

## Visible-near infrared spectroscopy based citrus greening detection: Evaluation of spectral feature extraction techniques

Sindhuja Sankaran\*, Reza Ehsani

Citrus Research and Education Center/IFAS, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA

### ARTICLE INFO

**Article history:**

Received 3 May 2011

Received in revised form

13 June 2011

Accepted 9 July 2011

**Keywords:**

Citrus greening

Spectral band selection

Vegetation indices

Feature extraction

Visible-near infrared spectroscopy

### ABSTRACT

Citrus greening is a serious disease affecting citrus production in Florida and different parts of the world. This disease is spread by an insect vector and the trees are killed several years after infection. There is no known treatment for the disease. Disease detection and removal of infected trees is a critical part of citrus greening disease management efforts. This paper reports the evaluation of spectral features extracted from visible-near infrared spectroradiometer spectra for their potential to detect citrus greening disease. The extraction of spectral features is an effort to lower the cost of the optical sensor while maintaining their performance. Spectral features: (i) spectral reflectance bands and (ii) vegetation indices (VIs) were derived from 350–2,500 nm spectral reflectance data using two feature extraction methods: stepwise discriminant analysis and stepwise regression analysis. Following the selection of spectral features, the features were assessed using two classifiers, quadratic discriminant analysis (QDA) and soft independent modeling of classification analogies (SIMCA) to determine the overall and individual class classification accuracies. The classification results indicated that both the spectral features (spectral bands and VIs) yielded good overall (higher than 80%) and healthy class (higher than 85%) classification accuracies using the QDA-based algorithm. The SIMCA-based algorithm yielded good average citrus greening class classification accuracy (higher than 83%) using selected spectral features. Thus, the present study demonstrates the applicability of utilizing spectral features for detection of greening in citrus.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

*Candidatus Liberibacter asiaticus* (CLas), the phloem-limiting bacterium is the putative causal agent of citrus greening or Huan-glongbing (HLB), a devastating citrus disease. HLB is spread by a vector, the Asian citrus psyllid, *Diaphorina citri* Kuwayama. The vector was first found in Florida in 1998, followed by the discovery of the disease in 2005 (Li et al., 2006; Gottwald, 2010; Hawkins et al., 2010a). Since 2005, citrus greening has spread to all citrus-producing counties in Florida, reducing citrus productivity and affecting the economics of citrus orchard management. This disease is known to infect all cultivars of citrus, ultimately resulting in the death of the trees. The HLB-associated bacteria, i.e. the CLas, disrupt the carbohydrate (food) transfer to the other regions of the tree, resulting in twig die-back and tree decline. Symptoms such as leaf chlorosis, blotchy mottle, and asymmetric fruits appears about six months to two years after the onset of disease. The juice from the

infected tree fruit is bitter in taste and acidic in flavor (Gottwald, 2010), resulting in very poor fruit juice quality. There is no known treatment for citrus greening.

Some of the citrus greening management practices recommended by the researchers for orchards with low-infection levels are inoculum reduction through disease detection and tree removal, and psyllid control through pesticide application (Manjunath et al., 2008; Gottwald, 2010; Spann et al., 2010). Some of the issues associated with these greening management practices are the increased production costs. For example, orange production costs in Southwest Florida have increased from about \$3,079/ha to \$4,448/ha in 2008 (Muraro and Morris, 2009). With about 2,225 sq. km of citrus production in Florida, the estimated increase in production costs cannot be ignored. The U.S. citrus industries are taking great efforts to control and contain the disease, and to arrive at economical solutions to reduce the impact of this disease. In citrus orchards with low-infection rates, one of the major challenges in controlling the spread of the disease is that the infected trees serve as possible internal source of inoculum, resulting in further spread of the disease (Tatineni et al., 2008; Gottwald, 2010). In addition, external source of inoculum (HLB-infected trees)

\* Corresponding author. Tel.: +1 863 956 1151x1290; fax: +1 863 956 4631.

E-mail address: [sindhu@ufl.edu](mailto:sindhu@ufl.edu) (S. Sankaran).

resulting from abandoned/unmanaged orchards, residential sites or outdoor nurseries also affects the spread of disease. While the external source control largely depends on the regulation policies, the internal source can be controlled by the grower (Manjunath et al., 2008).

Disease detection is critical for the management of citrus greening. Currently, manual scouting is widely used for detecting greening symptoms. During the scouting process, a team of trained inspectors pass along the rows of citrus orchards either by walking or in mobile vehicles such as all-terrain vehicles and elevated platforms attached to vehicles to look for symptoms indicating citrus greening. The HLB-infected trees could be identified with an average detection rate of about 47–59% (Futch et al., 2009). Some of the challenges during scouting are: uneven distribution of disease symptoms within a tree, effect of nutritional deficiencies, environmental factors and other stress conditions resulting in similar symptoms (Futch et al., 2009). Although there are laboratory-based polymerase chain reaction (PCR) methods that can accurately detect greening (Teixeira et al., 2005; Li et al., 2006, 2009; Lacava et al., 2006), there is a need for a sensing or screening system that would provide reliable detection of citrus greening symptoms in real-time and under field conditions. Since PCR technique is expensive and time-consuming (Hawkins et al., 2010a), having a prescreening technique that can identify the infected trees will reduce the demand for scouting. Also it will reduce the number of samples for PCR test which could result in lower disease management costs and more effective disease detection.

Visible-near infrared spectroscopy provides a rapid, non-invasive sensing tool for disease detection in plants (Sankaran et al., 2010b). Visible-near infrared spectroscopy has been applied in varied applications, including detection of biotic and abiotic stresses in plants (Polischuk et al., 1997; Spinelli et al., 2006; Naidu et al., 2009; Sankaran et al., 2011). Some of the spectroscopic techniques such as visible-near infrared spectroscopy, hyperspectral imaging, and fluorescence spectroscopy have been applied to leaves and fruit peel for detection of citrus canker (Belasque et al., 2008; Qin et al., 2008, 2009; Balasundaram et al., 2009) under laboratory and greenhouse conditions. Other spectroscopic methods such as mid-infrared spectroscopy have also been applied for detecting greening in citrus leaves (Hawkins et al., 2010a, 2010b; Sankaran et al., 2010a). Balasundaram et al. (2009) studied the application of visible-near infrared spectroscopy in detecting citrus canker and other similar disease conditions in the peel of the grapefruit under laboratory conditions. Similarly, Qin et al. (2009) studied hyperspectral imaging for detecting canker and other diseases or stresses such as greasy spot, insect damage, melanose, scab and wind scar in grapefruits. A spectral information divergence (SID) method was applied to classify the spectral signature (450–930 nm) of the fruits with overall classification accuracy of 95–96%. Although some studies on citrus disease detection using visible-near infrared spectroscopy have been conducted, further research is essential to explore the applicability of visible-near infrared spectroscopy for the detection of citrus greening.

Our previous study (Sankaran et al., 2011) demonstrated the potential of visible-near infrared spectral signatures in the range of 350–2,500 nm for detection of citrus greening in leaves. When features extracted from the entire visible-near infrared spectra were used for classification, it was found that among the studied classifiers, quadratic discriminant analysis (QDA) and soft independent modeling of classification analogies (SIMCA) yielded good HLB-class classification accuracies of higher than 89% and 93%, respectively. However, in order to develop a simple and portable sensor, it is important to understand the characteristics of the spectral reflectance signature.

The present study focuses on the selection of spectral features specific to citrus greening, which will serve as a fundamental

knowledge for detecting citrus greening. Selection of robust spectral features will also enable the development of a rugged, cost-effective sensor system. The major objectives of this research were: (i) to identify spectral reflectance bands and evaluate these bands for their ability to classify citrus greening, and (ii) to select vegetation indices (VIs) and evaluate their ability to discriminate citrus greening using two classification algorithms. First, stepwise discriminant analysis and stepwise regression analysis were used for the selection of spectral features based on their discriminatory power. Then, the two classification algorithms, QDA and SIMCA were applied on the selected spectral features extracted from the spectral signature of healthy and diseased leaves under laboratory and field conditions. The overall and individual class classification accuracies of various tests have been reported in this paper.

## 2. Materials and methods

### 2.1. Data collection

A field portable spectroradiometer (SVC HR-1024, Spectra Vista Cooperation, NY) with a 4° field of view (Fig. 1) was used to collect the spectral reflectance data in the range of 350 to 2,500 nm. The reference spectra were acquired using a white panel (Spectralon Reflectance Target, CSTM-SRT-99-100, Spectra Vista Cooperation, NY). The spectral reflectance data were collected from healthy and diseased leaves under field and laboratory conditions. The citrus leaves were from a 21.8 sq. km citrus orchard (Devil's Garden, Southern Gardens, Clewiston, FL). The field data collection procedure used is described in Sankaran et al. (2011).

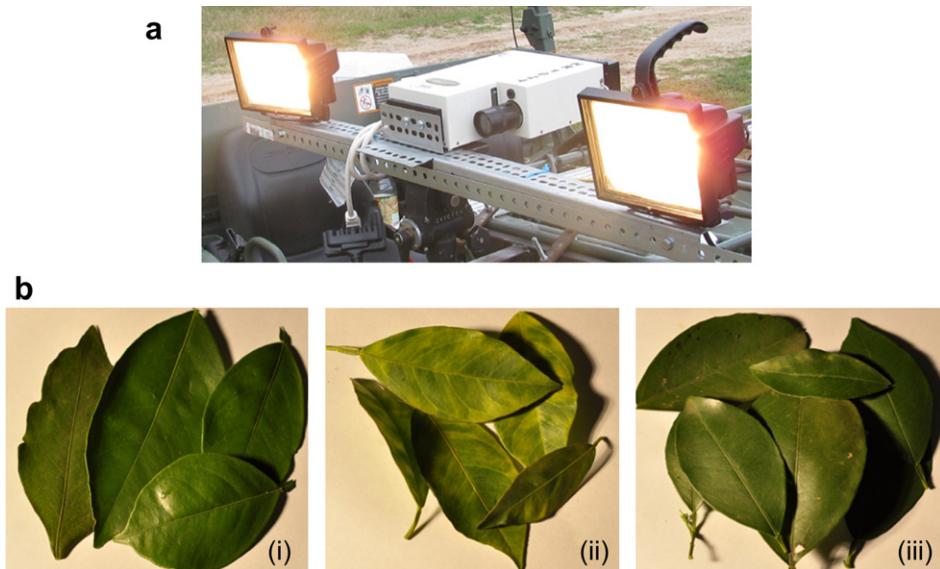
For the laboratory data collection, a portable halogen work light (500 W) was used as a light source. The sensor was placed on a stand at about 0.6 m away from the leaf samples. Five spectra were collected from each sample. In addition to the healthy and symptomatic diseased leaves, data were collected from non-symptomatic leaves from trees known to be infected with citrus greening. The non-symptomatic leaves were from the same branch as the HLB-infected symptomatic leaves. All the leaves were tested for HLB-related bacteria using the PCR method. A sequence of primers and probes were used for the amplification of target gene 16S rDNA. The primer–probe combination used for the detection of CLas was HLBaspr. COXfpr was used for positive internal control to confirm the DNA extraction quality (Li et al., 2006). The disease was assessed based on the threshold cycle ( $C_t$ ) values acquired from PCR analysis. A  $C_t$  value of 30 and lower indicates the presence of bacteria and the sample was marked “PCR positive”; while a value of 32 and higher indicates the absence of bacteria and sample was marked “PCR negative”.

The spectral reflectance data collected in the field and in the laboratory are henceforth referred as Dataset-I and -II, respectively. The two datasets were combined as Dataset-III. The number of samples in each of the datasets is summarized in Table 1. The datasets were preprocessed prior to spectral feature extraction using customized statistical algorithm written in MATLAB® (ver. 7.6, The MathWorks Inc., Natick, MA). The five spectral replicates were averaged for each sample and the averaged spectra were normalized based on Euclidean norm (Sankaran et al., 2011). The spectral reflectance data were then averaged for every 25 nm, finally resulting in 86 spectral reflectance data for each sample.

### 2.2. Feature extraction

#### 2.2.1. Selection of spectral reflectance bands

Two different feature extraction methods, stepwise discriminant analysis (SAS®9.2, SAS Institute Inc., Cary, NC, USA) and stepwise regression analysis (MATLAB® 7.6) were used for selecting



**Fig. 1.** (a): Sensor set-up showing spectroradiometer and light source. (b): Healthy (i), symptomatic (ii), and non-symptomatic (iii) diseased leaves.

the useful spectral reflectance bands from 86 spectral reflectance bands. In stepwise discriminant analysis (*proc stepdisc*), a model was built with the addition of one spectral band (variable) at each step. The model was examined at each step to verify whether the added spectral band contributes toward discriminating the classes. Finally, the spectral bands (variables) were selected based on the level of significance (5%) of an *F*-test through analysis of covariance and squared partial correlation. In stepwise regression analysis (*stepwisefit*), spectral bands (variables) were added or removed from a multilinear model based on their statistical significance in a regression. Unlike stepwise discriminant analysis, in stepwise regression analysis, an initial model was built and based on the computed *F*-statistics with and without the chosen spectral band; the chosen spectral band is added or removed.

The two feature extraction methods were applied to identify the spectral bands from each of the three datasets. After the extraction process, the spectral bands that were identified as important spectral bands using both the methods, i.e. the common spectral bands, were selected from each dataset, which were used in further analysis. The set of selected spectral reflectance bands from Dataset-I, -II and -III are labeled as SB-1, SB-2 and SB-3, respectively. In addition to these sets of spectral bands, a set of new spectral bands (SB-4) was identified such that the number of spectral bands was limited to six, and the wavelength was restricted to 1,500, and evaluated.

## 2.2.2. Selection of vegetation indices

In addition to the spectral reflectance bands, twenty-two vegetation indices (VIs) representing plant health were obtained. The vegetation indices computed were: normalized difference vegetation index (NDVI), renormalized difference vegetation index

(RDVI), modified simple ratio (MSR), triangular vegetation index (TVI), modified chlorophyll absorption ratio index (MCARI), modified chlorophyll absorption ratio index (MCARI1 and MCARI2), modified triangular vegetation index (MTVI1 and MTVI2) (Haboudane et al., 2004); transformed chlorophyll absorption ratio index (TCARI), structure intensive pigment index (SIPPI), Greenness index (GI), simple ratio pigment index (SRPI), photochemical reflectance index (PRI and PRI1), normalized pigment chlorophyll index (NPCI and NPCI1) (Zarco-Tejada et al., 2005); simple ratio index (SRI) (Stenberg et al., 2004); Enhanced vegetation index (EVI) (Nagler et al., 2005; Jiang et al., 2008); Red edge shift (RES) (Carroll et al., 2008); and Nitrogen reflectance index (NRI), Plant senescence reflectance index (PSRI) (Devadas et al., 2009). The spectral bands that were used for estimating VIs were blue, green, red, and infrared region were 450 nm, 550 nm, 670 nm and 800 nm, respectively. Computation of some vegetation indices such as MCARI, TCARI and NPCI1 used red spectral band of 700 nm wavelength; while PRI and PRI-1 used green bands 531 nm, 570 nm and 550 nm, 570 nm, respectively. The VIs indicate the physiological status of the plant such as chlorophyll variation, leaf cellular structure, and leaf green area, among others. After the computation of vegetation indices, the important vegetation indices were identified using stepwise discriminant and stepwise regression analysis. The VIs that were not selected in any of the three datasets using both methods during the feature extraction process were eliminated, and the remaining VIs were used as inputs in the classification algorithms.

## 2.3. Classifiers

After the spectral feature extraction, the selected features were evaluated using classification algorithms. In our previous study, it was found that the classifiers, QDA and SIMCA resulted in high overall classification accuracies (higher than 80%) when features from the entire visible-near infrared spectra were used for classification (Sankaran et al., 2011). Therefore, in this study, QDA and SIMCA-based algorithms were used to assess the selected spectral features. The selected spectral features were extracted from each of the three datasets, randomized, separated into training and testing datasets, and classification accuracies were computed using QDA and SIMCA. The ratio of training to testing datasets was 75:25. The ratio of training: testing dataset samples in Dataset-I, Dataset-II,

**Table 1**  
Details of the datasets used in this study.

Dataset	Healthy/Symptomatic HLB-infected trees			Non-symptomatic (HLB-infected tree)		
	Healthy	Diseased	Total	Healthy (PCR negative)	Diseased (PCR positive)	Total
Dataset-I	100	93	193	48	24	72
Dataset-II	98	93	191	48	24	72
Dataset-III	198	186	384	96	48	144

**Table 2**

Results from stepwise discriminant analysis (sDA) and stepwise regression analysis (sRA).

Dataset	Feature extraction method	Wavelength (nm)
Dataset-I	sDA	537, 612, 688, 713, 763, 998, 1066, 1120, 1148, 1296, 1472, 1546, 1597, 1647, 1822, 1873, 1898, 2073, 2121, 2172, 2296, 2348, 2493
	sRA	537, 612, 688, 713, 763, 998, 1066, 1120, 1148, 1296, 1472, 1524, 1597, 1672, 1923, 2023, 2121, 2172, 2348, 2398, 2493
Dataset-II	sDA	437, 513, 638, 662, 713, 813, 963, 1222, 1346, 1445, 1524, 1622, 1696, 1848, 2471
	sRA	638, 662, 713, 813, 1445, 1622, 2172, 2471
Dataset-III	sDA	413, 537, 638, 662, 713, 763, 963, 1245, 1272, 1546, 1597, 1672, 1746, 1898
	sRA	437, 537, 563, 688, 713, 738, 1120, 1191, 1546, 1597, 1647, 1746, 1898, 2147, 2172, 2422, 2493

**Table 3**

Selective spectral bands used for the data analysis.

Set of spectral bands	Wavelengths (nm)
Spectralband-1* (SB-1)	537, 612, 688, 713, 763, 998, 1066, 1120, 1148, 1296, 1472, 1597, 2121, 2172, 2348, 2493
Spectralband-2† (SB-2)	638, 662, 713, 813, 1445, 1622, 2471
Spectralband-3‡ (SB-3)	537, 713, 1546, 1597, 1746, 1898
Spectralband-4# (SB-4)	537, 662, 713, 813, 1120, 1472

Note: \*derived from Dataset-I, †derived from Dataset-II, ‡derived from Dataset-III, #derived from feature selection results.

and Dataset-III was 145:48, 143:48, and 288:96, respectively. The classifier was trained and tested five times on each dataset. This was done by randomizing the feature dataset and then drawing train and test datasets. Thus, the reported overall and individual class classification accuracies are an average of five test runs on each dataset. Hereafter, the healthy and diseased (HLB) leaf classes will be referred as class-1 and class-2, respectively. In addition to testing the three datasets, the spectral features extracted from the non-symptomatic datasets were also evaluated to determine the classification accuracies.

### 3. Results

#### 3.1. Spectral reflectance bands

The spectral reflectance bands exhibiting discriminatory power, which were determined using stepwise discriminant analysis and stepwise regression analysis, are summarized in Table 2. Table 3 summarizes the three sets of selected spectral bands (SB-1 to SB-3) derived from each dataset. These three sets primarily consisted of few bands from the green (537 nm) and red region (630–760 nm) of the visible spectra, and many from the near-infrared region (with two to six bands having wavelengths greater than 1,500 nm) of the electromagnetic spectra. Therefore, a fourth set of spectral bands (SB-4) was identified using the selected spectral bands from all the datasets, with six bands of wavelengths less than 1,500 nm. Balasundaram et al. (2009) reported that the spectral regions between 500 and 800 nm exhibited the maximum discriminatory power with upper threshold of useful wavelengths up to 1,100 nm for detecting canker in citrus peel. However, in this study, in addition to the visible regions, near-infrared regions with wavelengths higher than 1,100 nm contributed to the discriminatory power.

The selected spectral band sets were evaluated for their ability to detect greening in citrus leaves. The average overall and individual class classification accuracies are summarized in Table 4. In general, the QDA-based algorithm resulted in higher average overall and class-1 classification accuracy than those of SIMCA. On the other hand, the SIMCA performed better than the QDA, comparing the average class-2 classification accuracy using all four spectral bands sets, with the exception being SB-1 while classifying Dataset-II. A similar trend was observed in our previous study, when the features from the entire spectral reflectance data (350–2,500 nm) were used as input features for classification (Sankaran et al., 2011).

Among the four spectral bands sets (SB-1 to SB-4), in most cases, classifiers using the selected spectral bands from SB-1 set resulted in higher overall and individual class classification accuracies than other sets. This trend could be anticipated as the number of spectral bands representing SB-1 set (16 spectral reflectance bands) was much higher than the other sets (6–7 spectral reflectance bands). Comparing this aspect, six spectral bands in the SB-4 set having a wavelength of less than 1,500 nm were able to yield an average overall classification accuracy of about 84–87% and 76–81% using QDA- and SIMCA-based algorithms, respectively. A high average overall and class-1 classification accuracy of 84–91% was achieved using the QDA-based algorithm; while high average class-2 classification accuracy of 87–95% was achieved using the SIMCA-based algorithm. Compared to the average scouting efficiency of 47–61% (Futch et al., 2009) for greening detection, these classification accuracies are promising.

#### 3.2. Vegetation indices

Based on the feature selection methods, the selected vegetation indices (15 VIs) were as follows: RDVI, MSR, MCARI, MCARI1, MCARI2, TVI, SIPI, GI, SRPI, PRI, PRI1, NPCI, NPCI1, RES, and Plant

**Table 4**

Average overall and individual class classification accuracies using selected spectral bands.

Class	Classification accuracy (%)					
	Dataset-I		Dataset-II		Dataset-III	
	QDA	SIMCA	QDA	SIMCA	QDA	SIMCA
<b>SB-1</b>						
Overall	89 ± 6	82 ± 4	90 ± 4	83 ± 5	91 ± 1	83 ± 2
Class-1 (Healthy)	91 ± 9	71 ± 9	89 ± 11	82 ± 9	93 ± 4	71 ± 4
Class-2 (HLB)	87 ± 5	95 ± 6	93 ± 5	83 ± 4	88 ± 4	93 ± 2
<b>SB-2</b>						
Overall	85 ± 7	74 ± 7	87 ± 4	83 ± 6	86 ± 1	78 ± 2
Class-1 (Healthy)	87 ± 7	54 ± 12	87 ± 8	79 ± 10	90 ± 4	60 ± 7
Class-2 (HLB)	81 ± 14	98 ± 2	86 ± 10	87 ± 11	81 ± 3	95 ± 3
<b>SB-3</b>						
Overall	90 ± 3	86 ± 6	81 ± 3	77 ± 6	88 ± 1	82 ± 2
Class-1 (Healthy)	89 ± 5	78 ± 10	85 ± 8	71 ± 12	89 ± 5	68 ± 8
Class-2 (HLB)	91 ± 6	96 ± 4	78 ± 5	83 ± 5	86 ± 6	94 ± 4
<b>SB-4</b>						
Overall	84 ± 6	76 ± 8	84 ± 5	78 ± 5	87 ± 4	81 ± 3
Class-1 (Healthy)	91 ± 8	61 ± 17	85 ± 7	71 ± 11	90 ± 5	68 ± 10
Class-2 (HLB)	75 ± 12	95 ± 8	83 ± 2	87 ± 4	84 ± 5	93 ± 5

**Table 5**

Average classification accuracies using selected vegetation indices as input features.

Class	Classification accuracy (%)					
	Dataset-I		Dataset-II		Dataset-III	
	QDA	SIMCA	QDA	SIMCA	QDA	SIMCA
Overall	83 ± 5	68 ± 5	83 ± 7	84 ± 2	80 ± 4	75 ± 4
Class-1 (Healthy)	88 ± 7	45 ± 12	92 ± 3	85 ± 3	89 ± 6	59 ± 8
Class-2 (HLB)	77 ± 3	96 ± 4	73 ± 11	84 ± 5	70 ± 7	92 ± 3

**Table 6**

Average classification accuracies achieved using Dataset-III as the training dataset and non-symptomatic dataset as the testing dataset.

Class	Classification accuracy (%)					
	Dataset-I		Dataset-II		Dataset-III	
	QDA	SIMCA	QDA	SIMCA	QDA	SIMCA
<i>Spectral bands (SB-4)</i>						
Overall	53 ± 2	41 ± 1	61 ± 2	41 ± 3	57 ± 1	41 ± 2
Class-1 (Healthy)	50 ± 3	25 ± 3	68 ± 3	21 ± 6	59 ± 2	23 ± 4
Class-2 (HLB)	60 ± 4	74 ± 3	49 ± 3	81 ± 6	55 ± 3	78 ± 3
<i>Vegetation Indices</i>						
Overall	49 ± 2	43 ± 2	67 ± 1	60 ± 3	58 ± 2	51 ± 1
Class-1 (Healthy)	49 ± 3	25 ± 3	76 ± 1	52 ± 5	62 ± 2	36 ± 3
Class-2 (HLB)	51 ± 2	79 ± 6	50 ± 0	78 ± 2	52 ± 3	82 ± 5

senescence reflectance index (PSRI). The average classification accuracies using VIs as spectral features are reported in Table 5. QDA resulted in an average overall classification accuracy of about 80–83%, with average class-1 classification accuracy of about 88–92%. Similar to the spectral band features, the SIMCA-based algorithm yielded higher average class-2 classification accuracies (84–96%). The selected vegetation indices also showed some potential in greening detection in citrus.

#### 4. Discussion

The selected spectral features were evaluated using non-symptomatic datasets and results are reported in Table 6. Both the sets of spectral features (SB-4 bands and VIs) yielded average overall classification accuracies in the range 41–67%. In general, the SIMCA-based overall and class-1 classification accuracies

were lower than those of the QDA-based algorithm. However, the average class-2 classification accuracies of SIMCA were higher (74–82%) than those of QDA. The average class-2 classification accuracies in this study were not as high (90%) as in our previous study (Sankaran et al., 2011) when the features from the entire spectra were used to classify non-symptomatic samples. The probable reason for this could be that spectral reflectance bands other than the ones selected in this study may better represent the non-symptomatic condition. Considering the limitation of visual inspection of disease symptoms, the classification accuracy was encouraging for the further exploration visible-near infrared spectroscopy for detecting non-symptomatic diseased samples.

The overall and individual class classification accuracies, determined using QDA and SIMCA, acquired from the entire spectral data features and selected spectral features were compared. The results presented in Fig. 2 are obtained from analysis of the field dataset. Comparing the QDA-based classification accuracies, the overall and class-2 classification accuracies were lower using spectral features than those using the entire spectra. When comparing the SIMCA-based classification accuracy, although the spectral features yielded low overall and class-1 classification accuracies, the class-2 classification accuracies were comparable to that of entire spectral feature dataset.

Analysis of variance (ANOVA) was performed using the SAS® program at 5% level of significance to compare the classification accuracies derived from different spectral features. During ANOVA, Duncan's multiple range test (DMRT) was also performed to compare the treatment means while controlling the comparison-wise error rate. ANOVA revealed that the class-2 classification accuracies of QDA-based accuracies, and overall and class-1 classification accuracies of SIMCA-based accuracies were not equal. Further evaluation of the classification accuracies using DMRT revealed that among the three datasets (entire spectral reflectance data features, selected spectral bands and selected VIs), the classification accuracies resulting from entire spectral reflectance data were different. Although the vegetation indices showed some potential to retain some of the discriminating power to classify diseased from healthy leaves as that of spectral bands, further data collection and analysis to confirm whether the VIs represented the overall stress in plants or diseased condition.

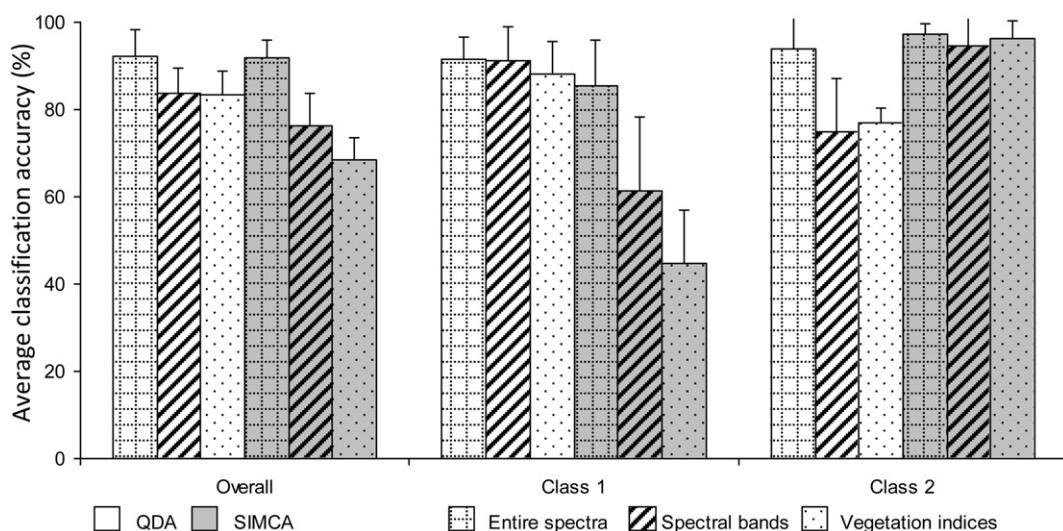


Fig. 2. Comparison of average QDA- and SIMCA-based classification accuracies achieved using spectral features and entire spectral reflectance data.

## 5. Conclusions

The present study evaluates the applicability of using features extracted from 350 to 2,500 nm spectroradiometer reflectance data for detecting citrus greening in leaves. Few spectral features will enhance the disease detection process by lowering the cost of the sensor, increasing the speed of detection, and enhancing the portability of the sensor. Therefore, two different statistical tools: stepwise discriminant analysis and stepwise regression analysis were used for selecting the spectral features with some discriminatory power while representing the diseased and healthy leaf classes. It was found that the selected spectral bands with a few visible bands (537, 662, 713 nm) and near-infrared bands (813, 1120, 1472 nm) resulted in an average overall classification accuracy of about 84–87%, with an average diseased class classification accuracy of about 75–84% using a QDA-based algorithm. Analysis of selected vegetation indices also resulted in a good average overall classification accuracies (80–83%), although the average diseased class classification accuracies were lower (70–77%) than spectral band features. Thus, the current study demonstrates the applicability of utilizing spectral features for detection of greening in citrus.

Future work will involve further evaluation of spectral features representative of other citrus stress conditions such as nutrient deficiency, including analysis of non-symptomatic diseased leaves. Further studies will be performed with a larger number of data for validating the results.

## Acknowledgments

The authors would like to thank the United States Department of Agriculture (USDA) – National Institute of Food and Agriculture (NIFA) and the Citrus Research and Development Foundation (CRDF) for their funding for this research. We would like to express our gratitude to Dr. Joao Camargo Neto, Dr. Joe Mari Maja, Dr. Jose Gonzalez-Mora, and Ms. Sherrie Buchanon. We would also like to thank Mr. Tim Gast and Mr. Michael Irey from U.S. Sugar Corporation and Southern Gardens Citrus for their assistance during HLB scouting and PCR analysis.

## References

- Balasundaram, D., Burks, T.F., Bulanon, D.M., Schubert, T., Lee, W.S., 2009. Spectral reflectance characteristics of citrus canker and other peel conditions of grapefruit. *Postharvest Biol. Technol.* 51, 220–226.
- Belasque, L., Gasparoto, M.C.G., Marcassa, L.G., 2008. Detection of mechanical and disease stresses in citrus plants by fluorescence spectroscopy. *Appl. Opt.* 47, 1922–1926.
- Carroll, M.W., Glaser, J.A., Hellmich, R.L., Hunt, T.E., Sappington, T.W., Calvin, D., Copenhaver, K., Fridgen, J., 2008. Use of spectral vegetation indices derived from airborne hyperspectral imagery for detection of European Corn Borer infestation in Iowa corn plots. *J. Econ. Entomol.* 101, 1614–1623.
- Devadas, R., Lamb, D.W., Simpfendorfer, S., Backhouse, D., 2009. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precis. Agric.* 10, 459–470.
- Futch, S., Weingarten, S., Irey, M., 2009. Determining HLB infection levels using multiple survey methods in Florida citrus. *Proc. Fla. State Hortic. Soc. (FSHS)* 122, 152–158.
- Gottwald, T.R., 2010. Current epidemiological understanding of citrus Huanglongbing. *Annu. Rev. Phytopathol.* 48, 119–139.
- Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90, 337–352.
- Hawkins, S.A., Park, B., Poole, G.H., Gottwald, T., Windham, W.R., Lawrence, K.C., 2010a. Detection of citrus Huanglongbing by Fourier Transform infrared-attenuated total reflection (FTIR-ATR) spectroscopy. *J. Appl. Spectrosc.* 64, 100–103.
- Hawkins, S.A., Park, B., Poole, G.H., Gottwald, T., Windham, W.R., Albano, J., Lawrence, K.C., 2010b. Comparison of FTIR spectra between Huanglongbing (citrus greening) and other citrus maladies. *J. Agric. Food Chem.* 58, 6007–6010.
- Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* 112, 3833–3845.
- Lacava, P.T., Li, W.B., Araújo, W.L., Azevedo, J.L., Hartung, J.S., 2006. Rapid, specific and quantitative assays for the detection of the endophytic bacterium *Methylobacter mesophilicum* in plants. *J. Microbiol. Methods* 65, 535–541.
- Li, W., Abad, J.A., French-Monar, R.D., Rascoe, J., Wen, A., Gudmestad, N.C., Secor, G.A., Lee, I.M., Duan, Y., Levy, L., 2009. Multiplex real-time PCR for detection, identification and quantification of '*Candidatus Liberibacter solanacearum*' in potato plants with zebra chip. *J. Microbiol. Methods* 78, 59–65.
- Li, W., Hartung, J.S., Levy, L., 2006. Quantitative real-time PCR for detection and identification of *Candidatus Liberibacter* species associated with citrus Huanglongbing. *J. Microbiol. Methods* 66, 104–115.
- Manjunath, K.L., Halbert, S.E., Ramadugu, C., Webb, S., Lee, R.F., 2008. Detection of '*Candidatus Liberibacter asiaticus*' in *Diaphorina citri* and its importance in the management of citrus huanglongbing in Florida. *Phytopathol.* 98, 387–396.
- Muraro, R.P., Morris, R.A., 2009. The Dynamics and Implications of Recent Increases In Citrus Production Costs. Fact sheet FE-793. Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida.
- Nagler, P.L., Scott, R.L., Westenborg, C., Cleverly, J.R., Glenn, E.P., Huete, A.R., 2005. Evapotranspiration on western U.S. rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers. *Remote Sens. Environ.* 97, 337–351.
- Naidu, R.A., Perry, E.M., Pierce, F.J., Mekuria, T., 2009. The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two red-berried wine grape cultivars. *Comput. Electron. Agric.* 66, 38–45.
- Polischuk, V.P., Shadchina, T.M., Kompanetz, T.I., Budzanivskaya, I.G., Boyko, A.L., Sozinov, A.A., 1997. Changes in reflectance spectrum characteristic of *Nicotiana debneyi* plant under the influence of viral infection. *Archiv. Phytopathol. Plant Prot.* 31, 115–119.
- Qin, J., Burks, T.F., Kim, M.S., Chao, K., Ritenour, M.A., 2008. Citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method. *Sens. Instrum. Food Qual. Saf.* 2, 168–177.
- Qin, J., Burks, T.F., Ritenour, M.A., Bonn, W.G., 2009. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. *J. Food Eng.* 93, 183–191.
- Sankaran, S., Mishra, A., Ehsani, R., Davis, C., 2010a. A review of advanced techniques for detecting plant diseases. *Comput. Electron. Agric.* 72, 1–13.
- Sankaran, S., Ehsani, R., Etxeberria, E., 2010b. Mid-infrared spectroscopy for detection of Huanglongbing (greening) in citrus leaves. *Talanta* 83, 574–581.
- Sankaran, S., Mishra, A., Maja, J.M., Ehsani, R., 2011. Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. *Comput. Electron. Agric.* 77, 127–134.
- Spann, T.A., Atwood, R.A., Dewdney, M.M., Ebel, R.C., Ehsani, R., England, G., Futch, S.H., Gaver, T., Hurner, T., Oswalt, C., Rogers, M.E., Roka, F.M., Ritenour, M.A., Zekri, M., Boman, B.J., Chung, K.-R., Danyluk, M.D., Goodrich-Schneider, R., Morgan, K.T., Morris, R.A., Muraro, R.P., Roberts, P., Rouse, R.E., Schumann, A.W., Stansly, P.A., Stelinski, L.L., 2010. IFAS guidance for Huanglongbing (greening) management. Publication no. HS1165. Horticultural Sciences Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida.
- Spinelli, F., Noferini, M., Costa, G., 2006. Near infrared spectroscopy (NIRs): perspective of fire blight detection in asymptomatic plant material. Proceeding of 10th International Workshop on Fire Blight. *Acta Hortic.* 704, 87–90.
- Stenberg, P., Rautiainen, M., Manninen, T., Voipio, P., Smolander, H., 2004. Reduced simple ratio better than NDVI for estimating LAI in Finnish pine and spruce stands. *Silva Fenn.* 38, 3–14.
- Teixeira, D.C., Danet, J.L., Eveillard, S., Martins, E.C., Junior, W.C.J., Yamamoto, P.T., Lopes, S.A., Bassanezi, R.B., Ayres, A.J., Saillard, C., Bové, J.M., 2005. Citrus huanglongbing in São Paulo State, Brazil: PCR detection of the '*Candidatus*' *Liberibacter* species associated with the disease. *Mol. Cell. Probes* 19, 173–179.
- Tatineni, S., Sagaram, U.S., Gowda, S., Robertson, C.J., Dawson, W.O., Iwanami, T., Wang, N., 2008. In planta distribution of '*Candidatus Liberibacter asiaticus*' as revealed by polymerase chain reaction (PCR) and real-time PCR. *Phytopathology* 98, 592–599.
- Zarco-Tejada, P.J., Ustin, S.L., Whiting, M.L., 2005. Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *Agron. J.* 97, 641–653.