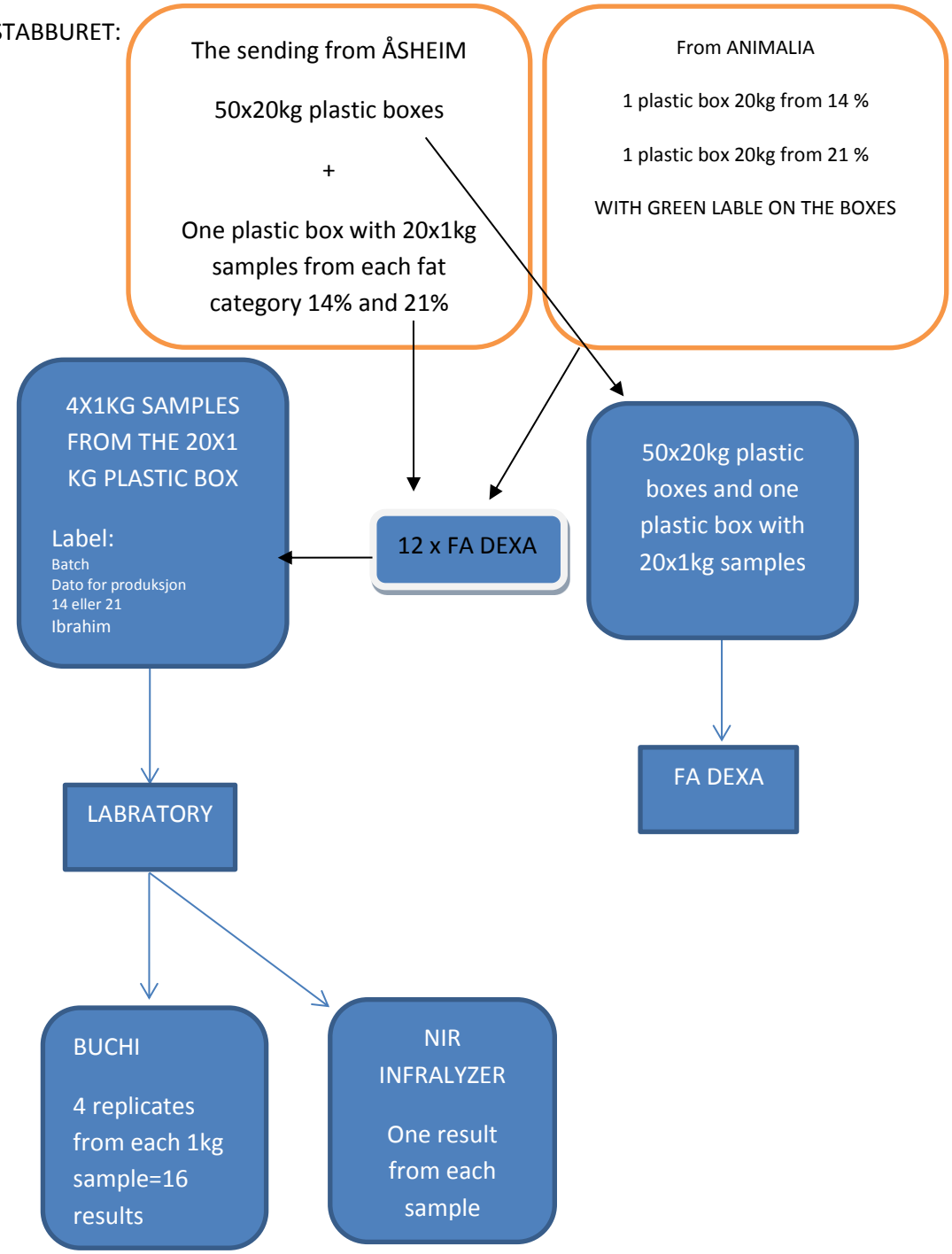
and 20 or 30 samples for ANIMALIA

and ERUFINS from 14% and

The same from 21 %.



2. STABBURET:


The sending from ÅSHEIM
50x20kg plastic boxes
+
One plastic box with 20x1kg samples from each fat category 14% and 21%
From ANIMALIA
1 plastic box 20kg from 14 %
1 plastic box 20kg from 21 %
WITH GREEN LABLE ON THE BOXES
4X1KG SAMPLES FROM THE 20X1 KG PLASTIC BOX
Label:
Batch
Dato for produksjon
14 eller 21
Ibrahim
12 x FA DEXA
50x20kg plastic boxes and one plastic box with 20x1kg samples
FA DEXA
LABRATORY
BUCHI
4 replicates from each 1kg sample=16 results
NIR INFRALYZER
One result from each sample

3.2.2 THE SECOND ROUND

On the 18th of April 2013, I represent the results from the first round in a meeting with representatives from STABBURET, AASHEIM KJØTT and TOMRA SORTING SULUTIONS (the seller of Q-monitor and Q-vision). I represented some tables and figures to explain the variance for each instrument and the strengths and the weaknesses of each instrument.

The conclusion was that six batches not enough to give complete view of the problem, so the decision was to collect new sample units from new batches.

The plan was to collect 50 sample units each unit 400g, so we have 20kg representative sample for each batch. This 20kg package should be homogenized properly be the bowl chopper (picture 3.2.7). The next step is to send four sample units to the laboratory at STABBURET to analyze those samples with INFRAALYZER instrument and BÜCHI instrument, and send five sample units to be analyzed be FOODSCAN at ANIMALIA. The entire batch has been analyzed by FA DEXA.

On the week 21 and 22, sample units from seventeen batches have been taken. Eleven batches were 14% fat content and six batches 21% fat content.

Picture 3.2.7

CHAPTER 4

THEORETICAL STATISTIC

In this chapter I will discuss briefly about some important multivariate statistics principals like the statistical models, estimation, prediction, validation and reduction in dimensions etc. We will discuss Analysis of Variance which we used to analyze my model in addition we will discuss multivariate calibration and its applications.

4.1 MULTIVARIATE STATISTIC

The objective of this thesis is to understand the principles of the multivariate analysis. For this reason we need to speak briefly about important topics like the estimation, prediction, validation, dimension's redaction and Partial Least Square (PLS). The most ideas in this part is taken from PETER J. BICKEL and KJELL A. DOKSUM, MATHEMATICAL STATISTICS (11).

In the modern technology, the instruments provide a large amount of information in short time. The experimenter needs only to push a button to obtain a big matrix of data where each column represent different variable. In the food industry there are a lot such instruments like Q-monitor and FOODSCAN. These instruments use the NIR technology to analyze different elements like fat, protein, water and etc. in the fish, meat, milk and etc.

The importance of multivariate analysis is that it deals with few observations made on many variables. The objective is to study the relationship between the variables or the columns of the data matrix, and how those variables work in combination and use this information to predict new response values.

4.1.1 NOTATIONS

In this thesis, we will use the following notation and syntax:

- Element is given with lowercase letter, e.g. y.
- Vector is given with lowercase bold letter, e.g. $\mathbf{y}$.
- Matrix is given with uppercase bold letter, e.g. $\mathbf{X}$.
- Transposed matrix/vector is given as $\mathbf{X}'$.
- Inverse is given as $\mathbf{X}^{-1}$.
- Parameters are given with Greek letters, e.g. α, β .
- Estimation of parameters and prediction of observations are given with hat symbol, e.g.$\hat{\beta}$.

4.1.2 STATISTICAL MODEL

A statistical model is a mathematical equation which describes the relation between the response $\mathbf{y}$ and one or more explanatory variables. The response could be continuous like height, weight and fat content or categorical like the opinion about different kind of food or people from different regions. In this thesis, we will discuss the continuous situation and our response will be fat content vector.

In general we have the response $\mathbf{y}$ which is a (nx1) vector. This response is a function of $\mathbf{x}$ and a parameter θ and the random noise term. Both the function and $\mathbf{x}$ are known, but the parameter θ is unknown.

$$\mathbf{y}_i = h(\mathbf{x}_i, \theta) + \epsilon_i$$

And we assume that:

$$\epsilon_i \sim N(0, \sigma^2) \ for \ i = 1, 2, \cdots, n$$

All residuals are assumed to be independent.

The simplest situation is:

$$y_i = \mu + \epsilon_i \ for \ i = 1, \cdots, n$$

Here the function of $h(\mathbf{x}_i, \theta)$ is (μ). This is one sample situation. And if we have one explanatory variable ($\mathbf{x}$), then we could have the model:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \ for \ i = 1, \cdots, n$$

We can extend this model to the model with p explanatory variables:

$$y_i = \beta_0 + \beta_1 \mathbf{x}_{1i} + \beta_2 \mathbf{x}_{2i} + \cdots + \beta_p \mathbf{x}_{pi} \ for \ i = 1, \cdots, n$$

The last model could be written in matrix form:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \qquad (4.1.1)$$

Here the response $\mathbf{y}$ is ($n \times 1$) vector. The $\mathbf{X}$ matrix is known ($n \times p + 1$) matrix where p is the number of the columns in this matrix and it represent the number of the explanatory variables, and n represent the number of the rows, hence the number of the observations. The (β) is the unknown parameters ($p + 1 \times 1$) vector. The random noise term (ϵ) is now ($n \times 1$) vector with expected zero $E(\epsilon) = 0$ and variance ($\sigma^2 I$) where (I) is the ($n \times n$) identity matrix. The identity matrix is a matrix which has ones in the diagonal and zeros everywhere else.

The error is the difference between the response vector and the expected of this vector.

$$\epsilon = \mathbf{y} - E(\mathbf{y}) = \mathbf{y} - \mathbf{X}\beta$$

4.1.3 ESTIMATION

The next step is to estimate the parameters in the model namely the parameters vector (β) using the sample units measurements. For correct understanding of the estimation we have to discuss some optimality theory.

4.1.3.1 FOR THE ONE PARAMETER SITUATION

In general we have an estimator $\hat{\theta}$ which is a function of date we observed. The question is how good is ($\hat{\theta}$), this depends on what we mean by the term "good". We have some criteria for estimation. The error or the distance between the estimator and the true value ($\mid \hat{\theta} - \theta \mid$) is one obvious measure of how good the estimator is and it is infinitive to minimize this distance. The problem that (θ) is unknown, the estimator ($\hat{\theta}$) is unknown until we have data and absolute values are difficult computationally. There is no general solution for these problems, but we have to consider what will happen with different values of (θ), we can also make some assumptions regarding the distribution of data ($\mathbf{x}$) and then take expectations, that is we assume e.g. normal distribution and we can have idea about the expected values and the error and we can work with squared errors.

A usual measure used in estimation theory is the mean squared error ($\mathbf{MS}_E$)

$$R(\hat{\theta}, \theta) = E(\hat{\theta} - \theta)^2 = Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2$$

The term $(E(\hat{\theta}) - \theta)^2$ is the squared bias of ($\hat{\theta}$). That is if ($\hat{\theta}$) is unbiased estimator ($E(\hat{\theta}) = \theta$) then:

$$R(\hat{\theta}, \theta) = Var(\hat{\theta})$$

Now if we want to compare between two estimators ($\hat{\theta}_1 \; and \; \hat{\theta}_2$), and the mean squared error of ($\hat{\theta}_1$) is at least as or maximally as big as the mean squared error of ($\hat{\theta}_2$), that is:

$$R(\hat{\theta}_1, \theta) \leq R(\hat{\theta}_2, \theta)$$

For all values of (θ) in the parameter base ($\theta \in \Theta$)

And for some values of (θ):

$$R(\hat{\theta}_1, \theta) < R(\hat{\theta}_2, \theta)$$

Then we choose ($\hat{\theta}_1$) over ($\hat{\theta}_2$), and ($\hat{\theta}_2$) is called inadmissible.

There is no uniformly best estimator which is better than all others in terms of minimizing

$R(\hat{\theta}, \theta)$, such estimator would have to have: $R(\hat{\theta}, \theta) = 0 \; for \; all \; \theta \; values$

That means the mean squared error for the estimator should be equal to zero to count it as a best estimator, but if I select incorrectly a number to be my estimator, this estimator would have mean squared error equal to zero only when this estimator be equal to the true value of the parameter and not correct otherwise.

For this reason we need other criteria to choose between estimators and we need to avoid "foolish" estimators.

The most important criteria are that the estimator should be unbiased. The estimator is unbiased if:

$$E(\hat{\theta}) = \theta$$

And then:

$$R(\hat{\theta}, \theta) = Var(\hat{\theta})$$

The best estimator in this class is the one with the smallest variance. In the class of unbiased estimators for (θ), there may be one estimator which is better than all other estimators. This estimator called the UMVU, which is Uniform Minimum Variance Unbiased.

If we have an estimator ($\hat{\theta}^*$) is UMVU then the variance of this estimator is equal or smaller than any other unbiased estimator. That is:

$$R(\hat{\theta}^*, \theta) = Var(\hat{\theta}^*) \leq Var(\hat{\theta}) = R(\hat{\theta}, \theta)$$

There are three problems for this UMVU estimator. The first problem is the UMVU estimator may not exist. The second problem is the UMVU may be inadmissible, that is biased estimator has smaller mean squared error ($R(\hat{\theta}, \theta)$). The third problem is unbiasedness is not invariant under transformations, if we have:

$$E(\hat{\theta}) = \theta$$

Then:

$$E(q(\hat{\theta})) \neq q(\theta)$$

A good example for this situation is if we have unbiased estimator for the variance (σ^2), the same estimator is not unbiased if we take the square root for this estimator. That is:

$$E(\hat{\sigma}^2) = \sigma^2$$

But:

$$E(\sqrt{\hat{\sigma}^2}) \neq \sigma$$

4.1.3.2 NULL MODEL ESTIMATION AND LEAST SQUARE ESTIMATION

The aim of the estimation methods, is to find estimators for the unknown parameter (β) in the model equation (4.1.1). These estimators are necessary for prediction (we will discuss in the next part). This process is very important because according to the estimation method, we chose to use, we will obtain different prediction error (we will discuss in the next part). There is several kinds of methods to estimate the parameters. We have the null model estimator, the Least Square estimator ($\hat{\beta}_{LS}$), the Principle Component estimator ($\hat{\beta}_{PCR}$) and the Partial Least Square estimator ($\hat{\beta}_{PLSR}$) (PCR and PLS will be discussed later).

The most simple estimation for the (β) is the null model estimation. In this estimator we assume that the parameter is equal to zero ($\beta = 0$). In this case for the centered model, we have:

$$\hat{y} = \overline{y}$$

The next type of estimation is the Least Square estimator. This has been the most common estimator, but it is impossible to apply on the big matrixes from the modern instrument. The least square estimation is defined in the follow equation (Montgomery, Peck and Vining 14).

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

This estimator is unbiased estimator, that is:

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

This method has extremely large variation if we have Multicollinearity problem or impossible to apply if we have ($n < p$) problem. The Multicollinearity problem happens when two or more explanatory variables are highly correlated to each other. We can take a simple example to illustrate this problem. Assume that we have two correlated variable. The variance for his estimator is:

$$Var(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2 \sum_{i=1}^{2} \lambda_i^{-1}\mathbf{e}_i\mathbf{e}_i'$$

Where (λ_i) is the ($\mathbf{i}'th$) eigenvalue for the eigenvector ($\mathbf{e}_i$).

If the two variables are highly correlated then the second eigenvalue will be very small and hence the elements of ($\hat{\beta}$) will have very big variance. That issue makes the estimator useless because if we repeat the experiment twice, will we obtain extremely different values of ($\hat{\beta}$).

The second problem is when ($n < p$) there we cannot find the inverse of the matrix ($\mathbf{X}'\mathbf{X}$) because this matrix has not full rank.

4.1.4 PREDICTION

Prediction is a statement about a random variable in the future depending on experience or knowledge. Prediction is the next step after the estimation. When we decide which estimator we will use, we will use new data matrix ($\mathbf{X}_{new}$) with the selected estimator to predict the observations one by one. For this reason the prediction is related to the estimation and good estimator provides a good prediction.

Generally if we have a parameter (θ) which could be anything ($\mu,\ \sigma^2,\ ect.$). This parameter is telling something about the population and hence the estimation is used when we want to say something about the population, while the prediction is when we want to say something about one individual response. This is the basic difference between the estimation and the prediction. The response and the data matrix are needed to find the model estimator while these estimators and the data matrix needed to predict a future response.

An important question is how good this prediction is. The answer is how close our guesses is to the true ($\mathbf{y}$) values. The average square distance between the guesses and the true values is called the prediction error and this error is defined in the following equation:

$$\theta^2 = E(y - \hat{y})^2$$

The error according to this equation is in square unite like squared centimeter and squared gram, hence if we take the square root for the sides of the equation, the error unit would be more reasonable. The equation would be:

$$\theta = \sqrt{E(y-\hat{y})^2}$$

The best predictor is defined as:

$$\hat{y} = E(y/x)$$

The idea of the best predictor is to minimize the prediction error. The lower bound is:

$$min\ \theta^2 = E(y-\hat{y})^2 = Var(y/x)$$

The prediction error could be separate into three parts, the first part is the variance of the response (y) given the data (x), the second part is the variance of the predictor ($\hat{y}$) given the data (x) and the third part is the squared difference between the expectation for the predictor and the expectation for the response. The third part is the predictor bias. The equation will be:

$$E(y-\hat{y})^2 = Var(y/x) + Var(\hat{y}/x) + (E(\hat{y}/x) - E(y/x))^2 \qquad (4.1.2)$$

The first part of this equation is the independent from the experiment for given model, hence it is the lower bound for prediction error, it is based on the or the nature, while the second and the third part are depend the method we use to estimate the parameters in the model.

If we have a sample with (n) sample units ($y_1, y_2, \cdots, y_n$) which are normally distributed with mean (μ) and variance (σ^2) and we need to predict the element (y_{n+1}). The most logical solution is to use the population mean (μ) if it is known or the sample average ($\overline{y}$) if it is unknown.

If we use the population mean (μ) as predictor, then the prediction error is the variance (σ^2):

$$\theta^2 = E(y-\mu)^2 = \sigma^2$$

If we use a known parameter like the population mean as a predictor ($\hat{y} = \mu$), the second and the third part of the equation (4.1.2) will be zero:

$$E(y-\hat{y})^2 = \sigma^2 + 0 + 0$$

But if we use the sample average as a predictor ($\hat{y} = \overline{y}$), then the equation will be:

$$E(y-\hat{y})^2 = Var(y/x) + Var(\overline{y}) + (E(\overline{y}) - E(y/x))^2$$

This is:

$$E(y-\hat{y})^2 = \sigma^2 + \frac{\sigma^2}{n} + 0$$

In this thesis the response is a vector of fat content observations, and the data matrix is either the light reflections from different light channels with different strength for each channel in the NIR instruments or the lengths of the x-ray waves which gone throw the analyzed meet amount in the FA DEXA instrument.

For this reason we need to understand the matrix form for the prediction error. If we want a linear predictor, then the best predictor using the new ($p \times 1$) data vector ($\mathbf{x}$) is:

$$\hat{\mathbf{y}} = \beta' \mathbf{x}$$

This is what is optimal or what the best predictor is and the lower bound. But if the parameter vector (β) is unknown, we have to estimate this parameter vector. Our predictor now is:

$$\hat{\mathbf{y}} = \hat{\beta}' \mathbf{x}$$

Now we are depending on our sample and we have to add something to our lower bound that is the variance of the predictor and the bias. The prediction error now is:

$$\theta^2 = Var(\mathbf{y}) + Var(\hat{\mathbf{y}}) + (E(\hat{\mathbf{y}}) - E(\mathbf{y}))^2$$

$$= \sigma^2 \mathbf{I} + \mathbf{x}' Var(\hat{\beta}) \mathbf{x} + E(\hat{\beta} - \beta)' \mathbf{x}\mathbf{x}' E(\hat{\beta} - \beta)$$

It is worth mentioning that we do not know the real fat content because each instrument provide prediction of the fat content which is depend on another prediction from some chemical methods or another NIR method like FOODSCAN. In this case the prediction error is:

$$\theta^2 = E(\hat{y} - \hat{\hat{y}})^2$$

After the first instrument calibration, the instrument has to be calibrated from one time to other using some observations from the same instrument to recalibrate itself.

4.1.5 CRITERIA FOR MODEL VALIDATION

According to the different estimators ($\hat{\beta}$), we will obtain different prediction models. There are several methods or criteria to validate these different models. The most common criteria are the Root Mean Squared Error of Prediction ($\mathbf{RMSEP}$) and the ($R^2_{prediction}$).

4.1.5.1 THE ROOT MEAN SQUARED ERROR OF PREDICTION (RMSEP)

We discuss later the prediction error and we defined this error as a squared distance between the observation in the response vector and the prediction corresponding to this observation. The ($\mathbf{RMSEP}$) is the squared root for the sum of the squared difference between the observations and the predictions divide by the number of the sample unites in the response vector, that is:

$$\mathbf{RMSEP}_m = \hat{\theta}_m = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_{im})^2}$$

In this equation, (m) the method has been used to estimate the parameter vector ($\hat{\beta}$) like least square estimation, principal component analysis (PCA), partial least square (PLS) and etc. , (n) is the number of the predicted observations in the response vector ($\mathbf{y}$) and of course is the number of the

predictions in the same time, (y_i) is the value in the row (i) in the response vector and ($\hat{y}_i$) is the predicted value for the same row.

The most important idea is to find the estimator vector ($\hat{\beta}$) which provide predictors as similar as possible to the response observations, that is the prediction model with the smallest distance between the predictors and the response elements is the best model, hence a model with the smallest ($\mathbf{RMSEP}$) is the best model.

4.1.5.2 PRESS and ($R^2_{prediction}$)

The second way to examine the quality of the prediction is ($R^2_{prediction}$). This statistic depends on the prediction error sum of square ($PRESS$). The way to find ($PRESS$) is by cross validation leave one out (will be discussed later).

$$PRESS = \sum_{i=1}^{n} (y_i - \hat{y}_{im})^2$$

$$SS_{Total} = \sum_{i=1}^{n} (y_i - \overline{y})^2$$

The equation which is defined ($R^2_{prediction}$) according to Montgomery (15) is:

$$R^2_{prediction(m)} = 1 - \frac{PRESS}{SS_{Total}}$$

"This statistic gives some indication of the predictive capability of the regression model" (14), that means this statistic gives overview over how much of the variability could we expect this model to explain in predicting new observations.

We can use ($PRESS$) and hence ($R^2_{prediction}$) to compare between two models. A model with small ($PRESS$) value is better than mode with big ($PRESS$) value and hence a model with big ($R^2_{prediction}$) is better than model with small ($R^2_{prediction}$).

It is worth mentioning that there is one to one correspondence between ($RMSEP$) and ($PRESS$) so we do not need to use both of them.

4.1.6 VALIDATION OF PREDICTION QUALITY

After the parameter estimation and using these estimations to predict the new observation, the next step is to validate the prediction model to check the quality of this model. The basic idea is find how good is the parameter estimator ($\hat{\beta}$) by use the entire data matrix (X) and choose a method to estimate this parameter vector and make prediction model to predict new response values. The next step is delete one part of the data matrix and the corresponding response observation or observations and use estimator vector and the remaining data matrix to predict the deleted value or values. For each predicted value, the distance between the deleted response value and predicted

value should be found. When the entire data and the corresponding response values be deleted and new values be predicted and the distance between the response values and the predicted values, these distances can be used to find the Root Mean Square Error of Prediction (RMSEP). The best prediction model provides the smallest (RMSEP) hence we have to choose the estimator vector which provides the smallest (RMSEP) to build the prediction model.

There are several methods to validate the prediction models. In this thesis, we will discuss three of these methods:

1. Leave one out cross validation:
   This is the first and the most simple kind of cross validation technique. A usual, we have a response vector ($\mathbf{y}$) with dimension ($n \times 1$), could be anything (fat content in this thesis) and a data matrix ($\mathbf{X}$) with dimension ($n \times p$) which contain all the explanatory variables and each column is one variable, in this thesis this matrix contain the light reflection for the different light strength lamps .This technique simply is:
   - To leave the first element or observation in the response vector and the first row from the data matrix which are different values for different variables corresponding to the leaved or deleted value of the response.
   - Use the remaining data matrix with dimension ($n-1 \times p$) and the remaining response vector with dimension ($n-1 \times 1$) as training data to find the estimator vector ($\hat{\beta}$). We can use any preferred method to find this estimator vector.
   - Use the estimator vector to predict the leaved or deleted value of the response ($\hat{y}_i$) and find the distance between the deleted value and the predicted value ($e_i = y_i - \hat{y}_i$).
   - Repeat the same process for all the elements in the response vector and for all the rows in the data matrix one by one and until the last row in the data matrix and in the response vector.
   - Find the (RMSEP) and validate the quality of prediction according to this measurement.

   Because that leave one out technique takes almost the entire data matrix, so the bias will be low but the variance can be high Hastie (15).

2. K-fold cross-validation:
   The process for this technique is:
   - In this technique the entire data set have to be divided to equal k folds. Commonly five or ten rows each fold.
   - Use the k-1 as a training data to predict the leaved row. Repeat the process for all the rows in this fold. Find ($RMSEP$) for this fold.
   - Use the same process for all the folds and find ($RMSEP$) for all the folds.
   - To obtain the overall prediction error estimation, the average of the ($RMSEP$) for all folds should be considered.

   In this technique the variance is low but the bias can be a problem.

3. Calibration and test sets:

   The most idyllic situation is to have two data sets one for the calibration and another one for the test. But as usually the experimenter either has not enough economic resources to collect new data set or the research subject is destroyed because of production or weather condition or another reasons. The solution for this problem is to divide the data set into two parts. The first part is the calibration set and the second test set. The basic idea is to use the calibration set to develop a prediction model by estimating first the parameters vector ($\hat{\beta}_{cal.}$) by using different methods to find the method which provides the smallest ($RMSEP_{cal.}$) by the cross validation and compare this estimation with the complexity of the model, that is when we decide the best method to estimate the parameters, we start to add the components one by one and find the level with the smallest cumulative error (16).

   The calibration set model is:

   $$\mathbf{y}_{cal.} = \mathbf{X}_{cal.}\beta_{cal.}$$

   Using the calibration response vector, the calibration data matrix and the preferred estimation method to find the calibration estimator vector ($\hat{\beta}_{cal.}$). The next step is to use this estimator vector with the test response vector and the test data matrix to build the prediction model. This model is:

   $$\hat{\mathbf{y}}_{test} = \mathbf{X}\hat{\beta}_{test}$$

   Use this prediction model and a validation technique like leave one out or k-fold cross validation to find ($RMSEP_{test}$). If we compare ($RMSEP_{test}$) with ($RMSEP_{cal.}$) and the two values are far from each other, so the prediction model is not robust. We cannot use this model to predict new value close to the "true" values of the response vector ($\mathbf{y}$).

4.1.7 REDUCTION IN DIMENSIONS

The dimension's reduction techniques are the most important Principle in the multivariate calibration. We use these techniques when we have many correlated variables or when we have the case of ($n < p$) that is the number of the explanatory variables is much bigger the number of the observations. Many of these variables are correlated and very big part of the variation is in few variables, so it is reasonable to study only the variables which contain the big amount of variation. Sometimes this matrix has many variables and few observations like in case of genetic analysis where we have few observations and thousands of explanatory variables. In this case it is impossible to find the data matrix inverse.

By plotting the prediction error against the model complexity the point with the smallest cumulative error is the point where the curve of the estimation error and the curve of the model error. The figure (4.1.1) (16) illustrates the relationship between the error prediction and the complexity of model. It is clear that as many variables as we add to the model, the estimation error increase and the model error decrease and we select the model with the number of variables which gives the combination of the smallest estimation error and the smallest model error.

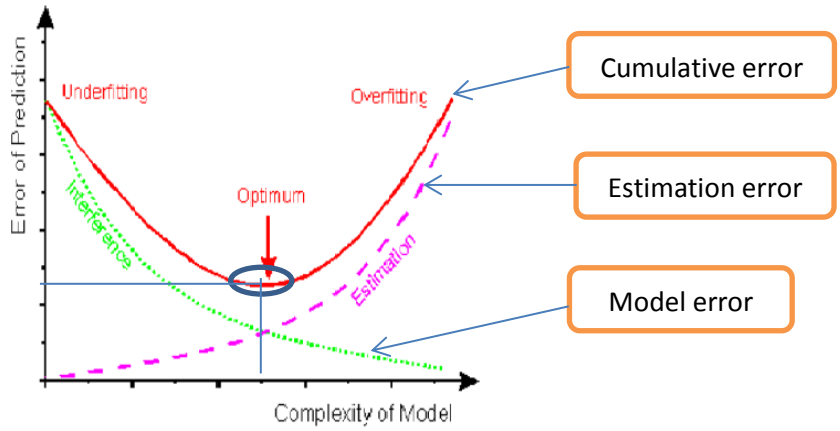


Figure 4.1.1: The relationship between the error prediction and the complexity of model

As we discussed in the previous part, the balance between numbers of the explanatory variables and hence the estimation error and the model error is required to obtain an optimal model. Reducing the number of the explanatory variables makes the entire experiment easier because there are few parameters to estimate and hence much smaller estimation error. By reducing the dimension of the data matrix, the experimenter earns much time and effort. The problem of finding the balance between the prediction error and the model complexity can be solved by several methods. The most common methods are the principal component analysis and regression (PCA, PCR) and the partial least analysis and regression (PLS, PLSR). In the rest of this part, we will briefly discuss the eigenvalues and eigenvectors and the dimension reduction methods.

The basic idea is to reduce the dimension of the data matrix ($\mathbf{X}$) by finding a new matrix (Z) such as:

$$\underbrace{Z}_{n\times q} = \underbrace{X}_{n\times p}\underbrace{R}_{p\times q} \tag{4.1.3}$$

Here we see that we reduced the dimension of the data set from ($n \times p$) to ($n \times q$), where q often much smaller than p. The matrix R differ according to the used method. The next step is to use this reduced matrix to find the Least Square estimator.

It is worth mentioning that sometimes we need to return to the estimator in the data matrix space ($\hat{\beta}_X$). We can find this estimator by using the estimator from the dimension reducing method and the data reducing matrix (R). The equation is:

$$\underbrace{\hat{\beta}_X}_{p\times 1} = \underbrace{R}_{p\times q}\underbrace{\hat{\beta}_Z}_{q\times 1} = R(Z'Z)^{-1}Z'\mathbf{y} = R(R'X'XR)^{-1}R'X'\mathbf{y}$$

The estimator vector here ($\hat{\beta}_Z$) refer to estimator from PCR or PLSR.

4.1.7.1 EIGENVALUES AND EIGENVECTORS

We are interested on the eigenvalues for the covariance matrix of the data matrix ($\mathbf{X}$) which is ($\mathbf{X}^{'}\mathbf{X}/n-1$) when data matrix is centered. Call this matrix ($\hat{\Sigma}$) with dimension ($p \times p$). If we can find a ($p \times 1$) nonzero vector (e) and a scalar (λ) such as:

$$\hat{\Sigma}e = \lambda e$$

So (λ) is eigenvalues and (e) is eigenvector for the covariance ($\hat{\Sigma}$).

The solutions ($\lambda_1, \lambda_2, \cdots, \lambda_p$) of the polynomial equation ($\mid \hat{\Sigma} - \lambda I \mid = 0$) are the eigenvalues for the covariance matrix ($\hat{\Sigma}$).

The main reason that these eigenvalues and eigenvectors are so important is because if we can find a linear combination of the explanatory variables that maximize the variance, that is:

$$max[Var(a^{'}X)] = max(a^{'}\hat{\Sigma}a)$$

That is we want to find (a) that maximize the variance of ($\mathbf{X}$). And we need one restriction of course some restriction on (a) that is ($a^{'}a = 1$). The solution is to use the eigenvector which corresponding to the largest eigenvalues instead for (a), then we have:

$$a = e_1$$

And:

$$Var(e_1^{'}X) = e_1^{'}\hat{\Sigma}e_1$$

From the definition of the eigenvalue and eigenvector:

$$Var(e_1^{'}X) = e_1^{'}\hat{\Sigma}e_1 = e_1^{'}\lambda_1 e_1 = \lambda_1 e_1^{'}e_1 = \lambda_1$$

So we found the linear combination which kept the most variation of the system. If this (λ) is large compared to the trace of the matrix so we can use only one variable instead for the entire matrix.

We can express the total variance by the eigenvalues. If we assume that we have a matrix with the eigenvectors as columns:

$$E = (e_1, e_2, \cdots, e_p)$$

And if we have a diagonal matrix with the eigenvalues on the diagonal and zeros everywhere else:

$$\Lambda = diag(\lambda_i)$$

Then the variance matrix is:

$$\Sigma = E^{'}\Lambda E$$

$$EE^{'} = I$$

Where (I) is the ($p \times p$) identity matrix. The trace of the variance matrix is the total variance and it is the trace for the diagonal matrix (Λ). the trace for diagonal matrix is the sum of the diagonal elements. That is:

$$trace(\Sigma) = trace(\Lambda EE^{'}) = \sum_{i=1}^{p} \lambda_i$$

If we divide any number of (λ) by the total variance ($\sum_{i=1}^{p} \lambda_i$), we obtain how much this or these (λ's) capture from the total variance and according to how much of the variance we satisfy with, we can chose the number of the variables. If we have a data matrix with one thousand variables and only three variables capture more than 90% of the variance, so we need to focus only on three variables and that save a lot of time and effort.

We can use the eigenvalues to indicate if there is multicollinearity in the ($X^{'}X$) matrix by examine the condition number of this matrix, which defined as:

$$k = \frac{\lambda_{max}}{\lambda_{min}}$$

If this condition number is less than 100, then there is no serious problem with multicollinearity. Condition number between 100 and 1000 indicate a multicollinearity problem and condition number exceeds 1000 indicate severe multicollinearity problem.

4.1.7.2 THE PRINCIPAL COMPONENT ANALYSIS (PCA) AND THE PRINCIPAL COMPONENT REGRESSION (PCR)

The principal component analysis and regression is an effective method to reduce the data matrix dimension. This method depends on the eigenvalues and eigenvectors which we discuss previously. The first component should be in the direction of the highest variance and the second should be orthogonal on the first component and so on.

The first process in the principal component analysis is to center the response vector ($\mathbf{y}$) and the columns on the data matrix ($\mathbf{X}$) by subtracting the average of the response from each observation and the average of each column from the observations in this column.

The ($n \times p$) matrix (X) can be expressed and replaced by the matrix of (g) vectors

($Z = [Z_1, Z_2, \cdots, Z_g]$) where $g < p$ (often much smaller) and each vector is with dimension ($n \times 1$). Each vector in this matrix represents principle component and the first component (Z_1) spans in the direction of the highest variation in the matrix ($\mathbf{X}$). The second component (Z_2) is orthogonal on the first component and so on until the entire variation explained by the components. The figure 4.1.2 (18) illustrates the first and the second component.

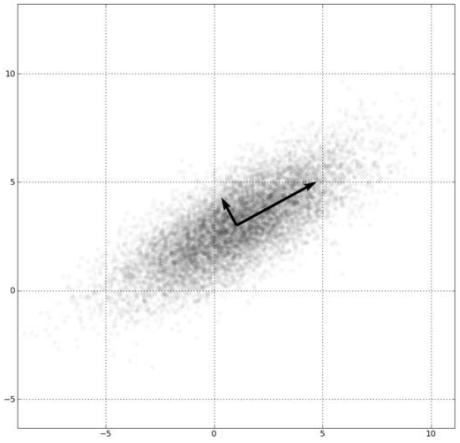


Figure 4.1.2: The first and the second component

The matrix (Z) provide the property that the columns are orthogonal to each other. That is:

$$Z_1^{'}.Z_2 = 0$$

The relation between the centered data matrix ($\mathbf{X}$) and the principal component matrix (Z), can be expressed by the follow equation (19.Harald Martens, Tormod Næs. Multivariate Calibration):

$$X = ZE^{'} + F$$

Where (E) is the loadings matrix which represent the eigenvectors matrix of the covariance matrix ($X^{'}X$). The residuals (F) refer to the unexplained variation in ($\mathbf{X}$). These residuals can be due to measurement mistake, operator mistake etc.

We can express the principal component in the follow equation, (19.Harald Martens, Tormod Næs. Multivariate Calibration):

$$Z_g = XE_g \qquad (4.1.4)$$

The number of the component depends on the experiment and the experimenter. That is the experimenter has to decide self, according to how much of the variability he or she want to explain by this experiment. In prediction, it is cross validation or test set which determines the number of components.

To find the percentage of the variability each component, we take the ratio of the first eigenvalue on the sum of the all the eigenvalues for the matrix ($X^{'}X$). To find how much of the variation represent the first two component, we take the ratio of the sum of the first and the second eigenvalues on the sum of the all the eigenvalues. By the same way, we add the eigenvalues one after one until we capture the entire amount of the variation. This ratio is called the conditional number and it is the ratio between the explained and the unexplained variation of (X) by g number of components. The ratio can be expressed by the follow equation:

$$\psi = \frac{\lambda_1+\lambda_2+\cdots+\lambda_g}{\lambda_1+\lambda_2+\lambda_3+\cdots+\lambda_p}$$

Now we can use these principle components to predict a set of new response observations. The used model called the principle component regression (PCR) which gives us ($\hat{\beta}_{PCR}$). In this model, we construct a new set of variables instead for the original variables in (X) and its keep as much as possible information from the matrix (X) and also hopefully highly correlated to the response vector, let this matrix be the matrix (Z) in the equation (4.1.4).

Now we do the linear regression on the matrix (Z), the model is:

$$\mathbf{y} = Z\beta + \epsilon$$

This model is the principle component regression (PCR). We do not have the constant part because the data matrix is centered.

As we discussed previously, the variance of the estimator is the most determined issue for the performance for this estimator. If we have an estimator with large variance, this estimator will be useless because this it depends on the data hence can provide estimators with large difference from one experiment to other.

The variance of the estimator vector now is reduced but we do systematic error and this is what the experimenter has to "pay" to have estimator with small variance which is the most important issue in the multivariate calibration. That is:

$$E(\hat{\beta}_{PCR}) \neq \beta_{PCR}$$

And $E(\hat{\beta}_{LS}) = \beta_{LS}$ this is the least square estimator.

But $Var(\hat{\beta}_{PCR}) < Var(\hat{\beta}_{LS})$

With another word the least square estimation is unbiased but it has a large variance due to the multicollinearity and the principle component regression estimator is biased but it has small variance and hence good performance in the prediction.

In general the PCR method focuses on the variation in the data matrix (X). If we want to reduce the dimension using the PCR, we need to find a new matrix Z from equation (4.1.3) which has much fewer columns.

The next step is to use the least square method on this matrix. The benefit is to reduce the number of variables from e.g. 100 variables to one or two. So what we did is to remove the correlated variables and reduced the variance of the estimator. In the PCR the (R) is the matrix of the eigenvectors. The estimator vector now is:

$$\hat{\beta}_{PCR} = (Z'Z)^{-1}Z'\mathbf{y}$$

Or $\hat{\beta}_{PCR} = (E'X'XE)^{-1}E'X'\mathbf{y}$

The PCR is mainly prediction method, therefore we use the estimator vector ($\hat{\beta}_{PCR}$) to create a prediction model to predict new observations. This model is:

$$\hat{\mathbf{y}}_{PCR} = X\hat{\beta}_{PCR}$$

We can make different estimator vectors with different number of component. The last step is to find which one of these estimators provides the best predictor model. We use these estimators with different models and for each model, we make cross validation to find the ($RMSEP$) for each model and select to use the model with the smallest ($RMSEP$) value.

4.1.7.3 The Partial Least Square (PLS) and the Partial Square Regression (PLSR)

The second method of the data reducing is the partial least square regression. The basic idea for this method is the same in the principal component regression, which is to find a new smaller set of variables to reduce the dimension of the data matrix (X). The purpose of this method is to find new estimator vector ($\hat{\beta}_{PLSR}$) to build a prediction model.

The main difference between the principle component regression and the partial least square regression is that the (PCR) use the information about the variation only from the data matrix (X), and focus on finding combinations having the largest variance among the variables in this matrix, and ignore the information in the response ($\mathbf{y}$). While the partial least square regression focus on finding the largest covariance between the response ($\mathbf{y}$) and the data matrix (X) that is if we do (PLS), the first component is more or less the covariance between the response and the data matrix and this is the reason that many scientists believe that the (PLSR) is doing better than the (PCR) because the covariance between the variables from the data matrix (X) could be not relevant for the response vector ($\mathbf{y}$) which is the aim of the prediction model. Hence for the new data matrix Z from equation (4.1.3) the difference between the two methods is in the way to find the matrix which reduces the dimension (R).

We can put the process to find the PLS in several steps algorithm. Before starting the process to find the PLS, we have to center the response vector ($\mathbf{y}$) and the data matrix (X) as we described in the PCR part and called the centered matrixes ($\mathbf{y}_0$) and (X_0).

The algorithm according to Martens and Næs (19) is:

1) The first step is to find the covariance between ($\mathbf{y}_0$) and (X_0) and normalize it by divide it by the length this is what we call the loading weights. That is:
$$\hat{w}_1 = X_0^{'}\mathbf{y}_0$$
$$\hat{w}_1 = \frac{\hat{w}_1}{\|\hat{w}_1\|} \text{ (Normalizing the loading weights)}$$
2) The next step is to find the scores by use the data matrix with loadings weights. This scores vector defined as:
$$\hat{t}_1 = X\hat{w}_1$$
3) The third step is to use the scores from the previous step to find the ($\mathbf{y}$-loadings) using the following equation:
$$\hat{q}_1 = \mathbf{y}_o^{'}\hat{t}_1(\hat{t}_1^{'}\hat{t}_1)^{-1}$$
And the (X-loadings) by the following equation:
$$\hat{p}_1 = X_o^{'}\hat{t}_1(\hat{t}_1^{'}t_1)^{-1}$$

4) In this step, we use all the information from the previous steps to find the deflation of ($\mathbf{y}$ and X) by the two following equations:

$$y_1 = \mathbf{y}_o - \hat{t}_1.\hat{q}_1^{'}$$
$$x_1 = X_o - \hat{t}_1.\hat{p}_1^{'}$$

5) To extract the loading weights $[\hat{w}_2, \hat{w}_3, \cdots, \hat{w}_q]$ and the scores $[\hat{t}_2, \hat{t}_3, \cdots, \hat{t}_q]$, we repeat the steps from 1 to 4, $(q-1)$ times to obtain the following matrix:

The loadings weights matrix	$W = [\hat{w}_1, \hat{w}_2, \cdots, \hat{w}_q]$
The scores matrix	$T = [\hat{t}_1, \hat{t}_2, \cdots, \hat{t}_q]$
The X-loadings matrix	$P = [\hat{p}_1, \hat{p}_2, \cdots, \hat{p}_q]$
The Y-loadings matrix	$Q = [\hat{q}_1, \hat{q}_2, \cdots, \hat{q}_q]$

When all these matrixes detected, we are ready to estimate the parameter vector using the PLS method. This estimation is:

$$\hat{\beta}_{PLS} = W(P^{'}W)^{-1}Q$$

Now, we can use this estimator to illustrate the regression depend on the PLS components. This regression is the PLSR. One more time, we can this regression model to create the prediction model depending on the PLSR method to predict a set of new observations. After the prediction process cross validation is needed to validate the quality of this predictor. One more time, we have to find the ($RMSEP$). And like the PCR, we can make different prediction models with different number of component and find the ($RMSEP$) for each model and select the model with smallest ($RMSEP$) value.

4.2 THE ANALYSIS OF VARIANCE (ANOVA)

The analysis of variance (ANOVA) is probably one of the most useful techniques in the field of statistical inference. The purpose of (ANOVA) is estimate effects and to test for significant difference between these. ANOVA provides a statistical test of whether or not the effects of two or more groups are all equal, and therefore generalizes t-test to more than two groups. The most of the ideas in this part are taken from Douglas C. Montgomery, Design and Analysis of Experiments, seventh edition (13).

If we want to analyze the fat content measured by different instrument, we need a model to describe the observations of this experiment.

$$y_{ij} = \mu_i + \epsilon_{ij} \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, n \end{cases} \qquad (4.2.1)$$

Here y_{ij} the $j'th$ observation on factor level i, μ_i is the mean for the factor level or treatment i or group i, and ϵ_{ij} is random error.

This model could be written by using the treatment or group effect. The model in this case would be:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, n \end{cases} \quad (4.2.2)$$

Where the μ is parameter common to all treatments or groups and we can call it the overall mean, while τ_i is a parameter unique to the treatment *i* and it is the effect of treatment or group *i* which make this treatment mean different from the overall mean. This model called the single –factor analysis of variance (ANOVA) because it has only one factor but this factor may have many groups.
The basic idea of the analysis of variance is to divide the total variability into different parts, and investigate if there is any difference between the groups and if so, where are the differences. The total corrected sum of square is:

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 \quad (4.2.3)$$

The sum of square (4.2.3) may be decomposed into:

$$\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 \quad (4.2.4)$$

This sum of square identity is for the balanced experiment and it is for both the fixed and the random effect model (will be explained later). The difference in the assumptions and restrictions will provide different estimators for the model parameters.
The equation (4.2.4) may be written symbolically as:

$$SS_{Total} = SS_{Treatments} + SS_{error} \quad (4.2.5)$$

Degrees of freedom:
$(N-1)$ for $SS_{Total,}$ N is the total number of the observations.
$(a-1)$ for $SS_{Treatment}$ and a is the number of the treatment levels

$a(n-1) = (N-a)$ for SS_{Error}

The degrees of freedom are the number of the independent elements which are connected to the sum of squares.
The equation (4.2.5) means the total variation may be split into two parts the first part is the variation we have in the model due to the effect of the treatments or the groups, and it is the part which is explained by the model.
The second part is the unexplained part, and it is the variation within the treatments which is due to only random error. The equation (4.2.5) provides us two estimators for the variance σ^2, the first based on the variance between the treatments and the second based on the variance within the treatments.
These two estimates are called mean squares and we can formalize them as:

$$MS_{Treatment} = \frac{SS_{Treatment}}{a-1} \quad (4.2.6)$$

And $$MS_{Error} = \frac{SS_{Error}}{N-a} \quad (4.2.7)$$

We have also $$E(MS_{Error}) = \sigma^2 \quad (4.2.8)$$

That is the MS_{Error} is unbiased estimator for the noise variance
The observations within a treatment or group are called replicates.

In the analysis of the variance (ANOVA) or objective will be test hypotheses about the treatments means and to estimate these means. For the hypotheses testing we should have some assumptions. First the model errors are assumed to be normally and independently distributed variables with mean zero and constant variance for all levels of the factor, that is:

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

That is:

$$E(\epsilon_{ij}) = 0$$
$$Var(\epsilon_{ij}) = \sigma^2$$
$$Cov(\epsilon_{ij}, \epsilon_{kl}) = 0 \ (for \ ij \ \neq \ kl)$$

4.2.1 THE F-TEST

In general, if we have two independent Chi-square distributed variables u and v degrees of freedom, then

$$F = \frac{\chi_u^2/u}{\chi_v^2/v} \sim \ F_{u,v} \qquad (4.2.9)$$

Then F follows a Fisher-distribution with u degrees of freedom for nominator and v degrees of freedom for the denominator.

It may be shown that SS_{Tr}/σ^2 and SS_E/σ^2 are two independent Chi-square distributed random variables with (a-1) and (N-a) degrees of freedom under H_0.

From the equation (4.5), we can find the test statistic for the hypothesis no different in treatment means. This test is:

$$F_0 = \frac{SS_{Tr}/(a-1)}{SS_E/(N-a)} = \frac{MS_{TR}}{MS_E}$$

We reject the H_0 if the test statistic is large and conclude that there are differences in the treatments means if:

$$F_0 > F_{\alpha,(a-1),(N-a)}$$

Where: α is the significant level, (a-1) is the degrees of freedom for the SS_{Tr} and (N-a) is the degrees of freedom for the SS_{Error}.

Or we reject if the p-value < the level of significance (α).”the p-value is the probability that the test will take a value that is at least as extreme as the observed value of the statistic when the null hypothesis is true”(12)

4.2.2 MODEL ADEQUACY CHECKING

The model tests and confidence interval depend on the model assumptions. The first assumption that the data adequately described by the model, the second assumption that the error is normally and independent distributed and the third assumption that the variance is constant.

We may check these assumptions by using residual analysis. The residual is the deviation between the observed value and the fitted value.

A fitted value for each observation is:

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i = \overline{y}_{..} + (\overline{y}_{i.} - \overline{y}_{..}) = \overline{y}_{i.}$$

So the fitted value is the treatment level sample mean. And the residual is:

$$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \overline{y}_{i.} \qquad (4.2.10)$$

Through the analysis of residuals, different kind of model inadequacies and violations of the underlying assumptions can be discovered. Therefore a plot like the plot in (figure 4.2.1) which is the plot for one of the models I used in my experiment can help us to investigate if there are some problems with the model fitting.

An extremely useful plot to check the assumption of the normality is the normal probability plot for the residuals. The residuals at this plot will look like a straight line if we have good fit. If this line bends down slightly on the left side and upward slightly on the right side, it is indexing that the largest residuals are not quite as large as expected.

Sometimes we have one or more residuals much larger than any of the others. These residuals are representing what we called outliers observations. This kind of residuals can damage the analysis of the variance therefore if we have one or more of these outliers, we have to investigate carefully the reason back these outliers. Usually the cause of the outlier is a human mistake such mistake in calculations or a data coding or copying error. If there is no such mistake, these outliers can be more informative than the rest of the data. Therefore we have to be careful with these outliers and not throw them out or ignore them unless we have reasonable nonstatistical grounds for doing so.

There are several statistical procedures to detecting outliers. A common way to check for outliers is to use the standardized residual. If the errors ε_{ij} are N (0, σ^2), the standardized residual is:

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}} \qquad (4.2.11)$$

The standardized residuals should be approximately normally distributed with mean zero and variance equal to 1. Hence approximately 68% of the standardized residuals should fall within the limit ±1, approximately 95% of the within ±2, and nearly all should fall within ±3. Standardized residuals bigger than 3 or 4 is a potential outlier.

Another informative plot is the residuals versus fitted values plot. This plot is makes it possible to check the linearity and the constant variance assumptions. If the model is correct and the assumptions are satisfied, the residuals in this plot should be structureless. An obvious pattern is may be index of nonconstant variance. If the error in the experiment was constant percentage of the size of the observations, then the variance increases as the number of the observations increases. In this case the plot of residuals versus fitted values would look like megaphone. We have also problem of nonconstant variance if the data follow a nonnormal, skewed distribution because in the skewed distribution, the variance seem to be function of the mean.

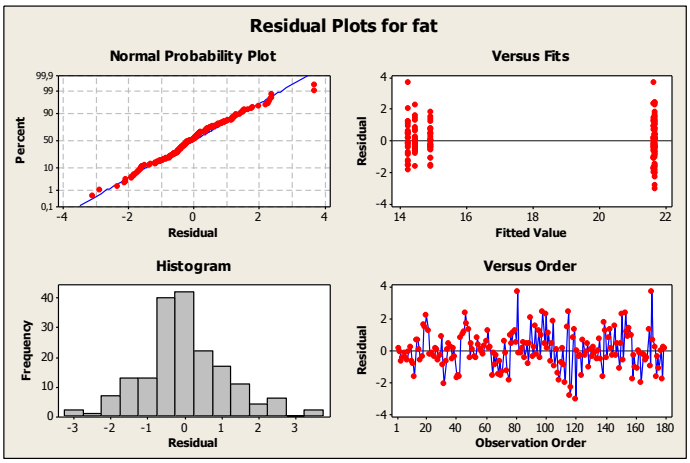


Figure 4.2.1: Residual plots for one of the models in this thesis

4.2.3 FIXED FACTOR

The Fixed effect model is the situation when we chose specific levels of the factor. In this situation we are interested in the effect of these specific levels, and our conclusions will be about only these specific treatment levels. In my experiment the effect of the instrument and the effect of the fat category are fixed because we were interested in the effect of the specific instruments (FA DEXA, SOODSCAN, BÜCHI, INFRAALYZER and SBR) and the specific two fat categories (14% and 21%). In The model (4.2.2), the factor effect τ is fixed factor, and then we want to draw conclusion about only the selected levels of this factor.

Now we have:

$$E(y_{ij}) = \mu + \tau_i$$

Usually we think of

$$\mu = \frac{1}{a}\sum_{i=1}^{a}\mu_i$$

This implies the restriction

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

$$E(MS_{Error}) = \sigma^2 \qquad \text{and} \qquad E(MS_{Tr}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

Under the null hypothesis (see testing fixed effects)of no treatment effect, both MS_{Error} and $MS_{Treatment}$ are unbiased estimates of σ^2. If H_0 is wrong, then $E(MS_{Tr}) > \sigma^2$.

4.2.3.1 TESTING THE FIXED EFFECTS

To test for a treatment effect, we will test that the treatment effects (τ_i) are zero, that is:

$H_0 = The\ means\ are\ equal\ (the\ means\ model)$, or
$H_0 = No\ treatment\ effect\ (the\ effects\ model)$

We can formalize this as:

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a = \mu$ Versus
$H_1 : At\ least\ two\ means\ different\ (the\ means\ model)$, or

$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$ Versus
$H_1 : At\ least\ two\ different\ from\ zero\ (the\ effects\ model)$

4.2.3.2 THE PARAMETERS ESTIMATORS

For the model in the equation (4.2.2) the unknown parameters are: $\mu,\ \tau_1,\ \tau_2,\ \cdots,\ \tau_a\ and\ \sigma^2$ The unbiased estimator for the error variance σ^2 is MS_E. The estimators for the expectation parameters are:

$$\hat{\mu} = \overline{y}_{..}$$

$$\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$$

We have also the mean for the i'th treatment is:

$$\mu_i = \mu + \tau_i$$

The estimator is:

$$\hat{\mu}_i = \overline{y}_{..} + (\overline{y}_{i.} - \overline{y}_{..}) = \overline{y}_{i.}$$

The variance of this estimator is:

$$Var(\overline{y}_{i.}) = Var(\tfrac{1}{n}\sum_{j=1}^{n} y_{ij}) = \tfrac{1}{n^2} n\sigma^2 = \tfrac{\sigma^2}{n}$$

Hence the stander error is $\sqrt{MS_E/n}$

And a (1-α) 100% Confidence Interval for μ_i is:

$$\overline{y}_{i.} \pm\ t_{\alpha/2,(N-a)} \cdot \sqrt{MS_E/n}$$

We can also find a (1-α) 100% confidence interval for σ^2:

$$\left[\ \frac{SS_E}{\chi^2_{1-\alpha/2,N-a}},\ \frac{SS_E}{\chi^2_{\alpha/2,N-a}}\ \right]$$

4.2.4 RANDOM FACTOR

In the fixed factor model we want to make conclusions from hypothesis tests about specific levels of the factor, e.g. two fat categories or some specific fat measuring instruments. If we chose randomly some levels from a population of possible level and we want to draw conclusion for the entire population and not for the chosen levels, then we say that the factor is random and the model called

the components of variance or random effect model. The random factor model is the same model as the fixed model (4.2.2).

The effect factor τ_i is random. The difference with the fixed effect model is in the assumptions. We assumed at the fixed effect model, that the treatment effects sum to zero ($\sum_{i=1}^{a} \tau_i = 0$). Know, in the random factor model, we replace this assumption with the assumption that the treatment effects come from a population with expectation equals zero ($E(\tau_i) = 0$).

$$\tau_i \sim N(0, \sigma_\tau^2)$$

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

That is both τ_i and ϵ_{ij} are random variables and we assume that τ_i and ϵ_{ij} are independent for all I and j.

The total variance is:

$$Var(y_{ij}) = Var(\mu + \tau_i + \epsilon_{ij}) = \sigma_\tau^2 + \sigma^2$$

The total variance describes the total variability across all the random levels and replicates. The two population parameters σ_τ^2 and σ^2 are called variance components. The variance σ^2 describes the variability between the replicates made on the same level of the random variable. This is the variance which we cannot explain. In the experiment in this thesis this variance is the variance for the sample units which belong to the same fat category, have been analyzed by the same instrument and belong to the batch. The variance σ_τ^2 describes the variability between the random factor levels, this would be the variance between the batches in my experiment.

4.2.4.1 TEST FOR RANDOM EFFECTS

It is meaningless to test hypothesis about included treatment effect because we are not interested in the effect of these specific treatment levels. Our test know that there is no effect of the random factor, all $\tau's$ must be zero, not only the $\tau's$ in our data ($\tau_1, \cdots, \tau_a$) but all in the population. In our experiment this means that there is no batch effect.

The hypothesis test:

$$H_0 : \sigma_\tau^2 = 0 \quad \text{Versus}$$

$$H_1 : \sigma_\tau^2 > 0$$

The sum of squares and the degrees of freedom is as for the fixed model.

We still have:

$\frac{SS_E}{\sigma^2} \sim \chi^2_{N-a}$ Where N is the total number of the observations (N=a.n).

And under the null hypothesis ($H_0 : \sigma_\tau^2 = 0$), we have:

$\frac{SS_{Tr}}{\sigma^2} \sim \chi^2_{a-1}$

And under the null hypothesis, the statistic test will be:

$$F_0 = \frac{MS_{Tr}}{MS_E} \sim F_{(a-1),(N-a)}$$

We reject the null hypothesis if the ($F_0 > F_{\alpha,(a-1),(N-a)}$), this means that we have factor effect. Or we reject if the p-value < the level of significance (α).

4.2.4.2 THE PARAMETERS ESTIMATORS

The expectation of the mean square plays important role in estimation of the variance component and to setup the correct test. It can be shown that:

$$E(MS_{Tr}) = \sigma^2 + n\sigma^2_\tau$$

And further:

$$E(MS_E) = \sigma^2$$

Hence logical estimators are found by setting:

$$\hat{\sigma}^2 = MS_E$$

$$\hat{\sigma}^2 + n\hat{\sigma}^2_\tau = MS_{Tr}$$

This gives:

$$\hat{\sigma}^2_\tau = \frac{MS_{Tr} - MS_E}{n}$$

This estimation method is known as the analysis of variance method or the method of moments.

If we have different number of replicates in the different levels of the treatments then we have unbalanced model. In such situation we use n_0 instead of n.

We define n_0 as:

$$n_0 = \frac{1}{a-1}\left(\sum_{i=1}^{a} n_i - \frac{\sum_{i=1}^{a} n_i^2}{\sum_{i=1}^{a} n_i}\right)$$

In some cases, the method used for variance component estimation gives negative estimates if ($MS_{Tr} > MS_E$). To deal with this problem we could replace the estimate by zero or take it as an indication that the model is wrong and redefine the model or use another estimation method which does not give negative estimates (e.g. Maximum Likelihood Estimation).

The estimation for the (μ) is:

$$\hat{\mu} = \overline{y}_{..} = \frac{1}{na}\sum_{i=1}^{a}\sum_{i=1}^{n} y_{ij}$$

$$= \frac{1}{na}\sum_{i=1}^{a}\sum_{i=1}^{n}(\mu + \tau_i + \epsilon_{ij})$$

$$\hat{\mu} = \overline{y}_{..} = \mu + \overline{\tau}_{.} + \overline{\epsilon}_{..}$$

The variance of $\hat{\mu}$ is:

$$\begin{aligned} Var(\hat{\mu}) &= Var(\mu + \overline{\tau}_{.} + \overline{\epsilon}_{..}) \\ &= Var(\overline{\tau}_{.}) + Var(\overline{\epsilon}_{..}) \\ &= \frac{\sigma_\tau^2}{a} + \frac{\sigma^2}{an} \\ &= \frac{1}{na}(n\sigma_\tau^2 + \sigma^2) \end{aligned}$$

The standard error:

$$SE(\hat{\mu}) = \sqrt{\frac{1}{na}(n\hat{\sigma}_\tau^2 + \hat{\sigma}^2)} = \sqrt{\frac{1}{na}MS_{Tr}}$$

We can use the standard error to find a (1-α) 100% Confidence interval to test the hypothesis $H_0 : \mu = \mu_0$:

$$\left(\hat{\mu} \pm \tau_{\alpha/2,(a-1)} \sqrt{\frac{1}{na}MS_{Tr}} \right)$$

And we can reject the H_0 if the μ_0 not in the C.I at the significant level α.

We can also find a (1-α) 100% confidence interval for σ^2:

$$\left[\frac{SS_E}{\chi^2_{1-\alpha/2,N-a}}, \frac{SS_E}{\chi^2_{\alpha/2,N-a}} \right]$$

4.2.4.3 PREDICTION OF RANDOM EFFECTS

The two random variables of the model are τ_i and ϵ_{ij}. I we want to predict the values of these two variables according to our model.

The predictor values for the error are the residuals:

$$e_{ij} = (y_{ij} - \hat{y}_{ij}) = (y_{ij} - (\hat{\mu} + \hat{\tau}_i))$$

The random effects τ_i are predicted by:

$$\hat{\tau}_i = (\overline{y}_{i.} - \overline{y}_{..})$$

4.2.5 THE TWO-FACTOR MIXED MODEL

In this part, we will discuss the mixed model. This model has two factors one of them fixed (A) and the other on is random (B). In this case we want to draw conclusion about the specific levels of the fixed model and conclusion about the general effect of the random factor.

In this kind of models, we have the term interaction. The basic idea of the interaction is that the effect of one factor depends on the level of the second factor. That is if we have two fat categories and two or more instruments to analyze the fat content, the effect of the instrument depend on the fat category that means some instruments could be better to analyze specific fat category.

Figure 4.2.2 is a figure from the web net (12) to show how the interaction plot looks like. We see if there is no significant interaction effect, the plot will represent two parallel lines like the parts A, B and C from (figure 4.2.2), while the part D from the same figure show nonparallel lines, that is index of significant interaction effect.

(Figure 4.2.3) is from the experiment in this thesis here I investigated the potential interaction effect between fat category and the fat analysis instrument. The plot showed us two parallel lines because we do not have significant interaction effect.

If we have two factors, one fixed and second random, the interaction will be random variable also.

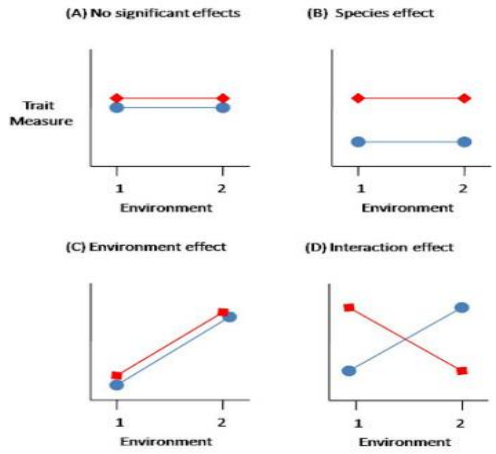


Figure 4.2.2: Interaction plots from the internet show the idea of the interaction

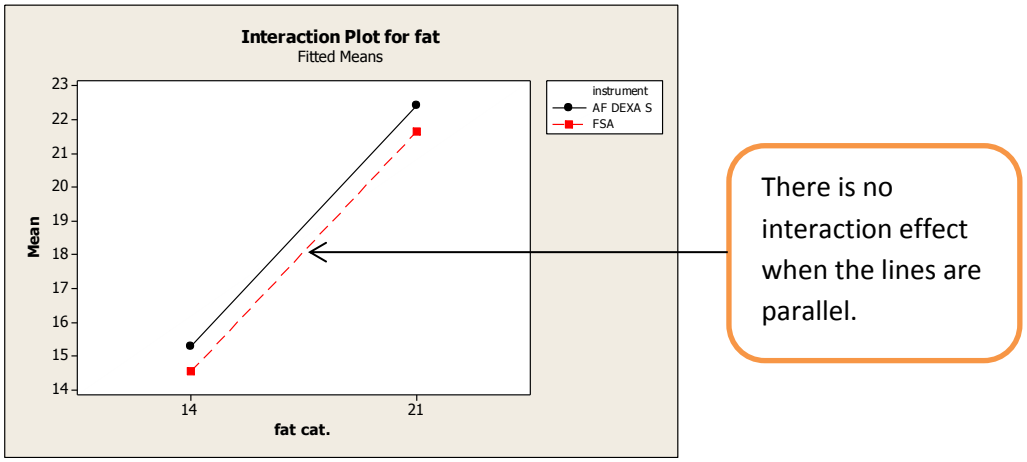


Figure 4.2.3: Interaction plot between two instruments in this thesis

4.2.5.1 THE RESTRICTED MIXED MODEL

We call the model in this part the restricted model because the sum of the interaction effect over the levels of the fixed factor equals zero.

The model now is:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, b \\ k = 1, 2, \cdots, n \end{cases} \quad (4.2.12)$$

The model assumptions are:

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

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

$$\beta_j \sim N(0, \sigma_\beta^2)$$

$$(\tau\beta)_{ij} \sim N[0, ((a-1)/a)\sigma_{\tau\beta}^2]$$

Because we have the restricted model, we assume that the interaction effects sum to zero across the fixed effects levels:

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

The expected mean sums of squares for the restricted model are:

$$E(MS_A) = \sigma^2 + n\sigma_{\tau\beta}^2 + \frac{bn\sum_{i=1}^{a}\tau_i^2}{a-1}$$

$$E(MS_B) = \sigma^2 + an\sigma_\beta^2$$

$$E(MS_{AB}) = \sigma^2 + n\sigma_{\tau\beta}^2$$

$$E(MS_E) = \sigma^2$$

We will use these expected mean of squares to set the test statistic and to find the estimators.

4.2.5.2 TESTING IN THE RESTRICTED MIXED MODEL

The first test is to check if there is any effect for the fixed factor. The hypothesis in this situation is:

$H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$ Versus

$H_1 : at\ least\ one\ of\ \tau's\ not\ equal\ to\ zero$

Under the H_0, the statistic test is:

$$F_0 = \frac{MS_A}{MS_{AB}} \sim F_{\alpha,(a-1),(a-1)(b-1)}$$

We see that to obtain correct test statistic we have to divide the MS_A by the MS_{AB} and not the MS_E as we used to do.

The second test is for the random factor and the hypothesis for this test is:

$H_0 : \sigma_\beta^2 = 0$ Versus

$H_1 : \sigma_\beta^2 \neq 0$

Under the H_0 the test statistic is:

$$F_0 = \frac{MS_B}{MS_E} \sim F_{\alpha,(b-1),ab(n-1)}$$

The third test is for the interaction term and the hypothesis is:

$H_0 : \sigma_{\tau\beta}^2 = 0$ Versus

$H_1 : \sigma_{\tau\beta}^2 \neq 0$

And the test statistic is:

$$F_0 = \frac{MS_{AB}}{MS_E} \sim F_{\alpha,(a-1)(b-a),ab(n-1)}$$

From the expected means of square we could obtain the variance component estimators for the random factor the interaction term and the error. The fixed factor has no variance.

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_E}{an}$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n}$$

$$\hat{\sigma}^2 = MS_E$$

4.2.5.3 THE UNRESTRICTED MIXED MODEL

This model is the same model for the restricted mixed model. The difference is we have no restriction on the random interactions. To distingue this model from the restricted model we will use different Greek letters for the effects.

$$y_{ijk} = \mu + \alpha_i + \gamma_j + (\alpha\gamma)_{ij} + \epsilon_{ijk} \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, b \\ k = 1, 2, \cdots, n \end{cases} \qquad (4.2.13)$$

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

$\epsilon_{ij} \sim N(0, \sigma^2)$

$\gamma_j \sim N(0, \sigma_\gamma^2)$

$(\alpha\gamma)_{ij} \sim N(0, \sigma_{\alpha\gamma}^2)$

The expected mean sums of squares are:

$$E(MS_A) = \sigma^2 + n\sigma^2_{\alpha\gamma} + \frac{bn\sum_{i=1}^{a}\alpha_i^2}{a-1}$$

$$E(MS_B) = \sigma^2 + an\sigma^2_{\gamma}$$

$$E(MS_{AB}) = \sigma^2 + n\sigma^2_{\alpha\gamma}$$

$$E(MS_E) = \sigma^2$$

The statistic tests according to the expected mean sums of squares will be now:

The first test is to check if there is any effect for the fixed factor. The hypothesis in this situation is:

$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$ Versus

$H_1 : at\ least\ one\ of\ \alpha's\ not\ equal\ to\ zero$

Under the H_0, the statistic test is:

$$F_0 = \frac{MS_A}{MS_{AB}} \sim F_{\alpha,(a-1),(a-1)(b-1)}$$

We see that to obtain correct test statistic we have to divide the MS_A by the MS_{AB} and not the MS_E as we used to do.

The second test is for the random factor and the hypothesis for this test is:

$H_0 : \sigma^2_{\gamma} = 0$ Versus

$H_1 : \sigma^2_{\gamma} \neq 0$

Under the H_0 the test statistic is:

$$F_0 = \frac{MS_B}{MS_{AB}} \sim F_{\alpha,(b-1),(a-1)(b-1)}$$

The third test is for the interaction term and the hypothesis is:

$H_0 : \sigma^2_{\alpha\gamma} = 0$ Versus

$H_1 : \sigma^2_{\alpha\gamma} \neq 0$

And the test statistic is:

$$F_0 = \frac{MS_{AB}}{MS_E} \sim F_{\alpha,(a-1)(b-a),ab(n-1)}$$

The variance component estimates in the unrestricted mixed model are:

$$\hat{\sigma}^2_{\gamma} = \frac{MS_B - MS_{AB}}{an}$$

$$\hat{\sigma}^2_{\alpha\gamma} = \frac{MS_{AB} - MS_E}{n}$$

$$\hat{\sigma}^2 = MS_E$$

Good part of the authors prefers the restricted model. These authors believe that the restricted model is slightly more general than the unrestricted model. But the unrestricted model is easier to use with the unbalanced data set. Many computer programs have the unrestricted model as a default, like SAS, MINITAB and R.

4.2.6 NESTED FACTORS MODEL

If we have two factors A and B, sometimes the levels of a factor B are similar but not identical for different levels of the factor A. this gives the nested or hierarchical design and we say the levels of factor B are nested under the levels of factor A. this type of model are widespread in industrial and agricultural studies. The models which we discussed until know were models with crossed factors that mean the levels of the factor B are identical for all levels of A. but if the levels of the factor B which are tested at the first level of the factor A, are different from the levels of B which are tested at the second level of A.

The two-stage model has two factors on nested within the other:

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{(ij)k} \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, b \\ k = 1, 2, \cdots, n \end{cases}$$

This model is for the balanced nested design. The notation $\beta_{j(i)}$ indicates that the level j of B is nested within level I of A. We have also the replicates nested within the level i of A and level j of B, but the replicates are always nested within the levels of the factors in the model, so we should have the notation $\epsilon_{(ij)k}$ for all models. In the nested model, we do not have interaction effect since not all levels of B is tested under every level of A.

The assumptions for this model are dependent on the kind of the factors A and B. We can have the both factor fixed, the both factors random or A fixed and B random.

The assumptions:

If A is fixed we assume $\sum_{i=1}^{a} \tau_i = 0$

If B is fixed we assume $\sum_{j=1}^{b} \beta_{j(i)} = 0$

If A is random we assume $\tau_i \sim \ N(0, \sigma_\tau^2)$

If B is random we assume $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$

We can split the total sum of squares into two parts, the explained part which is the effect of the factors and the unexplained part which is the error or the noise.

$$SS_T = SS_A + SS_{B(A)} + SS_E$$

We have abn-1 degrees of freedom for SS_T, a-1 for SS_A, a(b-1)for $SS_{B(A)}$ and ab(n-1) for SS_E.

4.2.6.1 THE EXPECTED MEAN SUM OF SQUARES

The expected mean values depend on the combination of the factors. The table 4.2.1 summarized the expected values.

$E(MS)$	A Fixed B Fixed	A Fixed B Random	A Random B Random
$E(MS_A)$	$\sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\tau^2$
$E(MS_{B(A)})$	$\sigma^2 + \frac{n\sum\sum \beta_{j(i)}^2}{a(b-1)}$	$\sigma^2 + n\sigma_\beta^2$	$\sigma^2 + n\sigma_\beta^2$
$E(MS_E)$	σ^2	σ^2	σ^2

Table 4.2.1: The expected mean sum of square

4.2.6.2 THE TESTING

From the table (1) we can find the test statistic for all the combinations for the two factors. If the both factors A and B are fixed then both tests are based on SS_E and the null hypotheses are:

$$H_0 : \tau_i = 0 \; and \; H_0 : \beta_{j(i)} = 0 \; where \; i = 1, 2, \cdots, a \; and \; j = 1, 2, \cdots, b$$

If A is a fixed factor and B is a random factor, then to test the effect of the factor A, we use the $MS_{B(A)}$ as test error and the null hypothesis is: $H_0 : \tau_i = 0$. For the random factor, we use the MS_E as test error and the null hypothesis is: $H_0 : \sigma_\beta^2 = 0$.

If both factors are random, to test the factor A, the null hypothesis is: $H_0 : \sigma_\tau^2 = 0$ and we use $MS_{B(A)}$ as test error. For the factor B, the null hypothesis is: $H_0 : \sigma_\beta^2 = 0$ and we use MS_E as test error.

4.2.6.3 VARIANCE COMPONENTS

Dependent on the assumptions, the variance components will be:

- A fixed, B fixed: σ^2
- A fixed, B random: $\sigma_\beta^2 \; and \; \sigma^2$
- A random, B random: $\sigma_\tau^2, \; \sigma_\beta^2 \; and \; \sigma^2$

From the equations in table (4.2.1), we can obtain the following estimates:

$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{B(A)}}{nb}$$

$$\hat{\sigma}_\beta^2 = \frac{MS_{B(A)} - MS_E}{n}$$

$$\hat{\sigma}^2 = MS_E$$

4.2.6.4 THE ESTIMATION OF μ AND THE VARIANCE OF THIS ESTIMATION

$$\hat{\mu} = \overline{y}_{...} = \mu + \overline{\tau}_{.} + \overline{\beta}_{.(.)} + \overline{\epsilon}_{...}$$

The variance of this estimator is:

$$Var(\overline{y}_{...}) = Var(\mu) + Var(\overline{\tau}_{.}) + Var(\overline{\beta}_{.(.)}) + Var(\overline{\epsilon}_{...})$$

This gives:

For A and B fixed

$$Var(\overline{y}_{...}) = \frac{1}{abn}\sigma^2$$

For A fixed and B random

$$Var(\overline{y}_{...}) = \frac{1}{ab}\sigma_{\beta}^2 + \frac{1}{abn}\sigma^2$$

For A and B random

$$Var(\overline{y}_{...}) = \frac{1}{a}\sigma_{\tau}^2 + \frac{1}{ab}\sigma_{\beta}^2 + \frac{1}{abn}\sigma^2$$

These variances of the estimator are very important to find the confidence interval for μ and the sample size for the model.

4.2.7 THE USED MODEL IN THIS THESIS, COMPARING PAIRS OF TREATMENT MEANS AND TUKEY'S TEST

The model which we use in the first round of the data collection in this thesis is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2, \cdots, 6 \\ k = 1, 2, \cdots, 6 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.10)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_{\gamma}^2)$ we assume that the batches are independent.
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that the residuals are independent.
We assume that the batches and the residuals are independent from each other.

y_{ijkl}: Is the fat content for sample unit from the fat category (i), instrument (j), batch (k) and the residual (l)

The measured mean (μ) is the fat content for all the meat which have been sent from AASHEIM KJØTT to STABBURET in the 14% and 21% fat category.

τ_i : Is the effect of the fat category (i).

β_j: Is the effect of the instrument (j).

γ_k: Is the effect of the batch (k).

$\epsilon_{(ijk)l}$: Is the error term.

In this model, we have the fat category and the instrument crossed and the batch effect nested within the fat category. We will discuss more about this model in the chapter with the results. The model for the second round of the data collection is the same model with the difference in the batch number (17 instead of 6).

The basic aim of this thesis is check if there is any difference between the instruments. If there are differences, where these differences are? We interested most at the differences between Q-monitor at AASHEIM KJØTT and FA DEXA at STABBURET from one side and FOODSCAN at ANIMALIA from the other side. That is we used FOODSCAN as reference instrument.

We want to determine which instrument means differ by testing the difference between all pairs of instruments means. The mean for the instruments is define as :

$$\mu_{instrument} = \mu + \tau_{fat.cat} + \beta_{instrument}$$

The hypotheses that we wish to test for one fat category are:

$H_0 : \mu_{instrument.i} = \mu_{instrument.j}$ or $H_0 : \mu_{instrument.i} - \mu_{instrument.j} = 0$ $for\ all\ i \neq j$ versus

$H_0 : \mu_{instrument.i} \neq \mu_{instrument.j}$ or $H_0 : \mu_{instrument.i} - \mu_{instrument.j} \neq 0$

In this thesis, we will use Tukey's method to test the hypotheses. The benefit of this method is that it controls the experiment error rate, this total error rate should not exceed the significance level α.

Tukey's test declares two means significantly different if the absolute value of their sample differences exceeds (13):

$$T_\alpha = q_\alpha(a, df)\sqrt{\frac{MS_{Error}}{n}}$$

Where α is the family error rate, df is the degree of freedom of SS_{Error} and a is the number of factor levels. The q-value is a critical value from the studentized range statistic found in tables.

When sample sizes are not equal as in this thesis, the privous equation become:

$$T_\alpha = \frac{q_\alpha(a,df)}{\sqrt{2}}\sqrt{MS_{Error}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

4.2.8 THE PROBLEMS WITH THE UNBALANCED DATA SET

In this thesis, we investigated six instruments, each instrument return different number of results. That means we have unbalanced data set. The unbalanced data leads to several problems. One of the most important problem that the unbalanced data provides correlated estimators. It is difficult mathematically to deal with unbalanced data. It is difficult also to calculate the variance components.

To make this more understandable we assume that we have three observations from two different instruments. These observations are uncorrelated, but the estimators are correlated.

$$Y = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}}_{X} \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \end{bmatrix}$$

So we have $X'X = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and $(X'X)^{-1} = \frac{1}{9-1}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$

And the correlation matrix for the estimators will be:

$corr\begin{pmatrix} \hat{\mu} \\ \hat{\alpha} \end{pmatrix} = \begin{pmatrix} 1 & -3/8 \\ -3/8 & 1 \end{pmatrix}$ the correlation matrix not diagonal any more. It means that the estimators are correlated.

4.2.9 THE PROBLEM WITH THE CORRELATED OBSERVATIONS

In the start of the sample units' collection and in the first analysis by FOODSCAN at ANIMALIA, we wanted to check if it is useful to take more than one replicates for some of the sample unites. From the first twenty sample units, I analyzed two replicates for ten of these samples.

If we have for instance two replicates from the same sample unit, then we have:

$$\text{var}(\overline{Y}) = \frac{\sigma^2}{2}(1+\rho) \ .$$

Where $\overline{Y}$ is the average for the replicates.

As we see the variance of the estimated average is dependent on the correlation (ρ). The value ($\frac{\sigma^2}{2}\rho$) it is what we add to variance when we move from the independent observations to the correlated replicates situation. In our situation we have very small variance and hence big correlation and this is reasonable because we toke the replicates from the same sample.

CHAPTER 5

RESULTS

5.1 RESULTS FROM MULTIVARIATE CALIBRATION

In this we will illustrate what we discussed in the theoretical part about multivariate calibration in chapter 4.1. For this purpose, we borrowed two data sets from Dr. Jens Petter Wold at NOFIMA. These data sets are from using Qvision-500 on meat samples. The mechanism of this instrument is that the sample illuminated by to two light spots. Light that has travelled (15-20 mm) through the sample is detected in a small, non-illuminated area between the light spots (figure 5.1.1). This instrument provides spectral images of fifteen wavelengths between 760nm and 1040nm with a spectral resolution of 20nm.Different peaks and broad peaks on different wavelengths refer for different absorption levels. For instance the peak about 930nm is absorption by fat and the peak at 980nm is absorption by water while the variation at the lower wavelengths is due to color variations.

The different elements in the sample like fat, protein and water reflex different wavelengths. Each sample provides a spectrum. The figure 5.1.2 shows the spectra for the 35 samples in data set number one. In the Y axis is the reflection and in the X axis is the wavelengths. Each sample represents one line in this plot. The shape of the sample line is dominated by the color and the content of fat, water, and protein etc. the peak in each sample line is according to what this sample content most. For instance the sample with high fat content will create peak at the wavelength which correspond to light absorption for fat.

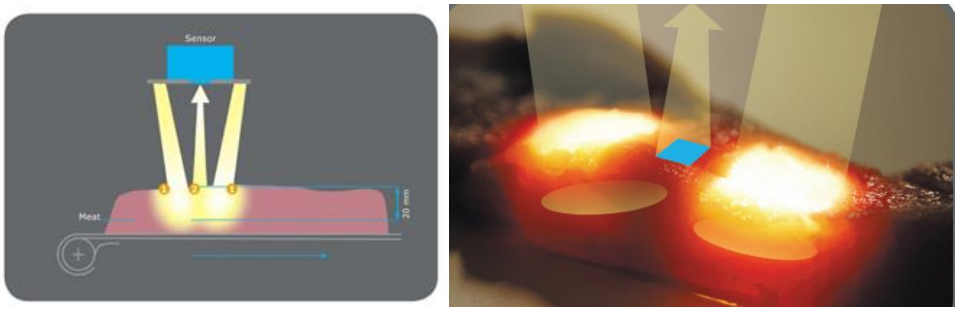


Figure 5.1.1: The light reflection mechanism for Near Infrared spectroscopy

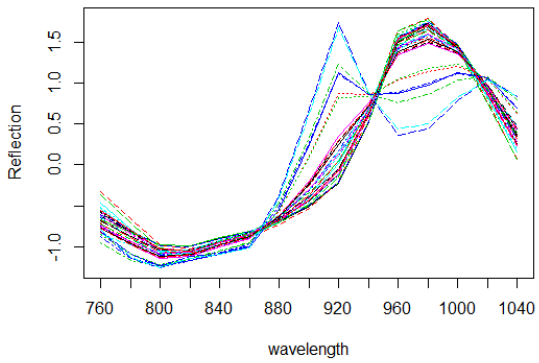


Figure 5.1.2: Near Infrared spectrum of data set 1

5.1.1 PRINCIPAL COMPONENT ANALYSIS (PCA)

We used PCA on data set 1. The following results are what we obtain from this analysis.

	PC1	PC2	PC3	PC4
X760	-0.11962038	0.66986144	0.01704039	0.37736258
X780	-0.14038241	0.34589222	-0.16847984	0.20460181
X800	-0.09428203	0.09617766	-0.33066013	0.02008453
X820	-0.04458733	0.02552274	-0.32223753	-0.13531671
X840	-0.06082022	-0.03706795	-0.24768990	-0.26823886
X860	-0.05991392	-0.07257437	-0.18615432	-0.36516362
X880	0.09407167	-0.01859671	-0.11240985	-0.37520451
X900	0.36895906	0.08399104	0.03808298	-0.23734187
X920	0.62411859	0.15273780	0.31409113	-0.02106914
X940	0.10768900	-0.02138646	0.47935761	-0.01574157
X960	-0.37518930	-0.07305671	0.38812759	-0.09138489
X980	-0.40937576	-0.15034990	0.25467553	-0.10007131
X1000	-0.19492238	-0.23460482	0.16851357	0.19350864
X1020	0.08474560	-0.32353552	-0.01764268	0.34852833
X1040	0.21950982	-0.44301018	-0.27461488	0.46544675

Table 5.1.1: The loadings on the first 4 components data set 1

The loadings provide the eigenvectors for the data matrix (x). The first column (PC1) is the first eigenvector corresponding to the largest eigenvalue. According to the wanted amount of variance, we select the number of the columns we should include in the loadings matrix (Z) from the equation (4.1.3). By this, we reduce the dimension of the data matrix from 15 variables in the original matrix to much fewer variables. If we satisfy with the variance included in the first component, we could obtain the one dimension variable.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Standard deviation	0.8572	0.14415	0.08477	0.03349	0.01008	0.009373	0.006624
Proportion of Variance	0.9616	0.02719	0.00940	0.00147	0.00013	0.000110	0.000060
Cumulative Proportion	0.9616	0.98876	0.99816	0.99963	0.99976	0.999880	0.999940

Table 5.1.2: Proportions of variance (Importance of components) for the first seven components of data set 1

This part of the analysis results could be the most important part. The first row which is standard deviation provides the square root for the eigenvalues from the largest to the smallest. These eigenvalues corresponding to the variances of the components, and the sum of these eigenvalues is the total variance (the sum of the 15 eigenvalues). The cumulative proportions provide how much of the variance each principal component explains. That is the first principle component capture (96.16%) of the variance. If we choose to use four principle components, we capture (99.96%) of the variance. The first number in the proportion of variance row is the first eigenvalue divide by the sum for all the eigenvalues and so on for all the components.

The scree plot is a plot of the eigenvalues. This plot is useful to select the number of the components.

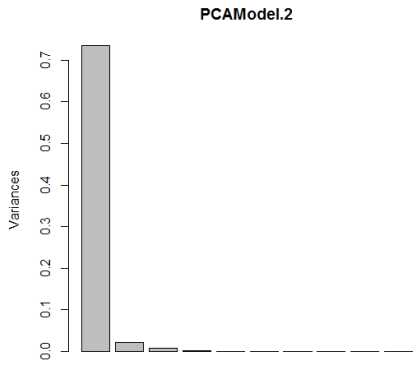


Figure 5.1.3: Scree plot for PCA on data set 1

Figure 5.1.3 show the number of the components in the X axis and in the Y axis, the amount of the variance captured by the eigenvalue for each component. In this figure we noticed that the first eigenvalue capture huge amount of the variance.

5.1.2 PRINCIPLE COMPONENT REGRESSION (PCR)

We tried the PCR with all possible 15 components. The results are captured in two tables. The first table (5.1.3) shows the number of component and RMSEP corresponding for each one of these component. The second table (5.1.4) shows the number of the component and in the first row how much of the variance captured by each component. The second row in this table is the (R^2) values. The first number in this row shows the amount of the variability explained by the model if we use the

first component as the reduced data matrix (Z). The second number is the (R^2) if we decide to use one or two or more components. R^2 is statistical measure of how close the data observations from the regression line. We obtain this value by dividing the explained variance by the total variance, so it show us how much of the variability of the response explained by the model.

	(Intercept)	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps
CV	22.82	3.835	3.897	2.733	2.791	2.847	3.070

Table 5.1.3: Cross-validated using 35 leave-one-out segments (RMSEP) for the PCR model

	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps	7 comps	8 comps
X	96.16	98.88	99.82	99.96	99.98	99.99	99.99	100
fett	97.39	97.49	98.80	98.86	98.87	98.94	99.00	99

Table 5.1.4: The amount of variability captured by the component and (R^2) for each component

From table 5.1.4, we noticed that the first component which captures 96.16% of the variability in X has ($R^2 = 97.39$). That is the model with one principle component explains 97.39% of the total

variance. We noticed also that the value of R^2 almost stop to increase after the first three components.

From table 5.1.3, we noticed that the lowest value for RMSEP is obtained by using 3 components. From the fourth component the value of the RMSEP starts to increase.

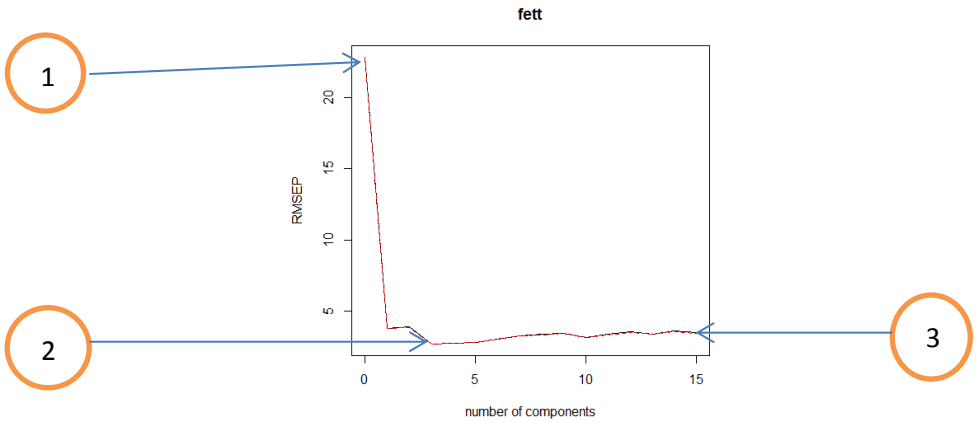


Figure 5.1.4: RMSEP against number of component plot

When we plot the RMSEP values from the table 5.1.3 against the number of the component, we noticed that there are three important points. The point number one is RMSEP for the model without any explanatory variables (the so called null model). The second point is the minimum RMSEP for the model with three principal components as explanatory variables. The third point is RMSEP if we use all the components as explanatory variables matrix, this would be RMSEP if we use the Least Square model. That is because the estimator $\hat{\beta}$ for PCR and LS are equal when we use all components.

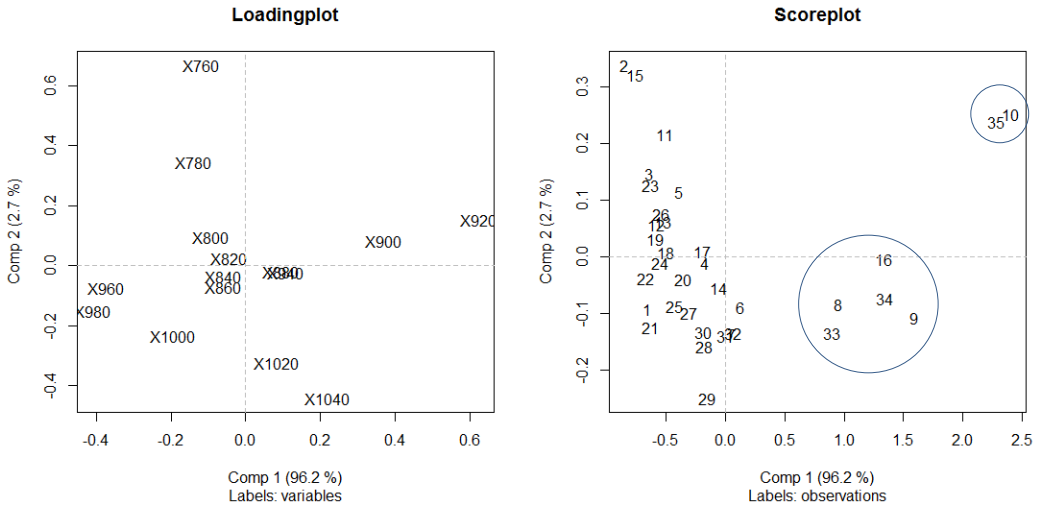


Figure 5.1.5: The loading plot and the score plot for PCR on data set 1

In the first plot in this figure we plot pairs of loadings vectors or the first and the second eigenvectors. This plot distinguish between the wave lengths which refer to the high fat content and the wave lengths which refer to low fat content. The variables or wave lengths which provide low fat

content are to the left side of the plot and the wave length which provide high fat content to the right.

The second plot is the score plot; we plot pairs of score vectors. This plot is useful for detecting groups in the data. This plot confirms the result from the first plot. We noticed that the observations 10 and 35 which are the observations with highest fat content in the data are to the farthest right. The second group is the observations (8, 9, 16, 33 and 34) are also with high fat content but lower than the two observations we refer to previously.

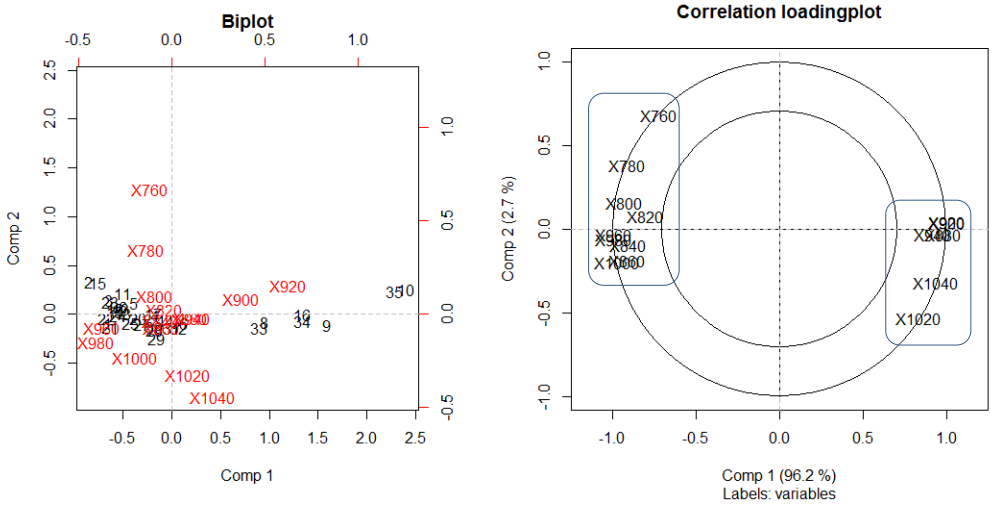


Figure 5.1.6: The biplot and the correlation plot for PCR on data set 1

The first plot in this figure is the biplot. This plot is a combination of the plots in the figure 5.1.5. The fourth plot is correlation loadings plot. This is plot of the correlations between the original variables and the principal components. This plot is useful for checking the contributions of each variable to the components. In the correlation plot, we noticed clearly two groups of wave lengths. The first to the left is for the wave lengths back the observations with low fat content and the second to the right is for the wave lengths back the observations with high fat content.

5.1.3 PARTIAL LEAST SQUARE REGRESSION (PLSR)

We fitted the PLSR model and the results are captured in the next two tables.

	(Intercept)	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps
CV	22.82	3.815	3.098	2.727	2.828	3.216	3.460

Table 5.1.4: Cross-validated using 35 leave-one-out segments (RMSEP) for the PLSR model on data set 1

	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps	7 comps	8 comps
X	96.16	97.79	99.81	99.96	99.97	99.98	99.99	99.99
fett	97.42	98.60	98.82	98.89	99.05	99.17	99.26	99.30

Table 5.1.5: The amount of variability captured by the component and (R^2) for each component

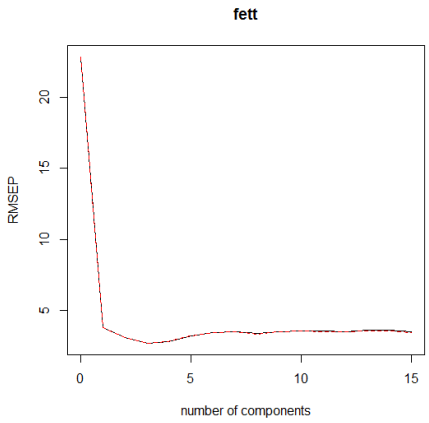


Figure 5.1.6: RMSEP against number of component plot for data set 1 using PLSR

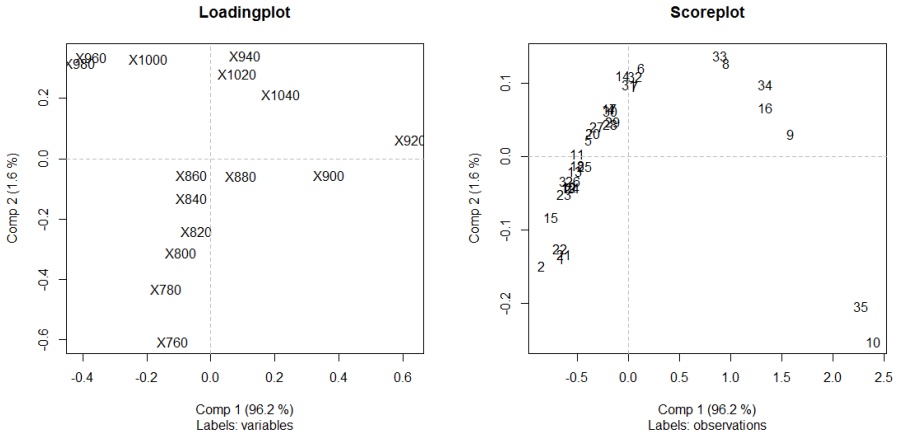


Figure 5.1.7: Loading plot and score plot for data set 1 using PLSR

The score plot for data set 1 using PLSR does not look like the score plot for the same data set using PCR. Actually it is the same plot for the both methods, the difference in the second component provides ostensibly different plot. The second component is not important in both methods.

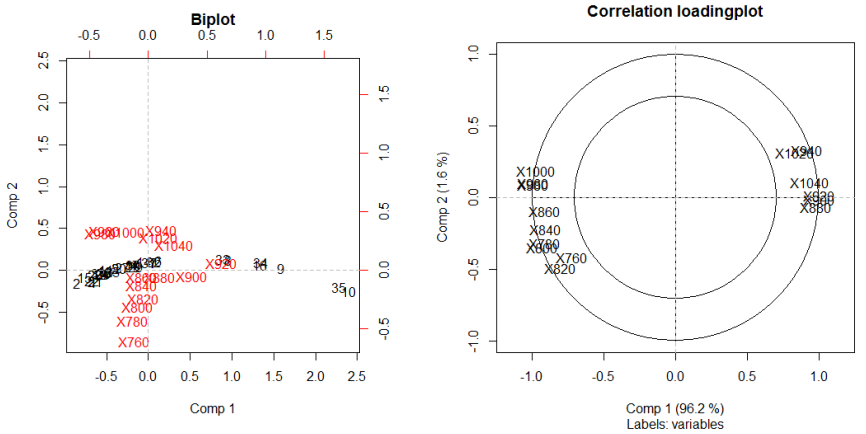


Figure 5.1.8: Biplot and corrlation plot

The results from the PLSR show almost the same results from PCR. That is because the first component explain almost all the variability (96%) in X, and this component is highly correlated to the fat content response. In this situation there is no difference which method we use. It is easier to understand, interpret and use PCR therefore we preferred to use PCR. If the most variability explained by two components and the second component not correlated to the fat content response, we should use PLSR instead for PCR. The reason is that PCR concentrate on the variance inside the data matrix while PLSR consider the covariance between the fat content response and the data matrix.

When we used PCR with three components on the data, we obtained RMSEP equal to 2.73%. The next step is to plot the cross validated predicted fat content values against the observed fat content values.

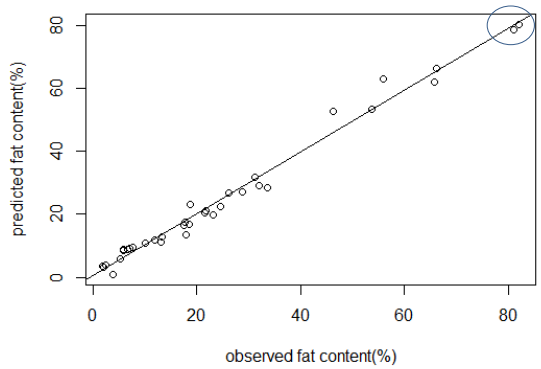


Figure 5.1.9: Plot of cross validated predicted values against observed values for data set 1, when PCR with 3 components is used

When we make a simple leaner model for the predicted values ($\hat{\mathbf{y}}$) as response and the observed values ($\mathbf{y}$) as explanatory variable, this model will be:

$$\hat{\mathbf{y}} = \beta_0 + \beta_1 \mathbf{y} + \epsilon \qquad \text{with the common assumption on the error} \quad (5.1.1)$$

We obtain ($R^2 = 0.985$) and the follow coefficients:

Coefficients:

$\hat{\beta}_0$	$\hat{\beta}_1$
0.4115	0.9824

Table 5.1.6: Estimated Regression Coefficients for model 5.1.1

The results from this table show as that we systematically over predict due to value of $\hat{\beta}_0$. This value should be zero at the optimal situation that means the entire observation will lie on the regression line. If we look to the plot in the figure 5.1.14, we will see that some values over the line. These observations are the over predicted observations. That is the prediction model provides predicted value larger than the observed value. By the same explanation, the observations under the line are the under predicted values. For instance the two observations in the top of the plot which marked

with circle, are over 80% fat content, but the prediction model predict these values under 80% fat content. That is the model under predicted these two observations.

Due to the high R^2 this prediction model seems to be suitable to use with second data set, the test data set. Which is contains 386 observations.

5.1.4 CALIBRATION AND TEST SETS

When we used the PCR model with three components on the test data set, we obtained RMSEP value equal to (6.586%). This number is suspiciously high compared to the cross validation result in data set 1, therefore we will plot the predicted values against the observed values and we will plot the NIR spectrums.

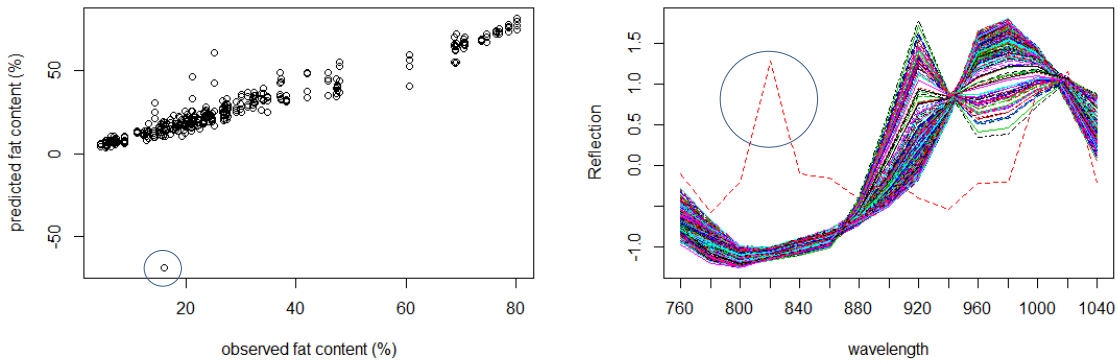


Figure 5.1.10: Predicted values vs observed value plot and NIR spectrums plot for test set

From these two plots in figure 5.1.10, we noticed that there is one extreme observation. These kinds of observations are known as outliers. Any outlier should be investigated very carefully. These kinds of observations could be very informative. Therefore we cannot delete them before we are sure that they are not result of a mistake or just noise. I this case, I contact Dr. Wold and he explained for me that is only mistake back this observation and just noise. Therefore we will refit the model without this observation.

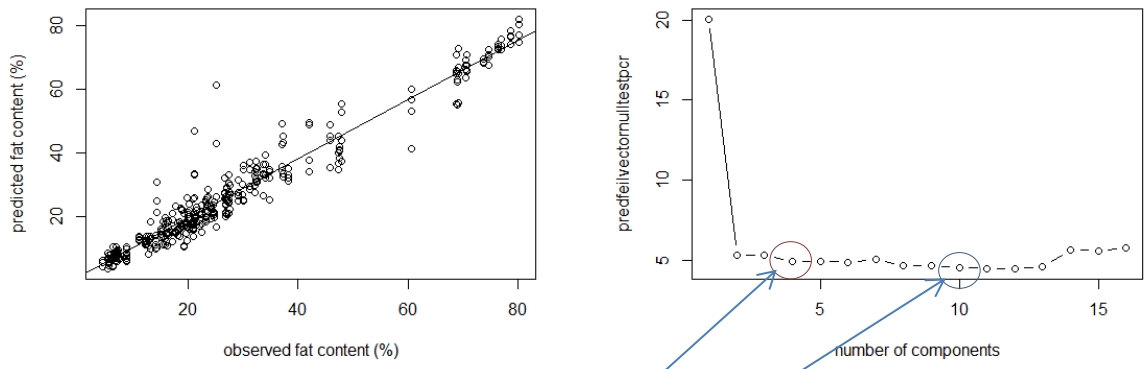


Figure 5.1.11: Predicted values vs observed values and RMSEP plot for all components

```
20.066667  5.336269  5.325347  4.894802  4.937096  4.883855  5.073242
4.641739  4.641631  4.562053  4.463710  4.464856  4.591994  5.648811
5.583525  5.783757
```

Table 5.1.7: The RMSEP values for the models from null to 15 components

The PCR model provides a RMSEP value equal to (4.89%). This value still high and is almost the double of the RMSEP when we used this model on the calibration data set. When we use the model 5.1.2 in this situation, we obtain ($R^2 = 0.94$).

$$\hat{\mathbf{y}}_{test} = \beta_0 + \beta_1 \mathbf{y}_{test} + \epsilon \qquad \text{with the common assumption on the error} \quad (5.1.2)$$

The estimated coefficients are:

```
Coefficients:
      β̂0          β̂0
  0.8159      0.9316
```

Table 5.1.8: Estimated Regression Coefficients for model 5.1.2 using PCR

This high RMSEP is logical if we know that the variance inside the test set is very large. The reason that the variance is so large is that 90 sample units (pork trimmings) have been scanned, each sample have been scanned 4 times from 4 different surfaces. The different in the fat content for the same sample unit and for different surfaces could be huge. For the same sample unit, one surface could be very lean and the other surface is only fat. Dr. Wold from NOFIMA explained to me that the average for the 4 replicates would decrease RMSEP.

5.2 RESULTS FROM ANOVA

Part of the aim of this thesis is to learn how to build different models. Therefore I started to analyze the collected data step by step to fit different models with different techniques. This gradually data analysis helps to discuss different problems. The estimated means according to the model will be represented in the final model when we select the results from all the instruments. Some simple results will be represented to make the results understandable for the representatives from the involved companies in this thesis.

5.2.1 THE DATA FROM FOODSCAN AT ANIMALIA

In the 17th of January-2013, I visited AASHEIM KJØTT in Drammen for the first time. It was planned to send two batches the first 14% fat content and the second 21% fat content. From each batch, I collected ($20 \times 1kg$) sample units to be tested by FOODSCAN NIR instrument at ANIMALIA. The same number of sample units from the same batches has been sent to be tested at EUROFINS and at STABBURET. The entire batches have been sent to STABBURET to be tested with x-ray instrument there.

I transferred the sample units with cooler from Drammen to ANIMALIA in Oslo. The next day I used the FOODSCAN instrument to analyze the units. These sample units are numbered from one to twenty. From each sample with odd number, I toke two replicates. For this reason, I have thirty observations from each fat category. We wanted to test if it is useful to have several replicates instead for several sample units. The first model for the data from the FOODSCAN is:

$$fat_{ijk} = mean + fat.cat._i + sample_{j(i)} + Error_{(ij)k} \qquad (5.2.1)$$

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{(ij)k} \begin{cases} i = 1, 2 \\ j = 1, 2, \cdots, 20 \\ k = 1 \ or \ 2 \end{cases}$$

Where

$\sum_{i=1}^{2} \tau_i = 0$

$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$ we assume that samples are independent.

$\epsilon_{(ij)k} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that sample and error are independent from each other.

This model is a mixed nested model because fat category is fixed and samples are random. Samples are nested within fat category.

1. for the 14% fat category:

mean	sd	var	n
14.431	0.82	0.67	30

2. for the 21% fat category:

mean	sd	var	n
21.684	1.01	1.03	30

The sample means for each of the two groups 14% and 21%

Table 5.2.1: Sample mean and stander deviation and variance for each fat category

It is worth mentioning that the fat content we reported from Q-monitor instrument at ÅSHEIM KJØTT for the first two batches was 14.1% and 20.9%.

So that means:

14.431-14.1= 0.33 the difference between the measurement from Q-monitor and the average of the observation which have been tested with FOODSCAN from ANIMALIA in the 14% fat category

21.684-20.9= 0.78 the difference between the measurement from Q-monitor and FOODSCAN from ANIMALIA in the 21% fat category

It is worth mentioning also that the variance and stander deviations above are not correct. The 10 replicates are correlated to the 10 observations, so we will find the variance by other calculations.

Response: fat

	Sum Sq	Df	Mean Sq	F value	Pr(>F)
fat.cat.	702.86	1	702.86	571.51	0.0000
fat.cat.:sample	46.73	38	1.23	10.00	0.0000
Residuals	2.46	20	0.12	-	-

$MSE+1,5.(MS_{sample})=0.12+1.5(1.23)=1.97$

1. This is the variance between the replicates

2. The estimated sample variance

Table 5.2.2: Results from the variance analysis for model 5.2.1

	Expected mean squares
fat.cat.	(3) + 1.5 (2) + 30 Q[1]
fat.cat. sample	(3) + 1.5 (2)
Residuals	(3)

from those variance component we can find the total variance

Table 5.2.3:Variance components for model 5.2.1

From table 5.2.3 we can find the total variance. From these two table, we noticed that:

1. From (1 and 2) in table 5.2.2, it is easy to see that the variance between the samples is more than 10 times bigger than the variance between the replicates. For this reason I think it is reasonable to take 30 samples instead of 20 samples with 2 replicates for the 10 of the 20 samples.

2. At the part 4.2.3 in this thesis, we discussed that replicates for some of the sample units weaken the model. For this reason I recommend to remove the replicates to have better estimators and more stable model.

3. The calculation to find the estimated total variance for the fat content is:

$$\widehat{Var}(fat) = 0 + \widehat{Var}(sample) + \widehat{Var}(residuales)$$

$$\widehat{Var}(residuales) = 0.12$$

$$\widehat{Var}(sample) = \widehat{Var}(residuales) + (1.5).(\widehat{Var}(sample))$$

$$= 0,12 + 1,5.1,23 = 1,965$$

$\widehat{Var}(fat) = 1,965 + 0,12 = 2,085$ this is the estimated total variance.

I visited AASHEIM KJØTT two more times, in the 24th and 31st of January. These times from each batch, I selected ($30 \times 700gr$) sample units. I followed the same process from the first visit to analyze these sample units by the FOODSCAN at ANIMALIA. The data from the FOODSCAN are the first data have obtained because I did the process by myself. The results from the other instruments take time to obtain them because the x-ray instrument at STABBURET was out of order. the instruments depend on the chemical methods take long time to analyze this number of sample units. The befit of the NIR instruments is that they are fast, cheap, and accurate. After the three visits, I obtained 80 sample units from each fat category.

The model I used for the full data from FOODSCAN at ANIMALIA is nested model with fixed fat category effect, random batch effect and random sample effect:

$$fat_{ijkl} = \mu + fat.cat._i + batch_{j(i)} + sample_{k(j)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(j)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, \cdots, 6 \\ k = 1, 2, \cdots, 20 or 30 \\ l = 1 \ or \ 2 \end{cases} \quad (5.2.2)$$

Where:

$\sum_{i=1}^{2} \tau_i = 0$

$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$ we assume that batches are independent.

$\gamma_{k(j)} \sim N(0, \sigma_\gamma^2)$ we assume that samples are independent.

$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batches, samples and errors are independent from each other.

The model I used for the entire data set is mixed nested model with fixed fat category effect, random batch effect and random sample effect. In this model still have the replicates, which mean the error (ϵ) now is the variance between the replicates.

I used the MINITAB program to analyze the data and I got the following result.

Source	DF	Seq SS	Adj SS	Adj MS	F	P	
fat cat.	1	2309,674	2218,834	2218,834	1192,82	0,000	x
batch(fat cat.)	4	7,495	7,421	1,855	1,26	0,288	x
sample(fat cat. batch)	154	227,642	227,642	1,478	12,02	0,000	
Error	20	2,459	2,459	0,123			
Total	179	2547,269					

S = 0,350621 R-Sq = 99,90% R-Sq(adj) = 99,14%

Table 5.2.4: Variance analysis for model 5.2.2

Source	Estimated Value
batch(fat cat.)	0,01331
sample(fat cat. batch)	1,20874
Error	0,12293

This is the error variance and it is the variance between the replicates

Table 5.2.5: Variance components for model 5.2.2

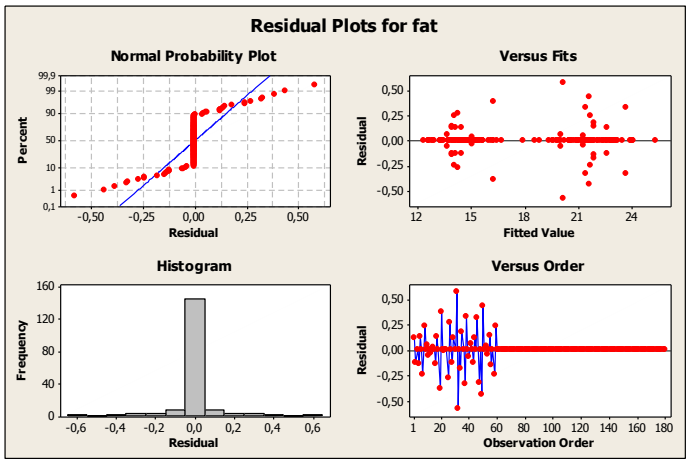


Figure 5.2.1: Residual plots for model 5.2.2

Form the plot we can easy see that there is problem with this model and the reason is the correlation between the replicates. If we look to the first plot (Normal probability plot), this plot should show that the error distribution is approximately normal, the observations should lie on the line but they are not and this is the first index that there is problem with this model.

When we look to the second plot, we can see obvious pattern. If the model is correct and the assumptions are satisfied, the residuals should be structure less. This shape is index to nonconstant variance because of high correlation between the replicates. We can find the estimated correlation between the replicates by the following formula:

$$\widehat{corr.}(rep.) = \frac{\hat{\sigma}^2_{Batch}+\hat{\sigma}^2_{Sample}}{\hat{\sigma}^2_{Batch}+\hat{\sigma}^2_{Sample}+\hat{\sigma}^2_{Error}} = \frac{0.01+1.21}{0.01+1.21+0.12} = \frac{1.22}{1.34} = \boxed{0.91}$$

This is the estimated correlation between the replicates if we have the model 5.2.2

When we have replicates in the data the estimated error (MS_{Error}) will refer to the variance between the replicates. Every sample unit with only one replicate will has residual zero. From the last three of the residual plots we can see that the most residuals are zero. That is because we have only 20 replicates which differ from zero otherwise we have only one replicate which mean zero variance.

If we assume that we don't have replicates, so we treat the replicate as sample unites and the sample unites will be the replicates. We will get much better model but in this situation we underestimate the error because we ignore the correlation between the replicates and the observations.

The model if we consider the replicates as sample units is.

$$fat_{ijkl} = \mu + fat.cat._i + batch_{j(i)} + Error_{(ij)K}$$

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{(ij)k} \begin{cases} i = 1,2 \\ j = 1,2,3 \\ k = 1,2,\cdots,30 \end{cases} \qquad (5.2.3)$$

Where:

$$\sum_{i=1}^{2} \tau_i = 0$$

$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$ we assume that batches are independent.

$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	2309,67	2309,67	2309,67	1232,66	0,000
batch(fat cat.)	4	7,49	7,49	1,87	1,42	0,230
Error	174	230,10	230,10	1,32		
Total	179	2547,27				

S = 1,14996 R-Sq = 90,97% R-Sq(adj) = 90,71%

Table 5.2.6: Variance analysis for the model 5.2.3

Source	Estimated Value
batch(fat cat.)	0,01838
Error	1,32242

Now the error is actually the variance between the samples inside each batch

Table 5.2.7: Variance components for model 5.2.3

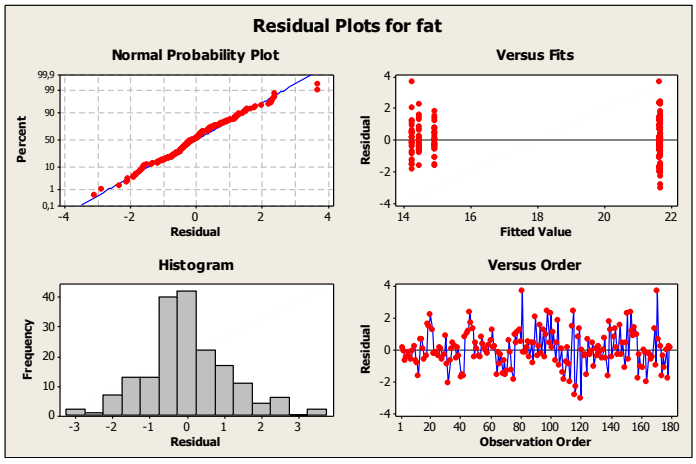


Figure 5.2.2: Residual plots for model 5.2.3

The plots look very nice and good but this is not correct because we ignored the high correlation between the sample units and the replicates that we assumed them as sample units. If we ignore the first two batches we get better balanced model but we lose a lot of information.

Removing the replicates:

Now, we have to remove the replicates randomly from the excel file because they are much correlated. If we have these replicates in the model as replicates, the model will be incorrect. If we conceder the replicates as sample units, we underestimate the variance because we ignore the

correlation between the replicates. I used the R program to make a vector of ones and twos and remove the replicate according to these numbers in the vector.

```
 [1] 1 1 1 2 2 1 1 1 2 2 1 2 1 1 2 1 1 2 1 2
```

I refitted the model without the replicates. The model will be:

$$fat_{ijkl} = \mu + fat.cat._i + batch_{j(i)} + Error_{(ij)K}$$

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{(ij)k} \begin{cases} i = 1,2 \\ j = 1,2,3 \\ k = 1,2,\cdots,20 \ or \ 30 \end{cases} \quad (5.2.4)$$

Where:

$\sum_{i=1}^{2} \tau_i = 0$

$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$ we assume that batches are independent.

$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	2039,61	1970,61	1970,61	1085,00	0,000
batch(fat cat.)	4	7,30	7,30	1,83	1,31	0,268
Error	154	214,12	214,12	1,39		
Total	159	2261,03				

S = 1,17914 R-Sq = 90,53% R-Sq(adj) = 90,22%

Source	Estimated Value
batch(fat cat.)	0,01656
Error	1.39038

Table 5.2.8: Variance analysis results and variance component for model 5.2.4

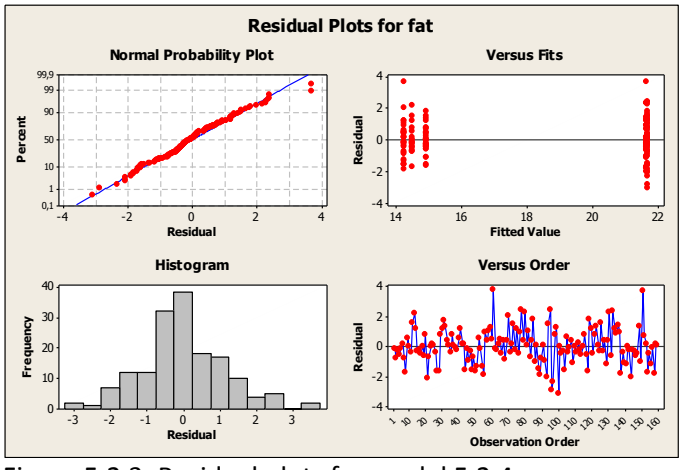


Figure 5.2.3: Residual plots for model 5.2.4

According to the results from the FOODSCAN NIR instrument, I made a table to show the difference between the FOODSCAN instrument which is the reference method as AASHEIM and STABBURET agreed, and the Q-monitor at AASHEIM KJØTT. This tables shows that there are positive differences for all the batches. That is FOODSCAN shows fat content measurements larger than Q-monitor. If we look to the variance column, we see that there are large variances between the samples, because the mechanism of the Q-monitor instrument, hence very careful sample units collecting required to obtain the most representative sample units.

Batch	date	N	Mean	SE Mean	StDev	Variance	Q-MONITOR	differance
1	17.01.2013	20	14,47	0,20	0,88	0,78	14,10	0,37
2	17.01.2013	20	21,68	0,22	0,98	0,96	20,90	0,78
3	24.01.2013	30	14,20	0,22	1,19	1,42	14,10	0,10
4	24.01.2013	30	21,68	0,28	1,52	2,32	21,00	0,68
5	31.01.2013	30	14,90	0,15	0,84	0,70	14,00	0,90
6	31.01.2013	30	21,65	0,24	1,34	1,79	20,90	0,75

Table 5.2.9: The difference between the FOODSCAN and Q-monitor

5.2.2 THE VISIT TO STABBURET 19.2.2013

At the 19th of February 2013, I visited Stabburet at Råbakken in Fredrikstad. I wanted to follow the process to test the fat content using the FA DEXA x-ray machine. Together with Mr.Steinar Schie, we registered all values from the FA DEXA for all batches from Åsheim Kjøtt AS, (pictures 5.2.1, 5.2.2).
The model used for this data will have both cross and nested factors. We analyzed the both fat category (14%, 21%) with all the instruments. This is the reason why fat category is crossed with instruments. Of course the batches are nested inside the fat category.
When we have crossed factors, we can have model with interaction between the factors. The interaction is when the effect of a one factor depends on the level of the other factor. That is the performance of one instrument dependent of the fat category. In such situation, the researcher should advice the producer to use specific instrument for specific fat category. The analysis of the data show us that the interaction term not significant, so we can ignore the interaction effect and refit the model without the interaction term.

Picture 5.2.1

Picture 5.2.2

We fit a model with the fixed factor instrument and the interaction term between the fat category and the instrument. The model we fit is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + fat.cat. * instrument_{ij} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \tau\beta_{ij} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2 \\ k = 1, 2, \cdots, 6 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.5)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\sum_{i=1}^{2} \sum_{j=1}^{2} \tau\beta_{ij} = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent.

$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	5427,06	5089,98	5089,98	630,38	0,000
instrument	1	51,42	56,62	56,62	76,19	0,000
fat cat.*instrument	1	0,08	0,03	0,03	0,04	0,835
batch(fat cat.)	4	34,41	34,41	8,60	11,57	0,000
Error	427	317,34	317,34	0,74		
Total	434	5830,30				

S = 0,862075 R-Sq = 94,56% R-Sq(adj) = 94,47%

Source	Estimated Value
batch(fat cat.)	0,1103
Error	0,7432

The interaction effect not significant

Table 5.2.10: Variance analysis results and variance component for model 5.2.5

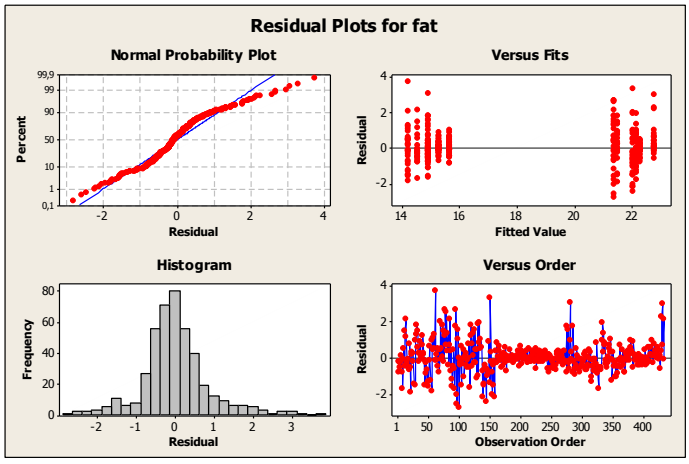


Figure 5.2.4: Residual plots for model 5.2.5

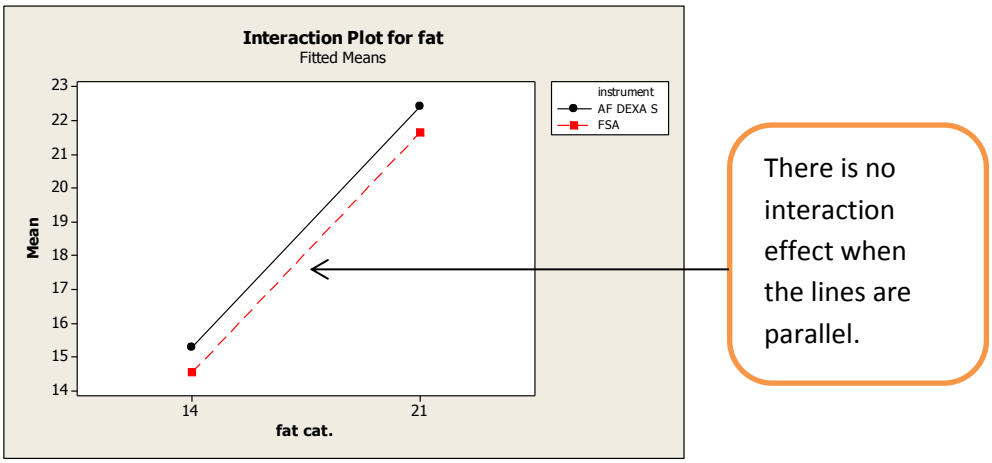


Figure 5.2.5: Interaction plot between fat category and instruments

From the results in table 5.2.10 and figure 5.2.5, we can see that the interaction is not significant. We can ignore the interaction term and refit the model without this term.

It is worth mentioning that halve part of the last 21% batch went to the production wrongly before we analyzed it with FA DEXA. The worker at STABBURET used the halve part of the in the production before the analysis. This is an example of the problems during the experiment and the experimenter has to live with it.

The model without the interaction term is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1,2 \\ j = 1,2 \\ k = 1,2,\cdots,6 \\ l_j = 1,\cdots,20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.6)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent.
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	5427,06	5441,07	5441,07	637,37	0,000
instrument	1	51,42	56,64	56,64	76,38	0,000
batch(fat cat.)	4	34,45	34,45	8,61	11,61	0,000
Error	428	317,37	317,37	0,74		
Total	434	5830,30				

S = 0,861111 R-Sq = 94,56% R-Sq(adj) = 94,48%

Source	Estimated Value
batch(fat cat.)	0,1098
Error	0,7415

Table 5.2.11: Variance analysis and variance component for model 5.2.6

5.2.3 FIT THE MODEL WITH THE DATA FROM EUROFINS

AASHEIM KJØTT uses the chemical instrument at EUROFINS to check the accuracy of the Q-monitor by sending some sample units from some batches. The used model is model 5.2.6. The difference is in the number of the used instruments. Now we have three instruments instead for two.

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2, 3 \\ k = 1, 2, \cdots, 6 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.7)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent.
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	7497,1	7481,9	7481,9	457,21	0,000
instrument	2	54,6	61,5	30,8	22,23	0,000
batch(fat cat.)	4	65,7	65,7	16,4	11,87	0,000
Error	587	812,5	812,5	1,4		
Total	594	8430,0				

S = 1,17654 R-Sq = 90,36% R-Sq(adj) = 90,25%

Source	Estimated Value
batch(fat cat.)	0,1534
Error	1,3842

Table 5.2.12: Variance analysis and variance component for model 5.2.7

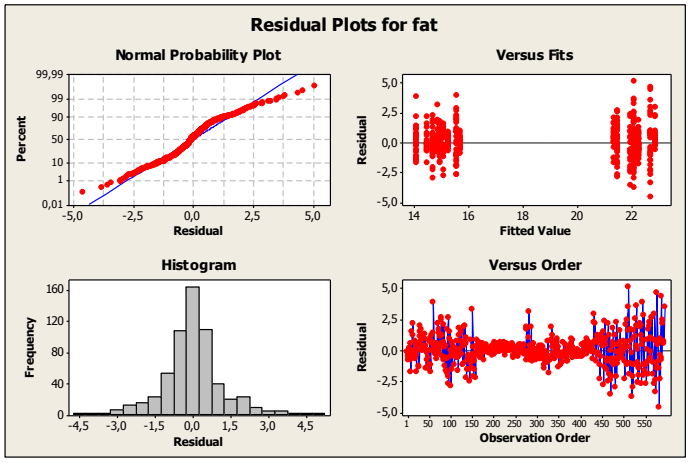


Figure 5.2.6: Residual plots for model 5.2.7

Results for instrument = EUROFINS

Variable	fat cat.	N	Mean	SE Mean	StDev	Variance
fat	14	80	15,126	0,169	1,511	2,283
	21	80	22,299	0,234	2,093	4,380

Results for instrument = FA DEXA

Variable	fat cat.	N	Mean	SE Mean	StDev	Variance
fat	14	150	15,270	0,0487	0,597	0,356
	21	125	22,356	0,0706	0,789	0,623

Results for instrument = FSA

Variable	fat cat.	N	Mean	SE Mean	StDev	Variance
fat	14	80	14,530	0,115	1,029	1,059
	21	80	21,671	0,148	1,320	1,743

Table 5.2.13: Descriptive statistics for each instrument

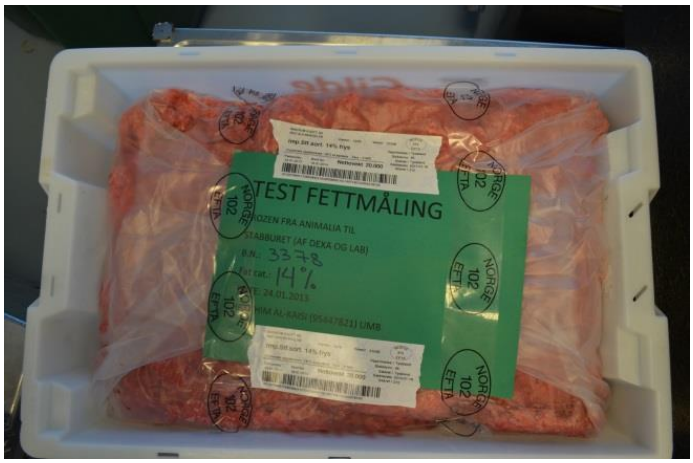
The table 5.2.13 contains some descriptive statistics. These statistics help the representatives from STABBURET and AASHEI KJØTT to understand the difference between the instruments. We could find the estimates from the model and we will do when we have the final model with all the instruments. In this table, we ignore the variance of the batches hence we underestimate the variance.

This descriptive statistic shows the performance of the EUROFINS instrument is not so good. The variance is so high and the average is over the average from the Q-monitor. It could be useless to use this instrument as a reference for the Q-monitor.

5.2.4 THE ANALYSIS OF FA DEXA WITH REPLICATES:

In the end of this part of the experiment, we obtained six plastic packages which contained the sample units from AASHEIM KJØTT and I analyzed at ANIMALIA these packages are marked with green labels (picture 5.2.1). There are another six packages contained the sample units which have been send from AASHEIM KJØTT to STABBURET, these are marked with yellow labels. The sample unites from the boxes with the green labels and the yellow labels are identical with only difference that the sample unites with the green labels are homogenized with the meat blender while the sample units with the yellow labels are in the original form from AASHEIM KJØTT.

To test the accuracy of the AF DEXA x-ray instrument at STABBURET, we tested these twelve boxes with this instrument with replicates. Each box has been tested six times and then let the instrument to calibrate itself and tested the same box more six times hence we obtain twelve replicates for each box. We analyzed the data the results as following:



Picture 5.2.1

To test the accuracy of the AF DEXA x-ray instrument at STABBURET, we analyzed these twelve packages with this instrument with replicates. Each package has been tested six times and then let the instrument to calibrate itself and analyzed the same package more six times hence we obtain twelve replicates for each package. We analyzed the data the results as following:

$$fat_{ijkl} = \mu + fat.cat._i + batch_{j(i)} + sample_{k(j)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_{j(i)} + \gamma_{k(j)} + \epsilon_{(ijk)l} \begin{cases} i = 1,2 \\ j = 1,\cdots,6 \\ k = 1,2,\cdots,12 \\ l = 1,2,\cdots,12 \end{cases} \quad (5.2.8)$$

Where:

$\sum_{i=1}^{2} \tau_i = 0$

$\beta_{j(i)} \sim N(0, \sigma_\beta^2)$ we assume that batches are independent.

$\gamma_{k(j)} \sim N(0, \sigma_\gamma^2)$ we assume that samples are independent.

$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch, sample and error are independent from each other.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat.cat	1	1826,85	1826,85	1826,85	168,81	0,000
batch(fat.cat)	4	43,29	43,29	10,82	7,92	0,014
sample(fat.cat batch)	6	8,19	8,19	1,37	18,27	0,000
Error	132	9,86	9,86	0,07		
Total	143	1888,20				

S = 0,273365 R-Sq = 99,48% R-Sq(adj) = 99,43%

Source	Estimated Value
batch(fat.cat)	0,39402
sample(fat.cat batch)	0,10757
Error	0,07473

This is the variance between the replicates

Table 5.2.14: Variance analysis and variance components for model 5.2.8

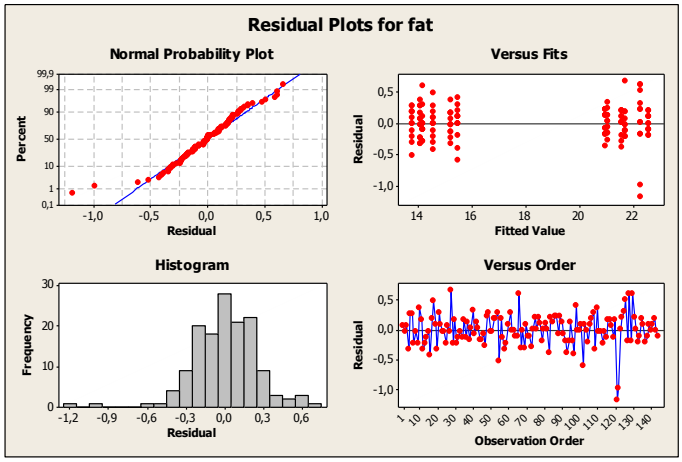


Figure 5.2.7: Residual plots for model 5.2.8

From table 5.2.14, we noticed that the variance between the replicates is very low. This can be index that this instrument is accurate instrument. Because of the high correlation between the replicates, we cannot consider this experiment as one of the instruments and have it in the variance analysis. This experiment (FA DEXA with replicates) are useful when compare the estimated mean for this experiment with the mean from the other instruments. The following two plots show that the FOODSCAN at ANIMALIA which is the neutral reference instrument and the AF DEXA with replicates are very near to each other.

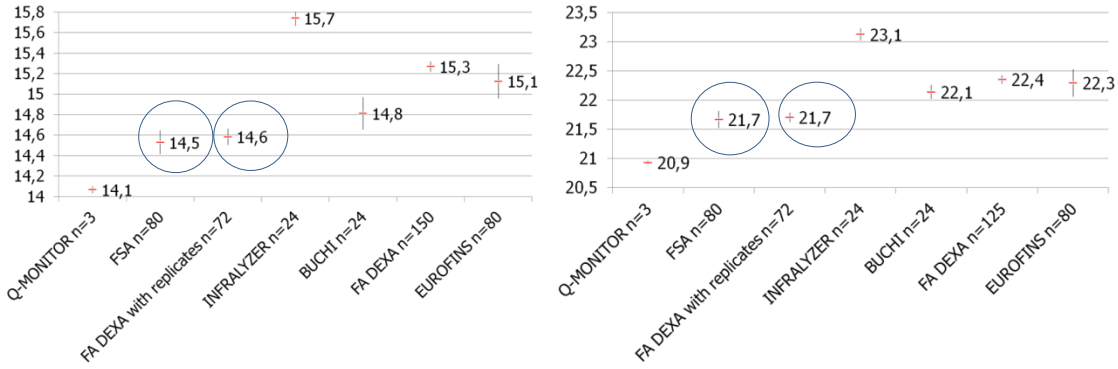


Figure 5.2.8: The estimated means for all instruments and for each fat category

Another informative analysis is to use this instrument with the packages with different labels. The difference between packages with different labels is that the packages with green labels are homogenized with the meat blender while the packages with yellow labels are grinded meat. In the next two plots, I analyzed the packages which have been analyzed and homogenized at ANIMALIA (green labels) and the packages which have been analyzed at STABBURET (yellow labels) with the FA DEXA x-ray instrument, we noticed that mean values for the both kind of packages are very close to each other, that means the homogenized level does not affect the AF DEXA measuring.

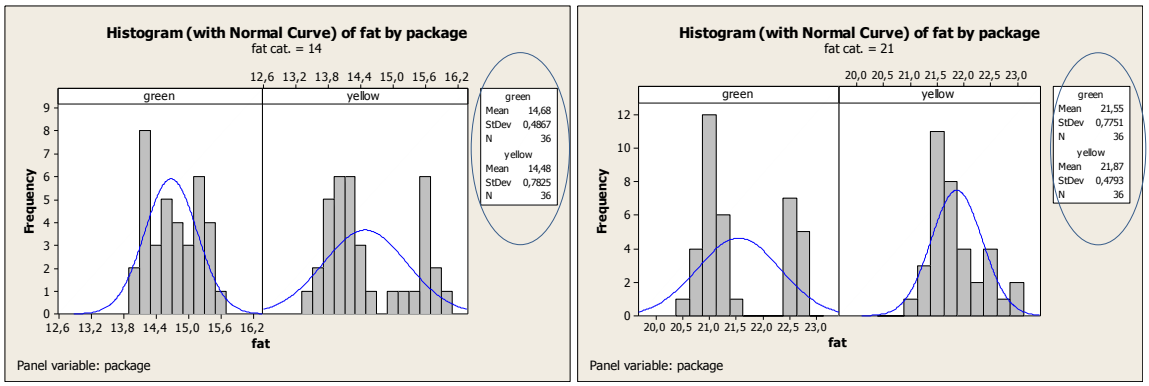


Figure 5.2.9: Histogram and normal curve of fat by packages

In next two plots and table6, the packages with green label have tested with FOODSCAN and FA DEXA. The results are very near to each other and this could be index that the FA DEXA instrument is closer to the FOODSCAN reference instrument than Q-monitor at AASHEIM KJØTT.

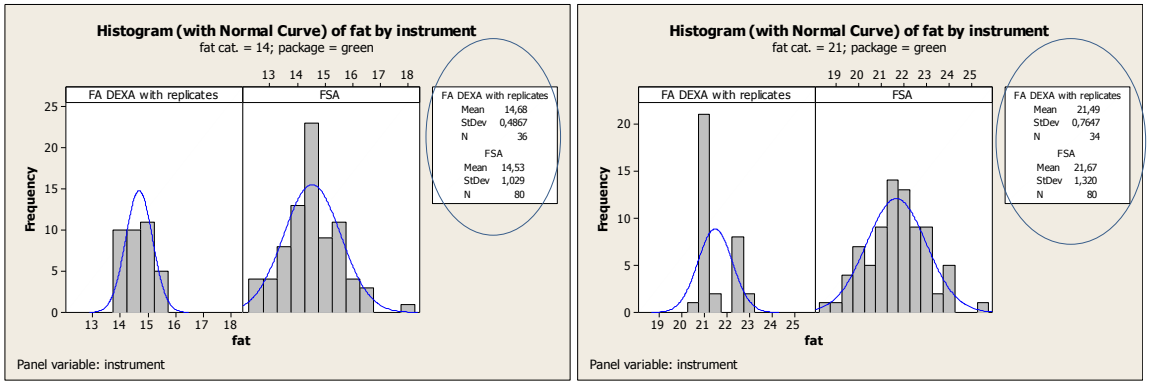


Figure 5.2.10: Histogram with normal curve of by instrument for the green package

Results for fat cat. = 14; package = green

Variable	instrument	N	Mean	SE Mean	StDev
fat	FA DEXA with replicates	36	14,683	0,0811	0,487
	FSA	80	14,530	0,115	1,029

Results for fat cat. = 21; package = green

Variable	instrument	N	Mean	SE Mean	StDev
fat	FA DEXA with replicates	34	21,494	0,131	0,765
	FSA	80	21,671	0,148	1,320

Table 5.2.15: Descriptive statistics for green packages

5.2.5 THE RESULTS FROM FIVE INSTRUMENTS

Now the results five instrument is ready to use. Two of these instruments are chemical instruments and three are NIR instruments, and we compared these instruments with the measurements from Q-monitor.

The model is unbalanced model because we have different number of measurements from each instrument. This model is mixed model because we have both fixed and random factors and both crossed and nested design. The fat category and the instrument are fixed factors because we want to say something about the specific fat category and the specific instruments, while the batches are selected randomly and hence this factor is random. We have the fat category and the instrument factors are crossed because we used all the instruments with each fat category, while the batch factor is nested within the fat category that is the batches 1, 3 and 5 belong to the 14% fat category and the batches 2, 4 and 6 belong to the 21% fat category. The model is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2, \cdots, 5 \\ k = 1, 2, \cdots, 6 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.9)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent.
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

y_{ijkl}: Is the fat content for sample unit from the fat category (i), instrument (j), batch (k) and the residual (l)

The mean (μ) is the fat content for all the meat which have been sent from AASHEIM KJØTT to STABBURET.

τ_i : Is the effect of the fat category (i).

β_j: Is the effect of the instrument (j).

γ_k: Is the effect of the batch (k).

$\epsilon_{(ijk)l}$: Is the error term.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	8806,58	8795,17	8795,17	440,47	0,000
instrument	4	87,69	94,27	23,57	19,34	0,000
batch(fat cat.)	4	80,08	80,08	20,02	16,43	0,000
Error	681	829,93	829,93	1,22		
Total	690	9804,2				

S = 1,10395 R-Sq = 91,53% R-Sq(adj) = 91,42%

Source	Estimated Value
batch(fat cat.)	0,1647
Error	1,2187

Table 5.2.16: Variance analysis and variance component results for model 5.2.9

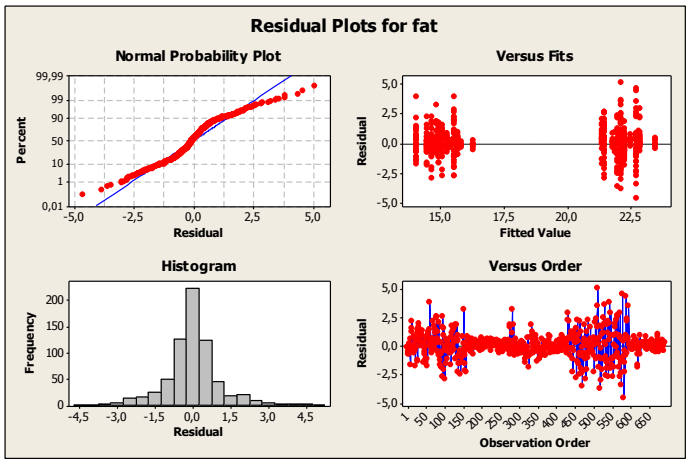


Figure 5.2.11: Residual plots for model 5.2.9

```
Results for instrument = BUCHI
            fat
Variable    cat.    N     Mean   SE Mean   StDev  Variance
fat         14     24   14,812     0,156   0,765     0,585
            21     24   22,143     0,119   0,585     0,342


Results for instrument = EUROFINS
            fat
Variable    cat.    N     Mean   SE Mean   StDev  Variance
fat         14     80   15,126     0,169   1,511     2,283
            21     80   22,299     0,234   2,093     4,380


Results for instrument = FA DEXA
            fat
Variable    cat.     N     Mean   SE Mean   StDev  Variance
fat         14     150   15,270    0,0487   0,597     0,356
            21     125   22,356    0,0706   0,789     0,623


Results for instrument = FSA
            fat
Variable    cat.    N     Mean   SE Mean   StDev  Variance
fat         14     80   14,530     0,115   1,029     1,059
            21     80   21,671     0,148   1,320     1,743


Results for instrument = INFRALYZER
            fat
Variable    cat.    N     Mean   SE Mean   StDev  Variance
fat         14     24   15,743    0,0758   0,371     0,138
            21     24   23,131     0,105   0,514     0,264
```

Table 5.2.17: Descriptive Statistics for five instruments

From the results in table 5.2.17, we notice that the results from the entire instruments show estimated mean higher than the mean from Q-monitor from AASHEIM KJØTT. One more important notice that the responsible in the laboratory at STABBURET does not believe that the INFRAALYZER NIR instrument is accurate. From the analyze results, we see that this instrument has the lowest variance, this could be index that this instrument is accurate but the average is much higher the other instruments. The conclusion could be that this instrument is accurate but need new recalibration.

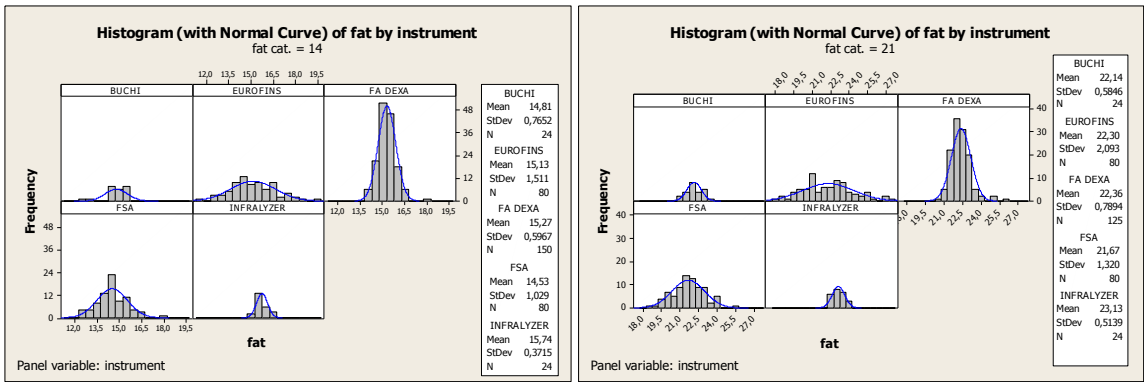


Figure 5.2.12: Histogram with the normal distribution for each fat category

The last two plots show the histogram with the normal distribution curve and we can see that the EUROFINS instrument has the widest curve corresponding to the largest variance and the INFRAALYZER has the narrowest curve corresponding to the lowest variance.

We wanted to see the average according to the fat category for all the samples and with all the instruments.

Variable	fat cat.	N	Mean	SE Mean	StDev	Variance
fat	14	358	15,073	0,0542	1,026	1,053
	21	333	22,218	0,0750	1,369	1,873

Table 5.2.18: Descriptive Statistics for each fat category over all instruments

The table (5.2.18) show the total sample mean for each fat category. STABBURET according to this table paied for more than one prosent extra in those six batches.

fat category	instrument	N	Mean	SEMean	StDev	Variance	Q-monitor	differance
14	BUCHI	24	14,81	0,16	0,77	0,59	14,07	0,74
14	EUROFINS	80	15,13	0,17	1,51	2,28	14,07	1,06
14	FA DEXA	150	15,27	0,05	0,60	0,36	14,07	1,20
14	FA DEXA with replicates	72	14,58	0,08	0,65	0,43	14,07	0,51
14	FSA	80	14,53	0,12	1,03	1,06	14,07	0,46
14	INFRALYZER	24	15,74	0,08	0,37	0,14	14,07	1,67
14	Q-MONITOR	3	14,07	0,03	0,06	0,00	14,07	0,00
21	BUCHI	24	22,14	0,12	0,58	0,34	20,93	1,21
21	EUROFINS	80	22,30	0,23	2,09	4,38	20,93	1,37
21	FA DEXA	125	22,36	0,07	0,79	0,62	20,93	1,43
21	FA DEXA with replicates	72	21,71	0,08	0,66	0,44	20,93	0,78
21	FSA	80	21,67	0,15	1,32	1,74	20,93	0,74
21	INFRALYZER	24	23,13	0,10	0,51	0,26	20,93	2,20
21	Q-MONITOR	3	20,93	0,03	0,06	0,00	20,93	0,00

Table 5.2.19: Results for all the instruments and the differences with Q-monitor

Table 5.2.19 shows the results for each instrument and the difference between these instruments and the Q-monitor. We notice that all instruments are over Q-monitor in the estimated mean.

The following two plots show the spreading of the observations which are the samples analysis for every instrument and for each fat category (14%, 21%). I noticed that the EUROFINS instrument has

the largest spreading or variance. In the other hand both AF DEXA and INFRAALYZER have the smallest spreading. The spreading for FA DEXA with replicates is not correct because much of the values are much correlated replicates.

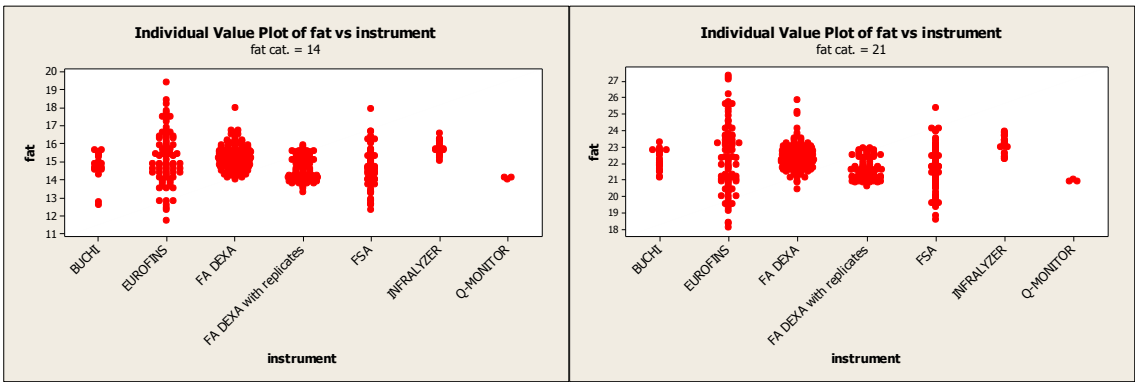


Figure 5.2.13: Values spreading plots for the all instruments and for each fat category

5.2.6 COMPARE Q-MONITOR WITH THE OTHER INSTRUMENTS

The used model is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2, \cdots, 6 \\ k = 1, 2, \cdots, 6 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.10)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

y_{ijkl}: Is the fat content for sample unit from the fat category (i), instrument (j), batch (k) and the residual (l)

The mean (μ) is the fat content for all the meat which have been sent from AASHEIM KJØTT to STABBURET.

τ_i : Is the effect of the fat category (i).

β_j: Is the effect of the instrument (j).

γ_k: Is the effect of the batch (k).

$\epsilon_{(ijk)l}$: Is the error term.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	8875,53	8865,80	8865,80	448,62	0,000
instrument	5	95,49	102,11	20,42	16,86	0,000
batch(fat cat.)	4	79,25	79,25	19,81	16,36	0,000
Error	686	830,90	830,90	1,21		
Total	696	9881,16				

S = 1,10056 R-Sq = 91,59% R-Sq(adj) = 91,47%

Term	Coef	SE Coef	T	P
Constant	18,5081	0,0869	212,89	0,000
fat cat.				
14	-3,58727	0,04193	-85,56	0,000
instrument				
BUCHI	-0,0304	0,1561	-0,19	0,846
EUROFINS	0,1862	0,1124	1,66	0,098
FA DEXA	0,3496	0,1028	3,40	0,001
FSA	-0,4261	0,1124	-3,79	0,000
INFRALYZER	0,9288	0,1561	5,95	0,000
(fat cat.)batch				
14 1	-0,01488	0,08474	-0,18	0,861
14 2	-0,42504	0,08102	-5,25	0,000
21 4	-0,27624	0,08658	-3,19	0,001
21 5	-0,17139	0,08260	-2,07	0,038

There is significant difference between the instruments

We can use these coefficients to estimate the means for each instrument

Table 5.2.20: Variance analysis and coefficients for model 5.2.10

We can use the coefficients to find the estimated mean for each instrument if we do not have interaction between the instrument and the fat category. We investigated the interaction between the instruments and the fat category and the results in the next table.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat cat.	1	10689,33	2714,88	2714,88	333,41	0,000
instrument	6	124,20	132,94	22,16	21,35	0,000
batch(fat cat.)	4	116,00	115,77	28,94	27,88	0,000
fat cat.*instrument	6	1,29	1,29	0,21	0,21	0,975
Error	823	854,23	854,23	1,04		
Total	840	11785,05				

The interaction term is not signifecant

Table 5.2.21: Variance analysis for the model with interaction term

From the variance analysis to the model with the interaction term, we found that the interaction term is not significant hence we can use the coefficients in the table 5.2.20 to estimate the mean for each instrument.

$$\hat{\mu}_{instrument} = \hat{\mu} + \hat{\tau}_{fat.cat} + \hat{\beta}_{instrument} \qquad (5.2.1)$$

We use this equation to find the estimated mean for each instrument and for 14% fat category.

$$\hat{\mu}_{BUCHI} = 18.51 - 3.59 - 0.03 = 14.89$$

$$\hat{\mu}_{EUROFINS} = 18.51 - 3.59 - 0.18 = 14.74$$

$$\hat{\mu}_{FADEXA} = 18.51 - 3.59 - 0.35 = 14.57$$

$$\hat{\mu}_{FSA} = 18.51 - 3.59 - 0.43 = 14.49$$

$$\hat{\mu}_{INFRAALYZER} = 18.51 - 3.59 + 0.93 = 15.85$$

$\hat{\mu}_{Q-monitor} = 18.51 - 3.59 - 1.01 = 13.91$

By the same way, we can find the estimated means for 21% fat category.

The most important result in table 5.2.20 is that there is differance between the instruments according to p-value for the factor instrument, which is smaller than (0.05) level of significance. The next step is where these differences are. We can use the next table 5.2.21 to investigate if there are differences between the instruments means. This table is the results from Tukey's method to test all possible instrument pairs. As we discussed in part 4.2.7, we preferred this method because the overall significance level at most α when the sample sizes are unequal.

instrument = BUCHI subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
EUROFINS	0,2166	0,1813	1,195	0,8396
FA DEXA	0,3800	0,1724	2,204	0,2356
FSA	-0,3957	0,1813	-2,183	0,2458
INFRALYZER	0,9592	0,2247	4,270	0,0003
Q-MONITOR	-0,9777	0,4766	-2,052	0,3131

instrument = EUROFINS subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
FA DEXA	0,163	0,1103	1,482	0,6761
FSA	-0,612	0,1230	-4,976	0,0000
INFRALYZER	0,743	0,1813	4,096	0,0006
Q-MONITOR	-1,194	0,4577	-2,609	0,0949

instrument = FA DEXA subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
FSA	-0,776	0,1103	-7,033	0,0000
INFRALYZER	0,579	0,1724	3,360	0,0102
Q-MONITOR	-1,358	0,4543	-2,989	0,0334

instrument = FSA subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
INFRALYZER	1,3548	0,1813	7,474	0,0000
Q-MONITOR	-0,5821	0,4577	-1,272	0,8006

instrument = INFRALYZER subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
Q-MONITOR	-1,937	0,4766	-4,064	0,0007

Table 5.2.22: Tukey Simultaneous Tests to compare the instruments

instrument = FA DEXA with replicates subtracted from:

instrument	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
FSA	-0,0658	0,1169	-0,563	0,9978
INFRALYZER	1,2917	0,1693	7,629	0,0000
Q-MONITOR	-0,6451	0,4233	-1,524	0,7303

Table 5.2.23: Tukey Simultaneous Tests to compare FA DEXA with replicate and FSA and Q-monitor

The first part of the table 5.2.22 is compare BÜCHI instrument with the other instruments. The numbers in the column "difference of means" are the results when we subtract the estimated mean

of BÜCHI from the estimated means of each instrument. From the p-value column, we noticed that BÜCHI is significantly different from only INFRAALYZER.

The most important process is to test the difference between FA DEXA, Q-monitor and FOODSCAN. This information is in the third part in the table 5.2.22. We noticed that the difference between FA DEXA and FOODSCAN is highly significant. As we discussed before, we used FOODSCAN with 80 selected sample units while the FA DEXA used to analyze the entire amount of meat in each batch. When we analyzed the same number of sample units with both FA DEXA and FOODSCAN, the results was much near to each other as we see in the table 5.2.23 in the experiment FA DEXA with replicate. Therefore the most important process in this kind of experiments is to select the most representative sample units.

The difference between FOODSCAN and Q-monitor according to the p-value is not significant. According to this data set FOODSCAN is nearest to Q-monitor from FA DEXA.

5.2.7 THE SECOND ROUND OF DATA COLLECTION

I represented the results from the first round of data collection to the representatives from AASHEIM KJØTT, STABBURET and TOMRA. The specialists in Q-monitor from TOMRA advised to collect more data because they claim that six batches are not enough to make correct conclusions. AASHEIM KJØTT, STABBURET and TOMRA cooperated and selected data from seventeen batches. Eleven batches from 14% fat content and six batches from 21% fat content. I obtained this data few days before I represent my thesis.
The used model is:

$$fat_{ijkl} = \mu + fat.cat._i + instrument_j + batch_{k(i)} + Error_{(ijk)l}$$

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_{k(i)} + \epsilon_{(ijk)l} \begin{cases} i = 1, 2 \\ j = 1, 2, \cdots, 5 \\ k = 1, 2, \cdots, 17 \\ l_j = 1, \cdots, 20 \ or \ 25 \ or \ 30 \ or \ 50 \end{cases} \quad (5.2.11)$$

Where

$\sum_{i=1}^{2} \tau_i = 0$
$\sum_{j=1}^{2} \beta_j = 0$
$\gamma_{k(i)} \sim N(0, \sigma_\gamma^2)$ we assume that batches are independent
$\epsilon_{(ijk)l} \sim N(0, \sigma^2)$ we assume that residuals are independent.

We assume that batch and error are independent from each other.

y_{ijkl}: Is the fat content for sample unit from the fat category (i), instrument (j), batch (k) and the residual (l)

The mean (μ) is the fat content for all the meat which have been sent from AASHEIM KJØTT to STABBURET.

τ_i : Is the effect of the fat category (i).

β_j: Is the effect of the instrument (j).

γ_k: Is the effect of the batch (k).

$\epsilon_{(ijk)l}$: Is the error term.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
fat.cat.	1	17954,87	17954,87	17954,87	712,28	0,000
instrument	4	160,24	160,24	40,06	50,84	0,000
batch(fat.cat.)	15	378,11	378,11	25,21	31,99	0,000
Error	1067	840,75	840,75	0,79		
Total	1087	19333,97				

S = 0,887671 R-Sq = 95,65% R-Sq(adj) = 95,57%

Term		Coef	SE Coef	T	P
Constant		18,8370	0,0571	330,04	0,000
fat.cat.					
14		-4,25034	0,02816	-150,95	0,000
instrument					
Büchi		0,9611	0,1007	9,54	0,000
FADEXA		-0,26620	0,06120	-4,35	0,000
Foodscan		-0,31068	0,09355	-3,32	0,001
NIR		0,6909	0,1007	6,86	0,000
(fat.cat.)batch					
14	1	-0,4626	0,1058	-4,37	0,000
14	2	-0,2590	0,1058	-2,45	0,015
14	3	-0,0085	0,1058	-0,08	0,936
14	4	-0,1944	0,1058	-1,84	0,066
14	5	0,3330	0,1058	3,15	0,002
14	6	0,3481	0,1058	3,29	0,001
14	7	-0,9438	0,1058	-8,92	0,000
14	8	-0,8493	0,1058	-8,03	0,000
14	9	0,9834	0,1058	9,30	0,000
14	10	0,7439	0,1058	7,03	0,000
21	12	-0,8623	0,1013	-8,51	0,000
21	13	0,4021	0,1013	3,97	0,000
21	14	0,8251	0,1013	8,15	0,000
21	15	-0,6299	0,1013	-6,22	0,000
21	16	0,3679	0,1013	3,63	0,000

Table 5.2.24: Variance analysis and coefficients for model 5.2.11

We can use the coefficients in the table 5.2.24 to estimate the mean for each instrument.

$$\hat{\mu}_{instrument} = \hat{\mu} + \hat{\tau}_{fat.cat} + \hat{\beta}_{instrument}$$

We use this equation to find the estimated mean for each instrument and for 14% fat category.

$$\hat{\mu}_{BUCHI} = 18.84 - 4.25 + 0.96 = 15.55$$

$$\hat{\mu}_{FADEXA} = 18.84 - 4.25 - 0.27 = 14.32$$

$$\hat{\mu}_{FOODSCAN} = 18.84 - 4.25 - 0.31 = 14.28$$

$$\hat{\mu}_{NIR} = 18.84 - 4.25 + 0.69 = 15.28$$

$$\hat{\mu}_{Q-monitor} = 18.84 - 4.25 - 1.08 = 13.51$$

```
instrument = Büchi  subtracted from:

            Difference       SE of           Adjusted
instrument    of Means  Difference  T-Value   P-Value
FADEXA          -1,227      0,1119   -10,97    0,0000
Foodscan        -1,272      0,1444    -8,81    0,0000
NIR             -0,270      0,1522    -1,78    0,3881
Qvision         -2,036      0,2407    -8,46    0,0000


instrument = FADEXA  subtracted from:

            Difference       SE of           Adjusted
instrument    of Means  Difference  T-Value   P-Value
Foodscan       -0,0445      0,1010   -0,440    0,9922
NIR             0,9571      0,1119    8,555    0,0000
Qvision        -0,8089      0,2174   -3,720    0,0019


instrument = Foodscan  subtracted from:

            Difference       SE of           Adjusted
instrument    of Means  Difference  T-Value   P-Value
NIR             1,0016      0,1444    6,935    0,0000
Qvision        -0,7645      0,2358   -3,241    0,0105


instrument = NIR  subtracted from:

            Difference       SE of           Adjusted
instrument    of Means  Difference  T-Value   P-Value
Qvision         -1,766      0,2407   -7,337    0,0000
```

Table 5.2.25: Tukey Simultaneous Tests to compare the instruments

An important part in this table is the part which compares FA DEXA with FOODSCAN and Q-monitor. According to the p-values the difference between FA DEXA and FOODSCAN is not significant hence is no difference between these two instruments.

The next important part is the part for comparing FOODSCAN with Q-monitor. We noticed that the difference between these two instruments is not significant according to the p-value. From these two results, we can see that FA DEXA is closer to the reference instrument FOODSCAN.

```
Results for fat.cat. = 14
Variable  instrument    N    Mean  SE Mean  StDev  Variance
fat       Büchi        44  15,607    0,121  0,803     0,645
          FADEXA      550  14,309   0,0398  0,932     0,869
          Foodscan     55  14,227    0,127  0,942     0,888
          NIR          44  15,304    0,118  0,782     0,611
          Qvision      11  14,009   0,0368  0,122    0,0149

Results for fat.cat. = 21
Variable  instrument    N    Mean  SE Mean  StDev  Variance
fat       Büchi        24  23,940    0,108  0,528     0,279
          FADEXA      300  22,843   0,0825  1,428     2,040
          Foodscan     30  22,866    0,103  0,564     0,318
          NIR          24  23,730   0,0886  0,434     0,188
          Qvision       6  21,100    0,159  0,390     0,152
```

Table 5.2.26: Descriptive statistics for two fat categories and all instruments

The results from this table confirm that FA DEXA and FOODSCAN are very close to each other. These two instruments provide almost the same results. The second notice that the sample mean for these three instruments (FOODSCAN, FA DEXA and Q-monitor) are very close to each other in the 14% fat

category. There are near to two present differences between Q-monitor and the other two instruments at 21% fat category.

CHAPTER 6

DISCUSSION

This thesis was divided in two parts, the first part is about multivariate calibration and the second part is about variance analysis. In the first part we used two data sets from Dr. Wold at NOFIMA to illustrate what we discussed at the theoretical part. The second part is the practical part where I selected the sample units by myself and where I followed the batches from the sender (AASHEIM KJØTT) to the production line at (STABBURET). I used FOODSCAN T at ANIMALIA to analyze the sample units.

6.1 DISCUSSION ABOUT MULTIVARIATE CALIBRATION

In chapter 4.1, we discussed theoretically the multivariate calibration issue. Throw the use of these data sets, we noticed some comments:

1. The first step was to use the Principle Component Analysis to analyze the first data set. The data set has 35 observations. All the variables of this data matrix are on the same scale (only reflections, no fat), therefore we used the covariance matrix instead for the correlation matrix to find the components. The covariance matrix provides more information about the variance, but we cannot use if the elements of the data matrix not from the same scale.

2. The results from PCA show that the first component captured 96.16% of the variability, and the first four components captured almost all the variability.

3. We tried PCR with all possible 15 components. We used leave one out cross validation to obtain RMSEP. We noticed that the lowest value for the RMSEP at 3 components and the (R^2) are almost does not change after 3 components. That was an index that we can use only 3 components instead for the original data set with 15 variables.

4. When we plotted the first 2 components, we noticed that the first component distinguished between the observations with high fat content and the observation with low fat content.

5. We fitted a PLSR model, and the results were very close to the results from the PCR method. The reason is that big amount of the variability in the data matrix captured by the first component. This component is highly correlated to the fat content response. In this situation there is almost no difference which method we use. It is easier to understand, interpret and use PCR therefore we preferred to use PCR.

6. We used the PCR model with 3 components and the leave one out cross validation. We fitted a linear model with the cross validated predicted values as response and the observed values as the explanatory variable. We obtained ($R^2 = 0.985$) and due to value

of $\hat{\beta}_0$ we systematically over predict. Although this prediction model seemed to be suitable to use with second data set.

7. When we used the PCR model with 3 components on the test set, we obtained RMSEP value equal to (6.586%). This number is suspiciously high compared to the cross validation result in Data set 1. Therefore we plotted the cross validated predicted values against the observed values and we will plot the NIR spectrums. We founded that there is extreme observation which affect the result. We removed this observation and we obtained better model but still with high RMSEP. RMSEP for using PCR model with 3 components on the test data set is (4.89%) and ($R^2 = 0.94$).The variance inside the test data set because each sample unit has been scanned 4 times from 4 different surfaces. The variance between the surfaces for the same sample units could be huge. One side could be lean meat and the other side only fat. This fact makes us to accept this high RMSEP.

6.2 DISCUSSION ABOUT VARIANCE ANALYSIS

I selected sample units from six batches, three from 14%and three from 21%. I started to analyze the sample units step by step to discuss different issues. First I analyzed the sample units at ANIMALIA and then I obtained the results from STABBURET and the last results from EROFINS. After the data analysis, we can note some comments.

1. Each sample units was from 700g to 1000g and each analysis to FOODSCAN instrument at ANIMALIA was approximately 100g. Hence we could take several replicates from each sample unit. I analyzed two replicates from the half part of the selected sample units from the first visit to AASHEIM KJØTT. We found that FOODSCAN is accurate in the meaning that the replicates have very low variance. The variance between the sample units are much larger than the variance between the replicates (almost fifteen times larger), therefore it is better to select more sample units instead of replicates.
   The second notes is: If we want to take replicates, we have to take replicates from all the sample units instead for part of them because the statistical analyze program will refer wrongly to the estimated error (MS_{Error}) as the variance between the replicates and to find the correct estimated error we have to make another calculations and the model adequacy checking plots show a lack of fit. If assume that the replicates are sample units as the other observations, we underestimate the variance because we ignore the high correlation between the replicates from the same sample units but we obtain much better model.

2. I tried to investigate if there is any interaction effect between the fat category and the instruments. That is instruments could perform with one fat category better than the other one. The effect of the interaction was not significant at the 5% level of significant. The p-value was so high and the interaction very weak, so we removed it from the model.

3. AASHEIM KJØTT sometimes sends some sample units to EUROFINS to analyze them by the chemical instrument SBR to check the accuracy of the Q-monitor. However we found that for 80 sample units, this instrument has the largest variance among all the instruments. Also the total average for each fat category is much higher than the average from the Q-monitor instrument, so I think it is waste of time and money to send few sample units to EUROFINS.

4. STABBURET has a NIR instrument (INFRAALYZER), this instrument has very low variance and lower than the BÜCHI instrument but the average is much higher than the Q-monitor and the other instruments. This could be evidence that this instrument is very good, but needs recalibration. I think STABBURET could trust this instrument and use it more than BÜCHI instrument. This is because it is much cheaper and easier to use. BÜCHI instrument analyze just few grams and needs a trained staff to follow the analyze process. In addition it needs to use dangerous chemicals and long time to obtain the wanted results, while NIR instrument safe, easy to use, take only some seconds to provide the results and has good accuracy according to the low variance it has.

5. We used FA DEXA two to analyze each batch. The first use is for the whole batch and the second use for only the collected sample units for the batch. We called the second use, FA DEXA with replicates. When we analyzed the same sample units with both FOODSCAN at ANIMALIA and FA DEXA at STABBURET (FA DEXA with replicates), we noticed that the results are very close to each other. That is FA DEXA is almost identical to the instrument where both AASHEIM KJØTT and STABBURET agreed to be the reference instrument. When we used the FA DEXA to analyze the entire batch there was a different between FOODSCAN and FA DEXA.

6. All the instruments gave fat content average larger than Q-monitor. This could be an evidence for that Q-monitor needs recalibration. I have been also told that AASHEIM KJØTT has a FOODSCAN instrument but it has not been used to check the results of Q-monitor. In my opinion FOODSCAN would be much better to use compared to sending some sample units to EUROFINS. The advantage of FOODSCAN is that it is very easy to use and it does not need specially trained staff and provides accurate results if few seconds. The Q-monitor seller (TOMRA SORTING SOLUTIONS) uses FOODSCAN to calibrate Q-monitor instrument.

7. After the meeting with the representatives from STABBURET, AASHEIM KJØTT and TOMRA the conclusion was to select data from larger number of batches. Sample units have been selected at the week 21 and 22 from seventeen new batches, eleven 14% and six 21%. The results from the analysis show that the difference between the FA DEXA and FOODSCAN is not significant, but the difference between these two instruments and Q-monitor is significant. That is the two instruments at STABBURET and ANIMALIA are very near to each other. It is worth mentioning that the instrument at ANIMALIA (FOODSCAN) is the reference instrument.

8. We notice also that when the selected sample units come from larger number of batches, the differences between FA DEXA, FOODSCAN and Q-monitor are very small at 14% fat content category. At the 21% fat content category, the difference between FA DEXA and FOODSCAN from one side and Q-monitor from the other side is around two present.

6.3 PROBLEMS

When we selected the sample units, we faced some problems.

1. The sample selection should be very fast, because the factory has to keep produce and sometimes it is not possible to select the sample units in the way we planned to do. In the start we planned to select one sample unit from each 20kg meat package going out of the Q-monitor instrument to obtain the most representative sample units from the entire 1000kg batch. This was impossible according to the production process at AASHEIM KJØTT therefore we selected some sample units from each 200kg meat container coming out of the 1000kg big container in the end of the Q-monitor instrument.

2. We lost one part of the data by a mistake in the first round of data selection. Data from half of one of the batches have been lost because it has been used in the production at STABBURET before the analysis by FA DEXA. I marked this part with red label with information about the batch and a note that this batch is not for production nevertheless the worker used this part in the production.

3. We have different instruments and each instrument returns different number of results for each batch hence we have a situation with unbalanced data. This unbalance makes the task of comparing the instruments with each other difficult. That is for one batch we have one result from Q-monitor, 20 or 30 results from FOODSCAN and SBR, 4 results from BÜCHI and INFRAALYZER and 50 results from FA DEXA.

4. The most important problem in my opinion is the organization part, because there are several companies involved in this experiment. Hence we should organize the date and the time which is appropriate for each place involved in this experiment to select the sample units. We should select one day where AASHEIM KJØTT would prepare and send batches to STABBURET. These batches should be in fresh form and from 14% and 21% fat category. The date should be appropriate for ANIMALIA to let me use the FOODSCAN instrument to analyze the selected sample units. When we achieved an appropriate date for both AASHEIM KJØTT and ANIMALIA, FA DEXA instrument at STABBURET was out of order for almost two months.
   These batches were kept in the freezer at STABBURET, marked with red labels with all the needed information to avoid send them to the production.
   The problem of disagreement between STABBURET and AASHEIM about fat content in the sent meat from AASHEIM KJØTT to STABBURET was the basic aim of this thesis. Back of this disagreement could be a huge amount of money annually. Each one of these two companies believe in the results from their own analyze instrument. There are several persons involved

in this thesis with different point of view for the problem that makes the organization of the sample selection plan and the illustration of this plan quit difficult.

6.4 CONCLUSIONS

In the end of this thesis, we can conclude.

1. The most important conclusion in my opinion is that the reference instrument (FOODSCAN) at ANIMALIA is closer to the FA DEXA at STABBURET than Q-monitor at AASHEIM KJØTT.

2. The results from the first and second data collection show that Q-monitor provides results with lower fat content than the other instruments. A new recalibration for Q-monitor according to the results from these two data collection, could improve the performance of this instrument.

3. The results from the first round of the data collection show that the INFRAALYZER NIR instrument at STABBURET is an instrument with very low variance hence high accuracy. The results from the second round of the data collection show that both BÜCHI and INFRALYZER are very close to each other. STABBURET could increase the use of this instrument after the appropriate re-calibration. NIR instruments are easy, fast, safer and cheaper to use.

4. The data analysis show that the small NIR instruments like FOODSCAN and INFRALYZER are good analysis instrument with low variance between the replicates. AASHEIM KJØTT has one FOODSCAN instruments which is not used to check the accuracy of Q-monitor, and prefer to send some samples to EUROFINS for this purpose. AASHEIM KJØTT could adopt this instrument as a reference instrument to check the measurements from Q-monitor after the appropriate re-calibration.

6.5 FURTHER STUDY

1. The results from the second round of data collection show that the differences between Q-monitor and the other instruments are higher at 21%fat category. For further study, I would like to include batches from higher fat content. I want to investigate if the differences become larger for the higher fat category.
2. I would like also to participate in any re-calibration process. I would like to use the knowledge from the multivariate calibration part in this thesis to learn to re-calibrate any instrument.
3. After the re-calibration process, I would like to collect more sample units to check if the re-calibration has changed the differences between the instruments.
4. I would like to work more with the unbalanced data set, outliers and investigate more the calibration method

APPENDIX

R-CODE

```
##Downloading the pachage for PLS##

library(pls)

## reading the data sets from clipboard ##

Dataset <- read.table("clipboard",dec=",")

Dataset_test <- read.table("clipboard")

## create a function to find RMSEP ##

rmsep <- function(y, ypredpcr){

sqrt(mean((y-ypredpcr)^2))

}

n <- dim(Dataset)[1]

## define the response (fat) Y and the wavelengths matrixes X for##

##the calibration data set                                        ##

Y <- as.matrix(Dataset[,1,drop=FALSE])

X <- as.matrix(Dataset[,-1])

###Ytest <- as.matrix(Dataset_test[,16])#### tra bort rad 242

###Xtest <- as.matrix(Dataset_test[,-16])####tar rad nummer 242

## define the response (fat) Y and the wavelengths matrixes X ##

## for the test data set                                       ##

Ytest <- as.matrix(Dataset_test[-242,16])#### tra bort rad 242

Xtest <- as.matrix(Dataset_test[-242,-16])####tar rad nummer 242

## finding the predictied error vector ##

ypredpcr <- matrix(0,n,15)

predfeilvectorpcr<- rep(0,15)

useobs <- (1:n)

for(j in 1:15){

  for(i in useobs){

    fit <- pcr(Y~X, ncomp=15, data=Dataset, subset=-i)
```

```
    ypredpcr[i,j] <- predict(fit, ncomp=j,newdata=X[i,,drop=FALSE])
  }
  predfeilvectorpcr[j]<- rmsep(ypredpcr[,j],Y)
}
predfeilvectorpcr
### to find the predicted error for the null model ####
yprednullpcr <- rep(0,n)
for(i in 1:n){
  yprednullpcr[i] <- as.matrix(mean(Y[-i,]))
}
predfeilnullpcr <- rmsep(yprednullpcr,Y)
predfeilnullpcr
predfeilvectornullpcr <- c(predfeilnullpcr,predfeilvectorpcr)
predfeilvectornullpcr
## to plot the RMSEP against the number of the components ##
plot(1:16, predfeilvectornullpcr, xlab="number of components", type="b")
## ploting the cross validated predicted values against the observed values ##
plot(Y,ypredpcr[,3],ylab="yhat")
## puting the 45 degree line on the plot ##
curve(1*x,add=TRUE)
## make a simple linear model for Y-hat as a response and Y as explanatory variable ##
model<- lm(ypredpcr[,3]~Y)
## plot the model line on the Y-hat against Y plot ##
abline(lm(ypredpcr[,3]~Y))
## finding the value of R-squared ##
summary(mod)[c("r.squared", "adj.r.squared")]
## we will use the PCR model with 3 components on the test data set##
fit <- plsr(Y~X, ncomp=15, data=Dataset)
## findeing the cross validated predicted values ##
testypred <- predict(fit, ncomp=3, newdata=Xtest)[,1,1]
```

```
## finding RMSEP ##

rmsep(Ytest,testypred)

plot(Ytest,testypred,ylab="predicted fat content (%)",xlab="observed fat content (%)")

## we faound one extreme observation and we removed it ####

which(testypred<0)

## simple linear model for Y-hat and Y ##

model<- lm(testypred~Ytest)

abline(lm(testypred~Ytest))

summary(model)[c("r.squared")]

## RMSEP for test set ##

ntest <- dim(Xtest)[1]

ypredtestpcr <- matrix(0,ntest,15)

predfeilvectortestpcr<- rep(0,15)

useobstest <- (1:ntest)

rmsep <- function(y, ypredpcr){

  sqrt(mean((y-ypredpcr)^2))

}

for(j in 1:15){

  for(i in useobstest){

    fit <- plsr(Y~X, ncomp=15, data=Dataset, subset=-i)

    ypredtestpcr[i,j] <- predict(fit, ncomp=j,newdata=Xtest[i,,drop=FALSE])

  }

  predfeilvectortestpcr[j]<- rmsep(ypredtestpcr[,j],Ytest)

}

predfeilvectortestpcr

## to find the predicted error for the null model##

yprednulltestpcr <- rep(0,n)

for(i in 1:ntest){

  yprednulltestpcr[i] <- as.matrix(mean(Ytest[-i,]))

}
```

```
predfeilnulltestpcr <- rmsep(yprednulltestpcr,Ytest)

predfeilnulltestpcr

predfeilvectornulltestpcr <- c(predfeilnulltestpcr,predfeilvectortestpcr)

predfeilvectornulltestpcr

## to plot the RMSEP values against the number of components ##

plot(1:16, predfeilvectornulltestpcr, xlab="number of components", type="b")
```

DATA SETS

THE FIRST DATA SET

This is the calibration data set from Dr. Wold at NOFIMA. This data set is 35 rows and 15 columns. The follow table is part of this data set and the entire date in the CD.

fett	760	780	800	820	840	860	880	900	920	940	960	980	1000	1020	1040
2,1099999	-0,63614511	-0,80701232	-0,97352678	-0,99383217	-0,89905357	-0,81748748	-0,69075173	-0,50565743	-0,22878551	0,53088546	1,53111696	1,72870481	1,46320379	0,92099458	0,37734661
3,81999993	-0,32949334	-0,66390055	-0,97412425	-1,03274596	-0,94599682	-0,87097406	-0,7370286	-0,52865809	-0,21801068	0,59044713	1,63157701	1,7919029	1,44863975	0,76334143	0,07502417
5,92999983	-0,52151358	-0,78309327	-1,02545583	-1,04716194	-0,93176144	-0,83454448	-0,67604089	-0,44145438	-0,10763612	0,6303606	1,59763849	1,75647271	1,42259848	0,80335122	0,15824032
18,6600018	-0,64348769	-0,89403939	-1,10172188	-1,10109663	-0,99058729	-0,88715553	-0,66297537	-0,30259022	0,164739	0,71913213	1,44282341	1,57925546	1,38510668	0,92692256	0,36567482
18,4499989	-0,53705239	-0,82484603	-1,07067025	-1,08863819	-0,98186684	-0,88329679	-0,68779355	-0,37506101	0,05574366	0,69270772	1,51836061	1,65993309	1,40484369	0,86269516	0,25494123
31,1200008	-0,73752457	-0,97665018	-1,14804876	-1,11840653	-0,9994759	-0,89552027	-0,626297	-0,19491279	0,34934846	0,75324208	1,34810126	1,48198235	1,35552168	0,96053594	0,4481039
33,6399994	-0,76212478	-0,97003061	-1,12210429	-1,10114884	-0,98735046	-0,88883454	-0,63297409	-0,22638948	0,28258961	0,73029119	1,35579538	1,50106823	1,37328255	0,97737169	0,47055829
53,7400017	-0,82275289	-1,08786392	-1,22619426	-1,15959835	-1,05786896	-0,95785213	-0,57079798	0,10555199	0,86875266	0,85562086	1,03124774	1,14396858	1,19881392	1,05295002	0,62602276
66,2300034	-0,94562525	-1,17296827	-1,2446419	-1,14054	-1,05092311	-0,94622046	-0,46876121	0,33561566	1,23167336	0,86673194	0,76180178	0,85987419	1,03898203	1,08141565	0,79358572
82,1200027	-0,81624466	-1,13995183	-1,23239899	-1,11535275	-1,07398713	-0,99263209	-0,37936082	0,67422003	1,74435937	0,87414736	0,36555815	0,43604669	0,785258	1,04362142	0,82671732
17,8399982	-0,46963778	-0,78128701	-1,04851091	-1,07952523	-0,97452515	-0,88862389	-0,69990307	-0,40283358	0,01184107	0,68519503	1,56717539	1,69647551	1,4123708	0,81681103	0,15497769
7	-0,57149327	-0,80775708	-1,03008139	-1,04471648	-0,92753309	-0,83012122	-0,67469949	-0,43586504	-0,10701343	0,61790043	1,55057395	1,73240578	1,44106519	0,85126352	0,23607141
13,1899996	-0,57475233	-0,81411868	-1,03395259	-1,04833829	-0,94389439	-0,85251957	-0,68203741	-0,42541385	-0,05671464	0,63963187	1,54980326	1,71207237	1,4244734	0,85444421	0,25131691
26,1999989	-0,71310049	-0,94633001	-1,12462378	-1,09963381	-0,98037797	-0,87137097	-0,63286418	-0,248495	0,24122779	0,73356563	1,41635478	1,55836594	1,37906933	0,92350835	0,36470422
5,28000021	-0,36750808	-0,70806068	-1,01380289	-1,05162418	-0,94996524	-0,85913712	-0,70555645	-0,48068365	-0,15126996	0,63048911	1,63938797	1,78125632	1,4192282	0,74541283	0,07183407
55,8800011	-0,81315619	-1,09396088	-1,23732996	-1,16859472	-1,09521425	-0,9992125	-0,5434106	0,2423033	1,12434828	0,86267447	0,87085599	0,96976703	1,12027478	1,06754673	0,6931085
24,6200008	-0,63625103	-0,89019865	-1,0956775	-1,09282088	-0,9855482	-0,89191741	-0,66284484	-0,30687353	0,16205132	0,71519774	1,45216513	1,59887278	1,38190699	0,91495991	0,33697817
11,96	-0,61754471	-0,84605378	-1,04076564	-1,04568481	-0,92915958	-0,83356541	-0,6639967	-0,4089019	-0,0626015	0,63591796	1,53456068	1,71070659	1,42542684	0,86962867	0,27203348
6,70999956	-0,58752555	-0,82093149	-1,02681148	-1,03470874	-0,92335653	-0,82620883	-0,66644508	-0,44251654	-0,12112969	0,61672056	1,56166852	1,72781646	1,43368924	0,85640144	0,25333753
17,5400009	-0,66587436	-0,88487452	-1,06532776	-1,06035221	-0,94482994	-0,84550256	-0,65828198	-0,36414757	0,03837213	0,66516215	1,49725521	1,65293109	1,41564333	0,89858651	0,32124034
2,42000008	-0,66543341	-0,81856942	-0,97358048	-0,9938336	-0,89239031	-0,81500804	-0,68867952	-0,49869925	-0,21753605	0,53231555	1,51417243	1,71811712	1,47126436	0,93452889	0,39333165
1,88999999	-0,60190785	-0,79881668	-0,98551452	-1,00730884	-0,903449	-0,81641769	-0,68402737	-0,49591455	-0,21725851	0,53863376	1,53724205	1,74874246	1,4655602	0,89828598	0,32215053
5,78000021	-0,51386058	-0,78207952	-1,02364254	-1,05019414	-0,93726969	-0,84273332	-0,68145323	-0,45011464	-0,11811929	0,61923665	1,57285035	1,746575	1,43731678	0,83303779	0,19045043

THE SECOND DATA SET

This is the test data set from Dr. Wold at NOFIMA. This data set is 368 rows and 15 columns. The follow table is part of this data set and the entire date in the CD.

-8.9770949e-001	-1.0299382e+000	-1.1117295e+000	-1.0546281e+000	-9.0837544e-001	-8.0255014e-001	-
5.7355267e-001	-2.3429041e-001	2.1176371e-001	7.1249282e-001	1.4048520e+000	1.5476280e+000	
1.3597853e+000	9.4661254e-001	4.2963967e-001	2.3375000e+001			
-8.3117038e-001	-9.8944390e-001	-1.1024202e+000	-1.0561827e+000	-9.2511529e-001	-8.1577080e-001	-
5.9435469e-001	-2.6872945e-001	1.6633852e-001	6.8943578e-001	1.4365977e+000	1.5815481e+000	
1.3749269e+000	9.3449229e-001	3.9984810e-001	2.3375000e+001			
-8.8529056e-001	-1.0173148e+000	-1.1041509e+000	-1.0493155e+000	-9.1058093e-001	-7.9300267e-001	-
5.7067168e-001	-2.5813320e-001	1.5345985e-001	6.9184673e-001	1.4137893e+000	1.5793029e+000	
1.3646545e+000	9.4350863e-001	4.4189811e-001	2.3375000e+001			
-8.7731081e-001	-1.0115666e+000	-1.0936937e+000	-1.0361452e+000	-8.9901435e-001	-7.8378081e-001	-
5.7943553e-001	-2.7877235e-001	1.2332141e-001	6.7684722e-001	1.4275719e+000	1.6025515e+000	
1.3931937e+000	9.4090402e-001	3.9532971e-001	2.3375000e+001			

THE THIRD DATA SET

The first data set is the results of the analysis for 80 sample units from 6 batches. This data set is 842 rows and 7 columns. The follow table is part of this data set and the entire date in the CD.

fat	batch	fat cat.	instrumer	sample	batch anal	package
14,31	1	14	FSA	1	1	green
13,73	1	14	FSA	2	1	green
14,23	1	14	FSA	3	1	green
13,96	1	14	FSA	4	1	green
14,32	1	14	FSA	5	1	green
14,64	1	14	FSA	6	1	green
13,72	1	14	FSA	7	1	green
12,78	1	14	FSA	8	1	green
15,03	1	14	FSA	9	1	green
14,44	1	14	FSA	10	1	green
14,05	1	14	FSA	11	1	green
16,00	1	14	FSA	12	1	green
16,66	1	14	FSA	13	1	green
15,68	1	14	FSA	14	1	green
14,16	1	14	FSA	15	1	green
14,22	1	14	FSA	16	1	green
13,99	1	14	FSA	17	1	green
14,44	1	14	FSA	18	1	green
13,83	1	14	FSA	19	1	green
15,26	1	14	FSA	20	1	green
19,59	4	21	FSA	1	2	green
20,97	4	21	FSA	2	2	green
21,71	4	21	FSA	3	2	green
21,87	4	21	FSA	4	2	green

HE FOURTH DATA SET

The second data set is the results from the analysis for sample units from 17 batches. This data set is 1089 rows and 6 columns. The follow table is part of this data set and the entire date in the CD.

fat.cat.	batch	sample.nu.	fat	instrument	pekode
14	1	1	14,00	Qvision	1
14	2	1	13,70	Qvision	1
14	3	1	13,90	Qvision	1
14	4	1	14,00	Qvision	1
14	5	1	14,10	Qvision	1
14	6	1	14,10	Qvision	1
14	7	1	14,10	Qvision	1
14	8	1	14,10	Qvision	1
14	9	1	14,00	Qvision	1
14	10	1	14,10	Qvision	1
14	11	1	14,00	Qvision	1
21	12	1	20,60	Qvision	1
21	13	1	21,70	Qvision	1
21	14	1	21,30	Qvision	1
21	15	1	20,80	Qvision	1
21	16	1	21,20	Qvision	1
21	17	1	21,00	Qvision	1
14	1	1	14,59	Foodscan	2
14	1	2	13,20	Foodscan	2
14	1	3	14,66	Foodscan	2
14	1	4	14,61	Foodscan	2
14	1	5	14,30	Foodscan	2
14	2	1	14,95	Foodscan	2

BIBLIOGRAPHY

1. http://aasheimkjott.no/index.php?option=com_content&view=article&id=46&Itemid=55
2. http://www.animalia.no/Footer/Om-oss/Om-Animalia/
3. http://www.eurofins.com/en/about-us/eurofins-fact-sheet.aspx
4. http://www.odenberg.com/process-analytics/
5. Yukihiro Ozaki, W. Fred McClure, Alfred A. Christy, NEAR-INFRARED SPECTROSCOPY IN FOOD SCIENCE AND TECHNOLOGY.
6. http://www.eaglepi.com/fat_analysis.php
7. http://www.gerstel.de/pdf/p-gc-an-2009-02.pdf
8. http://www.unityscientific.com.au/imagesDB/products/IA2000.pdf
9. http://www.omnilexica.com/?q=sbr
10. Richard L. Scheaffer, William Mendenhall lll, R. Lyman Ott, ELEMENTARY SURVEY SAMPLING, FIFTH EDITION
11. PETER J. BICKEL and KJELL A. DOKSUM. MATHEMATICAL STATISTICS http://www.amazon.com/Mathematical-Statistics-Basic-Selected-Topics/dp/0132306379
12. http://en.wikipedia.org/wiki/Degrees_of_freedom_(statistics)
13. Douglas C. Montgomery, Design and Analysis of Experiments, seventh edition.
14. http://skeetersays.wordpress.com/2008/08/12/demystifying-statistics-on-the-interpretation-of-anova-effects/
15. Montgomery, Peck and Vining http://www.amazon.com/Introduction-Regression-Solutions-Probability-Statistics/dp/0470258306/ref=sr_1_3?ie=UTF8&qid=1371910498&sr=8-3&keywords=introduction+to+linear+regression+analysis+montgomery.
16. Trevor Hastie, Robert Tibshirani and J. H. Friedman, The element of statistical learning. http://www.amazon.com/The-Elements-Statistical-Learning-Prediction/dp/0387848576/ref=sr_1_fkmr0_2?ie=UTF8&qid=1372862717&sr=8-2-fkmr0&keywords=15.%09Trevor+Hastie%2C+Robert+Tibshirani+and+J.+H.+Friedman%2C+The+element+of+statistical+learning.
17. http://www.frank-dieterle.de/phd/2_8_1.html
18. https://en.wikipedia.org/wiki/File:GaussianScatterPCA.png
19. Harald Martens, Tormod Næs. Multivariate Calibration. http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471930474.html
20. http://blog.minitab.com/blog/adventures-in-statistics/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit