The Canadian Society for Bioengineering
The Canadian society for engineering in agricultural, food, environmental, and biological systems.

CSBE | SCGAB

La Société Canadienne de Génie Agroalimentaire et de Bioingénierie
La société canadienne de génie agroalimentaire, de la bioingénierie et de l'environnement

Paper No. CSBE11-119

Near-Infrared Hyperspectral Imaging for Protein and Hardness Predictions of Bulk Samples of Western Canadian Wheat from Different Locations and Crop Years Using Multivariate Regression Models

Sivakumar Mahesh

Department of Biosystems Engineering, E2-376 EITC, University of Manitoba, Winnipeg, MB, Canada R3T 5V6.

Digvir S. Jayas

Department of Biosystems Engineering, E2-376 EITC, University of Manitoba, Winnipeg, MB, Canada R3T 5V6.

Jitendra Paliwal

Department of Biosystems Engineering, E2-376 EITC, University of Manitoba, Winnipeg, MB, Canada R3T 5V6.

Noel D.G. White

Cereal Research Centre, Agriculture and Agri-Food Canada, Winnipeg, MB, Canada R3T 2M9

Written for presentation at the
CSBE/SCGAB 2011 Annual Conference
Inn at the Forks, Winnipeg, Manitoba
10-13 July 2011

ABSTRACT The objective of this study was to develop predictive multivariate regression models for estimating protein contents and hardness of bulk samples of wheat that were separated by locations and crop years. In this study, Canada Western Red Spring (CWRS), Canada Western Hard White Spring (CWHWS), Canada Western Soft White Spring (CWSWS), and Canada Prairie Spring Red (CPSR) wheat samples were used. Wheat samples of different classes were obtained from nearby farms of different major wheat growing locations of Manitoba, Alberta, and Saskatchewan provinces and from 2007, 2008, 2009 crop years. The samples were conditioned to moisture levels of 13, 16, and 19% (wet basis) and pooled

together for model development purposes. Image and spectral information of bulk samples in the spectral region between 960 and 1700 nm were acquired using the NIR hyperspectral imaging system at every 10 nm intervals. Reference protein contents and hardness values were determined using the Dumatherm method and single kernel characterization system (SKCS), respectively. In the protein content prediction study, using the full data set, the ten-component partial least squares regression (PLSR), which had 1.76%, 1.33%, and 0.68 for the estimated mean square error of prediction (MSEP), standard error of cross validation (SECV), and correlation coefficient (r), respectively, gave better results than the ten-component PCR model. In the hardness prediction part, the estimated MSEP, SECV, and r values were 147.7, 12.15, and 0.82, respectively, for the ten-component PLSR model. The NIR hyperspectral imaging can be used to predict quality parameters of wheat based on locations and crop years.

Keywords: wheat classes, protein, hardness, hyperspectral

INTRODUCTION Agricultural materials, which contain complex chemical components such as protein, carbohydrates, and fat, are heterogeneous in nature and vary in chemical components from one sample to the other. In Canada, wheat, which is ranked at the top in the list of agricultural crops grown followed by canola and maize, was produced and exported in quantities of 28.6 and 15.7 Mt, respectively, in 2008 (FAOSTAT, 2011). Wheat grades determined by the Canadian Grain Commission (CGC) ensure a satisfactory performance in wheat quality and milling potential; and prove that grains are contaminant free (CWB, 2011). Classification of wheat is mainly based on its colour (red vs. white), hardness (soft vs. hard), and growing season (winter vs. spring). Eight major western Canadian wheat classes, which are produced throughout the prairie provinces (Manitoba, Saskatchewan, and Alberta) and exported to various countries around the world, are: Canada Western Red Spring (CWRS), Canada Prairie Spring Red (CPSR), Canada Western Extra Strong (CWES), Canada Western Red Winter (CWRW), Canada Prairie Spring White (CPSW), Canada Western Amber Durum (CWAD), Canada Western Soft White Spring (CWSWS), and Canada Western Hard White Spring (CWHWS).

Wheat has higher protein content, which varies between 10 and 18% of the total dry matter, than maize, paddy, or other cereal crops. Protein content of wheat is considered as one of the basic intrinsic properties and it always has an effect on functional properties of processed products. Also, the presence of specific type of protein in wheat affects the baking quality. MacRitchie (1987) described from the study to evaluate protein fractions of wheat for bread making and dough mixing that the relative amount of globulin-type to the glutenin-type proteins seemed to be influential in baking quality.

Protein content of single kernels and bulk samples of wheat were calculated from near infrared reflectance spectroscopic values of individual kernels (Delwiche, 1998, 2000). A near-infrared transmittance spectrophotometer was used to find out the practicability of measuring protein contents of intact kernels of wheat (Delwiche, 1995). Watson et al. (1977) developed regression models for protein content of wheat using near-infrared reflectance spectroscopy. They further reported that the near-infrared reflectance values were affected by and were based on the hardness of wheat which is the key determinant of end usage. Williams (1979) utilized near infrared reflectance spectroscopy to inspect wheat for protein and hardness. Slaughter et al. (1992) specified that protein and hardness were best suited for discriminating between hard red spring and hard red winter wheat classes.

Spectral imaging is a modern technology that combines spectroscopy with image processing. Hyperspectral imaging (imaging spectrometry), an extension of multispectral imaging, is becoming a popular research tool from which both spectral and spatial information of samples can be acquired simultaneously. Hyperspectral imaging provides a large data set, otherwise

called a data cube, which facilitates a complete and reliable analysis of intrinsic properties and external characteristics of samples. It is a recognized tool that permits spectroscopic image analysis of a sample or a point within the region of interest using chemical sensing and image processing techniques (Headwall, 2011). Typically, two to ten spectral images are collected in the multispectral imaging, but, up to several hundred images can be acquired by hyperspectral imaging systems (Lawrence et al., 2003). In the laboratory, hyperspectral imaging instruments use wavelengths in the near infrared, infrared, and Raman regions. The NIR hyperspectral imaging has been used for classification of wheat classes (Mahesh et al., 2008, 2011a, 2011b; Choudhary et al., 2009). Singh et al. (2007) detected different fungal species in wheat using the LWNIR hyperspectral imaging in the wavelength region of 1000-1600 nm. Xing et al. (2009) evaluated the α-amylase activity in single kernels of wheat using a short wavelength infrared (SWIR) hyperspectral imaging system in the wavelength region of 1000-2500 nm.

Considering the advantages of using the LWNIR hyperspectral imaging, the objectives of this study were to investigate hyperspectral image cubes of wheat for assessing quality parameters such as protein and hardness using partial least squares regression (PLSR) and principal components regression (PCR) models.

METHODS AND MATERIALS

Grain Samples Samples of four western Canadian wheat classes, two red and two white, of varying degrees of hardness and protein grown at different locations in the Prairie provinces (Manitoba, Saskatchewan, and Alberta of western Canada) in 2007, 2008, and 2009 crop years were collected and used for this study (Table 1). The wheat classes used were CWRS, CPSR, CWHWS, and CWSWS. Selection method used for identifying locations, sample conditioning, moisture determination, and imaging procedures were discussed in detail in Mahesh et al. (2011b).

Table 1. Growing locations (crop years) near the listed towns from where wheat samples were collected. CWRS - Canada Western Red Spring, CPSR - Canada Prairie Spring Red, CWHWS - Canada Western Hard White Spring, and CWSWS - Canada Western Soft White Spring, AB - Alberta, MB – Manitoba, and SK – Saskatchewan.

CWRS	CPSR	CWHWS	CWSWS
Corning, SK (2008, 2009)	Edmonton, AB (2008, 2009)	Churchbridge, SK (2008, 2009)	Corning, SK (2008, 2009)
Dauphin, MB (2008, 2009)	Rosemary, AB (2008, 2009)	Kenton, MB (2008, 2009)	Jansen, SK (2008, 2009)
Domain, MB (2008, 2009)	Viking, AB (2007, 2009)	Limerick, SK (2008, 2009)	Kenton, MB (2008, 2009)
Melfort, SK (2008, 2009)	Wainwright, AB (2008, 2009)	Mather, MB (2008, 2009)	Nokomis, SK (2008, 2009)
Tisdale, SK (2008, 2009)	Corning, SK (2009)	Shaunavon, SK (2008, 2009)	Wilkie, SK (2008, 2009)
	Unity, SK (2008)		

Near-infrared hyperspectral imaging system and image processing The near-infrared hyperspectral imaging system consisted of a near-infrared camera with two VariSpec liquid crystal tunable filters (LCTFs) (Model No. MIR06, Cambridge Research and Instrumentation Inc., Woburn, MA), a 25 mm F1.4 C-mount lens (Electrophysics Corp. Fairfield, NJ), a sample stage, and a light source controlled through a Dell Optiplex GX280 Intel(R) (Dell Inc., Round Rock, TX) computer (Figure 1). Mahesh et al. (2011b) provided the specific details of the system, spatial calibration, image standardization, image acquisition, and image analysis.

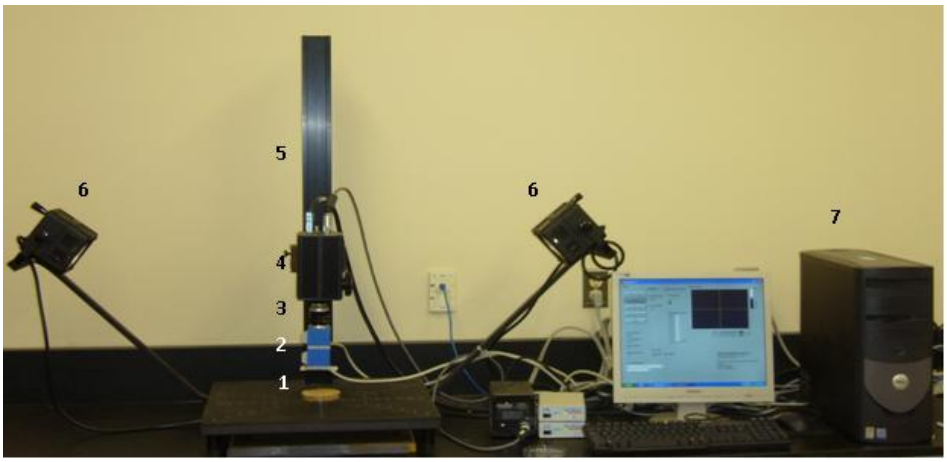

Fig. 1. The NIR hyperspectral imaging system. 1. Bulk wheat sample, 2. Liquid crystal tunable filter (LCTF), 3. Lens, 4. NIR camera, 5. Copy stand, 6. Illumination, 7. Data processing system.

Partial least squares regression (PLSR) model The PLSR was used to form the prediction model between spectral responses of tested samples of wheat and their quality parameters such as protein and hardness. The development PLSR algorithm, which was used in this study, was reported for determining quality parameters of strawberries (ElMasry et al., 2007). The PLSR analysis was conducted between the NIR reflectance intensities and the interested parameter (protein or hardness) using Matlab (Matlab R2009b, MathWorks, Natick, MA). Maximization of the spectral data, which was formed from computing the average spectra with 75 wavelengths in wavelength range of 960-1700 nm, was done using the PLSR. The transfer of highly correlated data, which was also large in size and often co-linear in nature, into independent latent variables was performed. A mathematical relationship between a set of independent variables, X matrix ($\mathrm{N}_{7200\ \mathrm{samples}} \times \mathrm{K}_{75\ \mathrm{wavelengths}}$), and the dependent variable, Y vector ($\mathrm{N}_{7200\ \mathrm{samples}} \times 1$) was obtained for protein and hardness models using an overall data set having 7200 spectra in total. The PLSR produced a regression model, PLS components that were optimal in number, and the predicted value of the parameter for each scanned sample. The number of PLS components, which were optimal for developing the model, were selected using the percent variance explained by the components and the minimum value of estimated mean square error of prediction (MSEP). In regression analysis, mean squared error, which was referred as the estimate of error variance, was calculated by dividing residual sum of squares by the number of degrees of freedom. The standard error of cross-validation (SECV), estimated MSEP, and r values were used for evaluating the PLSR model.

The optimal wavelengths were identified using the PLS β-coefficients. As maximum spectral information was explained, these wavelengths were used for on-line multispectral imaging applications. Non-significant wavelengths, which had the low absolute values of β-coefficients, could completely be rejected as they had zero involvement in the prediction of intrinsic attributes of wheat. The prediction performance was estimated using the 10-fold cross validation method (Cogdill et al., 2004). In the k-fold cross-validation, which can estimate model performances accurately (Refaeilzadeh et al., 2011), the original data set was randomly divided into 10 subsample groups (720 samples in each group). The PLSR model was developed using data from 9 subsample groups (720×9 = 6,480 samples, in total). The PLSR or PCR prediction models of each of the ten groups, which had 720 samples in a group, were validated using the calibration equations derived from the NIR spectra of the rest of the nine groups of wheat samples. The cross-validation was repeated ten times by which data from each of the

subsample group was used exactly once for validating the model. Average of ten iterations was reported for the model performance.

Principal components regression (PCR) model The PCR model, which was employed for developing protein and hardness models using the full spectral data set (960-1700 nm) of 7200 sample spectra in total, was developed to avoid prediction instabilities caused by co-linear nature of predictor data set (P). In this method, uncorrelated principal components, which had decreasing variances, were generated from the NIR reflectance intensities. Regression models were developed between principal component scores and the attributes of interest (protein and hardness). A ten-fold cross-validation, which was explained in detail in the previous section, was used for validating the PCR models. The estimated MSEP values were used for finding out the optimal number of principal components. The SECV, estimated MSEP, and r of the predictive model were used for evaluating the PCR model.

RESULTS AND DISCUSSION

Tables 2 and 3 show the summary of dataset statistics and results of grouping that was performed using the Scheffe's test for protein contents and hardness of wheat, respectively.

Table 2. Summary of protein dataset statistics for wheat (n = 30 per wheat class). CWRS - Canada Western Red Spring, CPSR - Canada Prairie Spring Red, CWHWS - Canada Western Hard White Spring, and CWSWS - Canada Western Soft White Spring

Wheat class	Protein content			
	Average (%)*	Maximum (%)	Minimum (%)	Standard deviation (%)
CWRS	14.62[a]	16.19	11.88	1.28
CPSR	13.26[b]	15.39	11.38	1.13
CWHWS	14.38[a]	16.45	12.12	1.44
CWSWS	11.25[c]	12.93	9.51	1.06

Table 3. Summary of hardness dataset statistics for wheat (n = 300 kernels per wheat class per moisture). CWRS - Canada Western Red Spring, CPSR - Canada Prairie Spring Red, CWHWS - Canada Western Hard White Spring, and CWSWS - Canada Western Soft White Spring.

Wheat class	Hardness			
	Average*	Maximum	Minimum	Standard deviation
CWRS	70.8[ab]	87	57	8.73
CPSR	66.1[b]	78.8	54.5	7.21
CWHWS	78.9[a]	86.6	66.2	7.87
CWSWS	27.9[c]	37.9	12.8	8.67

For both Tables 2 and 3, * = average values with same alphabet in the superscript are not significantly different.

Prediction of protein contents The PLSR and PCR models were developed using the average spectra of wheat for the full spectral range of 960-1700 nm (75 wavebands). A ten-fold cross-validation method was used for both PLSR and PCR models for selecting the optimal number of components that were involved in model development processes. The first 3 components of the PLSR and PCR accounted for 99.9% of variations in the input data set. Ten iterations were performed in the cross-validation and mean square errors of prediction (MSEP) were estimated for explaining model performances. The optimal number of components for the PLSR and PCR was selected by considering the component that had the minimum MSEP values. Figure 2 shows the estimated MSEP values for components used in the PLSR and PCR models for predicting protein contents of wheat samples. The minimum MSEP values were 2.02 and 1.76 for the PCR and PLSR with ten components, respectively.

The fitted protein contents against the observed protein contents of wheat samples for the ten-component PLSR and ten-component PCR models are shown in Figs. 3 and 4, respectively.

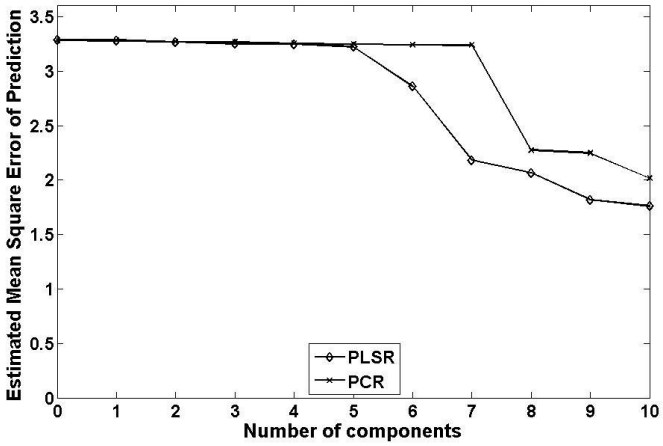

Fig. 2. Estimated mean square error of prediction values for the PLSR and PCR models for predicting protein contents of wheat. PLSR – Partial least squares regression, PCR – Principal components regression.

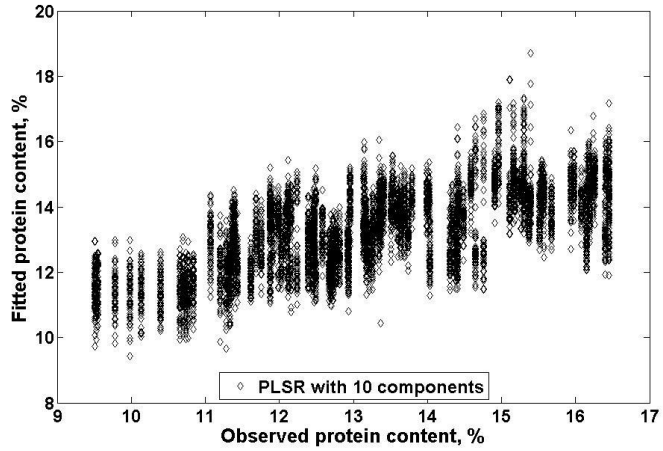

Fig. 3. Fitted protein contents against observed protein contents of wheat using the ten-component PLSR model (n = 7200, r = 0.68, and SECV = 1.33). PLSR – Partial least squares regression, SECV – Standard error of cross validation.

The PLSR, which included first 10 PLS components in the model, had an estimated MSEP of 1.76% with a SECV of 1.33% and an r of 0.68. The PCR, which had first 10 components in the model, gave 2.02%, 1.42%, and 0.62 for the estimated MSEP, SECV, and r, respectively. Cogdill et al. (2004) reported an SECV of 1.2% and an r^2 of 0.87 for predicting moisture concentrations in individual kernels of maize using the NIR hyperspectral imaging. ElMasry et al. (2007) observed that the PLSR method was superior for predicting moisture, total soluble solids content, and pH for strawberries using the Vis-NIR hyperspectral images acquired from the wavelength region of 400-1000 nm. Also, in the same study, the moisture contents of strawberries were precisely predicted using the PLSR with an r value 0.96 for the validation set.

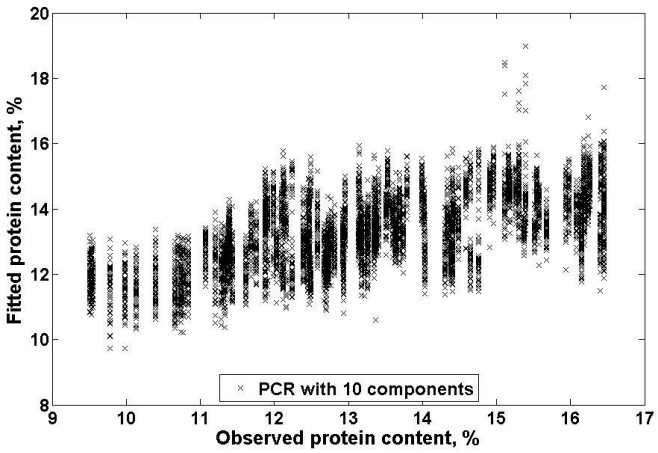

Fig. 4. Fitted protein contents against observed protein contents of wheat using the ten-component PCR model (n = 7200, r = 0.62, and SECV = 1.42). PLSR – Partial least squares regression, SECV – Standard error of cross validation.

Prediction of hardness Average NIR reflectance spectra of 960-1700 nm for 7200 wheat samples were used for predicting hardness using PLSR and PCR. The first three components of the PLSR and PCR accounted for 99.9% of variations in the input data set. The details of performing the ten-fold cross-validation were explained in the earlier section. Ten replications were performed in the cross-validation process and estimated MSEP were estimated. Figure 5 shows the estimated MSEP values for components of the PLSR and PCR models for predicting hardness of wheat samples. The minimum MSEP values were 193.1 and 147.7 for the tenth component of PCR and PLSR models, respectively.

Pomeranz et al. (1988) stated that spring wheat classes had more hardness values than winter wheat classes. The ten-component PLSR gave relatively better hardness predictions than the ten-component PCR for wheat. With the first 10 components included, the PLSR calibration had an estimated MSEP of 147.7 with a SECV of 12.15 and r of 0.82. Lu (2001) reported that the PLSR models had r values of 0.80 and 0.65 for predicting firmness of Hedelfinger and Sam cherries, respectively. The PCR, which had first 10 components in the model, had 193.1, 13.9, and 0.75 for the estimated MSEP, SECV, and r, respectively. The predicted hardness against the observed hardness of wheat samples for the ten-component PLSR and ten-component PCR models are shown in Figs. 6 and 7, respectively.

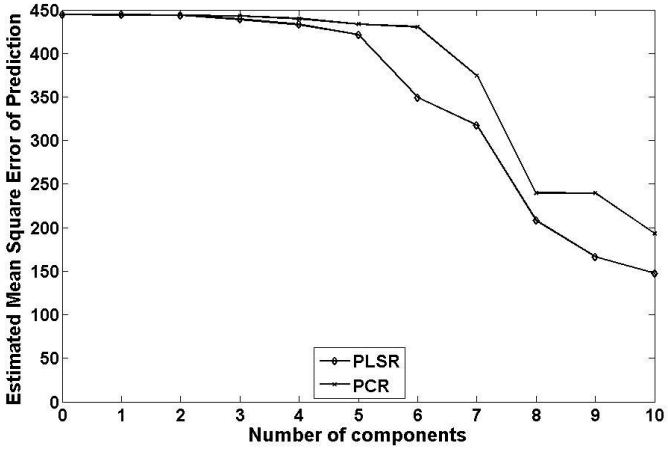

Fig. 5. Estimated mean square error of prediction values for the PLSR and PCR models for predicting hardness of wheat. PLSR – Partial least squares regression, PCR – Principal components regression.

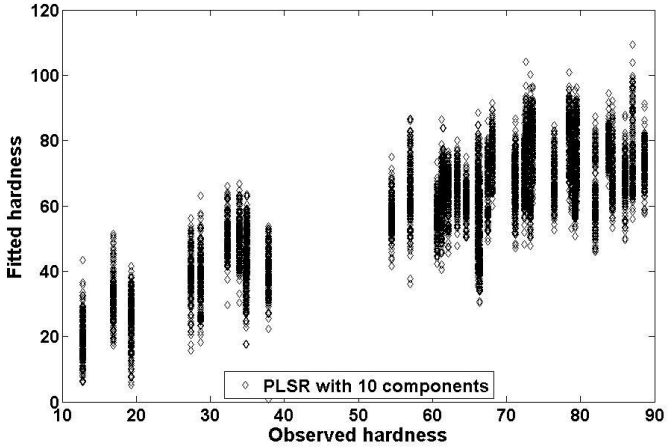

Fig. 6. Fitted hardness against observed hardness of wheat using the ten-component PLSR model (n = 7200, r = 0.82, and SECV = 12.15). PLSR – Partial least squares regression, SECV – Standard error of cross validation.

Based on the values of β-coefficients and the PC loading values of the PLSR and PCR models, the top 10 wavelengths for predicting protein contents and hardness of wheat were identified (Table 4). For predicting protein contents of wheat, using the PLSR, the NIR reflectance features from the wavelength regions of 1180-1200 and 1460-1500 nm were identified as important. The NIR absorptions of wheat in the wavelength 1480 nm are attributed to protein contents (Delwiche and Massie, 1996). Wavelengths from the regions of 960-1030 and 1670-1700 nm were identified crucial by the PCR model. In the hardness prediction study, the NIR reflectance features from the wavelength regions of 1180-1220, 1320-1400, and 1460-1490 nm were found very necessary by the PLSR. Wavelengths from the regions of 960-1030 and 1670-1700 nm were identified essential by the PCR. Maghirang and Dowell (2003) reported that the NIR wavelengths of 1100, 1200, 1380, 1450, and 1670 nm were contributed mainly in predicting hardness of wheat samples using the PLSR.

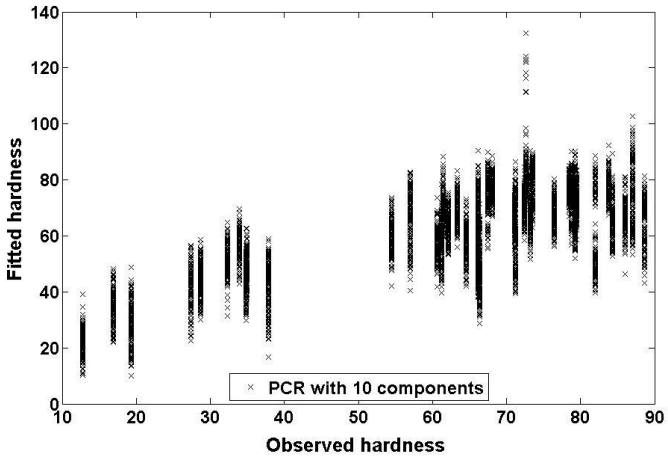

Fig. 7. Fitted hardness against observed hardness of wheat using the ten-component PCR model (n = 7200, r = 0.75, and SECV = 13.9). PLSR – Partial least squares regression, SECV – Standard error of cross validation.

The NIR absorptions of wheat in the wavelength regions of 960-1060, 1330-1480 nm, and at 1680 nm are likely associated to protein, starch, and hardness of samples (Maghirang and Dowell, 2003; Murray and Williams, 1987). These wavelengths can be added to a multispectral imaging system that will be used for conducting on-line quality assessments in wheat in the future.

Table 4. Top 10 wavelengths, in descending order, identified by the ten-component PLSR and PCR models for predicting protein contents and hardness of wheat samples. PLSR – Partial least squares regression, PCR – Principal components regression.

	Protein prediction		**Hardness prediction**	
Rank	**Top 10 wavelengths by PLSR**	**Top 10 wavelengths by PCR**	**Top 10 wavelengths by PLSR**	**Top 10 wavelengths by PCR**
1	1190	1700	1390	1700
2	1210	1690	1220	1690
3	1180	1670	1180	1670
4	1090	1680	1320	1680
5	1670	960	1480	960
6	1220	970	1190	970
7	1130	980	1210	980
8	1490	1030	1460	1030
9	1500	1020	1490	1020
10	1460	990	1400	990

CONCLUSION

The NIR reflectance features were extracted from hyperspectral image cubes of bulk samples of wheat. Ten-component PLSR and PCR models were developed for predicting protein contents

and hardness of wheat. The PLSR models had r values of 0.68 and 0.82, which were higher than those of PCR models, for predicting protein contents and hardness of wheat. Key wavelengths, which were mostly useful for protein content and hardness prediction, were identified. Overall, PLSR models demonstrated better prediction performances than the PCR models for predicting protein contents and hardness of wheat. The PLSR models are thus recommended for predicting protein contents and hardness of western Canadian wheat classes.

ACKNOWLEDGEMENTS

We are thankful to the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Canada Research Chairs (CRC) Program, the Canada Foundation for Innovation (CFI), Manitoba Research Innovation Fund, and the University of Manitoba Graduate Fellowship (UMGF) for partial funding of this study. We thank Mr. Colin Demianyk, Biologist, Agriculture and Agri-Food Canada for taking strenuous efforts for collecting samples from different locations and crop years and Mr. Jean-Christophe Habeck for assistance in image acquisition.

REFERENCES

CGC. 2011. Quality of western Canadian wheat 2009. http://www.grainscanada.gc.ca/wheat-ble/harvest-recolte/2009/hqww09-qrbo09-2-eng.htm (13/04/2011).

Choudhary, R., S. Mahesh, J. Paliwal, and D.S. Jayas. 2009. Identification of wheat classes using wavelet features from near infrared hyperspectral images of bulk samples. *Biosystems Engineering* 102(2): 115-127.

Cogdill, R.P., C.R. Hurburgh, Jr., and G.R. Rippke. 2004. Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE* 47(1): 311-320.

Delwiche, S.R. 1995. Single wheat kernel analysis by near infrared transmittance: protein content. *Cereal Chemistry* 72(1): 11-16.

Delwiche, S.R. 1998. Protein content of single kernels of wheat by near infrared reflectance spectroscopy. *Journal of Cereal Science* 27: 241-254.

Delwiche, S.R. 2000. Protein content of bulk wheat from near-infrared reflectance of individual kernels. *Cereal Chemistry* 77(1): 86-88.

Delwiche, S.R. and D.R. Massie. 1996. Classification of wheat by visible and near-infrared reflectance from single kernels. *Cereal Chemistry* 73(3): 399-405.

ElMasry, G., N. Wang, A. ElSayed, and M. Ngadi. 2007. Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering* 81: 98-107.

FAOSTAT. 2011. Food and Agriculture Organization of the United Nations, Rome. http://faostat.fao.org (14/04/2011).

Headwall. 2011. Spectral imaging capabilities of hyperspec™ imaging technology in pharmaceutical operations. http://www.headwallphotonics.com/downloads/hw_hyperspectral-in-pharma.pdf (13/04/2011).

Lawrence, K.C., B. Park, W.R. Windham, and C. Mao. 2003. Calibration of pushbroom hyperspectral imaging system for agricultural inspection. *Transactions of the ASAE* 46(2): 513-521.

Lu, R. 2001. Predicting firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. *Transactions of the ASAE* 44(5): 1265-1271.

MacRitchie, F. 1987. Evaluation of contributions from wheat protein fractions to dough mixing and bread making. *Journal of Cereal Science* 6(3): 259-268.

Maghirang, E.B. and F.E. Dowell. 2003. Hardness measurement of bulk wheat by single-kernel visible and near-infrared reflectance spectroscopy. *Cereal Chemistry* 80(3): 316-322.

Mahesh, S., A. Manickavasagan, D.S. Jayas, J. Paliwal, and N.D.G. White. 2008. Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering* 101: 50-57.

Mahesh, S., D.S. Jayas, J. Paliwal, and N.D.G. White. 2011a. Identification of wheat classes at different moisture levels using near-infrared hyperspectral images of bulk samples. *Sensing and Instrumentation for Food Quality and Safety* 5(1): 1-9.

Mahesh, S., D.S. Jayas, J. Paliwal, and N.D.G. White. 2011b. Neural network prediction of wheat classes and moisture contents using near-infrared (NIR) hyperspectral images of bulk samples from different locations and crop years. ASAE Paper No. 1110539. St. Joseph, MI: ASABE.

Murray, I. and P.C. Williams. 1987. Chemical principles of near-infrared technology. In *Near-infrared Technology in the Agricultural and Food Industries*, 17-34. St. Paul, MN: AACC.

Pomeranz, Y., Z. Czuchajowska, M.D. Shogren, G.L. Rubenthaler, L.C. Bolte, H.C. Jeffers, and P.J. Mattern. 1988. Hardness and functional (bread and cookie-making) properties of U.S. wheats. *Cereal Foods World* 33: 297-304.

Refaeilzadeh, P., L. Tang, and H. Liu. 2011. Cross-validation. http://www.public.asu.edu /~ltang9/papers/ency-cross-validation.pdf (24/04/2011).

Singh, C.B., D.S. Jayas, J. Paliwal, and N.D.G. White. 2007. Fungal detection in wheat using near-infrared hyperspectral imaging. *Transactions of the ASABE* 50(6): 2171-2176.

Slaughter, D.C, K.H. Norris, and W.R. Hruschka. 1992. Quality and classification of hard red wheat. *Cereal Chemistry* 69(4): 428-432.

Watson, C.A., W.C. Shuey, O.J. Banasik, and J.W. Dick. 1977. Effect of wheat class on near infrared reflectance. *Cereal Chemistry* 54(6): 1264-1269.

Williams, P.C. 1979. Screening wheat for protein and hardness by near infrared reflectance spectroscopy. *Cereal Chemistry* 56(3): 169-172.

Xing, J., P.V. Hung, S. Symons, M. Shahin, and D. Hatcher. 2009. Using a short wavelength infrared (SWIR) hyperspectral imaging system to predict alpha amylase activity in individual Canadian western wheat kernels. *Sensing and Instrumentation for Food Quality and Safety* 3: 211-218.