Food Analytical Methods (2026) 19:48
https://doi.org/10.1007/s12161-025-02931-7

RESEARCH

Improved Prediction of Wheat Quality and Functionality Using Near-Infrared Spectroscopy and Novel Approaches Involving Flour Fractionation and Data Fusion

Denise Ziegler[1] · Lukas Buck[2] · Katharina Anne Scherf[2,3,4] · Lutz Popper[5] · Alexander Schaum[1] · Bernd Hitzmann[1]

Received: 19 May 2025 / Accepted: 6 October 2025

Abstract

The accurate and rapid determination of wheat quality is of great importance for the wheat supply chain. Near-infrared (NIR) spectroscopy has become an established method for this purpose. So far, however, predictions for most wheat quality characteristics are not accurate enough to replace reference measurements, with the exception of protein content. This study investigates the potential to improve the prediction of 41 wheat quality parameters (protein- and starch-related parameters, solvent retention capacity, farinograph, extensograph, alveograph) based on a flour fractionation approach (sieve fractionation, dough preparation, gluten washing) and data fusion using the established techniques of NIR spectroscopy and chemometrics. Results show that preprocessing of flour significantly altered the composition of the samples, which reflected in spectral differences of their NIR spectra. This also led to a change in the prediction accuracy for many wheat quality parameters. Compared to the prediction using flour spectra, flour fractionation with or without data fusion improved the RMSECV between 5.6 and 28.6% for 35 out of the 41 quality parameters tested, leading to R^2_{CV} between 0.80 and 0.96 for many of them. Gluten, dough, and the 50–75 µm and the 75–100 µm fractions were particularly important for the improved predictions. The best predictions were often based on data fusion of spectra from different sample types, demonstrating the importance of using complementary information from different data sources to improve predictions. The results underline the potential of this novel approach to be established in the industry as an extension of conventional NIR spectroscopy to improve wheat quality prediction.

Keywords Chemometrics · Data fusion · Flour fractions · NIR spectroscopy · Rheology · Wheat quality

Introduction

Wheat is an important raw material for the production of many staple foods for humans. It is processed into a wide variety of products (Pojić and Mastilović 2013). As the number of products increased and processing methods became more advanced, the industry established more and more quality standards. This resulted in the need for more and more tests to analyze these quality standards within the wheat supply chain. For this reason, a variety of methods for determining wheat quality were established over the years (Miralbés 2004). These include not only analytical methods to examine the constitution of the samples, but also rheological methods to determine the functional properties as well as baking tests that reflect the end-product quality. However, these conventional methods are time-consuming, expensive and usually require a large amount of flour and equipment

✉ Denise Ziegler
denise.ziegler@uni-hohenheim.de

1 Department of Process Analytics, Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart, Germany

2 Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

3 Leibniz Institute for Food Systems Biology, Technical University of Munich, Freising, Germany

4 Technical University of Munich, TUM School of Life Sciences, Professorship of Food Biopolymer Systems, Freising, Germany

5 Mühlenchemie GmbH & Co. KG, Ahrensburg, Germany

Published online: 05 December 2025

(Miralbés 2004; Pojić and Mastilović 2013). For this reason, it has long been established that wheat quality is determined indirectly by measuring a parameter that can be determined fast and that reliably reflects wheat quality. In practice, this is the protein content, as the gluten proteins in particular strongly govern the functional and end-product properties of wheat. An established method for determining the protein content is Near-infrared (NIR) spectroscopy, which allows an accurate prediction within just a few minutes. However, many studies have shown that the protein content is not a suitable indicator of wheat quality, as the correlations with other properties (e.g., water absorption, dough development time, baking volume) are rather low. Nevertheless, in practice, the price of wheat is based on protein content, which means that farmers are not paid for the actual quality of their grain and also use high-nitrogen fertilization to achieve the highest possible protein content, which can lead to groundwater pollution (Gabriel et al. 2017; Nagel-Held et al. 2022, 2024). Therefore, it is of utmost importance for the entire wheat supply chain to find methods that enable an accurate and rapid determination of many different wheat quality characteristics with as little flour and effort as possible.

Spectroscopy is a suitable method for this purpose, as NIR spectroscopy in particular is already widely used to determine not only the protein content, but also, e.g., the water content, ash content, and the particle size of flour. It meets the requirements of speed, low flour quantity, and ease of handling and NIR spectrometers are already widely available (Pojić and Mastilović 2013). For this reason, several studies have tested NIR spectroscopy in particular for the prediction of other wheat quality traits, especially rheological and end-product characteristics. However, the reported predictions are often poor or only good enough for screening purposes (Dowell et al. 2006; Gabriel et al. 2017; Jirsa et al. 2008; Miralbés 2003, 2004; Nagel-Held et al. 2022, 2024), which is already a step forward for breeders but not good enough for the further wheat supply chain to actually replace the reference analyses. One possible new approach is to preprocess the flour into flour fractions and dough before spectroscopic analysis. The hypothesis behind this is that this preprocessing changes the composition of the samples by enriching and depleting certain components. Additionally, various reactions take place that further alter the composition, especially when the gluten network is formed during dough preparation. This could also change the signals in the spectra and reduce the problematic superimposition of signals in NIR spectroscopy, which might have a positive effect on the prediction accuracy of various wheat quality characteristics. Other studies have already shown that NIR spectroscopy can detect changes during dough preparation (Alava et al. 2001; Albanell et al. 2012; Wesley et al. 1998) and changes specifically in the gluten protein structure induced by heat and moisture (Bruun et al. 2007). Ziegler et al. (2025a) have shown that flour fractionation resulted in spectral changes of fluorescence spectra and significantly improved the prediction of rheological characteristics like the dough development time. Furthermore, Ziegler et al. (2025b) have presented promising results for an improved prediction of specific loaf volume using flour fractionation in combination with NIR spectroscopy. They have shown that data fusion is a valuable tool to further improve predictions, as complementary information for the prediction of baking quality was contained in spectra of different flour fractions.

The aim of this study is to test the potential of the novel approach of flour fractionation with and without data fusion in combination with NIR spectroscopy to improve the prediction of a wide variety of analytical and rheological measurements of wheat flour. Predicted are protein and wet gluten content, Osborne and SDSS-GMP (sodium dodecyl sulfate soluble proteins — glutenin macropolymer) fractionation, Hagberg falling number, starch damage, and solvent retention capacity (SRC) as well as farinograph, extensograph, and alveograph analyses. The practical relevance of the tested methods is ensured by using a very diverse sample set consisting of 50 commercially available wheat flour samples originating from ten countries and four harvest years.

Material and Methods

Wheat and Flour Samples

The sample set used in this study consisted of 50 commercially available wheat samples, representing mixtures of many different cultivars. It was a diverse sample set, because samples were harvested in different years (2019 – 2022) and were of different qualities although wheat classes were mostly unknown. Twenty-seven samples originated from Germany while others were grown in Australia, the USA, Latvia, Lithuania, Mexico, India, Poland, Romania, and Ukraine. The wheat samples were milled by Mühlenchemie GmbH & Co. KG (Ahrensburg, Germany) using a Buhler MLU 202 laboratory mill. After milling, the ash content of the flour samples was adjusted to approximately 0.60%.

Rheological and Analytical Measurements

All used reference analyses and methods as well as the obtained parameters are specified in Table 1. For Osborne and SDSS-GMP fractionation, detailed method descriptions are provided in the subsequent sections. The number of repetitions for every analysis varied between one and three, since the amount of flour was limited. Descriptive statistics for every obtained parameter can be found in Table 2. For

Table 1 List of reference analyses with used methods, number of replicates, and measured parameters

Reference analyses (abbreviation)	Method	Number of replicates	Measured parameters
Flour protein	ICC 159	2	Protein content
Wet gluten; Gluten Index	ICC 155	2	Wet gluten content; Gluten Index
Osborne fractionation	See section "Osborne Fractionation"	3	Osborne total extractable proteins; albumin and globulin; gliadin; glutenin; gliadin/glutenin ratio
SDSS-GMP fractionation	See section "SDSS-GMP Fractionation"	3	SDSS-GMP total proteins; SDSS; GMP; GMP-HMW; GMP-LMW
Hagberg falling number	ICC 107/1	2	Hagberg falling number
Starch damage	AACC 76-31	1	Starch damage
Solvent retention capacity (SRC)	Method based on AACC 56-11.02.02 using an automated system	1	Water; sucrose; lactic acid; sodium carbonate
Farinograph (FA)	ICC 115	2	Dough development time DDT; water absorption WAM; stability S; dough softening DS (10 min after start); quality number FQN
Extensograph (EX)	ICC 114/1	2	Energy (45 min); resistance (45 min); extensibility (45 min); maximum (45 min); ratio number (45 min); ratio number max. (45 min)
Alveograph (AL)	ICC 121	1	Maximum pressure C_{max}; tenacity P; extensibility L; swelling index G; deformation energy W; ratio P/L; elasticity index Ie; strength coefficient K; strain hardening index SH; minimum of first derivative D_{min}; maximum of first derivative D_{max}

parameters where the number of samples listed in this table is 49, the measurement of one sample was missing.

Osborne Fractionation

Flour (100 mg) was extracted in three stages according to Wieser et al. (1998). In the first extraction step, the albumin and globulin fraction was obtained. After the addition of 1 mL of salt solution (400 mmol/L NaCl and 67 mmol/L Na_2PO_4/KH_2PO_4 (pH 7.6)) to the flour, the suspension was vortexed for 2 min and stirred for 10 min at 22 °C. Afterwards, the suspension was centrifuged (25 min, 22 °C, 3550 rcf) and the supernatant was collected in a 2 mL volumetric flask. The procedure was repeated once more. Then, the residue from the albumin and globulin fraction was extracted three times using 0.5 mL of 60% ethanol, each time using the same procedure as for the albumin and globulin fraction. In the last extraction step, the glutenin fraction was obtained by extracting the residue from the gliadin fraction under reducing conditions using 10 mg/mL dithiothreitol (DTT) in buffer solution (0.1 mol/L TRIS-HCl (pH 7.5)/1-propanol (50%, v/v) with 12 g urea). After the addition of 1 mL of buffer solution to the residue, the suspension was vortexed for 2 min and stirred for 30 min in a water bath at 60 °C. Afterwards, the suspension was centrifuged (25 min, 22 °C, 3550 rcf) and the supernatant was collected in a 2 mL volumetric flask. The procedure was repeated once more. The collected supernatants in the three volumetric flasks of all three extraction steps were filled up to 2 mL with the respective solvents, filtered (0.45 µm), and separated by ultra-high-performance liquid chromatography (UHPLC). Specifications on UHPLC separation are provided in Supplementary Material S1.

SDSS-GMP Fractionation

To obtain the SDS-soluble proteins, 100 mg of flour was extracted using 1 mL of SDS solution (1% (w/v) SDS in 0.05 mol/L NaH_2PO_4 (pH 6.9)). After vortexing for 2 min and stirring for 30 min at 22 °C, the sample was centrifuged (25 min, 22 °C, 3550 rcf). The supernatant was collected in a 5 mL volumetric flask and the procedure was repeated one more time. Afterwards, the GMP fraction was extracted from the residue using 1 mL of GMP extraction solution (50% (v/v) propan-1-ol, 0.05 mol/L Na_2HPO_4/KH_2PO_4 (pH 7.5) and 1% (w/v) DTT). The suspension was vortexed for 2 min and stirred for 30 min in a water bath at 60 °C before centrifugation (25 min, 22 °C, 3550 rcf). The supernatant was collected in a 2 mL volumetric flask and the extraction was repeated one more time. The collected supernatants in the two volumetric flasks were filled up with the respective solvents, filtered (0.45 µm), and separated by UHPLC. Specifications on UHPLC separation are provided in Supplementary Material S1.

Table 2 Descriptive statistics and squared Pearson correlation coefficient of individual quality parameters and protein content

Parameter	$r^2_{\text{Pearson; protein}}$ (-)	Number of samples (-)	Mean ± SD	Minimum	Maximum	Range
Protein (%)	1.00	50	12.2 ± 1.6	9.2	16.5	7.3
Wet gluten (%)	0.83	50	27.9 ± 4.2	19.5	38.0	18.6
Gluten Index (-)	0.01	50	93 ± 8	59	100	41
Osborne total (g/100 g)	0.92	50	11.91 ± 1.63	9.38	16.54	7.17
Albumin and globulin (mg/g)	0.24	50	23.11 ± 2.12	18.76	28.55	9.80
Gliadin (mg/g)	0.93	50	65.30 ± 10.97	47.02	98.18	51.16
Glutenin (mg/g)	0.79	50	30.70 ± 4.45	23.47	41.96	18.49
Gliadin/glutenin (-)	0.14	50	2.13 ± 0.18	1.70	2.67	0.97
SDSS/GMP total (g/100 g)	0.83	50	10.82 ± 1.53	8.00	14.95	6.95
SDSS (mg/g)	0.72	50	93.14 ± 11.11	75.47	126.34	50.87
GMP (mg/g)	0.63	50	15.08 ± 5.70	2.83	28.31	25.48
GMP-HMW (mg/g)	0.48	50	3.75 ± 1.66	0.40	7.22	6.82
GMP-LMW (mg/g)	0.66	50	11.33 ± 4.18	2.43	21.47	19.04
Hagberg falling number (s)	0.16	50	411 ± 82	296	668	372
Starch damage (Ai%)	0.09	50	95.24 ± 0.50	94.34	96.58	2.24
SRC water (%)	0.10	50	66.3 ± 3.3	56.3	73.2	16.9
SRC sucrose (%)	0.46	50	106.0 ± 5.2	94.8	119.8	25.0
SRC lactic acid (%)	0.29	50	132.2 ± 13.8	98.1	157.5	59.4
SRC sodium carbonate (%)	0.00	50	89.8 ± 6.3	75.3	104.1	28.8
FA DDT (min)	0.50	50	2.70 ± 2.05	1.18	8.35	7.17
FA WAM (%)	0.30	50	58.2 ± 2.6	54.3	65.3	11.0
FA S (min)	0.55	49	7.31 ± 5.46	1.54	24.77	23.23
FA DS (FE)	0.63	50	47 ± 28	3	112	109
FA FQN (-)	0.49	49	80 ± 64	23	291	268
EX energy 45 min (cm^2)	0.66	49	100 ± 25	40	163	123
EX resistance 45 min (BU)	0.01	49	313 ± 44	195	408	213
EX extensibility 45 min (mm)	0.74	49	164 ± 25	125	231	106
EX maximum 45 min (BU)	0.30	49	447 ± 73	208	560	352
EX ratio number 45 min (-)	0.29	49	2.0 ± 0.4	1.2	3.0	1.9
EX ratio number (Max.) 45 min (-)	0.03	49	2.8 ± 0.5	1.7	3.7	2.1
AL C_{max} (-)	0.72	50	2542 ± 444	1834	3742	1908
AL P (mm H_2O)	0.20	50	92 ± 17	62	138	76
AL L (mm)	0.28	50	93 ± 23	47	141	94
AL G (-)	0.27	50	21.2 ± 2.7	15.2	26.4	11.2
AL W (10^{-4} J)	0.74	50	276 ± 82	145	530	385
AL P/L (-)	0.02	50	1.07 ± 0.44	0.52	2.94	2.42
AL Ie (%)	0.75	50	53.8 ± 5.7	37.4	69.8	32.4
AL K (-)	0.10	50	4766 ± 746	3445	6199	2754
AL SH (-)	0.46	50	1.78 ± 0.09	1.51	1.99	0.48
AL D_{min} (-)	0.10	50	−2.49 ± 0.48	−3.77	−1.82	1.95
AL D_{max} (-)	0.16	50	6.72 ± 0.70	5.44	8.41	2.97

Flour Fractionation and Dough Preparation

Air jet sieving of flour samples was performed using the 200LS-N Hosokawa Alpine AG (Augsburg, Germany) air jet sieve machine at 2500–2600 Pa. Five different sieves (mesh sizes: 32 µm, 50 µm, 75 µm, 100 µm) were used to sieve 70 g of flour into the following five sieve fractions: < 32 µm, 32–50 µm, 50–75 µm, 75–100 µm, > 100 µm. Each fraction was sieved for 10 min before the flour remaining on the sieve was transferred to the next largest sieve.

Dough was prepared in a 50 g farinograph using 30 g of flour (14% moisture) and 0.60 g of salt according to ICC No.

115 (30 °C, 63 rpm, optimum water absorption WAM) until the dough development time was reached. Two 15 g dough pieces were washed out with a glutomatic 2202 (Ing. Stefan Kastenmüller GmbH, Martinsried, Germany) according to ICC No. 155 using 80 µm metal sieves. The four obtained gluten pieces as well as the remaining dough were covered and rested in a temperature-controlled chamber for 10 min at 25 °C before they were frozen at −28 °C. This procedure was repeated three times in total to obtain three dough pieces and six gluten pieces for each flour sample. Furthermore, the collected starch slurry was centrifuged at 3046 rcf for 1 min using a VWR Mega Star 600R centrifuge (VWR international GmbH, Darmstadt, Germany). The precipitate was also frozen at −28 °C and the supernatant discarded. All frozen samples were freeze-dried for 24 h in a Christ Alpha 1–4 (Martin Christ Gefriertrocknungsanlagen GmbH, Osterode, Germany) on the day after preparation. The freeze-dried samples were milled using the Ultra Centrifugal Mill ZM 200 (0.5 mm sieve, 12,000 rpm) (Retsch GmbH, Haan, Germany).

Flours and sieve fractions were stored in airtight containers in the dark at around 15 °C to minimize sample changes. Freeze-dried dough, gluten, and starch samples were additionally stored in aluminum containers to prevent permeation of water vapor. They were brought to room temperature (21 ± 1 °C) 24 h prior to spectroscopic analyses.

Spectroscopic Analysis

Near-infrared spectra were recorded using the MPA (Bruker Optik GmbH, Ettlingen, Germany) in diffuse reflectance mode as absorbance spectra. A rotating cup (quartz glass, Ø 5.1 cm) was used to measure samples with a resolution of 8 cm^{-1}, a scanner velocity of 10 kHz, and 64 scans per measurement. Six measurements were performed for every sample. In between measurements, samples were thoroughly mixed. Before each measurement, a slight compaction was carried out by lightly tapping the filled cup on the table for ten times.

Chemometric Analysis

All analyses were performed using MATLAB (R2021b, The MathWorks, Inc). The workflow is presented in Fig. 1 and described in the following sections. As each of the 50 flour samples was fractionated into five sieve fractions, gluten, starch, and dough, spectra of a total of 450 samples were recorded. Due to the sixfold determination of spectra, a total of 2700 spectra were available for analysis.

NIR spectra were preprocessed using a number of different preprocessing routines including Savitzky-Golay smoothing (order, 2; frame length, 25) and differentiation (order, 2; frame length, 13) filters, detrending (order, 2), highpass filter (as described by Mburu et al. (2021)), standard normal variate transformation (SNV), multiplicative scatter correction (MSC), extended multiplicative scatter correction (EMSC), min-max normalization, 1. norm and 2. norm normalization, and autoscaling of variables. Different individual and combined preprocessing approaches of the methods named above were tested. In every case, mean spectra were calculated after preprocessing for each of the 50 samples of every sample type (flour, sieve fractions, gluten, starch, dough) and subsequently used for further analyses.

Principal component analysis (PCA) of spectra of flour, flour fractions, and dough was performed to analyze score and loading plots for spectral differences. Prediction models for all 41 parameters of different rheological and analytical measurements listed in Table 1 were calculated using NIR spectra of flour, flour fractions, and dough. Two different model-building approaches were followed after spectral preprocessing. In the first approach (a), the prediction models were calculated based on the individual spectra of flour, flour fractions, and dough. In the second approach (b), data fusion was carried out prior to the model-building process by concatenating mean preprocessed spectra of different sample types. Data fusion was only tested within a particular fractionation method. The reason is that the fusion of spectra across different fractionation methods (i.e., spectra of sieve fractions and spectra of dough, gluten, or starch) would drastically increase the amount of work and equipment needed to obtain all necessary spectra for prediction, making the procedure irrelevant for practical application.

After preprocessing and data fusion, principal component regression (PCR) and partial least squares regression (PLSR) were tested as prediction models for every of the 41 parameters. One to ten components were tested for modeling. For every model, a leave-one-out cross-validation (LOOCV) was performed and the predictive ability was assessed by calculating metrics according to Eqs. 1 to 3. The best models were determined by the minimum root mean squared error of cross-validation (RMSECV). To compare the predictive performance of different models for the same parameter, the percentage improvement or worsening of the RMSECV I_{RMSECV} was calculated according to Eq. 4. Positive values of I_{RMSECV} indicate that the RMSECV of the model data is improved compared to the model reference, while negative values indicate that the RMSECV is worsened. Especially in case of data fusion, it is always important to assess the effect of data fusion on the predictive performance, as data fusion is only beneficial if it actually leads to an improved prediction compared to models based on the individual data. To assess the effect of data fusion using Eq. 4, the minimum RMSECV of the two individual models was used for the $RMSECV_{model\ reference}$. For example, model 1 was based on data 1 and yielded RMSECV 1. Model 2 was based on data 2 and yielded RMSECV 2. Model 3 was based on the

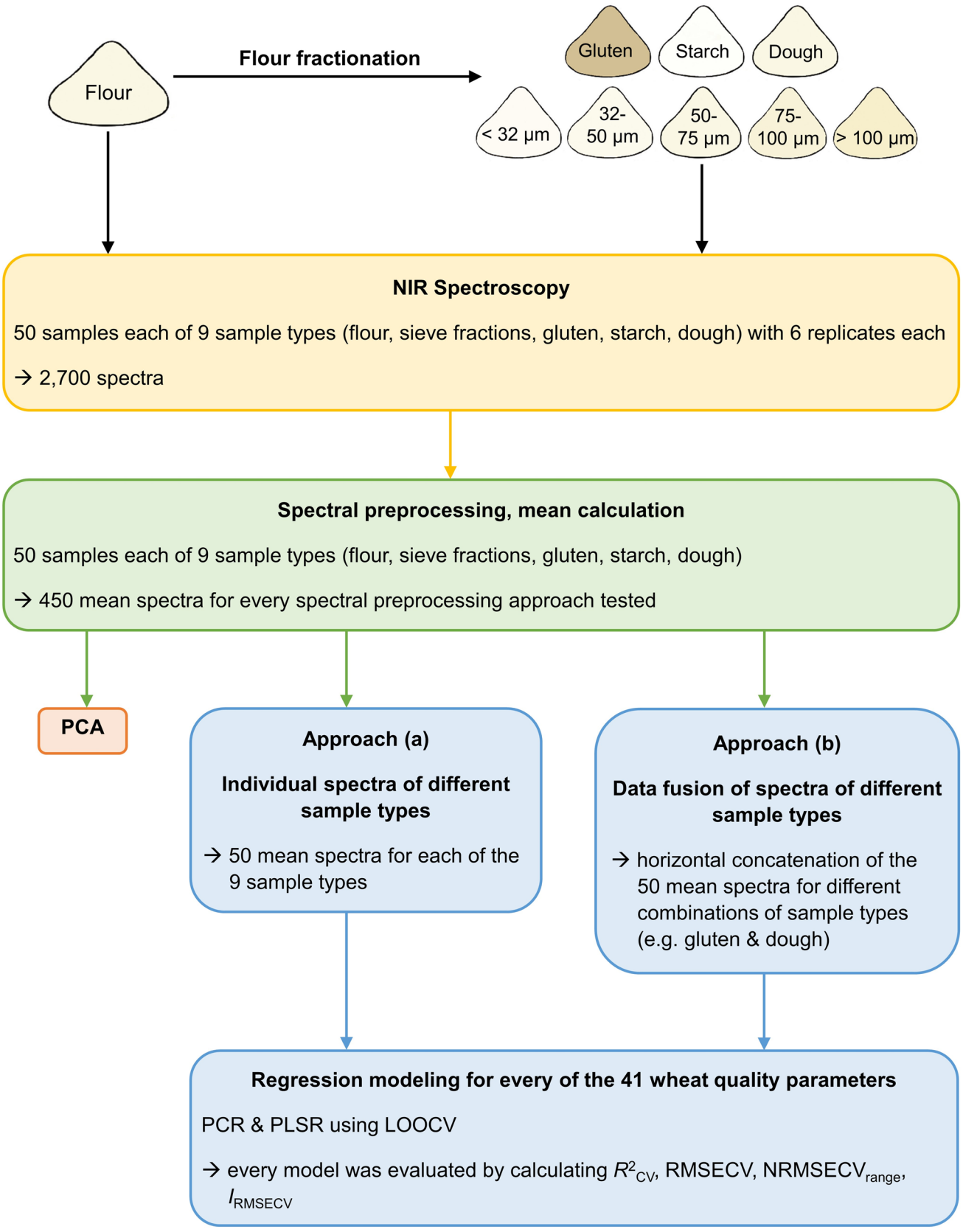

Fig. 1 Workflow for the chemometric analysis

concatenated data 1 and 2 and yielded RMSECV 3. Then, RMSECV 3 would be used for $RMSECV_{model\ data}$ and the minimum value of RMSECV 1 and RMSECV 2 would be used for $RMSECV_{model\ reference}$.

The following metrics were calculated for every model of all 41 wheat quality parameters listed in Table 2.

$$R^2_{CV} = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2} \quad (1)$$

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{n}} \quad (2)$$

$$NRMSECV_{range}\,[\%] = \frac{RMSECV}{y_{max} - y_{min}} \cdot 100\,\% \quad (3)$$

$$I_{RMSECV}\,[\%] = 100\% - \frac{RMSECV_{model\ data}}{RMSECV_{model\ reference}} \cdot 100\,\% \quad (4)$$

where R^2_{CV} is the coefficient of determination of cross-validation, RMSECV is the root mean squared error of cross-validation, $NRMSECV_{range}$ is the range normalized RMSECV, n is the total number of samples of the parameter that is predicted (49 or 50 samples according to Table 2), y_i is the measured value of the i-th sample, $\hat{y}_i$ is the predicted value of the i-th sample, $\bar{y}$ is the mean value of the sample set (listed for each of the 41 parameters in Table 2), y_{max} is the maximum value of the sample set (see Table 2), y_{min} is the minimum value of the sample set (see Table 2), I_{RMSECV} is the improvement of the RMSECV, and $RMSECV_{model\ data}$ is the RMSECV of the model that is to be compared to the RMSECV of the reference model $RMSECV_{model\ reference}$.

It should be emphasized that in this study, the R^2_{CV} was specified as the coefficient of determination according to Eq. 1. As an important difference compared to the squared Pearson correlation coefficient, which is also frequently reported in the literature, this metric can actually become negative. A negative coefficient of determination shows that the predictions based on the model are very poor, even worse than if the mean value of the sample set was simply used for the prediction.

Results and Discussion

Wheat quality predictions for a total of 41 parameters covering a broad spectrum of established analytical and rheological analyses for flour characterization (protein-related parameters, Hagberg falling number, starch damage, SRC, farinograph, extensograph, alveograph) were calculated based on spectra of flour, flour fractions, and dough with and without data fusion. The discussion of the obtained results is divided into four parts. In the first section, the spectral differences of flour, flour fractions, and dough are presented. In the second part, the potential of flour fractionation and data fusion for improved prediction of wheat quality is evaluated using the calculated improvement of the RMSECV I_{RMSECV}. Next, the individual quality parameters and their achieved predictive performances are discussed in more detail. In the final section, a general discussion of influencing factors on the predictions is presented.

Analysis of Spectral Difference of Flour Fractions and Dough

Example spectra of flour, sieve fractions, gluten, starch, and dough are presented in Fig. 2. Spectra were SNV transformed to remove the influence of particle size differences. It can be seen that the general characteristics of the spectra appear very similar to the naked eye and that only the gluten spectrum shows clear differences compared to the other spectra.

To investigate the spectral differences in more detail, two separate PCAs were performed for the two fractionation methods to analyze their effects on the composition and spectra of the obtained fractions separately. Score and loading plots are shown in Fig. 3. In the loading plots, four different spectral regions (I–IV) are marked for further discussion. Loadings were interpreted with the help of a literature review and related to possible differences in the composition of flour, flour fractions, and dough. However, NIR spectra are characterized by broad and overlapping peaks and one molecular vibration usually contributes to multiple peaks in the spectrum because of different combination and overtone vibrations (Workman and Weyer 2012). Consequently, clear peak assignments are difficult, which is why the assignments found in the literature are sometimes contradictory. For this reason, the discussion of peak assignments in the following sections of this study does not claim to be exhaustive and was kept more general.

PCA of Flour, Dough, Gluten, and Starch Spectra

A clear separation of scores can be observed for the spectra of flour, gluten, starch, and dough along the principal components PC1 and PC2. PC3, on the other hand, does not seem to separate the spectra of different fractions, but the spectra of different samples within each fraction.

The loadings of PC2, which clearly separates flour from the other fractions, show two interesting peaks in region II (around 5100 cm^{-1}) and in region IV (around 7100 cm^{-1}). Both peaks are known to be related to the OH vibrations of water (Bruun et al. 2007; Salgó and Gergely 2012; Workman

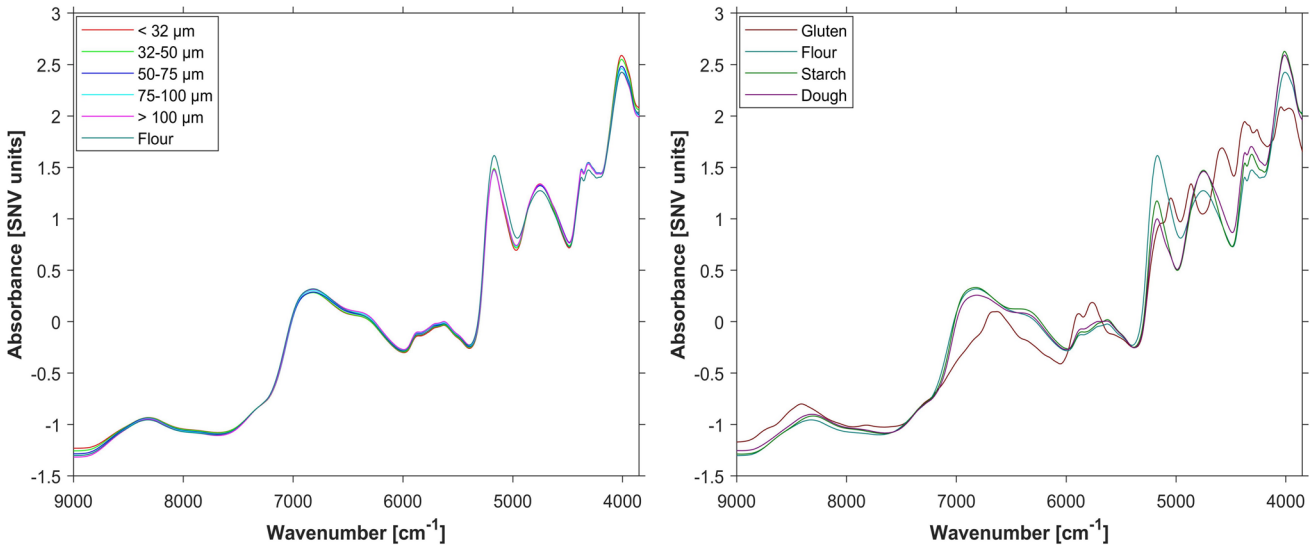

Fig. 2 Mean NIR spectra of all 50 flour samples for sieve fractions and flour (left) and dough, gluten, starch, and flour (right). Spectra were pre-processed by SNV transformation

and Weyer 2012) and therefore this separation along PC2 is thought to be caused by differences in the water content. Because gluten, starch, and dough have been freeze-dried, the water content of these fractions is much lower compared to the water content of flour. The scores of the gluten spectra are somewhat closer to flour, because gluten showed hygroscopic behavior during the measurements, which caused the water content to increase after freeze-drying.

The loadings of PC1 show many peaks related to protein vibrations. Peaks in region I (around 4200–4500 cm^{-1}) are known to be related to CH vibrations. They originate from proteins, carbohydrates, and also lipids. However, peaks in this region show high sensitivity to protein secondary structure (α-helix, β-sheet, random coil) and are also influenced by amino acid side chain vibrations and their microenvironment (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012). Peaks in region II (around 4900 cm^{-1}) originate from protein NH vibrations, but they are located at the flank of the water peak and are therefore masked by water signals (Workman and Weyer 2012). Signals in region III (around 5750–5900 cm^{-1}) are due to CH and SH vibrations. Again, many protein signals originating from different protein secondary structures (α-helix, β-sheet), cysteine, and amino acid side chain vibrations as well as their microenvironment and their interactions are responsible for signals in this region (Bruun et al. 2007; Workman and Weyer 2012). The separation of sample types along PC1 is in accordance with this observation. Gluten has a much higher protein content than the other samples, which is why its distance from the other scores is greatest. However, the protein content of dough is also higher than the protein content of flour, because its water content was reduced by freeze-drying, which is why the scores of dough are closest to the scores of gluten. Starch is protein depleted, which is why it has the most negative scores among the flour fractions. It is interesting to note that the scores of flour cannot be separated from the scores of starch along PC1, although the protein content of flour is clearly higher than the protein content of starch. This could indicate that the separation of dough and gluten from starch and flour along PC1 is not only caused by the increased protein content of dough and gluten. It might also be influenced to a large extent by the changes in protein structure and changes in the interactions of proteins with other flour constituents caused by the development of the gluten network during dough formation. This observation is supported by the fact that other researchers have successfully monitored differences in protein structure caused, e.g., by gluten hydration, denaturation, wheat maturation, and dough formation using NIR spectroscopy (Bruun et al. 2007; Salgó and Gergely 2012; Wesley et al. 1998; Workman and Weyer 2012). Despite the general non-specificity of NIR peaks, the results of the study show that NIR spectra contain a large amount of information about the protein structures and their changes in various processes. This information could in turn be useful for predicting rheological wheat dough characteristics in particular.

The loadings of PC3, whose scores distinguish between different samples rather than between fractions, show some similarities to PC1 for peaks especially in regions I, II, and III. This underlines that differences in the samples regarding water content, protein content, and structure, but also differences due to carbohydrates and lipids (e.g., region I, around 4200–4300 cm^{-1}) (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012), can be detected by NIR

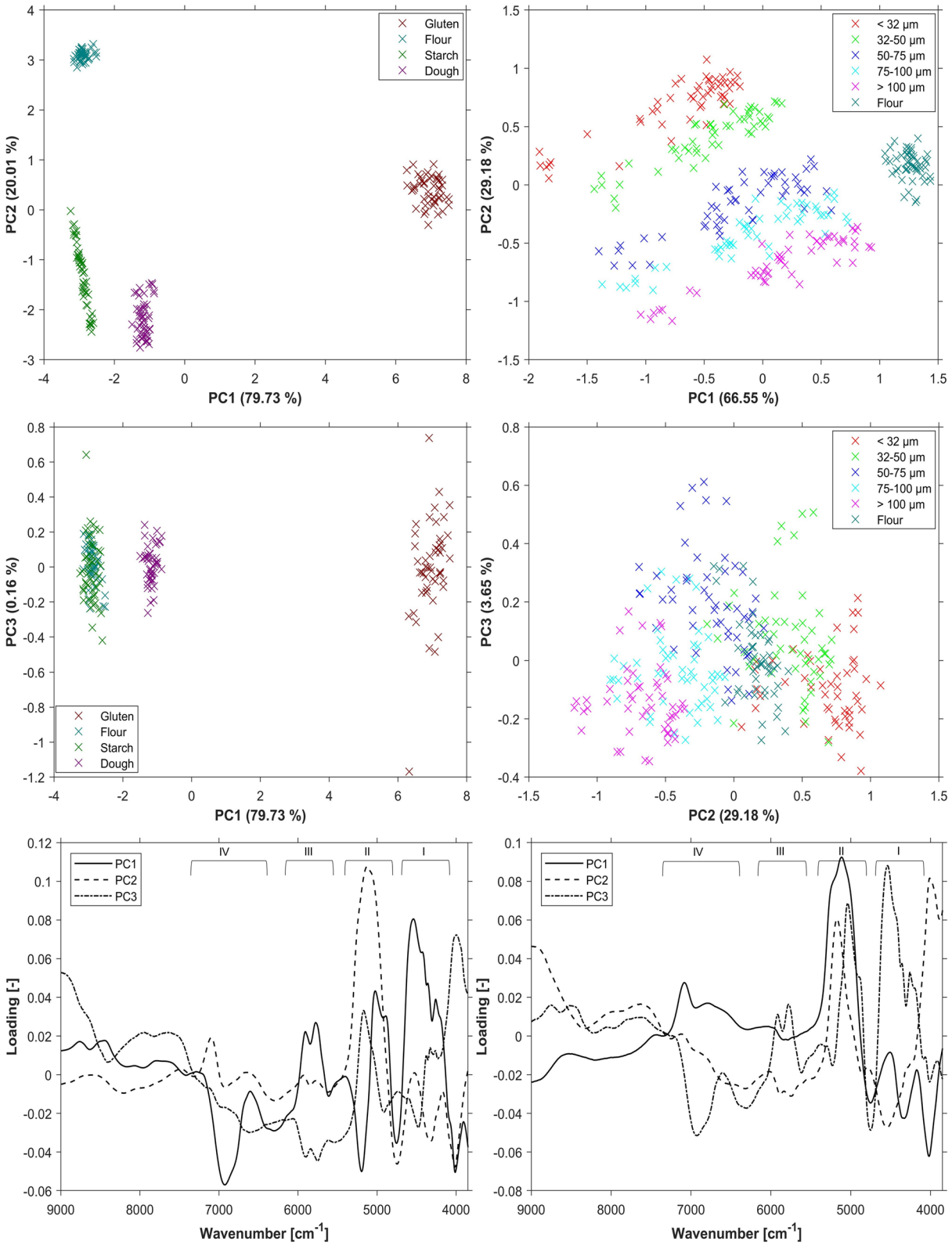

Fig. 3 PCA of spectra of sieve fractions and flour (left column) and of spectra of dough, gluten, starch, and flour (right column). Spectra were SNV transformed before calculation of PCA. Score plots are presented in the upper and middle rows and loading plots with indicated regions for discussion I–IV are shown in the lower row

spectroscopy and lead to a differentiation of samples. It is interesting to note that the PC3 scores, especially for gluten, but also to a certain extent for starch and dough, show greater scattering than the PC3 scores of flour. This could indicate that the differences between the samples could be emphasized and amplified by the fractionation procedure, which could also have a positive effect on wheat quality predictions if these differences are related to differences in the property of interest that is to be predicted.

PCA of Flour and Sieve Fractions

The scores of flour spectra can be clearly separated from the scores of the sieve fractions by PC1. PC2 and PC3, on the other hand, are responsible for a separation of different sieve fractions, although neighboring particle size ranges overlap to a certain extent (more for PC3 than for PC2). For PC2, the sieve fractions show a linear trend towards lower scores with increasing particle size. The scores of PC3 show a curvature trend for the sieve fractions, with the 50–75 µm fraction reaching the maximum with the highest score values. When comparing the loadings for PC1 to PC3 of this PCA to the loadings of the PCA of flour, dough, gluten, and starch spectra, it becomes clear that the general trend and the observed regions and peaks are quite similar, even if the order of the loadings is different.

For flour and sieve fractions, loadings of PC1 show peaks related to water vibrations in region II (around 5100 cm^{-1}) and in region IV (around 7100 cm^{-1}) (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012). This is in accordance with the reduced water content of the sieve fractions compared to flour, which is caused by the sieving operation. Furthermore, the score values of PC1 tend to become smaller with decreasing particle size. Also, some samples tend to have even lower scores than others, because groupings of samples with larger negative score values for PC1 are visible. This might be caused by differences in room temperature and moisture during sieving, because the samples with the most negative scores for PC1 were prepared on days where these conditions were the most different from the other preparation days. This could have resulted in the water content of these samples being even lower than samples that were prepared on other days.

The loadings of PC2 are more difficult to interpret because they do not show many clear peaks throughout the entire wavenumber range. More generally, it seems that differences in water content, proteins, carbohydrates, and lipids are responsible for the differentiation of sieve fractions along PC2. It appears that flour spectra are most similar to the 50–75 µm fraction based on the scores of PC2. However, the exact compositional differences contributing to this differentiation cannot be determined based on this PCA analysis.

The separation of sieve fractions based on PC3 is less clear, because the scores overlap to a greater extent. However, the loadings of PC3 are quite similar to the loadings of PC1 of the PCA of flour, dough, gluten, and starch spectra, especially for region I (around 4200–4500 cm^{-1}) and region III (around 5750–5900 cm^{-1}). As discussed in the corresponding section of the other PCA, these two regions are strongly influenced by protein vibrations and indicate differences in protein secondary structure (α-helix, β-sheet, random coil), amino acids, and their interactions and microenvironment. However, vibrations due to lipids and carbohydrates are also present, in particular in region I (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012). Further analysis showed that the sieve fractions are actually protein-enriched and that this protein enrichment follows the trend observed in the scores of PC3, meaning that the 50–75 µm fraction had the highest protein content (results not shown). This indicates that, besides possible other compositional differences, differences in protein content and possibly also differences in protein structure of these fractions are captured by NIR spectroscopy and contribute to the differentiation along PC3. It has already been observed before that the protein contents of different flour particle sizes differ and that different types of proteins are enriched in different particle size ranges. For example, free "wedge" proteins are milled to really small particle sizes and are therefore present in smaller particle size fractions, whereas "adherent" proteins retain larger particle sizes because they are connected to starch granules and covered by a lipid layer (Coulson and Sim 1965; Jones et al. 1959). It is possible that not only the protein enrichment but also the separation of proteins based on other characteristics (e.g., structural differences) by sieve fractionation can contribute to improved wheat quality predictions based on NIR spectra.

Lastly, it can again be observed that the scattering of scores along PC1, PC2, and PC3 is greater for the sieve fractions than for flour. For PC1, the main cause seems to be differences in water content, as discussed above, but for PC2 and PC3, it appears that other differences between the samples may be emphasized and amplified by the sieving procedure. This could also have a positive effect on the predictions if these differences relate to differences in the property of interest that is to be predicted.

Evaluation of the Potential of Flour Fractionation and Data Fusion to Improve Wheat Quality Predictions Based on the Improvement of the RMSECV I_{RMSECV}

For all 41 wheat quality parameters tested in this study, prediction models using spectra of flour and flour fractions obtained by the two fractionation methods (sieve fractionation; dough preparation followed by gluten washing)

with and without data fusion methods (approach (b) and (a), respectively) were investigated. The best predictions achieved by approach (a) using individual spectra of flour, flour fractions, and dough are presented in Table 3. Table 4 lists improved predictions obtained by data fusion in approach (b) and finally Table 5 gives an overview of the overall best predictions comprising methods from Table 3 and Table 4.

This part of the discussion is not intended to focus on the individual quality parameters and the corresponding predictions, but on the general ability of flour fractionation and data fusion as novel methods to improve wheat quality predictions. Consequently, the following sections discuss and use the improvement of the RMSECV I_{RMSECV} to compare models to demonstrate the general potential of the used methods for improved predictions. A more detailed discussion about individual parameters can be found in the section "Analysis of the Predictive Performance of the Best Models for Individual Wheat Quality Parameters."

The results in Table 3 show that individual spectra of flour fractions and dough significantly changed the predictive performance of models compared to flour spectra. Dough preparation and gluten washing, as one of the fractionation methods, enabled an improved prediction for 14 of the 41 parameters with the improvement of the RMSECV $I_{RMSECV} \geq 5.0\%$ compared to flour spectra, with I_{RMSECV} ranging between 5.0 and 25.0%. Especially, the spectra of dough and also the spectra of gluten were able to achieve the improvements. Starch spectra never achieved an improvement $\geq 5.0\%$ compared to flour spectra; in fact, the prediction usually deteriorated. In comparison, the individual spectra of the sieve fractionation method improved the prediction of 27 of the 41 parameters by at least 5.0% compared to flour spectra (I_{RMSECV} ranged from 5.0 to 17.9%). The spectra of the 50–75 µm and the 75–100 µm fractions achieved the best predictions most frequently here.

Table 4 shows that data fusion can be a valuable method to achieve further improvements in the predictions. Improvements of the RMSECV I_{RMSECV} of at least 5.0% compared to flour spectra and individual spectra of the corresponding fractionation method as well as compared to the corresponding models of the individual constituents of the fused data model were achieved for many different parameters and both fractionation methods. The dough preparation and gluten washing method combined with data fusion was able to improve the prediction of 20 parameters with $I_{RMSECV} \geq 5.0\%$, whereby in particular the combination of gluten and flour spectra as well as the combination of gluten and dough spectra contributed to this. In contrast, data fusion models of sieve fractions achieved improvements of at least 5.0% for 13 of the 41 parameters. In most cases, a combination of two spectra of sieve fractions also presented the greatest improvement, although the exact combination of spectra varied for different parameters.

Lastly, Table 5 allows a combined evaluation of the best results regarding the potential of both flour fractionation methods with and without data fusion to achieve improved predictions of wheat quality parameters compared to flour spectra. Only for six out of the 41 parameters, no improvement of the RMSECV $I_{RMSECV} \geq 5.0\%$ was achieved by flour fractionation with or without data fusion. For 35 parameters, the highest achieved improvements I_{RMSECV} compared to flour spectra were in the range of 5.6–28.6%, although I_{RMSECV} was actually $\geq 15.0\%$ for many of them. For 17 of these parameters, dough preparation followed by gluten washing was the fractionation method that yielded the best results, whereby this was achieved six times by single spectra (especially of dough) and 11 times by combined spectra after data fusion (especially the combinations of gluten and flour as well as gluten and dough). Sieve fractionation achieved the best predictions for 16 parameters, seven times by single spectra (especially 50–75 µm and 75–100 µm fraction) and nine times by combined spectra after data fusion (especially combinations of two sieve fractions). For the parameters "SRC sucrose" and "EX energy 45 min," both fractionation methods yielded equally good improvements.

In conclusion, both fractionation methods tested in this study contributed almost equally to the overall best models obtained. Spectra of sample types on which the best models were most frequently based were dough and gluten as well as the 50–75 µm and 75–100 µm fractions. The overall best models achieved were dominated by models based on combined spectra after data fusion. The observed predictive qualities are a result of the changes in composition caused by the fractionation methods (separation as well as chemical and enzymatic reactions), which in turn alter spectral signals, as also presented and discussed in the section "Analysis of Spectral Difference of Flour Fractions and Dough." As a result, relevant spectral signals for the prediction of the property of interest may be enhanced or the superimposition of irrelevant and interfering signals may be reduced in the spectra of flour fractions and dough compared to the spectra of flour, which could be a possible reason for improved predictions. For example, further analyses have shown that the 50–75 µm fraction is more protein-enriched compared to the other sieve fractions (results not shown). It is known from literature that proteins with different properties enrich in different particle size ranges (Coulson and Sim 1965; Jones et al. 1959). Possibly, the NIR signals of proteins enriched in this sieve fraction are more relevant for the prediction of certain wheat quality parameters. Furthermore, NIR spectra of gluten and dough could contain additional valuable information about rheological properties, because NIR spectroscopy is able to capture information related to protein secondary structure (changes) and interactions of proteins, water,

Table 3 Cross-validation results for overall best predictions using individual spectra of flour, sieve fractions and dough, gluten, and starch in approach (a). For every prediction model, the improvement of the RMSECV I_{RMSECV} compared to flour spectra is also presented. Highlighted in bold letters are results with $R^2_{CV} \geq 0.70$ for flour spectra and results with $R^2_{CV} \geq 0.70$ and $I_{RMSECV} \geq 5.0\%$ for fractions spectra. Best predictions for every of the nine individual sample types are presented in Supplementary Material S1

Parameter	Flour		Dough, gluten, starch				Sieve fractions			
	R^2_{CV} (-)	RMSECV	Spectrum (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV}* (%)	Spectrum (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV}* (%)
Protein (%)	**0.96**	**0.3**	Dough	0.96	0.4	−33.3	75–100 µm	0.95	0.4	−33.3
Wet gluten (%)	**0.93**	**1.2**	Dough	0.92	1.2	0.0	**50–75 µm**	**0.94**	**1.0**	**16.7**
Gluten Index (-)	0.03	7	Gluten	0.33	6	14.3	75–100 µm	0.41	6	14.3
Osborne total (g/100 g)	**0.86**	**0.60**	**Dough**	**0.88**	**0.57**	**5.0**	**75–100 µm**	**0.89**	**0.53**	**11.7**
Albumin and globulin (mg/g)	0.20	1.88	Dough	0.35	1.70	9.6	50–75 µm	0.43	1.59	15.4
Gliadin (mg/g)	**0.90**	**3.41**	Dough	0.90	3.44	−0.9	**> 100 µm**	**0.93**	**2.92**	**14.4**
Glutenin (mg/g)	**0.79**	**2.03**	Dough	0.73	2.28	−12.3	75–100 µm	0.79	2.02	0.5
Gliadin/glutenin (-)	0.47	0.14	Dough	0.15	0.17	−21.4	75–100 µm	0.25	0.16	−14.3
SDSS/GMP total (g/100 g)	**0.94**	**0.37**	Dough	0.94	0.39	−5.4	**75–100 µm**	**0.95**	**0.33**	**10.8**
SDSS (mg/g)	**0.85**	**4.32**	**Dough**	**0.91**	**3.24**	**25.0**	**75–100 µm**	**0.88**	**3.84**	**11.1**
GMP (mg/g)	**0.72**	**3.00**	**Dough**	**0.80**	**2.52**	**16.0**	**32–50 µm**	**0.77**	**2.71**	**9.7**
GMP-HMW (mg/g)	0.69	0.92	Dough	0.71	0.88	4.3	50–75 µm	0.72	0.88	4.3
GMP-LMW (mg/g)	**0.72**	**2.20**	**Dough**	**0.80**	**1.85**	**15.9**	**32–50 µm**	**0.78**	**1.95**	**11.4**
Hagberg falling number (s)	**0.72**	**44**	Gluten	0.49	58	−31.8	< 32 µm	0.64	49	−11.4
Starch damage (Ai%)	**0.87**	**0.18**	Dough	0.74	0.25	−38.9	50–75 µm	0.81	0.22	−22.2
SRC water (%)	0.66	2.0	Starch	0.60	2.1	−5.0	**< 32 µm**	**0.76**	**1.7**	**15.0**
SRC sucrose (%)	**0.73**	**2.7**	Dough	0.65	3.1	−14.8	**50–75 µm**	**0.77**	**2.5**	**7.4**
SRC lactic acid (%)	0.67	7.9	Dough	0.41	10.5	−32.9	32–50 µm	0.62	8.4	−6.3
SRC sodium carbonate (%)	0.61	3.9	Starch	0.48	4.5	−15.4	50–75 µm	0.68	3.6	7.7
FA DDT (min)	**0.80**	**0.91**	Dough	0.77	0.98	−7.7	75–100 µm	0.71	1.09	−19.8
							> 100 µm	0.71	1.09	−19.8
FA WAM (%)	**0.79**	**1.2**	Dough	0.66	1.6	−33.3	**50–75 µm**	**0.83**	**1.1**	**8.3**
FA S (min)	0.59	3.45	Dough	0.67	3.08	10.7	> 100 µm	0.65	3.19	7.5
FA DS (FE)	**0.77**	**14**	**Dough**	**0.81**	**12**	**14.3**	**50–75 µm**	**0.79**	**13**	**7.1**
FA FQN (-)	0.53	43	Dough	0.61	39	9.3	> 100 µm	0.58	41	4.7
EX energy 45 min (cm^2)	0.67	15	Dough	0.67	15	0.0	**32–50 µm**	**0.71**	**14**	**6.7**
EX resistance 45 min (BU)	0.05	43	Gluten	0.26	38	11.6	50–75 µm	0.15	40	7.0
EX extensibility 45 min (mm)	**0.75**	**12**	Dough	0.75	13	−8.3	> 100 µm	0.76	12	0.0
EX maximum 45 min (BU)	0.35	58	Gluten	0.48	52	10.3	50–75 µm	0.41	55	5.2
EX ratio number 45 min (-)	0.29	0.4	Dough	0.36	0.4	0.0	> 100 µm	0.39	0.4	0.0
EX ratio number (Max.) 45 min (-)	0.07	0.5	Gluten	0.22	0.5	0.0	50–75 µm	0.21	0.5	0.0
AL C_{max} (-)	**0.81**	**190**	Dough	0.80	197	−3.7	50–75 µm	0.82	187	1.6
AL P (mm H_2O)	0.68	10	Dough	0.60	11	−10.0	**75–100 µm**	**0.77**	**9**	**10.0**

Table 3 (continued)

Parameter	Flour		Dough, gluten, starch				Sieve fractions			
	R^2_{CV} (-)	RMSECV	Spectrum (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV}* (%)	Spectrum (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV}* (%)
AL L (mm)	0.45	17	Dough	0.50	16	5.9	75–100 µm	0.57	15	11.8
AL G (-)	0.44	2.0	Dough	0.45	2.0	0.0	< 32 µm	0.53	1.9	5.0
							75–100 µm	0.53	1.9	5.0
AL W (10^{-4} J)	**0.87**	**30**	**Dough**	**0.91**	**24**	**20.0**	50–75 µm	0.86	31	−3.3
AL P/L (-)	0.28	0.38	Starch	0.26	0.38	0.0	50–75 µm	0.46	0.33	13.2
AL Ie (%)	**0.75**	**2.9**	Dough	0.77	2.8	3.4	**75–100 µm**	**0.82**	**2.5**	**13.8**
AL K (-)	**0.78**	**350**	Dough	0.69	413	−18.0	**50–75 µm**	**0.83**	**309**	**11.7**
AL SH (-)	0.44	0.07	**Gluten**	**0.70**	**0.06**	**14.3**	75–100 µm	0.67	0.06	14.3
AL D_{min} (-)	**0.71**	**0.26**	Dough	0.52	0.34	−30.8	**75–100 µm**	**0.79**	**0.23**	**11.5**
AL D_{max} (-)	0.68	0.39	Starch	0.63	0.43	−10.3	**75–100 µm**	**0.79**	**0.32**	**17.9**

*Improvement of the RMSECV I_{RMSECV} compared to flour spectra

and starch (Alava et al. 2001; Albanell et al. 2012; Bruun et al. 2007; Wesley et al. 1998). Other studies have already shown that NIR spectra of dough can contribute to improved predictions of wheat baking quality (Gabriel et al. 2017; Ziegler et al. 2025b). However, further analyses of the exact compositions of the different flour fractions are necessary to draw more detailed conclusions from the composition to the predictive ability. Another factor influencing the prediction quality is the data fusion strategy applied. Especially the complex quality characteristics (e.g., rheological parameters) depend on many different flour components and their interactions in specific processes such as kneading. In some cases, data fusion of different spectra may yield improved predictions if the components and signals relevant for the prediction of a specific parameter have been separated into several fractions by the fractionation process. Thus, when complementary information contained in spectra of different fractions is combined in a prediction model by data fusion, the predictive performance is improved. Data fusion methods in general have already proven useful in the food context (Borràs et al. 2015) and also particularly for wheat quality predictions, as Nagel-Held et al. (2022) achieved improved predictions when flour spectra of different spectroscopic methods were combined. Ziegler et al. (2025b) showed the potential of different data fusion approaches for an improved prediction of wheat baking quality, also involving NIR spectra of flour fractions and dough. However, this study is the first to apply data fusion of spectra of flour fractions and dough for the prediction of a large number of analytical and rheological parameters. In summary, the results of this study underline the strong potential of flour fractionation and data fusion for the improvement of a wide range of wheat quality predictions compared to predictions based on flour spectra.

Analysis of the Predictive Performance of the Best Models for Individual Wheat Quality Parameters

Wheat can be analyzed for its quality in the grain stage as well as in the flour stage after milling using many different characterization methods. Wheat flour quality is not only influenced by the characteristics of the cereal grain, but for example also by milling conditions and flour extraction. It can be defined by various analytical and rheological parameters that characterize both flour composition and flour functionality. In this study, a total of 41 analytical and rheological wheat flour and dough quality parameters were predicted, including protein-related parameters, starch properties, solvent retention capacity analysis as well as the most commonly performed rheological analyses (farinograph, extensograph, alveograph). The best models for all parameters are presented in Table 5. The best predictions obtained with flour spectra are compared with the best predictions obtained with spectra of flour fractions with or without data

Table 4 Cross-validation results for overall best predictions achieved using data fusion for both fractionation methods in approach (b). Only results of data fusion models that achieved an improvement of the RMSECV I_{RMSECV} of ≥ 5.0% compared to flour, individual spectra of the corresponding fractionation method and also compared to the corresponding models of the individual constituents of the fused data model are presented (the latter is indicated in the table). Highlighted in bold letters are results with $R^2_{CV} \geq 0.70$

Parameter	Dough, gluten, starch				Sieve fractions			
	Spectra combination (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV} (%) (compared to)	Spectra combination (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV} (%) (compared to)
Protein (%)								
Wet gluten (%)	**Gluten + flour**	**0.94**	**1.1**	**8.3 (flour)**	**< 32 µm + 50–75 µm**	**0.95**	**0.9**	**10.0 (50–75 µm)**
Gluten Index (-)								
Osborne total (g/100 g)					**32–50 µm + 50–75 µm**	**0.92**	**0.47**	**17.6 (50–75 µm)**
Albumin and globulin (mg/g)								
Gliadin (mg/g)	**Flour + starch**	**0.92**	**3.17**	**7.0 (flour)**	**< 32 µm + 50–75 µm**	**0.94**	**2.58**	**16.0 (50–75 µm)**
					32–50 µm + 50–75 µm	**0.94**	**2.58**	**16.0 (50–75 µm)**
Glutenin (mg/g)	**Gluten + flour**	**0.83**	**1.82**	**10.3 (flour)**				
Gliadin/glutenin (-)								
SDSS/GMP total (g/100 g)					**< 32 µm + 32–50 µm + 50–75 µm**	**0.97**	**0.27**	**10.0 (< 32 µm + 50–75 µm)**
SDSS (mg/g)					**< 32 µm + 50–75 µm**	**0.91**	**3.38**	**12.7 (50–75 µm)**
GMP (mg/g)								
GMP-HMW (mg/g)								
GMP-LMW (mg/g)								
Hagberg falling number (s)								
Starch damage (Ai%)					**< 32 µm + 50–75 µm + flour**	**0.88**	**0.17**	**5.6 (< 32 µm + flour)**
SRC water (%)	Flour + starch	0.69	1.9	5.0 (flour)				
SRC sucrose (%)	**Flour + starch**	**0.77**	**2.5**	**7.4 (flour)**				
SRC lactic acid (%)								
SRC sodium carbonate (%)					**< 32 µm + 50–75 µm**	**0.71**	**3.4**	**5.6 (50–75 µm)**
FA DDT (min)	**Gluten + flour**	**0.87**	**0.73**	**19.8 (flour)**				
FA WAM (%)	**Flour + starch**	**0.83**	**1.1**	**8.3 (flour)**	**< 32 µm + 50–75 µm**	**0.86**	**1.0**	**9.1 (50–75 µm)**
					50–75 µm + 75–100 µm	**0.86**	**1.0**	**9.1 (50–75 µm)**
					50–75 µm + flour	**0.87**	**1.0**	**9.1 (50–75 µm)**
FA S (min)	**Gluten + dough**	**0.72**	**2.84**	**7.8 (dough)**				
FA DS (FE)								
FA FQN (-)	Gluten + dough	0.67	36	7.7 (dough)				
EX energy 45 min (cm^2)	**Gluten + flour**	**0.78**	**12**	**20.0 (flour)**	**< 32 µm + 50–75 µm + 75–100 µm + > 100 µm + flour**	**0.78**	**12**	**7.7 (< 32 µm + 50–75 µm + 75–100 µm + > 100 µm)**
	Gluten + dough	**0.78**	**12**	**20.0 (dough)**				
EX resistance 45 min (BU)	Gluten + flour	0.34	36	5.3 (gluten)				
EX extensibility 45 min (mm)	**Gluten + flour**	**0.83**	**10**	**16.7 (flour)**				
EX maximum 45 min (BU)								
EX ratio number 45 min (-)								
EX ratio number (Max.) 45 min (-)	Gluten + flour	0.42	0.4	20.0 (gluten)				
	Gluten + dough	0.36	0.4	20.0 (gluten)				
AL C_{max} (-)					**< 32 µm + > 100 µm + flour**	**0.86**	**165**	**6.3 (< 32 µm + > 100 µm)**
AL P (mm H_2O)	**Gluten + flour**	**0.78**	**9**	**10.0 (flour)**	**50–75 µm + 75–100 µm**	**0.79**	**8**	**11.1 (75–100 µm)**

Table 4 (continued)

Parameter	Dough, gluten, starch				Sieve fractions			
	Spectra combination (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV} (%) (compared to)	Spectra combination (-)	R^2_{CV} (-)	RMSECV	I_{RMSECV} (%) (compared to)
AL L (mm)	Gluten + flour	0.63	14	17.6 (flour)				
AL G (-)	Gluten + flour	0.62	1.7	15.0 (flour)	< 32 µm + 50–75 µm	0.55	1.8	5.3 (< 32 µm)
					50–75 µm + 75–100 µm	0.57	1.8	5.3 (50–75 µm)
AL W (10^{-4} J)					**50–75 µm + flour**	**0.90**	**26**	**13.3 (Flour)**
AL P/L (-)	Gluten + flour	0.45	0.33	13.2 (flour)				
AL Ie (%)	**Gluten + dough**	**0.85**	**2.3**	**17.9 (dough)**				
AL K (-)								
AL SH (-)	**Gluten + flour**	**0.77**	**0.05**	**16.7 (gluten)**				
	Gluten + dough	**0.80**	**0.05**	**16.7 (gluten)**				
AL D_{min} (-)								
AL D_{max} (-)	**Flour + dough**	**0.72**	**0.37**	**5.1 (flour)**				

fusion. For selected parameters, the predictions of the best models are compared in Fig. 4.

Protein-Related Parameters

The prediction of protein content from NIR flour spectra has been an established technique in the industry for many years because it can be predicted very well with R^2 values usually > 0.90, often > 0.95, as many studies have reported in the past (Dowell et al. 2006; Jirsa et al. 2008; Miralbés 2003, 2004; Mutlu et al. 2011; Nagel-Held et al. 2022, 2024; Williams 2020). Wet gluten content is often predicted as well, because the gluten content is especially relevant for flour quality due to its ability to form the viscoelastic gluten network. Its predictive quality is often slightly lower than protein content, but the R^2 is usually > 0.85 (Dowell et al. 2006; Nagel-Held et al. 2022; Williams 2020). The results of this study are consistent with these results, as protein content and wet gluten content were predicted from flour spectra with R^2_{CV} of 0.96 and 0.93, respectively. Flour fractionation did not improve the prediction of protein content, but the prediction of wet gluten content was improved to an R^2_{CV} of 0.95 by using the combination of spectra from the < 32 µm and the 50–75 µm fractions.

As the focus in the industry shifts more towards protein quality than total protein content, there is a growing interest in predicting protein composition from spectroscopic data to replace the time-consuming and expensive reference measurements (Schuster et al. 2023). However, not many studies have reported predictions for protein composition. In this study, both Osborne fractionation and SDSS-GMP fractionation were performed and predicted. The gliadin and glutenin contents were predicted acceptably to well with an R^2_{CV} of 0.90 and 0.79, respectively, using flour spectra and an R^2_{CV} of 0.94 and 0.83, respectively, using combinations of spectra of sieve fractions. However, the content of albumin and globulin and the gliadin/glutenin ratio were not predicted well with $R^2_{CV} < 0.50$ achieved with both flour and fractions spectra. These results are consistent with those of Dowell et al. (2006) and Schuster et al. (2023), who both showed similar trends for the predictive ability of the above protein fractions. No results for the prediction of SDSS-GMP fractionation were found in the literature. However, this study achieved good results for the prediction of SDSS, GMP, and GMP-LMW fractions using flour spectra (R^2_{CV} of 0.85, 0.72, and 0.72, respectively). Interestingly, for all three protein fractions, spectra of dough achieved the greatest improvement in the prediction (R^2_{CV} of 0.91, 0.80, and 0.80, respectively). The GMP-HMW fraction achieved only a moderate predictive quality using flour spectra with an R^2_{CV} of 0.69 without any improvement by flour fractionation.

In general, the good predictability of protein content is due to the fact that this is an analytical quality parameter for which classical absorbers exist in the NIR region (e.g., CH, SH, but especially NH and $CONH_2$) (Pojić and Mastilović 2013; Workman and Weyer 2012). In this study, the improved models for protein parameters based on spectra of flour fractions were often achieved by a combination of two sieve fractions. Apart from Gluten Index, the 50–75 µm fraction was always included in these models. First, this shows that complementary information about protein composition and concentration is contained in spectra of different fractions, which is why data fusion models often led to the best results. Second, the importance of the 50–75 µm fraction for protein parameters is in accordance with the results presented in the section "Analysis of Spectral Difference of Flour Fractions and Dough." The 50–75 µm fraction showed the highest protein enrichment among the sieve fractions (results not shown), which also reflected in the PCA score plot. One possible reason why the spectra of this fraction are especially relevant for the prediction is that the higher protein content of this fraction simply increased

Table 5 Summary of best predictions achieved using flour spectra and best improvements achieved using either individual spectra of flour fractions (approach (a)) or combined spectra by data fusion (approach (b)). The improvement of the RMSECV I_{RMSECV} compared to flour spectra is indicated. Highlighted in bold letters are results with $R^2_{CV} \geq 0.70$ for flour spectra and results with $R^2_{CV} \geq 0.70$ and $I_{RMSECV} \geq 5.0\%$ for improved predictions. Further model specifications (preprocessing method, model type, number of components) can be found in Supplementary Material S1

Parameter	Best prediction from flour spectra			Improved prediction using spectra of flour fractions and dough (individual spectra or combinations)				
	R^2_{CV} (-)	RMSECV	NRMSECVrange (%)	Spectrum or Combination (-)	R^2_{CV} (-)	RMSECV	$NRMSECV_{range}$ (%)	I_{RMSECV}* (%)
Protein (%)	**0.96**	**0.3**	**4.12**	No improvement ≥ 5.0%				
Wet gluten (%)	**0.93**	**1.2**	**6.46**	**< 32 µm + 50–75 µm**	**0.95**	**0.9**	**4.84**	**25.0**
Gluten Index (-)	0.03	7	17.08	75–100 µm	0.41	6	14.64	14.3
Osborne total (g/100 g)	**0.86**	**0.60**	**8.37**	**32–50 µm + 50–75 µm**	**0.92**	**0.47**	**6.56**	**21.7**
Albumin and globulin (mg/g)	0.20	1.88	19.19	50–75 µm	0.43	1.59	16.23	15.4
Gliadin (mg/g)	**0.90**	**3.41**	**6.67**	**< 32 µm + 50–75 µm**	**0.94**	**2.58**	**5.05**	**24.3**
				32–50 µm + 50–75 µm	**0.94**	**2.58**	**5.05**	**24.3**
Glutenin (mg/g)	**0.79**	**2.03**	**10.98**	**Gluten + flour**	**0.83**	**1.82**	**9.85**	**10.3**
Gliadin/glutenin (-)	0.47	0.14	14.44	No improvement ≥ 5.0%				
SDSS/GMP total (g/100 g)	**0.94**	**0.37**	**5.33**	**< 32 µm + 32–50 µm + 50–75 µm**	**0.97**	**0.27**	**3.89**	**27.0**
SDSS (mg/g)	**0.85**	**4.32**	**8.50**	**Dough**	**0.91**	**3.24**	**6.37**	**25.0**
GMP (mg/g)	**0.72**	**3.00**	**11.78**	**Dough**	**0.80**	**2.52**	**9.89**	**16.0**
GMP-HMW (mg/g)	0.69	0.92	13.50	No improvement ≥ 5.0%				
GMP-LMW (mg/g)	**0.72**	**2.20**	**11.56**	**Dough**	**0.80**	**1.85**	**9.72**	**15.9**
Hagberg falling number (s)	**0.72**	**44**	**11.83**	No improvement ≥ 5.0%				
Starch damage (Ai%)	**0.87**	**0.18**	**8.04**	**< 32 µm + 50–75 µm + flour**	**0.88**	**0.17**	**7.59**	**5.6**
SRC water (%)	0.66	2.0	11.84	**< 32 µm**	**0.76**	**1.7**	**10.06**	**15.0**
SRC sucrose (%)	**0.73**	**2.7**	**10.80**	**50–75 µm**	**0.77**	**2.5**	**10.00**	**7.4**
				Flour + starch	**0.77**	**2.5**	**10.00**	**7.4**
SRC lactic acid (%)	0.67	7.9	13.30	No improvement ≥ 5.0%				
SRC sodium carbonate (%)	0.61	3.9	13.55	**< 32 µm + 50–75 µm**	**0.71**	**3.4**	**11.81**	**12.8**
FA DDT (min)	**0.80**	**0.91**	**12.70**	**Gluten + flour**	**0.87**	**0.73**	**10.19**	**19.8**
FA WAM (%)	**0.79**	**1.2**	**10.91**	**50–75 µm + flour**	**0.87**	**1**	**9.1**	**16.7**
FA S (min)	0.59	3.45	14.86	**Gluten + dough**	**0.72**	**2.84**	**12.23**	**17.7**
FA DS (FE)	**0.77**	**14**	**12.85**	**Dough**	**0.81**	**12**	**11.01**	**14.3**
FA FQN (-)	0.53	43	16.05	Gluten + dough	0.67	36	13.44	16.3
EX energy 45 min (cm^2)	0.67	15	12.20	**Gluten + flour**	**0.78**	**12**	**9.76**	**20.0**
				Gluten + dough	**0.78**	**12**	**9.76**	**20.0**
				< 32 µm + 50–75 µm + 75–100 µm + > 100 µm + flour	**0.78**	**12**	**9.76**	**20.0**
EX resistance 45 min (BU)	0.05	43	20.19	Gluten + flour	0.34	36	16.91	16.3
EX extensibility 45 min (mm)	**0.75**	**12**	**11.33**	**Gluten + flour**	**0.83**	**10**	**9.44**	**16.7**
EX maximum 45 min (BU)	0.35	58	16.48	Gluten	0.48	52	14.78	10.3
EX ratio number 45 min (-)	0.29	0.4	21.06	No improvement ≥ 5.0%				
EX ratio number (Max.) 45 min (-)	0.07	0.5	23.82	Gluten + flour	0.42	0.4	19.05	20.0
AL C_{max} (-)	**0.81**	**190**	**9.96**	**< 32 µm + > 100 µm + flour**	**0.86**	**165**	**8.65**	**13.2**
AL P (mm H_2O)	0.68	10	13.16	**50–75 µm + 75–100 µm**	**0.79**	**8**	**10.53**	**20.0**
AL L (mm)	0.45	17	18.09	Gluten + flour	0.63	14	14.9	17.6
AL G (-)	0.44	2.0	17.86	Gluten + flour	0.62	1.7	15.18	15.0
AL W (10^{-4} J)	**0.87**	**30**	**7.80**	**Dough**	**0.91**	**24**	**6.24**	**20.0**
AL P/L (-)	0.28	0.38	15.71	50–75 µm	0.46	0.33	13.64	13.2
AL Ie (%)	**0.75**	**2.9**	**8.96**	**Gluten + dough**	**0.85**	**2.3**	**7.1**	**20.7**
AL K (-)	**0.78**	**350**	**12.71**	**50–75 µm**	**0.83**	**309**	**11.22**	**11.7**
AL SH (-)	0.44	0.07	14.59	**Gluten + dough**	**0.80**	**0.05**	**10.42**	**28.6**
AL D_{min} (-)	**0.71**	**0.26**	**13.34**	**75–100 µm**	**0.79**	**0.23**	**11.8**	**11.5**
AL D_{max} (-)	0.68	0.39	13.14	**75–100 µm**	**0.79**	**0.32**	**10.78**	**17.9**

*Improvement of the RMSECV I_{RMSECV} compared to flour spectra

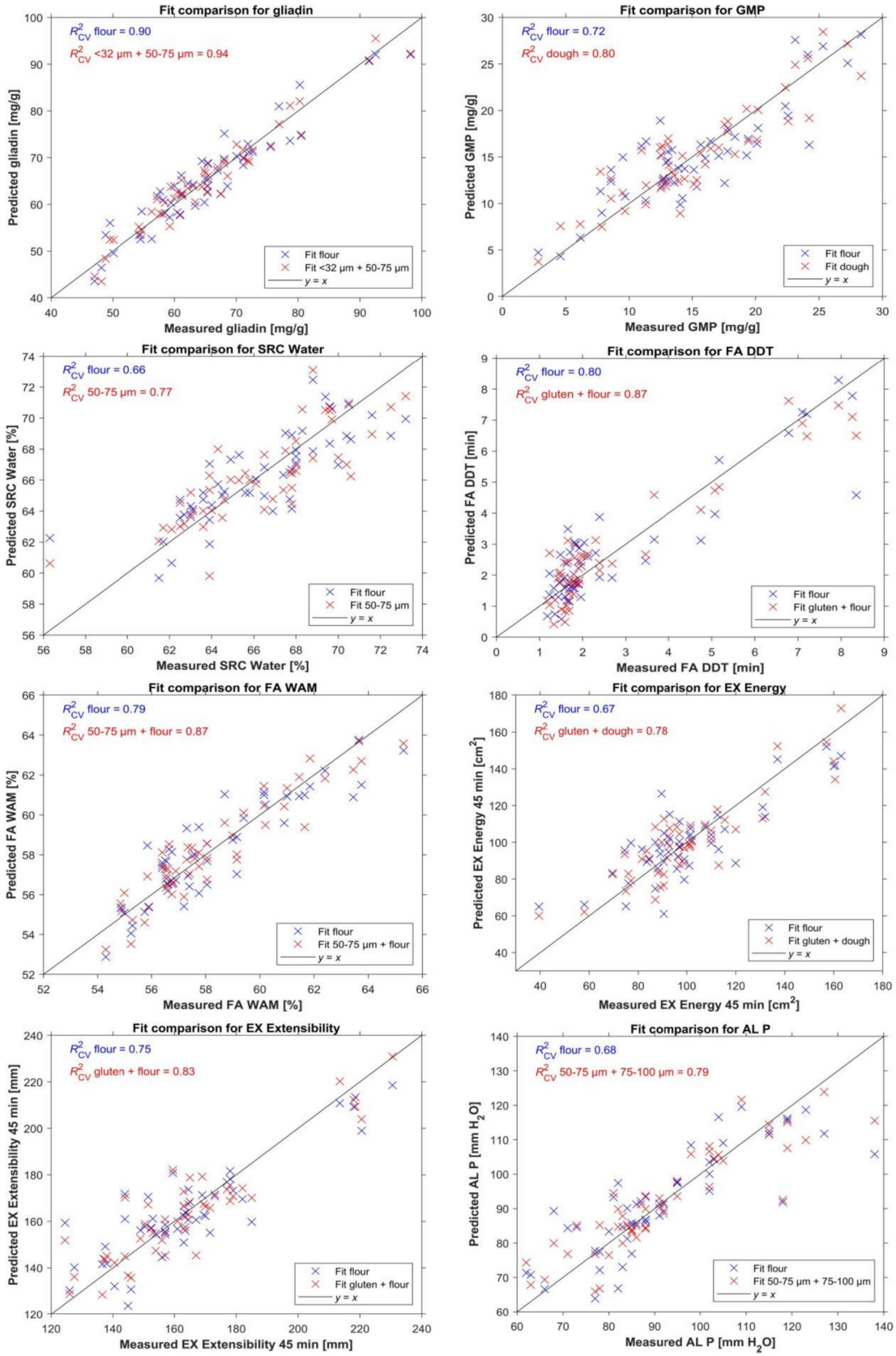

Fig. 4 Fit comparison for selected best models from Table 5

absorbance signals of vibrations originating from protein. Another (additional) reason could be that certain proteins have accumulated in this fraction due to the separation of particle sizes whose signals are more relevant for the prediction of protein fractions. Interestingly, the SDSS- and GMP-related fractions are best predicted using spectra of dough. Possibly, the proteins in these fractions show specific structural changes during dough formation which can be differentiated by NIR spectroscopy because NIR spectra contain information about amino acid composition and also about protein secondary structure (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012). This could be the reason for improved predictions of certain protein fractions using dough spectra.

Starch Properties (Hagberg Falling Number, Starch Damage)

Two important starch-related properties of interest are the Hagberg falling number and starch damage. Starch damage relates to the amount of starch granules damaged during milling, which depends on the grain hardness and the milling procedure. It is susceptible to α-amylase activity, which affects starch paste consistency and also the sugar supply in yeast dough fermentation (Cauvain 2015; Miralbés 2004). The Hagberg falling number is a measure of the α-amylase activity, because it measures the decreasing viscosity of a starch paste (Delwiche et al. 2018; Edwards 2007).

The prediction of both parameters by NIR spectroscopy of wheat has been tested in numerous studies, but only starch damage was successfully predicted with R^2 values often > 0.90 (Lancelot et al. 2021; Miralbés 2004; Pojić and Mastilović 2013; Williams 2020). The prediction accuracy achieved in this study using flour spectra is similar to these results. The improvement achieved by a combination of sieve fractions was only minor. As damaged and undamaged starch are not chemically different, it is unlikely that NIR spectroscopy can distinguish between both. It is more likely to be an indirect calibration caused by correlations of starch damage to other parameters that can be predicted from NIR spectra (Dowell et al. 2006). It could also be a result of the ability of NIR spectroscopy to distinguish differences in hardness, which relates to the content of damaged starch (Pojić and Mastilović 2013).

Hagberg falling number has not yet been successfully predicted in the literature (Delwiche et al. 2018; Dowell et al. 2006; Nagel-Held et al. 2024). Surprisingly, in this study, an R^2_{CV} of 0.72 was achieved by flour spectra. No improvement was observed by flour fractionation and data fusion. It is unlikely that NIR spectroscopy can distinguish α-amylase from other proteins (Delwiche et al. 2018). Consequently, the better prediction achieved in this study may be a result of an indirect calibration enabled by specific sample set characteristics. However, the highest correlation to the other parameters tested in this study was an $r^2_{Pearson}$ of 0.45 with the farinograph dough stability.

It may be an intuitive assumption that the prediction of starch parameters can be improved by NIR spectra of the starch fraction. However, this was not the case. One possible cause is that the starch fraction underwent such major changes during preparation that it is no longer possible to establish a relationship with the flour parameters. The reason for this is that the starch suspension was stored during the three production rounds of dough and gluten and only centrifuged at the end of the process for each flour sample. During this time, many (enzymatic) reactions could have taken place. Furthermore, not all starch particles were separated from the suspension by the centrifugation process, resulting in the loss of some material that could have been relevant for the prediction.

Solvent Retention Capacity

SRC analysis is based on the exaggerated swelling of different flour polymeric compounds (gluten proteins, damaged starch, solvent-accessible arabinoxylans/pentosans) in different solvents (water, 5% lactic acid solution, 5% sodium carbonate solution, 50% sucrose solution) without heat or shear. The pattern of SRC values is used to estimate flour and baking quality (Kweon et al. 2011).

Until now, the prediction of SRC values has not been a focus of studies predicting wheat quality parameters, as it is a technique less commonly used to characterize flour quality compared to the other flour characteristics tested in this study. Only Lancelot et al. (2021) tested the prediction of SRC values using NIR spectra of flour and obtained good predictions with $R^2_{Prediction} > 0.80$ for water, lactic acid, and sodium carbonate SRC and an $R^2_{Prediction}$ of 0.75 for sucrose SRC. In contrast, the prediction results in this study are worse and sucrose SRC is the parameter that can be predicted best. For flour spectra, R^2_{CV} ranged from 0.61 to 0.73 while the results were improved to R^2_{CV} between 0.71 and 0.77 for water, sucrose, and sodium carbonate SRC, mainly using the < 32 µm and the 50–75 µm fractions. No improvement was achieved for the lactic acid parameter. As the composition of the different flour fractions in relation to the polymers captured by SRC is largely unclear, no conclusion can be drawn as to why these particular fractions appear to be important for the predictions. A possible explanation for the better predictions of Lancelot et al. (2021) is that they only predicted the SRC parameters of the same flour after different storage conditions and times. As a consequence, there will have been greater similarities between the calibration and validation dataset than if they had used completely different flour samples, which may have improved the predictions.

Generally, the swelling behavior of different polymers relates to their chemical structure, as this structure is responsible for the solvent-polymer interactions (Kweon et al. 2011). Different structural characteristics contribute to different vibrations detected by NIR spectroscopy. From the results reported in the literature (Miralbés 2004; Pojić and Mastilović 2013; Williams 2020) and also in this study, it appears that NIR spectroscopy can detect and distinguish signals related to different major flour components. However, it is likely that the swelling behavior is not only influenced by the amounts of different polymers, but also by their specific structural characteristics (e.g., protein secondary structure, amylose and amylopectin ratios). Other studies suggest that NIR spectroscopy can capture at least some of this structural information and the interactions of flour components with water (Alava et al. 2001; Bruun et al. 2007; Salgó and Gergely 2012). However, it seems that not all relevant information for the prediction of the swelling behavior can be detected by NIR spectroscopy, neither from flour nor from fractions or dough spectra, which is why the achieved predictive quality in this study is only moderate to good.

Rheological Parameters (Farinograph, Extensograph, Alveograph)

There is a great interest in the prediction of rheological parameters, because the many empirical rheological measurements available to evaluate wheat quality are time-consuming and require a large amount of flour. Because they simulate the flour behavior in processes such as kneading or proofing, they provide important information on flour quality (Edwards 2007; Pojić and Mastilović 2013). The most commonly established methods are farinograph, extensograph, and alveograph measurements, which is why the parameters of these methods were predicted in this study.

Farinograph water absorption is the only parameter for which most studies agree on the prediction accuracy. Reported R^2 values usually range from around > 0.70 to < 0.90 (Dowell et al. 2006; Miralbés 2004; Mutlu et al. 2011; Nagel-Held et al. 2022; Pojić and Mastilović 2013; Williams 2020). This is in accordance with the results in this study, because an R^2_{CV} of 0.79 was achieved using spectra of flour and an R^2_{CV} of 0.87 was obtained using the combination of the 50–75 µm fraction and flour spectra. Miralbés (2004) attributed the good prediction to the fact that water absorption is mostly governed by macromolecules like proteins and damaged starch and that NIR spectra contain information about these. Possibly, the proteins (and other compounds) enriched in the 50–75 µm fraction strongly influence water absorption, which is why this fraction again contributed to an improved prediction.

Contrary to this, the reported predictions for some of the rheological parameters are consistently not good enough for screening purposes. This is the case, for example, for the extensograph ratio number and resistance, where R^2 values are usually < 0.50 (Nagel-Held et al. 2022, 2024), which is also in accordance with the results of this study. It seems that some parameters simply cannot be related to signals captured by NIR spectroscopy and that the novel approach of flour fractionation and data fusion can also not achieve an acceptable predictive quality for these parameters.

For most rheological parameters, the reported prediction accuracies in the literature vary and range from R^2 values indicating that they cannot be predicted from NIR spectra of wheat (often between 0 and 0.50) to R^2 values that show an acceptable prediction by NIR spectroscopy (> 0.70; sometimes even > 0.80). This is for example the case for the farinograph dough development time, stability, and dough softening; for the extensograph energy and extensibility; and for the alveograph parameters W, P, L, and P/L (Dowell et al. 2006; Jirsa et al. 2008; Miralbés 2003, 2004; Mutlu et al. 2011; Nagel-Held et al. 2022, 2024; Pojić and Mastilović 2013). Interestingly, acceptable to good predictions with $0.75 \leq R^2_{CV} \leq 0.87$ were achieved from flour spectra for multiple parameters in this study, including farinograph dough development time and dough softening; extensograph extensibility; and alveograph C_{max}, W, Ie, and K. Flour fractionation and data fusion improved the results of these parameters to $0.81 \leq R^2_{CV} \leq 0.91$, often involving spectra of gluten and dough. Furthermore, the predictions of the parameters farinograph stability, extensograph energy and alveograph P, SH, D_{min}, and D_{max} were improved by flour fractionation and data fusion to R^2_{CV} values between 0.72 and 0.80, most often using spectra of gluten and dough. Sometimes, sieve fractions were also involved in the best models for the predictions of these parameters, including most frequently the 50–75 µm and the 75–100 µm fractions, among others. Data fusion proved to be an important contributor to the improved predictions, as in many cases the best models were achieved using a combination of spectra. This again indicates that complementary information is contained in spectra of flour, flour fractions, and dough.

In summary, for many of the rheological parameters predicted in this study, higher prediction accuracies were already achieved using NIR spectra of flour compared to the results reported in the literature. Furthermore, the novel approach of flour fractionation and data fusion was able to improve the predictions of many of the tested parameters. However, the general predictive quality of rheological parameters was limited to $R^2_{CV} < 0.90$. Various reasons for the inconsistency in the predictive qualities reported for rheological parameters in numerous studies are discussed. This involves especially the characteristics of the samples used (varieties, growing locations, harvest years, number of samples) (Dowell et al. 2006; Nagel-Held et al. 2024; Pojić and Mastilović 2013). Many studies have shown that the predictions are generally better

when samples are classified according to these characteristics and then separate models are calculated for different classes (Dowell et al. 2006; Nagel-Held et al. 2024; Miralbés 2003). Although the sample set in this study is small and a true validation of results has to be performed, the sample set is very diverse, because the wheat samples were commercially available mixtures originating from ten countries and four harvest years. The results therefore show that NIR spectroscopy has the potential to predict many rheological parameters with acceptable accuracy without sample classification if the sample set is diverse enough to include many possible variations in the calibration dataset. The results could be further improved by including many more samples from more locations and harvest years in the model-building process. However, the predictive quality of rheological parameters may generally be limited, because the errors of the reference analyses are generally much larger compared to analytical determinations of, e.g., protein content (Nagel-Held et al. 2024; Pojić and Mastilović 2013).

General Discussion of Factors Influencing the Predictive Quality

The prediction accuracy is affected by various general influencing factors. An overview is presented in Fig. 5. Some of these factors have already been outlined in the previous discussion.

Two main factors that have already been discussed in previous sections are the properties of the samples used and the accuracy of the reference analyses (Dowell et al. 2006; Nagel-Held et al. 2024; Pojić and Mastilović 2013). When establishing prediction models, the number of samples should be large and the samples should cover a wide range of the property of interest (e.g., low to high protein content, short to long dough development times). However, as samples cannot be custom made, this is difficult to control. Also, the distribution of sample values in this study was not ideal for some wheat quality parameters, as can be seen in Fig. 4. Nevertheless, the results of this study show that prediction models can be successfully established based on a diverse sample set (varieties, growing location, harvest years). This suggests that the predictions may further be improved if even more samples are used for modeling, from even more countries of origin, harvest years, etc. For some wheat quality parameters, the achievable prediction accuracy may be limited in advance by the accuracy of the reference analyses, which is especially true for rheological analyses (Nagel-Held et al. 2024; Pojić and Mastilović 2013). It is likely that the limited prediction accuracies of rheological parameters achieved in this study in Tables 3, 4, and 5 are (at least partly) a result of this. A larger number of repetitions of these reference analyses could contribute to reduced measurement errors and consequently to increased prediction accuracies in future studies.

The characteristics of NIR spectroscopy strongly govern the achieved prediction accuracies. These include specific challenges that NIR spectroscopy faces (e.g., the broad peaks and the superimposition of strong water signals originating from OH vibrations in multiple regions of the

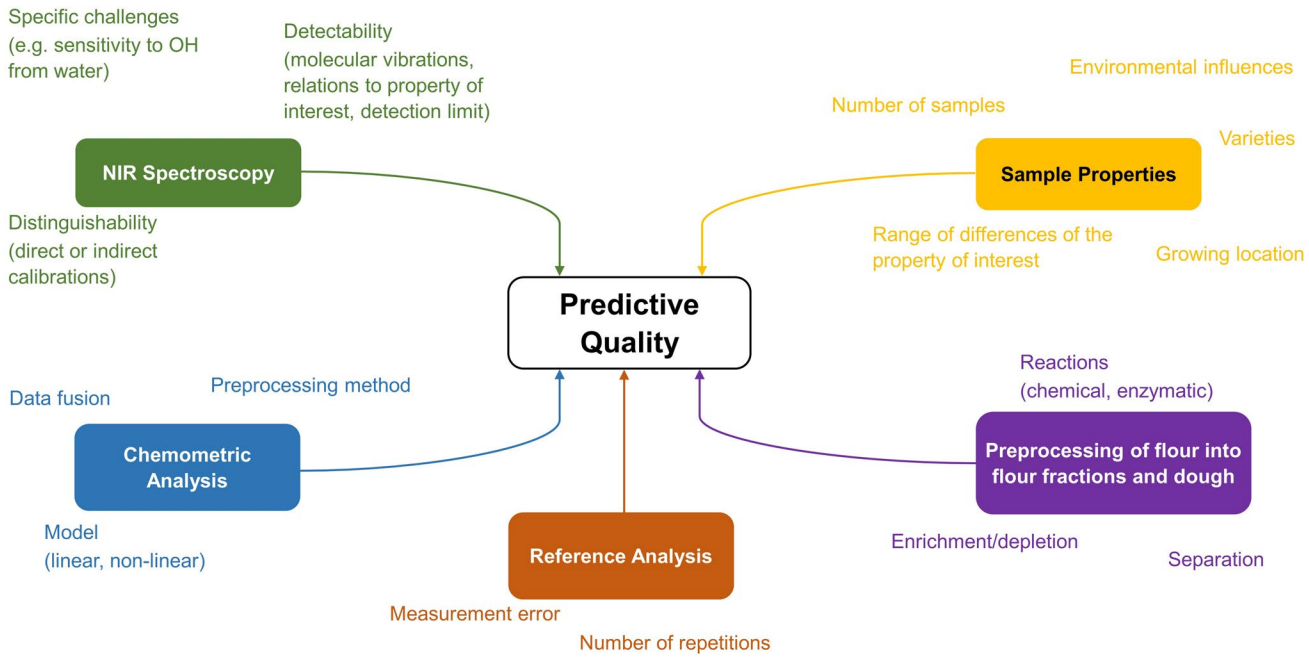

Fig. 5 Overview of factors influencing the predictive quality

spectrum) (Workman and Weyer 2012) as well as the general detectability and distinguishability of signals that relate to the property of interest that is to be predicted (Delwiche et al. 2018). This means that the property of interest must be related in some way to molecules present in the sample which show signals in NIR spectra in a measurable quantity. Furthermore, it is necessary that NIR spectroscopy is able to distinguish these signals from signals of other molecules that contribute to the same vibrations due to similar or identical molecular structures. As some researchers have noted in the past, it is likely that NIR spectroscopy can only achieve this to a limited extent. For example, damaged and undamaged starch are not chemically different and also the spectra of amylose and amylopectin are very similar (Dowell et al. 2006; Pojić and Mastilović 2013). Regarding protein-related vibrations, NIR spectroscopy can distinguish between different protein conformations (Bruun et al. 2007; Salgó and Gergely 2012; Workman and Weyer 2012), but it is unlikely that it can, for example, detect a specific protein (e.g., α-amylase) because of this (Delwiche et al. 2018). On top of this, it has been suggested multiple times that some signals in NIR spectra originate from the interactions of different molecules (e.g., water, carbohydrates, proteins) (Alava et al. 2001; Wesley et al. 1998), but it is unclear to what extent NIR spectroscopy can distinguish between the exact carbohydrates or proteins that participate in the interactions. Consequently, it has been suggested before that many of the achieved prediction accuracies may be results of indirect calibrations, especially regarding correlations to the protein content as a property that can be predicted very well from NIR spectra due to the characteristic protein vibrations (Dowell et al. 2006; Miralbés 2004; Schuster et al. 2023). As presented in Table 2, the protein content is also correlated to multiple wheat quality parameters which showed good prediction accuracies in this study shown in Table 5 (especially protein fractions and rheological parameters). However, the achieved prediction accuracies using flour spectra are often higher than the simple correlations to protein content (e.g., for wet gluten, SDSS-GMP fractionation, some farinograph and alveograph parameters). Furthermore, the predictions of many of these and additional parameters were improved using flour fractionation with or without data fusion. In summary, it is therefore likely that NIR calibrations of wheat quality are to some extent indirect, but it appears that for some wheat quality parameters, additional information can be captured by NIR spectroscopy to improve the predictions compared to the simple correlations to protein content.

The chemometric method used to establish the prediction model also has a major influence on the achieved prediction accuracy. The interactions of the chosen preprocessing algorithm(s) and the data fusion method (if applicable) as well as the model type (linear, non-linear) and the exact model itself are crucial for the success or failure of the modeling. Consequently, it is necessary to optimize all of the above points to achieve the highest possible prediction accuracy using the exact methods tested. However, this is still a question of trial and error, which makes the model optimization a computationally intensive process, depending on how extensive the tested methods are (Borràs et al. 2015; Du et al. 2022; Pojić and Mastilović 2013). In this study, a variety of different preprocessing algorithms were tested. Different methods yielded the best models, as shown in Supplementary Material S1. Data fusion successfully improved the prediction of many parameters, as presented in Tables 4 and 5, although multiple different combinations of spectra yielded the best models. Only two linear regression models (PCR, PLSR) were tested in this study. Non-linear models could further improve the predictions, because it is likely that some of the wheat quality parameters have non-linear relationships with the signals captured by NIR spectroscopy (Pojić and Mastilović 2013). However, successful optimization of non-linear models often requires a larger number of flour samples and even more computing time, as many different hyperparameters have to be optimized. Consequently, the results of this study can serve as a starting point to decide for which wheat quality parameters it might be worthwhile to test non-linear models in future studies.

Lastly, in this particular study, the flour fractionation procedure strongly influenced the model accuracy. Flour fractionation separates flour components with different properties. In sieving, this is achieved because different flour components were milled to different particle sizes, as they exhibit different properties in milling (Jensen et al. 1982; Jones et al. 1959). In dough preparation and gluten washing, the separation is due to the gluten network development and the inherent property of certain molecules to remain together in this network when the dough is washed out, while others are separated by the washing solution (Schopf et al. 2021). This separation of different flour components changes the composition of the obtained fractions compared to flour by enriching and depleting different components in different fractions. Additionally, different chemical and enzymatic reactions take place during flour fractionation and dough preparation, which further alter the properties of the obtained fractions. As shown in this study in Figs. 2 and 3, this results in spectral differences, which in turn have a major influence on the model accuracies presented in Tables 3, 4, and 5 and Supplementary Material S1. The results of this study emphasize that flour fractionation is a valuable new method that allows better prediction of wheat quality compared to flour spectra for many quality parameters. However, from the analyses performed in this study, no exact conclusion can be drawn as to why especially the 50–75 µm and the 75–100 µm fractions as well as gluten and dough were the most important sample types for the prediction of many parameters with or without data fusion.

Further studies could yield more insights into the relationship between wheat quality parameters and flour fractions by analyzing the detailed composition of the flour fractions. Additionally, further fractionation methods (e.g., triboelectric separation) that make use of other properties of flour components for the separation could also be tested for the prediction of wheat quality.

Conclusion

This study investigated the potential to improve the prediction of wheat quality based on a flour fractionation approach (sieve fractionation, dough preparation, and gluten washing) and data fusion using the established techniques of NIR spectroscopy and chemometrics and a diverse sample set consisting of 50 commercially available wheat flour samples with many different qualities, countries of origin and harvest years. A variety of quality characteristics including protein-related parameters (protein content, gluten content, Osborne and SDSS-GMP fractionation methods), Hagberg falling number, starch damage, SRC, and parameters of rheological analyses (farinograph extensograph, alveograph) were predicted.

Flour fractionation and dough preparation altered the composition of the obtained fractions and dough compared to flour, which reflected in spectral differences of their NIR spectra and enabled a differentiation by PCA. This change in the information content of the NIR spectra led to a change in the prediction accuracy for many wheat quality parameters when predictions were based on spectra of flour fractions and dough instead of flour spectra. In this way, the RMSECV was improved between 5.6 and 28.6% for 35 out of the 41 quality parameters tested. Dough preparation and gluten washing as well as sieve fractionation each achieved the best predictions for about half of these parameters. In a majority of cases, the best models were based on data fusion of spectra from different sample types. The fractions that were the most relevant for the improved predictions were the 50–75 µm and the 75–100 µm fractions as well as gluten and dough. For several of the parameters tested, prediction accuracies of $0.80 \leq R^2_{CV} \leq 0.96$ were achieved using spectra of flour (especially for protein-related parameters, starch damage, farinograph DDT, alveograph C_{max} and W). In contrast, flour fractionation with or without data fusion improved the prediction of all of these parameters (apart from protein content) to $0.87 \leq R^2_{CV} \leq 0.97$ and also achieved prediction accuracies of $0.80 \leq R^2_{CV} \leq 0.87$ for multiple additional parameters (for protein-related parameters, farinograph WAM and DS, extensograph extensibility, alveograph Ie and K and SH). In addition, the SRC parameters and several parameters of the rheological analyses were predicted with $0.70 < R^2_{CV} < 0.80$ by flour fractionation with and without data fusion (flour spectra achieved $R^2_{CV} < 0.70$ here).

The results of this study show that more relevant information for the prediction of wheat quality can be generated when flour fractionation precedes NIR spectroscopy, as different flour components are enriched and depleted by the fractionation procedures and chemical and enzymatic reactions further alter the composition. Additionally, data fusion can be a valuable approach to improve the prediction of many wheat quality parameters by combining complementary information that is present in the spectra of different sample types. However, it seems that the general predictive ability of some parameters using NIR spectroscopy and chemometrics is limited, especially regarding Hagberg falling number, SRC parameters, and rheological analyses in general. Possible reasons for this are high measurement errors of reference analyses, the inability of NIR spectroscopy to detect and distinguish signals relevant for the prediction of these parameters, and possible non-linear relationships that cannot be described by the linear models tested in this study. However, since many different factors affect the prediction quality simultaneously, it is difficult to determine the exact reasons for the achieved prediction accuracies.

In summary, for many of the quality parameters tested, good to very good prediction accuracies with large improvements in the RMSECV were achieved for this diverse sample set using flour fractionation with or without data fusion compared to flour spectra. Further studies can test this approach for the prediction of wheat quality using a larger number of flour samples and possibly also other types of spectroscopic analyses (e.g., fluorescence, Raman), fractionation methods (e.g., triboelectric separation), and regression models (e.g., neural networks). Further studies are also needed to analyze the composition of these sample types in order to draw conclusions as to why specific fractions are particularly important for the prediction of wheat quality. These could also yield insights into the relationship between flour components and flour quality.

The results emphasize that flour fractionation and data fusion have the potential to be used in the industry as an extension of conventional NIR spectroscopy and chemometric techniques to predict many wheat quality parameters. To ensure reliability and scalability of the method for practical application in industrial settings, multiple aspects should be considered. This includes standardization and automation of fractionation and measurement procedures to increase throughput and enhance reproducibility as well as a cost-benefit analysis that weighs the effort for production of flour fractions against the achieved improvement in prediction accuracy. Furthermore, the sensitivity of predictions based on flour fractions for various influencing factors should be taken into account. Further studies could analyze

the influences of grain characteristics (e.g., cultivar, harvest season, moisture content) and flour milling (milling conditions, flour extraction) as well as flour fractionation and NIR measurements (equipment, fractionation protocol, measurement parameters) on the prediction accuracies. These influences should be controlled as best as possible. They necessitate periodic re-training of the models, which is generally common practice for industrial spectroscopic applications (e.g., for the prediction models of flour protein content) in order to adapt the models to changes in product matrices and processes over time. Also, robust calibration-transfer methods and validation of models need to be implemented.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12161-025-02931-7.

Acknowledgements This IGF Project of the FEI (01IF21711N) was supported within the program for promoting the Industrial Collective Research (IGF) of the Federal Ministry of Economic Affairs and Energy (BMWE), based on a resolution of the German Parliament. The authors would like to thank the project partners Mühlenchemie GmbH & Co. KG (Ahrensburg, Germany) for providing flour samples and performing wheat quality reference analyses and CHOPIN Technologies (Villeneuve-la-Garenne, France) for performing further wheat quality reference analyses. Open Access funding enabled and organized by Project DEAL.

Author Contribution Ziegler D: idea, methodology, investigation, formal analysis, visualization, writing and editing – original draft. Buck L: investigation (Osborne & SDSS-GMP fractionation). Scherf KA: project administration, writing – review. Popper L: project administration, writing – review. Schaum A: supervision, project administration, writing – review. Hitzmann B: supervision, project administration, writing – review.

Funding Open Access funding enabled and organized by Projekt DEAL. This IGF Project of the FEI (01IF21711N) was supported within the program for promoting the Industrial Collective Research (IGF) of the Federal Ministry of Economic Affairs and Energy (BMWE), based on a resolution of the German Parliament. Open Access funding was enabled and organized by Project DEAL.

Data Availability Data will be made available on request.

Declarations

Competing interest The authors declare no competing interests.

References

Alava JM, Millar SJ, Salmon SE (2001) The determination of wheat breadmaking performance and bread dough mixing time by NIR spectroscopy for high speed mixers. J Cereal Sci 33(1):71–81. https://doi.org/10.1006/jcrs.2000.0341

Albanell E, Miñarro B, Carrasco N (2012) Detection of low-level gluten content in flour and batter by near infrared reflectance spectroscopy (NIRS). J Cereal Sci 56(2):490–495. https://doi.org/10.1016/j.jcs.2012.06.011

Borràs E, Ferré J, Boqué R, Mestres M, Aceña L, Busto O (2015) Data fusion methodologies for food and beverage authentication and quality assessment – a review. Anal Chim Acta 891:1–14. https://doi.org/10.1016/j.aca.2015.04.042

Bruun SW, Søndergaard I, Jacobsen S (2007) Analysis of protein structures and interactions in complex food by near-infrared spectroscopy. 1. Gluten powder. J Agric Food Chem 55(18):7234–7243. https://doi.org/10.1021/jf063680j

Cauvain S (2015) Principles of dough formation. In: Cauvain S (ed) Technology of breadmaking. Springer International Publishing, Cham, pp 303–337

Coulson CB, Sim AK (1965) Proteins of wheat flour. Nature 208(5010):583–584. https://doi.org/10.1038/208583a0

Delwiche SR, Higginbotham RW, Steber CM (2018) Falling number of soft white wheat by near-infrared spectroscopy: a challenge revisited. Cereal Chem 95(3):469–477. https://doi.org/10.1002/cche.10049

Dowell FE, Maghirang EB, Xie F, Lookhart GL, Pierce RO, Seabourn BW, Bean SR, Wilson JD, Chung OK (2006) Predicting wheat quality characteristics and functionality using near-infrared spectroscopy. Cereal Chem 83(5):529–536. https://doi.org/10.1094/CC-83-0529

Du Z, Tian W, Tilley M, Wang D, Zhang G, Li Y (2022) Quantitative assessment of wheat quality using near-infrared spectroscopy: a comprehensive review. Compr Rev Food Sci Food Saf 21(3):2956–3009. https://doi.org/10.1111/1541-4337.12958

Edwards WP (ed) (2007) The science of bakery products. Royal Society of Chemistry, Cambridge, pp 139–153. https://doi.org/10.1039/9781847557797

Gabriel D, Pfitzner C, Haase NU, Hüsken A, Prüfer H, Greef J-M, Rühl G (2017) New strategies for a reliable assessment of baking quality of wheat – rethinking the current indicator protein content. J Cereal Sci 77:126–134. https://doi.org/10.1016/j.jcs.2017.08.002

Jensen SA, Munck L, Martens H (1982) The botanical constituents of wheat and wheat milling fractions. Quantification by autofluorescence. Cereal Chem 59(6):477–484

Jirsa O, Hrušková M, Švec I (2008) Near-infrared prediction of milling and baking parameters of wheat varieties. J Food Eng 87(1):21–25. https://doi.org/10.1016/j.jfoodeng.2007.09.008

Jones CR, Halton P, Stevens DJ (1959) The separation of flour into fractions of different protein contents by means of air classification. J Biochem Microbiol Technol Eng 1(1):77–98. https://doi.org/10.1002/jbmte.390010108

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 Chem 88(6):537–552. https://doi.org/10.1094/CCHEM-07-11-0092

Lancelot E, Fontaine J, Grua-Priol J, Le-Bail A (2021) Effect of long-term storage conditions on wheat flour and bread baking properties. Food Chem 346:128902. https://doi.org/10.1016/j.foodchem.2020.128902

Mburu M, Komu C, Paquet-Durand O, Hitzmann B, Zettel V (2021) Chia oil adulteration detection based on spectroscopic measurements. Foods. https://doi.org/10.3390/foods10081798

Miralbés C (2003) Prediction chemical composition and alveograph parameters on wheat by near-infrared transmittance spectroscopy. J Agric Food Chem 51(21):6335–6339. https://doi.org/10.1021/jf034235g

Miralbés C (2004) Quality control in the milling industry using near infrared transmittance spectroscopy. Food Chem 88(4):621–628. https://doi.org/10.1016/j.foodchem.2004.05.004

Mutlu AC, Boyaci IH, Genis HE, Ozturk R, Basaran-Akgul N, Sanal T, Evlice AK (2011) Prediction of wheat quality parameters using near-infrared spectroscopy and artificial neural networks. Eur Food Res Technol 233(2):267–274. https://doi.org/10.1007/s00217-011-1515-8

Nagel-Held J, Kaiser L, Longin F, Hitzmann B (2022) Prediction of wheat quality parameters combining Raman, fluorescence, and near-infrared spectroscopy (NIRS). Cereal Chem 99(4):830–842. https://doi.org/10.1002/cche.10540

Nagel-Held J, El Hassouni K, Longin F, Hitzmann B (2024) Spectroscopy-based prediction of 73 wheat quality parameters and insights for practical applications. Cereal Chem 101(1):144–165. https://doi.org/10.1002/cche.10732

Pojić MM, Mastilović JS (2013) Near infrared spectroscopy—advanced analytical tool in wheat breeding, trade, and processing. Food Bioproc Tech 6(2):330–352. https://doi.org/10.1007/s11947-012-0917-3

Salgó A, Gergely S (2012) Analysis of wheat grain development using NIR spectroscopy. J Cereal Sci 56(1):31–38. https://doi.org/10.1016/j.jcs.2012.04.011

Schopf M, Wehrli MC, Becker T, Jekle M, Scherf KA (2021) Fundamental characterization of wheat gluten. Eur Food Res Technol 247(4):985–997. https://doi.org/10.1007/s00217-020-03680-z

Schuster C, Huen J, Scherf KA (2023) Prediction of wheat gluten composition via near-infrared spectroscopy. Curr Res Food Sci 6:100471. https://doi.org/10.1016/j.crfs.2023.100471

Wesley IJ, Larsen N, Osborne BG, Skerritt JH (1998) Non-invasive monitoring of dough mixing by near infrared spectroscopy. J Cereal Sci 27(1):61–69. https://doi.org/10.1006/jcrs.1997.0151

Wieser H, Antes S, Seilmeier W (1998) Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. Cereal Chem 75(5):644–650. https://doi.org/10.1094/CCHEM.1998.75.5.644

Williams PC (2020) Application of chemometrics to prediction of some wheat quality factors by near-infrared spectroscopy. Cereal Chem 97(5):958–966. https://doi.org/10.1002/cche.10318

Workman J, Weyer L (2012) Practical guide and spectral atlas for interpretive near-infrared spectroscopy, 2nd edn. CRC Press, Boca Raton, p 326. https://doi.org/10.1201/b11894

Ziegler D, Buck L, Scherf KA, Popper L, Hitzmann B (2025) Fluorescence spectroscopy of flour fractions and dough: analysis of spectral differences and potential to improve wheat quality prediction. Cereal Chem. https://doi.org/10.1002/cche.10881

Ziegler D, Buck L, Scherf KA, Popper L, Schaum A, Hitzmann B (2025) Improved prediction of wheat baking quality by three novel approaches involving spectroscopic, rheological and analytical measurements and an optimized baking test. J Food Meas Charact 19(3):1673–1692. https://doi.org/10.1007/s11694-024-03063-y