Prediction of Wheat Milling Characteristics by Near-Infrared Reflectance Spectroscopy

Jaroslav BLAŽEK, *Ondřej* JIRSA *and Marie* HRUŠKOVÁ

Department of Carbohydrate Chemistry and Technology, Faculty of Food and Biochemical Technology, Institute of Chemical Technology Prague, Prague, Czech Republic

Abstract

Blažek J., Jirsa O., Hrušková M. (2005): **Prediction of wheat milling characteristics by near-infrared reflectance spectroscopy**. Czech J. Food Sci., **23**: 145–151.

The aim of this study was to explore the use of NIR spectroscopy of laboratory milled flour to predict the milling characteristics of wheat. Quantitative traits of the milling process of wheat were predicted by analyses of NIR spectra of six sets consisting of 94 samples. Reference data were obtained by grinding the samples on the laboratory mill Chopin CD1-auto (France), spectral data were measured on spectrograph NIRSystem 6500. Commercial spectral analysis software WINISI II was used to collect spectra, develop calibration equations and evaluate calibration performance. The quality of prediction was evaluated by coefficients of correlation between the measured and the predicted values from cross and independent validation. MPLS/PLS regression and ANN methods were used. A statistically significant dependence (at the probability level of 99%) was determined for all traits studied in the case of cross-validation. Satisfactory accuracy of the prediction models by independent validation was achieved only for semolina extraction rate, models for other characteristics did not show acceptable precision.

Keywords: wheat flour; prediction; NIR spectroscopy; NIRSystem 6500; milling characteristics

Wheat flour as the most important recipe ingredient is a key factor in the baking process determining the quality of the baker's products. The production of smooth light flour from alimentary wheat is a complex process whose effectivity can be described by many quantitative traits of the milling process such as semolina and flour extraction rates, semolina extractability, and Mohse yield. The values of these parameters are determined by the milling quality of wheat which is affected by the physico-mechanical properties of the grain endosperm. Practical assessment of these parameters is inconvenient due to the long time and acquisition costs (especially the cost of a cylindrical laboratory mill), and the problems associated with the achievement of the standard milling process. It is thus sensible to explore the prospects of indirect (screening) methods such as the application of instruments working on the principle of near infrared spectroscopy.

Near infrared (NIR) spectroscopy provides a method for a rapid, non-destructive, and accurate analysis of the composition of a sample. It allows the discrimination between various organic compounds and can be used to acquire both the qualitative and quantitative information. It not only supplies chemical information, but also the information of whether the physical properties of a sample can be obtained (Gottlieb *et al.* 2004). Nielsen *et al.* (2001) explored the use of multi-way models to provide a better understanding and translation of multivariate measurements into the operational routines of the mill operators. It is suggested that multi-way models can be used on a wide range of multivariate data containing both chemical and physical characteristics together with NIR spectra to assure a high degree of interpretability.

Almost all NIR applications on the agriculture foods to date have been developed by exploiting the empirical relationships between the spectral data and reference analytical data. This approach has proved highly successful in many cases but can lead to calibrations that are critically dependent on the characteristics of the samples used for the calibration (Osborne 2000). In such cases, the calibration requires frequent adjustments to maintain accuracy, a situation that undermines the benefits of NIR technique.

NIR has long been a recognised method for an accurate prediction of the protein content of wheat to assess its breadmaking potential. The baking quality of flour, however, relates to both the amount and the quality of the gluten proteins and is also determined by the complex interactions of all the biochemical constituents present in flour. Providing a measure of all the primary constituents simultaneously, NIR should have the potential for the determination of this quality. Various biochemical and physical properties of dough, relating to the baking quality, have been reasonably estimated by NIR, but a strong correlation between the measured property and the total protein content can lead to wrong conclusions. NIR is, however, sensitive to some degree not only to the protein content but also to the protein quality.

Although the total protein content is the primary factor in determining the final use of wheat, there is often a need to measure the properties that are indicative of the protein quality (Wesley *et al.* 2001). The measurements such as those obtained from the Extensograph, Mixograph, and Farinograph are widely used to assess the quality of wheat flour, both by end-users and by plant breeders. These measurements require relatively large amounts of material and are time-consuming which makes them unsuitable for use in the early stages of the breeding programmes and at the grain receival stations. Wesley *et al.* (1999) suggests that the curve-fitting methodology of measuring the gliadin and glutenin contents offers a simple alternative to partial least squares and multiple linear regression techniques. Although the curve fitting method does not predict the 'correct' answer for the gliadin and glutenin contents as measured by SE-HPLC, the R^2 values and the rank correlation values for samples grown in separate years do suggest that the method will rank the samples sufficiently for the discrimination into high, medium, and low content for each component (Wesley *et al.* 2001). Most importantly, the results are independent of the total protein content.

Osborne (1984) investigated NIR for predicting SDS sedimentation volume as a measurement of protein quality and concluded that there was no significant correlation with the NIR spectral data when the contribution of the protein content to the calibration was removed. Rubenthaler and Pomeranz (1987) studied water absorption, mixing time, and loaf volume, achieving reasonable predictions for these characteristics, and suggested that NIR indeed measures more than the simple total protein content. Williams *et al.* (1988) and Pawlinsky and Williams (1998) investigated a wide range of quality parameters with a view to develop whole grain NIR calibrations concerning the determination of wheat strength and bread-baking functional factors. Maghirang and Dowell (2003) investigated the hardness measurement of bulk wheat by single-kernel visible and near-infrared reflectance spectroscopy. Comparing predicted values with the reference method, the model developed correctly classified mixed samples with 72–100% accuracy. The results confirm the potential of using visible and NIR spectroscopy of whole single wheat kernels as a rapid and nondestructive measurement of bulk wheat grain hardness.

Automated colour classification of single wheat kernels using visible and near-infrared reflectance was investigated by Dowell (1998). It is important to determine the colour class of wheat accurately because the functionality traits of genetically red and white wheats cause them to mill and bake differently. The prediction results showed > 99% correct classification for single kernels when using the visible and NIR regions. The averaging of single kernel classifications resulted in 100% correct classification of bulk samples.

Near infrared spectroscopy has the potential for providing the information on the chemical processes that occur during the dough development in relation to the flour strength and mixing action and intensity (Wesley *et al.* 1998). Variations in two specific absorbance wavelengths in the second derivative spectrum (1160 nm and 1200 nm) as the dough is mixed follow the same trend as the mixer power consumption. Wesley *et al.* (1999) suggests that NIR spectroscopy may provide useful insights into the chemical changes that occur during the dough development and raises questions as to the definition of the dough development in terms of physical and chemical measurements. Delwiche and Weaver (1994) noted that the ability of NIR to robustly predict parameters such as dough mixing time, mixing tolerance, and overall bake score was low due to the complexity of the interactions between protein, starch, and lipid. A point to consider is that the physical properties of interest themselves do not have NIR spectra. It is the individual chemical structures in a wheat grain or flour that give rise to the absorption features on which NIR measurements are based (Wesley *et al.* 2001). Unfortunately, the exact nature of the relationship between the chemical composition and dough rheological properties is not known.

The use of chemometric procedures with near infrared spectroscopic data to produce calibration equations for analytical chemistry achieved success in many applications. A large increase in the prediction errors is observed when the calibration equation developed on one instrument is directly applied on another. Since many spectral differences can exist between two spectrometers, a standardisation procedure is a requirement for the long-term use of quantitative or qualitative models. Duponchel *et al.* (1999) studied the standardisation of near infrared spectrometers using artificial neural networks. In this way, the time-consuming step of recalibration for the second spectrometer is avoided and the initial error prediction level is retrieved. Osborne *et al.* (1999) confirmed that the single sample standardisation method can be successfully used to optically match NIR reflectance monochromator instruments of the same brand.

The aim of this study was to explore the use of NIR spectroscopy of laboratory milled flour to predict the milling characteristics of wheat. It is possible to expect that NIR has the potential to predict the milling traits because the efficiency of the milling process is dependent on the grain structure. Chemical and physical properties of the grain endosperm and covers (outer and inner pericarp, germ), separability of covers from the endosperm, starch damage, protein content, ash content (or the purity of the flour), moisture and granulation are some of the factors influencing the milling process and its efficiency.

MATERIAL AND METHODS

Samples. The milling characteristics of wheat were predicted by analyses of NIR spectra of six sets of wheat consisting of 94 samples. Variety wheat (54 samples) was supplied by SELGEN Stupice (crop years 2002 and 2003) and the Research Institute of Crop Production in Prague-Ruzyně (crop year 2003). Commercial wheat (40 samples) was obtained from the industrial mill Delta Prague (crop years 2002 and 2003). All wheat samples were milled on the laboratory mill Chopin CD1-auto (France) at the standard milling test – protocol Chopin. Eight samples of commercial wheat (crop year 2003) were used as a validation set.

References analysis. Samples were milled on the laboratory mill and individual milling intermediates and products were weighed. The milling characteristics were obtained by appropriate calculations. The ash content and moisture content needed for the calculation of Mohse yield (value) were assessed using Inframatic 8600 (Perten, Sweden). Semolina extraction rate is the weight of flour fractions and milling outlet divided by the sample weight. Flour extraction rate is the weight of flour fractions divided by the sample weight adjusted to the moisture of 14%. Semolina extractability is the flour extraction rate divided by the sum of flour extraction rate and the weight of the milling outlet. Mohse yield represents the difference between the assessed flour extraction rate and the ideal extraction rate according to the Mohse ash table.

NIR hardware. A wavelength scanning instrument NIRSystem 6500 (Foss NIRSystems, Inc., USA) was used to measure NIR diffuse reflectance spectra from 400 to 2500 nm at every 2 nm step. The analysis was carried out using small ring cup cells. The NIR spectra of flours were collected in five scans for each and were recorded as $\log(1/R)$.

Calibration and validation. Commercial spectral analysis software (WINISI II – ver. 1.00, Foss NIRSystems, Infrasoft International) was used to collect the spectra, to develop the calibration equations, and to evaluate the calibration performance. Prior to the calibration, the absorbance spectra were transformed mathematically by a range of scatter corrections to minimise the nonlinear effect of the light scatter due to the particle size differences between the samples (none, standard normal variate + detrending, standard normal variate only, detrending only and multiplicative scatter correction). The second data transformation is derivative mathematics that reduces the inter-correlation between the spectral data points of a spectrum. Two different treatments were applied during the calibration development – 1, 4, 4, 1, and 1, 8, 8, 1, where the first number indicates the order of the derivative (0 represents no derivative, 1 the first derivative and 2 the second derivative of log 1/*R*), the second is the gap in the data points over which the derivative is calculated, the third and fourth numbers refer to the number of the data points used in the first and second smoothing, respectively.

The calibration was performed using partial least square (PLS) and modified partial least square (MPLS) regressions available in WINISI software. The optimal number of terms was determined by cross-validation of the calibration samples. With full cross-validation, each sample is removed one at a time from the sample set, a new calibration performed and the predicted score calculated for the sample removed (Osborne 1984; Miralbés 2004). This procedure is repeated until every sample has been left out once. However, in the case of MPLS/PLS no samples were eliminated on account of their higher deviation and one-to-one cross-validation was used. The performance of the model was determined by the following statistics: standard error of calibration (SEC), standard error of cross validation (SECV), standard error of performance (SEP), coefficient of determination (R^2), linear correlation coefficient (r) between the reference values and the values estimated by the prediction models.

In addition, Artificial Neural Network method (ANN) was applied. The number of terms was limited to four. Outliers were eliminated and the starting set was divided into calibration and validation parts in the ratio of 75:25. The best calibration equations for mPLS/PLS and ANN were selected according to their maximum coefficient of correlation or minimum standard error of cross-validation or standard error of validation (SEV; for ANN). Following the completion of the calibration, the models were validated using an independent set of wheat samples from a commercial mill.

RESULTS AND DISCUSSION

Milling characteristics

The means, ranges, and coefficients of variations of wheat milling characteristics are summarised in Table 1. The calibration set consisted of variety and commercial subsets. Variety wheat subsets (data not shown) showed a lower semolina extraction rate (average 51.2%, range 22.1–63.4%) and a higher semolina extractability (average 86.1%, range 77.0–91.8%) in comparison with the commercial subsets (average values 53.3% and 70.3%, respectively). The flour extraction rate was com-

Table 1. Milling characteristics (in %)

Parameters	Mean	Range		Coefficient of variation (%)
		min.	max.	
Calibration set (94 samples)				
Semolina extraction rate	52.1	22.1	63.4	15.44
Semolina extractability	79.4	55.3	91.8	14.59
Flour extraction rate	68.5	60.0	73.7	3.94
Mohse extraction rate	–0.1	–9.1	8.9	–
Validation set (8 samples)				
Semolina extraction rate	47.0	22.7	56.3	26.6
Semolina extractability	87.2	85.1	92.2	3.0
Flour extraction rate	67.2	61.2	70.6	4.1
Mohse extraction rate	–1.2	–9.7	2.2	–

parable in all sets (average of variety wheat 67.8%, commercial wheat 69.4%), while Mohse yield was higher with variety wheat (average 0.9%) than with commercial wheat (average –1.6%). The validation set is characterised by a lower average semolina extraction rate (47.0%) and a higher semolina extractability (87.2%).

Prediction according to the cross-validation

The statistical evaluation of calibration and cross-validation of wheat milling characteristics are summarised in Table 2. The subsets of commercial (40) and variety (54) samples showed a statistically significant dependence between the predicted and the reference values with probability higher than 99% for all the traits studied (data not shown). A statistically significant dependence at the probability level of 99% described by correlation coefficients was detected by cross validation using the combined calibration set consisting of both commercial and variety wheat samples (94 samples) for all the characteristics studied. The highest statistical correlation (r = 0.633) in the combined set was achieved for semolina extractability. Similar results were found in the case of ANN method, higher r (and lower SEV) was achieved.

Prediction according to the independent validation

The correlation coefficient r and standard error of prediction SEP for the milling characteristics are shown in Table 3. The best calibration equations (selected according to SECV/SEV) were verified by an independent validation, for which the independent set of eight wheat samples was used. The performance of the validation was tested

Table 2. Prediction of milling characteristics

Parameters	Calibration				Cross-validation		
	n	Terms	SEC	r	Groups	SECV	r
mPLS/PLS							
Semolina extraction rate	94	1	6.03	0.444	94	6.34	0.391
Semolina extractability	94	4	6.05	0.730	94	7.10	0.633
Flour extraction rate	94	4	1.84	0.544	94	2.03	0.449
ANN							
Semolina extraction rate	68	4	7.59	0.187	21	2.46	0.882
Semolina extractability	85	4	0.94	0.994	8	0.81	0.995
Flour extraction rate	74	4	1.34	0.768	18	1.12	0.814

Table 3. Independent validation of calibration equations

Parameters	Calibration				Independent validation		
	n	Terms	SEC	r	n	SEP	r
mPLS/PLS							
Semolina extraction rate	94	1	6.03	0.444	8	6.71	0.920
Semolina extractability	94	4	6.05	0.730	8	2.04	0.505
Flour extraction rate	94	4	1.84	0.544	8	2.57	0.313
ANN							
Semolina extraction rate	68	4	7.59	0.187	8	7.70	0.852
Semolina extractability	85	4	0.94	0.994	8	2.96	0.020
Flour extraction rate	74	4	1.34	0.768	8	2.54	0.265

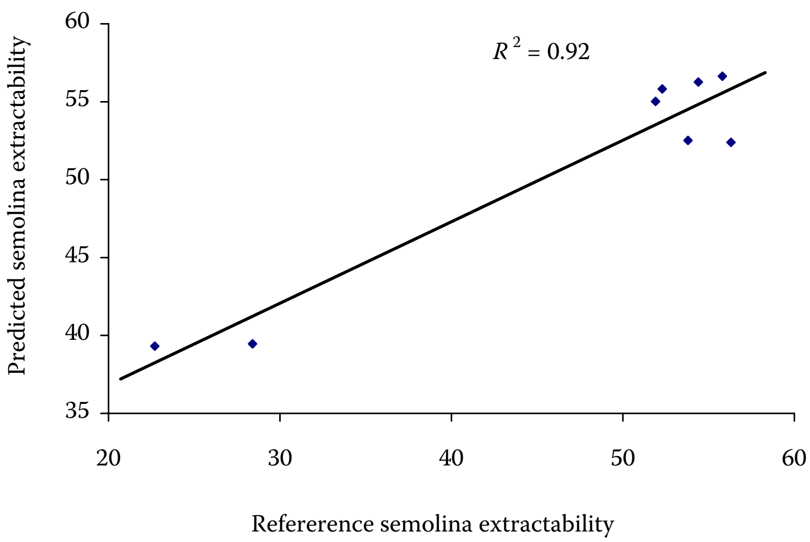

Figure 1. Comparison of semolina extractability determined by prediction model and by reference method

based on the standard error of prediction (SEP) and the linear correlation coefficient of prediction (r). mPLS/PLS regressions gave SEP values ranging from 2.04% to 6.71% and r values from 0.313 to 0.920. The best result of the independent validation was achieved for semolina extraction rate, for which the correlation coefficient between the reference values and the values predicted by the calibration equation indicates a statistically significant prediction model at the probability level of 99% (Figure 1). The independent validation using the ANN method confirmed the significance of prediction of semolina extraction rate by achieving $r = 0.852$ (SEP = 7.70%). Apart from this milling characteristic, low coefficients of correlations were found for other two traits studied. The results achieved by mPLS/PLS and ANN methods were comparable for the flour extraction rate (mPLS/PLS: $r = 0.313$, ANN: $r = 0.265$), while the outcomes for semolina extractability obtained by the two methods showed a lesser resemblance.

CONCLUSION

The results of this work demonstrate the use of NIR technique for the assessment of the milling characteristics of wheat. Near infrared spectroscopy of laboratory prepared wheat flour combined with MPLS/PLS and ANN regression is a simple and rapid technique for determining the wheat milling characteristics. Statistically significant dependence between the predicted and the reference values at the probability level of 99% was found by calibration for all milling characteristics studied in both variety and commercial subsets. Semolina extractability showed the strongest statistical correlation. Similar results were found in the case of ANN method, higher r (and lower SEV) was achieved. The potential of NIR spectroscopy to predict the milling characteristics was examined by an independent validation by a set of flours but the accuracy of the validation results was not fully satisfactory. A statistically significant prediction model was achieved for semolina extraction rate, prediction models for other traits of these sets did not demonstrate acceptable results.

To the best of our knowledge, this is the introduction to demonstrate the potential of NIR to predict quantitative milling parameters of wheat. Further examination is needed to extend the calibration and validation sets with a higher number of samples, particularly in the case of ANN. Since NIR is a predictive method, it has inherently built in errors associated with the reference method plus other errors that could occur (sample presentation, sampling, scatter effects, storage).

As the obtained dependences and prediction models are based on NIR spectroscopy of wheat flour, standard laboratory mill must be available to enable to assess the studied quantitative traits of the milling process by NIR technique. Developing an NIR calibration-and-prediction model using the spectral data of whole grain might be the next appropriate step.

References

Delwiche S.R., Weaver G. (1994): Bread quality of wheat flour by near-infrared spectrophotometry: feasibility of modelling. Journal of Food Science, **59**: 410–415.

Dowell F.E. (1998): Automated color classification of single wheat kernels using visible and near-infrared reflectance. Cereal Chemistry, **75**: 142–144.

Duponchel L., Ruckebusch C., Huvenne J.P., Legrand P. (1999): Standardisation of near infrared spectrometers using artificial neural networks. Journal of Near Infrared Spectroscopy, **7**: 155–166.

Gottlieb D.M., Schultz J., Bruun S.W., Jacobsen S., Søndergaard I. (2004): Multivariate approaches in plant science. Phytochemistry, **65**: 1531–1548.

Maghirang E.B., Dowell F.E. (2003): Hardness measurement of bulk wheat by single-kernel visible and near-infrared reflectance spectroscopy. Cereal Chemistry, **80**: 316–322.

Miralbés C. (2004): Quality control in the milling industry using near-infrared transmittance spectroscopy. Food Chemistry, **88**: 621–628.

Nielsen J.P., Bertrand D., Micklander E., Courcoux P., Munck L. (2001): Study of NIR spectra, particle size distributions and chemical parameters of wheat flours: a multi-way approach. Journal of Near Infrared Spectroscopy, **9**, 275–284.

Osborne B.G. (1984): Investigations into the use of near infrared reflectance spectroscopy for the quality assessment of wheat with respect to its potential for bread baking. Journal of the Science of Food and Agriculture, **35**: 106–110.

Osborne B.G., Kotwal Z., Wesley I.J., Saunders L., Dardenne P., Shenk J.S. (1999): Optical matching of near infrared reflectance monochromator instruments for the analysis of ground and whole wheat. Journal of Near Infrared Spectroscopy, **7**: 167–178.

Osborne B.G. (2000): Recent developments in NIR analysis of grains and grain products. Cereal Foods World, **45**: 11–15.

Pawlinsky T., Williams P. (1998): Prediction of wheat bread-baking functionality in whole kernels, using near infrared reflectance spectroscopy. Journal of Near Infrared Spectroscopy, **6**: 121–127.

Rubenthaler G.L., Pomeranz Y. (1987): Near-infrared reflectance spectra of hard red winter wheats varying widely in protein content and breadmaking potential. Cereal Chemistry, **64**: 407–411.

Wesley I.J., Larsen N., Osborne B.G., Skerritt J.H. (1998): Non-invasive monitoring of dough mixing by near infrared spectroscopy. Journal of Cereal Science, **27**: 61–69.

Wesley I.J., Uthayakumaran S., Anderssen R.S., Cornish G.B., Bekes F., Osborne B.G., Skerritt J.H. (1999): A curve-fitting approach to the near infrared reflectance measurement of wheat flour proteins which influence dough quality. Journal of Near Infrared Spectroscopy, **7**: 229–240.

Wesley I.J., Larroque O., Osborne B.G., Azudin N., Allen H., Skerritt J.H. (2001): Measurement of gliadin and glutenin content of flour by NIR spectroscopy. Journal of Cereal Science, **34**: 125–133.

Williams P.C., Jaby El-Haramein F., Ortiz-Fereira G., Srivastava J.P. (1988): Preliminary observations on the determination of wheat strength by near infrared spectroscopy. Cereal Chemistry, **65**: 109–114.

Received for publication November 9, 2004
Accepted after corrections February 15, 2005

Corresponding author:

Ing. Marie Hrušková, CSc., Vysoká škola chemicko-technologická v Praze, Fakulta potravinářské a biochemické technologie, Ústav chemie a technologie sacharidů, Technická 5, 166 28 Praha 6, Česká republika
tel.: + 420 220 443 206, fax: + 420 233 199 990, e-mail: marie.hruskova@vscht.cz