Received: 9 June 2023 | Revised: 23 January 2024 | Accepted: 1 February 2024

DOI: 10.1002/cche.10772

CEREALS & GRAINS ASSOCIATION

RESEARCH ARTICLE

Prediction models to evaluate baking quality instruments for commercial wheat flour

Louise Selga[1] | **Eva Johansson**[2] | **Roger Andersson**[1]

[1]Department of Molecular Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

[2]Department of Plant Breeding, Swedish University of Agricultural Sciences, Lomma, Sweden

Correspondence
Louise Selga, Department of Molecular Sciences, BioCenter, Swedish University of Agricultural Sciences, P.O. Box 7015, SE-750 07 Uppsala, Sweden.
Email: louise.selga@slu.se

Funding information
Ministry of Enterprise and Innovation, Grant/Award Number: N201703895; Sveriges Lantbruksuniversitet, Grant/Award Number: SLU.ua.2017.1.1.1-2416; Lantmännen

Abstract

Background and Objectives: Loaf volume is the main indicator of wheat flour quality, but test baking has major limitations. Here, prediction models were used to evaluate which methodology best captured the baking quality in Swedish commercial wheat flour and if the chemical composition of flour increased prediction accuracy.

Findings: Flour type (e.g., winter vs. spring wheat) affected prediction model results significantly. Thus, separate prediction models should be developed for each flour type. Combining data from alveograph, farinograph, and glutomatic tests with protein and damaged starch gave the best prediction results. The main loaf volume predictors were dough strength for winter wheat, stability for spring wheat, and extensibility for flour blends. The composition of protein and arabinoxylan influenced several quality parameters but did not improve loaf volume predictions.

Conclusions: Best predictions were obtained for winter wheat. Spring wheat and flour blend models contained only one latent variable, indicating that protein content was the main determinant for loaf volume in these samples.

Significance and Novelty: This study is one of few using prediction models to evaluate instrument suitability to determine loaf volume. Instruments suitable for predicting quality were determined for commercial winter wheat flour, which is the main product of Swedish mills.

KEYWORDS
flour, gluten, loaf volume, partial least square regression, rheology, wheat

1 | INTRODUCTION

Wheat is the most traded crop globally (Food and Agriculture Organization Corporate Statistical Database, 2020), and the grain is used for a wide range of products. To secure the most suitable wheat quality for their products, mills and bakeries use specifications when purchasing wheat grain or flour. However, these do not always capture the parameters with the most significant effect on baking quality. This results in unexpected differences in the baking quality of commercially milled wheat flour. Flour quality appears to be mainly determined by the gluten protein content and composition (Hamer et al., 2009), but other flour components also have an impact on the baking quality of bread (Goesaert et al., 2005). Thus, differences in bread loaf volume may arise as a result of variations in protein quality despite the same protein content in the flour (Carson & Edwards, 2009). In industrial bakeries, consistent performance in the production and of the end-use consumer product is highly preferable, as flour with varying and unpredictable functionality could lead to variation in end-product quality, process interruptions, and food waste (Carson & Edwards, 2009). Many consider dough strength to be the most important quality trait, and it is often used to specify which flour is suitable for a baking application. Flours of weak gluten strength, such as winter wheat flours, are often blended with stronger flours in Swedish mills to achieve an appropriate and stable quality.

Dough strength is primarily determined by the protein content and molecular size distribution (Huebner & Wall, 1976; Johansson et al., 2013; Zhang et al., 2008). The content and composition of the gluten proteins, and in particular of the high-molecular-weight glutenin subunits, have an impact on dough strength (Guzmán et al., 2022). Size-exclusion high-performance liquid chromatography (SE-HPLC) is commonly used to rapidly estimate the amount and size distribution of protein in wheat flour (Helguera et al., 2020; Johansson et al., 2001). The percentage of unextractable polymeric protein in total polymeric protein (%UPP) has been shown to correlate to gluten strength (Malik et al., 2011). Several different solvents and extraction methods have been applied for quantifying gluten proteins (Shewry & Lafiandra, 2022). Glutenin polymers are stabilized by SS bonds (Tatham et al., 1990) and interact strongly with other polymers and gliadins through arrays of hydrogen bonds (Belton, 1999; Wellner et al., 2005). There is no consensus on which extraction method is best for separating polymers without disrupting SS bonds (Shewry & Lafiandra, 2022).

Dough strength can be tested by several empirical rheological tests, including alveograph, extensograph, farinograph, and mixolab tests. Alveograph and extensograph tests also indicate dough elasticity and extensibility, which need to be balanced for air to be retained during fermentation, and a too-low extensibility leads to lower ovenspring (Carson & Edwards, 2009). Additionally, farinograph tests measure water absorption, that is, the amount of water needed to reach a predetermined dough resistance. The water absorption of the dough is affected not only by the gluten content and composition but also by the amount of arabinoxylan (AX), damaged starch, and the flour particle size distribution (Posner & Hibbs, 2011).

The type of flour and dough quality control tests used in mills vary between countries, and the system is partially influenced by tradition. In addition to the use of empirical rheological tests, wheat flour quality is also evaluated using baking tests in Swedish mills. However, test baking has major limitations, being laborious and results varying with the method used (Thanhaeuser et al., 2014). To the Swedish milling industry, producing flour with a stable baking quality is more important than achieving a high baking quality. The Swedish milling industry finds their current system of empirical rheological tests and test baking not satisfactory in ensuring stable flour quality. By evaluating the instruments currently used within the Swedish milling industry and additional instruments available, predictions of the outcome in terms of loaf volume and quality might be improved. The present study aimed at (1) using prediction models to evaluate which available tests are most appropriate to the Swedish milling industry, (2) evaluating how the stability of baking quality is affected by the chemical compounds in the flour, such as content and composition of protein, damaged starch, and AXs, and (3) evaluating the applicability of loaf volume prediction models for various types of flour. A large sample set was collected from mills in Sweden to ensure relevant variation in the studied parameters. The research questions were addressed using principal component analysis (PCA) and partial least square regression (PLS), as these methods are suitable for the evaluation of dependent variables. Unlike multiple linear regression, PLS can analyze *X*-variables that are strongly correlated and noisy, which is expected from several of the parameters studied (Wold et al., 2001).

2 | MATERIALS AND METHODS

2.1 | Materials

The material used in the present study has been described previously by Selga et al. (2023). Thus, a total of 197 commercial sieved wheat flours were collected continuously from two mills (Lantmännen Cerealia) in Malmö (southern Sweden) and Strängnäs (central Sweden). The material was collected during harvest years 2018 and 2019, and there was a high-temperature drought during 2018 (Lama et al., 2022). The material included four flour types, composed of (i) Swedish winter wheat, (ii) Swedish spring wheat, (iii) a blend of 0%–15% Swedish winter wheat, 15%–35% Swedish spring wheat, and 50%–70% of a high-protein German winter wheat, and (iv) a high stability (*S*) blend of Swedish spring and winter wheat. The flours did not contain any additions, such as malt or ascorbic acid. After the milling, samples were stored at −20°C until further analysis.

2.2 | Flour quality analysis

The protein content (dry matter basis, %dm) of the samples was determined by near-infrared transmittance spectroscopy (NIT; Infratec, FOSS) following the manufacturer's procedure. A wide array of tests were performed according to the AACC approved methods of analysis, that is, wet gluten analysis (AACC 38-12.02), dough proofing analysis (AACC 89-01.01, Rheo F4; KPM Analytics), farinograph tests (AACC 54-21.02), alveograph tests (AACC 54-30.02), starch damage (AACC 76-33.0, SDmatic; KPM Analytics), solvent retention capacity (SRC) (AACC 56-15.01, SRC-Chopin 2; KPM Analytics), and mixolab tests (AACC 54-60-01). The analysis of AX composition has been described previously (Selga et al., 2023).

Test baking was performed by the mill according to their standard procedures. Test baking used flour (2 kg), the water of controlled temperature was added according to farinograph absorption, yeast (100 g), sugar (35 g), lard (35 g), salt (35 g) and ascorbic acid (3 g), and barley malt was added to correct the falling number to 270 s. Doughs were mixed in a Wendel mixer (Diosna). A fixed mixing time was used, which had been predetermined by the mill for each flour type. The mixing time was 4 min for winter wheat, 6 min for flour blends, and 6.5 min for spring wheat. Six freeform loaves with a weight of 420 g were mechanically shaped. A fixed proofing time was used, which was slightly shorter for winter wheat than for the other flour types. After baking, the loaf volume was measured using a Volscan profiler (Stable micro systems) and the average volume was reported for each sample. The standard procedure followed by the mill did not include measuring the weight of baked loaves; thus, specific volume was not reported.

2.3 | SE-HPLC

The amount and size distribution of monomeric and polymeric protein was analyzed according to Gupta et al. (1993) with modifications according to Lan et al. (2023). The analysis was performed in duplicates on 16.5 mg freeze-dried sieved wheat flour. The samples were suspended in 1.4 mL 0.5% sodium dodecyl sulfate (SDS) phosphate buffer (pH 6.9), then vortexed for 10 s, shaken at 2000 rpm for 5 min, and centrifuged at 10,000g for 30 min. The supernatant was used for the quantification of extractable proteins, and the pellet was used for the determination of unextractable proteins. The supernatant was heated to 80°C for 2 min to inactivate proteases, in accordance with Larroque et al. (2000). The pellet was resuspended in 1.4 mL 0.5% SDS-phosphate buffer (pH 6.9) and sonicated for 45 s at amplitude 5, using a 3 mm exponential microtip (Soniprep 150; Tamro). The samples were centrifuged for 30 min at 10,000g, and the supernatants were collected and heated at 80°C for 2 min.

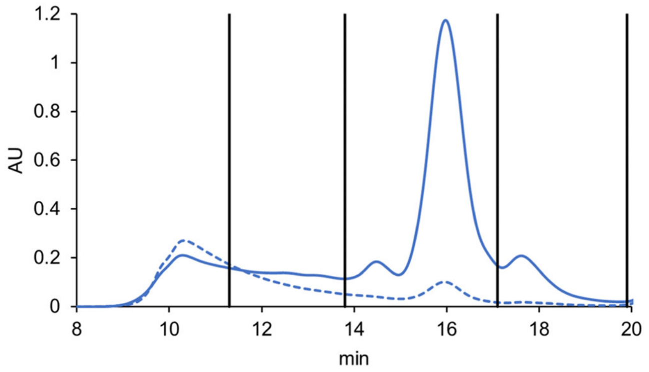

FIGURE 1 SE-HPLC chromatograms of SDS-extractable proteins (solid) and SDS-unextractable proteins (dashed). Chromatogram areas were divided at the following retention times: 8–11.3 min for large polymeric proteins (LPPs), 11.3–13.8 min for smaller polymeric proteins (SPPs), 13.8–17.1 min for large monomeric proteins (LMPs), and 17.1–19.9 min for smaller monomeric proteins (SMPs). [Color figure can be viewed at wileyonlinelibrary.com]

All samples of extractable and unextractable proteins were quantified by SE-HPLC (Waters) on a BioSep-SEC s4000 Phenomenex column, according to Lan et al. (2023). The samples were separated according to molecular size distribution in 30 min and detected by UV absorbance at 210 nm. In accordance with Johansson et al. (2001), the resulting chromatograms were divided according to molecular size into large polymeric proteins (LPPs), smaller polymeric proteins (SPPs), large monomeric proteins (LMPs), and smaller monomeric proteins (SMPs) (Figure 1). These chromatogram areas were used to calculate five protein parameters, as described in Table 1. These are total extractable protein (TOTE), total unextractable protein (TOTU), the percentage of unextractable polymeric protein in total polymeric protein (%UPP), the percentage of large unextractable monomeric protein in total large monomeric protein (%LUMP), and the ratio of monomeric protein to polymeric protein (Mon/Pol).

2.4 | Statistical analysis

Significant differences were calculated at 95% confidence by Tukey's honestly significant difference test, using Minitab (Minitab). Tukey's honestly significant difference test was used as it is the most suitable post hoc comparison when testing a large number of pairs

TABLE 1 Calculations of parameters describing protein molecular size distribution, based on SDS-extractable (e) and SDS-unextractable (u) polymeric proteins (PP = LPP + SPP) and monomeric proteins (MP = LMP + SMP).

Parameter	Definition	Calculation
TOTE	Total SDS-extractable proteins	ePP + eMP
TOTU	Total SDS-unextractable proteins	uPP + uMP
%UPP	Percentage of unextractable polymeric protein in total polymeric protein	$\frac{\text{uPP}}{\text{uPP} + \text{ePP}}$
%LUMP	Percentage of large unextractable monomeric protein in total large monomeric protein	$\frac{\text{uLMP}}{\text{uLMP} + \text{eLMP}}$
Mon/Pol	Ratio of monomeric protein to polymeric protein	$\frac{\text{eMP} + \text{uMP}}{\text{ePP} + \text{PP}}$

(International Business Machines Corporation, 2021). PCA and PLS were performed using SIMCA 17 (Sartorius). Separate models were built for each flour type. Test sets composed of 33% of the samples were generated for each flour type by distributing samples according to loaf volume and selecting every third sample. The test sets were excluded when building the PLS models and were used to calculate the root mean squared error of prediction (RMSEP). The PLS models were optimized to minimize Q^2 by stepwise removal of parameters with low variable importance in projection (VIP) scores and the removal of parameters which were highly unstable across cross-validations. Quality instruments which yielded parameters with low overall VIP scores were excluded to obtain a high Q^2 based on a limited set of instruments. Linear regressions were performed in Excel using the calibration sets and test sets described above.

3 | RESULTS AND DISCUSSION

3.1 | The effect of flour composition on baking quality

Significant differences were obtained between winter and spring wheat flour for all parameters evaluated, with the exception of water absorption (Table 2) and total AX content (Selga et al., 2023). PCA analysis identified that the high variation between flour types mainly originated from differences in protein content, loaf volume, alveograph extensibility (L) and TOTE (results not shown). Separate PCA models for flour blends, winter, and spring wheat flours (Figure 2a–c) displayed stronger internal correlations and a higher explained variation for winter wheat compared to the spring wheat and flour blends due to a higher overall variance (Table 2). The variation captured in principal components (PC) 1-2 ranged between 32% and 43% for the different PCA models (Figure 2), which is not uncommon when modeling data with large measurement errors and is to be expected from the empirical rheological tests used. The variance covered by each PC in PCA depends on many aspects of the raw data, such as the level of covariance, measuring error, and variance in the variables included (Esbensen et al., 2002). When PCA, as in this case, is applied for exploring the data rather than to make predictive classification models, it is very important to evaluate the number of relevant PCs in relation to knowledge of the context-specific problem studied (Esbensen et al., 2002). The scores in Figure 2 clearly show that both PC 1 and 2 contribute to the separation of the sample categories.

Several of the utilized methodologies intend to measure the gluten strength of the flour/dough: alveograph strength (W), lactic acid SRC (LAc SRC), gluten index (GI), farinograph stability, mixolab stability, and %UPP. All these parameters measuring gluten strength were well correlated in the spring and winter wheat PCA models, respectively (Figure 2a,b). The gluten strength of the samples from harvest year 2018 was higher than in samples from 2019, as determined by SE-HPLC and the alveograph (Table 2) due to the drought during 2018 (Lama et al., 2022). Loaf volume was placed differently in the PCA models for the different flour types, and a correlation to protein content was only seen for the spring wheat flours ($R = 0.49$, $p < .01$). Loaf volume was negatively correlated to dough strength for both the spring and winter wheat (Figure 2a,b), which corresponded with previous studies and is a result of the relatively gentle mixing with a fixed mixing time used in test baking in Sweden (Johansson et al., 2001). The correlation between loaf volume and dough strength varies depending on the test baking method (Dupuis & Fu, 2017), although a recent study on CIMMYT wheat varieties pointed out gluten strength as the most important parameter for bread-making (Guzmán et al., 2022). However, the overall variation in

TABLE 2 Mean values, with coefficients of variation (%) displayed in parenthesis, for parameters evaluated in spring wheat (SW) flour, winter wheat (WW) flour, and flour blends (Blends). Gluten composition parameters are defined in Table 1.

	Flour type			Milling location		Harvest year	
	SW	Blends	WW	Malmö	Strängnäs	2018	2019
Test baking							
Loaf volume, mL (Loaf *V*)	2468a (6)	2464a (6)	1901b (6)	2170a (15)	2149a (13)	2144a (15)	2185a (14)
Composition							
Protein, %dm (PC)	14a (4)	13b (3)	12c (4)	12a (10)	13a (6)	13a (9)	12a (9)
Damaged starch, %dm (DS)	5.7b (8)	5.8b (7)	6.6a (8)	6.1b (10)	6.5a (10)	6.4a (9)	6.1b (10)
Ash, %dm	0.65a (5)	0.60b (6)	0.58c (5)	0.58b (8)	0.64a (5)	0.60a (8)	0.59a (8)
Wet gluten							
Wet gluten, %dm (WG)	36a (6)	36a (5)	32b (7)	34a (9)	35a (5)	34a (8)	34a (9)
Gluten index (GI)	94a (4)	94a (4)	88b (9)	92a (6)	87b (11)	92a (6)	90a (9)
Farinograph							
Water absorption, % (14%mb) (WA)	60a (2)	60a (2)	59a (3)	59b (2)	61a (2)	60a (2)	59a (3)
Dough development time, min (DDT)	4.9a (20)	4.3b (19)	3.0c (24)	3.6a (34)	4.0a (20)	3.6a (32)	3.8a (30)
Stability, min (*S*)	8.4a (27)	7.9a (26)	5.9b (25)	7.3a (30)	5.6b (28)	7.2a (30)	6.7a (33)
Degree of softening, BU (DoS)	72b (21)	73b (21)	84a (20)	75b (21)	90a (17)	74b (22)	82a (20)
Alveograph							
Tenacity, mmH_2O (*P*)	80b (14)	80b (12)	88a (15)	82b (15)	93a (11)	89a (13)	80b (16)
Extensibility, mm (*L*)	107c (14)	103b (13)	75a (17)	92a (23)	80b (26)	88a (23)	91a (24)
Strength, 10^{-4} J (*W*)	258a (14)	254a (11)	208b (14)	229a (18)	240a (16)	245a (13)	220b (20)
P/*L*	0.8b (28)	0.8b (23)	1.2a (29)	0.9b (35)	1.3a (35)	1.1a (34)	0.9b (39)
Elasticity index, % (EI)	54a (5)	54a (4)	47b (7)	50a (9)	50a (11)	51a (7)	50b (11)
Dmin	−2.1a (12)	−2.2b (11)	−2.8c (15)	−2.4a (18)	−2.7b (21)	−2.6b (18)	−2.4a (20)
Solvent retention capacity (SRC)							
Lactic acid, % (14%mb) (Lac SRC)	135b (4)	137a (4)	122c (5)	128a (7)	125b (6)	130a (6)	126b (7)
Sucrose, % (14%mb) (Suc SRC)	106a (3)	106a (3)	100b (3)	103a (4)	103a (3)	104a (3)	102b (4)
Na_2CO_3, % (14%mb) (CO_3 SRC)	84a (4)	84a (5)	85a (6)	85a (5)	85a (5)	87a (4)	83b (5)
Water, % (14%mb) (H_2O SRC)	65a (4)	65a (4)	66a (5)	65b (4)	67a (5)	67a (4)	65b (5)
Gluten performance index	0.70b (4)	0.72a (3)	0.66c (5)	0.68a (6)	0.66b (6)	0.68a (6)	0.68a (6)
Mixolab (ML)							
C1, Nm (consistency at absorption)	1.10a (2)	1.10a (3)	1.10a (3)	1.10a (3)	1.10a (2)	1.10a (2)	1.10a (3)
Cs, Nm (consistency at 8 min)	1.03a (3)	1.02a (3)	0.98b (4)	1.00a (4)	1.00a (5)	1.01a (3)	1.00b (5)
C2, Nm (protein weakening)	0.54b (7)	0.58a (6)	0.57a (8)	0.57a (7)	0.55b (8)	0.58a (6)	0.55b (7)
C3, Nm (starch gelatinization)	1.64b (4)	1.77a (3)	1.79a (6)	1.76a (6)	1.75a (6)	1.74b (4)	1.77a (7)
C4, Nm (hot gel stability)	1.47a (9)	1.43a (11)	1.38b (7)	1.40b (8)	1.46a (11)	1.36b (7)	1.45a (9)
C5, Nm (starch retrogradation)	2.66c (9)	2.87a (7)	2.78b (7)	2.80a (7)	2.71b (11)	2.76a (6)	2.80a (9)
WA, % (14%mb) (WA [ML])	59a (2)	59a (2)	59a (3)	59b (2)	60a (2)	59a (2)	59a (3)
DDT, min (DDT [ML])	4.1a (36)	3.3b (42)	2.2c (41)	3.0a (49)	2.8a (49)	2.7a (47)	3.0a (49)
S, min (*S* [ML])	9.3a (4)	9.2a (6)	8.2b (15)	8.7a (12)	8.8a (11)	9.0a (12)	8.4b (7)

(Continues)

TABLE 2 (Continued)

	Flour type			Milling location		Harvest year	
	SW	Blends	WW	Malmö	Strängnäs	2018	2019
Rheo F4							
Proofing volume, mL (Proof *V*)	1573a (4)	1541b (4)	1509c (5)	1528b (5)	1579a (4)	1531a (4)	1546a (5)
CO_2 retention, % (CO_2 ret.)	80b (3)	80a (3)	81a (3)	81a (3)	78b (3)	80a (3)	80a (3)
Gluten composition							
TOTE, %dm	10.1a (5)	9.4b (5)	8.7c (5)	8.7a (10)	8.7a (8)	9.4a (10)	9.4a (9)
TOTU, %dm	3.8a (10)	3.7a (9)	3.1b (11)	3.4a (15)	3.5a (12)	3.5a (13)	3.3b (16)
%UPP, %dm	50a (4)	50a (5)	47b (6)	48b (6)	49a (5)	49a (5)	47b (7)
%LUMP, %dm	12a (19)	12a (20)	11a (18)	11a (15)	12a (27)	12a (20)	11a (16)
Mon/Pol	1.36a (4)	1.34b (3)	1.32c (4)	1.34a (4)	1.33a (5)	1.33a (4)	1.34a (4)

Note: Different letters (a, b, c) indicate significant differences ($p < .05$) in the parameter within sample categories flour type, milling location, and harvest year. %dm = dry matter basis, %mb = moisture basis.

commercial wheat flours is low compared to wheat grown in field trials and this also influenced parameter interactions. By using a fixed mixing time during test baking, wheat of exceptional quality risks being undermixed. Thus, the test baking method would not be ideal for wheat of exceedingly high or low quality. However, the method applied here is the one being used commercially in Swedish mills, where the purpose behind mill test baking is to identify flours of deviating quality, not to rank low- and high-quality flour, as the goal is to ensure stable flour quality.

Flours milled at the Strängnäs mill in central Sweden differed significantly from flours milled in the Malmö mill, in the south of Sweden, for several of the parameters evaluated (Table 2, Figure 2d,e). The total AX content was reported in a previous study to be significantly higher in samples from Strängnäs compared to samples from Malmö (Selga et al., 2023). This might be caused both by differences in milling practices and differences in growing conditions. Samples from Strängnäs also had a significantly higher ash content and damaged starch content (Table 2), indicating quality differences arising from differences in milling practices between the two locations. The Strängnäs flour mainly had higher levels of WE-AX (Selga et al., 2023), which originates from the endosperm (Marion & Saulnier, 2020), so WE-AX levels were not caused by bran inclusion during milling. Thus, differences between Malmö and Strängnäs were also likely caused by differences in growing conditions.

Water absorption was the largest loading on PC 1 for winter wheat (Figure 2a) due to a positive correlation between fiber content, damaged starch, and protein content. These parameters have been reported to correlate with water absorption in previous studies as well (Courtin et al., 1999; Greer & Stewart, 1959). For the spring wheat and blends, water absorption appeared mostly influenced by WU-AX and damaged starch (Figure 2b,c). It is notable that damaged starch did not correlate to protein content in all the sample sets, even though protein and damaged starch showed similar coefficients of variation (Table 2). Damaged starch differed significantly between harvest years and locations, while protein content did not (Table 2).

Alveograph tenacity (*P*) and damaged starch were placed close to each other in the PCA plots for all three wheat types evaluated here (Figure 2), which correspond to previous studies (Preston et al., 1987). Following the methodology for alveograph analyses, the amount of water was not adjusted when the dough was prepared. However, the gluten network requires water to form (Belton, 1999), and as both damaged starch and WU-AX bind water, these components have an impact on free water being available in the dough. Shortage of water results in a short dough in the alveograph with high tenacity. The AX composition varied distinctly in all the studied flour types, and WU-AX showed large positive PC 1 values for all flour types (Figure 2). The SRC-Chopin 2 was used to evaluate if the flour composition could be estimated by the flour SRC. Each of the SRC parameters corresponded to %UPP, damaged starch, and total AX in winter wheat, as visualized by the PCA plot (Figure 2a). However, this relationship was not noted for the spring wheat and blend samples. Thus, the results of the winter wheat might be hampered by the correlation among damaged starch, total AX, and protein content in these samples. Fermentation volume, tested with a Rheo F4, did not correlate significantly to loaf volume.

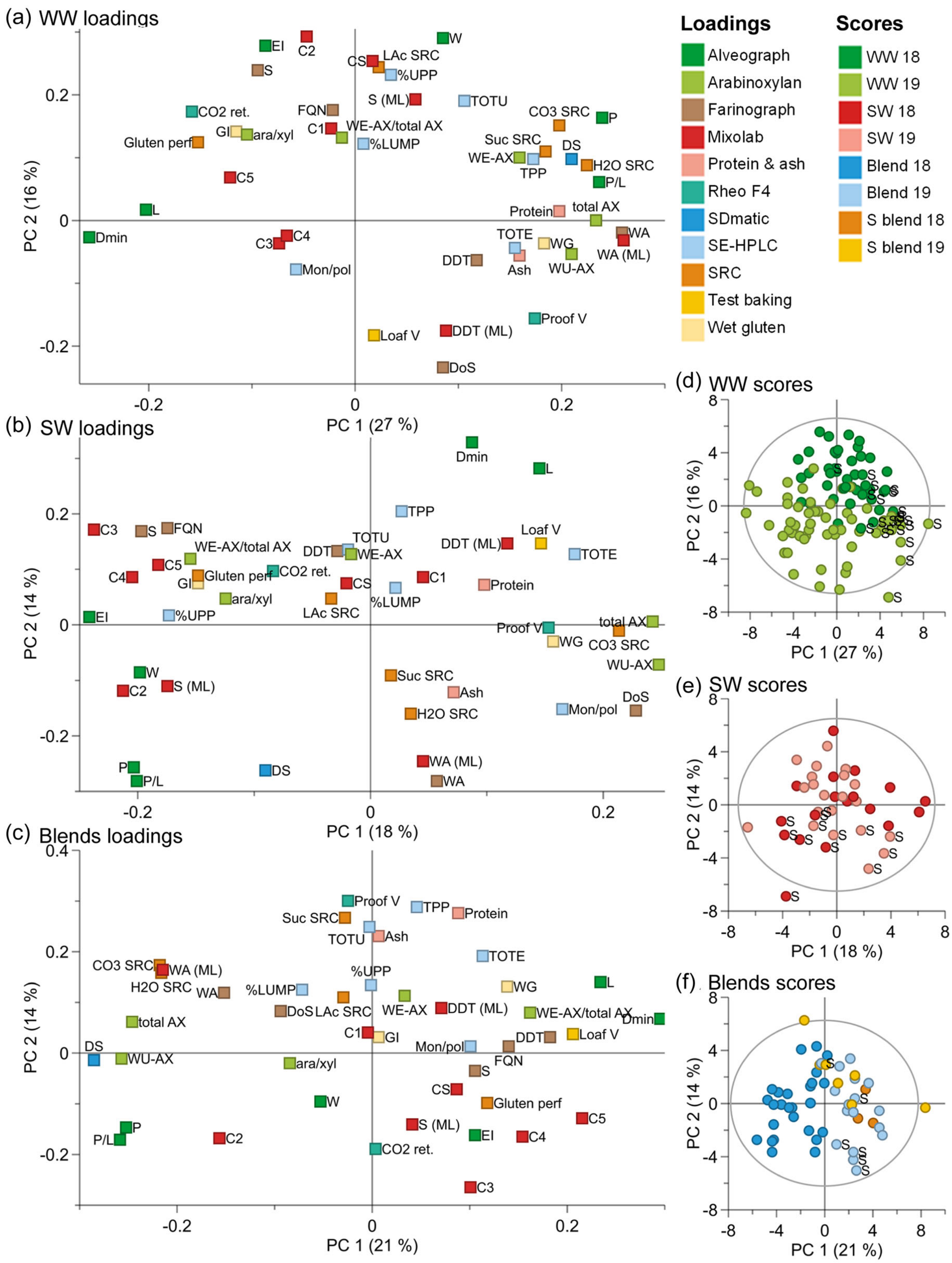

FIGURE 2 Loadings and scores from PCA of (a, d) winter wheat (WW), (b, e) spring wheat (SW), and (c, f) flour blends. Scores are colored darker for harvest year 2018 (18) and lighter for 2019 (19), with samples from Strängnäs (S) labeled. Circle indicates 95% Hotelling's T2 distribution. Loadings are colored according to methods, and parameter abbreviations are defined in Table 2. [Color figure can be viewed at wileyonlinelibrary.com]

However, its positive placement to protein and negative placement to *W* were consistent for all sample types.

3.2 | Predicting loaf volume

One PLS model was selected for each flour type based on their Q^2Y values (Table 3) and the lower number of instruments used (Figure 4). Similar sets of instruments were selected for winter and spring wheat PLS models, with alveograph, farinograph, glutomatic wet gluten, and NIT included in both models. Additionally, damaged starch, as quantified by SDmatic was included in the spring wheat PLS model (Figure 4b). However, the parameters included in each instrument differed between models (Figure 4), and none of the models could reliably predict the other flour types. This difference between flour types was caused by low variation in the quality of the commercial samples used in this study. While overall trends between quality measurements and loaf volume have been seen in previous studies using wheat from field trials (Guzmán et al., 2022), these trends may not always be relevant in a commercial setting, where the variations in quality are much lower. Here, the focus is rather on identifying deviating samples. PCA and PLS models may identify parameter patterns leading to a deviating quality, and this pattern may differ between flour types.

TABLE 3 PLS models predicting loaf volume (mL).

Sample set	*n*	*A*	*n X*	R^2Y	Q^2Y	RMSEP
Winter wheat	70	3	15	0.62	0.50	74 mL
Spring wheat	26	1	10	0.55	0.38	113 mL
Flour blends	35	1	9	0.35	0.30	126 mL

Note: R^2Y (cumulative) measures model fit to *Y*, and Q^2Y (cumulative) measures the total variation of *Y* that can be predicted, as estimated by cross-validation.

Abbreviations: *A*, number of latent variables (e.g., PLS factors); *n*, number of samples in calibration set; *n X*, number of parameters included in model; RMSEP, root mean squared error of prediction.

The flour blend PLS model included alveograph, mixolab, SDmatic, and NIT parameters (Figure 4c). The interpretability of the mixolab curve would need to improve for it to be used for routine quality control in Swedish mills.

Except for protein content, damaged starch, and ash content, flour composition data did not improve PLS predictions. While %UPP yielded a very high VIP score when included in the winter wheat PLS model, both Q^2Y and RMSEP remained unchanged when %UPP was excluded. The relevant information carried by %UPP appeared to be fully captured through the combination of other parameters estimating gluten strength, mainly gluten index, elasticity index, and farinograph stability. This reaffirms the value of empirical rheological methods for quality control.

The winter wheat PLS model (Figures 3a and 4a) had the lowest RMSEP and was the only PLS model with more than one latent variable, for example, PLS factor (Table 3). This indicated a more complex relationship between the parameters than only covariance to protein content. For comparison, linear regression of loaf volume based only on protein content gave an RMSEP of 102 mL. The winter wheat PLS model did not appear to be overfitted, as the prediction error remained in the same range when the number of parameters or latent variables was changed. Winter wheat displayed the lowest range in measured loaf volume (Figure 3), which may have contributed to these RMSEP values being relatively low. Both protein and wet gluten content had high positive

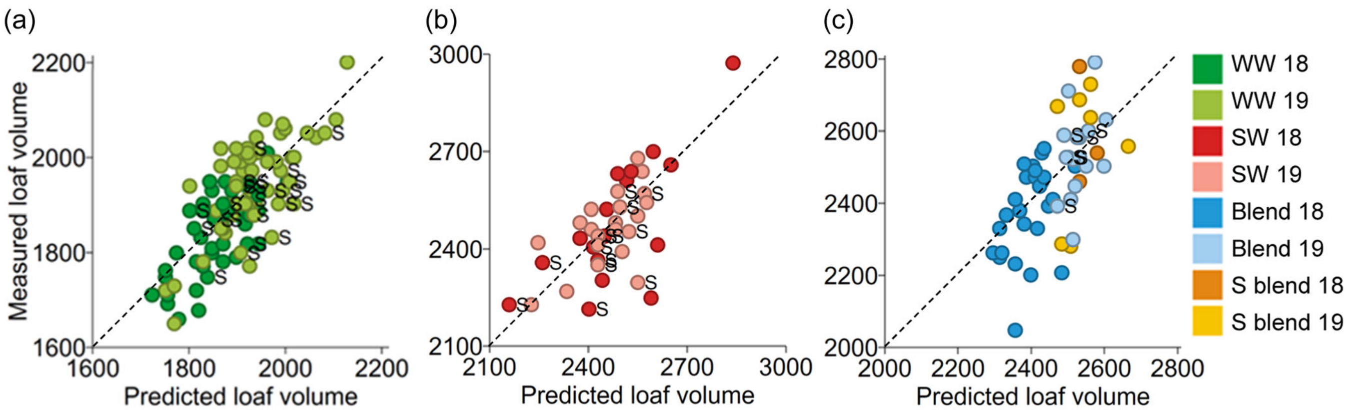

FIGURE 3 Measured loaf volume (mL) versus PLS predicted loaf volume (mL), with dashed unity line, for winter wheat (WW) flour (a), spring wheat (SW) flour (b), and flour blends (Blend and high stability blend, S blend) (c), from 2018 harvest (18) and 2019 harvest (19). Samples from Strängnäs (S) are labeled. Both test sets and calibration sets are plotted. Note that the axis differs between plots. [Color figure can be viewed at wileyonlinelibrary.com]

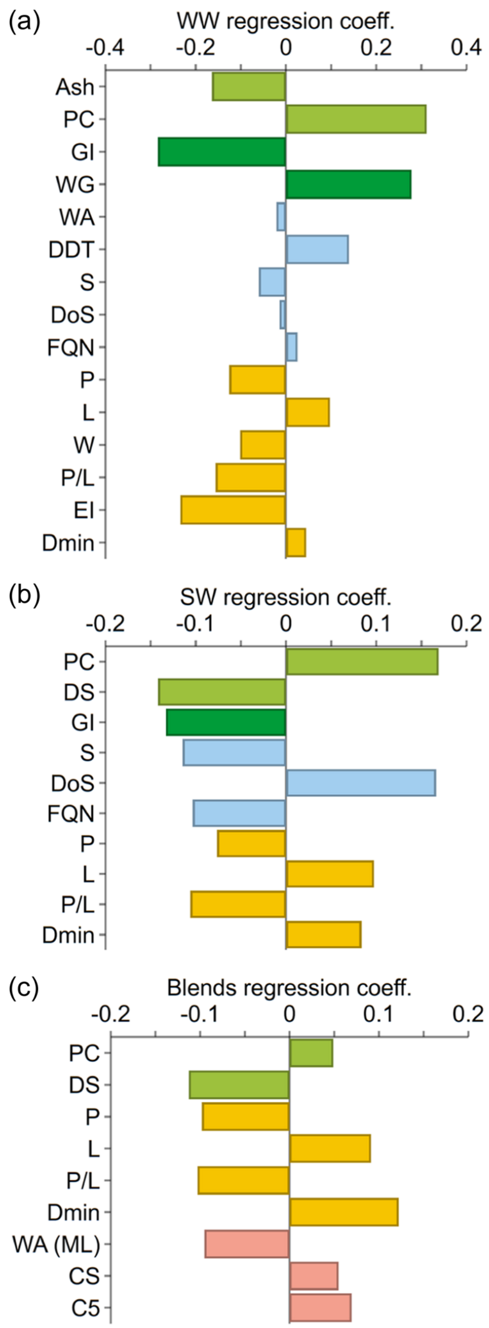

FIGURE 4 Regression coefficients for winter wheat PLS model (a), spring wheat PLS model (b), and flour blends PLS model (c). Abbreviations are defined in Table 2. Coefficients are colored according to the data source: flour components (light green), wet gluten (dark green), farinograph (light blue), alveograph (yellow), or mixolab (pink). [Color figure can be viewed at wileyonlinelibrary.com]

loading weights while gluten index (GI) and elasticity index (EI) had high negative loading weights across all latent variables. Additionally, the first latent variable had high positive loading for the degree of softening and high negative loading for farinograph stability, and the second latent variable displayed a high positive loading weight for extensibility and high negative loading weights for water absorption and ash. Overall, a low gluten strength and high protein content, coupled with a high extensibility and low water absorption, produced the highest loaf volumes for winter wheat flour. While a high gluten strength is generally beneficial to loaf volume (Guzmán et al., 2022), this differs between the test baking method used (Dupuis & Fu, 2017), and previous studies using a similar Swedish test baking method observed similar trends (Johansson et al., 2001). This model used instruments currently used in the mill, making it highly commercially viable.

The spring wheat PLS model (Figures 3b and 4b) yielded a higher RMSEP and a lower Q^2Y compared to the winter wheat PLS model. For comparison, linear regression of protein content vs. loaf volume gave an RMSEP of 147 mL, and by this comparison, spring wheat PLS modeling gave the biggest improvement in RMSEP. However, the spring wheat PLS model had the most unstable parameters over cross-validations, which was also reflected in the low Q^2Y (Table 3). Additionally, spring wheat flour had the lowest number of samples and the highest range in measured loaf volume (Figure 3) of the different flour types. The cross-validation results indicated that parameter variation patterns differed a lot between samples, making it challenging to generate a representative test set and the RMSEP should therefore be interpreted accordingly. Protein content and DoS were the largest positive loading weights, and damaged starch was the largest negative loading. TOTE yielded a high VIP score when added to the spring wheat PLS model, and improved Q^2Y and RMSEP somewhat. TOTE was however excluded from the final model, due to the complexity of performing SE-HPLC to evaluate flour in a mill context. There were differences in loaf volume between the harvest years which were not well captured by any measured parameters, except for differences in damaged starch between the 2 years. This could be seen in the spring wheat PCA scores (Figure 2e), which was the only model where the samples did not assort according to harvest year (Figure 2d,e). Barrera et al. (2007) observed a negative correlation between loaf volume and damaged starch, when obtaining high damaged starch levels by excessive milling. This could arise from gluten proteins and damaged starch competing for water (Barrera et al., 2007). However, Dhaka et al. (2012) saw no correlation between loaf volume and damaged starch when comparing wheat varieties with a range in damaged starch closer to those observed in this study. Damaged starch alone did not seem likely to cause the differences in loaf volume between harvest years, as the differences between harvest years in damaged starch were larger for winter wheat flours, and damaged starch did not improve winter wheat predictions. There was a difference in loaf volume between Malmö and Strängnäs during 2018 but not during

2019, while the differences in damaged starch between Malmö and Strängnäs were present both during 2018 and 2019. Differences in milling practices between the two mills are therefore not likely to be a major cause of differences in loaf volume. Considering the low production in the mills and high variability in baking performance, using traditional quality control for spring wheat flour may be preferable over implementing prediction models.

The blends PLS model (Figures 3c and 4c) yielded the largest RMSEP and lowest Q^2Y. Unlike the other models, protein content had the lowest VIP score out of the parameters included in the model. The correlation between protein content and loaf volume was only 0.27 ($p < .05$) for the flour blends. When using linear regression between protein content and loaf volume, the RMSEP was 158 mL. The alveograph parameter Dmin was the highest positive loading weight and damaged starch was the highest negative loading weight. Both these parameters differed significantly between harvest years (Table 2), leading to distinctly different predictions for the two harvest years (Figure 3c). The PLS model did not appear to take in information related to the composition of spring and winter wheat in the blends. Including parameters that differed with wheat composition, such as WU-AX content (Selga et al., 2023), did not improve the RMSEP. The inadequate performance of the PLS model was probably caused by the flour blends displaying the lowest variance in the measured parameters (Table 2). The lack of variation appeared to prevent prediction models from being a useful tool for quality control of this flour type.

4 | CONCLUSIONS

Each flour type yielded unique PCA results, and separate PLS models for each flour type yielded the best predictions. While gluten and AX composition facilitated interpretability of the quality tests, they did not improve predictions of loaf volume, except for TOTE in spring wheat flour. Notably, %UPP performed well when modeling winter wheat loaf volume but was ultimately excluded as the relevant dough strength was fully captured by the empirical rheology tests included. The mills in Sweden do not currently utilize all quality tests presented here. The alveograph contributed to all prediction models. However, the results were influenced by the WU-AX content and damaged starch content and should be interpreted accordingly. The farinograph is well adapted for Swedish conditions with differences in water absorption; however, %UPP correlated more to *W* than stability in this sample set. SDmatic rapidly measures damaged starch, the levels of which differed between locations and harvest years. This test may therefore be useful, especially when starting milling during new harvests. While the mixolab contributed to loaf volume predictions for flour blends, these loaf volume predictions remained poor. Finally, while the SRC Chopin gave promising results in winter wheat, these results were not reproducible in spring wheat and not relevant for predicting loaf volume. PCA and PLS models are great for combining dependent variables and excluding measurement errors, which are prevalent in empirical rheological measurements, and especially in test baking. PLS modeling has potential to be used to predict loaf volume for winter wheat flour in Swedish mills, as this model performed best overall. This is the main product of Swedish mills and modeling would reduce test baking costs drastically simply by using data that can already be collected in the mills.

ACKNOWLEDGMENTS

The authors want to thank Assar Sundholm, Edvin Nåbo, Josefine Dahlman, and Stephanie Laurent-Hedlund for their technical assistance. Funding: This study was financed through an industry PhD program of the Swedish Government (governmental decision N201703895), SLU (grant number SLU. ua.2017.1.1.1-2416) and Lantmännen. Louise Selga is employed by Lantmännen.

ORCID

Louise Selga http://orcid.org/0000-0002-0264-9255

REFERENCES

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

Belton, P. S. (1999). Mini review: On the elasticity of wheat gluten. *Journal of Cereal Science*, *29*(2), 103–107. https://doi.org/10.1006/jcrs.1998.0227

Carson, G. R., & Edwards, N. M. (2009). Criteria of wheat and flour quality. In K Khan & P Shewry (Eds.), *Wheat: Chemistry and technology* (pp. 97–118). American Association of Cereal Chemists, Inc (AACC).

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

Dhaka, V., Gulia, N., & Khatkar, B. (2012). Application of mixolab to assess the bread making quality of wheat varieties. *Journal of Food Processing & Technology*, *01*(2), 1–8. https://doi.org/10.4172/scientificreports.183

Dupuis, B., & Fu, B. X. (2017). A new lean no time test baking method with improved discriminating power. *Journal of Cereal Science*, *74*, 112–120. https://doi.org/10.1016/j.jcs.2017.01.017

Esbensen, K. H., Guyot, D., Westad, F., & Houmoller, L. P. (2002). *Multivariate data analysis—In practice: An introduction to* *multivariate data analysis and experimental design* (5th ed). CAMO Process AS.

FAOSTAT. (2020). *FAOSTAT: Statistical database of production quantity from crops and livestock products*. FAO. Retrieved December 14, 2022, from: https://www.fao.org/faostat/en/#compare

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*(1–3), 12–30. https://doi.org/10.1016/j.tifs.2004.02.011

Greer, E. N., & Stewart, B. A. (1959). The water absorption of wheat flour: Relative effects of protein and starch. *Journal of the Science of Food and Agriculture*, *10*(4), 248–252. https://doi.org/10.1002/jsfa.2740100409

Gupta, R. B., Khan, K., & Macritchie, F. (1993). Biochemical basis of flour properties in bread wheats. I. Effects of variation in the quantity and size distribution of polymeric protein. *Journal of Cereal Science*, *18*(1), 23–41. https://doi.org/10.1006/jcrs.1993.1031

Guzmán, C., Crossa, J., Mondal, S., Govindan, V., Huerta, J., Crespo-Herrera, L., Vargas, M., Singh, R. P., & Ibba, M. I. (2022). Effects of glutenins (Glu-1 and Glu-3) allelic variation on dough properties and bread-making quality of CIMMYT bread wheat breeding lines. *Field Crops Research*, *284*, 108585. https://doi.org/10.1016/j.fcr.2022.108585

Hamer, R. J., MacRitchie, F., & Weegels, P. L. (2009). Structure and functional properties of gluten. In K Khan & P Shewry (Eds.), *Wheat: Chemistry and technology* (pp. 153–178). American Association of Cereal Chemists, Inc (AACC).

Helguera, M., Abugalieva, A., Battenfield, S., Békés, F., Branlard, G., Cuniberti, M., Hüsken, A., Johansson, E., Morris, C. F., Nurit, E., Sissons, M., & Vazquez, D. (2020). Grain quality in breeding. In G. Igrejas, T. M. Ikeda, & C. Guzmán (Eds.), *Wheat quality for improving processing and human health* (pp. 273–307). Springer International Publishing.

Huebner, F., & Wall, J. (1976). Fractionation and quantitative differences of glutenin from wheat varieties varying in baking quality. *Cereal Chemistry*, *53*(2), 258–269.

IBM. (2021). *SPSS statistics 25.0: GLM post hoc comparisons*. IBM. Retrieved September 25, 2023, from: https://www.ibm.com/docs/en/spss-statistics/25.0.0?topic=analysis-glm-post-hoc-comparisons

Johansson, E., Malik, A. H., Hussain, A., Rasheed, F., Newson, W. R., Plivelic, T., Hedenqvist, M. S., Gällstedt, M., & Kuktaite, R. (2013). Wheat gluten polymer structures: The impact of genotype, environment, and processing on their functionality in various applications. *Cereal Chemistry*, *90*(4), 367–376. https://doi.org/10.1094/CCHEM-08-12-0105-FI

Johansson, E., Prieto-Linde, M. L., & Jönsson, J. Ö. (2001). Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality. *Cereal Chemistry*, *78*(1), 19–25. https://doi.org/10.1094/CCHEM.2001.78.1.19

Lama, S., Vallenback, P., Hall, S. A., Kuzmenkova, M., & Kuktaite, R. (2022). Prolonged heat and drought versus cool climate on the Swedish spring wheat breeding lines: Impact on the gluten protein quality and grain microstructure. *Food and Energy Security*, *11*(2), e376. https://doi.org/10.1002/fes3.376

Lan, Y., Kuktaite, R., Chawade, A., & Johansson, E. (2023). Diverse wheat lines to mitigate the effect of drought on end-use quality. *Frontiers in Food Science and Technology*, *3*, 1–15.

Larroque, O. R., Gianibelli, M. C., Sanchez, M. G., & MacRitchie, F. (2000). Procedure for obtaining stable protein extracts of cereal flour and whole meal for size-exclusion HPLC analysis. *Cereal Chemistry*, *77*(4), 448–450. https://doi.org/10.1094/CCHEM.2000.77.4.448

Malik, A. H., Prieto-Linde, M. L., Kuktaite, R., Andersson, A., & Johansson, E. (2011). Individual and interactive effects of cultivar maturation time, nitrogen regime and temperature level on accumulation of wheat grain proteins. *Journal of the Science of Food and Agriculture*, 2192–2200. https://doi.org/10.1002/jsfa.4439

Marion, D., & Saulnier, L. (2020). Minor components and wheat quality: Perspectives on climate changes. *Journal of Cereal Science*, *94*, 103001. https://doi.org/10.1016/j.jcs.2020.103001

Posner, E. S., & Hibbs, A. N. (2011). *Wheat flour milling* (2nd ed.). American Association of Cereal Chemists.

Preston, K. R., Kilborn, R. H., & Dexter, J. E. (1987). Effects of starch damage and water absorption on the alveograph properties of Canadian hard red spring wheats. *Canadian Institute of Food Science and Technology Journal*, *20*(2), 75–80. https://doi.org/10.1016/S0315-5463(87)71093-1

Selga, L., Andersson, A. A. M., Moldin, A., & Andersson, R. (2023). Determining levels of water-extractable and water-unextractable arabinoxylan in commercial Swedish wheat flours by a high-throughput method. *Journal of Cereal Science*, *109*, 103608. https://doi.org/10.1016/j.jcs.2022.103608

Shewry, P. R., & Lafiandra, D. (2022). Wheat glutenin polymers 1. Structure, assembly and properties. *Journal of Cereal Science*, *106*, 103486. https://doi.org/10.1016/j.jcs.2022.103486

Tatham, A. S., Shewry, P. R., & Belton, P. S. (1990). Structural studies of cereal prolamins, including wheat gluten. *Advances in Cereal Science and Technology*, *10*, 1–78.

Thanhaeuser, S. M., Wieser, H., & Koehler, P. (2014). Correlation of quality parameters with the baking performance of wheat flours. *Cereal Chemistry*, *91*(4), 333–341. https://doi.org/10.1094/CCHEM-09-13-0194-CESI

Wellner, N., Mills, E. N. C., Brownsey, G., Wilson, R. H., Brown, N., Freeman, J., Halford, N. G., Shewry, P. R., & Belton, P. S. (2005). Changes in protein secondary structure during gluten deformation studied by dynamic Fourier transform infrared spectroscopy. *Biomacromolecules*, *6*(1), 255–261. https://doi.org/10.1021/bm049584d

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Zhang, P., He, Z., Zhang, Y., Xia, X., Chen, D., & Zhang Y. (2008). Association between % SDS-Unextractable polymeric protein (%UPP) and End-Use quality in Chinese bread wheat cultivars. *Cereal Chemistry*, *85*(5), 696–700. https://doi.org/10.1094/CCHEM-85-5-0696

How to cite this article: Selga, L., Johansson, E., & Andersson, R. (2024). Prediction models to evaluate baking quality instruments for commercial wheat flour. *Cereal Chemistry*, *101*, 681–691. https://doi.org/10.1002/cche.10772