

DETECTION OF FUNGUS-INFECTED CORN KERNELS USING NEAR-INFRARED REFLECTANCE SPECTROSCOPY AND COLOR IMAGING

J. G. Tallada, D. T. Wicklow, T. C. Pearson, P. R. Armstrong

ABSTRACT. Contamination of grain products by fungus can lead to economic losses and is deleterious to human and livestock health. Detection and quantification of fungus-infected corn kernels would be advantageous for producers and breeders in evaluating quality and in selecting hybrids with resistance to infection. This study evaluated the performance of single-kernel near-infrared reflectance spectroscopy (NIRS) and color imaging to discriminate corn kernels infected by eight fungus species at different levels of infection. Discrimination was done according to the level of infection and the mold species. NIR spectra (904 to 1685 nm) and color images were used to develop linear and nonlinear prediction models using linear discriminant analysis (LDA) and multi-layer perceptron (MLP) neural networks. NIRS was able to accurately detect 98% of the uninfected control kernels, compared to about 89% for the color imaging. Results for detecting all levels of infection using NIR were 89% and 79% for the uninfected control and infected kernels, respectively; color imaging was able to discriminate 75% of both the control and infected kernels. In general, there was better discrimination for control kernels than for infected kernels, and certain mold species had better classification accuracy than others when using NIR. The vision system was not able to classify mold species well. The use of principal component analysis on image data did not improve the classification results, while LDA performed almost as well as MLP models. LDA and mean centering NIR spectra gave better classification models. Compared to the results of NIR spectrometry, the classification accuracy of the color imaging system was less attractive, although the instrument has a lower cost and a higher throughput.

Keywords. Classification, Discriminant analysis, Maize, Multi-layer perceptron, NIR spectroscopy.

Extensive fungal contamination of grain products can lead to substantial economic losses for farmers, traders, grain handlers, and millers. The quality grade has to be reduced when a significant number of mold-damaged kernels persist in the bulk of the grain; consequently, the grain is penalized with a lower price, and its target end-use becomes limited. More importantly, the presence of fungal infection in food or grain is a major health and safety issue because of the potential debilitating diseases in humans and livestock associated with ingesting high levels of mycotoxins. In the case of corn, aflatoxin and fumonisin are the most common and most toxic mycotoxin compounds produced by fungi (Bruns, 2003). Certain *Aspergillus* species can invade host plants during their growth and may lead to accumulation of aflatoxins in the grains, while fumonisins

are produced by certain *Fusarium* species. Exposure to aflatoxin through the diet can lead to liver cancer (Williams et al., 2004), and fumonisins are a potential risk factor for neural tube and other birth defects associated with the consumption of tortillas during the first trimester (Marasas et al., 2004). Another common fungus, *Stenocarpella maydis* (syn. *Diplodia maydis*), causes a neuromuscular disease called diplodiosis, a common nervous disorder of cattle and sheep grazing on infected maize crop residue in southern Africa (Prozesky et al., 1994) and more recently in Argentina (Odriozola et al., 2005).

Numerous factors such as seasonal, biotic, and abiotic constraints, together with the kind of cultural practices adopted for different hybrids, can synergistically influence the incidence of mycotoxins in corn. Changes in management practices are often found to be effective in controlling contamination through timely planting, better irrigation water control, pest and disease control, and improved plant nutrition, especially at certain specific stages of crop growth. However, the most highly recommended approach for reducing the levels of mycotoxins in food products is to develop better hybrids that are resistant to fungi (Betran and Isakeit, 2004; Menkir et al., 2006).

Screening methods have been developed to evaluate promising breeding lines of corn for fungal resistance, including a kernel-screening assay (Brown et al., 1999), enzyme-linked immunosorbent assay (ELISA) analysis of aflatoxin and fumonisin (Abbas et al., 2006), and the ergosterol test (Dowell et al., 1999). These methods require considerable sample preparation and chemical processing, which often take a considerable amount of time and resources. Optical methods have been explored to significantly

Submitted for review in April 2010 as manuscript number IET 8515; approved for publication by the Information & Electrical Technologies Division of ASABE in May 2011.

Mention of a trademarked or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

The authors are **Jasper G. Tallada**, Associate Professor, Department of Electrical Engineering, University of the Philippines, Los Baños, Philippines; **Donald T. Wicklow**, Microbiologist, USDA-ARS National Center for Agricultural Utilization Research, Peoria Illinois; **Tom C. Pearson**, ASABE Member, Research Engineer, and **Paul R. Armstrong**, ASABE Member, Research Engineer, USDA-ARS Center for Grain and Animal Health Research, Engineering and Wind Erosion Research Unit, Manhattan, Kansas. **Corresponding author:** P. R. Armstrong, USDA-ARS Center for Grain and Animal Health Research, Engineering and Wind Erosion Research Unit, 1515 College Ave., Manhattan, KS 66502; phone: 785-776-2728; e-mail: paul.armstrong@ars.usda.gov.

reduce the amount of time required to detect fungus-infected kernels in soybeans, corn, and wheat and to identify the specific species causing the infection. Ng et al. (1998) used a color camera to evaluate mechanical and mold damage in corn kernels by digitally measuring the infected areas. Wang et al. (2004) investigated fungus-infected soybeans with four common species using a DA7000 spectrometer in the range of 400 to 1700 nm, and they achieved very high classification accuracies of 99% between healthy and infected kernels. Gordon et al. (1999) utilized transient near-infrared spectroscopy and were able to show the potential of the technique for on-line evaluation. Their early tests achieved 85% or 95% success rates for grains infected with *A. flavus*. Berardo et al. (2005) used NIRS in the range of 400 to 1100 nm to rapidly detect kernel rot and mycotoxins in bulk samples of maize. They were able to derive a reasonable calibration model to estimate the percentage fungal infection of *Fusarium verticillioides* and the quantity of ergosterol and fumonisin B1 in the meal. Pearson and Wicklow (2006) investigated the detection of corn kernels infected with a number of different fungal species using NIR reflectance, visible and NIR transmittance imaging, color reflectance imaging, and x-ray imaging. They achieved high discrimination accuracies using stepwise discriminant analysis, particularly for the extensively damaged kernels. A critical factor in the success of these methods is the observable changes in the color and shape of the seeds brought about by metabolic processes caused by the impregnation and growth of fungus on the kernels. While these changes lead to lower kernel densities and reduced seed hardness, the effectiveness of sorting methods to identify infected kernels largely depends on the virulence of the fungus and the extent of damage it caused (Wicklow, 1995). In addition, most studies on corn have focused on stationary presentation of the kernels. In order to achieve higher sorting rates, especially for simultaneous evaluation of many hybrid lines by the plant breeder, technologies that allow for continuous kernel movement must be developed.

Pearson et al. (2008) designed a color image-based sorter that performed well in separating red and white wheat at a moderate feed rate of about 30 kernels s⁻¹, achieving sorting accuracies of 96% and 95% for “easy” and “difficult” samples, respectively. Armstrong (2006) developed a rapid single-kernel sorting instrument that receives infrared reflectance spectra from kernels tumbling down a light tube. The instrument can measure some chemical constituents such as protein and oil contents in corn and soybeans, and sort the kernels into classes suitable for breeding work. Both of these instruments could be used to detect fungus-contaminated corn kernels, but they have not yet been tested for their ability to do so.

OBJECTIVES

The objectives of this study were to evaluate the performance of rapid single-kernel NIRS and color imaging instruments for detection and identification of fungus-infected corn kernels. Classification models used for discrimination were also compared based on linear discriminant analysis and artificial neural networks. Several spectrometric data and normalized histogram data pretreatments were also compared for achieving robust discrimination models.

Table 1. Grain sample codes, fungal cultures, and accession numbers.^[a]

Label	Fungus Species	Accession No.
ASP1	<i>Aspergillus flavus</i> ^[b]	NRRL 32354
BIP1	<i>Bipolaris zeicola</i>	NRRL 47238
DIP1	<i>Diplodia maydis</i>	NRRL 43670
FUS1	<i>Fusarium oxysporum</i>	NRRL 37597
PEN1	<i>Penicillium oxalicum</i>	NRRL 58759
PEN2	<i>Penicillium funiculosum</i>	NRRL 58760
TRI1	<i>Trichoderma harzianum</i>	NRRL 54022
Symptomless kernels from un-inoculated ears		
CON1	Irrigated plot	
CON2	Non-irrigated plot	

^[a] Grain from wound-inoculated ears of Burrus 794sRR.

^[b] Non-aflatoxin producing strain

MATERIALS AND METHODS

CORN SAMPLES

Samples for this study were obtained from a field-grown commercial corn hybrid, Burrus 794sRR, in Kilbourne, Illinois, in 2007. The corn ears were wound-inoculated during the late milk to early dough stage of kernel development with one of the fungal species and accessions as listed in table 1 according to a procedure by Wicklow (1999). The ears were allowed to field dry and were then hand-harvested, manually shelled, packed into separate plastic bags, and placed in cold storage.

A four-point grading scale according to level of severity of fungal damage was adopted according to the following criteria:

- Level 1: Asymptomatic; no visible signs of damage except for very minor discoloration.
- Level 2: Mildly infected; tiny blotches of fungal growth on kernel surfaces; slightly discolored.
- Level 3: Moderately infected; visible fungal growth on 30% to 70% of the kernel surface; heavily discolored.
- Level 4: Severely infected; fungal growth over the entire surface of severely discolored kernels.

The resulting samples for the study had kernels of varying levels of damage according to the infecting fungal species. The study assumed that each sample contained the entire possible range of observable infection severity in order to aid in kernel selection. Two samples, infected by different species of *Penicillium* and exhibiting significant differences in the severity scales, were treated independently.

Three of the samples were so severely rotted that it was not possible to identify grains exhibiting varying levels of infection and were excluded from the study. Severely diseased, shrunken, and light kernels were also removed from the sample pool as they can be easily aspirated out by grain cleaners. Two un-inoculated control samples that were obtained from irrigated and non-irrigated plots were also used.

Corn kernels belonging to levels 1 and 4 were first selected to establish the range of fungal damage. Asymptomatic kernels of level 1 were physically intact with no visible mechanical or physiological damage, and they appeared to be free of any fungal infection, although they possibly came in contact with the fungus. At the other end of the range, level 4 kernels had visible symptoms of heavy infection almost consuming the entire kernel body. Kernels for levels 2 and 3 were then selected based on this extreme observable range of infection. The classification of kernels was thus based on the levels of infection displayed by the specific fungal species present. A representative sample of graded kernels is shown in figure 1 for BIP1 (*Bipolaris zeicola*) kernels.

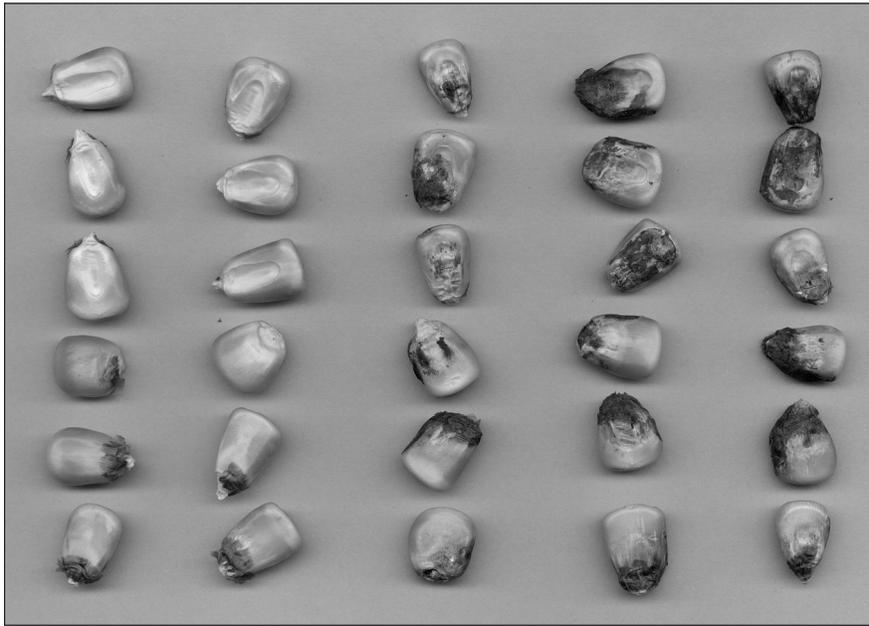


Figure 1. Representative sample of six randomly picked kernels from control 1 and graded BIP1 (*Bipolaris zeicola*) kernels. Columns, from left to right, are control 1 and levels 1 through 4. Top three rows show the germ side; bottom three show germs down.

A total of seven inoculated samples and two control samples were used. Control samples were not inoculated, and no observable mold infection was observed. Twenty-four kernels were manually selected from a sample for each infection level and fungus species and placed in separate paper packets, for a total of 96 kernels for each sample or a total of 672 infected kernels. Ninety-six kernels were also selected from each of the two control samples, and thus a grand total of 864 kernels were used for NIRS and image analysis. The samples were equilibrated for moisture at laboratory room conditions for more than seven days prior to NIRS or color imaging data collection.

SINGLE-KERNEL NIR REFLECTANCE MEASUREMENT

The NIR instrument was primarily designed to rapidly sort kernels of hybrids according to a target constituent such as protein and oil contents. The instrument consisted of a spectrometer, a light tube assembly, a control circuit, and a computer. The CDI spectrometer (NIR256L-1.7T1, Control Development, Inc., South Bend, Ind.) had a thermoelectrically cooled 256-pixel InGaAs sensor with a spectral range of 904 to 1685 nm. The light tube assembly had 48 miniature tungsten lightbulbs (Part 1150, 5 V, 0.115 A, Gilway Technical Lamp, Woburn, Mass.) arranged equidistantly in six rows along the tube periphery. A photo-electric switch was used to monitor the passage of the kernels in the upper portion of the instrument. When triggered by a moving kernel, this switch sends an electronic signal to the spectrometer to collect a scan and send the data directly through a USB interface to the computer. A Microsoft Visual C++ program processed the spectral data. A complete description of the construction and operation of the assembly is provided by Armstrong (2006). One difference between the instrument used by Armstrong (2006) and that used in this study is that a bifurcated fiber collected spectra from both ends of the tube rather than the single fiber previously used.

The instrument was allowed to warm up for at least an hour in the morning to stabilize the light and spectrometer

electronics. Background dark current and light reference reflectance spectra from a white Spectralon panel (99% diffuse reflectance) were collected at constant intervals during the spectral data collection. While a kernel moved approximately midway through the light tube, its raw diffuse reflectance spectrum was collected using a 43 ms integration time. This integration time was experimentally selected to avoid saturation of the diode array sensor. The absorbance spectrum was automatically computed using the dark and reference spectra and then individually saved for later processing. The scanning of kernels was randomly sequenced by feeding the kernels using a vibratory feeder. The feeder provided kernel singulation and avoided hand-contact with the kernels. The kernels from a packet were loaded into the feeder and sequentially scanned until 50 spectral scans were collected. A total of 1800 spectra were collected. This procedure was repeated for the packet, with replacement, to produce another 1800 spectra.

SINGLE-KERNEL COLOR IMAGE ACQUISITION

A color imaging system was recently developed to separate red and white winter wheat and is adaptable for other grain types. Briefly, the system consisted of a hopper, a two-stage feeder, a photo-electric trigger light sensor at the lower end of the chute, an area scan color camera, two halogen light sources, an air nozzle ejector device, two collection bins, and a computer. Seeds were placed in the hopper and were transported into the vibratory feeder through a belt feeder seed pickup. This arrangement was necessary for more effective singulation of seeds. The seeds then slid down an inclined chute, and as they exited the lower end, a photo-electric sensor detected the presence of a kernel and signaled the color camera to capture a 640 × 480 color image. The image was sent to the computer via an IEEE 1394 Firewire connection, and the processing was done using a Microsoft Visual C++ program. A specific description of the working parts and performance can be found in Pearson et al. (2008). The major difference between the system used in this study and that de-

scribed in the above reference is that the mirrors used for multi-view image capture were removed.

The procedure for corn kernel image acquisition was similar to the procedure for NIRS. Prior to data collection, the lamps and color camera were allowed to warm up for at least an hour. Kernels in a packet were then placed into the feeder to individually present them to the camera for image capture and storage. Each packet was randomly pre-sequenced by level of infection, and the packets were fed several times until 60 kernel images were collected for each packet. The extra 10 images were kept as alternative images because some kernels were occluded in the images. All images were stored using lossless 320×240 pixel bitmap (BMP) file format. This apparent scaling down of the image size was a result of extraction of the raw image pixels from the Bayer patterned image. Two rounds of image acquisition were performed for a total of 120 kernel images for each packet.

DATA ANALYSIS

NIR Spectroscopy

Prior to statistical analysis, datasets were prepared by merging the spectrum files and the sample data using Microsoft Visual Basic 6.0. The spectra were constrained to the range of 950 to 1650 nm to eliminate regions with higher spectral noise due to lower sensor sensitivity. Spectral data pretreatments were applied using the same program, and the datasets were saved in separate files.

Mean centering (MC) and standard normal variate (SNV) are two of the many pretreatments commonly used for spectral data. The MC pretreatment brought the spectra to a common zero absorbance axis using the following formula:

$$Xmc_{o,\lambda} = X_{i,\lambda} - \bar{X} \quad (1)$$

where \bar{X} is the mean of the absorbance values in the clipped range of a spectrum, and $Xmc_{o,\lambda}$ and $X_{i,\lambda}$ are the mean-centered and original absorbance values, respectively, at a particular wavelength (λ , nm). This pretreatment essentially eliminated shifts in the spectra caused by varying distances of the kernels from the fiber optic cable as they traveled down the light tube. Moreover, SNV pretreatment was applied to the absorbance data to minimize the effects of scattering and variable geometry of spectral acquisitions using the following formula:

$$Xsnv_{o,\lambda} = \frac{X_{i,\lambda} - \bar{X}}{SD} \quad (2)$$

where \bar{X} and SD are the mean and standard deviation, respectively, of the absorbance values in the clipped range of a spectrum, and $Xsnv_{o,\lambda}$ and $X_{i,\lambda}$ are the transformed and original absorbance values, respectively, for a particular wavelength (λ , nm). This study aimed to minimize further data pretreatments to achieve robust classification models.

The software, Tanagra version 1.4 (Rakotomalala, 2005), provided a graphical user interface to easily compare different models for machine learning. The instance selection tool in this program was used to randomly divide the dataset into a training set for model development at 50% of the total, with the remaining 50% as a validation set for model performance evaluation. To reduce the dimensional complexity of the spectrometric data, the five principal components (PCs) with the highest eigenvectors were computed from the training set and used as predictor variables. These five PCs were suffi-

cient enough to explain at least 99% of the variation in the spectra and were made consistent throughout the analyses. Linear discriminant analysis (LDA) was used to find classification models to discriminate infected from uninfected kernels and to discriminate between infecting fungus types. Similarly, a multi-layer perceptron (MLP) artificial neural network model with one hidden layer containing five neurons was also explored. The models were developed using the training set and were applied to both the training and validation datasets.

Color Imaging

Matlab 7.0.4 release 14 and its Image Processing Toolbox (The Mathworks, Inc., Natick, Mass.) were used to process the images as follows. Using an M-program script, a corn image file was loaded and broken down into its red, blue, and green planes. The pixel values in each color plane were bit-shifted to the right by two and counted into separate arrays. This step assigned the values into 64 gray levels (from 0 to 63), thereby achieving more meaningful frequency histograms. Since the background was essentially black, the count for the first level (zero value) was reset to zero. The frequency histograms were normalized by dividing the counts in each gray level by the sum of the remaining counts in all levels as follows:

$$NF_i = \frac{F_i}{\sum_{i=1}^{64} F_i} \quad (3)$$

where NF_i is the normalized frequency at level i , and F_i is the cumulative frequency count at level i . Furthermore, a cumulative normalized histogram was derived by subsequently adding up the values of the normalized frequency histogram from one to the target gray level i as follows:

$$CNF_i = \sum_{j=1}^i NF_j \quad (4)$$

where CNF_i is the cumulative normalized frequency at the target level i . An output file was finally generated by merging the cumulative histograms of all the images in the dataset with the details of the kernel samples.

LDA and MLP models were also used within the Tanagra software to discriminate infected from uninfected kernels and to discriminate between fungus types. However, only the data coming from the red and green planes were utilized in the analysis, as the blue plane contained little information to allow for the discrimination of infection levels. This study compared models derived from five PC scores of the cumulative histogram data against the models using raw cumulative histogram data. The dataset was also divided into 50% training and 50% validation sets, similar to the procedures adopted for NIR data.

RESULTS AND DISCUSSION

NIR SPECTROSCOPY

Figure 2 shows a plot of the averaged, mean-centered spectra for kernels with level 3 infections by the different fungi. The spectra of the two control samples had a wider span of values when compared to those of the fungus-infected kernels, which substantially overlap one another, except for the

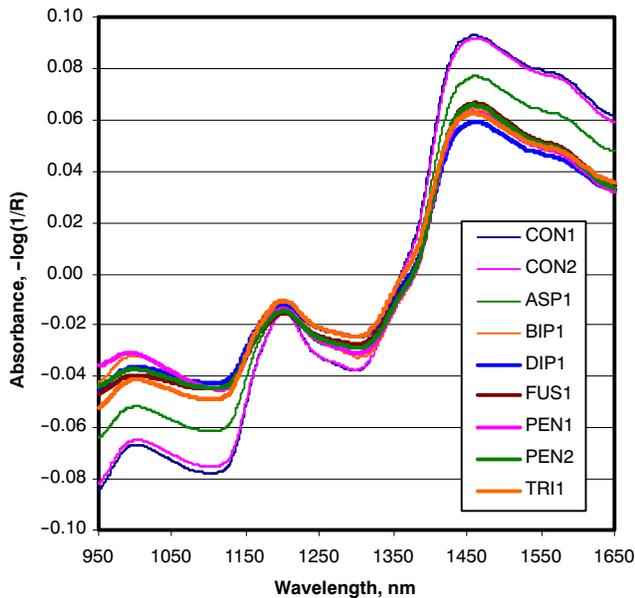


Figure 2. Average mean-centered absorbance spectra of control corn kernels (CON1 and CON2) and fungus-infected kernels at level 3 (ASP1 = *Aspergillus flavus*, BIP1 = *Bipolaris zeicola*, DIP1 = *Diplodia maydis*, FUS1 = *Fusarium oxysporum*, PEN1 = *Penicillium oxalicum*, PEN2 = *Penicillium funiculosum*, and TRI1 = *Trichoderma harzianum*).

Aspergillus flavus sample (ASP1), which lies about midway between the control kernel plots and the rest of the fungus species and proved more difficult to classify. Previous studies of NIRS for fungus detection utilized stationary kernels with a fixed distance from the collecting fiber optic probe (Wang et al., 2004; Pearson and Wicklow, 2006). As a result, the general shape of the NIR spectra of the kernels infected with different fungi generally looked similar, and the main difference was seen as shifts in the baselines. In this study, the kernels were in motion, contributing to an unavoidable baseline shift that required preprocessing treatment of the spectra to bring them to a common axis. Mean centering (MC) is a feasible and common approach; however, some spectroscopic variations that can be useful for the classification models will be lost in the data pretreatment process. Another approach is the standard normal variate (SNV). This brings the curves even closer to one another as it accounts for differences in the spans of absorbance values in the course of correcting for spectral scattering. This will leave the smaller spectral variations caused directly by scattering or adsorption of NIR light due to the different effects of the different types of fungi.

Intuitively, it was expected to be easier to discriminate kernels at rather advanced stages of infection (levels 3 and 4) from the control kernels compared to discriminating kernels at early stages of infection (levels 1 and 2) from the control kernels. Analyses for each level versus the control confirmed this expectation, but for brevity, the classification accuracies for only two cases are shown in table 2. Classifications by fungus were generally more accurate compared to combined fungus data. Combined data show reasonable classifications for the control and level 4 infected kernels, but other levels of infection were poorly classified. For further analysis, the kernels at levels 3 and 4 were combined to represent kernels at an advanced stage of infection. Similarly, levels 1 and 2 were combined to represent kernels in early stages of infection. This is similar to the approach of Pearson and Wicklow (2006).

Table 2. Classifications by infection level using LDA of NIR, mean centered, spectra for all fungus species and FUS1 (all values are percentages). Shaded cells show correct classifications.

	Control	Level 1	Level 2	Level 3	Level 4	Total Kernels
Classifications of all infected kernels						
Control	76	14	8	2	0	400
Level 1	36	26	22	4	0	350
Level 2	8	14	44	20	2	350
Level 3	2	3	22	38	23	350
Level 4	0	0	1	17	69	350
Classifications of FUS1 infected kernels						
Control	78	18	0	4	2	50
Level 1	20	56	20	4	0	50
Level 2	0	12	84	2	2	50
Level 3	0	0	8	70	22	50
Level 4	0	2	4	20	74	50

The NIR spectrometric data were analyzed using a two-stage approach. In the first stage, the main goal was to discriminate between the control and fungal-infected kernels. The second stage was aimed at discriminating between kernels infected with different fungus species. Table 3 shows the results of the linear discriminant analysis (LDA) and multi-layer perceptron neural network (MLP) models for discriminating infected kernels using the NIR spectrometric data receiving the two pretreatments for kernels in the validation sets. The numbers of kernels for each fungus-infected set were re-sampled to keep the number of kernels in each class as balanced as possible. All subsequent analyses used this method for model development.

The results in table 3 show that, on average, slightly better classification rates were obtained with MC compared to SNV for discrimination between uninfected and infected kernels. However, both the LDA and MLP models tended to classify the kernels as infected at the expense of false positives of uninfected control kernels. For levels 3 and 4, LDA can classify the uninfected kernels better than infected ones (96% versus 74%), while MLP had almost similar classification rates of 92% and 91% for uninfected and infected kernels, respectively. As expected, discriminating infected and uninfected kernels was much easier on average in this group

Table 3. Classification accuracies for control and infected kernels using LDA and MLP neural network models on the validation sets by NIR spectrometry (all values are percentages).

Level of Infection	LDA		MLP	
	Control	Infected	Control	Infected
Mean centering				
Levels 1, 2	81	73	64	80
Levels 3 and 4 ^[a]	96	74	92	91
All levels	89	79	84	83
Levels 3 and 4 ^[b]	98	63	98	72
Standard normal variate				
Levels 1 and 2	53	100	66	98
Levels 3 and 4 ^[a]	97	75	88	91
All levels	85	59	82	82
Levels 3 and 4 ^[b]	98	42	91	69

^[a] The model calibration dataset was comprised of 50% of the control and infected kernels at levels 3 and 4. The remaining 50% were used for validation for this case only.

^[b] The model calibration dataset was comprised of 50% of the control and infected kernels at levels 3 and 4. The remaining 50% were used for validation including kernels at levels 1 and 2. See text for explanation.

Table 4. Classification accuracies for control and infected kernels by fungus type using LDA and MLP neural network models on the validation sets using mean centered NIR spectrometry data (all values are percentages).

Fungus Species	LDA			MLP		
	Control	Levels 1 and 2	Levels 3 and 4	Control	Levels 1 and 2	Levels 3 and 4
ASP1	93	76	74	81	86	68
BIP1	93	65	90	94	70	93
DIP1	94	67	86	91	65	89
FUS1	82	80	88	80	76	88
PEN1	100	94	83	100	90	89
PEN2	96	80	92	94	71	90
TRI1	85	52	81	78	66	84

because of more pronounced differences in the observable severity of infection. For example, LDA had an average of 85% (mean of 96% and 74%) for discriminating level 3 and 4 infected kernels from uninfected kernels versus an average of 77% for levels 1 and 2. One approach to model development is to focus on the extreme cases of infection and use them for the entire span of infection. This approach yielded poorer results (table 3, mean centering, levels 3 and 4), with high classification rates for uninfected kernels (98%) and lower rates (63% and 72% for LDA and MLP, respectively) for infected kernels. By properly using all levels in model development (table 3, all levels), one can achieve a more robust model with a balanced classification accuracy. The LDA model still discriminated uninfected kernels better than infected ones (89% versus 79%), while MLP achieved balanced results at 83% versus 84%.

The results varied according to fungus type when using different modeling procedures to determine infection level, as shown in table 4. In general, control kernels were better recognized than infected kernels, and most infected kernels were detected well at levels 3 and 4, with the exception of kernels infected with ASP1. Most likely, the span of spectra for ASP1 could have significantly overlapped those of normal kernels, making any distinction difficult, as can be seen in figure 2. Kernels infected with PEN1 were detected well at levels 1 and 2, while other species had mixed results at this level of infection. Detection differences between PEN1 and PEN2 are attributed to the narrower range of infection in PEN2, along with the fact that PEN1 showed signs of severe sporulation and darker kernels, particularly for levels 3 and 4, which would be likely due to difference in species.

The accuracy of identifying the infecting fungus species is higher at more advanced levels of infection because of better scattering of light, as explained by Pearson and Wicklow (2006). Table 5 shows the results for the classification accuracy of identifying the fungal species using the LDA and MLP models at different levels of infection and data pretreatment. Relatively better results were obtained for the MLP model when using either the MC or SNV pretreatments. At an advanced stage of fungal infection, the species BIP1 and PEN1 were most accurately identified, followed by the control and FUS1. Both BIP1 and PEN1 had darker kernels than usual, which could have aided in their identification. However, the results for MLP with MC pretreatment indicated that PEN1 had higher levels of classification accuracy, even when all levels of infection were considered. BIP1 classification accuracies decreased markedly, from 99% for levels 3 and 4 to 63% for all levels. The low accuracies for ASP1 and DIP1 were likely caused by close similarity of their spectra to other species and were thus misclassified as other species.

Using artificial neural network models instead of the LDA models did not give any clear advantage for discrimination between healthy and fungus-infected kernels. Both types of models used the same PC scores for model development, which could have eliminated any nonlinear variations and thus led MLP to perform as well as the LDA models. When the entire sample set was sorted, less than 20% false negatives for fungal infection were observed using the MLP models. This rate of false negatives is most likely caused by inclusion of asymptomatic kernels that are actually healthy as long as their pericarps remain intact (Gembeh et al., 2001). Additionally, the kernels were continuously moving, such that when spectrometric data were taken by the instrument, sound areas of the kernels may strongly influence the spectra. This would further complicate the sorting process. A good approach to overcome this is to re-run the sound kernels until all infected kernels are separated.

Identifying the infecting fungus was difficult, probably because of the removal of the spectral baselines that could help in the models. As expected, the more severe the infection is, the easier the identification became. In addition, the number and types of fungi that are incorporated in the model development certainly affect the overall performance of the model. A lesser number of species would enable more robust separation of spectra, consequently decreasing the chances of classifying a fungus-infected kernel to the wrong species class.

Table 5. Classification accuracy for identifying the fungus type using LDA and MLP neural network models on the validation data sets by NIR spectrometry (all values are percentages).

	Levels 1 and 2				Levels 3 and 4				All Levels			
	LDA		MLP		LDA		MLP		LDA		MLP	
	MC	SNV	MC	SNV	MC	SNV	MC	SNV	MC	SNV	MC	SNV
CON	72	32	72	28	90	34	92	77	88	69	88	45
ASP1	20	35	0	25	27	17	42	12	34	14	23	12
BIP1	38	36	45	50	83	83	98	99	41	54	63	63
DIP1	18	14	1	1	60	33	68	53	6	4	13	12
FUS1	58	54	57	39	69	69	63	78	61	51	61	47
PEN1	84	81	86	94	94	79	90	99	68	66	93	97
PEN2	57	59	53	49	52	57	56	36	42	22	61	54
TRI1	51	41	82	69	59	72	58	65	63	45	75	73
Mean	50	44	50	44	67	56	71	65	50	40	59	50

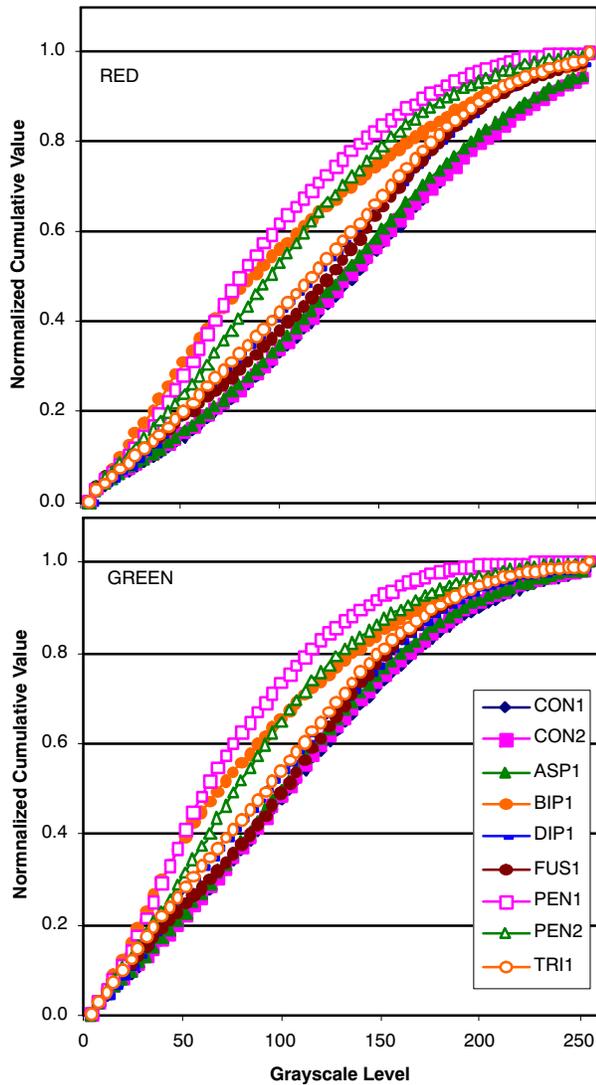


Figure 3. The normalized cumulative frequency plots of the red and green image planes of corn control kernels (CON1 and CON2) and fungus-infected kernels (ASP1 = *Aspergillus flavus*, BIP1 = *Bipolaris zeicola*, DIP1 = *Diplodia maydis*, FUS1 = *Fusarium oxysporum*, PEN1 = *Penicillium oxalicum*, PEN2 = *Penicillium funiculosum*, and TRI1 = *Trichoderma harzianum*) also showed substantial overlap.

Table 6. Classification accuracies for control and infected kernels using LDA and MLP neural network models on the validation sets by color imaging (all values are percentages).

Level of Infection	LDA		MLP	
	Control	Infected	Control	Infected
With PCA				
Levels 1 and 2	62	65	68	79
Levels 3 and 4	89	83	86	82
All levels	75	75	56	85
No PCA				
Levels 1 and 2	59	64	37	84
Levels 3 and 4	86	83	89	81
All levels	68	71	69	72

COLOR IMAGING

Normalized cumulative histogram plots for the red and green color channels are shown in figure 3 for the different fungus types at advanced stages (levels 3 and 4) of infection. The line plots for the control and ASP1 appear at the bottom, while BIP1, PEN1, and PEN2 species compose the upper region. This separation could be explained by the relatively higher abundance of darker pixels that skew their histograms to the left, leading them to accumulate pixels faster than the images of the other fungus species.

Table 6 shows the classification accuracies of various models for identifying control and infected kernels. Similar to the trend of the NIRS results, higher levels of infection had better classification accuracies (81% to 89%). When the entire data set (all levels) was used, classification accuracies of 75% were achieved with LDA. The use of principal component analysis on the data did not improve the classification results, while LDA performed almost as well as the MLP models, with or without the use of principal component analysis. Compared to the NIRS results, the results of classification accuracy of color imaging was probably less attractive because of the better interrogation of the physical state and chemical constituent contents of kernels by NIRS. There was no clear dominant result in the classification of the infecting fungal species, except that slightly better results were again observed for the BIP1 and PEN1 species (table 7).

One plausible constraint on the use of color imaging is the limited electromagnetic range in which optical data can be obtained. Physical discoloration in the kernels is the only source of variation, and this could possibly work well for discrimination of the levels of infection for the same fungus species or between two species at a rather advanced stage of

Table 7. Classification accuracy for identifying the fungus type using LDA and MLP neural network models on the validation data sets by color imaging (all values are percentages).^[a]

	Levels 1 and 2				Levels 3 and 4				All Levels			
	LDA		MLP		LDA		MLP		LDA		MLP	
	+PCA	-PCA	+PCA	-PCA	+PCA	-PCA	+PCA	-PCA	+PCA	-PCA	+PCA	-PCA
CON	66	34	35	20	43	32	31	34	46	41	34	44
ASP1	1	22	4	13	33	25	4	23	30	24	6	19
BIP1	28	34	13	39	42	43	38	38	43	38	27	36
DIP1	12	16	0	8	7	13	0	7	17	15	2	15
FUS1	19	16	11	17	27	19	7	6	32	24	0	15
PEN1	21	28	33	19	28	32	5	23	33	24	27	22
PEN2	13	15	36	14	21	16	15	24	21	16	15	26
TRI1	0	21	26	63	11	22	84	54	10	18	85	39
Mean	25	23	21	24	26	25	23	26	29	25	24	27

^[a] +PCA = principal component analysis applied to the histogram data; -PCA = no principal component analysis applied.

infection. Improvements to the vision system could include the incorporation of multiple views of a kernel.

CONCLUSION

In general, NIRS and imaging methods have good recognition of heavily mold-infected and uninfected kernels. The rapid single-kernel NIRS instrument performed better than the color imaging system in discriminating between infected and healthy corn kernels. This instrument also seems to perform better at identifying the BIP1 and PEN1 fungal species on heavily infected kernels, but it performs poorly for other types and infection levels. In most cases, MLP models and LDA models had similar levels of accuracy using any of the data pretreatments. Despite the limited accuracy of the instruments to discriminate lesser infected kernels, both were able to discriminate between the uninfected and the more heavily infected kernels. As such, this method could be useful as a screening tool for quality control and studying hybrid resistance to infection or agronomic conditions leading to significant incidences of mold.

REFERENCES

- Abbas, H. K., R. D. Cartwright, W. Xie, and W. T. Shier. 2006. Aflatoxin and fumonisin contamination of corn (maize, *Zea mays*) in Arkansas. *Crop Protection* 25(1): 1-9.
- Armstrong, P. R. 2006. Rapid single-kernel NIR measurement of grain and oil-seed attributes. *Applied Eng. in Agric.* 22(5): 767-772.
- Berardo, N., V. Pisacane, P. Battilani, A. Scandolaro, A. Pietri, and A. Marocco. 2005. Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. *J. Agric. Food Chem.* 53(21): 8128-8134.
- Betran, F. J., and T. Isakeit. 2004. Aflatoxin accumulation in maize hybrids of different maturities. *Agron. J.* 96(2): 565-570.
- Brown, R. L., Z.-Y. Chen, T. E. Cleveland, and J. S. Russin. 1999. Advances in the development of host resistance in corn to aflatoxin contamination by *Aspergillus flavus*. *Phytopathology* 89(2): 113-117.
- Bruns, H. A. 2003. Controlling aflatoxin and fumonisin in maize by crop management. *J. Toxicology* 22(2-3): 153-173.
- 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.
- Gembeh, S. V., R. L. Brown, C. Grimm, and T. E. Cleveland. 2001. Identification of chemical components of corn kernel pericarp was associated with resistance to *Aspergillus flavus* infection and aflatoxin production. *J. Agric. Food. Chem.* 49(10): 4635-4641.
- Gordon, S. H., R. W. Jones, J. F. McClelland, D. T. Wicklow, and R. V. Greene. 1999. Transient infrared spectroscopy for detection of toxigenic fungi in corn: Potential for on-line evaluation. *J. Agric. Food. Chem.* 47(12): 5267-5272.
- Marasas, W. F., R. T. Riley, K. A. Hendricks, V. L. Stevens, T. W. Sadler, J. Gelineau-van Waes, S. A. Missmer, J. Cabrera, O. Torres, W. C. Gelderblom, J. Allegood, C. Martinez, J. Maddox, J. D. Miller, L. Starr, M. C. Sullards, A. V. Roman, K. A. Voss, E. Wang, and A. H. Merrill Jr. 2004. Fumonisin disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: A potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. *J. Nutr.* 134(4): 711-716.
- Menkir, A., R. L. Brown, R. Bandyopadhyay, Z.-Y. Chen, and T. E. Cleveland. 2006. A USA-Africa collaborative strategy for identifying, characterizing, and developing maize germplasm with resistance to aflatoxin contamination. *Mycopathologia* 162(3): 225-232.
- Ng, H. F., W. F. Wilcke, R. V. Morey, and J. P. Lang. 1998. Machine vision evaluation of corn kernel mechanical and mold damage. *Trans. ASAE* 41(2): 415-420.
- Odriozola, E., A. Odeon, G. Canton, G. Clemente, and A. Escande. 2005. *Diplodia maydis*: A cause of death of cattle in Argentina. *New Zealand Veterinary J.* 53(2): 160-161.
- Pearson, T. C., and D. T. Wicklow. 2006. Detection of corn kernels infected by fungi. *Trans. ASABE* 49(4): 1235-1245.
- Pearson, T. C., D. Brabec, and S. Haley. 2008. Color image based sorter for separating red and white wheat. *Sensing Instrum. Food Qual. Safety* 2(4): 280-288.
- Prozesky, L., T. S. Kellerman, D. P. Swart, B. P. Maartens, and R. A. Schultz. 1994. Perinatal mortality in lambs of ewes exposed to cultures of *Diplodia maydis* (= *Stenocarpella maydis*) during gestation: A study of central nervous system lesions. *Onderstepoort J. Vet. Res.* 61(3): 247-253.
- Rakotomalala, R. 2005. TANAGRA: Un logiciel gratuit pour l'enseignement et la recherche. In *Actes de EGC 2005*, 2: 697-702. RNTI-E-3. Berlin, Germany: Springer-Verlag.
- Wang, D., F. E. Dowell, M. S. Ram, and W. T. Schaupaugh. 2004. Classification of fungal-damaged soybean seeds using near-infrared spectroscopy. *Intl. J. Food Prop.* 7(1): 75-82.
- Wicklow, D. T. 1995. The mycology of stored grains: An ecological perspective. In *Stored Grain Ecosystems*, 197-249. D. S. Jayas, N. D. G. White, and W. E. Muir, eds. New York, N. Y.: Marcel Dekker.
- Wicklow, D. T. 1999. Influence of *Aspergillus flavus* strains on aflatoxin and bright greenish yellow fluorescence of corn kernels. *Plant Dis.* 83(12): 1146-1148.
- Williams, J. H., T. D. Phillips, P. E. Jolly, J. K. Stiles, C. M. Jolly, and D. Aggarwal. 2004. Human aflatoxicosis in developing countries: A review of toxicology, exposure, potential health consequences, and interventions. *American J. Clin. Nutr.* 80(5): 1106-1122.