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

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

CSBE | SCGAB

Paper No. 08-196

Identification of western Canadian wheat classes at different moisture levels using near-infrared (NIR) hyperspectral imaging

S. Mahesh

Biosystems Engineering, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada

D.S. Jayas

Biosystems Engineering, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada

J. Paliwal

Biosystems Engineering, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada

N.D.G. White

Agriculture and Agri-Food Canada, 195 Dafoe Road, Winnipeg, MB, R3T 2M9, Canada

Written for presentation at the
CSBE/SCGAB 2008 Annual Conference
North Vancouver, British Columbia
July 13 - 16, 2008

Abstract

Wheat class identification is one of the major tasks in grain handling facilities. Though some wheat classes have similar appearance, they can possibly have different chemical compositions. Wheat classes at different moisture levels need to be identified in order to accurately segregate, properly dry, and safely store before processing. In this study, a near-infrared (NIR) hyperspectral imaging technique was used to identify five western Canadian wheat classes at different moisture levels. Wheat bulk samples were scanned in the 960 – 1700 nm wavelength region at 10 nm intervals using an Indium Gallium Arsenide (InGaAs) NIR camera. Seventy five relative reflectance intensities were extracted from the scanned images to develop classification models using a statistical classifier. Using linear discriminant analysis (LDA), classification accuracies of 81 – 100% were obtained in identifying the wheat classes with each wheat class at five different moisture levels. Classification accuracies were generally 60 – 89% using quadratic discriminant analysis (QDA) for identifying wheat classes at different moisture levels. Near-

infrared hyperspectral imaging technique can be used to identify western Canadian wheat classes at different moisture levels.

Keywords: Near-infrared hyperspectral imaging, wheat classes, statistical classifier.

INTRODUCTION

Canadian wheat production and export were 26.7 and 14.0 Mt, respectively, in 2005 (FAOSTAT 2008). Though it is advisable to perform wheat harvesting at a moisture level of 13 – 15%, it could be done at higher moisture levels (above 15%) immediately followed by an effective drying process (McNeill and Overhults 2008). In general, wheat is procured from farmers at different moisture levels by grain handling facilities. High moisture wheat should be dried to an optimal moisture level (12 – 13%) to store safely and prevent spoilage and/or sprouting prior to processing. In Canada, wheat is classified based on its color (red or white), hardness (soft or hard), and growing season (winter or spring). A specific wheat class is used as a primary raw material for products such as bread, pasta, noodles, and flat bread. Visual method is commonly used to identify wheat classes in the grain handling facilities in Canada. Various other methods such as polyacrylamide gel electrophoresis (PAGE), high performance liquid chromatography (HPLC) have also been used for identifying wheat classes by measuring wheat proteins (CGC 2008). A machine vision approach has been used to distinguish two wheat classes (Canada Western Red Spring (CWRS) and Canada Western Amber Durum (CWAD)), barley, oats, and rye (Paliwal et al. 1999). Wheat lots will not be at a uniform moisture level when they reach primary/terminal elevators or other processing facilities. In this scenario, the visual method can not be used for identifying wheat at different moisture levels because of the subjectivity of the method. It creates an immense need in the Canadian grain industry to develop a rapid and consistent method to identify wheat classes at different moisture levels.

Near-infrared spectroscopy is used in various fields such as animal husbandry, agriculture, and pharmaceuticals. In agriculture, it has been used to determine quality parameters such as protein, starch, moisture content, and oil content of different agricultural commodities such as whole (Delwiche 1998) and ground wheat (Wang et al. 2004b); deoxynivalenol levels in wheat (Petterson and Aberg 2003) and barley (Ruan et al. 2002). Identification of waxy wheat varieties and differentiating them from partially waxy wheat varieties and wild wheat phenotypes have been performed using this method (Delwiche and Graybosch 2002). It has also been used to determine four different life stages of *Sitophilus oryzae* (L.) (rice weevil) at four different infestation levels in artificially infested CWRS wheat (Paliwal et al. 2004). Armstrong (2006) determined moisture and protein contents of soybean and moisture content of corn by developing a partial least squares (PLS) model using spectra obtained from single kernels.

In the current study, near-infrared hyperspectral imaging technique that combines two important techniques, i.e., imaging and NIR spectroscopy, was used. It provides data information in the form of a hypercube that includes spatial and spectral information of samples. Effective models have already been developed using this technique to determine moisture and oil contents from single kernels of maize (Cogdill et al. 2004), firmness and soluble solids content of strawberries (Nagata et al. 2005), and to identify bitter pit lesions in apples (Nicolai et al. 2006).

Multivariate statistical tools such as partial least squares regression (PLSR), principal components analysis (PCR), multiple linear regression (MLR), linear and quadratic discriminant models, or artificial neural network (ANN) models have been used to develop classification models to detect different types of damage (Wang et al. 2002) and different types of fungal damages (Wang et al. 2004a) in soybean. These tools have been also used to identify heat damage (Wang et al. 2001) and waxy wheat (Delwiche and Graybosch 2002); and to detect insect fragments (Perez-Mendoza et al. 2003) in wheat. Classification of spoiled beef from unspoiled was done using linear and quadratic discriminant functions of a statistical classifier (Panigrahi et al. 2006). Wang et al. (2004a) used neural network models for the classification of fungal-damaged soybean seeds using NIR spectroscopy. In the current study, a statistical classifier was chosen for developing models to classify western Canadian wheat classes at different moisture levels.

The objective of this study was to determine the feasibility of NIR hyperspectral imaging for differentiating western Canadian wheat classes at different moisture levels using a statistical classifier.

METHODS AND MATERIALS

Near-infrared hyperspectral imaging system

The near-infrared hyperspectral imaging system used in this study consisted of a near-infrared camera with a liquid crystal tuneable filter (LCTF), a lens, a sample stage, and a light source controlled through a Dell Optiplex GX280 Intel(R) (Dell Inc., Round Rock, TX) computer (Fig. 1). An Indium Gallium Arsenide (InGaAs) camera (Model No. SU640-1.7RT-D, Sensors Unlimited Inc., Princeton, NJ) was used for acquiring images in the 960 – 1700 nm NIR wavelength region incremented by 10 nm intervals.

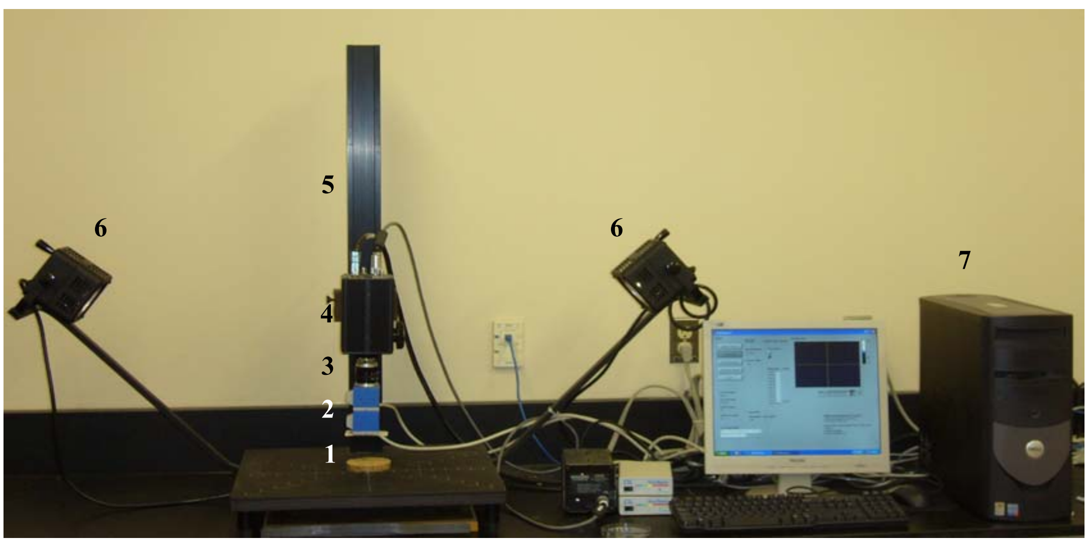


Fig. 1 NIR hyperspectral imaging system: 1. Bulk wheat sample, 2. Liquid crystal tunable filter (LCTF), 3. Lens, 4. NIR camera, 5. Copy stand, 6. Illumination, 7. Data processing system.

The electronically tunable LCTF (VariSpec model No. MIR 06, Cambridge Research and Instrumentation Inc., Woburn, MA) had a 20 mm aperture size and a 10 mm transmission bandwidth. This high quality interference filter helped to rapidly select a wavelength in the NIR region without any vibration. This filter was attached to the camera which ultimately produced 12 bit multispectral images. The data acquisition board (NI PCI-1422, National Instruments Corp., Austin, TX) was attuned to RS-422 signals generated from the camera system for image acquisition.

The sample was illuminated by a pair of 300 W alternating current (AC) halogen lights (Ushio Lighting Inc., Cypress, CA) fitted on either side of the copy stand that supported the NIR imaging system. These halogen bulbs had the capacity to emit light in a wavelength range of 400 – 2500 nm. Alternating current lights were chosen because AC sources are more convenient and can achieve higher throughputs than direct current (DC) sources. Moreover, since hyperspectral images were collected at fairly high integration times (i.e., > 250 ms), there is no advantage in using a DC source at such high integration times. A high integration time is similar to averaging several spectra in a point spectroscopy system.

Near-infrared hyperspectral images were acquired with the help of a control program written in LabView environment (Version 1, National Instruments Corp., Austin, TX). The camera was aligned to the centre wavelength of 1330 nm in the NIR camera's usable wavelength region of 960 – 1700 nm using LabView. Hyperspectral images were acquired and saved after performing the necessary system calibrations (Martens and Naes 1992).

Grain samples

For this study, five wheat class samples, viz., Canada Western Red Spring (CWRS), Canada Western Extra Strong (CWES), Canada Western Red Winter (CWRW), Canada Western Soft White Spring (CWSWS), and Canada Western Hard White Spring (CWHWS) were obtained from seed distributors of Manitoba and Saskatchewan. About 10 kg composite samples were conditioned to five different moisture levels (12, 14, 16, 18, and 20% wet basis). In Canada, wheat of the same class from different growing regions is not segregated and is marketed as a composite class. Wheat samples for imaging were prepared by randomly taking a petri dish (0.090 m in diameter and 0.011 m in depth) full of samples from the 10 composite class sample. One hundred near-infrared hyperspectral images were collected from a wheat class at each moisture level. In total, 187,500 (100 images/class/moisture × 75 wavelength slices/image × 5 wheat classes × 5 moisture levels) images were taken and analyzed.

Image analysis

Image acquisition Hyperspectral images of wheat bulk samples were collected at the equipment's usable wavelength range of 960 – 1700 nm incremented by a 10 nm interval. The NIR wavelength region was segmented into 75 slices, resulting in an NIR hyperspectral image cube consisting of 75 images (an image per slice) with the first image at 960 nm. Dark current images were taken each time prior to acquire NIR hyperspectral images. Near-infrared reflectance intensities were recorded for every pixel of the scanned image of the sample taken at each wavelength slice. The reflectance intensities mainly depend on NIR absorbance properties of the main constituents (protein, starch, oil content, and moisture content) of the sample. Near-infrared reflectance intensity of a sample will be high when it absorbs a minimum amount of NIR radiation and vice versa.

Image cropping Cropping at the center of the image was done to avoid pixels with poor reflectance intensities along the four edges of the image. An area of 200 × 200 pixels around the centre pixel was cropped from each image. Near-infrared reflectance intensity of each pixel of the cropped region in each slice of the NIR hyperspectral image was extracted. Mean NIR reflectance intensities (referred to as NIR reflectance intensities) were calculated for all wavelength slices of the NIR hyperspectral images. Single median spectrum was found from a region of interest mask of a hyperspectral image and used for developing PLS calibration models (Burger and Geladi 2006). Statistical mean centering methods used in a region of interest of a hyperspectral image help to reduce the data volume of hyperspectral images and to analyze the data easily.

Relative reflectance intensity at each wavelength slice of an NIR hyperspectral image was considered as a feature for developing classification models to differentiate wheat classes. It was determined using the following formula:

$$R = \left(\frac{S - D}{W - D} \right) \quad (1)$$

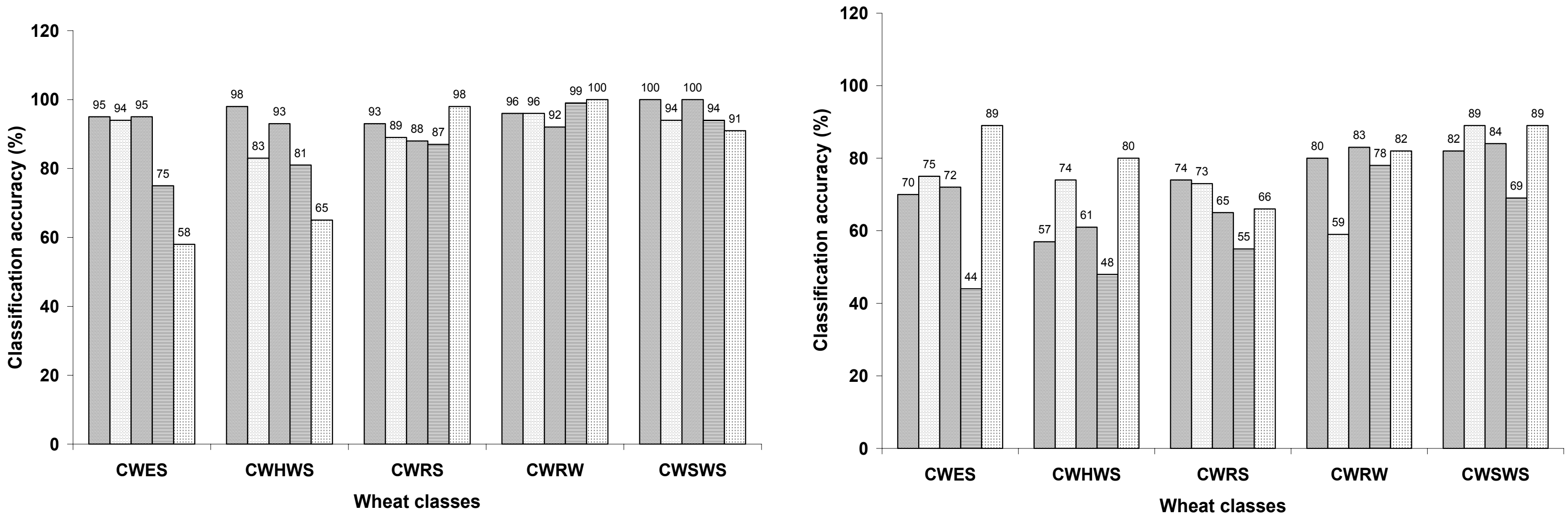
where R = relative reflectance intensity of each slice of the NIR hyperspectral image of wheat; S = reflectance intensity of each slice of the NIR hyperspectral image; D = reflectance intensity of the dark current; W = reflectance intensity of a 99% reflectance standard white panel (Labsphere, North Sutton, NH).

In total, seventy five relative reflectance intensity features were extracted from a NIR hyperspectral image of a wheat sample. Image cropping and feature extraction were done in MATLAB (Version 7, The Mathworks, Inc., Natick, MA).

Development of classification models PROC DISCRIM of SAS (Version 9.1.3, SAS Institute Inc., Cary, NC) was used for developing classification models using LDA and QDA with a leave-one-out cross validation method. PROC STEPDISC was used to identify top 10 wavelengths based on their contribution to the identification of wheat classes. The level of contribution of each wavelength was identified from the values of partial R^2 and average squared canonical correlation (ASCC). Wavelengths with the highest level of contribution were identified and subsequently removed from further analysis to find the next best wavelength. This analysis was continued until the 10^{th} rank wavelength was found.

RESULTS AND DISCUSSION

Classification accuracies of identifying western Canadian wheat classes at different moisture levels using LDA and QDA models are shown in Fig. 2. Using the LDA model, 81 – 100% classification accuracies were obtained in identifying wheat classes at different moisture levels with a leave-one-out cross validation method (Fig. 2a). Classification accuracies were ≥ 91% for all moisture levels of CWRW and CWSWS wheat classes. The LDA results showed that high protein wheat classes at high moisture levels (18% and 20%) were misclassified with other high protein wheat class(es) at high or medium (16%) moisture levels. Also, lower moisture level (12 and 14%) wheat classes were misclassified with other lower moisture level wheat class(es). Misclassifications were not seen typically among the low protein wheat classes.



a) LDA b) QDA

12% m.c. 14% m.c. 16% m.c. 18% m.c. 20% m.c.

Fig. 2 Classification accuracies of wheat classes at different moisture levels using a) LDA b) QDA models with a leave-one-out cross validation method (n = 100 samples/class/moisture): CWES – Canada Western Extra Strong, CWHWS – Canada Western Hard White Spring, CWRS – Canada Western Red Spring, CWRW – Canada Western Red Winter, CWSWS – Canada Western Soft White Spring.

The QDA model did not perform as good as the LDA model in identifying wheat classes at different moisture levels. In the QDA model, classification accuracies were 60 – 89% to identify wheat classes and their moisture levels with a leave-one-out cross validation method (Fig. 2b). Poor classification accuracies of < 50% were reported for CWES and CWHWS wheat classes at 18% moisture level. Classification accuracies were ≥ 82% to identify CWSWS wheat at all moisture levels except 18%. Also, in the QDA model, misclassifications among wheat classes were mainly based on their protein and moisture levels.

The top ten wavelengths of NIR hyperspectral images based on their contribution to identifying wheat classes at different moisture levels using STEPDISC with their partial R^2 and ASCC values are shown in Table 1. The STEPDISC results showed that all input features should be necessary for identifying wheat classes at different moisture levels. Also, any reduction in the number of input features would lead to further reduction in classification accuracies of wheat classes.

Table 1. Top ten wavelengths of NIR hyperspectral images based on their contribution to identify wheat classes at different moisture levels using STEPDISC.

No.	Wavelength (nm)	Partial R^2	ASCC
1	1310	0.66	0.03
2	1450	0.80	0.06
3	1060	0.76	0.09
4	1700	0.72	0.12
5	1330	0.55	0.13
6	1200	0.33	0.14
7	1160	0.33	0.15
8	1090	0.29	0.16
9	1490	0.28	0.16
10	1070	0.26	0.18

ASCC = Average squared canonical correlation

In this study, different amounts of relative reflectance intensities were observed for the wheat samples at different moisture levels. The amount of protein, starch, moisture, and oil content present in wheat could be different for different wheat classes and each wheat class at different moisture levels. Wang et al. (1999), Delwiche and Massie (1996), and Murray and Williams (1987) reported that NIR absorptions at 960 and 1420 nm wavelengths were related to water content; and 1470, 1480, and 1500 nm to protein content; 1200, 1230, 1310, 1360, 1610 and 1700 nm to carbohydrate content; 960, 1060, 1330, 1390, 1480, and 1680 nm to kernel hardness; and 1390 nm to oil content of wheat. The results of STEPDISC showed that some of the above mentioned wavelengths contributed more in identifying wheat classes at different moisture levels.

Delwiche and Graybosch (2002) reported classification accuracies of 42 – 71% in identifying waxy wheat using a linear discriminant function with 1 – 10 principal component scores as input. They also reported classification accuracies of 46 – 71% in identifying waxy wheat using a quadratic discriminant function with 1 – 10 principal component scores as input. Panigrahi et al. (2006) found that meat samples stored at both 4 and 10°C had classification accuracies of > 93% using a quadratic discriminant analysis with a bootstrapping validation method. In another study, a maximum classification accuracy of 86.6% was achieved in classifying barley based on the ergosterol levels using linear and quadratic discriminant functions with a leave-one-out cross validation method (Balasubramanian et al. 2006). Mean classification accuracies of 89.1 and 99.1% were reported in classifying seven cereal grains using linear parametric classifier with the top two and five reflectance features in the visible region (Mohan et al. 2005). In the present study, classification accuracies were 81 – 100% and 60 – 89% in identifying wheat classes at different moisture levels using LDA and QDA models, respectively. Near-infrared hyperspectral imaging technique has the potential to be used to identify wheat classes at different moisture levels.

CONCLUSIONS

Classification accuracies were 80 – 100% and 60 – 89% for LDA and QDA models, respectively, in identifying five major western Canadian wheat classes at different moisture levels. Near-infrared hyperspectral imaging was found useful to identify different moisture level wheat classes with the extracted relative NIR reflectance intensities as input for developing statistical classification models. This current technique could be used to identify classes of wheat and their moisture levels and develop an automatic grain quality assessment tool. The major limitation of this technique is the production of a large volume of information which requires appropriate data processing techniques to interpret the results accurately. Proper calibration methods are also needed to remove the inherent and external noises of the system during imaging. In the future, wheat samples from different crop years and locations could be included in the sample space to improve robustness and classification efficiency of the models.

ACKNOWLEDGEMENTS

We thank the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chairs program, and the Canada Foundation for Innovation (CFI) for funding this study.

REFERENCES

Armstrong, P.R. 2006. Rapid single-kernel NIR measurement of grain and oil-seed attributes. *Applied Engineering in Agriculture* 22: 767-772.

Balasubramanian, S., S. Panigrahi, B. Kottapalli and C.E. Wolf-Hall. 2007. Evaluation of an artificial olfactory system for grain quality discrimination. *LWT-Food Science and Technology* 40: 1815-1825.

Burger, J. and P. Geladi. 2006. Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples. *Analyst* 131: 1152-1160.

CGC. 2008. Variety identification monitoring. http://grainscanada.gc.ca/grl/variety_id/vid_research-e.htm (2008/05/15).

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: 311-320.

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. and R.A. Graybosch. 2002. Identification of waxy wheat by near-infrared reflectance spectroscopy. *Journal of Cereal Science* 35: 29-38.

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

FAOSTAT. 2008. Food and Agriculture Organizations of the United Nations. http://faostat.fao.org (2008/05/15).

Martens, H. and T. Naes. 1992. Multivariate calibration. Chichester, UK: Wiley.

McNeill, S. and D. Overhults. 2008. A comprehensive guide to wheat management in Kentucky. http://www.ca.uky.edu/agc/pubs/id/id125/10.htm (2008/05/15).

Mohan, L.A., C. Karunakaran, D.S. Jayas and N.D.G. White. 2005. Classification of bulk cereals using visible and NIR reflectance characteristics. *Canadian Biosystems Engineering* 47: 7.7-7.14.

Murray, I. and P.C. Williams. 1987. Chemical principles of near-infrared technology. In *Near-infrared Technology in the Agricultural and Food Industries,* eds. P.C. Williams and K.H. Norris, 17-34. St. Paul, MN: American Association of Cereal Chemists Inc.

Nagata, M., J.G. Tallada, T. Kobayashi and H. Toyoda. 2005. NIR hyperspectral imaging for measurement of internal quality in strawberries. ASAE Paper No. 053131. St. Joseph, MI: ASABE.

Nicolai, B.M., E. Lotze, A. Peirs, N. Scheerlinck and K.I. Theron. 2006. Non-destructive measurement of bitter pit in apple fruit using NIR hyperspectral imaging. *Postharvest Biology and Technology* 40: 1-6.

Paliwal, J., N.S. Sashidhar and D.S. Jayas. 1999. Grain kernel identification using kernel signature. *Transactions of the ASAE* 42: 1921-1924.

Paliwal, J., W. Wang, S.J. Symons and C. Karunakaran. 2004. Insect species and infestation level determination in stored wheat using near-infrared spectroscopy. *Canadian Biosystems Engineering* 46: 7.17-7.24.

Panigrahi, S., S. Balasubramanian, H. Gu, C.M. Logue and H. Marchello. 2006. Design and development of metal oxide based electronic nose for spoilage classification of beef. *Sensors and Actuators B* 119: 2-14.

Perez-Mendoza, J., J.E. Throne, F.E. Dowell and J.E. Baker. 2003. Detection of insect fragments in wheat flour by near-infrared spectroscopy. *Journal of Stored Products Research* 39: 305-312.

Petterson, H. and L. Aberg. 2003. Near infrared spectroscopy for determination of mycotoxins in cereals. *Food Control* 14: 229-232.

Ruan, R., Y. Li, X. Lin and P. Chen. 2002. Non-destructive determination of deoxynivalenol levels in barley using near-infrared spectroscopy. *Applied Engineering in Agriculture* 18: 549-553.

Wang, D., F.E. Dowell and R.E. Lacey. 1999. Single wheat kernel color classification using neural networks. *Transactions of the ASAE* 42: 233-240.

Wang, D., F.E. Dowell and D.S. Chung. 2001. Assessment of heat-damaged wheat kernels using near-infrared spectroscopy. ASAE Paper No. 01-6006. St. Joseph, MI: ASAE.

Wang, D., M.S. Ram and F.E. Dowell. 2002. Classification of damaged soybean seeds using near-infrared spectroscopy. *Transactions of the ASAE* 45: 1943-1948.

Wang, D., F.E. Dowell, M.S. Ram and W.T. Schapaugh. 2004a. Classification of fungal-damaged soybean seeds using near-infrared spectroscopy. *International Journal of Food Properties* 7: 75-82.

Wang, W., J. Paliwal and D.S. Jayas. 2004b. Determination of moisture content of ground wheat using near-infrared spectroscopy. ASAE Paper No: MB04-200. St. Joseph, MI: ASABE.