



ARTICLE

Detection and Discrimination of Tea Plant Stresses Based on Hyperspectral Imaging Technique at a Canopy Level

Lihan Cui¹, Lijie Yan¹, Xiaohu Zhao¹, Lin Yuan², Jing Jin³ and Jingcheng Zhang^{1,*}

¹College of Artificial Intelligence, Hangzhou Dianzi University, Hangzhou, 310018, China

²Zhejiang University of Water Resources and Electric Power, Hangzhou, 310018, China

³Zhejiang Agricultural Technology Extension Center, Hangzhou, 310020, China

*Corresponding Author: Jingcheng Zhang. Email: zhangjcrs@hdu.edu.cn

Received: 23 December 2020 Accepted: 21 January 2021

ABSTRACT

Tea plant stresses threaten the quality of tea seriously. The technology corresponding to the fast detection and differentiation of stresses is of great significance for plant protection in tea plantation. In recent years, hyperspectral imaging technology has shown great potential in detecting and differentiating plant diseases, pests and some other stresses at the leaf level. However, the lack of studies at canopy level hampers the detection of tea plant stresses at a larger scale. In this study, based on the canopy-level hyperspectral imaging data, the methods for identifying and differentiating the three commonly occurred tea stresses (i.e., the tea leafhopper, anthrax and sun burn) were studied. To account for the complexity of the canopy scenario, a stepwise detecting strategy was proposed that includes the process of background removal, identification of damaged areas and discrimination of stresses. Firstly, combining the successive projection algorithm (SPA) spectral analysis and K-means cluster analysis, the background and overexposed non-plant regions were removed from the image. Then, a rigorous sensitivity analysis and optimization were performed on various forms of spectral features, which yielded optimal features for detecting damaged areas (i.e., YSV, Area, GI, CARI and NBNDVI) and optimal features for stresses discrimination (i.e., MCARI, CI, LCI, RARS, TCI and VOG). Based on this information, the models for identifying damaged areas and those models for discriminating different stresses were established using K-nearest neighbor (KNN), Random Forest (RF) and Fisher discriminant analysis. The identification model achieved an accuracy over 95%, and the discrimination model achieved an accuracy over 93% for all stresses. The results suggested the feasibility of stress detection and differentiation using canopy-level hyperspectral imaging techniques, and indicated the potential for its extension over large areas.

KEYWORDS

Hyperspectral imaging technology; tea plant; diseases and pests; sunburn; spectral analysis

1 Introduction

Tea is one of the most popular healthy drinks. Owing to the healthy benefits such as lowering the risks from heart disease, cancer, and diabetes, the demand of tea is rising around the world. As the largest tea producing country, China grew 2.93 million hectares of tea in 2018 [1,2]. Given that the quality of fresh tea leaves is decisive to the quality of the final tea product, management of the tea plantation is critical.



Stresses of tea plants such as diseases, pests, and sunburns are some important restrictive factors that severely threaten the tea quality [3]. Therefore, the rapid and effective detection and differentiation of these tea plant stresses is of urgent demand to guide proper prevention measures in tea plantations.

At present, the investigation of tea plant stresses is mainly based on field inspection, which is inefficient, prone to be subjective, and difficult to be applied over large areas. As a nondestructive detection technique, hyperspectral sensing is able to respond the changes of plant morphology, material composition, physiological and biochemical status, which thereby has a great potential in detecting diseases, pests and other stresses in tea plantations. Some attempts have shown the capability of hyperspectral observations in detecting plant diseases and pests. Jone et al. [4] used hyperspectral data in determining the disease severity of leaf spot (*Xanthomonas perforans*) on tomato leaves. Combining the partial least squares (PLS) regression, correlation analysis, and stepwise multiple linear regression (SMLR) procedure, significant wavelengths were identified to construct the prediction models. The best result produced a relatively high prediction accuracy with an R^2 of 0.82. Tian et al. [5] found that the hyperspectral imaging observations could be used to detect cucumber downy mildew, and the accuracy reached 90%. In detecting one biotic stress (*Venturia inaequalis*) in apple trees, Delalieux et al. [6] applied logistic regression, partial least squares logistic discriminant analysis, and tree-based modeling on the hyperspectral data, and developed robust stress spectral indicators. Mahlein et al. [7] successfully used hyperspectral signals in differentiating three leaf diseases (i.e., *Cercospora* leaf spot, sugar beet rust and powdery mildew) on sugar beet plants. A comprehensive spectral analysis was conducted with a RELIEF-F algorithm, which yielded the optimal bands and constructed specific disease indexes for discrimination of diseases. To detect tea diseases, Yuan et al. [8] selected three sensitive bands (542, 686 and 754 nm) for tea anthrax based on hyperspectral imaging technology, and constructed two vegetation indexes: The tea anthracnose ratio index (TARI) and the tea anthracnose normalization index (TANI). Based on the features above, they propose a discrimination method that combined the unsupervised classification and two-dimensional adaptive threshold methods. The overall accuracy of the models reached up to 98%.

Currently, studies about plant stress detection with hyperspectral data are mainly carried out at a leaf scale [9–11]. However, to achieve the large-scale inspection of tea stresses, it is required to develop the detection methods at a canopy scale. However, different from the relatively pure background in the leaf images, the canopy images include a large amount of variation about plant morphology, canopy structure, and the interaction between light and the plants. The complexity in the canopy images poses a big challenge in the stress detection. Huang et al. [12] compared the spectral responses of the rice leaf folder (*Cnaphalocrocis medinalis*) disease at a leaf and canopy scales, and found that despite the spectral response at the canopy scale is weaker than that at the leaf scale, some spectral bands and features still exhibited sensitivity at the canopy scale. Based on the canopy spectral measurements, Zhang et al. [13] discriminated two different stresses with similar symptoms (i.e., the stripe rust disease and nitrogen stress) in winter wheat. Besides, Behmann et al. [14] used hyperspectral imaging data to analyze the early drought stress in barley. The results suggested that by integrating image analysis and spectral analysis, the method can not only detect the stress at an early stage, but also identify the specific positions of the damaged areas in the plants' canopy. Given that the canopy level stress detection methods can potentially be used in near ground UAV sensing platforms and achieve the fast scan and inspection of plant stresses in tea gardens, it is necessary to conduct experimental analysis based on canopy level hyperspectral imaging data. However, relevant works on this matter are still very rare at present.

As an important tea producing region in China, Zhejiang province is famous for its high quality green tea. However, some famous varieties such as Longjing-43 and Anji Baiye-1 are severely affected by plant diseases, pests and sunburn stress. In this work, based on the canopy level hyperspectral imaging data and several statistical and machine learning approaches, we attempt to develop methods for identifying

and differentiating plant damages caused by three frequent tea plant stresses [i.e., anthracnose disease (AH), green leafhopper (GL) and sunburn (BU)]. The objectives of this study includes: (1) To identify the optimal spectral bands and features for detecting and differentiating the different stresses; (2) To construct a stepwise protocol for detecting stress damage areas differentiating their types under the complicated canopy scenario; (3) To validate the established method against the experimental data.

2 Materials and Methods

2.1 Materials

The target tea stresses of this study (i.e., the tea leafhoppers, anthrax and sunburns) often occur in tea gardens in Zhejiang province. The three stresses have relatively similar symptoms such as the grayish-brown lesions on leaves. However, given that the prevention of the different stresses require different procedures, it is important to differentiate them for precise management. This experiment was carried out in the experimental tea field of the Chinese Academy of Agricultural Sciences (Zhejiang) in June 2019. The disease, pest and sunburn stresses occurred naturally in the studied field. Based on a field survey, three locations corresponding to the three stresses were picked by an experienced plant pathologist to conduct *in-situ* canopy hyperspectral imaging measurements.

2.2 Hyperspectral Imaging Measurements at a Canopy Scale

The hyperspectral images were obtained using an UHD185 imaging spectrometer developed by Cubert Company in Germany (<http://cubert-gmbh.com/>). The imaging spectrometer is a full-frame, snap-shot device, which have 126 bands within 450–950 nm. The spectrometer has a spectral sampling interval of 4 nm, and an image resolution of 1000 × 1000 pixels. The spectral measurements were made between 10:00 AM and 2:00 PM. The spectral images were taken at a distance of 50 cm above from the top of the tea canopy. The image radiometric correction was performed to convert the reflectance using the reference whiteboard and blackboard. For each type of stress, 40–50 spectral images were taken. After quality control, 30 hyperspectral images were used for subsequent analysis for each stress (Fig. 1).

2.3 Development of the Stresses Detection Procedure

2.3.1 Selection of ROIs

To study the spectral characteristics of the different tea stresses, it is necessary to select the region of interest (ROI) to extract spectral information for analysis. Considering the complexity of the canopy image scenario, three types of ROIs were defined, including: Background region (BR), healthy leaf region and scab region (SR). Given that the wax coat of tea leaves tends to cause overexposure in the images, classes of overexposure regions (OR) were specifically defined. Besides, due to the leaf inclination and mutual occlusion, the healthy leaf region was under varied illumination conditions, which thus hampered the detection of stresses. To address this issue, within the healthy leaf region (HR), three types of ROIs were defined, including the (1) Highly illuminated leaf region (HLR), (2) Shadow leaf region (SLR) and (3) Normal-level illuminated leaf region (NLR). All the ROIs have the size of 20 × 20 pixels. For each type of the ROI classes, 3 ROIs were selected from each image.

2.3.2 Selection of Spectral Features for Stress Detection

To facilitate the stepwise detection and differentiation of the tea stresses, three sets of spectral features were selected corresponding to each step, including: (1) Spectral bands for differentiating plant areas and the background; (2) Spectral features for identifying anomalous areas in the tea canopy; (3) Spectral features for differentiating stresses.

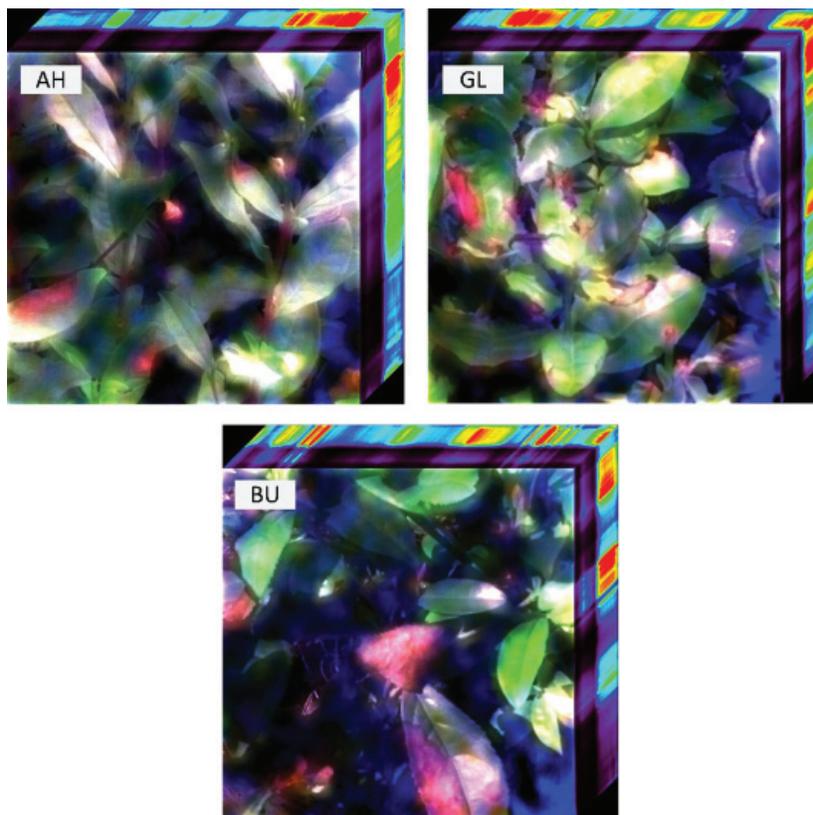


Figure 1: Hyperspectral images of the three tea plant stresses (AH, anthrax disease; GL, leafhopper; BU, sun burn)

(1) Selection of spectral bands for differentiating plant areas and the background

The selection of spectral bands applied the successive projection algorithms (SPA). The SPA is able to generate a combination of wavelengths with the least information redundancy, which thus minimize the collinearity among variables and reduce the dimension of features [15]. In this study, the number of bands in one combination was set as 5–10, and the accuracy of the classification was used as the criteria for bands selection.

(2) Selection of spectral features for identifying anomalous areas in the tea canopy

In this step, the three types of spectral features were included to form a candidate spectral feature set, which included the spectral derivative features, continuum removal features and vegetation indexes. These spectral features were designed to emphasize the spectral response to the plant physiological and biochemical changes (Tab. 1). The selection of spectral features adopted the *t*-test and autocorrelation analysis. Firstly, a *t*-test was used to test the difference between stressed (including all AH, GL, and BU stresses) and healthy samples, with the criteria of *p*-value < 0.001. Then, a pairwise cross-correlation was applied to eliminate features highly correlated with each other. In this process, a criteria of $R^2 < 0.8$ was applied to further reduce the information redundancy among the sensitive spectral features.

Table 1: Introduction of spectral features used in this study

Spectral Feature	Definition	Position/Formula	References
Spectral derivative features			
BMV	Maximum differential value	Blue edge (490–540 nm)	Gong et al. 2002 [16]
BPMV	Position of the maximum differential value		
BSV	Sum of differential values		
YMV	Maximum differential value	Yellow edge (540–620 nm)	
YPMV	Position of the maximum differential value		
YSV	Sum of differential values		
RMV	Maximum differential value	Red edge (660–780 nm)	
RPMV	Position of the maximum differential value		
RSV	Sum of differential values		
Continuum removal features			
Depth	The minimal value around the concave region of the normalized curve within 530–770 nm	Near infrared (530–770 nm)	Pu et al. 2003 [17]
Width	The width of the spectral band at the half of the absorption depth		
Area	The area of the absorption feature continuum removed spectral curve		
Vegetation indexes			
ATSAVI	Adjusted Transformed Soil-Adjusted VI	$a(R_{800} - aR_{670} - b) / [aR_{800} + R_{670} - ab + X(1 + a^2)]$ $X = 0.08, a = 1.22, \text{ and } b = 0.03$	Baret and Guyot. 1991 [18]
GI	Greenness Index	R_{554} / R_{677}	Zarco-Tejada et al. 2005 [19]
NBNDVI	Narrow Band Normalized Difference Vegetation Index	$(R_{850} - R_{680}) / (R_{850} + R_{680})$	Rouse et al. 1973 [20]
NDVI	Normalized Difference Vegetation Index	$(R_{NIR} - R_{red}) / (R_{NIR} + R_{red})$	Rouse et al. 1973 [20]
OSAVI	Optimized Soil-Adjusted Vegetation Index	$1.16(R_{800} - R_{670}) / (R_{800} + R_{670} + 0.16)$	Rondeaux et al. 1996 [21]
PRI	Photochemical Reflectance Index	$(R_{531} - R_{570}) / (R_{531} + R_{570})$	Gamon et al. 1992 [22]
CARI	Chlorophyll Absorption Ratio Index	$\left(\frac{(670a + R_{670} + b)}{(a^2 + 1)^{1/2}} \right) (R_{700} / R_{670})$ $a = (R_{700} - R_{550}) / 150, b = R_{550} - 550a$	Kim et al. 1994 [23]
CI	Chlorophyll Index	$(R_{750} - R_{705}) / (R_{750} + R_{705});$ $R_{750} / (R_{700} + R_{710}) - 1$	Gitelson and Merzlyak. 1996 [24] Gitelson et al. 2005 [25]
CRI	Carotenoid Reflectance Index	$CR_{I550} = (R_{510}) - 1 - (R_{550}) - 1;$ $CR_{I700} = (R_{510}) - 1 - (R_{700}) - 1$	Gitelson et al. 2002 [26]
LCI	Leaf Chlorophyll Index	$(R_{850} - R_{710}) / (R_{850} + R_{680})$	Datt. 1999 [27]
MCARI	Modified Chlorophyll Absorption in Reflectance Index	$[(R_{701} - R_{671}) - 0.2(R_{701} - R_{549})] (R_{701} / R_{671})$	Daughtry et al. 2000 [28]
PSRI	Plant Senescence Reflectance Index	$(R_{680} - R_{500}) / R_{750}$	Merzlyak et al. 1999 [29]

(3) Selection of spectral features for differentiating three plant stresses

To select spectral features that could distinguish the three types of stresses, t-tests were conducted between each of any of the two stress type pairs (i.e., GL vs. AH, BU vs. AH and GL vs. BU), respectively. Using the criteria of p -value < 0.05 , the spectral features that passed all three t-tests were selected, ensuring the features that have the capability in differentiating all three stresses. Then, a cross-correlation analysis was performed to remove spectral features with a high level of information redundancy, and finally form a feature set for distinguishing the stresses.

2.3.3 Construction of a Stepwise Stress Detection and Discrimination Procedure

Considering the complexity of the tea plant canopy images, a stepwise procedure was developed for detecting and discriminating tea stresses based on the hyperspectral imaging data. The procedure includes the steps as: Removal of non-plant background, identification of plant damaged areas and discrimination of the different stresses. The schematic diagram of the workflow is shown in Fig. 2.

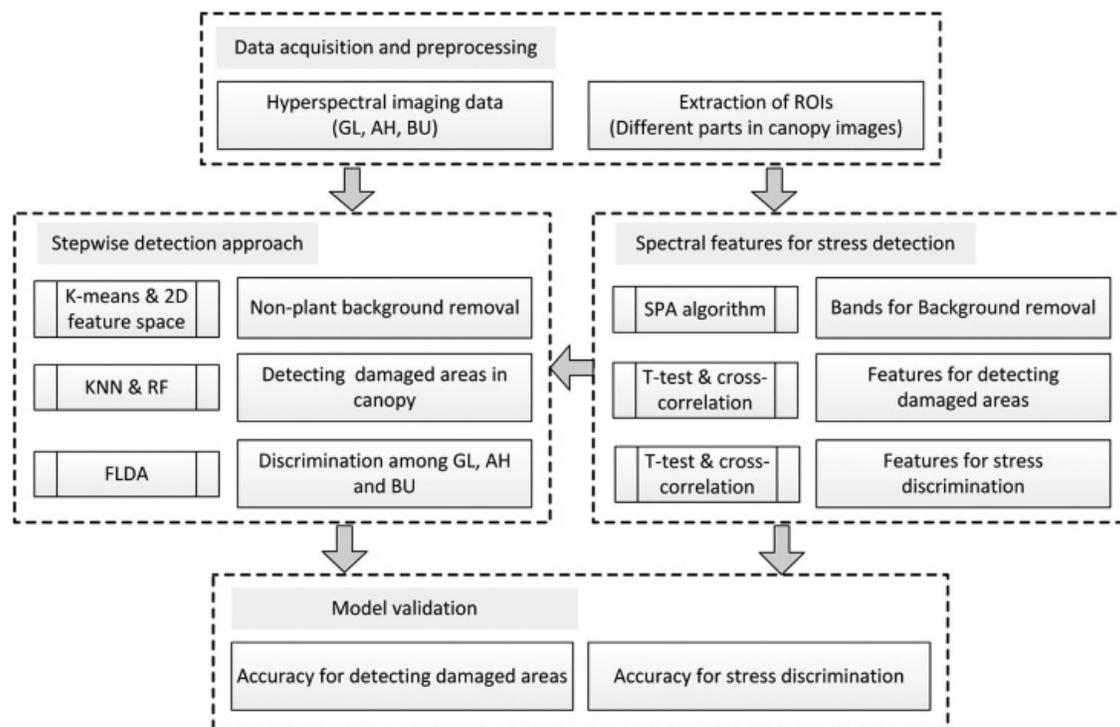


Figure 2: Workflow of data analysis

Firstly, to remove the non-plant background area, the K-means method was applied on a couple of sensitive spectral bands to generate the clustering result of the image. By analyzing the clustering results in the Red-NIR spectral feature space, a threshold method was used to remove those non-leaf areas [8]. To determine the most appropriate threshold value for Red and NIR bands, twenty evenly spaced intervals were traversed from the minimum to maximum. The threshold corresponding to the highest accuracy was determined and applied for background removal.

Secondly, a model for detecting damaged areas was established based on the corresponding optimal spectral features and two classic machine learning classifiers, the K-Nearest Neighbor (KNN) and Random Forest (RF). As a non-parametric learning algorithm, the KNN assigns samples to the class to which the majority of the K data points belong based on feature distance. The algorithm for RF is a supervised

approach that made up the decision trees. It uses bagging and feature randomness when building individual decision trees to try to create an uncorrelated forest whose prediction is more accurate than that of any individual tree. The performance of the two classifiers in detecting damaged areas was compared.

Finally, a discrimination model was established based on the corresponding optimal spectral features and the Fisher linear discrimination model (FLDA). As a classic classification approach, the FLDA is able to generate linear combination of spectral features and construct a linear equation for classification. Within the damaged areas, the type of stress was thus separated with the FLDA.

2.4 Accuracy Evaluation

In this study, all hyperspectral images were randomly divided into two parts: 50% of data were used for training, and the other 50% for verification. According to the confusion matrix, the overall Accuracy (OA) and kappa coefficient were calculated as accuracy indicators. In this study, hyperspectral image data processing was completed in ENVI 5.3 (Exelis Visual Information Solutions, USA), and data analysis and modeling was completed in Matlab 2014 (MathWorks, USA).

3 Results and Discussion

3.1 Spectral Characteristic of Tea Plant Diseases and Pests

By observing and comparing the spectral curves that were extracted from the ROIs of GL, AH, BR stressed hyperspectral images, it was found that the spectral characteristic follow the same pattern across different types of ROIs. In all types of stress images, the overexposed region (OR) and the background region (BR) have the highest and lowest spectral reflectance in most bands, respectively. The reflectance of highly illuminated leaf region (HLR) and shadow leaf region (SLR) for the healthy leaf area (HR) were slightly higher and lower, respectively, than those of the normal-level illuminated leaf region (NLR) (Fig. 3).

In comparing the spectral reflectance between healthy and damaged leaf regions, it was found that the reflectance significantly increased in the red band (620–660 nm) and decreased in the near infrared platform (780–900 nm) for the damaged regions, and the shape of the entire spectral curves tend to be relatively flat. Such spectral changes might be associated with the damage to the pigment system and destruction of cellular structure [30]. Among the three stresses, GL and AH showed similar spectral curves, whereas the spectral curve of BU deviated from that of the healthy leaf regions at a larger extent, which indicates the different extent between biological and physical destruction. As a physical damage, BU usually causes relatively wide and severe damage to plant tissues, while the damage caused by disease infection and pest infestation usually presented a gradual damage process with a relatively limited extent. The spectral difference among the different ROIs forms an important basis for subsequent detection of canopy damaged areas and discrimination of stresses.

3.2 Selection of Optimal Spectral Features for Tea Stress Detection

For separating tea plants from the background in the canopy hyperspectral images, a total of seven spectral bands were selected through the SPA algorithm, including: 460 nm, 550 nm, 670 nm, 706 nm, 742 nm, 893 nm and 920 nm. These bands were distributed at some important positions in the spectral curve of plants (Fig. 4). For example, 460 nm, 550 nm and 670 nm were located around the blue edge, green peak, and red valley regions. These locations are sensitive to changes of some important biochemical components such as the content of chlorophyll, which may have a significant difference between healthy and stressed plant regions. The 706 nm and 742 nm located around the red-edge region, which are sensitive to the stress status of plants. As the SPA algorithm adopted the combinatorial optimization strategy, the optimized bands have a relatively low level of correlation among each other, which is carried out in the background removal process.

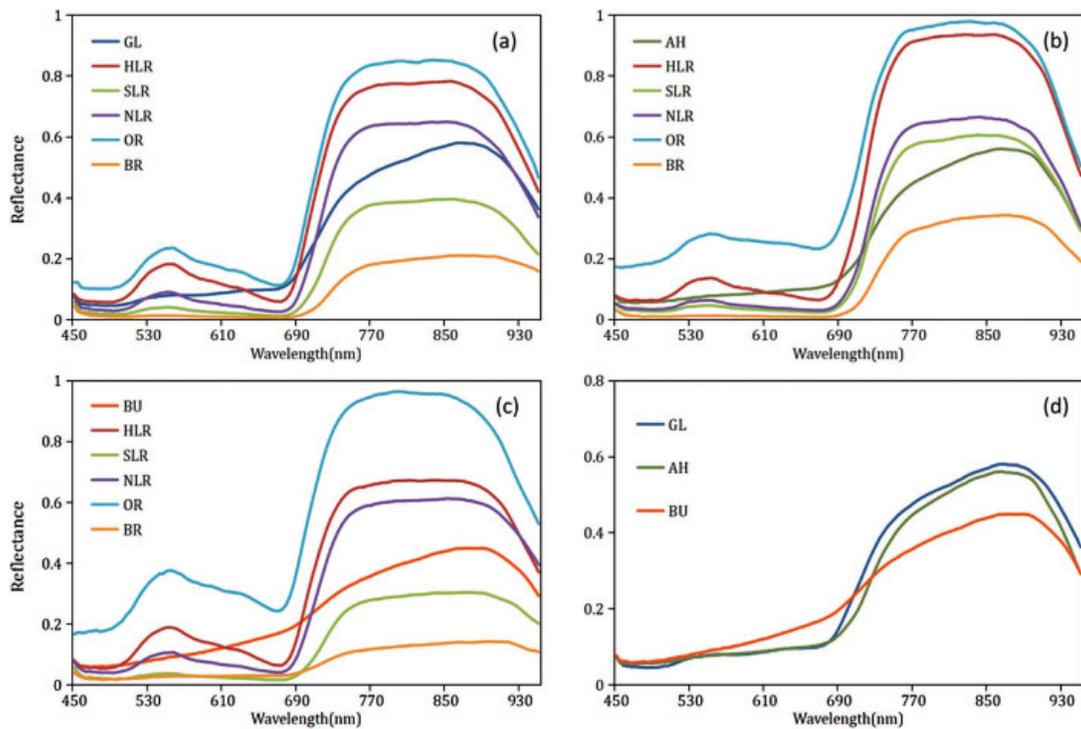


Figure 3: Reflectance spectral curves of different regions in images of 3 tea stresses. (a) Comparison of averaged spectral curves among different ROIs in GL images; (b) Comparison of averaged spectral curves among different ROIs in AH images; (c) Comparison of averaged spectral curves among different ROIs in BU images; (d) Comparison of averaged spectral curves among the damaged ROIs of the 3 stresses (i.e., GL, AH, BU)

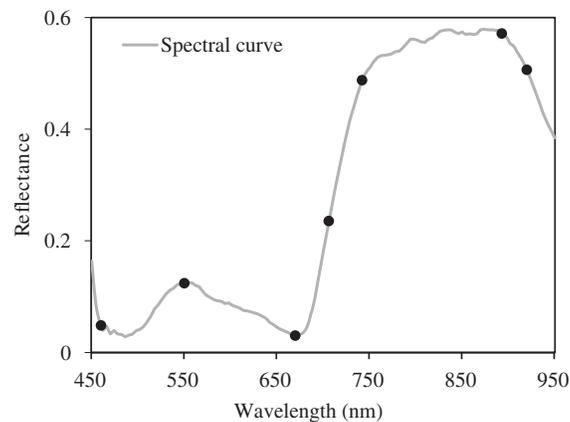


Figure 4: The locations of the optimal bands selected by SPA algorithm

In selecting features for detecting damaged areas, the *t*-test analysis and the cross-correlation analysis were applied on the three types of candidate spectral features (i.e., spectral derivative features, continuum removal features and vegetation indexes). This yielded a total of 5 spectral features: YSV, Area, GI, CARI and NBNDVI. Among them, the spectral derivative feature YSV captures the degree of spectral variation within 540–620 nm, while the continuum removal feature Area indicates the intensity of

spectral absorption within 530–770 nm. These two spectral shape-based features mainly reflect the variation of pigment contents. The other three vegetation indexes indicated the greenness and the chlorophyll content of plants. Given that the content of these pigments will significantly decrease when plants are under stress, the identified spectral features can (1) Serve as indicators of the health status of plants, and (2) Be used for detecting the damaged areas in the hyperspectral images (Tab. 2).

Table 2: Spectral features for detecting and discriminating tea stresses

Spectral index	<i>T</i> -test (<i>p</i> -value < 0.001)			Spectral index	<i>T</i> -test (<i>p</i> -value < 0.05)		
	GL & HR	AH & HR	BU & HR		GL & AH	GL & BU	AH & BU
BMV				BMV			
BPMV				BPMV			
BSV				BSV		+	
YMV				YMV			
YPMV		+		YPMV			+
YSV	+	+	+	YSV	+		+
RMV		+	+	RMV	+		+
RPMV				RPMV		+	
RSV			+	RSV	+		+
Depth	+	+	+	Depth	+		+
Width	+	+	+	Width	+		+
Area	+	+	+	Area	+		+
GI	+	+	+	GI	+		+
PRI	+	+	+	PRI	+		+
CARI	+	+	+	CARI	+		+
MCARI		+		MCARI	+	+	+
PSRI	+	+	+	PSRI	+		+
OSAVI	+	+	+	OSAVI	+		+
ATSAVI	+	+	+	ATSAVI	+		+
NDVI	+	+	+	NDVI	+		+
CI	+	+	+	CI	+	+	+
CRI				CRI	+	+	
LCI				LCI	+	+	+
MTCI				MTCI		+	+
RARS		+	+	RARS	+	+	+
RGR	+	+	+	RGR	+		+
TCI				TCI	+	+	+
VOG1				VOG1	+	+	+
NBNDVI	+	+	+	NBNDVI	+		+

Note: The “+” indicates the indexes satisfying the significant difference in the corresponding *t*-test. The indexes in dash box indicate the final selected indexes with cross-correlation analysis.

In addition, the same strategy was applied on the selection of spectral features for discriminating different stresses, which yielded 6 optimal spectral features: MCARI, CI, LCI, RARS, TCI and VOG. These features are all vegetation indexes that associated with the contents of chlorophyll and carotenoid (Tab. 2). Given that different stresses can cause subtle differences of damage to the pigment system in plant leaves, the spectral responses are able to be captured by the features and used for discriminating these stresses.

3.3 Detection and Discrimination of Tea Plant Stresses

Firstly, and based on the optimal bands generated by the SPA algorithm, the non-plant background needs to be removed from the tea canopy hyperspectral images. The K-means algorithm was applied on the images of the selected bands to conduct the clustering on the image pixels. The pixels belonging to each cluster were then projected to the two-dimensional feature space that consisted in the reflectance of the RED (630–690 nm) and near-infrared (NIR; 760–900 nm) bands. It was observed that the background area was located in the part with relatively low RED and NIR values, whereas the overexposed area was located in the part with relatively high RED and NIR values. Through an iterative-optimized thresholding approach, the background and overexposed area were removed so that the subsequent analysis was focused on the plant area in the images (Fig. 5). Comparing with the traditional pixel-based classification methods, the clustering-based method effectively mitigated the edge irregularities and the “salt and pepper” phenomenon in the classification results [8].

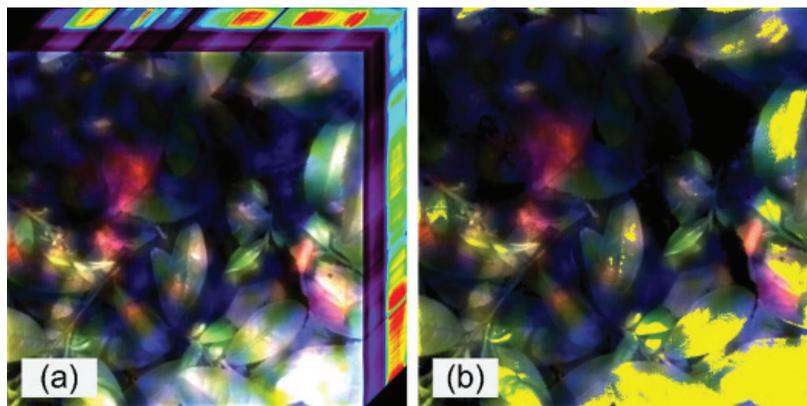


Figure 5: Demonstration of the non-plant background removal process (In the result image, the overexposed and background areas are highlighted in yellow and black)

Within the plant area of the image, based on the selected optimal spectral features that were suitable for damage detection, both KNN and RF were applied for constructing the model for detecting the damaged areas on tea plants (Tab. 3). In this process, unlike the usual pixel-based analysis, the modeling process also took the K-means clustering results as basic units. The detecting models were validated with the validation image data, and the results showed that the KNN outperformed the RF model in general. The KNN achieved relatively high accuracy for all three stresses, with an OA over 95%. While for RF, the performance differed significantly among the stresses, with a relatively high accuracy for GL (OA = 98%), and moderate accuracies for AH (OA = 86%) and BU (OA = 83%). The difference of the model performance can also be observed from a couple of examples as shown in Fig. 6. It is noted that the results of the RF models included several over-classified regions, which was the main reason for their relatively low accuracy. The significant spectral difference between the stressed and healthy samples may provide a solid basis for the classification using the KNN model, which adopted the straightforward

feature distance criteria. The simple structure and low level of computational complexity makes it a promising solution in practice.

Table 3: The accuracies of detecting damaged areas in GL images

Image code	KNN		RF	
	OA	Kappa	OA	Kappa
GL	98.72%	0.58	98.26%	0.55
AH	98.50%	0.64	85.93%	0.17
BU	95.36%	0.67	83.15%	0.31

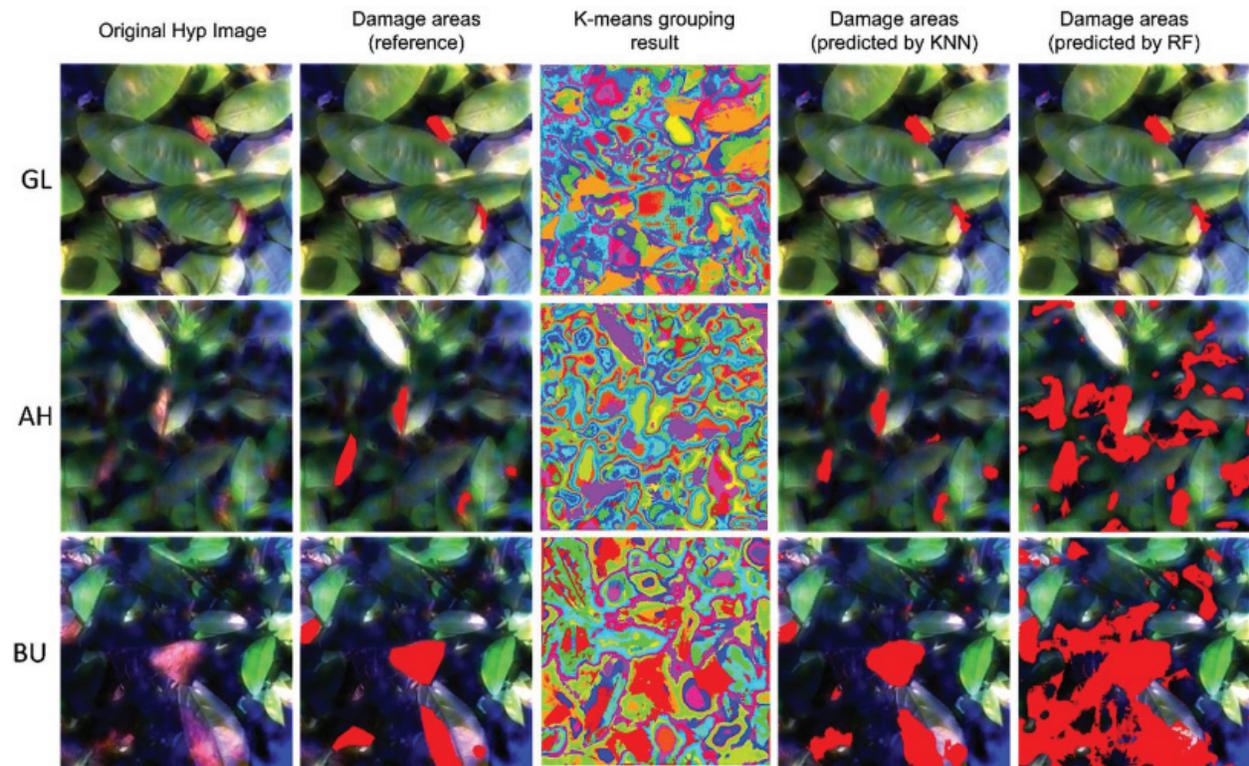


Figure 6: Demonstration of the damage detection on different plant stresses and different classifiers

For stress discrimination, after comparing with the visual recognition results, the Fisher discriminant model achieved relatively high accuracies for all the three types of stresses. In them, the discrimination accuracy reached 98.53% and 98.29% for GL and AH, respectively. The accuracy was slightly lower for BU, with an OA of 93.94%. Despite the three stresses were similar on symptoms, it is encouraging that the subtle differences were effectively captured by the spectral signals (Fig. 7). When comparing the present study with some studies that were conducted at a leaf level, the complicated background scenarios at the canopy level posed a big challenge in detecting and differentiating the diseases and pest. However, by adopting the conception of stepwise detection, the final discriminant analysis can be carried out in a relatively pure scenario (i.e., the damaged area), which is beneficial for improving the performance and stability of the model. It is of great importance to incorporate this canopy-level tea

stress detecting and discriminating method in some autonomous driving platforms such as Unmanned Aerial Vehicle (UAV) that are equipped with hyperspectral or low-cost customized-bands cameras. More studies and experiments need to be carried out to verify the effectiveness of corresponding methods in large-scale detection of stresses in tea plantations.

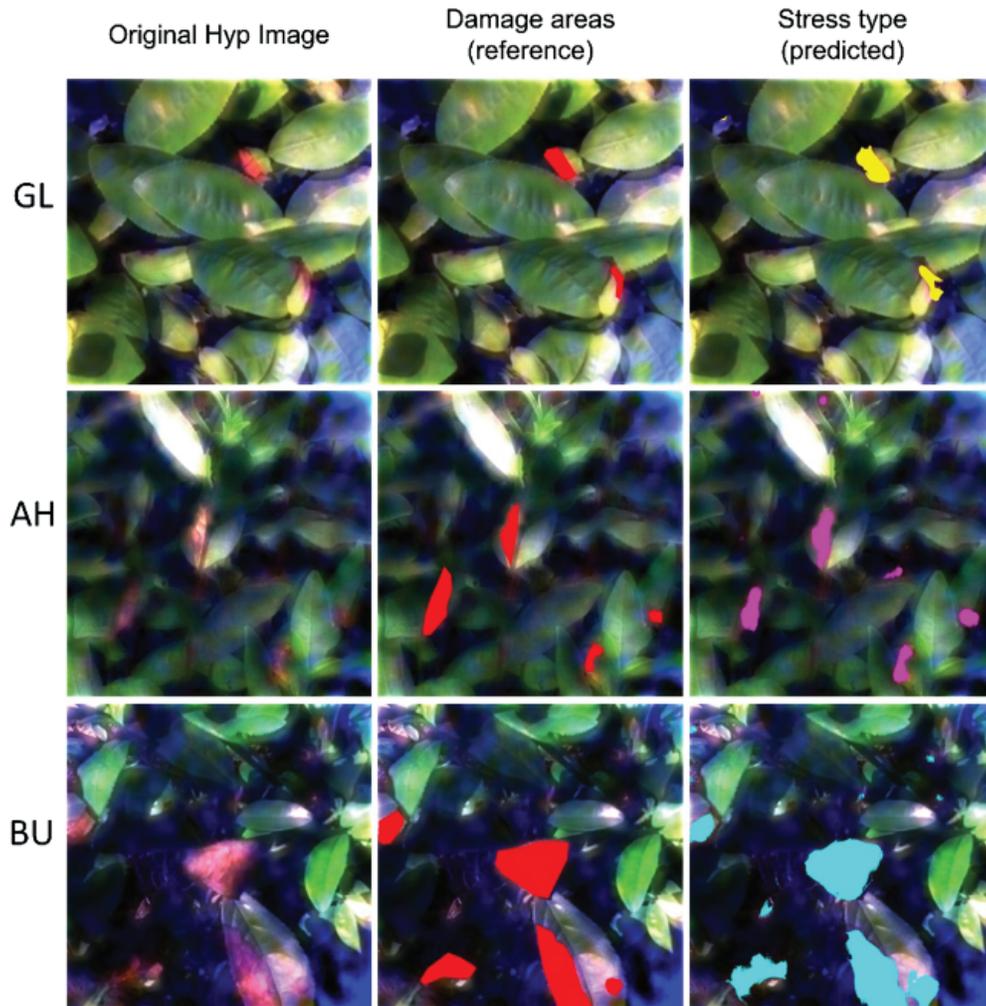


Figure 7: Demonstration of the tea plant stress discrimination (The areas highlighted in red in the second column indicate damaged areas; the third column showed the discriminated results, with GL in yellow, AH in pink and BU in lake blue)

4 Conclusion

Aiming at detecting and discriminating among three commonly occurring stresses on tea plants (i.e., plant disease, pest and sunburn) at a canopy scale, this study proposed a stepwise procedure based on the canopy level hyperspectral imaging data and spectral analysis, image processing, and machine learning approaches. The main findings included:

1. Despite the structural complexity of the canopy images of tea plants, significant spectral differences were evident among different regions. Through comprehensive spectral sensitivity analysis, the optimal bands for background removal were identified at 460 nm, 550 nm, 670 nm, 706 nm,

742 nm, 893 nm and 920 nm; the optimal spectral features for damaged area detection were identified as YSV, Area, GI, CARI, and NBNDVI; and the spectral features for stress discrimination were identified as MCARI, CI, LCI, RARS, TCI, and VOG.

2. Based on the selected spectral features, a non-plant background removal method was proposed combining K-means and two-dimensional (Red-NIR) spectral feature space analysis. Then, a machine learning based model was constructed to identify the damaged areas in the tea canopy. Finally, a stress discrimination model was established with the Fisher linear discriminant method.
3. The validation results suggested that the proposed procedure is able to achieve a relatively high accuracy. The accuracy of damage area detection reached 95%, and the stress discrimination accuracy reached 98%. Our results confirmed the feasibility in detecting and differentiating among GL, AH and BU stresses in tea plants at a canopy scale.

Different from the stress detecting studies that were conducted at a leaf scale, the present study showed a possibility of using canopy-level hyperspectral imaging data in detecting plant stresses. The proposed method may provide a basis for detecting anomalous areas in tea gardens or orchards over large areas.

Acknowledgement: The authors are grateful to Dr. Huawei Guo and Prof. Wenyan Han for their assistance in experimental design and data collection.

Funding Statement: This work was supported by Zhejiang Public Welfare Program of Applied Research (LGN19D010001), Zhejiang Agricultural Cooperative and Extensive Project of Key Technology (2020XTTGCY04-02; 2020XTTGCY01-05), and the National Key R&D Program of China (2017YFE0122500).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Zheng, G. J., Gao, H. Y. (2015). The status analysis of tea quality safety in China. *Journal of Food Safety & Quality*, 6(7), 2869–2872.
2. Xu, Y. M., Lu, W. J., Lv, W. X. (2019). An empirical analysis of the impact of China's tea export trade on the development of China's tea industry. *Journal of Tea*, 45(4), 185–192.
3. Qi, L. C., Zhang, Y. D. (2016). Major tea plant diseases and their control. *Journal of Tea*, 42(1), 10–12.
4. Jones, C. D., Jones, J. B., Lee, W. S. (2010). Diagnosis of bacterial spot of tomato using spectral signatures. *Computers and Electronics in Agriculture*, 74(2), 329–335. DOI 10.1016/j.compag.2010.09.008.
5. Tian, Y. W., Zhang, L. (2012). Study on the methods of detecting cucumber downy mildew using hyperspectral imaging technology. *Physics Procedia*, 33(1), 337–354. DOI 10.1016/j.phpro.2012.05.130.
6. Delalieux, S., van Aardt, J., Keulemans, W., Schrevens, E., Coppin, P. (2007). Detection of biotic stress (*Venturia inaequalis*) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications. *European Journal of Agronomy*, 27(1), 130–143. DOI 10.1016/j.eja.2007.02.005.
7. Mahlein, A., Rumpf, T., Welke, P., Dehne, H., Plümer, L. et al. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment*, 128(4), 21–30. DOI 10.1016/j.rse.2012.09.019.
8. Yuan, L., Yan, P., Han, W. Y., Huang, Y., Wang, B. et al. (2019). Detection of anthracnose in tea plants based on hyperspectral imaging. *Computers and Electronics in Agriculture*, 167, 105039. DOI 10.1016/j.compag.2019.105039.
9. Mahlein, A. K., Rumpf, T., Welke, P., Dehne, H. W., Plümer, L. et al. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment*, 128(4), 21–30. DOI 10.1016/j.rse.2012.09.019.
10. Baranowski, P., Jedryczka, M., Mazurek, W., Babula-Skowronska, D., Siedliska, A. et al. (2015). Hyperspectral and thermal imaging of oilseed rape (*Brassica napus*) response to fungal species of the genus *alternaria*. *PLoS One*, 10(3), e0122913. DOI 10.1371/journal.pone.0122913.

11. Xie, C., Yang, C., He, Y. (2017). Hyperspectral imaging for classification of healthy and gray mold diseased tomato leaves with different infection severities. *Computers and Electronics in Agriculture*, 135(2), 154–162. DOI 10.1016/j.compag.2016.12.015.
12. Huang, J., Liao, H., Zhu, Y., Sun, J., Sun, Q. et al. (2012). Hyperspectral detection of rice damaged by rice leaf folder (*Cnaphalocrocis medinalis*). *Computers and Electronics in Agriculture*, 82(3), 100–107. DOI 10.1016/j.compag.2012.01.002.
13. Zhang, J., Pu, R., Huang, W., Yuan, L., Luo, J. et al. (2012). Using *in-situ* hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. *Field Crops Research*, 134, 165–174. DOI 10.1016/j.fcr.2012.05.011.
14. Behmann, J., Steinrücken, J., Plümer, L. (2014). Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93(4), 98–111. DOI 10.1016/j.isprsjprs.2014.03.016.
15. Milanez, K. D. T. M., Nóbrega, A., César, T., Silva Nascimento, D., Galvo, R. K. H. et al. (2017). Selection of robust variables for transfer of classification models employing the successive projections algorithm. *Analytica Chimica Acta*, 984, 76–85. DOI 10.1016/j.aca.2017.07.037.
16. Gong, P., Pu, R. L., Heald, R. C. (2002). Analysis of *in situ* hyperspectral data for nutrient estimation of giant sequoia. *International Journal of Remote Sensing*, 23(9), 1827–1850. DOI 10.1080/01431160110075622.
17. Pu, R. L., Ge, S., Kelly, N. M., Gong, P. (2003). Spectral absorption features as indicators of water status in coast live oak (*Quercus agrifolia*) leaves. *International Journal of Remote Sensing*, 24(9), 1799–1810. DOI 10.1080/01431160210155965.
18. Baret, F., Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35(2–3), 161–173. DOI 10.1016/0034-4257(91)90009-U.
19. Zarco-Tejada, P. J., Berjón, A., López-Lozano, R., Miller, J. R., Martín, P. et al. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99(3), 271–287. DOI 10.1016/j.rse.2005.09.002.
20. Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, 1, 48–62.
21. Rondeaux, G., Steven, M., Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2), 95–107. DOI 10.1016/0034-4257(95)00186-7.
22. Gamon, J. A., Penuelas, J., Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 41(1), 35–44. DOI 10.1016/0034-4257(92)90059-S.
23. Kim, M. S., Daughtry, C. S. T., Chappelle, E. W., McMurtrey, J. E. (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (APAR). *Proceedings of the 6th International Symposium on Physical Measurements and Signatures in Remote Sensing*, pp. 299–306. France: Val d'Isere.
24. Gitelson, A. A., Merzlyak, M. N. (1996). Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology*, 148(3–4), 494–500. DOI 10.1016/S0176-1617(96)80284-7.
25. Gitelson, A. A., Viña, A., Ciganda, C., Rundquist, D. C., Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, 32(8), L08403. DOI 10.1029/2005GL022688.
26. Gitelson, A. A., Kaufman, Y. J., Stark, R., Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80(1), 76–87. DOI 10.1016/S0034-4257(01)00289-9.
27. Datt, B. (1999). A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using Eucalyptus leaves. *Journal of Plant Physiology*, 154(1), 30–36. DOI 10.1016/S0176-1617(99)80314-9.
28. Daughtry, C. S. T., Walthall, C. L., Kim, M. S., Colstoun, E. B., McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, 74(2), 229–239. DOI 10.1016/S0034-4257(00)00113-9.
29. Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, 106(1), 135–141. DOI 10.1034/j.1399-3054.1999.106119.x.
30. Pu, R. (2017). *Hyperspectral remote sensing: Fundamentals and practices*, pp. 490. Boca Raton: CRC Press.