Classification of Scab– and Other Mold–Damaged Wheat Kernels by Near–Infrared Reflectance Spectroscopy

S. R. Delwiche

Abstract. *Scab (Fusarium head blight) is a disease that causes wheat kernels to be shriveled, underweight, and difficult to mill. Scab is also a health concern because of the possible concomitant production of the mycotoxin deoxynivalenol. Current official inspection procedures entail manual human inspection. A study was undertaken to explore the possibility of detecting scab–damaged wheat kernels by a near–infrared (NIR) diode array spectrometer. Wheat kernels from three categories (sound, scab–damaged, and mold–damaged) were visually inspected by and furnished by the USDA Grain Inspection, Packers, and Stockyards Administration (GIPSA). The reflectance spectrum of each intact kernel was collected at 6–nm increments over a 940 to 1700 nm region. Exhaustive searches of the best combination of individual wavelengths, best difference of wavelengths, best ratio, and combinations thereof were performed on a set of 100 kernels from each of the three categories. The best modeling classification occurred for precisely aligned kernels using a combination of kernel mass and the difference in log(1/R) at two wavelengths, 1182 and 1242 nm. When applied to a test set, the classification model correctly identified the intended category with 95% accuracy. The most difficult to classify conditions were scab–damaged vs. mold–damaged and mold–damaged vs. sound. When scab–damaged and mold–damaged kernels were combined into one category, the overall accuracy for a two–category (sound vs. damaged) classification model was between 95% and 98%, depending on the kernel orientation scheme and inclusion of kernel mass. The achieved accuracy levels demonstrate the feasibility of using NIR reflectance spectroscopy with as few as two wavelengths to assist in wheat grading and commercial sorting.*

Keywords. *Fungi, Grading, Near–infrared, Scab, Spectroscopy, Wheat.*

Fusarium head blight, also known as scab, is a fungal disease caused by species of *Fusarium* (particularly *F. graminearum*) that occur in small grains. In the U.S., it is most problematic in wheat (hard red spring, durum, and soft red winter classes), barley, oats, and corn. In wheat, the condition is caused by humid conditions during stages of flowering or early kernel development, which causes the kernels to be shrunken, chalky white, or pink in appearance. *Fusarium* species may produce a metabolite called deoxynivalenol (DON), also known as vomitoxin, which is toxic to non–ruminant animals. The U.S. Food and Drug Administration (FDA) advisory–level specification for DON in finished wheat products (e.g., flour, semolina) destined for human consumption is 1 ppm, with higher allowable levels (5 to 10 ppm) in livestock and poultry feeds. Chemical tests for DON level concentration are typically based on HPLC or immunoassay (ELISA), for which commercial test kits are available. However, what is not available is a machine vision or spectroscopy–based system that can mimic the visual analysis performed by the grain inspector for wheat kernel damage. Although scab damage does not necessarily indicate the presence of DON, it is considered a physical defect that can result in the downgrading of wheat and barley during official inspection.

Currently, USDA Grain Inspection, Stockyards, and Packers Administration (GIPSA) inspectors examine samples for scab and other forms of damage (i.e., heat, frost, immaturity, insect, sprout, black tip fungus, mold) on wheat that is exported from the U.S. Inspectors must manually pick through a 15–g sample to separate the scab–damaged from non–scab damaged kernels. Scab, along with other forms of damage, can lower the official grade, starting with a threshold of 2% by weight, which reduces the grade from No. 1 to No. 2. Grade is progressively reduced (to No. 5) with increase in level of damaged kernels until the weight concentration reaches 15%, whereupon a lot is assigned the grade of sample grade (USDA–GIPSA, 1997). Additionally, sales contracts may specify more stringent criteria for the maximum tolerable levels of scab–damaged kernels. Inspections are labor intensive and can typically take more than 10 minutes per sample. The USDA is currently seeking new methods for grain inspection that are rapid and objectively based.

Much of the research on instrumentation for wheat kernel analysis has been based on digital image analysis (Neuman et al., 1987; Shatadal et al., 1995a, 1995b; Symons and Fulcher, 1988; Zayas et al., 1985, 1986). Most of this research was based on extraction of kernel morphological features for

Article was submitted for review in September 2002; approved for publication by the Food & Process Engineering Institute Division of ASAE in January 2003.

Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply endorsement or recommendation by the USDA.

The author is **Stephen R. Delwiche,** Agricultural Engineer, USDA–ARS, Beltsville Agricultural Research Center, Instrumentation and Sensing Laboratory, Building 303, BARC–East, Beltsville, MD 20705–2350; phone: 301–504–8450; fax: 301–504–9466; e–mail: delwiche@ba.ars.usda.gov.

the purpose of identifying non–wheat species from wheat or for classifying wheat. Sapirstein (1995) reviewed the literature on image analysis of grain through the mid 1990s. Detection of diseased kernels by image analysis is a relatively recent endeavor, with some published works on corn (Ng et al., 1998) and wheat (Ruan et al., 1998; Luo et al., 1999) available. In Ruan's study, neural networks were used on color and textural features of bulk samples to develop a model for wheat scab detection. Results indicated detection of visibly scab–damaged kernels by machine was equivalent to or better than an expert human panel.

Preliminary work by Williams (1997) demonstrated moderate success with NIR modeling of DON in bulk samples. Of the numerous studies that have used near–infrared (NIR) spectroscopy on single wheat grains, such as for protein content (Delwiche, 1995, 1998), wheat class (Delwiche and Massie, 1996; Dowell, 1997, 1998; Wang et al., 1999), insect infestation (Chambers and Ridgeway, 1996; Dowell et al., 1998), vitreousness (Dowell, 2000), and hardness (Delwiche, 1993), a very limited number have examined fungal disease. Pearson et al. (2001) were able to classify single corn kernels for low and high concentrations of aflatoxin, based on either NIR reflectance or transmittance. Similarly, Dowell et al. (2002) demonstrated that such classification modeling is also possible for detection of fumonisin in corn. The only published work on single kernel wheat scab detection by NIR appears to be that by Dowell et al. (1999), who used partial least squares regression to relate the spectra (400 to 1700 nm) to DON concentration, as measured by HPLC.

The objective of the current study was to determine the feasibility of using NIR reflectance to classify single wheat kernels according to their presence or absence of wheat scab, as identified by the federal inspector. An additional condition of damaged wheat, identified by the inspector as "mold–damaged" and often a precursor to the scab condition, was also studied in its context to the sensitivity of NIR to differentiate sound and scab–damaged wheat.

Materials and Methods

Wheat

Hard red spring wheat kernels were obtained from the USDA–GIPSA Technical Services Division (Kansas City, Mo), having originated as several commercial samples that were submitted to GIPSA for official inspection. Each kernel was visually examined by the GIPSA inspectors of the Board of Appeals and Review and categorized into one of three categories: sound, mold–damaged, and scab–damaged. This procedure is routinely performed on wheat samples in order to assign class and grade. Although mold and scab are both considered to be cases of damage for which wheat grade is detrimentally affected, the mold condition, which is often caused by *Fusarium*, is physically less severe than the scab condition. Scab, on the other hand, is exclusively attributed to *Fusarium* and is visually identified as such by the inspector when a kernel has a very noticeable chalky, shriveled appearance. A minimum of 138 kernels per category was used in single kernel analysis. Kernel mass (±0.01 mg) was measured within a week after spectral measurement.

Equipment

A Zeiss MCS511 diode array spectrometer, consisting of a thermoelectrically cooled 128–element indium gallium arsenide array (nominal wavelength range = 940 to 1700 nm), was controlled by the equipment manufacturer's software operating within the Grams/32 (Galactic Industries, Salem,

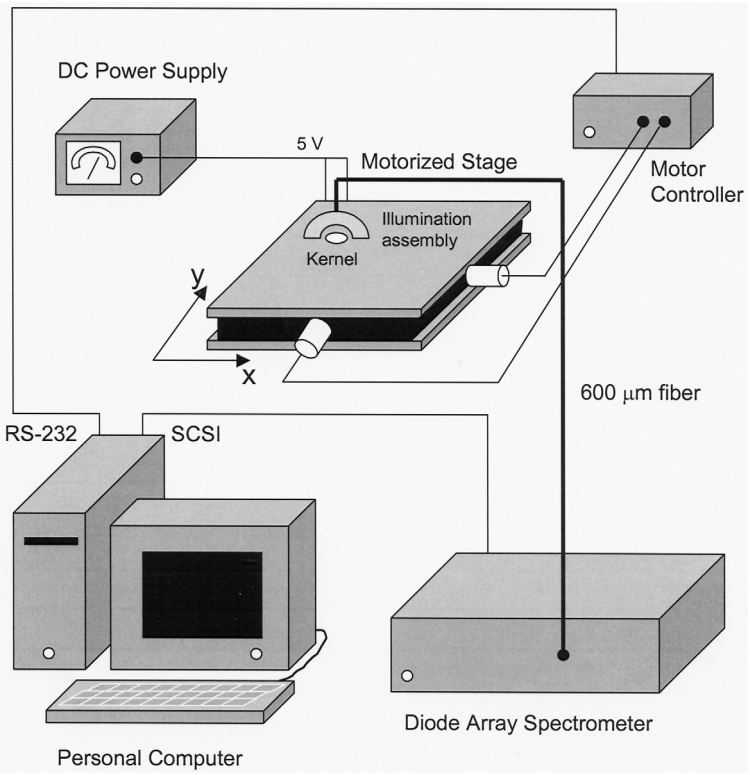

Figure 1. Schematic of single kernel optical system.

N.H.) spectral analysis software. Kernels were illuminated by an external illumination assembly (Control Development, South Bend, Ind.) that consisted of two opposing 5 V, 150 mA tungsten filament lamps with gold–coated parabolic reflectors. These lamps were oriented 45° from the horizontal at a distance of approximately 5 mm from the surface of the kernel. A 600 μm diameter fiber, oriented directly above the kernel (90° with respect to the horizontal) and positioned approximately 5 mm from the kernel surface, conveyed diffusely reflected light from the kernel to the diode array. A schematic of the system is shown in figure 1.

Procedure

The three categories were handled separately during spectral data acquisition. Within each category, the kernels were randomly drawn and placed on any of the 49 (7 rows × 7 columns) milled slots of a custom–machined Bakelite plate that was pin–mounted to a two–directional motorized stage (Velmex, Inc., Bloomfield, N.Y.). The size of each slot was sufficiently large to accommodate a wheat kernel, such that the kernel's upper surface projected approximately 1 mm above the surface of the plate. The plate, including slots, was painted flat black to reduce specular reflection. Before collecting spectra on each row of seven kernels, a reference spectrum was collected from a 7–mm thick piece of sintered polytetrafluoroethylene (Labsphere, Sutton, N.H.) of approximately 98% absolute reflectance throughout the wavelength region examined. For each kernel, the stage was positioned to align the kernel's centroid with the center of the illumination assembly. With an integration time of 6.8 ms, 32 spectra were co–added, and then averaged before the log(1/R) spectrum was stored in a computer file. To determine the sensitivity of kernel orientation on classification ability, each kernel was scanned in two configurations. For the first configuration (hereafter, the precise orientation), kernels were placed crease down with the germ end of the kernel always to one side. For the second configuration (hereafter, the random orientation), kernels were placed randomly in the milled slots, though by the geometric similarity between kernel and slot, the long axis of the kernel was always in line with that of the slot. Separate models were developed for the two types of kernel orientation. Each kernel was subsequently dried (130°C, 19 h), and its dry weight was recorded for moisture content determination.

Classification Modeling

The raw, stored spectral data, which possessed wavelength differences between adjacent points ranging from 5.85 to 6.23 nm were adjusted by linear interpolation to produce a uniform 6.00–nm wavelength spacing. Because of the relatively poor quality of the readings below 1000 nm, only points in the 1002 to 1704 nm region (118 points total) were used in the modeling procedures. The first 100 samples within each category were used as the cross–validation set. The remaining samples (n = 38, 43, and 96 for scab, mold, and sound categories) became the test set. Modeling procedures fell within two general types: one–basis and three–basis. These are separately described below.

For the one–basis modeling, linear discriminant analysis using the SAS (SAS Institute Inc., Cary, N.C.) procedure DISCRIM was applied with one–sample–out cross–validation. It was assumed that the independent variables of each model were multivariate normally distributed in each category (scab, mold, and sound) and that the covariance matrices of the categories were identical, thus allowing for the pooling of the covariance matrices across categories (Næs et al., 2002). With the application of Bayes' theorem, a sample was assigned to the category whose center in n–dimensional space was closest to the sample, which is analogous to selecting the category with the corresponding highest posterior probability for that sample. To maintain model simplicity, most models consisted of one or two independent variables; however, as many as four wavelengths were used to form an independent variable. Using A_x to represent log(1/R) at wavelength x, the following model structures were examined during cross–validation:

A_1, A_2

A_1, A_2, A_3, A_4

A_1/A_2

$A_1 - A_2$

$(A_1 - A_2)/A_3$

$A_1/(A_2 - A_3)$

$(A_1 - A_2)/(A_3 - A_4)$.

Before each cross–validation trial, the SAS procedure STEPDISC was used to perform an exhaustive search of all combinations of the (118) wavelengths to determine the optimal simple difference ($A_1 - A_2$) and simple quotient (A_1/A_2) structures (6,903 possible combinations). The $A_1/(A_2 - A_3)$ structure was formed by fixing A_1 at the single best wavelength determined from the A_1, A_2 model and then allowing the denominator difference to vary in search of an optimum structure (6,903 searches). Likewise, the $(A_1 - A_2)/A_3$ structure was formed by fixing the numerator at the best difference [from the $(A_1 - A_2)$ structure] and then allowing for variation of the A_3 denominator value (128 searches). For the ratio of differences structure, $[(A_1 - A_2)/(A_3 - A_4)]$, wavelengths at 11 locations (1002, 1104, 1200, 1302, 1374, 1386, 1398, 1416, 1470, 1566, and 1668 nm) plus a dummy wavelength for which $A_2 = 0$ or $A_4 = 0$ (to permit single–value numerators or denominators) were used in an exhaustive search (4,290 possible combinations). The 11 locations were selected by visually examining the log(1/R) spectra and selecting the points that were at or near a local maximum or minimum, a baseline region (1080 to 1120 nm), or an extended sloped region (1350 to 1450 nm). Each best combination was subsequently subjected to cross–validation. Additionally, kernel mass by itself and in combination with the best $(A_1 - A_2)$ was examined. Lastly, a linear discriminant analysis model utilizing the scores from the best four principal components (as determined by STEPDISC) was examined. Because the principal components were selected based on their collective ability to differentiate the samples into the three classes, these components were not necessarily in sequential order of their contribution to their explanation of overall spectral variance (e.g., the four PC model was not limited to PCs 1 through 4).

Three–basis modeling was performed using a SIMCA (soft independent modeling of class analogy) approach. A principal component analysis of each category's spectra was used to reduce the spectral dimension to 2, 4, or 6 components, whereupon the category's model was used to calculate the distance, in units of Mahalanobis distance, between an

"unknown" sample and the model center. For the circumstance in which the sample was known to belong to the same category as the category whose model was under development, the distance was evaluated as that between the sample and the center of the model developed from the $n - 1$ other samples in the category. For the other circumstance (that is, when a sample did not belong to the category whose model was being used), the distance between model center and sample was evaluated in the same manner as when the sample was from the test set or was a true unknown. SIMCA analysis was performed using commercial software (PLSPlus, Galactic Industries, Salem, N.H.).

Results and Discussion

Kernels infected with mold or scab were generally lower in mass than sound kernels, as shown by cross–validation set averages of 15.0, 25.9, and 36.6 mg for the scab, mold, and sound kernels, respectively. Aside from the potential to produce DON, *Fusarium* infection causes reduction in test weight and specifically kernel mass by its colonization of the seed (Cunfer, 1987). A model that used mass alone produced a cross–validation accuracy of 86.0% and a test set accuracy of 82.6% (table 1). The log(1/R) spectra of the single kernels from each of the three categories were similar in overall shape, as shown by the 100–kernel averages for each category (fig. 2). The most noticeable difference among categories is the overall level of spectral absorption, with sound kernels possessing higher log(1/R) values, on average, throughout the entire wavelength region. However, the within–category spectral variation, as depicted by the $\pm 1\sigma$ envelope surrounding each average spectrum in figure 2, was sufficiently large to cause spectral overlap between categories, thus precluding the use of a single–wavelength classification model. The exhaustive search of the best two wavelengths for the condition of precise kernel orientation resulted in the selection (in order of importance) of 1470 and 1332 nm, which yielded a cross–validation accuracy of 83.7% and a test set accuracy of 88.9%. The first wavelength corresponds to the region of greatest absorption, while the second wavelength is at the beginning of a monotonically increasing broad region that culminates near the first wavelength. Greater spectral absorption of sound kernels throughout the entire wavelength region is attributed to a combination of greater kernel mass (hence, greater optical density) and a higher moisture content (11.4%, 9.9%, and 10.2%, dry basis averages, for the sound, mold, and scab cross–validation set categories, respectively). Inclusion of two additional wavelengths (1320 and 1404 nm) resulted in the improvement of the cross–validation accuracy (89.0%); however, this was to the detriment of the general model, as seen by a decline in test set accuracy to 86.3%.

Of the other fixed wavelength models examined (table 1), the best classification accuracies were obtained with a model that was the difference in two log(1/R) values at 1182 and 1242 nm [$\log(1/R)_{\lambda = 1182\ nm} - \log(1/R)_{\lambda = 1242\ nm}$, hereafter designated as $(A_{1182} - A_{1242})$], which is mathematically equivalent to minus the logarithm of the ratio of the detected energies from these respective wavelengths. In this case, the cross–validation and test set accuracies were 91.3% and 92.7%, respectively. Spectral explanation for the selection of these two wavelengths is difficult because they lie on opposite sides of a broad absorption band that has maximum absorption near 1200 nm. Generally, this difference is smallest for scab–damaged kernels, intermediate for mold–damaged kernels, and greatest for sound kernels (fig. 3). When this wavelength difference was combined with kernel mass, model accuracy improved even further (95.7% and 95.0% for cross–validation and test set accuracies, respectively). This combination produced the highest accuracy of all models examined. A breakdown of the predictions by category for this model [$(A_{1182} - A_{1242})$, kernel mass] is shown in table 2. From this table, it is seen that misclassifications were usually between the scab and mold categories.

Table 1. Classification accuracies of linear discriminant analyses.[a]

	Precise Orientation			Random Orientation		
Model[b]	$\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (nm)	Cross Validation[c] (%)	Test[d] (%)	$\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (nm)	Cross Validation (%)	Test (%)
Mass	—	86.0	82.6	—	86.0	82.6
A_1, A_2	1470, 1332	83.7	88.9	1470, 1290	84.3	81.3
A_1, A_2, A_3, A_4	1470, 1332, 1320, 1404	89.0	86.3	1470, 1290, 1338, 1194	91.3	65.4
A_1/A_2	1314, 1296	83.7	80.7	1314, 1290	83.3	64.8
$A_1 - A_2$	1182, 1242	91.3	92.7	1224, 1218	84.3	88.7
$(A_1 - A_2)$, mass	1182, 1242	95.7	95.0	1224, 1218	92.3	88.9
$(A_1 - A_2)/A_3$	1182, 1242, 1002	68.3	75.7	1224, 1218, 1212	67.0	72.3
$A_1/(A_2 - A_3)$	1470, 1386, 1374	81.0	83.8	1470, 1680, 1566	75.0	74.7
$(A_1 - A_2)/(A_3 - A_4)$	1398, 1374, 1416, 1302	76.7	72.5	1398, 1386, 1470, 1302	76.3	71.2
PCA	PCs 1, 2, 7, 3	89.3	86.4	PCs 1, 2, 3, 6	84.7	86.4
SIMCA PCA	2–factor	85.3	85.8	2–factor	83.3	83.7
SIMCA PCA	4–factor	86.3	85.6	4–factor	77.3	74.8
SIMCA PCA	6–factor	86.7	83.6	6–factor	82.0	73.4

[a] Each value is the average of the percentages of correctly classified kernels belonging to three categories: scab–damaged, mold–damaged, and sound.

[b] Models consisting of more than one term have terms separated by a comma. Some single–term models actually consist of more than one wavelength, as noted by the mathematical expression (difference, ratio, or combination of both) displayed. A_x refers to log(1/R) at a wavelength specified in the adjoining column; PCA = principal component analysis; and SIMCA PCA = soft independent modeling of class analogy applied to the principal components scores of each category.

[c] Cross–validation set consisted of 100 kernels from each of the categories: sound, mold–damaged, and scab–damaged.

[d] Test set consisted of 96 sound, 43 mold–damaged, and 38 scab–damaged kernels.

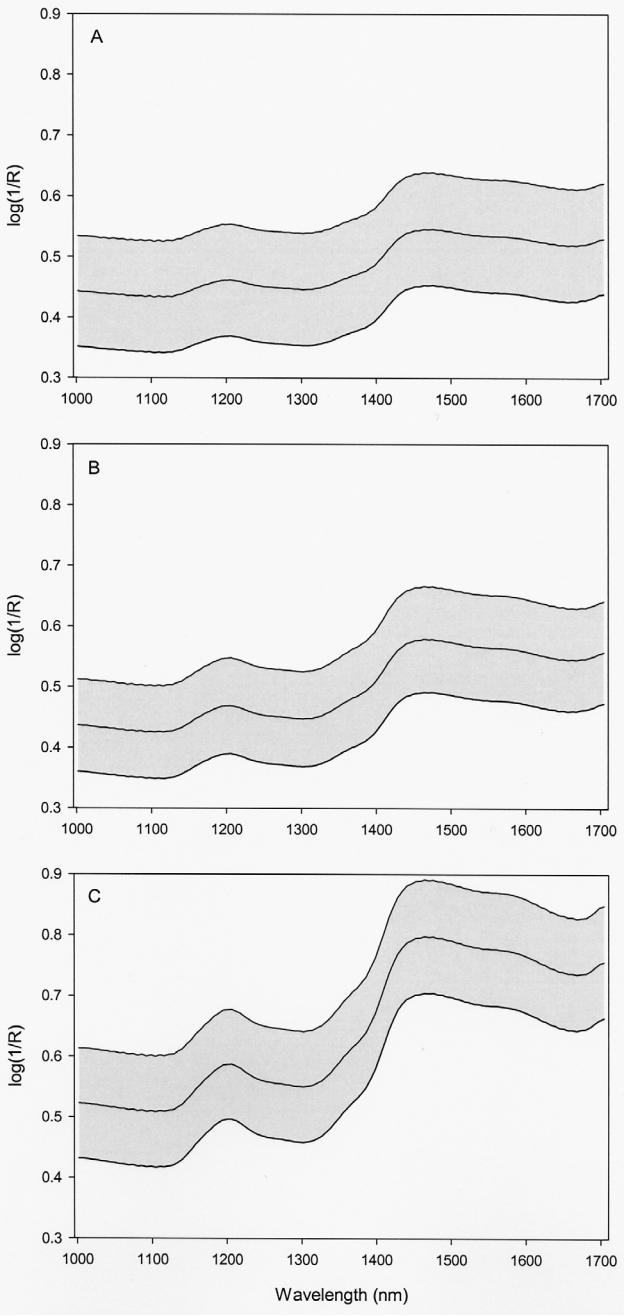

Figure 2. Mean spectrum ± one standard deviation envelope for each category of cross–validation samples (n = 100 spectra per category): A = scab, B = mold, C = sound.

Misclassified sound kernels were assigned to the mold category, but never to the scab category. Thus, the behavior of the classification model was in keeping with the human inspector's judgment — that mold–damaged kernels often arise from *Fusarium* infection that is not so severe as to be labeled as scab–damaged. A plot of ($A_{1182} - A_{1242}$) vs. kernel mass of the cross–validation samples is shown in figure 4, with kernels that were erroneously classified during the cross–validation procedure indicated by filled–in symbols. Misclassified kernels were congregated at two places in the graph, with each place lying in an interface between two established categories: scab–mold or mold–sound.

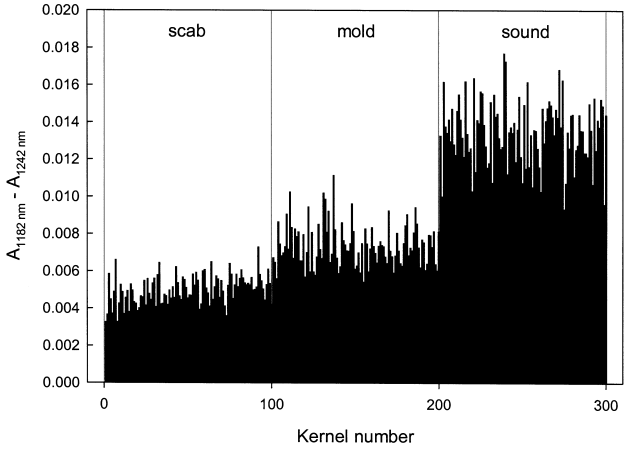

Figure 3. Two–wavelength spectral difference [$\log(1/R)_{\lambda = 1182\ nm} - \log(1/R)_{\lambda = 1242\ nm}$] of the precisely oriented cross–validation samples (n = 100 kernels per category, contiguously arranged).

Table 2. Classification results for samples in cross–validation and test sets precise kernel orientation. Model: $A_{1182} - A_{1242}$, kernel mass.

Actual Category	Number of Kernels Assigned to Category[a]			
	Scab	Mold	Sound	Total
Cross–validation				
Scab	**97**	3	0	100
Mold	5	**94**	1	100
Sound	0	4	**96**	100
Total	102	101	97	300
Test				
Scab	**37**	1	0	38
Mold	4	**39**	0	43
Sound	0	3	**93**	96
Total	41	43	93	177

[a] Diagonal values (in **bold**) represent correct assignments.

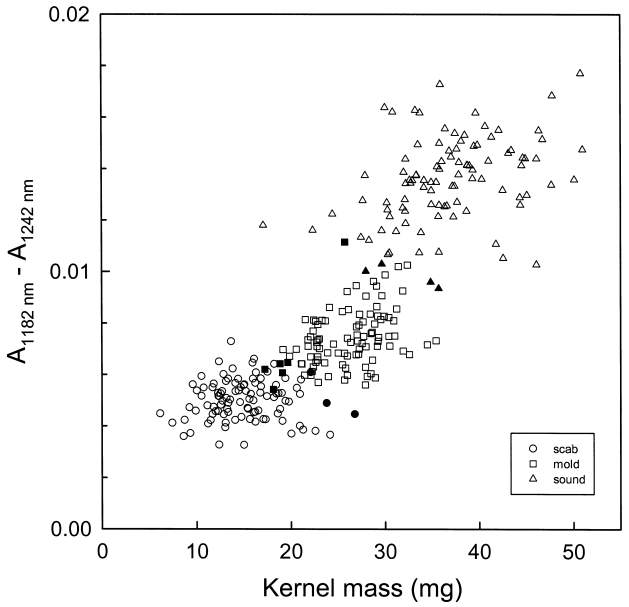

Figure 4. Plot of two–wavelength spectral difference [$\log(1/R)_{\lambda = 1182\ nm} - \log(1/R)_{\lambda = 1242\ nm}$] versus kernel mass of the precisely oriented cross–validation samples (n = 100 kernels per category: scab, mold, sound). Kernels that were misclassified (by the model that used these terms) during cross–validation are indicated as filled–in symbols. True category is indicated by symbol shape. Misclassification never occurred between scab and sound kernels.

Wavelengths: Their Ratios or Differences vs. Principal Components

A four–factor principal component model, in which the scores from PCs 1, 2, 7, and 3 (as determined by the stepwise procedure in SAS to be most useful in distinguishing the categories) were used in determining the linear discriminant function, achieved a cross–validation accuracy of 89.3% and a test set accuracy of 86.4% (table 1). These values are actually lower than the accuracies of the best wavelength difference model, ($A_{1182} - A_{1242}$). Thus, reasonably high classification accuracy does not require the collection of reflectance readings from a broad wavelength region (1000 to 1700 nm) followed by spectral dimension reduction through PCA. Instead, a less expensive instrument consisting of two interference filters could ostensibly be used to classify scab–damaged, mold–damaged, and sound wheat kernels.

SIMCA vs. One–Basis PCA

Accuracies of the 2–, 4–, and 6–factor SIMCA PCA models ranged from 85.3% to 86.7% for the cross–validation set, and from 83.6% to 85.8% for the test set (table 1). Interestingly, the 6–factor model, while demonstrating the greatest of the SIMCA PCA cross–validation set classification accuracies, had the lowest test set accuracy. Thus, this tendency for overfitting as the number of factors increases suggests that the 2–factor model was optimal for a SIMCA modeling approach; however, it was still not as accurate as the simpler 2–factor, one–basis PCA linear discriminant function model.

Effect of Kernel Orientation

Classification accuracies for all model structures that utilized spectra of precisely oriented kernels are also shown for randomly oriented kernels in table 1. Generally, classification accuracy declined for randomly oriented kernels, as seen by the fact that test set accuracies of random orientation models were always less than or equal to their corresponding values of the precise orientation models. Some of the wavelengths identified during the stepwise selection procedures for the randomly oriented kernel spectra were the same as those for the precisely oriented kernel spectra. For example, the region of highest absorption (1470 nm) was chosen for inclusion in the best two–wavelength and four–wavelength models, regardless of kernel orientation. Similarly, three principal components (PCs 1, 2, and 3) of the four–component PCA models were common between orientation types, with the models yielding the same test set overall accuracy (86.4%). As with the precise orientation models, the best random orientation model, with a test set accuracy of 88.9%, was that formed from the difference of two wavelength absorptions ($A_{1218} - A_{1224}$) combined with kernel mass, although the model that used the same absorption difference, but without kernel mass, had nearly the same accuracy (88.7%). Unlike the precise orientation model, the wavelengths chosen to form this difference accentuate the slope of a broad band, whose peak is near 1200 nm. A breakdown of the best random orientation model by category is shown in table 3. Similar to the results for precise orientation, classification error rarely arose from confusion between scab and sound kernels. Rather, misclassifications were typically between scab and mold categories or between mold and sound categories.

Two–Category Classification

Although official inspections result in the separation of mold–damage and scab–damage, these categories are eventually combined when wheat grade is determined. Further, commercial processors are likely to be concerned with retaining the sound kernels, irrespective of the nature of damage of the non–healthy kernels. Therefore, in a short series of modeling trials (patterned after the procedures described earlier) whose results are described in table 4, the mold– and scab–damaged categories were combined into one category designated as "damaged." Using two of the most effective model structures from the three–category modeling, that is ($A_1 - A_2$), with and without kernel mass, cross–validation accuracy averages ranged between 96.0% and 98.0%, depending on orientation scheme and whether kernel mass was included in the model. Test set accuracies, ranging from 94.6% to 98.4%, were also very high. Interestingly, for the random orientation spectra, the same

Table 3. Classification results for samples in cross–validation and test sets random kernel orientation. Model: $A_{1224} - A_{1218}$, kernel mass.

Actual Category	Number of Kernels Assigned to Category[a]			
	Scab	Mold	Sound	Total
Cross–validation				
Scab	**92**	8	0	100
Mold	7	**92**	1	100
Sound	0	7	**93**	100
Total	99	107	94	300
Test				
Scab	**36**	2	0	38
Mold	8	**35**	0	43
Sound	0	9	**87**	96
Total	44	46	87	177

[a] Diagonal values (in **bold**) represent correct assignments.

Table 4. Classification accuracies of linear discriminant analyses two–category (sound vs. damaged) models. [a]

	Precise Orientation			Random Orientation		
Model[b]	λ_1, λ_2 (nm)	Cross Validation[c] (%)	Test[d] (%)	λ_1, λ_2 (nm)	Cross Validation (%)	Test (%)
Mass	—	89.0	90.5	—	89.0	90.5
$A_1 - A_2$	1698, 1362	97.5	98.4	1224, 1218	96.0	94.6
$(A_1 - A_2)$, mass	1698, 1362	98.0	98.4	1224, 1218	96.2	96.8

[a] Each value is the average of the percentages of correctly classified kernels belonging to two categories: damaged and sound.
[b] Model wavelength structure is explained in footnote [b] of table 1.
[c] Cross–validation set consisted of the same kernels from the three–class cross–validation set. The original scab–damaged and mold–damaged categories were combined to form the damaged category of the present table (100 sound kernels, 200 damaged kernels).
[d] Test set was formed in the same manner as the cross–validation set (96 sound kernels, 81 damaged kernels).

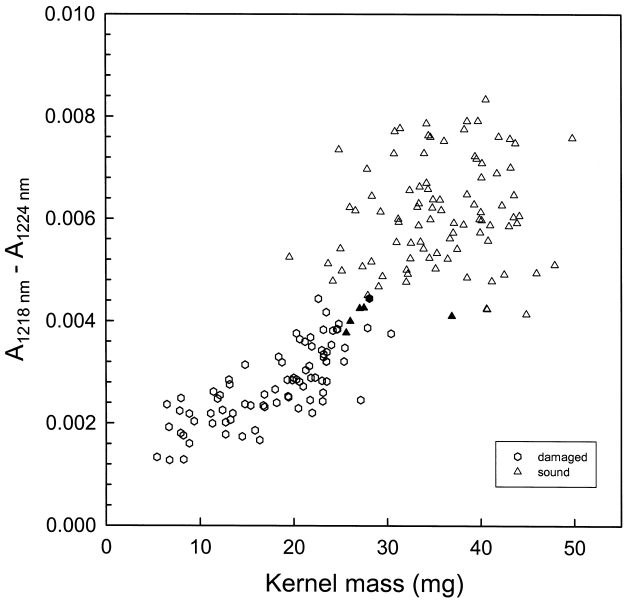

Figure 5. Plot of two–wavelength spectral difference [$\log(1/R)_{\lambda = 1218\ nm} - \log(1/R)_{\lambda = 1224\ nm}$] versus kernel mass of the randomly oriented test set samples (n = 96 sound kernels, 81 damaged kernels). Kernels that were misclassified (by the two–category model that used these terms) are indicated as filled–in symbols. True category is indicated by symbol shape.

wavelength difference ($A_{1224} - A_{1218}$) that was determined as optimal in three–category modeling was also selected in two–category modeling. The application of a model using this difference and mass is shown in figure 5, in which one damaged and five sound kernels were incorrectly classified, yielding an average accuracy of 96.8%. Although the accuracy improved slightly (to 98.4%) with precise orientation, the high level of accuracy with random orientation, which incidentally was 94.6% when kernel mass was excluded, suggests the feasibility of a high output commercial sorting instrument.

Conclusions

The use of NIR reflectance for the sorting of individual wheat kernels according to their state of health (scab–diseased, mold–diseased, and sound) appears feasible, with classification accuracies as high as 95%. Simple wavelength difference models, ($A_{\lambda 1} - A_{\lambda 2}$), combined with kernel mass, were found to be very effective at classification. Thus, the detection of mold or scab damage on wheat kernels by NIR reflectance can be of benefit to not only the wheat inspector, but also to the milling industry where the use of commercial sorters (e.g., ScanMaster II, Satake USA, Houston, Texas) can be used to cull individual wheat kernels that could otherwise contaminate grain lots at levels of DON exceeding FDA advisory levels, although not at such a high concentration as to be detected by bulk sample NIR analysis. Classification accuracy is enhanced with the consistent orientation of the kernel; however, even with random kernel orientation, accuracy was nearly 90%. The use of full wavelength range, mathematically condensed (by principal component analysis) spectral data did not improve classification accuracy. Further, a three–basis principal component decomposition and classification model, such as that obtained by implementation of a SIMCA design, did not offer any advantages over a one–basis model. Combining the mold–damaged and scab–damaged kernels into one category and developing a two–wavelength, with or without kernel mass, two–category (sound vs. damaged) model resulted in accuracies of 95% and higher for conditions similar to commercial sorting.

Acknowledgments

Eurvin Williams of USDA–GIPSA (Kansas City, Mo.) is thanked for kindly providing samples and reference inspection services. Joyce Shaffer (USDA–ARS, Beltsville, Md.) is thanked for collecting spectral and weight measurements.

References

Chambers, J., and C. Ridgway. 1996. Rapid detection of contaminants in cereals. In *Near Infrared Spectroscopy: The Future Waves*, 484–489. A. M. C. Davies and P. C. Williams, eds. Chichester, U.K: NIR Publications.

Cunfer, B. M. 1987. Bacterial and fungal blights of the foliage and heads of wheat. In *Wheat and Wheat Improvement*, 528–541. 2nd ed. E. G. Heyne, ed. Madison, Wisc.: American Society of Agronomy.

Delwiche, S. R. 1993. Measurement of single–kernel wheat hardness using near–infrared transmittance. *Trans. ASAE* 36(5): 1431–1437.

_____. 1995. Single wheat kernel analysis by near–infrared transmittance: Protein content. *Cereal Chem.* 72(1): 11–16.

_____. 1998. Protein content of single kernels of wheat by near–infrared reflectance spectroscopy. *J. Cereal Sci.* 27(3): 241–254.

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

Dowell, F. E. 1997. Effect of NaOH on visible wavelength spectra of single wheat kernels and color classification efficiency. *Cereal Chem.* 74(5): 617–620.

_____. 1998. Automated color classification of single wheat kernels using visible and near–infrared reflectance. *Cereal Chem.* 75(1): 142–144.

_____. 2000. Differentiating vitreous and nonvitreous durum wheat kernels by using near–infrared spectroscopy. *Cereal Chem.* 77(2): 155–158.

Dowell, F. E., J. E. Thorne, and J. E. Baker. 1998. Automated nondestructive detection of internal insect infestation of wheat kernels by using near–infrared reflectance spectroscopy. *J. Econ. Entomology* 91(4): 899–904.

Dowell, F. E., M. S. Ram, and L. M. Seitz. 1999. Predicting scab, vomitoxin, and ergosterol in single wheat kernels using near–infrared spectroscopy. *Cereal Chem.* 76(4): 573–576.

Dowell, F. E., T. C. Pearson, E. B. Maghirang, F. Xie, and D. T. Wicklow. 2002. Reflectance and transmittance spectroscopy applied to detecting fumonisin in single corn kernels infected with *Fusarium verticillioides*. *Cereal Chem.* 79(2): 222–226.

Luo, X., D. S. Jayas, and S. J. Symons. 1999. Identification of damaged kernels in wheat using a colour machine vision system. *J. Cereal Sci.* 30(1): 49–59.

Næs, T., T. Isaksson, T. Fearn, and T. Davies. 2002. *A User–Friendly Guide to Multivariate Calibration and Classification*, 221–259. Chichester, U.K.: NIR Publications.

Neuman, M., H. D. Sapirstein, E. Shwedyk, and W. Bushuk. 1987. Discrimination of wheat class and variety by digital image analysis of whole grain samples. *J. Cereal Sci.* 6: 125–132.

Ng, H. F., W. F. Wilcke, R. V. Morey, and J. P. Lang. 1998. Machine vision evaluation of corn kernel and mechanical and mold damage. *Trans. ASAE* 41(2): 415–420.

Pearson, T. C., D. T. Wicklow, E. B. Maghirang, F. Xie, and F. E. Dowell. 2001. Detecting aflatoxin in single corn kernels by transmittance and reflectance spectroscopy. *Trans. ASAE* 44(5): 1247–1254.

Ruan, R., S. Ning, A. Song, A. Ning, R. Jones, and P. Chen. 1998. Estimation of *Fusarium* scab in wheat using machine vision and a neural network. *Cereal Chem.* 75(4): 455–459.

Sapirstein, H. D. 1995. Variety identification by digital image analysis. In *Identification of Food–Grain Varieties*,91–130. C. W. Wrigley, ed. St. Paul, Minn.: American Association of Cereal Chemists.

Shatadal, P., D. S. Jayas, and N. R. Bulley. 1995a. Digital image analysis for software separation and classification of touching grains: I. Disconnect algorithm. *Trans. ASAE* 38(2): 635–643.

_____. 1995b. Digital image analysis for software separation and classification of touching grains: II. Classification. *Trans. ASAE* 38(2): 645–649.

Symons, S. J., and R. G. Fulcher. 1988. Determination of wheat kernel morphological variation by digital image analysis: I. Variation in Eastern Canadian milling wheats. *J. Cereal Sci.* 8: 211–218.

USDA–GIPSA. 1997. Chapter 13: Wheat. In *Grain Inspection Handbook, Book II*. Release date 1 June 1997. Washington, D.C.: USDA Grain Inspection, Packers, and Stockyards Administration.

Wang, D., F. Dowell, and R. Lacey. 1999. Single wheat kernel color classification by using near–infrared reflectance spectra. *Cereal Chem.* 76(1): 30–33.

Williams, P. C. 1997. Recent advances in near–infrared applications for the agriculture and food industries. In *Proc. International Wheat Quality Conference*, 109–128. J. L. Steele and O. K. Chung, eds. Manhattan, Kansas: Grain Industry Alliance.

Zayas, I., Y. Pomeranz, and F. S. Lai. 1985. Discrimination between Arthur and Arkan wheats by image analysis. *Cereal Chem.* 62(6): 478–480.

Zayas, I., F. S. Lai, and Y. Pomeranz. 1986. Discrimination between wheat classes and varieties by image analysis. *Cereal Chem.* 63(1): 52–56.