Evaluation of Novel Quality Control Methods for Wheat Flour

DEPARTMENT OF FOOD TECHNOLOGY, ENGINEERING AND NUTRITION | LUNDS UNIVERSITY

STEPHANIE LAURENT-HEDLUND & JOSEFINE DAHLMAN

KLGM10 MASTER'S THESIS IN FOOD TECHNOLOGY | SPRING 2021

Master's Thesis in Food Technology

Evaluation of Novel Quality Control Methods for Wheat Flour

by

Stephanie Laurent-Hedlund & Josefine Dahlman

Supervisor Lund University: Stephen Burleigh
Supervisor Lantmännen Cerealia: Louise Selga
Examiner: Lars Nilsson

Abstract

The quality of a flour can vary greatly depending on the wheat variety and on the conditions while growing the wheat. Currently, the quality is mainly assessed by test baking the flour to observe the volume of the final product. The aim of this study was to investigate if three instruments (Alveolab, Mixolab and Rheo F4) analyzing rheological behaviours connected to dough preparation and baking properties, can provide additional valuable information to current flour quality controls. To achieve this, Mixolab and Rheo F4 analyses were conducted on a large set of flour samples, and the results were processed with the multivariate analysis methods PCA and PLS. Ultimately, predictive models for baking volume were created and evaluated. Furthermore, different preparation methods of doughs were tested in Mixolab and Alveolab to investigate how well these methods translate to the industrial bakery process. The results from the study indicate that the examined instruments alone are not able to give an accurate prediction of the baking volume. However, they seemed to provide a minor positive impact when creating a model where other quality control parameters and parameters connected to the chemical composition were included. This model obtained a predictive power, Q^2, of 0.834, suggesting that additional information is needed in order to create a predictive model that is valid for industrial use. It was further seen that the standard protocols in Mixolab and Alveolab provide results vastly different from those obtained when preparing the analyzed doughs according to industry standards. To conclude, the methods do not seem to give enough additional valuable insights on the final quality to justify the extra cost and time it would take to include it in quality controls.

Popular Science Abstract

The continuous search of the perfect quality control for wheat flours

Flour, water and yeast - how complex can it be? Although bread is one of the oldest and most basic staple foods in the world, we have actually not yet been able to fully understand the underlying mysteries of bread making. The main components of all wheat flour are protein, carbohydrates and water but also small amounts of fat and minerals. What makes the flours different from each other is the ratios between these components, which in turn is strongly dependent on how the conditions were while the wheat was grown. These factors all contribute to the complexity of wheat flours.

For a company working with a large scale production of flour, such as Lantmännen Cerealia, it is of utmost importance to know how well a flour will bake. Consequently, several quality controls are conducted on the flour to give indications of how it will behave, and ultimately the flour is test baked and the final bread volume is examined. Many of these control methods can be quite time demanding and may not always give an accurate prediction of the final bread volume. This leaves room for improvement of current quality controls and an optimal solution would be to find a way to predict the baking volume. Therefore, this project aimed to evaluate the contribution of the three instruments Mixolab, Rheo F4 and Alveolab from Chopin Technologies could have in predicting baking volume. These equipment are designed to provide information regarding dough behaviour during mixing and proofing, starch and protein quality as well as rheological parameters such as tenacity, extensibility and elasticity.

Similar studies have been conducted in other countries, but what makes this one special is that it contains a large sample set and a large number of different analyses conducted on the flours. Mixolab and Rheo F4 were run with a set of samples and the results were analyzed by multivariate analysis. It turned out that Rheo F4 and Mixolab had a minor positive effect on prediction of final bread volume. Alveolab and Mixolab were further used to examine how well the methods correlate to the reality in industrial bakeries. The results imply that there are some differences that could be interesting to further investigate. Overall, the investigated methods show potential of providing valuable information on dough behaviour. Although they may not add enough new insights in order to be incorporated into routine quality controls.

Individual Contribution

Both students have been involved in performing the experimental work, analyzing the results and writing all parts of the report. Additionally, both have attended weekly meetings with the supervisors from Lund University and Lantmännen Cerealia.

Acknowledgements

During this process, we have received valuable support from a special group of people. A big thank you to Louise Selga, Jan Poulsen and Thony Hedin from Lantmännen Cerealia for your continuous involvement and for guiding us through this process. Also, thank you to our supervisor Stephen Burleigh at Lund university for providing us with valuable insights and a deeper understanding in multivariate analysis. We are grateful to Chopin Technologies, to Georges Tawail and Hanna Zhyhunova, for letting us borrow the equipment and for helping us with technical guidance and support. Thank you to Lars Nilsson for your role as the examiner. Finally, thank you to all colleagues at the PUTS-Q department of Cerealia in Malmö for giving us insights into your work and for all the delicious bread and cookies.

Table of Contents

1. Introduction

A short background and the aim of the report is provided to get a better understanding of the project.

1.1 Project background

Lantmännen has its base in an agriculture cooperative and is one of the leading actors in northern Europe within several areas, such as agriculture and foods. The operations of the company mainly include growing and refining cereals. Further, the company has two different divisions within the food sector, namely Lantmännen Cerealia and Lantmännen Unibake. In both these divisions, flour is of great importance. Industrial bakeries rely on knowledge about flour behaviour in order to ensure products of consistent quality. However, the mills receive grains of varying quality depending on factors such as when the wheat was harvested and on the wheat variety. The main current method for assessing the flour quality is by test baking followed by measuring the bread volume. This method is however both time consuming and costly which leaves a demand for a new and better quality control method to be developed. As a complement to test baking, various analyses of the composition and behaviour are currently a part of the quality control, giving an indication of specific properties related to flour quality.

This master's thesis is a part of a PhD project, a collaboration between Lantmännen and Swedish University of Agricultural Science (SLU), that aims to create a predictive method based on correlations between flour composition and baking properties.

1.2 Aim

This project aims to investigate if the three instruments Alveolab, Mixolab and Rheo F4 from Chopin Technologies (Villeneuve-la-Garenne, France) can provide additional valuable information to current flour quality controls. This includes finding correlations between the chemical composition of wheat flour samples, baking properties and rheological parameters provided by the Chopin instruments.

To achieve this goal, the first part of the project was dedicated to doing rheological experiments. Thereafter, the results were analysed with multivariate analysis against previously obtained data on the chemical composition and important quality control parameters. The obtained results were further studied to find which components influence the flour quality, and ultimately see if it was possible to create a valid predictive model for baking volume. The results from the analyses were combined with a literature study into the following master’s thesis report.

2. Background

To understand the complexity of baking with wheat flours, it is important to have some knowledge regarding its chemical and rheological behaviour.

2.1 Wheat flour

Wheat is the most cultivated cereal both in Sweden and worldwide. With a great variety of end uses it has become a staple ingredient in many cuisines. Winter- and spring wheat are two classes of the cereal that are, as the names imply, sown during different seasons, resulting in differences in composition. Winter wheat is the most commonly grown variety in Sweden and it accounted for 89% of the acreage used for wheat production in 2020 (Swedish Board of Agriculture, 2020). Spring wheat has a higher protein content than winter wheat, making it suitable for bread making, which requires strong gluten networks. The lower protein content in winter wheat makes it more suitable for products with a crumbly texture, such as cookies and cakes. Commercial wheat flour generally consists of 63-72% starch (with an approximate amylose to amylopectin ratio of 25:75), 7-15% protein, 14% moisture and 2-3% lipids (Finnie et al, 2016, 31).

2.2 Dough development and baking

During the bread making process, several crucial phenomena take place. When flour and water are mixed, a gluten network starts to form by sulfhydryl groups establishing intramolecular disulfide bonds and cross-links the peptides (Sahi, 2014). While mixing, air bubbles are folded into the dough and get trapped within the gluten network resulting in a foam structure. Further, a starch-protein matrix is formed where the starch granules are embedded into the gluten network. Added yeast ferments sugars in the flour to carbon dioxide, which inflates the previously formed bubbles and causes an expansion of the dough. Ultimately, when the bubbles expand further, they eventually ruptures and forms an interconnecting network, giving rise to a sponge-like structure. During the baking stage, there is a net movement of water from the hydrated gluten to the starch which gelatinizes and settles the structure. (Rosentrater & Evers, 2018) (Mills et al., 2003)

2.2.1 Starch

There are several important quality parameters for wheat flour that are linked to starch. These parameters quantify and describe the amount of damaged starch, alpha-amylase activity and starch gelatinization and retrogradation properties. The functionality of starch in food systems mainly derives from its ability to bind large amounts of water upon gelatinization. The temperature at which gelatinization occurs can be affected by the presence of solutes, such as salts. Dilute salt solutions have been found to raise the gelatinization temperature (Nicol et al., 2019).

Starch granules that have been physically damaged and altered from their native form, for example due to shear forces and pressure during milling, are referred to as damaged starch. Damaged starch is more susceptible to enzymatic hydrolysis by amylase and absorbs more water than native starch granules. Therefore, the amount of damaged starch will strongly influence the final product and rheological properties of the dough, which demonstrates the importance of having it as a quality parameter for flour (Arya et al., 2015).

Amylases are hydrolytic enzymes found in flour that target the bonds between glucose units in starch. The hydrolysis results in smaller fragments in the form of dextrins, and fermentable sugars such as maltose, which can act as an additional substrate source for yeast that ferments it into leavening gases. Excessive activity may however result in sticky and dense breads. Due to aforementioned consequences, amylase activity in flour heavily influences the baking properties and often needs to be regulated by addition of malt. Native starch granules are hydrolyzed slowly, while damaged starch and gelatinized starch are more available for the enzyme, resulting in a greater extent of hydrolysis. Two common methods for obtaining information regarding amylase and starch behaviour is to conduct falling number and amylograph analyses. (Finnie et al, 2016, 42-43, 71-72)

2.2.2 Formation of a gluten network

Monomeric gliadins and polymeric glutenins pose a majority of the flour proteins, making up 80-85% of the total amount. As the wheat flour is combined with water, the proteins within the flour get hydrated and start to form a gluten network. Generally, the glutenin proteins are considered to build the network and affect the elasticity and cohesiveness of the dough. Further, the function of the gliadins is to lubricate the network and contribute to the extensibility and viscosity of the dough. (Ooms, N. & Delcour, J. A. 2019) The ratio of glutenin versus gliadin has been suggested to have additional effect on the viscoelastic properties where a higher ratio of glutenin results in a stronger dough (Shewry et al., 2003). A common improver added to form strong doughs is ascorbic acid, a vitamin widely used in the baking industry. The compound facilitates the formation of SS-bonds between the proteins and consequently give rise to a more stable network (Sahi, 2014).

There are several useful methods used in flour quality controls that provide parameters related to the quantity and functional properties of gluten in doughs. Some examples of these are gluten index, wet gluten and various farinograph parameters.

2.2.3 Fermentation and gas retention

One of the most contributing factors to a bread's volume and texture, and thereby the overall quality, is the size and number of gas cells. Air is incorporated into the dough upon mixing in the form of gas cell nuclei, leaving a foam structure where gas cells are dispersed in the starch-protein matrix. Energy used during mixing and added pressure, together with the viscosity of the dough will determine the number and size of the bubbles. Generally, many and small bubbles are desirable (Mills et al., 2003). As the yeast ferments, carbon dioxide is produced in the dough's aqueous phase and diffuses to the formed gas cell nuclei. The carbon dioxide then causes an expansion of the cells, and ultimately of the dough. In the early stages of the process, the loss of gas is usually slow since the continuing production of carbon dioxide by the yeast saturates the aqueous phase, preventing it from disappearing from the gas cells. Nevertheless, there is some gas loss which is thought to be attributed to the diffusion of gas to the external surface of the dough where it evaporates into the surroundings (Gan et al., 1995).

As in most foams, the gas bubbles experience coalescence, and potentially disproportionation, while expanding. These phenomena impact the stability of the foam, but it is not yet fully understood to what extent. The stability is influenced by the composition and surface properties of the lamella, connecting two bubbles, including proteins, lipids and other compounds soluble in dilute salt solutions. Proteins of small size diffuse to the surface where they create a network of low molecular weight surfactants (Gan et al., 1995).

Arabinoxylan (AX), a non-starch polysaccharide found in the cell wall in cereals, has also been shown to affect the stability of the foam. It has mainly been found that AX stabilizes the protein films in the foam due to an increased viscosity of the dough's aqueous phase, but also by mediating interactions between the proteins in the adsorbed layer. Arabinoxylans may be categorized into WE-AX (water extractable) and WU-AX (water unextractable), which have different impacts on the dough. WE-AX is thought to stabilize the foam by increasing the viscosity of the dough's aqueous phase and may thus result in higher bread volumes. On the other hand, WU-AX has been shown to have a negative effect on loaf volumes. This could potentially be attributed to its disturbing effect on the gluten network formation caused by the steric hindrance of the large WU-AX molecules and its ability to bind large amounts of water, leading to less water available for gluten formation. Moreover, WU-AX is thought to reduce the gas retention capacity of the dough by perforating the gas cells which promotes coalescence (Courtin & Delcour, 2002).

2.3 Theory behind the analysis methods

During the past century, several methods have been developed and refined to measure different rheological properties in doughs. The three main instruments investigated in this project are Rheo F4, Mixolab and Alveolab.

2.3.1 Mixolab

Mixolab from Chopin Technologies was developed based on the combined principles of the farinograph and viscoamylograph. It is used to analyze the quality of starch and proteins in flours and provides measurements of dough behavior during mixing. This is done by characterizing the rheological behaviour of dough subjected to a heating and cooling cycle alongside kneading by dual mixing blades. The dough's resistance to the mixing action of the blades gives rise to a torque, which is measured in real time and plotted together with the temperature versus time. A typical curve from a Mixolab test is shown in *Figure 1*. The obtained curve is divided into five sections where each of the different sections provides information on the measured parameters.

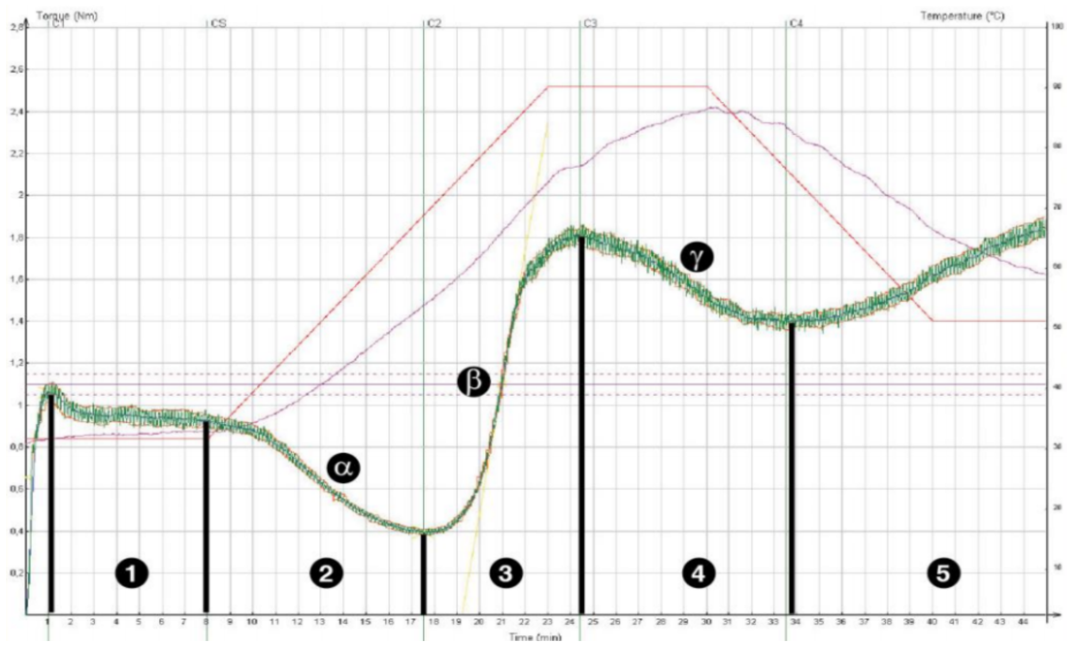

Figure 1. Example of a Mixolab curve with interpretive values illustrated. The left Y-axis displays the measured torque and the right Y-axis shows temperature. The red line represents the target temperature profile and the purple line visualizes the actual temperature during the measurement. Figure taken from Mixolab manual. (Chopin Technologies, 2016a).

The first section evaluates the dough behavior during mixing at constant temperature. This stage determines the water absorption capacity of the flour by measuring the amount of water that needs to be added in order to achieve a target consistency of 1.1 Nm+-0.05. Additionally, the parameter *C1* (torque at the target consistency), the dough development time and dough stability are obtained. The dough development time is the time it takes for the dough to reach the target consistency, while the dough stability is the time this torque is kept before decreasing. When the temperature program starts, the torque decreases and the curve enters the second section. This section is related to the quality of the gluten protein network and its ability to withstand heat and mechanical work.

As the temperature of the dough continues to rise further according to the temperature program, it eventually reaches the point where the starch starts to gelatinize. The swelling of the starch granules during gelatinization gives rise to a viscosity increase and thereby an increase in torque. This behaviour is shown in the third section of the curve and gives an indication of the starch quality and quantity. The fourth section is constituted by the degradation of starch, dependent on the amylase activity in the flour. Depending on the amount of amylase, the torque and viscosity will decrease with different intensities from the peak seen in the third section. In the fifth and last part of the curve, the dough cools down and increases in consistency again as a result of the re-association and crystallization of the starch, i.e. retrogradation (Chopin Technologies, 2016a).

2.3.2 Rheo F4

Rheo F4 from Chopin Technologies analyzes the proofing properties of doughs under set conditions. The instrument measures the volume of a dough during dough development and CO_2-production and release. It also gives an indication of the porosity of the dough and proofing tolerance over time. The results from the Rheo F4 analysis are shown as three curves in two graphs (*Figure 2a* and *Figure 2b*), with the first graph showing the dough development and the second showing both the amount of gas produced and released. From these curves, several key values are extracted and used to interpret the results. An explanation of the different parameters gathered from the analysis can be seen in *Table 1*.

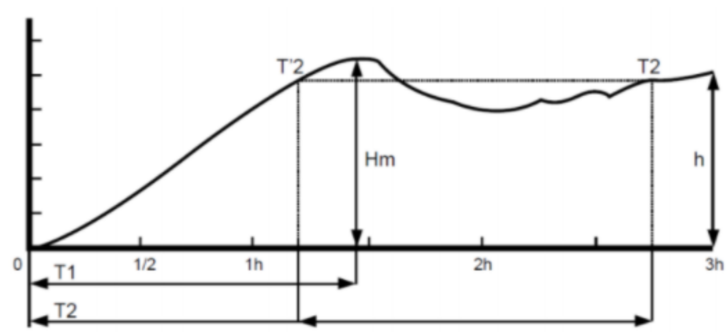

Figure 2a. Example of a dough development curve with associated parameters illustrated. Figure taken from Rheo F4 manual (Chopin Technologies, 2013).

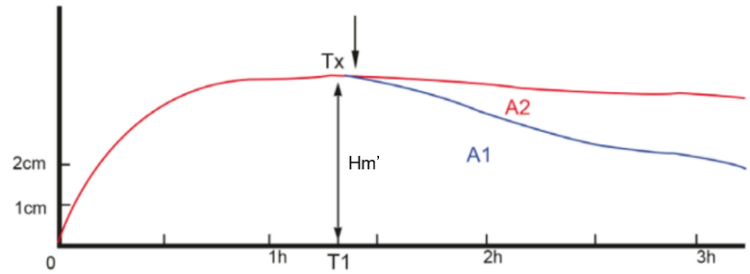

Figure 2b. Example of a gas production/gas release curve with associated parameters illustrated.The red line demonstrates the amount of produced CO_2 while the blue line shows the release of CO_2. Figure taken from Rheo F4 manual (Chopin Technologies, 2013).

Table 1. Explanation of parameters connected to Rheo F4 measurements.

Parameter	Explanation
Hm	Maximum height during the dough development
h	The dough development height at the end of the test
(Hm-h)/Hm	Describes the drop in development height from its maximum (Hm) to the height after 3 hours (h) in percentage.
T1	The time where the dough height is at its maximum
Hm'	Maximum height of the gas production curve
T'	The time when the gas production curve reaches its maximum
Tx	The time it takes until the dough starts releasing CO_2 (dough porosity time)
A1	The volume of gas retained in the dough
A2	The amount of gas released during fermentation

A1 will further be referred to as Retention volume (RetVol) and A2 as released amount of CO_2 (*VolCO2*). By adding these two together, a total volume of produced gas (*TotVol*) is obtained. Furthermore, by dividing the volume of gas retained within the dough with the total volume, the retention coefficient (*RetCoeff*) can be determined. A high retention coefficient is desirable and indicates that the flour has its origin in healthy grains. Higher retention coefficients can usually be obtained if *Tx* appears later (Chopin Technologies, 2013).

2.3.3 Alveolab

Alveolab is an alveograph from Chopin Technologies used to investigate rheological characteristics such as tenacity, elasticity, baking strength and extensibility of the dough. The instrument creates a dough, based on a constant or adapted water addition protocol, and measures the deformation of a dough bubble while inflating gas.

The results are recorded in an alveogram where the pressure is plotted versus length of the bubble. A few parameters of interest gathered are *P*, *L*, *G*, *Ie*, *P/L* and *W*. *P* represents the pressure the dough resists before it is inflated and can be interpreted as the tenacity of the dough and its ability to resist deformation. *L* is the abscissa obtained when the dough ruptures and gives an indication of the dough extensibility and the maximum volume of air the bubble can contain. Further, *G* and *Ie* are the extensibility and elasticity index, respectively. *W* represents the dough baking strength and is obtained by calculating the area under the curve (Chopin Technologies, 2016b)

2.4 Multivariate analysis

Multivariate analysis is a hypernym for techniques used to analyze data containing more than one variable. The methods are designed to explore interrelationships between several variables and elucidate features of the dataset.

Principal Component Analysis, PCA, is a method used for reducing the dimensionality of large datasets while minimizing the loss of information. This makes it possible to further analyze and visualize the data, allowing for easier detection of patterns and correlations between variables. The drawbacks of dimensionality reduction by PCA is that it comes at the expense of accuracy (some information always gets lost) and that the new variables, the PCs, can be harder to interpret.

The results from a PCA are often visualized in the form of score plots, loading plots and biplots. The score plot shows how the samples are distributed in relation to each other and their new dimensions. Clustering of samples in this kind of plot can be an important key to trace back what it is that influences the differences among clusters. A loading plot shows the relative positions of the initial variables in the new setup. This demonstrates how much, and in what way, the principal components are influenced by each variable. Variables placed close to the PC axes are less influential than those further out on the axes. Samples are considered to be positively correlated with each other if they are placed close together in the plot, while samples forming large angles around 180° to each other are negatively correlated. Biplots gather the information from both the score and loading plot by superimposing them onto one another. Since these are displayed at the same time, it provides further insight into the relationship between observations and variables. (Jaadi, 2021)(Jolliffe, 2002)

The PCA may be combined with a Partial Least Square (PLS) regression, a method widely used to examine the relationship between two different data matrices. PLS regression can be used to create models where explanatory variables from the X-matrix are used to predict response variables in the Y-matrix. The method creates a linear multivariate model that projects both X and Y data in a new space. Consequently, it can analyze large and noisy X- and Y matrices (Wold et al., 2001). PLS modelling usually includes randomly dividing the data into a larger training set and a smaller sample set. The training set is then used to create the predictive model which is tested on the remaining data in the sample set, providing an unbiased evaluation of the model. The performance of the model can be described by parameters such as R^2 and Q^2 which often are termed "goodness of fit" and "goodness of prediction", respectively (Eriksson et al., 2013).

3. Materials and Methods

All of the materials and equipment used during this project were provided by Lantmännen Cerealia and Chopin Technologies. Several parameters connected to the flour samples had been determined previously to the project. These are presented in *Table A* in *Appendix 1*. The table further states by whom the data have been gathered, the range of values and how they will be referred to later in the report.

3.1 Flours

The subjects investigated were 207 samples of wheat flour. The wheat was harvested during 2018-2019 and refined between August 2018 - August 2020 in Lantmännen's mills in Malmö and Strängnäs. Furthermore, the flour samples are categorized within five different groups; Spring 1, Spring 2, Winter 1, Winter 2 and Blend.

3.2 Rheological analysis

Three different analysis methods were used to evaluate the given set of flours. The methods required a predetermined moisture content of the flour samples, which was measured with Near Infrared Transmittance (NIT) technology.

All 207 samples were examined with Mixolab and later six of these were further analyzed as adjusted flour and optimally worked dough. Due to initial problems with the Rheo F4 instrument, data could only be obtained for 165 samples. Alveolab data was collected on six samples of adjusted flour and optimally worked dough.

3.2.1 Method reproducibility

Before starting the experiments, a duplicate study on the three instruments was conducted to ensure the reproducibility of the methods. For this study, five flour samples were analyzed in duplicates by two different operators on each of the instruments. The Rheo F4 and Mixolab results were then compared to limit of reproducibility values provided by Chopin Technologies.

3.2.2 Mixolab

The Mixolab analyses were conducted according to the predefined Chopin+ protocol. For the analyses, 75 g dough was prepared with a given flour to water ratio dependent on the moisture content and water absorption capacity of the specific flour. In order for the test to be deemed acceptable, a target consistency of 1.1 +/- 0.05 Nm had to be reached within the first 8 minutes. If the dough did not reach the target, the test was restarted with another water absorption value as suggested by the software. The resulting curves were visually inspected to see if the torque maximums and minimums were placed accurately. If not, the markers were manually corrected.

3.2.3 Rheo F4

Preparation of the dough was done according to the Chopin protocol. Initially, an adjusted amount of water with dissolved fresh yeast (7g) was added to a constant amount of flour (250g). The amount of water to add was calculated based on moisture content and a previously measured P-value (from Alveolab CH). To knead the dough, the mixing chamber in Alveolab was used. Further, the equipment was run with the standard protocol for 3 hours at 28.5℃.

3.2.4 Analyses with adjusted flour: Alveolab and Mixolab

Six flour samples were adjusted with ascorbic acid, salt and malt to mimic the flour mixes used when creating doughs in industrial bakeries. Some of the adjusted flour was used to conduct Alveolab and Mixolab analyses directly. The Alveolab analyses for these samples were done according to the Alveolab Constant Hydration protocol with adjusted flour and water amounts. The quantity of flour was always 250 g and the amount of water was calculated according to previously determined farinograph water absorption values. Distilled water was used for the analysis. The rest of the adjusted flour was further used to create an optimally kneaded dough, using a spiral mixer from Kemper, for additional Mixolab and Alveolab analyses. After mixing, the optimally kneaded dough was examined with the Alveolab bubble blowing procedure and with the Mixolab Chopin+ protocol.

3.3 Data presentation and multivariate analysis

To analyze the datasets, Matlab and Simca were used. All data was gathered in several X and Y matrices, where the composition was dependent on what parameters that were of interest while assessing the results.

3.3.1 Matlab

All data used in the multivariate analysis was pre-processed in Matlab. Pre-processing was done by first using the `ismissing` function, where the script could detect if there was any data missing in the matrices. If a value was detected as missing, it was replaced with an average value for that specific parameter. If several values were missing for the same sample, it was removed from the dataset. This was followed by a Grubbs test to detect outliers within the different columns. If all parameters in a sample were detected as outliers, it was removed from the dataset. If only single parameters were pointed out as outliers, these were replaced with an average for that specific parameter. Matlab was further used as a confirmation method to ensure that the results given in Simca were somewhat close to the results that Matlab yielded.

3.3.2 Simca: PCA, PLS-regression and plots

Simca was used to produce PCA plots and PLS models. All PLS models were made with baking volume as a single Y-variable to be predicted. The PLS models were made by randomly selecting 75% of the samples as a training set for the model, while the remaining 25% was used as a sample set. The randomization was made by organizing the samples according to values generated by Excel. Further, the number of principal components used in the model was chosen based on where the highest predictive power, Q^2, was achieved. These models were then improved by removing variables that did not significantly influence baking volume in the original PLS model. Additionally, the program was used to produce plots to see if there seemed to be any correlations between interesting variables.

4. Results and Discussion

All data gathered during this project, and the previously obtained data, is shown in *Table A, Appendix 1*. The table shows the range of data and the abbreviations used further throughout the statistical analysis and modelling.

4.1 Method reproducibility

The duplicate study for the Mixolab and Rheo F4 analyses showed that the results were reproducible to 100% for all parameters except *Temp C3*, which gave 80% reproducibility. All parameters assessed can be found in *Table B, Appendix 2*.

4.2 Mixolab

The Mixolab data was further analyzed in order to investigate the relationship between the obtained parameters and previously gathered quality data that was deemed to be relevant.

Figure 3 illustrates a PCA biplot of the Mixolab dataset. The first two principal components explain 48% of the dataset. Further, the samples are mainly spread along the PC 1 axis which is connected to the parameters *C3*, *DiffC23*, *Doughdev* and *Stability*. PC 1 seems to separate the samples with Spring 1 to the right side and the rest of the flours further left. As mentioned earlier, Spring 1 is theoretically classified as a flour of higher quality. It is therefore possible that high values of *Doughdev*, *Stability* and *Cs* might be indicators of better wheat quality, while high *C3*, *DiffC23* or *DiffC45* values might indicate lower quality.

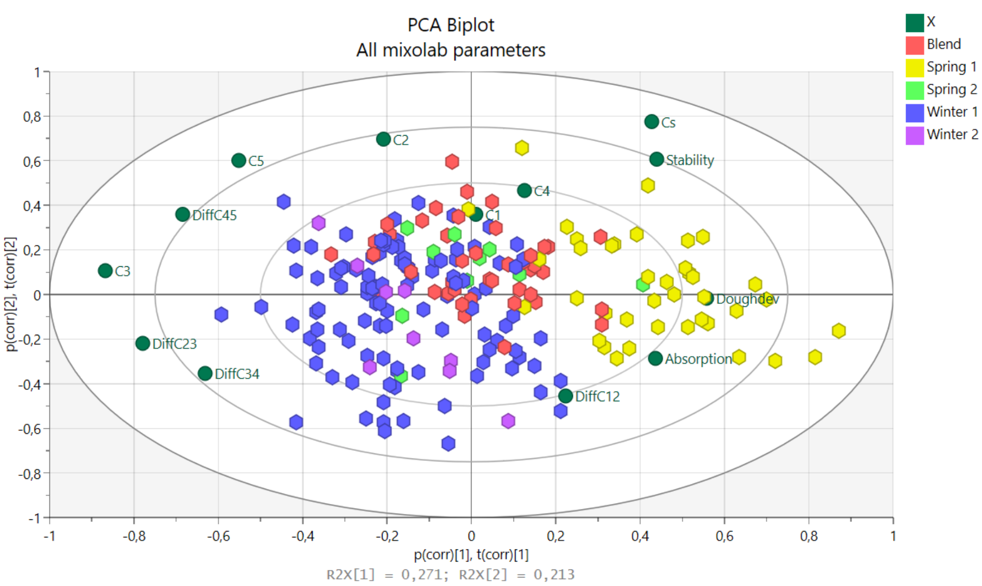

Figure 3. Biplot for a PCA model constituted by all gathered Mixolab data. Mixolab parameters are denoted as X in the figure legend and colored green.

Figure 4 depicts a PCA loading plot for the Mixolab dataset and selected relevant quality control parameters. The first two principal components only make up for 45% of the total variance in the dataset, which means that a considerable amount of information gets lost when reducing the dimensionality to two new variables. Most of the parameters related to protein quality and protein content are somewhat clustered with *BakingVol*. The placement of these parameters along the positive end of PC 1 indicates that PC 1 is mainly described by various protein properties. On the other side, PC 2 seems to be more related to the incorporation of water and starch properties.

An interesting topic to examine is how the various Mixolab parameters are placed in relation to other quality parameters that are supposed to measure similar properties. For example, as seen in *Figure 4*, the water absorption parameter obtained from Mixolab (*Absorption*) is very closely correlated with the farinograph (*FWaterAbs14*), suggesting that they may be interchangeable. Furthermore, *FStability* from the farinograph, and *Cs* and *Stability* from Mixolab seem to be quite well correlated and the same goes for *FDevTime* and *Doughdev*. The retrogradation parameters *C5* and *RVAFinVisc* also correlate closely with each other. These connections indicate that there might be a possibility of substituting some of the current quality control analyses with Mixolab and getting equivalent results. However, since the PCA plot only captures 45% of the variance, it is very hard to draw any definitive conclusions. It would perhaps be possible to explore this more using a correlations matrix in further studies. Mixolab parameters that seem to correlate the most with *BakingVol* are *Cs, Stability*, *Doughdev*, (positive correlations) and *Diff C34* and *Diff C23* (negative correlations).

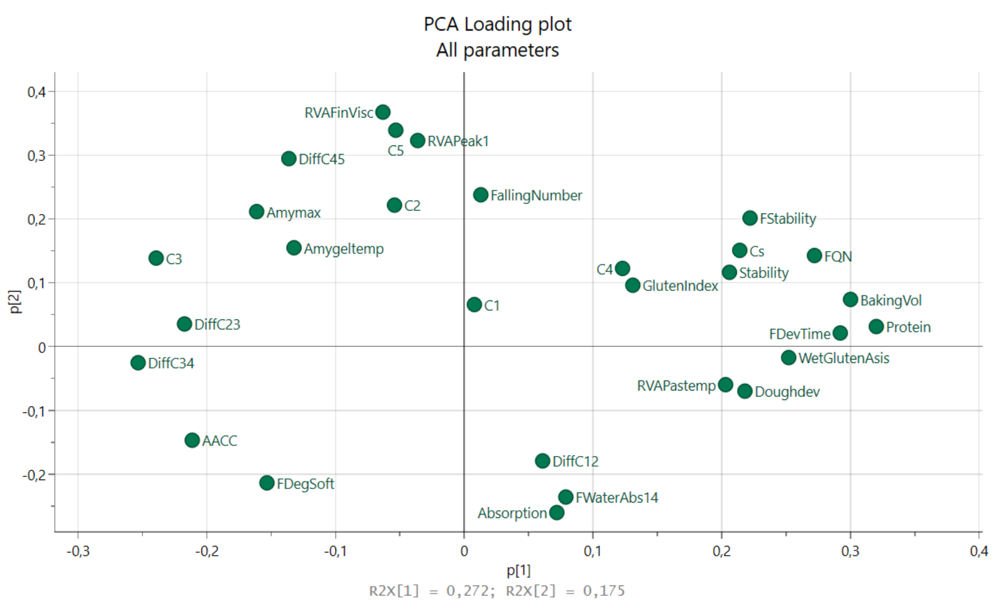

Figure 4. PCA loading plot for all Mixolab parameters and selected quality control parameters.

4.3 Rheo F4

During the pre-processing step it was observed that some samples missed fiber content data and these were removed from the dataset. After pre-processing, 151 samples remained.

A PCA was performed with all of the gathered Rheo F4 parameters and the resulting biplot is presented in *Figure 5*. Together, PC1 and PC2 make up for 61,7% of the variance in the dataset. Some observations made are that the height of the gas release curve (*Hm'*) and total volume (*TotVol*) cluster closely together to the far right on PC 1. Retention volume *(RetVol*) and volume of released carbon dioxide (*VolCO2*) also places quite close to these two parameters with regards to PC 1. The figure thus implies that these parameters are positively correlated. On the other hand, the retention coefficient places far to the left which contradicts the theory. The dough development parameters, *Hm* and *h*, cluster close to each other in the first quadrant. An explanation to this appearance could be that many of the tests had their peak close to the end of the test and it therefore falls naturally that the two indicators are close to each other.

The various flour types separate some along the PC 1 axis, with Winter 1 and Blend mostly in the middle and to the left. All Spring 1 samples are placed on the right side of the axis while Spring 2 is quite spread between the different groups. As expected, the plot shows that Spring 1 seems to be more positively correlated with parameters *Hm*, *h*, *RetVol*, *Hm'* and *TotVol* than Winter 1 samples. A potential trend that may be noted is that Spring 1 tends to place further up on the PC 2 while the larger mass of Winter 1 samples places further down the axis. It is however not possible to draw any similar conclusions for Blend and Spring 2.

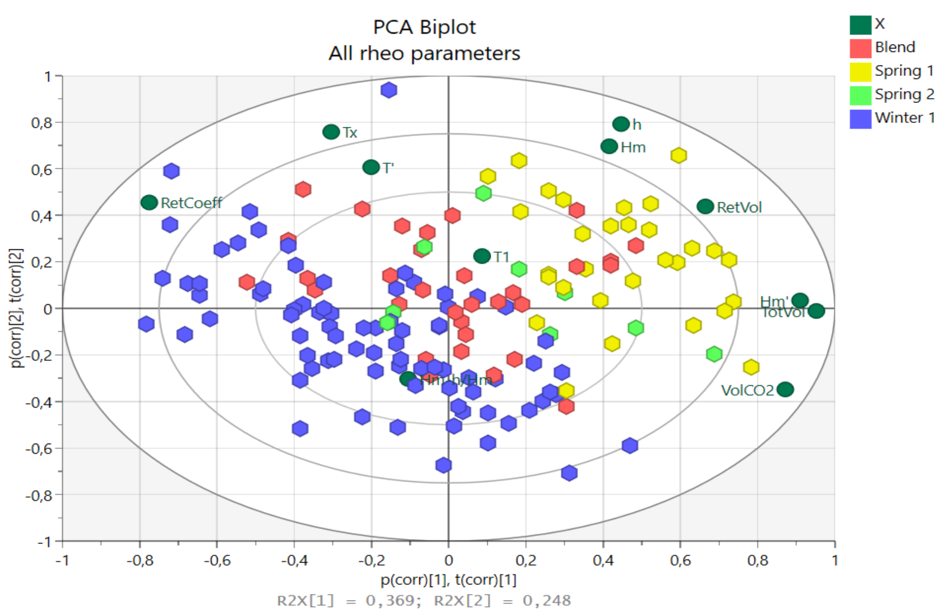

Figure 5. PCA biplot for all Rheo F4 data. Rheo F4 parameters are denoted X in the legend and are colored green.

Figure 6 shows a PCA loading plot of all gathered Rheo F4 parameters, together with previously collected data that could be of interest. The figure shows that the Rheo F4 parameters *h*, *Hm* and *RetVol*, are primarily correlated positively to *Protein*, *L*, *W*, *Ie* and water extractable arabinoxylan (*WEAX*). Furthermore, the same Rheo F4 parameters are placed close to *BakingVol*, which was expected since a greater dough development would give rise to larger bread volumes.

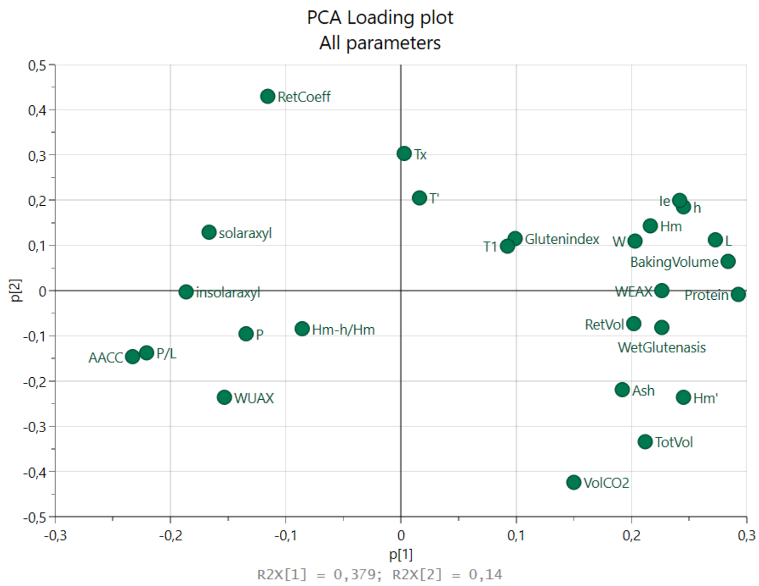

Figure 6. PCA loading plot for all Rheo F4 parameters and selected quality control parameters.

The relationship between baking volume and maximum fermentation height (*Hm*) was further investigated with a scatter plot, shown in *Figure 7*, where the flours are colored according to product type. The plot shows no indication of a strong correlation between the two variables. Although, it displays two clusters, where Winter 1 forms a cluster of lower *Hm* values and baking volumes in the left corner. The three remaining groups of flour, Blend, Spring 1 and Spring 2, are more spread along the axes. Therefore, *Hm* seems to have a different impact on the various types of flour.

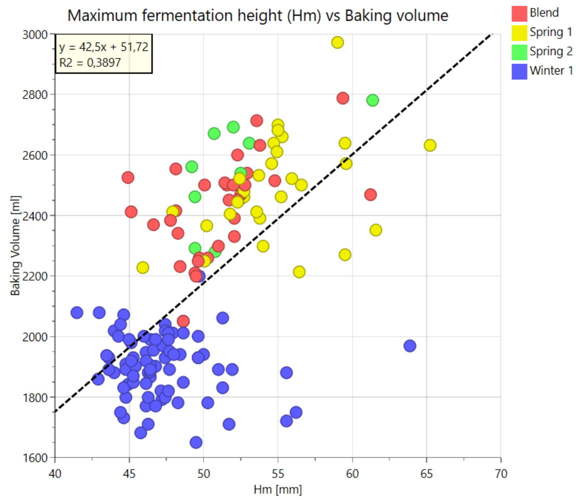

Figure 7. Maximum height during fermentation (Hm) plotted versus baking volume where the samples are colored according to product type.

The placement of damaged starch (*AACC*) in the loading plot could indicate a negative correlation with the fermentation parameter indicating the total volume of gas produced (*TotVol*). According to the theoretical background, it is implied that higher amounts of damaged starch would make more substrate available for the yeast to ferment, and therefore result in a positive correlation. The plot in *Figure 8* investigates the relationship between the two variables further and does not indicate any correlation between the parameters. A possible explanation for the lack of correlation could be that all of the samples contained a substantial amount of damaged starch. The sample with the lowest amount still had 4.8 g/100g (see Appendix 1) which potentially could provide more additional substrate than the yeast can ferment during the three hours of analysis.

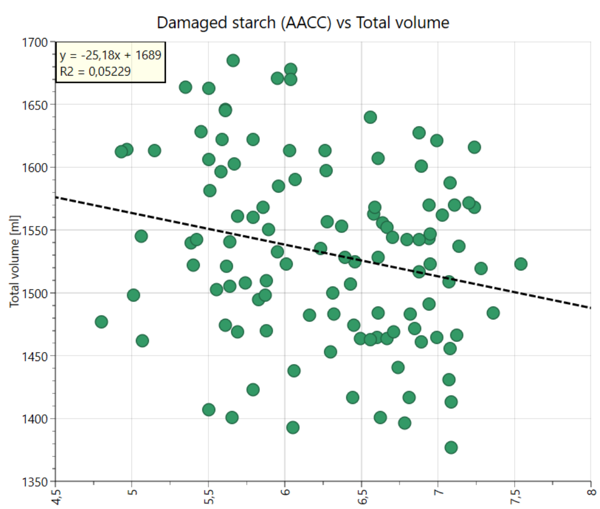

Figure 8. Scatter plot of Damaged starch versus Total volume.

The two types of arabinoxylan, *WE-AX* and *WU-AX*, should in theory be positively, respectively negatively, correlated with higher retention volumes due to their impacts on foam stability. The loading plot in *Figure 6* seems to somewhat confirm this hypothesis since *WE-AX* places close to retention volume and other parameters indicative of good baking quality, while *WU-AX* is placed on the opposite side of the plot. However, when plotting *WE-AX* and *WU-AX* against retention volume, as shown in *Figure 9*, no clear correlations can be observed.

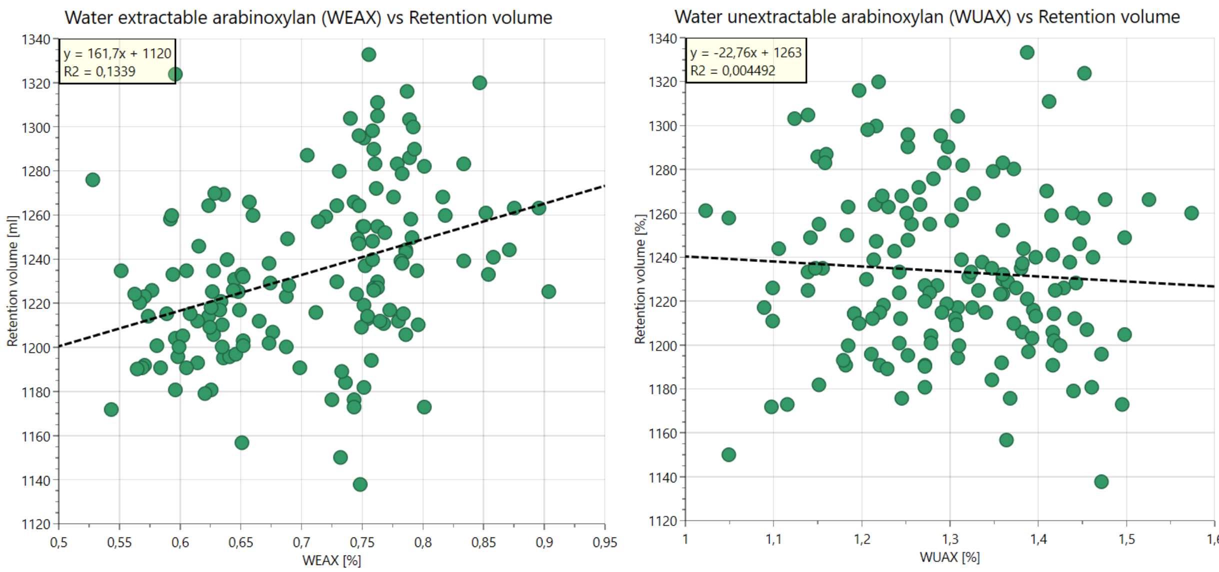

Figure 9. Scatter plots of retention volume versus water extractable arabinoxylan respectively water unextractable arabinoxylan.

4.4 Predicting baking volume with PLS regression modelling

In order to investigate the possibility of either replacing or complementing test baking with a predictive model based on other test methods, several PLS regression models were made. First, models were made using only Mixolab parameters and then with only Rheo F4 parameters. The last models were based on Mixolab, Rheo F4 and various quality control parameters. A summary of all PLS models can be found in *Table 2*.

Table 2. Summary of PLS regression models for predicting baking volume. TS stands for training set and SS for sample set. RMSEP is the Root Mean Square Error of Prediction.

Model	X variables	Y variables	PLS components	Included samples	R^2Y	Q^2	$R^2Y_{Predicted}$	RMSEP
Mixolab	*Absorption, C1, Cs Doughdev, C2, C3 DiffC12, C4, C5*	Baking volume	2	n_{TS}=154 n_{SS}=51	0.565	0.495	0.566	211

	DiffC23, DiffC34, DiffC45, Stability							
Mixolab Improved	*Doughdev, Cs, C4, DiffC34, Stability*	Baking volume	2	n_{TS}=154 n_{SS}=51	0.541	0.518	0.436	238
Rheo F4	*Hm, h, (Hm-h)/Hm, T1, Hm', T', Tx, TotVol, VolCO$_2$, RetVol, RetCoeff*	Baking volume	1	n_{TS}=113 n_{SS}=38	0.541	0.499	0.502	213
Rheo Improved	*Hm, h, Hm-h/Hm, Hm', TotVol, VolCO$_2$, RetVol, RetCoeff*	Baking volume	2	n_{TS}=113 n_{SS}=38	0.568	0.514	0.556	201
All combined	*All parameters (see Table A, Appendix 1)*	Baking volume	2	n_{TS}=113 n_{SS}=38	0.845	0.777	0.825	126
All combined Improved	*AACC, Protein, WetGlutenasis, WEAX, Insolaraxyl, Solaraxyl, L,G,Dmin, FDevTime, Amymax, RVABreakdown,Hm,h RVAPeaktime, Cs, Absorption,Doughdev*	Baking volume	1	n_{TS}=113 n_{SS}=38	0.842	0.834	0.811	130
Reference 1: All parameters except Mixolab and Rheo F4	*AACC, Protein, WetGlutenasis, WEAX, Insolaraxyl, Solaraxyl,RVAPeak1, RVABreakdown, RVAPeaktime, L, G, P/L,Dmin,FDevTime,*	Baking volume	3	n_{TS}=113 n_{SS}=38	0.857	0.828	0.775	142
Reference 2: All parameters except Mixolab	*AACC, Protein, WetGlutenasis, WEAX, Insolaraxyl, Solaraxyl, h, L, G, P/L, Dmin,FDevTime, RVAPeak1,RVAPeaktime, RVABreakdown*	Baking volume	2	n_{TS}=113 n_{SS}=38	0.859	0.834	0.785	138
Reference	*AACC, Protein,*	Baking	3	n_{TS}=113	0.862	0.827	0.762	146

3: All parameters except Rheo F4	*WetGlutenasis, WEAX, Insolaraxyl, Solaraxyl, FQN, FWaterAbs14, FDevTime, L, G, W, P/L, Ie, K, Dmin, RVABreakdown, RVAPeaktime, RVAPastemp, Cs, Absorption,Doughdev*	volume	n_{SS}=38

4.4.1 Mixolab parameters

A PLS model with all Mixolab parameters as X-variables and baking volume as the Y-variable, named *Mixolab*, can be seen as a biplot in *Figure 10*. This model had a Q^2 value of 0.495. As seen in the figure, some of the parameters seem to be more correlated with baking volume than others. For example, *Stability*, *Cs* and *Doughdev* seems to be quite closely positively correlated with *BakingVol* and *DiffC34* seems to correlate quite strongly negatively. On the other hand, parameters such as *C1*, *C2*, *DiffC12*, *Absorption*, *C3*, *DiffC23*, *DiffC45* and *C5* seem to be less correlated, placing either quite close to the middle or perpendicular to baking volume.

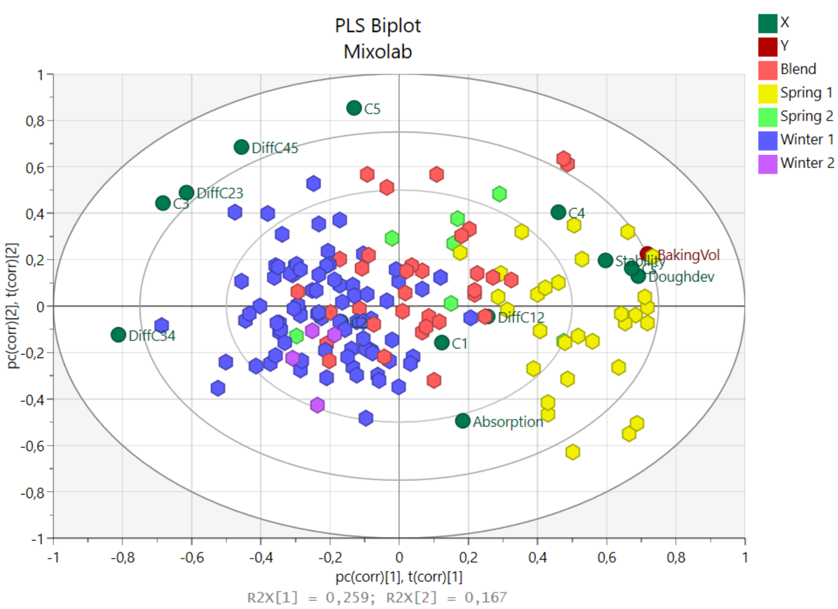

Figure 10. Biplot for a PLS model with Mixolab parameters as X and baking volume as Y.

When further investigating the influence of the different parameters on baking volume, it was confirmed that the above mentioned parameters did not have a significant influence. This is illustrated in *Figure 11* with effects plots from both Simca and Matlab. The effects plot from

Simca is based on the training set while the one from Matlab includes the complete set of samples. When comparing the two plots, it can be seen that the parameters with significant influence in the training set also are significant for the complete set. The training set can thus be interpreted to be representative for the full set, from that perspective.

Further, a decision was made to create a new model, named *Mixolab Improved*, containing only the significant variables from Simca and thus *C1*, *C2*, *DiffC12*, *Absorption*, *C3*, *DiffC23*, *DiffC45* and *C5* were removed. As a result, the predictive power (Q^2) improved to 0.518. However, it should be said that none of the models possess a good predictive power.

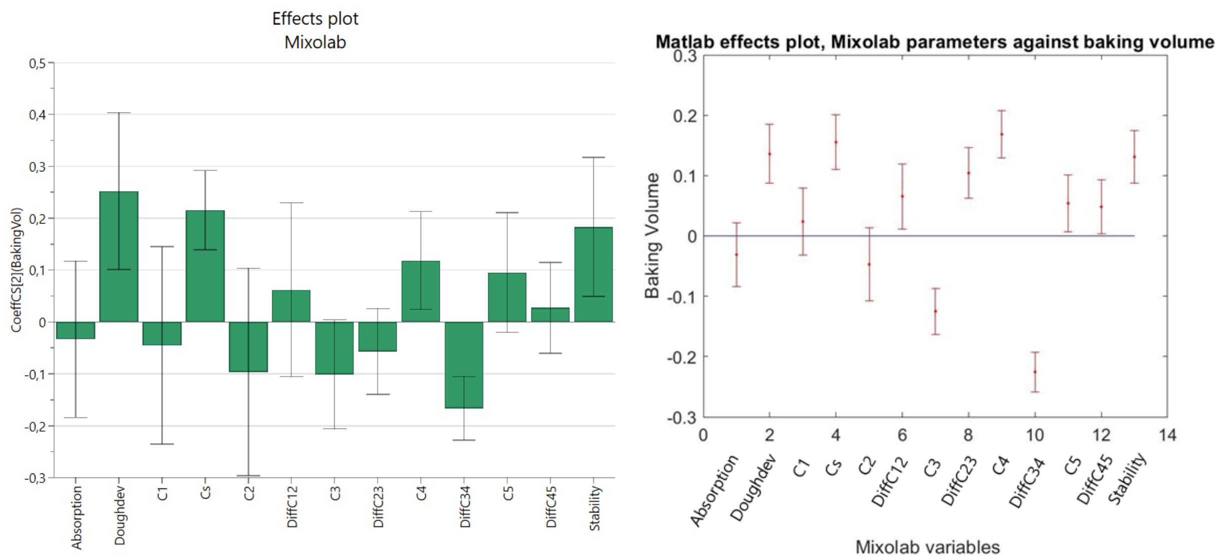

Figure 11. Effects plot for the model *Mixolab*, showing if the variables have a significant impact on baking volume. Left: An effects plot from Simca on the training set. Right: An effects plot from Matlab on the whole dataset.

The improved model was tested by predicting the baking volume of the sample set. As presented in *Figure 12*, the prediction resulted in a $R^2Y_{Predicted}$ value of 0.436 and a RMSEP (root mean squared error of prediction) value of 238 ml. These results indicate a poor fit of the model, which can be observed quite clearly in the figure. Furthermore, the fit worsened in comparison with the *Mixolab* model, where the $R^2Y_{Predicted}$ value was 0.566 and the RMSEP value was 211 ml. However, it must be noted that these results only indicate that the first model, *Mixolab*, fits this certain sample set better. The increased Q^2 value still suggests that the latter model should be more accurate in general.

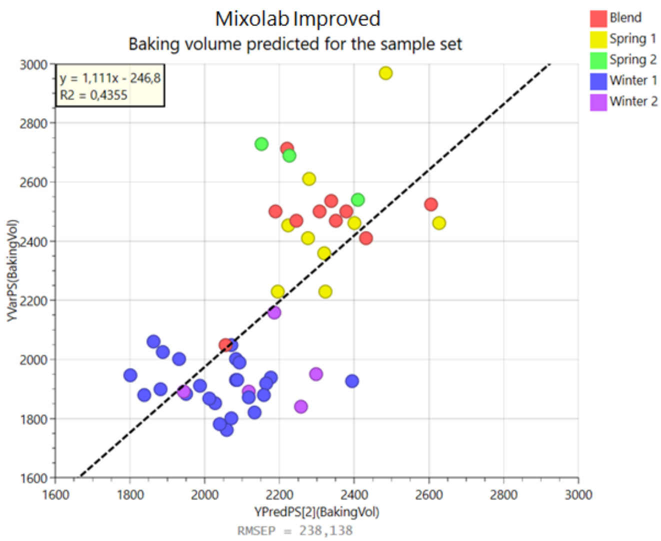

Figure 12. Predicted baking volume for the sample set using the *Mixolab Improved* model with the correlated linear equation, $R^2Y_{Predicted}$ and root mean squared error of prediction (RMSEP).

4.4.2 Rheo F4 parameters

A PLS regression model, *Rheo*, was constructed where all Rheo F4 variables were used as predictors for baking volume. A biplot for this model is presented in *Figure 13*. From the figure it can be observed that *T1*, *T'*, *Tx* and *(Hm-h)/h* seem to place fairly uncorrelated to baking volume since they all are quite centered in the plot. Furthermore, the product types display a similar relationship with baking volume as they do in the PLS biplot for the *Mixolab* model (*Figure 10*).The predictive power of the model was 0.499, $R^2Y_{Predicted}$ was 0.504 and a root mean squared error of prediction (RMSEP) of 213 ml was obtained.

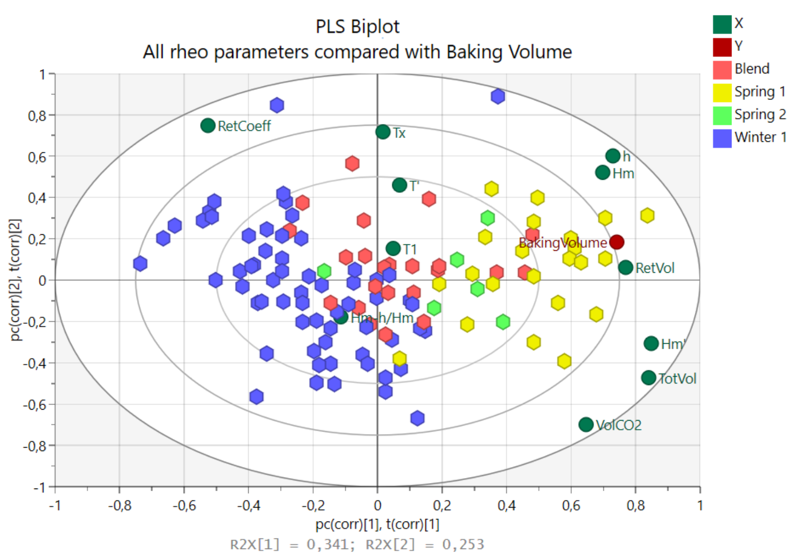

Figure 13. Biplot for a PLS model with Rheo F4 parameters as X and baking volume as Y.

Moreover, an effects plot of the initial model can be seen in *Figure 14*, where it demonstrates how the different parameters affect the baking volume. As a consequence of removing variables without significant influence, *T1*, *T'* and *Tx* were excluded for the improved model, *Rheo Improved*.

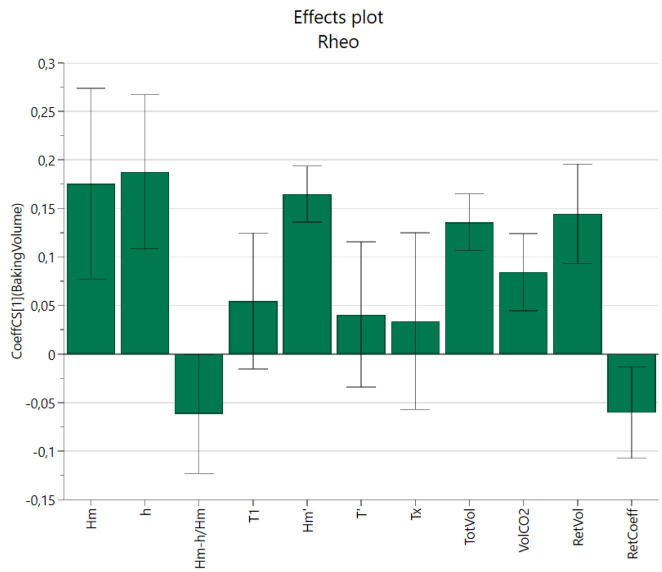

Figure 14. Effects plot for the model, *Rheo*, showing if the variables have a significant impact on baking volume.

Figure 15 shows the result of *Rheo Improved* predicting the baking volume of the sample set. The predictive power of the improved model only improved slightly to a Q^2-value 0.514, while the $R^2Y_{Predicted}$ increased to 0.556 and the RMSEP decreased to 201 ml.

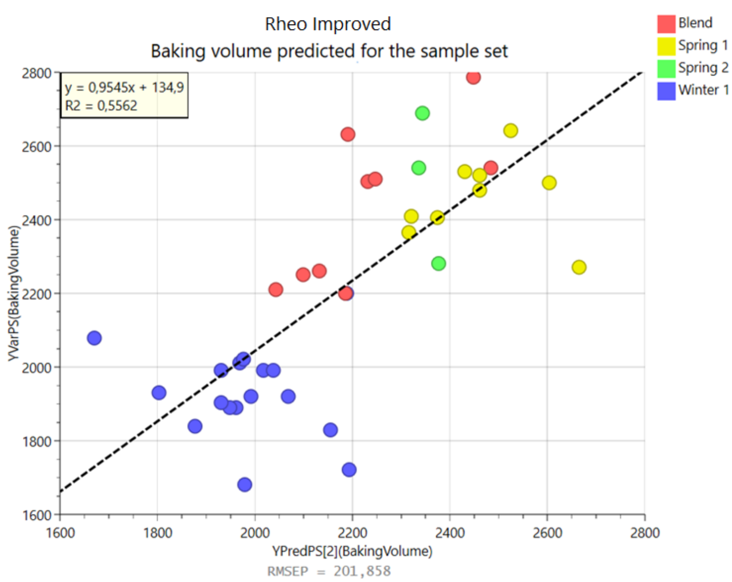

Figure 15. Predicted baking volume for the sample set using the *Rheo Improved* model with the correlated linear equation, $R^2Y_{Predicted}$ and root mean squared error of prediction (RMSEP).

4.4.3 Combined models

All of the available parameters were gathered in a model called *All Combined* and the resulting PLS biplot is presented in *Figure 16*. The model displays a cumulative R^2X value of 0.33 where the first component makes up 0.27 of the value and the second component 0.06. Consequently, the first component seems to have the most impact on the results and is also contributing to a separation of baking volume to the right side of the plot. Furthermore, it can be observed that the flour types are placed in similar clusters as previous PLS models shown above. As seen in *Figure 16* and throughout the report, the protein parameters consistently cluster close to baking volume. These findings confirm the great importance protein content in flour has for final bread quality. However, no data regarding the glutenin and gliadin ratios was available during the project and could thus not be examined. It is possible that this information could provide further insights and perhaps be a key to creating a better prediction of the baking volume.

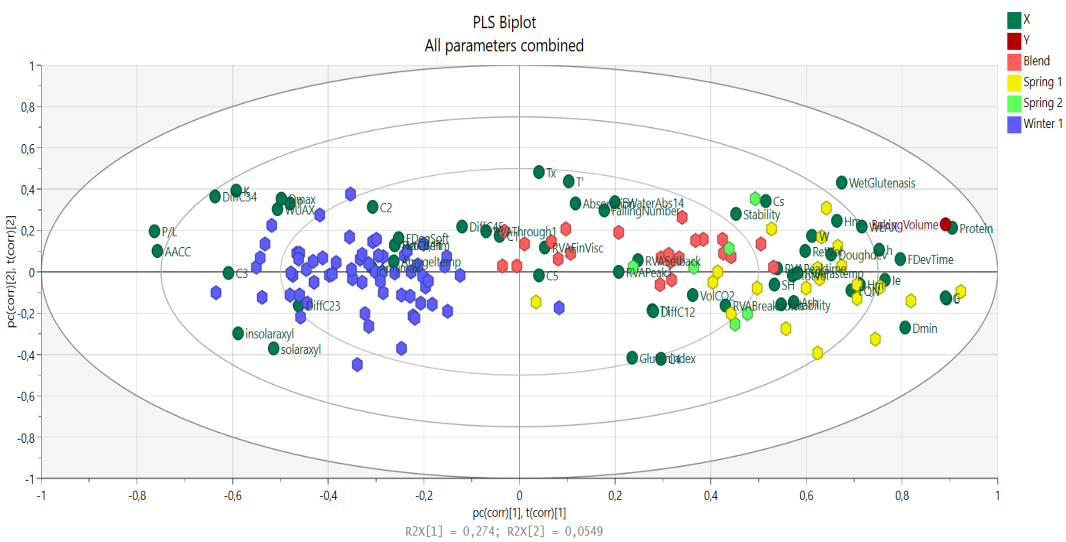

Figure 16. Biplot for a PLS model with all parameters combined as X and baking volume as Y.

The model was improved by removal of insignificant parameters and the new composition of variables are shown in *Table 2*. It can be noted that there still are a few Rheo F4 and Mixolab parameters remaining after the improvement. The following parameters had a significant impact on the first model containing all parameters; *Hm*, *h*, *Absorption*, *Doughdev* and *Cs*.

The formed model *All combined improved* had a Q-value of 0.834, which is the highest value obtained in this report. When the model was tested to predict the baking volume of the sample set as seen in *Figure 17*, it can be observed that the model indeed fits the data quite well. However, even though the Q-value is quite high and indicates a fairly good predicting power, it is probably still insufficient from an industrial perspective. The RMSEP value means that the model on average will predict the baking volume roughly 130 ml off from the measured value.

However, it is important to keep in mind that there is an initial error deriving from the test baking method that the prediction model is built on. Thus the accumulated error will be somewhat larger than RMSEP values presented throughout the report. Considering this amount of uncertainty, and the fact that the industries need consistently reliable results, it would be hard to put this model into practical use.

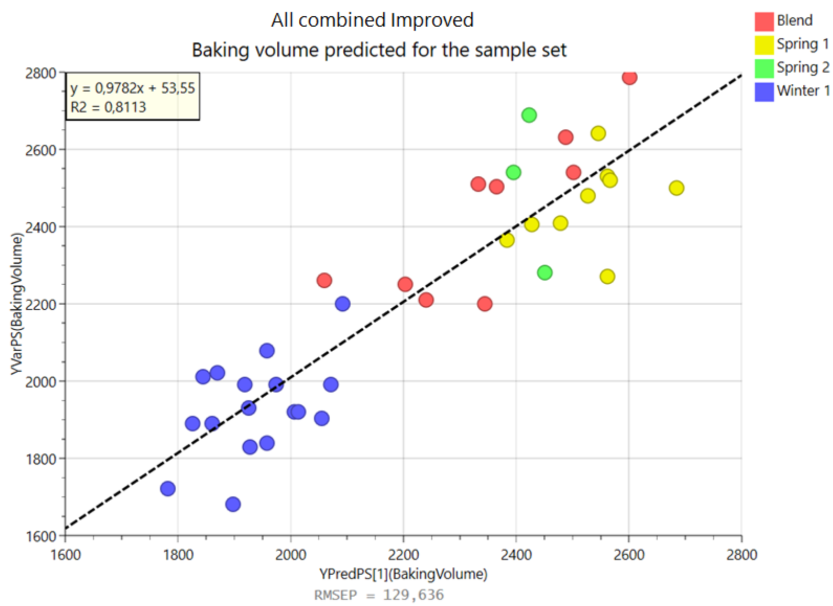

Figure 17. Predicted baking volume for the sample set using the *All Combined Improved* model with the correlated linear equation, $R^2Y_{Predicted}$ and root mean squared error of prediction (RMSEP).

Apart from the previously mentioned models, three reference models were created based on different combinations of datasets in order to get a clearer understanding of their individual contribution to the baking volume prediction. These models were constructed using the same methodology as previously, but only the improved models are presented in *Table 2*. When comparing the performance of the models, *All combined improved* was found to be the best. This could indicate that parameters from Mixolab and Rheo F4 possibly could add some value to a predictive model. However, it should be noted that *Reference model 2*, where Mixolab parameters were excluded, received the exact same Q-value, but fitted the sample set a bit worse. Overall it can be seen that the differences in predictive power of the reference models and *All combined improved* are very small. *Reference model 3* excluded Rheo F4 parameters and got the lowest Q-value, although it only differed by 0.007. So, after comparing the performances it seems evident that the time and effort it takes to perform Mixolab and Rheo F4 analyses is hard to justify with this minimal increase in predictive power of the model.

4.5 Optimally worked dough and adjusted flour

The results from the investigation of the influence of flour additives, regulators and optimal kneading are presented in the following two segments. By adding these elements, the results will give a more representative perspective on the bread making process in the industry. It is important to keep in mind throughout these results that the number of flour samples tested were significantly less than in the previously described experiments.

4.5.1 Mixolab

Two representative Mixolab curves obtained with unregulated flour, adjusted flour and optimally worked dough are shown in *Figure 18* and *19*, while the rest of the results are presented in Appendix 3. These curves show that, in general, the adjusted flour and optimally worked dough provided very similar curves. An exception, however, is that the optimally worked dough tends to lie a bit below the adjusted flour during the dough development part of the curve (up to *C2*). The biggest differences in all the analyses between unregulated flour and the adjusted flour and optimally worked dough, is that the *C3* peak appears at lower temperatures and that the *C2* minimums are higher. Additional differences can be seen in the analyses with the Blend and Winter 1 samples, where the unregulated flour curves showed significantly higher *C3* peaks along with greater drops from *C3* to *C4*.

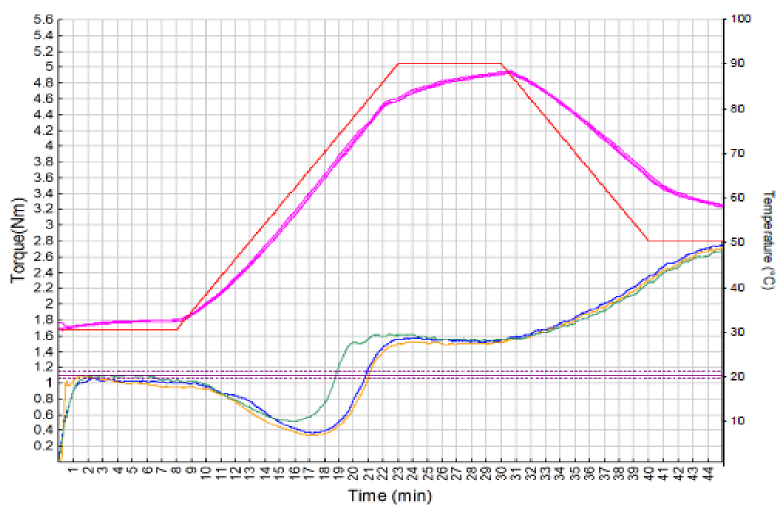

Figure 18. A comparison of curves from the unadjusted flour (green line), adjusted flour (blue line) and the optimally worked dough (yellow line) for a Spring 1 flour.

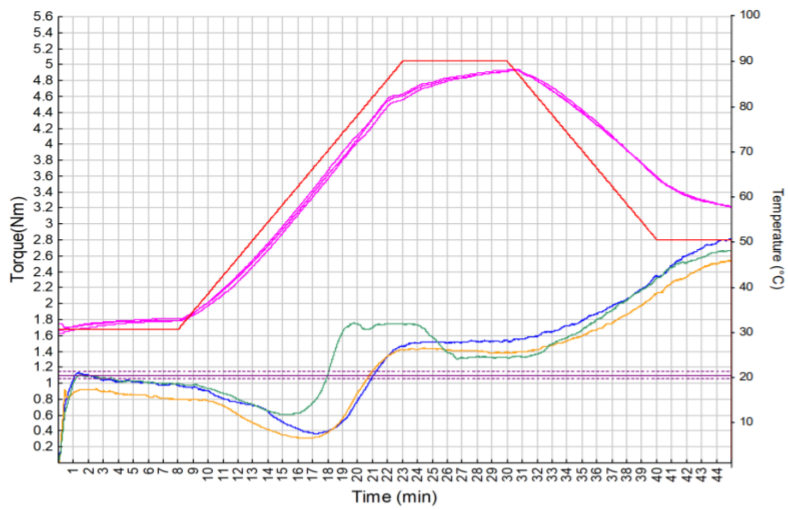

Figure 19. A comparison of curves from the unadjusted flour (green line), adjusted flour (blue line) and the optimally worked dough (yellow line) for a Blend flour.

The similarities between the adjusted flour and optimally worked dough curves indicates that the main factor affecting the Mixolab analysis is the flour composition and not how the dough is worked. However, since the optimally worked dough already has been through the dough development phase before entering the Mixolab equipment, parameters obtained from this stage are not comparable.

The large difference in observed gelatinization temperature between the unadjusted and adjusted flour could be attributed to the addition of salt in the latter. Since the thermal breakdown of the gluten network and starch gelatinization are two processes that may overlap and occur simultaneously, the observed higher *C2* minimums could perhaps be an effect of the lower gelatinization temperature. This would mean that the decrease in torque due to protein denaturation is masked by the increase associated with swelling of the starch granules. Consequently, it may be suggested that the *C2* parameter might be less correlated with protein quality than intended and instead more related to starch properties. Interestingly enough, it can be seen in *Figure 4* that *C2* indeed clusters more with starch property parameters than the protein quality ones. Unfortunately, data regarding the onset temperature of gelatinization for the different flours has not been obtained and it is therefore not possible to confirm that *C2* correlates with earlier gelatinization. It is however something that might be interesting to investigate further.

Furthermore, the difference in gelatinization peak heights (*C3*) is presumably related to the addition of malt, which increases the amylolytic activity and thus enhances the degradation of starch. Consequently, less starch is gelatinized, resulting in lower peaks. The lower content of starch could also explain why the *C3* peak height difference can not be seen for the spring wheat since it usually contains lower amounts of starch. It is hard to explain why there is no clear drop from *C3* to *C4* in the case of the regulated flours and Spring 1. Another consequence of adding malt in the adjusted flours should theoretically result in larger decreases, which is the opposite of what can be seen in the figures. The reason for this behaviour could be worth investigating further.

To summarize, Mixolab analyses of unadjusted flour seem to provide results in the first and last part of the curve that reflects the reality in the baking industry. Meanwhile, parameters obtained from the middle part of the curve, such as *C2* and *C3*, change drastically upon flour regulation and are therefore not directly transferable to reality.

4.5.2 Alveolab

The position of Alveolab parameters in relation to baking volume in *Figure 6* and *16* motivated further studies of the method. This was done in a similar manner as with Mixolab in section 4.5.1. Correlations between the different Alveolab experiments were compiled in *Table 3*. The table shows the strength of the interrelationships between the different ways of conducting the analysis. High correlation values were obtained between the adjusted flour and the optimally worked dough (column 2) and show that these two adjustments yield comparable results. Further, it would indicate that the mixing in the Alveolab is equivalent to the mixing conducted in the industry. Column 3 and 4 show poor correlations between the standard protocol and the two methods examined for the adjusted flour. Consequently, the results obtained using the standard protocol does not seem to give an accurate indication of the rheological behaviours of doughs in the industry.

Table 3. The correlations coefficients between an optimally worked dough, an adjusted flour, and an unadjusted flour with standard protocol in regards to P, L, W and P/L.

Parameter	Corr. coeff (adjusted flour and optimally worked dough)	Corr. coeff (optimally worked dough and standard protocol)	Corr. coeff (adjusted flour and standard protocol)
P	0.94	-0.41	-0.30
L	0.65	0.32	0.35
W	0.89	0.53	0.79
P/L	0.81	-0.4	-0.23

The obtained results were further compared with the baking volume, where the correlations can be seen in *Table 4*. All parameters are positively correlated with baking volume for the analysis with the modified flours (adjusted flour and optimized worked dough). Only two parameters were positively correlated to baking volume while using unregulated flour and the standard protocol. Generally, the results from the constant hydration protocol had stronger correlations to baking volume.

Table 4. The correlations between baking volume and Alveolab values for an adjusted flour, an optimally worked dough, and an unadjusted flour with standard protocol.

Correlations with baking volume					
P_adj.flour	0.435	P_opti.dough	0.660	P_SP	-0.551
L_adj.flour	0.382	L_opti.dough	0.032	L_SP	0.825
W_adj.flour	0.811	W_opti.dough	0.712	W_SP	0.844
P/L_adj.flour	0.122	P/L_opti.dough	0.584	P/L_SP	-0.743

With the result in *Table 3*, it is inferred that there are more similarities between the adjusted flour and optimally worked dough analyses than with standard protocol. Although, *Table 4* shows that the results using the standard protocol might be interesting to investigate in order to predict the baking volume. Therefore, analysis of a larger amount of samples is recommended to give a more certain prediction.

5. Conclusions

To summarize the findings of this report, it can be said that the Chopin instruments Mixolab and Rheo F4 had a minor positive impact in the creation of a PLS regression model for predicting baking volume. The best performing model included the Mixolab parameters Absorption, Dough development time and Cs and Rheo F4 parameters maximum dough height and dough height after three hours (*Hm* and *h*), together with 13 previously determined quality control parameters, since all of these were found to have significant influence on baking volume. However, since the predictive power of this model was only negligibly better than models where the Chopin instruments were excluded, there does not seem to be sufficient reasons for adding these into routine quality controls.

PCA plots revealed some correlations between parameters obtained from Mixolab and corresponding parameters from other currently used methods. These correlations may be interesting to further investigate in order to explore the possibility of using Mixolab as an alternative to some of the current quality control methods.

Furthermore, it was found that results obtained with unadjusted flours in Mixolab and Alveolab, are not directly translatable to the flour mixtures and dough preparations used in industrial bakeries. However, the mixing performed in the two equipment seem to be roughly equivalent to how doughs are kneaded in the industry. It should be noted that the measurements were only conducted on six flour samples and one should therefore be careful to draw any definite conclusions.

With this observation in mind, it would be interesting to further investigate how the predefined protocols of the equipment relates to industrial processes. It would be especially intriguing since it is somewhat problematic to base baking volume predictions on data obtained with unadjusted flour. Moreover, it could be a good idea to include other quality parameters than baking volume, such as texture analysis of the final bread, to get a broader quality perspective.

References

Arya, Shalini; Sadawarte, Prajakta; Ashish, Waghmare. 2015. *Importance of damaged starch in bakery products - a review*. Institute of Chemical Technology, Mumbai. Available at: https://www.researchgate.net/publication/277634923_Importance_of_damaged_starch_in_bakery_products_-_a_review

Chopin Technologies. 2013. *RHEO F4. User's Manual*. [manual] Villeneuve-La-Garenne.

Chopin Technologies. 2016 (a). *Mixolab 2. User's Manual*. [manual] Villeneuve-La-Garenne.

Chopin Technologies. 2016 (b). *Alveolab. User's Manual*. [manual] Villeneuve-La-Garenne.

Courtin, C.M.; Delcour, J.A. 2002. Arabinoxylans and Endoxylanases in Wheat Flour Bread-making. *Journal of Cereal Science* 35(3): 225-243. DOI: doi.org/10.1006/jcrs.2001.0433

Eriksson, L; Byrne T; Johansson, E; Trygg, J; Vikström, C. 2013. *Multi- and Megavariate Data Analysis : Basic Principles and Applications*. 3rd edition. MKS Umetrics AB, 42.

Finnie, Sean; Atwell, William A. 2016. *Wheat flour.* 2nd edition. AACC International. DOI: doi.org/10.1016/C2015-0-06189-5

Gan, Z.; Ellis, P.R.; Schofield, J.D. 1995. Gas Cell Stabilisation and Gas Retention in Wheat Bread Dough. *Journal of Cereal Science* 21 (3): 215-230. DOI: doi.org/10.1006/jcrs.1995.0025

Jaadi, Zakaria. 2021. *A step-by-step explanation of Principal Component Analysis (PCA).* https://builtin.com/data-science/step-step-explanation-principal-component-analysis (Accessed: 2021-04-12)

Jolliffe, I. T. 2002. *Introduction*. Principal Component Analysis. 2nd edition. Springer, New York, NY, 1-9. DOI: doi-org.ludwig.lub.lu.se/10.1007/b98835

Mills, E.N.C.; Wilde, P.J.; Salt, L.J.; Skeggs, P. 2003. Bubble Formation and Stabilization in Bread Dough. *Food and Bioproducts Processing* 81(3): 189-193. DOI: doi.org/10.1205/096030803322437956

Nicol, Thomas W.J.; Isobec, Noriyuki; Clark, James H; Matubayasi, Nobuyuki; Shimizu, Seishi. 2019. The mechanism of salt effects on starch gelatinization from a statistical thermodynamic perspective. *Food Hydrocolloids* 87: 593-601. DOI: doi.org/10.1016/j.foodhyd.2018.08.042

Ooms, N.; Delcour, J. A. 2019. How to impact gluten protein network formation during wheat flour dough making. *Current Opinion in Food Science*, 25, 88–97. DOI: doi.org/10.1016/j.cofs.2019.04.001

Rosentrater, Kurt A; Evers, A.D. 2018. Chapter 8: Bread baking technology. *Kent's technology of cereals: An Introduction for Students of Food Science and Agriculture.* 5th edition. Woodhead Publishing, 565-622. DOI: doi.org/10.1016/C2014-0-02847-X

Sahi, S. 2014. Ascorbic Acid and Redox Agents in Bakery Systems. Bakery Products Science. 2nd edition. Wiley Blackwell, 183-196. DOI: doi.org/10.1002/9781118792001.ch10

Shewry, Peter; Halford, Nigel; Lafiandra Domenico. 2003. Genetics of Wheat Gluten Proteins. *Advances in Genetics* 49, 111-184 DOI: doi.org/10.1016/S0065-2660(03)01003-4

Swedish Board of Agriculture. 2020. *Jordbruksmarkens användning 2020. Preliminär statistik*. https://jordbruksverket.se/download/18.5b7c91b9172c01731758cd12/1592494021096/JO10SM2001.pdf

Wold, Svante; Sjöström, Michael; Eriksson, Lennart. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58 (2): 109-130. DOI: doi.org/10.1016/S0169-7439(01)00155-1

Appendix 1: Overview of data

Table A displays the data used in the study using only unadjusted flour (not optimal worked dough). It visualises by whom the analysis was performed, the parameters measured, how these parameters were used in the multivariate analysis, within what range the data was and the unit.

Table A. All data used in this project gathered before and during the project.

Analysis method	Performed by	Parameters	Referred to in report as	Min	Mean	Max	Unit
Mixolab	Report authors	Absorption	*Absorption*	55.3	59.0	63.1	%
		Dough development	*Doughdev*	1.4	2.9	6.5	Min
		Torque at C1	*C1*	1.05	1.1	1.15	Nm
		Torque during stability time	*Cs*	0.89	1.00	1.11	Nm
		Torque at C2	*C2*	0.48	0.57	0.65	Nm
		Torque at C3	*C3*	1.52	1.76	1.99	Nm
		Torque at C4	*C4*	1.19	1.40	1.83	Nm
		Torque at C5	*C5*	2.14	2.78	3.42	Nm
		Stability	*Stability*	5.5	8.7	10	Min
		Difference between C1 and C2	*DiffC12*	0.44	0.53	0.66	Nm
		Difference between C2 and C3	*DiffC23*	0.94	1.19	1.48	Nm
		Difference between C3 and C4	*DiffC34*	0.05	0.36	0.57	Nm
		Difference between C4 and	*DiffC45*	0.76	1.37	2.11	Nm

		C5					
Rheo F4	Report authors	Maximum height - dough development	*Hm*	41.5	49.8	65.2	mm
		Dough height after 3h	*h*	32.2	47.5	59.6	mm
		Drop in dough height	*Hm-h/Hm*	0	3.9	15.8	%
		Time to reach maximum dough height	*T1*	87	156.3	180	min
		Maximum height - gas release	*Hm'*	58.4	67.5	76.3	ml
		Time to reach Hm'	*T'*	48	68.3	97.5	min
		Dough porosity time	*Tx*	43.5	66.8	97.5	min
		Total Volume	*TotVol*	1377	1535	1685	ml
		Volume of CO_2 released	*VolCO2*	155	301	414	ml
		Retention Volume	*RetVol*	1138	1234	1333	ml
		Retention Coefficient	*RetCoeff*	75.3	80.5	87.5	%
Alveolab	Previous Master's student	Tenacity	*P*	56	83.7	114	mm H_2O
		Extensibility	*L*	50	90.9	146	mm
		Extensibility index	*G*	15.7	21	26.8	-
		Dough baking strength	*W*	135	232.3	346	10^{-4} J

		Curve configuration ratio	*P/L*	0.4	1.0	2.28	-
		Elasticity index	*Ie*	37.9	50.4	60.5	%
		Strength coefficient	*K*	28147	42003	55003	-
		Strain hardening	*SH*	1.51	1.7	1.86	-
		Maximum derivative	*Dmin*	-3.74	-2.5	-1.61	-
		Minimum derivative	*Dmax*	4.76	6.4	7.8	-
Rapid Visco Analyzer	Report authors	First peak	*RVAPeak1*	1003	1370	1679	RVU
		Final viscosity	*RVAFinalVisc*	1289	1781	2154	RVU
		Pasting temperature	*RVAPastingtemp*	85.0	87.4	90.1	℃
SDmatic	Previous Master's student	Damaged starch	*AACC*	4.80	6.28	7.67	%
Farinograph	Lantmännen Cerealia	Water absorption (14%)	*FWaterAbs14*	55.0	59.5	65.5	%
		Development time	*FDevTime*	1.2	3.7	7.4	min
		Stability	*FStability*	3.4	6.7	12.5	min
		Degree of softening	*FDegSoft*	36.0	79.5	121	BU
		Farinogram quality number	*FQN*	16.0	75.2	126	-
Amylograph	Lantmännen Cerealia	Amylogram max	*Amymax*	692	1329	2081	AU

		Gelatinisation temperature	*Amygeltemp*	84.5	89	93.9	°C
Test baking	Lantmännen Cerealia	Baking volume	*BakingVol*	1650	2158	2970	ml
Wet gluten	Lantmännen Cerealia	Gluten index	*GlutenIndex*	72.6	91.4	99.6	%
		Wet gluten as is	*WetGlutenAsis*	22.9	29	35.5	%
Falling Number	Lantmännen Cerealia	Falling number	*FallingNumber*	304	394	471	s
Fibre	SLU	Water unextractable arabinoxylan	*WUAX*	Data not yet published			
		Water extractable arabinoxylan	*WEAX*				
		Insoluble arabinoxylan	*insolaraxyl*				
		Soluble arabinoxylan	*solaraxyl*				
FOSS NIT	Lantmännen Cerealia	Protein content	*Protein*	10.5	12.4	16.2	%
		Ash content	*Ash*	0.48	0.60	0.72	%

Appendix 2: Reproducibility data

Table B displays results from the reproducibility study conducted for Mixolab and Rheo F4, where the accepted ranges are based on limit of reproducibility values provided by Chopin Technologies.

Table B. Reproducibility data for Mixolab and Rheo F4.

Instrument	Parameter	Reproducibility
Mixolab	Water absorption	100%
	Torque C2	100%
	Torque C3	100%
	Torque C3	100%
	Torque C4	100%
	Torque C5	100%
	Stability	100%
	Time	100%
	Temp C1	100%
	Temp C2	100%
	Temp C3	80%
	Temp C4	100%
	Temp C5	100%
Rheo F4		
	Hm	100%
	h	100%
	(Hm-h)/Hm	100%
	T1	100%

T2	100%
T2-T'2	100%
H'm	100%
T'1	100%
Tx	100%
Total volume	100%
Volume CO2	100%
Retention volume	100%
Retention coefficient	100%

Appendix 3: Mixolab curves for optimal worked dough

Figures 1-4 displays the raw curves, from Mixolab 2, for the remaining flours assessed in section *4.5 Optimal worked dough and adjusted flour*. The green line represents unadjusted flour, the blue line is the adjusted flour and the yellow line is connected to the pre-worked dough. Underneath each figure it is stated what type of flour the measurement is based on.

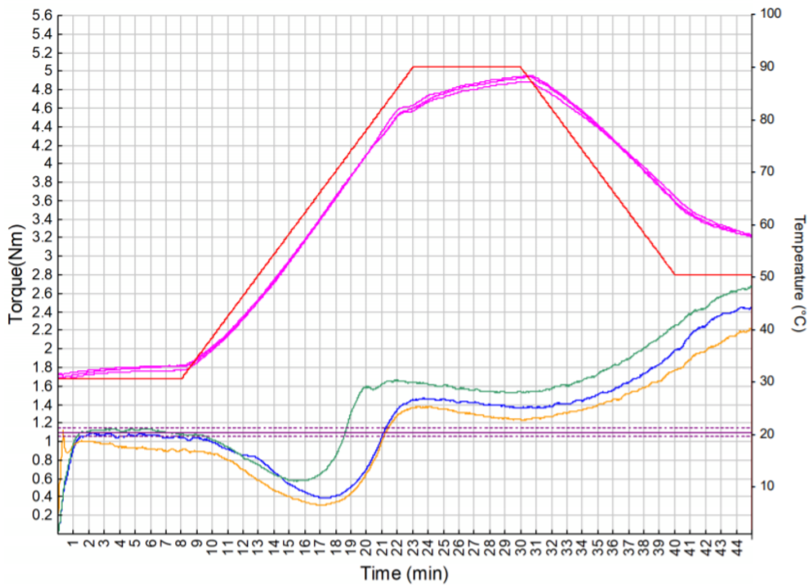

Figure 1. Mixolab raw curve of a Spring 1 flour.

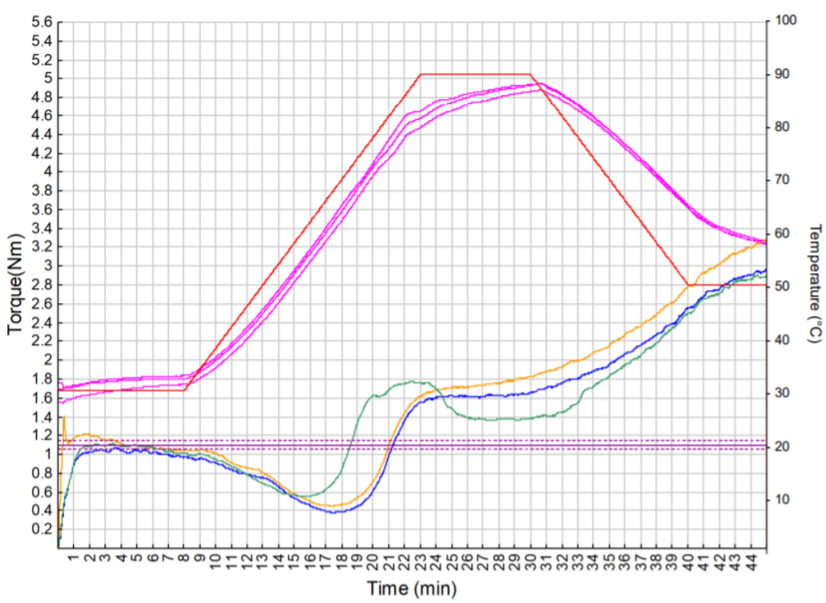

Figure 2. Mixolab raw curve of a Blend flour.

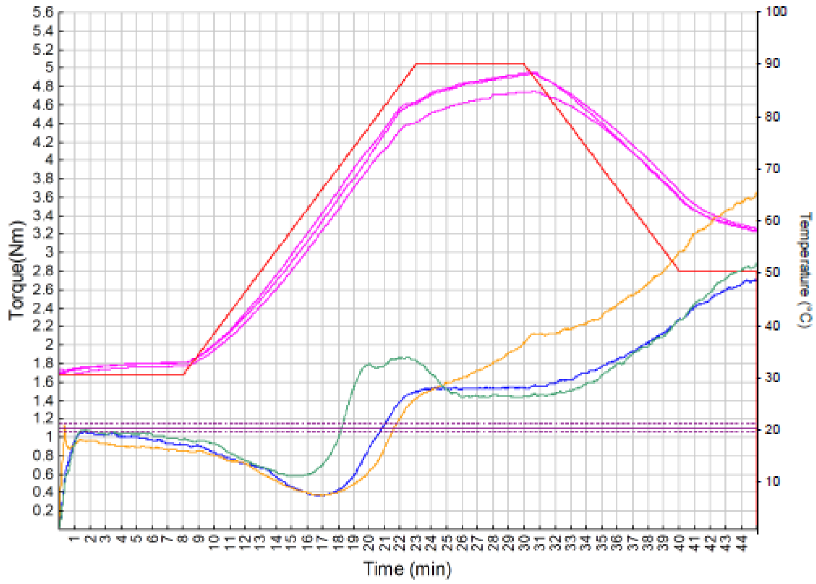

Figure 3. Mixolab raw curve of a Winter 1 flour.

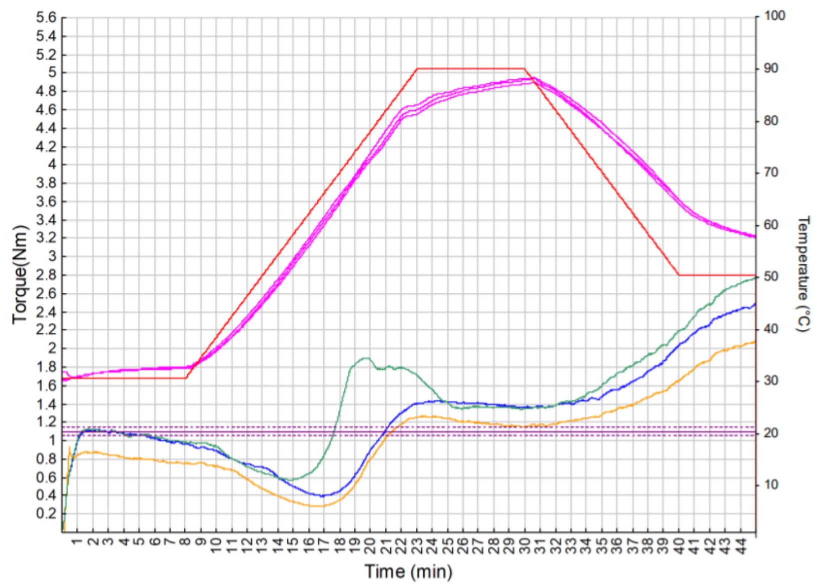

Figure 4. Mixolab raw curve of a Winter 1 flour.