Measurement of Reduction Efficiency in Green Liquor Using a NIR Spectrometer

Josef Persson

Sustainable Process Engineering, masters level
2016

Luleå University of Technology
Department of Civil, Environmental and Natural Resources Engineering
Measurement of reduction efficiency in green liquor using a NIR spectrometer

Josef Persson
Preface
While writing the last words of this report I realized how many people have been involved in this project and how they have been crucial for the realization of the results. Without them I could not have finished my work!

I would like to thank my supervisor at Domsjö Fabriker AB, Mathias Dahlbäck, for the guidance and support given during the project. His help with planning the project were much needed, and I also appreciated to always have someone there for bouncing ideas off.

Krister Lundgren has also been a great help for me with many of the practical aspects of the project as well as being a great source of experience and insight for me to draw upon.

Ann-Christine Westin has also been helpful in teaching me how to do the manual chemical analysis and her advice concerning the practical aspects in the lab.

My thanks also go to the many process technicians and lab technicians that have kindly answers all my questions and helped me in my work.

My thanks also go to Elias Sundvall and Niklas Berg at Pulpeye for all the information, advice and technical support concerning the instrument, the Extract data program and the data base. They were also very helpful in pointing me in the correct direction for the literary research I would need to do.

I would also like to everyone at DomInnova who took me in as one of their own and made sure that I would feel at home with them. Special thanks go to Magnus Hörnsten who believed in me when I first contacted him about my thesis project and did a great deal of work in finding a project that would suit me.

Luleå, July 2016

Josef Persson
Abstract
Domsjö Fabriker has earlier installed a Near Infrared (NIR) spectrometer after one of their recovery boilers. The purpose is to monitor the reduction efficiency of the boiler and later do process optimization.

In this work calibration models for the instrument have been created. 108 green liquor samples have been extracted. A NIR spectrum was recorded for each sample and the samples were subsequently analyzed in laboratory for total alkali, sulfide and total sulfur. Several calibration models were created with multivariate data analysis and their performance and robustness were compared. The best model was able to predict reduction efficiency with a RMSEP of 2.75 percent units.

Moreover, models were created for prediction of total alkali with a RMSEP of 0.108 mol/l, sulfides with a RMSEP of 1.95 g/l, total sulfur with a RMSEP of 2.83 g/l and S/Na\textsubscript{2} ratio with a RMSEP of 0.022. The result is good enough that the instrument could be used to optimize the process and monitor process disturbances.
Sammanfattning
Domsjö fabriker har tidigare installerat ett så kallat Near Infrared (NIR) spektrometer efter en utav sina sodapannor. Syftet är att kunna övervaka reduktionsgraden på pannan för att i senare steg kunna optimera processen.


Förutom detta skapades modeller med RMSEP på 0,108 mol/l för prediktering av totalt alkali, 1,95 g/l för sulfider, 2,83 g/l för totalt svavel och 0,022 för S/Na₂ kvoten. Resultaten är tillräckligt bra för att instrumentet ska kunna användas i syfte att optimera processen och övervakning av processtörningar.
Content
Preface ........................................................................................................ i
Abstract .................................................................................................... ii
Sammanfattning ....................................................................................... iii
List of Figures .......................................................................................... v
List of Tables ............................................................................................. vi
Acronyms .................................................................................................. vi
1 Introduction ........................................................................................... 1
   1.1 Background ..................................................................................... 1
   1.2 Purpose ........................................................................................... 1
   1.3 Process overview of Domsjö Fabriker AB ....................................... 1
2 Literature study .................................................................................... 3
   2.1 The recovery process ...................................................................... 3
   2.2 The recovery boiler ......................................................................... 3
   2.3 Reduction of chemicals in the recovery boiler ............................... 4
   2.4 The sodium/sulphur ratio .............................................................. 5
   2.5 Total alkali ..................................................................................... 6
   2.6 Green liquor dregs ......................................................................... 7
   2.7 Generation of new pulping chemicals ......................................... 7
   2.8 Vibrational spectroscopy ............................................................... 9
   2.9 Near infrared spectroscopy ........................................................... 11
   2.10 NIRS in liquor analysis ............................................................... 12
   2.11 Earlier work .................................................................................. 13
   2.12 Sampling for developing a NIR-based analytical method .......... 13
   2.13 Multivariate data analysis and Principal Component Analysis (PCA) ........................................................................ 14
   2.14 Partial Least Squares method .................................................... 15
   2.15 PLS for NIR applications ............................................................ 17
   2.16 Diagnostic tools of PLS ............................................................... 19
   2.17 Validation ..................................................................................... 19
3 Method .................................................................................................. 21
   3.1 Sampling equipment ...................................................................... 21
   3.2 The automatic sampling sequence ............................................. 22
   3.3 Manual sampling method ............................................................ 22
   3.4 Sampling design ............................................................................ 22
List of Figures
Figure 1. Overview of the Domsjö Fabriker AB process. .............................................. 2
Figure 2. Titration curve of green liquor. .................................................................. 7
Figure 3. Five different measurement modes........................................................... 10
Figure 4. Example of interpretation of PLSR in the case of spectroscopically analysis of green liquor. 16
Figure 5. A dataset of a number of spectral curves before (top) and after (bottom) normalization. ... 18
Figure 6. Schematic diagram of sampling equipment.................................................. 21
Figure 7. Actual installation of the NIRS control system at mill................................... 21
Figure 8. Reduction efficiency over two years............................................................ 24
Figure 9. Total alkali versus total sulphur over two years.......................................... 25
Figure 10. Score plot (left) and corresponding loadings plot (right) of model 2 ............ 27
Figure 11. Score plot (left) and corresponding loadings plot (right) of the two first PCs of model 6... 27
Figure 12. Online measurements using model 2 from 21st May to 9th June. ............... 28
Figure 13. Online measurements using model 2 from 21st May to 9th June. ............... 29
Figure 14. Online measurement of total alkali using model 2 from 21st May to 9th June.. 29
Figure 15. Online measurements of reduction efficiency using model 6 from 21st May to 9th June.. 30
Figure 16. Online measurements of reduction efficiency using model 6 from 21st May to 9th June.. 30
Figure 17. Score plot of the two first PC:s................................................................................. 32
Figure 18. The abnormal samples classified according to the first two PCs. ................................. 33
Figure 19. Proposal of reconstruction of existing sampling equipment. ........................................ 37
Figure 20. Example of two samples and their corresponding NIR spectra.. ............................... 38

List of Tables

Table 1. Chemical properties shown to be measurable by NIRS................................................... 12
Table 2. Spectral regions chosen for each analysed component.................................................... 13
Table 3. Model descriptions .......................................................................................................... 22
Table 4. The normal variance of the green liquor.......................................................................... 25
Table 5. Summaries of the models................................................................................................... 26
Table 6. RMSEP of the first five models. ........................................................................................ 26
Table 7. PLS statistics of the models based on model 2................................................................. 31
Table 8. RMSEP between measured results and prediction of the other time period. .................. 31
Table 9. PLS statistics of the models based on model 6................................................................. 31
Table 10. RMSEP between measured results and prediction of the other time period. ............... 31

Acronyms

IR Infrared
NIR Near infrared
RE Reduction efficiency
NIRS Near infrared spectroscopy/spectrometer
PLS Partial least squares
PLSR Partial least squares regression
PC Principal component
RMSEP Root mean squared standard error of prediction
SEKAB Svensk Etanolkemi AB
1 Introduction

1.1 Background
The recovery boiler in paper mills and cellulose production plants is one of the most important parts of the chemical recovery process, without it the recovery of sulphur in active form would not be possible. Thus it is desirable to monitor its performance. This is often done by manual sampling and analysis of the product of the boiler, the green liquor. The reduction efficiency is one of the interesting parameters of the performance since it gives us an estimate of what percentage of the sulphur can be recovered and how much will remain as inactive dead load. Optimization of the reduction efficiency may be tricky to do if the manual analysis is done only once a day. An on-line measurement of the reduction efficiency is better for optimization since it is quicker and trends can be seen more easily.

Domsjö Fabriker AB has installed a near infrared spectroscopy instrument for on-line analysis of the green liquor from one of their boilers. The purpose is to investigate if the reduction efficiency can be measured reliably by this method. But yet no robust calibration model has been produced for the instrument.

1.2 Purpose
The purpose of this project is to create a well working and robust calibration model for measuring the reduction efficiency in a recovery boiler with a near infrared spectroscopy instrument. Also an evaluation of the performance is made and compared with the demands of Domsjö Fabriker.

Also included in the project was doing a literary survey of the chemical recovery and generation of pulping chemicals, NIR-spectroscopy and multivariate data analysis.

1.3 Process overview of Domsjö Fabriker AB
Domsjö consider itself to be a bio refinery because wood as the main raw material is refined into three main products: cellulose, bioethanol and lignin. An overview of the process can be found in Figure 1.
The raw material is debarked, chipped and then cooked in batch digesters in a two-stage sodium sulphite process. The resulting pulp is washed, bleached and then washed again. The bleaching is done in a closed loop, totally chlorine free process. The pulp is then dried, cut and packaged into bales and then sold as speciality cellulose. The washing liquid is evaporated and the sugar is fermented to produce bioethanol. The bioethanol is distilled out and sold to a neighbouring plant, SEKAB. A part of the remaining liquid is used for lignin production while the rest is combusted in the recovery boilers to produce steam. The unburnt smelt containing the pulping chemicals is tapped at the bottom and sent to the cooking liquor preparation where new cooking liquor is produced and thus the chemical recovery cycle is complete.
2 Literature survey

2.1 The recovery process

In sulphite pulping sodium sulphite, sodium hydrogen sulphite (often called sodium bisulphite) and sulphur dioxide are added to the pulp to aid in liberating the fibres from the wood matrix. These chemicals oxidize during the cooking into sulphate and thiosulphate. Reaction (1) shows the reaction of the oxidation of the hydrogen sulphite ion.

\[ 4\text{HSO}_3^- \leftrightarrow 2\text{SO}_4^{2-} + S_2\text{O}_3^{2-} + 2H^+ + H_2O \] (1)

Recovering the pulping chemicals is necessary from an economical point of view. The recovery includes extracting the sulphur from the black liquor as a sulphide, reducing it into oxidation state IV and react it with air into sulphur dioxide. (Larsson, 1992)

2.2 The recovery boiler

After the digesters the cellulose is washed out of the pulp. In the case of Domsjö fabriker the resulting liquor is fermented and ethanol is distilled out before the pulping chemicals are recovered. The fermentation utilizes the glucose content present in the liquor. After the fermentation the liquor is first evaporated and then burned in the recovery boiler to make use of the energy content of the liquor as well as extracting the pulping chemicals. In the bottom of the boiler a smelt is removed and dissolved in water to produce the so called green liquor (Larsson, 1992). The smelt consists mainly of sodium sulphide and sodium carbonate (Hupa, 2008) but also low concentrations of sulphates, thiosulphates and polysulphides.

The recovery boiler serves two purposes: to utilize the energy content in the black liquor and to reduce sulphur compounds into molten sulphide. The matter of recovery of the cooking chemicals makes the design of the recovery boiler more complex than a simple burner. The energy is utilized by burning the organic molecules and using the heat to produce high pressure steam. The steam in turn can be used to produce electricity and to heat up other process flows.

The recovery boiler is an enormous piece of equipment with a height of 60 to 70 meters and is often both the biggest and the most expensive unit in the paper plant. Liquor guns spray droplets of the black liquor into the fireplace. The droplets will partially evaporate and combust before landing in a smelt bed in the bottom of the fireplace where the rest of the combustion happens. The smelt is tapped from the bottom via smelt spouts and transported to a dissolver tank where the smelt is dissolved in water. The inside of the fireplace is lined with tubes where water is heated and evaporated. When the flue gases from the fireplace raises they first meet the superheater tubes, where steam is heated. Higher up is the boiler tubes where the rest of the water is evaporated. About half of the water evaporates in the boiler tubes and half in the fireplace tubes. Finally the flue gases meet the economizer where the feed water is preheated (Theliander, 2006).

The distribution of air in the recovery boiler is essential for the reduction of the pulping chemicals. That is why recovery boilers are equipped with three or four sets of air ports: primary air is added at the bottom, secondary air is added just below the liquor guns, tertiary above the liquor guns and
quaternary air is added even higher up. Thus it is possible to create reducing conditions in the smelt by restricting the air added at primary and secondary air ports. At the same time oxidizing conditions can be maintained above the liquor guns by adding enough tertiary and quaternary air. This makes it possible to completely oxidize all combustible gases such as carbon monoxide and hydrogen and thus utilize most of the energy content in the black liquor. Since the primary air guns are aimed more or less directly at the smelt bed they can be used to control the size and shape of the smelt bed. (Theliander, 2006).

In the smelt bed the organic material is turned into char and the reducing gases carbon monoxide and hydrogen gas. A quite large part of the chemicals is carried away by the flue gases as dust or vapour. Therefore, an electric precipitator is usually installed to collect the dust and feed it back into the boiler via the black liquor. However, a small part of the sulphur is lost into the vapour phase and is emitted as mostly sulphur dioxide and dihydrogen sulphur (Hupa, 2008). The sulphur lost this way is usually much less than one percent of the total sulphur and is more of an environmental concern than a problem for the process.

Dust that collects on the heat exchanger tubes can be removed with a soot-blowing system. These consist of movable lances that are moved into the tube banks. Steam is blown through nozzles in the lances which cleans the tubes. In the economizer a shot cleaning system can be utilized. Small metal balls are dropped onto the tubes, removing the dust. The dust together with the balls are then collected and separated. The dust can be recycled into the process via the black liquor and the balls are used again. (Theliander, 2006)

From the time that the black liquor droplets are introduced into the boiler until they hit the smelt bed four important reactions happens: drying, pyrolysis and combustion of the pyrolysis gases, combustion of char and combustion of sodium sulphide. The distribution of droplets is of great influence on the performance of the recovery boiler. Too big droplets will carry moisture to the smelt bed since the drying has not been complete. This may lead to a temperature drop in the smelt bed and risks that the combustion ceases (Theliander, 2006). If the temperature goes below about 800 degrees C solids starts to form and the bed “freezes” (Hupa, 2008). Too small droplets on the other hand may be carried away with the flue gases. This leads to too much recirculation of fly ashes and low efficiency (Theliander, 2006). The liquor gun nozzles are designed to optimize the spread and sizes of the droplets.

2.3 Reduction of chemicals in the recovery boiler

The sulphates are turned into sulphides by a reduction reaction with the char (Theliander, 2006):

\[ 4C + Na_2SO_4 \leftrightarrow Na_2S + 4CO \]  \hspace{1cm} (2)

Also a minor part of the sulphates are reduced by the reducing gases carbon monoxide and hydrogen gas:

\[ 4CO + Na_2SO_4 \leftrightarrow Na_2S + 4CO_2 \]  \hspace{1cm} (3)

\[ 4H_2 + Na_2SO_4 \leftrightarrow Na_2S + 4H_2O \]  \hspace{1cm} (4)
The reduction of sulphur compounds is mainly achieved by the combustion of char because the reduction rate of char is about two orders of magnitude faster than reduction by the reducing gases (Hupa, 2008).

The percent reduction efficiency is a measurement on how much of the sulphur have been reduced to sulphide in the recovery boiler. Only the sulphur in sulphide form is useful in the generation of new pulping chemicals so it is important to monitor and keep the reduction efficiency high. It is calculated from the molar concentration of all sulphide ions and total sulphur according to the formula

\[
\text{Reduction efficiency} \% = \frac{\text{Sulphide}}{\text{Total sulphur}} \times 100\% \quad (5)
\]

Note that in the green liquor the hydrogen sulphide dominates over other sulphide compounds. However, sulphide in all forms is counted.

Most of the reduction happens in the smelt bed so in order to achieve a good reduction large enough amounts of char must be carried with the droplets into the bottom of the furnace. Thus the combustion of char must still be somewhat incomplete when the black liquor droplets hit the smelt bed. Also the temperature of the bed is important. The reduction rate is roughly doubled if the temperature is increased by 50 to 60 degrees C (Hupa, 2008).

Complete reduction of the smelt can in principle be achieved when the local air factor is about 0.8 or lower. 0.8 air factor means that the combustion is feed 80 per cent of the oxygen needed for complete oxidation of the hydrocarbons. However, the reduction rates are fairly slow so it is hard to come near complete reduction (Hupa, 2008). Modern paper mills may achieve reduction efficiency around 98 percent in the recovery boiler. Most mills may achieve around 95 percent while poorly operated mills achieve around 85 to 90 percent. Low reduction efficiency results in high deadload content in the green liquor. This in turn can result in scaling within the evaporators which has to be removed with costly washing (Trung, Osmond, Allison, Uloth, & Porter, 2010).

Poor reduction efficiency is commonly caused by bad operating conditions that cause an oxidizing environment in the smelt bed. Air that leaks in from the smelt spouts or other parts of the boiler may also lead to low reduction efficiency. The reduction efficiency may vary between smelt spouts due to local fluctuations in performance or air infiltration (Trung, Osmond, Allison, Uloth, & Porter, 2010).

### 2.4 The sodium/sulphur ratio

The sodium sulphur ratio (S/Na\text{2}) of the green liquor is most important in the regeneration of new pulping liquor. The new pulping liquor must have the correct balance between sodium and sulphur or else the best pulp quality cannot be achieved. Also a too low sodium concentration may lead to higher chemical losses as sulphur leaves as gaseous compounds (most H\text{2}S and SO\text{2}) throughout the process. Historically the S/Na\text{2} ratio were often balanced by adding Na\text{2}SO\text{4} (in kraft pulping) in excess and let the SO\text{2} excess leave in the recovery boiler (Moreira Saturnino, 2012). Due to environmental reasons this is not possible today.

The S/Na\text{2} ratio can be calculated with the following formula if the molar concentrations of total alkali, total sulphur and sulphide are known:
Since the NaHS is included in both the measurement of total alkali and total sulphur it has to be subtracted from the denominator since it would otherwise be counted twice.

Sodium and sulphur compounds are also present in the flue gases from the recovery boiler. This can be an environmental problem since emission of sulphur dioxide contributes to the acid rain and therefore emissions are strictly regulated by law. It is also a problem for the process since acidic sulphates may form which cause fouling of the economizer and corrosion in the boiler banks. However, if the sodium content in the flue gases is high enough the sulphur dioxide content may drop below the detection limit and all these problems are eliminated (Hupa, 2008). That is why the S/Na\textsubscript{2} ratio of the flue gases is important.

If the S/Na\textsubscript{2} ratio is low the sodium compounds will dominate the furnace gases. The sodium compounds will react with almost all sulphur dioxides and can be collected in the electrostatic precipitator. Acidic sulphates will not form and the emission of sulphur dioxide will be completely eliminated.

The factors that affect the S/Na\textsubscript{2} ratio the most is the sulphidity of the black liquor and the temperature of the smelt bed. A low sulphidity gives a lower S/Na\textsubscript{2} ratio and vice versa. A high bed temperature gives a lower S/Na\textsubscript{2} ratio and vice versa. The S/Na\textsubscript{2} ratio is typically monitored by taking a sample of the flue dust, dissolving it in water and measuring the pH. A high pH indicates a low S/Na\textsubscript{2} ratio and vice versa but the optimal value varies between boilers (Hupa, 2008).

The recovery boiler is in steady state seen over a long time. All sodium and sulphur that enters with the black liquor must exit sooner or later via either the green liquor, flue gases or in the precipitator. Thus monitoring and observing drastic changes in the S/Na\textsubscript{2} ratio in the green liquor may give an idea of the S/Na\textsubscript{2} ratio in the flue gases.

### 2.5 Total alkali

The amount of alkali in the green liquor is important for the regeneration of new pulping chemicals since the concentration of the liquor determines how much of the liquor should be used when mixing new pulping liquid.

During titration of green liquor with a strong acid the pH-curve usually exhibits three inflexion points. From the inflexion points three different measurements can be derived: effective alkali, active alkali and total alkali.

The total alkali is defined as the total concentration of all alkaline constituents in the green liquor. The active alkali is defined as the total concentration of all alkaline constituents except carbonates. In practice this consists mainly of hydroxide and hydrosulphide ions. The effective alkali is defined as the concentration of all strongly alkaline constituents determined by titration to the first inflexion point. In practice this means the concentration of hydroxide ions (Scandinavian Pulp, Paper and Board Testing Committee, 1985).

Figure 2 shows an example of a titration curve where inflexion points corresponding to the effective and total alkali has been marked. The results is usually reported as either the molar consumption of
acid per litre of sample at these points or converted to g/l of Na₂O. The consumption of acid between
the first and the second inflexion points is associated with the carbonates. Thus the active alkali is
calculated with the formula

\[
Active \ alkali = \frac{(2a - 2b - c)}{v}
\]  

(7)

where \(a\), \(b\) and \(c\) is the molar consumption of acid up to the first, second and third inflexion point
respectively and \(v\) is the volume of the sample.

Figure 2. Titration curve of green liquor. pH-curve in black line and first derivative in blue dashed line. Inflexion points
marked with red dashed lines.

2.6 Green liquor dregs
Usually a considerable amount of solid particles are present in the green liquor. This is called green
liquor dregs or sludge. These particles are removed via clarifiers or filters installed after the smelt
dissolver and are usually disposed of in landfills. Before the clarifier the green liquor is called raw
green liquor while after the clarifier it is called clarified green liquor.

The dregs can be composed of unburned organic material from the recovery boiler and trace
elements from the wood. Some trace elements, like calcium, aluminium and barium, pose potential
problems to the process and their removal is necessary.

2.7 Generation of new pulping chemicals
The pulping chemicals used in sodium sulphite pulping is sodium sulphite, sodium hydrogen sulphite
and sulphur dioxide. In the first part of cooking the wood chips are treated in a solution of sodium
sulphite and sodium hydrogen sulphite. In the second step about two thirds of the liquid is removed
and sulphur dioxide is added making the cooking liquor acidic with a pH of about 1.5 (Larsson, 1992).
The cooking chemicals regeneration plant needs to be able to recycle most of these chemicals for the
cooking to be economically feasible. Also it is in this part of the factory where the cooking chemicals
are mixed together in a suitable mix. Thus it is the aim of this part of the recovery process to use the
green liquor from the recovery boiler, which is a mix of sulphides, polysulphides, sulphates,
carbonates and solid particles, and transform it into a suitable mix of sodium sulphite and sodium
hydrogen sulphite. Also a side flow of sulphur dioxide has to be created to be used in the second part
of the cooking.
The smelt coming from the recovery boiler turns into green liquor by dissolving it in water. The green liquor still contains some solid particles which are removed in the filters and clarifiers. The clarified green liquor goes to a pre-carbonisation step where all hydroxide is reacted with carbon dioxide to hydrogen carbonate (often called bicarbonate) according to formula:

\[ CO_2(g) + OH^- \leftrightarrow HCO_3^- \]  

(8)

It is important to neutralize the hydroxide in this way since otherwise hydrogen sulphide ions may form in the following reaction column according to formula:

\[ OH^- + H_2S \leftrightarrow HS^- + H_2O \]  

(9)

The reactor column serves two purposes: to form sulphite in the bottom part and to extract hydrogen sulphide in the upper part. The reactor column is a tray column with bubble caps in the bottom and valves in the upper part. The bottom part is about a third of the trays. Here a solution of hydrogen sulphite (often called bisulphite) is added which reacts with hydrogen carbonate to make sulphite according to formula:

\[ HCO_3^- + HSO_3^- \leftrightarrow CO_2(g) + SO_3^{2-} \]  

(10)

During this process gaseous carbon dioxide forms which travels upwards in the column. In the bottom a solution of sulphite and hydrogen sulphite is tapped.

Carbon dioxide forms in the reaction column but excess carbon dioxide from the Claus reactors are also added. The carbon dioxide’s role is to oxidize hydrogen sulphide ions and form hydrogen carbonate in the upper part of the column. The formula:

\[ CO_2 + H_2O + HS^- \leftrightarrow H_2S + HCO_3^- \]  

(11)

The hydrogen sulphide is lead to Claus reactors where it is partially oxidized over an aluminium oxide catalyst. Three Claus reactors in series can usually achieve about 95-97 per cent conversion of hydrogen sulphide into elemental sulphur (Mouljin, Makkee, & van Diepen, 2013). The overall reaction looks like:

\[ 2H_2S + SO_2 \leftrightarrow S_2 + 2H_2O \]  

(12)

The elementary sulphur is combusted in a furnace forming sulphur dioxide. The sulphur dioxide is stored as a water solution. Part of the sulphur dioxide is lead back to the Claus reactors while the rest is used to make the cooking liquid.

Thiosulphate is a catalyst for the decomposition of hydrogen sulphite:

\[ 4HSO_3^- \leftrightarrow 2SO_4^{2-} + S_2O_3^{2-} + 2H^+ + H_2O \]  

(13)

This is detrimental for the pulping process. To lower the concentration of thiosulphate a side flow is removed from the reaction column to a thiosulphate destruction reactor. Here oxygen is added under 15 bars pressure oxidizing the thiosulphate into sulphate. Also polysulphides and hydrogen sulphide oxidizes (Larsson, 1992).
As highlighted above, the chemistry in the recovery process is both complex and therefore tricky to control. Still, in order to produce a good pulp quality in an economical and environmental friendly way it is necessary to have good control over the composition of the cooking liquor. Vibrational spectroscopy may be a suitable tool for monitoring the quality of the green liquor.

2.8 Vibrational spectroscopy

The field of vibrational spectroscopy consist of Raman and infrared spectroscopy. The infrared spectroscopy is usually divided into the near, mid and far infrared regions. The general principle for all spectroscopy is to let light or other electromagnetic radiation hit a sample and then measure at which frequencies the light can either pass through the sample or is reflected by the sample. Many different chemical and physical properties can be determined by these methods.

In any molecule the chemical bond between the atoms can be vibrating. The vibration can be in the form of stretching the bond back and forth, as a bending of the molecule, as a twisting motion, etc. The number of different ways a molecule can vibrate, \( N_{\text{vib}} \), can be calculated by the formula

\[
N_{\text{vib}} = 3 \times N - 6
\]  

in the general case and in case of linear molecules by the formula

\[
N_{\text{vib}} = 3 \times N - 5
\]  

where \( N \) is the number of atoms in the molecule (Steele, 2002).

The frequency of the vibration, and thus its energy, is defined by the strength of the bond and the mass of the individual atoms in the molecule. If a photon of the same energy as the energy of the vibration hits the molecule the photon may be absorbed. However, it will only be absorbed if the vibration causes a change in dipole moment of the molecule (Pasquini, 2003). Any chemical compound has a unique set of frequencies at which photons can be absorbed and the compound can be identified by the absorption spectra. Even positional and rotational isomers can be distinguished from each other even though they have the same set of bonds and atoms. This is because the vibration of one bond slightly affects adjacent bonds (Steele, 2002).

In reality bands of frequencies are absorbed rather than single frequencies. The molecules exhibit some non-ideal behaviour by absorbing some photons with slightly higher or lower frequency. This is caused by the repulsive force between the electron clouds when the atoms approach each other and also caused by the variation of the bond force when the atoms move apart. This also gives rise to the possibility of overtones and combination bands (Pasquini, 2003).

Overtones are absorbed frequencies that are a multiple of the frequency defined by the bond. For example, the fundamental frequency for the carbon-hydrogen bonds stretch band is about 290 kHz. The first overtone is around 580 kHz, the second around 870 kHz and so on. The overtones are weaker in intensity: each following overtone is about ten to a hundred times less intense (Workman, 2014).

Combination bands appear when two or more fundamental vibrations are excited simultaneously. The combination bands present in the NIR region is from addition of two frequencies. If a vibration does not absorb because the absorption would not change the dipole moment it can still contribute
to combination bands if the other vibration does change the dipole moment. This can result in some vibrations, which fundamental frequency is not detected in the mid-IR region, to be detected in the NIR region as combination bands (Pasquini, 2003).

The intensity of absorption for a certain species is determined by the magnitude of the dipole moment change caused by the vibration but also on the degree of anharmonicity. Many substances in solutions follow Beer’s law

$$A = ecd$$

where $A$ is the absorption, $c$ is the concentration, $d$ is the path length and $e$ is the molar attenuation coefficient. This gives a linear relation between the concentration and the absorption which simplifies quantitative analyses (Pasquini, 2003). The molar attenuation coefficient is a measurement of how strongly the substance decreases the light intensity. Both the absorption and the molar attenuation coefficient are usually measured on a logarithmic scale.

Figure 3 depicts five different ways to measure with NIRS. Transmittance mode lets the NIR light pass through the sample and the light that is not absorbed can be detected. Transflectance mode is similar but the light is reflected with a mirror and thus passes through the sample twice. These two modes are used for fluid samples. When solid samples are analysed diffuse reflectance mode can be used where the reflected light is measured instead of the transmitted light. In the interactance mode the detector is placed a little more distant from the light source. Thus the beam has a higher probability to penetrate into and interact with the sample. The resulting detected beam contains more information on the composition of the sample. Transmittance mode can also be used for solid samples and has been used on, for example, pharmaceutical tablets (Pasquini, 2003).

![Figure 3. Five different measurement modes. (a) transmittance (b) transflectance (c) diffuse reflectance (d) interactance and (e) transmittance through a scattering medium (Pasquini, 2003).](image)

A dispersive NIR spectrometer uses a stationary dispersing element to disperse the NIR beam into its frequency components. The dispersing element is commonly either a prism or a grating. The
dispersed spectrum then falls on an array of detectors that captures the intensity of the different frequencies. (Saptare, 2003)

A Fourier transform NIR instrument operates slightly differently since it uses two NIR beams, one that interacts with the sample and one that travels a different optical path via a moving mirror. The two beams are recombined to create an interference signal that can be measured with an interferometer. By moving the mirror the interference signal can produce spectral information over a range of frequencies. Finally the interference signal is analysed with Fourier transform to gain information of the individual frequencies. FT-NIR spectrometers offer some advantages over grating technology NIR such as higher signal-to-noise ratio, extremely high resolution and fast and accurate frequency determinations. (Armstrong, Maghirang, Xie, & Dowell, 2006)

A recent trend is to use an acousto-optical tuneable filter (AOTF) as a monochromator (Trung & Allison, 2015). A monochromator is a device that emits light of only one specific frequency. An AOTF consists of a crystal bonded to a piezoelectric transducer. When an oscillating electric signal is applied to the transducer the refractive index changes in the crystal. Thus certain frequency ranges can be made to diffract while the rest of the light travels through the crystal undiffracted. By changing the frequency in the electric signal the region that diffracts can be changed very quickly (Bertie, 2002). AOTF equipped NIR instruments offers fast measurements but poor spectral resolution. They are also associated with higher maintenance costs since recalibration is necessary during use and after servicing. Problems with persistent fouling have been reported when used in the pulp and paper industry (Trung & Allison, 2015).

2.9 Near infrared spectroscopy

The near infrared spectral region is defined as the wavelengths from 780 to 2500 nanometres for convenience. This region is important to spectroscopical analysis since many common molecular bonds give rise to absorption bands in this region, for example C-H bonds and O-H bonds (Workman, 2014). NIRS have found use in many different fields. It can measure the amount of proteins, hydrocarbons, oil and fats and other properties of agricultural and food products. It can detect seed germination and fruit ripeness. It can be used to measure octane number, sulphur content and additives in motor fuels. In environmental science it has been used to predict pH in lakes and detect petrol contamination in soils. The textile industry has used it to identify textile fibres and detect contaminations. It can be used in medicine to diagnose cancer and measure blood flows in the brain. Within the field of pharmaceuticals both active ingredients and excipients has been quantitatively measured in tablets and solutions (Pasquini, 2003) (Workman, 2014).

NIRS differs from IR spectroscopy in a useful way because of the low absorption of most species in this region. This enables the use of longer path lengths through samples and often eliminates the need for sample preparation (Workman, 2014). When using IR spectrometers a sample thickness around 10 micrometres is preferable which may be inconvenient or even impossible to prepare. A NIR spectrometer on the other hand may be able to measure a sample with a thickness between 0.1 to 5 millimetres (Griffiths, 2002).

The NIR region contains mostly the overtone and combination bands. These are weak and often overlapping making the interpretation quite hard. On the other hand mid-IR spectroscopy uses mostly the fundamental frequencies of vibration to identify species. These are usually more separated and thus easy to analyse. This is partly why the NIR region was ignored for a long time by
many scientists. Mid-IR frequencies have been used to identify substances since about year 1900 while NIR spectroscopy gained wide use first in the eighties (Pasquini, 2003).

In practical applications measurements of liquids may be affected by the presence of solid particles in the liquid. Solid material absorb nearly all infrared light uniformly but if the particle size is comparable to the NIR wavelength the absorption may differ slightly with the wavelength due to quantum effects.

2.10 NIRS in liquor analysis

A lab-based NIRS can be used in daily analysis of the liquor from the recovery plant, replacing the titration methods that are often used for the same analyses. NIRS based methods are often quicker and less sensitive to operator bias compared to titration analysis (Hodges, Cullinan, & Krishnagopalan, 2006).

Making online measurements comes down to delivering a sample to a sample cell since no or little sample preparation is needed. Low cost fibre optics with low absorbance in the NIR region is available to be used so the sample cell can be placed far from the spectrometer. Standard mill water can be used as a reference since NIRS is not sensitive to trace substances (Hodges, Cullinan, & Krishnagopalan, 2006). The NIR spectrometer characteristics such as their rugged design, high throughput and the ability to measure concentration of several components simultaneously makes them well suited for online measurements. Also, they require less sample preparation, no reagents and consist of very few moving parts. The use of fibre optic cables makes it possible to use one spectrometer for several different sample points. The sample cells can be placed up to 300 meters from the spectrometer (Trung & Allison, 2015).

Table 1. Chemical properties shown to be measurable by NIRS. References: 1 (Hodges, Cullinan, & Krishnagopalan, 2006); 2 (Trung & Allison, 2015).

<table>
<thead>
<tr>
<th></th>
<th>White liquor</th>
<th>Green liquor</th>
<th>Black liquor</th>
<th>Weak wash</th>
<th>ClO₂ generator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective alkali</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active alkali</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total titratable alkali</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Total dissolved solids</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Total dissolved deadload</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Organic solids</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inorganic solids</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solids content</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Causticizing efficiency</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Reduction efficiency</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Residual effective alkali</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Residual active alkali</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Lignin content</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbonate (Na₂CO₃)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulphide (Na₂S)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulphate (Na₂SO₄)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiosulphate (Na₂S₂O₃)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
</tr>
<tr>
<td>Hydroxide (NaOH)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acid (H₂SO₄)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorate (NaClO₃)</td>
<td>X¹</td>
<td>X¹</td>
<td>X¹</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NIRS has many useful applications related to liquor analysis. Hodges et al and Trung and Allison list several chemical properties that can be measured with NIR equipment and can be seen in Table 1 (Trung & Allison, 2015) (Hodges, Cullinan, & Krishnagopalan, 2006).

2.11 Earlier work
Leclerc used transmission NIRS to measure sulphidity in Kraft pulping green liquor. PLS calibration was made on synthetic green liquor samples and the concentrations of sulphide, the combined concentrations of carbonate and hydroxide, total titratable alkali and chloride were measured. A two-component PLS calibration was made and measurements were made on green liquor samples from five mills and compared with chemical analysis. The RMSEP was found to be ±1.0 g/l as Na₂O for sulphide, ±1.7 g/l as Na₂O for combined carbonate and hydroxide and ±1.4 g/l as Na₂O total titratable alkali. Recalculated for comparability with this work the RMSEP for the total titratable alkali is 0.045 moles/l and for the sulphides 1.25 g/l.

Also, a three-component PLS calibration was made to measure sulphide, combined carbonate and hydroxide and chlorides and then measurements were made on green liquor samples from four mills. The RMSEP was found to be ±0.5 g/l as Na₂O for sulphide, ±1.5 g/l as Na₂O for combined carbonate and hydroxide and ±1.6 g/l as Na₂O for chlorides. The RMSEP for chlorides is not negligible since the average chloride concentration was 3.9 g/l. In these experiments certain spectral regions were chosen for building the PLS models. These regions are presented in Table 2. (Leclerc, 1997). Recalculated for comparability with this work the RMSEP for the sulphides is 0.63 g/l.

Table 2. Spectral regions chosen for each analysed component. Values are presented as wavenumbers in cm⁻¹.

<table>
<thead>
<tr>
<th>Two-component PLS calibration</th>
<th>Three-component PLS calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaHS</td>
<td>NaOH+Na₂CO₃</td>
</tr>
<tr>
<td>5728-6060</td>
<td>6116-6243</td>
</tr>
<tr>
<td>6744-6900</td>
<td>6400-6630</td>
</tr>
<tr>
<td>6930-7030</td>
<td></td>
</tr>
</tbody>
</table>

Trung et al. used an online FT-NIR spectrometer to measure sulphide, sulphate, thiosulphate, TTA and carbonate in green liquor in four paper mills and compared with laboratory analysis. They found the root mean squared standard error of prediction (RMSEP) to be ±0.2 g/l and ±0.6 g/l as Na₂O for thiosulphate and sulphate measurements and ±0.8 g/l for sulphide, TTA and carbonate (Trung, Osmond, Allison, Uloth, & Porter, 2010). Recalculated for comparability with this work the RMSEP for the total titratable alkali is 0.026 moles/l and for the sulphides 1.01 g/l.

2.12 Sampling for developing a NIR-based analytical method
In practical applications of NIR technology it is often nearly impossible to produce artificial calibration samples of known concentrations. Instead, the results from the NIRS analysis of the samples are compared to a standard method of analysis that is used as a reference method. The quality of the reference method must be known to accurately validate the NIRS performance (Pasquini, 2003).

In order to make a good model without unnecessary economical expenses related to analysis of samples the way in which the samples are gathered must be considered. The samples analysed
should include the full spectrum of variability of the component of interest, preferably in a uniformly spread out fashion over the total range. Also, other factors that may affect the NIR spectrum must also be included in the sample set, for example other interfering chemical components or temperature differences (Pasquini, 2003). A sufficient number of samples should be taken to be able to define the relationships between the spectrum and the component of interest with statistical significance (Cao, 2013). All this often translates into a considerable amount of samples needed. Pasquini recommends that about 50 to 100 samples are used in the calibration in the case of natural samples such as agricultural produce. More complex samples may require hundreds of samples for accurate calibration (Pasquini, 2003).

Building a model is easier if the data contain good information. However, increasing the number of samples or the amount of variables may not necessarily provide more information. For example, broadening the spectral range of a NIR instrument will provide more variables, but these may be uninformative if no component of interest has spectral absorption bands in this range. Likewise, a not to large but representative sample set is the goal since adding redundant samples to the set risks adding more noise than information. Ideally the properties responsible for making a sample informative are identified in advance and sampling are planned to give the best information (Cao, 2013).

2.13 Multivariate data analysis and Principal Component Analysis (PCA)

Chemometrics is the use of mathematical and statistical techniques for extracting relevant information from analytical data, for example a NIR spectrum (Pasquini, 2003). NIRS relies heavily on chemometrics and Principal Component analysis (PCA) and Partial Least Squares Regression (PLSR) are two commonly used methods used when dealing with spectroscopic data. Let us start with PCA.

Principal Component Analysis forms the basis for multivariate data analysis. It is commonly used to classify data or to see what makes different objects or variables differ from each other. It can also be used to see whether different variables within the data table correlate to each other. The most important method is to make a representation of a multivariate data table into a low-dimensional space, such as a two-dimensional graph. This simplifies overview of the data and conclusions can be drawn either through visual analysis or through simple geometric measurements. (Eriksson, Johansson, Kettnan-Wold, Trygg, Wikström, & Wold, 2006)

In PCA all data is put into a single data table which we can call the X-space. All data points can now be plotted into a space which has as many dimensions as there are variables in the data table. If the table has two variables this would be like a simple diagram on a paper but usually it has more and thus the space is hyper dimensional. A straight line is now calculated that best approximates the swarm of data points. This is called a Principal component (PC). Another way of saying this is that the PC is made so that the data point is most spread out along this line, or more correctly; the variance in the least square sense is largest. After the first PC has been calculated all variance in the data that can be explained by this first PC can be subtracted from the data table and what remains can be put in a new data table. A new PC is now calculated on what remains of the data. Several new PC:s can be produced this way, each new one usually explaining less of the original data material. PC:s are usually produced until their statistical significance is deemed too low.

For any data point a projection on the line can be made to which gives a value called the score, t. The indices on t denote which PC and which data point it corresponds to, for example observation A
projects onto the first PC on score $t_1$. All scores from the whole set of data points can be collected in a vector $t$. The score vector reflects the information in the original X-variables that is of relevance for modelling the response variables by this first PC.

When the PCs have been produced a new kind of plot can be made: the score plot. A score plot uses two PCs as axes and plots the data points accordingly. This plot can be useful for classification. For examples if the data points shows up not evenly spread but in clusters it is likely that these clusters do differ in some important aspect.

2.14 Partial Least Squares method

Partial least squares regression (PLSR) is one of the most popular regression methods used for analysing spectral data (Hodges, Cullinan, & Krishnagopalan, 2006). PLS originally was considered to be the acronym for “partial least squares” but the more descriptive interpretation “projection to latent structures” is sometimes used instead (Eriksson, Johansson, Kettaneh-Wold, Trygg, Wikström, & Wold, 2006).

The variables are split into X-variables and Y-variables. All measured variables that are used to predict are called predictors or X-variables. The variables that are a result are called response variables or Y-variables. In the case of NIRS each measured frequency is an X-variable and each compound of interest is a Y-variable. All measured frequencies forms the X-space and all compounds of interest forms the Y-space. A single spectrum is a data point in the X-space and corresponds to a data point in the Y-space.

Similarly to PCA the PLS method creates PCs. However, the PLS method creates PCs that correlate against the response variables. A principal component (PC) is a line through the X-space which well approximates the swarm of data points in the X-space but also has good correlation to the Y-space. If there is more than one response variable the Y-space is multidimensional. The PCs in the X-space will correlate against corresponding lines in the Y-space. Each data point in the Y-space can be projected onto this line to produce scores, similarly as is done in the X-space. The scores in the Y-space are usually denoted as $u$ and the score vector as $u$.

The PC is an approximation of the data point swarm so not all data points will be situated on this line. There will usually be a distance between the predictions of the response variables based on the X-variables. This difference is called residual. A model that predicts the responses well has low residuals. Adding more PCs increases the descriptive ability of the model. A new PC is always calculated so that it is orthogonal to all other PC, passes through the origin and improves the description of the X-variables while still having a good correlation to the y-residuals.

There are a few commonly used key numbers to investigate the models performance and explainability. The $R^2_X$ expresses the how well the variation in the X-space is explained by the model. Similarly, $R^2_Y$ expresses how well the variation in the Y-space is explained by the model.

The statistical significance of a PC is can be tested with an eigenvalue limit. A PC is considered significant if its eigenvalue is larger than 2. The significance can also be measured via cross validation. Cross validation works by excluding part of the data material and making a PLS model from the remaining data. The excluded part is than predicted from this new PLS model and the difference is measured. The cross validation results in a number $Q^2$ called the predictability. It signifies the
fraction of the variation of the X-variables that can be predicted from the component. (Eriksson, Johansson, Kettaneh-Wold, Trygg, Wikström, & Wold, 2006)

Figure 4. Example of interpretation of PLSR in the case of spectroscopically analysis of green liquor.

When dealing with PLSR on spectroscopic data a different and perhaps easier to understand interpretation can be done. Each measured spectra consist of several measured frequencies. Each frequency is an X-variable and a whole spectrum is a data point in the X-space. A PC can be described by a vector that contains a value for each frequency. This vector can be represented by a spectrum and thus each PC represents a separate spectrum. Each calibration spectra can be constructed by adding the PCs onto each other in different extent. The extent of which each PC is added is the score. However, there will always be a little difference between this constructed spectrum and the real measured spectrum. This difference is the residuals. A concentration of a certain compound is a response variable. All compounds together forms the Y-space. The PCs are calculated to give a good
correlation to this compound of interest. Thus each PC is associated with an amount of each compound.

When a prediction is calculated from a new measured spectrum the amount of each PC needed to produce this new spectrum is noted as the spectrum’s X-score. These X-scores are then used together with the amount of compounds each PC is associated with to calculate the predictions. See Figure 4.

2.15 PLS for NIR applications
The resulting spectra from a NIR analysis of an unknown sample may include other information than the absorption peaks caused by the compound of interest. The solvent and any other compounds present may also result in peaks. Also the sampling cell, stray light, temperature differences, solid particles and random noise may have some influence on the spectra. Moreover, absorption peaks can be overlapping while other frequencies in the spectra may not contain any relevant information at all. All this things makes the use of standard regression methods difficult if not impossible to use.

A multivariate approach to spectrometry may be a little more complicated than using standard curves but it serves several advantages in comparison to the traditional approach where usually only one frequency is used. Firstly, multivariate analysis can use several different spans of frequencies and therefore it is less sensitive to interference. An interfering substance or physical interference may result in changed intensities in certain frequencies used in the analysis but no or little change in others. Moreover, these interferences can often be detected as changed spectral features and may be recognized as an outlier in the scores or residuals. Secondly, random noise problems have a lower significance. Random noise contributes to a higher intensity in some frequencies and a lower intensity in other frequencies. Using several different frequencies for the analysis “cancels out” some of the noise and gives a higher precision. Thirdly, a standard curve is valid only for samples that are similar to the calibration samples. Differences in composition or changes in the process may make new samples different enough to make the standard curve unusable. However, with multivariate analysis outliers and nonlinearities can be detected and a more robust calibration can be done instead (Eriksson, Johansson, Kettaneh-Wold, Trygg, Wikström, & Wold, 2006).

There are a few different methods used to pre-process the signal before PLS analysis: scaling, mean centring and derivation. These methods aim to remove features from the spectrum that are either irrelevant or unrelated to the compounds of interest.

Scaling is used when the variables have considerably different variance since the variance influences the perceived “importance” of a variable. For example, a certain frequency may express large changes in intensity and may seem important to the analysis even though it is not. Scaling changes the numerical values so that each frequency has equal variance. Also, scaling may be employed such as to give some variables larger variance than others if these variables are known beforehand to have a larger “importance”. However, usually no scaling is needed in spectrometric analysis since all variables are expressed in the same unit. Mean-centring adds a vector to each variable so that the mean of all variables are the same, usually zero. In the case of spectrometry this means that each frequency would have its mean value at zero (Eriksson, Johansson, Kettaneh-Wold, Trygg, Wikström, & Wold, 2006).

Derivation uses the slope of the spectrum curve. The spectrometer may be experiencing a baseline displacement if the empty cell or an internal reference beam is employed as reference. In that case
the baseline is not at zero absorbance. This can be corrected by using the derivative of the spectra in analysis (Pasquini, 2003). However, derivation may in some cases reduce the signal and increase the noise. Care must be taken when derivation is used. Also, the second derivative can be used. It will remove the effect of a linear baseline displacement (Eriksson, Johansson, Kettaneh-Wold, Trygg, Wikström, & Wold, 2006). Linear baseline displacement effects may happen if particles close in size to the NIR wavelength is present in the samples.

Savitzky-Golay smoothing aims at reducing the amount of noise in the spectrum while maintaining the shape and height of the waveform peaks. For each data point in the spectrum a polynomial function is calculated by fitting the function as close as possible to the adjacent data points. A new value for the data point is set from the polynomial function and the process is repeated for the next data point (Schafer, 2011). The number of adjacent data points used and the order of the polynomial function should be set according to what is appropriate for the type and quality of the data. If the number of data points used is too high it risks distorting important peaks but if the number of data points is too low some noise is let through the filter. If a high order of the polynomial function is chosen more noise is let through but narrower peaks in the spectrum is preserved and vice versa if a low order polynomial is chosen (Orfanidis, 2010). In quantitative analysis the distortion of peaks is less important. Since the same smoothing filter is applied both to the calibration and later test samples the distortion effects should cancel out (Zeaiter, Roger, & Bellen-Maurel, Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods, 2005).

Normalisation of spectra can be useful when the data set shows large variation in intensity between spectra. This can be caused by the presence of sludge, by variations in the solvent or other effects that are uninteresting to the analysis. Normalisation tries to neutralize these variations. Usually normalisation is done by dividing all measured values in the spectrum with a scalar. The scalar can be the maximum intensity value in the spectrum, the area under the spectrum curve or the sum of all intensities. This reduces the intensity variation between spectra and gives all spectra in the data set comparable intensities. Figure 5 shows an example of normalisation.

![A dataset of a number of spectral curves before (top) and after (bottom) normalization.](image-url)
2.16 Diagnostic tools of PLS
There are several tools and methods that can help identify problems with a PLS model and to determine its accuracy and reliability.

Outliers in the data set can either be mistakes or important values. If the outlier is caused by a mistake in the sampling or analyse it does not belong to the data set and should be eliminated. However, if the outlier is caused by a real variation in the sample it provides important information and should not be removed. Instead, more samples of this class should be obtained and included into the model. (Pasquini, 2003)

The distance between a data point and the calibration model is a useful tool for identifying outliers. It is called the DmodX for the distance in the X-space and DmodY for the distance in the Y-space. The distance is usually expressed in standard deviations. A high DmodX or DmodY for a calibration sample signifies that the sample may be dissimilar to the rest of the sample set and the reason for dissimilarity should be investigated. A high DmodX for a prediction sample likewise signifies dissimilarity from the calibration set and the predicted result should be less accurate than what usually is the case.

2.17 Validation
Obviously, the accuracy of a NIR spectrometer cannot be better than the accuracy of the reference method used for calibration. The accuracy of the test of NIRS calibration is composed of the total variance defined by the standard error of prediction (SEP):

\[ SEP^2 = S_r^2 + S_{NIR}^2 + S_e^2 \]  

(17)

where \( S_r \) is the repeatability of the reference method, \( S_{NIR} \) the repeatability of the NIRS method and \( S_e \) is the lack of fit of the calibration model (Cao, 2013).

In this work the laboratory analysis were based on the standard methods SCAN-N 30:85, SCAN-N 31:94 and SCAN-N 5:83. The reported relative standard deviation for SCAN-N 31:94 is 1.2% (Scandinavian Pulp, Paper and Board Testing Committee, 1994). The reported relative standard deviation for SCAN-N 5:83 is 1.2% (Scandinavian Pulp, Paper and Board Testing Committee, 1983). No relative standard deviation is reported for SCAN-N 30:85 (Scandinavian Pulp, Paper and Board Testing Committee, 1985).

There is some confusion to what exactly the robustness of a model means and several authors has tried to give an all-encompassing definition even though different definitions can be made depending on in which field the term is used. Consensus seems to be around that robustness of a model has to do with the models ability to retain its predictive capabilities even though the measurements conditions varies. Variations in temperature, power, process and other parameters that are not measured directly should not affect the measurements too much. A robust model should also remain accurate and precise even over long time periods. The root mean standard error of prediction (RMSEP) is the most commonly used criterion for comparing model robustness. RMSEP is calculated:

\[ RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{i,ref})^2} \]  

(18)
where $n$ is the number of samples, and $y_i$ and $y_{i, ref}$ is the measured value and the measured value from the reference method respectively (Zeaiter, Roger, Bellon-Maurel, & Rutledge, 2004).
3 Method

3.1 Sampling equipment

Samples of raw green liquor were diverged from the main stream after the dissolving tank of boiler 9 and analysed with a Redeye analyser from Pulpeye. It contains a dispersive spectrometer with a slit size of 50 µm and a grating with 200 lines per mm. It analyses in transmittance mode at a wavelength interval of 900 to 1750 nm. The light source and spectrometer is connected via fibre optics to a flow-through sample cell. Samples for manual analysis can be taken through a sampling outlet. A schematic diagram of the sampling equipment can be seen in Figure 6.

Figure 6. Schematic diagram of sampling equipment.

Figure 7. Actual installation of the NIRS control system at mill.
3.2 The automatic sampling sequence
The NIRS control system follows a set sequence when a sample is taken. First the sampling pipe is flushed with water and then the sample cell is flushed with water into the return pipe. A reference spectrum with no beam is taken and another reference spectrum with water is taken. Green liquor is then led through the pipes to push out remaining water. Then the sample outlet is opened and a NIR measurement is done at the same time. After a sample has been taken all pipes are again flushed with water. The sample is scanned ten times during 35 seconds in total.

When an online measurement is done the same sequence is followed except that the sample outlet is never opened.

3.3 Manual sampling method
Before a sample was taken the sampling sequence was always run once first to remove remaining water and liquor from the sample piping. Samples were collected in a 100 millilitre container and care was taken to not let the sample oxidize by overfilling the container and sealing it.

To check for unwanted water dilution, sludge disturbances and other irregularities two samples were always taken immediately after each other. If the samples differed considerably both samples were discarded.

In total 54 paired samples were taken giving in total 108 data points.

3.4 Sampling design
An experimental region was determined by studying the daily lab analysis of hydrogen sulphide and total sulphur in the green liquor. 583 data points from 2014-02-14 to 2016-02-10 were used. Calibration samples were taken with the aim to span the entire variation of this time period.

3.5 Laboratory analysis
Samples were left to sediment and cool for about an hour. The samples were then analysed by titration for total alkali, hydrogen sulphide and total sulphur with Domsjö methods 2, 17 and 13 respectively. These analysis methods are based on the standards SCAN-N 30:85, SCAN-N 31:94 and SCAN-N 5:83 respectively. The automatic titration was made with a Mettler T50 titroprocessor.

<table>
<thead>
<tr>
<th>Table 3. Model descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
</tr>
<tr>
<td><strong>Model 4</strong></td>
</tr>
<tr>
<td><strong>Model 5</strong></td>
</tr>
<tr>
<td><strong>Model 6</strong></td>
</tr>
<tr>
<td><strong>Model 7</strong></td>
</tr>
</tbody>
</table>

3.6 Data analysis
Multivariate data analysis is done with the ExtractEye data program from Pulpeye. Spectra is normalized by dividing with the sum of all intensities, smoothed with Savitsky-Golay smoothing and differentiated before analysis. Several PLS models were made slightly differently. The models and the idea behind each can be seen in Table 3. In all models a quarter of the data points where randomly
excluded and used for calculation of RMSEP. Also comparisons of online measurement against daily samples were done.

3.7 Robustness test
The data material was split into two time periods and a calibration model was made for each period. The model of the first period was then used for prediction of the second time period and the model of the second period was used for prediction of the first period. The difference between the predicted results and the measurement were noted.

Two robustness tests were done: one were the model were created by the same method as model 2 and one were the model were created by the same method as model 6.

3.8 Classification of abnormal samples
Some online measurements were found to be abnormal because the calibration models predicted very low reduction efficiency even though the boiler operated normally at the time. To analyse the cause of this abnormalities a PCA model were made. Both spectra of process water and spectra of green liquor were used to create the PCA model. The abnormal samples were then classified according to this model.
4 Results

4.1 Experimental region
A diagram over the daily samples of sulphide and total sulphur over two years is presented in Figure 8. The daily samples are overlaid with the calibration samples analysed in this work. Most of the data points of both the calibration samples and the daily samples are lower than 90 per cent reduction efficiency. A similar diagram over the daily samples of total sulphur and total alkali over two years is presented in Figure 9. In both cases the calibration samples spans fairly well the normal concentration variations in the green liquor with the exception of a few outliers.

![Figure 8. Reduction efficiency over two years. Diagonal lines signify the reduction efficiency.](image-url)
Based on the analyses of the daily samples the normal variation of the concentration was calculated. It is presented in Table 4.

Table 4. The normal variance of the green liquor.

<table>
<thead>
<tr>
<th></th>
<th>Standard deviation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total alkali</strong></td>
<td><strong>Sulphide</strong></td>
<td><strong>Total sulphur</strong></td>
</tr>
<tr>
<td>(moles/l)</td>
<td>(g/l)</td>
<td>(g/l)</td>
</tr>
<tr>
<td>0.50</td>
<td>8.46</td>
<td>6.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard deviation as percentage of average value</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.84</td>
</tr>
</tbody>
</table>

4.2 Precision

A summary of the statistics of the first five models can be found in Table 5. Eight samples were excluded from model 3 and twelve samples were excluded from model 4. The spectral regions selected for model 5 were 1009.7-1200.3; 1300.5-1397.5; 1404.6-1432.8; 1443.4-1446.9; 1630.2-1661.9; 1672.5-1707.8 nm;
The RMSEP of the models can be found in Table 6. Since the models only predict the concentrations of total alkali, sulphide and total sulphur the RMSEP of reduction efficiency and S/Na$_2$ ratio is calculated from these concentrations.

A second derivation seems to give slightly better results of all measurement than only a single derivation. Excluding samples from process start-ups gives slightly better prediction of RE but slightly worse prediction of S/Na$_2$ ratio. Model 4 gives the best results for prediction of the concentrations. Limiting the spectral regions seems to only benefit the measurements of RE and only do so just slightly. Model 6 gives the best results for measurement of RE. Limiting the responses to only the S/Na$_2$ ratio seems to not benefit the modelling since it gives only average results.

Table 5. Summaries of the models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Significant PCs</th>
<th>R2X cum</th>
<th>R2Y cum</th>
<th>Q2 cum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>3</td>
<td>0.877</td>
<td>0.854</td>
<td>0.827</td>
</tr>
<tr>
<td>Model 2</td>
<td>3</td>
<td>0.805</td>
<td>0.907</td>
<td>0.890</td>
</tr>
<tr>
<td>Model 3</td>
<td>3</td>
<td>0.771</td>
<td>0.859</td>
<td>0.816</td>
</tr>
<tr>
<td>Model 4</td>
<td>5</td>
<td>0.983</td>
<td>0.898</td>
<td>0.861</td>
</tr>
<tr>
<td>Model 5</td>
<td>2</td>
<td>0.924</td>
<td>0.744</td>
<td>0.728</td>
</tr>
<tr>
<td>Model 6</td>
<td>4</td>
<td>0.930</td>
<td>0.872</td>
<td>0.802</td>
</tr>
<tr>
<td>Model 7</td>
<td>2</td>
<td>0.836</td>
<td>0.522</td>
<td>0.493</td>
</tr>
</tbody>
</table>

Table 6. RMSEP of the first five models.

<table>
<thead>
<tr>
<th></th>
<th>Total alkali (moles/l)</th>
<th>Sulphide (g/l)</th>
<th>Total sulphur (g/l)</th>
<th>RE (% unit)</th>
<th>S/Na$_2$ (no unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.162</td>
<td>2.68</td>
<td>3.38</td>
<td>3.65</td>
<td>0.026</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.159</td>
<td>2.28</td>
<td>3.19</td>
<td>3.63</td>
<td>0.025</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.212</td>
<td>2.80</td>
<td>2.90</td>
<td>2.89</td>
<td>0.030</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.108</td>
<td>1.95</td>
<td>2.83</td>
<td>3.75</td>
<td>0.022</td>
</tr>
<tr>
<td>Model 5</td>
<td>0.195</td>
<td>3.27</td>
<td>4.18</td>
<td>3.34</td>
<td>0.026</td>
</tr>
<tr>
<td>Model 6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.75</td>
<td>-</td>
</tr>
<tr>
<td>Model 7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.027</td>
</tr>
</tbody>
</table>

RMSEP as percentage of average predicted value

<table>
<thead>
<tr>
<th></th>
<th>Total alkali (moles/l)</th>
<th>Sulphide (g/l)</th>
<th>Total sulphur (g/l)</th>
<th>RE (% unit)</th>
<th>S/Na$_2$ (no unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>3.75</td>
<td>5.06</td>
<td>5.38</td>
<td>4.34</td>
<td>3.26</td>
</tr>
<tr>
<td>Model 2</td>
<td>3.67</td>
<td>4.29</td>
<td>5.08</td>
<td>4.32</td>
<td>3.10</td>
</tr>
<tr>
<td>Model 3</td>
<td>4.83</td>
<td>5.13</td>
<td>4.55</td>
<td>3.39</td>
<td>3.79</td>
</tr>
<tr>
<td>Model 4</td>
<td>2.56</td>
<td>3.82</td>
<td>4.62</td>
<td>4.53</td>
<td>2.83</td>
</tr>
<tr>
<td>Model 5</td>
<td>4.50</td>
<td>6.16</td>
<td>6.65</td>
<td>3.97</td>
<td>3.34</td>
</tr>
<tr>
<td>Model 6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.27</td>
<td>-</td>
</tr>
<tr>
<td>Model 7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.48</td>
</tr>
</tbody>
</table>

At this point the main work continued on investigating only two models: model 2 and model 6. Model 2 were chosen since it has the best Q2 and model 6 since it has the best RMSEP for predicting the reduction efficiency.
4.2.1 Diagnostics
From the score plot in Figure 10 we can see that the data points are fairly well spread. However a few outliers can be seen in the upper left corner. Studying the dates of these data points reveals that they are samples taken during process start-ups and have low reduction efficiency. Therefore they may bring valuable information to the model and are not necessarily valid exclusions.

Figure 11 shows a score plot for model 6. The same outliers can be seen here as in model 2.

![Figure 10. Score plot (left) and corresponding loadings plot (right) of model 2. Point 2 is total alkali, 3 is sulphide and 4 is total sulphur.](image)

![Figure 11. Score plot (left) and corresponding loadings plot (right) of the two first PCs of model 6. The reduction efficiency is the only response variable seen in the loadings plot.](image)

4.2.2 Online measurements
Online measurement of the green liquor using model 2 can be seen in Figure 12. Several noticeable features can be seen in the diagram. Firstly, some conspicuous vertical lines can be seen in the diagram. I call these “abnormal samples” and will be addressed later in the report. When the online
measurements started the boiler had been recently started up. At boiler start-up the general concentration of the green liquor and the reduction efficiency is rather low and increases over time. This seems to be captured well by the measurements as all three curves increases quite sharply in the beginning of the period. At the end of the period the boiler were shut-down which is usually followed by a decrease in green liquor concentration and reduction efficiency. This seems also to be well captured by the measurement as all curves decrease sharply after the boiler shut-down. The curves look fine until the reduction efficiency falls under about 70 per cent where the measurements seem to behave strange. During the measurement period the pressure gauge sprung leak and was not repaired until a couple of days later. This seems to have resulted in a number of faulty measurements.

![Figure 12. Online measurements using model 2 from 21st May to 9th June. Purple line is reduction efficiency, green line is total sulphur and red line is sulphide. The squares are corresponding daily lab analyses.](image)

The online measurement of total alkali and S/Na$_2$ ratio can be seen in Figure 13. The same features mentioned above can also be seen in this diagram.

The S/Na$_2$ ratio varies very little during this period so it is hard to tell whether the measurements follow the variation or not. For the alkali all of the daily samples are located above the online measurements. The average difference was about 0.4 moles/l. A diagram where the online measurement curve has been shifted up can be seen in Figure 14. Here the online measurements seem to capture the variation of the daily samples quite well.
Figure 13. Online measurements using model 2 from 21st May to 9th June. Red line is total alkali and blue line is S/Na$_2$ ratio corresponds to the secondary axis. The squares are corresponding daily lab analyses.

Figure 14. Online measurement of total alkali using model 2 from 21st May to 9th June. A constant +0.4 moles/l have been added.
Figure 15. Online measurements of reduction efficiency using model 6 from 21st May to 9th June. The squares are corresponding daily lab analyses.

Figure 16. Online measurements of reduction efficiency using model 6 from 21st May to 9th June. A constant -4 percent units has been added. The squares are corresponding daily lab analyses.

The online measurements using model 6 can be seen in Figure 15. The same features mentioned earlier can be seen here as well. However, almost all daily samples are located below the online measurement curve. The difference was on average about 4 per cent units. In Figure 16 the curve has been shifted down 4 per cent units. Now the online measurements seem to capture the variation in reduction efficiency very well.
4.3 Robustness test

4.3.1 Model 2

Table 7. PLS statistics of the models based on model 2.

<table>
<thead>
<tr>
<th>Significant PCs</th>
<th>R2X cum</th>
<th>R2Y cum</th>
<th>Q2 cum</th>
</tr>
</thead>
<tbody>
<tr>
<td>First period</td>
<td>3</td>
<td>0.867</td>
<td>0.916</td>
</tr>
<tr>
<td>Second period</td>
<td>3</td>
<td>0.798</td>
<td>0.835</td>
</tr>
</tbody>
</table>

The PLS statistics for the two models created for the robustness test of model 2 can be found in Table 7. The values were somewhat comparable to the original model 2, but for the first period it was slightly better and for the second period it was slightly worse.

The results from the robustness test based on model 2 can be found in Table 8. The prediction of reduction efficiency and $S/Na_2$ ratio is somewhat better than the predictions of concentrations. This is because the models usually overestimates all concentrations or underestimates all concentrations at the same time making the quotients less divergent.

Table 8. RMSEP between measured results and prediction of the other time period.

<table>
<thead>
<tr>
<th>RMSEP</th>
<th>Total alkali (moles/l)</th>
<th>Sulphide (g/l)</th>
<th>Total sulphur (g/l)</th>
<th>RE (% unit)</th>
<th>$S/Na_2$ (no unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First period</td>
<td>4.67</td>
<td>6.86</td>
<td>7.77</td>
<td>4.97</td>
<td>4.72</td>
</tr>
<tr>
<td>Second period</td>
<td>6.01</td>
<td>10.07</td>
<td>7.65</td>
<td>5.46</td>
<td>4.39</td>
</tr>
</tbody>
</table>

4.3.2 Model 6

The PLS statistics for the two models created for the robustness test of model 6 can be found in Table 9. Here the Q2 value is lower for the second period as well and slightly higher for the first period. However, all values are comparable to the original model 6.

Table 9. PLS statistics of the models based on model 6.

<table>
<thead>
<tr>
<th>Significant PCs</th>
<th>R2X cum</th>
<th>R2Y cum</th>
<th>Q2 cum</th>
</tr>
</thead>
<tbody>
<tr>
<td>First period</td>
<td>4</td>
<td>0.955</td>
<td>0.900</td>
</tr>
<tr>
<td>Second period</td>
<td>5</td>
<td>0.981</td>
<td>0.811</td>
</tr>
</tbody>
</table>

Table 10. RMSEP between measured results and prediction of the other time period.

<table>
<thead>
<tr>
<th>RMSEP</th>
<th>Total alkali (moles/l)</th>
<th>Sulphide (g/l)</th>
<th>Total sulphur (g/l)</th>
<th>RE (% unit)</th>
<th>$S/Na_2$ (no unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First period</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.47</td>
<td>-</td>
</tr>
<tr>
<td>Second period</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4.64</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentage of average predicted value</th>
<th>Total alkali (moles/l)</th>
<th>Sulphide (g/l)</th>
<th>Total sulphur (g/l)</th>
<th>RE (% unit)</th>
<th>$S/Na_2$ (no unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First period</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.95</td>
<td>-</td>
</tr>
<tr>
<td>Second period</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.55</td>
<td>-</td>
</tr>
</tbody>
</table>
The results from the robustness test based on model 6 can be found in Table 10. The RMSEP is a somewhat better than for model 2. The RMSEP for the second period is almost twice as large as for the first period.

4.4 Classification of abnormal samples
The PCA analysis yielded a model with two significant principal components. The cumulative R2X were 0.990 and the cumulative Q2 were 0.990. According to the score plot seen in Figure 17 there is a clear divide between the water samples and green liquor samples on the first component. The water samples are gathered in a small area on the left side of the plot and the green liquor is spread out along the second component on the right side.

Figure 17. Score plot of the two first PC:s. All of the water samples are gathered closely on the left side of the plot while the green liquor samples are spread out vertically on the right hand side.
When the abnormal samples are plotted into the model they all lie near where the water samples are situated. See Figure 18. Thus it is likely that these samples are either just water or green liquor heavily diluted by water.
5 Discussion and conclusions

As has been stated earlier it is important for the sampling to cover all of the natural variation in the process. When seen to each individual concentration the samples spans the normal variations well enough. However, as seen in Figure 8, there is an area that is not covered when the sulphide is about 50 g/l and total sulphur is about 70 g/l. A better calibration model would include samples from this area as well. It is in general not economically feasible to manipulate the recovery boiler process to give green liquor with certain properties so the sampling had to be done with whatever the process gives.

The sampling was done during a relatively short period, about twelve weeks, so there might be some potential variation that has not been included into the model, for example seasonal variation, which would require sampling over a whole year.

All models are quite close to each other concerning RMSEP. Since the accuracy of the RMSEP measurement is unknown it seems plausible that they actually are all within the same confidence interval. Thus there would be no measurable difference between them.

Each one of the models compared has its own benefits and drawbacks and the model used should be selected with its use in mind. If the measurements of RE is deemed the highest priority either model 3 or model 6 should be used since they have the best RMSEP for RE. Model 6 though probably has somewhat lower robustness. On the other hand, if the measurement of total alkali, sulphide and total sulphur are interesting enough then model 4 should be considered. If the preferred use is to spot minute changes in reduction efficiency during optimisation model 6 should be preferred since it seems to capture variations very well.

When choosing a calibration model for a NIR instrument, often the accuracy of the model has to be weighed against the robustness and applicability of the model. If process start-ups are excluded from the model some accuracy are gained but this likely limits the usefulness of the online measurement during start-ups. Likewise, model 6 shows some potential for accurate measurement but its accuracy is probably sensitive to changes in total sulphur. This is because the PLS approach is good at finding linear relationships but the RE is a quotient. Thus there is an inverse linear relationship between the total sulphur content and the absorption caused by the total sulphur. Likely this model will give good results as long as the sulphur concentration is within a certain interval but the accuracy decreases when outside this interval.

The RMSEP:s reported here are in general higher than has been reported earlier (Leclerc, 1997) (Trung & Allison, 2015). However slightly different methods were used in these reports. Leclerc used synthetic calibration samples not actual process samples. Trung and Allison used an FT-NIR instrument which is known to be more precise than dispersive instruments.

The PCA shows that the equipment for supplying the instrument with a sample does not always function satisfactory because the equipment delivers water or heavily diluted green liquor. It is probably unavoidable that the sampling equipment does not function when a mechanical failure happens, as in the case when the pressure gauge leaked. But during the testing period there were some other times when a sample was not delivered probably. This is most probably a caused by too low pressure in the main green liquor line, because of low pump load. What happens is that the
cleaning water used to flush the sample cells stays and the pump does not fully replace the water in the sample cell with green liquor. The pump load required to fill the sample cell with green liquor probably varies somewhat with the viscosity of the green liquor.

The results from the robustness test shows that in the general case a calibration model is applicable to other time periods, even though the process and the properties of the green liquor may be very different. The predictions are still accurate within a few per cent. It is worth noting that the calibration models used in the robustness test uses calibration samples from a relatively short time period, about five weeks. Thus these models are expected to be less robust than a model using samples from the whole time period.

The use of a NIR instrument to monitor the reduction efficiency may lead to both economic and environmental benefits. It may enable easier optimisation of the reduction efficiency and operation at higher reduction efficiency. Higher reduction efficiency leads to less inactive chemicals as dead load and better delignification selectivity in the cooking. The dead load requires some energy to transport around the process and takes up some space in all containers. A higher selectivity means that less material goes to waste and more material can be made into the cellulose product. Thus less resources and energy is consumed.

6 Recommendations to Domsjö Fabriker

6.1.1 Choosing a model
The main purpose is to use the NIR instrument to monitor reduction efficiency. Model 6 would be best suited for this since it has the best RMSEP. This is also confirmed by the visual examination of the curves. However, continued manual sampling should be done to monitor the performance of the model. If the models performance starts degrading it should be considered to either change it for another model or complement the model with more calibration samples.

6.1.2 The problem of pressure
Getting a representative sample from the green liquor line into the sample cell or a sampling container requires a high enough pressure in the green liquor line to overcome the resistance of the small sampling pipe. This is usually achieved if the green liquor pump operates at a high enough load. This is however not always the case since the pump is regulated by the liquor level in the dissolving tank. An interlock is installed so that the measurement does not start until the control signal to the pump is at least 25 percent. If Domsjö considers using the NIRS for online measurements a method of online sampling that is reliable at all times is needed.

As a first fix I recommend increasing the condition for the interlock with a few per cent. The maximum pump load usually varies between 25-30 per cent during normal operation. Increasing the condition for the interlock will give measurements that are more reliable but may at the same time decrease the regularity of the measurements since the pump load may not be sufficiently high at times. However, since this is the quickest and easiest change to implement it should be tried first.

During the sampling done in this work two ways to achieve higher pressure has been tried, either to manually override the regulation and set the pump load higher or to manually close the valve after the sampling pipe branch.
Manually setting the pump to a higher load increases the pressure quickly but risks draining the tank since flow from the recovery boiler can be very low at times. The tank can be drained in minutes.

Closing the valve increases the pressure before the valve where the sampling pipe branch is located. If the valve is closed fully this stops the flow and the pump should be at full load in a few minutes. Testing showed that samples could be taken during this time. However, this method poses a risk of overflowing the tank since the flow is stopped. This is a serious risk since the pump at times can be at full load even during normal operation and closing the valve at such time could overflow the tank fairly quickly.

Closing the valve partially may pose less risk to overflowing the tank since the pump still can maintain a flow through the valve. It may however take longer time until the pump reaches a high enough load.

Widening the sample pipe may not fully solve the problem since the sample cell will be the smallest part of the pipe. Widening the sample cell cannot be done easily.

An ideal solution for online measurements would be a pressure sensor and a control valve after the sampling branch connected to the NIRS control system. When a sample is needed the control system checks for pressure. If the pressure is too low the control valve is slowly closed until high enough pressure is achieved and a sample can be taken.

6.1.3 Multiple sampling cells
The NIR instrument allows for a second sample cell to be installed. I suggest installing a sample cell for measurements of the green liquor of boiler 8 as well. This makes possible to optimize the reduction efficiency of boiler 8 as well. Also, since the green liquor from both boilers are sent into the same storage tank before going to the chemical regeneration plant (“kokvätskeberedningen”) the current online measurements gives only limited information for the chemical regeneration plant to use. However, the composition of the green liquor is crucial to the operation of the chemical regeneration plant, since the concentration of chemicals determines the optimal mixing. Measurements of the green liquor from boiler 8 would give more complete information.

The calibration models developed in this work is only applicable to green liquor that is sufficiently similar in composition. It is uncertain whether the green liquor from boiler 8 needs a separate calibration model or not.

The current installation only allows for measurement while the green liquor line 1 is used. During washing of the pipeline sampling should be done from green liquor line 2. A simple T-connection with valves installed before the pressure gauge should suffice to connect line 2 to the existing sample cell as proposed in Figure 19.
6.1.4 Recalibration

The calibration models developed in this work give accurate results if applied to green liquor that is sufficiently similar to the calibration samples. If the composition of the green liquor should change due to reconstructions or major process changes the calibration models may become less accurate. I suggest that manual samples are taken regularly according to the method used in this work and compared to the result of the NIR instrument. If the disparity becomes too large a new calibration model should be developed.

If the DModX temporarily shows high values it signifies that the green liquor may be different from the calibration samples in some way. Thus the DModX of the measurements should be monitored and samples should be taken when the DModX is high. The samples could be included in the model to create a more robust model over time. However, care must be taken to make sure that the high DModX is not caused by some fault of the sampling equipment, for example if water remains in the sampling cell when the measurement is done. Including such samples in the calibration model would make the calibration model inconsistent.

6.2 Continued work

As stated earlier, NIR instruments has been shown to be able to measure active alkali, effective alkali and carbonate concentration in green liquor (Hodges, Cullinan, & Krishnagopalan, 2006) (Trung & Allison, 2015). These chemical properties can be derived from the titration curves when measuring total alkali of the samples in this work. Thus a large data material of calibration samples already exists. However, the values have to be manually extracted from the titration curves in the Labex database and PLS-models needs to be produced and tested.

There seems to be a correlation between the amount of dregs in the green liquor and the overall intensity of the spectrum. If that is the case the amount of dregs in the green liquor can be easily measured with the NIR-spectrometer. An example of this can be seen in Figure 20 where sample two
were taken immediately after sample one. The two spectra shows a large difference in intensity and the two samples shows a large difference in the amount of sedimented dregs in the bottom of the containers.

![NIR Spectra](image)

Figure 20. Example of two samples and their corresponding NIR spectra. Sample 1 corresponds to spectrum at 08:49:21 and sample 2 corresponds to spectrum at 08:51:07.

Thiosulphate content has a detrimental influence in the cooking step since it catalyses the degradation of the cooking chemicals (Larsson, 1992). However, NIR-instruments have been shown to be able to measure thiosulphates in green liquor (Trung & Allison, 2015). Detection of high thiosulphate content in green liquor gives the recovery plant an early warning and may help them to limit the consequences. This may be something to investigate in future studies.
7 References


