Near Infrared Spectroscopy – An Introductory Study On Measurement Techniques And Tools For Analysing Moisture Content In Pulp And Paper Media

Master of Science
30 Hp

Term: VT 2012
Supervisor: Helena Håkansson
Niclas Andersson
Examiner: Lars Järnström
Abstract

The growing technique of near infrared spectroscopy is used to build fundamental knowledge of how the method might be applied for moisture determination in pulp and paper media. Filter paper was used to investigate the angle and distance dependence, otherwise obscured by the uneven surface of pulp or dry fibers. For contact measurements both filter paper and CTMP-fibers were used. The angle and distance evaluation is based on visual spectral interpretation to later incorporate preprocessing methods of the data followed by a multivariate analysis using PLS. With the knowledge built before the experimental startup, the acquired spectra are known to have many co-linearities and hence, for the purpose of determining the moisture content abundant information which might enable the reduction of data points for evaluation. The distance and angle experiments gave results indicating that the intensity of the reflected light change the information gathered and must be accounted for in calibration. Turning the measuring probe 30 degrees or more at 5 mm distance reduced the effect of high intensity specular reflection. Temperature displayed differences in the spectra, cold samples giving less response than warm samples although no clear correlation was found. Multivariate analysis of the contact measurement samples gave most accurate result after multiple signal correction preprocessing. Distance measurements on fiber rendered models with high RMSEC values and gave no reasonable base for prediction.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTMP</td>
<td>Chemical thermo mechanical pulp</td>
</tr>
<tr>
<td>FC</td>
<td>Fiber concentration</td>
</tr>
<tr>
<td>FP</td>
<td>Filter paper</td>
</tr>
<tr>
<td>FPC</td>
<td>Filter paper concentration</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared</td>
</tr>
<tr>
<td>MLR</td>
<td>Multiple linear regression</td>
</tr>
<tr>
<td>NIR</td>
<td>Near infrared</td>
</tr>
<tr>
<td>NIRS</td>
<td>Near infrared spectroscopy</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>Principal component regression</td>
</tr>
<tr>
<td>PLS</td>
<td>Partial least squares</td>
</tr>
<tr>
<td>RMSEC</td>
<td>Root mean square error of calibration</td>
</tr>
<tr>
<td>RMSEP</td>
<td>Root mean square error of prediction</td>
</tr>
</tbody>
</table>
# Table of Contents

1 Introduction.......................................................................................................................... 6
  1.1 The purpose of this thesis ................................................................................................. 7
  1.2 Theory .............................................................................................................................. 7
    1.2.1 Electromagnetic radiation ....................................................................................... 7
    1.2.2 Interaction with matter ......................................................................................... 8
    1.2.3 Infrared .................................................................................................................... 8
    1.2.4 Near infrared .......................................................................................................... 9
    1.2.5 Overtones and combination bands ......................................................................... 9
    1.2.6 Modes of vibration ............................................................................................... 10
    1.2.7 Sampling ............................................................................................................... 11
    1.2.8 Forms of water ....................................................................................................... 12
  1.3 Multivariate analysis tools ............................................................................................ 12
    1.3.1 Multiple linear regression .................................................................................... 12
    1.3.2 Principal component analysis ............................................................................. 13
    1.3.3 Principal component regression .......................................................................... 13
    1.3.4 Partial least squares ............................................................................................ 13
    1.3.5 Preprocessing ....................................................................................................... 14
    1.3.6 Multiple signal correction ................................................................................... 14
    1.3.7 Derivative ............................................................................................................. 14
    1.3.8 Wavelength selection ......................................................................................... 14
  1.4 Sample homogeneity ....................................................................................................... 14
  1.5 Reflection ....................................................................................................................... 15

2 Material and method........................................................................................................... 16
  2.1.1 Light source ............................................................................................................ 16
  2.1.2 Fiber optics ............................................................................................................. 16
  2.1.3 Spectrometer .......................................................................................................... 16
  2.1.4 Probe ....................................................................................................................... 17
  2.1.5 Loss on drying ....................................................................................................... 17
  2.1.6 Software .................................................................................................................. 18
  2.2 Experimental section .................................................................................................... 18
    2.2.1 Sample preparation ............................................................................................... 18
    2.2.2 Angular dependence measurements .................................................................... 19
    2.2.3 Direct contact measurement ............................................................................... 19
    2.2.4 Distance dependence measurement ................................................................... 19
2.2.5 Surface vs bulk moisture experiments .......................................................... 20  
2.2.6 Water absorption at various temperatures ...................................................... 20  
2.2.7 Temperature dependence measurement ........................................................ 20  
2.2.8 Contact measurements on fiber .................................................................... 20  
2.2.9 Distance measurements on fiber .................................................................... 20  
2.3 List of experiments ........................................................................................... 22  

3 Result and discussion ......................................................................................... 24  
3.1 Angle dependence ......................................................................................... 24  
3.2 Contact measurements ................................................................................... 26  
3.3 Distance dependence .................................................................................... 27  
3.4 Automated weighing ..................................................................................... 28  
3.5 Temperature dependence of wet fibers ........................................................... 30  
3.6 Separating the optical fibers .......................................................................... 30  
3.7 Fibers against the fiber optics ........................................................................ 31  
3.8 Applying multivariate analysis preprocessing techniques to the spectral data .... 33  
3.8.1 Contact measurement .................................................................................. 33  
3.8.2 Distance measurement ................................................................................ 40  
3.9 Result summary .............................................................................................. 47  

4 Conclusion .......................................................................................................... 48  

5 Acknowledgements ............................................................................................. 50  

6 References .......................................................................................................... 52
1 Introduction

Water is an abundant resource and readily used in many parts of the production industry. The uses differ from coolant to solvent and many times it is a large part of the finished product. The monitoring of moisture content is many times crucial to the final product quality. To determine the moisture content can be a tedious process with offline detection such as loss on drying (LOD). Inline determination can give a more rapid response and a growing measurement technique for inline determination is near infrared spectroscopy (NIRS). Spectra from NIRS and especially reflectance measurements are rarely free from interferences and may need preprocessing to accurately describe the different samples. The extensive information in a NIR spectrum is difficult, if even possible to interpret by visual inspection and hence powerful methods of multivariate analysis techniques are applied to remove highly correlated data and create models where the most prominent variances in the data are used. [1, 2]

The use of near infrared measurements has proven to be a valid method for quantification of chemical constituents in many different fields, ranging from protein determination of wheat to the ocean skin surface temperature [2,3].

Thanks to rapid data collection in the form of NIR spectroscopy, large amounts of valuable information can be obtained in the matter of seconds instead of tedious offline measuring techniques such as loss on drying (LOD) or Karl-Fischer titration. Other measurement techniques available today use microwave radiation, nuclear magnetic resonance, conductivity or similar properties to find correlation to the moisture content. [1] Except the clear downside of time consuming techniques these methods rely on homogeneity of the sample since small changes such as a decrease in alkalinity will change the electric conductivity without a change in the moisture content. NIRS is hence more selective in measuring moisture. [4]

To process the data acquired multivariate techniques are in common use. [ A reduction from the full spectrum may be done by using a wavelength selection procedure, reducing the amount of data needed for calibration. Simple calibration is often utilized and accepted as an accurate measurement technique. The small amount of information needed for simple calibration is an advantage but when many factors affect the result consideration of additional data might be needed. However these methods can be adequate when only one observation is studied and other factors are rejected. As done in previous studies factors such as distance, angles and packing procedures are held constant and hence they are not considered as a large contributor to the result. In measuring fiber these factors need to be considered due to the shear nature of the uneven surface of fibers bundles [5,12,10].

Offline determination requires the removal of representative samples from the production line and transport to laboratory facilities for analysis. This time-consuming process along with analyzing processes like LOD or titration will render the determination outdated before it is finished. The total throughput in many line processes is far too rapid to respond to analysis data coming several hours after the batch is done. The ability of rapid process control is hence completely diminished. In worst case entire batches might go to waste since they are outside the specification range.

Atline analysis is a bit more effective since transportation and such is removed but the measuring technique still requires working up of the samples and the method of analysis might still be time consuming. Online measurements are an improvement of the atline techniques but automated measuring equipment are attached to the production line which gives a more rapid response and reducing measurement uncertainty thanks to an automated process. The sample is removed from the process.
Inline measurements are the most rapid type of equipment since it assesses the products on the production line by direct measurement without the need for sample removal. [1]

Many of these measurement techniques require calibration, often by the use of multivariate techniques further described in section 1.3 [1].

1.1 The purpose of this thesis

- To develop fundamental knowledge of how the moisture content of fibers could be interpreted by means of investigating the reflected spectrum of the incidence light.
- To evaluate the physical layout of the measuring probe and how the light is reflected in the different angels for use in further studies.
- To evaluate where interesting wavelengths regions to use in calibration with reduced amount of data points can be found.

1.2 Theory

1.2.1 Electromagnetic radiation
The electromagnetic spectrum consists of all wavelengths of electromagnetic radiation. There is wide acceptance that radiation and its interaction with matter contain large amounts of information.[5] Depending on the receiving module, spectrometer or single diodes of various semiconducting materials, interactions from the different regions of the spectrum may be analyzed. An example of a complex yet common detector is the human eye. Electromagnetic radiation of wavelengths between 400-750 nm allows the distinguishing of colors. [6] Other detectors such as semiconductors assist in collecting data from parts of the spectra, invisible to the human eye. [2]

The propagating radiation, known as photons is described both by a wave motion or a particle, known as the wave-particle duality. [6] For the purpose in this thesis the wave description is more suitable over the particle description. Using the wave theory the motion is described by a sinusodial wave as seen in Figure 1. Orthogonal to the propagation of the radiation is the corresponding magnetic field caused by the moving photon. This magnetic field will cause a response from molecules with a dipole moment, permanent or induced. Because of the charge difference of a polar molecule it will try to align to the magnetic field and start to stretch, bend or rotate. This induced motion will oscillate back and forth when the direction of the magnetic field is changed and hence whatever the type of movement energy with a corresponding frequency will be absorbed to facilitate the motion. [7]

![Electric and corresponding magnetic field of propagating light](image)

Figure 1 Electric and corresponding magnetic field of propagating light
Figure 2 shows the electromagnetic spectra and what interaction with matter the different regions is connected with, the narrow range used in this thesis is emphasized.

Figure 2 Electromagnetic spectra with the small fraction of visible and NIR radiation emphasized.

The different interactions are used in a wide range of applications. Many of which are encountered on a daily basis such as radio broadcasts, microwave oven heating, x-ray machines to name but a few.

### 1.2.2 Interaction with matter

Radiation interaction with matter is not a continuous process but follows the laws of quantum mechanics. Energy packages, quanta, must fit an energy gap to interact; this is a quantum mechanics selection rule. Imagine radio waves traveling through walls and even more dense materials being only slightly distorted. It does so because no energy gap in the solid wall is likely to absorb this amount of energy, even though exceptions exists, explaining the attenuation that even radio waves experience.

The energy, $E$, of a quanta is described by the Bohr equation, where $\nu$ is the frequency and $h$ is the Planck constant

$$E = h \nu$$  \hspace{1cm} (1)

All different processes of electromagnetic interaction with matter will be governed by a selection rule. Electron configuration change, excitation, energies must fit into the electron excitation gap whilst the infrared absorption must fit the selection rules and energy gaps of bond vibrational energies.

The energy difference, gap size, $\Delta E$ is given by Equation 2

$$\Delta E = E_2 - E_1$$  \hspace{1cm} (2)

The selection rules are not exact and the absorption band will be broadened for a number of different reasons. For instance physical effects such as the Doppler effect will affect radiation through gases as well as colliding atoms. The lifetime of the excited state is finite and changes with time, hence the absorption energy will also change, affecting the band broadening [6,7].

### 1.2.3 Infrared

Infrared (IR) absorption occurs at wavelengths between 0.75 – 200 $\mu$m. Interaction with matter in this region is due to atomic bond absorption. The infrared range can be subdivided into three major parts, near-infrared, mid-infrared and far-infrared. The major differences being what energy transition is
prominent. Near infrared radiation deals with overtones from the fundamental vibrational wavelengths around 6000 nm [7].

The basic idea in bond absorption is that a bond is never fully rigid but will oscillate around a given equilibrium point. This oscillation may be enhanced to a higher energy state when irradiated with the corresponding wavelength [5].

1.2.4 Near infrared

Near infrared (NIR) radiation is the segment of radiation just above the range of visual light, approximately 750-2500 nm. In this region, as mentioned before, the overtone bands of prominently O-H, N-H, C-H and S-H and combination bands are useable out of analytical point of view. The response from inter- and intramolecular bonds is different depending on the surroundings as explained previously. Even though the similar bonds mentioned exist in the cellulose, hemicellulose major constituents of wooden fibers [12], these will only have a minor contribution to the absorption bands in the spectra. The abundant response of water molecules will render the spectra of reflected light to contain most information about the presence of water and this study focus solely on moisture. Other constituents might alter the spectra and calibration for any application will need to consider this fact. From interaction point of view, it must be recognized that it is moisture that might disturb the spectra considering other constituents. [8,19]

In the NIR region, water absorption bands are expected at 744, 980, 1458 nm [5] from theory and at 1940/1945 nm from previous studies [2]. These are both combination bands and overtones. The contributions are shown in Table 1, adopted from [2].

<table>
<thead>
<tr>
<th>Theory</th>
<th>Observed</th>
<th>Type of absorption</th>
<th>Vibrational assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>744</td>
<td>760</td>
<td>O-H third overtone</td>
<td>Stretch</td>
</tr>
<tr>
<td>980</td>
<td>970</td>
<td>O-H second overtone</td>
<td>Stretch</td>
</tr>
<tr>
<td>1458</td>
<td>1450</td>
<td>O-H first overtone</td>
<td>Stretch</td>
</tr>
<tr>
<td></td>
<td>1940</td>
<td>O-H Combination</td>
<td>Stretch and deformation</td>
</tr>
</tbody>
</table>

Deviations from these theoretical values of absorption often occur due to phenomenon such as the Doppler effect but knowing the theoretical value can help in peak detection and identification since the region where to look for a specific compound is identified.

To utilize NIR in measuring solid material a common method is to record the reflected light from a surface [7]. There is a possibility to measure the radiation passing through solids but the samples need to be very thin since the radiation in the NIR region only penetrates the sample surface. The penetration depth change along the NIR spectrum and at longer wavelengths the penetration depth range from a few microns to a several millimeters [2, 4].

1.2.5 Overtones and combination bands

In the case the energy transition is from the ground state and up to energy level 3 or above, or whenever $\Delta n > 1$, the energy of absorption is known as an overtone. Overtones are between 10-1000 times less intense than the fundamental and overtones past n=4 is of little use for analytical purposes [2].
Absorption bands might arise because of the additive absorption from different energy transitions and hence create bands that do not correspond to a specific energy transition but two different. For example ν₁ + ν₂ can give a third band, ν₃, except ν₁ and ν₂ alone. Water absorption at 1940 nm is one such combination band. Since combination bands are governed by at least two absorption frequencies, it makes the band a little more selective. Hence it is easier to assign a specific compound to a combination band [5].

The absorption considered here is the first and second overtone and combination band of stretch and deformation as seen in Table 1.

### 1.2.6 Modes of vibration

Different modes of vibration will be prominent in different molecules. There are major differences in the modes of vibration of a linear and a non-linear molecule and hence the total number of translational, vibrational and rotational modes is different. The oscillating movement is easily understood by looking at the configuration of the water molecule. The small hydrogen atoms are bound to the electron-drawing oxygen molecule. This creates the electric dipole. The angles and length of the bond is changing with that of the bond characteristic wavelength, giving each molecule a specific response to radiation.

Possible movements are translational, rotational and vibrational. Translating and rotating movements are readily understood but the vibrating modes are more intricate. Vibrations are a group consisting of stretching and bending movements. Each bond might stretch and contract either in a symmetric way or in opposite directions (asymmetric). Deformation inducing a change in the bond angle is known as bending. The different types of bending are scissoring, rocking, wagging and twisting is depicted in Figure 3.

![Figure 3 Modes of vibration for a water molecule. a, Asymmetric stretch b, Symmetric stretch c, scissoring d, rocking e, wagging f, twisting](Image)

Water is also a very special kind of molecule with its ability to form hydrogen bonds. This ability renders water highly adsorbable to surfaces which alter the bond response to radiation. Water in gases, liquids and solids is said to exist in different forms, meaning that even though they are similar water
molecules will exhibit slightly different properties depending on the surroundings. This is also mentioned in the section regarding free and bound water [9].

With this in mind one can imagine the bond as a harmonic oscillator with its own vibrational frequency and hence its own wavelength. This sinusoidal wave will be able to absorb incoming energy with of the same frequency and become amplified. Since the new energy state of the bond is not favorable from energy equilibrium point of view the emission of a quanta will spontaneously occur and the bond will return to its more stable state. The most stable state is considered to be at room temperature or a given reference temperature. The harmonic oscillator is used predict fundamental frequency, calculated from a function deduced from Hooke's law and displayed in Equation 3[5]

\[ v = \frac{\pi \sqrt{k}}{2\mu} \]  

(3)

where \( k \) is the force constant of the spring and \( \mu \) is the reduced mass.

The difference in reduced mass between O-H, N-H,C-H as these are 0.89, 0.87, 0.85 [5] respectively isn’t very different but the electron drawing or donating properties of the constituents are changing the force constant or tightness of the spring. This force constant, \( k \), is therefore altering the absorption frequencies of the different bonds.

Hooke's law is applicable in macro systems but when changing arena towards the nano- or Ångström or even smaller scales, the quantum mechanics model need to be applied. This is because the continuum of the macro scale no longer applies but the interaction energies are quantified. With starting point in the complex time-independent Schrödinger equation a simpler connection with high resemblance to Equation 3 may be deduced. This connection is displayed in Equation 4. In this equation the quantification is accounted for. For deduction and further information the reader is advised to read for example. [5]

\[ E_v = (v + \frac{1}{2})\hbar \nu \quad (v = 1, 2, 3, \ldots) \]  

(4)

This equation explains the energy absorption that occurs at the fundamental absorption frequency but also the overtone bands that will arise when absorption takes place over multiple energy levels, as \( v \) is allowed to take any real value.

1.2.7 Sampling
The major problem of moisture measurement is ensuring that the sampling will render accurate and precise results. Especially difficult is the collecting of samples when the component to be measured is both omnipresent in the surrounding air as well as being an easily vaporized component when the sample is taken from a hot production line. To evaluate the accuracy of the samples will help in evaluating how well the calibration performs. A common method of working up the samples is grinding or evenly packing a cell with powder using a constant pressure to reduce the effect of interference from multiple reflection centers and angles. In one study, [12] they used sanding to smoothen the surface of fresh sawn wood to reduce this effect. These preprocessing methods are not suitable for inline determination. To acquire samples for calibration at a given production line a sampling cyclone or other equipment must be installed in the same line as the measurement probe will be used. [17]
1.2.8 Forms of water
Water might be present in many forms in solid samples. It might be adhered to the surface, bound in the bulk or present as vapor along the surface. Measurements done at a distance will inevitably also respond to vapor in the surrounding air. Depending on the aggregation form the water will display slightly shifted absorption peaks in the spectra. It is possible to separate the different aggregation forms [ref] but since the absorption bands are very broad the differences are likely to be of minor importance when building fundamental knowledge. Due to this absorption differences the distinction between free and bound water should be made to be incorporated into further studies to increase predictive capability of the models built.

1.2.8.1 Free water
Water that is not adsorbed to a surface or bulk but is free to move is referred to as free. This water is in excess to the equilibrium moisture content of the material. Minor or totally unaffected by the surroundings the free water molecules absorb radiation at the fundamental frequencies characteristic to O-H bonds. [10]

1.2.8.2 Bound water
Water molecules that through adsorption and capillary condensation [9] or by any other mechanism adheres to the surface or diffuse into the bulk of the material is experiencing a different surrounding and hence the absorption might be shifted due to the forces acting on the molecule, changing the conditions for the molecule movement. [10]

1.3 Multivariate analysis tools
In this section standard linear algebra notations will be used where applicable to explain the procedures in which the mathematics works when applying the multivariate analysis tools. Bold capital letters denotes matrices and minor bold letter denotes a vector. For deeper descriptions of the techniques the interested is referred to [11]

1.3.1 Multiple linear regression
Multiple linear regression (MLR) is a technique most suited for systems with linear response, no interference, low noise and no co-linearites. Another downside with MLR is the absolute use of all the available data. This means that even data that is totally uncorrelated to the response is used to build the model. An example of this downside is the adding of totally unrelated data or noise may create a predictive model that with even greater accuracy fit the response. Mathematically MLR is described by the creation of an additional matrix t with coefficients that in the best way uses X to reproduce the response vector y while at the same time minimizing the error vector E using a least square approach. Mathematically this is described by Equation 5. To enable this procedure it is required to have more samples than unknowns, since from a pure mathematical point of view, this is required to solve such a matrix without introduction of free variables. [14]

\[ X t + e = y \] (5)

When using large datasets this approach might not be suitable due to the large chance of creating a model with almost a zero e vector but very low predictive ability with unknown samples. Unless a
solid wavelength selection procedure is done, this overfitting will be almost inevitable. How such a selection is done must depend on the data at hand so that a correct interpretation is made possible. How such a procedure should be performed on a given dataset is up to the analyst after inspection of the raw data. More about how such a procedure may be done can be found in the section about wavelength selection.

Spectral differences that are not correlated to what is investigated must not be chosen as a significant contributor since this will make the prediction of an unknown sample inaccurate since the difference apparently is explained by another variable.

High correlation in the original variables may cause unreliable results in MLR.[15]

In Simca P+, a software widely used in multivariate analysis, there is no function for direct MLR calculation but by utilizing a hierarchical model of a PCA X-block analysis and using fewer wavelengths than samples the unscaled regression coefficients are equivalent to MLR.[16]

1.3.2 Principal component analysis
Principal component analysis (PCA) is the first step in conducting a principal component regression (PCR). The eigenvectors for a dataset, $X$, is stored in a new matrix, $V$. To find the score matrix these are multiplied. This gives the score matrix $U$. No consideration of how well these eigenvectors correlates with the response has been made and hence, better models could have been created by taking this into account. However this method is important when studying the raw data and the contributions from the different variables. [11]

1.3.3 Principal component regression
Principal component regression (PCR) is based on the findings of a PCA. It is basically a MLR calibration but now based on the new variables from the PCA which explains most of the variance in the calibration data. In SIMCA, PCR is performed by doing a PCA on the X-block and set this to a hierarchical base model and then a PLS built on these new variables will be equivalent to a PCR. Continuing from the previous section on PCA, the score matrix $U$ is now used to regress $Y$ via the creation of regression coefficient matrix, $S$. The model is now able to make prediction of a new response set $Y_2$ by multiplying the new dataset $X_2$ with the eigenvectors found before, $V$, to obtain $U_2$. $U_2$ is now multiplied by $S$ to obtain $Y_2$. In creating $S$, the calibration $Y$ data is used, even though no consideration whether this is the best way to predict $Y$ using the ingoing variable set $X$ has been made. [11]

1.3.4 Partial least squares
Partial least squares (PLS) is a method that utilizes the response information in building the model and hence it is known to perform better in predictions as it reduces the potential use of data totally uncorrelated to the response. Score and loading matrices are built to represent the $X$ data and at the same time also doing the same thing for the $Y$ data. In short the $X$ data is used in estimate the factors (scores and loadings) to $Y$ and vice versa. The result of this is given in Equations 6 and 7.

\[
X = TP + E \quad (6)
\]

\[
Y = UQ + F \quad (7)
\]
E and F are the errors in the calibration. Minimizing these are important but overfitting, the use of unimportant variables, must be avoided to obtain models with accurate predictions. T and U are the scores for each matrix and P and Q are the corresponding loadings.[11,13]

1.3.5 Preprocessing
Data acquired from spectrophotometric measurement are rarely free from noise or physical interactions and hence some kind of preprocessing scheme could be applied to remove such variations. The easiest way to preprocess data, often already applied in the spectrophotometer software is some kind of averaging. Averaging might remove important information in the raw data since each point will be an average value of the closest points.

Data from NIR spectroscopic measurements are often displayed as log(1/R) values, thus transformed into absorbance since the same mathematics is used for transmittance, equation (8), as for reflectance as well as removing some nonlinear relationships in the data. [8]

\[ A = \log \left( \frac{1}{R} \right) \]  

(8)

To evaluate the model in terms of how well it performs by validation the root mean square error of calibration (RMSEC) and the root mean square error of prediction (RMSEP) was used. The equation for calculating RMSEC and Rmsep is displayed in Equation (9).

\[ \text{RMSE(C,P)} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{(c,p)i} - y_i)^2} \]  

(9)

1.3.6 Multiple signal correction
A common preprocessing technique is the multiple signal correction (MSC). This procedure averages all the spectra and calculates the deviation from this mean spectrum and creates new datasets based upon the differences. In this way the information gathered even though slightly different intensities may be used in the same calibration set. This is in common use in reflectance measurements. MSC is often utilized when different intensities or slopes might be present among the spectra.

1.3.7 Derivative
If baseline or intensity differences are prominent the use of derivative might be useful since this technique take advantage of the slopes of the spectra without giving the original data of intensity any influence. A downside with this procedure is the splitting of peaks into two smaller peaks since the start of the peak will have a positive derivative the end of the peak will have negative values.

1.3.8 Wavelength selection
If an expected large contributor to the spectra are variables not correlated to what is of measurement interest, careful selection of wavelengths, highly correlated and non-correlated wavelengths might be chosen to represent the data set. Knowledge such as where the absorption bands are prominent and where none would be expected is of interest. This procedure is applicable when other constituents or interactions may interfere with the resulting spectrum and must be neglected in calibration.[18]

1.4 Sample homogeneity
In the case of inhomogeneous samples the results from calibration and prediction are less likely to be accurate compared to a homogenous sample. Controlling the moisture content of the samples is needed to verify the sample moisture content and allow for correct evaluation.
1.5 Reflection

Reflection from a surface can be due to both specular and diffuse reflection. Both will contain information about the constituents of the sample. The radiation that has travelled the furthest and interacted with the sample most, should contain even more distinct differences. [Ref] The downside is that the intensity of this radiation will be very low and hence noise levels will be high.

Diffuse reflection may enable radiation to travel very long distances within a sample and the spectra obtained from this will be more affected by the constituents since it has interacted with the sample several times rather than only reflected once. However the intensity of this radiation is very low and will suffer severely from specular reflection interaction. [2]

**Figure 4** Specular and diffuse reflection
2 Material and method

2.1.1 Light source
The light source used was an ocean optics tungsten halogen bulb, HL 2000. This light source gives continuous spectra from 360-2000 nm. Outside of this range the intensity is very low but detectable. This source of light is of great use together with the InGaAs spectrometer with the peak sensitivity within the same region.

2.1.2 Fiber optics
The fiber optics used are 600 µm in diameter NIR transparent use in UV-VIS. The bonds in the material does not respond to the wavelengths used in the experiments in any significant way and hence they are called IR-transparent. Studying water low fiber response to OH bonds is important and fibers used for this purpose are known as low OH fibers The light traveling through fiber optics is undergoing total reflection when it hits the fiber wall and hence the light might travel large distances without any loss in intensity. For the interested the physics behind total reflection might be found in for instance [6]. The same rules governing total reflection also dictates how the incoming and outgoing light will be collected and diffracted. For distance and angle dependence measurement this is very interesting since the illuminated area and the intensity of the illumination will change with altered distances and angles. As Figure 5 shows, the illumination beam will reflect intensely at short distances and low angles. This will decrease and drop quickly when moving outside of the beam boundary. Radiation traveling through an optical fiber is bouncing against the walls, this will create a light beam exiting the fiber in the shape of a cone. This cone represents the most intense radiation even though radiation will be scattered in all direction to a minor degree. Measuring the direct reflectance from this cone will render the highest intensity due to the specular effect of the sample surface. The probe used in the experiments has the receiving fiber very close to the light source fiber and hence a large angle is expected to move outside of this high intensity reflection area.

2.1.3 Spectrometer
The spectrometer used was an ocean optics NIR-Quest 256 2.1. This spectrometer is an InGaAs multiple array detector. The response is between 900 and 2150 with the useful range of 900-2100. The peak response is at 1900 nm. (Food and beverages) The radiation entrance is through an optical fiber connected by an SMA 905 connector. The light is collected using a convex mirror and passed through a grating with 100 l/mm. The scattered photons will hit the semiconductor material and induce an electrical current which will be measured and sent to the software to be interpreted and presented. The resolution for the spectrometer is <8 nm. Each spectrum consists of 256 data points between 900 nm and 2150 nm. The signal to noise ratio is 10000:1 at 100 ms integration time. Mostly longer integration times will be used thus giving even higher signal to noise ratio.
2.1.4 Probe

2.1.4.1 Reflection
The probe used in the experiments is a stainless steel cylinder with the optical fibers running in from the back and out at a 30 degrees angle at the top. The use of three fibers on a row enables the introduction of a “near” and a “far” fiber according to Figure 5. These are hence used to allow two different ways of measure, the major difference being the intensity of the reflected light due to light path lengths.

![Figure 5 Source and receiving near and far fiber](image)

2.1.4.2 Transmittance
The probe used for transmittance is similar to the reflectance probe but instead of measuring the intensity reflected back from the surface its measures the radiation travelling through the material. Hence there is a very important difference in the sense that the sending and receiving fiber is aligned to allow the beam to travel straight into the receiving end after it has traveled through the sample. This technique is more suitable in liquid sample measurement since solid material easily could clog the slit between sending and receiving fiber.

2.1.4.3 Separating the optical fibers
To investigate how the response would change if the receiving fiber had a different angle than the light source to reduce the amount of specular reflectance two new fibers where constructed. The ends stabilized by steel tubing and SMA 905 connectors in the ends connected to the light source and spectrometer. Instead of the probe with closely connected fibers the new fibers were independent to each other to enable the change of splitting angle. The receiving fiber was attached to a protractor to allow the angle to be changed in a controlled manner. Two different radial distances 3 and 6 mm were tested in increasing angle, from 20 up to 70 degrees. 0 to 20 and 70-90 was excluded due to physical problems in the setup making these measurements. More tilt should equal less specular reflectance according to Figure 4, even more in this case since the fibers are separated. The idea of this setup is to determine how the change of angles alters the water response with less intensity of the refracted radiation with an angular difference between the source and the receiver.

2.1.5 Loss on drying
The method used to determine the moisture content of the calibration and validation samples is the loss on drying (LOD) technique. This simple but time consuming scheme is routinely used for moisture content determination. The method is not applicable when other volatile solvents might be present thus giving a false estimation of the water content. In this thesis the only solvent used is water and other drying compounds are neglected.
2.1.6 Software
SpectraSuite is a software for interpretation of the acquired data. This program is provided by Ocean Optics together with the spectrometer.

Simca P+ is a software for multivariate analysis. Throughout the thesis it will be used for analysis of the spectral data if nothing else stated.

Excel is used for basic data handling and graphical presentation of the data.

2.1.6.1 Automatic weighing
To enable spectral measurement on various moisture levels, keeping variables such as angle and distance at a constant a scale was purchased. The Sagitta 100/0,01 scale was connected to a RS323 connector via a 9-pin to 25-pin converter according to the resellers instructions. To record the scale measurement a logging software also provided by the reseller was used. The downside of this setup was that the scale didn’t send measurements at a given rate but instead when it experienced a change of at least a few milligrams. To be able to utilize the setup a programmable logic controller (PLC) was used to control a magnetic valve which in turn controlled the flow of pressurized air into a cylinder controlling a piston. A 50 g weight was attached to the piston head. The PLC was set at an interval of 900/8 s/s, meaning that when 900 seconds passed the piston placed the weight on the scale for 8 seconds and the reading would easily be spotted among the other small changes and acquired from the logged values. This weight was used to calculate the moisture content of the sample at specific times, correlated to the time the spectrum was taken.

2.2 Experimental section
All experiments were performed at the facilities of BTG Instruments, Säffle.

In the measurements with a reference the raw data are given in percentage of intensities. The reflection is hence given by \( R = \frac{I_{\text{sample}}}{I_{\text{Reference}}} \times 100 \). The logarithmic treatment hence gives values in excess of total absorption by a term of 2 since \( \log (100) = 2 \). The absorption is corrected by adding this term, creating a graph with total absorbance at the value 1 and zero being totally unaffected by the radiation interaction with the sample.

2.2.1 Sample preparation
Filter paper (Munktell’s, L3 – 100 – 1000) was prepared by weighing and adding the required amount of water to reach the preset moisture level ranging from 30 to 95 % dry material by weight. After weighing and between measurements the samples were kept in airtight zipper bags to avoid drying.

To produce the fiber samples used for this study bleached CTMP-market pulp from Rottneros was dried and weighed before water was added. Seven samples of various solid fractions, 30-90% were created and homogenized by mixing the moist pulp with a stick mixer. The samples were left for 5 days to reach equilibrium before the experiments were performed. During the equilibrium process and during the experiments the samples were kept in airtight zipper bags.

To check for deviation from homogeneity seven samples were taken from the calibration set and oven dried and left to dry in a desiccator. The average and standard deviations were calculated and a confidence interval was established. The moisture contents of the samples are given in Table 2.
Table 2 Fiber concentration calibration and prediction fiber samples with 95% confidence interval

<table>
<thead>
<tr>
<th>Calibration set</th>
<th>n = 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>Fiber concentration [%]</td>
</tr>
<tr>
<td>1</td>
<td>30,369 ±0,544</td>
</tr>
<tr>
<td>2</td>
<td>41,636 ±0,149</td>
</tr>
<tr>
<td>3</td>
<td>55,898 ±1,921</td>
</tr>
<tr>
<td>4</td>
<td>65,389 ±3,314</td>
</tr>
<tr>
<td>5</td>
<td>72,298 ±0,964</td>
</tr>
<tr>
<td>6</td>
<td>83,665 ±1,069</td>
</tr>
<tr>
<td>7</td>
<td>91,282 ±0,630</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prediction set</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33,245 ±1,542</td>
</tr>
<tr>
<td>2</td>
<td>48,709 ±3,010</td>
</tr>
<tr>
<td>3</td>
<td>44,318 ±3,127</td>
</tr>
<tr>
<td>4</td>
<td>58,761 ±2,895</td>
</tr>
</tbody>
</table>

2.2.2 Angular dependence measurements
Dry filter paper at zero angle was used as the reference and background correction was used. The integration time was set at 400 ms. At this integration time the noise is very low and only one spectrum was acquired. Spectrum for both near and far receiving fiber was acquired in close succession. The spectrum was normalized at 1100 nm to allow interpretation of the spectral differences. 1100 nm was used as the normalizing base since no spectral information regarding water is expected here in between the absorption bands at 970 and 1450 nm. The distance between the fibers and the samples was held constant at 5 mm.

2.2.3 Direct contact measurement

2.2.3.1 Filter paper
The probe was rotated 180 degrees and the samples were put straight on to the lense and a metallic weight was put on top to secure the sample and ensuring that the same pressure was applied to each sample. Both receiving fibers were used in close succession. The spectra were normalized at 1100 nm.

2.2.3.2 Fiber
The fibers used in direct contact with the fiber optics were held by hand and rubbed against the glass protecting the fiber optics to avoid scratching. This procedure was conducted with a steady pace and 2 spectra were acquired each second for a total of 180 spectra divided in three rounds. Integration time was set to 400 ms corrected with the dark spectra and the reference used was dry filter paper inserted into a sample holder.

2.2.4 Distance dependence measurement
Measuring the intensities at different distances was conducted using filter paper only. The spectrum was recorded at the given distance and the plate was successively lowered to increase the distance between the probe and the sample. The integration time used was 400 ms and the reference used was dry paper in a sample holder.
2.2.5 Surface vs bulk moisture experiments
To establish whether there is a moisture fraction gradient (there undoubtubly is one) that renders spectra not really equivalent to the real moisture fraction in the sample a study using filter paper was conducted as follows. 7 pieces of filter paper was stacked on top of each other after weighin each paper separately. The average thickness of the filter paper was 0.267 mm, found by measuring the 7 piece bundle with a caliper and averaging the result. The thickness of each paper indicated the depth at which the moisture content was investigated. Water was added to the sample bundles and left to dry. Two bundles in the open air and one bundle beneath a hood, creating a more controlled environment as would be expected measuring inline. The samples were left to dry for one hour before the first weighing. After weighing the bundles was identically recreated using the same papers and left for another 3 hours before weighed again.

2.2.6 Water absorption at various temperatures
For this experiment a transmittance probe was used. The reference was air. Integration time was 1 ms. The probe was sunken down into a water bath after the reference and dark spectra was taken. A heater increased the temperature of the bath and new spectrum was taken at various temperatures. Before each measurement the gap between source and receiving fiber was cleaned from bubbles. In comparing temperatures the 8.5 °C spectra was used as reference.

2.2.7 Temperature dependence measurement
The temperature dependence of pure water absorption was investigated through transmittance tests where the reference was air. The temperature will have an effect on the absorption wavelength approximately 3 % according to [2]. The point of interest here is to see that with appropriate wavelength selection, the absorption peaks may still be valid since only the energy of the absorbed radiation change, and not the extent to which it absorbs.

The temperature dependence measurements on wooden fibers were performed by using the same samples as the contact and distance measurements but pretreating the closed bags with lowered and elevated temperatures. A fridge and an oven were used. The fridge was set at 8 °C and the oven at 70 °C. The same procedure was performed for each measurement. 20*3 scans were averaged to calculate the average spectra. The scans were performed by moving a bundle of fibers onto the fiber optic probe.

2.2.8 Contact measurements on fiber
Fibers were rubbed against the optical fibers and 2 spectra per second was acquired. A total of three sets with 20 measurements per set were attained with samples in random order. Integration time used was 400 ms and the reference was dry filter paper.

2.2.9 Distance measurements on fiber
To acquire spectra from the fiber samples at a distance the plate with the sample was positioned below the probe giving the sample a gap of approximately 10 mm. This distance will vary as much as a few mm between the spectra when the plate is moved during recording as the sample is uneven and will render spectral data for a number of different distances and angles. The reference used for these experiments was dry filter paper as the reflecting surface. Dark spectrum was used. To compensate for the reduced intensity in the reflected light the integration time used was 700 ms. 3*60 spectra was acquired by recording 2 spectra each second while moving the sample in front of the optics. Severe outliers detected in the raw data by visual inspection were not used for further analysis.

Investigation of the moisture content by measuring at a distance, enables the a larger area to be illuminated and hence give a more representative measurement at each given point in time. The
altering of distance +2 mm and the angle difference will cause a disturbance in the form of intensity differences. The information gathered from incoming light need to be visually inspected to prevent obvious outliers to be part of the calibration. 60*3 scans were obtained for each of the 7 samples. Using integration time of 400 ms enables even spectra to be acquired and smoothing operations will not be needed.
### 2.3 List of experiments

Table 3 below summarizes all the experiments with pretreatment, integration time and more variables reported.

**Table 3 List of experiments**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Preprocessing</th>
<th>Integration time [ms]</th>
<th>Samples</th>
<th>Sets</th>
<th>Repetitions</th>
<th>Probe</th>
<th>Recieving optical fiber</th>
<th>Variabel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water absorption Angle</td>
<td>Water</td>
<td>Log(1/R)+2</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>-</td>
<td>Temperature</td>
</tr>
<tr>
<td>Contact</td>
<td>Filter paper</td>
<td>Log(1/R)+2, Normalized</td>
<td>400</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>R</td>
<td>Near/Far</td>
<td>Angle</td>
</tr>
<tr>
<td>Distance</td>
<td>Filter paper</td>
<td>Log(1/R)+2</td>
<td>300</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>R</td>
<td>Near/Far</td>
<td>FPC</td>
</tr>
<tr>
<td>Automated weighing Temperature</td>
<td>Fiber</td>
<td>Log(1/R)+2</td>
<td>1000</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>R</td>
<td>Far</td>
<td>Distance</td>
</tr>
<tr>
<td>Contact</td>
<td>Fiber</td>
<td>Log(1/R)+2, MSC, 1st derivative</td>
<td>600</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>R</td>
<td>Far</td>
<td>FC</td>
</tr>
<tr>
<td>Distance</td>
<td>Fiber</td>
<td>Log(1/R)+2, MSC, 1st derivative</td>
<td>400</td>
<td>7</td>
<td>1</td>
<td>40</td>
<td>R</td>
<td>Far</td>
<td>Temperature/FC</td>
</tr>
</tbody>
</table>

Experiments used for multivariate analysis

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Preprocessing</th>
<th>Integration time [ms]</th>
<th>Samples</th>
<th>Sets</th>
<th>Repetitions</th>
<th>Probe</th>
<th>Variabel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact</td>
<td>Fiber</td>
<td>Log(1/R)+2, MSC, 1st derivative</td>
<td>400</td>
<td>7</td>
<td>3</td>
<td>20</td>
<td>R</td>
<td>FC</td>
</tr>
<tr>
<td>Distance</td>
<td>Fiber</td>
<td>Log(1/R)+2, MSC, 1st derivative</td>
<td>700</td>
<td>7</td>
<td>3</td>
<td>60</td>
<td>R</td>
<td>FC</td>
</tr>
</tbody>
</table>

* T = Transmission

* R = Reflection
3 Result and discussion
The results are presented in the order given in Table 3.

3.1 Angle dependence
The difference in the information gathered at different angles is displayed below in Figure 10-13.

**Figure 6** Angle dependence graph. FPC 30 % Near fiber

**Figure 7** Angle dependence graph. FPC 30% Far fiber
Differences in the moisture content do obstruct the graphs as the surface of the filter paper will change due to the presence of water. The altered conditions for the reflection, such as light scattering due to altered paper structure will change how the acquired spectra look from different angles. A difference is obvious when the moisture content of the paper is increased. The paper is becoming increasingly saturated by water and rather than swelling the fibers creating the uneven surface the water will start to fill the cavities and create a smooth surface. This change might induce problems in calibration since measurement of the surface might carry a large input from for instance a water droplet. The smooth surface will reflect the light in a uniform way and hence the information gathered in the spectrum will be less altered than if acquired from a rough surface. This is obvious from both the receiving optical fibers but even more prominent when using the near fiber. A critical point seems to be around 30° where the angle dependence decreases. This is most likely because of the lowered intensity of the reflected light outside the immediate beam. Hence the angle will have a great effect of the spectra gathered, as was recognized in previous studies where the sample preparation of solid samples consisted of ensuring a smooth sample in the form of powder or grinded surfaces. Since there is a difference in the information gathered at different angles further measurement and spectra
interpretation must consider this to be affecting the results whether the measurements are performed inside or outside of the most intense specular reflectance beam.

3.2 Contact measurements
To reduce the effect of angle on the spectra the measurements were conducted with the sample filter paper in direct contact with the lens, like a contact probe used in other studies. The two different fibers, near and far results are displayed in Figure 10 and 11.

Figure 10 Contact measurement FP Near

Figure 11 Contact measurement FP Far

The Near optical fiber seems to suffer from significant spectral interferences. The smoothening effect of water reduce the reflection of the entire spectra and obscures the absorptions bands expected as light coming from the source and only reflected on the surface directly will enter the reading fiber. This light will be relatively unaffected by the sample and hence distort the spectra depending on the surface structure of the fiber at the given measuring point. For contact measurement the far fiber gives a more easily interpreted difference between the samples according to their moisture content. In both of the figures the 50 % dry sample is deviating from the expected pattern. Even though the filterpaper
were difficult to keep at a constant moisture content level, the difference here seems to indicate a
difference in the surface giving values closely to the even more dry samples.

### 3.3 Distance dependence

The far fiber displays a maximum at 3 mm distance as seen in Figure 12. The intensity in the
absorption bands does not display the same intensity change as the wavelengths without the expected
absorption. This result in information change due to the distance difference against the reference, it is
clear that the spectra change as the distance increase, also changing the information gathered in the
interesting regions around the water bands and even more prominent change can be spotted in regions
where water isn’t supposed to alter the spectra. This indicates that the distance between the probe and
sample is a large contributor to the spectra and to estimate moisture content, this variable must be
controlled or accounted for in calibration. Choosing spectra with the same intensity in regions not
affected by water is plausible when dealing with this change but this requires further studying.

![Figure 12 Measuring at increased distance using the far fiber](image)

![Figure 13 Far fiber distance measurement against the reference](image)
The same increase in distance was conducted at a 30 degree angle. This removes the 3 mm maxima and instead the intensity decreases with distance displayed in fig 14. The intensity reduction doesn’t affect the wavelengths equally as in the previous case at 0 degree angle. The absorption in the water characteristic regions is high and hence light reflects poorly. Against the reference, the 30 deg angle experiment doesn’t display the same maximum at 3 mm, most likely so because the specular reflection is low and the spectra will consist of more diffusely reflected light. This radiation contains more linear information about the moisture content. The intensity will decrease with distance and hence this variable need to be accounted for and at an angle seems more straight forward rather than the apparent nonlinear response from the specular reflectance dominated sample from 0 degree measurements.

**Figure 14** Distance measurement at 30 degrees using far fiber.

**Figure 15** Distance measurement using far fiber compared to the reference

### 3.4 Automated weighing

The setup with a scale was run overnight and recorded. The dryness varies from 74-84% according to scale reading when a given spectra is recorded. The distance was set to 5 mm in a 30 degree angle. In Figure 16 the raw spectral data are presented. The intensity difference seen when the sample is dried ought to contain information regarding moisture content but the increase in intensity over the entire
spectra reveals that direct reflection from the surface increase as the sample becomes more dry. Using the reference to compare the spectra reveals that large differences are present through large parts of the spectra, not only restricted to the characteristic water absorption bands, displayed in Figure 17. The change in moisture content should generate larger differences in these areas but change in the structure of the drying fiber and hence the altering of the specular reflection seems to dominate the result.

**Figure 16** Automated weighing raw data. FC 74-84%. Measured with a 5 mm distance in an angle of 30 degrees. The black line is the reference. More dark red equals lower fiber concentration.

**Figure 17** Automated weighing data. After reference and log 1/R processing applied.
3.5 Temperature dependence of wet fibers
How the spectra change with temperature when measuring fiber heated, cooled and room tempered are displayed in Figure 18.

![Figure 18: Fiber with various temperatures and moisture levels](image)

The effect of elevated temperatures is less stable spectra, especially in the upper range of wavelengths. The differences in absorbance for the wavelengths are strong around the absorption band (1450 and 1940 nm). A difference can also be seen in between the bands, an effect that is accounted for by the wide absorption band of water. The room tempered samples and the heated samples show high resemblances whilst the cooler sample shows all through the spectra lower absorption. This is expected since the heated samples are cooled very rapidly and the difference between the samples is likely fairly low even if the samples were removed from the oven and measured at once. Constant temperature chambers should be used to achieve accurate temperature measurement above room temperature. The cool samples were measured straight from the refrigerator and hence had very little time to achieve temperature equilibrium with the surrounding room temperature. Since this creates large differences in the calibration for any kind of preprocessing method or simple calibration the temperature seems to be a very important factor and the reflectance of even higher temperature is of interest for future studies.

3.6 Separating the optical fibers
The results from these measurements were left out due to experimental data giving spectra that could not be correlated to the moisture content at all. Possible errors with the equipment are fiber optic heads of poor quality and rigging problems. Contribution from distance and angles especially as they were both changed simultaneously to some extent was far too great to incorporate the spectra into any meaningful comparison.
3.7 Fibers against the fiber optics

The spectra acquired displays responses with varying intensity, expected since light scattering effects will affect the result. This effect is increased as the moisture content decrease as seen comparing Figures 19, 20 and 21 with increasing FC. The smoothing effect of water seems to create more stable spectra. The difference between spectra from the same sample indicates the need to use average results. The spectra are averaged as seen in Figure 22 followed by a normalization procedure in Figure 23. After the normalization the different moisture contents are prominent around the absorption bands as predicted by the theory even though some distortion might have been induced by the normalization procedure.

**Figure 19** Raw data contact measurement FC 30.4% Receiving fiber Far.

**Figure 20** Raw data contact measurement FC 41.6% Receiving fiber Far.
Figure 21 Raw data, contact measurement FC 83.7%. Receiving fiber Far.

Figure 22 Averaged raw data for various FC. Receiving fiber Far.

Figure 23 Averaged and normalized spectral data for different moisture content. Data acquired using the far fiber.
3.8 Applying multivariate analysis preprocessing techniques to the spectral data

3.8.1 Contact measurement

3.8.1.1 MSC

As seen in Figure 24 the MSC pretreatment do show the water absorption bands in the expected order. The same is not true along the entire spectra, where other effects than moisture content such as light scattering change the spectra. Below the 1950 water band the same is true and the more narrow absorption here is displayed by the little part of the spectra where this is true.

3.8.1.2 Building a model based on the MSC values using PLS

The loading for the first component is displayed in Figure 25. The influence of the characteristic water absorption bands are clear since peaks both around 1450 and 1940 are spotted. This shows how these wavelengths contribute to the resulting model. The second components loading also contribute a lot to the result and are displayed in Figure 26. The water absorption bands are clearly important here as well but also a large contributor is found at around 1880 nm.

Figure 24 MSC treated data, contact measurement. Varying FC

Figure 25 P-loading component 1, contact measurement
Together with the t-scores as seen in Figure 27 the connection found by Simca shows separation of the groups of samples with equal moisture content. The overlapping of groups does present a problem seen in the observed vs predicted plot in Figure 28. The predicted values spread widely which makes the results of prediction highly uncertain. Nevertheless, to use average values seems like a plausible idea even though any single measurement likely will end up outside any acceptable limit of prediction error.
Figure 218 Observed vs Predicted, MSC calibration data, contact measurement
3.8.1.3 Derivative

**Figure 229** 1st derivative preprocessing

Using the first derivative pretreatment as displayed in Figure 29 creates spectra that show high correlation to the moisture content for many wavelengths. The derivative pretreatment considers the slopes of the spectra rather than the absolute value and hence signals with different intensity might be used in the same calibration. But this result must be restricted not to cover very large differences in signal intensity since previous results have indicated that the information change according to distance and angle of reflection. This spectrum however may easily be interpreted for moisture content by looking at the derivative around the absorption bands. A steep slope displayed as a peak of the derivative spectra indicates more moisture.

**3.8.1.4 Building a model with the 1st derivative treated data using PLS**

The first component in the derivative model shows a few wavelengths that contribute to the result, seen in Figure 30. The large contributions from the wavelengths close to 2000 nm may present problems in calibration as both the light source and the spectrometer has decreased intensity and sensitivity in this region making the spectra more sensitive to noise.

**Figure 23** P-loadings 1st component 1st derivative preprocessing
The second component loading shows large contribution from wavelengths just below the 1940 nm band as showed in Figure 31. Since this combination band previously has shown to be less wide than the 1450 nm band it is plausible that information regarding the moisture content might be present here even though seemingly lot more prominent than the 1450 nm band in the first component.

![Contact measurement 1st derivative preprocessing PLS component 2](image)

**Figure 24** P-loadings second component 1st derivative preprocessing

The third component shown in Figure 32 displays a peak at the same wavelengths as the second but moving down the list of components other wavelengths not considered to have direct connection to moisture content get taken into account. This is likely to be starting point for model overfitting and more components should not be added even though apparent better fits might be achieved. Further analysis of the third component is done in the interpretation of the scoreplot below.

![Contact measurement 1st derivative preprocessing PLS component 3](image)

**Figure 32** P-loadings 3rd component 1st derivative preprocessing

The scores of the first vs the second component can be seen in Figure 33. The grouping found previously is again apparent but in this scoreplot the samples are spread like a hand fan. The different scores with regards to the moisture content are apparently found in both first components. This cannot be said for the third component portrayed in Figure 34. The spreading of the samples in the y-direction (3rd component) shows no clear separation between the different moisture contents and like suspected before this component is not explaining differences induced by sample moisture content.
The models created show very high correlation coefficients and explained variance as seen in Table 4. The seemingly perfect correlations show that data representing the moisture content levels can be found, not surprising with one response and 256 variables. When using the model in predictions problem seems to arise. RMSEP is so high that relevant predictions cannot be made for the untreated and derivative models but the MSC preprocessing seems to generate a model more suitable for prediction since the RMSEP doesn’t increase in the same magnitude as for the untreated and derivative preprocessing method.
<table>
<thead>
<tr>
<th>Pretreatment</th>
<th>PC:s</th>
<th>R2X</th>
<th>R2Y</th>
<th>Q2(cum)</th>
<th>Rmsec</th>
<th>Rmsep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>4</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>6.57</td>
<td>18.93</td>
</tr>
<tr>
<td>MSC</td>
<td>2</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>5.54</td>
<td>8.60</td>
</tr>
<tr>
<td>1st Der</td>
<td>3</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>7.07</td>
<td>37.98</td>
</tr>
</tbody>
</table>

Table 4 Contact measurement comparison of PLS model result
3.8.2 Distance measurement
The calibration raw data from a FC 30 % sample is displayed in Figure 35, collected from a distance and hence varying intensities are displayed.

![Graph showing absorption vs. wavelength for FC 30% samples](image)

**Figure 35** FC 30% calibration raw data acquired from measuring at 10 mm distance

It is obvious that some samples are severe outliers. Different explanations can be found such as altering distances and angles or light scattering effects when the fiber aligned in a way creating a surface more suitable for direct reflection. These are disregarded. The calibration data for FC 30 % will look like Figure 36.

![Graph showing absorption vs. wavelength for FC 30% calibration data after spectral selection](image)

**Figure 36** FC 30% Calibration data after spectral selection

After spectral selection every set of solid fraction calibration samples will consist of between 15-25 spectra with the appearance as above in Figure 36. To average each set of these will create a graph
looking like Figure 37. Each run is averaged and hence each moisture content fraction will be represented by three spectra.

![Figure 37](image)

**Figure 37** FC 30% calibration data averaged in each set rendering 3 spectra per

### 3.8.2.1 MSC preprocessing

After MSC the spectra will look as presented in Figure 38. The contact measurement where the spectra actually lined up in a way that could be interpreted as differences in the moisture content is not repeated for the distance measurement to the same extent but still correlations to moisture content may be seen.

![Figure 38](image)

**Figure 38** FC 30% Calibration data after MSC preprocessing

### 3.8.2.2 Building a model based on the MSC values using PLS

The loading for the first component presented in Figure 39 do show tendencies to utilize spectral information around the absorption bands in building the model. But still large contribution is found throughout the entire spectra. The second component loading, seen in Figure 40 use very little
information where the known absorption is prominent. The resulting model will hence be of very little use in prediction since apparently other spectral differences than moisture content likely is prominent in the calibration.

![Distance measurement MSC preprocessing PLS component 1](image)

**Figure 39** P-loading component 1 from PLS model built on the MSC corrected values from distance measurement

![Distance measurement MSC preprocessing PLS component 2](image)

**Figure 40** P-loadings second component from PLS model built on the MSC corrected values from distance measurement.

Looking at the score plot in Figure 41 showing component 1 vs 2, groupings can be seen within the same moisture content fractions and the groups do appear in the order expected thanks to moisture content. But this is partially expected as the Y data is used in the calibration so even with this apparently logic correlation the model fit and prediction should not be given credibility knowing that data uncorrelated to the moisture content has been used to build the model.
In Figure 42 the third component shows no clear distinction of moisture fractions groupings. This component doesn’t separate many of the samples at all and without the spreading thanks to the first component differences between the samples would be hidden. This is true for all the data except the 91.3% moisture content.

Figure 42 Score plot component 1 vs 3 distance measurements MSC corrected
3.8.2.3 Derivative

The spectral data after derivative preprocessing is displayed in Figure 43. The result of this preprocessing method seems to create spectral data that can easily be interpreted as moisture content differences. The darker the line the more absorption seems to occur, especially in the regions where it is expected to change due to moisture content.

![Figure 43 Contact measurements after 1st derivative preprocessing](image)

3.8.2.4 Building a model based on the derivative values using PLS

The first component loading from the derivative measurement is displayed in Figure 44. As could have been guessed by the preprocessed spectra a large contribution comes from the regions around the water absorption bands. Unwanted is the contribution from the region above the 1940 absorption band. The very unstable region is used in the model building creating models not correlated to the moisture content. Without changing spectrometer for some receiver with higher response in this region these wavelengths should not be used. The second component, seen in Figure 45 also show contribution from the water absorption bands. Finding this contribution here gives the model higher credibility than if none would be found as is the case of the third component as seen in Figure 46. This component doesn’t contain much information that can directly be correlated to the moisture content.

![Figure 254 P-loading 1st component after 1st derivative pretreatment](image)
The score plot of component 1 vs 2 of the derivative spectra is displayed in Figure 47. The grouping of moisture levels can be seen but already in the second component score there is a lack of separation. As expected from the loading plot the third component as seen in Figure 48 does not separate the moisture content at all and this and the fourth component is likely used to overfit the model.
Score plot component 1 vs 2 distance measurement after 1st derivative preprocessing and PLS

**Figure 47** Score plot component 1 vs 2

Score plot component 1 vs 3 distance measurement after 1st derivative preprocessing and PLS

**Figure 48** Scoreplot component 1 vs 3

**Table 5** Distance measurement Simca evaluation table

<table>
<thead>
<tr>
<th>Pretreatment</th>
<th>PC:s</th>
<th>R²X</th>
<th>R²Y</th>
<th>Q²(cum)</th>
<th>Rmsec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>4</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>8.11</td>
</tr>
<tr>
<td>MSC</td>
<td>3</td>
<td>1.00</td>
<td>0.99</td>
<td>0.98</td>
<td>6.73</td>
</tr>
<tr>
<td>1st Der</td>
<td>4</td>
<td>0.95</td>
<td>0.95</td>
<td>0.82</td>
<td>14.83</td>
</tr>
</tbody>
</table>
3.9 Result summary

From the investigation of angular dependence the intensity decreases at approximately 30 degrees, within this angle the intensity of the reflected radiation is high. Outside the boundaries of intense specular reflection the spectra gathered are more uniformed but still differences can be seen.

In contact measurement the far fiber is the best to use since the near fiber is suffering from being very sensitive to the surface of the sample, distorting the spectra reducing correlation to the moisture content. For contact measurement the far fiber is recommended to use since the near fiber is suspected to suffer from intense specular reflection depending largely on the surface being measured.

The far fiber receives the most intense radiation at a distance of 3 mm, but at the angle of 30 degrees this peak is no longer present. The largest difference with regards to the reference in the received radiation is at wavelengths below the region of interest but the information in the regions of interest is also changing. Since the spectrometers peak sensitivity is high in the region of interest around the water band the change in spectra might still be of importance in moisture determination measurements. Measuring at an angle to avoid large changes in the spectra because of uneven surfaces is hence recommended if measurements are to be done at a distance.

To measure drying fiber will change the spectra for a large range of wavelengths, likely much of the changes will be due to the change in reflecting surface these changes are devious to explain only by moisture content change. The reflecting surface and how this change need to be investigated to say for sure that it is the single largest contribution to the altered spectra or what effect it might have.

The temperature change the absorption of free water by shifting parts of the spectra and the width of the absorption band at 1450 is decreased. The temperature have different effects for the different wavelengths. The largest differences in the absorption is found at 1393, 1505 and 1874 nm. The effect is spread over many wavelengths around the 1505 nm peak, the other peaks being less wide.

Measuring using direct contact renders more even spectra for samples containing more moisture. The drier the sample gets the more it will be affected by other effects such as light scattering. Averaging does render spectra that reveal differences between the moisture content but between adjacent samples the similarities is sometimes very small and averaging procedures need to be applied in order to show this differences. Normalizing the averaged values divides the different moisture fraction samples with an apparent connection to moisture content. Hence giving averaging and normalizing procedures increased credibility. Knowing that the spectra change rapidly with distance and angle averaging and normalization should only be used if stable spectra might be acquired since a technique such as normalizing might distort the spectra severely if large normalization coefficients are used.

Using preprocessing techniques such as MSC and 1st derivative spectral information is more easily visually interpreted. But even with these preprocessing techniques full spectra PLS render models with poor predictive power seen by the RMSEP values. MSC outperforms 1st derivative in both RMSEC and RMSEP and is the preprocessing method seemingly best in fitting the data gathered in the experiments.

The loadings and scores of the PLS models do show tendencies towards using the information around the water absorption bands, at least in the first couple of components, strengthening the possibility for successful calibration and prediction also after wavelength selection. Notice should be taken of the 1940 nm absorption band as this is more narrow than the 1450 nm band and appears slightly shifted in some spectra.
4 Conclusion

The results show that angle and distance are large contributors to the type of spectra acquired. The uneven surface of pulp and paper media do generated large differences due to the varying spectral reflections. A single spectrum may be misleading if used in analysis of moisture content.

Averaging many spectra however, show results in line with theory. The rapid process of acquiring a spectrum using NIR spectroscopy enables averaging of large data sets while at the same time keeping a rapid measuring process suitable for inline determination.

Preprocessing the data may create models with apparently lower fit to the calibration but displays improvement in the prediction of unknown samples. The preprocessing method that performed the best in this work was the MSC preprocessing tool.

When determining the moisture content of a sample both by visual inspection and by multivariate calibration the wavelength regions around 1450 nm and 1940 nm is of use since differences correlated to the moisture content is prominent. The 1450 nm absorption band is very wide and will also affect a large range of wavelengths. The 1940 nm absorption band is slightly shifted to shorter wavelengths in some experiments and this shift must be further studied at especially elevated temperature since the spectra acquired in temperature measurements shows large variations in this wavelength region.

Future studies could involve on-site measurement to see how spectra would appear in a real application. The possibility of steam disturbing the measurements must be considered and hence separating the absorption of water and steam should be further studied.

Keeping both temperature and moisture at a constant is needed for more exact calibration and constant temperature and moisture environments need to be created to facilitate these experiments.

A larger illuminated area could generate more averaged data already in the data acquisition process and experiments investigating this possibility could be undertaken.

4.1 Future studies

To acquire data useful for evaluating the technique field test in a paper mill is important. Since the measuring environment is difficult to simulate. As seen in this study the variables such as distance and angle and surface roughness are distorting the results. In a pipe the environment is not controlled but at least the impact of distance and angles can be reduced into a more narrow range within witch the light reaching the detector is more uniform.
Acknowledgements
I wish to thank all who assisted me in performing the experiments, analyzing the result and giving feedback on the written work. This foremost includes my supervisors Helena Håkansson (KaU) and Niclas Andersson (BTG instruments), who’s patience must have been tested many times. But also many other people at BTG who helped me in a number of different ways should not be forgotten. This includes Bengt Skålen (BTG) for fruitful discussions about sample preparation and experimental setup overall by sharing experiences, Johnny Weng (BTG) for help with experimental equipment eg fiber optics, Hans Pettersson (BTG) for programming the PLC and Dennis Groning for help with connecting the scale to the computer enabling logging of the weights. This is to name but a few.

Lastly I want to thank my family and friends for support during the time this work was performed, this is aimed especially to my girlfriend Annica Wingqvist who encouraged me and assisted in scrutinizing of the final report.
6 References


