Identifying Stored-Grain Insects Using Near-Infrared Spectroscopy

F. E. DOWELL, J. E. THRONE, D. WANG, AND J. E. BAKER

Grain Marketing and Production Research Center, USDA–ARS, 1515 College Avenue, Manhattan, KS 66502

ABSTRACT   Proper identification of insects in grain storage facilities is critical for predicting development of pest populations and for making management decisions. However, many stored-grain insect pests are difficult to identify, even for trained personnel. We examined the possibility that near-infrared (NIR) spectroscopy could be used for taxonomic purposes based on the premise that every species may have a unique chemical composition. Tests were conducted with 11 species of beetles commonly associated with stored grain. Spectra from individual insects were collected by using a near-infrared diode-array spectrometer. Calibrations were developed by using partial least squares analysis and neural networks. The neural networks calibration correctly identified >99% of test insects as primary or secondary pests and correctly identified >95% of test insects to genus. Evidence indicates that absorption characteristics of cuticular lipids may contribute to the classification of these species. We believe that this technology could be used for rapid, automated identification of many other organisms.

KEY WORDS   beetles, near-infrared spectroscopy, stored grain, taxonomy, neural networks

SELECTING MANAGEMENT STRATEGIES for insect pests in stored grain requires decisions that are based on measurable, interacting factors (Flinn and Hagstrum 1990). These factors include type of grain, initial storage time, grain temperature, grain moisture content, storage duration, prior insecticide use, degradation rate of protectant insecticides, and number and species of pest insects present. Correct identification of pest insects can be critical for management decisions. For example, the lesser grain borer, *Rhyzopertha dominica* (F.) (Coleoptera: Bostrichidae), and the rusty grain beetle, *Cryptolestes ferrugineus* (Stephens) (Coleoptera: Laemophloeidae), are major pests of wheat in the Midwest grain belt (USDA 1986). *R. dominica* is a primary pest and can damage intact grain kernels, whereas *C. ferrugineus* is a secondary feeder and normally does not damage whole kernels. Fumigation may be necessary or advised when high populations of *R. dominica* are present, whereas fumigation may not be required when high populations of *C. ferrugineus* are present. Unfortunately, correct identification of many pest species found in grain, particularly the beetles, can be difficult, time consuming, and can require advanced taxonomic training. Thus, a rapid, automated system for identifying insects could be useful in making pest management decisions.

Previous research shows that near-infrared spectroscopy (NIRS) can rapidly and automatically detect the presence of hidden insect larvae or external adult insects in wheat samples (Ridgway and Chambers 1996, Dowell et al. 1998). In this report, we provide evidence that NIRS may be used to classify rapidly adult beetles likely to be present in grain.

NIRS should be able to differentiate between insect species based on their absorbance characteristics because the cuticle of each insect species may have a unique chemical composition (Lockey 1988). This unique chemical composition causes molecules to vibrate at unique frequencies and absorb NIR energy corresponding to these frequencies and overtones of these fundamental frequencies (Murray and Williams 1990). For example, a molecule of water vibrates at fundamental frequencies that result in absorption of NIR energy at ≈2,760 nm. This absorption also can be measured at 1st and 2nd overtones located around 1,400 and 950 nm. Thus, moisture content can be determined by measuring the intensity of absorption at any of these wavelengths. This same principle applies to other organic compounds such as protein, lipids, starch, and others. Gibbs and Crowe (1991) showed Fourier transform infrared spectroscopy (>2,500 nm) could detect differences in hydrocarbons from lipid extracts obtained from different regions (e.g., head, ventral abdomen) of a single insect species, but they did not use NIRS to differentiate between species.

The objective of our research was to determine if NIRS (400–1,700 nm) could differentiate between primary and secondary insect pests of stored products, differentiate between genera within primary and secondary insects, differentiate species within a genus, and differentiate between species across genera by using spectra collected from individual, whole insects.

Table 1. Stored-grain beetles (Coleoptera) used in the neural network and partial least squares analysis of NIR spectra when comparing species, genera, and primary versus secondary pests

Code	Common name	Scientific name	Family
		Secondary Pests	
FGB	Flat grain beetle	*Cryptolestes pusillus* (Schönherr)	Laemophloeidae
RGB	Rusty grain beetle	*Cryptolestes ferrugineus* (Stephens)	Laemophloeidae
SGB	Sawtoothed grain beetle	*Oryzaephilus surinamensis* (L.)	Silvanidae
MGB	Merchant grain beetle	*Oryzaephilus mercator* (Fauvel)	Silvanidae
CFB	Confused flour beetle	*Tribolium confusum* Jacquelin du Val	Tenebrionidae
RFB	Red flour beetle	*Tribolium castaneum* (Herbst)	Tenebrionidae
		Primary Pests	
LGB	Lesser grain borer	*Rhyzopertha dominica* (F.)	Bostrichidae
GB	Larger grain borer	*Prostephanus truncatus* (Horn)	Bostrichidae
GW	Granary weevil	*Sitophilus granarius* (L.)	Curculionidae
RW	Rice weevil	*Sitophilus oryzae* (L.)	Curculionidae
MW	Maize weevil	*Sitophilus zeamais* Motschulsky	Curculionidae

Materials and Methods

About 20 adult insects were selected from each of 11 beetle species commonly found in stored grain (Table 1). These species were obtained from stock colonies reared at 25°C and 50–60% RH. Composition of laboratory diets for the different species were as follows: diets for the flat grain beetle, *Cryptolestes pusillus* (Schönherr) (Coleoptera: Laemophloeidae); rusty grain beetle; sawtoothed grain beetle, *Oryzaephilus surinamensis* (L.) (Coleoptera: Silvanidae); and merchant grain beetle, *Oryzaephilus mercator* (Fauvel) (Coleoptera: Silvanidae), consisted of oatmeal: whole wheat flour:Brewer's yeast:wheat germ (60:30: 5:5), (vol:vol); diets for the confused flour beetle, *Tribolium confusum* Jacquelin du Val (Coleoptera: Tenebrionidae), and the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae), consisted of whole wheat flour and Brewer's yeast (95:5), (vol:vol); the lesser grain borer was reared on whole wheat lightly dusted with wheat flour; the larger grain borer, *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae), and the maize weevil, *Sitophilus zeamais* Motschulsky (Coleoptera: Curculionidae), were reared on whole-kernel corn; and the rice weevil, *Sitophilus oryzae* (L.) (Coleoptera: Curculionidae), and granary weevil, *Sitophilus granarius* (L.) (Coleoptera: Curculionidae), were reared on hard red winter wheat.

The live insects were individually placed in a black V-shaped trough (12 mm long, 10 mm wide, 5 mm deep) and illuminated with white light via a fiber bundle (8 mm diameter) positioned 13 mm from the top of the trough and oriented 45° from vertical. A reflectance probe (2-mm diameter) was oriented vertically 9.5 mm from the top of the trough. The reflectance probe carried the reflected energy to a spectrometer (DA7000, Perten Instruments, Springfield, IL). The diode-array spectrometer measures visible (400–750 nm) and NIR (750–1,700 nm) reflectance at a rate of 30 spectra per second. Procedures included collecting a baseline, collecting 8 spectra from each of the insects, and averaging the 8 spectra for each insect. This resulted in spectra from ≈275 individual insects. Collecting and averaging the 8 spectra from each in-

sect took <1 s. The baseline consisted of collecting a spectrum of the empty trough to use as a reference. A new baseline was collected after each group of ≈40 insects.

Data were analyzed using partial least squares regression (PLS) (Galactic Industries 1996) and a back-propagation neural network (NeuralWare 1995). Correlation plots (r^2) and factors indicating which wavelengths contribute to classifications in 2-way comparisons were obtained using PLS (Murray and Williams 1990). Two-way comparisons were made using PLS by assigning a value of 1 or 2 to the comparisons of interest (for example, primary = 1, secondary = 2).

The neural network analysis gave classification percentages for all 2-way as well as higher-order comparisons (up to 11 species compared). The neural network had 1 hidden layer, 100 input nodes, and 2, 3, or 11 outputs. The 100-input nodes corresponded to the absorbance values at 10 nm increments from 700 to 1,700 nm. The outputs corresponded to the comparisons being tested. The learning rate, momentum, and learning events used in the neural network were ≤0.6, 0.4–0.5, and 10,000–30,000, respectively. The learning rate, momentum, learning events, and number of hidden layers affect the neural network accuracy and speed (Hecht-Neilsen 1989). For both the PLS and neural network analyses, even-numbered samples served as training or calibration sets, whereas odd samples were used for testing calibration models.

Table 2 shows the comparisons tested. The classification of paramount interest was primary versus secondary insects because management strategies for these 2 types of insects may be quite different and because knowing whether an insect is a primary or secondary pest is usually sufficient for making a management decision. In addition, insects were classed into their respective genera (2 or 3 comparisons) within the primary and secondary groupings, and classed into species within a genus (2 or 3 comparisons). A final comparison sought to classify insects into their respective species independent of previous groupings (11 comparisons).

Table 2. Accuracy of classifying insects species with calibrations developed by using partial least squares (PLS) regression and a neural network (NN)

Comparison[a]	n	% correct		no. PLS factors
		NN	PLS	
Primary vs Secondary Insects				
(FGB, RGB, SGB, MGB, RFB, CFB) vs (GW, RW, MW, LGB, GB)	110	99.1	96.4	8
Families or genera within primary or secondary				
(GW, RW, MW) vs (LGB, GB)	60	100	100	10
(FGB, RGB, SGB, MGB) vs (RFB, CFB)	60	95	100	6
(FGB, RGB) vs (SGB, MGB) vs (RFB, CFB)	60	96.7	—	—
Species within genera or family				
GW vs RW vs MW	30	83.3	—	—
LGB vs GB	20	100	100	6
FGB vs RGB	20	90	90	5
SGB vs MGB	20	55	60	3
RFB vs CFB	20	80	100	13
GW vs RW	20	100	85	3
GW vs MW	20	95	100	5
RW vs MW	20	75	95	5
All Species	110	71	—	—

[a] See Table 1 for species codes.

We reasoned that the insect cuticle likely would absorb most of the NIR energy. To test this, the spectral absorbance of rice weevil cuticular lipids, extracted with a chloroform ($CHCl_3$) rinse and impregnated onto filter paper, was measured and analyzed for comparison with PLS correlations and factors. Lipids were extracted by placing 26.4 g of rice weevil adults in a 500-ml flask, adding 200 ml $CHCl_3$, and gently swirling for 3 min at room temperature. The extract was filtered through Whatman no. 1 filter paper, concentrated in a rotoevaporator, and an aliquot containing 1 mg of lipid was applied to a Whatman no. 1 filter paper disk (6.4 mm diameter). The lipid concentration was 0.78 mg/cm^2 of filter paper.

Spectra of ground samples of cuticle from 5th-instar tobacco hornworm larvae, *Manduca sexta* (L.) (Lepidoptera: Sphingidae), and the β(1–4)-linked hexasaccharide of 2-acetamido-2-deoxy-D-glucopyranoside derived from crab chitin (Sigma, St. Louis, MO) were obtained to determine if their absorbance peaks are identical or similar to the wavelengths found useful for detection.

To test the robustness of NIR calibrations developed using the laboratory stock colonies, insects from each of 3 species—the foreign grain beetle, *Ahasverus Advena* (Waltl) (Coleoptera: Silvanidae); *C. ferrugineus;* and *R. dominica*—were obtained from samples of wheat from bins in Kansas and classed using the NIR system. This testing occurred 6 mo after the original calibrations.

Results and Discussion

The neural network classified 99.1% of primary and secondary insects correctly, whereas PLS had a lower classification percentage (96.4%) (Table 2). Both calibrations classified insects by genus within primary and secondary groups with an accuracy of ≥95%. Classifying insect species within a genus resulted in correct classifications ranging from 55 to 100%. In an 11-way classification among all species and using the neural network calibration, classification accuracies ranged from 30 to 100% (Table 3). The worst classifications were the sawtoothed grain beetle versus merchant

Table 3. Neural network results from an 11-way classification

Actual species	Predicted species[a]										
	FGB	RGB	SGB	MBG	CFB	RFB	LGB	GB	GW	RW	MW
FGB	90	10	—	—	—	—	—	—	—	—	—
RGB	40	50	10	—	—	—	—	—	—	—	—
SGB	—	—	30	30	—	—	40	—	—	—	—
MGB	—	—	20	70	—	—	10	—	—	—	—
CFB	—	—	—	—	60	40	—	—	—	—	—
RFB	—	—	—	—	20	80	—	—	—	—	—
LGB	—	—	—	—	—	—	100	—	—	—	—
GB	—	—	—	—	—	—	—	90	10	—	—
GW	—	—	—	—	—	—	—	—	100	—	—
RW	—	—	—	—	—	—	—	—	10	70	20
MW	—	—	—	—	—	—	—	—	—	10	90

See Table 1 for species codes.
[a] Results are the percentage of actual species classified into each of the 11 species categories.
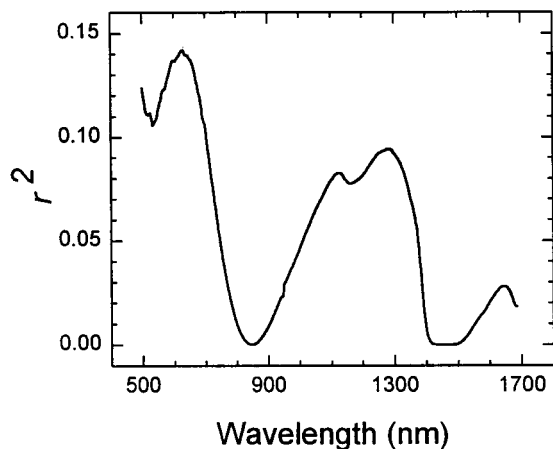
Fig. 1. Correlation spectra for primary versus secondary stored-grain insects.

Fig. 2. PLS factor plots indicating wavelengths contributing to insect classifications.

grain beetle (55% correct) in the pair-wise comparison (Table 2) and the sawtoothed grain beetle versus all others (30% correct) in the 11-way classification (Table 3). Thus, the calibrations developed using neural network and broadest groupings (primary versus secondary) resulted in the highest classification accuracy.

When examining the effect of wavelength regions on classification accuracies, the PLS results showed that data within the visible wavelength region did not improve classifications or reduce the number of factors needed for calibrations (data not shown). Fewer factors are desirable because less information is needed to explain variability in the data. For example, for the primary versus secondary comparison, the calibration models developed using either the NIR or NIR plus visible wavelengths both resulted in 96.4% of insects correctly classed. The number of factors needed for the NIR and NIR plus visible region calibrations were 8 and 17 factors, respectively. Fewer factors when using only the NIR region likely occurred because there is little visible difference between many of the insect categories. Thus, including the visible region contributed no additional useful information.

Further information about wavelengths contributing to classifications can be derived from PLS correlation plots and factor weights. Correlation plots show that wavelengths of 450–700 nm, 900–1,400 nm, and 1,500–1,700 nm were more highly correlated to insect species than other wavelengths (Fig. 1).

The first 2 factors of the PLS comparisons showed that wavelengths with the most weight occurred at ≈1,130, 1,325, and 1,670 nm. The 3rd and 4th factors had peaks around 1,420 nm. Fig. 2 shows plots for 1 and 3 PLS factors. When comparing the wavelengths from the first 2 factors with absorbances of various functional groups, the absorbances correspond closely to the 1st and 2nd overtones of CH_3, and to a lesser extent CH_2 (Murray and Williams 1990). The CH combination overtones correspond with the 3rd and 4th fac-
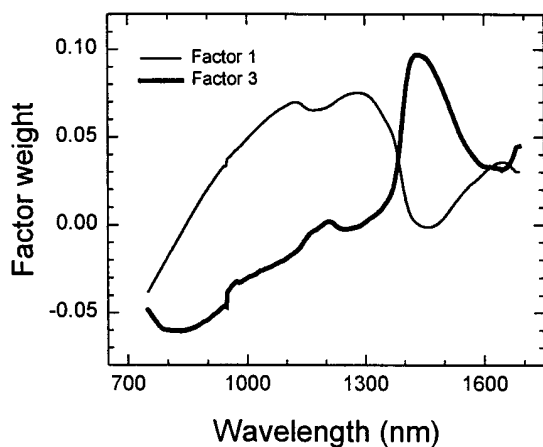
tors. These absorbance regions also generally agree with those reported by Ridgway and Chambers (1996), but at overtones found at longer wavelengths (1,700 nm) when detecting the presence of insects in grain.

CH_3 and CH_2 are common chemical moieties in components that make up the epicuticular lipids in insects. Insect cuticular lipids are composed mainly of fatty acids, alcohols, esters, glycerides, sterols, aldehydes, ketones, and hydrocarbons (Lockey 1988). Long-chain hydrocarbons often are major components of cuticular lipids in insects, but their concentration can vary widely, from 3 to nearly 95% of the total lipid. Hydrocarbons make up ≈32% of the total surface lipid in rice weevils (Baker et al. 1984). Fig. 3 shows the difference spectra calculated by subtracting a spectra of rice weevil cuticular lipid on filter paper (0.78 mg/cm²) from the spectra of filter paper treated with solvent only. This difference spectra with peaks
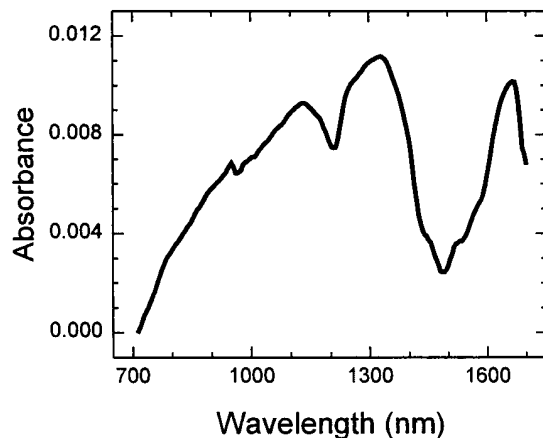

Fig. 3. Spectra of cuticular lipids extracted from adult rice weevils.

occurring at the CH_3 overtones ($\approx$1,130 and 1,670 nm) supports the conclusions indicated by the PLS factors. Thus, each insect species appears to have molecules with unique vibrational characteristics that may be caused by the unique mixture of hydrocarbon molecules and other lipid classes.

Spectra of the chitin hexamer and ground insect cuticle had absorbance peaks around 1,400–1,500 nm. Neither the correlation plots nor the factor weight plots indicated that this wavelength range was useful. Thus, it appears that the NIR system may have detected differences in cuticular lipids between species, but that the chitin within the cuticle did not contribute to classifications. Other compounds contained in insect cuticle that could be contributing to classification include protein, catachols, pigments, and oxalates (Kramer et al. 1995).

When classifying field insects using calibrations developed from laboratory stock colonies, 100% of the *C. ferrugineus* were correctly classed as secondary insects and in the correct genus. The *A. advena* were not included in the original calibration; however, 80% were correctly classed by the model as secondary insects. For the *R. dominica,* 67% were correctly classed as primary insects. Of the *R. dominica* correctly classed as primary, 83% were placed in the correct genus. Of those placed in the correct genus, 100% were placed in the correct species. The calibration developed using laboratory colonies appears to classify field insects with reasonable accuracy, especially considering that the configuration for illuminating and viewing insects had changed between the original calibration and subsequent testing of field insects.

Insects that enter traps can be counted electronically (Shuman and Weaver 1996). However, integration of filtered NIR sensors within an insect trap could provide an automated means of not only counting the trapped insects but also identifying the insect to type or species. This type of timely information concerning pest insect populations in stored grain would certainly be an advantage when implementing control strategies.

Computer vision, which uses digitized images from cameras and provides information about object color, shape, and size, could provide an alternate means of identifying insect species. Although computer vision could likely be integrated into an insect trap, it poses additional problems of proper lighting, shadows, insect presentation, image segmentation (Zayas and Flinn 1997).

In summary, our results showed that NIR spectroscopy coupled with PLS or neural network spectral analysis techniques can be used to classify the 11 insect species examined in this study, with primary and secondary insects being classed with >99% accuracy. The unique composition of cuticular lipids in the different beetles may be partially responsible for the classifications achieved with this system. Although we were not able to classify all tested insects to the species level with high accuracy, identification to species is not necessarily required for making pest management decisions in grain storages. Identification to genus or identification as a primary or secondary pest is usually sufficient. In addition to stored grain insects, we believe that this technology could be used for rapid, automated identification of many other organisms.

References Cited

Baker, J. E., S. M. Woo, D. R. Nelson, and C. L. Fatland. 1984. Olefins as major components of epicuticular lipids of three *Sitophilus* weevils. Comp. Biochem. Physiol. B 77: 877–884.

Dowell, F. E., J. E. Throne, and J. E. Baker. 1998. Automated nondestructive detection of internal insect infestation of wheat kernels using near-infrared reflectance spectroscopy. J. Econ. Entomol. 91: 899–904.

Flinn, P. W., and D. W. Hagstrum. 1990. Stored grain advisor: a knowledge-based system for management of insect pests of stored grain. AI Applications Nat. Res. Manage. 4: 44–52.

Galactic Industries. 1996. Grams/32 user's guide, version 4.0. Galactic, Salem, NH.

Gibbs, A., and J. H. Crowe. 1991. Intra-individual variation in cuticular lipids studied using fourier transform infrared spectroscopy. J. Insect Physiol. 37: 743–748.

Hecht-Neilsen, R. 1989. Neural computing. Addison-Wesley, New York.

Kramer, K. J., T. L. Hopkins, and J. Shaefer. 1995. Applications of solids NMR to the analysis of insect sclerotized structures. Insect Biochem. Mol. Biol. 25: 1067–1080.

Lockey, K. H. 1988. Lipids of the insect cuticle: origin, composition and function. Comp. Biochem. Physiol. B 89: 595–645.

Murray, I., and P. C. Williams. 1990. Chemical principles of near-infrared technology, pp. 17–34. *In* P. C. Williams and K. H. Norris [eds.], Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists. St. Paul, MN.

NeuralWare. 1995. Reference guide. NeuralWare, Pittsburgh, PA.

Ridgway, C., and J. Chambers. 1996. Detection of external and internal insect infestation in wheat by near-infrared reflectance spectroscopy. J. Sci. Food Agric. 71: 251–264.

Shuman, D., and D. Weaver. 1996. Innovations in electronic monitoring of stored-grain insects, pp. 57–1 to 57–4. *In* Proceedings, Annual International Research Conference On Methyl Bromide Alternatives and Emissions Reductions, 4–6 November 1996, Orlando, FL. Methyl Bromide Alternatives Outreach, Fresno, CA.

U.S. Department of Agriculture. 1986. Stored-grain insects. U.S. Dep. Agric. Agric. Res. Serv. Agric. Handb. 500.

Zayas, I. Y., and P. W. Flinn. 1997. Detection of insects in bulk wheat samples with machine vision, paper no. 973149. Presented at the Annual American Society of Agricultural Engineers meeting, 10–14 August 1997, Minneapolis, MN, ASAE, St. Joseph, MI.