Estimating the severity of apple mosaic disease with hyperspectral images

Songtao Ban1, Minglu Tian2,3, Qingrui Chang1*

(1. College of Natural Resource and Environment, Northwest A&F University, Yangling 712100, China; 2. Agricultural Information Institute of Science and Technology, Shanghai Academy of Agricultural Sciences, Shanghai 201403, China; 3. Shanghai Engineering Research Centre for Digital Agriculture, Shanghai 201403, China)

Abstract: Soil Plant Analysis Development (SPAD) Chlorophyll Meter reading was used to effectively characterize chlorophyll content, which is an important indicator of the health status of plant leaves. In this study, the hyperspectral images of apple leaves infected by apple mosaic virus (ApMV) were captured, and their SPAD values were measured. The spectral reflectance of leaves with varying disease infection is significantly different. Several hyperspectral vegetation indices were highly correlated with the SPAD values of apple leaves (correlation coefficient > 0.9). Models were established to estimate apple foliar SPAD values based on these vegetation indices. Among the models, the multivariate regression model with partial least square regression (PLSR) method achieved the highest accuracy. The SPAD value of a whole apple leaf was calculated from its SPAD distribution image and used as a quantitative index to represent the health status of an apple leaf. Furthermore, the SPAD value of a whole apple leaf could also be estimated rapidly and accurately by extracting the spectral average value of the whole leaf using a simple model. It can be used as a rapid detection method of SPAD values of apple leaves to monitor and describe the health conditions of apple leaves quantitatively.

Keywords: hyperspectral image, apple leaf, mosaic disease, SPAD, plant health detection

DOI: 10.25165/ijabe.20191204.4524


1 Introduction

Apple mosaic disease is a common viral disease on apple trees caused by Apple mosaic virus (ApMV). Irregular pale-yellow spots or bands along major veins can be found on affected leaves[1]. Damages in mesophyll cells would cause a decrease in chlorophyll content, and thus decline the photosynthetic capacity of leaves. As a result, the apple yield will reduce considerably[2]. The health status and the physiological performance of a plant can be assessed accurately by measuring the amount of chlorophyll content in leaves[3-5]. Thus, the severity of mosaic diseases can be estimated by determining the chlorophyll content in leaves[6].

SPAD Chlorophyll Meter is widely used to measure foliar chlorophyll content non-destructively, in vivo. By measuring the transmittance of two light radiation at 650 nm and 940 nm, the instrument calculates a relative SPAD (Soil Plant Analysis Development) value that is highly correlated with chlorophyll content. In previous researches, the coefficient of determination of the relationship between chlorophyll content and SPAD value were between 0.8-0.95[7-10]. Therefore, it is feasible to use SPAD value as a diagnostic indicator to reflect leaf health conditions, and the level of disease on apple leaves can be determined by their SPAD values[11,12]. Although the SPAD Chlorophyll Meter is able to acquire instant SPAD data, it can only measure the chlorophyll content of a single point on one leaf each time. It cannot obtain the SPAD value or its distribution of a whole leaf. However, mosaic spots or bands are distributed unevenly on leaves, and single-point measurement cannot accurately evaluate the disease level of a whole leaf. As leaves are the basic units of photosynthesis of an individual plant, the health condition of a whole leaf should be quantitatively described.

A hyperspectral image has the characteristic of image-spectrum integration. Every pixel on the hyperspectral image contains abundant spectral information of the target object. In theory, the specific property of a target object on each pixel can be predicted by using a proper remote sensing inversion model based on their spectral signatures. With the development of hyperspectral imaging technology, imaging spectrometers are widely used in agriculture-related processes, such as estimating chlorophyll[13], water[14,15], and nutrient[16,17] contents of crop leaves, monitoring damages caused by pests and diseases in farmland[18], and examining disease spots on crop fruits and seeds[19,20]. Huang et al.[21] quantified the disease index of yellow rust in wheat with photochemical reflectance index derived from aerial hyperspectral images. Delalieux et al.[22] used both fluorescence and hyperspectral images to detect apple scab infections on leaf scale and found that the index p1480/p2135 had the potential to identify scab disease before visible symptoms appeared. Mehrabeglu et al.[23] applied the support vector machine classifier on hyperspectral images to identify red blotch disease on grape leaves. Xie et al.[24] used image classification technique to distinguish healthy tomato leaves and leaves with gray mold of different severities according to spectral characteristics.

Hyperspectral imagery technology has great potential to detect plant diseases. However, there were few researches on apple mosaic disease by hyperspectral imaging. Besides, most
researches focused on image classification technique when using hyperspectral images to identify plant diseases and used the percent of diseased leaf area to estimate disease severity, which is not objective or accurate enough\cite{23}. Therefore, the objective of this paper is to develop a method that can quantitatively describe the severity of apple mosaic disease on leaf scale through hyperspectral images by detecting the SPAD value.

2 Materials and methods

2.1 Leaf sampling and SPAD value measurement

A total of 160 infected (in varying severity) and healthy apple leaves from the cultivar ‘Gala’ were randomly collected from an orchard in Shaozhai Village (34°22′12″ N, 108°00′45″ E), Xinglin Town, Fufeng County, Shaanxi Province, China on June 7, 2016, when apple trees were in their fruit filling stage. The leaves were sealed in a plastic bag and then placed in an insulation box with ice bags to maintain freshness.

The leaves were immediately transported to the laboratory for measurement. With the leafstalk at the bottom, each leaf was divided into four sections according to the symptom, and the distribution and severity of disease spots within each section were similar. Since the chlorophyll content of each section for each leaf needed to be measured, it is impossible to use the chemical method because some of the sections weren’t weighted enough for the extraction. Therefore, a SPAD-502 Plus Chlorophyll Meter (Minolta Camera Co., Ramsey, NJ, Japan) was used to measure foliar chlorophyll content, which could measure an area as small as 2 mm × 3 mm. In order to cover the whole section, each section was measured ten to twenty times according to its size, and the average of these readings was used as the SPAD value of the section. In this way, the SPAD values of 640 sections were recorded.

2.2 Acquisition and processing of hyperspectral images of leaves

SOC710-VP portable hyperspectral imager (Surface Optics Corp., San Diego, CA, USA) is an embedded push-broom imaging spectrometer that does not require an external scanning platform and possesses a visual focusing function. The hyperspectral image consisted of 128 wavebands within the range of 400-1000 nm and 520×696 pixels. After measuring the hyperspectral image of the leaves, the leaves were acquired under sunlight in an outdoor space by using SOC710-VP. The measurement was conducted between 12:00 pm to 2:00 pm in sunny weather and no wind or clouds.

With the support of SRAnal 710 software, raw images were converted into reflectance images that could be processed by the remote sensing image software ENVI 5.1 (EXELIS VIS, Boulder, CO, USA). The sections corresponding to the measured SPAD on each leaf were selected by using the ROI tool in ENVI 5.1, and the weighted average of spectral reflectance in each ROI was used as the spectrum of this section. Therefore, the spectra of 640 samples were obtained.

2.3 SPAD-sensitive vegetation indices

As the linear or nonlinear combinations of reflectance in several wavebands, vegetation indices (VIs) are widely used in remote sensing to build relationships between spectral properties and plant biophysical parameters. VIs can be used to retrieve foliar pigments, water content, cell structure, biomass, and other parameters in plant highlights\cite{26}. A number of narrowband VIs that are sensitive to SPAD value or chlorophyll content were used, and the correlation coefficients, denoted by \( r \), between these indices and SPAD values were calculated in Excel 2016 (Microsoft, Redmond, Wash, USA). Five VIs with absolute \( r \) values higher than 0.9 were selected to establish an SPAD value estimation model of apple leaves (Table 1).

### Table 1 Vegetation indices and formulas used in this paper

<table>
<thead>
<tr>
<th>Name</th>
<th>Formula</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green Normalized Difference Vegetation Index</td>
<td>( \rho_{790-550}/\rho_{774-740} )</td>
<td>\cite{27}</td>
</tr>
<tr>
<td>Pigment Specific Simple Ratio for Chl a (PSSRa)</td>
<td>( \rho_{680}/\rho_{550} )</td>
<td>\cite{28}</td>
</tr>
<tr>
<td>Pigment Specific Simple Ratio for Chl b (PSSRb)</td>
<td>( \rho_{680}/\rho_{635} )</td>
<td></td>
</tr>
<tr>
<td>Vegmolleng Red Edge Index 1 (VORI)</td>
<td>( \rho_{750}/\rho_{710} )</td>
<td>\cite{29}</td>
</tr>
<tr>
<td>Simple Vegetation Index Ratio (SR)</td>
<td>( \rho_{740}/\rho_{667} )</td>
<td></td>
</tr>
</tbody>
</table>

2.4 Modelling methods

First, the 640 samples were sorted in ascending order of SPAD values. Then, using systematic sampling method, taking the third sample as the start point and taking the next sample with every five-sample interval, the data of the 640 samples were divided into a calibration subset containing 512 samples, which were used to build the regression models, and a validation subset containing 128 samples, which were utilized to evaluate the regression models. Table 2 shows the descriptive statistics of the datasets.

Single-element regression analyses between vegetation indices and SPAD values of leaves were carried out in Origin 9.0 (OriginLab Cooperation, Northampton, USA), and the optimal results were selected to establish the single-element regression models for foliar SPAD estimation based on single VI.

Collinearity diagnostics of VIs was carried out in SPSS 20 (IBM Corp., Armonk, NY, USA) and the results are seen in Table 3. Four eigenvalues are close to 0 and two condition indices are larger than 30, indicating a serious problem with collinearity. In such a case, multivariate linear regression will not work well. Partial least squares regression (PLSR) is a multivariate statistical analysis method. Compared with traditional multiple regression method, PLSR comprehensively considers principal component analysis, classic correlation analysis, and multiple linear regression in the modeling process, effectively eliminating the collinear effect of independent variables\cite{30}. Therefore, a SPAD estimation model based on multiple VIs was established using SIMCA-P12.0 (Umetrics AB, Umeå, Sweden) and the PLSR was used to make full use of the information contained in the VIs that is highly correlated with foliar SPAD values.

### Table 2 Descriptive statistics of the apple leaves SPAD values

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration subset</td>
<td>512</td>
<td>27.2</td>
<td>23.7</td>
<td>0</td>
<td>69.6</td>
</tr>
<tr>
<td>Validation subset</td>
<td>128</td>
<td>27.1</td>
<td>23.5</td>
<td>0</td>
<td>69.1</td>
</tr>
<tr>
<td>All samples</td>
<td>640</td>
<td>27.2</td>
<td>23.6</td>
<td>0</td>
<td>69.6</td>
</tr>
</tbody>
</table>
3 Results

3.1 SPAD values and hyperspectral signatures of apple leaves with different disease degrees

With the aggravation of the mosaic disease, leaves show small-scaled yellow and pale spots in early stages, large-scaled yellow and pale regions in moderate degrees, and complete whitening symptoms in severe cases. Apple mosaic virus destroys leaf tissue structures and causes chlorophyll decomposition in infected regions\[31\]. As a result, SPAD values whitening symptoms in severe cases. Apple mosaic virus yellow and pale regions in moderate degrees, and complete small-scaled yellow and pale spots in early stages, large-scaled yellow and pale spots in severe cases.苹果花叶病毒在受侵染区域导致叶绿素分解，SPAD值变化显著。Zhang et al. \[33\] and Devadas et al. \[34\] found similar phenomena in the visible regions. However, the reflectance reduction in the NIR regions is not as evident as that in the visible regions because of cell structure destruction.

Figure 1 SPAD values of infected regions with varying severity on an apple leaf with mosaic disease

In Figure 2, the spectral reflectance in regions with different disease degrees, represented by various SPAD values, varies significantly in the visible wavelength (400-700 nm), where the reflectance increases evidently as SPAD values decrease. This phenomenon occurs because the absorption spectrum of chlorophyll is mainly in the visible light range, and the chlorophyll decomposition in the infected regions caused by the disease weakens the absorption of visible light but enhances the reflectance\[32\], and that is why the disease spots are in pale-yellow color. In general, the spectral reflectance of the infected regions in the near-infrared (NIR) wavebands (700-1000 nm) is lower than that in normal regions because of cell structure destruction. However, the reflectance reduction in the NIR regions is not as evident as that in the visible regions. Similar phenomena were found by Zhang et al. \[33\] and Devadas et al. \[34\]. The possible reason is that chlorophyll does not interact with NIR radiation, and the reflectance and transmission of NIR radiation is mainly determined by the internal structure of the leaves.

3.2 Establishment and verification of VIs-based apple foliar SPAD value estimation model

Based on simple regression analysis, five single-element apple foliar SPAD estimation models were established for 512 calibration subsets by using each of the VIs as independent variable. These five models were marked as SPAD-GNDVI, SPAD-PSSRa, SPAD-PSSRb, SPAD-VOG1, and SPAD-SR. A multivariate SPAD estimation model was established using the five VIs based on PLSR method as independent variables and recorded as SPAD-PLSR. The coefficients of determination ($R^2$) and the root-mean-square error (RMSE) of the models were used to evaluate the accuracy of different models. In Table 4, SPAD-PLSR model shows the highest accuracy because all spectral information contained in the VIs is used. Among the five single-element SPAD estimation models, SPAD-GNDVI shows the highest accuracy, and the $R^2$ and RMSE values are similar to the SPAD-PLSR because of the significant correlation between the spectral reflectance at 550 nm and the SPAD value, as described in Section 3.1.

### Table 3 Collinearity diagnostics of VIs

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Eigenvalue</th>
<th>Condition Index (constant)</th>
<th>GNDVI</th>
<th>PSSRa</th>
<th>PSSRb</th>
<th>VOG1</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.810</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.160</td>
<td>6.025</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>0.020</td>
<td>17.124</td>
<td>0.03</td>
<td>0.02</td>
<td>0.04</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>4</td>
<td>0.008</td>
<td>26.702</td>
<td>0.06</td>
<td>0.47</td>
<td>0.03</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>0.001</td>
<td>78.637</td>
<td>0.12</td>
<td>0.31</td>
<td>0.70</td>
<td>0.20</td>
<td>0.43</td>
</tr>
<tr>
<td>6</td>
<td>0.001</td>
<td>95.340</td>
<td>0.76</td>
<td>0.19</td>
<td>0.23</td>
<td>0.79</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Note: $y$ refers to SPAD, $x_1$ refers to GNDVI, $x_2$ refers to PSSRa, $x_3$ refers to PSSRb, $x_4$ refers to VOG1, and $x_5$ refers to SR.

The prediction abilities of all models were evaluated by a validation subset. Fitting analysis between the predicted SPAD values based on different models and the measured SPAD values were conducted. $R^2$, RMSE, relative error (REP), and the slope of the fitting equations were used as indices to evaluate the estimation accuracy (Table 5). The prediction result of SPAD-PLSR model...
showed the highest accuracy and thus can be used as the optimal SPAD estimation model. In addition, SPAD-GNDVI model also presented high prediction accuracy and a relatively simple structure.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2$</th>
<th>RMSE</th>
<th>REP (%)</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAD-GNDVI</td>
<td>0.8686</td>
<td>3.8559</td>
<td>14.9464</td>
<td>0.9437</td>
</tr>
<tr>
<td>SPAD-PSSRa</td>
<td>0.8043</td>
<td>5.6747</td>
<td>19.1126</td>
<td>0.8895</td>
</tr>
<tr>
<td>SPAD-PSSRb</td>
<td>0.8451</td>
<td>4.3341</td>
<td>14.5974</td>
<td>0.9344</td>
</tr>
<tr>
<td>SPAD-VOGI</td>
<td>0.8151</td>
<td>5.5851</td>
<td>18.8109</td>
<td>0.8963</td>
</tr>
<tr>
<td>SPAD-SR</td>
<td>0.8114</td>
<td>5.4679</td>
<td>18.4165</td>
<td>0.8882</td>
</tr>
<tr>
<td>SPAD-PLSR</td>
<td>0.8712</td>
<td>3.8439</td>
<td>13.1101</td>
<td>0.9515</td>
</tr>
</tbody>
</table>

### 3.3 SPAD value distribution in apple leaves

In the ENVI software, mask tool was used to remove the background and extract the hyperspectral images of leaves. The distribution of SPAD values of apple leaves was obtained by calculating hyperspectral images of apple leaves pixel-by-pixel using the SPAD-PLSR model. Figure 3 shows the true color images (Figure 3a) and SPAD value distribution images (Figure 3b) of three leaves in severe (leaf No.1), medium (leaf No.2), and mild (leaf No.3) disease level. Each pixel value of the resulting image is the SPAD value of the leaf pixel. The area percentage of different level of SPAD values in each leaf could be calculated from the frequency distribution of SPAD value (Figure 4). This statistical result would help elucidate the damage accurately and quantitatively. In Figure 4, the SPAD values of more than 77% area of the whole leaf No.1 are lower than 10, indicating that chlorophyll in most area of the leaf was degraded because of the infection. Conversely, more than 90% area of leaf No.3 shows SPAD values exceeding 50, which implied that this leaf is slightly infected and worked well.

### 3.4 Estimation of SPAD values of a whole leaf

Leaf is the basic unit of plant photosynthesis function. The SPAD value of the whole leaf can be used as an indicator of its photosynthesis ability and can provide a general overview of its health condition quantitatively. A weighted average of all pixels values on the SPAD value distribution image of the leaf could be calculated as the SPAD value for the whole leaf. The SPAD values of the three whole leaves are 7.12 (leaf No.1), 17.06 (leaf No.2), and 58.97 (leaf No.3) respectively, which accurately reflect the health status and photosynthesis ability of each leaf.

Figure 4 Area statistics (expressed as a percentage) of the predicted SPAD value distribution of each leaf in Figure 3. The SPAD values were divided into 7 classes, and the area of each class was counted

### 4 Discussion

Hyperspectral remote sensing has been an effective tool for monitoring the health status of vegetation and determining the severity of crops diseases. Non-imaging field spectrometers, such as Analytical Spectral Devices (Boulder, CO, USA), have been widely used to detect plant diseases. These spectrometers provide several advantages, including a wide wavelength range and high spectral resolution. However, there are still several disadvantages. For instance, the background cannot be eliminated effectively when the measurement is performed in open space. In an enclosed space with an integrating sphere, the field of view is fixed and limited to almost a point. Moreover, the infected regions on the leaves are scattered and uneven. As such, it is hard for non-imaging spectrometers to acquire a pure spectrum of a region with diseases. By comparison, hyperspectral imagers add these limitations through their photography function. The pure spectra of leaf regions with different disease severities can be precisely extracted from hyperspectral images, and the disease distribution can be mapped using image recognition and classification techniques. With these advantages, hyperspectral imagers are considered suitable instruments for studying plant diseases at the leaf and canopy scales. In our research, a portable hyperspectral imaging system was used to collect the hyperspectral images of apple leaves with mosaic disease. The regions where the spectra were extracted and where SPAD measurements were taken corresponded highly in the images. As a result, the spectral signatures of the specific infected region were analyzed (Figure 1) and accurate regression models were established (Table 5).

The determination of severity is an essential purpose of disease detection by remote sensing. The criteria used to grade disease severity in previous studies were usually based on leaf color, morphology or area percentage of infected regions within a leaf, which were determined via a visual estimation method. However, these methods depended on the experiences of observers and lacked a constant and objective standard. Therefore, misjudgement is inevitable. Our observation also revealed that, in certain cases, a large area of the infected region in a leaf was estimated visually, but the disease was in its early stage and the damage to the leaf was mild. While in other cases, a small area of the infected region was visually detected, but the disease was in a later stage and the damage to the leaf was serious. Figure 5 shows two leaves with obviously different area percentages of infected regions, as determined by visual estimation, but their leaf SPAD

![Figure 3](image-url)
values are 21.54 (leaf No.4) and 20.73 (leaf No.5) respectively, which indicate that their photosynthesis abilities as whole leaves are very similar. In such case, if the area percentage of infected regions is used as the only disease indicator, the true influence of the disease can be revealed partially but not comprehensively, and the physiological and biochemical mechanisms cannot be quantitatively described. Therefore, the SPAD value and the area percentage of the infection severity of the whole leaf, should be considered as vital indicators to determine plant disease severity.

![Figure 5](Image)

Figure 5 A comparison of two leaves with different infected patterns but similar average SPAD values

5 Conclusions

SPAD values can accurately indicate the chlorophyll content of plant leaves and thus can be used as a quantitative indicator for the leaf health. In this study, the hyperspectral images of apple leaves with mosaic disease were acquired by using a portable imaging spectrometer. Spectral analysis results revealed that leaf regions with different disease levels significantly differed in the visible wavebands of the hyperspectral images. VIs were selected to establish the regression models for apple leaf SPAD values. GNDVI, a normalized combination of near-infrared and green-light wavebands, yielded the highest correlation with the SPAD values of apple leaves. SPAD estimation models with a single VI and multiple VIs as independent variables were prepared through single-element regression and PLSR. Among the models, the SPAD-PLSR model with five VIs as independent variables achieved the highest estimation accuracy. As such, this model was used as an optimal model for analyzing the hyperspectral images pixel by pixel, and obtained the SPAD distribution images of the leaves. Furthermore, the average of all pixel values on the SPAD value distribution image was considered to be the SPAD value of the whole leaf. In this manner, the disease level, health conditions, and photosynthesis of apple leaves could be quantitatively described. This approach could also be used in other instances in which plants are exposed to stress caused by diseases, insects, or water and nutrients.

In this paper, the severity of mosaic disease could be accurately estimated on a leaf scale by using the hyperspectral image. On the basis of this research, by considering the scale effect, it is possible to extend this method to a larger scale to evaluate the health status of a whole apple tree in future studies.

Acknowledgments

This work was supported by the National High Technology Research and Development Program of China (Grant No. 2013AA102401) and the Growth plan for young talents in Shanghai municipal agricultural system (Hu Nong Qing Zi (2018) No. 1-29). The authors are grateful to Xumei Wu, Qi Wang, Jiaoyang Yu, Lili Luo, Mingming You and Zhuruan Zhang for their help with the experiment.

[References]


Roberts P L, Wood K R. Effects of a severe (P6) and a mild (W) strain of cucumber mosaic virus on tobacco leaf chlorophyll, starch and cell ultrastructure. Physiological Plant Pathology, 1982; 21(1): 31, N3, 32–33, N5, 37.


