Visible/Near Infrared Reflectance (VNIR) Spectroscopy for Detecting Twospotted Spider Mite (Acari: Tetranychidae) Damage in Strawberries

AIMEE B. FRAULO,1,2 MATTHEW COHEN,3 AND OSCAR E. LIBURD4


ABSTRACT The twospotted spider mite, Tetranychus urticae Koch, is among the most economically important pests in strawberries (Fragaria spp.). As T. urticae feeds, it ingests mesophyll cells that contain pigments essential for physiologic function and alters radiant energy use of the leaf tissue, severely compromising plant health and productivity. In our study, diffuse reflectance spectroscopy in the visible and near infrared (VNIR) portions of the spectrum was used to identify specific spectral regions altered by T. urticae feeding and to quantitatively assess T. urticae density. During the 2006–2007 growing season, 80 strawberry leaflets with varying levels of T. urticae infestation were collected. Spectral classification of both mite density (continuous) and mite density class (categorical) were developed. Spider mite density classes were low infestation (0–20 mites/leaflet), moderate infestation (20–50 mites/leaflet), and high infestation (≥50 mites/leaflet). Continuous spectral prediction for leaf infestation was developed using partial least squares (PLS) regression. Classification trees were used to train spectra to categorical levels of infestation. Both models were calibrated with 67% of the samples, and accuracy was evaluated using the remaining 33%. Categorical validation accuracy was 81%, with odds ratios for correctly predicting extreme categories (low and high) of 33 and 47.7, respectively. Continuous validation efficiency was also high, with an $r^2$ between predicted and observed of 0.85 and a root-mean-squared error (RMSE) of 12.2 mites per leaf. Developing a spectral pest monitoring system would provide a diagnostic tool allowing early and effective intervention for precision management of T. urticae in strawberry.

KEY WORDS spider mites, precision pest management, visible and near infrared spectra, Fragaria spp.
ternal scattering at the cell wall–air interface (Pinter et al. 2003). Steep gradients in reflectance between red (650–700 nm) and NIR, known as the “red edge,” indicates plant stress frequently caused by dehydration and cellular damage (Pinter et al. 2003, Lillesand et al. 2004).

Studies by Fitzgerald et al. (2004) and Landeros et al. (2004) showed that pest damage can be detected from spectral reflectance changes in large agricultural field crops like cotton. Correlations between *T. urticae* damage and reduced photosynthetic function of strawberry plants has been documented extensively (Sances et al. 1979, 1981; Kielkiewicz 1985; Iatrou et al. 1995; Reddall et al. 2004). However, quantitative assessment of *T. urticae* infestation on strawberry using spectral information has not been published. If spectral information provides robust diagnosis of infestation, the implications for early detection and spatial pest management are significant. Our objective in this study is to identify regions of the visible/near infrared (VNIR) spectrum that are affected by *T. urticae* infestation and develop statistical assessment of *T. urticae* infestation from foliar spectra. The primary benefit of spectral pest monitoring is rapid diagnosis. With improved diagnostic measures, early intervention is possible, which is likely to decrease economic and ecological costs associated with *T. urticae* infestation (Dayang and Kamaruzaman 1999).

**Materials and Methods**

**Field Plots.** Sixteen 7.3 by 7.3-m with 11-m buffer between plots were planted with strawberry variety Festival at the University of Florida Plant Science Research and Education Unit, near Citra, FL (82.17° W, 29.41° N). Each plot contained six double rows of strawberries each containing transplants 35 cm apart within and between rows. Strawberries were planted in early October (2006) on raised beds over black plastic mulch and fertilized, weeded, and sprayed with fungicides using standard commercial practices (Brown 2003).

**Sampling and Assessment of *T. urticae* Infestation.** Twenty leaflets were collected from each of the experimental plots once per week for 4 wk between December 2006 and January 2007 totaling 80 leaflets. Leaflets were taken randomly from each plot and stored individually in Zipper Seal Storage Bags (American Value; Dolgencorp, Goodlettsville, TN) and transported back to the Small Fruits and Vegetable integrated pest management (IPM) Laboratory at the University of Florida, Gainesville, FL. Leaflets were examined under a dissecting microscope to determine the number of *T. urticae* motiles per leaflet. Each individual leaflet was placed back into its original storage bag, labeled with the number of mites found on the leaflet, and scanned within 2 h after collection before significant dehydration and foliar damage occurs.

**Spectral Scanning and Data Processing.** Leaflet diffuse spectral reflectance was obtained using a Field-spec FR spectro-radiometer (Analytic Spectral Devices, Boulder, CO), with Spectron (LabSphere, Hutton, NH) as a white reference. This instrument has a spectral resolution of 3 nm in the visible spectrum (full width at half-maximum [FWHM]), 10 nm in the NIR, and a range of 350–2,500 nm. Internal interpolation results in spectra with 1-nm resolution across the entire spectra.

Two areas (3.2 cm²) of the adaxial side of each collected leaflet were scanned with a contact probe containing a high temperature tungsten filament lamp, a fiber optic collector oriented at 15° off nadir for collecting diffuse spectra over a 10-mm-diameter spot, and leaf-clip assembly to minimize stray light interference. After the first scan, the leaf was rotated 90° and rescanned; spectral precision was maintained by rescanning leaves that failed to satisfy a <1% error criterion between the first and second scans. White reference scans were obtained every 10 min to control for sensor drift.

High-resolution spectra at 1-nm wavelength intervals were resampled using moving window averaging to 10-nm bands to expedite statistical analysis and to better reflect instrument capabilities. All analyses were performed on both raw relative reflectance and on data following first derivative transformation with second-order smoothing (Fearn 2000); overall, predictions using raw reflectance were more accurate and homoscedastic so only these are reported. Principal components analysis of reflectance spectra was used to visualize the high-dimensional data and to identify spectral outliers before predictive modeling; in this data set, no spectral outliers were detected.

**Data Analysis**

**Spectral Determination of *T. urticae* Infestation Levels.** Processed data were analyzed using partial least squares (PLS) regression (Beebe et al. 1998) to assess the efficacy of spectral prediction of mite numbers. PLs have been widely used to draw chemometric inference from high-dimensional spectral data (Reeves et al. 1999, Cozzolino and Moron 2003). We chose PLS over principle component regression because, in PLS, dimensionality reduction is based on covariance of spectral predictors with response variables (e.g., mite density), whereas in principal components regression, the dimensionality reduction is based on the spectral data alone (Chang et al. 2001). Spectral training was done using 62% of the samples and hold-out validation data using the remaining 38% of the samples. Hold-out validation was done to reduce the influence overfitting caused by data set dimensionality and limited sample size; sample allocation to calibration or validation sets was done randomly. Prediction efficiency was measured using the coefficient of determination (*r²*), root-mean squared error (RMSE), and the relative performance determination (RPD) (Dunn et al. 2002); we computed values for both calibration and validation but focus on validation results. Relative performance determination is the preferred diagnostic because it scales model error by population dispersion (RPD = SD/RMSE); relative performance determination val-
ues exceeding 2.0 are generally considered sufficient for prediction (Chang et al. 2001, Dunn et al. 2002).

**Spectral Determination of Categories of *T. urticae* Infestation.** Because pest management decisions are generally made based on categorical levels of infestation (Pedigo et al. 1986, Bode and Calvin 1990), we examined the spectral separability of practical *T. urticae* categories by dividing the leaves into three categories: low density (0–20 motile mites/leaflet); moderate density (20–50 motile mites/leaflet); and high density (>50 motile mites/leaflet). We used classification tree (CT) analysis to calibrate and validate categorical prediction of mite density class from spectral response data; classification trees are comprised of recursive binary partitions in the data based on predictor variable threshold values (Breiman et al. 1984). Each partition maximizes the purity of the subsequent nodes with respect to the a priori categorical targets (mite infestation). To avoid model overfitting, the calibrated tree was developed using v-fold cross-validation (v = 10), and its prediction efficiency was evaluated using the hold-out validation data (Steinberg and Colla 1997). CT models were developed in the classification and regression tree (CART) software environment (v 5.0; Salford Systems, LA Jolla, CA).

CT model predictions were evaluated using two diagnostic measures. First, overall, producer, and user accuracy defined, respectively, as the total number of correctly classified samples divided by the total number of samples, the number of correct assignments in a category divided by the total number of observations in that category, and the number of correct assignments in a category divided by the total number of assignments to that category. Second, the odds ratio (OR) of correct classification provides a unitless metric of model concordance with observations (Agresti 1990). Models with large statistically significant OR values (i.e., >10.0) can be used for diagnostic purposes; for this three-category model, ORs and associated SEs were computed for each class in relation to the other two. Given the need for early detection, we are particularly interested in the OR values for spectral diagnosis of low mite infestation.

**Results**

**Spectral Data.** The population of samples displayed diagnostic variability most strongly at the “red edge” (∼760 nm) and throughout the short-wave NIR (800–1,300 nm), with slightly less variability in the green visible region (520–580 nm; Fig. 1). Although concordance in spectral pattern is strong across all mite density levels, particularly for the short wave NIR and most of the visible range, differences in both albedo and absorbance feature shape among classes of *T. urticae* infestation are clear.

Principal component analysis condensed raw reflectance data into five principal factors encompassing nearly 98% of the total variance. A bi-plot of the first two principal components (explaining 65.6 and 17.5% of the variance, respectively; Fig. 2) shows qualitatively the global separation of *T. urticae* infestation categories. Spectral waveband loading on these components is shown in Fig. 1 as the absolute value of Pearson’s correlations between component and waveband; PC1 seems to be strongly associated with sample albedo, whereas loading on PC2 indicates strong concordance with regions exhibiting spectral separation between mite density classes.

**Accuracy of *T. urticae* Density Prediction.** The predictions of *T. urticae* density from partial least squares
regression correlated strongly with observed values for the holdout validation data set (Fig. 3). The validation $r^2 = 0.85$, and the RPD (2.75) suggest the potential for prediction of continuous *T. urticae* infestation using spectral methods. The RMSE value is 12.2 mites per leaflet, suggesting low model expected error rates; this rate is comparable to the rate associated with mite detection and enumeration. Predictor importance for the PLS model (Fig. 1) suggests that those areas of the spectrum that are qualitatively useful for discriminating classes of *T. urticae* infestation are selected by the PLS model for quantitative prediction of mite density.

**Fig. 2.** Bi-plot of principal components 1 and 2 from strawberry leaflet spectra at multiple *T. urticae* infestation densities; mean values for an infestation class ($\pm 2$ SE) are shown with enlarged symbols.

**Fig. 3.** Regression of predicted versus observed raw *T. urticae* numbers/leaflet.

**Discussion**

Leaflet spectra seem to be strongly affected by mite density; principal components analysis, an unsupervised classification technique, seems to extract measured categories of infestation relatively well. Moreover, spectral change caused by mite infestation did not occur evenly across the visible and NIR spectrum. Because *T. urticae* feed primarily on cells containing chloroplasts in the spongy mesophyll layer, the observation of maximum change in the green ($\approx 510$ nm) and red ($\approx 700$ nm) regions is expected. However, the
differences in the NIR (800–1,300 nm) may be the result of changes in leaf water content caused by *T. urticae* feeding (Wedding et al. 1958). Discrimination of high infestation density from low and moderate can be achieved in several regions (800–1,300, 1,600–1,500, and 2,200–2,400 nm), but discriminating low from moderate infestation is really only possible in the 800- to 1,300-nm range. Because the objective is to obtain early warning detection (i.e., before high infestation), this spectral range is the most likely to be useful for diagnostic purposes.

Fitzgerald et al. (2004) noted that the strawberry spider mite (*Tetranynchus turkestani* Ugarov and Nikolks) damage in cotton is observable in the 850-nm wavelength, and subsequent work by Fitzgerald et al. (2005) showed that spectral variation in the green region was difficult to detect in the field from commercial satellite data without spectral unmixing. Our findings suggest that ground-based sensors with finer spectral resolution and less atmospheric interference can detect *T. urticae* with accuracy and precision in both the green bands and elsewhere in the visible and near infrared portions of the spectrum. Further studies are needed to validate our findings under field conditions, particularly given the field portability of the instrument used for this work.

Strawberry plants are most susceptible to dehydration and nutrient deficiency associated with *T. urticae* damage during critical vegetative growth period early in the season (Oatman and Voth 1972, Sances et al. 1981, Rhodes et al. 2006). Early detection of *T. urticae* is essential in achieving control as they have a high level of fecundity and a rapid life cycle. At temperatures >35°C a female can lay as many as 20 eggs per day and the lifecycle can be as short as 7 d (Shanks and Doss 1959, Grostal and Dicke 2000, White and Liburd 2004). When infestation reaches a high level, chemical and biological controls are unable to reduce *T. urticae* populations below the economic threshold, which in strawberries is considered to be between 5 and 20 motile mites per leaflet (Oatman and Voth 1972, Harman et al. 2005). Laboratory and field experiments show that effective biological control agents must be released early in the season when there is a low incidence of spider mite infestation (Greco et al. 2005, Fraulo and Liburd 2007).

The ultimate objective of our study was to detect *T. urticae* in strawberry leaflets before infestations are high enough to preclude management and before physiological damage is visible to the human eye. Currently, growers monitor their fields by manually scouting with a hand lens to inspect individual leaflets for *T. urticae* presence. Using VNIR spectroscopy allows efficient and effective detection. Large scale technology in conjunction with infrared sensor systems is already popular for detection of agricultural pests (Zhang et al. 2002, Fitzgerald et al. 2005, Reisig and Godfrey 2007). This technological method enables growers to detect and monitor growing pest populations and identify “hot-spots” for precision pest management. Several options exist for integrating this technology. A spectrometer can be mounted on a tractor or any GPS enabled field equipment. With fore-optics with a conical field of view of 25°, the image is enlarged with respect to the height of the sensor, allowing for deployment of as high as 20 m above the target with no detectable effect of signal to noise ratio. Similar technology has been used extensively in soil and nutrient assessment in agricultural fields (Zhang et al. 2002, Pinter et al. 2003, Pydipati et al. 2005, Shaw 2005), so integration into pest management protocols is feasible.

### References Cited


Dayang, A. I., and J. Kamartuzaman. 1999. Geospatial information technologies for Malaysian agriculture in the next millennium. Seminar on Repositioning Agriculture In-

### Table 1. Summary of classification tree model of mite density categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Observed</th>
<th>User accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Calibration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted</td>
<td>Low</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Producer</td>
<td>100%</td>
</tr>
<tr>
<td>Validation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted</td>
<td>Low</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Producer</td>
<td>83%</td>
</tr>
<tr>
<td>Odds ratio (low category)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odds ratio (mod. category)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odds ratio (high category)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Odds ratios are computed for validation data only.