Damaged starch and dietary fibre content in Swedish wheat flour

– PLS-modelling to predict baking volume

Skadad stärkelse och kostfibrer i svenskt vetemjöl
- PLS-modellering för att förutse brödvolym

Edvin Nåbo

Degree project/Independent project – 30 hp
Swedish University of Agricultural Sciences, SLU
Molecular Sciences, 2021:02
Agricultural Programme – Food Science
Uppsala, 2021

Damaged starch and dietary fibre content in Swedish wheat flour – PLS-modelling to predict baking volume

Skadad stärkelse och kostfibrer i svenskt vetemjöl
- PLS-modellering för att förutse brödvolym

Edvin Nåbo

Supervisor:	Roger Andersson, SLU, Department of Molecular Sciences
Assistant supervisor:	Louise Selga, Lantmännen and SLU, Department of Molecular Sciences
Examiner:	Annica Andersson, SLU, Department of Molecular Sciences
Year of publication:	2021
Credits:	30 hp
Level:	Second cycle, A2E
Course title:	Master Thesis in Food Science, A2E – Agriculture Programme - Food Science
Course code:	EX0877
Programme/education:	Agricultural Programme – Food Science (Livsmedelsagronomprogrammet)
Course coordinating dept:	Department of Molecular Sciences
Title of series:	Molecular Sciences
Part number	2021:02

Keywords: **damaged starch, dietary fibre, bread volume, PLS**

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file.

If you are more than one author you all need to agree on a decision. Read about SLU's publishing agreement here: https://www.slu.se/en/subweb/library/publish-and-analyse/register-and-publish/agreement-for-publishing/.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.

Abstract

SDmatic, SRC-CHOPIN 2 and Alveolab are used to evaluate flour, but not widely used in Sweden. This study aimed to evaluate the machines and see if they could be used to predict baking volume for bread baked on Swedish wheat flour. PLS-models were built with baking volume as the Y-variable. It was noticed that baking volume of breads made on winter wheats and spring wheats were explained by different parameters and as a result building separate PLS-models for these groups gave the best results. Damaged starch had negative impact on baking volume for spring wheats but not for bread baked on winter wheats. All PLS-models were optimised for Q2 by removal of X-variables. Variables from SDmatic, SRC-CHOPIN 2 or Alveolab were left in all PLS-models. The most promising model in this study was built on winter wheats and had a Root Mean Square Error of Prediction (RMSEP) at 75 ml, which can be compared to the average bread with a volume of 2032 ml. This model had only one parameter from these machines and it is thus unclear how useful these machines are when predicting baking volume of bread baked on Swedish wheat flour. Glucomannan was the most important parameter for this model based on Variable Importance in Projection (VIP)-scores and was positively correlated with baking volume. Baking volume was the only predicted quality parameter and future studies should analyse how these machines can predict other quality parameters, such as crumb structure, bread staling and consumer acceptability.

Keywords: damaged starch, dietary fibre, bread volume, PLS

Sammanfattning

SDmatic, SRC-CHOPIN 2 och Alveolab används ofta för att bestämma egenskaper hos vetemjöl, men inte i Sverige. I den här studien undersöktes om maskinerna kunde användas i modeller för att förutse bakvolym av bröd bakat på svenskt vetemjöl. PLS-modeller gjordes med bakvolym som Y-faktor. Olika parametrar förklarade bakvolymen för bröd bakat på vårvete och höstvete, och därför gjordes separata PLS-modeller för vårvete och höstvete. Skadad stärkelse hade en negativ inverkan på bakvolymen för bröd bakat vårvete men inte för bröd bakat på höstvete. Parametrar i PLS-modellerna togs bort för att optimera Q2. Minst en variabel från SDmatic, SRC-CHOPIN 2 eller Alveolab var kvar i alla PLS-modeller efter optimeringen. Den mest lovande modellen var byggd på höstvete och hade Root Mean Square Error of Prediction (RMSEP) på 75 ml vilket kan jämföras med den genomsnittliga bakvolymen på 2032 ml för bröd bakat på höstvete. Denna modell hade enbart en parameter från Alveolab och inga från SDmatic eller SRC-CHOPIN 2. Det är därför svårt att säga om dessa maskiner är användbara i PLS-modeller som ska förutse bakvolym av bröd bakat på svenskt vetemjöl. Glucomannan var den viktigaste parametern för denna modell baserat på Variable Importance in Projection (VIP) och var dessutom positivt korrelerat med bakvolym. Bakvolym var den enda kvalitativa parametern i den här studien. Framtida studier skulle kunna se om dessa maskiner kan användas för att förutse parametrar som strukturen av inkråmet, brödets hållbarhet och konsumentacceptans.

Nyckelord: skadad stärkelse, kostfibrer, brödvolym, PLS

Table of contents

List of tables

List of figures

Abbreviations

Ara/xyl	Quota of arabinose and xylose residues in arabinoxylan
AX	Arabinoxylan
BW	Baker´s wheat
Dmax	Highest point in the derivative curve of the Alveolab curve
Dmin	Lowest point in the derivative curve of the Alveolab curve
ES	Extra strong spring wheat
Insol man	Insoluble mannose (originating from glucomannan), analysed with an adjusted Uppsala method
PCA	Principal Component Analysis
PLS	Partial Least Squares Regression
RMSEP	Root Mean Square Error of Prediction
SH	Strain Hardening
Sol man	Soluble mannose (originating from glucomannan), analysed with an adjusted Uppsala method
SW	Spring wheat
UCD	Chopin Dubois Unit
UCDc	Adjusted UCD
VIP	Variable Importance in Projection
WE-AX	Water-extractable arabinoxylan
WU-AX	Water-unextractable arabinoxylan
WW	Winter wheat
Xyl sol	Soluble xylose

1. Introduction

Wheat is the largest cereal crop in Sweden with an estimate of 452 700 hectares cultivated? of which 89 % was winter wheat for the production year 2020 (Swedish Board of Agriculture 2020). Wheat flour typically consists of 70-75 % starch, 14 % water, 10-12 % protein, 2-3 % non-starch polysaccharides and 2 % lipids (Goesaert et al. 2005). Starch in wheat usually consists of 25-28 % amylose and 72-75 % amylopectin. Starch is located in the endosperm of wheat kernels and stored in granules. Pieces of starch detach from the granules as they are being crushed in the milling process and are called damaged starch. Flours with smaller particle sizes have more damaged starch (Wang & Flores 2000; Ma et al. 2020). Damaged starch has different processing properties, compared to native starch, and affects water absorption of flours and mixing properties of doughs. The water absorption increases since damaged starch swell at room temperature and bind more water, which native starch only does at higher temperatures (Morrison & Tester 1994). It also contributes to late stage fermentation of dough by promoting β-amylase hydrolysis which degrades starch into maltose (Delcour & Hoseney 2010). Yeast consumes maltose and the by-product, carbon dioxide, makes the dough rise. A content of 4.5-8 % damaged starch has been found to yield flour with the highest baking quality (Arya et al. 2015).

SDmatic is an amperometrical method for estimating starch damage by measuring how much iodine a flour absorbs in a water-iodine solution. Amylose from damaged starch granules form complexes with iodine and the result is thus affected by amylose content of the flour. Enzymatic tests for damaged starch are more accurate but take longer time and demands highly trained operators (McAllister et al. 2011).

Wheat flour contains polysaccharides other than amylose and amylopectin and are referred to as non-starch polysaccharides. Arabinoxylan (AX), β-glucan, cellulose and arabinogalactan-peptides are part of this group. Dry matter of wheat endosperm cell walls consist of 75 % non-starch polysaccharides of which 85 % are AX. AX are chains of β-1,4-linked D-xylopyranosyl residues substituted on O-2, O-3 or O-2,3 with α-L-arabinofuranoside residues. A quota of 0.5-0.6 of arabinose to xylose (A/X) is normal in wheat AX. Sifted wheat flour consists of 1.3 and 2.8% AX (Mendis et al. 2013). AX are either water-extractable (WE-AX) or water-unextractable (WU-AX). Water extractability of AX increases as the molecular weight decreases and the degree of substitution increases (Goesaert et al. 2005). WU-AX impair proper gluten formation during dough mixing, either by physical hindrance or by extensive water absorption yielding less water for the development of the gluten network (Courtin et al. 1999; Courtin & Delcour 2002). Gas cells within dough might collapse when WU-AX perforate them and thus result in a lower bread volume (Courtin & Delcour 2002). WE-AX are associated with breads with better quality (Goesaert et al. 2005). An explanation might be the increased viscosity of the aqueous phase in the dough which stabilises air bubbles (Kaur et al. 2019), an effect which also increases oven-spring and yields breads with better crumb structure and increased volume (Gan et al. 1995). Too much AX give stiff doughs and breads with lower volume. If AX are degraded by enzymes the dough becomes sticky and yields a bread with lower volume (Delcour et al. 1991). WE-AX and WU-AX have also both been shown to slow down staling of bread (Courtin & Delcour 2002; Kaur et al. 2019).

In addition to AX, wheat flour contains several polymers that affect the viscoelasticity of doughs when mixed with water. Solvent Retention Capacity, SRC, measures swelling of flour using four different solutions that enhances the swelling of specific polymers. The process of SRC involves weighing of flour, dissolving in excess solution, shaking and centrifugation of the tubes and finally weighing the remaining pellet (Kweon et al. 2011). The SRC-CHOPIN 2 is an automated version of this test which enables more efficient testing and is less operator demanding (Dubat et al. 2019). Huen et al. (2018) used PCA to show that values from SRC-CHOPIN, AlveoLab and SDmatic correlated with values from traditional flour analyses such as Farinograph and Extensograph. The Alveolab analyses rheological behaviour of dough. In combination with SRC one could explain which polymer affects the rheological behaviour in the Alveolab. P-value from an alveograph is a measurement of the dough's tenacity and ability to resist deformation. P-values generally increases with increased level of damaged starch. Since damaged starch is mainly affected by milling, a combination of Alveolab and SRC might give insights into how to adapt the milling process to achieve the best flour possible (Kweon et al. 2011).

Baking quality is often measured as in volume of the final bread. Several tests are applied in Sweden (e.g. Falling number, Farinograph, Amylograph etc.) for flour quality control. These tests can predict bread volume but not in a perfect way and therefore test baking is used as a reference. The aim of this report is therefore twofold. The first aim is to study if SRC-CHOPIN 2, SDmatic and Alveolab can be used to predict flour quality of Swedish wheat, and the second aim is to evaluate how user friendly these tests are.

2. Materials and methods

This study tested 199 wheat flours with SDmatic, AlveoLab and SRC-CHOPIN 2. 15 of the 199 flours were tested as duplicates. Amylograph, wet gluten, falling number, Farinograph, baking volume, ash content and protein content were previously determined by Lantmännen Cerealia. Fibre composition using a modified Uppsala method (Andersson et al. 1999) was analysed by the Swedish University of Agricultural Sciences, SLU. The dietary fibre results are reported as water-extractable arabinoxylan (WE-AX), water-unextractable arabinoxylan (WU-AX) and insoluble mannose residues from the fibre analysis (man insol). The ratio between arabinose and xylose residues (ara/xyl) is also included in the analysis.

2.1. Flours

Five types of wheat flour were tested. All wheat, except for the German E-wheat, was harvested in Sweden during either 2018, 2019 or 2020.

2.1.1. Winter wheat

The name winter wheat (WW) originates from time of sowing since this type of wheat is sown in the winter half of the year and harvested the next year. A wheat with typically a protein content of 10-12 % and considered to be weaker than spring wheat. Ideal for biscuits or cakes but sometimes used for breadmaking as well (Yngveson 2015). WW has a slightly higher exchange (1-3 %) during milling than SW and ES.

2.1.2. Spring wheat

Spring wheat (SW) is sown during the spring and harvested in the autumn the same year. This type of wheat is used for breadmaking and has usually a protein content of 12-14 %.

2.1.3. Extra strong spring wheat

Extra strong spring wheat (ES) is spring wheat with a high protein content, 13-14 % and considered to be an extra strong flour. This classification is done by Lantmännen Cerealia.

2.1.4. Baker's wheat

Baker´s wheat (BW) is a mixture of WW, SW and German E-wheat. In Germany, wheat can be classified as E, A, B or K. E-wheat is the highest class, with at least a protein content of 13.8 %, a falling number of 285 s and a bread volume of 710 ml/100 g. E-wheat is rarely used on its own and instead mixed with other flours to improve their quality (Lásztity & Salgó 2002).

2.1.5. Julius

Julius flour is a winter wheat flour and consists only of one cultivar? named Julius. This was the only analysed flour that was cultivar specific and harvested year 2020.

2.2. Chemicals

All chemicals used in this report, listed in Table 1, were bought at VWR, part of Avantor**.**

Table 1. Chemicals used in this report

Chemical	**Concentration or purity**	**CAS-number**
Citric acid	99 %	77-92-9
Sodium thiosulphate	1 mol/l	7772-98-7
Potassium iodine	≥99.5 %	7681-11-0
Lactic acid	80-85 %	50-21-5
Sodium carbonate	99 %	497-19-8
D(+)-Sucrose	≥99 %	57-50-1

2.3. SDmatic

Based on the principles of Medcalf & Gilles (1965) the SDmatic measures damaged starch based on iodine absorption in a solution at 35°C. SDmatic runs an electric current through an iodine solution. Flour is added and will absorb iodine and the current will drop. Damaged starch absorbs more iodine compared to native starch and SDmatic may therefore estimate the amount of damaged starch. SDmatic measures the drop in current as percentage of iodine absorption (Ai %). SDmatic presents the result as five different values that are all based on the iodine absorption (Ai %, UCD, UCDc, AACC and Farrand). Ai % is the actual iodine absorption with values often between 93 and 95 %. The four other units are calculated based on the Ai %, Appendix 1. Chopin Dubois Unit (UCD) scales the Ai % and can be calculated directly from Ai %, while the adjusted Chopin Dubois Unit (UCDc) adjusts for the moisture content and protein content of the flour. AACC76-31 and Farrand are enzymatic methods for determination of damaged starch. SDmatic gives an estimate of these values.

2.3.1. SDmatic's procedure

The SDmatic User's manual explains how to prepare one sample (CHOPIN Technologies 2019). The manual instructs how to perform one test which requires 120 ml distilled water. To speed up the process and to lower the impact of the operator for each test, a batch of SDmatic solution was prepared by scaling up the solution from 120 ml to 2 litres. 2000 ml ± 0.1 ml distilled water, 50 g ± 0.2 g potassium iodine, 25 g ± 0.1 g citric acid and 17 drops sodium thiosulfate 0.1 mol/l were mixed in a DURAN® glass bottle by shaking for 1 minute. Solution was then dosed into a reaction beaker for each analysis and weighed to 124.5 g ± 0.1 g. The reaction beaker was placed in the machine and the SDmatic arm was lowered to submerge the probe in the solution. 1 g ± 0.1 g flour was weighed onto an SDmatic spoon and placed in the machine. Exact flour weight, moisture content and protein content of the flour were added in the test instructions of the machine.

The measurement cycle has six phases:

1. SDmatic heats the solution to 35° C.

2. & 3. Iodine is produced electrochemically by the probe according to the mass of flour indicated at the beginning of the test.

4. No more iodine is produced and the electrical current which previously increased is kept at a plateau.

5. The flour is introduced in the solution. The current decreases as the iodine is absorbed by the flour.

6. The final current is measured by the probe 180 seconds after the addition of flour.

2.4. SRC-CHOPIN 2

SRC-CHOPIN 2 is an automated version of the Solvent Retention Capacity test. It tests how different polymers in wheat contribute to the swelling of flour in selective solvents. The constituents and solvents are, 5 % (w/w) lactic acid for gluten proteins, 5 % (w/w) sodium carbonate for damaged starch, 50 % (w/w) sucrose for pentosans and distilled water as a reference (Kweon et al. 2011).

2.4.1. Procedure of SRC-CHOPIN 2

Weights of tubes, flours and gels are saved by the machine and no calculations on the side are needed.

Empty tubes with caps on were weighed. 5 g ± 0.05 g flour was weighed and added to the tubes. Tubes with flours were weighed and the actual amount of flour added to each tube was measured by the machine. Syringes with 27 ml distilled water, 23 ml 50 % sucrose, 27 ml 5 % lactic acid or 26 ml 5 % sodium carbonate solution and tubes with flour were put into the machine. The SRC-CHOPIN 2 then started a 65 minutes automated program which adds solvents to the specific tubes and shakes, rests and centrifuges before finally discarding the supernatant. The tubes with the gels were weighed by the operator and an SRC % for each solvent was calculated by the machine, Equation 1.

Equation 1. Used for calculating SRC Water, SRC Sucrose, SRC Lactic acid and SRC Sodium carbonate

$$100 \cdot \left(\frac{Gel\ weight}{Flour\ weight} \cdot \frac{86}{100 - Flour\ moisture\ content} - 1\right)$$

2.5. Alveolab Chopin

Alveolab from Chopin Technologies measures the resistance in a piece of dough, called a patty, when it is blown into a bubble. The measurement ends when the dough bubble bursts. 5 patties are inflated, and the average is calculated by the machine. The pressure measured in mmH_2O is plotted against the extensibility of the bubble measured in mm (Figure 1). Various parameters can be deduced from an alveograph and the most used are explained in Table 2. Blowing a dough bubble is a way to estimate how good the dough would be able to hold air in bubbles formed during fermentation.

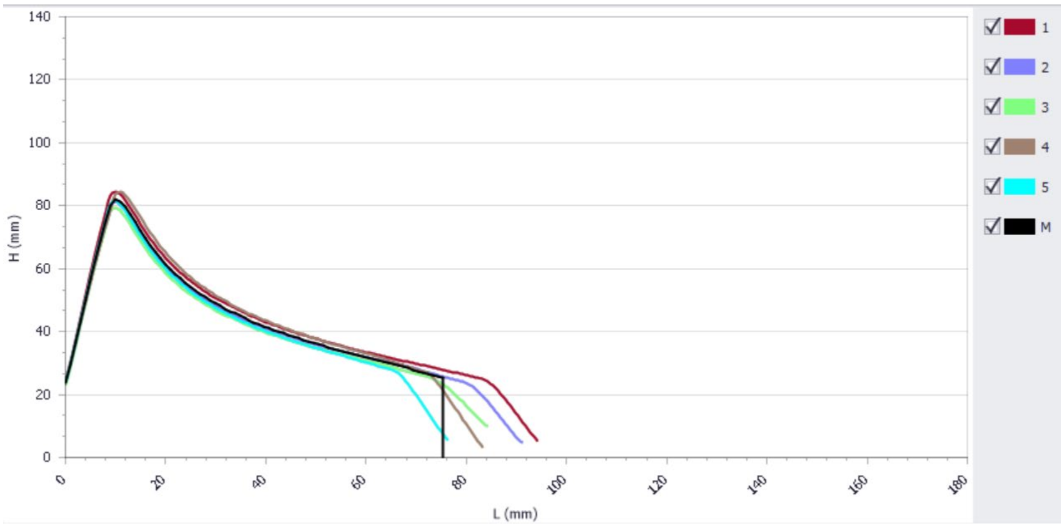

Figure 1. Example of an Alveolab graph, where the black line is the mean value of the five replicates measured (1-5).

Table 2. Description of Alveolab parameters

Parameters	Description	Related with
P	Maximum pressure	Tenacity of the dough
L	Bubble diameter upon bursting	Extensibility of the dough
P/L	P divided by L	Viscoelasticity of the dough
Ie	Pressure after 200 ml of air has been blown into the bubble divided by P (Ie = elasticity index)	The dough's tendency to retract after being stretched
W	Area under the curve	The strength of the flour and the energy needed to inflate the bubble

2.6. Reproducibility study

SDmatic, SRC-CHOPIN 2 and Alveolab data in this report were collected by two operators. 15 flours were tested in duplicates to test the operators' reproducibility. For each SDmatic, Alveolab and SRC parameter there is a Reproducibility Limit. The operators were reproducible for a parameter if the difference between them were lower than the Reproducibility limit. Time of training before the reproducibility study was roughly two weeks.

2.7. Statistical analysis

The multivariate software SIMCA® 16 was used for Principal Component Analysis, PCA, and Partial Least Square Regression, PLS.

3. Results

199 flours were analysed with SDmatic, SRC-CHOPIN 2 and Alveolab in this report. Other data presented was previously analysed and supplied by Lantmännen Cerealia except for the dietary fibre data analysed with a modified Uppsala-method (Andersson et al. 1999) by the Department of Molecular Sciences at the Swedish University of Agricultural Sciences, SLU, in Uppsala.

3.1. Overview of data

Table 3 gives an overview of the data with minimum, median and maximum values for some of the parameters for the flours.

Table 3. Parameters analysed in this report with minimum, median and maximum values shown as well as unit and the origin of the different analyses

Analysis	Performed by	Output	Min	Median	Max	Unit
Alveolab	Within this report	P	56	84	114	mmH_2O
		L	50	88	146	m
		W	135	227	346	10^{-4}J
		P/L	0.4	0.96	2.28	-dl
		Ie	37.9	50.65	60.5	% dl
SRC-CHOPIN 2	Within this report	SRC Lactic acid	110	126	147	% (dI)
		SRC Sodium carbonate	75.9	84.6	95.7	% (dI)
		SRC Sucrose	92.7	103	111	% (dI)
		SRC Water	58.7	65.7	74.0	% (dI)
SDmatic	Within this report	Ai	92.5	94.3	95.8	%
Ash	Lantmännen Cerealia	Ash content dm	0.48	0.59	0.72	%
Foss Infratec™	Lantmännen Cerealia	Protein 5,7 ts NIT	10.5	12.3	16.2	%

Falling number	Lantmännen Cerealia	Falling number	304	395	471	s
Amylograph	Lantmännen Cerealia	Amylogram max	774	1150	1580	s
		Amylogram gelatinisation temp	86.0	88.3	90.6	°C
Wet gluten	Lantmännen Cerealia	Gluten index	62.3	92.6	99.6	%
		Wet gluten dm	26.6	34.3	46.2	%
Farinograph	Lantmännen Cerealia	Farinogram water absorption	55.0	59.6	65.5	%
		Farinogram development time	1.2	3.0	4.7	min
		Farinogram Stability	3.2	6.4	17.5	min
		Farinogram degree of softening	36	79	121	BU
Test baking	Lantmännen Cerealia	Baking volume, spring wheats	2050	2460	2970	ml
		Baking volume, winter wheats	1650	1910	2200	ml
Uppsala-method	SLU	WU-AX	1.022	1.307	1.574	%
		WE-AX	0.528	0.729	1.064	%
		man insol	0.03	0.08	0.11	%

3.1.1. Principal Component Analysis

Principal component analysis (PCA) is an exploratory method used to summarise multidimensional data into two dimensions. It does so by fitting a line through the data that covers the most variance. This line is called principal component 1, PC 1. 90 degrees perpendicular to PC 1 another line, PC 2, can be drawn that covers the second most variance. By plotting PC 1 and PC 2 against each other the multidimensional data can be visualised in a two-dimensional plot. PCA can be used to see which observations in a dataset are alike and also explain why (Karamizadeh et al. 2013). Acore plot shows how observations relate to each other where each dot represents an observation. A loading plot shows how the variables relate and where each dot represents a variable.….

The score plot of the PCA for all flours and all parameters shows that the flours form two groups along PC 1 (Figure 2). One group is WW and Julius, which are winter wheats. The other group is SW, ES and BW and will be referred to as spring wheats, although BW is a mixture of winter wheat, spring wheat and German E-wheat.

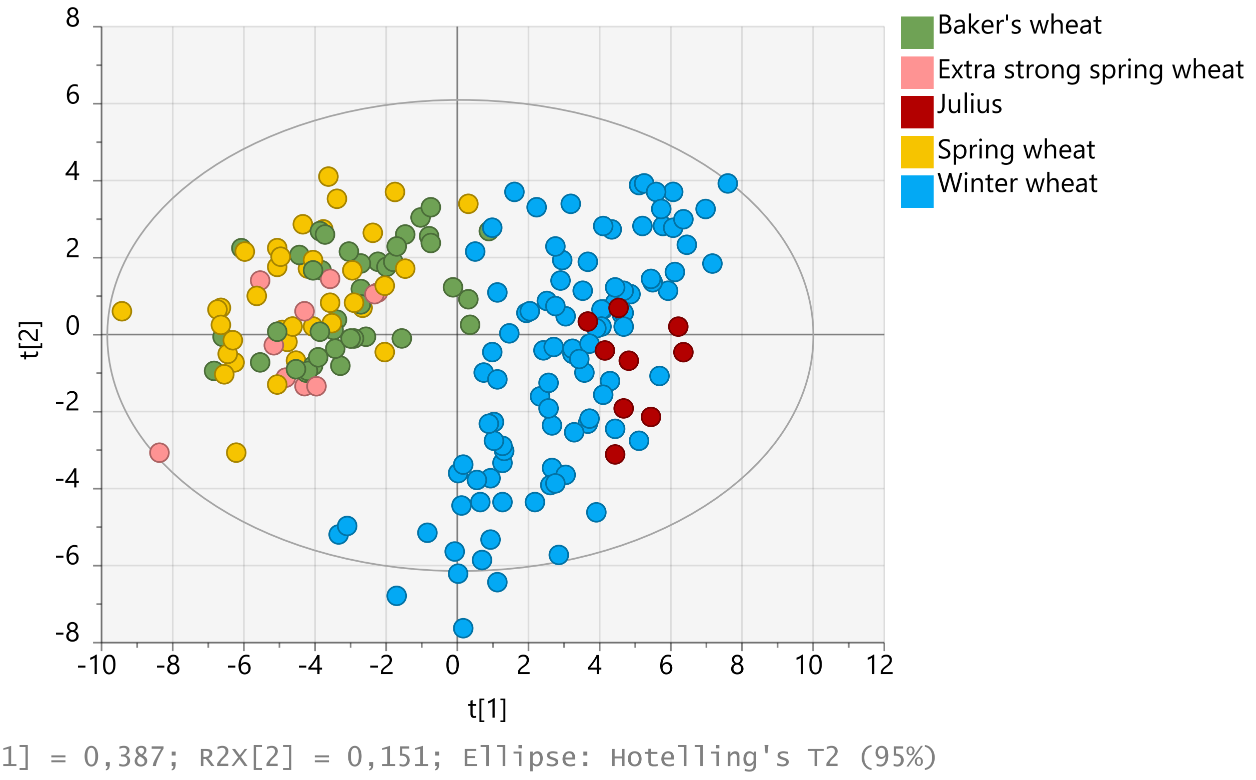

Figure 2. PCA score plot for 198 flours and all parameters in Table 2.

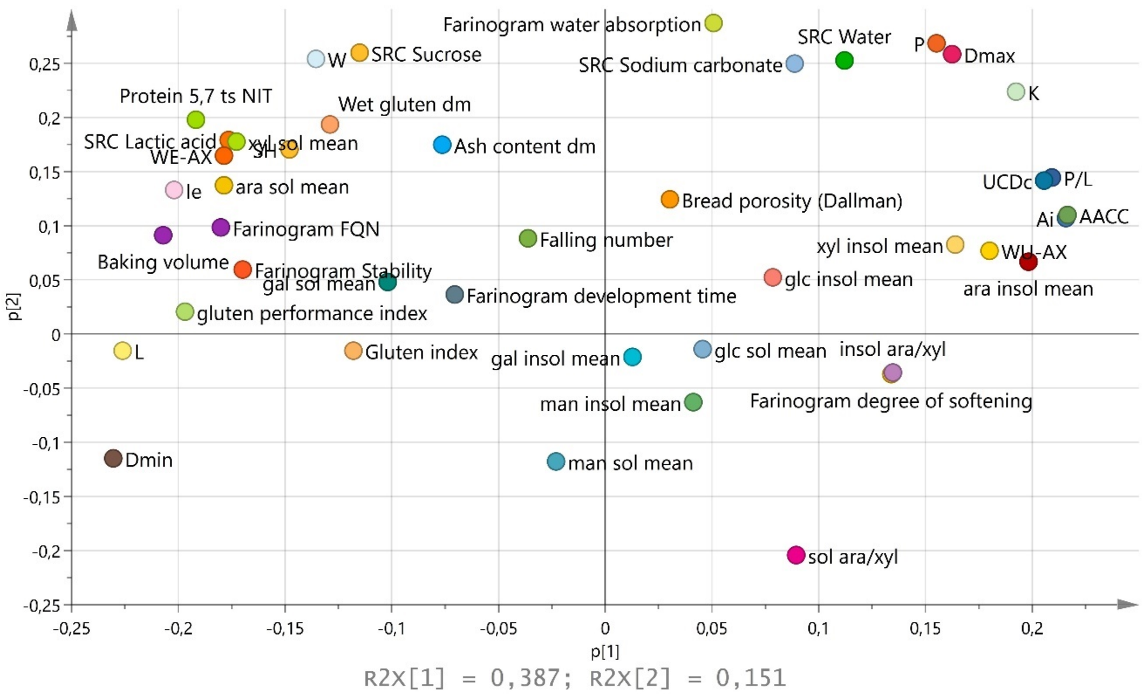

Figure 3. PCA loading plot for 198 flours and all parameters in Table 2.

3.2. Duplicate study

The two operators were reproducible for 100 % of the SDmatic parameters (results not shown). Reproducibility for Alveolab and SRC-CHOPIN 2 were roughly the same with 78 % and 80 % respectively. Alveolab and SRC-CHOPIN 2 involves a lot of manual work, so it is logical that the operators were the most reproducible for the SDmatic which only involves weighing of flour.

3.3. PLS-modelling of the dataset

Partial Least Squares, PLS, is a method used to predict response variables, Y, from a set of explanatory variables, X. Principal Component Analysis, PCA, fits principal components to the data to maximise their variance, but without taking the predictive power into account. Thus, PCA might neglect important parameters with high predictive power. PLS is instead focused on predictive power and the PLS-factors are drawn to maximise the predictive power (Pirouz 2006).

All PLS-models presented in this report, except for Figure 14, have baking volume as the Y-variable. The models are summarised and given a descriptive name in Table 4. All models were optimised for Q^2 by removal of X-variables.

Table 4. PLS-models summarised and given a descriptive name. Root mean square error of prediction, RMSEP, is only given for models used to predict unknown data. E.g. *SW80%* is built on 80% of the spring wheats and predicts the remaining 20 %. Variable Importance for the Projection, VIP, with parameters ordered from highest to lowest. n is the number of flours included in the model

Name	VIP	Flours	In the model	R^2Y	$R^2Y_{predicted}$	Q^2	RMSEP
***Allflours&years* (3 PLS-factors)**	Protein, Ai, Dmin, SRC Sucrose, insol ara/xyl, SRC sodium carbonate	All flours n=196	All years	0.853		0.841	
***WWallyears* (3 PLS-factors)**	Protein, fibre composition, gluten index, SH, W, farinograph, Ash, SRC Sodium carbonate	Winter wheats n=108	All years	0.658		0.563	

SWallyears **(2 PLS-factors)**	Ai, P/L, Dmin, K, P, Protein, Gluten index, SRC Water, Amylogram gelatinisation temp.	Spring wheats n=88	All years	0.432		0.376	
Allflours2018 **(1 PLS-factor)**	Protein, Ai, Dmin, insol ara/xyl, SRC Sucrose, SRC sodium carbonate	All flours n=86	2018	0.846	0.753	0.838	154 ml
WW2018 **(I PLS-factor)**	Wet gluten, Protein, Farinograph development time, UCDc, Farinograph	Winter wheats n=42	2018	0.555	0.103	0.517	169 ml
SW2018 **(1 PLS-factor)**	Protein, AACC, Ai, P/L, P, fibre composition, SRC water, Gluten index	Spring wheats n=44	2018	0.539	0.047	0.471	142 ml
Allflours80% **(4 PLS-factors)**	Protein, AACC, SRC Lactic acid, SRC Sucrose, insoluble ara/xyl, SH	All flours n=157	80 % of flours	0.863	0.856	0.848	122 ml
WW80% **(2 PLS-factors)**	Man insol , gluten index, SH, protein, Ash	Winter wheats n=85	80 % of flours	0.623	0.467	0.51	75 ml
SW80% **(1 PLS-factor)**	Ai, P/L, Protein, SRC Water, Wet gluten, Fibre composition	Spring wheats n=71	80 % of flours	0.455	0.323	0.361	124 ml

3.3.1. *Allflours&years*

Loadings for *Allflours&years* (a PLS-model with all flours from all years) can be seen in Figure 4. Notable is that parameters from all three Chopin instruments, the Uppsala method and the Foss Infratec^{TM} (near infra-red analysis) remain in the PLS after optimising the model.

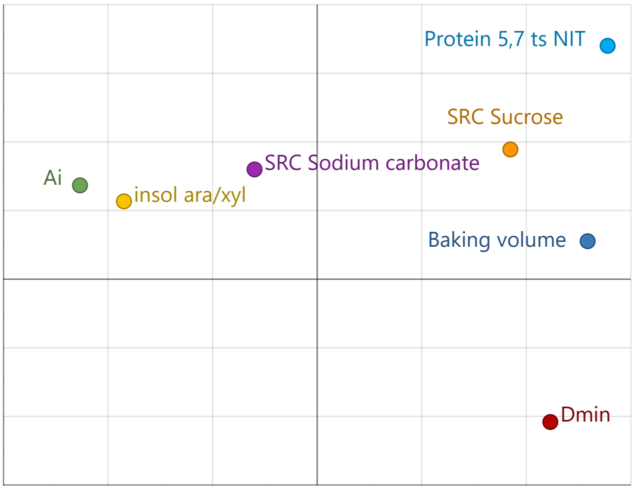

Figure 4. PLS-loadings for PLS model *Allflours&years*, which is based on the whole data set. Baking volume is the Y-variable.

The rigidity of the *Allflours&years* can be visualised by plotting observed, predicted and cross validation Y-predictions (Figure 5). As seen, predicted Y and cross validation Y-predictions are close to each other which mean that the model is rigid and removal of one observation wouldn't mean a large change of the model. The iodine absorption, Ai, should according to Figure 4 be negatively correlated with baking volume. This correlation can be studied by plotting baking volume against Ai (Figure 6). The correlation seems to be true for spring wheats but not for winter wheats. The strong negative correlation suggested by Figure 4 seems to originate out of the difference between the two groups of wheats.

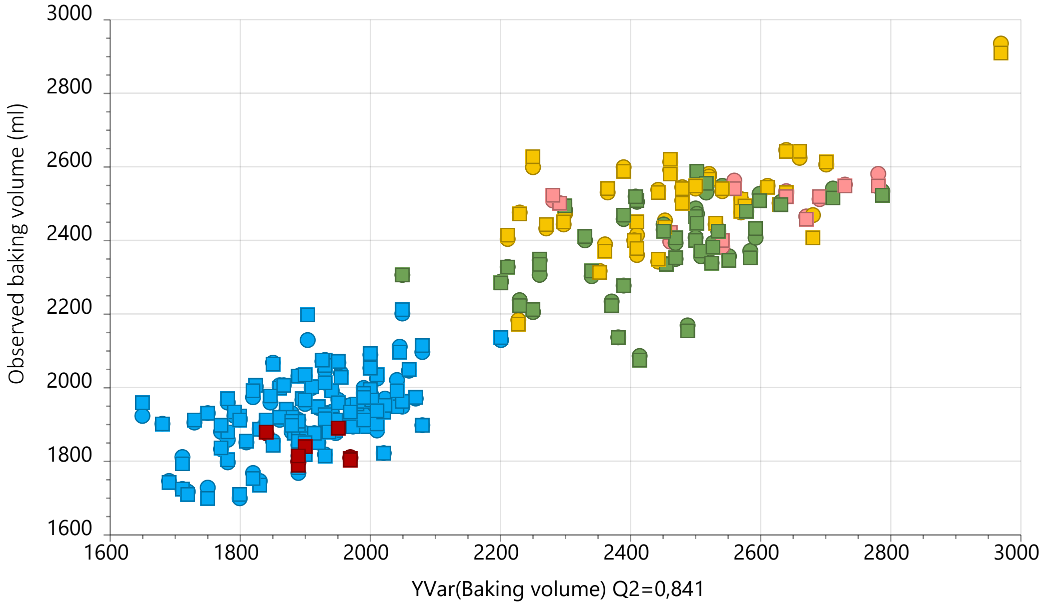

Figure 5. Observed Y plotted against predicted Y (circles), and cross validation Y-predictions (squares), (on the whole data set). BW-green, ES-pink, Julius-red, SW-yellow, WW-blue.

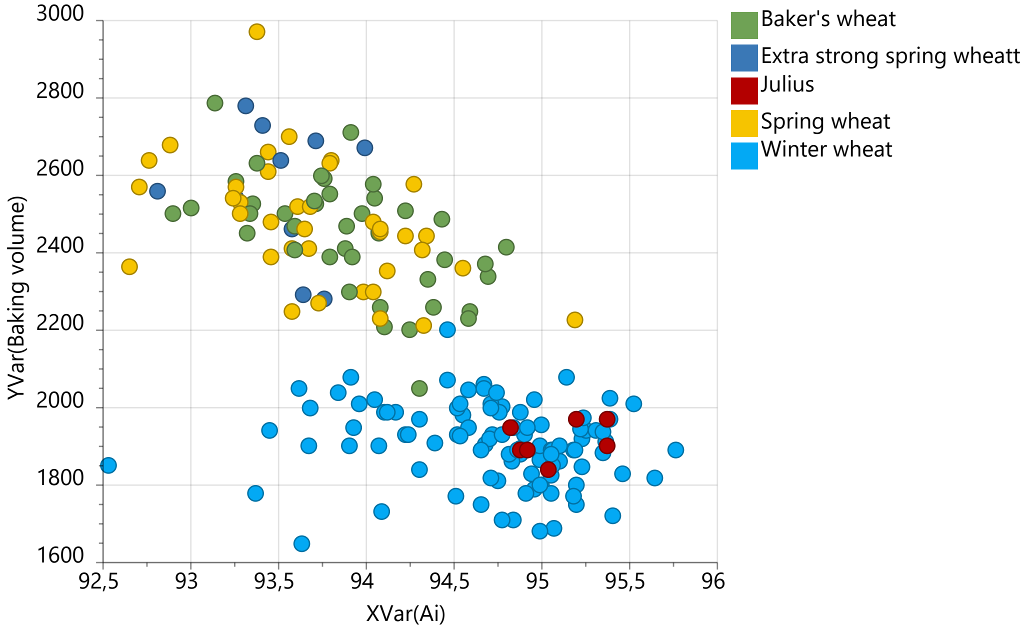

Figure 6. Baking volume as a function of iodine absorption, Ai, for all flours.

3.3.2. *WWallyears*

Since the flours formed two groups, winter wheats and spring wheats, in the PCA (Figure 2) one PLS for each group was made. Figure 7 shows the PLS-loadings for *WWallyears* (a PLS-model with WW from all years). Less variables could be removed this time, compared with *Allflours&years*, and with Q2=0.563 the model was less good at predicting the baking volume. The model was also less rigid, as seen in Figure 8 where the predicted Y variables are further off the cross validation Y-predictions. No parameters from SDmatic were left after optimising *WWallyears* for Q2 (Figure 7).

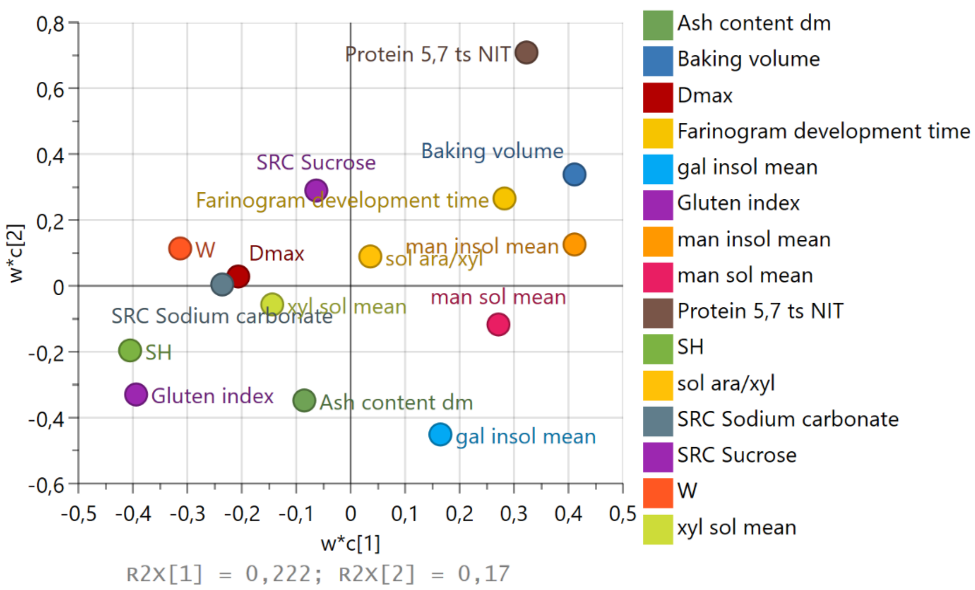

Figure 7. PLS-loadings for *WWallyears* (a model of WW and Julius after removal of X-variables to optimise Q^2). Baking volume is the Y-variable.

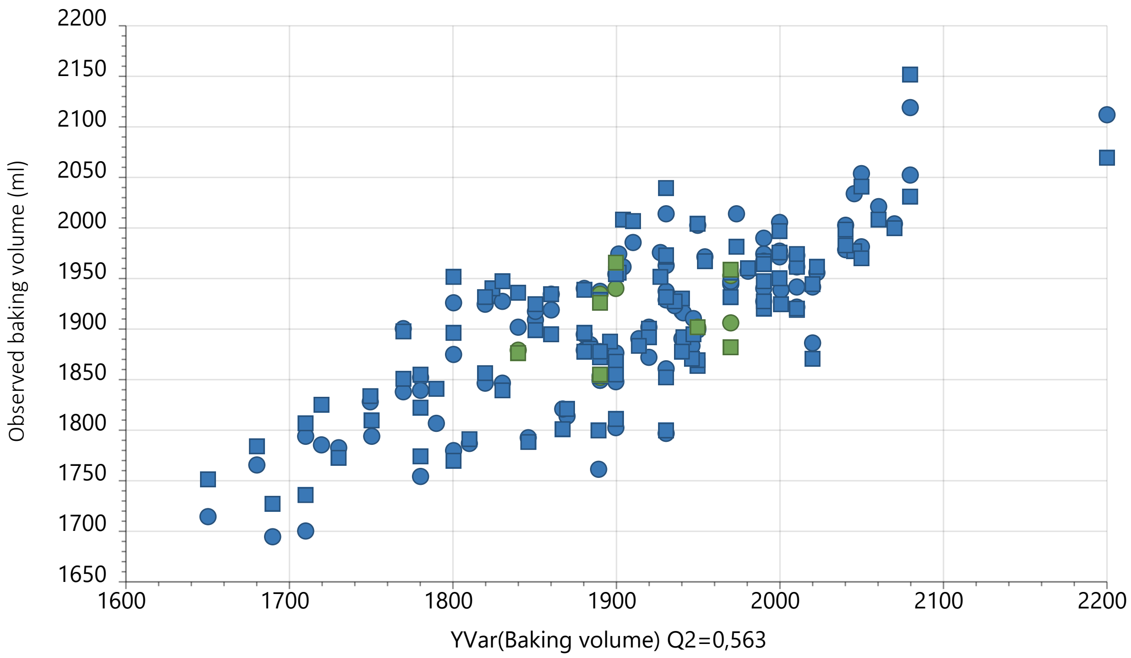

Figure 8. Observed Y plotted against predicted Y (circles), and cross validation Y-prediction (squares), (WW plus Julius). Julius-green, and WW-blue. Circles are predicted Y values and squares are predicted Y values from the cross validation.

3.3.3. *SWallyears*

SWallyears (a PLS-model built on SW from all years) had $R^2Y=0.432$ and $Q^2=0.376$ which is worse than both *Allflours&years* and *WWallyears*, Table 4. Even if this model is not good at predicting baking volume, it shows that different parameters explain baking volume for winter wheats and spring wheats (Figure 9). Baking volume for spring wheats seems to be positively correlated with protein content and Dmin (from Alveolab), and negatively correlated with Ai% and SRC water. Baking volume for winter wheats, in *WWallyears*, were positively correlated with insoluble mannose and protein content and negatively correlated with gluten index (Figure 7).

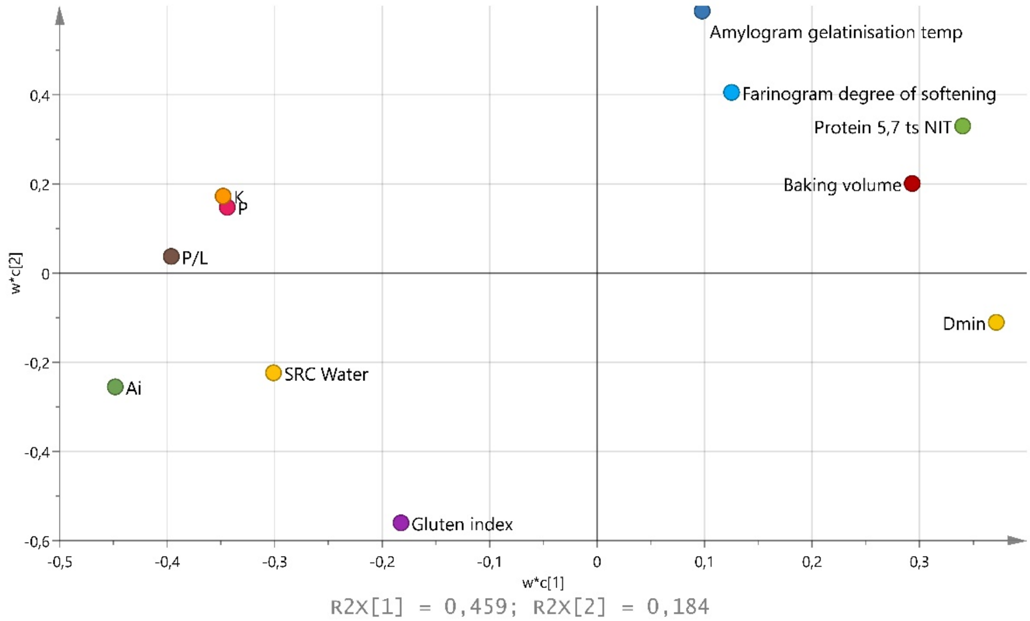

Figure 9. PLS-loadings for *SWallyears* (SW, ES and BW for all years).

3.4. Predicting baking volume

Baking volume is an important quality parameter with regards to consumer preferences, economy and industrial application. Being able to predict baking volume would therefore be desirable. A test set may be used to test a PLS-model's predictiveness. The model is built on a training set and then has to predict the baking volume for the test set. *Allflours&years, WWallyears and SWallyears* have no test sets since all observations were included in the models. Instead new models (*Allflours80%, WW80%* and *SW80%*) were made where 20 % of the observations were left out and used as test sets. To see how well models built on only data from 2018 could predict the harvest of 2019, another set of models were made (*Allflours2018, WW2018* and *SW2018*).

3.4.1. Predictions based on all flours

Allflours80% has a Root Mean Square Error of Prediction, RMSEP=122 ml which means that the average prediction deviated with 122 ml. With $R^2Y_{predicted}$=0.856 the model could fit the predicted data well, especially well for the winter wheats (Figure 10). *Allflours2018* has a RMSEP=154 ml. Baking volume for the whole dataset spanned from 1650 ml to 2970 ml. A RMSEP=154 ml seems acceptable at this point. But since each group, winter wheats and spring wheats, had a span of approximately 300 ml in baking volume it means that this model might predict an average sample as one of the highest or lowest, or vice versa. *Allflours2018* could therefore not be used to predict baking volume. The model could however fit the data well and had a $R^2Y_{predicted}$=0.753. A possible explanation why *Allflours80%* and *Allflours2018* could fit the data well is that the predicted data has a large span in baking volume, and that makes it easier to fit a line through the data with a high $R^2Y_{predicted}$-value.

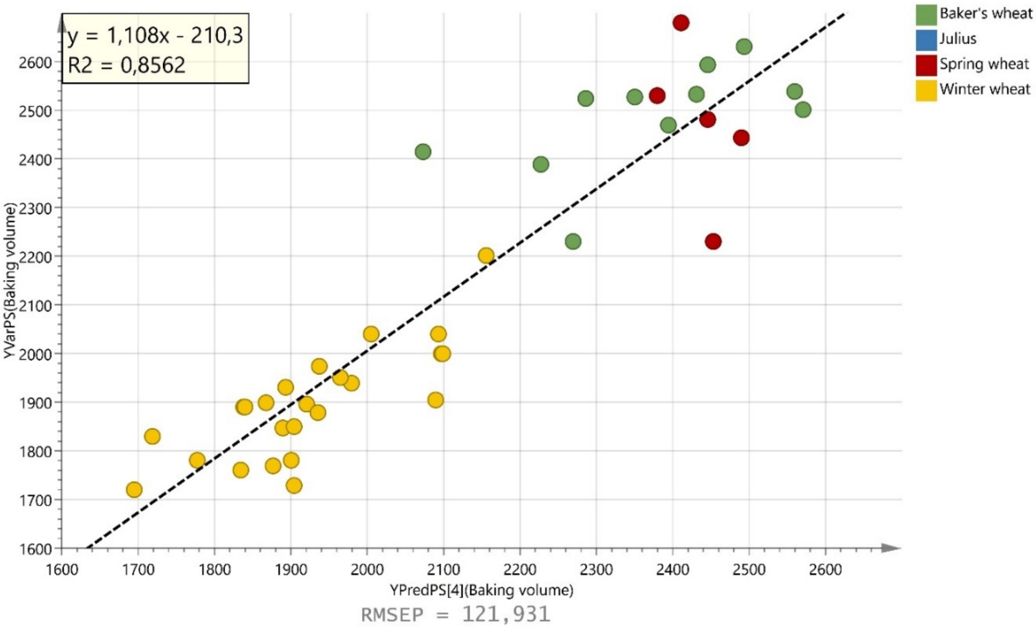

Figure 10. PLS-model based on 80 % of all flours, *Allflours80%*, and predicting the remaining 20 %. Included observations were selected at random and the model were optimised for Q2.

3.4.2. Predictions made on winter wheats

WW80% (RMSEP=75 ml and $R^2Y_{predicted}$=0.467) is quite good at predicting the baking volume (Figure 11). The RMSEP is the lowest for all models. This can be compared with the RMSEP=169 ml for *WW2018*. The baking volume for WW is roughly between 1800 and 2150 ml and a RMSEP of 169 ml means therefore that the model can't predict the next year's harvest. Even the $R^2Y_{predicted}$ =0.103 for *WW2018* is low. Important to remember is that WW80% is based on 85 flours which makes it better to predict baking volume than WW2018 does, which is based on only 42 flours. The same parameters were not selected for *WW80%* and *WW2018* (Table 4). This might be due to that different parameters were important for the baking volume for the two years and thus explain the poor estimation of 2019 year's harvest.

The same parameters were not selected for *WW80%* and *Allflours80%*, and more parameters were included in *WW80%* (Figure 12). This indicates that different parameters can explain the baking volume for winter wheats and spring wheats.

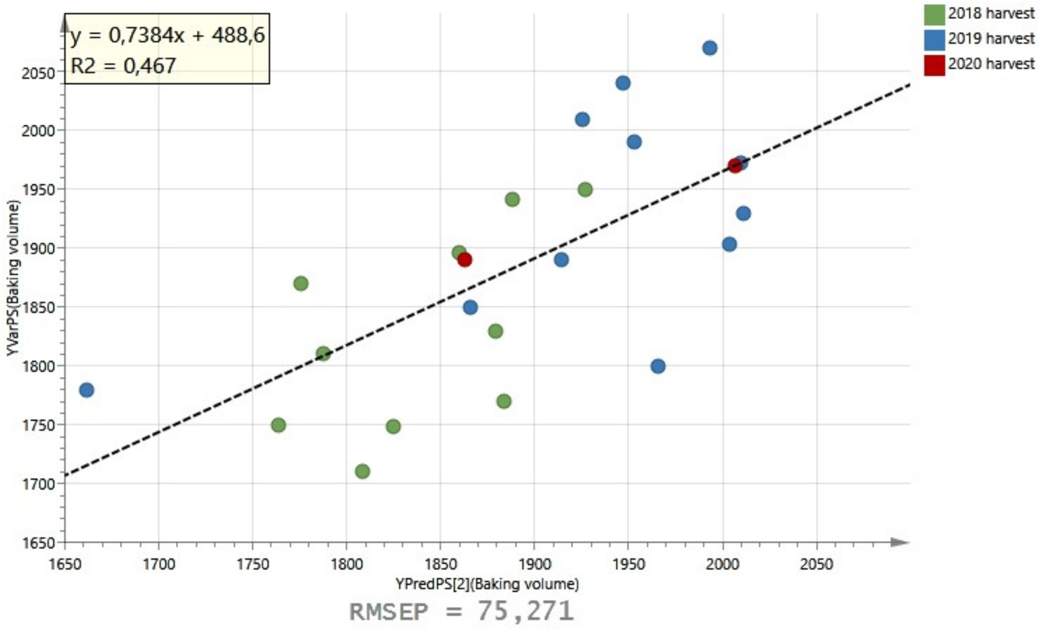

Figure 11. PLS-model based on 80 % of the winter wheats, *WW80%*, and predicting the remaining 20 %. Included observations were selected at random and the model were optimised for Q2.

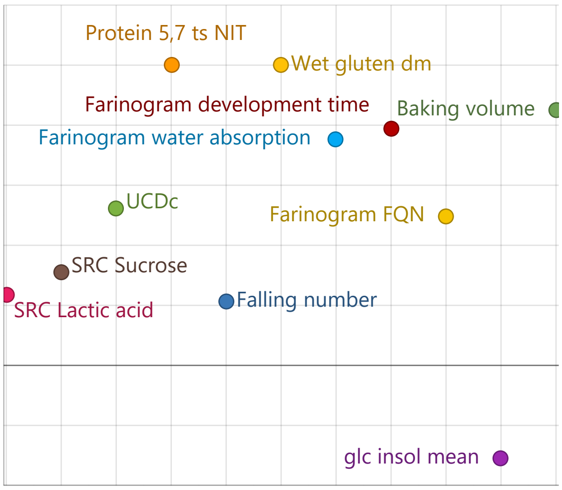

Figure 12. PLS-loadings for a model of WW of 2018 years harvest, *WW2018*, after removal of X-variables to optimise Q^2. Baking volume is the Y-variable.

UCDc is positively correlated with baking volume in *WW2018* (Figure 12). This is not the case for the loadings-plot for the whole dataset (Figure 3). A low correlation, although positive for UCDc and baking volume for *WW2018* may be seen in Figure 13.

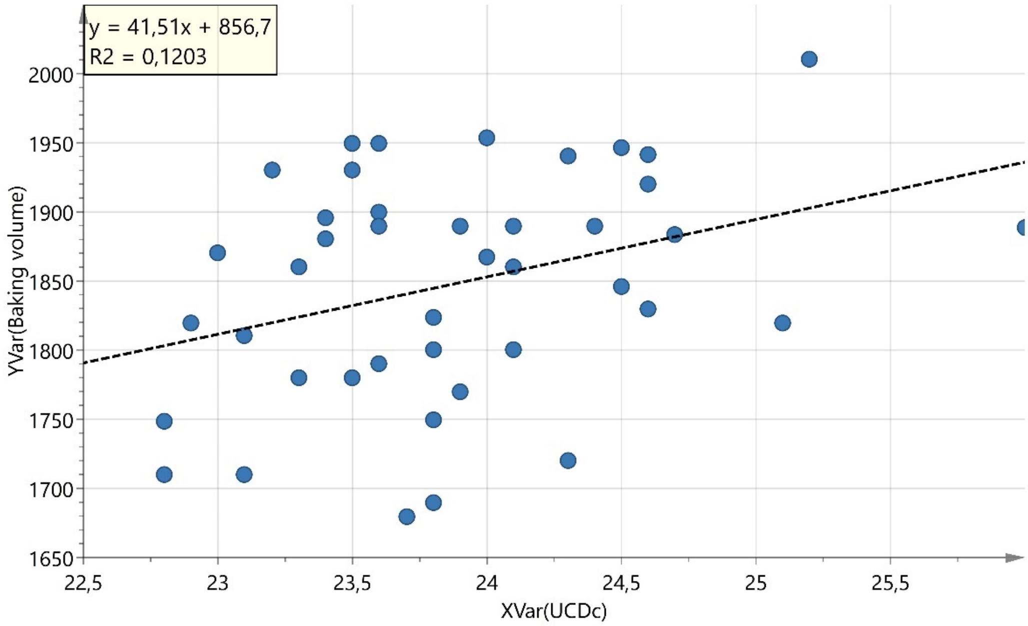

Figure 13. UCDc for *WW2018* plotted against baking volume.

3.4.3. Predictions made on spring wheats

SW80% (RMSEP=124 ml and $R^2Y_{predicted}$=0.323) has a higher prediction ability than *SW2018* (RMSEP=142 ml and $R^2Y_{predicted}$=0.047). This difference in predictiveness might be since the models are built on a different number of observations, 71 and 44 for SW80% and SW2018 respectively. And therefore, the 80%-model outperformed the 2018-model, as also seen for the WW-models.

Some different variables were selected for the two models, although the most important variables were similar (Table 4). The variables with highest VIP-values for *SW2018* (Protein, AACC, Ai and P/L) were similar to the ones for *SW80%* (Ai, Protein and P/L). Ai and AACC are both measurements of damaged starch. This suggests that the baking volume for spring wheats from 2018 and 2019 can be explained with roughly the same variables.

3.5. Predicting damaged starch content

Damage starch correlates with other tested parameters (Figure 3). Since SDmatic only tests for damage starch it could be left out if one could predict the content of damage starch. Figure 14 shows an attempt to make a PLS-model with AACC from

SDmatic as the Y-variable. AACC is an estimate of the AACC76-31 method. The model is based on 80 % of the whole dataset and predicts the remaining 20 %.

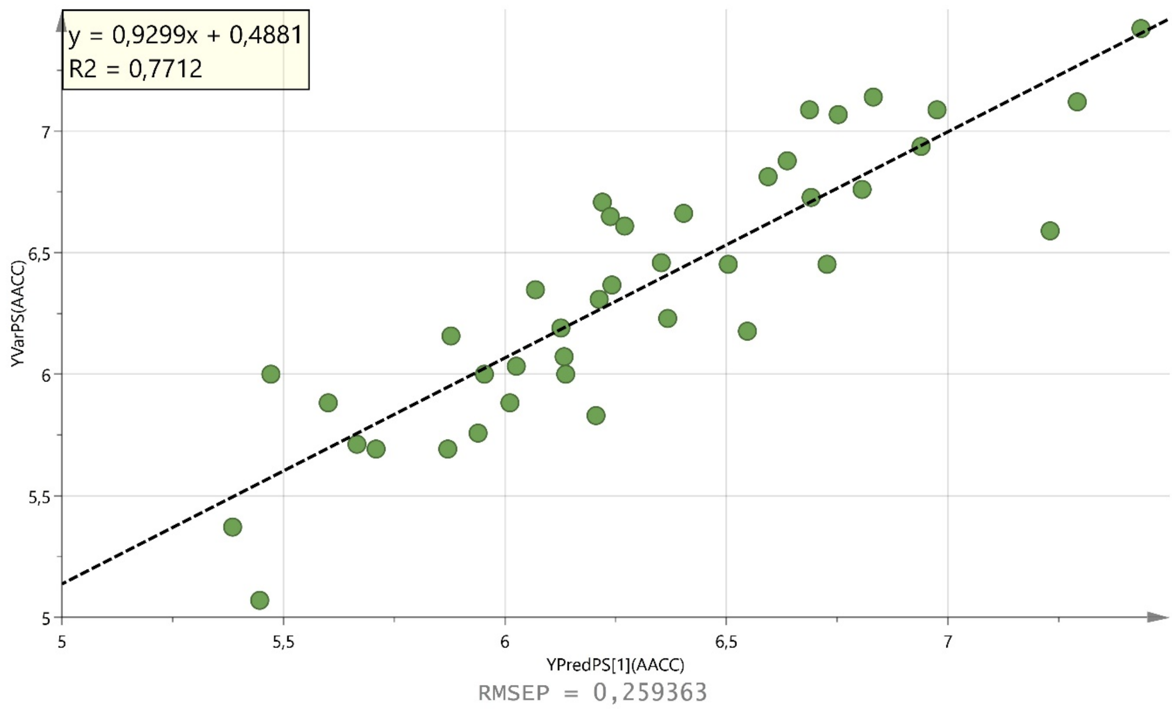

Figure 14. PLS-model based on 80% of the whole dataset, predicting the remaining 20 %. AACC is set as the Y-variable.

4. Discussion

The flours formed two clear groups, winter wheats and spring wheats, in the PCA in Figure 2. Based on this it was logical to make individual PLS-models for the two groups since they probably were explained by different variables. Models built on 80 % of the flours or flours from only 2018 were built for winter wheats, spring wheats and all flours together. This gave several PLS-models based on varying numbers of flours. PLS-models based on approximately the same number of flours will be compared first. The discussion will then continue with PLS-models based on winter wheats and then spring wheats.

4.1. PLS-models with based on approximately 86 flours

SWallyears (n=88)*, Allflours2018* (n=86) and *WW80%* (n=85) are based on approximately the same number of flours. This makes it easier to compare these models. *SWallyears* has an R^2=0.432 and Q^2=0.376. These values are low and means that the model does not fit the data well and that the model has a low predictiveness. *SWallyears* was not used to predict any test set and has therefore no $R^2Y_{predicted}$.

Allflours2018 can fit its predicted data quite well ($R^2Y_{predicted}$=0.753). The span of baking volume for winter wheats and spring wheats are about 250 ml and 350 ml respectively, Figure 5. Since *Allflours2018* has an RMSEP=154 ml it means that it can’t predict the baking volume. *WW80%* (RMSEP=75 ml and $R^2Y_{predicted}$=0.467) has a worse fit but a more acceptable RMSEP. The $R^2Y_{predicted}$-value alone might thus not be enough to decide if a model is good or not. The baking volume varied more for the spring wheats than the winter wheats, Table 4 and Figure 5. This makes it even more important to have low RMSEP for a model that will predict winter wheats. Because a model won’t be able to predict the baking volume if the RMSEP covers the whole span in baking volume.

4.2. PLS-models for all flours

The highest $R^2Y_{predicted}$=0.863 was found in the PLS-model *Allflours80%*. An explanation for this might be the broader span of baking volume which makes it easier to fit a line to the data and therefore gives a higher $R^2Y_{predicted}$.. The RMSEP=122 ml is more explanatory, which is the third highest for all the PLS-models. *Allflours80%* (RMSEP=122 and $R^2Y_{predicted}$=0.863) is better than *Allflours2018* (RMSEP=154 ml and $R^2Y_{predicted}$=0.753). There might be two reasons for this. One is that *Allflours80%* (157 flours) were built on more flours than *Allflours2018* (86 flours). The other is that the model included flours from all years and thus the parameters should be relevant for the predicted flours.

4.3. PLS-models for winter wheats

WW80% (RMSEP=75 ml) had a lower RMSEP than *WW2018* (RMSEP=169 ml). This might have two reasons. The reasons are the same as for the Allflours-models; that *WW80%* are based on more flours and flours from all years which mean that the parameters should be relevant for the predicted flours. It is however interesting that *WW2018* had the highest RMSEP=169 ml. This indicates that all of this year's parameters were not applicable for the other two years' flours.

Both *WW2018* and *WW80%* had protein content as an important parameter. That protein content is positively correlated with baking volume was also seen by Bockstaele *et al.* (2008). *WW2018* was the only of these two models with a parameter for damaged starch, namely UCDc. Damaged starch might increase late stage fermentation (Delcour & Hoseney 2010) and this could increase the baking volume. Baking volume and UCDc is negatively correlated in the loadings-plot for the whole data set, Figure 3. A low positive correlation between baking volume and UCDc for *WW2018* was seen in Figure 13. No positive correlation could be seen when Ai was plotted against baking volume for the whole data set (Figure 6). This low positive correlation is probably of low importance.

4.4. PLS-models for spring wheats

The *SW2018* predicted 2019 harvest year quite badly (RMSEP=142 ml) and fitted the data poorly ($R^2Y_{predicted}$=0.047). The RMSEP is still lower than *WW2018* (RMSEP=169 ml) but in combination with the low $R^2Y_{predicted}$ makes this model unable to use for prediction of baking volume. One thing that *SW2018* shows are the parameters important for spring wheats, and that they differ from winter wheats.

SW80% had $R^2Y_{predicted}$=0.323 which was higher than *SW2018*'s ($R^2Y_{predicted}$=0.047) but is still lower compared to *WW80%* ($R^2Y_{predicted}$=0.467). The low $R^2Y_{predicted}$-values for the spring wheat models indicate that the included parameters might not explain the variation in the data. This is interesting since the Swedish spring wheat is more similar to wheats grown in southern European countries, than the Swedish winter wheat is. The Alveolab is widely used in these countries to evaluate wheat performance (Lásztity & Salgó 2002). It is therefore interesting with the low predictiveness of the spring wheat models since machines like Alveolab are used for evaluation of wheats alike the Swedish spring wheat.

4.5. Damaged starch and protein content

Both *SW2018* and *SW80%* have damaged starch and protein content as important parameters as well as P/L. P is related to the dough's stiffness which increases with increased content of damaged starch since more damaged starch means a higher water absorption of the flour. P were close to UCDc, Ai and AACC (all parameters from SDmatic) in PC1 in the PCA (Figure 3) which shows that they correlated. L is on the other side of PC1 and is instead more positively correlated with protein content. P/L relates to both protein content and damaged starch and this is an explanation for why it was important for the PLS-models of spring wheats. The correlation between damaged starch and P was also noticed by Huen et al. (2018). Figure 3 shows a correlation between damaged starch and Farinograph water absorption as observed in other litterature (Bockstaele et al. 2008). Tipples *et al.* (1978) suggested that Farinograph water absorption could be predicted based on damaged starch and protein content. One could flip this argument and speculate that Farinograph data in combination with protein content could estimate damage starch content.

Figure 14 shows an attempt to estimate damaged starch (AACC from SDmatic). AACC is an estimate of the enzymatic method, AACC76-31, for determining the content of damage starch. The AACC spanned between 5 and 7 % damaged starch for the whole dataset. With RMSEP=0.26 %-units and $R^2Y_{predicted}$=0.77, the model might predict damaged starch quite well. Parameters included in the model came from the Alveolab, SRC-CHOPIN 2 and Foss Infratec™. One might thus exclude SDmatic from the set of analyses and instead estimate it. Parameters from SDmatic were however important when creating the PLS-models for predicting baking volume. Excluding SDmatic would therefore make it harder to predict baking volume.

Figure 6 shows a clear negative correlation between iodine absorption and baking volume for spring wheats. This relationship between damaged starch and baking volume was also noticed by Barrera *et al.* (2007). They have two theories to explain this. Either that too much damaged starch binds water and hinder optimum gluten formation or that too much damaged starch gives a high initial water absorption followed by a loss in viscosity due to enzymatic hydrolysis of the damaged starch. The test baking procedure in this study was adapted for water absorption of the flour so it is more likely that the later reason is applicable for this study. Since spring wheat flours have high protein contents the effect of damaged starch might be greater than for winter wheats. That winter wheats have a higher damaged starch content than spring wheats is also seen in Figure 6. Winter wheat has a 1-3 % higher exchange than spring wheat, which is a plausible explanation for the observed difference in damaged starch content.

4.6. Dietary fibre composition

Some parameters from the Uppsala-method were included in 7 out of the 9 PLS-models, Table 4. Insoluble mannose (determined by the Uppsala-method) was a parameter with high VIP-values for *WW80%* and *WWallyears*, and also positively correlated with baking volume. Insoluble mannose originates from glucomannan in the cell walls of the starchy endosperm. Glucomannan is a hemicellulose with emulsifying capacity. Li *et al.* (2020) showed that addition of konjac glucomannan improved the stability of gluten proteins in wheat dough. This could be an explanation for the positive correlation with baking volume.

WU-AX has been linked with lower baking volume (Courtin & Delcour 2002) and WE-AX with higher baking volume (Goesaert et al. 2005). WU-AX or WE-AX were however not included in any PLS-model predicting the baking volume in this report. It indicates that these were not useful parameters when predicting the baking volume for the samples in this study.

Insol ara/xyl (the quota of arabinose-residues and xylose-residues in WU-AX) was negatively correlated with baking volume for all PLS-models based on all flours. In other words, a higher degree of substitution for WU-AX were found to be negatively correlated with the baking volume.

Less than one percent of wheat flour consist of arabinogalactan. It has however an effect on baking volume and has been seen to increase it when dough was fortified with extra arabinogalactan (Saeed et al. 2015). Insoluble galactose, originating from insoluble arabinogalactan, was correlated with a lower baking volume for *WWallyears* and with a higher baking volume for *SW2018.* Insoluble galactose had a VIP>1 for *WWallyears* and this correlation would be interesting to study further.

5. Conclusion

PLS-modelling of wheat flours can be used to some extent to predict baking volume. The most promising model in this study was *WW80%*. Separate PLS-models for winter wheat and spring wheat should be made since they are explained by different parameters. Protein content were important for all models and damage starch were particularly important when explaining the spring wheat models. Although glucomannan constitutes less than 1 % of wheat flour, it was an important parameter for *WW80%* and *WWallyears.*

The duplicate study showed that the two operators were reproducible for SDmatic, Alveolab and SRC-CHOPIN 2, although only completely reproducible for SDmatic. The fact that they had limited training time is a good indication for the user-friendliness of the machines.

Some parameters from either SDmatic, SRC-CHOPIN 2 or Alveolab were left in all PLS-models after optimising Q2. *WW80%* (the most promising model) had only one variable from Alveolab and no other parameters from SDmatic or SRC-CHOPIN 2. How good these machines are at predicting baking volume for bread baked on Swedish wheat flour is therefore still unclear.

Future studies should analyse how these machines can predict other quality parameters, such as crumb structure, bread-staling and consumer acceptability.

References

Andersson, A.A.M., Merker, A., Nilsson, P., Sørensen, H. & Åman, P. (1999). Chemical composition of the potential new oilseed crops Barbarea vulgaris, Barbarea verna and Lepidium campestre. *Journal of the Science of Food and Agriculture*, 79 (2), 179–186. https://doi.org/10.1002/(SICI)1097-0010(199902)79:2<179::AID-JSFA163>3.0.CO;2-N

Arya, S., Prajakta, S. & Ashish, W. (2015). Importance of damaged starch in bakery products - a review. 2015. https://www.researchgate.net/publication/277634923_Importance_of_damaged_starch_in_bakery_products_-_a_review [2020-09-13]

Barrera, G.N., Pérez, G.T., Ribotta, P.D. & León, A.E. (2007). Influence of damaged starch on cookie and bread-making quality. *European Food Research and Technology*, 225 (1), 1–7. https://doi.org/10.1007/s00217-006-0374-1

Bockstaele, F.V., Leyn, I.D., Eeckhout, M. & Dewettinck, K. (2008). Rheological Properties of Wheat Flour Dough and the Relationship with Bread Volume. I. Creep-Recovery Measurements. *Cereal Chemistry*, 85 (6), 753–761. https://doi.org/10.1094/CCHEM-85-6-0753

CHOPIN Technologies 2019*SDmatic User's manual 01/2019*.

Courtin, C.M. & Delcour, J.A. (2002). Arabinoxylans and endoxylanases in wheat flour bread-making. *Journal of Cereal Science*, 35 (3), 225–243. https://doi.org/10.1006/jcrs.2001.0433

Courtin, C.M., Roelants, A. & Delcour, J.A. (1999). Fractionation-reconstitution experiments provide insight into the role of endoxylanases in bread-making. *Journal of Agricultural and Food Chemistry*, 47 (5), 1870–1877. https://doi.org/10.1021/jf981178w

Delcour, J., Vanhamel, S. & Hoseney, R. (1991). Physicochemical and Functional-Properties of Rye Nonstarch Polysaccharides .2. Impact of a Fraction Containing Water-Soluble Pentosans and Proteins on Gluten-Starch Loaf Volumes. *Cereal Chemistry*, 68 (1), 72–76

Delcour, J.A. & Hoseney, R.C. (2010). *Principles of Cereal Science and Technology*. Third edition. St. Paul, Minn. : AACC International, [2010].

Dubat, A., Berra, M. & Baik, B.-K. (2019). Collaborative Study Report: Automated Measurement of Wheat Flour Solvent Retention Capacity with the CHOPIN-SRC Instrument (AACCI Approved Method 56-15.01). *Cereal Foods World*, 64 (3), 0033. https://doi.org/10.1094/CFW-64-3-0033

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. https://doi.org/10.1006/jcrs.1995.0025

Goesaert, H., Brijs, K., Veraverbeke, W.S., Courtin, C.M., Gebruers, K. & Delcour, J.A. (2005). Wheat flour constituents: how they impact bread quality, and how to impact their functionality. *Trends in Food Science & Technology*, 16 (1), 12–30. https://doi.org/10.1016/j.tifs.2004.02.011

Huen, J., Boersmann, J., Matullat, I., Boehm, L., Stukenborg, F., Heitmann, M., Zannini, E. & Arendt, E.K. (2018). Wheat flour quality evaluation from the baker's perspective: comparative assessment of 18 analytical methods.

European Food Research and Technology, 244 (3), 535–545. https://doi.org/10.1007/s00217-017-2974-3

Karamizadeh, S., Abdullah, S.M., Manaf, A.A., Zamani, M. & Hooman, A. (2013). An Overview of Principal Component Analysis. 3

Kaur, A., Yadav, M.P., Singh, B., Bhinder, S., Simon, S. & Singh, N. (2019). Isolation and characterization of arabinoxylans from wheat bran and study of their contribution to wheat flour dough rheology. *Carbohydrate Polymers*, 221, 166–173. https://doi.org/10.1016/j.carbpol.2019.06.002

Kweon, M., Slade, L. & Levine, H. (2011). Solvent Retention Capacity (SRC) Testing of Wheat Flour: Principles and Value in Predicting Flour Functionality in Different Wheat-Based Food Processes and in Wheat Breeding—A Review. *Cereal Chemistry*, 88 (6), 537–552. https://doi.org/10.1094/CCHEM-07-11-0092

Lásztity, R. & Salgó, A. (2002). QUALITY ASSURANCE OF CEREALS - PAST, PRESENT, FUTURE. *Periodica Polytechnica Chemical Engineering*, 46 (1–2), 5–13. https://doi.org/N/A

Li, S., Qu, Z., Feng, J. & Chen, Y. (2020). Improved physicochemical and structural properties of wheat gluten with konjac glucomannan. *Journal of Cereal Science*, 95, 103050. https://doi.org/10.1016/j.jcs.2020.103050

Ma, S., Wang, C., Li, L. & Wang, X. (2020). Effects of particle size on the quality attributes of wheat flour made by the milling process. *Cereal Chemistry*, 97 (2), 172–182. https://doi.org/10.1002/cche.10230

McAllister, J.T., Walker, C.K. & Panozzo, J.F. (2011). Influence of Starch Composition on Starch Damage Values Determined by Megazyme and SDmatic Methods. *Cereal Chemistry*, 88 (4), 349–351. https://doi.org/10.1094/CCHEM-10-10-0142

Medcalf, D. & Gilles, K. (1965). Determination of Starch Damage by Rate of Iodine Absorption. *Cereal Chemistry*, 42 (6), 546-

Mendis, M., Ohm, J.-B., Delcour, J.A., Gebruers, K., Meinhardt, S. & Simsek, S. (2013). Variability in Arabinoxylan, Xylanase Activity, and Xylanase Inhibitor Levels in Hard Spring Wheat. *Cereal Chemistry*, 90 (3), 240–248. https://doi.org/10.1094/CCHEM-08-12-0103-R

Morrison, W.R. & Tester, R.F. (1994). Properties of Damaged Starch Granules. IV. Composition of Ball-milled Wheat Starches and of Fractions Obtained on Hydration. *Journal of Cereal Science*, 20 (1), 69–77. https://doi.org/10.1006/jcrs.1994.1046

Pirouz, D.M. (2006). An Overview of Partial Least Squares. *SSRN Electronic Journal*,. https://doi.org/10.2139/ssrn.1631359

Saeed, F., Arshad, M.U., Pasha, I., Suleria, H. a. R., Arshad, M.S., Qamar, A., Ullah, A. & Sultan, S. (2015). Effect of Arabinoxylan and Arabinogalactan on Textural Attributes of Bread. *Journal of Food Processing and Preservation*, 39 (6), 1070–1088. https://doi.org/10.1111/jfpp.12322

Swedish Board of Agriculture (2020). *Jordbruksmarkens användning 2020. Preliminär statistik.* [text]. https://jordbruksverket.se/om-jordbruksverket/jordbruksverkets-officiella-statistik/jordbruksverkets-statistikrapporter/statistik/2020-06-22-jordbruksmarkens-anvandning-2020.--preliminar-statistik [2020-11-15]

Tipples, K., Meredith, J. & Holas, J. (1978). Factors Affecting Farinograph and Baking Absorption .2. Relative Influence of Flour Components. *Cereal Chemistry*, 55 (5), 652–660

Wang, L. & Flores, R.A. (2000). Effects of Flour Particle Size on the Textural Properties of Flour Tortillas. *Journal of Cereal Science*, 31 (3), 263–272. https://doi.org/10.1006/jcrs.2000.0302

Yngveson, N. (2015). Proteinhalten ger en anvisning om avkast- ningsnivån i höstvete. 3

Acknowledgements

I would like to show my appreciation to the following:

- Lantmännen Cerealia for giving me this opportunity to dive into the world of wheat flour analysis.
- Louise Selga, industrial PhD student at SLU, for discussing theories and helping me in the writing process.
- Roger Andersson, professor at SLU, for your help with the PCA- and PLS-analyses

Appendix

Appendix 1. Equations for SDmatic units. All units are based on the iodine absorption, Ai. UCDc is adjusted by the moisture content and protein content of the flour.

$$UCD = -266.82 \cdot (1 - AI) + 36.962$$

$$UCDc = (100 \cdot UCD) \,/\, (126 - moisture\ content - protein\ content).$$

$$AACC\ 76\text{-}31.01 = 0.0533 \cdot (AI\%)^2 - 9.1482 \cdot (AI\%) + 394.94.$$

$$Farrand = 0.16100099 \cdot (Ai\%)^2 - 107.3818674 \cdot (Ai\%) + 4728.4965460$$

Appendix 2. Popular scientific summary.

Imagine never having to test bake. If you are a baker at a large bakery or a cereal scientist, you might be shocked over this drastic idea. If you are not, you are probably like most people. There is no way in hiding it that the following text is mostly aimed for people in the baking industry, but it might be even more insightful if you are not so stick along.

Let´s take it from the beginning. Bakeries test bake the flour as a quality measurement. The bread is baked under strict controlled circumstances and the final bread volume is measured. This process takes a lot of time and demands highly trained bakers. Time and money (salaries is one of the major outcomes in food companies) could be saved if we instead could predict the baking volume. This report tested if three machines from Chopin Technologies together with statistical models could do this.

The three machines from Chopin Technologies were used to test 200 flour samples. The machines were SDmatic, SRC-CHOPIN 2 and Alveolab. Wheat flour is made from wheat kernels and the kernels contain starch. Some starch is damaged during the milling process. SDmatic estimates the damaged starch content of flour. Damaged starch can act as feed for the yeast but too much will make the dough absorb too much water and become sticky. Where the sweet spot is depends on the product.

SRC-CHOPIN 2 tests how much gluten proteins and dietary fibres the flour has. Gluten proteins are important for baking volume. Fluffy white bread would not be possible without these proteins. Dietary fibre is healthy for us but might have negative impact on the baking volume. Yet again, the sweet spot depends on the product.

Lastly, the Alveolab. Have you ever blown a bubble with chewing gum? Have you ever tested how large of a bubble you could make? The Alveolab does that but with wheat dough. Larger bubbles that are easy to inflate are related to higher baking volume. These were the tests performed in this study but previous data on these flours were also available. If you are new to cereal science and thinks that three tests are enough then I will just mention that ash content, protein content, Faringraph data and fibre composition were previously analysed on these flours.

So how could we predict baking volume? We would use something called PLS, Partial Least Squares Regression. This a method to summarise several dimensions into two dimensions with a focus on predictive power. Does it sound a little bit complicated? Imagine that you draw a two-dimensional graph with baking volume on the Y-axis and protein content of the X-axis. Now let us add the damaged starch content on the Z-axis. We could add more and more axes but after three the plot is hard to read. PLS is there to help us. By setting baking volume as the Y variable and all the other test parameters as the X variables, PLS will try to see which parameters are important for explaining the baking volume. Unimportant parameters will be discarded and left out of the model.

This study analysed both winter wheat and spring wheat (named after the time of sowing). PLS-models for each group of wheat were made. The model for winter wheat could predict the baking volume with a difference of either plus or minus 75 ml compared to the observed baking volume. Since the average bread volume for winter wheats were 2030 ml it shows that this model could predict the baking volume quite well. The spring wheat model performed not so well, which is interesting since machines like Alveolab are commonly used to test flours more similar to spring wheat than winter wheat. But let us focus on the winter wheat model. The most important parameter for this model was the level of glucomannan, a type of dietary fibre, in the flour. The more the glucomannan, the higher the baking volume. Other studies have seen that addition of glucomannan to wheat dough could stabilise gluten proteins and therefore increase the bread volume. This large effect is interesting since glucomannan makes up than less than 1 % of the total flour.

The damaged starch content had no impact on the baking volume for winter wheats but had a clear negative impact for spring wheats. This might be since more damaged starch means a higher initial dough consistency but damaged starch is sensitive to enzymatic hydrolysis. If the damaged starch is degraded the dough will then loose its consistency and air cells in the dough might collapse and decrease the baking volume.

Three take home messages

- The model for winter wheat were more promising.
- Damaged starch had a negative impact on baking volume for spring wheats.
- Flour has less than 1 % glucomannan but it still affected the bread volume for winter wheats.

Final conclusion: Data from SDmatic, SRC-CHOPIN 2 and Alveolab might be used to predict baking volume but not yet to the point where we could stop test baking.