remote sensing

MDPI

Article

Finite Difference Analysis and Bivariate Correlation of Hyperspectral Data for Detecting Laurel Wilt Disease and Nutritional Deficiency in Avocado

Jeanette Hariharan [1,*], John Fuller [2], Yiannis Ampatzidis [3,*], Jaafar Abdulridha [3] and Andrew Lerwill [2]

1 Bioengineering Department, Florida Gulf Coast University, 10501 FGCU Blvd., Fort Myers, FL 33965, USA
2 Physics Department, Ave Maria University, 5050 Ave Maria Blvd., Ave Maria, FL 34142, USA
3 Agricultural and Biological Engineering Department, Southwest Florida Research and Education Center, University of Florida, 2685 FL-29, Immokalee, FL 34142, USA
* Correspondence: jhariharan@fgcu.edu (J.H.); i.ampatzidis@ufl.edu (Y.A.); Tel.: +1-239-658-3451 (Y.A.)

Received: 31 May 2019; Accepted: 23 July 2019; Published: 25 July 2019

Abstract: Laurel wilt (Lw) is a very destructive disease and poses a serious threat to the commercial production of avocado in Florida, USA. External symptoms of Lw are similar to those that are caused by other diseases and disorders. A rapid technique to distinguish Lw infected avocado from healthy trees and trees with other abiotic stressors is presented in this paper. A novel method was developed to analyze data from hyperspectral data using finite difference approximation (FDA) and bivariate correlation (BC) to discriminate Lw, Nitrogen (N), and Iron (Fe) deficiencies from healthy avocado plants. Several combinatorial methods were used in preprocessing the data, such as standard normal transformation of data, smoothing of the data, and polynomial fit. The FDA technique was derived using a Taylor Polynomial finite difference approximation. This FDA accentuates inflection points in the spectrum. These, in turn, reveal variance in the data that can be used to identify spectral signature associated with healthy and diseased states. By statistical correlation using the bivariate correlation coefficient of these enhanced spectral patterns, an algorithm (FDA-BC) for distinguishing Lw avocado leaves from all other categories of healthy or mineral deficient avocado leaves is achieved with an overall accuracy of 100%.

Keywords: Laurel wilt; hyperspectral data analysis; bivariate correlation; spectral signature; disease detection

1. Introduction

The avocado crop in Florida is considered the second most economically important crop after citrus. Avocados account for approximately 7500 acres in Miami-Dade County, with an economic impact of more than $54 million [1]. Since 2011, the Florida avocado industry has lost thousands of trees due to a deadly disease named Laurel wilt (Lw). Laurel wilt was identified for the first time in Savannah Georgia in 2002 [2]. In 2011, the ambrosia beetle attacked the commercial production of citrus in Florida [3]. The redbay ambrosia beetle, *Xyleborus glabratus* Eichhoff (Coleoptera, Curculionidae, Scolytinae), is related with fungal symbionts such as *Raffaelea lauricola* [4]; a fungus that blocks the flow of water to other branches, causing wilting. In a matter of a few weeks, a tree can die [5].

It is not easy to recognize the infected tree because Lw disease has similar symptoms of other disorders that appear in early and late stages [6]. Phytophthora root rot, salt damage, freezing, nutrient deficiency has the same symptoms in the early stages. Therefore, it is necessary to identify the disease by a qualified disease diagnostician.

At an early stage, leaves turn to a yellowish color, and at late stages, they turn reddish to purplish. The crown of the tree sometimes partially wilts, requiring an expert to distinguish which disease might be affecting the tree. The expert inspects the fungus after removing the bark to confirm the deep fungal infection in the sapwood. If infected by Lw, the grower should remove the infected tree to disrupt the life cycle of the ambrosia beetle [7]. Currently, there is no known chemical treatment for Lw. The best method for sterilization is to remove the tree, including the root, and burn them in the same grove [8]. Time is of utmost concern in disrupting the disease and improving the odds of survival because Lw kills the tree in a few weeks.

Timely and accurate detection methods are essential to inhibit and prevent the disease from spreading to another area or state [9]. Several methods have been utilized to detect diseases in the field [10–12]. Precision and sustainable applications in agriculture require the use of modern and rapid techniques and technologies for assessing crop health and stress status [13,14], detecting and precisely treating pests [15,16] and diseases, developing site-specific traceability systems [17,18], etc. Nondestructive and accurate methods are necessary to improve early disease detection. Abdulridha et al. and Sankaran et al. [6,7,19] developed rapid techniques to detect Lw in avocado and distinguish it from other diseases and deficiencies, which produce similar symptoms, utilizing hyperspectral and multispectral data and several classification algorithms (neural networks).

Hyperspectral data has been used in recent years to analyze, detect and classify plant disease and environmental factors related to plant stresses. Hyperspectral data analysis was used by Moshou et al. [20] to detect yellow rust (*Puccinia striiformis*) disease of winter wheat crops in the asymptomatic stage by utilizing hyperspectral and fluorescence imaging (450–900 nm). The spectral data were compared with multi-spectral fluorescence images. After comparing the 550 and 690 nm, they discovered that it was possible to detect the disease in its early stages. Varpe et al. [21] monitored the variance in chlorophyll content in *Syzygium cumini* plant species using spectral indices derived from hyperspectral data. The result of regression models demonstrates a good correlation between spectral indices (e.g., mSR, $R^2 = 0.157$; BIG2, $R^2 = 0.069$; ARI, $R^2 = 0.454$) and photosynthetic color contents in all species. Ahmadi et al. [22] utilized spectroradiometer techniques to detect early-stage Ganoderma basal stem rot on Malaysian oil palm. Neural networks were utilized to distinguish infected plants from healthy ones. Satisfactory results were obtained (83%–100%) by using artificial neural networks and first derivative spectral data in ranges of 540 nm to 550 nm. Corti et al. [23] developed a technique to monitor the nitrogen level and water stress in various canopy geometry (rice and spinach), using three detection methods, multispectral, hyperspectral and thermal imaging, in greenhouses. Multivariate regression analysis was applied successfully between spectral wavelengths. Specific wavelengths were proposed and selected to be utilized in field detection of the same crops. Bravo et al. [24] mounted a spectrograph on spray boom height to successfully detect yellow rust in wheat crops. A classification model based on quadratic discrimination was built on a selected group of wavebands obtained by stepwise variable selection. The classification error dropped from 12% to 4%. Franke and Menz [25] tracked the growth of two diseases, powdery mildew (*Blumeria graminis*) and leaf rust (*Puccinia recondita*), in wheat crops for three different periods. Two classification methods were applied; the mixture-tuned matched filtering and decision trees, using the Normalized Difference Vegetation Index (NDVI). The results varied based on the developmental stages of the disease (from 56.8% to 88.6%). Pérez-Bueno et al. [26] detected white root rot in avocado trees utilizing remote sensing. Calderon et al. [27] utilized airborne hyperspectral and thermal camera to detect Verticillium wilt in olive trees. Their contribution distinguished physiological parameters in the hyperspectral data that indicate disease.

Our approach has been to use the hyperspectral data along with enhanced spectral methods and statistical correlation to detect and classify healthy, Lw-diseased, Fe-Deficient and N-Deficient avocados. Improvement of the gradient spectra of various plant leaves was accomplished by using spatial derivatives and curve fitting as well as statistical correlation of the data after the enhancement process, to congregate various specimens (diseased stages, Fe,N-deficiencies, and healthy plants) of the avocado plant into their corresponding group. The spectral radiance data can be transformed

into percent reflectance and resampled in correspondence to the spectral band configuration used by the sensors. Using finite differences, clarification to higher order spectrograms were achieved. The characteristic signatures were then used in the statistical correlation process. After obtaining the divided difference spectra for second and fourth order spatial differences, and deriving a correlation process of the resultant spectra, the algorithm presented is able to clearly distinguish healthy and disease categories of the plant samples.

2. Materials and Methods

2.1. Plant and Sample Selection

Avocados were inoculated at the University of Florida's (UF) Tropical Research and Education Center (TREC) in Homestead, near Miami-Dade Florida's commercial avocado production area (CAPA) and moved to the Citrus Research and Education Center (CREC) in Lake Alfred. All potted '*Simmonds*' avocado trees were healthy and grew up in greenhouse conditions in 25 °C [19,28]. Irrigation, fertilization and other crop maintenance were conducted according to UF recommendations [19,28]. Ten trees (1 to 1.5 m) randomly selected and inoculated with laurel wilt *Raffaelea lauricola*, the culture media was malt extract agar (MEA). The holes (2 mm) were made by portable drill at a 45 downward angle 10 cm above the graft union into which pip 25 µL of inoculum from MEA, prepared with a hemocytometer at a concentration of 30,000 CFUs/mL, was injected and wrapped with para-film [28]. This was achieved by drilling four holes around the trunk and 25 µL of inoculum, prepared with a hem cytometer at a concentration of 30,000 CFUs/mL, was injected and wrapped with para-film. The first set of asymptomatic stage samples were collected 14 days after infection (DAI). The second stage of late development samples was collected 40 DAI. In the late stage, the leaves turned from green to yellowish and brown color with wilting appearing in most parts of the tree.

To obtain iron (Fe) deficiency, 10 potted avocado trees received 1 L Hoagland's solution once a week containing all the major minerals needed by the plants except Fe. A similar procedure was followed for 10 other plants that were not provided with nitrogen (N) to obtain N deficiency. The other trees were received all mineral even N and Fe. The nutrient deficiency symptoms were very similar to laurel wilt in the early stage (Figure 1). Forty leaves samples (four leaves × 10 plants) were collected for each treatment; each leaf was scanned five times, so a total of 200 spectral values were collected. The leaves were measured immediately (within 30 min) in the lab after collected from the greenhouse, so no delay occurred during the measurements.

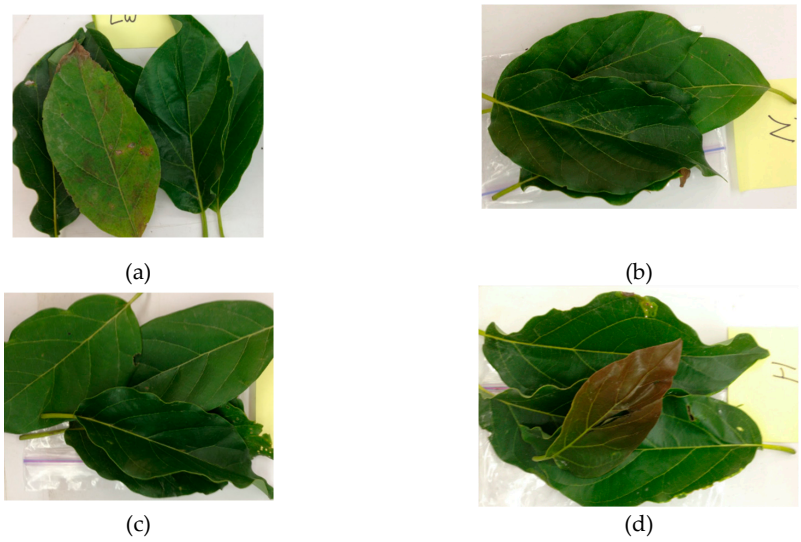

Figure 1. Examples of leaves at different early stress stages: (**a**) Laurel wilt (Lw); (**b**) N deficiency; (**c**) Fe deficiency, and (**d**) healthy leaves. Images were taken 14 days after infection.

2.2. Spectral Data Collection

Spectral data was collected by a spectroradiometer (SVC HR-1024, Spectra Vista Cooperation, NY, USA) with 4° fields of view in laboratory conditions. The range of wavelengths was between 400–970 nm. Two halogen lights were installed on two sides, to improve the ideal conditions for performing the scans and reducing errors. The SVC device was situated so that the lens was 50 cm above the sample, pointing down at it. Spectral signatures were calibrated with a barium sulfate standard reflectance panel (Spectral Reflectance Target, CSTM-SRT-99-100, Spectra Vista Cooperation, NY, USA). Five scans per leaf (five locations per leaf) for each healthy (H), Lw, Fe and N deficiencies were collected.

2.3. Higher Order Spectra Enhancement Methodology

Hyperspectral data of power density reflectance spectra, obtained from healthy and diseased avocado leaves, were used in this analysis. The first step of this methodology involves smoothing of the data using a standard normal transformation. The Standard Normal transformation was used to reduce the effects of outliers, restructure the data into a relational domain, improve data integrity and provide smoothing of the data.

There are four populations of data, and the goal is to examine variance phenomena among them; normalization of the data sets also allows for this comparison method with integrity. The probability density function that was used for this standard normal transformation is given by:

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\zeta} \tag{1}$$

where

$$\zeta = \frac{(X-\mu)^2}{2\sigma} \tag{2}$$

represents the standard normal distribution parameter with mean, μ, and standard deviation, σ. The task then becomes investigating the hyperspectral data to realize slope changes that may be of interest. Second order divided differences will reveal where inflection points occur in healthy data samples, and how these points vary with the various categories of infected/deficient sets that are tested. These inflection point variants prove to be unique to the signatures presented in this work and are distinguished in the correlation process later developed in this work. Considering how these higher-order data functions correlate at regions of interest in the spectrum allow for the categorization of plants into their respective states (i.e., healthy, deficient, diseased).

2.3.1. Higher-Order Spectral Analysis

Resolution enhancement was achieved for this by curve fitting hyperspectral data into a polynomial and then using finite difference approximation with optimal step size. Another common method is found by using Richardson's extrapolation. In both cases, an approximation of the polynomial is the first step in obtaining the numerical difference formulas. The process used for the avocado data set was to first formulate a Taylor interpolation polynomial and then use forward, central or backward divided difference approximations to resolve the data into higher order resolution. The second order forward finite differences are found for the avocado data set. These are compared for resolution enhancement according to the following algorithm.

The spectral data is divided into equally spaced points $\{x_0, x_1, x_{2,\ldots,} x_n\}$ on the interval [a, b], with x_0 = a and x_n = b. Then

$$x_i = x_0 + ih,\ h = \frac{b-a}{n}, \qquad 0 \le i \le n$$

For equivalently spaced data, divided differences can be simplified somewhat. For example:

$$x_{i+1} - x_i = h, \quad x_{i+2} - x_i = 2h, \quad x_{i+3} - x = 3h \tag{3}$$

Then using forward difference quotients, the finite difference method can be developed using a Taylor series approximation of the polynomial:

$$f[x+h] = f[x] + f'[x]h + \frac{f''[x]}{2} h^2 + \frac{f'''^{[x]}}{6} h^3 + \frac{f^{iv}[x]}{24} h^4 + \ldots \frac{f^n[x]}{n!} h^n \tag{4}$$

Now replacing h by 2h:

$$f[x+2h] = f[x] + 2f'[x]h + 2f''[x]h^2 + \frac{4f'''[x]}{3} h^3 + \frac{2f^{iv}[x]}{3} h^4 \tag{5}$$

Eliminating f'[x]h term:

$$2\,f[x+h] - f[x+2h] = f[x] - f''[x]h^2 - f'''[x]h^3 - \frac{7f^{iv}[x]}{12} h^4 + O\left(\xi^3\right) \tag{6}$$

where ξ is a reflectance value between steps, $x_0 - h < \xi < x_0 + h$ with order of error $O\left(\xi^3\right)$. The aim here is to develop this error term to the highest degree reasonable, so that the error decays faster, and our higher order spectra estimate is more accurate. A five-point formula can also be used, which would increase the degree of error to $O\left(\xi^4\right)$.

Solving for $f''[x]$ from Equation (2):

$$f''[x] = (f(x) - 2f(x+h) + f(x+2h))/h^2 - f'''[x]h - \left(7f^{iv}[x]\right)/12\ h^2 + O\left(\xi^3\right) \tag{7}$$

The term $f(x) - 2f(x+h) + f(x+2h)$ will be denoted as $\Delta^2 f(x)$

$$f''(x) \approx \frac{\Delta^2 f(x)}{h^2} - f'''[x]h \tag{8}$$

So, for any function of fixed step size, h, and continuing inductively, the forward divided differencing equations can be defined by the difference operator, Δ

$$f^{(k)}(x) \approx \frac{\Delta^k f(x)}{h^k} - f^{(k+1)}[x]h \tag{9}$$

A pattern emerges for the coefficients of the forward difference operations which is given by the binomial coefficients:

$$\Delta^k f(x) = \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} (-1)^i\, f(x + (k-i)h) \tag{10}$$

where $\begin{pmatrix} k \\ i \end{pmatrix}$ represents the binomial coefficient matrix:

$$\begin{pmatrix} k \\ i \end{pmatrix} = \frac{k!}{i!(k-i)!} \tag{11}$$

By using this binomial expression, a formula is established for the forward difference equation by induction:

$$f[x_i,\ x_{i+1}, x_{i+2},\ \ldots\ x_{i+k},] = \frac{\Delta^k f(x_i)}{h^k} \tag{12}$$

The general form of Equation (9) is used to obtain the second order finite difference quotients of the spectral data. The step size used was h = 2.1854 nm, interpolated and averaged for wavelengths used.

Other higher order formulations for finite differences can be derived by substituting these $O(\xi^k)$ formulas into a Taylor approximation of *f(x+h)*. For example, starting with Equation (2) and substituting in for the $f''(x)$ of Equation (6), the expression:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}\left[\frac{f(x) - 2f(x+h) + f(x+2h)}{h^2} - hf'''(x)\right] - \frac{h^3}{6}f'''(x) + O\left(\xi^4\right) \quad (13)$$

Combining terms and substituting $\frac{\Delta^3 f(x)}{h^3} + O\left(\xi^4\right)$ for $f'''(x)$ and solving for $f'(x)$:

$$f'(x) \approx \frac{1}{6h}[-11f(x) + 18f(x+h) - 9f(x+2h) + 2f(x+3h)] \quad (14)$$

With error of order $O\left(\xi^4\right)$, still using a convention of $f_0 = f(x)$, $f_1 = f(x+h)$, *etc*. These higher order finite differences provide higher accuracy (i.e., less error). For the second and third order, equations can be found to be:

$$f''(x) \approx \frac{1}{12h^2}[-f_{-2} + 16f_{-1} - 30f_0 + 16f_1 - f_2] + O\left(\xi^3\right) \quad (15)$$

$$f'''(x) \approx \frac{1}{8h^3}[-f_{-3} - 8f_{-2} - 13f_{-1} + 0f_0 - 13f_1 + 8f_2 - f_3] + O\left(\xi^4\right) \quad (16)$$

For our analysis, the second order finite difference approximation with error $O\left(\xi^3\right)$ was used.

2.3.2. Bivariate Correlation Process

After obtaining the finite difference spectra for various reflectance of diseased plant leaves, using the Taylor approximation and finite difference method outlined in Section 2.3.1, discernible patterns in the enhanced spectra at various wavelength regions emerged, and were of much greater magnitude among the diseased leaves' population than for the healthy leaves. By applying a population covariance method, the bivariate correlation coefficients (or Pearson's correlation coefficients) were calculated against populations of enhanced leaves' spectra for both the auto-correlated data and cross-correlation of diseased versus healthy plant specimens. The formula used in calculating the bivariate correlation coefficients is given by:

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y} \quad (17)$$

where $\rho_{x,y}$ represents the bivariate correlation coefficient, *Cov*(X,Y) the covariance taken between the two selected enhanced leaf spectra, σ_x the standard deviation for the enhanced spectra test case, and σ_y the standard deviation of the correlation spectrum. The bivariate correlation coefficients were taken for each Region of Interest (ROI) in this study. ROIs that showed inflection points or increased variance in spectral peaks for the enhanced spectra were regarded as areas that could be correlated more accurately. Both the cross-correlated and auto-correlated data were used to find the bivariate correlation coefficients in these specified ROIs. All formulations, calculations, and figures provided were computed using Matlab v.2019a (The MathWorks, Inc., Natick, MA, USA).

3. Results

A set of reflectance data collected by the sensor were obtained for healthy and diseased avocado leaves from the TREC and CREC. After obtaining the data, a standard normalized transformation was applied in order to reduce error of the approximation due to scattering and particle size differences, embedded in the reflectance data. Finite difference approximation (FDA) and bivariate correlation were used to distinguish diseased or nutrient-deficient plants from healthy specimens. Normalized finite difference approximations of each leaf's reflectance values with respect to the wavelength were calculated.

3.1. Data Analysis

After careful examination of the data and noting specifically the variance in the data between 750 and 950 nm, it was determined that a higher order FDA could discern inflection point differences in this region. Figure 2 displays a graph of the averaged reflectance data (plant reflectance signatures) in the complete spectrum and in the 750 to 950 nm region for each leaf category. Taking a fourth order FDA, the zero crossing regions in the 750–950 nm region show the inflection point variance in Figure 3. Figure 4 displays the second order FDA for each group between 750 and 950 nm, after normalizing and transforming the data Each group's FDA reveals categorical spectrums indicative of inflection point variations in this region. The variations of phase, amplitude and peak data are significant. These variants together with the correlation method defined in Section 2.3.2 were used to classify leaves belonging to the categories of Fe-Deficient, N-Deficient, Laurel wilt diseased, or healthy plants.

The interval [750, 950] nm, was selected for the particular ROI, for approximating the bivariate correlation coefficient because of the inflection point variance displayed in the fourth order zero crossing data and spectral variance defined in the second order FDA.

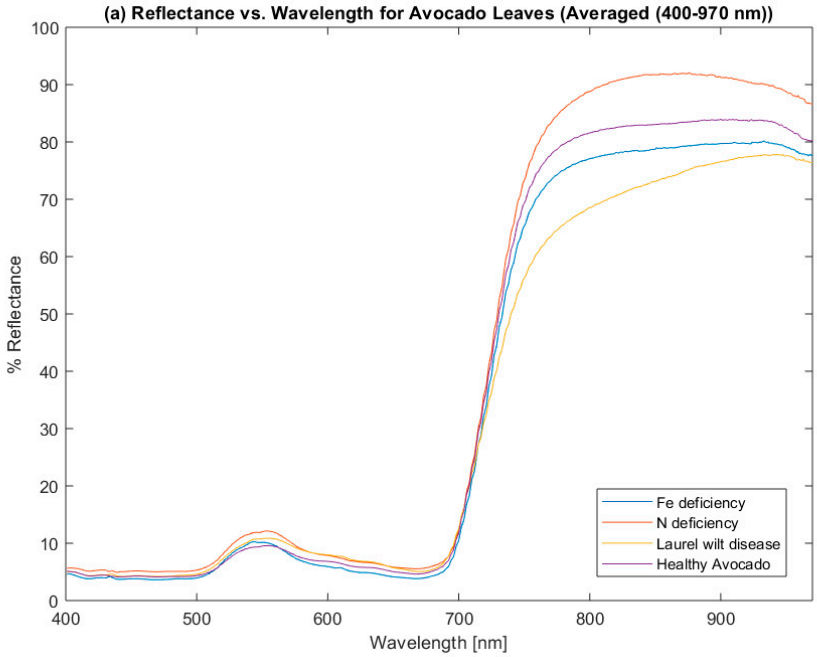

Figure 2. *Cont.*

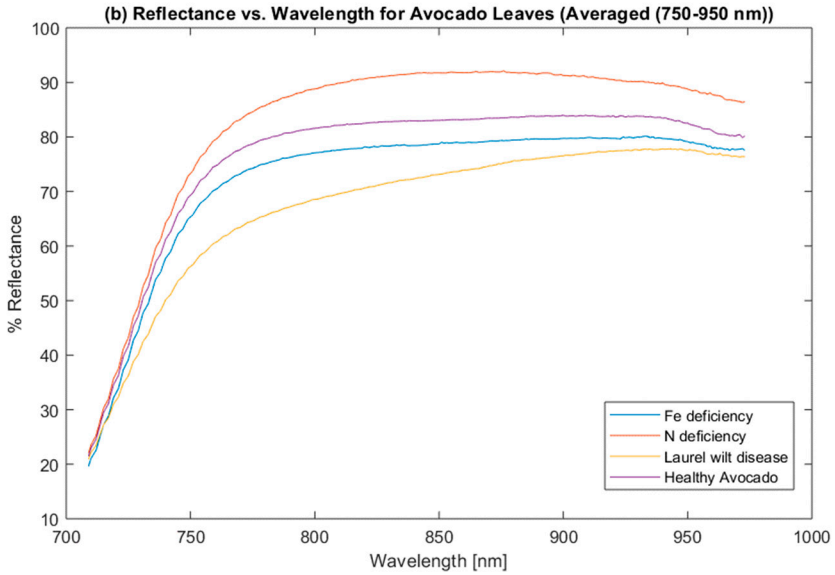

Figure 2. Reflectance signature (data) for Fe-Deficient, Laurel wilt, N-Deficient, and Healthy avocado: (**a**) 400–970 nm Range (**b**) 750–950 nm. Note variance in spectrum and inflection point variations distinct between leaf groups.

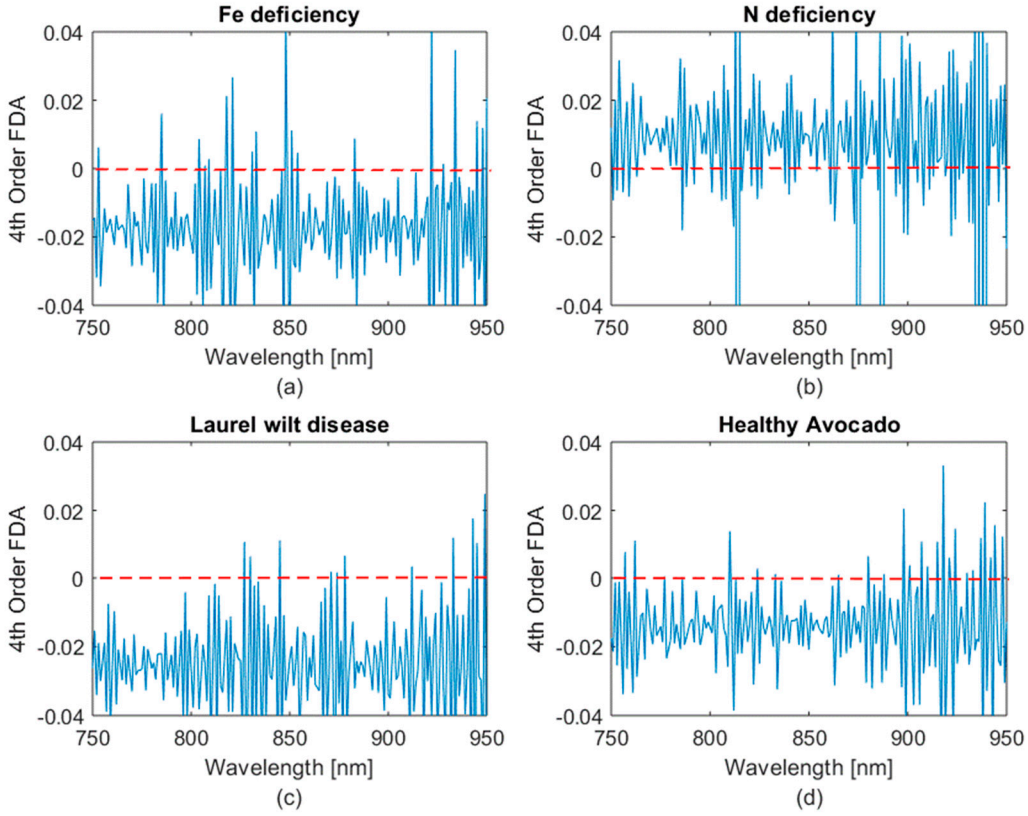

Figure 3. Fourth order finite difference approximations (FDAs) for each group of avocados in [750, 950] nm, normalized and transformed: (**a**) Fe-Deficient FDA; (**b**) N-deficient FDA; (**c**) Lw disease FDA; and (**d**) Healthy plants FDA; Note zero crossings' variations between groups.

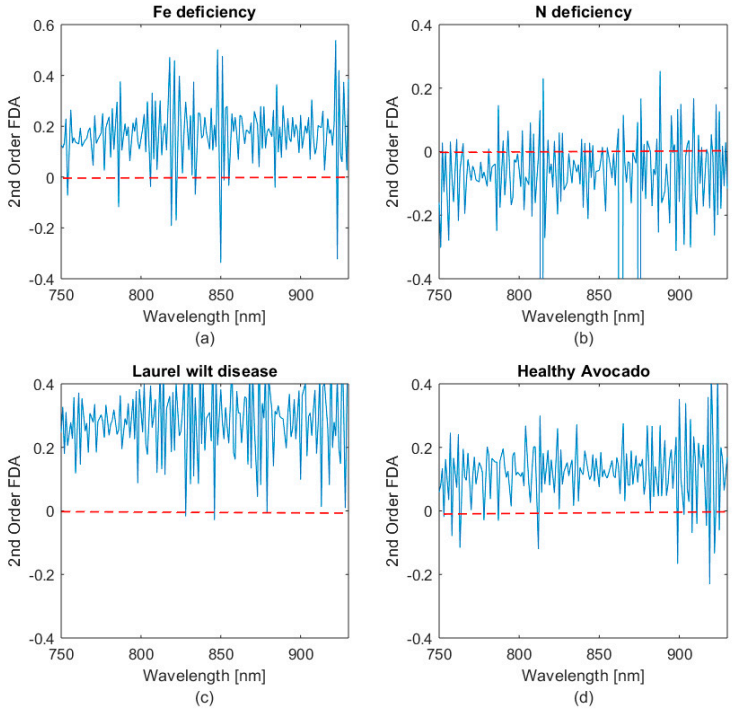

Figure 4. Finite Difference Approximations for each group of avocados in [750, 950] nm: (**a**) Fe-Deficient FDA; (**b**) N-deficient FDA; (**c**) Lw disease FDA; and (**d**) Healthy plants FDA; normalized, transformed.

3.2. *Applied FDA and Bivariate Correlation Process*

After normalization and smoothing of the data, a second order FDA was performed on each leaf's data set. The bivariate correlation coefficients were then calculated between each leaf's second order FDA and the second order FDA of leaves in the other categories. For the calculation of the bivariate correlation coefficient between leaves of the same group, randomly sampled leaves from each group were used to create the average spectra and these were correlated against only the remaining leaves. Since over 300 data sets for numerous avocado leaves were available for each respective leaf group, thirty randomly sampled data sets were used to create the averages. Results follow below for each group of avocados.

3.2.1. Fe-Deficient Classification Results

As shown in Figure 5, all Fe-Deficient specimens produced a correlation greater than 0.5 only with the average of the 30 randomly selected Fe-deficient specimens. Correlations with other groups center around zero. The Fe-Deficient correlation with the Lw-diseased FDA clarifies the most variance but were still below the 0.4–0.5 correlation coefficient point which is being established as the line of distinction categorically.

3.2.2. N-Deficient Classification Results

All N-Deficient specimens produced a correlation greater than 0.5 only with the average of the 30 randomly selected N-Deficient specimens (Figure 6). Correlations with other groups centered around zero. Again, The N-Deficient correlation with the Lw-diseased FDA show the most variance. In both cases, for the N-Deficient and the Fe-Deficient correlations, it can be noted that the variance is the greatest with Lw correlation. The Lw-N-Deficient to Healthy-N-Deficient correlation spectrum differs in an almost 180° out of phase relationship.

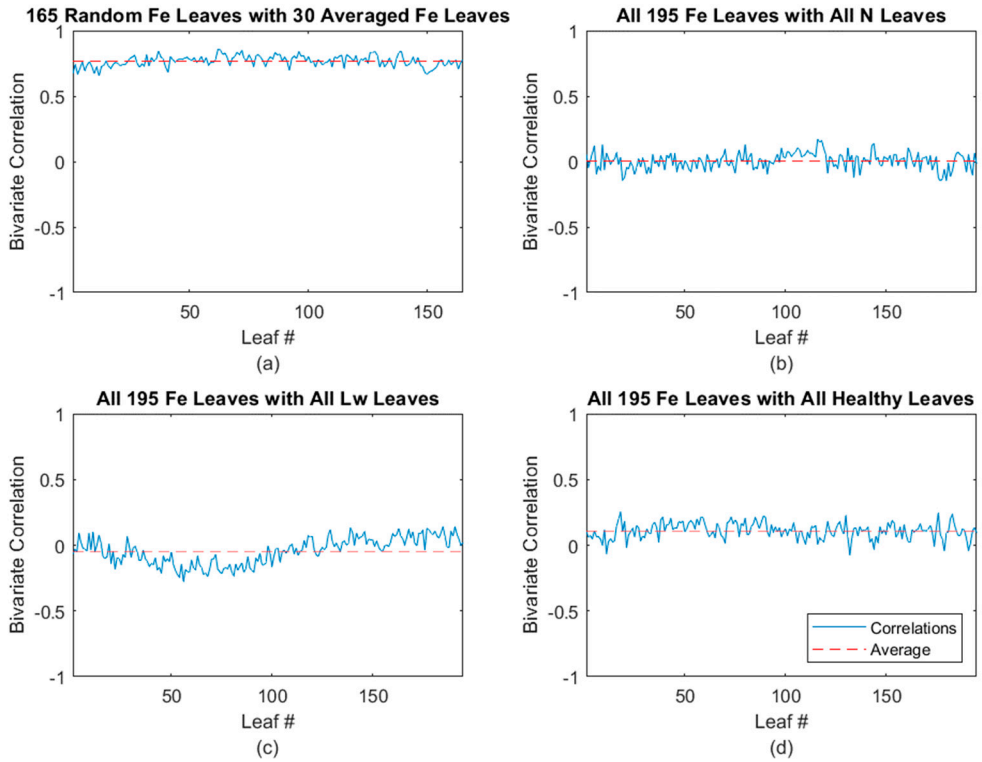

Figure 5. FDA- bivariate correlation (BC) algorithm applied for Fe-deficient avocado leaves: (**a**) Fe-deficient correlated with Fe-deficient; (**b**) Fe-deficient correlated with N-deficient; (**c**) Fe-deficient correlated with Lw; and (**d**) Fe-deficient correlated with healthy.

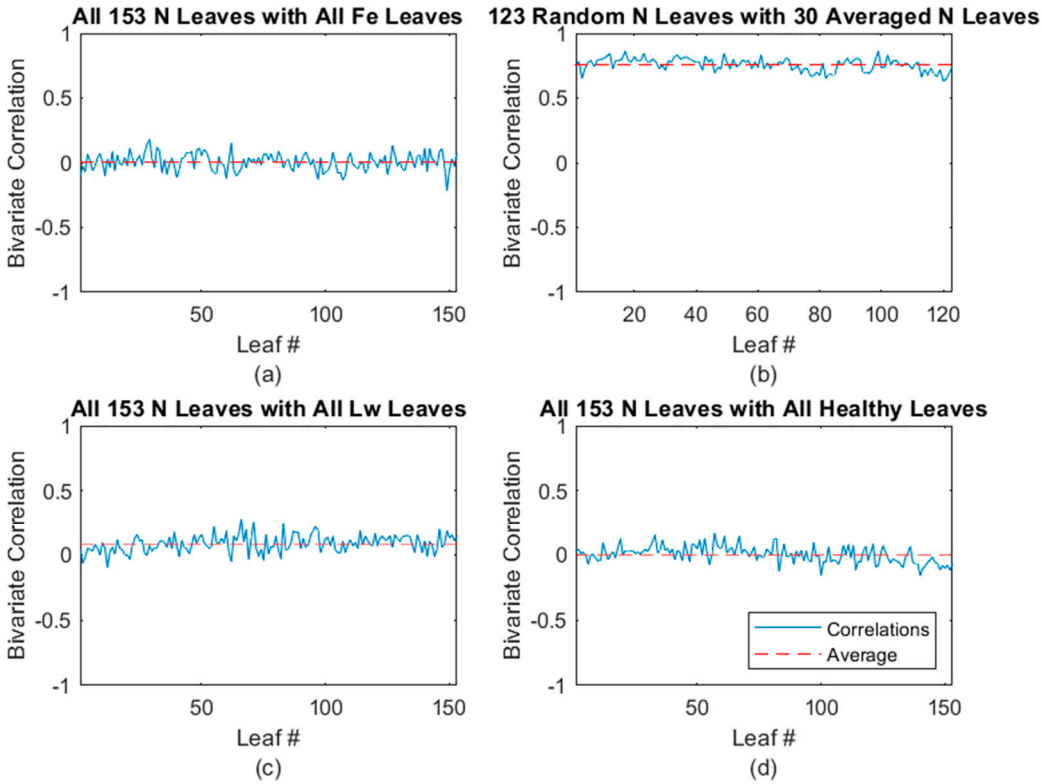

Figure 6. FDA-BC algorithm applied for N-deficient avocado leaves: (**a**) N-deficient correlated with Fe-deficient; (**b**) N-deficient correlated with N-deficient; (**c**) N-deficient correlated with Lw; and (**d**) N-deficient correlated with healthy.

3.2.3. Lw-Disease Classification Results

As seen in Figure 7, all 170 Lw specimens exceed 0.48 correlations with other groups centered near or below zero. This correlation set shows the highest degree of collinearity. This is particularly why it is difficult to distinguish Lw from other abiotic stressors or healthy specimens. Establishing a line of distinction categorically at the 0.4 correlation coefficient mark, to distinguish Lw from the other deficiencies and healthy plants, is considered to be the most applicable from these results.

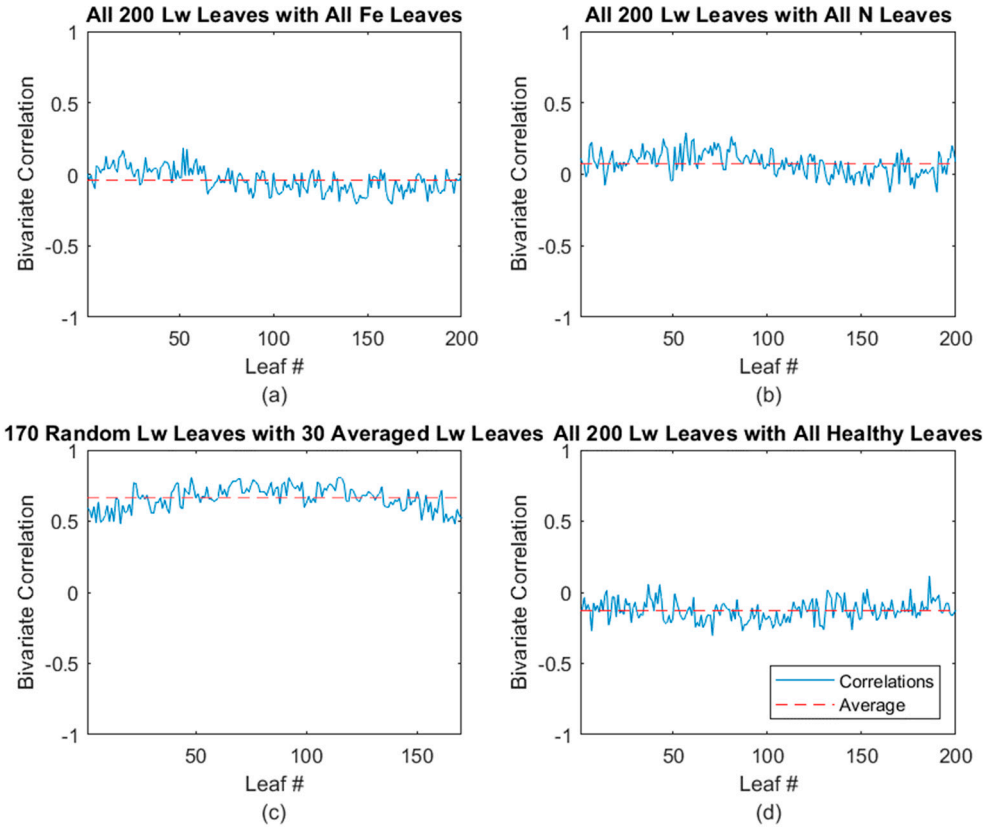

Figure 7. FDA-BC algorithm applied for Lw-diseased avocado leaves: (**a**) Lw correlated with Fe-deficient; (**b**) Lw correlated with N-deficient; (**c**) Lw correlated with Lw; and (**d**) Lw correlated with healthy.

3.2.4. Healthy Avocado Classification Results

All healthy specimens produced a correlation greater than 0.5 only with the average of the 30 randomly selected healthy specimens. Correlations with other groups centered around or below zero, as seen in Figure 8. Again, the pattern of inverse phase relationship occurs when correlated with Lw in this statistical process. The most variance is seen in the healthy to N-Deficient correlation.

3.3. Analysis of Classification Results

The results show that the FDA-BC algorithm is extremely efficient at distinguishing Lw disease and N, Fe deficiencies from healthy avocado plants. Each enhanced spectra group of avocado leaves correlated above 0.4 on a $-1 < \rho_{x,y} < 1$ bivariate range, with its own condition, and correlated near or below zero with any other group. Because of these results, an algorithm for distinguishing between the four groups of avocados was realized. The algorithm for each leaf consisted of:

(i) Normalization of the hyperspectral data.
(ii) Polynomial fitting of data.
(iii) Smoothing the data by moving median with absolute deviation.

(iv) Obtaining the second and fourth order finite difference approximation (FDA).
(v) Establishing regions of interest (ROI) through FDA inflection point analysis of the spectra.
(vi) Detecting and correctly categorizing the leaf sample into one of the healthy or diseased/deficient categories based on the correlation coefficient result.
(vii) For correlations greater than 0.4, the leaf specimen was classified with the specimen group correlated against.

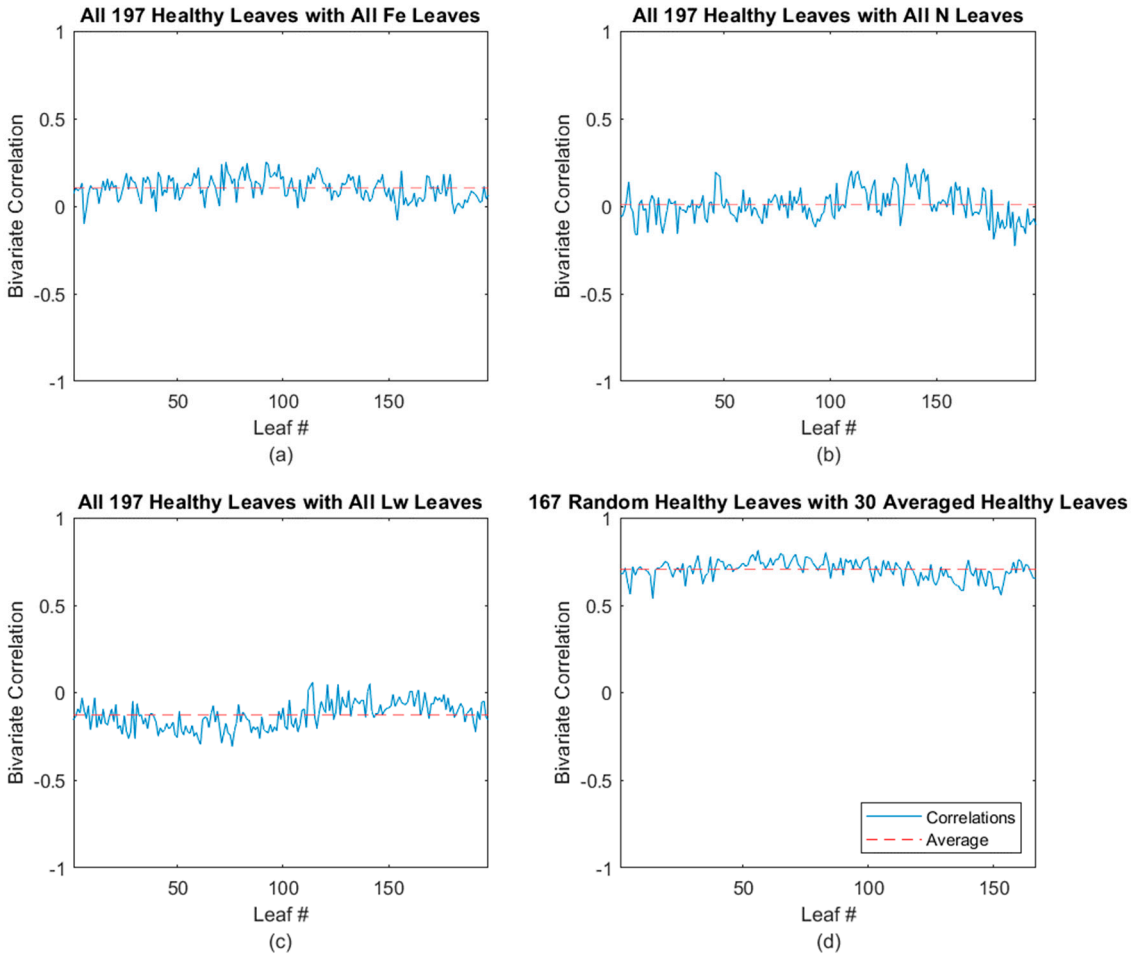

Figure 8. FDA-BC algorithm applied for healthy avocado leaves: (**a**) Healthy correlated with Fe-deficient; (**b**) Healthy correlated with N-deficient; (**c**) Healthy correlated with Lw; and (**d**) Healthy correlated with healthy.

The bivariate correlation of the enhanced spectra for the avocado data distinguishes each leaf into the correct disease or deficiency group of avocados, with significant accuracy. The accuracy of this FDA-BC algorithm is summarized for the categorization of various avocados and their deficiencies in Table 1. As can be noted, all deficiency classification using this method is highly accurate with the population data that was used for our data set. Healthy and disease classifications were also able to be detected with greater than 99% accuracy in most cases. The FDA-BC method is consolidated in Figure 9 to more concisely show the results of the detection and classification method. The results vary only slightly in repeated testing (within ~0.41%), due to the random sampling and averaging of the pattern signature for the autocorrelation process.

Table 1. Summary of Bivariate Correlations for Avocados. Diagnostic accuracy of FDA-BC method of disease plant classification for various avocado deficiencies.

Leaves Tested	Average Correlated With	Average ρ *	% (ρ) > 0.5	% (ρ) > 0.4
Healthy	**Healthy**	0.698	100.00	100.00
	Fe-deficient	0.071	0.00	0.00
	N-deficient	−0.027	0.00	0.00
	Laurel wilt	−0.181	0.00	0.00
Fe-deficient	Healthy	0.076	0.00	0.00
	Fe-deficient	0.778	100.00	100.00
	N-deficient	−0.021	0.00	0.00
	Laurel wilt	−0.090	0.00	0.00
N-deficient	Healthy	−0.032	0.00	0.00
	Fe-deficient	−0.021	0.00	0.00
	N-deficient	0.762	100.00	100.00
	Laurel wilt	0.027	0.00	0.00
Laurel wilt	Healthy	−0.186	0.00	0.00
	Fe-deficient	−0.078	0.00	0.00
	N-deficient	0.038	0.00	0.00
	Laurel wilt	0.645	100.00	100.00

* Bivariate Correlation Coefficient

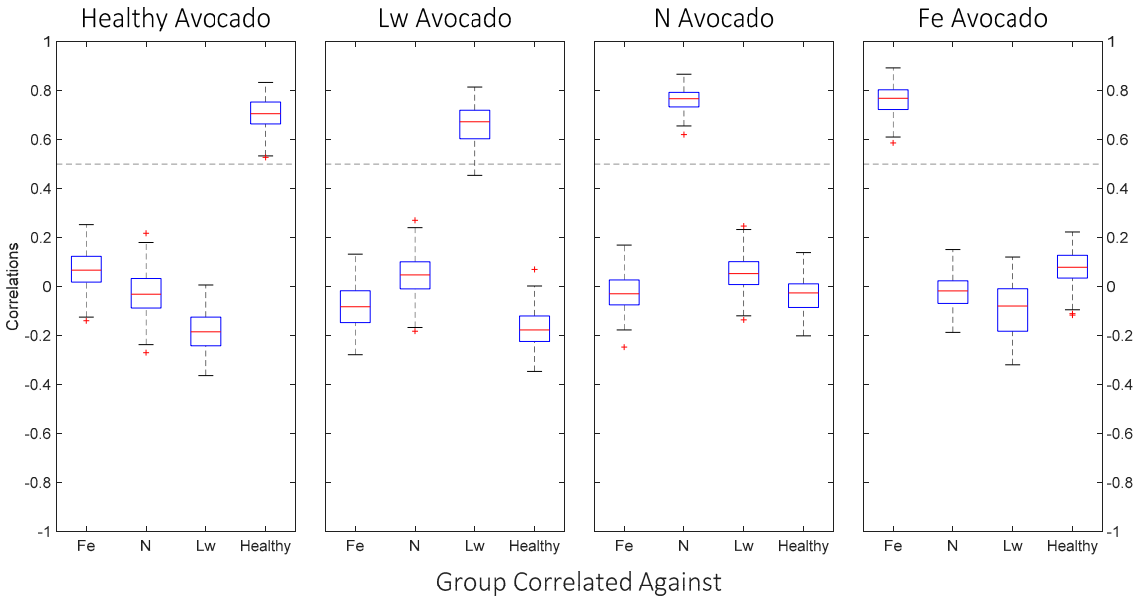

Figure 9. Results of FDA-BC algorithm; Results of Bivariate Correlations of Finite Difference Approximations for all leaf groups using autocorrelation for leaves of the same category, and cross-correlation with other categories.

Confusion Matrix as Figure of Merit

The Confusion Matrix is used as a measure of accuracy in predictive analysis. For the purpose of verifying the FDA-BC results, four rows of correlated data results represent true positives, true negatives, false positives and false negatives for each data set. The first row contains correct correlation results (true positives) for "leaf being tested" that fall above the bivariate division line of 0.4 for leaves in the same category. The second row contains correlated leaf results of the same category that fall below 0.4; this signifies that the categorization was incorrect or a false negative (e.g., An Lw leaf should have correlated above 0.4 for its category of Lw). The graphs in Figure 10 show the first two rows of the correlated data of the Confusion Matrix. Figure 10 gives the Confusion Matrix graphs of the data for the

[True Positives]/[False Negatives] prediction outcomes. Figure 10a shows the classification accuracy for Fe-Deficient plants; Figure 10b gives the results for the N-Deficient classification; Figure 10c shows the Lw classification results, and Figure 10d verifies the healthy avocado leaf identification. There are no false negatives in the Confusion Matrix (data row 2), thus verifying the 100% accuracy of the FDA-BC method in the ROI [750–950 nm].

The Accuracy (ACC) figure of merit represents the actual accuracy of the prediction, using the Confusion Matrix. Table 2 presents the Confusion Matrix for Healthy avocado predictions using the FDA-BC algorithm. All other ACC data is given in Table 2. These ACC results verify that the FDA-BC method used for classification of healthy, diseased and deficient avocado plants is extremely accurate. For the set of randomly sampled and averaged FDA correlations in this analysis, 100% accuracy was achieved, verifying the robustness of the FDA-BC algorithm.

Table 2. Summary of Confusion Matrix; Accuracy: 100%.

		Confusion Matrix Results				
		Actual Class				
		Healthy	Fe-Deficient	N-deficient	Lw-Diseased	
Predicted Class	Healthy	197	0	0	0	
	Fe-Deficient	0	195	0	0	
	N-Deficient	0		153	0	
	Lw-Diseased	0		0	200	
						Totals
True Positives (TP)		197	195	153	200	745
False Negatives (FN)		0	0	0	0	0
True Negatives (TN)		548	550	592	545	2235
False Positives (FP)		0	0	0	0	0
		ACC = (TP+TN)/(P+N) = 1.00				

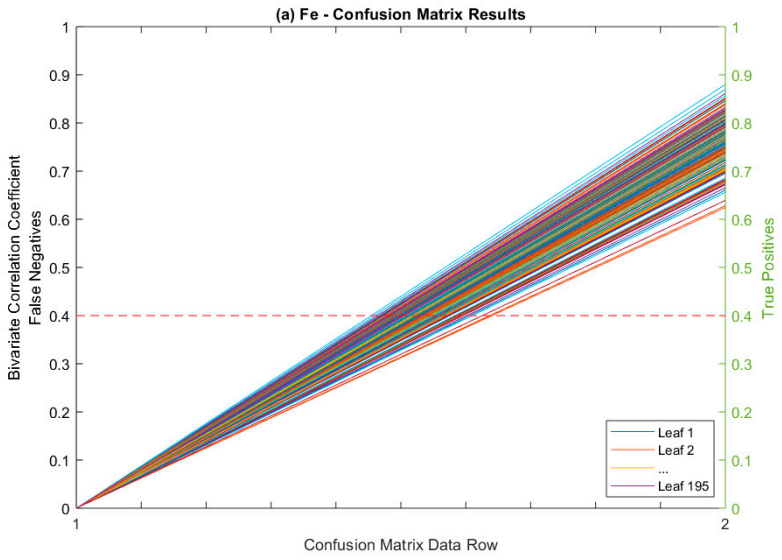

Figure 10. *Cont.*

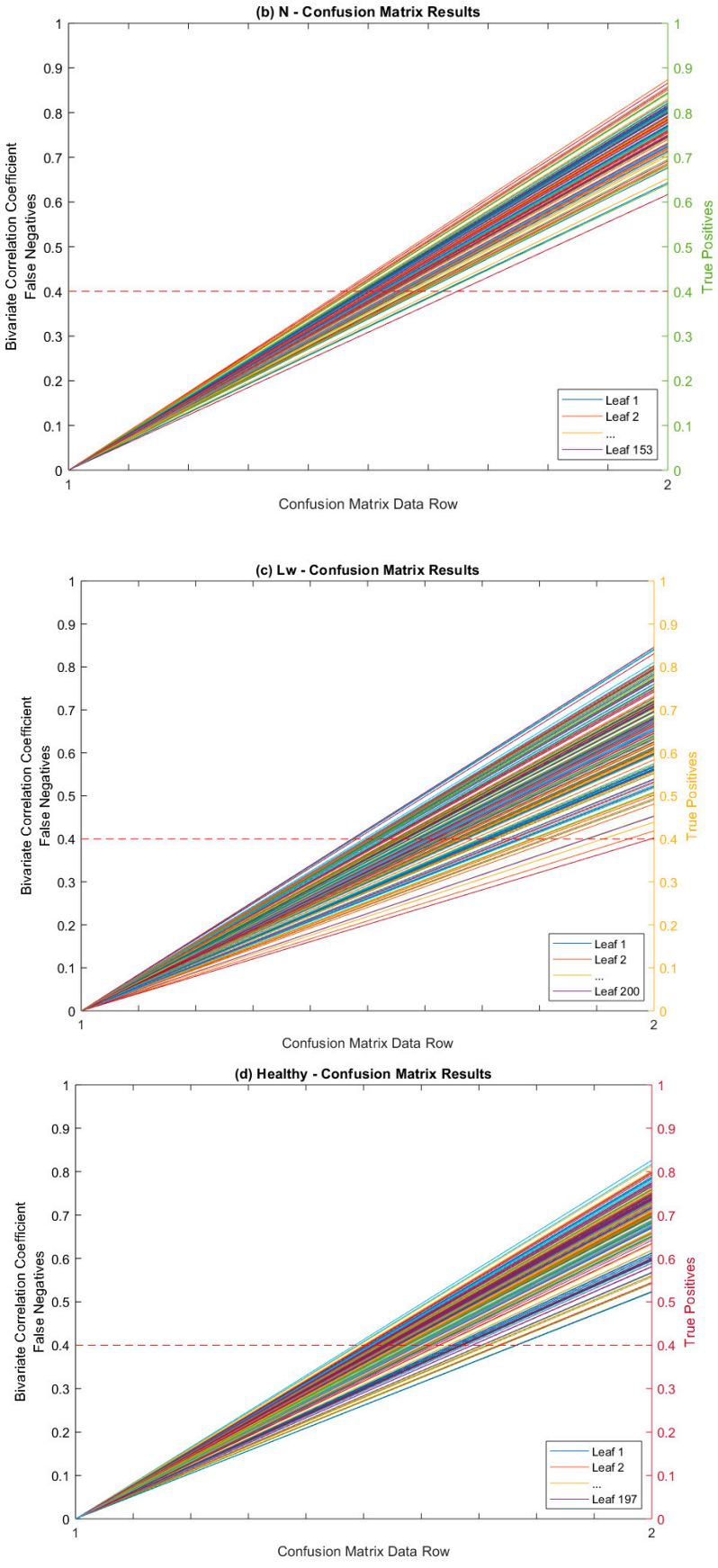

Figure 10. FDA-BC algorithm results of Confusion Matrix.

4. Discussion

Techniques in hyperspectral data analysis have grown rapidly in recent years. Similar work is being achieved in the field of hyperspectral data analysis for plant disorders' detection and classification [6,7,11,19–30]. Perez-Bueno et al. [25] have used normalized difference vegetative index (NVDI) and canopy temperature to form a binary regression model for classification of hyperspectral data in determining white root rot in avocados. Ye et al. [29] used diverse signal processing methods of hyperspectral imaging to distinguish bruised potatoes (at three various levels) from unbruised. Preprocessing methods included Savitzky-Golay smoothing of the data, averaging and transforming data using normal variate transformation to reduce scattering effects (clutter), first and second order derivitave analysis, and multiplicative scatter correction. By using dimensional reduction and a three dimensional state vector machine for decision analysis, a classification decision algorithm, resulted in 100% accuracy. While this algorithm is very accurate, its complexity and strategies used to implement the objective function are difficult to apply for a variety of features in other similar type hyperspectral data.

Other recent methods for detection and classification of diseased species using hyperspectral imagery analysis has been used to identify nutrient-deficient abiotic and biotic stressors in tomato plants [30]. Hyperspectral imaging analysis was used for early detection of root knot nematodes caused by biotic stress to distinguish from abiotic stresses caused by drought in tomato plants. This study chose to analyze data using partial least squares with support vector machine (PLS-SVM) algorithms and matched filter analysis to detect and classify biotic vs. abiotic stressors in the plants. Again, similar preprocessing methods of the data were used by Ye at al. [29]. This method then applies supervised classification using spectral divergence and discriminaton analysis to classify the data. Because of the variability in the data sets, reliable models were only able to be built for specific test sets.

Hyperspectral imaging techniques have been utilized recently to distinguish diseases in the asymptomatic and early stages of Lw disease in avocado. Abdulridha et al. [6,7] and Sankaran et al. [6,7,19] developed rapid techniques to detect Lw in avocado and distinguish it from other diseases and disorders, which produce similar symptoms, utilizing hyperspectral and multispectral data and several classification algorithms (neural networks). Two data sets were collected at 10 nm and 40 nm spectral resolution, and 23 vegetation indices (VIs) were calculated to detect Lw-affected trees by using two classification methods: decision tree (DT) and multilayer perceptron (MLP) neural networks. Additionally, the optimal wavelengths and VIs to discriminate healthy, Lw-infected and avocado trees with iron and nitrogen deficiencies were identified. The work focused on the comparative analysis of using several types of neural networks for classifying diseased, deficient, and healthy avocados. The research proved that the Multi-layer perceptron neural network was superior to the Decision Tree type network in classifying the diseased vs. deficient and healthy.

The FDA-BC method described in this article is more general than these recent methods. It is found to be successful for over 1,300 leaf samples tested. It uses a few well known preprocessing steps (polynomial fit, finite difference approximation with zero crossing identification, normalization, smoothing), followed by statistical correlation. The FDA-BC provides a robust algorithm for detection and classification of diseased plants using hyperspectral data analytics. This technique can be a reliable, verifiable algorithm to be implemented for use in other discriminatory data from hyperspectral data.

We have presented a numerical and statistical method to detect Lw-infected avocado trees (in early and late stage) and discriminate them from N, Fe-Deficient and healthy avocado plants. The method has proved to be highly accurate (ACC = 100%) in the ROI, for the data sets tested. It is a vigorous and versatile method that can be used to detect and classify other hyperspectral data of plant diseases or related stress factors.

Future work to improve this algorithm would be to use more advanced smoothing methods, define spectral pattern signatures for each diseased or deficient category of leaf as well as for the healthy plant to consolidate the correlation processing method, and to design an automatic zero crossing detection scheme. The smoothing would increase the accuracy when afforded with noisy,

in-field data. The signature pattern consolidation would deliver a more rapid processing of the data. The zero-crossing detection would provide immediate and automatic interpretation of the pertinent ROIs to be tested. This being stated, the accuracy of the result would not necessarily improve, but the implementation would be of more value in adding this algorithm as part of a sequence to the remote sensing, hyperspectral data processing system.

5. Conclusions

A method for distinguishing disease factors in plants has been introduced and designed. By developing higher order spectra with Finite Difference Approximation, and statistically correlating the results of these enhanced spectra, a method (FDA-BC) for detecting variations in spectra between healthy and diseased avocado plants has been achieved. The formulation of the problem, the analyses, algorithm development, implementation, and results have been shown in this paper. The FDA-BC method is a more generalized algorithm than some of the more recent methods that have been undertaken to provide solutions for the identification and classification of diseased plant data through hyperspectral data processing. FDA-BC involves fitting the data to a fourth order Taylor Polynomial with specific fit properties using the finite difference approximation to determine regions of interest in the data, smoothing of the FDA by moving average estimate, and normalization of the data. A bivariate correlation coefficient algorithm was then applied to classify the preprocessed hyperspectral data. This novel method has proven to be over 99% accurate on the ROIs tested for avocado Lw, Fe/N deficiencies and healthy plant species. The results are evidence that by means of finite difference approximation to enhance hyperspectral data, and analyzing these enhanced spectra through statistical correlation, the FDA-BC algorithm has proven to be highly accurate in the categorization of diseased vs healthy samples. For the sets of data considered, avocado diseased/deficient/healthy, 100% accuracy was achieved for all leaf samples tested. Other plant species' data are being considered for further verification of this numerical and statistical approach. By further investigation and refining of this algorithm, higher precision can be achieved for disease detection and classification in plant species using this methodology. It is, therefore, recommended that variations of this fundamental algorithm be applied to other hyperspectral data analyses as a general solution that is highly accurate and can be further designed to optimize or verify a neural network or support vector machine approach to detection and classification.

Author Contributions: Conceptualization, J.H. and Y.A.; methodology, J.H., and J.F.; validation, J.H., J.F., and J.A.; formal analysis, J.H. and J.F.; investigation, J.H., J.A., J.F., and Y.A.; resources, J.H. and Y.A; data curation, J.H., J.F., and J.A.; writing—original draft preparation, J.H., J.F., and J.A.; writing—review and editing, J.H., J.F., J.A., Y.A., and A.L.; visualization, J.H., J.A., and J.F.; supervision, J.H. and Y.A.

Funding: This research was partially funded by the Florida Department of Agriculture and Consumer Services, USDA Specialty Block Grant No. 019730.

Acknowledgments: The authors thank Dr. Reza Ehsani, Dr. Randy Ploetz, for their support and assistance in conducting this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Evans, E.A.; Crane, J.; Hodges, A.; Osborne, J.L. Potential economic impact of laurel wilt disease on the Florida avocado industry. *Horttechnology* **2010**, *20*, 234–238. [CrossRef]
2. Smith, J.A.; Dreaden, T.J.; Mayfield, A.E., III; Boone, A.; Fraedrich, S.W.; Bates, C. First report of laurel wilt disease caused by Raffaelea lauricola on sassafras in Florida and South Carolina. *Plant Dis.* **2009**, *93*, 1079. [CrossRef] [PubMed]
3. Ploetz, R.C.; Pena, J.E.; Smith, J.A.; Dreaden, T.J.; Crane, J.H.; Schubert, T.; Dixon, W. Laurel wilt, caused by *Raffaelea lauricola*, is confirmed in Miami-Dade county, center of Florida's commercial avocado production. *Plant Dis.* **2011**, *95*, 1589. [CrossRef] [PubMed]

4. Carrillo, D.; Duncan, R.E.; Ploetz, J.N.; Campbell, A.F.; Ploetz, R.C.; Pena, J.E. Lateral transfer of a phytopathogenic symbiont among native and exotic ambrosia beetles. *Plant Pathol.* **2014**, *63*, 54–62. [CrossRef]
5. Mayfield, A.E., III; Barnard, E.L.; Smith, J.A.; Bernick, S.C.; Eickwort, J.M.; Dreaden, T.J. Effect of propiconazole on laurel wilt disease development in redbay trees and on the pathogen in vitro. *Arboric. Urban For.* **2008**, *34*, 317–324.
6. Abdulridha, J.; Ehsani, R.; Ampatzidis, Y.; de Castro, A. Evaluating the performance of spectral features and multivariate analysis tools to detect Laurel Wilt Disease and Nutritional Deficiency in Avocado. *Comput. Electron. Agric.* **2018**, *155*, 203–211. [CrossRef]
7. Abdulridha, J.; Ehsani, R.; de Castro, A. Detection and Differentiation between Laurel Wilt Disease, Phytophthora Disease, and Salinity Damage Using a Hyperspectral Sensing Technique. *Agriculture* **2016**, *6*, 56. [CrossRef]
8. Fraedrich, S.W.; Harrington, T.C.; Bates, C.A.; Johnson, J.; Reid, L.S.; Best, G.S.; Leininger, T.D.; Hawkins, T.S. Susceptibility to laurel wilt and disease incidence in two rare plant species, pond berry and pond spice. *Plant Dis.* **2011**, *95*, 1056–1062. [CrossRef]
9. Luvisi, A.; Ampatzidis, Y.G.; De Bellis, L. Plant Pathology and Information Technology: Opportunity for Management of Disease Outbreak and Applications in Regulation Frameworks. *Sustainability* **2016**, *8*, 831. [CrossRef]
10. Ampatzidis, Y.; De Bellis, L.; Luvisi, A. iPathology: Robotic Applications and Management of Plants and Plant Diseases. *Sustainability* **2017**, *9*, 1010. [CrossRef]
11. Cruz, A.C.; Luvisi, A.; De Bellis, L.; Ampatzidis, Y. X-FIDO: An Effective Application for Detecting Olive Quick Decline Syndrome with Deep Learning and Data Fusion. *Front. Plant Sci.* **2017**, *8*, 1741. [CrossRef] [PubMed]
12. Cruz, A.; Ampatzidis, Y.; Pierro, R.; Materazzi, A.; Panattoni, A.; De Bellis, L.; Luvisi, A. Detection of grapevine yellows symptoms in *Vitis vinifera* L. with artificial intelligence. *Comput. Electron. Agric.* **2019**, *157*, 63–76. [CrossRef]
13. Ampatzidis, Y.; Partel, V. UAV-based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence. *Remote Sens.* **2019**, *11*, 410. [CrossRef]
14. Ampatzidis, Y.; Kiner, J.; Abdolee, R.; Ferguson, L. Voice-Controlled and Wireless Solid Set Canopy Delivery (VCW-SSCD) System for Mist-Cooling. *Sustainability* **2018**, *10*, 421. [CrossRef]
15. Partel, V.; Kakarla, C.; Ampatzidis, Y. Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Comput. Electron. Agric.* **2019**, *157*, 339–350. [CrossRef]
16. Partel, V.; Nunes, L.; Stansley, P.; Ampatzidis, Y. Automated Vision-based System for Monitoring Asian Citrus Psyllid in Orchards utilizing Artificial Intelligence. *Comput. Electron. Agric.* **2019**, *162*, 328–336. [CrossRef]
17. Ampatzidis, Y.; Tan, L.; Haley, R.; Whiting, M.D. Cloud-based harvest management information system for hand-harvested specialty crops. *Comput. Electron. Agric.* **2016**, *122*, 161–167. [CrossRef]
18. Ampatzidis, Y.G.; Vougioukas, S.G. Field experiments for evaluating the incorporation of RFID and barcode registration and digital weighing technologies in manual fruit harvesting. *Comput. Electron. Agric.* **2009**, *66*, 166–172. [CrossRef]
19. Sankaran, S.; Ehsani, R.; Inch, S.A.; Ploetz, R.C. Evaluation of visible-near infrared reflectance spectra of avocado leaves as a non-destructive sensing tool for detection of laurel wilt. *Plant Dis.* **2012**, *96*, 1683–1689. [CrossRef]
20. Moshou, D.; Bravo, C.; Oberti, R.; West, J.; Bodria, L.; McCartney, A.; Ramon, H. Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real Time Imaging* **2005**, *11*, 75–83. [CrossRef]
21. Varpe, A.B.; Surase, R.R.; Vibhute, A.D.; Gaikwad, S.V.; Rajendra, Y.D.; Kale, K.V.; Mehrotra, S.C. *Synygium cumini* Plant Photosynthetic Pigment Detection from Hyperspectral data sets using Spectral Indices. In Proceedings of the 2nd Intenational Conference on Man and Machine Interfacing (MAMI), Bhubaneswar, India, 21–23 December 2017.
22. Ahmadi, P.; Muharam, F.M.; Ahmad, K.; Mansor, S.; Abu Seman, I. Early Detection of Ganoderma Basal Stem Rot of Oil Palms Using Artificial Neural Network Spectral Analysis. *Plant Dis.* **2017**, *101*, 1009–1016. [CrossRef] [PubMed]

23. Corti, M.; Masseroni, D.; Marino Gallina, L. Use of Spectral and Thermal Imaging Sensors to Monitor Crop Water and Nitrogen Status. In *First Conference on Proximal Sensing Supporting Precision Agriculture*; EAGE: Houten, The Netherlands, 2015; Volume 3997/2214.
24. Bravo, C.; Moshou, D.; West, J.; McCartney, A.; Ramon, H. Early disease detection in wheat fields using spectral reflectance. *Biosyst. Eng.* **2003**, *84*, 137–145. [CrossRef]
25. Franke, J.; Menz, G. Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precis. Agric.* **2007**, *8*, 161–172. [CrossRef]
26. Perez-Bueno, M.L.; Pineda, M.; Vida, C.; Fernandez-Ortuno, D.; Tores, J.A.; de Vicente, A.; Cazorla, F.M.; Baron, M. Detection of White Root Rot in Avocado Trees by Remote Sensing. *Plant Dis.* **2019**, *103*, 1119–1125. [CrossRef] [PubMed]
27. Calderon, R.; Navas-Cortes, J.A.; Lucena, C.; Zarco-Tejada, P.J. High-resolution airborne hyperspectral and thermal imagery for early, detection of *Verticillium* wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sens. Environ.* **2013**, *139*, 231–245. [CrossRef]
28. Ploetz, R.C.; Perez-Martinez, J.M.; Smith, J.A.; Hughes, M.; Dreaden, T.J.; Inch, S.A.; Fu, Y. Responses of avocado to laurel wilt, caused by *Raffaelea lauricola*. *Plant Pathol.* **2012**, *61*, 801–808. [CrossRef]
29. Ye, D.D.; Sun, L.J.; Tan, W.Y.; Che, W.K.; Yang, M.C. Detecting and classifying minor bruised potato based on hyperspectral imaging. *Chemom. Intell. Lab. Syst.* **2018**, *177*, 129–139. [CrossRef]
30. Susic, N.; Zibrat, U.; Sirca, S.; Strajnar, P.; Razinger, J.; Knapic, M.; Voncina, A.; Urek, G.; Stare, B.G. Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging. *Sens. Actuators B Chem.* **2018**, *273*, 842–852. [CrossRef]