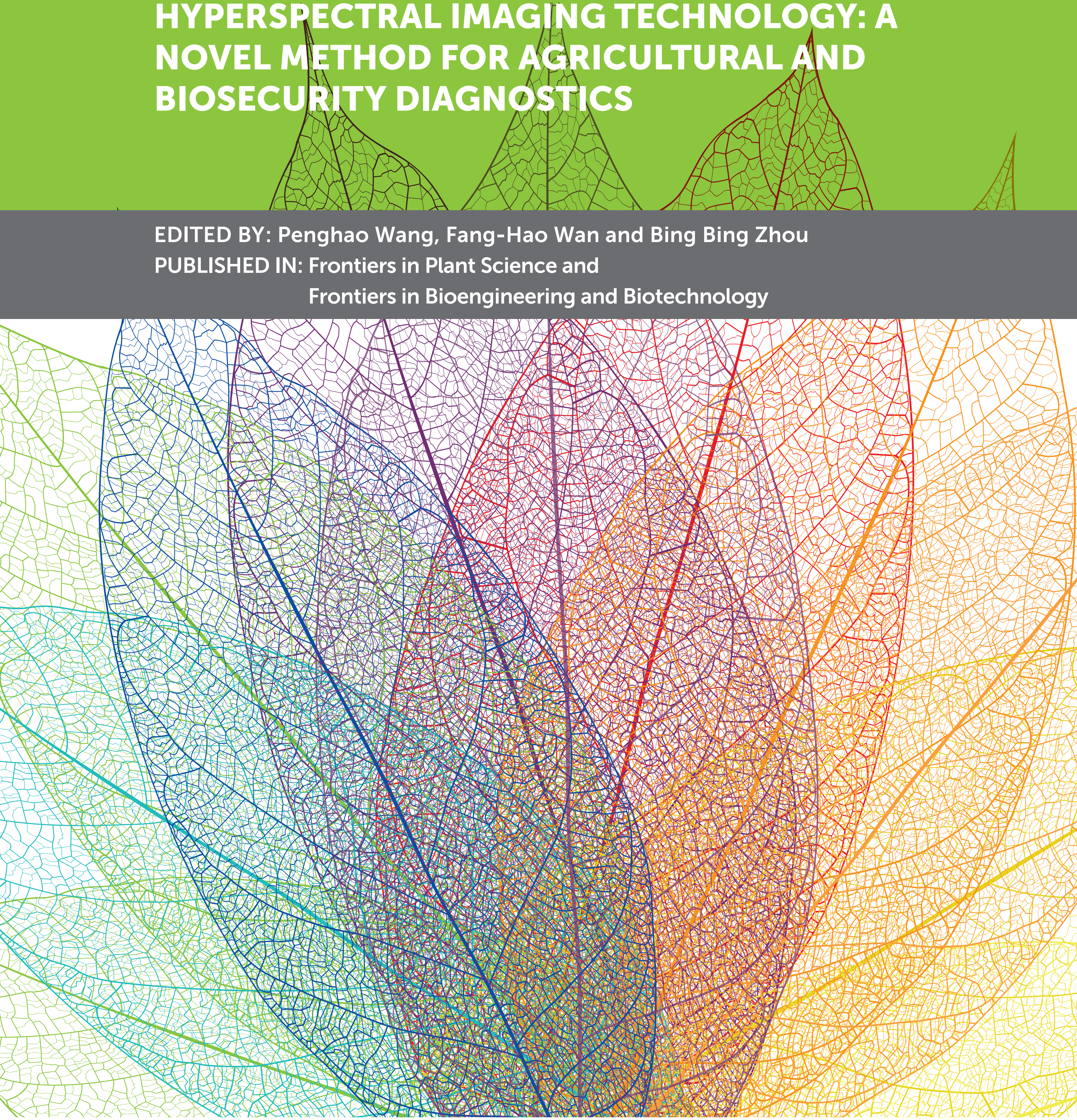
HYPERSPECTRAL IMAGING TECHNOLOGY: A NOVEL METHOD FOR AGRICULTURAL AND BIOSECURITY DIAGNOSTICS

EDITED BY: Penghao Wang, Fang-Hao Wan and Bing Bing Zhou

PUBLISHED IN: Frontiers in Plant Science and
Frontiers in Bioengineering and Biotechnology



frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88971-841-2
DOI 10.3389/978-2-88971-841-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

HYPERSPECTRAL IMAGING TECHNOLOGY: A NOVEL METHOD FOR AGRICULTURAL AND BIOSECURITY DIAGNOSTICS

Topic Editors:
Penghao Wang, Murdoch University, Australia
Fang-Hao Wan, Chinese Academy of Agricultural Sciences, China
Bing Bing Zhou, The University of Sydney, Australia

Prof. Zhou holds patents on the management of agricultural pests. All other Topic Editors declare no competing interests with regard to the Research Topic subject.

Citation: Wang, P., Wan, F.-H., Zhou, B. B., eds. (2021). Hyperspectral Imaging Technology: A Novel Method for Agricultural and Biosecurity Diagnostics. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-841-2

Table of Contents

frontiers
in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 12 January 2021
doi: 10.3389/fbioe.2020.616943

Rapid Detection of Pomelo Fruit Quality Using Near-Infrared Hyperspectral Imaging Combined With Chemometric Methods

*Huazhou Chen [1,2], Hanli Qiao [1,2], Quanxi Feng [1,2], Lili Xu [3], Qinyong Lin [4] and Ken Cai [4]**

[1] College of Science, Guilin University of Technology, Guilin, China, [2] Center for Data Analysis and Algorithm Technology, Guilin University of Technology, Guilin, China, [3] College of Marine Sciences, Beibu Gulf University, Qinzhou, China, [4] College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou, China

Pomelo is an important agricultural product in southern China. Near-infrared hyperspectral imaging (NIRHI) technology is applied to the rapid detection of pomelo fruit quality. Advanced chemometric methods have been investigated for the optimization of the NIRHI spectral calibration model. The partial least squares (PLS) method is improved for non-linear regression by combining it with the kernel Gaussian radial basis function (RBF). In this study, the core parameters of the PLS latent variables and the RBF kernel width were designed for grid search selection to observe the minimum prediction error and a relatively high correlation coefficient. A deep learning architecture was proposed for the parametric scaling optimization of the RBF-PLS modeling process for NIRHI data in the spectral dimension. The RBF-PLS models were established for the quantitative prediction of the sugar (SU), vitamin C (VC), and organic acid (OA) contents in pomelo samples. Experimental results showed that the proposed RBF-PLS method performed well in the parameter deep search progress for the prediction of the target contents. The predictive errors for model training were 1.076% for SU, 41.381 mg/kg for VC, and 1.136 g/kg for OA, which were under 15% of their reference chemical measurements. The corresponding model testing results were acceptably good. Therefore, the NIRHI technology combined with the study of chemometric methods is applicable for the rapid quantitative detection of pomelo fruit quality, and the proposed algorithmic framework may be promoted for the detection of other agricultural products.

Keywords: near-infrared hyperspectral imaging (NIRHI), pomelo fruit quality, agricultural product, chemometric method, partial least squares (PLS), Gaussian radial basis function (RBF)

OPEN ACCESS

Edited by:
Fang-Hao Wan,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Hang Zhang,
Tianjin Agricultural University, China
Zhongyu Sun,
Guangzhou Institute of Geography, China

Correspondence:
Ken Cai
kencaizhku@foxmail.com

Specialty section:
This article was submitted to Biosafety and Biosecurity, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *13 October 2020*
Accepted: *07 December 2020*
Published: *12 January 2021*

Citation:
Chen H, Qiao H, Feng Q, Xu L, Lin Q and Cai K (2021) Rapid Detection of Pomelo Fruit Quality Using Near-Infrared Hyperspectral Imaging Combined With Chemometric Methods.
Front. Bioeng. Biotechnol. 8:616943.
doi: 10.3389/fbioe.2020.616943

INTRODUCTION

Pomelo is one of the special agricultural fruit products that is popular in southern China. Its scientific name is *Citrus maxima (Brum.) Merr*. Ripe pomelo fruits are picked, stored, and served for eating. The fruit peel has functional curative effects in traditional Chinese medicine (Jiang et al., 2014). The flesh is edible and tastes delicious, sweet, and slightly sour; it is rich in sugar, vitamin C, and organic acids, which provide a variety of nutrients for the human body (Sirisomboon and Lapcharoensuk, 2012). People's health can be partially improved from the consumption of good-quality pomelo fruits. Eating pomelo can help maintain good stomach digestion ability and exerts an auxiliary effect of preventing influenza (Anlamlert et al., 2015). Thus, the pomelo fruit quality should be determined during the picking and storage process. Conventional laboratory methods for detecting chemical contents are tedious and time consuming. Rapid detection technology is appreciably on demand (Xu et al., 2020).

Hyperspectral imaging (HI) is regarded as an emerging advanced analytical technology for the non-destructive rapid determination of agricultural product quality (ElMasry et al., 2012; Barreto et al., 2018). HI generates two-dimensional spatial digital imagery accompanied with spectroscopic records for the analysis of spectral features in the ultraviolet, visible, near-infrared, or infrared regions. It technically supports signal processing in the field of computer vision (Lorente et al., 2012). The near-infrared (NIR) spectral region (around 800–2,500 nm) provides a versatile range of light frequencies to analyze the molecule structure and quantify their substantial contents (Pojić and Mastilović, 2013). The recognition of informative features from the natural overlapping signals requires the investigation of smart chemometric methods in the modern intelligent world (Sciutto et al., 2014; Cheng and Sun, 2015). On this basis, HI technology originating from the NIR region (denoted as NIRHI for short) facilitates the combination analysis of the imaging pixels and the NIR-range spectral data (Costa et al., 2011; Cheng et al., 2014). This technology is used as an advanced tool for qualitative and quantitative analyses in the fields of agriculture, food, and industry (Wu and Sun, 2013; Verdú et al., 2016; Arendse et al., 2018). Research on the quality detection of bakery food, meat, and fresh vegetables (Kamruzzaman et al., 2016; Erkinbaev et al., 2017) has been published, but quality assessments for fruits are a brand-new emerging application (Munera et al., 2017).

In NIRHI analysis, the spatial pixels include rich spectral information, and the spectral signals can be used for the rapid quantitative determination of any nutrient content in agricultural products. The selection of the spatial region of interest (ROI) and studies on chemometric methods to extract informative latent variables in the spectral dimension are both significant for NIRHI technology. Given that the selection of ROI has been studied extensively (Chen et al., 2019), feature extraction in the spectral dimension is the main focus of this study.

For the analysis of spectral data, partial least square (PLS) regression is a classical method in finding the latent variables that reflect most of the information of the target analytes. PLS performs principle component extraction, followed by linear regression on the component variables (Wold et al., 2001; Jin and Wang, 2019). However, for NIRHI analysis of complex objects such as pomelo fruit, the spectral dimension contains signal responses from all chemical compositions. The regression model does not stand as a linear formula for a few target analytes. A non-linear kernel function should then be introduced as an algorithmic embedment of PLS (Kim et al., 2005). The Gaussian radial basis function (RBF) is most commonly used for mapping data into a higher dimensional data space for linear fitting (Sandberg, 2003). Its effectiveness and fast tuning of the kernel width ensure that RBF is a successful algorithm application in the kernel PLS method (Shariati-Rad and Hasani, 2013; de Almeida et al., 2018).

In this work, NIRHI technology was applied for the rapid detection of the pomelo fruit quality during its picking and storage processes. The RBF-implemented PLS (RBF-PLS) method was investigated as an advanced chemometric method for the quantitative determination of the sugar (SU), vitamin C (VC), and organic acid (OA) contents. A deep learning architecture was built for parametric scaling optimizations. The model training procedure was launched by automatically tuning the PLS parameters in combination with the machine learning of the RBF kernel width, and the optimal model was tested based on the assumed pseudo-unknown samples. In this way, NIRHI may be considered a modern popular technology for detecting the fruit quality of agricultural products.

MATERIALS AND METHODS

Hyperspectral System

As shown in **Figure 1**, the NIRHI optical system was constructed under laboratory conditions with constant temperature and humidity (25 ± 1°C and 47 ± 1%RH). The NIR lights were originally generated from a 500 W-powered halogen light source. The halogen light was transformed into a series of parallel lights via a convex lens (with 30 mm focal length). The ImSpector N25E hyperspectral imager (Spectral Imaging Ltd., Oulu, Finland) is the main optical part, which splits the source halogen light into a single frequency and produces a batch of NIR wavelengths according to the system pre-settings. The N25E imager generates the full-length NIR waveband of 1,000–2,500 nm with the common resolution of 8 nm.

The NIR lights are further delivered to the pomelo samples through a pushbroom scanner. The scanner includes a flat mirror and a transflective mirror as its main optical parts. The pushbroom scanner uses a horizontally movable back-and-forth motion to form the spatial dimensions of the hyperspectral image. It is steadily set 20 cm away from the surface plane of the sample pool in the vertical direction. The reflectance lights that come out from the sample enter an MT-CT image detector, which includes a CCD unit and some necessary fundamental optical parts. The NIRHI spectral data are finally recorded at the data output segment, where there is always a high-performance computer.

Sample Preparation and Data Acquisition

A total of 300 mature pomelo fruits were collected from a pomelo forest in southern China. An elementary pre-experimental selection was made before measuring contents. Some fruits with a homogeneous peel surface were reserved for further detection. Some fruits with flesh that had minimal moisture were removed from the experiments. A total of 248 pomelo fruits were selected as the target samples for NIRHI measurements and conventional chemical detection.

Each of the 248 target fruits was cut into two halves along its central longitudinal surface. One half of each fruit was sent to quantify its contents of SU, VC, and OA contents. These three analytes should be detected on the cutting interface via different

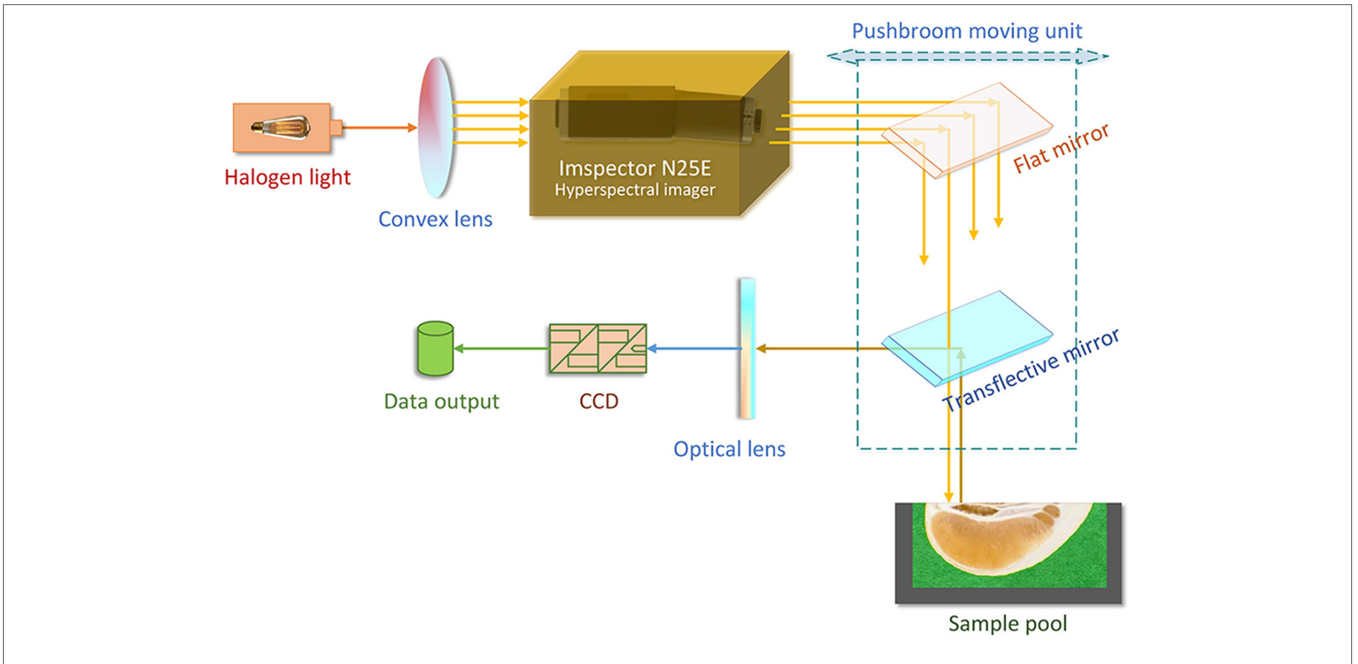


FIGURE 1 | The construction of the NIRHI optical system.

TABLE 1 | The descriptive statistics for the SU, VC, and OA of the 248 pomelo fruit samples.

	Maximum	Minimum	Average	Standard deviation
Sugar (SU, %)	14.52	8.63	11.51	1.78
Vitamin C (VC, mg/kg)	623.02	347.51	487.71	82.91
Organic acid (OA, g/kg)	13.68	8.91	11.12	1.39

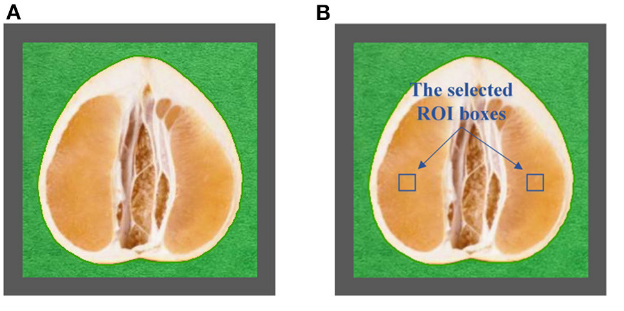


FIGURE 2 | The central longitudinal cut view of the equipped pomelo sample **(A)** and the selected ROI areas **(B)**.

chemical experiments. These chemically measured contents were supposed to be the reference values for NIRHI modeling because the analyte fruit flesh samples were from the same cutting surface. The SU content was identified by 3,5-dinitrosalicylic acid colorimetry (China's agricultural industry standard, NY/T 2742-2015). The VC content was determined by 2,6-dichloro-indophenol titration (China's national standard, GB 5009.86-2016). The OA content was detected by ion chromatography (China's national standard, GB 5009.157-2016). The descriptive statistics for the SU, VC, and OA contents of the 248 pomelo fruit samples are shown in **Table 1**.

The half was equipped in the sample pool accessory. As shown in **Figure 2A**, the halved pomelo fruit was placed into the cuboid sample pool (the gray frame). The interspace between the fruit and the box was filled with plasticine (the green part). The filled sample pool was equipped to the constructed hyperspectral system, and the NIRHI spectral data of this sample were collected by pushbroom scanning. The NIRHI data have two spatial dimensions and one spectral dimension. In the spatial dimensions, the selection of ROI was studied in our previous work (Chen et al., 2019), which reported that the 5 × 5 square-size data extracted from the core spatial pixel area provide optimal spectroscopy calibration results. Thus, we selected two ROIs of 5 × 5 pixels from the main flesh areas around the fruit-shaped equatorial plane (see the two blue boxes in **Figure 2B**). The spectral data within these two ROI areas were extracted from the NIRHI spatial-spectral data cube. Fifty pixels of NIR spectra were acquired for each pomelo fruit sample. The average of these 50 spectral data was calculated as the spectral information of each sample for further chemometric modeling. Finally, the average spectra of all 248 pomelo samples were obtained, and the spectral curves are illustrated in **Figure 3**.

The RBF-PLS Method

The RBF-implemented PLS method is a provoked kernel PLS regression algorithm extended from the common PLS regression. It uses the RBF kernel function to transform raw non-linear complex data into a new defined feature space, in which the data can be linearly predicted with the tuning of the number of PLS latent variables (Chakraborty, 2012; Goudarzi, 2016). The RBF kernel is defined as follows (Ring and Eskofier, 2016):

$$K\left(x_i, x_j\right) = \exp\left(\frac{-\left\|x_i - x_j\right\|^2}{\sigma^2}\right), \quad i, j = 1, 2 \ldots n$$

where σ represents the kernel width. Different values of σ would lead to diverse kernel mapping results in the new data space. For a fixed value of σ, the function $K\left(x_i, x_j\right)$ obtains different computing values for varying training data of x_i and x_j, thereby generating the kernel matrix for the n training samples, which is constructed as

$$K = \begin{bmatrix} K\left(x_1, x_1\right) & K\left(x_1, x_2\right) & \cdots & K\left(x_1, x_n\right) \\ K\left(x_2, x_1\right) & K\left(x_2, x_2\right) & \cdots & K\left(x_2, x_n\right) \\ \vdots & \vdots & \ddots & \vdots \\ K\left(x_n, x_1\right) & K\left(x_n, x_2\right) & \cdots & K\left(x_n, x_n\right) \end{bmatrix},$$

Successively, the matrix K is transformed to M,

$$M = K - \frac{1}{n} I_n K - \frac{1}{n} K I_n + \frac{1}{n^2} I_n K I_n,$$

where I_n is an n-dimensional square all-one matrix. To make the method smart and data-driven, the algorithm of RBF-PLS training can be operated in an iteration process as follows:

```
Step 1: $E = M, F = Y$;
Step 2: Randomly initialize $U$ (a matrix consists of $s$ latent variables);
Step 3: $V = KU$, $V \leftarrow V/\|V\|$;
Step 4: $C = Y^{\mathrm{T}} V$;
Step 5: $U = YC$, $U \leftarrow U/\|U\|$;
Step 6: Repeat Steps 3–6 until convergence occurs;
Step 7: Residual matrix $E$ and $F$ were computed,
$E \leftarrow (I - VV^T)E(I - VV^T)$, $F \leftarrow F - VV^T Y$,
where $I$ is an $n$-dimensional identity matrix;
Step 8: Turn to Step 3 until the convergence occurs for the residuals $E$ and $F$.
```

The predicted data of training set are evaluated by the equation

$$Y' = MU(V^T MU)^{-1} V^T Y,$$

where V is formed by the columns of latent vector v; U is formed by the columns of latent vector u; and Y is the predictor matrix. The training process shows that the optimization of the RBF-PLS calibration model is mainly controlled by tuning the RBF kernel width (i.e., σ) and the number of PLS latent variables (i.e., s). The combined optimization of σ and s should be an applicable machine learning mode for advanced parameter training.

Furthermore, for the testing sample set, the kernel matrix K_{test} is computed and constructed similar to constructing K, and K_{test} is $(t \times n)$-dimensional. Each element of K_{test} is obtained by computing the kernel function between the t testing samples and the n training samples. Successively, we will have K_{test} transformed to M_{test}:

$$M_{\text{test}} = K_{\text{test}} - \frac{1}{n} I_t K - \frac{1}{n} K_{\text{test}} I_n + \frac{1}{n^2} I_t K I_n,$$

where I_t is a $t \times n$ all-one matrix. The algorithm of the testing part is similar to that of the training part, and the prediction equation of the testing set has the same structure as that of the training set.

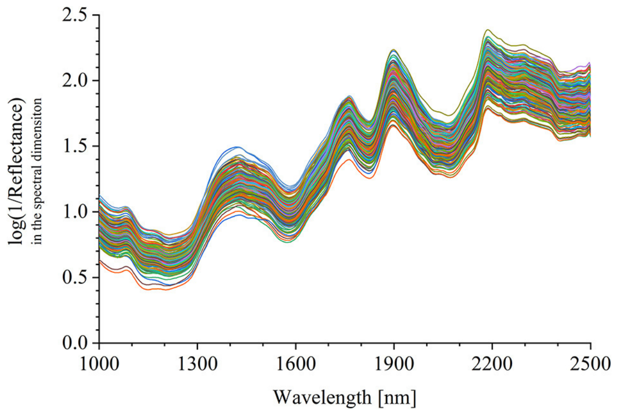


FIGURE 3 | The extracted average spectra of the 248 pomelo samples in the NIRHI spectral dimension.

The RBF-PLS model is developed by regression of the response matrix X against the predictor matrix Y. The model based on experimental data is established to quantitatively estimate the pseudo-unknown samples based on their measured features. RBF-PLS regression and prediction were carried out using the MATLAB coding platform (ver. R2018a) accompanied with its toolboxes. The parametric scaling on the kernel function can be launched in a deep learning mode, and the selection of latent variables can be embedded for deep combined optimization.

Model Evaluation Indicators

The chemometric study for the NIRHI analytical model requires the method to be intelligently adjusted to the detected data. Thus, the data knowledge should be recognized in a self-adaptive machine learning mode. On this basis, the NIRHI spectral data of all 248 pomelo fruit samples should be primarily divided into the training set and the testing set. The training samples are used to establish and optimize the model. By contrast, the testing samples, which are not involved in the model training process, aim to evaluate the best-trained calibration model. The model optimization effects need to be validated during the training process, so the training sample set should be further divided into two subsets: the calibration set and the validation set. The calibration set is for model establishment, and the validation set is for model optimization.

Experimental evidence showed that the samples divided for calibration, validation, and testing are usually in the ratio of 2:1:1

(Chen et al., 2015). Of the total 248 samples, we randomly chose 120 samples for calibration, 64 samples for validation, and 64 samples for testing.

The model prediction effects are generally quantified using two important indicators. One is the root mean square error (RMSE), which is used to estimate the model prediction bias. The other one is the correlation coefficient (CC), which is a statistical metric representing the closeness of the NIRHI predicted values to the chemically measured reference contents. These two indicators are formulated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}(y_i - y_{\mathrm{ave}})(\hat{y}_i - \hat{y}_{\mathrm{ave}})}{\sqrt{\sum(y_i - y_{\mathrm{ave}})^2\sum(\hat{y}_i - \hat{y}_{\mathrm{ave}})^2}},$$

where $\hat{y}_i$ and y_i are the NIRHI predicted value and its chemically measured reference value of the i-th sample, respectively. $\hat{y}_{\mathrm{ave}}$ and y_{ave} are the average predicted value and average reference value of n samples, respectively. n is the total number of target samples.

As the best optimal model was identified by the validation samples and evaluated by the testing samples, the model indicators were denoted as $\mathrm{RMSE_V}$ and $\mathrm{CC_V}$ for the validation sample set and denoted as $\mathrm{RMSE_T}$ and $\mathrm{CC_T}$ for the testing sample set.

RESULTS AND DISCUSSIONS

Parametric Scaling Deep Learning Results of the RBF-PLS Model

The extracted NIRHI spectra of the 248 pomelo fruit samples were used to establish calibration models by using the proposed

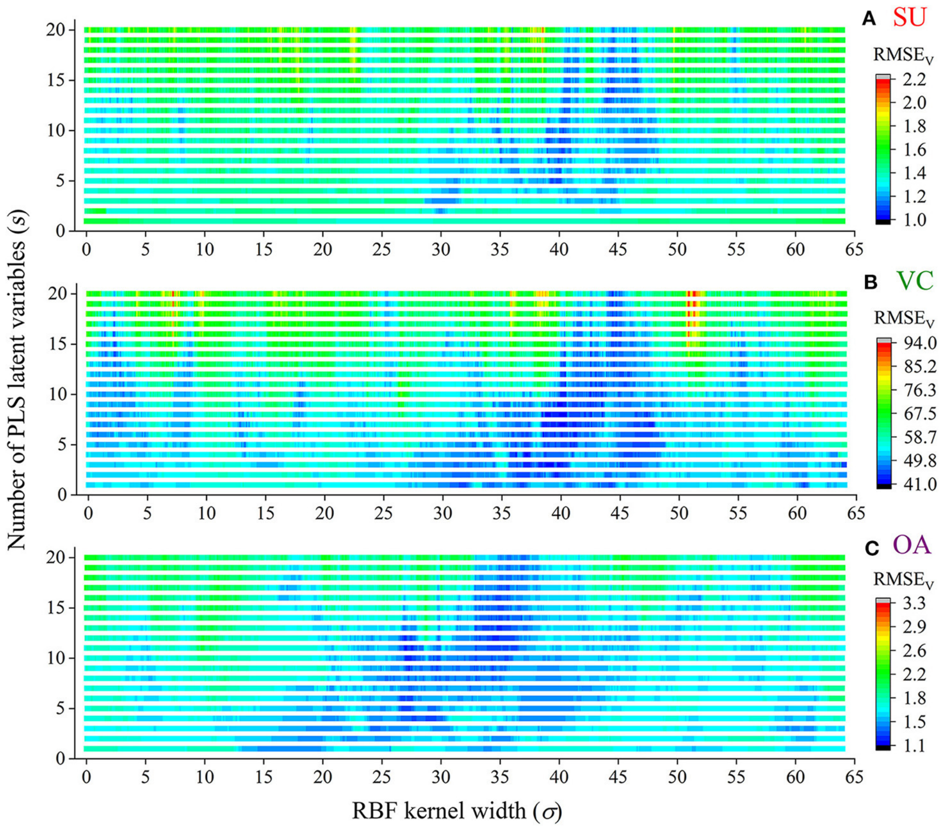


FIGURE 4 | The combined deep tuning of RBF kernel width (σ) and the number of PLS latent variables (s) for the optimization of the NIRHI calibration model (**A–C** are for the prediction of SU, VC, and OA, respectively).

TABLE 2 | The optimal RBF-PLS models for NIRHI prediction of SU, VC, and OA contents in pomelo fruit samples.

	RBR-PLS model			PLS model	
	Kernel parameters	**RMSE$_V$**	**CC$_V$**	**RMSE$_V$**	**CC$_V$**
SU (%)	$\sigma = 38.97$; $s = 8$	1.076%	0.921	1.361%	0.895
VC (mg/kg)	$\sigma = 44.38$; $s = 14$	41.381 mg/kg	0.913	50.672 mg/kg	0.862
OA (g/kg)	$\sigma = 27.20$; $s = 11$	1.136 g/kg	0.902	1.475 g/kg	0.875

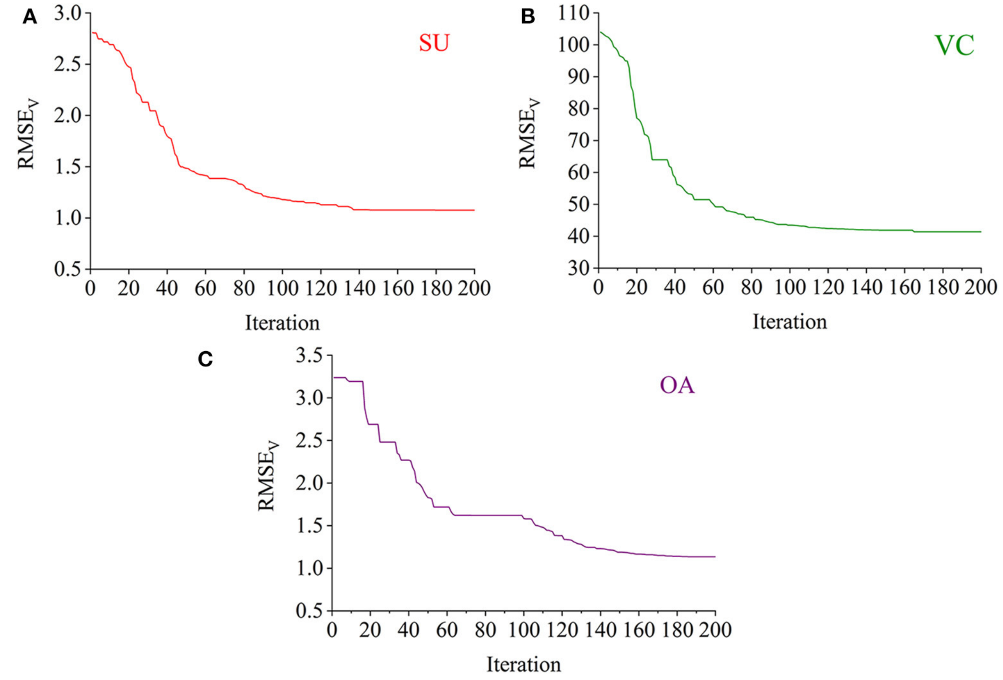


FIGURE 5 | The iterative optimization of the RBF-PLS models for the NIRHI data training (**A–C** are for the prediction of SU, VC, and OA, respectively).

RBF-PLS method. A deep learning architecture was built for parametric scaling search of the optimal RBF-PLS parameters. The models for predicting the SU, VC, and OA contents were trained based on the 128 calibration samples, and the modeling parameters were tunable for deep searching of their optimal combination values. For the RBF kernel, the kernel width (σ) is commonly set as 2^i ($i = 0, \pm1, \pm2 \ldots$) (Menezes et al., 2019). For statistical convenience, we set σ to change from 0.01 to 64 with the step of 0.01, which included the close estimation of 2^i with $i = 0, \pm1, \pm2, \pm3, \pm4, \pm5, \pm6$. Thus, there were 6,400 candidate values of σ for the kernel width scaling. Meanwhile, the PLS latent variable queues in the front were considered the most informative for spectral data interpretation (Shariati-Rad and Hasani, 2010). The number of latent variables (s) was set as integers from 1 to 20, which indicated that the most important latent variables were used for model optimization. The predictive RMSE_V of the validation samples was used as the main indicator to identify the appreciating model with its optimal parameter combination of (σ, s). The grid search of the RMSE_V corresponding to each combination is shown in **Figure 4**. In **Figure 4**, the two-dimensional axes represent the parametric tuning of σ and s, respectively. The predictive RMSE_V values of each model were demonstrated as contour color mappings. **Figures 4A–C** show the validation results for the SU, VC, and OA contents, respectively. The most optimal training results could be found at the dark blue digit locations, so the optimal combination of (σ, s) was identified (see **Table 2**). The corresponding modeling results (RMSE_V and CC_V) were also listed in the table. For comparison, the classical PLS model was established in model training, and the results are listed in **Table 2**. The prediction results in **Table 2** indicated that the RBF-PLS models performed better than the PLS model during the training process. Therefore, the RBF-PLS models are feasible for the NIRHI quantitative determination of the designated contents related to the quality of pomelo fruits.

Iteration Progress for the Selection of PLS Latent Variables

The RBF-PLS model was optimized by iterative updating of the matrix of latent variables (i.e., the matrix U). For a fixed value of s, the applied latent variables were randomly initialized and then gradually alternated. The progress of updating the latent variables was iteratively set for 200 times. For example, the optimal model for the prediction of the SU content was determined with eight latent variables. The eight latent variables were randomly chosen at the beginning, and they gave the initial predictive RMSE_V of 2.807%. The 200-time iteration model made the prediction more accurate as the RMSE_V curve went down and became stable (see **Figure 5A**). Finally, the prediction on the SU content with eight latent variables observed its optimal result of RMSE_V equal to 1.076%, which was captured within the 200 iteration time. Similarly, the iterative optimization trends of the RBF-PLS models for the prediction of the VC and OA contents are shown in **Figures 5B,C**. As shown in **Figure 5**, the iteration mechanism during the PLS process is feasible to enhance the optimization ability of the RBF-PLS calibration model for the NIRHI spectral analysis of pomelo fruit samples.

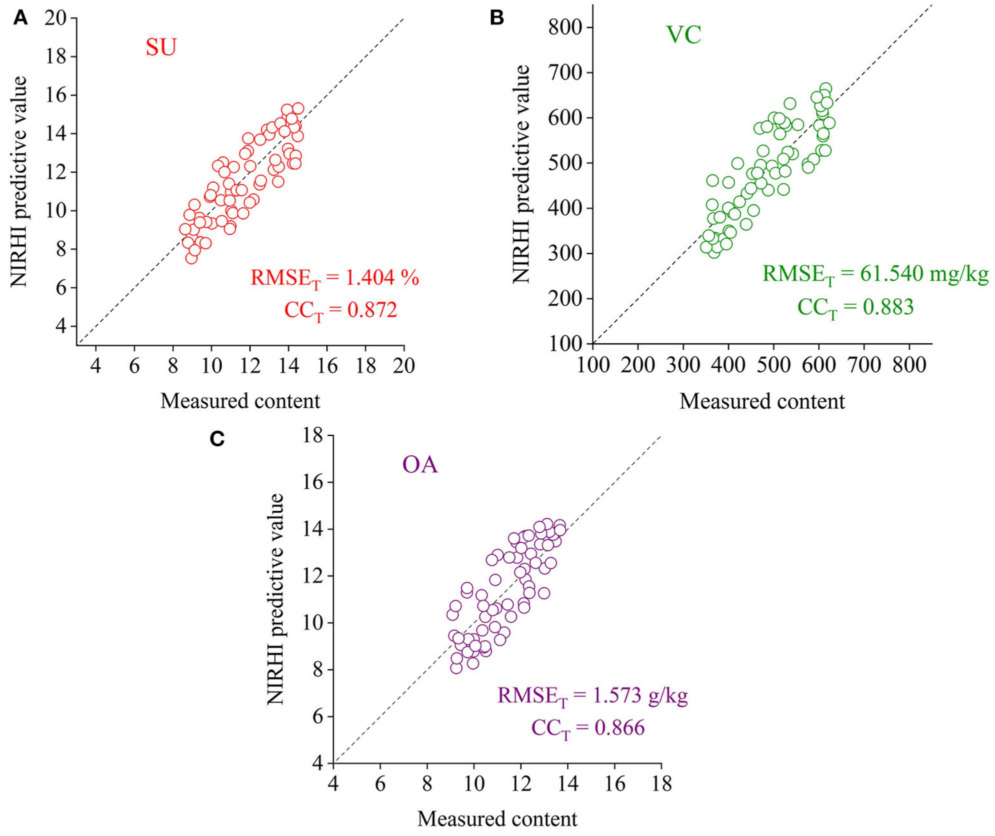


FIGURE 6 | The regression plots of the NIRHI predictions and the reference chemical measurements for the testing samples (**A–C** are for the prediction of SU, VC, and OA, respectively).

Model Evaluation Based on the Testing Samples

To verify the effectiveness of the RBF-PLS model applied to the NIRHI spectral analysis of pomelo fruit samples, the well-trained models for the prediction of the SU, VC, and OA contents were evaluated by the testing samples, which were not involved in the modeling process. The testing models were re-established by using the optimally selected parameter combination of (σ, s), as shown in **Table 2**. The regression plots of the NIRHI predictions and the reference chemical measurements are shown in **Figure 6**. The predicted RMSE$_T$ values were obtained as 1.404%, 61.540 mg/kg, and 1.573 g/kg for the model testing on the SU, VC, and OA contents, respectively, which were under 15% of their reference chemical measurements. The acquired CC$_T$ was larger than 0.85, which seemed to be acceptable for model evaluation of agricultural products.

CONCLUSIONS

The RBF-PLS method was proposed to extract the spectral features from the NIRHI data for the quantitative determination of the SU, VC, and OA contents in pomelo samples. The NIRHI spatial properties were pre-determined based on previous research results. The spectral calibration models were trained in the deep search of the combined parameters (σ, s), where σ was screened from 6,400 possible candidate values changing from 0.01 to 64 with a step of 0.01, and s was changed as an integer from 1 to 20. To observe the minimum RMSE$_V$ and CC$_V$, the grid values of (σ, s) were all tested, and the optimal parameters were identified. The optimal models were found during the calibration and validation processes, with the predictive results of RMSE$_V$ equal to 1.076% for SU, 41.381 mg/kg for VC, and 1.136 g/kg for OA. All of the three CC$_V$ exceeded 0.9. The selected models were evaluated based on the testing samples, and the prediction results were also appreciable. The experimental results indicated that the proposed parametric scaling RBF-PLS method is feasible to determine some pomelo fruit quality targeting contents in combination with the NIRHI technology. Studies on NIRHI chemometric methods are essential to improve the calibration models in the rapid determination of agricultural products.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

HC: conceptualization, methodology, writing–original draft, and writing–review & editing. HQ: data curation and formal analysis. QF: investigation and validation. LX: writing–original draft. QL: software and visualization. KC: resources, supervision, and writing–review & editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Natural Science Foundations of China (Grant Nos. 61763008 and 62003379), Natural Science Foundation of Guangxi Province (Grant No. 2018GXNSFAA050045), Guangzhou Science and Technology Program (Grant No. 202002030246) and Guangxi Science and Technology Base Foundation (Grant No. AD18281039).

REFERENCES

Anlamlert, W., Sermsappasuk, P., Yokubol, D., and Jones, S. (2015). Pomelo enhances cyclosporine bioavailability in healthy male thai volunteers. *J. Clin. Pharmacol*. 55, 377–383. doi: 10.1002/jcph.430

Arendse, E., Fawole, O. A., Magwaza, L. S., and Opara, U. L. (2018). Non-destructive prediction of internal and external quality attributes of fruit with thick rind: a review. *J. Food Eng*. 217, 11–23. doi: 10.1016/j.jfoodeng.2017.08.009

Barreto, A., Cruz-Tirado, J. P., Siche, R., and Quevedo, R. (2018). Determination of starch content in adulterated fresh cheese using hyperspectral imaging. *Food Biosci*. 21, 14–19. doi: 10.1016/j.fbio.2017.10.009

Chakraborty, S. (2012). Bayesian multiple response kernel regression model for high dimensional data and its practical applications in near infrared spectroscopy. *Comput. Stat. Data Anal*. 56, 2742–2755. doi: 10.1016/j.csda.2012.02.019

Chen, H., Qiao, H., Lin, B., Xu, G., Tang, G., and Cai, K. (2019). Study of modeling optimization for hyperspectral imaging quantitative determination of naringin content in pomelo peel. *Comput. Electron. Agric*. 157, 410–416. doi: 10.1016/j.compag.2019.01.013

Chen, H.-Z., Shi, K., Cai, K., Xu, L.-L., and Feng, Q.-X. (2015). Investigation of sample partitioning in quantitative near-infrared analysis of soil organic carbon based on parametric LS-SVR modeling. *RSC Adv*. 5, 80612–80619. doi: 10.1039/C5RA12468A

Cheng, J. H., and Sun, D. W. (2015). Recent applications of spectroscopic and hyperspectral imaging techniques with chemometric analysis for rapid inspection of microbial spoilage in muscle foods. *Compr. Rev. Food Sci. Food Saf*. 14, 478–490. doi: 10.1111/1541-4337.12141

Cheng, J. H., Sun, D. W., Pu, H., and Zeng, X. A. (2014). Comparison of visible and long-wave near-infrared hyperspectral imaging for colour measurement of grass carp (*Ctenopharyngodon idella*). *Food Bioprocess Technol*. 7, 3109–3120. doi: 10.1007/s11947-014-1325-7

Costa, C., Antonucci, F., Pallottino, F., Aguzzi, J., Sun, D. W., and Menesatti, P. (2011). Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision. *Food Bioprocess Technol*. 4, 673–692. doi: 10.1007/s11947-011-0556-0

de Almeida, V. E., de Araújo Gomes, A., de Sousa Fernandes, D. D., Goicoechea, H. C., Galvão, R. K. H., and Araújo, M. C. U. (2018). Vis-NIR spectrometric determination of brix and sucrose in sugar production samples using kernel partial least squares with interval selection based on the successive projections algorithm. *Talanta* 181, 38–43. doi: 10.1016/j.talanta.2017.12.064

ElMasry, G., Kamruzzaman, M., Sun, D. W., and Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Crit. Rev. Food Sci. Nutr*. 52, 999–1023. doi: 10.1080/10408398.2010.543495

Erkinbaev, C., Henderson, K., and Paliwal, J. (2017). Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique. *Food Control* 80, 197–203. doi: 10.1016/j.foodcont.2017.04.036

Goudarzi, N. (2016). Free variable selection QSPR study to predict 19F chemical shifts of some fluorinated organic compounds using Random Forest and RBF-PLS methods. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc*. 158, 60–64. doi: 10.1016/j.saa.2016.01.023

Jiang, J., Shan, L., Chen, Z., Xu, H., Wang, J., Liu, Y., et al. (2014). Evaluation of antioxidant-associated efficacy of flavonoid extracts from a traditional Chinese medicine Hua Ju Hong (peels of *Citrus grandis* (L.) Osbeck). *J. Ethnopharmacol*. 158, 325–330. doi: 10.1016/j.jep.2014.10.039

Jin, J., and Wang, Q. (2019). Evaluation of informative bands used in different PLS regressions for estimating leaf biochemical contents from hyperspectral reflectance. *Remote Sens*. 11, 11–15. doi: 10.3390/rs11020197

Kamruzzaman, M., Makino, Y., and Oshita, S. (2016). Parsimonious model development for real-time monitoring of moisture in red meat using hyperspectral imaging. *Food Chem*. 196, 1084–1091. doi: 10.1016/j.foodchem.2015.10.051

Kim, K., Lee, J. M., and Lee, I. B. (2005). A novel multivariate regression approach based on kernel partial least squares with orthogonal signal correction. *Chemom. Intell. Lab. Syst*. 79, 22–30. doi: 10.1016/j.chemolab.2005.03.003

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., and Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technol*. 5, 1121–1142. doi: 10.1007/s11947-011-0725-1

Menezes, M. V. F., Torres, L. C. B., and Braga, A. P. (2019). Width optimization of RBF kernels for binary classification of support vector machines: a density estimation-based approach. *Pattern Recognit. Lett*. 128, 1–7. doi: 10.1016/j.patrec.2019.08.001

Munera, S., Amigo, J. M., Blasco, J., Cubero, S., Talens, P., and Aleixos, N. (2017). Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *J. Food Eng*. 214, 29–39. doi: 10.1016/j.jfoodeng.2017.06.031

Pojić, M. M., and Mastilović, J. S. (2013). Near infrared spectroscopy-advanced analytical tool in wheat breeding, trade, and processing. *Food Bioprocess Technol*. 6, 330–352. doi: 10.1007/s11947-012-0917-3

Ring, M., and Eskofier, B. M. (2016). An approximation of the Gaussian RBF kernel for efficient classification with SVMs. *Pattern Recognit. Lett*. 84, 107–113. doi: 10.1016/j.patrec.2016.08.013

Sandberg, I. W. (2003). Gaussian radial basis functions and the approximation of input-output maps. *Int. J. Circuit Theory Appl*. 31, 443–452. doi: 10.1002/cta.242

Sciutto, G., Prati, S., Bonacini, I., Oliveri, P., and Mazzeo, R. (2014). FT-NIR microscopy: an advanced spectroscopic approach for the characterisation of paint cross-sections. *Microchem. J*. 112, 87–96. doi: 10.1016/j.microc.2013.09.021

Shariati-Rad, M., and Hasani, M. (2010). Selection of individual variables versus intervals of variables in PLSR. *J. Chemom*. 24, 45–56. doi: 10.1002/cem.1266

Shariati-Rad, M., and Hasani, M. (2013). Linear and nonlinear quantitative structure-property relationships modeling of charge-transfer complex formation of organic donors with iodine and iodine monochloride using partial least squares and radial basis function-partial least squares. *J. Iran. Chem. Soc*. 10, 1247–1256. doi: 10.1007/s13738-013-0265-x

Sirisomboon, P., and Lapcharoensuk, R. (2012). Evaluation of the physicochemical and textural properties of pomelo fruit following storage. *Fruits* 67, 399–413. doi: 10.1051/fruits/2012034

Verdú, S., Vásquez, F., Grau, R., Ivorra, E., Sánchez, A. J., and Barat, J. M. (2016). Detection of adulterations with different grains in wheat products based on the hyperspectral image technique: the specific cases of flour and bread. *Food Control* 62, 373–380. doi: 10.1016/j.foodcont.2015.11.002

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst*. 58, 109–130. doi: 10.1016/S0169-7439(01)00155-1

Wu, D., and Sun, D. W. (2013). Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh. *Talanta* 111, 39–46. doi: 10.1016/j.talanta.2013.03.041

Xu, S., Lu, H., Ference, C., Qiu, G., and Liang, X. (2020). Rapid nondestructive detection of water content and granulation in postharvest "shatian" pomelo using visible/near-infrared spectroscopy. *Biosensors* 10, 1–13. doi: 10.3390/bios10040041

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Chen, Qiao, Feng, Xu, Lin and Cai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Bioengineering and Biotechnology

REVIEW
published: 13 January 2021
doi: 10.3389/fbioe.2020.623705

High-Throughput Plant Phenotyping Platform (HT3P) as a Novel Tool for Estimating Agronomic Traits From the Lab to the Field

Daoliang Li[1,2,3,4,5]*, **Chaoqun Quan**[1,2,3,4,5], **Zhaoyang Song**[1,2,3,4,5], **Xiang Li**[6], **Guanghui Yu**[1,2,3,4,5], **Cheng Li**[1,2,3,4,5] **and Akhter Muhammad**[1,5]

[1] National Innovation Center for Digital Fishery, China Agricultural University, Beijing, China, [2] Beijing Engineering and Technology Research Centre for Internet of Things in Agriculture, China Agricultural University, Beijing, China, [3] China-EU Center for Information and Communication Technologies in Agriculture, China Agriculture University, Beijing, China, [4] Key Laboratory of Agriculture Information Acquisition Technology, Ministry of Agriculture, China Agricultural University, Beijing, China, [5] College of Information and Electrical Engineering, China Agricultural University, Beijing, China, [6] Department of Psychology, College of Education, Hubei University, Wuhan, China

OPEN ACCESS

Edited by:
Fang-Hao Wan,
Chinese Academy of Agricultural Sciences, China

Reviewed by:
Yu Jiang,
Cornell University, United States
Haiyan Cen,
Zhejiang University, China

***Correspondence:**
Daoliang Li
dliangl@cau.edu.cn

Specialty section:
This article was submitted to Biosafety and Biosecurity, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *30 October 2020*
Accepted: *15 December 2020*
Published: *13 January 2021*

Citation:
Li D, Quan C, Song Z, Li X, Yu G, Li C and Muhammad A (2021) High-Throughput Plant Phenotyping Platform (HT3P) as a Novel Tool for Estimating Agronomic Traits From the Lab to the Field. Front. Bioeng. Biotechnol. 8:623705. doi: 10.3389/fbioe.2020.623705

Food scarcity, population growth, and global climate change have propelled crop yield growth driven by high-throughput phenotyping into the era of big data. However, access to large-scale phenotypic data has now become a critical barrier that phenomics urgently must overcome. Fortunately, the high-throughput plant phenotyping platform (HT3P), employing advanced sensors and data collection systems, can take full advantage of non-destructive and high-throughput methods to monitor, quantify, and evaluate specific phenotypes for large-scale agricultural experiments, and it can effectively perform phenotypic tasks that traditional phenotyping could not do. In this way, HT3Ps are novel and powerful tools, for which various commercial, customized, and even self-developed ones have been recently introduced in rising numbers. Here, we review these HT3Ps in nearly 7 years from greenhouses and growth chambers to the field, and from ground-based proximal phenotyping to aerial large-scale remote sensing. Platform configurations, novelties, operating modes, current developments, as well the strengths and weaknesses of diverse types of HT3Ps are thoroughly and clearly described. Then, miscellaneous combinations of HT3Ps for comparative validation and comprehensive analysis are systematically present, for the first time. Finally, we consider current phenotypic challenges and provide fresh perspectives on future development trends of HT3Ps. This review aims to provide ideas, thoughts, and insights for the optimal selection, exploitation, and utilization of HT3Ps, and thereby pave the way to break through current phenotyping bottlenecks in botany.

Keywords: crop improvement, high-throughput, phenomics, phenotyping platform, plant science, remote sensing, sensors

INTRODUCTION

The growth and development of plants, involving their photosynthesis, transpiration, flowering, and fruiting processes, are the basis of life on earth, and support 7.5 billion people (Pieruschka and Schurr, 2019). Unfortunately, the agriculture that sustains humanity is now facing three stark challenges at once: climate change, resource depletion, and population growth (Kim, 2020). In the next 30 years, the global population is expected to grow by 25% to 10 billion (Hickey et al., 2019). One of the greatest challenges in the twenty-first century will be to quickly expand crop production to meet this growing demand for food, clothing, and fuel. Salinization and erosion of agricultural land around the world, coupled to declining phosphate reserves, pose a grave threat to growth in the global production of crops. On April 21, 2020, the World Food Programme (WFP) announced that as new coronavirus pandemic spreads and batters the global economy, the number of people facing severe food crisis in the world could increase to 265 million within the year. In the past decade, cheaper and faster sequencing methods have fostered increasing crop yields and generated an enormous increase in plant genomic data. The costs of sequencing have fallen dramatically, from $0.52 per Mb of DNA sequence in 2010 to just $0.010 in 2019, while the cost of sequencing a human-sized genome has decreased from $46 774 to a relatively paltry $942 (National Human Research Institute). Although high-throughput genotyping is expanding exponentially, the collection and processing of plant phenotypes constrain our ability to analyze the genetics of quantitative traits and limit the use of breeding for crop yield improvement (Mccouch et al., 2013).

The phenotype arises from interactions between genotype and environment (Hickey et al., 2019), and the essence is the temporal expression of the plants' gene map in characteristic geographic regions (Zhao, 2019). Phenotyping applies specific methods and protocols to measure morphological structural traits, physiological functional traits, and component content traits of cells, tissues, organs, canopy, whole plants, or even populations. However, traditional breeders perform artificial phenotyping based on the appearance, taste, and touch of the crop, undoubtedly a time-consuming, labor-intensive, and even destructive method that requires immense human resources to sample large population of crop plants. The limitation of phenotyping efficiency is increasingly recognized as a key constraint of progress in applied genetics, especially the time interval for acquiring traits in different environments (Guzman et al., 2015). Further, conventional phenotyping methods also make it difficult to capture physiological and biochemical phenotypes at the level of plant basic mechanisms that reveal patterns of genetics and biology. So, to alleviate this bottleneck, since 2000 a variety of phenotyping platforms have been developed which are now common tools in commercial or research teams (Granier and Vile, 2014).

An image-based, high-throughput phenotyping platform (HT3P) is defined as a platform that can image at least hundreds of plants daily (Fahlgren et al., 2015b). Given that some HT3Ps currently not only rely on images but also are based on contact (albeit non-destructive), "HT3P" is defined here as a platform that can collect massive amounts of phenotypic data from hundreds of plants every day with a high degree of automation. HT3P is a novel and powerful tool allowing us to monitor and quantify crop growth and production-related phenotypic traits in a non-destructive, fast, and high-throughput manner, and then to achieve genomics-assisted breeding (GAB) through genomic approaches of quantitative trait loci (QTL) mapping, marker assisted selection (MAS), genomic selection (GS), and genome-wide association studies (GWAS), thereby assisting crop growers to adapt to changing climate conditions and market demand for yield. When genomics and high-throughput phenotypic data are robustly linked together, this fusion will also greatly promote the development of phenotyping. Furthermore, as **Figure 1** shows, various types of HT3Ps contribute to the phenotyping of plant morphological structure, physiological function, and fractional content, and they can further promote the developments of multi-omics and reveal the regulatory networks and biological patterns of plants' growth and development.

Nevertheless, because the large phenotyping platforms mostly are developed by professional commercial companies, the underlying hardware and software are protected by patents, so they cannot be modified to meet specific research needs (Czedik-Eysenberg et al., 2018). Consequently, a diversified range of commercial phenotyping platforms, as well as those either customized or self-developed, are continuously emerging. In this context, this paper reviews HT3Ps (root phenotyping not included) under three scenarios: (1) greenhouses and growth chambers under strictly controlled conditions; (2) ground-based proximal phenotyping in the field, and; (3) aerial, large-scale remote sensing, with an emphasis on platform novelties, sensor configurations, operation modes, and applications. Then, we innovatively propose ways to combine HT3Ps for their comparative validation or comprehensive analysis. Finally, we discuss some prevailing issues in current high-throughput phenotyping and also highlight the prospects for future development of HT3P. We hope this review enables researchers on plant phenotyping to make more informed choices when employing HT3P, provides fresh ideas and thoughts for intrepid developers of HT3P, and that ultimately hastens the next green revolution in crop breeding.

HT3P FOR INDOOR PHENOTYPING UNDER STRICTLY CONTROLLED ENVIRONMENTAL CONDITIONS

High-throughput plant phenotyping in the growth chamber or greenhouse entails the precise control of environmental factors—temperature, humidity, gas concentration, air volume, wind speed, light intensity, spectral range, photoperiod, and nutrient content—and a high-throughput, non-destructive, highly repeatable, fast, and accurate capture of the plant response to a specific environment. This can be done using model crops or representative plants as research objects, and the analysis of plants' structure, physiology, and biochemical characteristics

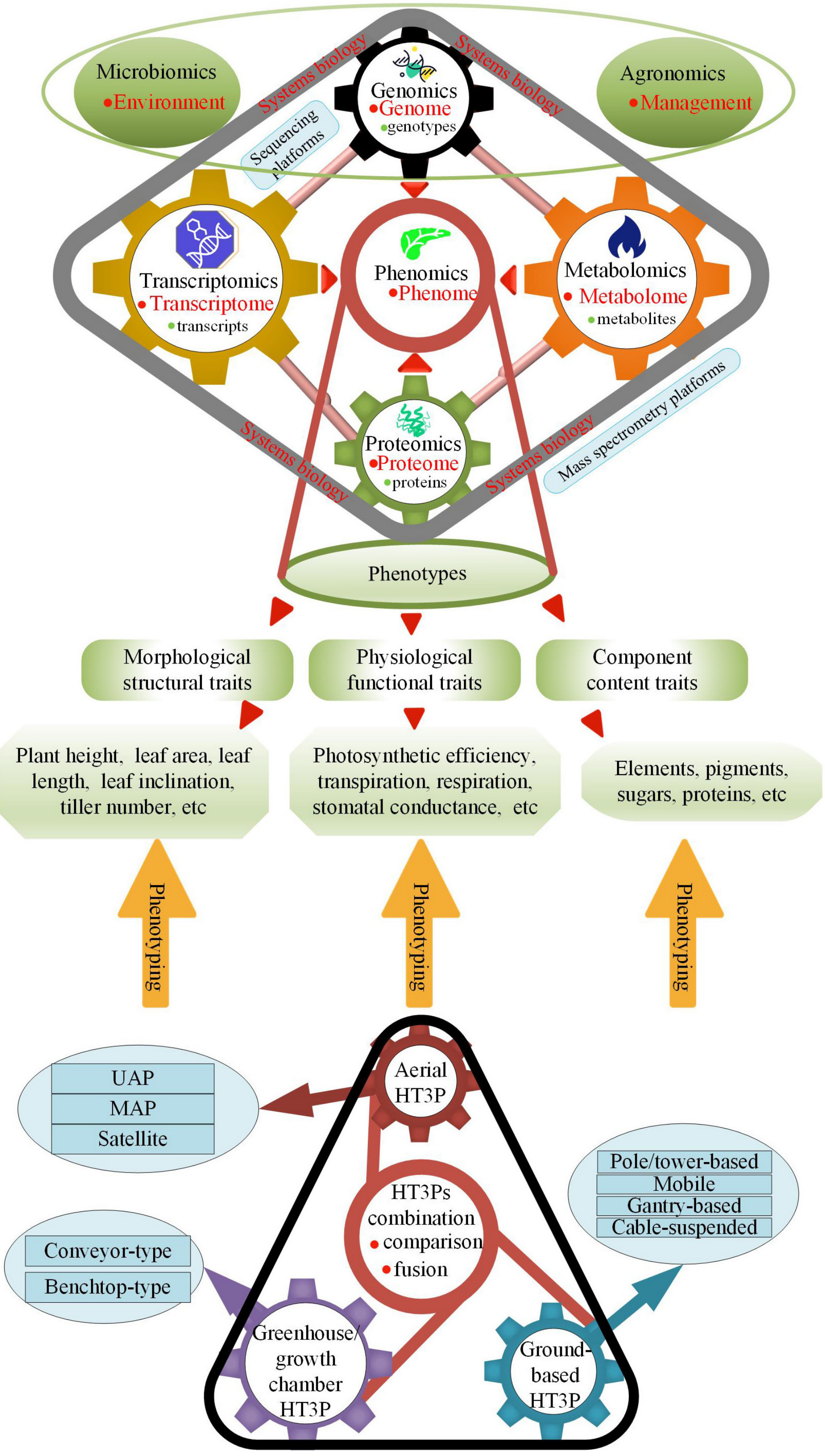


FIGURE 1 | HT3Ps employed for phenotyping plant phenotypes of genotype, environment, and management (G×E×M) interactions advance phenomics; sequencing platforms employed for researching genotypes and transcripts assist in genomics and transcriptomics; mass spectrometry platforms employed for researching proteins, and metabolites promote proteomics and metabolomics; -omics platforms further progress multi-omics in systems biology.

with assistance of HT3P can reveal adaptive mechanisms related to environmental signals, with a view to eventually elucidating their genetic control.

Given the mechanical structure of the platform and movement mode between the sensors and plants, an indoor HT3P can be categorized as either a benchtop type or a conveyor type. **Table 1** shows specific examples and details of these two types of HT3P. No matter which type it is, the phenotyping platform integrates common cameras, supplemental light sources, automatic watering, and weighing devices, to automatically collect plant phenotypic data. Available cameras include those capable of capturing RGB, infrared (IR), fluorescence (FLUO), near-infrared (NIR), multispectral, or hyperspectral images. For example, the FLUO imager is used to obtain chlorophyll or photosynthesis-related characteristics (Choudhury et al., 2019). Hyperspectral imaging in particular provides access to crucial metrics, such as those for photosynthesis, chlorophyll, and nitrogen content. See **Figure 2** for detailed information on the diverse sensors now available to monitor, quantify, and evaluate key agronomic traits.

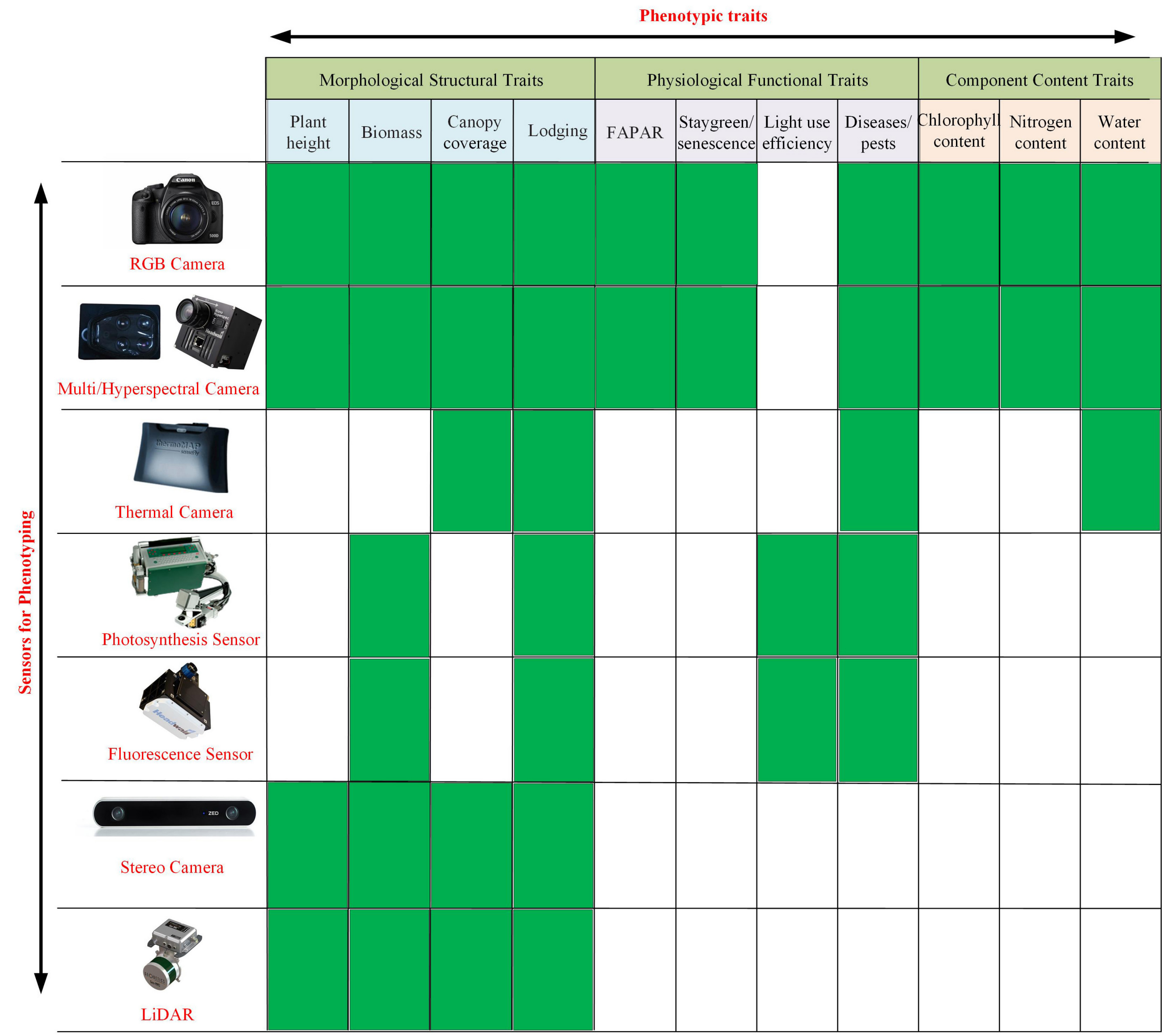
Compared with field conditions, although indoor experiments cannot provide the authenticity of soil system and the complexity of biological and abiotic stress for plants, the purpose of indoor HT3P experiment is to study qualitatively or quantitatively the response of representative or interesting plants to specific environment. Environmental control platform avoids the unpredictable phenotypic variation caused by the interaction between genotype and natural environment (G × E). Therefore, considering uncontrollable factors in the field, HT3P deployed in greenhouse or growth chamber is widely used to study the response of plants to specific growth conditions, and accurately capture the morphological structural, physiological functional or component content phenotypic indicators.

Conveyor-Type Indoor HT3P

The conveyor-type HT3P operates in the "plant-to-sensor" mode. Potted plants are transported into an imaging room with cameras, passing through an automatic door on the conveyor that is controlled by computer for automatic imaging, after which plants are returned to their original growth positions. Cameras are typically installed on the top and side of the

TABLE 1 | Overview of HT3Ps used in greenhouses and growth chambers under environmentally controlled conditions.

Indoor HT3P	Model	Sensors	Throughput (pots)	Plants	Traits	Location	References
Conveyor type	LemnaTec Scanalyzer 3D	RGB, NIR, FLUO	312	Barley	Biomass, plant height, width, compactness, drought stress	Germany	Chen et al., 2014; Neumann et al., 2015
	LemnaTec Scanalyzer 3D	RGB, NIR, FLUO, hyperspectral	672	Sorghum, maize, barley	Biomass, leaf water content	USA	Miao et al., 2020
	LemnaTec Scanalyzer 3D	RGB, NIR, FLUO, hyperspectral	2,400	Chickpea, wheat	Nutrient stress, salt stress, water content, nitrogen content	Australia	Neilson et al., 2015; Atieno et al., 2017; Bruning et al., 2019
	Bellwether	RGB, NIR, FLUO	1,140	Setaria	Plant height, biomass, water-use efficiency, water content	USA	Fahlgren et al., 2015a
	–	Hyperspectral	100	Maize	PLA, NDVI, perimeter, major axis length, minor axis length, eccentricity	USA	Ma et al., 2019
	HRPF	RGB, CT	5,472	Rice	Drought stress, tiller number	China	Yang et al., 2014; Duan et al., 2018
Benchtop type	Phenovator	Monochrome	1,440	*Arabidopsis thaliana*	PLA, PSII efficiency	The Netherlands	Flood et al., 2016
	Phenoscope	RGB	735	*Arabidopsis thaliana*	Rosette size, expansion rate, evaporation	France	Tisne et al., 2013
	–	RGB	350	*Arabidopsis thaliana*	Radiation dosage stress, projected area, convex hull area, perimeter length	Korea	Chang et al., 2020
	Phenoarch	RGB	–	Maize	Growth rate of ear and silk	France	Brichet et al., 2017
	Glyph	RGB	120	Soybean	Water use efficiency, drought stress	Argentina	Peirone et al., 2018
	LemnaTec Scanalyzer HTS	RGB, FPUO, NIR	–	*Arabidopsis thaliana*	Water stress	USA	Acosta-Gamboa et al., 2017



Sensors for Phenotyping	Morphological Structural Traits				Physiological Functional Traits				Component Content Traits		
	Plant height	Biomass	Canopy coverage	Lodging	FAPAR	Staygreen/ senescence	Light use efficiency	Diseases/ pests	Chlorophyll content	Nitrogen content	Water content
RGB Camera	🟩	🟩	🟩	🟩	🟩	🟩		🟩	🟩	🟩	🟩
Multi/Hyperspectral Camera	🟩	🟩	🟩	🟩	🟩	🟩		🟩	🟩	🟩	🟩
Thermal Camera			🟩	🟩				🟩			🟩
Photosynthesis Sensor		🟩		🟩			🟩	🟩			
Fluorescence Sensor		🟩		🟩			🟩	🟩			
Stereo Camera	🟩	🟩	🟩	🟩							
LiDAR	🟩	🟩	🟩	🟩							

FIGURE 2 | Sensors currently available to monitor, quantify, and estimate key morphological structural traits (e.g., plant height, biomass, canopy coverage, lodging), physiological functional traits (e.g., FAPAR, staygreen/senescence, light-use efficiency, disease/pests), and component content traits (e.g., chlorophyll content, nitrogen content, water content) of plants.

darkroom to perform this imaging, and/or the plants are rotated for data acquisition. The automatic door eliminates the interference of ambient light, and there are halogen lamps to provide illumination.

Scanalyzer 3D, a typical conveyor-type HT3P developed by LemnaTec GmbH (Aachen, Germany), has been adopted by some international organizations, covering the following versions (Yang et al., 2020). The Plant Accelerator of the Australian Plant Phenomics Facility (APPF) is a leading international plant phenotyping research institution. Its conveyor HT3P can handle 2,400 plants and is equipped with multiple imaging stations (RGB, NIR, FLUO, and hyperspectral), and this has been used successfully to study the nutrient deficiency of crops (Neilson et al., 2015) and salt tolerance of chickpea (Atieno et al., 2017). The four imaging chambers are separated and function independently of each other. Recently, Bruning et al. (2019) used just two hyperspectral imagers in its hyperspectral imaging room to evaluate the concentration and spatial distribution of water content and nitrogen level in wheat. In another example, the conveyor belt system in the Smarthouse (APPF, University of Adelaide) was used to study the effects of zinc (Zn) and an arbuscular mycorrhizal fungus upon tomato (Brien et al., 2020).

Similarly, Scanalyzer 3D, in the Greenhouse Innovation Center of University of Nebraska-Lincoln, allows the phenotyping of 672 plants with a height of up to 2.5 m, being able to collect RGB, FLUO, IR, NIR, and hyperspectral images from the top and side view of plants (Choudhury et al., 2016). Each imaging room is equipped with a rotating elevator that permits 360 side-views of a given plant (Choudhury et al., 2018). There are three watering stations with balance, which can apply watering to meet the target weight of the pot or in specific volume, for which the amounts of water added are recorded. Since the imaging chamber is self-contained, this HT3P unit allows the employed sensors to be adjusted according to research needs. For example, RGB, NIR, and FLUO cameras are used to analyze the spatiotemporal biomass accumulation of barley under drought stress (Neumann et al., 2015). In studying maize, Ge et al. (2016) used RGB and hyperspectral imaging rooms to analyze this crop's growth and water-use dynamics, in addition to quantifying its leaf water content. To measure the nutrient concentration and water content of plants, Pandey et al. (2017) relied solely on the hyperspectral imaging room of the Scanalyzer 3D, this being the first time hyperspectral data was used to detect the nutrient content of living plants *in vivo*. Furthermore, Miao et al. (2020) segmented the generated hyperspectral images of sorghum and maize, at the organ level, to identify genetic associations, which let them measure plant properties more broadly.

The bellwether phenotyping platform, at the Donald Danforth Plant Science Center, including a Convion (Winnipeg, Canada) growth chamber and an imaging station (LemnaTec Scanalyzer) (Fahlgren et al., 2015a). Plant barcodes on the pots are used for radiofrequency identification (RFID), to match up image data with the metadata. The 180-m-long conveyor belt system can accommodate 1,140 plants, which are transported into FLUO, VIS, and NIR imaging stations through dark adaptation channels. Interestingly, this conveyor system is divided into four modules that can run independently, or as a whole, which increases the research flexibility and scope of potential experiments. To sum up, as a widespread conveyor-belt HT3P for large-sized plants, Scanalyzer 3D is effective in studies of plant biology and plant breeding.

The purpose of the greenhouse is to provide a uniform, controlled environment. But since most conveyor HT3Ps often need to transport plants to the specific imaging room, this introduced microclimatic heterogeneity likely influences the plants' growth and response to environmental changes, rendering the phenotypic data collected inaccurate. Fortunately, the HT3P built by Purdue University overcomes this interference of a differential microclimate (Ma et al., 2019), in that plants are grown on cyclic conveyor belts throughout their whole growth cycle, thus exposing them to the same heat and radiation conditions. Huazhong University of Science and Technology and Huazhong Agricultural University (Wuhan, China) jointly developed a high-throughput rice phenotyping facility (HRPF) with an image analysis pipeline, able to perform color imaging and X-ray computed tomography (CT); it can monitor 15 agronomic traits of 1,920 rice plants (Yang et al., 2014). This HRPF was used to quantify the dynamic response of rice to drought (Duan et al., 2018). However, the investment cost of conveyor HT3Ps is high, and further improvement is needed to enhance flexibility.

The CT platform can high-throughput visualize and quantify external and internal geometric features, which offers the opportunity to collect morphological and anatomical characteristics of plants. There are two general types of CT platforms used in plant sciences: industrial CT scanners and medical CT. Industrial CT scanners with a higher resolution than medical CT, also known as micro-CT, μCT, or nano-CT, can be applied for the subtle phenotypic traits of plants. For example, Tracy et al. (2017) used μCT scanning to obtain detailed three-dimensional phenotypes of *Arabidopsis thaliana* and barley, which allows the measurement of spike size and further accurate staging at the flower and anther stages. This rapid and non-destructive method overcomes the traditional tedious steps in cytological microscopy, such as fixation, sectioning, staining, and microanalysis. Compared to μCT, medical CT can faster scan larger samples and larger numbers of samples, despite its lower resolution. Gomez et al. (2018) used a medical CT platform to study the geometric characteristics of sorghum stem, finding that medical CT estimates were highly predictive of morphological traits and moderately predictive of anatomical traits.

The conveyor-type HT3P (excluding CT) can carry samples of large size (e.g., sorghum and corn) and large capacity, but it may affect those plants with fragile stems due to shaking of the belt. And since spectral information is not collected *in situ*, there are environmental differences between the plants' growth location and the imaging room, which may lead to inaccurate phenotypic data. The CT platform enables the acquisition of meticulous morphological and anatomical traits of plant, which has great application prospects. The future conveyor-type HT3P will aim for high operational stability and environmental homogeneity, providing smooth plant transportation and accurate climate control for plant science research, and enabling complete precise phenotyping of plant traits throughout reproductive period.

Benchtop-Type Indoor HT3P

In measuring phenotypic traits susceptible to environmental changes (temperature, wind, to name a few), especially for small species with fragile stems, it is a wise choice to keep plants still while the sensors are moving about. This is exactly how a benchtop HT3P works, for which the operation mode is one of "sensor-to-plant." The imaging head is integrated with multiple sensors, driven by a computer-controlled mechanical arm, which automatically locates the position where a plant is growing and collects its phenotypic data *in situ*. In general, the benchtop HT3P also features a precisely controlled irrigation and weighing system, with supplemental light sources.

Arabidopsis thaliana is a prime model plant because of its wide distribution, fast life cycle, and relatively small genome (Kaul et al., 2000), making an ideal study species for the benchtop HT3P. The platform for measuring photosynthetic parameters (PSII, photosystem II), named Phenovator, can accommodate 1440 *A. thaliana* plants (Flood et al., 2016). Driven by an XY camera-movement system, the imaging head carrying the cameras measures photosynthesis and projected leaf area (PLA) at eight wavelengths, *via* an eight-position filter wheel installed on the monochrome camera. For the Phenoscope platform, its imaging system consists of digital camera only (Tisne et al., 2013), but it can carry 735 individual pots. Its ingenious feature is that it can continuously rotate each pot, so the sampled plant experiences the same external conditions, thus minimizing micro-environmental variation at the individual plant level and providing high spatial uniformity.

Similarly, the LemnaTec Scanalyzer HTS is equipped with a robotic arm that houses VIS, FLUO, and NIR cameras to take top views of small plants. It is has been used to study the time-dependent effect of water stress on *A. thaliana* (Acosta-Gamboa et al., 2017). Although the number of samples it can process is relatively small, it provides sufficiently rich spectral information. In mutation breeding, the phenotype discovery of many mutants is rapidly posing a limitation to molecular plant physiology research (Fraas and Lüthen, 2015). Recently, Chang et al. (2020) collected the growth images of 350 *A. thaliana* and compared the subtle morphological effects of different radiation dosages during its growing period, obtaining not only dynamic growth behavior information (such as the plant growth rate post-radiation) but also the phenotypic characteristics of dose effects.

Plant silk is normally difficult to detect and quantify because of its unique features. To overcome this, Phenoarch (INRA, Montpellier, France) made a breakthrough, when Brichet et al. (2017) used it to monitor the growth dynamics of corn ears and silks. That HT3P unit has two imaging cabins. First, the plants are rotated at a constant speed, and RGB cameras determine the spatial coordinates of the ears on them. Then the robotic arm assists in the automatic positioning of the camera at a 30-cm distance from the ear, to continue collecting of high-resolution images of silks. In this way, the daily growth of ear and silk of hundreds of plants can be tracked. Crop 3D, developed by the Chinese Academy of Sciences (Guo et al., 2016), takes LiDAR as the core sensor and integrates a high-resolution RGB camera with a thermal and hyperspectral imager, applying one key trigger to synchronously acquire multi-source data and extract plant morphology parameters. Specifically, the sensors adopt a vertical downward or overhead mode, to mount and shoot, carrying out single row scanning, multi-row scanning, and fixed-point positioning scanning.

Although a low-cost and non-commercial platform has a small sample capacity, it could also generate high-throughput phenotypic data. Glyph, as a representative, consists of four bridge-like structures, whose drip irrigation equipment and digital camera form a gantry that moves on the track between pair of rows. It has been successfully used for predicting the field drought tolerance in soybeans (Peirone et al., 2018). The SITIS platform, which consists of PVC pipes with irrigation points, was used to evaluate water stress tolerance of cotton cultivars (Guimarães et al., 2017), and it also enables the evaluation of plant roots. However, this platform is not an image-based, non-destructive one to measure plant traits, so it still requires much manual operation and experimental processing. The human-like robotic platform is a newer method to measure plant phenotypic traits instead of doing such manual operations. The *vivo* robotic system, consisting of a four degree of freedom (DOF) manipulator, a time-of-flight (TOF) camera, and a gripper integrated with an optical fiber cable and thermistor, can be used for the automatic measurement of maize and sorghum leaf traits (Atefi et al., 2019). More specifically, the TOF camera acts as the vision system, and the gripper can measure VIS-NIR spectral reflectance and temperature; however, its capture speed, as well as its capture success rate (78% for maize and 48% for sorghum), need further improvement. The rapid development of such a robot system can provide reference data and supplementary support for image-based plant phenotyping.

Strictly benchtop-type HT3Ps tend to focus on model plants of small size (e.g., *A. thaliana*) and allow for the collection of trait data associated with subtle phenotypic changes, and their situ extraction also ensures homogeneity of growth environment and undisturbed development. However, sophisticated commercial HT3Ps tend to be capital-intensive, while low-cost self-developed platforms have small sample capacity and low throughput, whose quality, credibility, and abundance of phenotypic data can be somewhat reduced. Fortunately, artificially intelligent plant growth chambers, plant factories, and rapid iterative breeding have opened new avenues for indoor HT3Ps. The future benchtop-type HT3P may be able to omnidirectionally monitor, capture, and track subtle morphological and physiological changes of multi-level traits in a wide range of model plants, with high-throughput and full-automation, to reveal functional gene expression and biogenetic regulation patterns.

HT3P FOR FIELD PHENOTYPING IN NOTORIOUSLY HETEROGENEOUS CONDITIONS AND RELATIVELY UNCONTROLLABLE ENVIRONMENTAL FACTORS

Plants that grow naturally in the field are affected by weather (e.g., rain, frost, snow), biotic and abiotic stresses (e.g., drought, low-temperature, low-nitrogen, pests), as well as soil properties (e.g., nutrient gradients, heterogeneity, micro-environment), all of which are extremely distinct from the environment of greenhouse and growth room, making the field crop phenotype an intricate one. Whereas controlled environment, image-based phenotyping platforms are almost universally popularized worldwide, the majority of crop breeding appears in the field with little if any selection in controlled environments (Furbank et al., 2019). And the indoor environment can only simulate but not recreate the real field setting. These factors stimulated the exponential increase of a wide variety of HT3Ps in fields.

Apparently, field HT3Ps operate in the "sensor-to-plant" mode. According to their usage scenarios and imaging distance, field HT3Ps can be categorized into ground-based and aerial platforms. The relationship between a platform's characteristics and field crop traits determines the efficiency of the phenotype platform to a certain extent (Kuijken et al., 2015). Based on this, ground-based platforms can be further classified as pole/tower-based, mobile, gantry-based, and cable-suspended. Likewise, aerial platforms could be categorized further, as the unmanned aerial platform (UAP), manned aerial platform (MAP), and satellite platform. **Figure 3** shows specific scenarios of various types of applied HT3Ps.

Despite the complex interactions of genotype, environment, and management (G×E×M), the proliferation of a wide variety of HT3Ps in recent years has greatly assisted researchers in understanding the genetic structure of crops, obtaining high-quality genetic gains, and improving the ability to genetically analyze crop traits related to yield and stress resistance. However, the diversity of HT3Ps also brings with considerations of availability, feasibility, standardization, big data, and reliability.

Ground-Based Field HT3P

Ground-based HT3P means proximal phenotyping that can provide higher resolution data than aerial remote sensing. It is convenient to collect phenotypic data of time series and analyze the dynamic response and time dependence of phenotypes. However, the ground-based HT3P is not suitable for large-scale phenotyping tasks. **Table 2** shows specific examples and details of these four types.

Pole/Tower-Based Field HT3P

The pole/tower-based HT3P is formed when sensors are mounted directly atop a pole or tower made of aluminum, steel, or plastic fibers, which can be of stationary or mobile

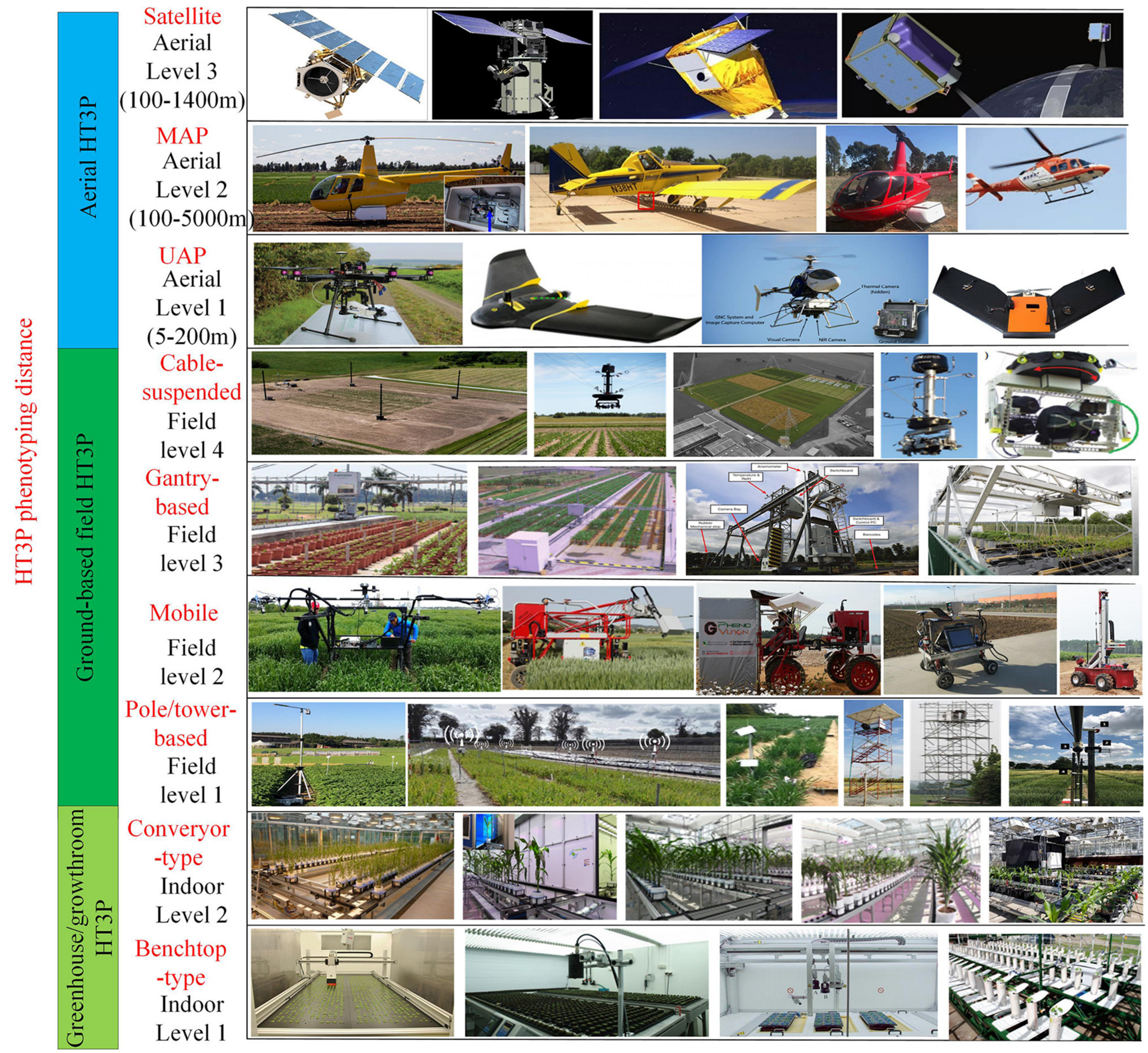


FIGURE 3 | The various types of HT3Ps mentioned in this review (partial display) including HT3Ps in the greenhouse/growth chamber (i.e., benchtop-type and conveyor-type), field ground-based HT3Ps (i.e., pole/tower-based, mobile, gantry-based, and cable-suspended) and aerial HT3Ps (i.e., UAP, MAP, and satellite).

TABLE 2 | Overview of ground-based HT3Ps used in the field under real uncontrolled environmental conditions.

Field HT3P	Designation	Sensors	Feature	Plants	Traits	Country	References
Pole/tower-based	–	Terrestrial laser scanner	Maximum load of 50 kg, 3.8 m high, covers 120 m	Maize, soybean, wheat	Canopy height	Switzerland	Friedli et al., 2016
	CropQuant	RGB, NIR	Combined with IoT	Crop	Crop growth rate	UK	Zhou et al., 2017
	PhenoCam	RGB	Large phenotyping network	Ecosystem	Canopy greenness	USA	Richardson et al., 2018
	–	Hemispherical video camera	Automatic camera track system	Wheat, oat, barley	Crop lodging	USA	Susko et al., 2018
	–	Laser-Induced Fluorescence Transient (LIFT)	Covers 50 m	Barley, Sugar Beet	Photosynthesis	Germany	Raesch et al., 2014
	–	RGB, NIR	Consists of two 8-m high towers	Rice	Shoot biomass, panicle number, grain weight	Colombia	Naito et al., 2017
Mobile	–	LiDAR, RGB, thermal IR, IR thermometer, hyperspectral	Speed of 1 m/s	Wheat	Canopy height, leaf angular distribution, leaf area, leaf volume, spike number, VIs, canopy transpiration	Australia	Deery et al., 2014
	–	Ultrasonic, NDVI, thermal IR, spectrometers, RGB	A "stop-measure-go" model	Soybean, wheat	Canopy height, NDVI, canopy temperature	USA	Bai et al., 2016
	Phenomobile Lite	LiDAR, RGB, NDVI	Three-wheeled buggy	Wheat	Plant height, biomass, ground cover	Australia	Jimenez-Berni et al., 2018
	GPhenoVision	RGB-D, thermal, hyperspectral	Modularity, customizability	Cotton	Canopy height, width, growth rate, projected leaf area, volume, yield	USA	Jiang et al., 2018
	–	LiDAR	Group observation on parcel, "stop-measure-go" model	Maize	Plant height	China	Qiu et al., 2019
	–	Ultrasonic, spectrometer, RGB, IR radiometer	Emergency stop and inspection	Soybean	Canopy height, canopy coverage, NDVI	USA	Murman, 2019
Gantry-based	LeasyScan	Planteye	Continuous phenotyping	Peanut, cowpea, pearl millet, maize	Canopy transpiration, plant height, 3D leaf area, water use efficiency	India	Vadez et al., 2015; Sunil et al., 2018
	Phénofield^{r}	RGB, VIS-NIR, LiDAR	Conditions controlled in the field	Wheat	Water stress resistance, nitrogen stress resistance	France	Beauchene et al., 2019
	Field Scanalyzer	RGB, FLUO, thermal IR, hyperspectral, 3D laser scanner	Fully automatic or remote manual operation	Wheat	Canopy height, spike number, canopy closure, canopy temperature, NDVI, photosynthesis	UK	Virlet et al., 2017
	Mini-Plot	Hyperspectral	Divided into open field and closed greenhouse areas	Barley	Disease severity	Germany	Thomas et al., 2018
Cable-suspended	NU-Spidercam	Multispectral, thermal IR, LiDAR, VIS-NIR spectrometer	Covered 0.4 ha, 30-kg payload, maximum speed of 2 m/s	Soybean, maize	Plant height, ground cover, canopy temperature	USA	Bai et al., 2019

(Continued)

TABLE 2 | Continued

Field HT3P	Designation	Sensors	Feature	Plants	Traits	Country	References
	FIP	Spectrometer, ultrasonic, DSLR, thermal, laser scanner, operator camera	Covers 1 ha, 2 to 5 m above the canopy, 12-kg payload, maximum speed of 2 m/s	Winter wheat, maize, soybean	Canopy cover, canopy height	Switzerland	Kirchgessner et al., 2017

type. Although this platform is simple in structure, similar to a small weather station, high throughput and complexity are not necessarily synonymous.

Friedli et al. (2016) applied a pole-based, terrestrial laser-scanning (TLS) platform, to monitor canopy height growth in maize, soybean, and wheat. The TLS is done *via* a laser scanner mounted upside down on a 3.8-m high aluminum elevator tripod. Its time and spatial resolution depend on the crop variety assessed and scanning distance. Combined with "Internet of Things" (IoT), CropQuant is equipped with RGB and No Infrared cameras to continuously monitor crop growth through high-resolution time-lapse photography (Zhou et al., 2017). It can be powered by batteries and solar panels and connected to an in-field WIFI network, as a mesh network node, to form an IoT-mode HT3P that quantifies crop growth and development. Furthermore, PhenoCam is a large phenotyping network, consists of a series of widely deployed digital cameras that automatically capture RGB images (typically, at 30 min intervals) to track biomes' vegetation phenology (Richardson et al., 2018). The camera is mounted on a pole, mast, or building. Although the data source is only visible images, such massive time series data sets can monitor the dynamic changes of an ecosystem.

A mobile handheld pole-based, Phenocorn, is integrated with a GreenSeeker (portable device), an IR thermometer, a web camera, and a global positioning system (GPS) receiver (Wei, 2017). This device can simultaneously collect normalized vegetation index (NDVI) and canopy temperature. Although the 12 kg hand-held Phenocorn can be carried on a person's back, it still requires much manual labor. Fortunately, it can modified for use in a cart that is manually pushed to collect phenotypic traits in the field (Crain et al., 2016). As is well-known, it is quite difficult to capture, track, and quantify crop lodging and crop movement. In tackling this, Susko et al. (2018) developed an automatic camera-tracking HT3P that consisted of a hemispherical video camera, a computer, and an industrial curve track system; its motor-driven camera moves along the track for phenotypic data acquisition. What makes this breakthrough so novel is that it can be used for dynamic imaging in video or static frequent imaging, thus allowing for the study of new plant phenotypes.

The tower-based HT3P is similar to the stationary pole-based platform, with sensors installed atop the tower, but the dimensions and height of a tower-based platform are generally higher and larger than that of pole-based. In a study on photosynthetic efficiency of barley and sugar beet, laser-induced fluorescence transient (LIFT) instruments were placed on the top of a 10-m high scaffold to measure the photosynthetic performance of agroecosystems (Raesch et al., 2014). However, the LIFT signal in the target area introduced noise from plant stems and the soil. In later work, Naito et al. (2017) installed an improved multispectral single-lens imaging system (i.e., VIS and NIR cameras), on two 8-m high towers, which could collect crop images from eight angles to estimate rice yield-related traits. Their results showed that the system has great potential for yield estimation during early crop development.

Both pole-based and tower-based HT3Ps are easy and low-cost to build and maintain, and are convenient for temporary use and multi-site deployments to form networks. However, phenotypic area of coverage and spectral information are extremely limited for a single unit, and multi-site trials increase the cost for large-scale field experiments. Looking ahead, being portable, scalable, rotatable, robust, and easy to install and remove are anticipated key features of pole/tower-based platforms, and phenotypic networks of economically-efficient distributed pole/tower-based HT3Ps will play a prominent role in the future of multi-site large-scale experiments and the calibration of high-dimensional phenotypic data.

Mobile HT3P

The mobile HT3P can move through the field and collect crop phenotypic traits in a semi-automatic or fully automatic manner, including refitted agricultural machinery (e.g., tractor, sprayer, or harvester), self-developed mechanical platform (e.g., cart or buggy), and commercial automatic platform. A mobile HT3P is generally composed of four subsystems: sensing system, data acquisition system, mechanical platform, and drive system. The sensing system covers GPS, environmental sensors (e.g., sunlight, wind speed), and phenotypic sensors (e.g., RGB, multispectral, hyperspectral, thermal). The data acquisition system generally is the data acquisition software on an onboard computer. Mechanical platform is for load-bearing and mobility, and the drive system can be classified as electricity, engine, and manpower. Theoretically, any agricultural mobile platform has the potential to be converted into a mobile HT3P.

The mobile HT3P converted from a tractor, sprayer, or harvester makes rational use of precision agricultural machinery that already exists. In the early stages, a pioneering field-based mobile was developed in Canberra's High-Resolution Plant Physics Facility (Deery et al., 2014). It was equipped with a height-adjustable sensor array, including LiDAR, RGB, thermal

IR, and hyperspectral cameras. Driven by the mobile, the sensors pick up traits' data along the surveyed plots. A cherry picker installed a linear scanning bar with spectral cameras has also been applied as a mobile HT3P (Pinto et al., 2016); its imaging cameras aim to collect canopy radiation in a linear push broom mode. The temporal change of canopy photochemical activity was tracked by generating sun-induced chlorophyll fluorescence (SIF) map. A multi-sensor system developed by Bai et al. (2016) consists of five sensor modules, and this platform was used to measure canopy traits of soybean and wheat. They adopted the "stop-measure-go" mode, to ensure the sensors precisely aligned with the plots, collecting plot-level phenotypic data that were free of blurring. But to control such a HT3P, a certain number of operators must participate. GPhenoVision is assembled from a high-clearance tractor, carrying RGB-D, thermal, and hyperspectral cameras for quantitative assessment of cotton canopy growth and development (Jiang et al., 2018). Customizability and modularity are the two key characteristics of GPhenoVision, which allows researchers to rapidly develop and upgrade sensing modules with specific phenotyping purposes, reduces development effort when adding or removing modules, and prevents malfunction of the entire system. Jimenez-Berni et al. (2018) installed a LiDAR, an NDVI sensor, and a digital camera on a low-cost three-wheeled buggy, called Phenomobile Lite, to evaluate canopy height, ground cover, and aboveground biomass. However, to change its direction, a "stop" button is need to be manually pressed, and almost the whole process requires operator follow-up.

To achieve the automatic measurement of yield-related traits, Bao et al. (2014) presented a mobile HT3P that could automatically obtain stereo images of sorghum plants. It was retrofitted from a garden tractor, equipped with six stereo camera heads on a vertical pole, which can be triggered synchronously. Three-dimensional images of two rows of crops can be collected in a single pass. The drawback is that before each bout of data collection, the platform must be driven manually through the field, stopping, and recording each sampling point to generate paths. Later, an agricultural mobile robot mounted with a 360°-view LiDAR, developed by Qiu et al. (2019), was used to efficiently calculate the row spacing and plant height of a maize field. Compared with in-row and one-by-one phenotyping methods, the 3D laser scanner sitting atop this robot obtains group observation of parcels. While it moves in a "stop-and-go" manner, the phenotypic data of parcel-level plant group could be simultaneously collected. Unlike the wheeled mobile platform, Stager et al. (2019) employed a modified crawler robot to collect sub-canopy traits at low elevation. But branches, tillers, or roots may hinder the movement of such crawler robots. Automated robots offer the prospect of unattended field operations, which likely be a major focus of future research of agricultural phenotyping platforms.

By letting the phenotyping height of the mobile HT3P vary, to adapt to different growing stages of crops, Flex-Ro was developed to identify differences in the emergence and maturity stages among soybean varieties (Werner, 2016). PhenoBox of Flex-Ro mainly integrates data acquisition hardware, while its height-adjustable PhenoBar, located at the front of this mobile platform, mainly consists of three sensor units, capable of covering 4.5-m swath (Murman, 2019). An operator controls the machine through the remote box or a MATLAB application called FlexRoRun. It is worth mentioning that a 3D smart sensor is incorporated, for obstacle detection, which successfully detects pedestrian-sized objects and triggers parking. This will be a critical security consideration for future robotic HT3Ps. Likewise, a multi-purpose field robot in combination with various apps can achieve different functions. For example, Bonirob robotic platform with the phenotyping app, penetrometer app, and precision spraying app can monitor plant growth, measure soil parameters, and apply chemical weeding, respectively (Bangert et al., 2013). Unlike the mobile HT3P for small-sized or early-growing crops, Robotanist can navigate autonomously in the fields of tall crops, such as corn or sorghum (Mueller-Sim et al., 2017). Interestingly, it has a three DOF manipulator that can touch and measure the strength of plant stalks, which is a phenotypic innovation based on contact.

A semi-automatic self-made mobile HT3P can reduce the development cost and soil compaction (because of lightweight architecture), but normally requires one or more operators to follow-up (Bai et al., 2016, 2018; Jimenez-Berni et al., 2018). Moreover, because of the "stop-measure-go" mode and slow response speed of low-cost sensors, the efficiency of crop traits' data acquisition cannot be guaranteed. Concerning the mobile HT3P based on the modified tractor, sprayer, or harvester, it often needs special personnel to drive. Its large volume and weight risk causing soil compaction and mechanical disturbance to the crops, which precludes the deployment in the field. And the faster travel speed than self-made cart also may not guarantee the quality of phenotypic data obtained. But the payload is large, so it can integrate diverse sensors to collect multi-source information. Robotic HT3P is capable of automatically navigating through the field and collecting data on crop traits, as well as doing continuous phenotyping throughout the day and night, but its development and maintenance costs are expensive. The future mobile HT3Ps will likely aim for lightness, automation, modularization, and customization, and the mechanical arm and emergency braking mechanism will offer potential applications that could greatly improve the flexibility, autonomy, and security.

Gantry-Based Field HT3P

As a gantry frame is equipped with a sensor box, and moves along the track and collects crop traits along XYZ directions, it becomes the gantry-based HT3P. When it moves back and forth on the track, repeated phenotyping done this way avoids the possibility of soil compaction and damage of crops' normal development.

LeasyScan, equipped with a set of scanners (PlantEye F300, Phenospex, Heerlen, the Netherlands), can perform continuous, synchronous and automated monitoring of plant water use and leaf canopy development, *via* linear movement above the surveyed plants (Vadez et al., 2015). The trigger and stop of each measurement are controlled by a mechanical barcode, which is also used for distance calibration from the scanner to the ground. LeasyScan was also utilized in a pre-breeding, genetic resource identification experiment with hybrid maize (Sunil et al., 2018), in which it measured plant height, 3D leaf area (i.e., total leaf area), and leaf area index (LAI). Similarly, Virlet et al. (2017) applied Field Scanalyzer (LemnaTec) to monitor and quantify morphological traits of wheat organs and canopy. Through the control software in the master computer, the gantry-based HT3P can operate in the full-automatic or manual mode. Unfortunately, though it can sample continuously, 24 h per day, but no two sensors can collect images simultaneously. Then, Field Scanalyzer also was used to determine wheat canopy height and further validate the potential of functional mapping analysis for detecting persistent quantitative trait loci (QTLs) (Lyra et al., 2020).

A phenotypic experiment that precise management is involved in the sufficiently realistic field environment is ideal for deciphering G×E×M interactions. PhénoField^R is the first field-based facility in the European Union, one equipped with high-throughput phenotyping devices on a gantry, automatic irrigating mobile rainout shelters on tracks, and environmental recording sensors. By controlling both irrigation and fertilization, this gantry-based HT3P was used to investigate water and nitrogen stress (Beauchene et al., 2019). Likewise, Mini-Plot, developed by Forschungszentrum Jülich, consists of a closed greenhouse area and an open-fenced area, containing 90 and 30 Mini-Plots, respectively; it was used to quantify the disease severity of barley varieties (Thomas et al., 2018). The measuring head consisting of a hyperspectral sensor, a mirror-based scanning system, and an automatic positioning system moves over the plant for imaging. Such gantry-based HT3Ps for phenotyping, which can integrate genetics, field environmental factors, and management practices, has greatly enhanced the understanding of genetic control pattern for breeders and plant biologist.

However, once the construction and installation of the gantry-based HT3P is completed, multi-site experiments can be expensive or even inoperative. High development, operation, and maintenance costs are also several factors that researchers should be aware of and bear in mind. Fortunately, gantry-based HT3Ps with a high degree of automation can collect time-series phenotypic data of high resolution. Owing to its large payload and continuous operation ability, which greatly improves expansibility, this approach does provide an opportunity to better understand the dynamics of plant circadian rhythms. Extending phenotypic area, shrinking volume, and reducing cost will imbue the gantry-based HT3Ps with better application prospects.

Cable-Suspended Field HT3P

A cable-suspended HT3P is mainly composed of sensing system, data acquisition system, mechanical transmission system, and drive system. Sensing system and data acquisition system are generally integrated in a sensor bar. Winches, cables, poles, and pulleys make up the drive and mechanical transmission system. Four poles are distributed in four corners of the field and a winch house for controlling the cables sits at the bottom of the pole. The sensor array with a GPS unit can be precisely positioned above the interested region for traits' data acquisition through cable-driven.

A typical representative of cable-suspended HT3P is the "FIP," located at the ETH research station in Switzerland. Its sensor heads integrated with a DSLR camera, a laser scanner, and a thermal camera can collect crop phenotypic data from 2 to 5 m above the canopy. The feasibility of FIP in monitoring canopy coverage and canopy height of winter wheat, maize, and soybean was verified by Kirchgessner et al. (2017). Furthermore, the Nu-Spidercam at the University of Nebraska has a multi-sensor systems, a subsurface drip irrigation (SDI) system and an automatic weather station (Bai et al., 2019), which is capable of accurately and flexibly capturing crop traits, such as plant height, canopy cover, and spectral reflection. Its sensor bar integrates numerous sensors, so software is designed for selecting available sensors to meet particular experimental needs. A drawback is that it only can operate continuously for 6–8 h when fully charged.

Phenotyping sensors of the cable-suspended HT3P is performing tasks over the crop canopy by cables during data collection of agronomic traits, meaning that the cable-suspended HT3P has lower dependence on soil conditions and less interference to plants than mobile HT3P. Phenotypic traits information of time series can still be collected within the established crop growth-monitoring period. And it can cover larger phenotypic region but has lower load than do gantry-based HT3P. In general, field HT3Ps are considerably weather-dependent and region-limited. Allowing stabilized imaging of plants clusters and continuous monitoring of crop growth at low elevation for a longer duration will be a likely future trend for cable-suspended HT3Ps.

Aerial HT3P Vulnerable to Weather Constraints and Aviation Regulatory

According to the imaging distance employed, aerial HT3Ps encompass the Unmanned Aerial Platform (UAP), Manned Aerial Platform (MAP), and satellite platform. UAP usually requires an operator at a ground station to remotely control or execute the flight task, by following a set planned path. Crop images are automatically acquired by onboard sensors. Nevertheless, MAP needs a dedicated person to pilot the aircraft and another passenger to manually capture the crop images while in flight. UAP and MAP typically carry global navigation satellite system (GNSS), but often this must be coupled with ground control points (GCPs) and calibration boards, for accurate georeferencing of the acquired data and for correcting its spatial resolution. A satellite platform collects large-scale field images from space, letting one access and download this satellite data as needed. **Table 3** lists the specific applications and details of the aerial HT3Ps. When compared with a ground-based platform, an aerial HT3P can collect broader-scale phenotypic regions in a shorter period. However, the aerial platform has strong weather dependence and relatively low spatiotemporal resolution because of its flight altitude and imaging distance.

Unmanned Aerial Platform (UAP)

In recent years, due to the reduction of Unmanned Aerial Vehicle (UAV) prices and the relaxation of air traffic regulations, the application of UAP in agricultural research has increased exponentially. This so-called UAP is a kind of aerial platform, for which UAV is the carrier, that integrated onboard sensors, a GPS unit, an inertial measurement unit (IMU), a battery and

TABLE 3 | Overview of aerial HT3Ps subject to air regulatory regime and weather constraints.

Aerial HT3P	Designation	Sensors	Flight altitude (m)	Plants	Traits	References
UAP	Phantom 4	RGB	20	Inbred lettuce lines	Carotenoid content	Mascarenhas Maciel et al., 2019
	3DR Solo quadcopter	Multispectral	45, 50	Maize	NDVI, chlorophyll red-edge index (CHL), hemispherical-conical reflectance factors (HCRF)	Fawcett et al., 2020
	Customized	Hyperspectral	80	Winter barley	NDVI, yield	Oehlschläger et al., 2018
	Self-developed octorotor	RGB, multispectral	25	Rice	Canopy height, VIs, canopy coverage	Wan et al., 2020
	Matrice 600 Pro	RGB, multispectral camera, infrared thermal	50	Cotton	Yield	Feng et al., 2020
	Tuffwing Mapper	RGB	120	Sorghum	Plant height	Han et al., 2018
	Ebee	Multispectral	50	Wheat	Yield	Hu et al., 2020
	Self-developed	Multispectral, thermal	150	Maize	Low-nitrogen stress resistance	Zaman-Allah et al., 2015
	Anaconda	RGB, multispectral	120	Sorghum, maize	Plant height, VIs	Shi et al., 2016a
MAP	Robinson R44 Raven helicopter	Radiometrically-calibrated thermal	60, 90	Wheat	Canopy temperature	Deery et al., 2016
	Air Tractor AT-402B	RGB	152–3,048	Crop	Pest severity	Yang and Hoffmann, 2015
	–	LiDAR	1,500	Maize	Biomass	Li et al., 2015
Satellite	GeoEye-1	Multispectral	684 k	Turfgrasses	Nitrogen content	Caturegli et al., 2015
	RapidEye	Multispectral	630 k	Wheat	Nitrogen stress	Basso et al., 2016
	Sentinel-1 and RADARSAT-2	Synthetic aperture radar (SAR)	700, 798 k	Wheat	Crop height, angle of inclination	Chauhan et al., 2020
	Fluorescence explorer (FLEX)	Fluorescence Imaging Spectrometer (FLORIS)	814.5 k	Terrestrial vegetation	Photosynthesis	Drusch et al., 2017

a crucial gimbal—for correcting the influence of pitch and roll motion—to collect phenotypic data at the plant canopy scale. To collect high-precision geographic position of plots, both GCPs and calibration boards are needed. Its successful phenotyping of plants depends on the characteristics of UAV and the properties of the deployed sensors (Sankaran et al., 2015). According to its most distinguishing feature, UAPs can be classified as multi-rotor or fixed-wing. UAV's flight, however, is weather-dependent, and the ideal conditions are clear, windless, and dry weather, similar to those required when applying agronomic inputs.

Because of its limited payload and endurance, UAP can only carry a finite number of phenotypic sensors (generally, no more than 3). A multi-rotor UAV can be a quadcopter, hexacopter, or octocopter. Compared with its fixed-wing counterpart, it has lower flight altitude and slower flight speed, but is capable of vertical takeoff and landing (i.e., it can hover). With respect to the multi-rotor UAP, the use of one sensor is most common. For example, RGB images acquired by multi-rotor UAPs have been widely used for researching the growth rate of wheat (Holman et al., 2016), carotenoid levels of inbred lettuce lines (Mascarenhas Maciel et al., 2019), vegetation index (Buchaillot et al., 2019), wheat height (Villareal et al., 2020), growth of different maize inbred lines (Wang et al., 2019), and canopy extraction of orchard (Wu et al., 2020). Parrot Sequoia (Parrot, France) and Micasense RedEdge (Micasense, US) are the more familiar multispectral cameras used with UAP. For example, Parrot Sequoia was used to evaluate the accuracy and spatial consistency of hemispherical-conical reflectance factors (HCRF) (Fawcett et al., 2020). Notably, cameras are fixed (at a 3-degree angle) to offset the average forward tilt in flight. The recent declining cost of thermal cameras has made airborne thermography more widely used. In addition, ICI camera was proven to be better than either the FLIR or thermomap camera for evaluating plants' physiological and biochemical characteristics (Sagan et al., 2019a). However, atmospheric and emissivity calibration are remained challenges to thermal imaging. Likewise, hyperspectral sensors on the multi-rotor UAP were used to predict the yield of winter barley (Oehlschläger et al., 2018). Interestingly, for that, they used a mirror to guide the ground images into horizontally-positioned sensors. Compared with mere RGB data, collecting additional spectral data enables more robust predictions. In such a setting, RGB and multispectral sensors are integrated into the UAP, as was done to evaluate the yield of rice (Wan et al., 2020), plant height of sorghum (Kakeru et al., 2017), ground cover of cotton (Duan et al., 2017), and senescence rate of wheat (Muhammad et al., 2018). To obtain multi-source phenotypic data, three sensors are also employed on the multi-rotor UAP. For instance, an RGB camera, a multispectral camera, and an thermal IR imager were used together for cotton yield estimation (Feng et al., 2020). However, due to the limited payload, two independent flights of a multi-rotor UAP are occasionally required. For example, in assessing genotypic differences in durum wheat production, RGB images were obtained on the first flight and multispectral canopy information later collected on the second pass (Gracia-Romero et al., 2019).

Compared to the multi-rotor, a fixed-wing UAP has longer flight time, higher flight altitude, greater payload, and faster flight speed. But the fixed-wing UAP lacks hovering capability and has certain requirements for its takeoff and landing (e.g., a runway). The flight speed may cause blurred images, a problem resolved by using imaging sensors with high frame rate and short exposure time (Shi et al., 2016b). Although the fixed-wing UAP does have a relatively large payload, it is almost always relied on a single sensor. For instance, an RGB camera mounted on Tuffing Mapper (Tuffing LLC, Boerne, USA) was used to evaluate the plant height of sorghum (Han et al., 2018). The fixed-wing UAP has a semi-autonomous horizontal takeoff and landing (HTOL), controlled by the Pixhawk controller. Similarly, an RGB camera, with an internal infrared filter removed for color infrared (CIR) detection, was used to assess the height and crown diameter of olive trees (Díaz-Varela et al., 2015). The eBee UAV (senseFly, http://www.geosense.gr/en/ebee/) is becoming a commonly used fixed-wing platform. A multispectral camera mounted on eBee was used to evaluate the yield of early wheat genotypes (Hu et al., 2020) and thermal camera performance (Sagan et al., 2019a), as well as for identifying, positioning, and mapping weedy patches of *Silybum marianum* (Tamouridou et al., 2017). Research on this platform shows that using its highest resolution fails to provide the highest accuracy for weed classification. To assess the spatial variation of maize under low nitrogen stress, Zaman-Allah et al. (2015) developed a fixed-wing UAP with a multispectral camera and a thermal camera, controlled by an automatic navigation system. Fixed-wing UAPs have enough of a payload to carry three sensors. The Anaconda (ReadyMadeRC, Lewis Center, Ohio), fixed-wing UAP equipped with two multispectral cameras and a high-resolution digital camera, was used to collect phenotype data of corn and sorghum (Shi et al., 2016a), for which it employed the traditional configuration of a twin boom thruster. Interestingly, an external GPS unit was added to the digital camera for accurate positional information. Although the 3DR Pixhawk autopilot system has been applied for autonomous takeoff and landing, during its flight the Anaconda must be controlled manually. Ingeniously, the unmanned helicopter is an alternative type of UAP. For example, the Pheno-Copter (gas-powered) with a visible camera, a NIR camera, and a thermal IR camera was applied to measure sorghum ground cover, sugarcane canopy temperature, and crop lodging (Chapman et al., 2014). However, Pheno-Copter's flight needs to be controlled from a ground station, and this often requires specialized training.

If accurate geo-registration accompanies the image acquisition process of UAP, precise micro-plot extractions should be greatly improved (Hearst, 2014). To achieve this goal, he sensor configuration pattern must also undergo improvement. The data collected by UAP requires a GNSS and an inertial navigation system (INS) for spatial matching and georeferencing. To facilitate the calculation of the camera's actual geographic location corresponding to a given image acquisition, a more than 60% overlap of adjacent images is generally needed (Jin et al., 2017). In this context, the spatial offset (lever arm) and angle relationship (boresight angles) between the GNSS/INS unit and sensors are of great significance. Furthermore, compared with the lever arm, the boresight angle matters more. In view of this, Habib et al. (2018) recently employed a UAP, with push-broom hyperspectral sensors, to establish three boresight angle-calibration methods (around GCPs, tie points, and approximate means). While spatial positioning information may be obtained from an airborne GNSS/INS unit, its resolution is not high enough. Therefore, in order to carry out high-precision geographical referencing, establishing a proper layout of ground control points (GCPs) and calibration boards is necessary (Yu et al., 2016). Further, multifunctional GCPs were suggested for calibrating geometric and reflectance, and their usage significantly improved the phenotyping accuracy and also reduced manpower (Thomasson et al., 2019). Likewise, Han and Thomasson (2019) developed an automatic mobile GCP equipped with two RTK-GPS units, a navigation computer, and an integrated driving controller. It can reliably recognize and predict the behavior and activities of the UAP during the flight instead of traditional fixed GCPs, which explores the potential of improving the accuracy and efficiency of data collection.

Since the relevant airspace regulations remain strict, the huge potential of UAP cannot yet be fully exploited. Despite being susceptible to weather, payload, endurance, and aviation regulatory constraints, UAPs characterized by large-scale phenotyping, efficiency, flexible flight plans, and relatively low cost have gradually shifted the trend in phenotypic missions from the ground into the air. In recent years, with policy adjustment, hardware optimization, commercial UAVs' price reduction, advances in battery technology, and operation simplicity, UAP has come to fully display its outstanding abilities in plant phenotyping and plant science. Developing strategies for crop phenotype by remote sensing (Yang et al., 2017), improving performance such as endurance, payload, and stability, reducing the cost of sensors, and promoting data processing capabilities are the future trends of UAP.

Manned Aerial Platform (MAP)

The MAP is converted from a manned helicopter or fixed-wing aircraft, by mounting to it the phenotype acquisition kit, placed in a cargo pod or directly installed on the step (via a bracket), with a passenger(s) evaluating the images and giving feedback on their quality to the pilot, in real time, *via* a video monitor in the cockpit. The phenotypic equipment used for this entail sensors, GPS unit, gyroscopes and/or inertial measurement units. Although the MAP is capable of carrying multiple sensors for complicated phenotypic tasks, this has not been fully exploited, likely because it would depend on much manual participation. For example, just one radiometric calibration thermal camera in the cargo pod of a helicopter was employed for collecting canopy temperature (Deery et al., 2016)—and a passenger must be present to perform the image assessment and provide feedback. For this, an RGB camera can be remotely controlled and triggered by an operator, to watch the "real situation" through a video monitor in the cockpit (Yang and Hoffmann, 2015). For estimating the aboveground biomass (AGB) and underground biomass (BGB) of maize, LiDAR was also installed on a MAP, to collect point-cloud data at the nominal height of 1,500 m (Li et al., 2015). High imaging heights and airspeeds certainly make the acquisition of phenotypic data at a high resolution and accuracy more challenging.

Similarly, MAP can also make effective use of pre-existing agricultural machinery, and can overcome challenging weather conditions to a certain extent because of flight stability. Nevertheless, conducting phenotypic experiments with MAP demands a specific amount of manpower and inevitably involves high costs. For example, MAP requires a trained pilot with relevant qualifications to operate the helicopter or plane, and a passenger on board doing the monitoring, assessing, communicating, and manual imaging. Moreover, considering the cost and technical issues, the advantages of MAP have not been fully exploited (i.e., flight altitude, flight speed, carrying capacity). Perhaps that's why MAP's prominence has dropped sharply in crop phenotyping, precisely because of UAP's unparalleled advantages, which enable it to complete the consistent phenotypic tasks and thus progressively supplant MAP.

Satellite Platform

Satellites can provide panchromatic imagery, multispectral imagery, or radio detection and ranging (RADAR) data. Panchromatic images of a single-band are displayed as gray-scale images with high resolution but limited spectral information, whereas multispectral images have a rich spectrum yet with relatively low resolution. Thus, these obtained panchromatic and multispectral images are usually merged by panchromatic sharpening or pan-sharpening, to obtain multispectral raster data having a high resolution. However, such optical satellite phenotyping is susceptible to uncooperative weather conditions, such as cloudy, rain, fog, and haze, and it also suffers from visible light saturation (Jin et al., 2015). In this case, RADAR and synthetic aperture radar (SAR) data are able compensate for this defect extremely well. This is due to the unique sensitivity of crop structure to microwaves, which effectively improves the availability of the satellite platform. WorldView series, RapidEye, GeoEye-1, SPOT series, QuickBird, Ikonos, Planet Scope, Pleiades series, KOMPSAT series, Satellite for Earth Observation series, Landsat series, Gaofen series, and SkySat series are currently the main satellite HT3Ps that can obtain color and multispectral images (Jin et al., 2020). Those satellite platforms able to obtain RADAR data are mainly Sentinel-1, RADARSAT-2, ENVISAT, TerraSAR-X/TanDEM-X, and RISAT-2 (Zhang et al., 2020).

In recent years, with the rapid development of satellite HT3Ps, there is increasingly more research on plant phenotyping done using satellite data. For example, multispectral data of GeoEye-1 satellite was used to evaluate the nitrogen status and spatial variability of different species of turfgrass (Caturegli et al., 2015), leading to an important guiding principle for turf fertilization management. The RapidEye satellite provides five bands with a spatial resolution of 5 m: blue, green, red, red-edge, and near infrared. Its multispectral images were used to study the variable rate of wheat nitrogen fertilizer effects (Basso et al., 2016). In evaluating the crop angle of indentation (CAI), Chauhan et al. (2020) relied on Sentinel-1 and radarsat-2 (multi-incidence angle) data and went on to evaluate the severity of crop lodging. An interesting satellite platform, FLuorescence EXplorer (FLEX), is equipped with a single payload fluorescence imaging spectrometer (FLORIS) that combines spectral and spatial resolution to retrieve and interpret the full chlorophyll fluorescence spectra emitted by terrestrial vegetation (Drusch et al., 2017). It is scheduled to launch in 2022 and will fly in the same orbit as Sentinel-3; hence, the availability and interoperability of auxiliary information.

The ability to process large-scale satellite phenotypic data at low cost according to international standard protocols is a key advantage of satellite phenotyping. Some satellites provide free data to would-be users, but the acquisition of high-precision commercial satellite data normally has a monetary charge. Fortunately, the cost of accessing to satellite data is now modest. Still, both panchromatic and multispectral imaging done by satellites remains susceptible to interference of atmospheric, clouds, and fog. In addition, the resolution issue—such as Pleiades-1a, 0.5 m; SPOT 6, 1.5 m; Planet Scope 3.0 m; Rapid Eye, 5.0 m, to name a few—and the data revisit period are also limiting factors. Fortunately, satellite resolution is making continuous progress as satellite technology advances. For instance, the Finnish technology start-up ICEYE has released commercial spaceborne SAR samples with the highest resolution (0.25 m) currently available worldwide. In the near future, low-orbiting nanosatellites and microsatellites with high spatiotemporal resolution may become join the prevailing aeronautical HT3Ps.

HT3PS' COMBINATION FOR COMPARATIVE VALIDATION OR COMPREHENSIVE ANALYSIS

A typical comprehensive HT3P consists of four components: sensor, platform, analysis, and visualization. Together, these should be able to perform high-throughput and non-destructive acquisition of a substantial amount of data on dynamic phenotypic traits of cultivated plants and their environmental characteristics, as well as providing multi-omics analyses, completing the entire phenotyping process from a holistic perspective, thereby ultimately facilitating crop improvement and molecular breeding of plants (as shown in **Figure 4**). Since the various types of HT3Ps each have their own unique merits—**Table 4** summaries the advantages and disadvantages of miscellaneous HT3Ps—with traditional manual measurements usually needed to take relative real data for comparison and verification, pursuing combinations of differing HT3Ps offers a way to break away from traditional phenotyping and obtain comprehensive high-precision phenotypic data. Several common and typical combinations of diverse HT3Ps are highlighted in **Table 5**.

HT3Ps' Combination for Comparative Validation

Some combinations of HT3Ps seek to compare and validate the phenotypic performance by applying different types of HT3Ps and thereby bypass traditional manual measurements entirely. For instance, Andrade-Sanchez et al. (2014) modified a sprayer to function as a mobile platform, to simultaneously collect canopy height, reflectance, and temperature in four adjacent rows in a cotton field; it can carry four sets of sensors, including sonar sensors, IR radiometers, and multispectral canopy sensors. Then, to verify the data authenticity of the mobile platform, a MAP (helicopter) was applied to collect visible-near-infrared (VNIR) and thermal IR data. In other work, to evaluate the plant height of wheat, a mobile platform with LiDAR and a UAP with RGB cameras were jointly used to generate 3D dense-point clouds (Madec et al., 2017). However, because of the low resolution of RGB images from UAP and the strong

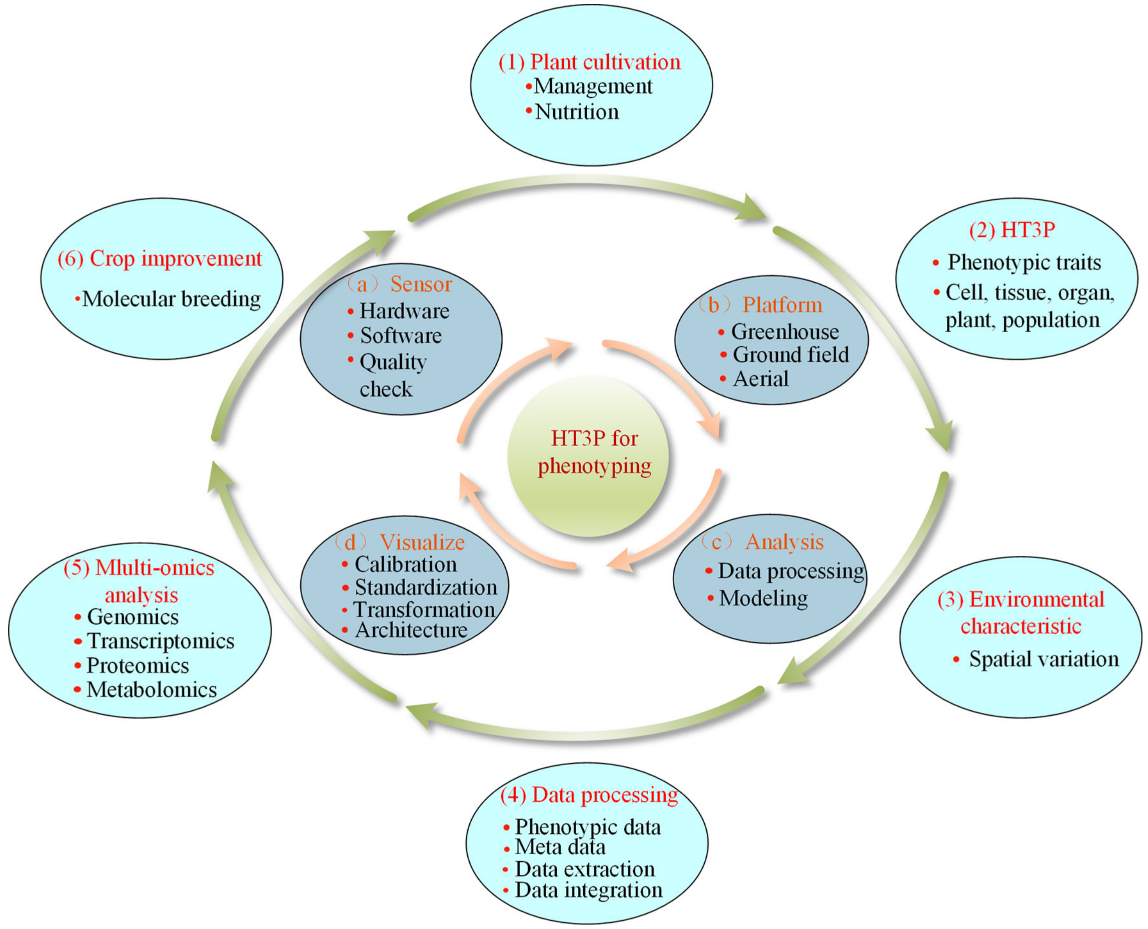


FIGURE 4 | Typical comprehensive HT3P components: (a) sensor, (b) platform, (c) analysis, (d) visualization; HT3P performs the entire workflow of phenotyping: (1) cultivation of plants, (2) extraction of phenotypic traits, (3) acquisition of environmental parameters, (4) data processing, (5) multi-omics analysis, and, ultimately, (6) aiding in crop improvement.

penetrability of LiDAR of the mobile platform, plant height measurement based on UAP was underestimated. Similarly, Khan et al. (2018) combined a mobile platform and a UAP to collect RGB images, which revealed that canopy height obtained from the mobile platform is more accurate than UAP, whereas the plant vigor evaluated from the UAP is more precise. A multi-rotor UAP, with an RGB camera and a fixed-wing UAP with a hyperspectral push-broom scanner, was devised by Habib et al. (2017) to verify the feasibility of using RGB-based orthophotos to improve the geometric features of hyperspectral orthophotos. In addition, the combination of a UAP and four satellites was implemented to compare the phenotypic capabilities of different resolutions in dry bean (Sankaran et al., 2019), whose results indicated that using sub-meter resolution satellites as HT3Ps holds promising application prospects for field crop phenotyping.

TABLE 4 | Overview of the comparison of advantages and disadvantages of the different types of HT3Ps.

HT3P type	Advantages	Disadvantages
Benchtop-type	Strong repeatability; continuous monitoring; precision phenotyping; high resolution	Expensive; for small plants only
Conveyor-type	Large sample size; strong repeatability; high resolution	Expensive; high operating costs
Pole/tower-based	Low cost; relatively simple structure; flexible movement	Small range; increased distance decreases resolution
Mobile	Semi-automatic or fully automatic; high resolution; expansive; high flexibility.	Affected by weather, soil conditions; soil compaction; mechanical interference; requires some manpower; long boom may cause sensor jittering and blurred images; safety mechanism needed
Gantry-based	Low weather dependency; continuous phenotyping all-day	Expensive; fixed limited area; high maintenance costs
Cable-suspended	Low weather dependency	Expensive; fixed limited area; limited endurance
UAP	Flexible flight plan; coverage a wide range of field plots; relatively low cost; GPS navigation;	Weather (light, rain, fog, etc.) dependence; limited payload; limited endurance; strict aviation regulation (altitude); flight training
MAP	Flexible payload; rapid coverage of large areas	Expensive; non-repeatable flight route; substantial manpower
Satellite	Coverage a wide range of field plots; relatively low cost	Low resolution; long return period; weather restrictions (except radar)

While some combinations of different types HT3Ps are still based on the time-consuming and laborious traditional field measurements, these will gradually disappear with the stabilization and improvement of the advanced HT3Ps. The combination of various types of HT3Ps for cross-validation is gradually moving forward, representing a landmark step in the field of plant phenotyping.

HT3Ps' Combination for Comprehensive Analysis

Some HT3Ps' performance aspects are combined to realize the fusion of multi-source data for collaborative and comprehensive phenotyping. For example, using both a mobile platform and a tower-based platform for canopy scale and single plant phenotyping has been proposed by Shafiekhani et al. (2017). As an autonomous mobile platform, Vinobot, with its stereo cameras installed on the robotic arm, can autonomously navigate in the field and collect data on individual plant traits. The tower-based Vinoculer can inspect a large-area canopy phenotype, and delegate specific regions to Vinobot for elaborate phenotyping, which greatly improves the flexibility and purposiveness. A robotic mobile platform and a UAP collected complementary multispectral data, which let investigators obtain comprehensive crop phenotypes (Ingunn et al., 2017). Specifically, that mobile platform provided detailed traits information of plants and the UAP obtained calibrated NDVI, and together they further predicted the heading date and yield. One multi-rotor UAP

TABLE 5 | Overview of typical applications of HT3Ps when used in combination.

Combination type (a+b)	Sensors^a	Sensors^b	Plants	Traits	References
Mobile + pole/tower-based	Stereo camera	RGB, infrared	Maize, sorghum	Plant height, leaf area index (LAI)	Shafiekhani et al., 2017
UAP + mobile	RGB	LiDAR	Wheat	Plant height	Madec et al., 2017
MAP + mobile	Monochromatic, thermal	Sonar, IR radiometer, multispectral	Cotton	Canopy height, canopy temperature	Andrade-Sanchez et al., 2014
MAP + pole/tower-based	Thermal IR	IR thermometer	Wheat	Canopy temperature	Deery et al., 2019
UAP + UAP	RGB	Hyperspectral	Crop	–	Habib et al., 2017
UAP + satellite	RGB, thermal, multispectral	Multispectral (VNIR, SWIR)	Soybean	Mean canopy temperature, water stress resistance, VIs	Sagan et al., 2019b

with an RGB camera, and another with a multispectral camera, were combined to monitor tomato crops, so as to formulate management measures and determine the best management scheme for specific fields (Marconi et al., 2019). Likewise, Anderson et al. (2020) employed a rotary-wing UAP and a fixed-wing UAP to monitor a track of plant height growth of an recombinant inbred line (RIL) maize population; hence, they could, for the first time, elucidate dynamic characteristics of quantitative trait loci (QTL) in real time under the field conditions. In addition, Deery et al. (2019) designed a MAP with airborne thermal IR cameras and a pole-based platform (Arducrop wireless IR thermometers) and used it to continuously measure crop CT (canopy temperature). To fill the temporal gap in the availability of satellite data and improve the usability of UAP, Sagan et al. (2019b) merged RGB, thermal, and multi-spectral images from UAP with VNIR and SWIR imagery taken by the WorldView-3 satellite, applying them to crop monitoring and early stress detection on a temporal scale, which contributed to form field-scale coordinated data of UAV and satellite virtual constellation.

The imaging distance of various HT3Ps will engender a differing spatial resolution and pixel size. Taking ground coverage (GC) as an example, the pixel size of an RGB image has a huge impact on the accuracy of GC's evaluation (Hu et al., 2019). In addition, there are significant disparities among different types of HT3Ps, such as their time resolution, weather dependence, experimental scale, and financial investment, to name a few. This means that combinations of HT3Ps ought to steer toward actual phenotypic requirements and concrete practical issues. Along with further refinement of plant phenotyping, the future HT3P portfolio is expected to integrate multi-site distributed platforms, single-point centralized platforms, and cloud-based platforms, to deeply mine and dissect multi-source phenotypic data from dynamic time series at multiple scales, for multiple species, and under multiple scenarios. Furthermore, to effectively integrate existing HT3Ps, phenotyping technology, phenotypic methods, data availability, and resources, to speed up the emergence of high-quality phenotypic achievements and accelerate crop breeding, while also reducing duplication of research and investment, more international and regional organizations, or initiatives (see **Table 6**) have come into being.

SIMULATION HT3P

The simulated HT3P aims to model plant growth, phenotypic expression, and phenotyping at various scales. It does this by integrating multi-source information in a modeling framework, such as that of germplasm resources, irrigation, fertilization, nutritional substance, spatial climate, soil environment, terrain properties and management records. For example, a digital plant phenotyping platform (D3P) would use environmental variables, crop management, and meteorological information as input, to generate 3D virtual canopy structure. Based on this, the collection of virtual canopy phenotypic traits can be performed by RGB, multi-spectral, and LiDAR simulators (Liu et al., 2019). For whole forests with large cover areas, long-lived cycles and high heterogeneity, Dungey et al. (2018) provided a prototype of a landscape-scale HT3P simulator, by consolidating remote sensing topography, environmental impacts, spatial abiotic information, management records and genomics into the modeling framework. It aimed to eliminate some traditional limitations in tree breeding programs and provide genetic gains in tree fitness.

By combining genomics, high-throughput phenotyping, and simulation modeling, we can obtain an adequate but sound understanding of phenotypic traits and their variation (Varshney et al., 2018). The application of various complex models to combine simulations with empirical methods will contribute markedly to accelerating the process of extracting ideal phenotypic traits for use in crop improvement.

FUTURE PROSPECTS FOR HT3P

The concept of HT3P is rather grand, such that the development and innovation of HT3Ps depends on the cross-disciplinary cooperation of agronomy, robotics, computer, automation, artificial intelligence, and big data, requiring the participation of experts—breeders, agronomists, plant scientists, mechanical engineers—and leadership from interdisciplinary talent of open innovation teams. Whether HT3P is phenotyping in the greenhouse or in the field, ground-based proximal phenotyping or aerial large-scale remote sensing, the future of HT3Ps lies in improving spatial-temporal resolution, sensor integration, turnaround time in data analysis, human-machine interaction, operational stability, throughput, automation, operability, and accessibility.

It is worth noting that the development, selection, and utilization of HT3Ps should be orientated by concrete project requirements, specific phenotypic tasks, and practical application scenarios, such as the field coverage (Kim, 2020), rather than assuming that more devices, technologies, and funds with which the HT3P is equipped, the better; partly because the collection of a large amount of data does not mean all of it is useful (Haagsma et al., 2020). Even in some cases, the experimental effects of applying single and multiple sensors are identical (Meacham-Hensold et al., 2020), and the data obtained from multiple devices are redundant. However, the combinations of various HT3Ps for comparative validation and comprehensive analysis could provide broad application prospects for inspection, extraction, and quantification of complex physiological functional phenotypes. Yet the technical issues of formulating standards and synchronizing calibrations for these multiple combinations are daunting tasks. Fortunately, the involvement of meta-analysis ensures the objectivity of HT3P development and selection. For example, Young (2019) applied meta-analysis method to develop an evaluation framework that can quantitatively and objectively assess the complexity and utility scores of high-throughput systems. As an effective analytical method of quantitative, scientific synthesis of research results (Gurevitch et al., 2018), meta-analysis may prove especially fruitful in the near future.

TABLE 6 | Overview of international organizations or regional initiatives contributing to plant phenotyping.

Organization acronym	Full name (or description)	URL
APPF	Australia Plant Phenomics Facility	https://www.plantphenomics.org.au/
APPN	Austrian Plant Phenotyping Network	https://appn.at/
CGIAR	Modernize breeding programs targeting the developing world	http://excellenceinbreeding.org/
CIMMYT	International Wheat and Maize Improvement Center	https://www.cimmyt.org/
CPPN	China Plant Phenotyping Network	–
CSISA	The Cereal Systems Initiative for South Asia	https://csisa.org/
DPPN	German Plant Phenotyping Network	https://dppn.plant-phenotyping-network.de/
EMPHASIS	European Plant Phenotyping Infrastructure	https://emphasis.plant-phenotyping.eu/
EPPN2020	European Plant Phenotyping Network (2020)	https://eppn2020.plant-phenotyping.eu
ESFRI	European Strategy Forum for Research Infrastructure	https://www.esfri.eu/
FPPN/PHENOME	French Plant Phenomic Network	https://www6.dijon.inrae.fr/umragroecologie_eng/Research-Programs/Investissement-Avenir/PHENOME
LatPPN	Latin American Plant Phenomics Network	–
LEPSE	Laboratory of Plant Ecophysiological Responses to Environmental Stresses	http://www1.montpellier.inra.fr/ibip/lepse/english/
NAPPN	The North American Plant Phenotyping Network	http://nappn.plant-phenotyping.org/
NPPN	Nordic Plant Phenotyping Network	https://nordicphenotyping.org/
NPEC	Netherlands Plant Eco-phenotyping Centre	https://www.wur.nl/en/product/TheNetherlands-Plant-Eco-phenotypingCentre-NPEC.htm
G2F	The Genomes to Fields Initiative	https://www.genomes2fields.org/
GCN	Green Crop Network	http://www.greencropnetwork.com/
IPPN	International Plant Phenotyping Network	https://www.plant-phenotyping.org/
JPPC	The Jülich Plant Phenotyping Centre	http://www.fz-juelich.de/ibg/ibg-2/EN/Research/Phenotyping/Phenotyping_article.html?nn=548814
MIAPPE	Minimum Information About a Plant Phenotyping Experiment	https://www.miappe.org/
PHEN-ITALY	Italian Plant Phenotyping Network	http://www.phen-italy.it/index.php
PhenomUK	Promotes an integrated, holistic view of the phenotyping process across the UK	https://www.phenomuk.net/
TERRAREF	Terraphenotyping Reference Platform	https://www.terraref.org/
Wheat Initiative	Endorsed by the G20 Agricultural Ministers, to contribute to improving world food security	https://www.wheatinitiative.org/our-vision

Specifically, the future conveyor-type HT3P requires consideration of operational stability and environmental homogeneity, and allowing phenotypic analysis for multi-level subtle traits of a wide variety of representative plants will be a key design factor to the development of the future benchtop-type HT3P. Scalability, rotatability and multi-site deployment will be the prospective features of pole/tower-based HT3Ps, and economically-efficient distributed ones will perform outstandingly in the calibration of high-dimensional phenotypic data. Mobile HT3Ps that can be transported to the experimental site are preferred by phenotypic researchers rather than the experiment coming to the limited platform (Roitsch et al., 2019). This means that the development of mobile HT3P needs to move toward flexibility and portability, and that modular and customizable design will be welcomed by the phenotyping community. The reduction of volume and cost is the major consideration for future gantry-based HT3P designs, and the new cable-suspended HT3P will have the ability to monitor continuously and consistently crop growth and development at low altitudes over long periods of time. As for UAP, the development of compact lightweight sensor configuration that is sensitive to plant-specific phenotypic traits will be a breakout (Xie and Yang, 2020). In addition, advanced battery technology is in dire need of a stage breakthrough, which can greatly improve the endurance, payload, and power of the UAV. For satellite phenotyping, improving image resolution and shortening the revisit cycle remain the focus of satellite platform development. Additionally, cost-effective platforms also warrant consideration, as smart phone, handheld portable instrument, backpack system, and wearable device are adopted and updated for utilization in phenotyping.

High-throughput data acquisition, data management, data interpretation, modeling, integration, and application together form the core and pillar of plant phenotyping. The main challenges faced by the new generation of phenotyping are data handling, images processing, and traits analyzing (Fahlgren et al., 2015a; Campbell et al., 2018; Hickey et al., 2019). Fortunately, the introduction of various software, web-based tools, pipelines, toolkits, deep learning tools, and online repository solutions to assist phenotypic researchers in processing phenotypic data will break through these technical bottlenecks. For example, a free multi-purpose software—Coverage Tool—can semi-automatically quantify a wide range of visual plant traits (Merchuk-Ovnat et al., 2019). Web-based image analysis tools, such as Field Phenomics (Guzman et al., 2015), are considered by us to be a hotspot for phenotypic solutions. Kar et al. (2020) developed an analysis pipeline with outlier detection, missing value imputation, and spatial adjustment for solving the problem of inaccurate and missing phenotypic data. Toolkits tend to be relatively specific, such as Plant 3D (P3D), which specializes in analyzing 3D point cloud data of plant structures (Ziamtsov and Navlakha, 2020). With the advancement of HT3P, their improving high-throughput and efficiency will produce increasingly big data. For huge datasets, deep learning tools are needed; however, only when large datasets that capture shared problems become available can the greatest benefit be gained from the application of deep learning tools. Online databases, such as http://www.plant-image-analysis.org, can effectively bridge the gap between developers and users, but still lack comprehensive management platforms that cover software, web-based tools, pipelines, toolkits, deep learning tools, and other phenotypic solutions, which will be a milestone breakthrough as well as a considerable challenge.

With the emergence of various HT3Ps, experimental designs, phenotyping methods, standardized management, both phenotype acquisition and its data analysis are becoming extremely prominent. Phenotypic data that costs substantial capital, labor, time, and energy, however, may 1 day be abandoned forever (Mir et al., 2019). Presently, a standard phenotyping agreement or data analysis methodology for plant phenotyping has yet to be established (Mahlein et al., 2019). The standardization of data and metadata from the HT3Ps contributes to an improved data utilization rate and it ensures the interoperability of data providers and experimental replication. Otherwise, data that is poor annotated and in a disorderly format may generate noise or disordered waves. Fortunately, relevant standard constraints are being proposed. For example, Krajewski et al. (2015) published a technical paper offering effective recommendations (at http://cropnet.pl/phenotypes) and initiatives (such as http://wheatis.org), making a further step toward establishing internationally practical solutions. Moreover, originating the relevant standardization of phenotyping can strengthen the comprehension and explanation of biological phenomena, contributing to the transformation of biological knowledge and establishment of a real coherent semantic network.

CONCLUSIONS

HT3P is a novel and powerful tool for obtaining plant-deep phenotypes (morphological structure, physiological function, component content) and dense phenotypes in complex field setting, which cannot be accomplished by traditional phenotyping approaches. This paper reviewed the application of HT3Ps in the growth chamber or greenhouse with strictly controlled environmental conditions and field phenotyping with notoriously heterogeneous conditions and uncontrollable environmental factors. Then, according to platform configuration and operation mode, further classifications were performed to provide comprehensive overview and description and assessment of the various types of HT3Ps currently available. The unique characteristics, applications, and strengths and weaknesses of various HT3Ps were emphasized. Going further, the simulation platform, various combinations of HT3Ps for comparative validation or comprehensive analysis, current phenotypic challenges, and the future development trends of HT3Ps were discussed.

With the assistance of powerful HT3Ps, phenomics has arguably entered a new stage (Tardieu et al., 2017). At this stage, the new and pressing challenge of next generation phenotyping will be to reasonably combine phenotypic experiments, various HT3Ps, models, data processing and handling scheme, meta-analysis, and visualization of phenotypic information for optimizing the allocation of research resources, efficiently accomplishing complex phenotypic tasks, and transforming massive multi-source phenotypic data into statistical and biological knowledge. Robust phenotyping is central to plant breeding (Hickey et al., 2019), and the development of satisfying crop varieties with high-yielding and strong stress resistance is the ultimate goal of crop breeding. High-throughput sequencing activity underpins the fast development of genomics (Shah et al., 2018). Likewise, HT3P as a novel and powerful phenotyping tool will explore a new period of rapid development in phenomics. Further, combining morphological, physiological, and elemental phenotyping with multi-omics methods from the perspective of holistic omics will usher in a new era of botany phenotyping.

AUTHOR CONTRIBUTIONS

DL and CL guided the writing of the article and reviewed the initial versions of it. CQ completed the manuscript's writing content and graphic production. ZS and GY gave comments on the framework and diagrams of the article. XL was involved in both language correction and graphic modification of the article. AM assisted in improving word accuracy and language expressiveness. All authors have made substantial contributions to the conception, drafting of the manuscript, read, and approved the submitted version of the article.

FUNDING

This work was supported by Shandong Key R&D Program (Grant no. 2019JZZY010703) and the Sino-British Cooperation Project (Grant no. 2017YFE0122100-1).

ACKNOWLEDGMENTS

The authors would like to thank the editors and reviewers for their valuable input, time, and suggestions for improving the overall quality of the manuscript. We would also like to thank the Charlesworth Group (https://www.cwauthors.com/frontiers/) for English language editing.

REFERENCES

Acosta-Gamboa, L. M., Liu, S., Langley, E., Campbell, Z., Castro-Guerrero, N., Mendoza-Cozatl, D., et al. (2017). Moderate to severe water limitation differentially affects the phenome and ionome of *Arabidopsis*. *Func. Plant Biol.* 44, 94–106. doi: 10.1071/FP16172

Anderson, S. L., Murray, S. C., Chen, Y., Malambo, L., and Jung, J. (2020). Unoccupied aerial system enabled functional modeling of maize height reveals dynamic expression of loci. *Plant Direct* 4:e00223. doi: 10.1002/pld3.223

Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., et al. (2014). Development and evaluation of a field-based high-throughput phenotyping platform. *Func. Plant Biol.* 41, 68–79. doi: 10.1071/FP13126

Atefi, A., Ge, Y., Pitla, S., and Schnable, J. (2019). *In vivo* human-like robotic phenotyping of leaf traits in maize and sorghum in greenhouse. *Comp. Electron. Agric.* 163:104854. doi: 10.1016/j.compag.2019.104854

Atieno, J., Li, Y., Langridge, P., Dowling, K., Brien, C., Berger, B., et al. (2017). Exploring genetic variation for salinity tolerance in chickpea using image-based phenotyping. *Sci. Rep.* 7:1300. doi: 10.1038/s41598-017-01211-7

Bai, G., Ge, Y., Hussain, W., Baenziger, P. S., and Graef, G. (2016). A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. *Comp. Electron. Agric.* 128, 181–192. doi: 10.1016/j.compag.2016.08.021

Bai, G., Ge, Y., Scoby, D., Leavitt, B., Stoerger, V., Kirchgessner, N., et al. (2019). NU-Spidercam: a large-scale, cable-driven, integrated sensing and robotic system for advanced phenotyping, remote sensing, and agronomic research. *Comp. Electron. Agric.* 160, 71–81. doi: 10.1016/j.compag.2019.03.009

Bai, G., Jenkins, S., Yuan, W., Graef, G. L., and Ge, Y. (2018). Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. *Front. Plant Sci.* 9:1002. doi: 10.3389/fpls.2018.01002

Bangert, W., Kielhorn, A., Rahe, F., Dreyer, A.-W., Kg, G., Co, et al. (2013). "Field-Robot-Based Agriculture: "RemoteFarming.1" and "BoniRob-Apps"," in *Tagung Land. TECHNIK - AgEng.*

Bao, Y., Nakami, A. D., and Tang, L. (2014). "Development of a field robotic phenotyping system for sorghum biomass yield component traits characterization," in *American Society of Agricultural and Biological Engineers Annual International Meeting 2014, ASABE.*

Basso, B., Fiorentino, C., Cammarano, D., and Schulthess, U. (2016). Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precis. Agric.* 17, 168–182. doi: 10.1007/s11119-015-9414-9

Beauchene, K., Leroy, F., Fournier, A., Huet, C., Bonnefoy, M., Lorgeou, J., et al. (2019). Management and characterization of abiotic stress via phenofield((R)), a high-throughput field phenotyping platform. *Front. Plant Sci.* 10:904. doi: 10.3389/fpls.2019.00904

Brichet, N., Fournier, C., Turc, O., Strauss, O., Artzet, S., Pradal, C., et al. (2017). A robot-assisted imaging pipeline for tracking the growths of maize ear and silks in a high-throughput phenotyping platform. *Plant Methods* 13:96. doi: 10.1186/s13007-017-0246-7

Brien, C., Jewell, N., Watts-Williams, S. J., Garnett, T., and Berger, B. (2020). Smoothing and extraction of traits in the growth analysis of noninvasive phenotypic data. *Plant Methods* 16:s36. doi: 10.1186/s13007-020-00577-6

Bruning, B., Liu, H., Brien, C., Berger, B., Lewis, M., and Garnett, T. (2019). The development of hyperspectral distribution maps to predict the content and distribution of nitrogen and water in wheat (*Triticum aestivum*). *Front. Plant Sci.* 10:1380. doi: 10.3389/fpls.2019.01380

Buchaillot, M. L., Gracia-Romero, A., Vergara-Diaz, O., Zaman-Allah, M. A., Tarekegne, A., Cairns, J. E., et al. (2019). Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. *Sensors* 19:1815. doi: 10.3390/s19081815

Campbell, Z. C., Acosta-Gamboa, L. M., Nepal, N., and Lorence, A. (2018). Engineering plants for tomorrow: how high-throughput phenotyping is contributing to the development of better crops. *Phytochem. Rev.* 17, 1329–1343. doi: 10.1007/s11101-018-9585-x

Caturegli, L., Casucci, M., Lulli, F., Grossi, N., Gaetani, M., Magni, S., et al. (2015). GeoEye-1 satellite versus ground-based multispectral data for estimating nitrogen status of turfgrasses. *Int. J. Remote Sens.* 36, 2238–2251. doi: 10.1080/01431161.2015.1035409

Chang, S., Lee, U., Hong, M. J., Jo, Y. D., and Kim, J.-B. (2020). High-throughput phenotyping (HTP) data reveal dosage effect at growth stages in *Arabidopsis thaliana* irradiated by gamma rays. *Plants* 9:557. doi: 10.3390/plants9050557

Chapman, S. C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M. F., et al. (2014). Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Chauhan, S., Darvishzadeh, R., Boschetti, M., and Nelson, A. (2020). Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sens. Environ.* 236:111488. doi: 10.1016/j.rse.2019.111488

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., et al. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26, 4636–4655. doi: 10.1105/tpc.114.129601

Choudhury, S. D., Bashyam, S., Qiu, Y., Samal, A., and Awada, T. (2018). Holistic and component plant phenotyping using temporal image sequence. *Plant Methods* 14:35. doi: 10.1186/s13007-018-0303-x

Choudhury, S. D., Samal, A., and Awada, T. (2019). Leveraging image analysis for high-throughput plant phenotyping. *Front. Plant Sci.* 10:508. doi: 10.3389/fpls.2019.00508

Choudhury, S. D., Stoerger, V., Samal, A., Schnable, J. C., Liang, Z., and Yu, J.-G. (2016). "Automated vegetative stage phenotyping analysis of maize plants using visible light images," in *KDD: Data Science for Food, Energy and Water.*

Crain, J. L., Wei, Y., Barker, J., Thompson, S. M., Alderman, P. D., Reynolds, M., et al. (2016). Development and deployment of a portable field phenotyping platform. *Crop Sci.* 56, 965–975. doi: 10.2135/cropsci2015.05.0290

Czedik-Eysenberg, A., Seitner, S., Güldener, U., Koemeda, S., Jez, J., Colombini, M., et al. (2018). The 'PhenoBox', a flexible, automated, open-source plant phenotyping solution. *New Phytol.* 219, 808–823. doi: 10.1111/nph.15129

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., and Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 4, 349–379. doi: 10.3390/agronomy4030349

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., Bovill, W. D., James, R. A., Condon, A. G., et al. (2019). Evaluation of the phenotypic repeatability of canopy temperature in wheat using continuous-terrestrial and airborne measurements. *Front. Plant Sci.* 10:875. doi: 10.3389/fpls.2019.00875

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., et al. (2016). Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography. *Front. Plant Sci.* 7:1808. doi: 10.3389/fpls.2016.01808

Díaz-Varela, R. A., De La Rosa, R., León, L., and Zarco-Tejada, P. J. (2015). High-resolution airborne UAV imagery to assess olive tree crown parameters using 3D photo reconstruction: application in breeding trials. *Remote Sens.* 7, 4213–4232. doi: 10.3390/rs70404213

Drusch, M., Moreno, J., Del Bello, U., Franco, R., Goulas, Y., Huth, A., et al. (2017). The fluorescence EXplorer mission concept—ESA's earth explorer 8. *IEEE Trans. Geosci. Remote Sens.* 55, 1273–1284. doi: 10.1109/TGRS.2016.2621820

Duan, L., Han, J., Guo, Z., Tu, H., Yang, P., Zhang, D., et al. (2018). Novel digital features discriminate between drought resistant and drought sensitive rice under controlled and field conditions. *Front. Plant Sci.* 9:492. doi: 10.3389/fpls.2018.00492

Duan, T., Zheng, B., Guo, W., Ninomiya, S., Guo, Y., and Chapman, S. C. (2017). Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by UAV. *Func. Plant Biol.* 44, 169–183. doi: 10.1071/FP16123

Dungey, H. S., Dash, J. P., Pont, D., Clinton, P. W., Watt, M. S., and Telfer, E. J. (2018). Phenotyping whole forests will help to track genetic performance. *Trends Plant Sci.* 23, 854–864. doi: 10.1016/j.tplants.2018.08.005

Fahlgren, N., Feldman, M., Gehan, M. A., Wilson, M. S., Shyu, C., Bryant, D. W., et al. (2015a). A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in setaria. *Mol. Plant* 8, 1520–1535. doi: 10.1016/j.molp.2015.06.005

Fahlgren, N., Gehan, M. A., and Baxter, I. (2015b). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* 24, 93–99. doi: 10.1016/j.pbi.2015.02.006

Fawcett, D., Panigada, C., Tagliabue, G., Boschetti, M., Celesti, M., Evdokimov, A., et al. (2020). Multi-scale evaluation of drone-based multispectral surface reflectance and vegetation indices in operational conditions. *Remote Sens.* 12:514. doi: 10.3390/rs12030514

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A., and Zhang, M. (2020). Yield estimation in cotton using UAV-based multi-sensor imagery. *Biosyst. Eng.* 193, 101–114. doi: 10.1016/j.biosystemseng.2020.02.014

Flood, P. J., Kruijer, W., Schnabel, S. K., Van Der Schoor, R., Jalink, H., Snel, J. F., et al. (2016). Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. *Plant Methods* 12:14. doi: 10.1186/s13007-016-0113-y

Fraas, S., and Lüthen, H. (2015). Novel imaging-based phenotyping strategies for dissecting crosstalk in plant development. *J. Exp. Bot.* 66, 4947–4955. doi: 10.1093/jxb/erv265

Friedli, M., Kirchgessner, N., Grieder, C., Liebisch, F., Mannale, M., and Walter, A. (2016). Terrestrial 3D laser scanning to track the increase in canopy height of both monocot and dicot crop species under field conditions. *Plant Methods* 12:9. doi: 10.1186/s13007-016-0109-7

Furbank, R. T., Jimenez-Berni, J. A., George-Jaeggli, B., Potgieter, A. B., and Deery, D. M. (2019). Field crop phenomics: enabling breeding for radiation use efficiency and biomass in cereal crops. *New Phytol.* 223, 1714–1727. doi: 10.1111/nph.15817

Ge, Y., Bai, G., Stoerger, V., and Schnable, J. C. (2016). Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. *Comp. Electron. Agric.* 127, 625–632. doi: 10.1016/j.compag.2016.07.028

Gomez, F. E., Carvalho, G. Jr., Shi, F., Muliana, A. H., and Rooney, W. L. (2018). High throughput phenotyping of morpho-anatomical stem properties using X-ray computed tomography in sorghum. *Plant Methods* 14:59. doi: 10.1186/s13007-018-0326-3

Gracia-Romero, A., Kefauver, S. C., Fernandez-Gallego, J. A., Vergara-Díaz, O., Nieto-Taladriz, M. T., and Araus, J. L. (2019). UAV and ground image-based phenotyping: a proof of concept with durum wheat. *Remote Sens.* 11:1244. doi: 10.3390/rs11101244

Granier, C., and Vile, D. (2014). Phenotyping and beyond: modelling the relationships between traits. *Curr. Opin. Plant Biol.* 18, 96–102. doi: 10.1016/j.pbi.2014.02.009

Guimarães, C. M., Stone, L. F., Brito, G. G. D., and Heuert, J. (2017). Evaluation of water-stress tolerance of Acala SJ 2 and Auburn 2 cotton cultivars in a phenotyping platform. *Rev. Ambient. Água* 12, 629–642. doi: 10.4136/ambi-agua.2105

Guo, Q., Wu, F., Pang, S., Zhao, X., Chen, L., Liu, J., et al. (2016). Crop 3D: a platform based on LiDAR for 3D high-throughputcrop phenotyping. *Sci. Sinica Vitae* 46, 1210–1221. doi: 10.1360/N052016-00009

Gurevitch, J., Koricheva, J., Nakagawa, S., and Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature* 55, 175–182. doi: 10.1038/nature25753

Guzman, H. A. R., Gonzalez-Navarro, F. F., Selvaraj-Gomez, M., Valencia, M., and Delgado, A. (2015). "Field phenomics: a web based image analysis platform using open source tools," in *2015 International Conference on Computational Science and Computational Intelligence* (CSCI).

Haagsma, M., Page, G., Johnson, J., Still, C., Waring, K., Sniezko, R., et al. (2020). "Is more data better? A comparison of multi-and hyperspectral imaging in phenotyping," in *EGU General Assembly Conference Abstracts.*

Habib, A., Xiong, W., He, F., Yang, H. L., and Crawford, M. (2017). Improving orthorectification of UAV-based push-broom scanner imagery using derived orthophotos from frame cameras. *IEEE J. Select. Topics Appli. Earth Observ. Remote Sens.* 10, 262–276. doi: 10.1109/JSTARS.2016.2520929

Habib, A., Zhou, T., Masjedi, A., Zhang, Z., Evan Flatt, J., and Crawford, M. (2018). Boresight calibration of GNSS/INS-assisted push-broom hyperspectral scanners on UAV platforms. *IEEE J. Select. Topics Appli. Earth Observ. Remote Sens.* 11, 1734–1749. doi: 10.1109/JSTARS.2018.2813263

Han, X., Thomasson, J. A., Bagnall, G. C., Pugh, N., Horne, D. W., Rooney, W. L., et al. (2018). Measurement and calibration of plant-height from fixed-wing UAV images. *Sensors* 18:4092. doi: 10.3390/s18124092

Han, X., and Thomasson, J. A. (2019). "Coordination and control for automatic mobile ground control points in agricultural remote sensing," in *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV*.

Hearst, A. A. (2014). *Automatic extraction of plots from geo-registered UAS imagery of crop fields with complex planting schemes* (Ph.D. Dissertations, Theses Gradworks). West Lafayette, IN: Purdue University; Purdue e-Pubs; ProQuest Dissertations Publishing.

Hickey, L. T., N., Hafeez, A., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9

Holman, F., Riche, A., Michalski, A., Castle, M., Wooster, M., and Hawkesford, M. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sens.* 8:1031. doi: 10.3390/rs8121031

Hu, P., Guo, W., Chapman, S. C., Guo, Y., and Zheng, B. (2019). Pixel size of aerial imagery constrains the applications of unmanned aerial vehicle in crop breeding. *ISPRS J. Photogramm. Remote Sens.* 154, 1–9. doi: 10.1016/j.isprsjprs.2019.05.008

Hu, Y., Knapp, S., and Schmidhalter, U. (2020). Advancing high-throughput phenotyping of wheat in early selection cycles. *Remote Sens.* 12:574. doi: 10.3390/rs12030574

Ingunn, B., Gunnar, L., Morten, L., Bleken, E., Grismstad, L., and Form, P. J. (2017). Exploring robots and UAVs as phenotyping tools in plant breeding. *IFAC Papersonline* 50, 11479–11484. doi: 10.1016/j.ifacol.2017.08.1591

Jiang, Y., Li, C., Robertson, J. S., Sun, S., Xu, R., and Paterson, A. H. (2018). GPhenovision: a ground mobile system with multi-modal imaging for field-based high throughput phenotyping of cotton. *Sci. Rep.* 8:1213. doi: 10.1038/s41598-018-19142-2

Jimenez-Berni, J. A., Deery, D. M., Rozas-Larraondo, P., Condon, A. T. G., Rebetzke, G. J., James, R. A., et al. (2018). High throughput determination of plant height, ground cover, and above-ground biomass in wheat with LiDAR. *Front. Plant Sci.* 9:237. doi: 10.3389/fpls.2018.00237

Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M. P., Hawkesford, M. J., Varshney, R. K., et al. (2020). "High-throughput estimation of crop traits: a review of ground and aerial phenotyping platforms," in *IEEE Geoscience Remote Sensing Magazine.* (Exeter; Devon, NH; Falmouth, MA; Cornwall: University of Exeter). doi: 10.1109/MGRS.2020.2998816

Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* 198, 105–114. doi: 10.1016/j.rse.2017.06.007

Jin, X., Yang, G., Xu, X., Yang, H., Feng, H., Li, Z., et al. (2015). Combined multi-temporal optical and radar parameters for estimating LAI and biomass in winter wheat using HJ and RADARSAR-2 Data. *Remote Sens.* 7, 13251–13272. doi: 10.3390/rs71013251

Kakeru, W., Wei, G., Keigo, A., Hideki, T., Hiromi, K. K., Masaaki, K., et al. (2017). High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. *Front. Plant Sci* 8:421. doi: 10.3389/fpls.2017.00421

Kar, S., Garin, V., Kholová, J., Vadez, V., Durbha, S. S., Tanaka, R., et al. (2020). SpaTemHTP: a data analysis pipeline for efficient processing and utilization of temporal high-throughput phenotyping data. *Front. Plant Sci.* 11:509. doi: 10.3389/fpls.2020.552509

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi: 10.1038/35048692

Khan, Z., Chopin, J., Cai, J., Eichi, V.-R., Haefele, S., and Miklavcic, S. (2018). Quantitative estimation of wheat phenotyping traits using ground and aerial imagery. *Remote Sens.* 10:950. doi: 10.3390/rs10060950

Kim, J. Y. (2020). Roadmap to high throughput phenotyping for plant breeding. *J. Biosyst. Eng.* 45, 43–55. doi: 10.1007/s42853-020-00043-0

Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A., et al. (2017). The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. *Func. Plant Biol.* 44, 154–168. doi: 10.1071/FP16165

Krajewski, P., Chen, D., Cwiek, H., Van Dijk, A. D., Fiorani, F., Kersey, P., et al. (2015). Towards recommendations for metadata and data handling in plant phenotyping. *J. Exp. Bot.* 66, 5417–5427. doi: 10.1093/jxb/erv271

Kuijken, R. C., Van Eeuwijk, F. A., Marcelis, L. F., and Bouwmeester, H. J. (2015). Root phenotyping: from component trait in the lab to breeding. *J. Exp. Bot.* 66, 5389–5401. doi: 10.1093/jxb/erv239

Li, W., Niu, Z., Huang, N., Wang, C., Gao, S., and Wu, C. (2015). Airborne LiDAR technique for estimating biomass components of maize:

a case study in Zhangye City, Northwest China. *Ecol. Indic* 57, 486–496. doi: 10.1016/j.ecolind.2015.04.016

Liu, S., Martre, P., Buis, S., Abichou, M., Andrieu, B., and Baret, F. (2019). Estimation of plant and canopy architectural traits using the digital plant phenotyping platform. *Plant Physiol.* 181, 881–890. doi: 10.1104/pp.19.00554

Lyra, D. H., Virlet, N., Sadeghi-Tehran, P., Hassall, K. L., Wingen, L. U., Orford, S., et al. (2020). Functional QTL mapping and genomic prediction of canopy height in wheat measured using a robotic field phenotyping platform. *J. Exp. Bot.* 71, 1885–1898. doi: 10.1093/jxb/erz545

Ma, D., Carpenter, N., Amatya, S., Maki, H., Wang, L., Zhang, L., et al. (2019). Removal of greenhouse microclimate heterogeneity with conveyor system for indoor phenotyping. *Comp. Electron. Agric.* 166:104979. doi: 10.1016/j.compag.2019.104979

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., et al. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. *Front. Plant Sci.* 8:2002. doi: 10.3389/fpls.2017.02002

Mahlein, A.-K., Kuska, M. T., Thomas, S., Wahabzada, M., Behmann, J., Rascher, U., et al. (2019). Quantitative and qualitative phenotyping of disease resistance of crops by hyperspectral sensors: seamless interlocking of phytopathology, sensors, and machine learning is needed! *Curr. Opin. Plant Biol.* 50, 156–162. doi: 10.1016/j.pbi.2019.06.007

Marconi, T. G., Oh, S., Ashapure, A., Chang, A., and Enciso, J. (2019). "Application of unmanned aerial system for management of tomato cropping system," in *Autonomous Air & Ground Sensing Systems for Agricultural Optimization & Phenotyping IV.* (College Station, TX: Texas A&M University).

Mascarenhas Maciel, G., Gallis, R. B. D. A., Barbosa, R. L., Pereira, L. M., Siquieroli, A. C. S., and Vitória Miranda Peixoto, J. (2019). Image phenotyping of inbred red lettuce lines with genetic diversity regarding carotenoid levels. *Int. J. Appl. Earth Observ. Geoinform.* 81, 154–160. doi: 10.1016/j.jag.2019.05.016

Mccouch, S., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., et al. (2013). Feeding the future. *Nature* 499, 23–24. doi: 10.1038/499023a

Meacham-Hensold, K., Fu, P., Wu, J., Serbin, S., Montes, C. M., Ainsworth, E., et al. (2020). Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging. *J. Exp. Bot.* 71, 2312–2328. doi: 10.1093/jxb/eraa068

Merchuk-Ovnat, L., Ovnat, Z., Amir-Segev, O., Kutsher, Y., Saranga, Y., and Reuveni, M. (2019). CoverageTool: a semi-automated graphic software: applications for plant phenotyping. *Plant Methods* 15:90. doi: 10.1186/s13007-019-0472-2

Miao, C., Pages, A., Xu, Z., Rodene, E., Yang, J., and Schnable, J. C. (2020). Semantic segmentation of sorghum using hyperspectral data identifies genetic associations. *Plant Phenomics* 2020:4216373. doi: 10.34133/2020/4216373

Mir, R. R., Reynolds, M., Pinto, F., Khan, M. A., and Bhat, M. A. (2019). High-throughput phenotyping for crop improvement in the genomics era. *Plant Sci.* 282, 60–72. doi: 10.1016/j.plantsci.2019.01.007

Mueller-Sim, T., Jenkins, M., Abel, J., and Kantor, G. (2017). "The robotanist: a ground-based agricultural robot for high-throughput crop phenotyping," in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore).

Muhammad, H., Mengjiao, Y., Awais, R., Xiuliang, J., Xianchun, X., Yonggui, X., et al. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* 10:809. doi: 10.3390/rs10060809

Murman, J. (2019). *Flex-Ro: a robotic high throughput field phenotyping system* (Ph.D. thesis). University of Nebraska-Lincoln, Lincoln, NE, United States.

Naito, H., Ogawa, S., Valencia, M. O., Mohri, H., Urano, Y., Hosoi, F., et al. (2017). Estimating rice yield related traits and quantitative trait loci analysis under different nitrogen treatments using a simple tower-based field phenotyping system with modified single-lens reflex cameras. *ISPRS J. Photogramm. Remote Sens.* 125, 50–62. doi: 10.1016/j.isprsjprs.2017.01.010

Neilson, E. H., Edwards, A. M., Blomstedt, C., Berger, B., Møller, B. L., and Gleadow, R. M. (2015). Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time. *J. Exp. Bot.* 66, 1817–1832. doi: 10.1093/jxb/eru526

Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., et al. (2015). Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. *Plant Cell Environ.* 38, 1980–1996. doi: 10.1111/pce.12516

Oehlschläger, J., Schmidhalter, U., and Noack, P. (2018). "UAV-based hyperspectral sensing for yield prediction in winter barley," in *2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* (Hanover).

Pandey, P., Ge, Y., Stoerger, V., and Schnable, J. C. (2017). High throughput *in vivo* analysis of plant leaf chemical properties using hyperspectral imaging. *Front. Plant Sci.* 8:1348. doi: 10.3389/fpls.2017.01348

Peirone, L. S., Pereyra Irujo, G. A., Bolton, A., Erreguerena, I., and Aguirrezábal, L. A. (2018). Assessing the efficiency of phenotyping early traits in a greenhouse automated platform for predicting drought tolerance of soybean in the field. *Front. Plant Sci.* 9:587. doi: 10.3389/fpls.2018.00587

Pieruschka, R., and Schurr, U. (2019). Plant phenotyping: past, present, and future. *Plant Phenomics* 2019:7507131. doi: 10.1155/2019/7507131

Pinto, F., Damm, A., Schickling, A., Panigada, C., Cogliati, S., Muller-Linow, M., et al. (2016). Sun-induced chlorophyll fluorescence from high-resolution imaging spectroscopy data to quantify spatio-temporal patterns of photosynthetic function in crop canopies. *Plant Cell Environ.* 39, 1500–1512. doi: 10.1111/pce.12710

Qiu, Q., Sun, N., Bai, H., Wang, N., Fan, Z., Wang, Y., et al. (2019). Field-based high-throughput phenotyping for maize plant using 3D LiDAR point cloud generated with a "Phenomobile". *Front. Plant Sci.* 10:554. doi: 10.3389/fpls.2019.00554

Raesch, A., Muller, O., Pieruschka, R., and Rascher, U. (2014). Field observations with Laser-Induced Fluorescence Transient (LIFT) method in barley and sugar beet. *Agriculture* 4, 159–169. doi: 10.3390/agriculture4020159

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., et al. (2018). Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. *Sci. Data* 5:180028. doi: 10.1038/sdata.2018.28

Roitsch, T., Cabrera-Bosquet, L., Fournier, A., Ghamkhar, K., Jiménez-Berni, J., Pinto, F., et al. (2019). Review: new sensors and data-driven approaches—a path to next generation phenomics. *Plant Sci.* 282, 2–10. doi: 10.1016/j.plantsci.2019.01.011

Sagan, V., Maimaitijiang, M., Sidike, P., Eblimit, K., Peterson, K., Hartling, S., et al. (2019a). UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue Pro R 640, and thermoMap cameras. *Remote Sens.* 11:330. doi: 10.3390/rs11030330

Sagan, V., Maimaitijiang, M., Sidike, P., Maimaitiyiming, M., Erkbol, H., Hartling, S., et al. (2019b). Uav/Satellite multiscale data fusion for crop monitoring and early stress detection. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W13, 715–722. doi: 10.5194/isprs-archives-XLII-2-W13-715-2019

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. *Eur. J. Agronomy* 70, 112–123. doi: 10.1016/j.eja.2015.07.004

Sankaran, S., Quirós, J. J., and Miklas, P. N. (2019). Unmanned aerial system and satellite-based high resolution imagery for high-throughput phenotyping in dry bean. *Comp. Electron. Agric.* 165: 104965. doi: 10.1016/j.compag.2019.104965

Shafiekhani, A., Kadam, S., Fritschi, F. B., and Desouza, G. N. (2017). Vinobot and vinoculer: two robotic platforms for high-throughput field phenotyping. *Sensors* 17:214. doi: 10.3390/s17010214

Shah, T., Xu, J., Zou, X., Cheng, Y., Nasir, M., and Zhang, X. (2018). Omics approaches for engineering wheat production under abiotic stresses. *Int. J. Mol. Sci.* 19:2390. doi: 10.3390/ijms19082390

Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., et al. (2016b). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE* 11:e0159781. doi: 10.1371/journal.pone.0159781

Shi, Y., Murray, S. C., Rooney, W. L., Valasek, J., Olsenholler, J., Pugh, N. A., et al. (2016a). "Corn and sorghum phenotyping using a fixed-wing UAV-based remote sensing system," in *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping: International Society for Optics and Photonics* (College Station, TX: Texas A&M University).

Stager, A., Tanner, H. G., and Sparks, E. E. (2019). Design and construction of unmanned ground vehicles for sub-canopy plant phenotyping. *arXiv preprint* arXiv:1903.10608.

Sunil, N., Rekha, B., Yathish, K., Sehkar, J., and Vadez, V. (2018). LeasyScan-an efficient phenotyping platform for identification of pre-breeding genetic stocks in maize. *Maize J.* 7, 16–22.

Susko, A. Q., Gilbertson, F., Heuschele, D. J., Smith, K., and Marchetto, P. (2018). An automatable, field camera track system for phenotyping crop lodging and crop movement. *HardwareX* 4:29. doi: 10.1016/j.ohx.2018.e00029

Tamouridou, A., Alexandridis, T., Pantazi, X., Lagopodi, A., Kashefi, J., and Moshou, D. (2017). Evaluation of UAV imagery for mapping *Silybum marianum* weed patches. *Int. J. Remote Sens.* 38, 2246–2259. doi: 10.1080/01431161.2016.1252475

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., and Bennett, M. (2017). Plant phenomics, from sensors to knowledge. *Curr. Biol.* 27, R770–R783. doi: 10.1016/j.cub.2017.05.055

Thomas, S., Behmann, J., Steier, A., Kraska, T., Muller, O., Rascher, U., et al. (2018). Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. *Plant Methods* 14:45. doi: 10.1186/s13007-018-0313-8

Thomasson, J. A., Han, X., Bagnall, G. C., Sima, C., and Rooney, W. L. (2019). "UAV data reliability improvements based on multifunctional GCPs (Conference Presentation)," in *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV*. (Starkville, MS: Mississippi State University, Mississippi State).

Tisne, S., Serrand, Y., Bach, L., Gilbault, E., Ben Ameur, R., Balasse, H., et al. (2013). Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *Plant J.* 74, 534–544. doi: 10.1111/tpj.12131

Tracy, S. R., Gómez, J. F., Sturrock, C. J., Wilson, Z. A., and Ferguson, A. C. (2017). Non-destructive determination of floral staging in cereals using X-ray micro computed tomography (μCT). *Plant Methods* 13:9. doi: 10.1186/s13007-017-0162-x

Vadez, V., Kholova, J., Hummel, G., Zhokhavets, U., Gupta, S. K., and Hash, C. T. (2015). LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. *J. Exp. Bot.* 66, 5581–5593. doi: 10.1093/jxb/erv251

Varshney, R. K., Thudi, M., Pandey, M. K., Tardieu, F., Ojiewo, C., Vadez, V., et al. (2018). Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy. *J. Exp. Bot*. 69, 3293–3312. doi: 10.1093/jxb/ery088

Villareal, M. K., Tongco, A. F., and Maja, J. M. J. (2020). Winter wheat crop height estimation Using Small Unmanned Aerial System (sUAS). *Agric. Sci.* 11, 355–368. doi: 10.4236/as.2020.114021

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., and Hawkesford, M. J. (2017). Field scanalyzer: an automated robotic field phenotyping platform for detailed crop monitoring. *Func. Plant Biol.* 44, 143–153. doi: 10.1071/FP16163

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (2020). Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer – a case study of small farmlands in the South of China. *Agric. Forest Meteorol.* 291:108096. doi: 10.1016/j.agrformet.2020.108096

Wang, X., Zhang, R., Song, W., Han, L., Liu, X., Sun, X., et al. (2019). Dynamic plant height QTL revealed in maize through remote sensing phenotyping using a high-throughput unmanned aerial vehicle (UAV). *Sci. Rep.* 9:3458. doi: 10.1038/s41598-019-39448-z

Wei, Y. (2017). *Handheld and Robotic Phenotyper Designs*. Kansas State University.

Werner, J. P. (2016). *Flex-Ro: design, implementation, and control of subassemblies for an agricultural robotic platform* (Ph.D. thesis), Biological Systems Engineering, Lincoln, NE, United States.

Wu, J., Yang, G., Yang, H., Zhu, Y., Li, Z., Lei, L., et al. (2020). Extracting apple tree crown information from remote imagery using deep learning. *Comp. Electron. Agric.* 174:105504. doi: 10.1016/j.compag.2020.105504

Xie, C., and Yang, C. (2020). A review on plant high-throughput phenotyping traits using UAV-based sensors. *Comp. Electron. Agric.* 178:105731. doi: 10.1016/j.compag.2020.105731

Yang, C., and Hoffmann, W. C. (2015). Low-cost single-camera imaging system for aerial applicators. *J. Appl. Remote Sens.* 9:096064. doi: 10.1117/1.JRS.9.096064

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., et al. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: current status and perspectives. *Front. Plant Sci.* 8:1111. doi: 10.3389/fpls.2017.01111

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., et al. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol. Plant* 13, 187–214. doi: 10.1016/j.molp.2020.01.008

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5:5087. doi: 10.1038/ncomms6087

Young, S. N. (2019). A framework for evaluating field-based, high-throughput phenotyping systems: a meta-analysis. *Sensors* 19:3582. doi: 10.3390/s19163582

Yu, N., Li, L., Schmitz, N., Tian, L. F., Greenberg, J. A., and Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* 187, 91–101. doi: 10.1016/j.rse.2016.10.005

Zaman-Allah, M., Vergara, O., Araus, J. L., Tarekegne, A., Magorokosho, C., Zarco-Tejada, P. J., et al. (2015). Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11:35. doi: 10.1186/s13007-015-0078-2

Zhang, C., Marzougui, A., and Sankaran, S. (2020). High-resolution satellite imagery applications in crop phenotyping: an overview. *Comp. Electron. Agric.* 175:105584. doi: 10.1016/j.compag.2020.105584

Zhao, C. (2019). Big data of plant phenomics and its research progress. *J. Agric. Big Data* 1, 5–18.

Zhou, J., Reynolds, D., Websdale, D., Le Cornu, T., Gonzalez-Navarro, O., Lister, C., et al. (2017). CropQuant: an automated and scalable field phenotyping platform for crop monitoring and trait measurements to facilitate breeding and digital agriculture. *BioRxiv [Preprints]*. 161547. doi: 10.1101/161547

Ziamtsov, I., and Navlakha, S. (2020). Plant 3D (P3D): a plant phenotyping toolkit for 3D point clouds. *Bioinformatics* 36, 3949–3950. doi: 10.1093/bioinformatics/btaa220

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Quan, Song, Li, Yu, Li and Muhammad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Plant Science

METHODS
published: 25 March 2021
doi: 10.3389/fpls.2021.642917

Hyperspectral Estimation Models of Winter Wheat Chlorophyll Content Under Elevated CO_2

*Yao Cai, Yuxuan Miao, Hao Wu and Dan Wang**

Department of Ecology, College of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing, China

Chlorophyll content is an important indicator of winter wheat health status. It is valuable to investigate whether the relationship between spectral reflectance and the chlorophyll content differs under elevated CO_2 condition. In this open-top chamber experiment, the CO_2 treatments were categorized into ambient (aCO_2; about 400 μmol·mol^{-1}) or elevated (eCO_2; ambient + 200 μmol·mol^{-1}) levels. The correlation between the spectral reflectance and the chlorophyll content of the winter wheat were analyzed by constructing the estimation model based on red edge position, sensitive band and spectral index methods, respectively. The results showed that there was a close relationship between chlorophyll content and the canopy spectral curve characteristics of winter wheat. Chlorophyll content was better estimated based on sensitive spectral bands and difference vegetation index (DVI) under both aCO_2 and eCO_2 conditions, though the accuracy of the models varied under different CO_2 conditions. The results suggested that the hyperspectral measurement can be effectively used to estimate the chlorophyll content under both aCO_2 and eCO_2 conditionsand could provide a useful tool for monitoring plants physiology and growth.

Keywords: elevated CO_2, hyperspectral estimation model, chlorophyll content, red edge position, sensitive band, spectral index, winter wheat

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Lingyan Zhou,
East China Normal University, China
Deepak Jaiswal,
University of Illinois
at Urbana-Champaign, United States
Xin Qiao,
ZJUT, China

****Correspondence:***
Dan Wang
wangdan@nuist.edu.cn

Specialty section:
This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 22 December 2020
Accepted: 04 March 2021
Published: 25 March 2021

Citation:
Cai Y, Miao Y, Wu H and Wang D (2021) Hyperspectral Estimation Models of Winter Wheat Chlorophyll Content Under Elevated CO_2. Front. Plant Sci. 12:642917. doi: 10.3389/fpls.2021.642917

INTRODUCTION

It is expected that the atmospheric CO_2 concentration will rise to 550 μmol·mol^{-1} in 2050 and reach or exceed 700 μmol·mol^{-1} at the end of the 21st century due to the increase of human population, energy production and utilization, deforestation and other intensive human activities (IPCC, 2013). Wheat is one of the world's most productive and important crops in the 21st century, and also the main source of food for human (Curtis and Halford, 2014). Under elevated CO_2, the physiology, growth and yield of wheat and other species are affected (Long et al., 2006; Wang et al., 2012).

Chlorophyll content was closely related to crop health, photosynthetic capacity and crop yield (Lukas et al., 2014). C_3 plants are more sensitive to elevated CO_2 than C_4 plants (Leakey et al., 2009). The chlorophyll content and photosynthetic rate of varieties of C_3 species, including crops and trees, was increased by elevated CO_2 (Zhang et al., 2013; Madhana et al., 2014; Fathurrahman et al., 2016; Choi et al., 2017). For wheat, previous studies had shown a positive (Dubey et al., 2015) or negative (Wang et al., 2013) CO_2 effects on the chlorophyll content and the difference might be resulted from the different experimental settings or CO_2 increasing levels used in different studies.

Given that systematic measurement of chlorophyll contents in elevated CO_2 condition is scarce, a detailed measurement of chlorophyll content of winter wheat throughout the growing season will be useful to understand the effect of elevated CO_2 on the physiology and growth of winter wheat.

Remote sensing methods could be used to accurately and rapidly relate variations in leaf optical properties with important plant characteristics, such as chlorophyll content and photosynthetic properties at the leaf and canopy scales (Ainsworth et al., 2014). Inversion of chlorophyll content by hyperspectral remote sensing was of great significance for crop growth status monitoring, yield estimation and agricultural planning (Liang et al., 2012; Flores-De-Santiago et al., 2013). Hyperspectral remote sensing had been used to monitor winter wheat chlorophyll content (He et al., 2018; Kasim et al., 2018). However, the application had been limited to specific test conditions (Serbin et al., 2012; Zhou et al., 2016) and there were few studies investigating hyperspectral remote sensing applications on winter wheat under elevated CO_2 conditions.

The hyperspectral estimation models could be determined through different techniques to extract hyperspectral characteristics, including reflectance spectrum and first derivative spectrum, absorption and reflection location and vegetation index (Li et al., 2014). Previous research studied the relationship between visible and near-infrared spectra and leaf chemical components and found out that the original spectral reflectance and the first and second derivatives of the spectra could be used to estimate crop agronomic parameters (Card et al., 1988). Red edge and sensitive bands based spectral models had been used to simulate chlorophyll and nitrogen content of many species (Hansen and Schjoerring, 2003; Chen et al., 2013; Clevers and Gitelson, 2013; Stratoulias et al., 2015). Identifying optimal hyperspectral estimation models of winter wheat under different CO_2 conditions is critical in crop growth monitoring and forecasting and requires further investigation.

In order to find an optimal estimation model for chlorophyll content and promote spectral analysis in the application of agriculture management under future global change conditions, an open top chamber (OTC) based CO_2 manipulation experiment was conducted for 2 years in this study. The objectives of this study were: (1) to establish statistical models to study the relationship between hyperspectral characteristics and chlorophyll content of winter wheat throughout the growing stages; (2) to investigate whether the relationship between hyperspectral characteristics and chlorophyll content varies under elevated CO_2 conditions.

MATERIALS AND METHODS

Experimental Site

The study site was located in the agrometeorological experimental station of Nanjing University of Information Science and Technology, in Nanjing city, Jiangsu province of China (32°16'N, 118°86'E). The climate in this region characterizes subtropical monsoon season, with annual average precipitation of 1,100 mm, the average temperature in recent years of 15.6℃ and the average annual frost-free period of 237 days. The soil texture in the tillage layer of winter wheat was loamy clay, and the clayey content was 26.1%. The bulk density of 0–20 cm soil was 1.57 $g \cdot cm^{-3}$, the pH (H_2O) value was 6.3, and the organic carbon and total nitrogen content were 11.95 and 1.19 $g \cdot kg^{-1}$, respectively.

Experimental Design

Open top chambers (OTC) were used in the experiment to manipulate CO_2 concentration. There were eight OTC chambers, all of which were octagonal prisms (opposite side diameter 3.75 m, height 3 m, bottom area 10 m^2) and equipped with aluminum alloy frames and toughened glass with high transmittance. There were two CO_2 treatments, ambient CO_2 (aCO_2) and elevated CO_2 (eCO_2, aCO_2 + 200 $\mu mol \cdot mol^{-1}$), each with four replicates. The treatment of elevated CO_2 started from regreening stage and lasted to the end of growing stage.

In order to avoid the rapid loss of CO_2 gas and reduce the experiment cost, the top opening of OTC was designed to tilt inward for 45°. The CO_2 concentration in the chambers was controlled with an automatic control platform, composed of CO_2 sensors, gas-supplying devices and automatic control system. Three wind-blowing fans were placed in each chamber to make the CO_2 gas in the chamber evenly distributed. The CO_2 sensor feeds back the CO_2 concentration information in the chamber to the automatic control system every two seconds. The CO_2 concentration averaged was 650 ± 58 $\mu mol \cdot mol^{-1}$ in elevated CO_2 chambers and 455 ± 42 $\mu mol \cdot mol^{-1}$ in ambient chambers across two growing seasons.

The local winter wheat variety of Ningmai 13 was selected in the study. The field measurement of spectrum and chlorophyll was conducted in 2018–2019 and 2019–2020 growing seasons. During the whole growing stages, fertilizer and water management were carried out in the local conventional way.

Spectrum Measurement

The spectral reflectance of winter wheat was measured by Field Spec4 of American analytical spectral device (ASD). The wavelength range was set at 350–2,500 nm. The sampling interval and resolution was set at 1.4 and 3 nm in the range of 350–1,000 nm; and 2 and 10 nm in 1,001–2,500 nm, respectively. The reflectance of winter wheat at five growth stages (jointing, booting, heading, filling and maturity stage) was measured on sunny days at 10:00 a.m.–2:00 p.m. Field Spec4 needed to be preheated 30 min before measurement. During the measurement, the sensor probe was placed vertically downward, the field of view angle was 10° and the probe was about 20 cm away from the top of the canopy. The measurement was carried out 10 times in different areas of an OTC. The reference white board was corrected immediately before and after the measurements in each chamber.

Measurement of Chlorophyll Content

At the same time as the spectral measurement, the chlorophyll content was measured by the portable chlorophyll meter SPAD-502. Relevant studies have shown that soil and plant analyzer developrnent (SPAD) value was positively correlated with the total chlorophyll content, with the correlation coefficient up to 0.99, and the SPAD value could be used to represent the chlorophyll content of plants (Costa et al., 2001; Uddling et al., 2007). When measuring the chlorophyll content, five wheat plants were selected at the corresponding position of canopy spectrum measurement, then SPAD values were measured for five times uniformly on the upper, middle and lower leaves of each plant, and the average value was taken as the chlorophyll content of this sample point. A total of 200 chlorophyll samples were measured in 2018–2019 and 120 samples in 2019–2020. Two years of data were combined together, among which 240 samples were selected to establish the models, and the remaining 80 samples were used to verify the models.

Statistical Analyses

View Specpro_6.0, Matlab_2017 and Origin_2018 were used to process and analyze the data. The spectral band range was set at 350–1,350 nm, and the wavelength corresponding to the largest first-order differential value in the red edge range (680–760 nm) was selected as the red edge position λr. The correlation analysis between canopy spectral reflectance and SPAD values of winter wheat was conducted, and the correlation coefficient was calculated to find out the sensitive bands. According to the original reflectance of winter wheat canopy, five common vegetation indexes were calculated. Each vegetation index had different characteristics. The normalized difference vegetation index (NDVI) was a common vegetation index and very sensitive to green vegetation. Ratio vegetation index (RVI) was sensitive to vegetation with high coverage. Difference vegetation index (DVI) and perpendicular vegetation index (PVI) were sensitive to the change of soil background. Optimizing soil and adjusting vegeta$\hat{y}_i$tion index (OSAVI) explained the changes in the optical characteristics of the background and corrected the sensitivity of NDVI to the soil background (Bannari et al., 2007; Yan et al., 2013). The calculation of each vegetation index was listed in **Table 1**.

Using the canopy spectral data of winter wheat, a regression estimation model with hyperspectral variables as independent variables and the chlorophyll content as dependent variables was established. Linear regression model was selected for all models:

$$Y = a + bx \tag{1}$$

In this study, the coefficient of determination (R^2) and the root mean square error (RMSE) were used to verify the linear regression model. The higher coefficient of determination R^2 and the smaller RMSE indicated a more accurate estimation model.

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{3}$$

Where $\hat{y}_i$ and y_i were the predicted values and measured values of the sample respectively, and i was the average value of the measured values of the sample, and n was the number of samples.

RESULTS

Chlorophyll Content

The chlorophyll content in winter wheat under aCO_2 and eCO_2 in five growth stages was listed in **Table 2**. The chlorophyll content was lowest in the maturing stage and highest in the heading stage and varied under different CO_2 treatments. At different growth stages, the effects of eCO_2 on chlorophyll content of winter wheat were different. In booting and heading stage, eCO_2 increased the chlorophyll content by 4.70–6.90%; in jointing, filling and maturity stage, eCO_2 decreased the chlorophyll content by 2.80–18.20%. During the whole growth stage, the chlorophyll content under aCO_2 was lower than that under eCO_2 (**Table 2**).

Canopy Spectral Reflectance

The original spectral band range was set at 350–1,350 nm and the canopy spectral reflectance under aCO_2 and eCO_2 at different

TABLE 1 | The calculation of the vegetation indexes.

Spectral index	Formulation	Authors
NDVI	$NDVI = (R_{NIR} - R_{RED})/(R_{NIR} + R_{RED})$	Rouse et al., 1973
RVI	$RVI = R_{NIR}/R_{RED}$	Jordan, 1969
DVI	$DVI = R_{NIR} - R_{RED}$	Richardson and Wiegand, 1977
PVI	$PVI = (R_{NIR} - 10.489 \times R_{RED} - 6.604)/\sqrt{1 + 10.489^2}$	Huete et al., 1985
OSAVI	$OSAVI = (1 + 0.16)(R_{NIR} - R_{RED})/(R_{NIR} + R_{RED} + 0.16)$	Rondeaux et al., 1996

TABLE 2 | The chlorophyll content of winter wheat.

Data composition	Jointing	Booting	Heading	Filling	Maturity	*SD*	CV%
aCO_2	48.04 ± 2.53^c	55.50 ± 2.15^b	57.42 ± 4.35^a	55.32 ± 3.78^b	44.28 ± 4.26^d	7.33	14.70
eCO_2	46.70 ± 3.53^d	58.10 ± 4.97^b	61.38 ± 3.75^a	51.98 ± 2.51^c	36.20 ± 7.36^e	10.65	21.38

SD is the standard deviation and the CV (%) is the coefficient of variation. Lowercase letters indicate significant levels ($p < 0.05$).

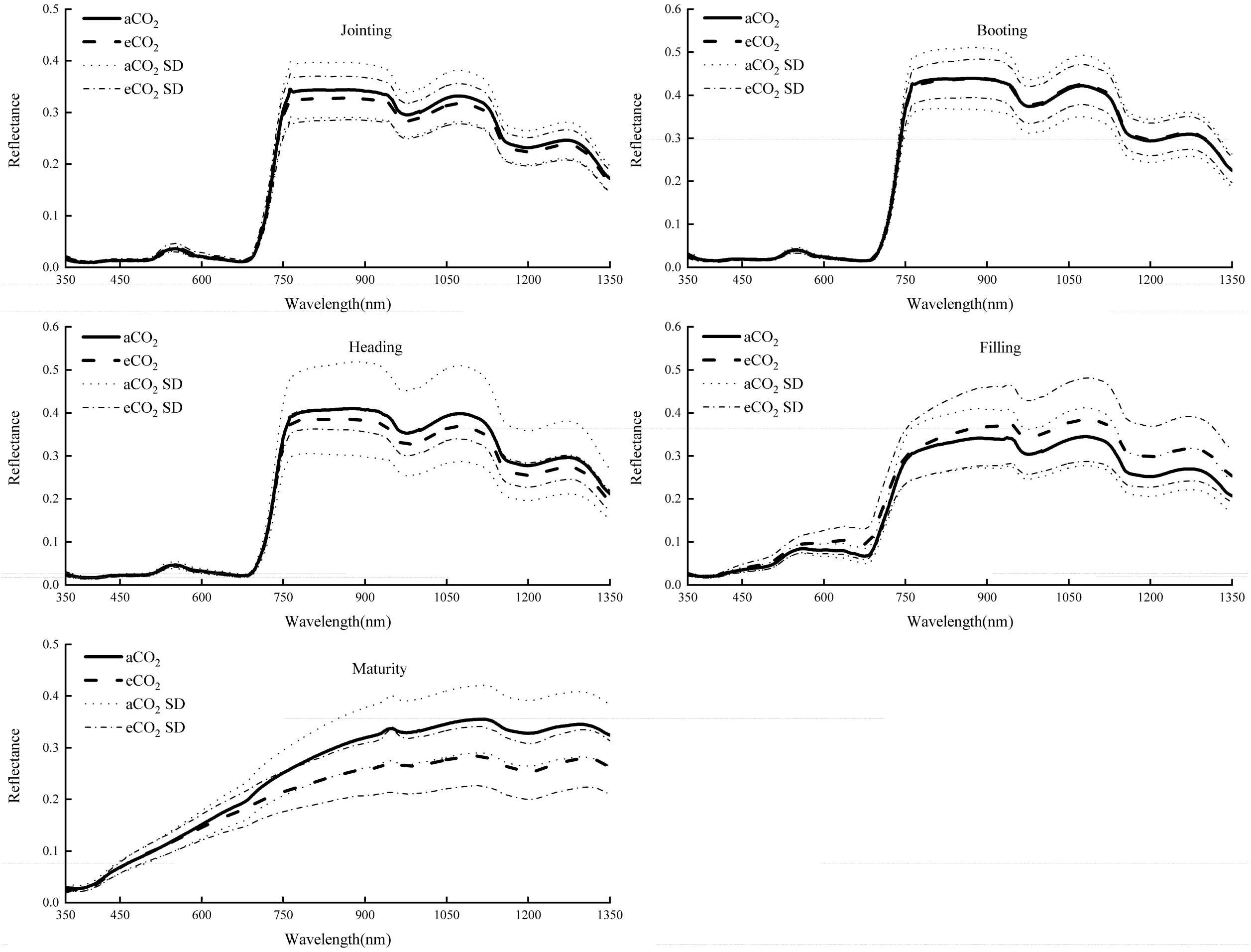


FIGURE 1 | Spectral reflectance of winter wheat canopy at different growing stages under aCO_2 and eCO_2.

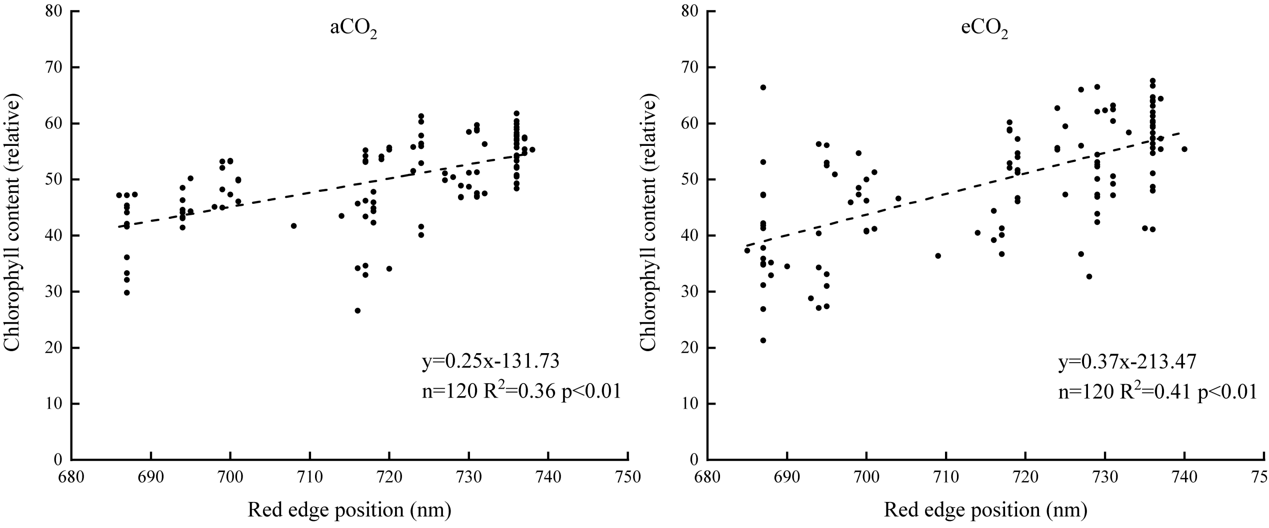


FIGURE 2 | The relationship between the chlorophyll content and the red-edge positions.

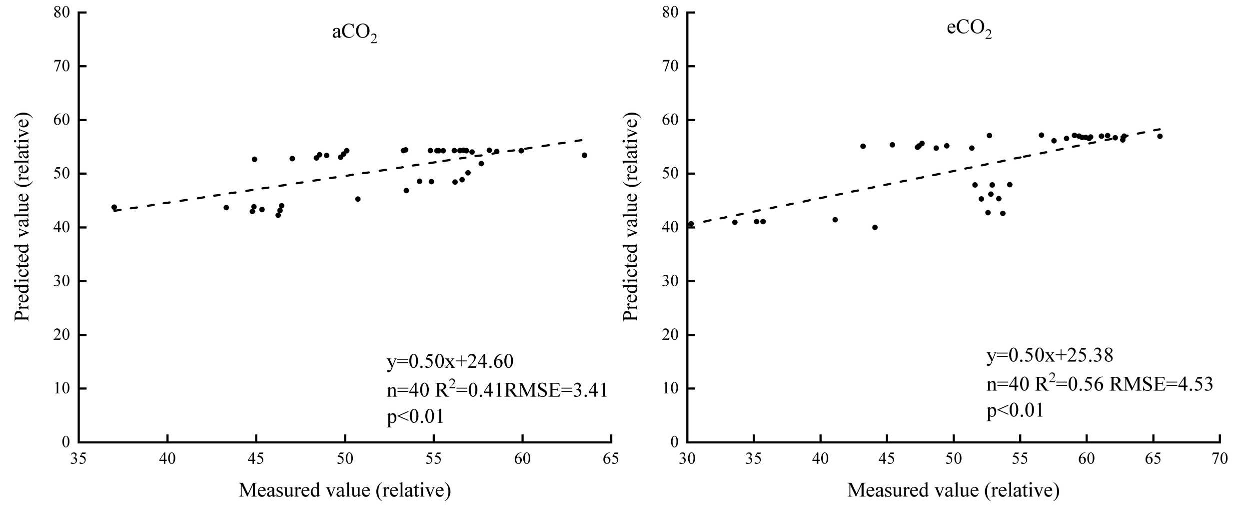


FIGURE 3 | The measured and estimated values of the chlorophyll content based on the red-edge position model under aCO$_2$ and eCO$_2$.

growth stages was shown in **Figure 1**. In each growth stage, the reflectance showed similar trend under aCO$_2$ and eCO$_2$, with an absorption band around 500 nm, an obvious "green peak" around 550 nm, the minimum value around 680 nm, and a "red edge" within the band range of 680–760 nm.

The spectral reflectance of winter wheat canopy under eCO$_2$ was slightly lower at jointing, heading and maturity stages, and was higher at filling stage than that at aCO$_2$, especially in the wavelength range of 760–1,350 nm. In the booting stage, the spectral reflectance of the two treatments was similar. Among the five growth stages, the spectral reflectance at the booting stage was the highest, reaching about 0.45.

Chlorophyll Content Estimation Models

The Red-Edge Position Model

Spectral reflectance rose rapidly at about 680 nm and slowly at about 760 nm. The band ranging between 680 and 760 nm was selected as the "red edge" spectrum. The linear regression model between the red edge position and the chlorophyll content was established to estimate chlorophyll content (**Figure 2**). The rest of the spectra and the chlorophyll content data were used to verify the model (**Figure 3**). The R^2 of the model was 0.36 and 0.41 under aCO$_2$ and eCO$_2$, respectively (**Figure 2**). The estimation model based on the red edge location estimated chlorophyll content slightly better under eCO$_2$ than under aCO$_2$ (**Figure 3**).

The Sensitive Band Spectral Model

The correlation coefficient of the spectral reflectance and chlorophyll content of winter wheat during the whole growth stage was analyzed (**Figure 4**). The correlation coefficient under aCO$_2$ was higher than that under eCO$_2$ between 350 and 1,350 nm. The canopy reflectance had the greatest correlation with the chlorophyll content at 740 and 749 nm under aCO$_2$ and eCO$_2$, respectively.

The sensitive bands of 740 and 749 nm were then selected under aCO$_2$ and eCO$_2$ respectively, and the linear model between the spectral reflectance and chlorophyll content of the sensitive bands was established to estimate the chlorophyll content of winter wheat (**Figure 5**). The model was validated using the rest of the sampling data (**Figure 6**). The R^2 of the linear model was 0.72 and 0.52 under aCO$_2$ and eCO$_2$, respectively (**Figure 5**) and the estimated values correlated well with the measured values of chlorophyll content under aCO$_2$ and eCO$_2$ (**Figure 6**).

The Spectral Index Model

Five different spectral indexes were extracted from the spectral reflectance curves (**Table 2**). Linear regression models of the five spectral indexes and chlorophyll contents of winter wheat were established. Under aCO$_2$, the rank of R^2 of the linear models was DVI > PVI > OSAVI > NDVI > RVI. Under eCO$_2$, the rank of the R^2 was DVI > OSAVI > NDVI > PVI > RVI

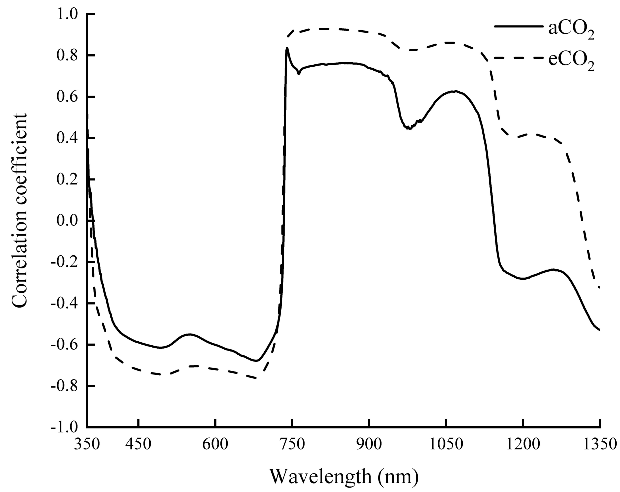


FIGURE 4 | The correlations between the chlorophyll content and the spectral reflectance of winter wheat during the whole growing season.

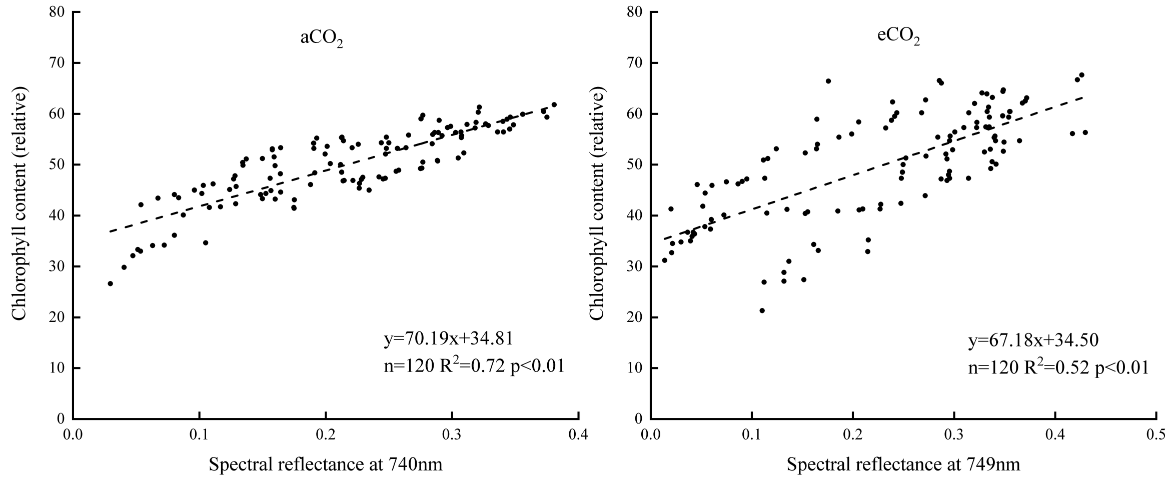


FIGURE 5 | The relationship between the chlorophyll content and the sensitive bands position of the spectral reflectance.

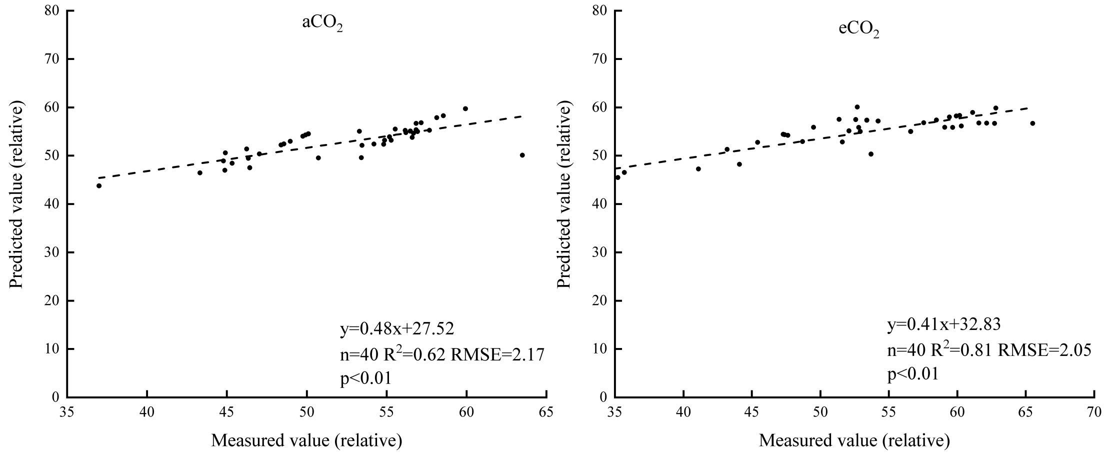


FIGURE 6 | The measured and estimated values of chlorophyll content based on the sensitive band position model under aCO$_2$ and eCO$_2$.

TABLE 3 | Estimation models of the chlorophyll content in wheat canopy based on different spectral indexes.

Treatment	Spectral index	Estimation equation	R^2	Significance
aCO$_2$	NDVI	y = 19.10x+38.22	0.56	$p < 0.01$
	RVI	y = 0.32x+46.38	0.21	$p < 0.01$
	DVI	y = 0.39x+40.80	0.67	$p < 0.01$
	PVI	y = 0.77x+34.28	0.57	$p < 0.01$
	OSAVI	y = 16.53x+38.23	0.56	$p < 0.01$
	NDVI	y = 27.83x+32.82	0.54	$p < 0.01$
	RVI	y = 0.50x+44.61	0.22	$p < 0.01$
eCO$_2$	DVI	y = 0.58x+37.31	0.60	$p < 0.01$
	PVI	y = 0.71x+36.72	0.30	$p < 0.01$
	OSAVI	y = 24.04x+32.89	0.54	$p < 0.01$

(**Table 3**). The DVI based estimation models was established using half of the measured data (**Figure 7**) and validated using the rest of the sampling data (**Figure 8**). The R^2 of the linear model established was 0.67 under aCO$_2$ and 0.60 at eCO$_2$ (**Figure 7**) and the estimated correlated well with the measured values of chlorophyll content under aCO$_2$ and eCO$_2$ (**Figure 8**).

DISCUSSION

In order to establish statistical models to study the relationship between the optical properties and chlorophyll content of winter wheat under elevated CO$_2$ conditions, we measured the chlorophyll content and spectral reflectance in winter wheat canopy under aCO$_2$ and eCO$_2$ conditions throughout the growing season for 2 years. The effects of elevated CO$_2$ on the chlorophyll content and spectral reflectance depended upon growing stages. The statistical models established in this study was effective under both ambient and elevated CO$_2$ conditions.

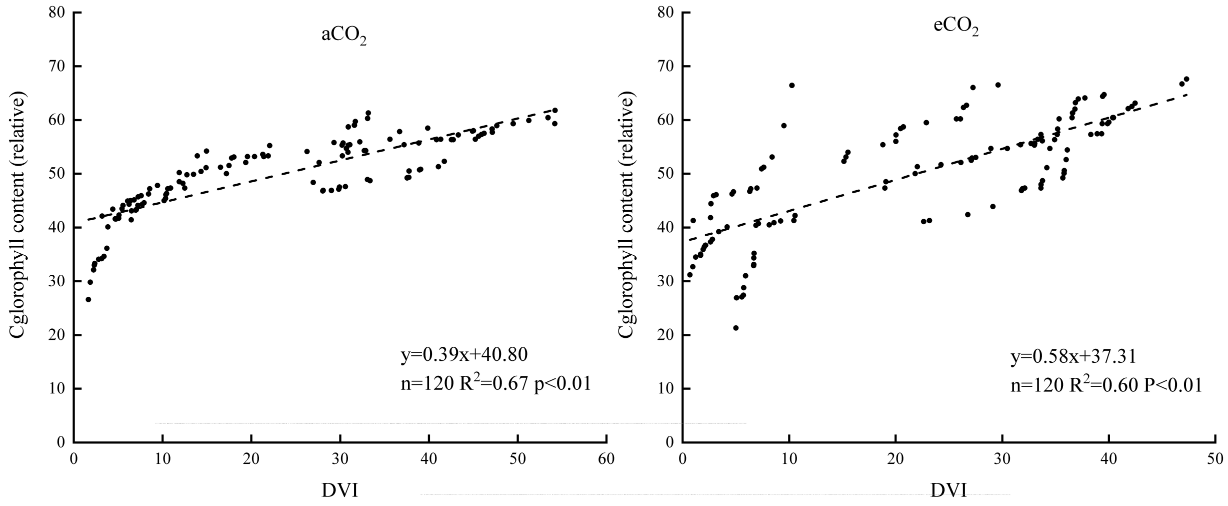


FIGURE 7 | The relationship between the chlorophyll content and the DVI.

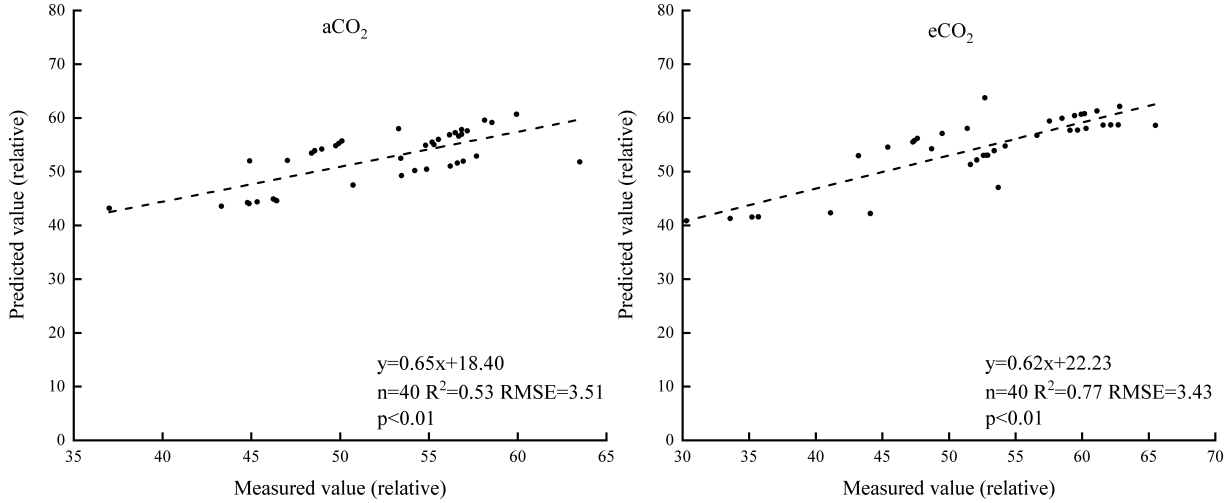


FIGURE 8 | The measured and estimated values of the chlorophyll content based on the DVI model under aCO$_2$ and eCO$_2$.

Elevated CO$_2$ increased the chlorophyll content of winter wheat at booting and heading stage, but decreased it at filling and maturity stage in this study. Elevated CO$_2$ usually had a positive effect on the chlorophyll content, but the specific effect depended on treatment duration and different species (Long et al., 2004). The increase of CO$_2$ concentration in the late growing stage might lead to the faster decline of chlorophyll concentration of wheat (Ommen et al., 1999). In this study, the senescence of winter wheat under eCO$_2$ was faster than that under aCO$_2$ and the chlorophyll content was decreased under eCO$_2$ at the later growing stages. The overall shapes of spectral curves did not change throughout the growing season, except in the maturity stage, the curves flattened due to the senescence of the leaves. The effect of elevated CO$_2$ on the spectral curves varied at different growing stages, with no impact in the earlier jointing and boosting stages, positive impact in the filling and negative impact in the heading and maturity stages. Though elevated CO$_2$ changed the maximum reflectance, it did not change the overall shape of the spectral curves of winter wheat at all the growing stages. The results were consistent with previous studies where the shapes of soybean canopy spectral curve did not change under different CO$_2$ treatments (Gray et al., 2010) and O$_3$ concentrations (Campbell et al., 2007).

Red edge position, sensitive band and vegetation index were effective means to retrieve crop chlorophyll content from the spectral curves (Dou et al., 2018; Kasim et al., 2018; Wang et al., 2019). Previous studies had shown that the position and reflectance of red edge were highly correlated with chlorophyll content of plant leaves and could be used as an indicator of chlorophyll content (Filella and Penuelas, 1994; Gitelson et al., 1996). The current study showed that the reflectance at the

680–740 nm wavelengths had a positive relationship with the content of chlorophyll under both ambient and elevated CO_2.

Sensitive bands could be used to calculate spectral indexes, which were sensitive to the difference of chlorophyll concentration in plant canopy (Hunt et al., 2011). The current results showed that the sensitive bands at 740 and 749 nm wavelength correlated with the chlorophyll content most, under aCO_2 and eCO_2, respectively (**Figures 4–6**), even though the established model fit slightly better under aCO_2 than at eCO_2 conditions. Vegetation indexes calculated from hyperspectral remote sensing technology had long been used to monitor the chlorophyll content of vegetation leaves (Meng et al., 2012; Guo et al., 2020). Among the five tested vegetation indexes, the DVI based model simulated the chlorophyll content best under both aCO_2 and eCO_2 conditions and the model using overall data from both the CO_2 treatments gave similar results (results not shown). Though the methods tested in the study proved effective to simulate winter wheat chlorophyll content under different CO_2 conditions, further investigations on how the spectral reflectance correlates with other biochemical contents and biophysical processes are still urgently needed for the purpose of guiding crop management and monitoring crop growth status in the future climate change situations.

In conclusion, the hyperspectral estimation models based on the red edge position, sensitive band and DVI vegetation index could all simulate the chlorophyll content of winter wheat. The accuracy of vegetation index and sensitive bands based models was higher than that of the red edge position model. The results suggested that the hyperspectral measurement can be effectively used to estimate the chlorophyll content under both aCO_2 and eCO_2 conditions and different equations should be established at specific CO_2 growing conditions based on the methods chosen. The findings in the study were useful in providing hyperspectral methods to monitor the growth status of winter wheat in the future global change situations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

DW designed and came up with the idea of the experiment. YC conducted the experiment, analyzed data, and wrote the manuscript. YM and HW helped in the field experiment and provided critical feedbacks on the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Nanjing University of Information Science and Technology (2013r115), Jiangsu Distinguished Professor Scholarship, Jiangsu six talent peaks (R2016L15), the Jiangsu Natural Science Foundation (BK20150894) and the National Natural Science Foundation of China (31500503 and 31770485) through DW.

ACKNOWLEDGMENTS

Wenhui Tao and Rui Kong help setting up and manage the experiment site.

REFERENCES

Ainsworth, E. A., Serbin, S. P., Skoneczka, J. A., and Townsend, P. A. (2014). Using leaf optical properties to detect ozone effects on foliar biochemistry. *Photosynth. Res.* 119, 65–76. doi: 10.1007/s11120-013-9837-y

Bannari, A., Khurshid, K. S., Staenz, K., and Schwarz, J. W. (2007). A comparison of hyperspectral chlorophyll indices for wheat crop chlorophyll content estimation using laboratory reflectance measurements. *IEEE Trans. Geosci. Remote Sens.* 45, 3063–3074. doi: 10.1109/TGRS.2007.897429

Campbell, P. K. E., Middleton, E. M., Mcmurtrey, J. E., Crop, L. A., and Chappelle, E. W. (2007). Assessment of vegetation stress using reflectance or fluorescence measurements. *J. Environ. Qual.* 36, 832–845. doi: 10.2134/jeq2005.0396

Card, D. H., Peterson, D. L., Matson, P. A., and Aber, J. D. (1988). Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sens. Environ.* 26, 123–147. doi: 10.1016/0034-4257(88)90092-2

Chen, B., Han, H. Y., Wang, F. Y., Liu, Z., and Xiao, C. H. (2013). Monitoring chlorophyll and nitrogen contents in cotton leaf infected by verticillium wilt with spectra red edge parameters. *Acta Agron. Sin.* 39:319. doi: 10.3724/SP.J.1006.2013.00319

Choi, D., Watanabe, Y., Guy, R. D., Sugai, T., Toda, H., and Koike, T. (2017). Photosynthetic characteristics and nitrogen allocation in the black locust (*Robinia pseudoacacia* L.) grown in a FACE system. *Acta Physiol. Plant.* 39:71. doi: 10.1007/s11738-017-2366-0

Clevers, J. G. P. W., and Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *Int. J. Appl. Earth Observ. Geoinform*. 23, 344–351. doi: 10.1016/j.jag.2012.10.008

Costa, C., Dwyer, L. M., Dutilleul, P., Stewart, D. W., Ma, B. L., and Smith, D. L. (2001). Inter-relationships of applied nitrogen, spad, and yield of leafy and non-leafy maize genotypes. *J. Plant Nutr.* 24, 1173–1194. doi: 10.1081/PLN-100106974

Curtis, T., and Halford, N. G. (2014). Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. *Ann. Appl. Biol.* 164, 354–372. doi: 10.1111/aab.12108

Dou, Z. G., Cui, L. J., Li, J., Zhu, Y. N., Gao, C. J., and Pan, X. (2018). Hyperspectral estimation of the chlorophyll content in short-term and long-term restorations of mangrove in Quanzhou bay estuary, China. *Sustainability* 10:1127. doi: 10.3390/su10041127

Dubey, S. K., Tripathi, S. K., and Pranuthi, G. (2015). Effect of elevated CO_2 on wheat crop: mechanism and impact. *Crit. Rev. Environ. Sci. Technol.* 45, 2283–2304. doi: 10.1080/10643389.2014.1000749

Fathurrahman, F., Nizam, M. S., Juliana, W. A. W., Doni, F., and Radziah, C. M. Z. C. (2016). Growth improvement of rain tree (Albizia saman Jacq. Merr) seedlings under elevated concentration of carbon dioxide (CO_2). *J. Pure Appl. Microbiol.* 10, 1911–1917.

Filella, I., and Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote Sens.* 15, 1459–1470. doi: 10.1080/01431169408954177

Flores-De-Santiago, F., Kovacs, J. M., and Flores-Verdugo, F. (2013). The influence of seasonality in estimating mangrove leaf chlorophyll-a content from hyperspectral data. *Wetl. Ecol. Manag.* 21, 193–207. doi: 10.1007/s11273-013-9290-x

Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *J. Plant Physiol*. 148, 501–508. doi: 10.1016/S0176-1617(96)80285-9

Gray, S. B., Dermody, O., and Delucia, E. H. (2010). Spectral reflectance from a soybean canopy exposed to elevated CO_2 and O_3. *J. Exp. Bot.* 61, 4413–4422. doi: 10.1093/jxb/erq244

Guo, Y. H., Wang, H. X., Wu, Z. F., Wang, S. X., Sun, H. Y., and Senthilnath, J. (2020). Modified red blue vegetation index for chlorophyll estimation and yield prediction of maize from visible images captured by UAV. *Sensors* 20:5055. doi: 10.3390/s20185055

Hansen, P. M., and Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* 86, 542–553. doi: 10.1016/S0034-4257(03)00131-7

He, R. Y., Li, H., Qiao, X. J., and Jiang, J. B. (2018). Using wavelet analysis of hyperspectral remote-sensing data to estimate canopy chlorophyll content of winter wheat under stripe rust stress. *Int. J. Remote Sens.* 39, 4059–4076. doi: 10.1080/01431161.2018.1454620

Huete, A. R., Jackson, R. D., and Post, D. F. (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote Sens. Environ.* 17, 37–53. doi: 10.1016/0034-4257(85)90111-7

Hunt, E. R., Daughtry, C., Eitel, J., and Long, D. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agron. J.* 103:1090. doi: 10.2134/agronj2010.0395

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology* 50, 663–666. doi: 10.2307/1936256

Kasim, N., Sawut, R., Abliz, A., Shi, Q. D., Maihmuti, B., and Yalkun, A. (2018). Estimation of the relative chlorophyll content in spring wheat Based on an optimized spectral index. *Photogramm. Eng. Remote Sens.* 84, 801–811. doi: 10.14358/PERS.84.12.801

Leakey, A. D. B., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., and Ort, D. R. (2009). Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *J. Exp. Bot.* 60, 2859–2876. doi: 10.1093/jxb/erp096

Li, X. C., Zhang, Y. J., Bao, Y. S., Luo, J. H., Jin, X. L., and Xu, X. G. (2014). Exploring the best hyperspectral features for LAI estimation using partial least squares regression. *Remote Sens.* 6, 6221–6241. doi: 10.3390/rs6076221

Liang, S., Zhao, G. X., and Zhu, X. C. (2012). Hyperspectral estimation models of chlorophyll content in apple leaves. *Spectrosc. Spectr. Anal.* 32, 1367–1370. doi: 10.3964/j.issn.1000-0593201205-1367-04

Long, S. P., Ainsworth, E. A., Leakey, A. D. B., Nösberger, J., and Ort, D. R. (2006). Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. *Science* 312, 1918–1921. doi: 10.1126/science.1114722

Long, S. P., Ainsworth, E. A., Rogers, A., and Ort, D. R. (2004). Rising atmospheric carbon dioxide: plants FACE the future. *Annu. Rev. Plant Biol.* 55, 591–628. doi: 10.1146/annurev.arplant.55.031903.141610

Lukas, V., Rodriguez-Moreno, F., Dryslova, T., and Neudert, L. (2014). Effectiveness of chlorophyll meter measurement in winter wheat at field scale level. *Agriculture* 60, 41–49. doi: 10.2478/agri-2014-0005

Madhana, S. K., Rachapudi, V. S., Mudalkar, S., and Reddy, A. R. (2014). Persistent stimulation of photosynthesis in short rotation coppice mulberry under elevated CO_2 atmosphere. *J. Photochem. Photobiol. B Biol.* 137, 21–30. doi: 10.1016/j.jphotobiol.2014.05.001

Meng, Q. Y., Dong, H., Qin, Q. M., Wang, J. L., and Zhao, J. H. (2012). MTCARI: a kind of vegetation index monitoring vegetation leaf chlorophyll content based on hyperspectral remote sensing. *Spectrosc. Spectr. Anal.* 32, 2218–2222. doi: 10.3964/j.issn.1000-0593201208-2218-05

Ommen, O. E., Donnelly, A., Vanhoutvin, S., Vanoijen, M., and Manderscheid, R. (1999). Chlorophyll content of spring wheat flag leaves grown under elevated CO_2 concentration and other environmental stresses within the ESPACE-wheat project. *Eur. J. Agron.* 10, 197–203. doi: 10.1016/s1161-0301(99)00011-8

Richardson, A. J., and Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogramm. Eng. Remote Sens.* 43, 1541–1552. doi: 10.1109/TGE.1977.294499

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. doi: 10.1016/0034-4257(95)00186-7

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1973). "Monitoring vegetation systems in the great plains with ERTS," in *Proceedings of the 3rd ERTS Symposium,* NASA SP-351 (Washington, D.C: NASA Special Publication), 309–317.

Serbin, S. P., Dillaway, D. N., Kruger, E. L., and Townsend, P. A. (2012). Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *J. Exp. Bot*. 63, 489–502. doi: 10.1093/jxb/err294

Stratoulias, D., Balzter, H., Zlinszky, A., and Tóth, V. (2015). Assessment ofecophysiology of lake shore reed vegetation based on chlorophyll fluorescence, field spectroscopy and hyperspectral airborne imagery. *Remote Sens. Environ.* 157, 72–84. doi: 10.1016/j.rse.2014.05.021

Uddling, J., Gelang-Alfredsson, J., Piikki, K., and Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynth. Res.* 91, 37–46. doi: 10.1007/s11120-006-9077-5

Wang, D., Heckathorn, S. A., Wang, X. Z., and Philpott, S. M. (2012). A meta-analysis of plant physiological and growth responses to temperature and elevated CO_2. *Oecologia* 169, 1–13. doi: 10.1007/s00442-011-2172-0

Wang, L., Feng, Z. Z., and Schjoerring, J. K. (2013). Effects of elevated atmospheric CO_2 on physiology and yield of wheat (Triticum aestivum L.): A meta-analytic test of current hypotheses. *Agric. Ecosyst. Environ.* 178, 57–63. doi: 10.1016/j.agee.2013.06.013

Wang, Y. M., Zhao, S. H., Cai, W. T., Heo, J., and Peng, F. C. (2019). A sensitive band to optimize winter wheat crop residue cover estimation by eliminating moisture effect. *Sustainability* 11, 1–18. doi: 10.3390/su11113032

Yan, F., Wu, B., and Wang, Y. J. (2013). Estimating aboveground biomass in Mu Us Sandy Land using Landsat spectral derived vegetation indices over the past 30 years. *J. Arid Land* 5, 521–530. doi: 10.1007/s40333-013-0180-0

Zhang, X. C., Yu, X. F., and Ma, Y. F. (2013). Effect of nitrogen application and elevated CO_2 on photosynthetic gas exchange and electron transport in wheat leaves. *Photosynthetica* 51, 593–602. doi: 10.1007/s11099-013-0059-5

Zhou, X. F., Huang, W. J., Kong, W. P., Ye, H. C., and Luo, J. H. (2016). Remoteestimation of canopy nitrogen content in winter wheat using airbornehyperspectral reflectance measurements. *Adv. Space Res.* 58, 1627–1637. doi: 10.1016/j.asr.2016.06.034

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Cai, Miao, Wu and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Plant Science

ORIGINAL RESEARCH
published: 30 March 2021
doi: 10.3389/fpls.2021.614417

Estimations of Water Use Efficiency in Winter Wheat Based on Multi-Angle Remote Sensing

Hai-Yan Zhang, Meng-Ran Liu, Zi-Heng Feng, Li Song, Xiao Li, Wan-Dai Liu, Chen-Yang Wang* and Wei Feng*

State Key Laboratory of Wheat and Maize Crop Science, National Engineering Research Center for Wheat, Henan Agricultural University, Zhengzhou, China

Real-time non-destructive monitoring of water use efficiency (WUE) is important for screening high-yielding high-efficiency varieties and determining the rational allocation of water resources in winter wheat production. Compared with vertical observation angles, multi-angle remote sensing provides more information on mid to lower parts of the wheat canopy, thereby improving estimates of physical and chemical indicators of the entire canopy. In this study, multi-angle spectral reflectance and the WUE of the wheat canopy were obtained at different growth stages based on field experiments carried out across 4 years using three wheat varieties under different water and nitrogen fertilizer regimes. Using appropriate spectral parameters and sensitive observation angles, the quantitative relationships with wheat WUE were determined. The results revealed that backward observation angles were better than forward angles, while the common spectral parameters Lo and NDDAig were found to be closely related to WUE, although with increasing WUE, both parameters tended to become saturated. Using this data, we constructed a double-ratio vegetation index (NDDAig/FWBI), which we named the water efficiency index (WEI), reducing the impact of different test factors on the WUE monitoring model. As a result, we were able to create a unified monitoring model within an angle range of −20–10°. The equation fitting determination coefficient (R^2) and root mean square error (RMSE) of the model were 0.623 and 0.406, respectively, while an independent experiment carried out to test the monitoring models confirmed that the model based on the new index was optimal, with R^2, RMSE, and relative error (RE) values of 0.685, 0.473, and 11.847%, respectively. These findings suggest that the WEI is more sensitive to WUE changes than common spectral parameters, while also allowing wide-angle adaptation, which has important implications in parameter design and the configuration of satellite remote sensing and UAV sensors.

Keywords: winter wheat, hyperspectral remote sensing, angle adaptability, water use efficiency, monitoring model

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Weiping Kong,
Chinese Academy of Sciences, China
Zitong Yu,
Kansas State University,
United States

***Correspondence:**
Chen-Yang Wang
xmzxwang@163.com
Wei Feng
fengwei78@126.com

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *06 October 2020*
Accepted: *08 March 2021*
Published: *30 March 2021*

Citation:
Zhang H-Y, Liu M-R, Feng Z-H, Song L, Li X, Liu W-D, Wang C-Y and Feng W (2021) Estimations of Water Use Efficiency in Winter Wheat Based on Multi-Angle Remote Sensing. Front. Plant Sci. 12:614417. doi: 10.3389/fpls.2021.614417

INTRODUCTION

Wheat is one of the important food crops in the world, and with recent economic development and population growth, the level of winter wheat production has become even more important for ensuring world food security. Meanwhile, due to global climate change, lack of water has become a key limiting factor in winter wheat production. Water use efficiency (WUE) is a broad agronomic concept that reflects the comprehensive effect of crops on water use. Leaf WUE is the main criterion used to measure drought tolerance and efficient water use in crops, and subsequent selection and screening of high WUE varieties is one of the most important goals of crop breeding (Richards, 2006; Zhang et al., 2007). Efficient use of limited water resources and increases in overall WUE has therefore become an urgent goal of winter wheat production.

Leaf WUE is defined as the ratio of net photosynthesis (P_N) to transpiration (Tr; Condon et al., 2002), which can be estimated using the carbon isotope ratio ($\delta^{13}C$; Hultine and Marshall, 2000). Recently, rapid development of remote sensing technology has provided an effective tool for large-scale analysis of water use in crops. Compared with traditional crop WUE monitoring and diagnostic tools, hyperspectral remote sensing technology has made it possible to obtain a huge amount of continuous large-scale data in a more efficient manner (Peñuelas et al., 1993b; Dong et al., 2011). Ground hyperspectral remote sensing technology selects sensitive bands using spectral characteristic information to obtain vegetation indexes, which are used to establish estimation models (Hatfield et al., 2008; Mistele and Schmidhalter, 2008).

An appropriate water content is the basis of vigorous plant growth and efficient water use. As early as 1971, Thomas et al. (1971) analyzed the relationship between the leaf water content (LWC) and spectral reflectance, and revealed a strong correlation with reflectance at 1450 and 1930 nm. Similarly, Carter (1991) suggested that the near-infrared absorption peak at 950–970 nm could be used to monitor plant moisture content, while Dawson et al. (1999) examined the performance of the moisture spectral index to estimate the canopy water content, revealing that indexes based on 970 and 1200 nm water absorption characteristics had a high coefficient of determination (Dawson et al., 1999). However, studies also suggest that reflectance at 970, 1200, and 1900 nm is easily affected by starch, protein, and nitrogen (Curran, 1989). For example, Sims and Gamon (2003) revealed that the best spectral bands for remote estimates of the plant water content at the canopy scale were 1150–1260 and 1520–1540 nm (Sims and Gamon, 2003), while in addition to the near infrared region, reflectance at 690 and 740 nm have also been shown to reflect water stress in plants (Dobrowski et al., 2005). Screening and analysis of spectral bands that are sensitive to water also provides a basis for the establishment of relevant vegetation indexes that reflect the water status. Peñuelas et al. (1993a) combined the water absorption band at 970 nm and the reference band at 900 nm as a ratio to establish a water index (WI) capable of tracking changes in water content. Similarly, Zarco-Tejada et al. (2003) used MODIS data to construct a plant water index (PWI) for monitoring vegetation moisture content, and revealed good consistency with the water content of ground crops (Zarco-Tejada et al., 2003). Moreover, the floating-position water band index (FWBI) has also been established, which uses the reflectance at 900 nm and minimum reflectance at 900–980 nm to represent the water status (Strachan et al., 2002). Yao et al. (2014) subsequently introduced a new water-sensitive band based on the normalized vegetation index NDSI (1429, 416) to construct a three-band vegetation index capable of estimating the leaf equivalent water thickness. However, these previous studies mainly used the sensor to obtain two-dimensional information of the crop in a vertical direction, and failed to include data from middle to lower parts of the canopy. The accuracy of remote sensing monitoring therefore requires further improvements.

Compared with vertical observation angles, the multi-angle observation method collects data from different directions, providing multi-dimensional information and representing a new method of remote sensing monitoring (Thenkabail et al., 2000; Pocewicz et al., 2007; Huang et al., 2011). A number of studies have been carried out to extract optical and structural information using multi-angle observations (Cierniewski et al., 2004; Rautiainen et al., 2004), suggesting that multi-angle remote sensing technology can improve the ability of a vegetation index to estimate crop canopy structure and distinguish between crop varieties (Shibayama and Wiegand, 1985; Diner et al., 1999). The photochemical reflectance index (PRI) is notably affected by the observation angle. For example, PRI values calculated using backward spectral data tend to be higher than those obtained with forward observation data (Drolet et al., 2008; Garbulsky et al., 2011; Middleton et al., 2011). Furthermore, Galvão et al. (2009) found that data collected in a backscattering direction was better at distinguishing between different soybean varieties, while Chopping et al. (2003) obtained canopy characteristics of desert grassland using multi-angle remote sensing data. Chen et al. (2005) and Leblanc et al. (2005) proposed a multi-angle index for measuring leaf aggregation based on hot and dark spot reflectivity. Multi-angle hyperspectral remote sensing has also made great progress in estimations of crop pigment content and nitrogen content (Stagakis et al., 2010; He et al., 2015). Meanwhile, He et al. (2016) constructed an angle insensitivity index (AIVI) based on analysis of different bands and vegetation indexes, improving the accuracy of plant nitrogen content estimations and expanding the scope of application.

In the field of remote sensing monitoring, in addition to analyses of vegetation canopy structure and physiological indicators, the performance indicators of crop production are also important. The P_N of field crops has been shown to be significantly correlated with physiological indicators, and studies suggest the use of the ratio vegetation index ($R_{810/680}$)

to directly estimate the P_N of rice leaves (Tian et al., 2005). In addition, based on hyperspectral data, Zhang et al. (2018) established an estimation model of nitrogen fertilizer use efficiency in winter wheat, while photosynthetic effective radiation (FPAR) captured by the canopy can also be directly obtained through multi-angle remote sensing (Chen et al., 2003).

Light use efficiency (LUE) is one of the most important traits in crops, and is usually reflected by the PRI (Zarco-Tejada et al., 2013). For example, Hall et al. (2008) proposed a method to obtain forest LUE directly from space by measuring the shadow component of the PRI based on multi-angle spectral information (Hall et al., 2008). Meanwhile, Bastiaanssen et al. (1999) used remote sensing data to estimate crop yield and crop transpiration (ETc), and then determined the WUE of crops in the Barkra region of India. Li et al. (2005) also used remote sensing observation means combined with meteorological data to invert crop WUE in the Haihe River Basin by estimating crop transpiration. Thus, while progress has been made in the use of remote sensing data to monitor crop WUE, the monitoring indicators, methods, and models remain inconsistent due to geographical differences and crop types, as well as differences in cultivation conditions. However, estimations of WUE utilization efficiency at the leaf scale based on hyperspectral remote sensing data are lacking, especially with regards the influence of different observation angles, and the angle range of model adaptation requires further clarification. The main goal of this study, therefore, was to create a model capable of estimating the instantaneous WUE of winter wheat leaves based on multi-angle hyperspectral remote sensing data. By clarifying the relationship between common spectral parameters and WUE at different vertical angles, a new vegetation index for estimating WUE was established. The new parameter was then compared with common vegetation indexes under different observation angles, and the optimal range of angles was determined, allowing establishment of a unified estimation equation. The findings provide a theoretical basis for real-time accurate monitoring of water use in winter wheat, supporting the screening of germplasm resources and efficient irrigation management.

MATERIALS AND METHODS

Experimental Design

Five experiments were carried out across 4 years at two different locations. Various water management, N rates, and cultivars of hexaploid winter wheat (*Triticum aestivum* L.) were studied, specific details are shown in **Table 1**. Experiments 1–4 were completed in the experimental station of Henan Agricultural University (35°51'N, 113°35'S), Zhengzhou, Henan Province, China, in fluvo-aquic soil. Experiments 1 and 4 were completed in 2016–17, and experiments 2 and 3 in 2017–18 and 2018–19, respectively. Experiment 5 was completed at Shangshui experimental station in Zhoukou, Henan Province (33°33'N, 114°37'S), in 2017–18, in lime concretion black soil. The experiments 1, 4, and 5 consisted of a only one irrigation regiments (twice irrigation, 750 m^{3}ha^{-1} at jointing plus anthesis stage), experiments 2–3 consisted of a three irrigation regiments (no irrigation, single irrigation of 750 m^{3}ha^{-1}, and irrigation750 m^{3}ha^{-1} at jointing plus anthesis stage). Three different winter wheat cultivars were examined, two erect (Yumai 49–198 and Zhoumai 27) and one horizontal (Zhengmai 9694).

TABLE 1 | Seasons, soil status, cultivars, nitrogen rates, irrigation frequency, and sampling dates for five experiments.

Exp. no.	Season, Site, and Cultivar	Soil characteristics	Treatments	Sampling stage
Exp. 1	2016-2017 Zhengzhou Yumai49-198	Type: fluvo-aquic soil, Organic-M: 20.7 g kg^{-1}, Soil pH ($CaCl_2$): 7.9, Total N: 1.9 g kg^{-1},AvailableP: 40.63 mg kg^{-1}, Available K: 116.2 mg kg^{-1}	Irrigated N: N rate (kg ha^{-1}), W_2: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)]. N: 50% prior to seeding and 50% at jointing. Irrigation frequencies: W_2 (twice at jointing and anthesis stage).	Booting Anthesis Mid-filling
Exp. 2	2017-2018 Zhengzhou Yumai49-198	Type: fluvo-aquic soil, Organic-M: 16.8 kg^{-1}, Soil pH ($CaCl_2$): 7.8, Total N: 0.92 g kg^{-1}, Available P: 18.90 mg kg^{-1}, Available K: 152.64 mg kg^{-1}	Water and nitrogen coupling: N rate (kg ha^{-1}), W_0: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)], W_1: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)], W_2: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)]. Irrigation frequencies: W_0(none), W_1(once at jointing stage), W_2 (twice at jointing and anthesis stage).	Booting Heading Anthesis Mid-filling
Exp. 3	2018-2019 Zhengzhou Yumai49-198	Type: fluvo-aquic soil, Organic-M: 16.8 kg^{-1}, Soil pH ($CaCl_2$): 7.8, Total N: 0.92 g kg^{-1}, Available P: 18.90 mg kg^{-1}, Available K: 152.64 mg kg^{-1}	Water and nitrogen coupling: N rate (kg ha^{-1}), W_0: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)], W_1: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)], W_2: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)]. Irrigation frequencies: W_0(none), W_1(once at jointing stage), W_2 (twice at jointing and anthesis stage).	Booting Heading Anthesis Mid-filling
Exp. 4	2016-2017 Zhengzhou Zhengmai9694	Type: fluvo-aquic soil, Organic-M: 16.8 kg^{-1}, Soil pH ($CaCl_2$): 7.8, Total N: 0.92 g kg^{-1}, Available P: 18.90 mg kg^{-1}, Available K: 152.64 mg kg^{-1}	Irrigated N: N rate (kg ha-1), W_2: [N_0(0), N_{12}(120), N_{18}(180), N_{24}(240)]. N: 50% prior to seeding and 50% at jointing. Irrigation frequencies: W_2 (twice at jointing and anthesis stage).	Booting Heading Anthesis Mid-filling
Exp. 5	2017-2018 Shangshui Zhoumai27	Type: lime concretion black soil, Organic-M: kg^{-1}, Soil pH (CaCl2): 7.0, Total N: kg^{-1}, Available P: 4.87 mg kg^{-1}, Available K: 176.52 mg kg^{-1}	Irrigated N: N rate (kg ha-1), W_2: [N_0(0), N_6(60), N_{12}(120), N_{18}(180), N_{24}(240)]. N: 50% prior to seeding and 50% at jointing. Irrigation frequencies: W_2 (twice at jointing and anthesis stage).	Heading Anthesis

Experimental plots 1–4 covered an area of 7 × 2.9 m, respectively, planted in a north-south direction, with 18 cm row spacing, while plot 5 covered an area of 9 × 6 m, planted in a north-south direction, with 20 cm row spacing. All experiments followed a completely randomized block design, and each treatment was repeated three times. All plots were managed according to local standard management practices.

Measurements of Agronomic Indicators

The P_N and Tr of the top leaf were determined in the field using a photosynthetic device (LI-6400 photosynthetic rate system; Li-Cor Inc., United States). Measurements were obtained in the open at a carbon dioxide concentration of approximately 385 μmol 1^{-1}. The built-in light source was set at 1600 μmol m^{-2} s^{-1}. The ratio of P_N and Tr was then used to reflect the WUE. Measurements were taken at the booting, heading, anthesis, initial-filling, and mid-filling stages.

Twenty representative plants from each treatment were then randomly selected and brought back to the laboratory where they were separated into stem and leaf samples. Leaf weight (FW) was recorded before drying the samples in an oven at 105°C for 30 min then to a constant weight at 70°C. The dry mass of the leaves (DW) was then determined and the LWC was calculated as follows:

$$\text{LWC} = (\text{FW-DW})/\text{FW}.$$

Plant leaf samples were simultaneously dried to a constant weight then crushed and passed through a sieve before determining the leaf nitrogen content (LNC) using the Kjeldahl method.

Canopy Spectrum Acquisition

At the same time as measuring the WUE, the spectral reflectance of the winter wheat canopy was also determined. A FieldSpec Pro FR 2500 back-mounted field hyperspectral radiometer (Analytical Spectral Device, American ASD Company) was used to sample 10 points per 1 m^2, which were then averaged as one point of data. Measurements were made on a sunny day with no cloud cover between 10:00 and 12:00 a.m. The field of view of the spectrometer was set at 25°, the spectral range was 350–1075 nm, and the sampling interval was 1.6 nm. Before sampling and during use, a 40 × 40 cm $BaSO_4$ whiteboard was used for calibration. To obtain multi-angle spectrum, a probe was fixed to the multi-angle observation frame according to the design of the field angle measurement system (FIGOS, **Figure 1**). A total of 13 observation angles were examined following the principal plane of the sun, with the sunny side representing backward observation angles (−60, −50, −40, −30, −20, and −10°, respectively), and angles on the opposite side representing forward observation angles (10, 20, 30, 40, 50, and 60°, respectively), with the vertical angle set at 0°.

Data Application

A self-developed computation program was used to optimize the sensitive band combinations and equations using MATLAB 7.0 software. Data from experiments 1–3 were used to construct the new vegetation index and WUE estimation model, while independent data from experiments 4–5 were used to test the model by comparing differences between the coefficient of determination (R^2), root mean square error (RMSE), and relative error (RE, %). A 1:1 scatter plot was then used to show the effect of the model. Some common spectral indices were calculated using the equation listed in **Table 2**. RMSE and RE were calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left(P_i - Q_i\right)^2} \quad (1)$$

$$\text{RE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left(\frac{P_i - Q_i}{Q_i}\right)^2} \times 100 \quad (2)$$

where P_i and Q_i represent the predicted and measured values, respectively, and n represents the number of samples.

RESULTS

Quantitative Relationships Between the Leaf Nitrogen Content, Water Content, and WUE

Based on the data from experiments 1–3, the relationships between the LNC, LWC, and ratio between LNC/LWC under different experimental conditions was analyzed in terms of the WUE (**Figure 2**). As shown in **Figure 2A**, when the irrigation treatment conditions are not distinguished, the relationship between LNC and WUE was generally poor (R^2 = 0.366). Under a single water treatment condition, the LNC showed a significant linear relationship with the WUE, and the correlation was best under W_1 (once water at jointing stage) conditions (R^2 = 0.869), followed by W_0 conditions (R^2 = 0.803). The worst correlation was observed under W_2 (twice water at jointing and anthesis stage) conditions (R^2 = 0.682). Similarly, the relationship between

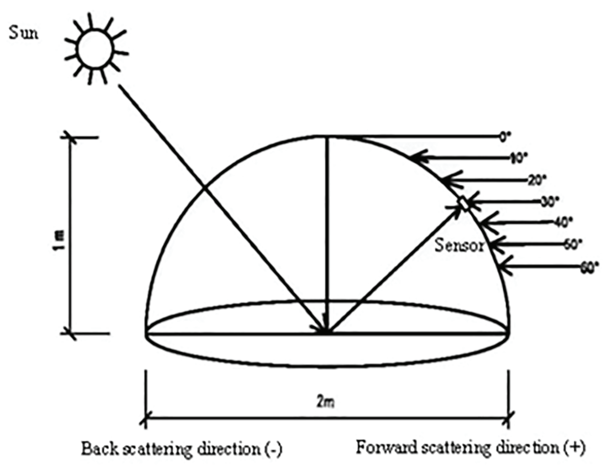


FIGURE 1 | Dimensions and design of the field goniometer system.

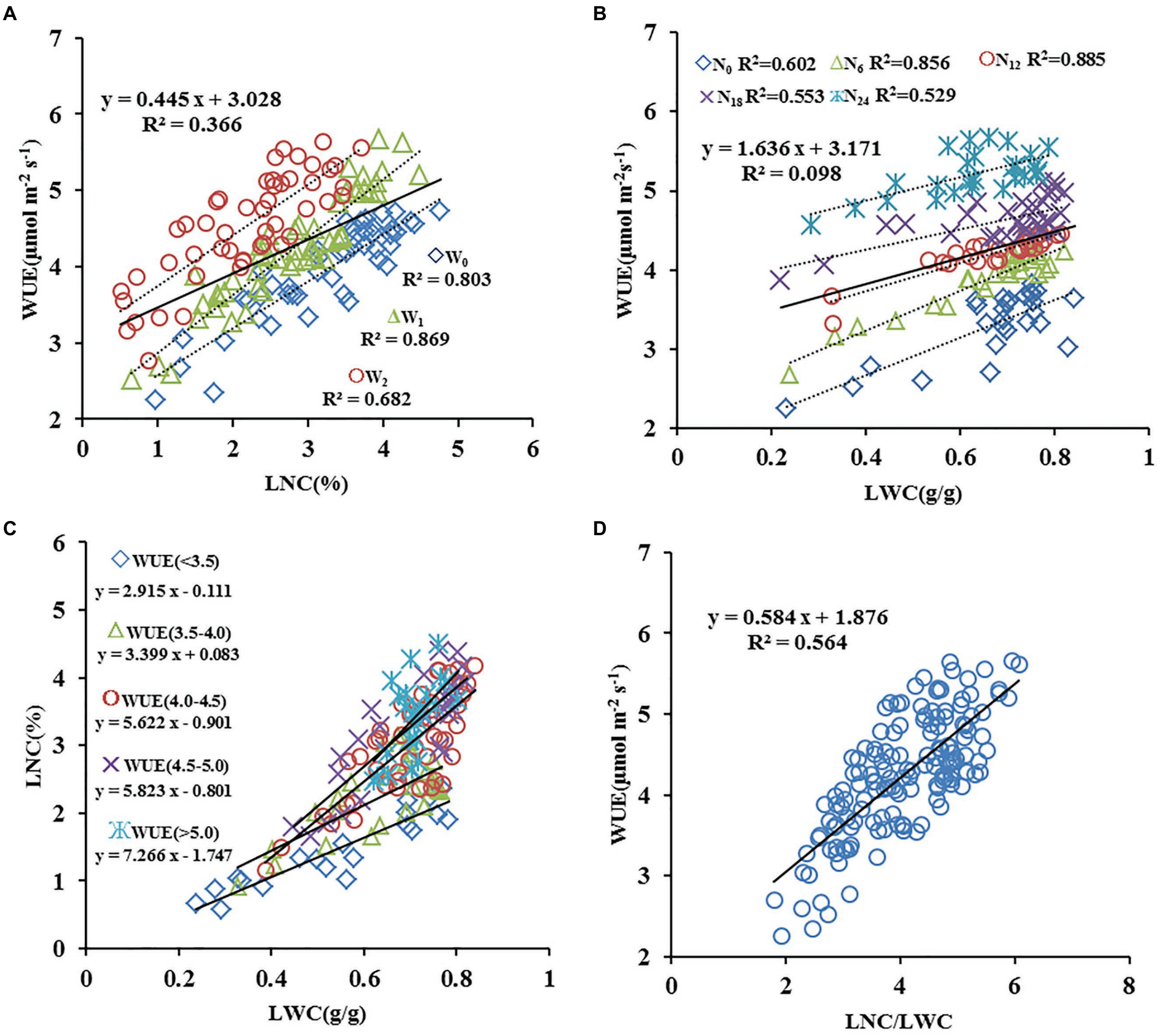


FIGURE 2 | Quantitative relationships between leaf nitrogen (N) content (LNC: **A**), water content (LWC: **B**), LNC/LWC **(D)** and water use efficiency (**C**: The relation between LWC and LNC. WUE; *n* = 140).

the LWC and WUE was also affected by nitrogen treatment. Overall, the R^2 between the LWC and WUE was only 0.098 (**Figure 2B**). Under different nitrogen treatment conditions, the relationship between LWC content and WUE was significant. The transition from high to low nitrogen treatment caused an increase followed by a decrease in R^2, with N_{12} (120 kg ha^{-1}) treatment giving the highest value ($R^2 = 0.885$). The relationships between the LNC and LWC content and WUE were therefore affected by each other, and these relationships were therefore analyzed further. Results revealed that the variation in WUE was closely related to the slope between LNC and LWC. With increasing WUE, the slope of the equation between LNC and LWC gradually increased (**Figure 2C**), and there was a significant positive correlation between the LNC/LWC ratio and WUE ($R^2 = 0.564$, **Figure 2D**). These findings suggest that the LNC/LWC ratio more accurately reflects the dynamic changes in WUE under different water and nitrogen conditions.

Relationship Between Common Spectral Parameters and WUE at the Vertical Observation Angle

The relationships between 330 previously reported vegetation indices and WUE were subsequently analyzed, then the best 12 were selected (**Figure 3**). As shown in the figure, the R^2 between WUE and only three of these parameters was greater than 0.4 (R^2 of Lo, EVI-1 and NDDAig: 0.439, 0.523, and 0.545; RMSE: 0.539, 0.511, and 0.500, respectively). To further

TABLE 2 | Summary of selected spectral parameters reported in the literature.

Vegetation indices	Formula	Reference
DVI(810,680)	$R_{810}-R_{680}$	Jordan, 1969
SRPI	R_{430}/R_{680}	Peñuelas et al., 1995a
WBI-1	R_{950}/R_{900}	Peñuelas et al., 1993a
WI	R_{900}/R_{970}	Peñuelas et al., 1997
Readone	R_{415}/R_{695}	Read et al., 2002
Lo	$\min(R_{680-780})$	Miller et al., 1990
PSRI	$(R_{680}-R_{500})/R_{750}$	Merzlyak et al., 1999
$R_{434}/(R_{496}+R_{401})$	$R_{434}/(R_{496}+R_{401})$	Tian et al., 2011
$R_{705}/(R_{717}+R_{491})$	$R_{705}/(R_{717}+R_{491})$	Tian et al., 2011
FWBI	$R_{900}/\min(R_{930-980})$	Strachan et al., 2002
PRI(570, 531)	$(R_{531}-R_{570})/(R_{531}+R_{570})$	Gamon et al., 1992
SIPI(800, 680, 445)	$(R_{800}-R_{445})/(R_{800}-R_{680})$	Peñuelas et al., 1995a
mSR705	$(R_{750}-R_{445})/(R_{705}-R_{445})$	Sims and Gamon, 2002
RES	$(R_{718}-R_{675})/(R_{755}-R_{675})$	Ju et al., 2010
NDVI(895, 675)	$(R_{895}-R_{675})/(R_{895}+R_{675})$	Santos and Negri, 1997
NDRE	$(R_{790}-R_{720})/(R_{790}+R_{720})$	Fitzgerald et al., 2006
NRI(570, 670)	$(R_{570}-R_{670})/(R_{570}+R_{670})$	Li et al., 2005
RDVI(800, 670)	$(R_{800}-R_{670})/\mathrm{sqrt}(R_{800}+R_{670})$	Roujean and Breon, 1995
NDDAig	$(R_{755}+R_{680}-2\times R_{705})/(R_{755}-R_{680})$	Feng et al., 2014
NDGI	$[R(_{520-560})-R(_{630-690})]/[R(_{520-560})+R(_{630-690})]$	Rouse et al., 1974
EVI-1	$2.5*(R_{860}-R_{645})/(1+R_{860}+6*R_{645}-7.5*R_{470})$	Huete et al., 2002
MCARI(700, 670, 550)	$[(R_{700}-R_{670})-0.2*(R_{700}-R_{550})]*(R_{700}/R_{670})$	Daughtry et al., 2000
Vari-GREEN	$(R_{520-560}-R_{630-690})/(R_{520-560}+R_{630-690}-R_{430-470})$	Gitelson et al., 2002
TSAVI(800, 670)	$1.4735*(R_{780}+1.4735*R_{650}-1.3681)/(-1.4735*R_{780}+R_{650}+1.4735*1.3681)$	Baret and Guyot, 1991
WEI	$[(R_{755}+R_{680}-2*R_{705})*\mathrm{Min}(R_{930-980})]/[(R_{755}-R_{680})*R_{900}]$	This study

improve the estimation accuracy using the relationship between LNC/LWC and WUE, LNC and LWC were converted using related vegetation indexes then the relationship with WUE was analyzed further. Ten indexes representing the LNC and seven representing the LWC were combined to create different ratios of the two vegetation indexes then the correlations with WUE were determined at a vertical angle (**Figure 4**). Ten combinations showed a R^2 greater than 0.40, and of these, three had an R^2 exceeding 0.50 and were considered optimal (NDDAig/FWBI, NDDAig/WBI-1, and SPRI/WBI-1). Overall, the combination of NDDAig/FWBI performed the best (R^2 = 0.624).

The quantitative relationships between the two common spectral parameters Lo and NDDAig and the optimized novel ratio vegetation index (NDDAig/FWBI) and WUE are shown as a scatter plot in **Figure 5**. A significant linear correlation was revealed between Lo and WUE (R^2 = 0.523 and RMSE = 0.501, **Figure 5A**), with an obvious saturation phenomenon. Compared with Lo, the NDDAig model showed improvement (R^2 = 0.543, RMSE = 0.486, **Figure 5B**), although the combined NDDAig/FWBI index gave the best results (R^2 = 0.624, 19.31 and 14.92% higher than that of Lo and NDDAig, respectively; RMSE = 0.441, 13.61, and 10.20% lower than that of Lo and NDDAig, respectively), with obvious weakening of the saturation phenomenon. Based on highest R^2 values, the new ratio vegetation index (NDDAig/FWBI) was constructed to generate a water efficiency index (WEI).

Relationships Between the Spectral Parameters and WUE at Different Zenith Angles

Based on the data from experiments1–3, the relationships between the 12 common vegetation indexes and new combined index WEI were analyzed in terms of WUE under different observation angles (**Table 3**). Overall, except for angles of −60 and −50°, all spectral parameters had a higher backward R^2 than forward R^2, especially in the range of −40–30°. Moreover, the optimal observation angles of the different vegetation indexes were inconsistent with the monitoring accuracy. Two of the 13 spectral parameters (GVI and TC_2) had an optimal observation angle of 0° with R^2 values of 0.375 and 0.365, respectively, while five spectral parameters [TSAVI (800, 670), RDVI (800, 670), PSRI, SIPI (800, 680, 445), and EVI-1] had an optimal angle of −20° with R^2 values of 0.424, 0.406, 0.407, 0.422, and 0.472, respectively. It is worth noting that the optimal observation angle of seven of the indexes was −10°, suggesting that observation angles of −10 to −20° are important when monitoring the WUE of winter wheat leaves.

Compared with the common spectral parameters, the new vegetation index WEI showed obvious advantages at specific observation angles, especially an angle of −10°. The R^2 and RMSE of the WEI model at different zenith angles are shown in **Figure 6**. The R^2 was highest in an angle range of −20–10° and the RMSE was relatively low.

Relationship Between the New Spectral Parameter and WUE Under Different Angle Ranges

From a single observation angle point of view, the monitoring accuracy of the new combined index was highest within a range of −20–10°, with highest precision at −10° (R^2 and RMSE: 0.635 and 0.441, respectively). By combining the data from different observation angles, equation fitting was further analyzed under five observation angle ranges according to the principle of adjacent observation angles. As shown in **Figure 7**, the monitoring accuracy of WEI was higher than that of Lo and NDDAig, and the RMSE value was lowest under different angle ranges. Compared with an observation angle of −10°, the R^2 of WEI decreased by 6.72% within a range of −20–20°, while the RMSE increased by 7.76% (**Figure 8A**). Meanwhile, the R^2 decreased by only 1.93% in a range of −20–10° and the RMSE increased by only 4.71% (**Figure 8B**). These findings suggest that within an angle range of −20–10°, the WEI model reduces the dependency on the observation angle, increasing the applicability and stability of the model.

Testing of the Estimation Model

The WUE estimation models were subsequently tested with the independent test data obtained in experiments 4–5 using

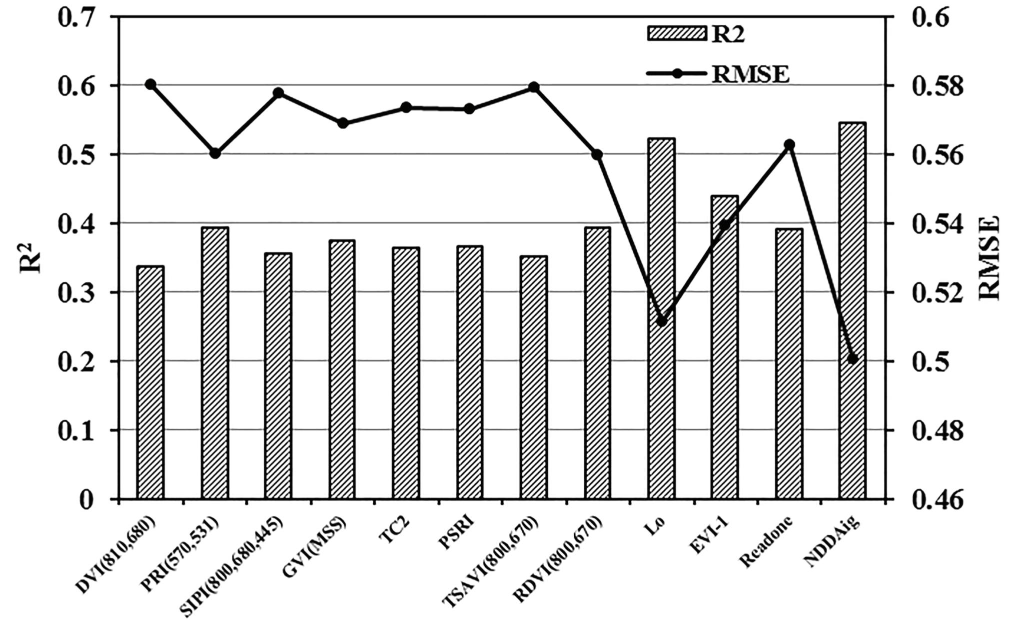


FIGURE 3 | Relationships between common vegetation indices and WUE (*n* = 140).

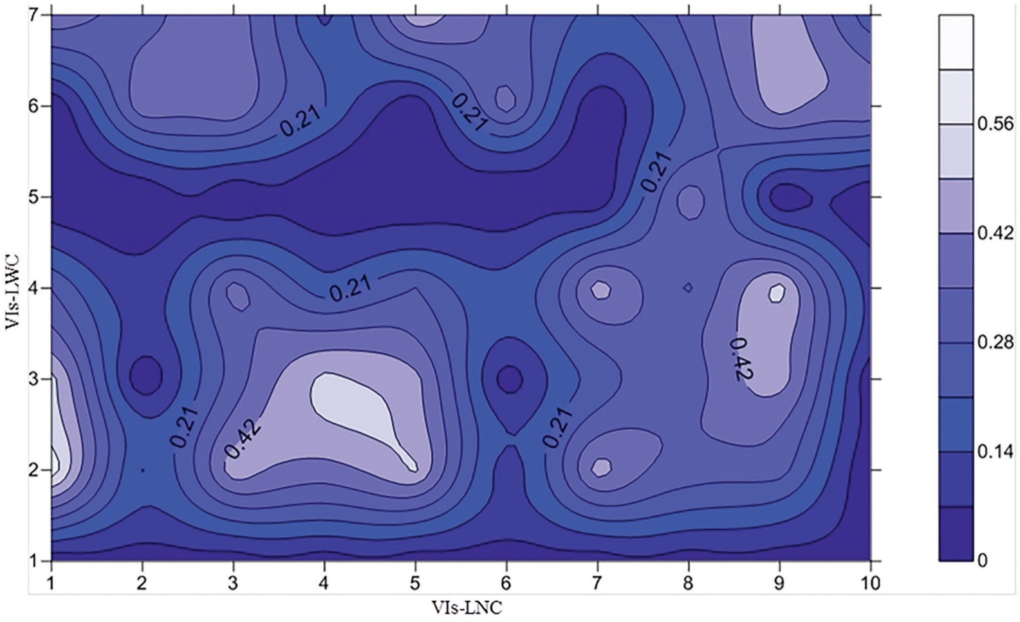


FIGURE 4 | Correlations between different parameter ratios and the WUE. [X-axis: VIs-LNC (vegetation indices related to LNC) 1–10 represent NDDAigig, $R_{434}/(R_{496} + R_{401})$, $R_{705}/(R_{717} + R_{491})$, SRPI, NDRE, Lo, NRI, NDGI, RES, MCART (700, 670, and 550), respectively; Y-axis: VIs-LWC (vegetation indices related to LNC) 1–7 represent PRI (570 and 531), FWBI, WBI-1, WI, Vari-GREEN, mSR705, NDVI (895 and 675), respectively; *n* = 140].

three indicators (R^2, RMSE, and RE). The prediction ability is shown as the ratio between predicted and observed values in **Figure 9**. The WUE model using NDDAig (R^2 = 0.602, RMSE = 0.474, and RE = 14.191%) as a variable was better than that using Lo (0.517, 0.523, and 18.012%, respectively); however, new parameter WEI gave better predictions, with

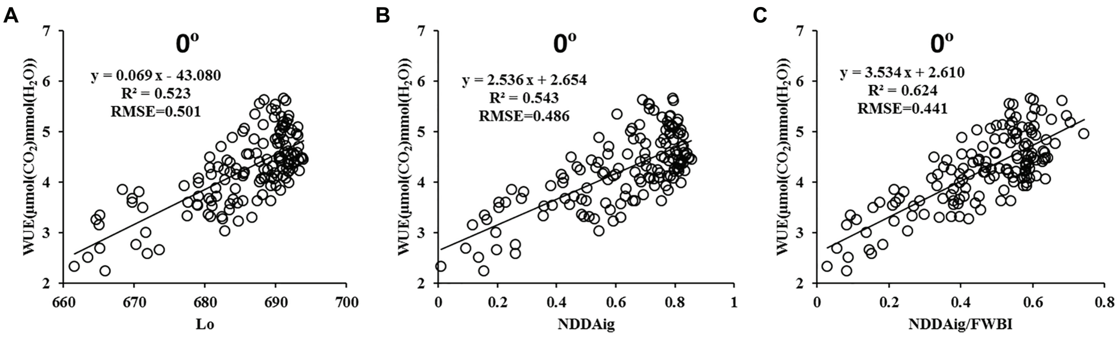


FIGURE 5 | Relationships between Lo **(A)**, NDDAig **(B)**, NDDAig/FWBI **(C)**, and the WUE at a 0° zenith angle (*n* = 140).

R^2 = 0.685, RMSE = 0.403, and RE = 11.147%. Overall, these findings suggest that the new combined index of WEI can be used to accurately monitor the WUE of winter wheat leaves.

DISCUSSION

Water and nutrients are not only the main stress factors affecting agricultural production, but they also interact with each other, playing individual as well as complimentary roles. The main aim of agricultural management, therefore, is to maximize the coupling effect of water and nitrogen, adjust water management according to nitrogen absorption, and use water management to promote nitrogen absorption (Sheshbahreh et al., 2019). Water stress significantly affects nitrogen absorption, while improved soil water conditions benefit nitrogen absorption and utilization. Increased uptake of nutrients under drought stress can also help improve drought resistance, while optimal increases in nitrogen fertilizer application can improve WUE and increase yield (Wolfe et al., 1988). As a result of the synergetic relationship between soil water and nitrogen, the changes in crop nitrogen and water contents are also synchronized. However, this relationship is also affected by irrigation and nitrogen fertilizer treatment. Meanwhile, this study confirmed that the relationship between the LNC and WUE is also affected by irrigation treatment, while at the same time, the relationship between the LWC and WUE is also affected by nitrogen fertilizer treatment. Thus, the use of LNC and LWC alone was relatively unreliable in characterizing the WUE. In contrast, the relationship between the ratio of LNC/LWC in terms of WUE was relatively less affected by irrigation and nitrogen fertilizer treatment. The coefficient of determination of the fitting equation was 0.564, suggesting that the LNC/LWC ratio is a good indicator of dynamic changes in WUE. Overall, the findings confirmed that an increase in nitrogen in line with an increase in the water content of crop leaves is beneficial to overall water absorption.

The WUE of a plant is genetically controlled as well as being affected by the environment, and can therefore be improved by both breeding and cultivation measures. A high WUE is beneficial in maintaining a certain yield under water stress, and therefore has important application value. In addition, WUE plays a significant role in estimations of net primary productivity (NPP = WUE × Tr) on a regional scale. However, recent studies have shown that the WUE is not constant, but rather it varies greatly with environmental conditions and the plant type (Yu et al., 2001). The use of remote sensing to rapidly and non-destructively determine the real-time WUE of a crop therefore provides important information for terrestrial ecosystem and water cycle models at different scales. The crop canopy spectrum provides mixed information, and is susceptible to factors such as plant coverage, soil type, and leaf area. Accordingly, a number of studies have aimed to construct new indexes that reduce noise and improve the estimation accuracy (Wang et al., 2012). For example, Peñuelas et al. (1997) and Pinol et al. (1998) successfully determined the humidity of combustibles using spectral data obtained in the field in the Mediterranean using a combined ratio spectral index (Peñuelas et al., 1997; Pinol et al., 1998). Studies have also shown that double-ratio vegetation indexes can reduce the effects of variables such as background and leaf area index (LAI), and provide more useful information for estimations of the vegetation canopy water content (Daughtry et al., 2000; Haboudane et al., 2002). Meanwhile, double-ratio vegetation indexes also include more sensitive bands, thus improving the estimation accuracy (Gitelson et al., 2017). Based on these studies, we therefore combined the close relationship between LNC/LWC and WUE to obtain a new spectral parameter, one indicate the change in nitrogen content of the leaf is selected, another parameter that is sensitive to the change in LWC is selected, the combination of these two vegetation indexes in the form of ratio provides an opportunity to invert WUE of leaves. To this end, we selected two spectral parameters of nitrogen content (NDDAig) and water content (FWBI), and combined the two (NDDAig/FWBI) to give a new index, WEI. Accordingly, estimations of the WUE of winter wheat leaves were greatly improved. Compared with common spectral parameters, the new WEI performed best at 13 angles, with optimal angle compatibility within the range of −20–10° (R^2 = 0.623). The use of independent test data also confirmed the accuracy of the model.

TABLE 3 | Coefficients of determination (R^2) of the linear relationships between leaf water use efficiency and the vegetation indices from different view zenith angles.

	−60°	**−50°**	**−40°**	**−30°**	**−20°**	**−10°**	**0°**	**10°**	**20°**	**30°**	**40°**	**50°**	**60°**
DVI(810,680)	0.246	0.322	0.336	0.316	0.342	0.351	0.337	0.334	0.306	0.285	0.263	0.237	0.237
PRI(570,531)	0.265	0.326	0.351	0.392	0.403	0.405	0.394	0.372	0.342	0.339	0.323	0.325	0.298
SIPI(800,680,445)	0.278	0.346	0.391	0.41	0.422	0.394	0.357	0.330	0.320	0.320	0.305	0.301	0.233
GVI(MSS)	0.244	0.324	0.344	0.317	0.346	0.359	0.375	0.355	0.315	0.287	0.263	0.228	0.231
TC_2	0.227	0.304	0.324	0.296	0.328	0.343	0.365	0.348	0.305	0.277	0.253	0.216	0.218
PSRI	0.251	0.311	0.346	0.392	0.407	0.393	0.366	0.346	0.321	0.314	0.293	0.283	0.229
TSAVI(800,670)	0.295	0.404	0.438	0.435	0.424	0.392	0.353	0.323	0.308	0.302	0.299	0.309	0.285
RDVI(800,670)	0.310	0.397	0.411	0.401	0.406	0.403	0.395	0.373	0.348	0.33	0.315	0.301	0.299
Lo	0.398	0.461	0.491	0.492	0.506	0.523	0.523	0.509	0.494	0.492	0.492	0.471	0.442
EVI-1	0.368	0.437	0.469	0.48	0.472	0.455	0.439	0.418	0.414	0.415	0.424	0.435	0.333
Readone	0.221	0.288	0.331	0.373	0.397	0.416	0.39	0.399	0.398	0.402	0.431	0.368	0.217
NDDAig	0.459	0.507	0.521	0.526	0.541	0.548	0.545	0.535	0.532	0.522	0.514	0.51	0.497
WEI	0.485	0.551	0.587	0.594	0.616	0.635	0.624	0.613	0.596	0.582	0.578	0.572	0.554
Average	0.295	0.366	0.392	0.398	0.411	0.416	0.407	0.387	0.367	0.356	0.348	0.331	0.298

The difference between a vegetation index in different observation directions depends on various factors, including the crop canopy structure, the shape, and angle of the sensor, shadows, and the soil type (Kimes et al., 1985). In this study, the relationship between WEI and WUE also varied depending on the angle of observation. Pocewicz et al. (2007) made full use of the hotspot effect of the backward observation angles to improve the estimation accuracy of LAI (Pocewicz et al., 2007). Meanwhile, a multi-angle observation method was used to collect image data of the soybean canopy at different growth stages in the field, revealing that a 40° zenith angle was the best angle for inverting chlorophyll density (Zhang et al., 2013). In this study, the R^2 between the spectral parameters and WUE decreased with increasing observation angle, possibly because a smaller angle results in more comprehensive information of the upper, middle, and lower canopies of the entire crop. The absorption of water by crop leaves is the result of interactions throughout the canopy; therefore, small-angle spectral information is important in determining an accurate WUE. In addition, the effects of backward observation angles were found to be better than those of forward observations. This is thought to be because when data is collected in a backward direction, the sensor is located on the same side as the sun, and data is mainly collected from canopy falling within the light. In contrast, the crop canopy with a larger shadow share is collected in a forward direction, although angles of −60° and −50° behave abnormally, possibly due to the decline in data quality at larger angles (Barnes and Hu, 2016).

Compared with the vertical angle, the wider angle range not only resulted in more information on the crop canopy, but also expanded the application range of the sensor, increasing overall efficiency. In addition to determining the best observation angle, it is also important to comprehensively model data from different angles to increase application accuracy (Guo et al., 2018; He et al., 2018). The WEI constructed in this study provided high estimation accuracy (R^2 = 0.623) within a range of −20–10°, and compared with the optimal angle (−10°), the estimation accuracy of WEI within a range of −20–10° decreased by only 1.93%, while the RMSE value increased by only 4.71%. These results suggest that the WEI reduces the sensitivity to observation angles within the range of −20–10°, helping establish a more unified monitoring model, and increasing the efficiency and applicability of portable monitors in the field. However, despite these findings, this experiment was carried out using only three winter wheat varieties under two ecological conditions, and therefore, further analysis of the monitoring model in other regions, and with different crop types and varieties is required.

CONCLUSION

Real-time monitoring of crop water use is of great significance in improving crop irrigation management and guiding water-saving agricultural production. Based on the WUE, this study adopted

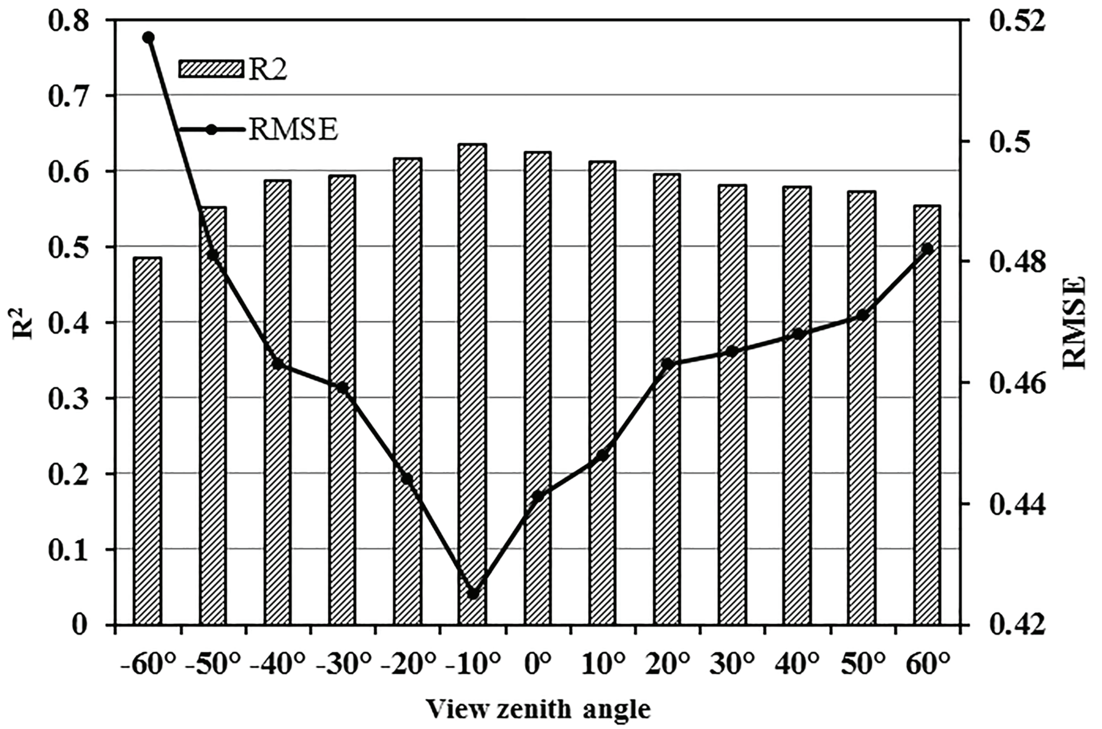


FIGURE 6 | Correlations between water efficiency index (WEI) and WUE at different zenith angles (*n* = 140).

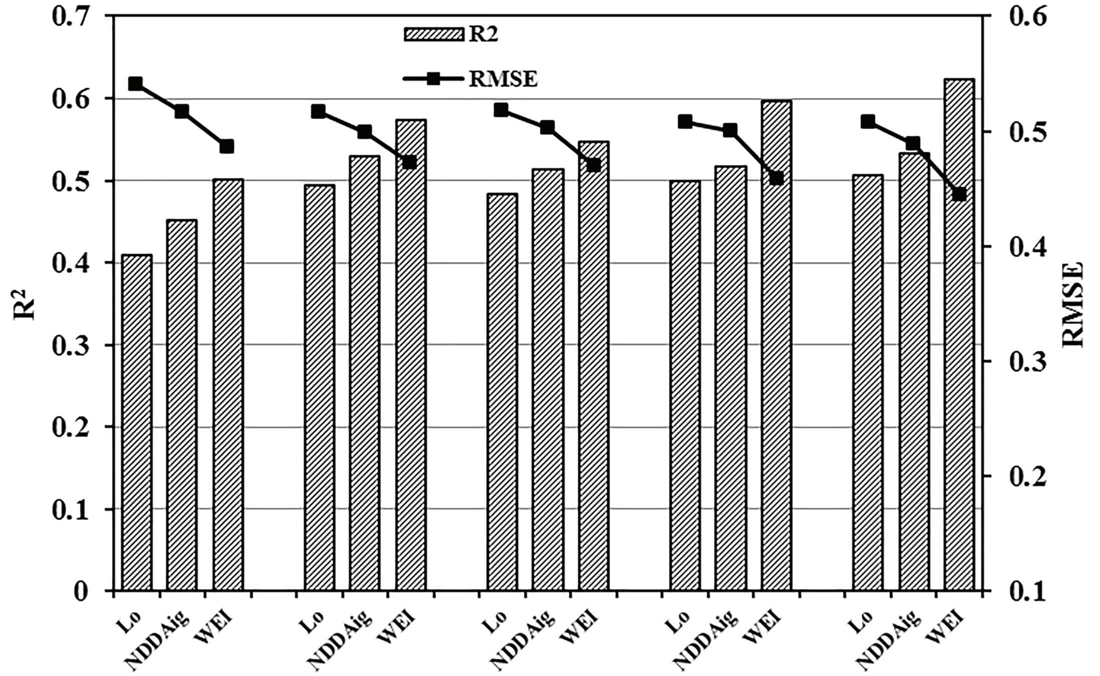


FIGURE 7 | Comparisons of the predictive abilities of Lo, NDDAig, and WEI within five zenith angle ranges (−60–60°, −60–0°, 0–60°, −20–20°, and −20–10°) with respect to WUE (*n* = 140).

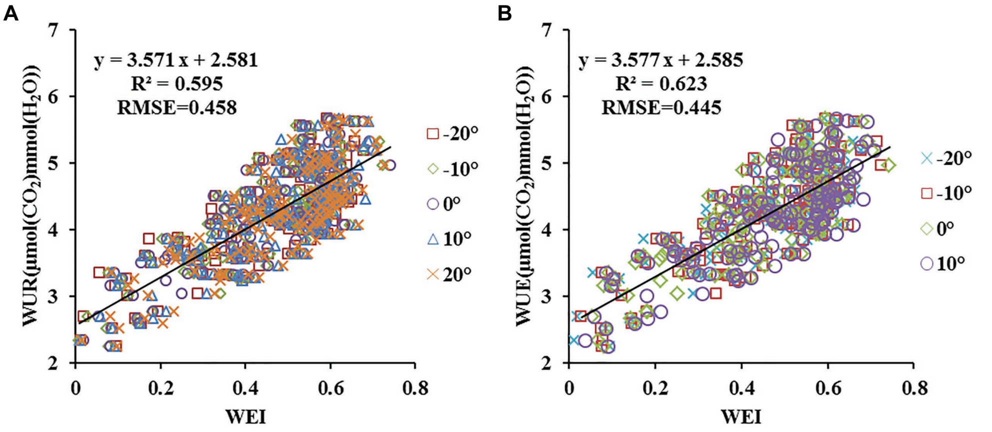


FIGURE 8 | Comparisons of the predictive power of WEI at different view zenith angles (VZAs) combinations in terms of WUE (**A**: n = 700; **B**: n = 560).

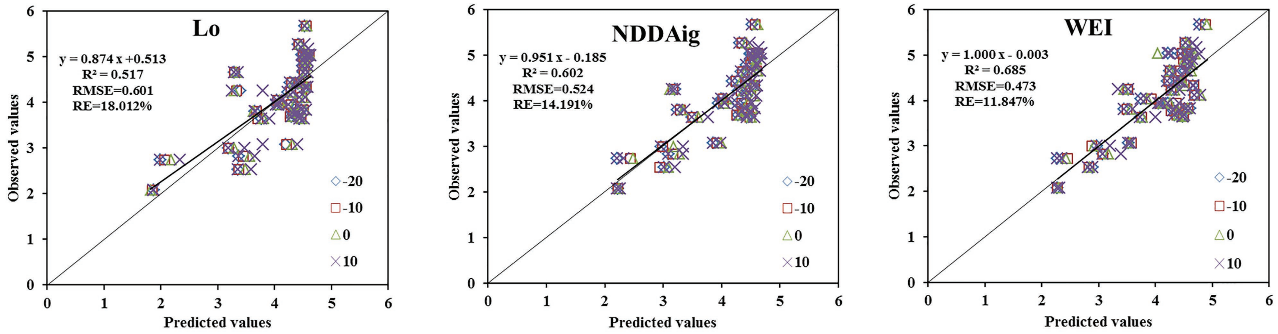


FIGURE 9 | Comparisons between predicted and measured WUE based on Lo, NDDAig, and WEI at a zenith angle of −20° to +10° (n = 120).

a combined vegetation index method, using NDDAig/FWBI to effectively determine the WUE of winter wheat leaves reflecting a new index named the WEI. The WEI performed better than other common vegetation indexes at 13 observation angles, with the most suitable observation range falling between −20 and 10°. Within this range, a unified estimation model was established with reduced dependency on observation angle limitations. These findings provide a basis for the selection of varieties with a high WUE as well as supporting water-saving cultivation management.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

H-YZ, C-YW and WF conceived the research concept; H-YZ, M-RL, LS and XL performed the experiments; H-YZ, C-YW wrote the paper; Z-HF and W-DL contributed to the results analysis and discussion.

FUNDING

This work was supported by the National Key Research and Development Program of China (Grant No. 2018YFD0300707) and the Fund for Modern Agro-industry System of Henan province (S2010-01-G07). National Natural Science Foundation of China (31671624).

REFERENCES

Baret, F., and Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173. doi: 10.1016/0034-4257(91)90009-U

Barnes, B. B., and Hu, C. (2016). Dependence of satellite ocean color data products on viewing angles: a comparison between SeaWiFS, MODIS, and VIIRS. *Remote Sens. Environ.* 175, 120–129. doi: 10.1016/j.rse.2015.12.048

Bastiaanssen, W. G. M., Thiruvengadachari, S., Sakthivadivel, R., and Molden, D. J. (1999). Satellite remote sensing for estimating productivities of land and water. *Int. J. Water Resour. Dev.* 15, 181–194. doi: 10.1080/07900629949005

Carter, G. A. (1991). Primary and secondary effects of water content on the spectral reflectance of leaves. *Am. J. Bot.* 78, 916–924. doi: 10.1002/j.1537-2197.1991.tb14495.x

Chen, J. M., Liu, J., Leblanc, S. G., Lacaze, R., and Roujean, J. L. (2003). Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sens. Environ.* 84, 516–525. doi: 10.1016/S0034-4257(02)00150-5

Chen, J. M., Menges, C. H., and Leblanc, S. G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sens. Environ.* 97, 447–457. doi: 10.1016/j.rse.2005.05.003

Chopping, M. J., Rango, A., Havstad, K. M., Schiebe, F. R., Ritchie, J. C., Schmugge, T. J., et al. (2003). Canopy attributes of desert grassland and transition communities derived from multiangular airborne imagery. *Remote Sens. Environ.* 85, 339–354. doi: 10.1016/S0034-4257(03)00012-9

Cierniewski, J., Gdala, T., and Karnieli, A. (2004). A hemispherical–directional reflectance model as a tool for understanding image distinctions between cultivated and uncultivated bare surfaces. *Remote Sens. Environ.* 90, 505–523. doi: 10.1016/j.rse.2004.01.004

Condon, A. G., Richards, R. A., Rebetzke, G. J., and Farquhar, G. D. (2002). Improving intrinsic water-use efficiency and crop yield. *Crop Sci.* 42, 122–131. doi: 10.2135/cropsci2002.1220

Curran, P. J. J. R. S. o. E. (1989). Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30, 271–278. doi: 10.1016/0034-4257(89)90069-2

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., and Mcmurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239. doi: 10.1016/S0034-4257(00)00113-9

Dawson, T. P., Curran, P. J., North, P. R. J., and Plummer, S. E. (1999). The propagation of foliar biochemical absorption features in forest canopy reflectance: a theoretical analysis. *Remote Sens. Environ.* 67, 147–159. doi: 10.1016/S0034-4257(98)00081-9

Diner, D. J., Asner, G. P., Davies, R., Knyazikhin, Y., Muller, J. P., Nolin, A. W., et al. (1999). New directions in earth observing: scientific applications of multiangle remote sensing. *B. Am. Meteorol. Soc.* 80, 2209–2228. doi: 10.1175/1520-0477(1999)080<2209:ndieos>2.0.co;2

Dobrowski, S. Z., Pushnik, J. C., Zarco-Tejada, P. J., and Ustin, S. L. (2005). Simple reflectance indices track heat and water stress-induced changes in steady-state chlorophyll fluorescence at the canopy scale. *Remote Sens. Environ.* 97, 403–414. doi: 10.1016/j.rse.2005.05.006

Dong, B., Shi, L., Shi, C. h., Qiao, Y. Z., Liu, M. Y., and Zhang, Z. B. (2011). Grain yield and water use efficiency of two types of winter wheat cultivars under different water regimes. *Agr. Water Manage.* 99, 103–110. doi: 10.1016/j.agwat.2011.07.013

Drolet, G. G., Middleton, E. M., Huemmrich, K. F., Hall, F. G., Amiro, B. D., Barr, A. G., et al. (2008). Regional mapping of gross light-use efficiency using MODIS spectral indices. *Remote Sens. Environ.* 112, 3064–3078. doi: 10.1016/j.rse.2008.03.002

Fitzgerald, G. J., Rodriguez, D., Christensen, L. K., Belford, R., Sadras, V. O., and Clarke, T. R. (2006). Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments. *Precis. Agric.* 7, 233–248. doi: 10.1007/s11119-006-9011-z

Feng, W., Guo, B.-B., Wang, Z.-J., He, L., Song, X., Wang, Y.-H., et al. (2014). Measuring leaf nitrogen concentration in winter wheat using double-peak spectral reflection remote sensing data. *Field Crop. Res.* 159, 43–52. doi: 10.1016/j.fcr.2014.01.010

Galvão, L. S., Roberts, D. A., Formaggio, A. R., Numata, I., and Breunig, F. M. (2009). View angle effects on the discrimination of soybean varieties and on the relationships between vegetation indices and yield using off-nadir Hyperion data. *Remote Sens. Environ.* 113, 846–856. doi: 10.1016/j.rse.2008.12.010

Gamon, J. A., Peñuelas, J., and Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* 41, 35–44. doi: 10.1016/0034-4257(92)90059-S

Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., and Filella, I. (2011). The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: a review and meta-analysis. *Remote Sens. Environ.* 115, 281–297. doi: 10.1016/j.rse.2010.08.023

Gitelson, A. A., Gamon, J. A., and Solovchenko, A. (2017). Multiple drivers of seasonal change in PRI: implications for photosynthesis 2. Stand level. *Remote Sens. Environ.* 190, 198–206. doi: 10.1016/j.rse.2016.12.015

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9

Guo, B. B., Zhu, Y. J., Feng, W., He, L., Wu, Y. P., Zhou, Y., et al. (2018). Remotely estimating aerial N uptake in winter wheat using red-edge area index from multi-angular hyperspectral data. *Front. Plant Sci.* 9:675. doi: 10.3389/fpls.2018.00675

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., and Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. doi: 10.1016/S0034-4257(02)00018-4

Hall, F. G., Hilker, T., Coops, N. C., Lyapustin, A., Huemmrich, K. F., Middleton, E. M., et al. (2008). Multi-angle remote sensing of forest light use efficiency by observing PRI variation with canopy shadow fraction. *Remote Sens. Environ.* 112, 3201–3211. doi: 10.1016/j.rse.2008.03.015

Hatfield, J. L., Gitelson, A. A., Schepers, J. S., and Walthall, C. L. (2008). Application of spectral remote sensing for agronomic decisions. *Agron. J.* 100, 117–131. doi: 10.2134/agronj2006.0370c

He, L., Coburn, C. A., Wang, Z. J., Feng, W., and Guo, T. C. (2018). Reduced prediction saturation and view effects for estimating the leaf area index of winter wheat. *IEEE. Trans. Geosci. Remote Sens.* 57, 1637–1652. doi: 10.1109/TGRS.2018.2868138

He, L., Song, X., Feng, W., Guo, B. B., Zhang, Y. S., Wang, Y. H., et al. (2016). Improved remote sensing of leaf nitrogen concentration in winter wheat using multi-angular hyperspectral data. *Remote Sens. Environ.* 174, 122–133. doi: 10.1016/j.rse.2015.12.007

He, L., Zhang, H. Y., Zhang, Y. S., Song, X., and Guo, T. C. (2015). Estimating canopy leaf nitrogen concentration in winter wheat based on multi-angular hyperspectral remote sensing. *Eur. J. Agron.* 73, 170–185. doi: 10.1016/j.eja.2015.11.017

Huang, W. J., Wang, Z. J., Huang, L. S., Lamb, D. W., Ma, Z. H., Zhang, J. C., et al. (2011). Estimation of vertical distribution of chlorophyll concentration by bi-directional canopy reflectance spectra in winter wheat. *Precis. Agric.* 12, 165–178. doi: 10.1007/s11119-010-9166-5

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

Hultine, K. R., and Marshall, J. D. (2000). Altitude trends in conifer leaf morphology and stable carbon isotope composition. *Oecologia* 123, 32–40. doi: 10.1007/s004420050986

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecolo. Soc. Am.* 50, 663–666. doi: 10.2307/1936256

Ju, C. H., Tian, Y. C., Yao, X., Cao, W. X., Zhu, Y., and Hannaway, D. J. P. (2010). Estimating leaf chlorophyll content using red edge parameters. *Pedosphere* 20, 633–644. doi: 10.1016/S1002-0160(10)60053-7

Kimes, D. S., Newcomb, W. W., Tucker, C. J., Zonneveld, I. S., Wijngaarden, V. W., Leeuw, D. J., et al. (1985). Directional reflectance factor distribution for cover types of northern Africa in NOAA 7/8 AVHRR bands 1 and 2. *Remote Sens. Environ.* 18, 1–19. doi: 10.1016/0034-4257(85)90034-3

Leblanc, S. G., Chen, J. M., White, H. P., Latifovic, R., Lacaze, R., and Roujean, J. L. (2005). Canada-wide foliage clumping index mapping from multiangular POLDER measurements. *Can. J. Remote. Sens.* 31, 364–376. doi: 10.5589/m05-020

Li, F., Liu, L. Y., Wang, J. H., Li, L. X., Zhao, C. J., and Cao, W. X. (2005) Detection of nitrogen status in FCV tobacco leaves with the spectral reflectance. IEEE International Geoscience & Remote Sensing Symposium. 1863–1866.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106, 135–141. doi: 10.1034/j.1399-3054.1999.106119.x

Middleton, E. M., Huemmrich, K. F., Cheng, Y.-B., and Margolis, H. A. (2011). "Spectral bioindicators of photosynthetic efficiency and vegetation stress" in *Hyperspectral remote sensing of vegetation.* eds. P. S. Thenkabail, J. G. Lyon, A. Huete (Boca Raton, FL, USA: CRC Press), 265–288.

Miller, J. R., Hare, E. W., and Wu, J. (1990). Quantitative characterization of the vegetation red edge reflectance 1. An inverted-gaussian reflectance model. *Int. J. Remote Sens.* 11, 1755–1773. doi: 10.1080/01431169008955128

Mistele, B., and Schmidhalter, U. (2008). Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. *Eur. J. Agron.* 29, 184–190. doi: 10.1016/j.eja.2008.05.007

Peñuelas, J., Baret, F., and Filella, I. (1995a). Semiempirical indexes to assess carotenoids chlorophyll-a ratio from leaf spectral reflectance. *Photosynthetica* 1995, 221–230. doi: 10.1007/BF00029464

Peñuelas, J., Filella, I., Biel, C., Serrano, L., and Save, R. (1993a). The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* 14, 1887–1905. doi: 10.1080/01431169308954010

Peñuelas, J., Filella, I., Lloret, P., Oz, F. M., and Vilajeliu, M. (1995b). Reflectance assessment of mite effects on apple trees. *Int. J. Remote Sens.* 16, 2727–2733. doi: 10.1080/01431169508954588

Peñuelas, J., Gamon, J. A., Griffin, K. L., and Field, C. B. (1993b). Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *Remote Sens. Environ.* 46, 110–118. doi: 10.1016/0034-4257(93)90088-F

Peñuelas, J., Pinol, J., Ogaya, R., and Filella, I. (1997). Estimation of plant water concentration by the reflectance water index WI (R_{900}/R_{970}). *Int. J. Remote Sens.* 18, 2869–2875. doi: 10.1080/014311697217396

Pinol, J., Filella, I., Ogaya, R., and Penuelas, J. (1998). Ground-based spectroradiometric estimation of live fine fuel moisture of Mediterranean plants. *Agric. For. Meteorol.* 90, 173–186. doi: 10.1016/S0168-1923(98)00053-7

Pocewicz, A., Vierling, L. A., Lentile, L. B., and Smith, R. (2007). View angle effects on relationships between MISR vegetation indices and leaf area index in a recently burned ponderosa pine forest. *Remote Sens. Environ.* 107, 322–333. doi: 10.1016/j.rse.2006.06.019

Rautiainen, M., Stenberg, P., Nilson, T., and Kuusk, A. (2004). The effect of crown shape on the reflectance of coniferous stands. *Remote Sens. Environ.* 89, 41–52. doi: 10.1016/j.rse.2003.10.001

Read, J. J., Tarpley, L., Mckinion, J. M., and Reddy, K. R. (2002). Narrow-waveband reflectance ratios for remote estimation of nitrogen status in cotton. *J. Environ. Qual.* 31, 1442–1452. doi: 10.2134/jeq2002.1442

Richards, R. A. (2006). Physiological traits used in the breeding of new cultivars for water-scarce environments. *Agr. Water Manage.* 80, 197–211. doi: 10.1016/j.agwat.2005.07.013

Roujean, J. L., and Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* 51, 375–384. doi: 10.1016/0034-4257(94)00114-3

Rouse, J. W., Haas, R. W., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (Greenwave Effect) of natural vegetation. *NASA/GSFCT Type III Final Rep.* 75–80.

Santos, P., and Negri, A. J. (1997). A comparison of the normalized difference vegetation index and rainfall for the amazon and northeastern Brazil. *J. Appl. Meteorol.* 36, 958–965. doi: 10.1175/1520-0450(1997)036<0958:ACOTND>2.0.CO;2

Sheshbahreh, M. J., Dehnavi, M. M., Salehi, A., and Bahreininejad, B. (2019). Effect of irrigation regimes and nitrogen sources on biomass production, water and nitrogen use efficiency and nutrients uptake in coneflower (*Echinacea purpurea* L.). *Agr. Water Manage.* 213, 358–367. doi: 10.1016/j.agwat.2018.10.011

Shibayama, M., and Wiegand, C. L. (1985). View azimuth and zenith, and solar angle effects on wheat canopy reflectance. *Remote Sens. Environ.* 18, 91–103. doi: 10.1016/0034-4257(85)90040-9

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Sims, D. A., and Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sens. Environ.* 84, 526–537. doi: 10.1016/S0034-4257(02)00151-7

Stagakis, S., Markos, N., Sykioti, O., and Kyparissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: an application on a *Phlomis fruticosa* Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sens. Environ.* 114, 977–994. doi: 10.1016/j.rse.2009.12.006

Strachan, I. B., Pattey, E., and Boisvert, J. B. (2002). Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote Sens. Environ.* 80, 213–224. doi: 10.1016/S0034-4257(01)00299-1

Thenkabail, P. S., Smith, R. B., and Pauw, E. D. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* 71, 158–182. doi: 10.1016/S0034-4257(99)00067-X

Thomas, J. R., Namken, L. N., Oerther, G. F., and Brown, R. G. (1971). Estimating leaf water content by reflectance measurements. *Agron. J.* 63, 845–847. doi: 10.2134/agronj1971.00021962006300060007x

Tian, Y. C., Yao, X., Yang, J., Cao, W. X., Hannaway, D. B., and Zhu, Y. (2011). Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crop Res.* 120, 299–310. doi: 10.1016/j.fcr.2010.11.002

Tian, Y. C., Zhu, Y., and Cao, W. X. (2005). Monitoring leaf photosynthesis with canopy spectral reflectance in rice. *Photosynthetica* 43, 481–489. doi: 10.1007/s11099-005-0078-y

Wang, W., Yao, X., Yao, X. F., Tian, Y. C., Liu, X. J., Ni, J., et al. (2012). Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crop Res.* 129, 90–98. doi: 10.1016/j.fcr.2012.01.014

Wolfe, D. W., Henderson, D. W., Hsiao, T. C., and Alvino, A. (1988). Interactive water and nitrogen effects on senescence of maize. II. Photosynthetic decline and longevity of individual leaves. *Agron. J.* 80, 865–870. doi: 10.2134/agronj1988.00021962008000060005x

Yao, X., Jia, W. Q., Si, H. Y., Guo, Z. Q., Tian, Y. C., Liu, X. J., et al. (2014). Exploring novel bands and key index for evaluating leaf equivalent water thickness in wheat using hyperspectra influenced by nitrogen. *PLoS One* 9:e96352. doi: 10.1371/journal.pone.0096352

Yu, G. R., Zhuang, J., and Yu, Z. L. (2001). An attempt to establish a synthetic model of photosynthesis-transpiration based on stomatal behavior for maize and soybean plants grown in field. *J. Plant Physiol.* 158, 861–874. doi: 10.1078/0176-1617-00177.

Zarco-Tejada, P. J., González-Dugo, V., Williams, L. E., Suárez, L., Berni, J. A. J., Goldhamer, D., et al. (2013). A PRI-based water stress index combining structural and chlorophyll effects: assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sens. Environ.* 138, 38–50. doi: 10.1016/j.rse.2013.07.024

Zarco-Tejada, P. J., Rueda, C. A., and Ustin, S. L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* 85, 109–124. doi: 10.1016/S0034-4257(02)00197-9

Zhang, D. Y., Coburn, C., Zhao, J. L., Wang, X., Wang, Z. J., and Liang, D. (2013) Chlorophyll Density Inversion of Soybean Canopy Based on Multi-angle Imaging Hyperspectral Data. 44, 205–213 Transactions of the Chinese Society for Agricultural Machinery (in Chinese with English abstract).

Zhang, H. Y., Ren, X. X., Zhou, Y., Wu, Y. P., He, L., Heng, Y. R., et al. (2018). Remotely assessing photosynthetic nitrogen use efficiency with in situ hyperspectral remote sensing in winter wheat. *Eur. J. Agron.* 101, 90–100. doi: 10.1016/j.eja.2018.08.010

Zhang, Z. B., Shao, H. B., Xu, P., Chu, L. Y., Lu, Z. H., and Tian, J. Y. (2007). On evolution and perspectives of bio-watersaving. *Colloid. Surface. B.* 55, 1–9. doi: 10.1016/j.colsurfb.2006.10.036

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Liu, Feng, Song, Li, Liu, Wang and Feng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Plant Science

REVIEW
published: 09 April 2021
doi: 10.3389/fpls.2021.616689

Applications of UAS in Crop Biomass Monitoring: A Review

Tianhai Wang[1], *Yadong Liu*[1], *Minghui Wang*[1], *Qing Fan*[2], *Hongkun Tian*[1], *Xi Qiao*[3,4]* and *Yanzhou Li*[1]*

[1] College of Mechanical Engineering, Guangxi University, Nanning, China, [2] College of Civil Engineering and Architecture, Guangxi University, Nanning, China, [3] Guangdong Laboratory of Lingnan Modern Agriculture, Shenzhen, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Area, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, [4] Guangzhou Key Laboratory of Agricultural Products Quality & Safety Traceability Information Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, China

Biomass is an important indicator for evaluating crops. The rapid, accurate and nondestructive monitoring of biomass is the key to smart agriculture and precision agriculture. Traditional detection methods are based on destructive measurements. Although satellite remote sensing, manned airborne equipment, and vehicle-mounted equipment can nondestructively collect measurements, they are limited by low accuracy, poor flexibility, and high cost. As nondestructive remote sensing equipment with high precision, high flexibility, and low-cost, unmanned aerial systems (UAS) have been widely used to monitor crop biomass. In this review, UAS platforms and sensors, biomass indices, and data analysis methods are presented. The improvements of UAS in monitoring crop biomass in recent years are introduced, and multisensor fusion, multi-index fusion, the consideration of features not directly related to monitoring biomass, the adoption of advanced algorithms and the use of low-cost sensors are reviewed to highlight the potential for monitoring crop biomass with UAS. Considering the progress made to solve this type of problem, we also suggest some directions for future research. Furthermore, it is expected that the challenge of UAS promotion will be overcome in the future, which is conducive to the realization of smart agriculture and precision agriculture.

Keywords: unmanned aerial systems, unmanned aerial vehicle, remote sensing, crop biomass, smart agriculture, precision agriculture

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Lea Hallik,
University of Tartu, Estonia
Jian Ma,
Sichuan Agricultural University, China

***Correspondence:**
Xi Qiao
qiaoxi@caas.cn
Yanzhou Li
lyz197916@126.com

Specialty section:
This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: *13 October 2020*
Accepted: *18 March 2021*
Published: *09 April 2021*

Citation:
Wang T, Liu Y, Wang M, Fan Q,
Tian H, Qiao X and Li Y (2021)
Applications of UAS in Crop Biomass
Monitoring: A Review.
Front. Plant Sci. 12:616689.
doi: 10.3389/fpls.2021.616689

INTRODUCTION

Agriculture plays an important role in maintaining all human activities. By 2050, population and socioeconomic growth are expected to double the current food demand (Niu et al., 2019). To solve the increasingly complex problems in the agricultural production system, the development of smart agriculture and precision agriculture provides important tools for meeting the challenges of sustainable agricultural development (Sharma et al., 2020). Biomass is a basic agronomic parameter in field investigations and is often used to indicate crop growth status, the effectiveness of agricultural management measures and the carbon sequestration ability of crops

(Bendig et al., 2015; Li W. et al., 2015). Fast, accurate and nondestructive monitoring of biomass is the key to smart agriculture and precision agriculture (Lu et al., 2019; Yuan et al., 2019).

Traditional biomass measurement methods are based on destructive measurements that require the manual harvesting (Gnyp et al., 2014), weighing and recording of crops, which makes large-scale, long-term measurements challenging and time-consuming, and these measurements are not only time-consuming and laborious but also difficult to apply over large areas (Boschetti et al., 2007; Yang et al., 2017). In other research areas, many studies have used satellite remote sensing to monitor biomass. Navarro et al. (2019) used Sentinel-1 and Sentinel-2 data to monitor the aboveground biomass (AGB) of a mangrove plantation. However, meteorological conditions have a great influence on satellite images, such as cloud and aerosol interference, surface glare and poor synchrony with tides (Tait et al., 2019). In addition, satellite data cannot provide sufficient data resolution for precision agricultural applications (Jiang et al., 2019; Song et al., 2020), and it is difficult to obtain timely and reliable data (Prey and Schmidhalter, 2019). Similar to satellite remote sensing, manned airborne equipment can cover a wide range, but the data are not detailed enough (Sofonia et al., 2019; ten Harkel et al., 2020). Meanwhile, although vehicle-mounted equipment can guarantee high accuracy, it has poor flexibility and slow speed (Selbeck et al., 2010; Tian et al., 2020). Unmanned aerial systems (UAS) represent a noncontact and nondestructive measurement method that can obtain the spectral, structural, and texture features of the target at different spatiotemporal scales (Jiang et al., 2019). These systems have the ability to obtain high spatial and temporal resolution data and have great application potential (Niu et al., 2019; Ramon Saura et al., 2019).

To date, most reviews of UAS in the field of agriculture are general reviews involving multiple fields in agriculture, and the description of biomass monitoring is not detailed enough (Hassler and Baysal-Gurel, 2019; Kim et al., 2019; Maes and Steppe, 2019). Reviews of remote sensing for crop biomass monitoring are rare and mainly introduce satellite remote sensing, while the application of UAS in crop biomass monitoring is rarely introduced (Chao et al., 2019). Therefore, the motivation of our study was to conduct a comprehensive review of almost all UAS-related studies in the field of crop biomass monitoring, including information on the equipment used in the field of crop biomass monitoring, biomass indices, and data processing and analysis methods. Finally, the relevant applications are reviewed according to different development directions.

THE COMPOSITION OF UAS

Unmanned aerial systems consist of unmanned aerial vehicle (UAV) platforms, autopilot systems, navigation sensors, mechanical steering components, data acquisition sensors, and other components (Jeziorska, 2019), among which the most important are the data acquisition sensors (Toth and Jozkow, 2016). Meanwhile, the type of UAV platforms and flight conditions will have a great impact on the data acquisition process of sensors, which need to be considered (Domingo et al., 2019; ten Harkel et al., 2020).

UAV Platforms

The most commonly used platforms in crop biomass monitoring are fixed-wing drones and rotor drones (Hassler and Baysal-Gurel, 2019). Hogan et al. (2017) summarized the characteristics of fixed-wing aircrafts and rotorcrafts. Fixed-wing aircrafts usually have a larger payload capacity, faster flight speed, longer flight time, and longer range than rotorcrafts. For these reasons, fixed-wing systems are particularly useful for collecting data over large areas. Fixed-wing aircrafts have poor mobility, need more space to land, and have more expensive prices than rotor UAVs. Rotor UAVs are very maneuverable and can hover, rotate and take pictures at almost any angle. Although there are also expensive models, more low-cost models have widely appeared in the market. Compared with fixed-wing aircrafts, the main disadvantage of rotor UAVs is their short range and flight time. **Figure 1** shows DJI Inspire 2 Rotor Drone[1] and eBee X Fixed-Wing Drone[2].

The flight planning of a fixed-wing UAV is very similar to that of a manned aircraft, while a rotor UAV can meet almost any trajectory requirements, including hover, slow motion and attitude control (Toth and Jozkow, 2016). These features enable rotor UAVs to perform extremely accurate tasks (Kim et al., 2019). Therefore, rotor UAVs are more commonly used in biomass monitoring than fixed-wing aircrafts.

Data Acquisition Sensors

Unmanned aerial systems usually obtain data through spectral sensors and depth sensors (Toth and Jozkow, 2016). Spectral sensors mainly include RGB sensors, multispectral sensors, and hyperspectral sensors, which can obtain color and texture information from the crop surface (Li et al., 2019). The difference between these three types of sensors is their ability to sense the spectrum (Shentu et al., 2018; Zhong et al., 2018; Kelly et al., 2019). Light detection and ranging (LiDAR) is a typical example of a depth sensor and can clearly obtain the three-dimensional structure and height information of crops (Wijesingha et al., 2019). **Figure 2** shows several UAV-mounted sensor types.

Spectral Sensors

Based on the same imaging principle, RGB, multispectral, and hyperspectral sensors all capture images by sensing spectral bands but have abilities to sense different spectral bands.

RGB sensors are a type of visible light camera that can detect three bands of color: red (R), green (G), and blue (B) (Shentu et al., 2018). The data from the three bands represent the intensity of R, G, and B in each pixel (Tait et al., 2019). Although RGB sensors have low accuracy compared with other sensors because they can collect spectral data from only three bands, the low-cost characteristic of RGB sensors is not possessed by others. Due to the need for low cost during the large-scale use of UAS in the monitoring of crop biomass, RGB sensors

---

[1] www.dji.com/inspire-2

[2] www.sensefly.com/drone/ebee-x-fixed-wing-drone/

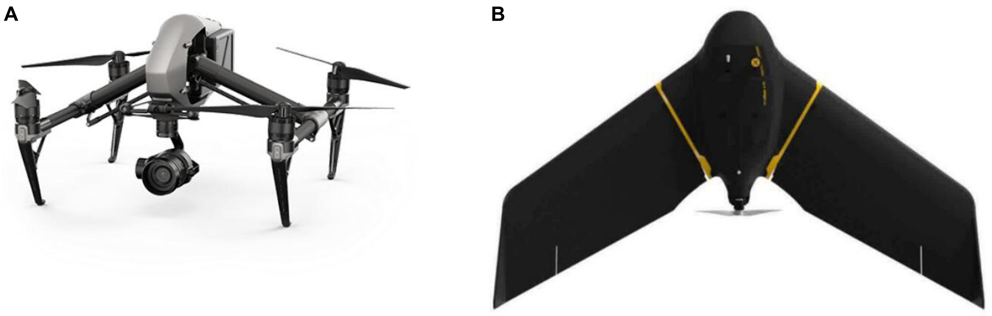


FIGURE 1 | (A) DJI Inspire 2 Rotor Drone and **(B)** eBee X Fixed-Wing Drone.

have received increasing attention because of their low-cost characteristics (Calou et al., 2019; Lu et al., 2019; Yue et al., 2019). The combination of RGB data with better biomass indices and advanced algorithms can obtain high accuracy at a low cost (Acorsi et al., 2019; Lu et al., 2019; Yue et al., 2019). Therefore, in the field of crop biomass monitoring by UAS, RGB sensors play an irreplaceable role.

Since spectral information is lost during the process of color image recording, the use of RGB input obviously limits the amount of information to extract from the highlighted area (Kelly et al., 2019). Compared with three-channel RGB imaging, multispectral images contain more imaging bands (Shentu et al., 2018).

Multispectral image data containing several near-infrared (NIR) spectral regions are superior to RGB data (Cen et al., 2019), but the disadvantage is that the cost of multispectral sensors is higher than that of RGB sensors (Costa et al., 2020). Different spectral bands can reflect the characteristics of different plants and can be used to effectively distinguish different crops (Xu et al., 2019). Song and Park (2020) used the RedEdge multispectral camera from MicaSense to analyze the spectral characteristics of aquatic plants and found that waterside plants exhibited the highest reflectivity in the NIR band, while floating plants had high reflectivity in the red-edge band.

Hyperspectral sensors can obtain more abundant spectral information than multispectral sensors (Zhong et al., 2018). Yue et al. (2018) used the UHD 185 Firefly (UHD 185 Firefly, Cubert GmbH, Ulm, Baden-Württemberg, Germany) hyperspectral sensor to collect panchromatic images with radiation records of 1000 × 1000 (1 band) and hyperspectral cubes of 50 × 50 (125 bands), with rich texture and spectral information. The disadvantage is that the cost of hyperspectral sensors is higher than that of multispectral sensors. In addition, the spatial resolution of hyperspectral images is lower than that of ordinary images, which may cause the loss of detail information for small targets. Meanwhile, more spectral information may not be useful in some cases. Tao et al. (2020) used hyperspectral sensors to study the correlation between different vegetation indices (VIs) and red-edge parameters and crop biomass. It was found that using too many spectral features as independent variables will lead to overfitting of the model, so it is necessary to use an appropriate number of spectral features that are highly related to biomass.

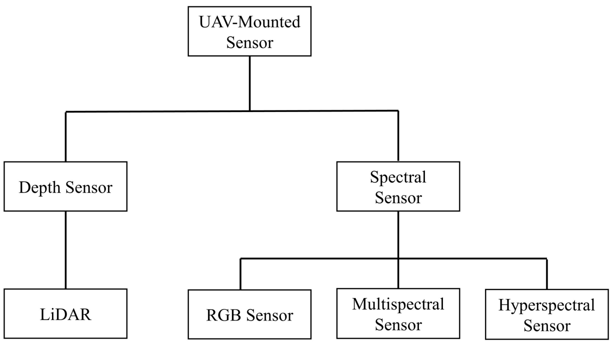


FIGURE 2 | UAV-mounted sensor types.

How to improve the reliability of spectral data is an unavoidable problem when using spectral sensors to collect data. First, the image resolution will affect the results of AGB monitoring, and the higher the image resolution is, the higher the prediction accuracy (Domingo et al., 2019). Yue et al. (2019) found that using image texture information to estimate the best image resolution for AGB monitoring depends on the size and row spacing of the crop canopy. Second, fisheye lenses may have an advantage over flat lenses. Calou et al. (2019) coupled a 16-megapixel plane lens with a 12-megapixel fisheye lens on a UAV for data collection, and the results showed that the fisheye lens estimation was the most accurate at an altitude of 30 m. Finally, at present, some applications using spectral data are processed without accurate or rough calibration. Guo et al. (2019) proposed a general calibration equation that is suitable for images under clear sky conditions and even under a small amount of clouds. The method needs to be further verified.

Although RGB sensors can only collect spectral data from the R, G, and B bands, the equipment is inexpensive. Although hyperspectral sensors can collect spectral information from many bands, the equipment is expensive. Unless there are special requirements for detailed hyperspectral images and the equipment is inexpensive, a multispectral sensor that balances the richness of spectral bands and equipment costs exhibits the highest cost performance and should be the default imaging choice (Hassler and Baysal-Gurel, 2019; Niu et al., 2019).

Light Detection and Ranging

Spectral data have poor robustness in the case of target overlap, occlusion, large illumination changes, shadows, and complex scenes. Depth data that do not change with brightness and color can provide additional useful information for complex scenes (Shuqin et al., 2016). At present, the depth sensors on UAV platforms are mainly LiDAR (Wallace et al., 2012; Qiu et al., 2019; Tian et al., 2019; Wang D.Z. et al., 2019; Yan et al., 2020). LiDAR has become an important information source for the evaluation of the vegetation canopy structure, which is especially suitable for species that limit artificial and destructive sampling (Brede et al., 2017). **Figure 3** shows a schematic illustration of the difference between LiDAR and spectral data (Zhu Y.H. et al., 2019).

Light detection and ranging is an active remote sensing technology that accurately measures distance by emitting laser pulses and analyzing the returned energy (Calders et al., 2015). With the development of global positioning system (GPS), inertial measurement unit (IMU), laser and computing technology, which make it possible to use LiDAR more inexpensively and accurately, a LiDAR system based on a UAV platform has become possible (Lohani and Ghosh, 2017). Compared with spectral sensors, LiDAR tends to provide accurate results of biomass prediction (Acorsi et al., 2019) because spectral data tend to be saturated in the middle and high canopy (Féret et al., 2017), and LiDAR can improve this through depth information (Hassler and Baysal-Gurel, 2019).

However, when it is difficult to estimate plant height, it is difficult to accurately monitor biomass through LiDAR. ten Harkel et al. (2020) used a VUX-SYS laser scanner to monitor the biomass of potato, sugar beet, and winter wheat. The researchers achieved good results in monitoring the biomass of sugar beet (R^2 = 0.68, RMSE = 17.47 g/m^2) and winter wheat (R^2 = 0.82, RMSE = 13.94 g/m^2), but the reliability for monitoring potato biomass was low. The reason for this result is that potatoes have complex canopy structures and grow on a ridge, and the other two crops have vertical structures and uniform heights. Therefore, for potatoes, it is difficult to visually determine the highest point of a specific position.

Finally, how to improve the data reliability by adjusting the LiDAR parameters is still lacking in more research. For instance, the sampling intensity of LiDAR has an impact on the accuracy of monitoring biomass (Wang D.Z. et al., 2020), but the current research in this area needs to be further verified.

Multisensor Fusion

The combination of data obtained from multiple sensors is an effective method to improve the accuracy of biomass estimation. On the one hand, the density of LiDAR point clouds has been improved with increased data resolution and penetrability (Wallace et al., 2012; Yan et al., 2020), which can improve the disadvantage that spectral data collected by RGB, multispectral and hyperspectral sensors are easily saturated in the middle and high canopy (Gitelson, 2004; Féret et al., 2017). On the other hand, the texture and spectral features that can be collected by RGB, multispectral, and hyperspectral sensors are also beyond the reach of LiDAR (Liu et al., 2019; Yue et al., 2019; Zheng et al., 2019). A variety of sensors with different characteristics are used to collect data, and the data that can reflect different characteristics of target crops are combined to provide more effective characteristics that are not cross-correlated that are needed for data analysis with a regression algorithm (Niu et al., 2019) to improve the accuracy of biomass estimation.

Wang et al. (2017) first evaluated the application of the fusion of hyperspectral and LiDAR data in maize biomass estimation. The results show that the fusion of hyperspectral and LiDAR data can provide better estimates of maize biomass than using LiDAR or hyperspectral data alone. Different from the previous methods of using LiDAR and optical remote sensing data to predict AGB separately or in combination, Zhu Y.H. et al. (2019) divided the estimation of maize AGB into two parts: aboveground leaf biomass (AGLB) and aboveground stem biomass (AGSB). AGLB was measuring with multispectral data, which are sensitive to the vegetation canopy. AGSB was measured with LiDAR point cloud data, which are sensitive to the vegetation structure. Compared with using LiDAR data alone or using multispectral data alone, the combination of LiDAR data and multispectral data can more accurately estimate AGB, in which the R^2 increases by 0.13 and 0.30, the RMSE decreases by 22.89 and 54.92 g/m^2, and the NRMSE decreases by 4.46 and 7.65%.

Other researchers have also carried out many studies in the field of multisensor data fusion and obtained the same results in studies on the monitoring of crop biomass, such as rice (Cen et al., 2019) and soybean (Maimaitijiang et al., 2020). In addition, different crops have different characteristics, and the same crop will show different characteristics under different growth conditions (Johansen et al., 2019; ten Harkel et al., 2020), which requires the use of different sensors to collect crop information comprehensively and screen out some information most related to biomass. The combination of data from multiple sensors is an effective method to improve the accuracy of biomass estimation.

Flight Parameters

To ensure the most accurate results for biomass monitoring, further tests should be carried out before experiments to determine the optimal flight parameters, such as altitude, speed, location of flight lines, and overlap (Domingo et al., 2019). Increasing the UAV flight height will reduce image resolution (Lu et al., 2019), and the sensitivity of the accuracy of biomass monitoring to the image spatial resolution is an important reference for the configuration of a UAV flight height. The estimation of the other flight parameters did not exhibit much different effects on the overall effect of biomass monitoring from that of standard flight parameters, but different flight parameters can lead to different point densities and distributions, which have a greater impact on biomass monitoring than altitude and velocity. A better crossover model and a closer flight path may improve biomass monitoring overall (ten Harkel et al., 2020). The images need to be overlapped sufficiently to improve the accuracy of biomass monitoring (Borra-Serrano et al., 2019). Therefore, the UAV flight plan should be wide enough. Domingo et al. (2019) found that reducing side overlap from 80 to 70% while

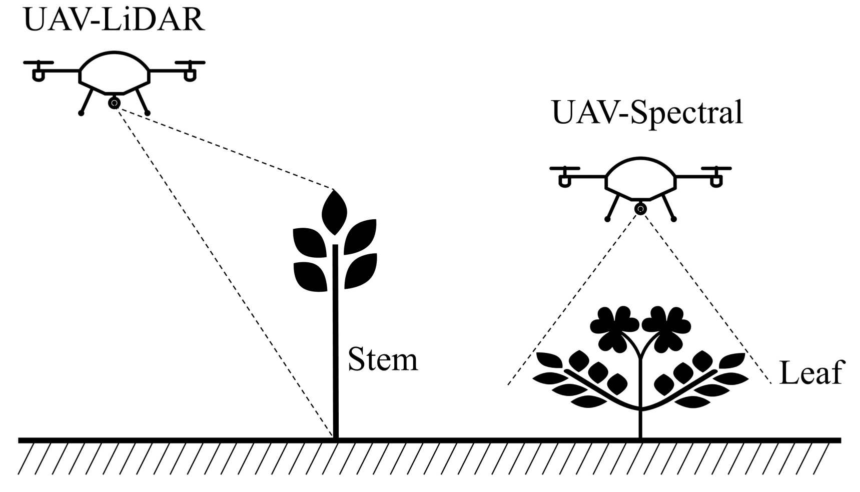


FIGURE 3 | A schematic illustration of the difference between LiDAR and spectral data.

maintaining a fixed forward overlap of 90% may be an option to reduce flight time and procurement costs. For specific species, such as rice, due to the physiological characteristics of rice, the analysis of the solar elevation angle during the creation of a flight plan is very important to avoid the influence of sun glint and hotspot effects (Jiang et al., 2019).

BIOMASS INDICES

It is a common method in biomass monitoring to use biomass indices to obtain data directly related to biomass. Common biomass indices include VIs and crop height (CH), which are extracted from images or three-dimensional point clouds. There are also relevant studies that do not use biomass indices but directly use images or three-dimensional point clouds to conduct correlation analysis with biomass (Nevavuori et al., 2019).

Vegetation Indices

A variety of VIs from remote sensing images can be used to monitor the state of vegetation on the ground. This method is also able to quantitatively evaluate the richness, greenness and vitality of vegetation. After years of development, VIs can be divided into various monitoring and calculation methods, among which the most commonly used is the normalized difference vegetation index (NDVI) proposed by Rouse et al. (1974). The NDVI is usually used to reflect information such as vegetation cover and growth, and its calculation formula is as follows:

$$NDVI = \frac{NIR - R}{NIR + R}$$

Near-infrared is the reflectance in NIR band, and R is the reflectance in red band. The value range of NDVI is (−1, 1). It is generally believed that an NDVI value less than 0 represents no vegetation coverage, while a value greater than 0.1 represents vegetation coverage (Li Z. et al., 2015). Since the index is positively correlated with the density of vegetation, the higher the NDVI value is, the higher the vegetation coverage will be.

Different VIs have unique characteristics, and more spectral features can be identified by using multiple VIs to obtain high monitoring accuracy. Marino and Alvino (2020) used the soil adjusted vegetation index (SAVI), NDVI and OSAVI to characterize 10 winter wheat varieties in a field at different growth stages and obtained optimal biomass monitoring results. Villoslada et al. (2020) combined 13 VIs to obtain the highest accuracy.

Vegetation indices can be built not only on the basis of spectral information but also on the basis of texture information. Texture is an important characteristic for identifying objects or image areas of interest. In several texture algorithms, the gray level co-occurrence matrix (GLCM), which includes variance (VAR), entropy (EN), data range (DR), homogeneity (HOM), second moment (SE), dissimilarity (DIS), contrast (CON), and correlation (COR), which are based on Haralick et al. (1973), is often used to test the effects of texture analysis from UAS data on biomass estimation potential (Zheng et al., 2019).

Vegetation indices based on image texture are usually combined with VIs based on spectral information to monitor crop biomass, and this combination can improve the accuracy of monitoring biomass significantly (Liu et al., 2019; Yue et al., 2019; Zheng et al., 2019). Zheng et al. (2019) predicted rice AGB using stepwise multiple linear regression (SMLR) in combination with VIs and image texture, and the results showed that the combination of texture information and spectral information significantly improved the accuracy of rice biomass estimations compared with the use of spectral information alone (R^2 = 0.78, RMSE = 1.84 t/ha).

Previously, as the required data were obtained by satellites, the spectral data collected would be affected by clouds. When there was cloud cover in the observation area, the information received by the satellite-borne sensor would be all cloud information, instead of reflecting the local vegetation cover (Feng et al., 2009).

The low altitude and flexibility of UAS solve this problem, making VIs more widely used. At present, VIs have become indispensable biomass indices for monitoring crop biomass. Common VIs are shown in **Table 1**.

Crop Height

Crop height is an important indicator to characterize the vertical structure, and CH is usually strongly correlated with biomass (Scotford and Miller, 2004; Prost and Jeuffroy, 2007; Salas Fernandez et al., 2009; Montes et al., 2011; Hakl et al., 2012; Alheit et al., 2014). The crop surface model (CSM) is an effective CH information extraction technique and has been widely used for different crops (Han et al., 2019).

Crop height data can be obtained using RGB sensors and multispectral sensors. Cen et al. (2019) established a CSM to determine the CH (Tilly et al., 2014) based on spliced RGB images. First, structure from motion (SfM) was used to generate a point cloud, and the specific steps can be found in the study of Tomasi and Kanade (1992). Point clouds consist of matching points between overlapping images such as crop canopies and topographic surfaces. A digital elevation model (DEM) and digital terrain model (DTM) were obtained by classifying the point clouds. The DEM was based on a complete dense point cloud representing the height of the crop canopy, while the DTM was developed from only the surface dense point cloud. The CSM could be obtained by subtracting the DTM from the DEM by importing the two models into ArcGIS software (ArcGIS, Esri Inc., Redlands, CA, United States). Hassan et al. (2019) used a Sequoia 4.0 multispectral camera with the same method to measure the CH of wheat, and the results showed that the correlation between the CH data from UAS and the actual height was very high ($\underline{R}^2 = 0.96$).

Crop height data can also be obtained using LiDAR. Zhu W.X. et al. (2019) used CloudCompare open-source software to construct CH raster data from the LiDAR point cloud and studied the effects of CH on monitoring the AGB of crops. The results showed that CH is a robust indicator that can be used to estimate biomass, and the high spatial resolution of the CH data set was helpful to improve the effect of maize AGB estimation.

The monitoring of crop biomass by a single biomass index is sometimes unreliable. On the one hand, it is difficult to obtain reliable CH data from LiDAR in some cases. Johansen et al. (2019) found that dust storms can cause tomato plants to flatten and that once the tomato fruits become large and heavy, the weight may cause the branches to bend downward, thereby reducing the height of the plants. ten Harkel et al. (2020) found that potatoes have complex canopy structures and grow on ridges, so it is difficult to visually determine the highest point of a specific position. In the above cases, VIs can achieve better results than other measurements. On the other hand, CH data can better reflect the three-dimensional information of crops and can more accurately reflect the biomass of crops in the scene of target overlap, occlusion, large changes in light, shadow, and complex scenes. In addition, the information collected by UAS includes not only target crops but also other interference information. If this interference information cannot be effectively eliminated, it will have a negative impact on the monitoring of crop biomass, which can be improved by the combination of multiple biomass indices (Niu et al., 2019). Therefore, the combination of multiple models for biomass estimation is an effective method to improve the accuracy of biomass estimation.

Multi-index Fusion

The combination of multiple models for biomass estimation is an effective method to improve the accuracy of biomass estimation. The combination of spectral and textural features to construct VIs or the combination of VIs and CH has been shown to improve the results of biomass estimation.

Based on the idea of combining VIs with CH, Cen et al. (2019) used a biomass model that combined VIs and CH to monitor rice biomass under different nitrogen treatments. The results showed that the CH extracted by the CSM exhibited a high correlation with the actual CH. The monitoring model that incorporated RGB and multispectral image data with random forest regression (RFR) significantly improved the prediction results of AGB, in which the RMSEP decreased by 8.33–16.00%, $R^2 = 0.90$, RMSEP = 0.21 kg/m^2, and RRMSE = 14.05%.

Relevant studies have proven that a biomass model combined with VIs and CH can also improve the biomass estimation accuracy for corn (Niu et al., 2019), wheat (Lu et al., 2019), ryegrass (Borra-Serrano et al., 2019), and other crops. These cases prove that the combination of VIs and CH is an effective way to build a biomass model. However, Niu et al. (2019) pointed out that the fusion of CH data derived from RGB images in the VIs model, which was based on MLR, did not significantly improve the estimation of the VI model, which may be caused by the clear correlation between VIs and CH in this crop (Schirrmann et al., 2016). Therefore, it is necessary to combine biomass indices reasonably for different crops.

According to the idea of combining spectral information with image texture to build VIs, Liu et al. (2019) used a linear regression model to convert the digital number (DN) of the original image into surface reflectance. The reflectivity obtained from the gain and offset values of each band was used to calculate the VIs and image texture. The results showed that the introduction of image texture into the partial least squares regression (PLSR) and RFR models could estimate winter rape AGB more accurately than a model based on VIs alone. The accuracy of the prediction of AGB by the RFR model using VIs and texture measurements (RMSE = 274.18 kg/ha) was slightly higher than that of the PLSR model (RMSE = 284.09 kg/ha). The same idea has also obtained good results in applications to winter wheat (Yue et al., 2019), rice (Zheng et al., 2019), soybean (Maimaitijiang et al., 2020), and other crops. Biomass models combined with VIs and image texture have great potential in the estimation of crop biomass.

DATA PROCESSING AND ANALYSIS METHODS

Data analysis is the key link to build the relationship between the data obtained from UAS and the actual crop biomass, and it is an important part of UAS. The data obtained from UAS often

TABLE 1 | Introduce the formulation and features of common VIs.

VIs	Formulation	Features	References
Ratio vegetation index	$RVI = NIR / R$	Monitor the photosynthetically active biomass of plant canopies.	Tucker, 1979
Green chlorophyll index	$GCI = (NIR/G) - 1$	Estimation of spatially distributed chlorophyll content in crops.	Gitelson et al., 2005
Red-edge chlorophyll index	$RECI = (NIR / RE) - 1$	Estimation of spatially distributed chlorophyll content in crops.	Gitelson et al., 2005
Normalized difference vegetation index	$NDVI = (NIR - R)/(NIR + R)$	Quantitative measurement of vegetation conditions over broad regions.	Rouse et al., 1974
Green normalized difference vegetation index	$GNDVI = (NIR - G)/(NIR + G)$	Nondestructive chlorophyll estimation in leaves.	Gitelson et al., 2003
Green-red vegetation index	$GRVI = (G - R) / (G + R)$	Monitor the photosynthetically active biomass of plant canopies.	Tucker, 1979
Normalized difference red-edge	$NDRE = (NIR - RE) / (NIR + RE)$	Increases the sensitivity of NDVI to chlorophyll content by approximately fivefold.	Gitelson and Merzlyak, 1997
Normalized difference red-edge index	$NDREI = (RE - G) / (RE + G)$	Estimation of senescence rate at maturation stages.	Hassan et al., 2018
Simplified canopy chlorophyll content index	$SCCCI = NDRE / NDVI$	Real-time detection of nutrient status.	Raper and Varco, 2015
Enhanced vegetation index	$EVI = 2.5 \times (NIR - R) / (1 + NIR - 2.4 \times R)$	The EVI remains sensitive to canopy variations while the NDVI is asymptotically saturated in high biomass regions.	Huete et al., 2002
Two-band enhanced vegetation index	$EVI2 = 2.5 \times (NIR - R) / (NIR + 2.4 \times R + 1)$	A 2-band EVI (EVI2), without a blue band, which has the best similarity with the 3-band EVI (EVI).	Jiang et al., 2008
Wide dynamic range vegetation index	$WDRVI = (a \times NIR - R) / (a \times NIR + R)$ $(a = 0.12)$	The sensitivity of the WDRVI to moderate-to-high LAI (between 2 and 6) was at least three times greater than that of the NDVI.	Gitelson, 2004
Soil adjusted vegetation index	$SAVI = (1 + L) (NIR - RE) / (NIR + RE + L)$	Almost eliminated soil-induced changes in vegetation index.	Huete, 1988
Optimized soil adjusted vegetation index	$OSAVI = (NIR - R) / (NIR - R + 0.16)$	Less sensitive to soil background and atmospheric effects.	Rondeaux et al., 1996
Modified chlorophyll absorption in reflectance index	$MCARI = [(RE - R) - 0.2 \times (RE - G)] \times (RE / R)$	Evaluate the nutrient variability over large fields quickly.	Daughtry et al., 2000
MCARI/OSAVI	$MCARI / OSAVI$	Evaluate the nutrient variability over large fields quickly.	Daughtry et al., 2000
Transformed chlorophyll absorption in reflectance index	$TCARI = 3 \times [(RE - R) - 0.2 \times (RE - G) \times (RE / R)]$	Minimizing LAI (vegetation parameter) influence and underlying soil (background) effects.	Haboudane et al., 2002
TCARI/OSAVI	$TCARI / OSAVI$	Minimizing LAI (vegetation parameter) influence and underlying soil (background) effects.	Haboudane et al., 2002

Generally, one or several kinds of VIs should be selected according to different situations. Abbreviate green, red, red-edge, and near-infrared to G, R, RE, and NIR in formulation.

contain different noises, and the information is highly correlated. Generally, effective data analysis methods are needed to interpret the data and establish a robust prediction model (Cen et al., 2019). Therefore, scientific and systematic data analysis methods often play an important role.

Data Preprocessing Methods

Since the data collected by UAS cannot be directly used to monitor biomass, a series of preprocessing steps is needed for the data. When spectral sensors are used, an indispensable step is geometric correction and mosaicking of the image. Common software includes Pix4DMapper and Agisoft Photoscan.

Pix4DMapper software (Pix4D, S.A., Lausanne, Switzerland) is UAS photography geometric correction and mosaic technology based on feature matching and SfM photogrammetry technology. Initially, images were processed in any model space to create three-dimensional point clouds. The point clouds could be transformed into a real-world coordinate system using either direct geolocation techniques to estimate the camera's location or GCP techniques for automatic identification within the point cloud. The point cloud was then used to generate the DTM required for image correction. Subsequent geographic reference images are linked together to form a mosaic of the study area (Turner et al., 2012).

Agisoft Photoscan software (Agisoft LLC, St. Petersburg, Russia) is also a common UAS data preprocessing software. The processing procedure is similar to that of Pix4DMapper. Finally, the UAS image is exported to TIFF image format for subsequent analysis (Acorsi et al., 2019; Lu et al., 2019). **Figure 4** shows RGB imagery datasets were processed using the software Agisoft PhotoScan (Sun et al., 2019).

(a) High-resolution proof images of the acquisition area
(b) Overall map of research area processed by Agisoft PhotoScan.

Data Analysis Methods

Machine learning algorithms are widely used to process biomass information. According to whether the input dataset is labeled, machine learning algorithms can be divided into supervised learning algorithms and unsupervised learning algorithms (Dike et al., 2018). Supervised learning algorithms depend on a labeled dataset. Classification algorithms and regression algorithms are the two output forms of supervised learning. In crop biomass monitoring, regression algorithms are more often used than classification algorithms. Because the expected biomass results are often continuous instead of discrete. Unsupervised learning does not rely on a labeled dataset. It is often used when the cost of labeled datasets is unacceptable (Sapkal et al., 2007). This is also a common method to reduce the dimensionality of the data. Most unsupervised learning algorithms are in the form of cluster analysis. **Figure 5** shows the types of machine learning algorithms.

Biomass monitoring is a typical regression problem, which can to be solved by supervised learning algorithms. Enough labeled datasets are the basis of supervised learning algorithms. Actual biomass data tend to obtain through destructive samplings (Jiang et al., 2019; Yue et al., 2019). Field trials are limited by the area of cropland and the crop growing season. Therefore, sufficiently large datasets are often not available. How to properly divide the datasets into training datasets and validation datasets is a challenge to train supervised learning algorithms. To solve this problem, Jiang et al. (2019) used fivefold cross validation, Han et al. (2019) used repeated 10-fold cross validation, Zhu W.X. et al. (2019) used leave-one-out-cross validation (LOOCV) to reduce generalization error. Fivefold cross validation and repeated 10-fold cross validation belong to *k*-fold cross validation. LOOCV is a special case of *k*-fold cross validation, in which the number of folds equals the number of instances (Wong, 2015). *k*-fold cross validation divides the datasets into *k* folds, treats each fold as a validation dataset and regards the other $k-1$ folds as a training dataset (Wong and Yang, 2017). The value of folds can be large and the value of replications should be small if *k*-fold cross validation is applied in the classification algorithms (Wong and Yeh, 2020).

Support Vector Regression

Support vector regression (SVR) is a boundary detection algorithm for identifying/defining multidimensional boundaries (Sharma et al., 2020), and the basis of this method is to solve the regression problem by using appropriate kernel functions to map the training data to the new hyperspace characteristics and transform the multidimensional regression problem into a linear regression problem (Navarro et al., 2019). Duan et al. (2019) in an analysis of VIs and image texture using SVR found that the SVR itself has the ability to find a suitable combination of different reflectance bands, which shown that SVR has strong adaptability to complex data and is suitable for data analysis in biomass monitoring. Yang et al. (2019) compared PLSR and SVR, and the results showed that the accuracy of SVR was higher than that of PLSR, and the SVR optimized by particle swarm optimization (PSO) could obtain more appropriate parameters and improve the accuracy of the model.

Random Forest Regression

Random forest regression is a data analysis and statistical method that is widely used in machine learning and remote sensing research (Viljanen et al., 2018). Compared with artificial neural networks (ANNs), RFR does not suffer from overfitting problems because of the law of large numbers, and the injection of suitable randomness makes them precise regressors (Breiman, 2001). The random forest algorithm makes full use of all input data and can tolerate outliers and noise (Jiang et al., 2019). This algorithm has the advantages of high prediction accuracy, no need for feature selection and insensitivity to overfitting (Tewes and Schellberg, 2018; Viljanen et al., 2018).

Artificial Neural Network

An ANN is an information processing paradigm that is inspired by the way biological nervous systems such as the brain process information (Awodele and Jegede, 2013). ANN is a nonparametric nonlinear model that uses a neural network to transmit between layers and simulates reception and processing of information by the human brain (Zha et al., 2020). In individual cases, the results of the algorithm are not better than those of the multiple linear regression (MLR) method. The reason for this difference may be that in these applications, a small sample set will not meet the needs of the artificial neural network (Han et al., 2019; Zhu W.X. et al., 2019; Zha et al., 2020), and compared with RFR, ANN needs large data sets and a large number of repetitions to generate appropriate nonlinear mapping and obtain the optimal neural network (Devia et al., 2019; Han et al., 2019); however, RFR can still be applied for a small amount of sample data (Han et al., 2019; Liu et al., 2019), which leads to more frequent biomass monitoring use of RFR. Therefore, before the development of a deep neural network (DNN), the remote sensing field, including UAS studies, shifted the focus of data analysis methods from ANN to SVR and RFR (Ma et al., 2019).

The appearance of DNN and a series of methods to solve overfitting improved the effect of ANN (Maimaitijiang et al., 2020). Nevavuori et al. (2019) used a convolutional neural network (CNN) to predict the biomass of wheat and barley. The researchers tested the influence of the selection of the training

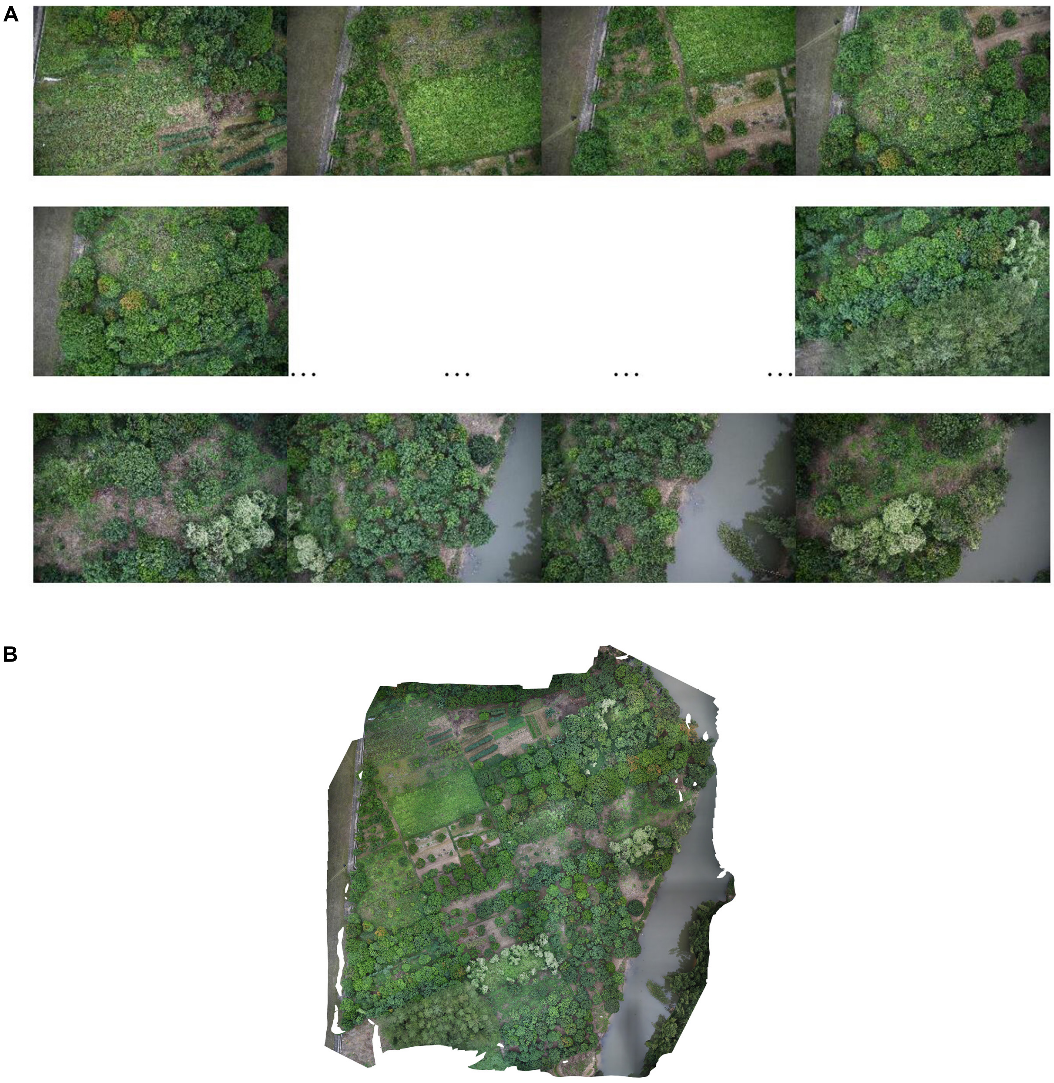


FIGURE 4 | RGB imagery datasets were processed using the software Agisoft PhotoScan. **(A)** High-resolution proof images of the acquisition area. **(B)** Overall map of research area processed by Agisoft PhotoScan.

algorithm, the depth of the network, the regularization strategy, the adjustment of super parameters, and other aspects of CNN on the prediction efficiency to improve the monitoring effect. This study proved that if enough information can be collected to increase the number of samples and solve the overfitting problem, ANN will perform no worse than RFR (Zhang et al., 2019).

Multiple Regression Techniques

Multiple linear regression (Borra-Serrano et al., 2019; Devia et al., 2019; Han et al., 2019; Zhu W.X. et al., 2019), SMLR (Lu et al., 2019; Zheng et al., 2019) and PLSR (Borra-Serrano et al., 2019; Liu et al., 2019; Yue et al., 2019) are also commonly used multiple regression algorithms. However, with the gradual progress of SVR, RFR, and ANN, these algorithms have gradually become references for SVR, RFR, and ANN and are no longer the main focus of data analysis.

Devia et al. (2019) described an MLR equation for monitoring rice biomass with VIs. In general, there was a linear relationship between the accumulation of biomass and VIs. However, the

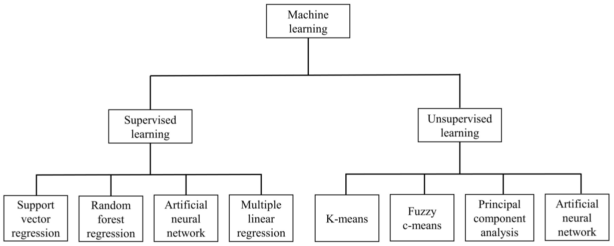


FIGURE 5 | The types of machine learning algorithms.

TABLE 2 | Summarize the equipment, methods, and important results of the studies cited in the body.

Crop	Platforms	Sensors	Biomass indices	Data analysis methods	Results	References
Wheat	DJI Phantom series	A digital camera	VIs, CH	RFR	R^2 = 0.78, RMSE = 1.34 t/ha, RRMSE = 28.98%	Lu et al., 2019
Rice	DJI S1000 DJI Phantom 4 Pro	Mini-MCA 12 multispectral camera DJI FC6310 digital camera	VIs, CH Meteorological feature	SER	R^2 = 0.86, RMSE = 178.37 g/m^2, MAE = 127.34 g/m^2	Jiang et al., 2019
Potato Sugar beet Winter wheat	RIEGL RiCOPTER	VUX-SYS laser scanner	CH	MLR	Potato: R^2 = 0.24, RMSE = 22.09 g/m^2 Sugar beet: R^2 = 0.68, RMSE = 17.47 g/m^2 Winter wheat: R^2 = 0.82, RMSE = 13.94 g/m^2	ten Harkel et al., 2020
Maize	DJI Phantom 2	Ricoh GR digital camera	CH	Statistical analysis	The estimated values were most accurate when using a fisheye lens at 30 m altitude.	Calou et al., 2019
Winter wheat	DJI S1000	DSC-QX100 digital camera	VIs	SMLR	R^2 = 0.89, MAE = 0.67 t/ha, RMSE = 0.82 t/ha	Yue et al., 2019
Rice	A lightweight octorotor UAV	An RGB camera A multispectral camera	VIs, CH	RFR	R^2 = 0.90, RMSEP = 0.21 kg/m^2, RRMSE = 14.05%	Cen et al., 2019
Winter wheat	DJI S1000	DSC–QX100 digital camera UHD 185 Firefly snapshot hyperspectral sensor	VIs	Exponential regression	R^2 = 0.67, MAE = 1.19, RMSE = 1.71	Yue et al., 2018
Winter wheat	DJI S1000	UHD 185-Firefly	VIs	PLSR	The results of AGB monitoring can be improved by combining the red-edge parameters with VIs.	Tao et al., 2020
Corn Wheat	DJI M600 Pro	Mini-MCA 6 multispectral camera	VIs	Linear regression	A systematical radiometric calibration method was proposed.	Guo et al., 2019
Rice	Mikrokopter OktoXL	Tetracam mini-MCA6 multispectral camera	VIs	SMLR	R^2 = 0.78, RMSE = 1.84 t/ha	Zheng et al., 2019
Winter oilseed rape	DJI S1000	Mini-MCA multispectral camera	VIs	PLSR RFR	RFR: RMSE = 274.18 kg/ha PLSR: RMSE = 284.09 kg/ha	Liu et al., 2019
Maize	DJI Phantom 4 Pro DJI M600 Pro	Parrot Sequoia multispectral camera DJI FC6310 digital camera RIEGL VUX-1UAV laser scanner	VIs, CH	MLR PLSR	MLR: R^2 = 0.82, RMSE = 79.80 g/m^2, NRMSE = 11.12% PLSR: R^2 = 0.86, RMSE = 72.28 g/m^2, NRMSE = 10.07%	Zhu Y.H. et al., 2019
Soybean	DJI S1000	Mapir Survey2 RGB camera Parrot Sequoia multispectral camera FLIR Vue Pro R 640 thermal imager	VIs, CH	DNN-F2	R^2 = 0.720, RMSE = 478.9 kg/ha, RRMSE = 15.9%	Maimaitijiang et al., 2020
Tomato	DJI Matrice 100	A RGB Zenmuse X3 sensor	VIs	RFR	R^2 = 0.85, RMSE = 0.052 m	Johansen et al., 2019
Ryegrass	Onyxstar HYDRA-12	RGB camera	VIs, CH Meteorological feature	MLR RFR	MLR: R^2 = 0.81, RMSE = 679 kg/ha, NRMSE = 21.3% RFR: R^2 = 0.70, RMSE = 769 kg/ha, NRMSE = 24.2%	Borra-Serrano et al., 2019
Wheat Barley	Airinov Solo 3DR UAV	Parrot's NIR-capable SEQUIOA-sensor	None	CNN	MAE = 484.3 kg/ha, MAPE = 8.8%	Nevavuori et al., 2019
Ten winter wheat cultivars	Ebee fixed-wing UAV	Canon Powershot S110 RGB camera Canon Powershot S110 NIR camera	VIs	Cluster analysis	Combination of multiple VIs can be a valid strategy.	Marino and Alvino, 2020

(Continued

TABLE 2 | Continued

Crop	Platforms	Sensors	Biomass indices	Data analysis methods	Results	References
Coastal meadows	Ebee fixed-wing UAV	Parrot Sequoia multispectral camera	VIs	RFR	Combination of multiple VIs can be a valid strategy.	Villoslada et al., 2020
Maize	DJI S1000	DSC-QX100 digital camera Parrot Sequoia multispectral camera	BIOVP (VIs, CH)	RFR	R^2 = 0.944, RMSE = 0.495, MAE = 0.355	Han et al., 2019
Bread wheat	DJI Inspires 1 model T600	Sequoia 4.0 multispectral camera	CH	Linear regression	R^2 = 0.96	Hassan et al., 2019
Maize	EWZ-D6 six-rotator UAV DJI M100 four-rotator UAV Ebee fixed-wing UAV	MultiSPEC-4C multispectral camera MicaSense RedEdge-M multispectral camera Alpha Series AL3-32 LiDAR sensor	CH	RFR	R^2 = 0.90, RMSE = 2.29, MRE = 0.22	Zhu W.X. et al., 2019
Rice	An UAV equipped with a Mini-MCA system	An array of 12 individual miniature digital cameras	VIs	SVR	SVR itself has the ability to find a suitable combination of different reflectance bands.	Duan et al., 2019
Winter wheat	Four-axis aerial vehicle UAV 3P	Sony EXMOR HD camera	VIs	SVR	R^2 = 0.9025, RMSE = 0.3287	Yang et al., 2019
Rice	UAV	Tetracam ADC-lite multispectral camera	VIs	MLR	R^2 = 0.76	Devia et al., 2019
Eggplant Tomato Cabbage	DJI 3 Pro	DJI FC300X RGB camera	CH	SVR RFR	R^2 ranging from 0.87 to 0.97 Bias ranging from −0.66 to 0.45 cm	Moeckel et al., 2018
Sorghum	Custom designed UAV platforms	Sony Alpha ILCE-7R Velodyne VLP-16 Two Headwall Photonics push-broom scanners	Four hyperspectral-based features and four LiDAR-based features	PLSR SVR RFR	The data source was more important than the regression method.	Masjedi et al., 2020
Rice	UAV platform	Tetracam ADC-lite multispectral camera	VIs	Multivariable regression	An average correlation of 0.76	Devia et al., 2019

This table covers plentiful case-studies from different regions for different crops.

relationship between biomass and VIs in other crops at maturity can be nonlinear. Therefore, MLR does not apply to these nonlinear relations of crops. Zheng et al. (2019) used SMLR to establish the relationship between rice biomass and remote sensing variables (VIs, image texture, and the combination of VIs and image texture). Although the estimation accuracy was high, the model was complex and difficult to generalize. Moeckel et al. (2018) tested the ability of PLSR, RFR, and SVR to predict the CH of eggplant, tomato and cabbage, and the results showed that PLSR did not exceed the performance of RFR and SVR, so it was excluded first.

The monitoring of biomass is a typical nonlinear problem (Zha et al., 2020). These regression techniques are more suitable for data showing linear or exponential relationships between remote sensing variables and crop parameters (Atzberger et al., 2010; Jibo et al., 2018; Lu et al., 2019). These methods are often not as good as SVR, RFR, and ANN in the monitoring of biomass.

The construction of a high-performance monitoring model based on advanced algorithms (such as machine learning algorithms) is a good method to improve the effect of crop biomass monitoring (Niu et al., 2019). The monitoring of biomass is a typical multi-feature nonlinear problem (Zha et al., 2020), and machine learning algorithms (such as SVR, RFR, and ANN) exhibit superior results in solving these types of problems (Breiman, 2001; Navarro et al., 2019; Maimaitijiang et al., 2020). During the study, by comparing with RFR, the researchers found that ANN was often superior to RFR when dealing with large sample sizes and complex, nonlinear, and redundant data sets (LeCun et al., 2015; Schmidhuber, 2015; Kang and Kang, 2017; Zhang et al., 2018; Maimaitijiang et al., 2020). However, in a small sample size, the lack of samples often leads to the phenomenon of overfitting, and RFR will achieve better results than ANN due to its stronger robustness and generalization ability (Zhang and Li, 2014; Yao et al., 2015; Yuan et al., 2017; Yue et al., 2017; Zheng et al., 2018; Zhu W.X. et al., 2019; Zha et al., 2020).

THE PROMOTION OF LARGE-SCALE UAS APPLICATIONS

Accuracy is an important indicator to evaluate the effects of UAS in the field of crop biomass estimation. In addition, reducing the cost to promote the large-scale application of UAS in this field is a difficult problem. From the perspective of improving the accuracy of crop biomass estimations, multisensor data fusion, multi-index fusion, the consideration of a variety of features not directly related to the monitoring of biomass, and the use of advanced algorithms are feasible directions (Maimaitijiang et al., 2020). Considering the promotion of large-scale applications, the use of low-cost sensors and the combination of suitable models and algorithms to improve the estimation accuracy of low-cost sensors, rather than the use of more expensive sensors, is an effective research path to promote the large-scale application of UAS in the field of crop biomass monitoring (Acorsi et al., 2019; Lu et al., 2019; Niu et al., 2019; Yue et al., 2019). RGB sensors are currently the most widely used low-cost sensor (Lussem et al., 2019), and studies based on RGB sensors are expected to promote the large-scale application of UAS in the field of crop biomass monitoring.

RGB sensors are not capable of providing NIR band data. Therefore, VIs associated with NIR bands cannot be used, which inhibits the enhancement of vegetation vitality contrast (Lu et al., 2019) and may affect the accuracy of biomass estimation. Lu et al. (2019) used a combination of advanced algorithms and multi-index fusion to compensate for this deficiency. Yue et al. (2019) fused the image texture and VIs to obtain the most accurate estimated value of AGB (R^2 = 0.89, MAE = 0.67 t/ha, RMSE = 0.82 t/ha). These study proves that the use of low-cost sensors can guarantee the accuracy of biomass estimation and is expected to promote large-scale applications.

Solar elevation angle (Jiang et al., 2019), meteorological conditions (Devia et al., 2019; Wang F. et al., 2019), rainfall (Liu et al., 2019; Rose and Kage, 2019), soil characteristics (Acorsi et al., 2019; Vogel et al., 2019), the spatial distribution of multiple plants in a block (Han et al., 2019), and other characteristics not directly related to biomass monitoring also affect the accuracy of biomass estimations. The monitoring accuracy of low-cost sensors can be improved by considering the characteristics that are not directly related to biomass monitoring.

Jiang et al. (2019) calculated the solar elevation angle to avoid sun glint and hotspot effects. In addition, growing degree days (GDD) was incorporated into the model to estimate rice AGB as a meteorological feature. Models incorporating meteorological features achieved better estimation accuracy (R^2 = 0.86, RMSE = 178.37 g/m^2, MAE = 127.34 g/m^2) than models that did not use these features (R^2 = 0.64, RMSE = 286.79 g/m^2, MAE = 236.49 g/m^2).

Other studies have also demonstrated the importance of considering features that are not directly related to monitoring biomass. Borra-Serrano et al. (2019) also took GDD as a meteorological feature and obtained the best estimate by combining CH, VIs and meteorological data variables in an MLR model (R^2 = 0.81) to monitor ryegrass dry-matter biomass. Devia et al. (2019) also mentioned the influence of solar elevation angle and indicated that weather conditions (sunny and cloudy) can affect the quality of the data, especially in lowland crops where moisture reflection changes the appearance of the image. The above studies showed that the accuracy of biomass estimation can be improved by considering meteorological characteristics and solar elevation angle. More sample points must be obtained from multiple research sites and under different environmental conditions in future studies to train a more robust multivariate model (Liu et al., 2019). Summary of relevant studies are shown in **Table 2**.

CONCLUSION AND FUTURE PERSPECTIVES

As a high precision, high flexibility and nondestructive remote sensing system, UAS have been widely used to monitor crop biomass. The application of UAS in the monitoring of crop biomass in recent years was reviewed in this article. Four kinds of data acquisition equipment (LiDAR, RGB sensor, multispectral sensor, and hyperspectral sensor), two biomass indices (VIs and CH) and three data analysis methods (SVR, RFR, and ANN) were introduced.

Despite the rapid progress in this area, difficulties remain. First, we need to improve the speed of data acquisition and processing. Although multisensor data fusion improves the accuracy of evaluation, it makes the process of data collection more complex, data sorting more difficult, and objectively reduces the speed of monitoring. In addition, although advanced algorithms improve the evaluation accuracy, they require a long training time. Second, there is no universal method that can be applied to all crops in all cases. Different crops, even the same crops in different environments, have different characteristics. This difference requires us to carefully distinguish the characteristics of crops, use appropriate sensors to collect characteristics, and test multiple indices to determine the best biomass indices. Third, the high cost of equipment hinders the large-scale use of UAS in crop biomass monitoring. Although research on low-cost sensors has appeared, the method that is needed to improve the estimation accuracy when using low-cost sensors still needs further research. It is predicted that adopting multi-index fusion, considering features not directly related to monitoring biomass, and the adoption of advanced algorithms can effectively improve the monitoring effect of low-cost sensors on crop biomass, which is the future development direction.

Because of its high precision, flexibility and nondestructive nature, UAS have the potential to become an important method for the monitoring of crop biomass. Crop biomass monitoring systems based on multisensor fusion and multi-index fusion, the consideration of features that are not directly related to biomass monitoring and the adoption of advanced algorithms are effective methods and development directions to improve the accuracy of crop biomass estimation by UAS. Because of their low cost, using RGB sensors have become an effective method to promote the large-scale application of UAS in the field of crop biomass monitoring. In the field of biomass monitoring, UAS still have great attraction, and there are an increasing number of studies on the monitoring of crop biomass based on UAS. Furthermore, it is expected that the challenges of UAS promotion will be overcome in the future, which is conducive to the realization of smart agriculture and precision agriculture.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The work in this article was supported by the Key Research and Development Program of Nanning (20192065), the National Natural Science Foundation for Young Scientists of China (31801804), the projects subsidized by special funds for Science Technology Innovation and Industrial Development of Shenzhen Dapeng New District (Grant No. PT202001-06), and the National Innovation and Entrepreneurship Training Plan for College Students (201910593010).

ACKNOWLEDGMENTS

The authors thank native English speaking expert from the editing team of American Journal Experts for polishing our article.

REFERENCES

Acorsi, M. G., Miranda, F. D. A., Martello, M., Smaniotto, D. A., and Sartor, L. R. (2019). Estimating biomass of black oat using UAV-based RGB imaging. *Agronomy Basel* 9:14. doi: 10.3390/agronomy9070344

Alheit, K. V., Busemeyer, L., Liu, W., Maurer, H. P., Gowda, M., Hahn, V., et al. (2014). Multiple-line cross QTL mapping for biomass yield and plant height in triticale (x *Triticosecale* Wittmack). *Theor. Appl. Genet.* 127, 251–260. doi: 10.1007/s00122-013-2214-6

Atzberger, C., Guerif, M., Frederic, B., and Werner, W. (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric.* 73, 165–173. doi: 10.1016/j.compag.2010.05.006

Awodele, O., and Jegede, O. (2013). "Neural networks and its application in engineering," in *Proceedings of the Insite: Informing Science + It Education Conference* (Macon, USA).

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Observ. Geoinform.* 39, 79–87. doi: 10.1016/j.jag.2015.02.012

Borra-Serrano, I., De Swaef, T., Muylle, H., Nuyttens, D., Vangeyte, J., Mertens, K., et al. (2019). Canopy height measurements and non-destructive biomass estimation of *Lolium perenne* swards using UAV imagery. *Grass Forage Sci.* 74, 356–369. doi: 10.1111/gfs.12439

Boschetti, M., Bocchi, S., and Brivio, P. A. (2007). Assessment of pasture production in the Italian Alps using spectrometric and remote sensing information. *Agric. Ecosyst. Environ.* 118, 267–272. doi: 10.1016/j.agee.2006.05.024

Brede, B., Lau, A., Bartholomeus, H. M., and Kooistra, L. (2017). Comparing RIEGL RiCOPTER UAV LiDAR derived canopy height and DBH with terrestrial LiDAR. *Sensors* 17:16. doi: 10.3390/s17102371

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., et al. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods Ecol. Evol.* 6, 198–208. doi: 10.1111/2041-210x.12301

Calou, V. B. C., Teixeira, A. D., Moreira, L. C. J., Neto, O. C. D., and da Silva, J. A. (2019). ESTIMATION OF MAIZE BIOMASS USING UNMANNED AERIAL VEHICLES. *Engenharia Agricola* 39, 744–752. doi: 10.1590/1809-4430-Eng.Agric.v39n6p744-752/2019

Cen, H. Y., Wan, L., Zhu, J. P., Li, Y. J., Li, X. R., Zhu, Y. M., et al. (2019). Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. *Plant Methods* 15:16. doi: 10.1186/s13007-019-0418-8

Chao, Z., Liu, N., Zhang, P., Ying, T., and Song, K. (2019). Estimation methods developing with remote sensing information for energy crop biomass: a comparative review. *Biomass Bioenergy* 122, 414–425. doi: 10.1016/j.biombioe.2019.02.002

Costa, L., Nunes, L., and Ampatzidis, Y. (2020). A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms. *Comput. Electron. Agric.* 172:13. doi: 10.1016/j.compag.2020.105334

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., and McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239. doi: 10.1016/S0034-4257(00)00113-9

Devia, C. A., Rojas, J. P., Petro, E., Martinez, C., Mondragon, I. F., Patino, D., et al. (2019). High-throughput biomass estimation in rice crops using UAV multispectral imagery. *J. Intell. Robot. Syst.* 96, 573–589. doi: 10.1007/s10846-019-01001-5

Dike, H. U., Zhou, Y., Deveerasetty, K. K., and Wu, Q. (2018). "Unsupervised learning based on artificial neural Network: a review," in *Proceedings of the 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, (Shenzhen: IEEE), 322–327.

Domingo, D., Orka, H. O., Naesset, E., Kachamba, D., and Gobakken, T. (2019). Effects of UAV image resolution, camera type, and image overlap on accuracy of biomass predictions in a tropical Woodland. *Remote Sens.* 11:17. doi: 10.3390/rs11080948

Duan, B., Liu, Y. T., Gong, Y., Peng, Y., Wu, X. T., Zhu, R. S., et al. (2019). Remote estimation of rice LAI based on Fourier spectrum texture from UAV image. *Plant Methods* 15:12. doi: 10.1186/s13007-019-0507-8

Feng, L., Yue, D., and Guo, X. (2009). A review on application of normal different vegetation index. *For. Inventory Plan.* 34, 48–52. doi: 10.3969/j.issn.1671-3168.2009.02.013

Féret, J. B., Gitelson, A. A., Noble, S. D., and Jacquemoud, S. (2017). PROSPECT-D: towards modeling leaf optical properties through a complete lifecycle. *Remote Sens. Environ.* 193, 204–215. doi: 10.1016/j.rse.2017.03.004

Gitelson, A., and Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697. doi: 10.1080/014311697217558

Gitelson, A., Viña, A., Ciganda, V., Rundquist, D., and Arkebauer, T. (2005). Remote estimation of canopy chlorophyll in crops. *Geophys. Res. Lett.* 32:L08403. doi: 10.1029/2005GL022688

Gitelson, A. A. (2004). Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *J. Plant Physiol.* 161, 165–173. doi: 10.1078/0176-1617-01176

Gitelson, A. A., Gritz †, Y., and Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282. doi: 10.1078/0176-1617-00887

Gnyp, M. L., Bareth, G., Li, F., Lenz-Wiedemann, V. I. S., Koppe, W., Miao, Y., et al. (2014). Development and implementation of a multiscale biomass model using hyperspectral vegetation indices for winter wheat in the North China Plain. *Int. J. Appl. Earth Observ. Geoinform.* 33, 232–242. doi: 10.1016/j.jag.2014.05.006

Guo, Y., Senthilnath, J., Wu, W., Zhang, X., Zeng, Z., and Huang, H. (2019). Radiometric calibration for multispectral camera of different imaging conditions mounted on a UAV platform. *Sustainability* 11, 1–24. doi: 10.3390/su11040978

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., and Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. doi: 10.1016/S0034-4257(02)00018-4

Hakl, J., Hrevušová, Z., Hejcman, M., and Fuksa, P. (2012). The use of a rising plate meter to evaluate lucerne (*Medicago sativa L.*) height as an important agronomic trait enabling yield estimation. *Grass Forage Sci.* 67, 589–596. doi: 10.1111/j.1365-2494.2012.00886.x

Han, L., Yang, G. J., Dai, H. Y., Xu, B., Yang, H., Feng, H. K., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* 15:19. doi: 10.1186/s13007-019-0394-z

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* SMC-3, 610–621. doi: 10.1109/TSMC.1973.4309314

Hassan, M. A., Yang, M., Fu, L., Rasheed, A., Zheng, B., Xia, X., et al. (2019). Accuracy assessment of plant height using an unmanned aerial vehicle for quantitative genomic analysis in bread wheat. *Plant Methods* 15:37. doi: 10.1186/s13007-019-0419-7

Hassan, M. A., Yang, M., Rasheed, A., Jin, X., Xia, X., Xiao, Y., et al. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* 10:809. doi: 10.3390/rs10060809

Hassler, S. C., and Baysal-Gurel, F. (2019). Unmanned aircraft system (UAS) technology and applications in agriculture. *Agronomy Basel* 9:21. doi: 10.3390/agronomy9100618

Hogan, S., Kelly, M., Stark, B., and Chen, Y. (2017). Unmanned aerial systems for agriculture and natural resources. *Calif. Agric.* 71, 5–14. doi: 10.3733/ca.2017a0002

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. doi: 10.1016/0034-4257(88)90106-X

Jeziorska, J. (2019). UAS for wetland mapping and hydrological modeling. *Remote Sens.* 11:39. doi: 10.3390/rs11171997

Jiang, Q., Fang, S. H., Peng, Y., Gong, Y., Zhu, R. S., Wu, X. T., et al. (2019). UAV-based biomass estimation for rice-combining spectral, TIN-based structural and meteorological features. *Remote Sens.* 11:19. doi: 10.3390/rs11070890

Jiang, Z., Huete, A. R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* 112, 3833–3845. doi: 10.1016/j.rse.2008.06.006

Jibo, Y., Feng, H., Guijun, Y., and Li, Z. (2018). A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sens.* 10:66. doi: 10.3390/rs10010066

Johansen, K., Morton, M. J. L., Malbeteau, Y., Aragon, B., Al-Mashharawi, S., Ziliani, M., et al. (2019). "Predicting biomass and yield at harvest of salt-stressed tomato plants using uav imagery," in *Proceedings of the 4th ISPRS Geospatial Week 2019, June 10, 2019 – June 14, 2019*, Vol. 42, (Enschede: International Society for Photogrammetry and Remote Sensing,), 407–411.

Kang, H. W., and Kang, H. B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLoS One* 12:19. doi: 10.1371/journal.pone.0176244

Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., et al. (2019). Challenges and best practices for deriving temperature data from an uncalibrated UAV thermal infrared camera. *Remote Sens.* 11:7098. doi: 10.3390/rs11050567

Kim, J., Kim, S., Ju, C., and Son, H. I. (2019). Unmanned aerial vehicles in agriculture: a review of perspective of platform, control, and applications. *IEEE Access* 7, 105100–105115. doi: 10.1109/access.2019.2932119

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Li, S. Y., Yuan, F., Ata-Ui-Karim, S. T., Zheng, H. B., Cheng, T., Liu, X. J., et al. (2019). Combining color indices and textures of UAV-based digital imagery for rice LAI estimation. *Remote Sens.* 11:21. doi: 10.3390/rs11151763

Li, W., Niu, Z., Huang, N., Wang, C., Gao, S., and Wu, C. Y. (2015). Airborne LiDAR technique for estimating biomass components of maize: a case study in Zhangye city, Northwest China. *Ecol. Indic.* 57, 486–496. doi: 10.1016/j.ecolind.2015.04.016

Li, Z., Hu, D., Zhao, D., and Xiang, D. (2015). Research advance of broadband vegetation index using remotely sensed images. *J. Yangtze River Sci. Res. Inst.* 32, 125–130. doi: 10.3969/j.issn.1001-5485.2015.01.026

Liu, Y. N., Liu, S. S., Li, J., Guo, X. Y., Wang, S. Q., and Lu, J. W. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Comput. Electron. Agric.* 166:10. doi: 10.1016/j.compag.2019.105026

Lohani, B., and Ghosh, S. (2017). Airborne LiDAR technology: a review of data collection and processing systems. *Proc. Natl. Acad. Sci. India Sec. A Phys. Sci.* 87, 567–579. doi: 10.1007/s40010-017-0435-9

Lu, N., Zhou, J., Han, Z. X., Li, D., Cao, Q., Yao, X., et al. (2019). Improved estimation of aboveground biomass in wheat from RGB imagery and point cloud data acquired with a low-cost unmanned aerial vehicle system. *Plant Methods* 15:16. doi: 10.1186/s13007-019-0402-3

Lussem, U., Bolten, A., Menne, J., Gnyp, M. L., Schellberg, J., and Bareth, G. (2019). Estimating biomass in temperate grassland with high resolution canopy surface models from UAV-based RGB images and vegetation indices. *J. Appl. Remote Sens.* 13:034525. doi: 10.1117/1.Jrs.13.034525

Ma, L., Liu, Y., Zhang, X. L., Ye, Y. X., Yin, G. F., and Johnson, B. A. (2019). Deep learning in remote sensing applications: a meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 152, 166–177. doi: 10.1016/j.isprsjprs.2019.04.015

Maes, W. H., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007

Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., and Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* 237:20. doi: 10.1016/j.rse.2019.111599

Marino, S., and Alvino, A. (2020). Agronomic traits analysis of ten winter wheat cultivars clustered by UAV-derived vegetation indices. *Remote Sens.* 12:16. doi: 10.3390/rs12020249

Masjedi, A., Crawford, M. M., Carpenter, N. R., and Tuinstra, M. R. (2020). Multi-temporal predictive modelling of sorghum biomass using UAV-based hyperspectral and lidar data. *Remote Sens.* 12:3587. doi: 10.3390/rs12213587

Moeckel, T., Dayananda, S., Nidamanuri, R. R., Nautiyal, S., Hanumaiah, N., Buerkert, A., et al. (2018). Estimation of vegetable crop parameter by multi-temporal UAV-borne images. *Remote Sens.* 10:18. doi: 10.3390/rs10050805

Montes, J. M., Technow, F., Dhillon, B. S., Mauch, F., and Melchinger, A. E. (2011). High-throughput non-destructive biomass determination during early plant development in maize under field conditions. *Field Crops Res.* 121, 268–273. doi: 10.1016/j.fcr.2010.12.017

Navarro, J. A., Algeet, N., Fernandez-Landa, A., Esteban, J., Rodriguez-Noriega, P., and Guillen-Climent, M. L. (2019). Integration of UAV, sentinel-1, and sentinel-2 data for mangrove plantation aboveground biomass monitoring in senegal. *Remote Sens.* 11:23. doi: 10.3390/rs11010077

Nevavuori, P., Narra, N., and Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* 163:9. doi: 10.1016/j.compag.2019.104859

Niu, Y. X., Zhang, L. Y., Zhang, H. H., Han, W. T., and Peng, X. S. (2019). Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sens.* 11:21. doi: 10.3390/rs11111261

Prey, L., and Schmidhalter, U. (2019). Simulation of satellite reflectance data using high-frequency ground based hyperspectral canopy measurements for in-season estimation of grain yield and grain nitrogen status in winter wheat. *ISPRS J. Photogram. Remote Sens.* 149, 176–187. doi: 10.1016/j.isprsjprs.2019.01.023

Prost, L., and Jeuffroy, M.-H. (2007). Replacing the nitrogen nutrition index by the chlorophyll meter to assess wheat N status. *Agron. Sustain. Dev.* 27, 321–330. doi: 10.1051/agro:2007032

Qiu, P. H., Wang, D. Z., Zou, X. Q., Yang, X., Xie, G. Z., Xu, S. J., et al. (2019). Finer resolution estimation and mapping of mangrove biomass using UAV lidar and worldview-2 data. *Forests* 10:21. doi: 10.3390/f10100871

Ramon Saura, J., Reyes-Menendez, A., and Palos-Sanchez, P. (2019). Mapping multispectral digital images using a cloud computing software: applications from UAV images. *Heliyon* 5:e01277. doi: 10.1016/j.heliyon.2019.e01277

Raper, T. B., and Varco, J. J. (2015). Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status. *Precis. Agric.* 16, 62–76. doi: 10.1007/s11119-014-9383-4

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. doi: 10.1016/0034-4257(95)00186-7

Rose, T., and Kage, H. (2019). The contribution of functional traits to the breeding progress of central-european winter wheat under differing crop management intensities. *Front. Plant Sci.* 10:1521. doi: 10.3389/fpls.2019.01521

Rouse, J., Haas, R., Schell, J., and Deering, D. (1974). Monitoring vegetation systems in the great plains with ERTS. *NASA Special Publication* 351:309.

Salas Fernandez, M. G., Becraft, P. W., Yin, Y., and Lubberstedt, T. (2009). From dwarves to giants? Plant height manipulation for biomass yield. *Trends Plant Sci.* 14, 454–461. doi: 10.1016/j.tplants.2009.06.005

Sapkal, S. D., Kakarwal, S. N., and Revankar, P. S. (2007). "Analysis of classification by supervised and unsupervised learning," in *Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, Vol. 1, (Sivakasi: IEEE), 280–284.

Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., and Dammer, K.-H. (2016). Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sens.* 8:706. doi: 10.3390/rs8090706

Schmidhuber, J. (2015). Deep learning in neural networks: an overview. *Neural Networks* 61, 85–117. doi: 10.1016/j.neunet.2014.09.003

Scotford, I. M., and Miller, P. C. H. (2004). Combination of spectral reflectance and ultrasonic sensing to monitor the growth of winter wheat. *Biosyst. Eng.* 87, 27–38. doi: 10.1016/j.biosystemseng.2003.09.009

Selbeck, J., Dworak, V., and Ehlert, D. (2010). Testing a vehicle-based scanning lidar sensor for crop detection. *Can. J. Remote Sens.* 36, 24–35. doi: 10.5589/m10-022

Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., and Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Comput. Oper. Res.* 119:17. doi: 10.1016/j.cor.2020.104926

Shentu, Y. C., Wu, C. F., Wu, C. P., Guo, Y. L., Zhang, F., Yang, P., et al. (2018). "Improvement of underwater color discriminative ability by multispectral imaging," in *Proceedings of the OCEANS 2018 MTS*, (Charleston, SC: IEEE).

Shuqin, T., Yueju, X., Yun, L., Ning, H., and Xiao, Z. (2016). Review on RGB-D image classification. *Laser Optoelectron. Progr.* 53, 29–42. doi: 10.3788/lop53.060003

Sofonia, J., Shendryk, Y., Phinn, S., Roelfsema, C., Kendoul, F., and Skocaj, D. (2019). Monitoring sugarcane growth response to varying nitrogen application rates: a comparison of UAV SLAM LIDAR and photogrammetry. *Int. J. Appl. Earth Observ. Geoinform.* 82:15. doi: 10.1016/j.jag.2019.05.011

Song, B., and Park, K. (2020). detection of aquatic plants using multispectral UAV imagery and vegetation index. *Remote Sens.* 12:16. doi: 10.3390/rs12030387

Song, P., Zheng, X., Li, Y., Zhang, K., Huang, J., Li, H., et al. (2020). Estimating reed loss caused by *Locusta migratoria* manilensis using UAV-based hyperspectral data. *Sci. Total Environ.* 719:137519. doi: 10.1016/j.scitotenv.2020.137519

Sun, Z., Jing, W., Qiao, X., and Yang, L. (2019). Identification and monitoring of blooming mikania micrantha outbreak points based on UAV remote sensing. *Trop. Geogr.* 39, 482–491.

Tait, L., Bind, J., Charan-Dixon, H., Hawes, I., Pirker, J., and Schiel, D. (2019). Unmanned aerial vehicles (UAVs) for monitoring macroalgal biodiversity: comparison of RGB and multispectral imaging sensors for biodiversity assessments. *Remote Sens.* 11:18. doi: 10.3390/rs11192332

Tao, H., Feng, H., Xu, L., Miao, M., Long, H., Yue, J., et al. (2020). Estimation of crop growth parameters using UAV based hyperspectral remote sensing data. *Sensors* 20:1296. doi: 10.3390/s20051296

ten Harkel, J., Bartholomeus, H., and Kooistra, L. (2020). Biomass and crop height estimation of different crops using uav-based lidar. *Remote Sens.* 12:18. doi: 10.3390/rs12010017

Tewes, A., and Schellberg, J. (2018). Towards remote estimation of radiation use efficiency in maize using UAV-based low-cost camera imagery. *Agronomy Basel* 8:15. doi: 10.3390/agronomy8020016

Tian, H., Wang, T., Liu, Y., Qiao, X., and Li, Y. (2020). Computer vision technology in agricultural automation —A review. *Inform. Process. Agric.* 7, 1–19. doi: 10.1016/j.inpa.2019.09.006

Tian, J. Y., Wang, L., Li, X. J., Yin, D. M., Gong, H. L., Nie, S., et al. (2019). Canopy height layering biomass estimation model (chl-bem) with full-waveform lidar. *Remote Sens.* 11:13. doi: 10.3390/rs11121446

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., et al. (2014). Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *J. Appl. Remote Sens.* 8:083671. doi: 10.1117/1.Jrs.8.083671

Tomasi, C., and Kanade, T. (1992). Shape and motion from image streams under orthography: a factorization method. *Int. J. Comput. Vis.* 9, 137–154. doi: 10.1007/BF00129684

Toth, C., and Jozkow, G. (2016). Remote sensing platforms and sensors: a survey. *ISPRS J. Photogram. Remote Sens.* 115, 22–36. doi: 10.1016/j.isprsjprs.2015.10.004

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0

Turner, D., Lucieer, A., and Watson, C. (2012). An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (sfm) point clouds. *Remote Sens.* 4, 1392–1410. doi: 10.3390/rs4051392

Viljanen, N., Honkavaara, E., Nasi, R., Hakala, T., Niemelainen, O., and Kaivosoja, J. (2018). A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. *Agric. Basel* 8:28. doi: 10.3390/agriculture8050070

Villoslada, M., Bergamo, T. F., Ward, R. D., Burnside, N. G., Joyce, C. B., Bunce, R. G. H., et al. (2020). Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecol. Indic.* 111:13. doi: 10.1016/j.ecolind.2019.105979

Vogel, S., Gebbers, R., Oertel, M., and Kramer, E. (2019). Evaluating soil-borne causes of biomass variability in grassland by remote and proximal sensing. *Sensors* 19:4593. doi: 10.3390/s19204593

Wallace, L., Lucieer, A., Watson, C., and Turner, D. (2012). Development of a UAV-LiDAR system with application to forest inventory. *Remote Sens.* 4, 1519–1543. doi: 10.3390/rs4061519

Wang, C., Nie, S., Xi, X., Luo, S., and Sun, X. (2017). Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* 9:11. doi: 10.3390/rs9010011

Wang, D. Z., Wan, B., Liu, J., Su, Y. J., Guo, Q. H., Qiu, P. H., et al. (2020). Estimating aboveground biomass of the mangrove forests on northeast Hainan Island in China using an upscaling method from field plots, UAV-LiDAR data and Sentinel-2 imagery. *Int. Journal of Appl. Earth Observ. Geoinform.* 85:16. doi: 10.1016/j.jag.2019.101986

Wang, D. Z., Wan, B., Qiu, P. H., Zuo, Z. J., Wang, R., and Wu, X. C. (2019). Mapping height and aboveground biomass of mangrove forests on Hainan Island using UAV-LiDAR sampling. *Remote Sens.* 11:25. doi: 10.3390/rs11182156

Wang, F., Wang, F., Zhang, Y., Hu, J., Huang, J., and Xie, J. (2019). Rice yield estimation using parcel-level relative spectra variables from UAV-based hyperspectral imagery. *Front. Plant Sci.* 10:453. doi: 10.3389/fpls.2019.00453

Wijesingha, J., Moeckel, T., Hensgen, F., and Wachendorf, M. (2019). Evaluation of 3D point cloud-based models for the prediction of grassland biomass. *Int. J. Appl. Earth Observ. Geoinform.* 78, 352–359. doi: 10.1016/j.jag.2018.10.006

Wong, T., and Yang, N. (2017). Dependency analysis of accuracy estimates in k-fold cross validation. *IEEE Trans. Knowl. Data Eng.* 29, 2417–2427. doi: 10.1109/TKDE.2017.2740926

Wong, T., and Yeh, P. (2020). Reliable accuracy estimates from k-fold cross validation. *IEEE Trans. Knowl. Data Eng.* 32, 1586–1594. doi: 10.1109/TKDE.2019.2912815

Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recogn.* 48, 2839–2846. doi: 10.1016/j.patcog.2015.03.009

Xu, R., Li, C., and Paterson, A. H. (2019). Multispectral imaging and unmanned aerial systems for cotton plant phenotyping. *PLoS One* 14:e0205083. doi: 10.1371/journal.pone.0205083

Yan, W. Q., Guan, H. Y., Cao, L., Yu, Y. T., Li, C., and Lu, J. Y. (2020). A self-adaptive mean shift tree-segmentation method using UAV LiDAR data. *Remote Sens.* 1214. doi: 10.3390/rs12030515

Yang, B. H., Wang, M. X., Sha, Z. X., Wang, B., Chen, J. L., Yao, X., et al. (2019). Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. *Sensors* 19:18. doi: 10.3390/s19204416

Yang, S., Feng, Q., Liang, T., Liu, B., Zhang, W., and Xie, H. (2017). Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. *Remote Sens. Environ.* 204, 448–455. doi: 10.1016/j.rse.2017.10.011

Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y., et al. (2015). Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. *Remote Sens.* 7, 14939–14966. doi: 10.3390/rs71114939

Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., et al. (2017). Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: analysis of rf, ann, and svm regression models. *Remote Sens.* 9:309. doi: 10.3390/rs9040309

Yuan, M., Burjel, J. C., Isermann, J., Goeser, N. J., and Pittelkow, C. M. (2019). Unmanned aerial vehicle-based assessment of cover crop biomass and nitrogen uptake variability. *J. Soil Water Conserv.* 74, 350–359. doi: 10.2489/jswc.74.4.350

Yue, J., Feng, H., Jin, X., Yuan, H., Li, Z., Zhou, C., et al. (2018). A comparison of crop parameters estimation using images from UAV-mounted snapshot hyperspectral sensor and high-definition digital camera. *Remote Sens.* 10:1138. doi: 10.3390/rs10071138

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., et al. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* 9:708. doi: 10.3390/rs9070708

Yue, J. B., Yang, G. J., Tian, Q. J., Feng, H. K., Xu, K. J., and Zhou, C. Q. (2019). Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. *ISPRS J. Photogram. Remote Sens.* 150, 226–244. doi: 10.1016/j.isprsjprs.2019.02.022

Zha, H. N., Miao, Y. X., Wang, T. T., Li, Y., Zhang, J., Sun, W. C., et al. (2020). Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. *Remote Sens.* 12:22. doi: 10.3390/rs12020215

Zhang, H., and Li, D. (2014). Applications of computer vision techniques to cotton foreign matter inspection: a review. *Comput. Electron. Agric.* 109, 59–70. doi: 10.1016/j.compag.2014.09.004

Zhang, L., Shao, Z., Liu, J., and Cheng, Q. (2019). Deep learning based retrieval of forest aboveground biomass from combined lidar and landsat 8 data. *Remote Sens.* 11:1459. doi: 10.3390/rs11121459

Zhang, N., Rao, R. S. P., Salvato, F., Havelund, J. F., Moller, I. M., Thelen, J. J., et al. (2018). MU-LOC: machine-learning method for predicting mitochondrially localized proteins in plants. *Front. Plant Sci.* 9:634. doi: 10.3389/fpls.2018.00634

Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., et al. (2018). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sens.* 10:2026. doi: 10.3390/rs10122026

Zheng, H. B., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y. C., et al. (2019). Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precis. Agric.* 20, 611–629. doi: 10.1007/s11119-018-9600-7

Zhong, Y. F., Wang, X. Y., Xu, Y., Wang, S. Y., Jia, T. Y., Hu, X., et al. (2018). Mini-UAV-borne hyperspectral remote sensing from observation and processing to applications. *IEEE Geosci. Remote Sens. Mag.* 6, 46–62. doi: 10.1109/mgrs.2018.2867592

Zhu, W. X., Sun, Z. G., Peng, J. B., Huang, Y. H., Li, J., Zhang, J. Q., et al. (2019). Estimating maize above-ground biomass using 3d point clouds of multi-source unmanned aerial vehicle data at multi-spatial scales. *Remote Sens.* 11:22. doi: 10.3390/rs11222678

Zhu, Y. H., Zhao, C. J., Yang, H., Yang, G. J., Han, L., Li, Z. H., et al. (2019). Estimation of maize above-ground biomass based on stem-leaf separation strategy integrated with LiDAR and optical remote sensing data. *PeerJ* 7:30. doi: 10.7717/peerj.7593

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Wang, Liu, Wang, Fan, Tian, Qiao and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 30 April 2021
doi: 10.3389/fpls.2021.626516

Hyperspectral Imaging for Identification of an Invasive Plant *Mikania micrantha* Kunth

Yiqi Huang[1]†, Jie Li[1,2]†, Rui Yang[1,2]†, Fukuan Wang[1,2], Yanzhou Li[1], Shuo Zhang[3], Fanghao Wan[2]*, Xi Qiao[1,2,4]* and Wanqiang Qian[2]*

[1] College of Mechanical Engineering, Guangxi University, Nanning, China, [2] Lingnan Guangdong Laboratory of Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Area, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, [3] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, China, [4] Guangzhou Key Laboratory of Agricultural Products Quality & Safety Traceability Information Technology, Zhongkai University of Agriculture and Engineering, Guangzhou, China

OPEN ACCESS

Edited by:
Kioumars Ghamkhar,
AgResearch Ltd, New Zealand

Reviewed by:
Yelin Huang,
Sun Yat-sen University, China
Shen Shicai,
Yunnan Academy of Agricultural Sciences, China
Abhimanyu Singh Garhwal,
AgResearch Ltd, New Zealand

Correspondence:
Wanqiang Qian
qianwanqiang@caas.cn
Xi Qiao
qiaoxi@caas.cn
Fanghao Wan
wanfanghao@caas.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *06 November 2020*
Accepted: *08 April 2021*
Published: *30 April 2021*

Citation:
Huang Y, Li J, Yang R, Wang F, Li Y, Zhang S, Wan F, Qiao X and Qian W (2021) Hyperspectral Imaging for Identification of an Invasive Plant Mikania micrantha Kunth. Front. Plant Sci. 12:626516. doi: 10.3389/fpls.2021.626516

Mile-a-minute weed (*Mikania micrantha* Kunth) is considered as one of top 100 most dangerous invasive species in the world. A fast and accurate detection technology will be needed to identify *M. micrantha*. It will help to mitigate the extensive ecologic and economic damage on our ecosystems caused by this alien plant. Hyperspectral technology fulfills the above requirement. However, when working with hyperspectral images, preprocessing, dimension reduction, and classifier are fundamental to achieving reliable recognition accuracy and efficiency. The spectral data of *M. micrantha* were collected using hyperspectral imaging in the spectral range of 450–998 nm. A different combination of preprocessing methods, principal component analysis (for dimension reduction), and three classifiers were used to analyze the collected hyperspectral images. The results showed that a combination of Savitzky-Golay (SG) smoothing, principal component analysis (PCA), and random forest (RF) achieved an accuracy (A) of 88.71%, an average accuracy (AA) of 88.68%, and a Kappa of 0.7740 with an execution time of 9.647 ms. In contrast, the combination of SG, PCA and a support vector machine (SVM) resulted in a weaker performance in terms of A (84.68%), AA(84.66%), and Kappa (0.6934), but with less execution time (1.318 ms). According to the requirements for specific identification accuracy and time cost, SG-PCA-RF and SG-PCA-SVM might represent two promising methods for recognizing *M. micrantha* in the wild.

Keywords: hyperspectral analysis, invasive plant, data preprocessing, dimension reduction, classification

INTRODUCTION

Mikania micrantha Kunth (*M. micrantha*), also known as "mile-a-minute," is one of the world's 100 most dangerous invasive species (Khadka, 2017). It is estimated that *M. micrantha* can produced between 90,000 and 210,000 seeds/m^2 (Macanawai et al., 2012; Day et al., 2016). The seeds are dispersed by wind, animals, and humans (Yang et al., 2005; Day et al., 2016). In China, *M. micrantha* achieved an average growth rate of 6–7 cm/day (Zhang et al., 2004; Day et al., 2016). The ecological environment has been seriously damaged, the biodiversity has been threatened, and the economy has been influenced by this weed (Shen et al., 2017). The yield losses of banana (*Musa* spp.), *Citrus* spp., and sugarcane (*Saccharum officinarum* L.) infested with *M. micrantha* ranged from 60 to 70% due to the twining which would block out sunlight (Shen et al., 2013). The economic losses were estimated at US$650,000–1.6 M/year on Neilingding Island (about 554 ha; Zhong et al., 2004). Therefore, identifying and monitoring *M. micrantha* are urgent, which would allow the plant to be controlled by providing accurate information about its geographical distribution (Tesfamichael et al., 2018).

Currently, monitoring *M. micrantha* mainly relied on manual inspection, which is labor intensive and inefficient (Day et al., 2012; Nath et al., 2019). Hyperspectral remote sensing is an efficient monitoring method that has been successfully used to monitor many alien invasive plants (Calvino-Cancela et al., 2014; Sabat-Tomala et al., 2020) and has shown great potential (Chance et al., 2016; Marcinkowska-Ochtyra et al., 2018). In these researches, researchers pay attention to analyzing the raw spectral data characteristics of target invasive plants, extracting spectral signature of the plants, and classifying the features of the plants (Masemola et al., 2019). Some methods, such as random forest (RF), support vector machine (SVM), and their improvements, have been applied for the classification of invasive plants and have achieved good results (Aneece and Epstein, 2017; Grosse-Stoltenberg et al., 2018; Tarantino et al., 2019). It is undeniable that the usage of some spectral wavebands from captured spectral wavebands of hyperspectral images may cause the loss of important spectral information. Nevertheless, the usage of full multispectral bands may cause information redundancy and interference. Therefore, it is necessary to find the balance between them by preprocessing the raw spectral data.

Some state-of-the-art spectral preprocessing detection methods have been proposed by researchers. Liu et al. (2019) used different preprocessing methods to extract hyperspectral reflectance characteristics. A Savitzky-Golay (SG) smoothing of the reflectance spectrum was performed, and the first derivative (FD), the second derivative (SD), and reciprocal logarithm transformation were performed on preprocessed reflectance data by multiple scatter correction and standard normal variate (SNV). The preprocessing methods above have enabled the optimal estimation model to gain better stability and higher precision. To effectively eliminate the noise and baseline hyperspectral drifting, Zhou et al. (2019) proposed a combination of FD, SD, and wavelet transform prepossessing on raw spectral data. Their model achieved 98.57% accuracy in prediction set. Yang et al. (2018) explored the effects of different pretreatment methods on the FT-MIR spectra detection of *Panax notoginseng*, where the best preprocessing combination for the collected spectra was a mix of baseline correction, SNV and FD with an 11 point smoothing. The above preprocessing methods were optimized based on the full-band raw spectral data set and retained all the information of the raw spectral data. However, the calculation workload and time will be increased (Xu et al., 2019). Therefore, this type of method needs to reduce the dimensions of feature sets and keeps most of the dataset information (Luo et al., 2019).

Currently, the methods for reducing the dimension of extracted spectral data from hyperspectral images mainly include feature extraction based on transformation (Du et al., 2018) [e.g., principal component analysis (PCA)] and feature selection based on non-transformation (Salimi et al., 2018; e.g., algorithms for selecting local feature bands). Peerbhay et al. (2015) used hyperspectral remote sensing for the detection and mapping of *Solanum mauritianum* located within commercial forestry ecosystems. This method, based on an RF and PCA, achieved a detection rate of 95% with a false positive rate of 6.39%. Orrillo et al. (2019) used PCA and a classification model preprocessed by an SNV and an SD to identify black pepper adulterated with common adulterant papaya seeds in near-infrared hyperspectral imaging and achieved 100% accuracy in the classification of berry samples. Aneece and Epstein (Aneece and Epstein, 2015) used PCA processed raw spectral data to distinguish among invasive-dominated successional plant communities in the wild. It indicates that different plant species could be identified using spectral information. The previous studies suggest that PCA has been effectively used to reduce raw spectral data dimension, thereby significantly increasing efficiency.

The literature review shows that hyperspectral identification is a potential method for accurate monitoring of *M. micrantha.* Generally, the level of hyperspectral identification can be generally improved only if preprocessing, the feature dimension reduction technique, and the classifier are all addressed (Qiao et al., 2018). Moreover, challenges are manifested in the variability of the raw spectral data of *M. micrantha* in a complex field environment, the lack of prior knowledge and background interference. To address these challenges, hyperspectral preprocessing algorithms [such as FD, SD, nine-point (9P) smoothing, SG smoothing, and SNV], a feature selection algorithm (PCA), and classification algorithms [such as RF, SVM, back propagation neural network (BPNN)] (Vetrekar et al., 2015; Qi et al., 2017) have been proposed, in combination, to recognize *M. micrantha* in wild environments, and an accurate and fast method will be chosen.

MATERIALS AND METHODS

Sample Preparation

A high-speed imaging spectrograph S185 manufactured by the German company Cubert was used to manually collect the *M. micrantha* hyperspectral images in the wild. The spectrometer weighs 470 g, uses DC12V power, and can obtain 138 spectral wavebands with a 4-nm sampling interval in the spectrum range of 450 to 998 nm. The collection site was a desolate field of farmland near the Xinnan subway station in Jiulong town, Guangzhou city, China (23°22′29.5′′ north latitude and 13°29′52.9′′ east longitude). The collection time was approximately 9:30 on November 21, 2018; the weather was cloudy. Before image acquisition dark reference (by closing the camera lens) and white reference (using a white plate) images

FIGURE 1 | Representative hyperspectral images of *Mikania micrantha* and background.

were collected to calibrate the spectrometer according to the following equation:

$$I_C = (I_R - I_D)/(I_W - I_D) \tag{1}$$

where I_C is the calibrated image, I_R is the raw image, I_W is the white reference, and I_D is the dark reference.

Then the lens of spectrograph were pointed directly toward at the surface of the plant, and manually focused on the middle of *M. micrantha* leaves. Eighteen hyperspectral images were collected over vegetation using the S185 spectrometer and used for this work. Six samples of the eighteen hyperspectral images contained the leaves and flowers of *M. micrantha*, other plants, and non-plant background are shown in **Figure 1**. An individual scan time was very short (less than 1 min), and all scans were basically carried out in an area of about 300 square meters. Therefore, it was made within half an hour, and illumination changes from scan to scan varied little.

The raw spectral data of *M. micrantha* and background were manually extracted from hyperspectral images by Cubeware software (**Figures 2A,B**), care was taken to avoid any cross-class contamination, and saved in ASCII format. 745 raw spectral data samples (*M. micrantha*: 377, background: 368) were collected and randomly divided into a training set (*M. micrantha*: 251, background: 245), a testing set (*M. micrantha*: 63, background: 62), and a validation set (*M. micrantha*: 63, background: 61). The labels of *M. micrantha* consisted of *M. micrantha* leaves and flowers, and the labels of background included leaves and flowers of other plants, as well as non-plant background. The training, testing, and validation sets were balanced to prevent bias in the classifiers and metrics.

Methods

Five preprocessing methods, one feature selection method and three classifiers were combined and implemented to process and classify extracted raw spectral data, respectively. The framework of the proposed methods to recognize *M. micrantha* and choose the optimal model is illustrated in **Figure 3**.

Preprocessing

Smoothing is widely used to eliminate the interference of high-frequency noise in raw spectral data and to improve the spectral signal-to-noise ratio (Saberioon et al., 2019). In this study, 9P smoothing, and SG smoothing were used to smooth the raw spectral data. 9P smoothing can reduce the noise by calculating the average value of a set of sample raw spectral data in the moving smoothing window. The smoothing procedure is as follows. First, the window size was determined to be nine in this article (Lawrence et al., 2006). Second, nine consecutive points on the raw spectral data ($x_{-4}, \cdots, x_{-1}, x_0,\ x_1, \cdots, x_4$) were selected. Then, the arithmetic mean was computed and assigned to x_0. Finally, the window was moved to the next point so that the center of the window traverses the whole raw spectral data. Similar to 9P smoothing, SG smoothing is a filtering method based on least squares polynomial fitting in a moving window. The window size was set as five in this article (Fu et al., 2018; Liu et al., 2019).

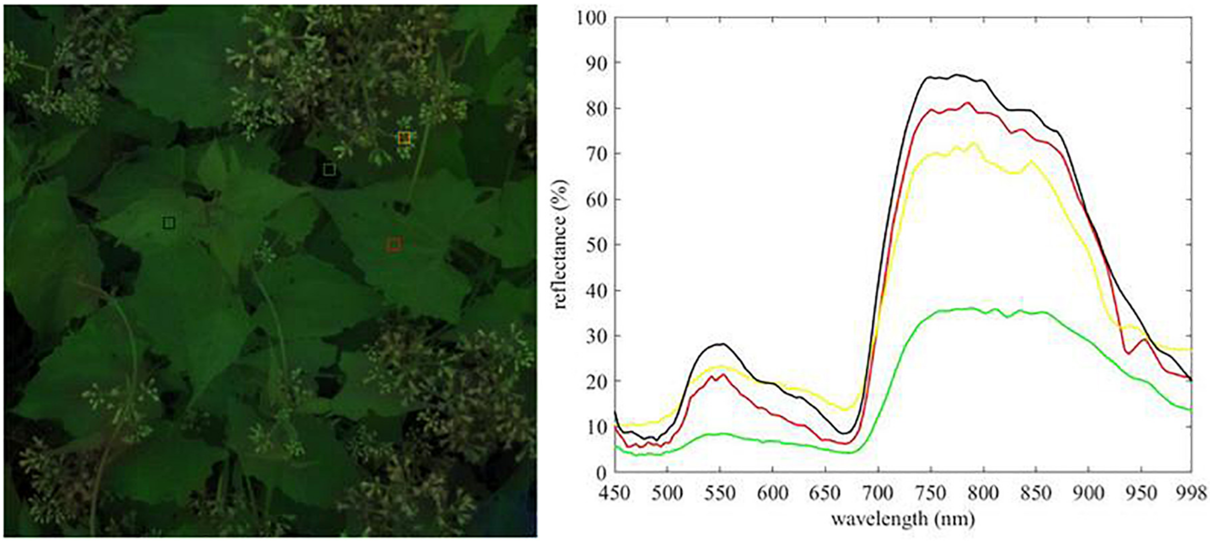


FIGURE 2 | Extraction of raw spectral data. **(A)** Pixels of hyperspectral images used for raw spectra extraction, **(B)** raw spectral data.

The derivative processing can correct the data far away from the zero plane (Saberioon et al., 2019), thus effectively reducing noise interference, suppressing useless information and highlighting the information of interest. This method is a good choice to deal with noise interference in raw spectral data. First derivative and SD are selected to preprocess extracted raw spectral data. The FD is sensitive to noise and can show the change of reflectivity (Golhani et al., 2019). The SD highlights subtle variations in the spectrum and is suitable for optimal wavelength selection (Wu et al., 2018).

The performance of SNV processing was tested on raw spectral data of *M. micrantha*. SNV assumes the reflection values of each wavelength in the raw spectral data to meet a certain distribution (Yang et al., 2018; Liu et al., 2019), thereby eliminating the errors caused by particle size difference between samples, spectral transformation and surface scattering (Asaari et al., 2018).

Dimension Reduction

Raw and preprocessed spectral data have 138 wavebands that makes the feature sets high-dimensional. If the set is directly used for *M. micrantha* target recognition, the calculation workload and time will be increased. Therefore, there was a need to reduce the dimension of the feature set and to keep most of the dataset information. PCA is a transform-based feature extraction method. In this work, PCA was used to transform the raw and preprocessed spectral data, the original high-dimensional raw and preprocessed spectral data were transformed into new comprehensive variable data, while keeping most of the information from the original spectral data (Jeyakumar and Sudha, 2019; Tian et al., 2020).

Classification

In the case of limited training samples, the robustness of SVM and RF in processing high-dimensional data makes them suitable for raw and preprocessed hyperspectral data (Tusa et al., 2020). SVM transforms low-dimensional linear inseparable samples into a high-dimensional feature space to make them linearly separable. Based on structural risk minimization, the optimal classification hyperplane is constructed in the feature space to obtain the global optimal solution (Cortes and Vapnik, 1995). RF begins by generating many trees and then votes for the most popular class. This method is an effective tool for classification because each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001).

A BPNN is a multi-layer feedforward network trained by error back propagation. The network takes the sum of error squares as the objective function, and the minimum value of the objective function is calculated by gradient descent method. The commonly used BPNN contains an input layer, an implicit layer, and an output layer. When sufficient training samples are available, the trained BPNN can identify complex objects with high accuracy (Vetrekar et al., 2015; Yao et al., 2019).

To choose the optimal model which is accurate and fast to identify the *M. micrantha* developed from spectral data of hyperspectral image, the three classifiers were tested via different combinations of PCA dimension reduction and other preprocessing methods.

Evaluation

To evaluate the performance of each of the proposed methods, four statistical parameters, namely, accuracy (A), average accuracy (AA), the Kappa value (Dash et al., 2019) and time, were considered. These parameters are frequently used for performance evaluation in classification problems (Xu et al., 2019). The parameters were calculated from below equations (2), (3), (4), and (5):

$$\text{Accuracy} \quad = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \qquad (2)$$

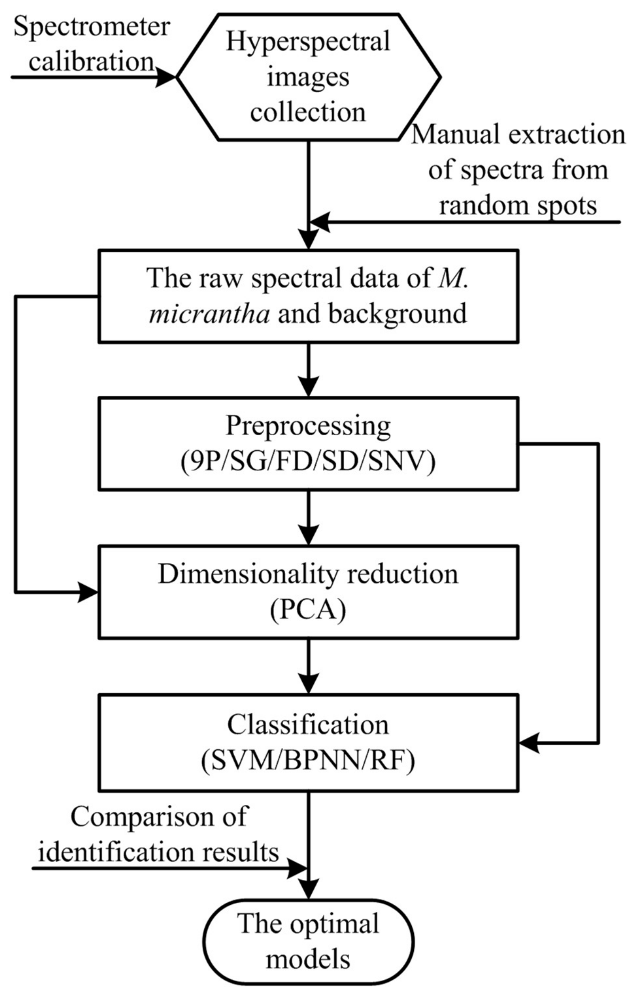


FIGURE 3 | Framework of the proposed methods implementation.

$$\text{Average Accuracy} = \text{TP}/\left[2\left(\text{TN}+\text{FP}\right)\right]+\text{TN}/\left[2\left(\text{TN}+\text{FP}\right)\right] \tag{3}$$

$$\text{Kappa} = \left(\text{Accuracy}-P_e\right)/\left(1-P_e\right) \tag{4}$$

$$P_e = \left[\left(\text{TP}+\text{FN}\right)\left(\text{TP}+\text{FP}\right)+\left(\text{TN}+\text{FP}\right)\left(\text{TN}+\text{FN}\right)\right]/\left(\text{TP}+\text{TN}+\text{FP}+\text{FN}\right)^2 \tag{5}$$

where TP is the number of samples correctly predicted to be *M. micrantha*, TN is the number of samples correctly predicted as the background, FP is the number of background samples incorrectly predicted as *M. micrantha*, and FN is number of *M. micrantha* samples incorrectly predicted to be the background.

Also, we summarized the computational time required by each method to recognize the samples in the validation set. All the aforementioned methods were coded and developed in MATLAB R2019a (The Math Works Inc., United States). The CPU of the PC is Intel(R) Core(TM) i7-7700, and the RAM is 16 GB.

RESULTS

Preprocessing

Raw and preprocessed spectral data of *M. micrantha* and background were presented in **Figures 4A–F**. In **Figure 4A**, *M. micrantha* spectral reflectance is slightly higher than the background in about 450–670 nm range, while parts of *M. micrantha* have the same reflectance as the background. In about 750–880 nm range, the reflectance of *M. micrantha* and background are scattered, and the reflectance distribution of the background basically overlapped with that of *M. micrantha*. The raw spectral data distribution of *M. micrantha* in the remaining spectral range is almost the same as the background. It indicates that the intra-class differences were more than inter-class differences of *M. micrantha* and background, and it is a challenging work for *M. micrantha* identification. In order to be more conducive to the recognition of raw spectral data, five kinds of preprocessing methods were used to eliminate data noise or highlight the distribution law of reflectance with wavelength. The raw spectral data preprocessed using the two smoothing methods are shown in **Figures 4B,C**. Compared with the raw spectral data in **Figure 4A**, the small fluctuations of reflectance over the entire wavelength range (e.g., 450–500 nm) are eliminated or changed more gently. The other three preprocessing methods remove other noises from raw spectral data. The direct analysis of the raw spectral data after derivation is illustrated in **Figures 4D,E**. The FD and SD were constant states (horizontal line) at both ends of the spectral band (450–470 nm and 978–998 nm), the relevant raw spectral data were obviously polluted by the system noise. In the other spectral range, the intra-class differences of the preprocessed spectral data were smaller than the raw spectral data. **Figure 4F** shows the preprocessed data of SNV. The intra-class differences of the preprocessed spectral data became smaller, especially the preprocessed spectral data in the 670–880 nm range. And the spectral noise was relatively reduced, too.

In summary, all the five preprocessing methods can eliminate part of the spectral noise. FD, SD, and SVN can significantly reduce the intra-class differences, however, the inter-class differences were not significantly improved by all pretreatments. Therefore, it is necessary to find the difference in the raw and preprocessed spectral data between *M. micrantha* and the background through subsequent processing. To determine the most suitable preprocessing method, the next step was to analyze the influence of each preprocessing method combining the dimension reduction and classifiers on the performance of *M. micrantha* identification.

Dimension Reduction by PCA

Principal component analysis was performed on the raw and preprocessed spectral data. In general, the first and second principal components have the maximum variation of the

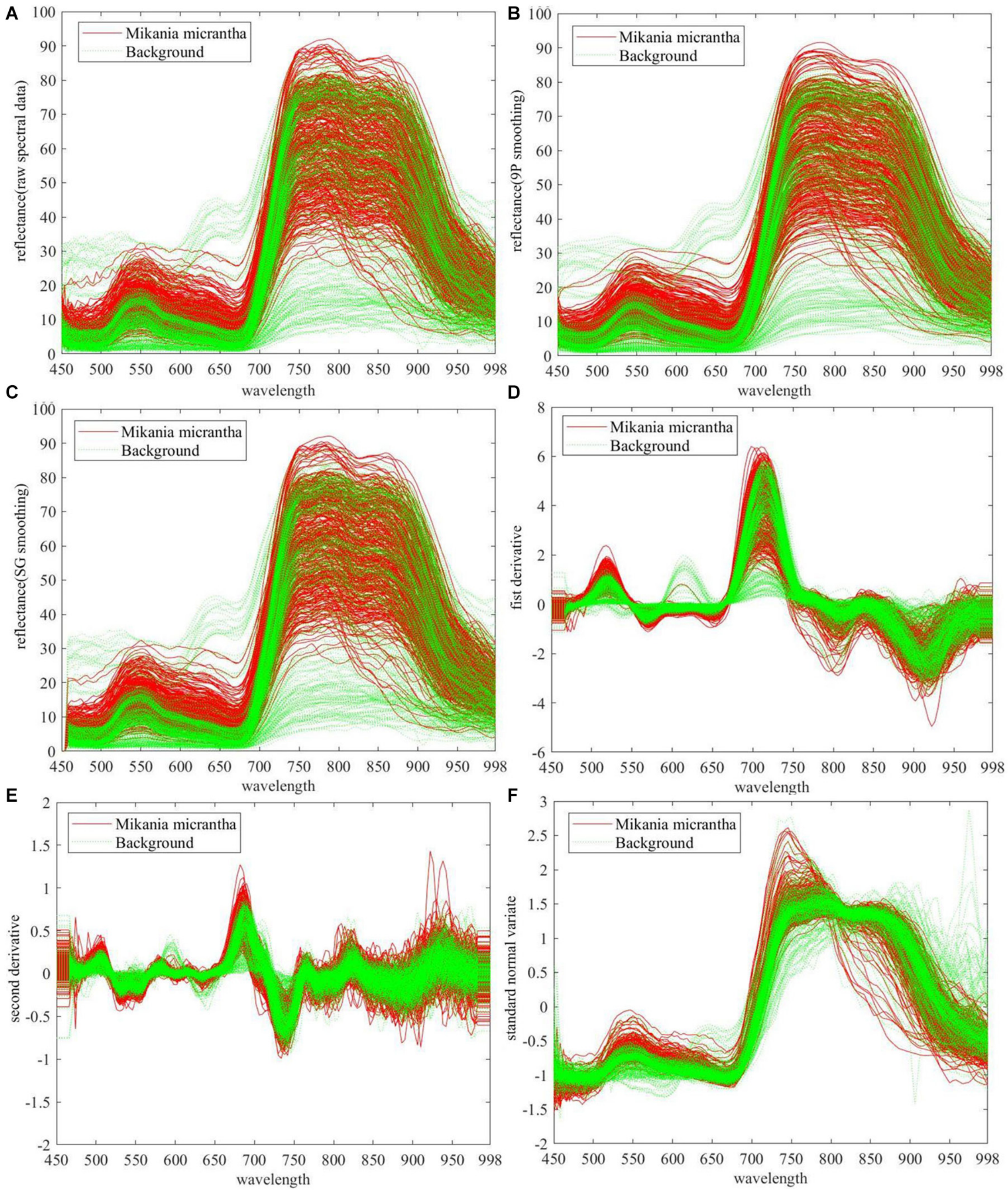


FIGURE 4 | Raw and preprocessed spectral data of 745 samples **(A)** raw spectral data, preprocessed by **(B)** 9P smoothing, **(C)** SG smoothing, **(D)** first derivative, **(E)** second derivative, and **(F)** standard normal variate.

original data. The first and second principal component scores of 745 samples were depicted as **Figure 5**.

Each of the datasets was clustered and distributed with the origin of the coordinates as the center. **Figure 5** shows the impact of each preprocessing method on sample clustering. Through comparison of the raw spectral data (**Figure 5A**) with the raw spectral data preprocessed by different methods, we found that the raw spectral data preprocessed by the SD (**Figure 5E**) showed the best clustering effect. In terms of clustering performance, the SNV (**Figure 5F**) ranked second, and the FD (**Figure 5D**) ranked third. Compared with the degree of clustering of the raw spectral data (**Figure 5A**), the clustering effect of the two smoothing treatments was the worst (**Figures 5B,C**). Nevertheless, the above results are predictable

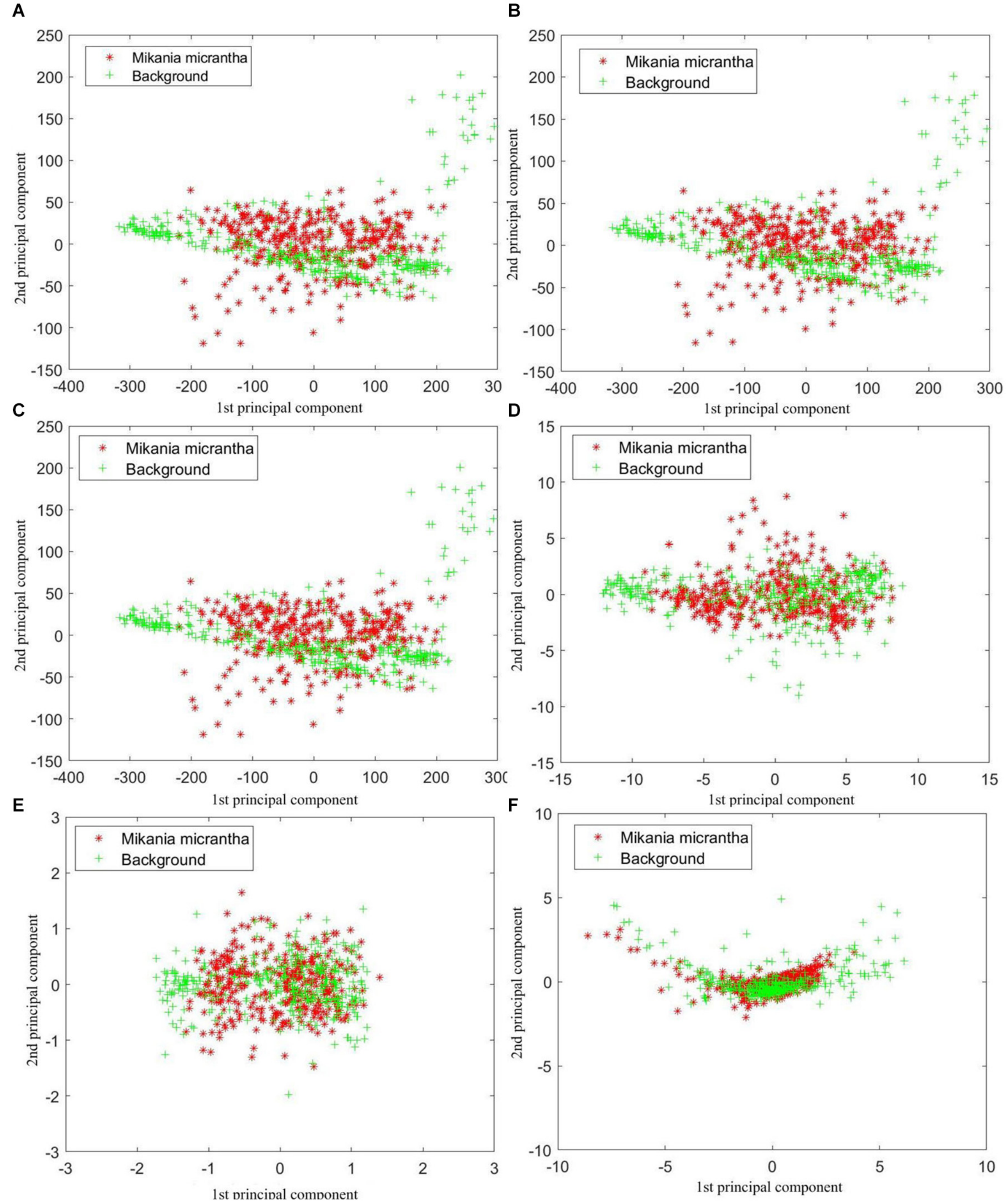


FIGURE 5 | First and second principal component scores of 745 samples based on **(A)** raw spectral data, preprocessed by **(B)** 9P smoothing, **(C)** SG smoothing, **(D)** first derivative, **(E)** second derivative, and **(F)** standard normal variate.

because the two smoothing treatments reduced the noise of the raw spectral data but did not change the details and the overall trend. In addition, the overlap between the two types of samples was obvious as shown in **Figure 5**. Thus, more principal components need to be taken into account.

As shown in **Figure 6**, the cumulative contribution rates of the first k $\left(k = 1, 2, \ldots, 138\right)$ principal components were also calculated. The raw spectral data, 9P smoothing, and SG smoothing had almost the same curves, and at approximately the first 5 principal components, all curves tended to be smooth

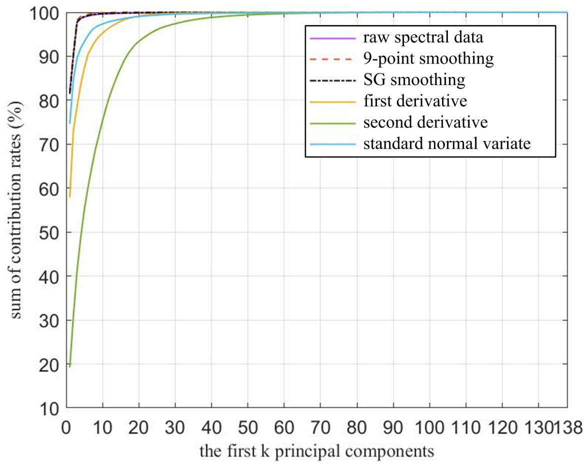


FIGURE 6 | Cumulative contribution rate curve of the principal components.

and close to 100%. Thus, all the curves are able to fully represent the information contained in the 138-dimensional raw and preprocessed spectral data. Although the contribution rate of the first principal component of the FD was approximately 20% lower than that of the SNV, the FD, and SNV did not have almost the same cumulative contribution rates until the first 18 principal components, where the values are close to 99%. The SD had the lowest contribution rate of the first principal component, and the cumulative contribution rates were not more than 99% until top 40 principal components. Therefore, the first *k* principal components based on different preprocessing contained most of the information. To achieve accurate *M. micrantha* identification, the first *k* principal components were selected as the input of the classification model. However, the *k* values were dependent on the preprocessing and classification algorithms combined with PCA.

Recognition Performance Assessment Based on Different Combinatorial Algorithms

Among the 745 samples, 496 samples were used to train the models, and 125 samples were used to test the trained models. The SVM, the BPNN, and the RF were separately trained based on the first k $(k = 1, 2, \ldots, 138)$ principal components, which were generated by employing PCA on raw and preprocessed spectral data. The recognition accuracies of the first *k* principal components are shown in **Figure 7**.

The results showed that the accuracy of the RF was significantly higher than the other two methods. When $k > 10$, the accuracies of the adjacent first *k* principal components fluctuated within the smallest range. The SVM was the second most accurate; its accuracy first increased and then decreased with the increase of *k*. The BPNN did not achieve good results, and the accuracies of adjacent *k* fluctuated within a large range, although the accuracy was higher than that of the SVM as the *k* increased to a certain degree. Using the same classification method, the accuracies of SG-PCA-RF, FD-PCA-RF, and SD-PCA-RF methods were more significantly improved than the accuracy of OR-PCA-RF, but 9P-PCA-RF did not achieve much improvement except for when the first *k* principal component was between 70 and 80. The accuracy of SNV-PCA-RF improved as the *k* increased, but there was little benefit in terms of dimension reduction. In **Figures 7C–E**, the maximum accuracy of the RF appeared at *k* between 10 and 20, and the accuracies of SD-PCA-RF were higher than those of SG-PCA-RF and FD-PCA-RF. In addition, compared with OR-PCA-SVM, the combinations of the other preprocessing methods with PCA and SVM did not improve accuracy obviously. And the *k* values corresponding to the maximum accuracy of the other combination methods were higher than that of OR-PCA-SVM, such as 9P-PCA-SVM, SG-PCA-SVM, and FD-PCA-SVM. For the BPNN algorithm, 9P smoothing and SG smoothing improved the accuracy and reduced the dimensions, while the *k* values corresponding to the maximum accuracy were located at between 1 and 10. However, the accuracy of other preprocessing methods was lower than the raw, thereby having even negative effects.

Therefore, not all combinations of preprocessing methods, PCA and classification methods improved accuracy while reducing the dimensions. To reduce the dimensions and improve accuracy, the first *k* principal components corresponding to the maximum accuracy of each combination method were separately confirmed to reduce the dimensions (**Table 1**) and to verify the recognition performance.

The remaining 124 samples were used as the validation set to verify the corresponding trained models at the first *k* principal components as shown in **Table 1**. There were over 10 runs for each method. The most frequent results are shown in **Table 2**. The best results for each quality index were highlighted in bold. The result shows that SG-PCA-RF yielded the best A, AA, and Kappa values, and 9P-PCA-SVM had the shortest execution time among all the methods (**Table 2**).

DISCUSSION

In the process of exploring the identification of *M. micrantha* based on hyperspectral technology, the combinatorial test of conventional spectral data processing methods was carried out. The results showed that RF and SVM based on homologous preprocessing spectral data maintained the advantages of accuracy and time, respectively. In terms of the recognition effect of *M. micrantha*, RF shows higher accuracy and recognition consistency than the other two classifiers. Certainly, if the time indicator is the most important in practical applications, SVM is also a good choice. After all, it also has a satisfactory accuracy and consistency.

When applied to the same classifier RF, SG smoothing yielded the best A, AA, Kappa, and time values, and SD yielded the second best A, AA, and Kappa values. When applied to SVM, SG smoothing yielded the best A, AA, and Kappa values, and 9P smoothing yielded the second best A, AA, Kappa values, and best time, but the improvement was not obvious compared with OR. The above results were basically consistent with the

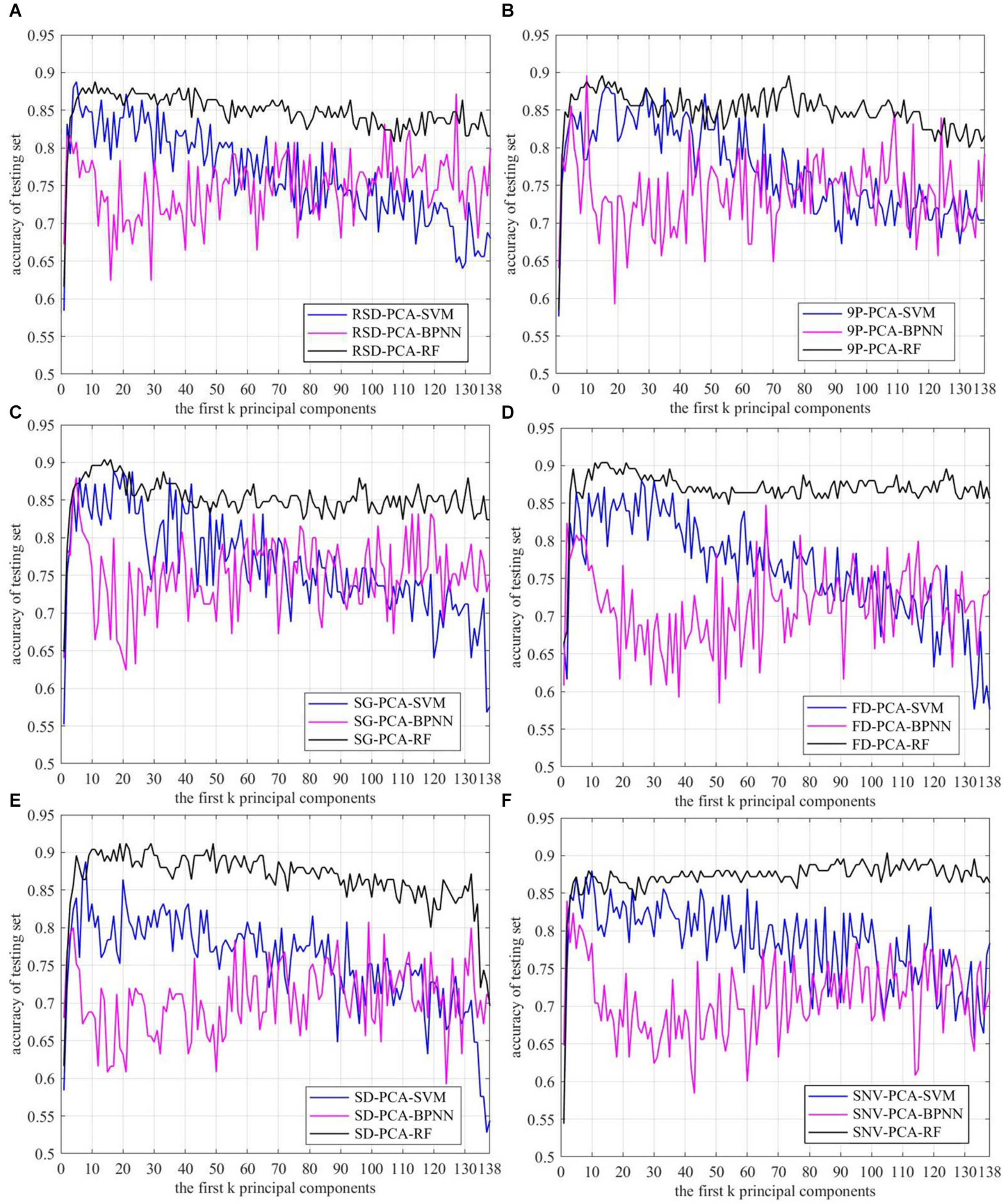


FIGURE 7 | Recognition accuracy curve of *Mikania micrantha* based on the first *k* principal components. **(A)** raw spectral data, preprocessed by **(B)** 9P smoothing, **(C)** SG smoothing, **(D)** first derivative, **(E)** second derivative, and **(F)** standard normal variate.

analysis results in **Figure 7** and **Table 1**. Overall, SG smoothing worked the best among the five common pretreatments tested during *M. micrantha* identification using hyperspectral image data. Moreover, the methods combining preprocessing with a classifier were also used for validation set recognition, without PCA. The results are shown in **Table 3**. Compared with **Table 2**, all indexes were worse in most of the cases as shown in **Table 3**. Although SNV-RF was better than SNV-PCA-RF in terms of A, AA, and Kappa, SNV-RF was still inferior to SG-PCA-RF in all indexes. Even for the BPNN classifier, which showed the

TABLE 1 | First *k* principal components corresponding to the maximum accuracy of each combination method.

***k*/max accuracy (%)**	**RSD**	**9P smoothing**	**SG smoothing**	**FD**	**SD**	**SNV**
PCA-SVM	5/88.80	16/88.00	17/88.80	26/88.00	8/88.80	10/88.00
PCA-BPNN	127/87.20	10/89.60	5/88.00	66/84.80	99/80.80	2/84.00
PCA-RF	11/88.80	15/89.60	14/90.40	11/90.40	19/91.20	105/90.40

RSD is raw spectral data.

TABLE 2 | Methods combining preprocessing, PCA, and a classifier for validation set recognition.

Methods	***k***	**Validation set**			
		A (%)	**AA (%)**	**Kappa**	**Time (ms)**
RSD-PCA-SVM	5	81.45	81.62	0.6302	0.963
RSD-PCA-BPNN	127	78.23	78.13	0.5636	6.169
RSD-PCA-RF	11	83.87	83.84	0.6772	10.275
9P-PCA-SVM	16	83.87	83.81	0.6770	**0.823**
9P-PCA-BPNN	10	75.81	75.77	0.5158	6.082
9P-PCA-RF	15	84.68	84.63	0.6932	10.028
SG-PCA-SVM	17	84.68	84.66	0.6934	1.318
SG-PCA-BPNN	5	81.45	81.46	0.6290	5.575
SG-PCA-RF	14	**88.71**	**88.68**	**0.7740**	9.647
FD-PCA-SVM	26	80.65	80.69	0.6132	1.115
FD-PCA-BPNN	66	70.97	70.99	0.4195	5.332
FD-PCA-RF	11	83.87	83.89	0.6775	10.571
SD-PCA-SVM	8	79.84	79.98	0.5978	1.014
SD-PCA-BPNN	99	72.58	72.50	0.4506	6.033
SD-PCA-RF	19	86.29	86.20	0.7252	10.653
SNV-PCA-SVM	10	81.45	81.49	0.6292	1.305
SNV-PCA-BPNN	2	82.26	82.31	0.6454	5.533
SNV-PCA-RF	105	85.48	85.51	0.7098	10.431

RSD is raw spectral data. The best results for each quality index were highlighted in bold.

TABLE 3 | Methods combining preprocessing with a classifier for validation set recognition.

Methods	**Validation set**			
	A (%)	**AA (%)**	**Kappa**	**Time (s)**
RSD-SVM	81.45	81.38	0.6285	**1.823**
RSD-BPNN	66.94	66.94	0.3387	5.963
RSD-RF	83.06	82.92	0.6602	12.665
9P-SVM	82.26	82.18	0.6445	3.146
9P-BPNN	71.77	71.86	0.4364	5.857
9P-RF	83.87	83.81	0.6770	9.616
SG-SVM	83.37	83.79	0.6768	1.900
SG-BPNN	71.77	71.86	0.4364	6.795
SG-RF	84.68	84.56	0.6927	9.350
FD-SVM	83.06	83.02	0.6609	2.049
FD-BPNN	77.42	77.39	0.5480	30.801
FD-RF	85.48	85.48	0.7096	9.329
SD-SVM	81.45	81.49	0.6292	21.212
SD-BPNN	80.65	80.64	0.6128	10.110
SD-RF	86.29	86.22	0.7254	10.605
SNV-SVM	82.26	82.33	0.6456	11.373
SNV-BPNN	71.77	71.78	0.4355	17.394
SNV-RF	**87.10**	**87.12**	**0.7420**	12.048

RSD is raw spectral data. The best results for each quality index were highlighted in bold.

worst comprehensive performance in *M. micrantha* recognition, PCA dimension reduction treatment can improve the recognition effect. However, the recognition effect fluctuated significantly with the change of the number of principal components. Therefore, PCA was able to improve the accuracy and efficiency of the algorithms in most cases.

In summary, the SG-PCA-RF (88.71% A, 88.68% AA, 0.7740 Kappa, and execution time of 9.647 ms) and SG-PCA-SVM (84.68% A, 84.66% AA, 0.6934 Kappa, and execution time of 1.318 ms) algorithms outperformed other methods for *M. micrantha* recognition. Therefore, the method should be selected according to the specific requirement for identification accuracy and time cost.

The recognition methods based on convolutional neural network (CNN) are very popular at present, however, it does not mean that these methods are applicable to all researches. Fernandes et al. (2019) used SVM and CNN to identify the hyperspectral image data of different grape vine varieties, and the test results showed that SVM achieved a recognition effect not inferior to CNN. Of course, we recognize that deep learning is a trend of image recognition. In order to further improve the recognition accuracy and consistency, it is necessary to expand the training set and employ the recognition method based on deep learning.

The main work of this research was to complete the identification of *M. micrantha* in a small field. The image samples used were hyperspectral images taken with a handheld spectrometer. In the future research, the hyperspectral images of other invasive plants will be collected to verify the generalization performance of the proposed method. In addition, the images acquired by the handheld spectrometer were mainly used to study the hyperspectral image data processing method, which can save time and cost on the basis of ensuring the reliability of the data. In practical applications, it is often necessary to identify invasive plants in a wide range, which requires the hyperspectral imager to be mounted on the UAV for image acquisition. Our study provides a reliable reference for hyperspectral image data processing of *M. micrantha*.

CONCLUSION

In this study, to determine the best methods for *M. micrantha* recognition based on hyperspectral technology, five preprocessing methods, one dimension reduction method, and three classifiers were separately combined to process the hyperspectral image data of *M. micrantha*. It was demonstrated that SG smoothing could eliminate the interference of high-frequency noise in raw spectral data and improved the spectral signal-to-noise ratio. Importantly, PCA reduced the dimensions of the feature set and kept most of the dataset information. Additionally, PCA improved the accuracy and calculation efficiency of the algorithm to some extent. In our study, the recognition accuracy and time after PCA dimension reduction were universally better than those without PCA processing. Finally, the dataset after dimension reduction was classified by classifiers, proving that RF had the most accurate and consistent result in our dataset, while SVM had the shortest execution time. In subsequent studies, SG-PCA-RF and SG-PCA-SVM algorithms, which performed well in this study, will be tested in the hyperspectral images of other invasive plants obtained by UAV.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YH and XQ: methodology, software, validation, writing – original draft, and writing – review and editing. JL: methodology, software, and validation. RY: methodology and software. FuW: software. YL and SZ: project administration. FaW: writing – review and editing. WQ: project administration, writing – original draft, and writing – review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

The work in this article was supported by the National Natural Science Foundation of China (31801804), projects subsidized by special funds for science technology innovation and industrial development of Shenzhen Dapeng New District (PT202001-06), and the Key Research and Development Program of Nanning (20192065).

ACKNOWLEDGMENTS

The authors thank the native English speaking experts from the editing team of American Journal Experts for polishing our article.

REFERENCES

Aneece, I., and Epstein, H. (2015). Distinguishing early successional plant communities using ground-level hyperspectral data. *Remote Sens.* 7, 16588–16606. doi: 10.3390/rs71215850

Aneece, I., and Epstein, H. (2017). Identifying invasive plant species using field spectroscopy in the VNIR region in successional systems of north-central Virginia. *Int. J. Remote Sens.* 38, 100–122. doi: 10.1080/01431161.2016.1259682

Asaari, M. S. M., Mishra, P., Mertens, S., Dhondt, S., Inze, D., Wuyts, N., et al. (2018). Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform. *ISPRS J. Photogramm. Remote Sens.* 138, 121–138. doi: 10.1016/j.isprsjprs.2018.02.003

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/a:1010933404324

Calvino-Cancela, M., Mendez-Rial, R., Reguera-Salgado, J., and Martin-Herrero, J. (2014). Alien plant monitoring with ultralight airborne imaging spectroscopy. *PLoS One* 9:9. doi: 10.1371/journal.pone.0102381

Chance, C. M., Coops, N. C., Crosby, K., and Aven, N. (2016). Spectral wavelength selection and detection of two invasive plant species in an urban area. *Can. J. Remote Sens.* 42, 27–40. doi: 10.1080/07038992.2016.1143330

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1023/A:1022627411411

Dash, J. P., Watt, M. S., Paul, T. S. H., Morgenroth, J., and Pearse, G. D. (2019). Early detection of invasive exotic trees using UAV and manned aircraft multispectral and LiDAR data. *Remote Sens.* 11:21. doi: 10.3390/rs11151812

Day, M. D., Clements, D. R., Gile, C., Senaratne, W. K. A. D., Shen, S., Weston, L. A., et al. (2016). Biology and impacts of pacific islands invasive species. 13. *Mikania micrantha* Kunth (Asteraceae). *Pacific Sci.* 70, 257–285. doi: 10.2984/70.3.1

Day, M. D., Kawi, A., Kurika, K., Dewhurst, C. F., Waisale, S., Saul-Maora, J., et al. (2012). *Mikania micrantha* Kunth (Asteraceae) Mile-a-Minute): its distribution and physical and socioeconomic impacts in Papua New Guinea. *Pacific Sci.* 66, 213–223. doi: 10.2984/66.2.8

Du, W., Xu, T. Y., Yu, F. H., and Chen, C. L. (2018). Measurement of nitrogen content in rice by inversion of hyperspectral reflectance data from an unmanned aerial vehicle. *Ciencia Rural* 48:10. doi: 10.1590/0103-8478cr20180008

Fernandes, A. M., Utkin, A. B., Eiras-Dias, J., Cunha, J., Silvestre, J., and Melo-Pinto, P. (2019). Grapevine variety identification using "Big Data" collected with miniaturized spectrometer combined with support vector machines and convolutional neural networks. *Comput. Electron. Agric.* 163:9. doi: 10.1016/j.compag.2019.104855

Fu, C. B., Gan, S., Yuan, X. P., Xiong, H. G., and Tian, A. H. (2018). Determination of soil salt content using a probability neural network model based on particle swarm optimization in areas affected and non-affected by human activities. *Remote Sens.* 10:19. doi: 10.3390/rs10091387

Golhani, K., Balasundram, S. K., Vadamalai, G., and Pradhan, B. (2019). Estimating chlorophyll content at leaf scale in viroid-inoculated oil palm seedlings (*Elaeis guineensis* Jacq.) using reflectance spectra (400 nm-1050 nm). *Int. J. Remote Sens.* 40, 7647–7662. doi: 10.1080/01431161.2019.1584930

Grosse-Stoltenberg, A., Hellmann, C., Thiele, J., Werner, C., and Oldeland, J. (2018). Early detection of GPP-related regime shifts after plant invasion by integrating imaging spectroscopy with airborne LiDAR. *Remote Sens. Environ.* 209, 780–792. doi: 10.1016/j.rse.2018.02.038

Jeyakumar, S., and Sudha, S. (2019). Hybrid hyperspectral image compression technique for non-iterative factorized tensor decomposition and principal component analysis: application for NASA's AVIRIS data. *Comput. Geosci.* 23, 969–979. doi: 10.1007/s10596-019-09855-2

Khadka, A. (2017). Assessment of the perceived effects and management challenges of *Mikania micrantha* invasion in Chitwan National Park buffer zone community forest, Nepal. *Heliyon* 3:14. doi: 10.1016/j.heliyon.2017.e00289

Lawrence, K. C., Windham, W. R., Park, B., Heitschmidt, G. W., Smith, D. P., and Feldner, P. (2006). Partial least squares regression of hyperspectral images for contaminant detection on poultry carcasses. *J. Near Infrared Spectrosc.* 14, 223–230. doi: 10.1255/jnirs.639

Liu, J. B., Han, J. C., Chen, X., Shi, L., and Zhang, L. (2019). Nondestructive detection of rape leaf chlorophyll level based on Vis-NIR spectroscopy. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 222:7. doi: 10.1016/j.saa.2019.117202

Luo, F. L., Du, B., Zhang, L. P., Zhang, L. F., and Tao, D. C. (2019). Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image. *IEEE Trans. Cybern.* 49, 2406–2419. doi: 10.1109/tcyb.2018.2810806

Macanawai, A. R., Day, M. D., Tumaneng-Diete, T., and Adkins, S. W. (2012). The impact of rainfall upon pollination and reproduction of *Mikania micrantha* in Viti Levu, Fiji. *Pak. J. Weed Sci. Res.* 18, 367–375.

Marcinkowska-Ochtyra, A., Jarocinska, A., Bzdega, K., and Tokarska-Guzik, B. (2018). Classification of expansive grassland species in different growth stages based on hyperspectral and LiDAR data. *Remote Sens.* 10:22. doi: 10.3390/rs10122019

Masemola, C., Cho, M. A., and Ramoelo, A. (2019). Assessing the effect of seasonality on leaf and canopy spectra for the discrimination of an alien tree species, *Acacia Mearnsii*, from co-occurring native species using parametric and nonparametric classifiers. *IEEE Trans. Geosci. Remote Sens.* 57, 5853–5867. doi: 10.1109/tgrs.2019.2902774

Nath, A., Sinha, A., Lahkar, B. P., and Brahma, N. (2019). In search of Aliens: factors influencing the distribution of *Chromolaena odorata* L. and *Mikania micrantha* Kunth in the Terai grasslands of Manas National Park. *India. Ecol. Eng.* 131, 16–26. doi: 10.1016/j.ecoleng.2019.02.012

Orrillo, I., Cruz-Tirado, J. P., Cardenas, A., Oruna, M., Carnero, A., Barbin, D. F., et al. (2019). Hyperspectral imaging as a powerful tool for identification of papaya seeds in black pepper. *Food Control* 101, 45–52. doi: 10.1016/j.foodcont.2019.02.036

Peerbhay, K. Y., Mutanga, O., and Ismail, R. (2015). Random forests unsupervised classification: random forests unsupervised classification: the detection and mapping of *Solanum mauritianum* infestations in plantation forestry using hyperspectral data. *IEEE J. Select. Topics Appl. Earth Observ Remote Sens.* 8, 3107–3122. doi: 10.1109/jstars.2015.2396577

Qi, H. J., Jin, X., Zhao, L., Dedo, I. M., and Li, S. W. (2017). Predicting sandy soil moisture content with hyperspectral imaging. *Int. J. Agric. Biol. Eng.* 10, 175–183. doi: 10.25165/j.ijabe.20171006.2614

Qiao, T., Yang, Z. J., Ren, J. C., Yuen, P., Zhao, H. M., Sun, G. Y., et al. (2018). Joint bilateral filtering and spectral similarity-based sparse representation: a generic framework for effective feature extraction and data classification in hyperspectral imaging. *Pattern Recognit.* 77, 316–328. doi: 10.1016/j.patcog.2017.10.008

Sabat-Tomala, A., Raczko, E., and Zagajewski, B. (2020). Comparison of support vector machine and random forest algorithms for invasive and expansive species classification using airborne hyperspectral data. *Remote Sens.* 12:21. doi: 10.3390/rs12030516

Saberioon, M., Cisar, P., Labbe, L., Soucek, P., and Pelissier, P. (2019). Spectral imaging application to discriminate different diets of live rainbow trout (*Oncorhynchus mykiss*). *Comput. Electron. Agric.* 165:6. doi: 10.1016/j.compag.2019.104949

Salimi, A., Ziaii, M., Amiri, A., Zadeh, M. H., Karimpouli, S., and Moradkhani, M. (2018). Using a feature subset selection method and support vector machine to address curse of dimensionality and redundancy in Hyperion hyperspectral data classification. *Egypt. J. Remote Sens. Space Sci.* 21, 27–36. doi: 10.1016/j.ejrs.2017.02.003

Shen, S., Shen, Z. R., and Zhao, M. (2017). Big data monitoring system design and implementation of invasive alien plants based on WSNs and WebGIS. *Wirel. Pers. Commun.* 97, 4251–4263. doi: 10.1007/s11277-017-4723-0

Shen, S., Xu, G., Zhang, F., Jin, G., Liu, S., Liu, M., et al. (2013). Harmful effects and chemical control study of *Mikania micrantha* H.B.K in Yunnan, Southwest China. *Afr. J. Agric. Res.* 8, 5554–5561.

Tarantino, C., Casella, F., Adamo, M., Lucas, R., Beierkuhnlein, C., and Blonda, P. (2019). *Ailanthus altissima* mapping from multi-temporal very high resolution satellite images. *ISPRS J. Photogramm. Remote Sens.* 147, 90–103. doi: 10.1016/j.isprsjprs.2018.11.013

Tesfamichael, S. G., Newete, S. W., Adam, E., and Dubula, B. (2018). Field spectroradiometer and simulated multispectral bands for discriminating invasive species from morphologically similar cohabitant plants. *Gisci. Remote Sens.* 55, 417–436. doi: 10.1080/15481603.2017.1396658

Tian, X., Fan, S. X., Huang, W. Q., Wang, Z. L., and Li, J. B. (2020). Detection of early decay on citrus using hyperspectral transmittance imaging technology coupled with principal component analysis and improved watershed segmentation algorithms. *Postharvest Biol. Technol.* 161:9. doi: 10.1016/j.postharvbio.2019.111071

Tusa, L., Kern, M., Khodadadzadeh, M., Blannin, R., Gloaguen, R., and Gutzmer, J. (2020). Evaluating the performance of hyperspectral short-wave infrared sensors for the pre-sorting of complex ores using machine learning methods. *Miner. Eng.* 146:10. doi: 10.1016/j.mineng.2019.106150

Vetrekar, N., Gad, R. S., Fernandes, I., Parab, J. S., Desai, A. R., Pawar, J. D., et al. (2015). Non-invasive hyperspectral imaging approach for fruit quality control application and classification: case study of apple, chikoo, guava fruits. *J. Food Sci. Technol. Mysore* 52, 6978–6989. doi: 10.1007/s13197-015-1838-8

Wu, N., Zhang, C., Bai, X. L., Du, X. Y., and He, Y. (2018). Discrimination of *Chrysanthemum* varieties using hyperspectral imaging combined with a deep convolutional neural network. *Molecules* 23:14. doi: 10.3390/molecules23112831

Xu, H. L., Zhang, H. Y., He, W., and Zhang, L. P. (2019). Superpixel-based spatial-spectral dimension reduction for hyperspectral imagery classification. *Neurocomputing* 360, 138–150. doi: 10.1016/j.neucom.2019.06.023

Yang, Q. H., Ye, W. H., Deng, X., Cao, H. L., Zhang, Y., and Xu, K. Y. (2005). Seed germination eco-physiology of *Mikania micrantha* HBK. *Bot. Bull. Acad. Sinica* 46, 293–299.

Yang, X. D., Li, G. L., Song, J., Gao, M. J., and Zhou, S. L. (2018). Rapid discrimination of Notoginseng powder adulteration of different grades using FT-MIR spectroscopy combined with chemometrics. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 205, 457–464. doi: 10.1016/j.saa.2018.07.056

Yao, Z. F., Lei, Y., and He, D. J. (2019). Early visual detection of wheat stripe rust using visible/near-infrared hyperspectral imaging. *Sensors* 19:16. doi: 10.3390/s19040952

Zhang, L. Y., Ye, W. H., Cao, H. L., and Feng, H. L. (2004). *Mikania micrantha* H B K in China–an overview. *Weed Res.* 44, 42–49. doi: 10.1111/j.1365-3180.2003.00371.x

Zhong, x, Huang, Z., Si, H., and Zan, Q. (2004). Analysis of ecological-economic loss caused by weed *Mikania micrantha* on Neilingding Island Shenzhen China. *J. Trop. Subtrop. Bot.* 12, 167–170.

Zhou, X., Sun, J., Tian, Y., Wu, X. H., Dai, C. X., and Li, B. (2019). Spectral classification of lettuce cadmium stress based on information fusion and VISSA-GOA-SVM algorithm. *J. Food Process Eng.* 42:9. doi: 10.1111/jfpe.13085

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Huang, Li, Yang, Wang, Li, Zhang, Wan, Qiao and Qian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fpls.2021.646173

Using Spectral Reflectance to Estimate the Leaf Chlorophyll Content of Maize Inoculated With Arbuscular Mycorrhizal Fungi Under Water Stress

Jinhua Sun[1], Liu Yang[1]*, Xitian Yang[1], Jie Wei[2], Lantao Li[3], Erhui Guo[1] and Yuhua Kong[1]

[1] College of Forestry, Henan Agricultural University, Zhengzhou, China, [2] Henan Ecological and Environmental Monitoring Center, Zhengzhou, China, [3] College of Resources and Environment, Henan Agricultural University, Zhengzhou, China

Leaf chlorophyll content is an important indicator of the growth and photosynthesis of maize under water stress. The promotion of maize physiological growth by (AMF) has been studied. However, studies of the effects of AMF on the leaf chlorophyll content of maize under water stress as observed through spectral information are rare. In this study, a pot experiment was carried out to spectrally estimate the leaf chlorophyll content of maize subjected to different durations (20, 35, and 55 days); degrees of water stress (75%, 55% and 35% water supply) and two inoculation treatments (inoculation with *Funneliformis mosseae* and no inoculation). Three machine learning algorithms, including the back propagation (BP) method, least square support vector machine (LSSVM) and random forest (RF) method, were used to estimate the leaf chlorophyll content of maize. The results showed that AMF increased the leaf chlorophyll content, net photosynthetic rate (A), stomatal conductance (gs), transpiration rate (E), and water use efficiency (WUE) of maize but decreased the intercellular carbon dioxide concentration (Ci) of maize and atmospheric vapor pressure deficit (VPD) regardless of the water stress duration and degree. The first-order differential spectral data can better reflect the correlation between leaf chlorophyll content and spectrum of inoculated maize when compared with original spectral data. The BP model performed bestin modeling the maize leaf chlorophyll content, yielding the largest R^2-values and smallest root mean square error (RMSE) values, regardless of stress duration. These results provide a reliable basis for the effective monitoring of the leaf chlorophyll content of maize under water stress.

Keywords: arbuscular mycorrhizal fungi, leaf chlorophyll content, spectral reflectance, machine learning algorithms, water stress

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Ana María Mendez-Espinoza,
Instituto de Investigaciones Agropecuarias (Chile), Chile
Xiancan Zhu,
Anhui Normal University, China

****Correspondence:***
Liu Yang
yangliutj@163.com

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *14 January 2021*
Accepted: *10 May 2021*
Published: *28 May 2021*

Citation:
Sun J, Yang L, Yang X, Wei J, Li L, Guo E and Kong Y (2021) Using Spectral Reflectance to Estimate the Leaf Chlorophyll Content of Maize Inoculated With Arbuscular Mycorrhizal Fungi Under Water Stress. Front. Plant Sci. 12:646173. doi: 10.3389/fpls.2021.646173

INTRODUCTION

Maize is one of the most important food crops in Asia. However, water stress is a main constraint of crop production at the global scale and is expected to increase in coming years (Lesk et al., 2016), and global crop yields are heavily affected by this constraint (Daryanto et al., 2016). Therefore, increasing the utilization efficiency of existing water resources is an effective way to mitigate agricultural water use limitations (Schewe et al., 2014).

Water stress affects the photosynthesis, transpiration, and water use efficiency (WUE) of plants and their absorption and utilization of water and nutrients and hinders the physiological and biochemical processes of plants (Kahil et al., 2015; Ortega-Farias et al., 2021). Plants can cope with water stress through their adaptive strategies (Pavithra and Yapa, 2018), and the symbiosis between arbuscular mycorrhizal fungi (AMF) and plants plays an important role in plants adaptations to water stress (Pavithra and Yapa, 2018). AMF canengage in mutualism with more than 80% of terrestrial plants in the world (Smith and Read, 2008). Plants provide carbohydrates to AMF, and AMF play a crucial role in the growth and nutrient uptake of host plants through beneficial physiological processes. Moreover, AMF improve plant performance under abiotic stresses such as drought (Pavithra and Yapa, 2018), pollution (Liu et al., 2018), or salinity (Arafat and He, 2011). AMF have been shown to improve plant absorption of water, increase photosynthetic pigments, stomatal conductance (Duc et al., 2018; Meddich et al., 2021), intercellular CO_2 concentration, transpiration rate, root volume and diameter, and stimulate H^+-ATPase activity and gene expression of plants in response to water stress (Cheng et al., 2021). *Bromus* species inoculated with AMF enhanced superoxide dismutase, peroxidase, catalase, and ascorbate peroxidase activities (Karimkhani et al., 2021) to help plants to tolerate water stress. In addition, inoculation with *F. mosseae* through the improvement of ionic and biochemical status of the plant can mitigate the detrimental effects of water-deficit stress on maize plants (Bahraminia et al., 2020). AMF increase water use and reduce oxidation damage by stimulating antioxidant activities (Pedranzani et al., 2016) and regulating water absorption and transport by aquaporin genes and endogenous hormones (Pedranzani et al., 2016; Quiroga et al., 2017; Cheng et al., 2021). Plant physiological phenomena are reflected in the healthy status of leaves. Leaf chlorophyll content is an important index for measuring plant photosynthesis and growth. However, the extraction and detection of leaf pigments by traditional chemical monitoring methods is tedious, destructive, discontinuous and time consuming. A spectro radio meter can be used to obtain plant spectral information through reflective rays at different wavelengths (700-1300 nm) with relatively high reflectivity observed in the near infrared region. The spectral reflectance of leaves in different wavebands represents different leaf characteristics of plants. Spectral reflectance in the visible and near infrared regions differs between plants under water stress and healthy plants. Spectro radio metric methods yield measurements more quickly, continuously and economically than traditional laboratory methods (Basayigit and Ozkul, 2015).

The vegetation indexes have been used to build inversion model of chlorophyll content of wheat, as the red edge parameters used to estimate chlorophyll content of plant under the stripe rust, water or salt stress have been investigated in many studies (Gu et al., 2008; Jiang et al., 2010; El-Hendawy et al., 2021). In the past, most models employed to estimate the leaf chlorophyll content have been based on linear regression method; however, this method cannot describe the complex relationship between modeling factors and dependent variables. Therefore, the modeling accuracy was not high. The use of machine learning algorithms is an effective way to express the complex relationship between the modeling factors and dependent variables (Saitta et al., 2011; Sonobe et al., 2021). The back propagation (BP) approach has been used to determine plant diseases (Saleem et al., 2019), and this methodology has potential in the analysis of hyperspectral reflectance data. The random forest (RF) method is a regression technique that combines numerous decision trees to classify or predict the value of a variable, and it has been used for estimating vegetation properties (Li et al., 2017; Chen et al., 2018) as well as for classification and regression (Biau and Scornet, 2015). Least square support vector machine (LSSVM) is a nonlinear system modeling method that has been proposed in recent years. LSSVM requires a small number of training samples and can approach nonlinear systems with high accuracy, and it has been widely used in many fields (Li et al., 2010; Mall and Suykens, 2015). However, studies involving the hyperspectral estimation of chlorophyll content in maize inoculated with AMF under drought stress are rare. In our research, pot experiments were conducted by controlling the amount of water, and the chlorophyll content of maize inoculated with AMF was monitored spectrally. Comparisons and analyses of the physiological characteristics and spectral response of AMF-inoculated and control maize under water stress were performed in this study. An inversion model of the chlorophyll content and multiple spectral variables was established. The objectives of this study were to (1) compare the chlorophyll content and other physiological characteristics of maize with or without inoculation under water stress, (2) determine the spectral response of AMF-inoculated and non-inoculated maize to water stress, and construct the best spectral estimation model of the chlorophyll content of maize under water stress.

METHODS

Materials

Maize seeds were obtained from the Henan Academy of Agricultural Sciences; and *Funneliformis mosseae* was obtained from the Research Institute of Plant Nutrition and Resources in the Beijing Academy of Agriculture and Forestry Sciences and cultured in a microbial laboratory at Henan Agricultural University. River sand with poor nutrients was collected as a substrate and then air dried, passed through a 1 mm sieve and mixed thoroughly. The basic properties of the soil are listed in **Table 1**. The soil was sterilized at 121°C for 2 h, air dried and prepared for the pot experiment.

TABLE 1 | Basic physical and chemical properties of sandy soil.

pH	AP (mg/kg)	AN (mg/kg)	TN (%)	TC (%)
7.78	1.66	28	0.107	0.891

AP, available phosphorus content; AN, available nitrogen content; TN, total nitrogen content; TC, total carbon content.

Experimental Design

The experiment was conducted in a greenhouse at the Forestry College at Henan Agricultural University on August 15, 2019. The inoculation treatments included inoculation with *F. mosseae* and no inoculation; and the water stress treatments included (a) no stress, irrigation to 75% of the maximum water holding capacity of the soil (WS75%; achieved with 150 mL water per plant); (b) moderate stress, wherein irrigation was provided to 55% of the maximum water holding capacity of the soil (WS55%; achieved with 110 mL water per plant), and (c) severe stress, wherein irrigation was provided to 35% of the maximum water holding capacity of the soil (WS35%; achieved with 70 mL water per plant). The maize plants were subjected to the water stress treatments for 20, 35, or 55 days after sprouting (on August 20, 2019), in other words, the water stress starts on September 10, September 25, and October 15, 2019, respectively. With 20, 35, and 55 days after sprouting representing the seedling stage, three leaf stage and jointing stage, respectively, of maize growth. Each treatment was performed with 5 replicates. Therefore, a total of 90 pots of maize were established (2 (inoculation treatments) × 3 (water stress degree treatments) × 3 (water stress duration treatments) × 5(replicates)).

Maize was inoculated with *F. mosseae*at 50 g/pot (20 spores/g), and the non-inoculated maize was inoculated with the same amount of sterilized microbial inoculum. NH_4NO_3, KH_2PO_4, and K_2SO_4 were added to achieve 100 mg N/kg, 30 mg P/kg, and 150 K mg/kg, respectively, for fertilization. Before water stress, the plants were irrigated with 75% of the maximum water holding capacity of soil during the growth process, and the soil moisture meter (LB9007, QingDao, in China) was used for real-time monitoring. Maize seedling was irrigated in time every day to ensure that the soil moisture was maintained between70% and 75%. Maize seeds were disinfected with 10% H_2O_2 for 10 min, rinsed with water 2-3 times and then washed with deionized water4-5 times. Seeds were cultured in an incubator at 25-28°C for 24 h. Two seeds were sown in each pot (caliber × height × bottom diameter = 12 × 13 × 9.5 cm) before germination, and the best growing seedling was kept. The light intensity was controlled by 10 high-pressure sodium lamps, and the air temperature and humidity were controlled at 25°C and 20%, respectively, by an air-conditioning unit.

Measurements

Leaf spectral collection and physiological measurement were carried out after each of the three water stress durations, and five pots were randomly selected from each water stress treatment.

Spectrometric Determination

Leaf spectra were acquired with a Field Spec 3 ground spectrometer (ASD Company, United States) under dark conditions. The wave band range of the ground spectrometer was 350-1025 nm, and the spectral resolution was 3 nm. To accurately reflect the whole plant growth status, each pot was placed on a black workbench, and two leaves collected from each of the upper, middle and lower layers of the plant were measured. The halogen lamp used for the indoor test matched the spectrometer and was placed above the sample at an angle of 15° (the angle between the halogen lamp and ground normal) and 20 cm from the target. The field angle of the probe was 25°when observed from10 cm above the target (with the imaged object formed an angle with two edges of the maximum probe range). Each leaf was measured to obtain spectral information, and a whiteboard correction was conducted every 30 min before and during the measurements. Each leaf was lit from four directions during measurement, and five spectra were collected for each lighting direction to avoid errors of leaf curling. Reflectance data were obtained from the average spectral values from all directions.

Physiological and Biochemical Determination

The chlorophyll contents of the leaves were determined using a SPAD-502 instrument (Konica Minolta, Japan). This device determines the relative chlorophyll content using dual wavelength optical absorbance measurements (at wavelengths of 620 and 940 nm) of leaf samples. The chlorophyll content was obtained from 5 selected points on each leaf during measurement, and the average value was taken to reflect the chlorophyll content of the whole leaf.

Photosynthetic parameters such as intercellular CO_2 concentration (Ci), net photosynthetic rate (A), stomatal conductance (gs), transpiration rate (E), and water use efficiency (WUE) of maize leaves and water vapor pressure (VPD) were measured by a LI-6400 photosynthesis instrument (LI-COR, United States).

The drying method was adopted to measure the aboveground and belowground biomass of maize: the roots and shoots of maize were heated at 105°C for 30 min, dried at 75°C for 3 days, and weighed to determine the dry weight of maize. Mycorrhizal colonization was determined by the Phillips and Hayman methods (Phillips and Hayman, 1970).

Determination of Soil Properties

The soil property analysis shown in **Table 1** was performed using standard soil test procedures. The soil pH values were measured at a 1:2.5 soil/water ratio (w/v) via a pHmeter (Gao et al., 2015). The soil total carbon content (TC) and total nitrogen content (TN) were measured by Dumas combustion using an ElementarVario MAX CN analyzer with the combustion chamber set at 900°C and an oxygen flow rate of 125 mL min^{-1}. The soil available phosphorus content (AP) was measured via the molybdenum antimony colorimetric method, and the soil available nitrogen content (AN) was determined by the alkali hydrolysis diffusion method (Bao, 1999).

Data Analysis

Data Smoothing

During spectral data acquisition, the spectral curve may contain noise because the energy response varies according to the waveband and environment. Therefore, the spectral data must be smoothed. In this study, the Savitzky-Golay (SG) method was used to smooth the spectral curve. This method not only removes the high-frequency components but also retains the characteristic trends of the original curve, and the denoising effect of the SG method is better than that of other methods. The SG algorithm is a least-squares convolution smoothing method that is synthetically applied according to polynomial fitting order and smoothing degree; however, the error of the polynomial fitting curve is calculated by the first derivative. In this study, the SG smoothing method was realized by MATLAB 2014.

First-Order Differential Processing Method

Derivative spectroscopy can reduce the influences of light, atmospheric scattering, absorption and background on the spectral characteristics of targets and there by obviously enhance the correlations between the first-order differential processing of the original spectral reflectance and biochemical indexes. In this study, the smoothed original spectral curves were processed with the first-order differential method.

Modeling Methods

The LSSVM, BP, and RF models were used to estimate the chlorophyll content in the leaves through reflectance spectra, and they were implemented in MATLAB 2014.

The LSSVM method is widely used for complex nonlinear modeling. If the training sample set is (x_i,y_i), with i = 1, 2, . . ., n, then $x \in R^d$ and $y \in R$. The main idea of the support vector machine (SVM) is as follows: first, a nonlinear mapping φ (·) is used to map the sample input space R^d to the feature space: $\varphi(x) = (\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n))$; next, the optimal decision function $y = w^T \cdot \varphi(x) + b$ is constructed in this high-dimensional feature space, and then the model parameters w and b are determined based on the principle of structural risk minimization. The calculation formula of structural risk is as follows:

$$R = c \cdot R_{emp} + 1/2\|w\|^2$$

Where c is the normalization parameter and R_{emp} is the loss function, which is also called the empirical risk. The common loss functions include the first loss function, the quadratic loss function and the Hubber loss function. Different loss functions represent different SVM models; LSSVM is the SVM with the quadratic loss function. That is, $R_{emp} = \sum_i \xi_i^2$, where ξ_i^2 is the prediction error of the model to the training samples.

The training of sample data is performed via the BP method, and the weights and thresholds of the network are constantly modified, so that the error function decreases along the negative gradient direction and approaches the expected output. BP is a widely used neural network model that is primarily applied for function approximation, model recognition and classification, data compression and time series prediction. The BP network is composed of an input layer, a hidden layer and an output layer. The hidden layer can have one or more layers. The network adopts the S-type transfer function: $f(x) = 1/(1+e^x)$. The error function $E = \sum_i (T_i+O_i)^2/2$ (where T_i is the expected output and O_i is the calculated output of the network) can be minimized by adjusting the weights and thresholds of the network.

The RF method builds multiple decision trees and fuses them to obtain a more accurate and stable model, which is represented by the combination of the bagging idea and random selection features. The RF method constructs multiple decision trees. When a certain sample needs to be predicted, the prediction results of each tree in the forest for the sample are counted, and the final result is selected from these prediction results by the voting method. Randomness is reflected in two aspects: the selection features and selection of samples. Thus, each tree in the forest has both similarities and differences from other trees.

Statistical Methods

Analysis of variance (ANOVA) was used to compare the effects of AMF among treatments on the mycorrhizal colonization, the growth and photosynthetic physiological progress of maize using the IBM SPSS 23.0 software program (SPSS Inc., Chicago, IL, United States). The least significant difference (LSD) test at the 0.05 probability level (p) was used to compare the means of the measured traits. Correlations between spectral and leaf chlorophyll content and two tailed test (significant level α was 0.05) were determined by the IBM SPSS 23.0 (SPSS Inc, Chicago, IL, United States). The figures were plotted using Origin 8.0 (OriginLab Corporation, Northampton, MA, United States).

RESULTS

Mycorrhizal Colonization

Mycorrhizal colonization represents an intimate association between fungi and host plants. With increasing stress duration and degree, mycorrhizal colonization decreased gradually (**Figure 1**). Mycorrhizal colonization of inoculated maize was

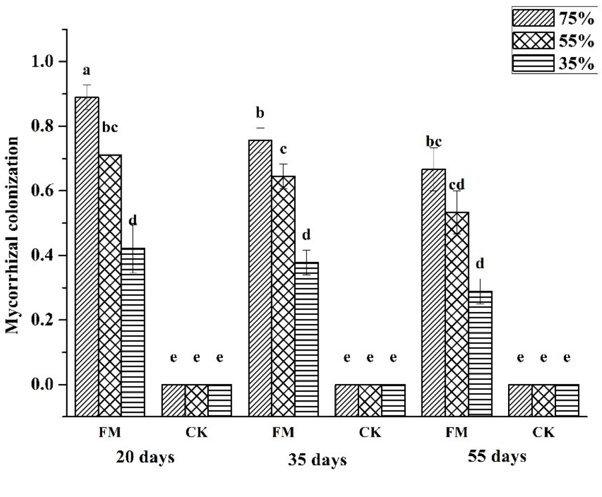


FIGURE 1 | Mycorrhizal colonization of maize under different treatments. FM, inoculated with *F. mosseae*; CK, without inoculation.

highest under the75% water supply regardless of stress duration and lowest under the35% water supply. Under the stress duration of 20 days, the mycorrhizal colonization of inoculated maize under a 75% and 55% water supply was increased significantly ($p < 0.05$) by 111.9% and 69% respectively, compared with that of inoculated maize under the 35% water supply. Under the stress duration of 35 days, the mycorrhizal colonization of inoculated maize under a 75% and 55% water supply was significantly increased ($p < 0.05$), by 100% and 29.2%, respectively, compared with that of maize under the 35% water supply; however, the mycorrhizal colonization of non-inoculated maize was zero, it may be because the maize grew on the sterile soil in an aseptic environment.

Maize Growth

With increasing water stress, the chlorophyll content of the leaves in both the inoculated and non-inoculated maize decreased gradually; however, when the water stress duration was 20 days, the chlorophyll content of AMF-inoculated maize was4.33, 3.28 and 4.66 more than that of non-inoculated maize under a 75%, 55%, and 35% water supply, respectively. When the water stress duration was 35 or 55 days, the differences in chlorophyll content between AMF-inoculated and non-inoculated maize were not significant when a normal water supply was used; however, when the water stress duration was 35 days, the chlorophyll content of inoculated maize was significantly higher than that of non-inoculated maize when the water supply was 55% or 35%. Under the stress duration of 20 days, the aboveground biomass of AMF-inoculated maize was increased significantly, by 25.7%, compared with that of non-inoculated maize when the water supply was 55%. Moreover, the aboveground biomass of inoculated maize was greater than that of maize without inoculation under the other stress durations, although the differences were not significant (**Table 2**). The independent effects of stress duration, water stress degree, and inoculation were significant; however, the interaction effect of any two or three factors was not significant (**Table 3**).

Photosynthetic Physiological Parameters of Maize

As shown in **Table 2**, water stress inhibited the photosynthesis and transpiration of maize, where as AMF inoculation alleviated the physiological damage of water stress to maize. Regardless of the water stress duration or watering conditions, Ci in the leaves of maize was significantly higher in non-inoculated maize than in AMF-inoculated maize ($p < 0.05$). As the degree of water stress increased, the Ci of the leaves of both inoculated and non-inoculated maize increased gradually. Under the stress duration of 20 days, the A of the inoculated maize was higher than that of the non-inoculated maize ($p > 0.05$). However, under a stress duration of 35 days, and a 75% and 55% water supply, the A of the inoculated maize was 1.6 and 1.28 times higher, respectively, than that of non-inoculated maize ($p < 0.05$). Under a stress duration of 55 days and a 55% water supply, the A of the inoculated maize was increased significantly, by 30%, compared with that of the non-inoculated maize ($p < 0.05$).

When the stress duration was 20 or 35 days, AMF significantly improved the gs of maize leaves regardless of the degree of water stress. The stomatal conductance of AMF-inoculated maize was significantly higher than that of non-inoculated maize

TABLE 2 | Physiological growth characteristics of maize under different treatments.

Stress duration	Water stress	Inoculation	SPAD	Ci	A	gs	E	VPD	WUE	Aboveground biomass
20 days	75%	FM	29.02 ± 0.49a	197.5 ± 2.2l	7.02 ± 0.21a	28.6 ± 0.95a	1.27 ± 0.08a	2.53 ± 0.07g	8.22 ± 1.41a	0.86 ± 0.14cd
		CK	24.69 ± 0.47d	258.5 ± 8.8j	6.85 ± 1.61ab	22.4 ± 1.1c	1.08 ± 0.33ab	2.78 ± 0.03f	6.32 ± 0.9ab	0.57 ± 0.18de
	55%	FM	27.76 ± 0.54b	210.2 ± 6.8k	6.35 ± 0.3b	25.6 ± 1.15b	1.13 ± 0.13ab	3.44 ± 0.45f	7.28 ± 0.85a	0.7 ± 0.08d
		CK	24.48 ± 2.97cd	276.7 ± 1.6i	6.42 ± 0.81ab	20.2 ± 1.53cd	0.88 ± 0.15 bc	4.38 ± 0.10e	5.8 ± 0.35b	0.52 ± 0.01e
	35%	FM	26.13 ± 0.49c	277.5 ± 2.5i	5.43 ± 1.43bc	21.2 ± 0.52c	0.72 ± 0.08cd	4.3 ± 0.22e	6.42 ± 1.28ab	0.52 ± 0.14de
		CK	21.47 ± 0.75d	304.7 ± 10.5h	3.9 ± 0.96cd	19.5 ± 0.75d	0.6 ± 0.05d	5.17 ± 0.19d	4.45 ± 0.6cd	0.35 ± 0.09e
35 days	75%	FM	24.92 ± 2.29cd	216.2 ± 4.2k	5.57 ± 0.35b	25.2 ± 0.74b	0.94 ± 0.03b	3.12 ± 0.06f	5.85 ± 0.48ab	1.57 ± 0.09b
		CK	21.52 ± 3.98de	265.2 ± 12.3ij	3.97 ± 0.20c	20.2 ± 1.53cd	0.72 ± 0.13d	4.12 ± 0.03ef	4.7 ± 0.2c	1.19 ± 0.25bc
	55%	FM	23.85 ± 1.05d	375.7 ± 21g	4.43 ± 0.15bc	21.3 ± 4.65c	0.81 ± 0.02b	4.43 ± 0.3e	4.28 ± 0.31cd	1.66 ± 0.3ab
		CK	20.34 ± 3.18e	387.5 ± 14.9g	3.15 ± 0.38d	17.9 ± 2.12d	0.72 ± 0.06cd	4.87 ± 0.04de	3.5 ± 0.41d	1.16 ± 0.78bc
	35%	FM	22.05 ± 0.92d	402 ± 19.5g	3.03 ± 0.73de	19.7 ± 1.61c	0.52 ± 0.03de	5.13 ± 0.03d	3.78 ± 0.51d	0.96 ± 0.05c
		CK	16.37 ± 1.7e	427.5 ± 14.1f	2.38 ± 0.32de	16.3 ± 1.26e	0.45 ± 0.09ef	6.35 ± 0.3c	2.57 ± 0.56e	0.93 ± 0.41cd
55 days	75%	FM	24.75 ± 0.44d	453.8 ± 1.5e	3.05 ± 0.15d	22.2 ± 0.58c	0.74 ± 0.05c	5.15 ± 0.05d	4.4 ± 0.41cd	2.37 ± 0.42a
		CK	21.76 ± 1.11d	483.2 ± 6.3b	2.52 ± 0.31de	18.3 ± 1.04de	0.61 ± 0.05d	6.15 ± 0.13c	3.37 ± 0.43d	1.58 ± 0.33ab
	55%	FM	22.65 ± 0.36d	459 ± 1.3d	2.25 ± 0.15e	19.3 ± 0.76d	0.6 ± 0.05d	6.27 ± 0.18c	3.62 ± 0.08d	1.71 ± 0.51ab
		CK	20.21 ± 3.58de	470.3 ± 1.8c	1.73 ± 0.25f	17.3 ± 1.53de	0.48 ± 0.06de	7.15 ± 0.05b	2.6 ± 0.20e	1.19 ± 0.33bc
	35%	FM	20.86 ± 2.34de	480.3 ± 7.4bc	1.48 ± 0.25fg	16 ± 1.32e	0.47 ± 0.03e	7.12 ± 0.03b	2.42 ± 0.46ef	1.5 ± 0.49bc
		CK	18.41 ± 1.25e	496.5 ± 2.5a	1 ± 0.26g	12.5 ± 2.00f	0.36 ± 0.05f	8.08 ± 0.08a	1.58 ± 0.24f	0.85 ± 0.7cd

The value in the table are the mean value ± standard deviation of three repeated samples. Different lowercase letters represent significant differences at the level of 0.05. FM, inoculated with F. mosseae; CK, without inoculation; SPAD, leaf color value; Ci, intercellular carbon dioxide concentration; A, net photosynthetic rate; gs, stomatal conductance; E, transpiration rate; VPD, atmospheric vapor pressure deficit; WUE, water use efficiency.

TABLE 3 | *F*-test and *p*-Values for different treatments.

	SPAD	Ci	A	gs	E	VPD	WUE	Aboveground biomass
Stress time	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**
F/n	26.99/2	223.46/2	169.50/2	45.73/2	65.00/2	128.9/2	132.77/2	31.78/2
Water stress degree	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**
F/n	15.64/2	340.13/2	41.13/2	45.61/2	56.31/2	619.59/2	41.11/2	8.59/2
Inoculation	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**	0.00**
F/n	47.43/1	150.43/1	17.62/1	71.69/1	23.98/1	311.66/1	52.48/1	15.40/1
Stress duration × Water stress	0.93	0.00**	0.42	0.51	0.04*	0.26	0.74	0.28
F/n	0.22/4	128.70/4	1.00/4	0.84/4	2.90/4	1.38/4	0.50/4	1.32/4
Stress duration × Inoculation	0.41	0.00**	0.24	0.49	0.63	0.08	0.12	0.18
F/n	0.92/2	12.81/2	1.50/2	0.74/2	0.47/2	2.74/2	2.23/2	1.80/2
Water stress × Inoculation	0.66	0.00**	0.77	0.16	0.51	0.04	0.79	0.70
F/n	0.42/2	6.67/2	0.26/2	1.96/2	0.69/2	3.41/2	0.24/2	0.36/2
Stress duration × Water stress × Inoculation	0.91	0.02*	0.19	0.39	0.80	0.00**	0.97	0.89
F/n	0.24/4	3.48/4	1.60/4	1.06/4	0.41/4	5.86/4	0.13/4	0.29/4

*SPAD, leaf color value; Ci, intercellular carbon dioxide concentration; A, net photosynthetic rate; gs, stomatal conductance; E, transpiration rate; VPD, atmospheric vapor pressure deficit;WUE, water use efficiency; F, F-test value; n, degrees of freedom. Superscript** indicates that the differences among treatments were significant at the level of 0.01; and superscript * indicates that the differences among treatments were significant at the level of 0.05.*

regardless of water stress duration or degree ($p < 0.05$). With increasing stress duration, the transpiration effect of AMF-inoculated maize became greater than that of non-inoculated maize. When the stress duration was 55 days, under a 75% and 55% water supply, the E of AMF-inoculated maize was significantly higher than that of maize without inoculation, by 0.22 and 0.09 times, respectively ($p < 0.05$). Under the stress durations of 20 days, the atmospheric VPD was significantly lower ($p < 0.05$)in the inoculated maize than in the non-inoculated maize. Under the stress durations of 20 days, the WUE of leaves did not significantly differ between the inoculated and non-inoculated maize with a normal water supply, however, under a 75% and 55% water supply, the leaf WUE of the inoculated maize was significantly higher than that of the non-inoculated maize ($p < 0.05$).

Each type of treatment (stress duration, water stress degree, and inoculation) had a significant effect ($p < 0.05$)on the photosynthetic parameters of maize leaves. However, the interaction effect of any two or three treatment parameters was not significant for any parameter except for Ci, VPD, and E (**Table 3**).

Spectral Characteristics of Maize Leaves

At the three water stress durations, the general trend of the spectral curve of maize with different treatments was similar. When the stress duration was 20 or 35 days, at the green peak (550 nm), and red edge (700 nm), the leaf reflectance of non-inoculated maize was significantly higher (except under the 35% water supply) than that of inoculated maize (**Figure 2**). In non-inoculated maize under a 75% water supply, leaf reflectance was highest at wavelengths of 550 nm and 700 nm at a stress duration of 20 days whereas in non-inoculated maize with a 35% water supply, leaf reflectance was highest at wavelengths of 550 nm and 700 nm at a stress duration of 35 days. When the stress duration was 55 days, the leaf reflectance of inoculated maize with a 35% water supply was highest at a wavelength of 550 nm; and that of non-inoculated maize with a 75% water supply was highest at a wavelength of 700 nm. The highest reflectance of maize leaves under a stress duration of 20 days was higher than that of leaves under 35 and 55 days of stress at the wavelength of 550 and 700 nm (**Figure 2**).

Correlation Analysis of Chlorophyll Content and First-Order Spectral Differential Variables in Maize

For the first-order derivative of spectral reflection, the spectral reflectance of inoculated maize leaves had larger correlation coefficients with the leaf chlorophyll content than that without inoculation under three water stress durations (**Figure 3**). When the stress duration was 20 or 55 days, the correlation coefficients between the first-order differential reflectance and the chlorophyll content of inoculated maize showed unstable performance at the range of 800-1000 nm especially, but during 400-800 nm, the maximum correlation coefficients values are in the inoculated maize at 550 nm and 700 nm, respectively ($\alpha < 0.05$). When the stress duration was 35 days, the correlation coefficients between the first-order differential reflectance and the chlorophyll content of inoculated maize showed larger than that without inoculation at two peak areas (the wavelengths of 550-700 nm and 850-950 nm; **Figure 3**) ($\alpha < 0.05$).

For the original spectral data, there was a negative correlation between leaf chlorophyll content and spectral reflectance of inoculated maize under three water stress durations, but a positive correlation between that of non-inoculated maize before 500 nm under the stress duration of 35 days and before 700 nm under the stress duration of 55 days. There are two peaks of correlation coefficient between leaf color value and spectral reflectance of inoculated maize at 550 and 700 nm were larger than that of non-inoculated maize under the stress duration of 35 or 55 days (**Figure 3**) ($\alpha < 0.05$).

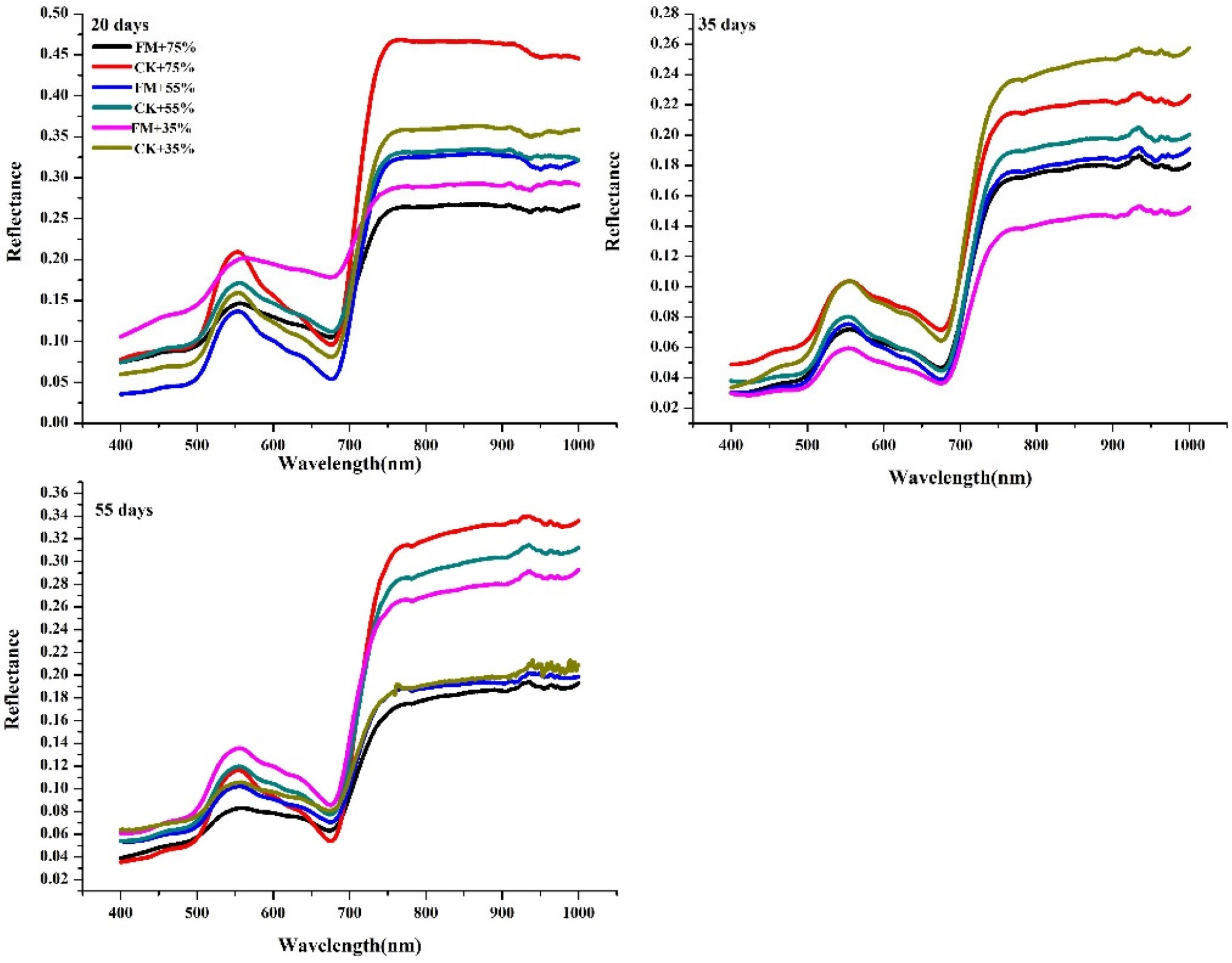


FIGURE 2 | Spectra of maize leaves at each stress level and different inoculation durations. FM, inoculated with *F. mosseae*; CK, without inoculation.

Compared with the original spectral data, the first-order differential processing can better reflect the correlation between leaf color value and spectrum of inoculated maize.

Hyperspectral Estimation Model of the Chlorophyll Content of Maize and Its Validation

To better reflect the chlorophyll content of inoculated maize under drought stress, the results of correlation analysis of the first-order differential spectral data were used to establish, an estimation model of the first-order differential values from sensitive bands corresponding to chlorophyll content. The LSSVM, RF, and BP nonlinear models based on machine learning were used. A regression analysis was carried out between the predicted and measured values, and the estimation accuracy of the models was evaluated by the coefficient of determination (R^2) values, root mean square error (RMSE)values and *p*-Value of the LSSVM, RF and BP models established with different spectral variables that were significant at the 0.001 level (**Figures 4A–C**). The results indicated that all of these models could be used to estimate the chlorophyll content of inoculated maize under drought stress. The BP model was better than the others achieving the largest R^2-values ($R^2 = 0.9796$ for the 20 days duration, $R^2 = 0.9951$ for the 35 days duration, and $R^2 = 0.9479$ for the 55 days duration) and smallest RMSEs (RMSE = 0.3875 for the20-day duration, RMSE = 0.2473 for the35-day duration, and RMSE = 0.6431 for the 55-day duration) regardless of the stress duration.

DISCUSSION

Water stress is a worldwide problem that affects the growth, yield and physiological processes of crops (Begum et al., 2019). AMF symbiosis has been reported to enhance plant tolerance to water stress (Duc et al., 2018; Zhang et al., 2018). Inoculation with *F. mosseae*has been shown to enhance plant dry weights compared to those of non-mycorrhizal plants under control and water-deficit stress conditions (Bahraminia et al., 2020). In this study, AMF colonization was negatively affected by water stress, possibly because of declines in spore germination and growth,

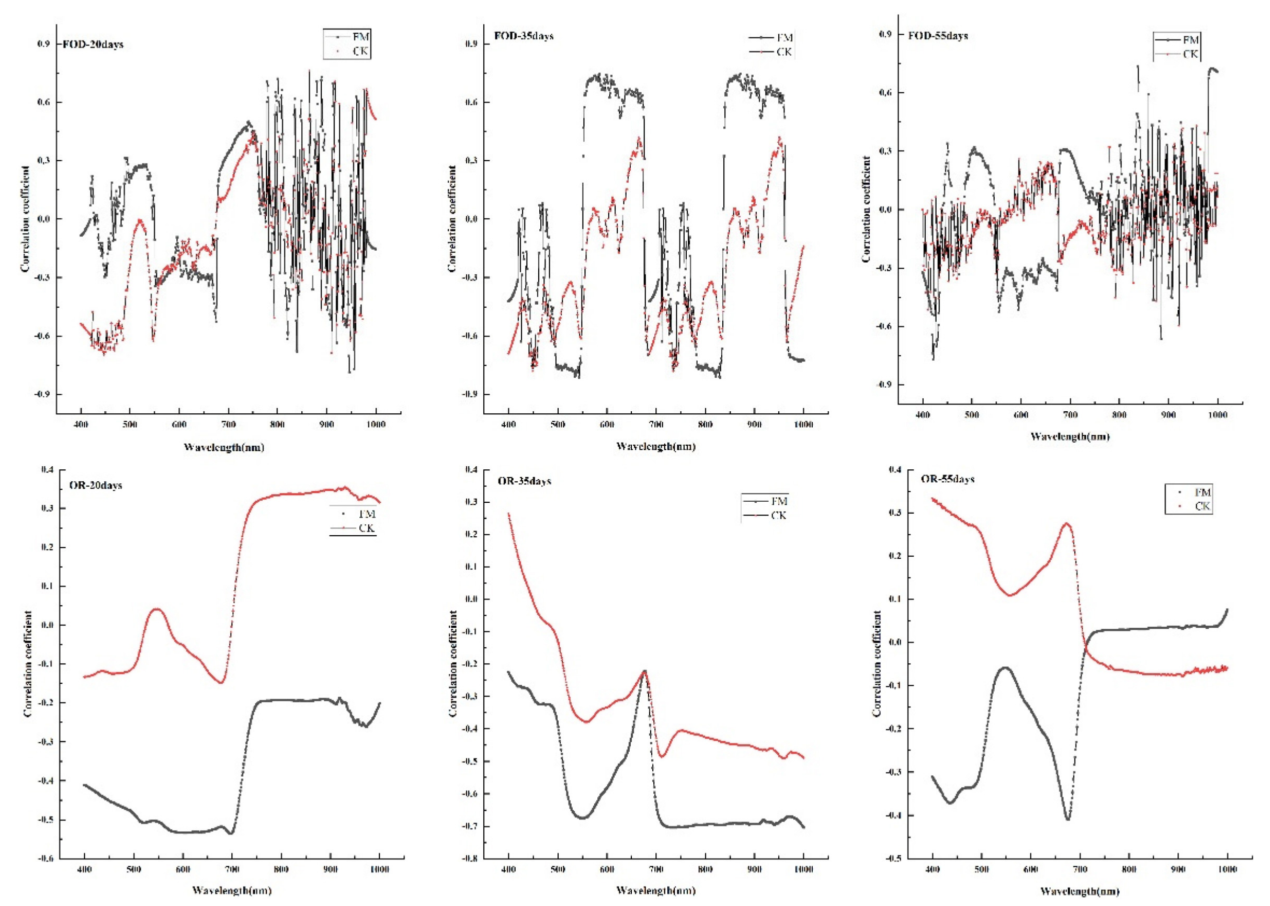


FIGURE 3 | Correlation between spectral reflectance and leaf color value under different treatments. FM, inoculated with *F. mosseae*; CK, without inoculation. OR-20 days:water stress duration was 20 days with the original spectral data; OR-35 days: water stress duration was 35 days with the original spectral data; OR-55 days: water stress duration was 55 days with the original spectral data; FOD-20 days:water stress duration was 20 days with the first-order derivative of spectral data; FOD-35 days: water stress duration was 35 days with the first-order derivative of spectral data; FOD-55 days: water stress duration was 55 days with the first-order derivative of spectral data.

reductions in the number of AMF structures, the inhibition of hyphal growth and expansion in soil (Salloum et al., 2018), or reductions in the supply of carbohydrate by host plants (Tyagi et al., 2017). However, AMF inoculation improved the growth of maize regardless of the stress degree and duration, possibly because the AMF facilitated maize absorption of available phosphorus and nitrogen under water stress (Ghorchiani et al., 2018). Moreover, with increasing degree of water stress, the stress-mitigation function of the AMF became more obvious (Pavithra and Yapa, 2018). The AMF may have reduced the damage of water stress to plant growth by changing plant physiological processes (Pavithra and Yapa, 2018) and radial root water transport (Quiroga et al., 2019).

Photosynthesis is an important indicator of physiological sensitivity to water stress (Chaves et al., 2009). In our study, water stress inhibited photosynthesis in maize. Similar results have been reported by others (Hazrati et al., 2016; Ahanger et al., 2017). Many studies have shown that water stress reduces A, gs and E (Yan et al., 2016; Hura et al., 2007). The reduced photosynthesis under water stress is due to reductions in the coupling factor and production of ATP (Tezara et al., 1999). AMF improved the photosynthesis of maize under the different treatments in this study as well as previous studies (Sánchez-Blanco et al., 2004; Khalvati et al., 2005). In this study, in maize under water stress, Ci and VPD decreased and the A, gs, E and WUE increased in AMF-inoculated maize relative to non-inoculated maize. Increased stomatal functioning enhances CO_2 entry into the leaf tissues and its successful assimilation was observed in AMF-inoculated plants, such results have also been reported in AMF-inoculated wheat (Zhou et al., 2015). The high photosynthetic rates of A MF-inoculated plants under water stress could be explained by the increased production of photosynthetic pigments and a higher carboxylation efficiency relative to that of non-inoculated plants (Sánchez-Blanco et al., 2004; Zhu et al., 2014).

Though the leaf chlorophyll content of maize was measured by SPAD instrument in a real time, fast and non-destructive way in this study, it was a relative value with the limitation that not taking into consideration of the biomass (Lei et al., 2019). In our study, AMF inoculation increased the leaf chlorophyll content of maize under water stress, consistent with there port

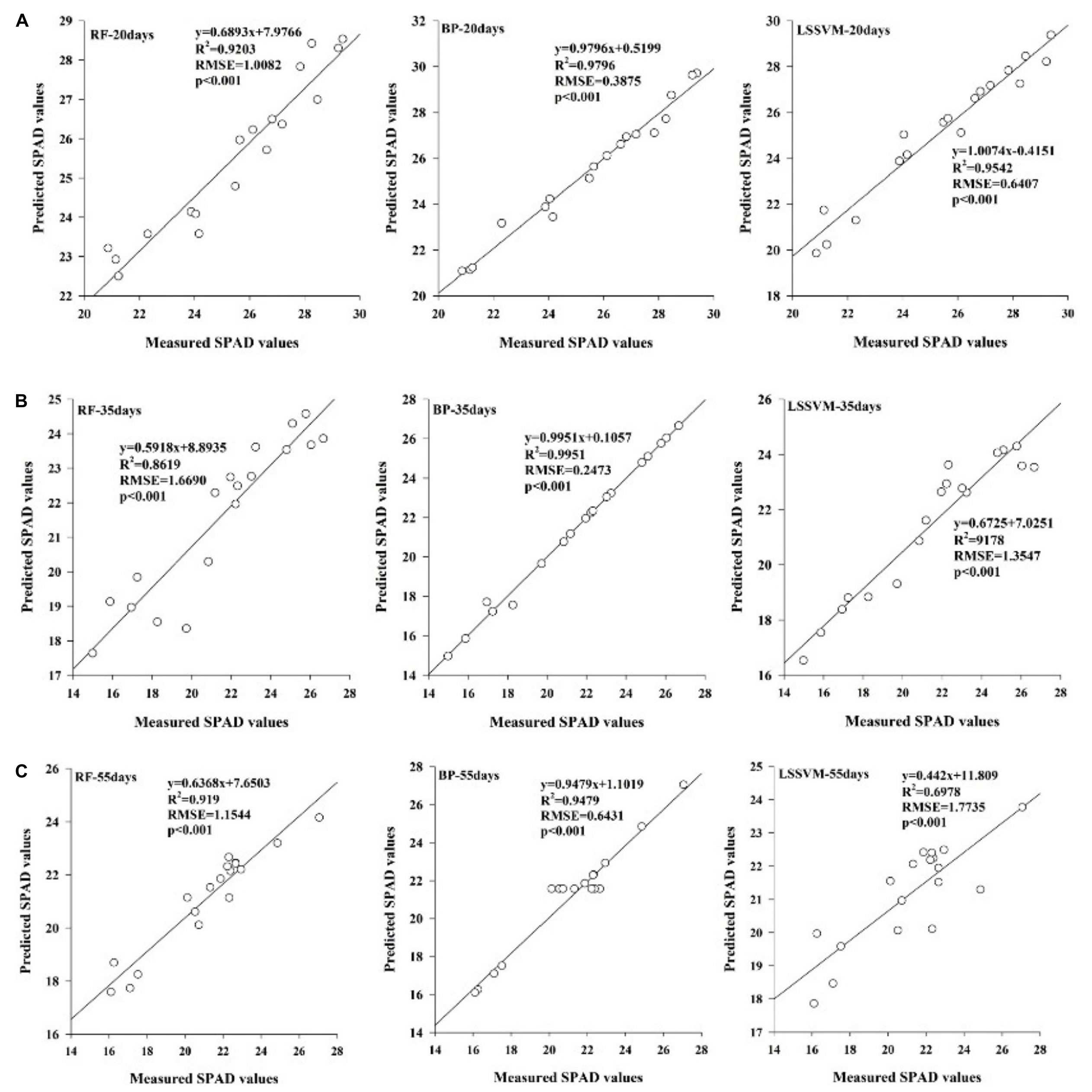


FIGURE 4 | Relationships of the measured SPAD values and predicted values based on the three machine learning algorithms under water stress durations of 20 days **(A)**, 35 days **(B)**, and 55 days **(C)**.

of other researchers (Begum et al., 2019). It may be due to the key role of Mg in chlorophyll synthesis, and Mg uptake increased in mycorrhizal plants (Mathur et al., 2018). The spectral reflectance of leaves is closely related to the leaf surface characteristics, leaf thickness, water content, and the contents of chlorophyll and other pigments (Feng et al., 2004). The wavelength region (500 to 900 nm) contains wavelengths with pigment absorption features (Merzlyak et al., 2003) as well as the red-edge (700 to 750 nm) (Mutanga and Skidmore, 2007). In our study, at the green peak (550 nm) and red edge (700 nm), The spectral reflectance of maize without inoculation under water stress was much higher than that of AMF-inoculated maize, it may be due to the more water utilization rate and chlorophyll content of inoculated maize compared to that without inoculation (Begum et al., 2019). The spectra lreflectance of maize leaves decreased with water stress duration, potentially due to greater damage of water stress to the leaves of maize at the seedling stage than at the later growth stages. The results of this study indicated that hyperspectral technology can be used to monitor changes in chlorophyll content in maize inoculated with mycorrhizae and exposed to water stress.

In the present study, compared with the original spectra, the first-order differential reflectance spectra better reflected the chlorophyll content of leaves at 400-800 nm. In other study, the correlation curve between the original spectrum and the chlorophyll relative content value was the best between the wavelengths 509-650 nm. The correlation between the first derivative spectrum and chlorophyll relative content value was the best and most stable at 450∼500 nm (Zhu et al., 2020). It can be explained by that the original spectrum can eliminate the influence of background noise on the spectrum, and reduce the scattering and absorption of light by atmosphere in the process of spectrum acquisition, after first-order differential processing (Mahlein et al., 2012; Manzo et al., 2013). The correlation coefficient between leaf color value and first-order differential spectral reflectance of inoculated maize at 550 and 700 nm were larger than that of non-inoculated maize under the stress duration of 35 or 55 days. The results varied in different conditions and environments. Hunt et al. (2013) found that the correlation between chlorophyll content and spectral reflectance at the canopy scale was strong in the visible region (400-760 nm). Some study showed that the red edge slope is the maximum reflectance in the red band, which can better reflect the chlorophyll content of plants and is often closely related to the photosynthetic rate of plants (Verrelst et al., 2010).

Our results demonstrated that all the models, constructed using the LSSVM, RF, and BP methods, could be used to estimate the chlorophyll content of inoculated maize under drought stress. The BP model showed the highest R^2 and slope and the lowest RMSE regardless of the water stress duration. BP model was more stable and accurate than others at estimation of leaf chlorophyll content (Shao et al., 2018). The best estimating model for the relationship between the leaf relative chlorophyll content and the reflectance spectra was the partial least squares (PLS) in the field experiment about winter wheat (Zhang et al., 2016). It may due to BP model was suitable for estimating the chlorophyll content in greenhouse experiment.

CONCLUSION

Arbuscular mycorrhizal fungi treatment increased the leaf chlorophyll content, A, gs, E, and WUE of maize but decreased the Ciofmaize and VPD regardless of the water stress duration or degree. The first-order differential reflectance of inoculated maize leaves was more significantly correlated with the chlorophyll content of maize than was the original spectral reflectance. Of the three machine learning models, the BP model achieved the largest coefficient of determination and smallest RMSE regardless of the stress duration. Thus, the BP model represented the optimal method of estimating the leaf chlorophyll content of inoculated maize.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LY supervised the work, contributed to data interpretation, and writing and revision of the manuscript. JS designed and performed the investigation, analyzed the data, contributed to data interpretation, and drafted the manuscript. XY provided the test site and revised the manuscript. JW and LL provided experimental instruments and modification of manuscript. EG and YK contributed to modification of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the School-level Innovation Fund (Nos. KJCX2020A06 and KJCX2020A05) and the Key R&D and Promotion Special Projects in Henan Province (Nos. 192102310184 and 202102310352).

ACKNOWLEDGMENTS

We thank the College of Resources and Environment of Henan Agricultural University for providing equipment.

REFERENCES

Ahanger, M. A., Tittal, M., Mir, R. A., and Agarwal, R. (2017). Alleviation of water and osmotic stress-induced changes in nitrogen metabolizing enzymes in *Triticum aestivum* L. cultivars by potassium. *Protoplasma* 254, 1953–1963. doi: 10.1007/s00709-017-1086-z

Arafat, A. H. A. L., and He, C. X. (2011). Effect of arbuscular mycorrhizal fungi on growth, mineral nutrition, antioxidant enzymes activity and fruit yield of tomato grown under salinity stress. *Sci. Hortic.* 127, 228–233. doi: 10.1016/j.scienta.2010.09.020

Bahraminia, M., Zarei, M., Ronaghi, A., Sepehri, M., and Etesami, H. (2020). Ionomic and biochemical responses of maize plant (*Zea mays* L.) inoculated with *Funneliformis mosseae* to water-deficit stress. *Rhizosphere* 16, 1–8. doi: 10.1016/j.rhisph.2020.100269

Bao, S. D. (1999). "Methods of soil phosphorus and potassium," in *Analysis of Soil Agriculture Chemistry*, 3rd Edn, ed. S. D. Bao (Beijing: China Agriculture Press). (in Chinese).

Basayigit, L., and Ozkul, M. (2015). Estimating Zn content of pepper under different fertilizer applications using a handheld spectrometer. *Int. J. Agric. Biol.* 17, 1–8.

Begum, H., Ahanger, M. A., Su, Y., Lei, Y., Mustafa, N. S. A., Ahmad, P., et al. (2019). Improved drought tolerance by AMF inoculation in maize (*Zea mays*) involves physiological and biochemical implications. *Plants* 8:579. doi: 10.3390/plants8120579

Biau, G., and Scornet, E. (2015). A random forest guided tour. *Test* 25, 197–227. doi: 10.1007/s11749-016-0481-7

Chaves, M. M., Flexas, J., and Pinheiro, C. (2009). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Ann. Bot.* 103, 551–560. doi: 10.1093/aob/mcn125

Chen, D. M., Shi, Y. Y., Huang, W. J., Zhang, J. C., and Wu, K. H. (2018). Mapping wheat rust based on high spatial resolution satellite imagery. *Comput. Electron Agric.* 152, 109–116. doi: 10.1016/j.compag.2018.07.002

Cheng, H. Q., Zou, Y. N., Wu, Q. S., and Kuca, K. (2021). Arbuscular mycorrhizal fungi alleviate drought stress in Trifoliate Orange by regulating H+-ATPase activity and gene expression. *Front. Plant Sci.* 12:659694. doi: 10.3389/fpls.2021.659694

Daryanto, S., Wang, L., and Jacinthe, P. A. (2016). Global synthesis of drought effects on maize and wheat production. *PLoS One* 11:e0156362. doi: 10.1371/journal.pone.0156362

Duc, N. H., Csintalan, Z., and Posta, K. (2018). Arbuscular mycorrhizal fungi mitigate negative effects of combined drought and heat stress on tomato plants. *Plant Phys. Biochem.* 132, 297–307. doi: 10.1016/j.plaphy.2018.09.011

El-Hendawy, S., Elsayed, S., Al-Suhaibani, N., Alotaibi, M., Tahir, M. U., Mubushar, M., et al. (2021). Use of Hyperspectral reflectance sensing for assessing growth and chlorophyll content of spring wheat grown under simulated saline field conditions. *Plants Basel* 10, 1–26. doi: 10.3390/plants10010101

Feng, X. W., Chen, W. X., and Bao, A. M. (2004). Analysis on the cotton physiological change and its hyperspectral response under the water stress conditions. *Arid Land Geogr.* 2, 121–126. (in Chinese), doi: 10.4324/9780203221099-25

Gao, Y. C., Wang, J. N., Guo, S. H., Hu, Y. L., Li, T. T., Mao, R., et al. (2015). Effects of salinization and crude oil contamination on soil bacterial community structure in the Yellow River Delta region, China. *Appl. Soil Ecol.* 86, 165–173. doi: 10.1016/j.apsoil.2014.10.011

Ghorchiani, M., Etesami, H., and Alikhani, H. (2018). Improvement of growth and yield of maize under water stress by co-inoculating an arbuscular mycorrhizal fungus and a plant growth promoting rhizobacterium together with phosphate fertilizers. *Agric. Ecosyst. Environ.* 258, 59–70. doi: 10.1016/j.agee.2018.02.016

Gu, Y. F., Ding, S. Y., Chen, H. S., Gao, Z. Y., and Xing, Q. (2008). Ecophysiological responses and hyperspectral characteristics of winter wheat (*Triticum aestivum*) under drought stress. *Acta Ecol. Sin.* 28, 2690–2697. (in Chinese),

Hazrati, S., Tahmasebi-Sarvestani, Z., Modarres-Sanavy, S. A. M., Mokhtassi-Bidgoli, A., and Nicola, S. (2016). Effects of water stress and light intensity on chlorophyll fluorescence parameters and pigments of *Aloe vera* L. *Plant Physiol. Biochem.* 106, 141–148. doi: 10.1016/j.plaphy.2016.04.046

Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., Perry, E. M., and Akhmedov, B. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Geoinformation* 21, 103–112. doi: 10.1016/j.jag.2012.07.020

Hura, T., Hura, K., Grzesiak, M., and Rzepka, A. (2007). Effect of long-term drought stress on leaf gas exchange and fluorescence parameters in C3 and C4 plants. *Acta Physiol. Plant* 29, 103–113.

Jiang, J. B., Chen, Y. H., and Huang, W. J. (2010). Using hyperspectral remote sensing to estimate canopy chlorophyll density of wheat under yellow rust stress. *Spectrosc. Spectral Anal.* 30, 2243–2247. (in Chinese)

Kahil, M. T., Dinar, A., and Albiac, J. (2015). Modelling water scarcity and droughts for policy adaptation to climate change in arid and semiarid regions. *J. Hydrol.* 522, 95–109. doi: 10.1016/j.jhydrol.2014.12.042

Karimkhani, S., Ardakani, M. R., Paknejad, F., and Manavi, P. N. (2021). Alleviation drought stress of *Bromus* species using mycorrhizal fungi contributed with drought-responsive biomarker. *Iran. J. Sci. Technol. Trans. Sci.* 45, 481–491. doi: 10.1007/s40995-021-01073-9

Khalvati, M. A., Hu, Y., Mozafar, A., and Schmidhalter, U. (2005). Quantification of water uptake by arbuscular mycorrhizal hyphae and its significance for leaf growth, water relations and gas exchange of barley subjected to drought stress. *Plant Biol.* 7, 706–712. doi: 10.1055/s-2005-872893

Lei, X. X., Zhao, J., Liu, H. C., Zhang, J. Y., Liang, W. Y., Tian, J. L., et al. (2019). Inversion of chlorophyll content and SPAD value of vegetable leaves based on PROSPECT model. *Spectrosc. Spectral Anal.* 39, 3256–3260.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature* 529, 84–87. doi: 10.1038/nature16467

Li, Y. J., Tang, X. J., and Liu, J. H. (2010). Application of least square support vector machine based on particle swarm optimization in quantitative analysis of gas mixture. *Spectrosc. Spectral Anal.* 30, 774–778. (in Chinese),

Li, Z. W., Xin, X. P., Tang, H., Yang, F., Chen, B. R., and Zhang, B. H. (2017). Estimating grassland LAI using the random forests approach and landsat imagery in the meadow steppe of Hulunber, China. *J. Integr. Agric.* 16, 286–297. doi: 10.1016/S2095-3119(15)61303-X

Liu, L., Li, J. W., Yue, F. X., Yan, X. W., Wang, F. Y., Bloszies, S. A., et al. (2018). Effects of arbuscular mycorrhizal inoculation and biochar amendment on maize growth, cadmium uptake and soil cadmium speciation in Cd-contaminated soil. *Chemosphere* 194, 495–503. doi: 10.1016/j.chemosphere.2017.12.025

Mahlein, A. K., Steiner, U., Christian, H., Dehne, H. W., and Oerke, E. C. (2012). Hyperspectral imaging for small-scale analysis of symptoms caused by di_erent sugar beet diseases. *Plant Methods* 8:3. doi: 10.1186/1746-4811-8-3

Mall, R., and Suykens, J. A. K. (2015). Very sparse lssvm reductions for large-scale data. *IEEE Trans. Neural Netw. Learn. Syst.* 26, 1086–1097. doi: 10.1109/TNNLS.2014.2333879

Manzo, C., Salvini, R., Guastaldi, E., Nicolardi, V., and Protano, G. (2013). Reflectance spectral analyses for the assessment of environmental pollution in the geothermal site of Mt, Amiata (Italy). *Atmospher. Environ.* 79, 650–665. doi: 10.1016/j.atmosenv.2013.06.038

Mathur, S., Sharma, M. P., and Jajoo, A. (2018). Improved photosynthetic efficacy of maize (*Zea mays*) plants with arbuscular mycorrhizal fungi (AMF) under high temperature stress. *J. Photochem. Photobiol. B* 180, 149–154. doi: 10.1016/j.jphotobiol.2018.02.002

Meddich, A., Rahou, Y. A., Boutasknit, A., Ait-El-Mokhtar, M., Fakhech, A., Lahbouki, S., et al. (2021). Role of mycorrhizal fungi in improving the tolerance of melon (*Cucumus melo*) under two water deficit partial root drying and regulated deficit irrigation. *Plant Biosyst.* 155, 1–12. doi: 10.1080/11263504.2021.1881644

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., Solovchenko, A. E., and Pogosyan, S. I. (2003). Application of reflectance spectroscopy for analysis of higher plant pigments. *Russ. J. Plant Physiol.* 50, 704–710. doi: 10.1023/A:1025608728405

Mutanga, O., and Skidmore, A. K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS J. Photogramm. Rem. Sens.* 62, 34–42. doi: 10.1016/j.isprsjprs.2007.02.001

Ortega-Farias, S., Espinoza-Meza, S., Lopez-Olivari, R., Araya-Alman, M., and Carrasco-Benavides, M. (2021). Effects of different irrigation levels on plant water status, yield, fruit quality, and water productivity in a drip-irrigated blueberry orchard under Mediterranean conditions. *Agric. Water Manag.* 249, 1–10. doi: 10.1016/j.agwat.2021.106805

Pavithra, D., and Yapa, N. (2018). Arbuscular mycorrhizal fungi inoculation enhances drought stress tolerance of plants. *Groundw. Sustain. Dev.* 7, 490–494. doi: 10.1016/j.gsd.2018.03.005

Pedranzani, H., Rodríguez-Rivera, M., Gutiérrez, M., Porcel, R., Hause, B., and Ruiz-Lozano, J. M. (2016). Arbuscular mycorrhizal symbiosis regulates physiology and performance of digitaria eriantha plants subjected to abiotic stresses by modulating antioxidant and jasmonate levels. *Mycorrhiza* 26, 141–152. doi: 10.1007/s00572-015-0653-4

Phillips, J. M., and Hayman, D. S. (1970). Improved procedures for clearing roots and staining parasitic and vesicular-arbuscular mycorrhizal fungi for rapid assessment of infection. *Trans. Br. Mycol. Soc.* 55, 158–161. doi: 10.1016/S0007-1536(70)80110-3

Quiroga, G., Erice, G., Aroca, R., Chaumont, F., and Ruiz-Lozano, J. M. (2019). Contribution of the arbuscular mycorrhizal symbiosis to the regulation of radial root water transport in maize plants under water deficit. *Environ. Exp. Bot.* 167, 1–10. doi: 10.1016/j.envexpbot.2019.103821

Quiroga, G., Erice, G., Aroca, R., Chaumont, F., and Ruiz-Lozano, J. M. (2017). Enhanced drought stress tolerance by the arbuscular mycorrhizal symbiosis in a drought-sensitive maize cultivar is related to a broader and differential regulation of host plant aquaporins than in a drought-tolerant cultivar. *Front. Plant Sci.* 8: 1056. doi: 10.3389/fpls.2017.01056

Saitta, L., Giordana, A., and Cornuejols, A. (2011). *Phase Transitions in Machine Learning*. Cambridge: Cambridge University Press.

Saleem, M. H., Potgieter, J., and Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants* 8:468. doi: 10.3390/plants8110468

Salloum, M. S., Menduni, M. F., and Luna, C. M. (2018). A differential capacity of arbuscular mycorrhizal fungal colonization under well-watered conditions and its relationship with drought stress mitigation in unimproved vs. Improved soybean genotypes. *Botany* 96, 135–144. doi: 10.1139/cjb-2017-0137

Sánchez-Blanco, M. J., Ferrández, T., Morales, M. A., and Alarcón, J. J. (2004). Variations in water status, gas exchange, and growth in *Rosmarinus officinalis* plants infected with *Glomus deserticola* under drought conditions. *J. Plant Phisiol.* 161, 675–682. doi: 10.1078/0176-1617-01191

Schewe, J., Heinke, J., and Gerten, D. (2014). Multi-model assessment of water scarcity under climate change. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3245–3250. doi: 10.1073/pnas.1222460110

Shao, D. G., Nong, X. Z., Tan, X. Z., Chen, S., Xu, B. L., and Hu, N. J. (2018). Daily water quality forecast of the South-To-North water diversion project of China based on the cuckoo search-back propagation neural network. *Water* 10:1471. doi: 10.3390/w10101471

Smith, S., and Read, D. (2008). *Mycorrhizal Symbiosis*, 3rd Edn. London: Academic Press.

Sonobe, R., Yamashita, H., Mihara, H., Morita, A., and Ikka, T. (2021). Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms. *Int. J. Rem. Sens.* 42, 1311–1329. doi: 10.1080/01431161.2020.1826065

Tezara, W., Mitchell, V. J., Driscoll, S. D., and Lawlor, D. W. (1999). Water stress inhibits plant photosynthesis by decreasing coupling factor and ATP. *Nature* 401, 914–917. doi: 10.1038/44842

Tyagi, J., Varma, A., and Pudake, R. N. (2017). Evaluation of comparative effects of arbuscular mycorrhiza (*Rhizophagus intraradices*) and endophyte (*Piriformospora indica*) association with finger millet (*Eleusine coracana*) under drought stress. *Eur. J. Soil Biol.* 81, 1–10. doi: 10.1016/j.ejsobi.2017.05.007

Verrelst, J., Schaepman, M. E., Malenovs'y, Z., and Clevers, J. G. P. W. (2010). _Effects of woody elements on simulated canopy reflectance: implications for forest chlorophyll content retrieval. *Rem. Sens. Environ.* 2010, 647–656. doi: 10.1016/j.rse.2009.11.004

Yan, W. M., Zhong, Y. Q., and Shangguan, Z. P. A. (2016). Meta-analysis of leaf gas exchange and water status responses to drought. *Sci. Rep.* 6:20917. doi: 10.1038/srep20917

Zhang, J. F., Han, W. T., Huang, L. W., Zhang, Z. Y., Ma, Y. M., and Hu, Y. M. (2016). Leaf chlorophyll content estimation of winter wheat based on visible and near-infrared sensors. *Sensors* 16, 1–11.

Zhang, T., Hu, Y. J., Zhang, K., Tian, C. Y., and Guo, J. X. (2018). Arbuscular mycorrhizal fungi improve plant growth of *Ricinus communis* by altering photosynthetic properties and increasing pigments under drought and salt stress. *Ind. Crops Prod.* 117, 13–19. doi: 10.1016/j.indcrop.2018.02.087

Zhou, Q., Ravnskov, S., Jiang, D., and Wollenweber, B. (2015). Changes in carbon and nitrogen allocation, growth and grain yield induced by arbuscular mycorrhizal fungi in wheat (*Triticum aestivum* L.) subjected to a period of water deficit. *Plant Growth Regul.* 75, 751–760. doi: 10.1007/s10725-014-9977-x

Zhu, J. Y., He, W. J., Yao, J. M., Yu, Q., Xu, C. Y., Huang, H. G., et al. (2020). Spectral reflectance characteristics and chlorophyll content estimation model of *Quercus aquifolioides* leaves at different altitudes in Sejila Mountain. *Appl. Sci.* 10:3636. doi: 10.3390/app10103636

Zhu, X. Q., Wang, C. Y., Chen, H., and Tang, M. (2014). Effects of arbuscular mycorrhizal fungi on photosynthesis, carbon content and calorific value of black locust seedlings. *Photosynthetica* 52, 247–252. doi: 10.1007/s11099-014-0031-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sun, Yang, Yang, Wei, Li, Guo and Kong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 23 July 2021
doi: 10.3389/fbioe.2021.696292

Rapid and Accurate Varieties Classification of Different Crop Seeds Under Sample-Limited Condition Based on Hyperspectral Imaging and Deep Transfer Learning

Na Wu[1], Fei Liu[1], Fanjia Meng[2], Mu Li[3], Chu Zhang[4]* and Yong He[1]*

[1] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China, [2] College of Information and Electrical Engineering, China Agricultural University, Beijing, China, [3] Maize Research Institute, Jilin Academy of Agricultural Sciences, Gongzhuling, China, [4] School of Information Engineering, Huzhou University, Huzhou, China

OPEN ACCESS

Edited by:
Fang-Hao Wan,
Institute of Plant Protection, Chinese Academy of Agricultural Sciences, China

Reviewed by:
Xi Qiao,
Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, China
Sumit Ghosh,
The Research Institute at Nationwide Children's Hospital, United States

****Correspondence:***
Chu Zhang
chuzh@zjhu.edu.cn
Yong He
yhe@zju.edu.cn

Specialty section:
This article was submitted to Biosafety and Biosecurity, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *16 April 2021*
Accepted: *23 June 2021*
Published: *23 July 2021*

Citation:
Wu N, Liu F, Meng F, Li M, Zhang C and He Y (2021) Rapid and Accurate Varieties Classification of Different Crop Seeds Under Sample-Limited Condition Based on Hyperspectral Imaging and Deep Transfer Learning. Front. Bioeng. Biotechnol. 9:696292. doi: 10.3389/fbioe.2021.696292

Rapid varieties classification of crop seeds is significant for breeders to screen out seeds with specific traits and market regulators to detect seed purity. However, collecting high-quality, large-scale samples takes high costs in some cases, making it difficult to build an accurate classification model. This study aimed to explore a rapid and accurate method for varieties classification of different crop seeds under the sample-limited condition based on hyperspectral imaging (HSI) and deep transfer learning. Three deep neural networks with typical structures were designed based on a sample-rich Pea dataset. Obtained the highest accuracy of 99.57%, VGG-MODEL was transferred to classify four target datasets (rice, oat, wheat, and cotton) with limited samples. Accuracies of the deep transferred model achieved 95, 99, 80.8, and 83.86% on the four datasets, respectively. Using training sets with different sizes, the deep transferred model could always obtain higher performance than other traditional methods. The visualization of the deep features and classification results confirmed the portability of the shared features of seed spectra, providing an interpreted method for rapid and accurate varieties classification of crop seeds. The overall results showed great superiority of HSI combined with deep transfer learning for seed detection under sample-limited condition. This study provided a new idea for facilitating a crop germplasm screening process under the scenario of sample scarcity and the detection of other qualities of crop seeds under sample-limited condition based on HSI.

Keywords: crop seeds, hyperspectral imaging, classification model, spectroscopic analysis, deep learning

INTRODUCTION

High-quality seeds are conducive to continue excellent species and guarantee crop yield and quality. Due to the significant differences in climate, soil, and water resources in different regions, breeders have pointedly developed many crop varieties to adapt to the local planting environment. Growth rules, stress resistance, and biochemical characteristics of different varieties of crops vary greatly

(Du et al., 2017; Zhang et al., 2020). For varieties that are still in the breeding stage, screening a variety with specific traits often needs to observe the phenotypic traits of the offspring plants, which is time-consuming and labor-intensive. As a seed carries all the genetic genes that develop into a plant, seed classification can be an alternative for screening variety with specific traits. For varieties that have been promoted widely, different varieties of seeds frequently circulate in the market, tending to be easily mixed, making the seed purity unable to be guaranteed. Conventionally, the manual vision inspection method based on the external phenotypic traits of seeds, like color, texture, and shape, is subjective and boring (Rashid and Singh, 2000). The more accurate methods based on the internal biochemical properties of seeds, such as DNA molecular markers (Ye et al., 2013) and protein electrophoresis techniques (Shuaib et al., 2007), rely on chemical agents and complex operation. Accordingly, it is necessary to develop a rapid and accurate method for the varieties classification of crop seeds.

As hyperspectral imaging (HSI) can obtain spectral and spatial location information simultaneously during one scan, it has the capability of probing the internal and external phenotypic traits of samples rapidly (Sendin et al., 2018). HSI has gained tremendous and continuous attention in breed screening (Feng et al., 2017), plant phenotyping (Qiu R. et al., 2018; Sun et al., 2019), and environment monitoring (Stuart et al., 2019). In seed-related tasks like determination of seed quality (Shrestha et al., 2016), diagnosis of seed diseases (Wu et al., 2020) and detection of seed components (Caporaso et al., 2018), HSI has been utilized as a rapid and accurate alternative. Since hyperspectral image contains a large amount of redundant collinear information, diverse linear and non-linear machine learning approaches, such as partial least squares discriminant analysis (PLS-DA), extreme learning machine (ELM), and least square support vector machines (LSSVM), were introduced to couple the relationship between seed spectra and a category label or component content (Caporaso et al., 2018; Kong et al., 2018; Weng et al., 2018).

In recent years, with the attention from academia and industry increasing, deep learning as the new state-of-the-art machine learning approach has also been applied in the spectral analysis field gradually (Jin et al., 2018; Wei et al., 2018; Yu et al., 2018). Compared with traditional approaches, deep learning can extract various low-level and high-level features automatically through a multilayered stack network structure (LeCun et al., 2015). This advantage can reduce the requirement of prior knowledge from specific tasks and human effort in feature engineering, which is very beneficial for analyzing redundant and high-dimensional spectral data.

However, typical deep learning models, such as deep networks, generally have serious big data dependencies. A high-performance deep network requires enough samples to adequately learn the feature patterns hidden in the massive and redundant spectral data. Unfortunately, in some tasks like seeds screening with specific traits during the breeding process or quality detection of precious agricultural products, it is challenging to establish a large-scale, high-quality dataset because of the high cost of obtaining and labeling samples (Lee et al., 2016; Xu et al., 2017; Sun et al., 2019). Besides, the precious data acquired at great expense is straightforward to be outdated and difficult to be reused in new tasks, which dramatically limits the rapid application of well-performing methods like the deep network in spectral analysis. Another problem is that the deep networks developed for different tasks are generally based on a common assumption, that is, training and testing data lie in the same feature space and have the same distribution (Weiss et al., 2016). Therefore, even for similar tasks, the tiny differences in the distribution of different datasets will make the network not reusable.

The emergence of transfer learning brings hope for solving the above two problems. The transfer learning method allows the training and testing data to lie in different feature spaces. It mainly investigates how to transfer useful knowledge from the relevant source domain to the target domain (Pan and Yang, 2010). This property not only relieves the demand for a large number of samples in the target task but makes reusing the shared knowledge like model structure and feature representation in the source domain possible. The target task can be expected completed, using limited samples and computation overhead. Deep transfer learning is the product of the combination of deep learning and transfer learning. It aims to study how to use the deep neural network to transfer knowledge and has been widely used in the computer vision field (Mohanty et al., 2016; Ghazi et al., 2017; Tan et al., 2018).

However, the deep transfer learning technique has not received much attention in the field of spectral analysis. Most studies perform task analysis at a pixel level based on remote sensing images, such as poverty mapping (Xie et al., 2016), image superresolution processing (Yuan et al., 2017), and crop yield prediction (Wang et al., 2018). For ground spectral images, Liu et al. (2018) showed the effectiveness of deep transfer learning in predicting soil clay content in different soils. For seeds of different crops, there are also certain similarities, for example, the structure of the seeds. Most seeds contain a seed coat, an embryo, and endosperm. These parts contain some common chemical components, like starch, fat, and enzymes, which are necessary for a seed to develop into a seedling (Beníteza et al., 2013; Zhao et al., 2018). This commonality may lead to the similarities among the spectral characteristics of different crop seeds. Therefore, when constructing a deep model for seed varieties classification of a specific crop based on HSI, the knowledge in the model is possible to be transferred to the classification tasks of other crop seeds. In this study, we aimed to investigate the feasibility of the deep transfer learning technique for the varieties classification of different crop seeds based on HSI.

The specific objectives were: (1) to develop a deep network model with excellent performance based on a sample-rich dataset; (2) to transfer common knowledge to the varieties classification of other crop seeds with sample-limited datasets through the deep network; (3) to evaluate the impact of training set size on the performance of transfer learning; and (4) to visualize the transferring process of deep network and the classification results. We hope to provide a common framework for rapid and accurate varieties classification of crop seeds under sample-limited condition through this study.

MATERIALS AND METHODS

Sample Collection and Dataset Description

This study investigated five kinds of crop seeds, including pea, rice, oat, wheat, and cotton. All images were obtained by the same line-scanning near-infrared HSI system, covering a spectral range from 874.41 to 1,733.91 nm with a resolution of 5 nm (Wu et al., 2018). An ImSpector N17E imaging spectrograph (Spectral Imaging Ltd., Oulu, Finland) and a Xeva 922 CCD camera (Xenics Infrared Solutions, Leuven, Belgium), configured with an OLES22 lens (Spectral Imaging Ltd., Oulu, Finland), were the critical components of this system. In addition, the illumination was provided by two 150 W tungsten halogen lamps (3900e Lightsource; Illumination Technologies Inc.; West Elbridge, NY, United States) set symmetrically under the camera. Multiple seed samples placed on a dark plate flowed a miniature conveyer belt to achieve batch detection. A hyperspectral image, containing 256 spectral channels, could be obtained through each scan by this system and then calibrated using the following formula.

$$I_c = \frac{I_r - I_d}{I_w - I_d}$$

where I_r and I_c represented the raw hyperspectral image and the corrected image, I_w and I_d represented the white and dark reference image. Each seed in the hyperspectral image was regarded as a region of interest (ROI). To get the mask of all the ROIs, simple threshold segmentation and morphological operation were performed on the channel image with the strongest contrast between the background and the seeds. Then the spectral vectors of all pixels within each ROI were extracted, and the bands in head and end ranges were removed to avoid noise introduced by the instability of the system. The reserved spectra with a range of 975–1,646 nm were further processed by wavelet transform (WT). The spectrum vector, representing a seed sample, was finally obtained by averaging all the transformed pixel spectra in one ROI.

Five spectra datasets with similar but different distributions were established in this study. Their detailed collection parameters and description information were summarized in **Table 1**. It should be noted that different parameters were set for imaging different crop seeds clearly since different seeds have different external phenotypes, such as size, height, and color. The most abundant dataset, the Pea dataset, contained a total of 10,420 samples from four varieties named Baiyan (2697), Heiyan (2,848), Changshouren (2,849), and Zhewan 1 (2,026), which were widely cultivated in southern China. Peas of the first two varieties generally need to be roasted before eating, while the latter two can be directly eaten due to the high water and sugar content. All the seeds were purchased from the Lvfeng seed company in Hangzhou, Zhejiang, China, in 2018. The dataset corresponding to each variety was randomly divided into a training set, a validation set, and a testing set at a ratio of 3:1:1. Then those independent subsets with the same category were merged and shuffled. Because of its large volume of data, the Pea dataset was selected as the source dataset.

TABLE 1 | Description of the source and target datasets.

Datasets		Parameters^1	#Variety	#Total	#Training	#Validation	#Testing
Source	Pea	(15.5, 3, 12)	4	10,420	6,252	2,084	2,084
Target	Rice	(9, 3, 11)	3	750	150	300	300
	Oat	(15.2, 3, 11.5)	4	1,000	200	400	400
	Wheat	(15, 3, 13)	5	1,250	250	500	500
	Cotton	(14, 3, 11.5)	7	1,750	350	700	700

Parameters^1 represents parameters of the hyperspectral imaging system, including the distance between the camera and the seed plate (cm), the exposure time of the camera (ms), and the speed along the x-axis of seeds movement (mm.s^{-1}).

The other four sample-limited datasets were used as the target datasets designed to contain different numbers of seed varieties for investigating their impact on the transferring effect. Each variety in these datasets contained 250 samples and was further divided into three subsets at a ratio of 1:2:2 to reflect sample-limited condition.

The first target dataset consisted of 750 spectral samples of three varieties of rice seeds, including Yongyou 9, Nuoyou 6211, and Zhongbaiyouhuazhan. These varieties are all hybrid rice with indica property and belong to hybrid indica-japonica, hybrid indica-glutinous, and hybrid indica rice, respectively. All seeds were collected by the College of Agriculture and Biotechnology, Zhejiang University in 2019.

The second dataset was the Oat dataset with the same number of varieties as the source dataset. It contained 1,000 samples from four varieties named Bayan 6, Dingyan 2, Muwang, and Jizhangyan 4, which were widely planted in the grasslands of northern China. The seeds harvested in 2017 were kindly provided by the Academy of Agricultural and Animal Sciences, Inner Mongolia, China.

A total of 1,250 samples from five varieties of wheat seeds, including Zhenmai 9, Annong 1,124, Longpingmian 6, Shannong 102, and Weilong 169, formed the Wheat dataset. These five varieties were extensively cultivated in the winter wheat regions of southern China. The seed samples were friendly provided by Anhui longping high-tech seed industry Co., Ltd., in Hefei, Anhui, China, in 2018.

The fourth dataset, the Cotton dataset, was consisted of 1,750 samples of seven varieties of cotton seeds. They were Jinxin 5, Jinxin 7, Shennongmian 1, Xinjiangzaomian 1, Xinluzaomian 29, Xinluzhong 52, and Xinluzhong 42. These varieties were mainly grown in Xinjiang Uyghur Autonomous Region, the largest cotton-producing region in China. And the cotton seeds were collected by Shihezi University in 2016.

In this study, multiple deep neural networks with different structures were first developed, using the source dataset. Then the optimal deep model was selected as the model to be transferred by comparing the classification accuracies. The transfer learning technique was investigated to transfer useful knowledge from the optimal deep model to the analysis of four target datasets. The training set of each target dataset was further transformed into 10 datasets to analyze the impact of sample size on transfer learning by randomly selected 10–100% samples from the original training set. Four commonly used multivariate analysis methods, including two linear methods: linear discriminant analysis (LDA) and PLS-DA, and two non-linear methods: multilayer perceptron

(MLP) and support vector machines (SVM), were introduced as the benchmarks.

Deep Classification Models Development

In the computer vision field, the huge image library, ImageNet, has spawned many excellent deep learning models like VggNet, InceptionNet, and ResNet (Krizhevsky et al., 2012). The specialness of VggNet is using small convolution kernels. The designers believed that using multiple convolution layers equipped with a 3 × 3 kernel to replace a convolution layer with a 5 × 5 kernel could reduce the network parameters and increase non-linear mapping, thereby increasing the representation capability (Simonyan and Zisserman, 2015). ResNet is also an outstanding network with many variations. What makes it unique is the introduction of residual learning. The residual module directly bypasses the input of a particular layer to the output, which makes ResNet only need to learn the residual between the input and the output (He et al., 2016). This manner solves the problem of performance degradation when the network depth increases. InceptionNet was born in the ILSVRC2014 competition. The most significant innovation of this network is introducing a module called "Inception" to replace the typical structure of the convolution layer, cascading the pooling layer (Szegedy et al., 2015). This Inception module contains four branches with different receptive fields to perceive the input patterns. By utilizing this module, InceptionNet can increase its width and learn more local features of different scales.

Inspired by these network structures, three one-dimensional deep neural networks were developed for the source dataset in this study, as shown in **Figure 1**.

The first one was VGG-MODEL. Two V blocks (**Figure 2**), containing two convolution layers equipped with a 1 × 3 kernel were designed to extract the feature patterns hidden in the spectral vectors. A batch normalization (BN) and an activation function, exponential linear unit (ELU), were inserted after each convolution to reduce the overfitting risk and speed up the convergence process. The number of convolution filters was set to 16 for the first V block and 32 for the second V block. A max-pooling layer was placed behind each V block to reduce the feature dimension. A flatten layer was set after the last max-pooling layer to convert its output feature into a one-dimensional vector form. Layer Fc1 and Fc2, consisting of 64 and 4 neurons, were used to perform the classification task like traditional neural networks. BN and ELU were also used behind Fc1. VGG-MODEL finally output the probability of the input spectral vector belonging to each category through a softmax function.

The second one was RES-MODEL. The first part of this network was similar to half of the V block, which contained a convolutional layer followed by BN, ELU, and a max-pooling layer. The difference was that the convolutional layer used 32 kernels, with a size of 1 × 7. The second part consisted of four cascaded residual modules, R block. This module was similar to the V block but added a transmission channel from input to output (**Figure 2**). The number of convolutional filters in the first R block was 32 and was doubled as the blocks going deeper. An average pooling layer was placed after the last R block to average the features in the spectral dimension. This layer could decrease the parameters in fully connected layers, thereby reducing the overfitting risk. The last part of RES-MODELDE was similar to that of VGG-MODEL but was equipped with one fully connected layer, Fc1, with four neurons.

The third one was INCEPTION-MODEL. Having the same structure as that of RES-MODEL, the first part of this network utilized 16 convolution filters, with a size of 1 × 3. It was followed by four I blocks (**Figure 2**), each of which cascaded a max-pooling layer except the last one. The number of filters in the first I block was 16 and was doubled as the blocks going deeper. As shown in **Figure 2**, the I block transmitted its input to four parallel branches. Three of them were convolution layers with 1 × 1, 1 × 3, and 1 × 5 kernels, respectively. They were employed to extract local spectral features at different scales. A 1 × 1 convolution was placed before 1 × 3 and 1 × 5 convolution to reduce the number of input channels. The last branch performed the max-pooling operation. The end of the INCEPTION-MODEL was similar to that of RES-MODEL.

To fairly compare the performance, these three deep networks employed cross-entropy as the objective function and used stochastic gradient descent (SGD) optimization algorithm. The learning rate and momentum were all set to 0.001 and 0.9, respectively. After debugging many times, the number of samples input into the network at one time, *batch_size*, was set to 128, and the number of training iterations, *epoch*, was set to 400. All networks were trained, using the training set of the source dataset. The model for each network that obtained the highest accuracy on the validation set was saved. The effectiveness of the model was evaluated on the testing set. The detailed parameters of these three networks were shown in **Supplementary Table 1**.

Transfer Learning Strategy

As an emerging tool in machine learning, transfer learning was proposed to remit the requirement of models for sufficient training data by transferring available knowledge from the relevant source domain to the target domain (Pan and Yang, 2010). The typical process of transfer learning was shown in **Figure 3A**. We defined a domain $\mathcal{D} = \{\mathcal{X}, \mathrm{P}(\mathcal{X})\}$ where $\mathcal{X}$ represented a feature space and $\mathrm{P}(\mathcal{X})$ represented its probability distribution, and a task $\mathcal{T} = \{y, f(.)\}$where y represented a label space and f represented a transformed function. When the task $\mathcal{T}$ was performed in the domain $\mathcal{D}$, f modeled $\mathrm{P}(y|\ x)$, where $y \in \mathcal{Y}$, $x \in \mathcal{X}$. In the transfer learning field, there are two domains: source domain $\mathcal{D}_S$ with task $\mathcal{T}_S$ and target domain $\mathcal{D}_T$ with task $\mathcal{T}_T$. The main goal of transfer learning is to improve the performance of the transformed function in the target domain $f_T(.)$, using the knowledge learned in $\mathcal{D}_S$ and $\mathcal{T}_S$, where $\mathcal{D}_S$ (or $\mathcal{T}_S$) and $\mathcal{D}_T$ (or $\mathcal{T}_T$) are different but relevant.

For deep transfer learning, $f(.)$ is various deep models designed for specific tasks. These deep models contain rich knowledge. Some knowledge is closely related to the specific task, while others can be shared between different tasks or objects. Deep transfer learning aims to transfer the common knowledge into the current target task to avoid learning this knowledge repeatedly, thus achieving rapid modeling. The structure of the model and the weight of the network are two important types of knowledge contained in deep models. In this study,

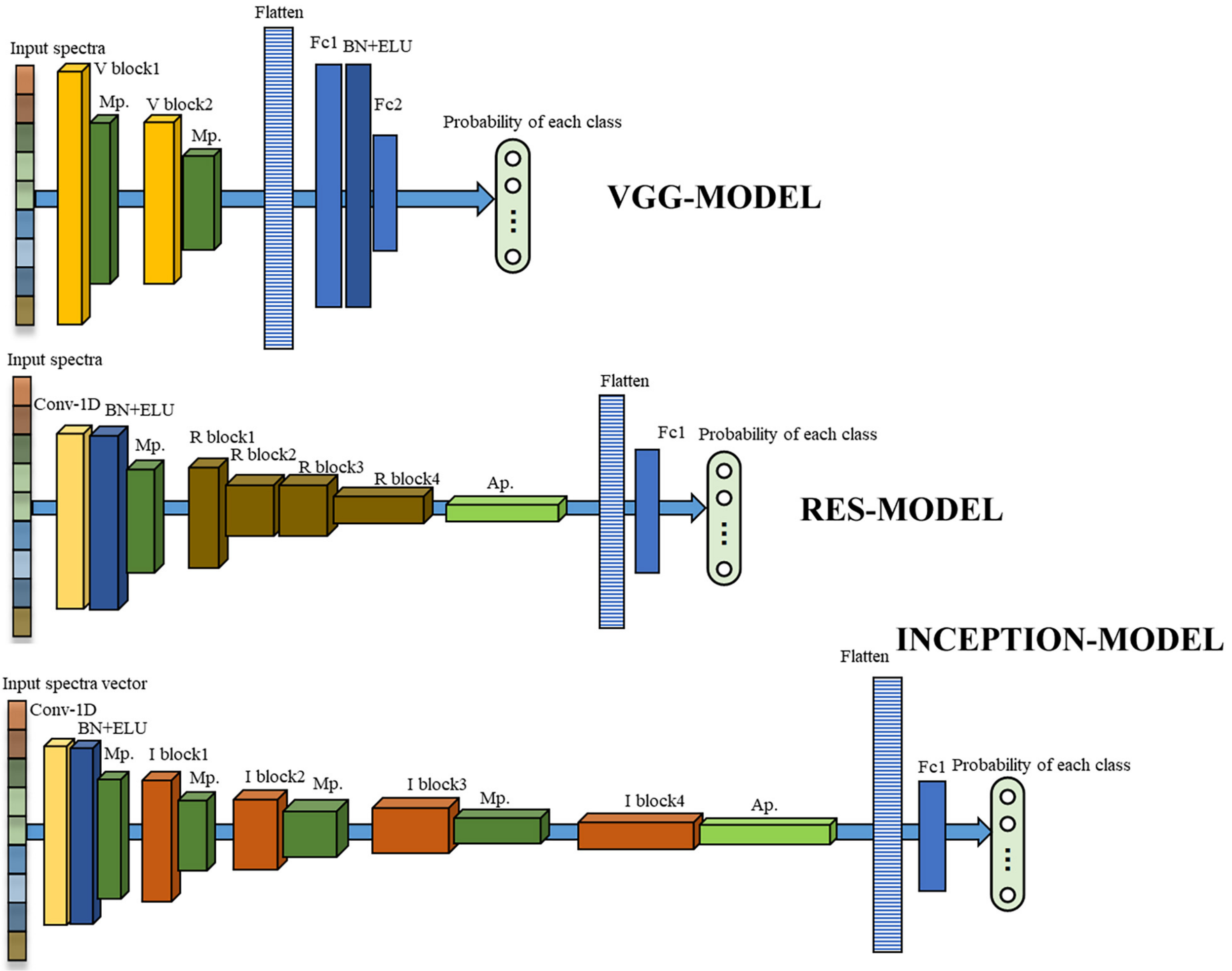


FIGURE 1 | The structures of three developed deep neural networks. The orange, brown, and red cubes represented the V block, R block, and I block in **Figure 2**, respectively. The dark green and light green cubes represented the max-pooling layer and the average pooling layer, respectively. The yellow and blue cubes represented a one-dimensional convolutional layer, BN cascade activation function ELU, respectively. The striped bars represent flattened layers, and the light blue bars represented fully connected layers. The length, width, and height of the cubes and the bars in the Figure were drawn according to the dimensional size of data in each layer for a more intuitive display.

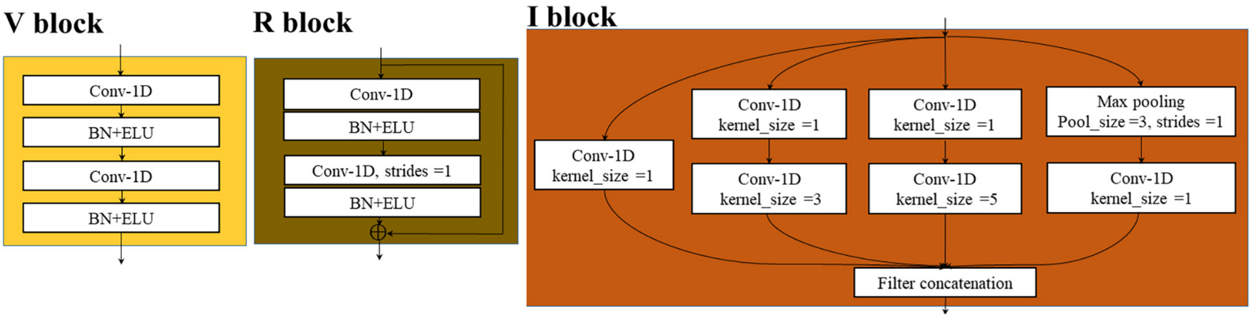


FIGURE 2 | The inner structures of three typical blocks.

the structure of the optimal deep model based on the source dataset was reused to simplify and shorten the modeling process. Since the number of seed varieties varies with different crops, the number of neurons in the output layer of the model was modified correspondingly. As the initial weights greatly influence the convergence speed and the final performance of the model,

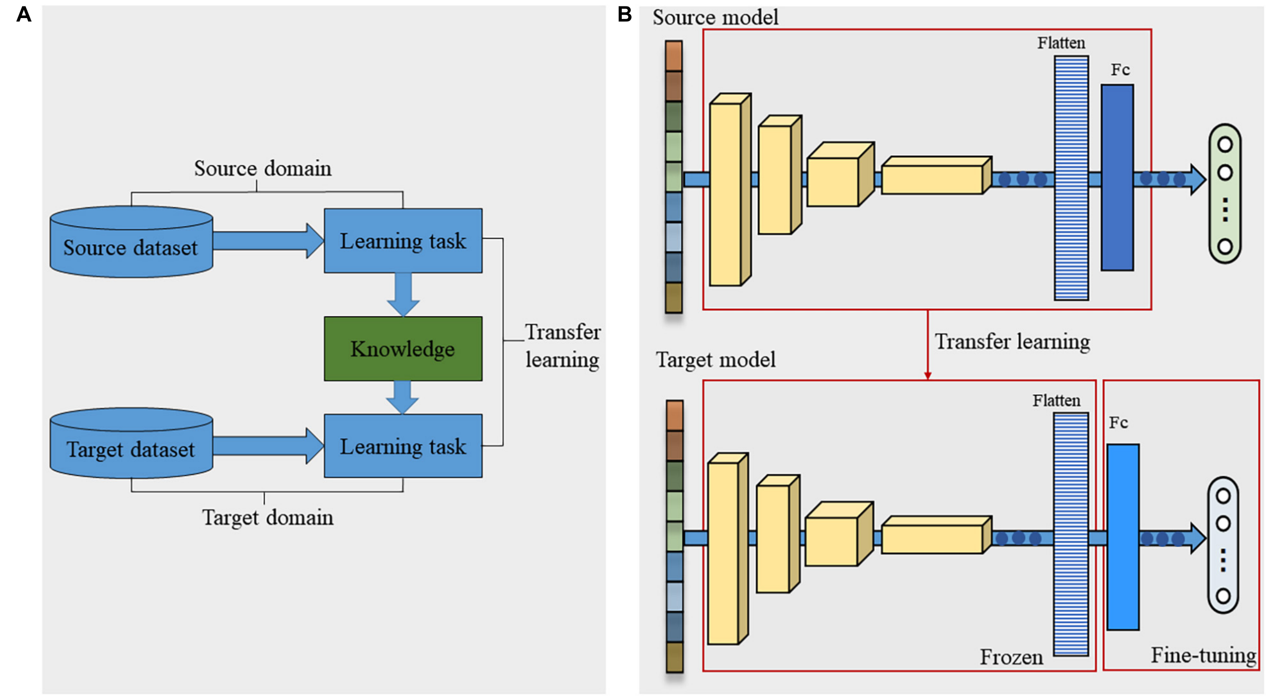


FIGURE 3 | Transfer learning strategy. **(A)** The typical process of transfer learning. **(B)** The deep transfer learning strategy in this study. The yellow cubes of different sizes represented multiple cascaded convolution layers. The striped bars and blue bars still represented the flattened layer and the fully connected layer.

this study transferred the weights of the optimal deep model based on the source dataset to the models based on the target datasets according to the network structure. Since the number of output neurons in the deep models based on the Rice, Wheat, and Cotton dataset differed from that in the source model, the weights of the last fully connected layer in these models needed to be randomly initialized.

During the transferring process, the weights of the layers before the flatten layer were frozen, and the target datasets were used to fine-tune the subsequent fully connected layers (**Figure 3B**). The first reason was that the target dataset was too small to retrain the entire network. The second reason was that the convolutional layers before the flatten layer might have extracted important feature patterns of the seed spectra, which could be reused in the target tasks. According to the size of the target datasets, the *batch_size* of the transferred network was set to 3, and the learning rate was set to 0.0001. The other configurations were the same as the source model.

Comparison Methods

In this study, the deep neural networks based on the source dataset and four target datasets were compared with conventional linear and non-linear multivariate analysis methods to confirm their validities in spectra analysis from both data-rich and data-poor sides.

LDA aims to find an optimal projected direction for raw variables. In the projected feature space, samples between classes hold maximal dispersion, while samples within classes hold minimal dispersion (Gerhardt et al., 2019). This projection manner facilitates transforming the samples into a linear separable state. The number of variables in the projected space, *n_lda*, is the only parameter that needs to be adjusted. We set *n_lda* to 1–20 and selected the optimal *n_lda* according to the classification performance of LDA.

The core principle of PLS-DA is also to conduct a linear transformation. Unlike LDA, the transformed latent variables (LVs) can carry the primary information hidden in the raw variables and maximize the correlation between the independent and the dependent variables (Kandpal et al., 2016). In spectral analysis, the number of LVs, *n_pls-da*, that minimize the sum of predicted residual error was usually selected. The range of *n_pls-da* was also set to 1∼20 in this study.

SVM can enable raw linear inseparable variables to transform into a linear separable space through a non-linear kernel function (Gerhardt et al., 2019). Radial basis function (RBF) kernel was often used with SVM in many spectral analysis tasks because of its ability to cluster samples with the same categories closely and make them linearly separable. In this study, SVM equipped with RBF kernel was introduced as a non-linear benchmark. Two parameters, penalty coefficient *c* and kernel parameter *g*, were set to {10, 100, 1,000, and 10,000} and {0.1, 0.01, 0.001, and 0.0001}, respectively.

MLP is a fully connected artificial neural network with one or more hidden layers (Taud and Mas, 2017). To obtain the optimal performance, a total of 32 structures were attempted to process the source dataset, which contained one to four hidden layers, and each was equipped with eight configurations for nodes in hidden layers. The number of nodes in hidden layers of the structure with four hidden layers was set to [(200-100-50-25), (180-90-45-23), (160-80-40-20), (140-70-35-18), (120-60-30-15), (100-50-25-13), (80-40-20-10), (60-30-15-8)], and was simplified as the number of the hidden layer decreases. For the target datasets, 24 same structures with one to three hidden layers were tried to get the optimal classifier.

In addition, to further verify the role of the shared features for effective transfer learning, a hyperspectral dataset in the remote sensing field, Indian-Pines[1] was introduced. It is a 145 × 145 × 224 cube, containing 10,249 effective pixels of 16 categories, whose size was similar to that of the Pea dataset. These pixels were also randomly divided into a training set, a validation set, and a testing set at a ratio of 3:1:1. The number of the bands for analysis was reduced to 200 by removing the bands absorbed by water. To eliminate the influence of factors, such as the deep model, the structure of the optimal deep model based on the source Pea dataset was used to train the Indian-Pines dataset and was recorded as Model 0. Then, Model 0 was transferred to the other four target datasets.

The parameters of all models in this study were adjusted toward the optimal states, using the corresponding validation set. All models were coded, using python language in Spyder 3.2.6 environment (Anaconda, Austin, TX, United States). The famous machine learning library, Sklearn[2], was introduced to implement the conventional models, and the popular deep learning framework, Keras, was employed to program deep models. A Win10 64-bit operating system with Inter (R) Core (TM) i5-8500 CPU and 8 GB RAM constituted the primary platform.

Model Visualization

Model visualization is significant for intuitively understanding the decision-making mechanism and clearly showing the computational result. In this study, visualization techniques were investigated from the perspective of the training process of the deep model and the classification results of the crop seeds. The raw seed spectra of different datasets and the feature representation of different layers in the optimal deep models based on the source Pea dataset and the deep model based on the Indian-Pines dataset were extracted. Their distributions were then expressed by t-distribution stochastic neighbor embedding (t-SNE). As an effective method for high-dimensional data visualization, t-SNE converts the similarity between sample points in high-dimensional space into Gaussian joint probability form and constructs a similar probability distribution in low-dimensional space (van der Maaten and Hinton, 2008). The ability to maintain the local structure of data is conducive to observing data patterns in low-dimensional space. Moreover, the advantage of HSI to obtain both spatial and spectral information was fully exploited in this study. The label of the sample predicted by the deep model was projected into the corresponding spatial position and represented by different colors to establish classification maps of crop seeds.

RESULTS AND DISCUSSION

Spectroscopic Analysis

The average spectra with the standard deviation of different varieties of seeds in five datasets were shown in **Figure 4**. Obviously, these spectral curves possessed similar fluctuation patterns and locations of peaks and valleys. The absorption bands at approximately 1,119.45 and 1,206.92 nm were caused by the second overtone of carbohydrates (C–H stretch) (Wu et al., 2019). The peak near 1,307.97 nm (in the range of 1,254 – 1,348 nm) was reported to be associated with the combinations of the first overtone of amide B (N–H stretch) and the fundamental vibrations of amide II and III (C–N stretch and N–H in-plane bend) (Daszykowski et al., 2008). The band at 1,469.95 nm (in the region of 1,410–1,502 nm) could be attributed to the first overtone of Amide A (N–H stretch), which might be the critical band for protein detection (Daszykowski et al., 2008; Ribeiro et al., 2011). The similar chemical components in different crop seeds led to the similarities between the spectral curves. This meant that certain shared features might be hidden in the spectral information of different crop seeds, which provided the possibility for effective transfer learning.

However, for different varieties of seeds of the same crop, some heterogeneities also existed between their spectral curves due to the content difference of chemical components. For example, the spectral curves of four varieties of pea seeds were naturally divided into two groups. Baiyan and Heiyan formed one group, while Changshouren and Zhewan 1 formed the other one. This trend was consistent with the classification results according to the edible manner resulted from the content difference of sugar and water. In addition, for the Rice dataset, it was because of the introduction of the japonica characteristic that the reflectance of variety Yongyou 9 was quite different from the other two varieties. Qiu Z. et al. (2018) also confirmed spectra differences existed between different varieties of rice seeds. Nie et al. (2019) found the optical characteristics of different varieties of hybrid okra and luffa seeds were very different. The metabolic analysis results showed that the content of components of different seeds varied greatly. The heterogeneity of the spectral features between different varieties laid the basis for using HSI to classify different varieties of crop seeds.

Classification Results on Source Dataset

The accuracies and the optimal parameters of all models on the training set and the testing set of the source dataset were summarized in **Table 2**. The overfitting phenomenon for all models was not serious due to the large-scale training set that might contain the spectral patterns in the testing set.

The accuracies of three deep models on the testing set were all above 99%, which was higher than most conventional methods. Owing to the convolution operation, the deep models could extract much discriminative information hidden in the raw redundant spectral data. Their performance superiorities were predictable. VGG-MODEL, with an accuracy of 99.57% on the testing set, was slightly conspicuous than the other two models. Generally speaking, the difficulty of improving

---

[1] http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

[2] https://scikit-learn.org/stable/

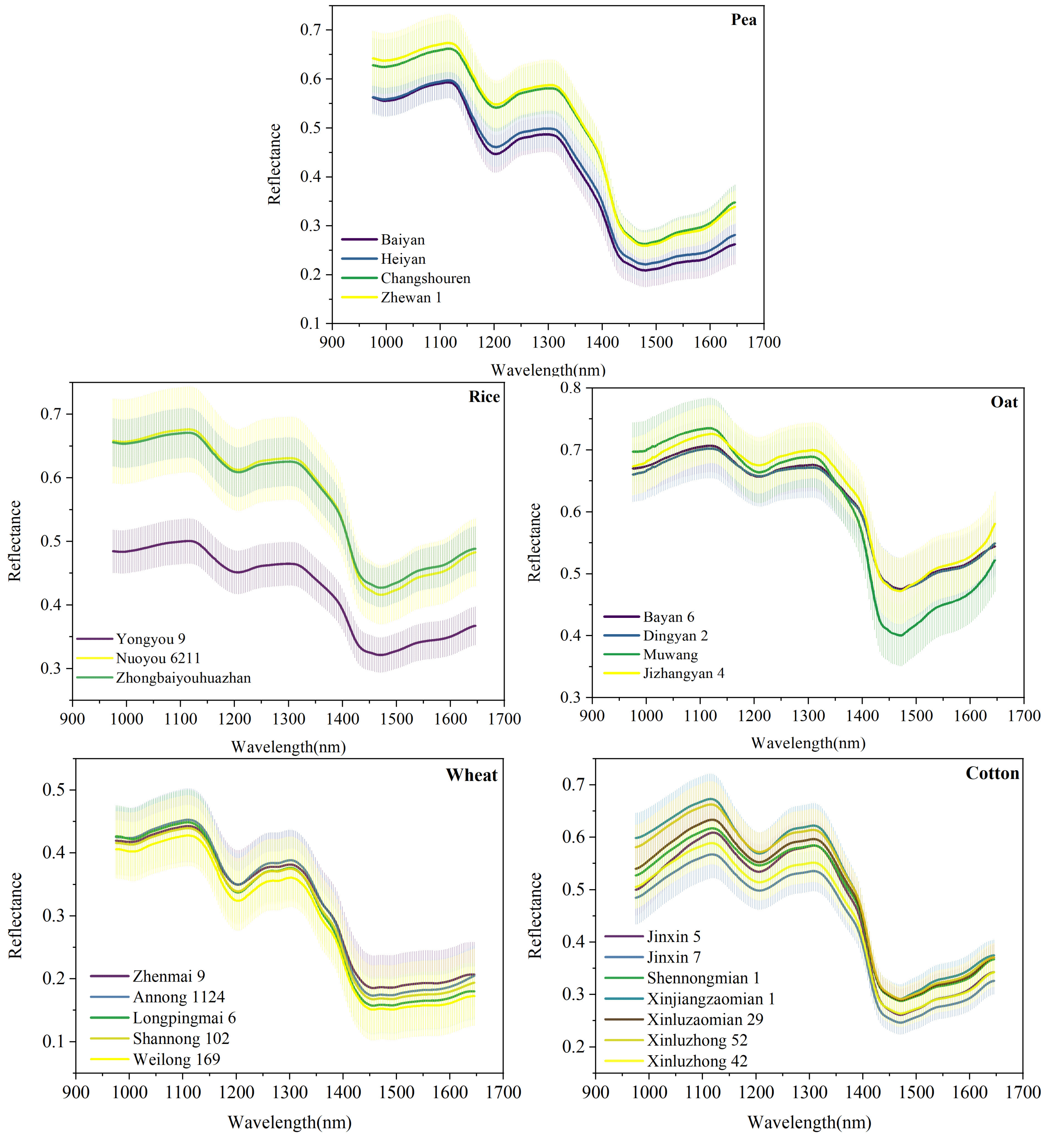


FIGURE 4 | The average spectra with the standard deviation of five crop seeds.

performance increases with the performance of the model being better. For example, in the 2014 ILSVRC competition, a 22-layer InceptionNet won the championship with a top-five error rate of 6.7% that was only 0.6% lower than the runner-up, VGGNet, with a 19-layer structure (Szegedy et al., 2015). In addition, the high version of Inception, Inception-v4, achieved a top-five error rate of 3.08% that was only 0.42% lower than the previous version, Inception V3 (Szegedy et al., 2016).

Since the structures of the three deep models were continuously adjusted to the optimal states according to the source dataset, they possessed different depths. In this study, INCEPTION-MODEL and RES-MODEL had a deeper structure than VGG-MODEL. In general, the deeper the model is, the richer the extracted features are. But this was based on an enough big dataset like ImageNet, and it should be guaranteed that the gradient would not disappear during model training.

TABLE 2 | The classification accuracies and optimal parameters of all the models on the source dataset.

Methods	Parameters[2]	Training (%)	Testing (%)
VGG-MODEL	(16, 32, 128, 201)	99.98	99.57
RES-MODEL	(32, 32, 64, 64, 128, 194)	99.76	99.14
INCEPTION-MODEL	(16, 32, 64, 128, 256, 349)	100	99.09
LDA	(1)	99.39	98.90
PLS-DA	(20)	87.14	86.90
SVM	$(10^4, 10^{-3})$	99.70	99.28
MLP	(200, 100, 50, 25)	93.81	93.52

Parameters[2] represents (number of major convolution filters, best epoch) for deep models, (n_lda) for LDA, (n_pls-da) for PLS-DA, (c, g) for SVM, (number of nodes in hidden layers) for MLP.

Zhang et al. (2019) developed a network with an inception structure that showed better performance than a comparison network, Model 3, with a similar structure to VGG-MODEL. However, the authors also pointed out that the superiority of the deep model was not in processing small datasets. In their study, Model 3 could not learn enough effective patterns from a few samples. The authors also indicated that the performance of Model 3 improved significantly when the size of the dataset increased slightly. The source dataset in this study was much larger than all the datasets in their study and was enough for VGG-MODEL training. For structures like ResNet, Zhu et al. (2019) compared the performance of a developed ResNet with a general deep convolutional neural network on a cotton dataset. Also, they found that ResNet was not as effective as the latter one.

The structure of the optimal model for a specific dataset was the result of a constant tradeoff and adjustment. It was closely related to the size and distribution of the sample set. A complex deep network could not always obtain higher performance than a simple one. In this study, for the source dataset, VGG-MODEL with the simplest structure and the shallowest depth won a small victory when faced with the relatively complex INCEPTION-MODEL and RES-MODEL. For conventional models, the accuracies of different methods on the testing set varied greatly. SVM performed best, followed by LDA. Thus, if we use traditional multivariate analysis methods, many models need to be tried and compared to determine the optimal one (Liu et al., 2017; Bao et al., 2019; Nie et al., 2019). Conversely, deep models will generally achieve satisfactory results if the training data are sufficient and the structure is designed reasonably. In the field of spectral analysis, deep learning is a very competitive and potential tool.

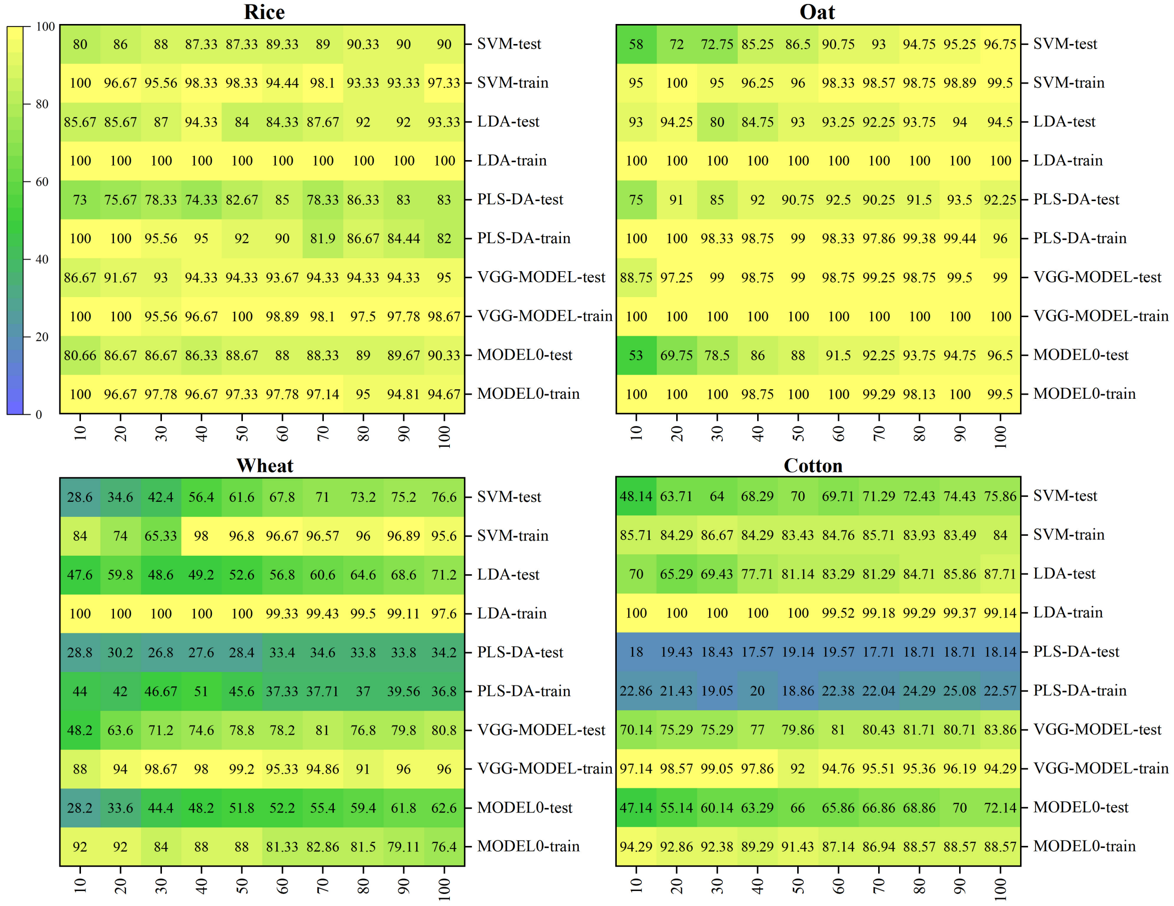
Classification Results on Target Datasets

Although the deep network might not perform well on a small dataset, its advantages would carry forward again after combining with transfer learning. To verify the effect of deep transfer learning, the slightly better-performing VGG-MODEL was used as the source model to be transferred in this study. Ten training sets with different sizes were built based on the original training sets to investigate the influence of training set size on the transferring effect. The classification results of the deep transferred model and the comparison methods were shown in **Figure 5**.

It could be seen that the deep transferred model was the only model that consistently performed well on the four datasets. For the 100% training set that was still very small compared with the training set of the source dataset, the deep model achieved accuracies of 95, 99, 80.8, and 83.86% on the testing sets of Rice, Oat, Wheat, and Cotton datasets, respectively. It was because of combination with transfer learning that the deep learning model could also obtain satisfactory results on these datasets. Transfer learning enabled deep learning to take advantage of itself and avoided the requirement for a mass of samples (Tan et al., 2018). As similar patterns existed among the spectra vectors of different crop seeds, varieties classification of different crop seeds belonged to different but similar tasks in the same domain. Thus, transfer learning was very suitable for varieties classification of different crop seeds in this study.

However, if the difference between the target dataset and the original dataset was too large, it might cause a negative transfer. In this study, Model 0, the deep model based on the Indian-Pines dataset, achieved accuracies of 98.31 and 90.05% on the training set and the testing set, respectively. But its performance was worse than the deep transferred model based on the source Pea dataset and most conventional multivariate analysis methods when transferred to the four target datasets (shown in **Figure 5**). This poor performance could be expected since the Indian-Pines dataset and the seed datasets in this study were quite different in sampling scenarios, spectral resolution, and spectral modes. This result illustrated the importance of the similarity between the features of the source dataset and the target dataset for effective transfer learning in this study. When there is a vast difference between these two datasets, the direct transfer may lead to undesirable results. More effective transfer learning methods need to be studied in the future.

For other conventional models, although they could also achieve good performance on some datasets, they could not always perform well on all. For example, for the 100% training set, LDA achieved an accuracy of 87.71% on the Cotton dataset, which was even higher than that achieved by the deep transferred model. However, it just obtained accuracies of 93.33, 94.5, and 71.2% on the Rice, Oat, and Wheat datasets. SVM performed relatively stable, just like previous research (Qiu Z. et al., 2018; Bao et al., 2019; Nie et al., 2019). It achieved accuracies of 90, 96.75, 76.6, and 75.86% on the four datasets when using the 100% training sets. As expected, MLP performed much worse than the deep neural network. For the Wheat dataset, it only got an accuracy of 32.4% on the testing set, which was just a little better than random guessing. The shallow neural network could not extract valuable discriminative information from redundant spectral data, which led to unsatisfactory results (Chen et al., 2014). PLS-DA, commonly used in spectral analysis, was also very unstable. Although it could obtain an accuracy of 92.25% on the Oat dataset, it performed rather severely on the Wheat and Cotton datasets, which contained more varieties. Its performance was consistent with the results of Nie et al. (2019). With the increase of the number of seed varieties, the possibility of samples being linearly separable became smaller, and the difficulty of



Rice

	10	20	30	40	50	60	70	80	90	100
SVM-test	80	86	88	87.33	87.33	89.33	89	90.33	90	90
SVM-train	100	96.67	95.56	98.33	98.33	94.44	98.1	93.33	93.33	97.33
LDA-test	85.67	85.67	87	94.33	84	84.33	87.67	92	92	93.33
LDA-train	100	100	100	100	100	100	100	100	100	100
PLS-DA-test	73	75.67	78.33	74.33	82.67	85	78.33	86.33	83	83
PLS-DA-train	100	100	95.56	95	92	90	81.9	86.67	84.44	82
VGG-MODEL-test	86.67	91.67	93	94.33	94.33	93.67	94.33	94.33	94.33	95
VGG-MODEL-train	100	100	95.56	96.67	100	98.89	98.1	97.5	97.78	98.67
MODEL0-test	80.66	86.67	86.67	86.33	88.67	88	88.33	89	89.67	90.33
MODEL0-train	100	96.67	97.78	96.67	97.33	97.78	97.14	95	94.81	94.67

Oat

	10	20	30	40	50	60	70	80	90	100
SVM-test	58	72	72.75	85.25	86.5	90.75	93	94.75	95.25	96.75
SVM-train	95	100	95	96.25	96	98.33	98.57	98.75	98.89	99.5
LDA-test	93	94.25	80	84.75	93	93.25	92.25	93.75	94	94.5
LDA-train	100	100	100	100	100	100	100	100	100	100
PLS-DA-test	75	91	85	92	90.75	92.5	90.25	91.5	93.5	92.25
PLS-DA-train	100	100	98.33	98.75	99	98.33	97.86	99.38	99.44	96
VGG-MODEL-test	88.75	97.25	99	98.75	99	98.75	99.25	98.75	99.5	99
VGG-MODEL-train	100	100	100	100	100	100	100	100	100	100
MODEL0-test	53	69.75	78.5	86	88	91.5	92.25	93.75	94.75	96.5
MODEL0-train	100	100	100	98.75	100	100	99.29	98.13	100	99.5

Wheat

	10	20	30	40	50	60	70	80	90	100
SVM-test	28.6	34.6	42.4	56.4	61.6	67.8	71	73.2	75.2	76.6
SVM-train	84	74	65.33	98	96.8	96.67	96.57	96	96.89	95.6
LDA-test	47.6	59.8	48.6	49.2	52.6	56.8	60.6	64.6	68.6	71.2
LDA-train	100	100	100	100	100	99.33	99.43	99.5	99.11	97.6
PLS-DA-test	28.8	30.2	26.8	27.6	28.4	33.4	34.6	33.8	33.8	34.2
PLS-DA-train	44	42	46.67	51	45.6	37.33	37.71	37	39.56	36.8
VGG-MODEL-test	48.2	63.6	71.2	74.6	78.8	78.2	81	76.8	79.8	80.8
VGG-MODEL-train	88	94	98.67	98	99.2	95.33	94.86	91	96	96
MODEL0-test	28.2	33.6	44.4	48.2	51.8	52.2	55.4	59.4	61.8	62.6
MODEL0-train	92	92	84	88	88	81.33	82.86	81.5	79.11	76.4

Cotton

	10	20	30	40	50	60	70	80	90	100
SVM-test	48.14	63.71	64	68.29	70	69.71	71.29	72.43	74.43	75.86
SVM-train	85.71	84.29	86.67	84.29	83.43	84.76	85.71	83.93	83.49	84
LDA-test	70	65.29	69.43	77.71	81.14	83.29	81.29	84.71	85.86	87.71
LDA-train	100	100	100	100	100	99.52	99.18	99.29	99.37	99.14
PLS-DA-test	18	19.43	18.43	17.57	19.14	19.57	17.71	18.71	18.71	18.14
PLS-DA-train	22.86	21.43	19.05	20	18.86	22.38	22.04	24.29	25.08	22.57
VGG-MODEL-test	70.14	75.29	75.29	77	79.86	81	80.43	81.71	80.71	83.86
VGG-MODEL-train	97.14	98.57	99.05	97.86	92	94.76	95.51	95.36	96.19	94.29
MODEL0-test	47.14	55.14	60.14	63.29	66	65.86	66.86	68.86	70	72.14
MODEL0-train	94.29	92.86	92.38	89.29	91.43	87.14	86.94	88.57	88.57	88.57

FIGURE 5 | The classification accuracies of all the models on the four target datasets.

distinguishing different varieties became greater. In a word, for traditional multivariate analysis methods, different datasets might correspond to different optimal models. Conversely, the deep transferred models based on the source Pea dataset could generally achieve satisfactory results.

In addition, it was worth mentioning that the deep transferred model based on the source dataset could also achieve good results when fine-tuned, using tiny datasets. For example, when using the 10% training set, which only contained five samples for each variety, it could achieve accuracies of 86.67, 88.74, and 70.14% on the Rice, Oat, and Co tton datasets. And the accuracy rose rapidly with the increase of training set size. Even on the Wheat dataset, where all models failed, the deep transferred model outperformed all the conventional methods. Deep transfer learning brings hope for scenarios with very limited samples. On the contrary, the accuracies of most conventional methods were very low when trained, using such a small dataset. The result that LDA got a high classification accuracy of 93% on the Oat dataset was unexpected. The reason might be that this small training set just fitted the classification rule of LDA because its accuracy dropped to 80% soon for the 30% training set and then slowly increased.

Moreover, it could be observed that almost all the models showed high accuracies on the Rice and Oat datasets but performed poorly on the Wheat and Cotton datasets. The sample distribution of a dataset with few categories was generally simple. Contrarily, the distribution of a dataset with more categories was relatively complicated, which was not conducive to classification. Thus, the dataset was an essential factor affecting the performance of models (Özdemir et al., 2019; Zhang et al., 2019). In addition, it could be seen that the deep transferred model based on the source Pea dataset got the best performance on the Oat dataset. Using the 20% training set, it obtained an accuracy of 97.25% on the testing set. Since the Oat dataset had the same number of varieties as the source dataset, all the weight parameters in the source model, including the weights in the last fully connected layers, could be transferred. This specialness allowed the maximum transferring of features in the source model.

Model Visualization

Visualizing the feature distribution at each layer of the deep network was an important channel to understand the training process of the deep model (Lin and Maji, 2016; Zintgraf et al., 2017; Zhang and Zhu, 2018). In this study, the t-SNE technique was used to visualize the original high-dimensional spectra and the features output by the flatten, Fc1, and Fc2 layers of the deep model in a two-dimensional space, as shown in **Figure 6** and **Supplementary Figures 1–8**. For the Pea dataset, the raw spectral samples were aggregated into two categories, consistent with the average spectral analysis. After passing the flatten layer, the spectra with easily confused categories like Baiyan and Heiyan, or Changshouren and Zhewan 1, gradually became distinguishable. As the layers deepened, the samples within a category were clustered closely, while those between different categories became discrete. The samples were clearly gathered into four categories after output by the Fc2 layer. It could be seen that the deep model gradually transformed the samples from a cluttered state to a distinguishable state. It was why the deep model could obtain better performance than the traditional methods.

For the four target datasets, the raw spectra in the Rice and Oat datasets, especially in the Rice dataset, were slightly more regular than those in the Wheat and Cotton datasets. The variety Yongyou9 was strongly distinguishable from the other two varieties. This phenomenon was also consistent with previous spectral analysis. Thus, most traditional methods performed better on the Rice and Oat datasets than on the other two datasets. Since all the weights before the flatten layer were transferred from the deep model based on the source Pea dataset or the Indian-Pines dataset directly, the features output by the flatten layer of the deep transferred model contained the spectral patterns learned from these two datasets. From **Supplementary Figures 1–4**, it could be seen that, for the Rice, Oat, and Cotton datasets, the features output by the flatten layer presented a more aggregated distribution pattern than the raw spectral samples. In **Supplementary Figures 5–8**, however, the distribution patterns of the features output by the flatten layer were not significantly improved compared with the raw spectral samples. These results intuitively illustrated the critical role of the shared features for transfer learning. Effective transfer learning was conductive to the classification of different varieties of seeds in this study. The spectral features learned from the source Pea dataset were reused through transferring, facilitating the classification of small target datasets. The Wheat dataset might be too cluttered so that the output of the flatten layer did not show distinguishability. The

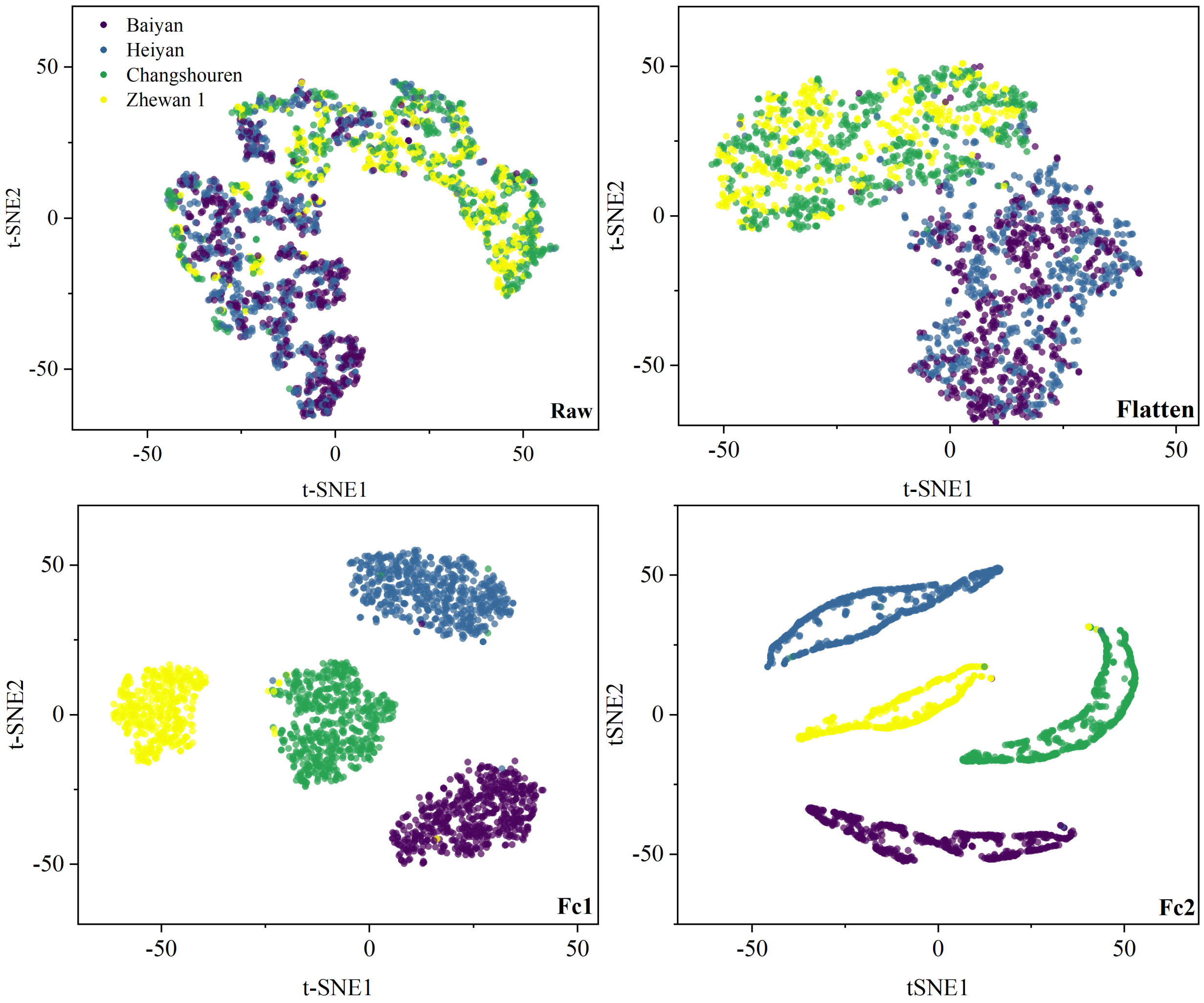


FIGURE 6 | Feature visualization of VGG-MODEL on the Pea dataset, using t-SNE.

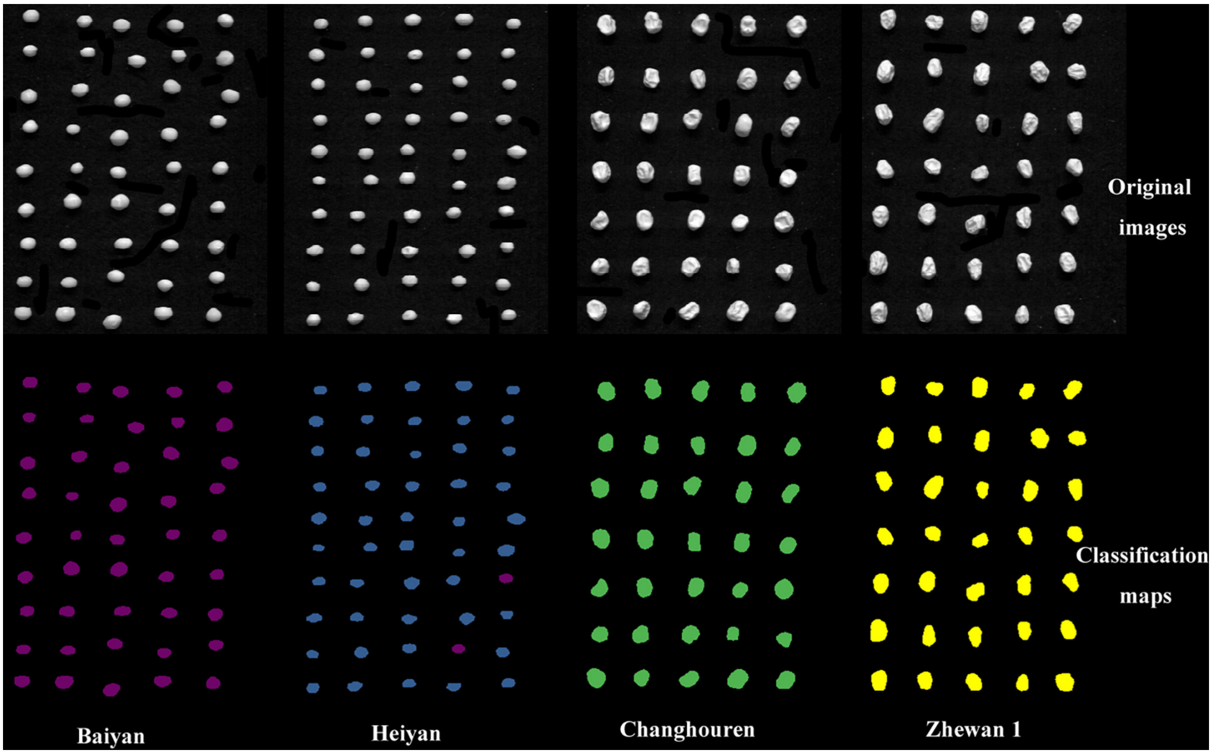


FIGURE 7 | The classification visual maps of pea seeds.

target datasets began to work from the Fc1 layer. The samples gradually showed strong separability with the layers deepened. After output by the final Fc2 layer, the rice and oat samples had been divided into three and four categories, respectively. Thus, the deep transferred model achieved two high accuracies of 95 and 99%. However, the wheat and the cotton samples still had some overlapping phenomenon, which led to relatively low accuracies of the deep transferred model. Since there were no effective features transferred, the features output by the Fc1 and Fc2 layer in **Supplementary Figures 5–8** showed a more discrete distribution pattern than those in **Supplementary Figures 1–4**. This was why the classification performance of the transfer model based on the Indian-Pines dataset was worse than that based on the source Pea dataset.

The classification visualization of crop seeds was helpful for breeders to select varieties that meet requirements and for market supervision authorities to check seed purity. In this study, the categories of pea seeds classified by the optimal model, VGG-MODEL, were visualized in a map. As shown in **Figure 7**, Baiyan and Heiyan showed similar smooth texture features in the original hyperspectral images. In contrast, Changshouren and Zhewan 1 showed rough texture due to water loss during the drying process. According to human vision, these four varieties were naturally divided into two categories, consistent with the visualization analysis of the distribution of the samples. Among the predicted 180 seeds, only two seeds of the variety Heiyan were misclassified into the similar Baiyan category. This accuracy was sufficient for variety selection during the breeding process or purity detection in actual production. The characteristics of batch detection of HSI combined with the capabilities of rapid analysis of deep transfer learning may provide a brand-new solution for identifying crop varieties under sample-limited condition. It is expected to help accelerate the process of crop variety screening.

CONCLUSION

This study attempted to use HSI and deep transfer learning to achieve accurate and rapid varieties classification of crop seeds under sample-limited condition. The VGG-MODEL based on the sample-rich dataset stood out from three deep neural networks with typical structures and was utilized as the deep source model to be transferred. The transfer results on the four small target datasets showed that the deep transferred model could fully use the shared spectral features of crop seeds extracted by the source deep model. The deep transferred model could achieve better performance than traditional multivariate analysis methods under sample-limited condition, especially when using tiny samples. Giving a glimpse into the process of deep transfer learning, the visualization of the feature distribution at each layer of the deep network further confirmed the portability of shared spectral features. It revealed why the deep network achieved high accuracy. The visualization of classification results provided an intuitive and convenient manner for varieties classification of crop seeds. In conclusion, HSI combined with deep transfer learning, was a great potential tool for the classification of seed varieties with limited samples, which will significantly accelerate the seed screening process in fields with scarce samples. This study also provided a new idea for detecting other qualities of crop seeds based on HSI under sample-limited condition.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

NW, FL, and CZ conceived the research concept. NW and CZ performed the experiments. NW wrote the manuscript. FL, FM, ML, and YH contributed to the results analysis and discussion. YH provided financial support. All the authors contributed to the study and approved the submitted version.

FUNDING

This work was supported by the National Key R&D Program of China (2018YFD0101002) and the National Natural Science Foundation (6196030055).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.696292/full#supplementary-material

REFERENCES

Bao, Y., Mi, C., Wu, N., Liu, F., and He, Y. (2019). Rapid classification of wheat grain varieties using hyperspectral imaging and chemometrics. *Appl. Sci.* 9:4119. doi: 10.3390/app9194119

Beníteza, V., Canteraa, S., Aguileraa, Y., Molláa, E., Estebana, R. M., Díazb, M. F., et al. (2013). Impact of germination on starch, dietary fiber and physicochemical properties in non-conventional legumes. *Food Res. Int.* 50, 64–69. doi: 10.1016/j.foodres.2012.09.044

Caporaso, N., Whitworth, M. B., and Fisk, I. D. (2018). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chem.* 240, 32–42. doi: 10.1016/j.foodchem.2017.07.048

Chen, Y., Lin, Z., Zhao, X., Wang, G., and Gu, Y. (2014). Deep learning-based classification of hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 2094–2107. doi: 10.1109/JSTARS.2014.2329330

Daszykowski, M., Wrobel, M. S., Czarnik-Matusewicz, H., and Walczak, B. (2008). Near-infrared reflectance spectroscopy and multivariate calibration techniques applied to modelling coffee beverages using near infrared spectroscopy the crude protein, fibre and fat content in rapeseed meal. *Analyst* 133, 1523–1531. doi: 10.1039/B803687J

Du, L., Xia, X., Lan, X., Liu, M., Zhao, L., Zhang, P., et al. (2017). Influence of arsenic stress on physiological, biochemical, and morphological characteristics in seedlings of two cultivars of maize (*Zea mays* L.). *Water Air Soil Pollut.* 228:55. doi: 10.1007/s11270-016-3231-2

Feng, X., Peng, C., Chen, Y., Liu, X., Feng, X., and He, Y. (2017). Discrimination of CRISPR/Cas9-induced mutants of rice seeds using near-infrared hyperspectral imaging. *Sci. Rep.* 7:15934. doi: 10.1038/s41598-017-16254-z

Gerhardt, N., Schwolow, S., Rohn, S., Pérez-Cacho, P. R., Galán-Soldevilla, H., Arce, L., et al. (2019). Quality assessment of olive oils based on temperature-ramped HS-GC-IMS and sensory evaluation: comparison of different processing approaches by LDA, kNN, and SVM. *Food Chem.* 278, 720–728. doi: 10.1016/j.foodchem.2018.11.095

Ghazi, M. M., Yanikoglu, B., and Aptoula, E. (2017). Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing* 235, 228–235. doi: 10.1016/j.neucom.2017.01.018

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ), 770–778.

Jin, X., Jie, L., Wang, S., Qi, H. J., and Li, S. W. (2018). Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field. *Remote Sens.* 10:395. doi: 10.3390/rs10030395

Kandpal, L. M., Lohumi, S., Kim, M. S., Kang, J. S., and Cho, B. K. (2016). Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds. *Sens. Actuators B Chem.* 229, 534–544. doi: 10.1016/j.snb.2016.02.015

Kong, W., Zhang, C., Huang, W., Liu, F., and He, Y. (2018). Application of hyperspectral imaging to detect *Sclerotinia sclerotiorum* on oilseed rape stems. *Sensors* 18:123. doi: 10.3390/s18010123

Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 84–90. doi: 10.1145/3065386

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Lee, H., Kim, M. S., Lim, H., Park, E., Lee, W. H., and Cho, B. K. (2016). Detection of cucumber green mottle mosaic virus-infected watermelon seeds using a near-infrared (NIR) hyperspectral imaging system: application to seeds of the "Sambok Honey" cultivar. *Biosyst. Eng.* 148, 138–147. doi: 10.1016/j.biosystemseng.2016.05.014

Lin, T. Y., and Maji, S. (2016). "Visualizing and understanding deep texture representations," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ: *IEEE*), 2791–2799.

Liu, D., Li, Q., Li, W., Yang, B., and Guo, W. (2017). Discriminating forchlorfenuron-treated kiwifruits using a portable spectrometer and Vis/NIR diffuse transmittance spectroscopy technology. *Anal. Methods* 9, 4207–4214. doi: 10.1039/c7ay00832e

Liu, L., Ji, M., and Buchroithner, M. (2018). Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. *Sensors* 18:3169. doi: 10.3390/s18093169

Mohanty, S. P., Hughes, D. P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 7:1419. doi: 10.3389/fpls.2016.01419

Nie, P., Zhang, J., Feng, X., Yu, C., and He, Y. (2019). Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning. *Sens. Actuators B Chem.* 296:126630. doi: 10.1016/j.snb.2019.126630

Özdemir, A., Yavuz, U., and Dael, F. A. (2019). Performance evaluation of different classification techniques using different datasets. *Int. J. Elec. Comp. Eng.* 9, 3584–3590. doi: 10.11591/ijece.v9i5.pp3584-3590

Pan, S. J., and Yang, Q. (2010). A survey of transfer learning. *IEEE Trans. Knowl. Data Eng.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191

Qiu, Z., Chen, J., Zhao, Y., Zhu, S., He, Y., and Zhang, C. (2018). Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. *Appl. Sci.* 8:212. doi: 10.3390/app8020212

Qiu, R., Wei, S., Zhang, M., Li, H., Sun, H., Liu, G., et al. (2018). Sensors for measuring plant phenotyping: a review. *Int. J. Agric. Biol. Eng.* 11, 1–17. doi: 10.25165/j.ijabe.20181102.2696

Rashid, M. A., and Singh, D. P. (2000). *A Manual on Vegetable Seed Production in Bangladesh*. Dhaka: Karshaf Printers (Pvt.) Ltd, 10.

Ribeiro, J. S., Ferreira, M. M. C., and Salva, T. J. G. (2011). Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta* 83, 1352–1358. doi: 10.1016/j.talanta.2010.11.001

Sendin, K., Williams, P. J., and Manley, M. (2018). Near infrared hyperspectral imaging in quality and safety evaluation of cereals. *Crit. Rev. Food Sci.* 58, 575–590. doi: 10.1080/10408398.2016.1205548

Shrestha, S., Knapič, M., Žibrat, U., Deleuran, L. C., and Gislum, R. (2016). Single seed near-infrared hyperspectral imaging in determining tomato (*Solanum lycopersicum* L.) seed quality in association with multivariate data analysis. *Sens. Actuators B Chem.* 237, 1027–1034. doi: 10.1016/j.snb.2016.08.170

Shuaib, M., Zeb, A., Ali, Z., Ali, W., Ahmad, T., and Khan, I. (2007). Characterization of wheat varieties by seed storage-protein electrophoresis. *Afr. J. Biotechnol.* 6, 497–500. doi: 10.4314/ajb.v6i5.56863

Simonyan, K., and Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv* [Preprint]. Available online at: https://arxiv.org/abs/1409.1556 (accessed April 10, 2015).

Stuart, M. B., McGonigle, A. J. S., and Willmott, J. R. (2019). Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems. *Sensors* 19:3071. doi: 10.3390/s19143071

Sun, D., Cen, H., Weng, H., Wan, L., Abdalla, A., El-Manawy, A. I, et al. (2019). Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. *Plant Methods* 15:54. doi: 10.1186/s13007-019-0432-x

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Piscataway, NJ: IEEE), 1–9.

Szegedy, C., Loffe, S., Vanhoucke, V., and Alemi, A. (2016). *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv* [Preprint]. Available online at: https://arxiv.org/abs/1602.07261 (accessed August 23, 2016).

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. (2018). "A survey on deep transfer learning," in *Proceedings of the International Conference on Artificial Neural Networks* (Berlin: Springer), 270–279.

Taud, H., and Mas, J. F. (2017). "Multilayer perceptron (MLP)," in *Geomatic Approaches for Modeling Land Change Scenarios*, eds M. Camacho Olmedo, M. Paegelow, J. F. Mas, and F. Escobar (Cham: Springer), 451–455.

van der Maaten, L., and Hinton, G. (2008). Viualizing data using t-SNE. *J. Mach. Learn Res.* 9, 2579–2605.

Wang, A. X., Tran, C., Desai, N., Lobell, D., and Ermon, S. (2018). "Deep transfer learning for crop yield prediction with remote sensing data," in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies* (New York, NY), 1–5.

Wei, W., Zhang, J., Zhang, L., Tian, C., and Zhang, Y. (2018). Deep cube-pair network for hyperspectral imagery classification. *Remote Sens.* 10:783. doi: 10.3390/rs10050783

Weiss, K., Khoshgoftaar, T. M., and Wang, D. D. (2016). A survey of transfer learning. *J. Big Data* 3:9. doi: 10.1186/s40537-016-0043-6

Weng, H., Lv, J., Cen, H., He, M., Zeng, Y., Hua, S., et al. (2018). Hyperspectral reflectance imaging combined with carbohydrate metabolism analysis for diagnosis of citrus Huanglongbing in different seasons and cultivars. *Sens. Actuators B Chem.* 275, 50–60. doi: 10.1016/j.snb.2018.08.020

Wu, N., Jiang, H., Bao, Y., Zhang, C., Zhang, J., Song, W., et al. (2020). Practicability investigation of using near-infrared hyperspectral imaging to detect rice kernels infected with rice false smut in different conditions. *Sens. Actuators B Chem.* 308:127696. doi: 10.1016/j.snb.2020.127696

Wu, N., Zhang, C., Bai, X., Du, X., and He, Y. (2018). Discrimination of chrysanthemum varieties using hyperspectral imaging combined with a deep convolutional neural network. *Molecules* 23:2381. doi: 10.3390/molecules23112831

Wu, N., Zhang, Y., Na, R., Mi, C., Zhu, S., He, Y., et al. (2019). Variety identification of oat seeds using hyperspectral imaging: investigating the representation ability of deep convolutional neural network. *RSC Adv.* 9, 12635–12644. doi: 10.1039/c8ra10335f

Xie, M., Jean, N., Burke, M., Lobell, D., and Ermon, S. (2016). "Transfer learning from deep features for remote sensing and poverty mapping," in *Proceedings of the 30th AAAI Conference on Artificial Intelligence* (Phoenix, AZ, United States: AAAI), 3929–3935.

Xu, J., Riccioli, C., and Sun, D. (2017). Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. *J. Food Eng.* 196, 170–182. doi: 10.1016/j.jfoodeng.2016.10.021

Ye, S., Wang, Y., Huang, D., Li, J., Gong, Y., Xu, L., et al. (2013). Genetic purity testing of F1 hybrid seed with molecular markers in cabbage (*Brassica oleracea var. capitata*). *Sci. Hortic.* 155, 92–96. doi: 10.1016/j.scienta.2013.03.016

Yu, X., Lu, H., and Wu, D. (2018). Development of deep learning method for predicting firmness and soluble solid content of postharvest Korla fragrant pear using Vis/NIR hyperspectral reflectance imaging. *Postharvest Biol. Technol.* 141, 39–49. doi: 10.1016/j.postharvbio.2018.02.013

Yuan, Y., Zheng, X., and Lu, X. (2017). Hyperspectral image superresolution by transfer learning. *IEEE J. Sel. Top Appl. Earth Observ.* 10, 1963–1974. doi: 10.1109/JSTARS.2017.2655112

Zhang, J., Yang, Y., Feng, X., Xu, H., Chen, J., and He, Y. (2020). Identification of bacterial blight resistant rice seeds using terahertz imaging and hyperspectral imaging combined with convolutional neural network. *Front. Plant Sci.* 11:821. doi: 10.3389/fpls.2020.00821

Zhang, Q. S., and Zhu, S. C. (2018). Visual interpretability for deep learning: a survey. *Front. Inform. Technol. Electron. Eng.* 19:27–39. doi: 10.1631/FITEE.1700808

Zhang, X., Lin, T., Xu, J., Luo, X., and Ying, Y. (2019). DeepSpectra: an end-to-end deep learning approach for quantitative spectral analysis. *Anal. Chim. Acta* 1058, 48–57. doi: 10.1016/j.aca.2019.01.002

Zhao, M., Zhang, H., Yan, H., Qiu, L., and Baskin, C. C. (2018). Mobilization and role of starch, protein, and fat reserves during seed germination of six wild grassland species. *Front. Plant Sci.* 9:234. doi: 10.3389/fpls.2018.00234

Zhu, S., Zhou, L., Zhang, C., Bao, Y., Wu, B., and Chu, H. (2019). Identification of soybean varieties using hyperspectral imaging coupled with convolutional neural network. *Sensors* 19:4065. doi: 10.3390/S19194065

Zintgraf, L. M., Cohen, T. S., Adel, T., and Welling, M. (2017). *Visualizing Deep Neural Networks Decisions: Prediction Difference Analysis. arXiv* [Preprint]. Available online at: https://arxiv.org/abs/1702.04595 (accessed February 15, 2017).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wu, Liu, Meng, Li, Zhang and He. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 19 August 2021
doi: 10.3389/fpls.2021.693854

Measurement of Environmentally Influenced Variations in Anthocyanin Accumulations in *Brassica rapa* subsp. *Chinensis* (Bok Choy) Using Hyperspectral Imaging

Hyo-suk Kim [1,2†], ***Ji Hye Yoo*** [3†], ***Soo Hyun Park*** [3], ***Jun-Sik Kim*** [4], ***Youngchul Chung*** [2], ***Jae Hun Kim*** [1*] ***and Hyoung Seok Kim*** [3*]

[1] Sensor System Research Center, Korea Institute of Science and Technology (KIST), Seoul, South Korea, [2] Department of Electronics and Communications Engineering, Kwangwoon University, Seoul, South Korea, [3] Smart Farm Convergence Research Center, Korea Institute of Science and Technology (KIST), Gangneung, South Korea, [4] Center for Intelligent and Interactive Robotics, Korea Institute of Science and Technology (KIST), Seoul, South Korea

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Patompong Johns Saengwilai,
Mahidol University, Thailand
Alexandr Muterko,
Russian Academy of Sciences, Russia

***Correspondence:**
Hyoung Seok Kim
hkim58@kist.re.kr
Jae Hun Kim
jaekim@kist.re.kr

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *12 April 2021*
Accepted: *22 July 2021*
Published: *19 August 2021*

Citation:
Kim H-s, Yoo JH, Park SH, Kim J-S, Chung Y, Kim JH and Kim HS (2021) Measurement of Environmentally Influenced Variations in Anthocyanin Accumulations in Brassica rapa subsp. Chinensis (Bok Choy) Using Hyperspectral Imaging.
Front. Plant Sci. 12:693854.
doi: 10.3389/fpls.2021.693854

Dietary supplements of anthocyanin-rich vegetables have been known to increase potential health benefits for humans. The optimization of environmental conditions to increase the level of anthocyanin accumulations in vegetables during the cultivation periods is particularly important in terms of the improvement of agricultural values in the indoor farm using artificial light and climate controlling systems. This study reports on the measurement of variations in anthocyanin accumulations in leaf tissues of four different cultivars in *Brassica rapa* var. *chinensis* (bok choy) grown under the different environmental conditions of the indoor farm using hyperspectral imaging. Anthocyanin accumulations estimated by hyperspectral imaging were compared with the measured anthocyanin accumulation obtained by destructive analysis. Between hyperspectral imaging and destructive analysis values, no significant differences in anthocyanin accumulation were observed across four bok choy cultivars grown under the anthocyanin stimulation environmental condition, whereas the estimated anthocyanin accumulations displayed cultivar-dependent significant differences, suggesting that hyperspectral imaging can be employed to measure variations in anthocyanin accumulations of different bok choy cultivars. Increased accumulation of anthocyanin under the stimulation condition for anthocyanin accumulation was observed in "purple magic" and "red stem" by both hyperspectral imaging and destructive analysis. In the different growth stages, no significant differences in anthocyanin accumulation were found in each cultivar by both hyperspectral imaging and destructive analysis. These results suggest that hyperspectral imaging can provide comparable analytic capability with destructive analysis to measure variations in anthocyanin accumulation that occurred under the different light and temperature conditions of the indoor farm. Leaf image analysis measuring the percentage of purple color area in the total leaf area displayed successful classification of anthocyanin accumulation in four bok choy cultivars in comparison to hyperspectral imaging and destructive analysis, but it also showed limitation to reflect the level of color saturation caused by anthocyanin accumulation under different environmental conditions in "red stem," "white stem," and "green stem." Finally, our hyperspectral imaging system was modified to be applied onto the high-throughput plant phenotyping system, and its test to analyze the variation of anthocyanin accumulation in four cultivars showed comparable results with the result of the destructive analysis.

Keywords: bok choy, health promotion, non-destructive, image analysis, indoor farming

INTRODUCTION

Anthocyanin, a group of water-soluble flavonoid pigments derived from the phenylpropanoid pathway, is responsible for the color of different plant tissues, such as flower, fruit, and leaf ranging from red to violet and blue (Strack and Wray, 1989). Recent evidence suggests that the dietary supplements of anthocyanin-rich vegetables are closely associated with the reduced risk of cardiovascular disease and cancer (Williams et al., 2004; Butelli et al., 2008; de Pascual-Teresa et al., 2010). Because of the eye-catching color and potential human health beneficial effects of anthocyanin, the improvement of anthocyanin accumulation in plants was attempted through the selection of high-anthocyanin germplasm and the optimization of cultivation environments including light and temperature (Bian et al., 2015; Passeri et al., 2016).

Since anthocyanin accumulations varied during the growing periods of the plant depending on the plant genetic background, environmental conditions, and physiological stress (Chalker-Scott, 1999; Sibley et al., 1999), the non-destructive method of anthocyanin content estimation would be extremely valuable and especially useful for the investigations of pigment changes in individual intact leaves over time. As the less expensive approach, color leaf images obtained by the commercial digital camera can also be considered as an alternative parameter for non-destructive anthocyanin estimation (Simko et al., 2016). The hyperspectral imaging technique is an advanced imaging technology that can combine the advantages of spectroscopic and imaging techniques to detect the continuous wavelengths from visible to near-infrared lights selectively. Hyperspectral imaging has been successfully employed in previous studies to measure anthocyanin contents in various vegetables and fruits in a non-destructive manner (Liu et al., 2017; Gabrielli et al., 2021). However, it has been rarely tested to measure differentially expressed anthocyanin in plants grown under different environmental conditions during the growing period of vegetables.

The objective of this study was to determine that hyperspectral imaging and commercially available color imaging techniques could be employed to analyze the variations in anthocyanin accumulation and be differentially expressed in leaf tissues by cultivar-dependent genetic effect, influence of environmental factors, and their genetic × environmental (G × E) associations. The target plant was a bok choy (*Brassica rapa* var. *chinensis*), an important dietary vegetable cultivated and consumed worldwide for its edible leaves, and cultivation experiments were performed in an indoor farm with LED artificial light climate control systems.

MATERIALS AND METHODS

Plant Materials and Experimental Design

The first experiment was conducted to select and calibrate proper anthocyanin reflectance indices, such as anthocyanin absorption index (Merzlyak et al., 2003) and three bands model (Gitelson et al., 2009) in bok choy. Four commercial cultivars that had the variation in anthocyanin accumulations, "green stem," "white stem," "red stem," and "purple magic" (Asia Seed Co., Republic of Korea) were used in this experiment. Seeds of each bok choy cultivar were germinated in small pots filled with a horticultural soil mix (Nongwoo Co., Republic of Korea). Seedlings of 3-week-old were transplanted into 10 cm pots and grown in a growth room at the Korea Institute of Science and Technology at Gangneung under a 28 /20°C and 14/10 h day/night temperature regime and with LED artificial light (110 μmol/m^2/s). After 6 weeks of growing, the third and fourth leaves from apical meristem in each plant were subjected to hyperspectral imaging and collected for chemical analysis.

The second experiment was conducted to determine whether hyperspectral imaging and commercially available color imaging techniques could be employed to analyze variations in anthocyanin accumulation, differentially expressed in leaf tissues by cultivar-dependent genetic effect, the influence of environmental factors, and their genetic × environmental (G × E) associations. For the second experiment, the same cultivars were subjected to two different environmental conditions (different light intensity, temperature, and cultivation period) after seedling. The environmental condition for relatively lower light intensity and slight temperature differences (NC; non-stimulation condition) was the same as the environmental condition for the first experiment. The environmental condition for relatively higher light and extreme temperature difference (SC; stimulation condition) was at a growth room under a 28/15°C and 14/10 h day/night temperature regime and with LED artificial light (300 μmol/m^2/s). The leaf samples were subjected after 2 (GS1) and 4 (GS2) weeks of growing for hyperspectral imaging and then collected for color image acquisition and chemical analysis. Leaf color images were taken together with the reference color card (CTrax 24ColorCard-2 × 3, Camera Trax, Las Vegas, NV, United States).

The third experiment was performed to confirm the possibility of hyperspectral imaging to be applied into the "sensor-to-plant"-type high-throughput phenotyping stage (Lee et al., 2018) for the measurement of anthocyanin accumulations in the whole plant level. The same cultivars were prepared under the stimulation condition for anthocyanin accumulation (SC), and 2 week-grown whole plant (GS1) was subjected to hyperspectral imaging.

The experimental design was a randomized complete block with four replicates of five plants.

Hyperspectral Imaging Data Acquisition

Bok choy leaves and whole plants were imaged with SOC710-VP hyperspectral imager (Surface Optics Corp., San Diego, CA, United States). The hyperspectral imager is composed of 520 lines, 696 samples, and 128 bands (the spectral resolution of 4.69 nm). For the first and second experiments, the target leaf was placed at the focal plane of a hyperspectral imaging system, consisting of a hyperspectral camera and halogen lamp in the same growth room. The distance from the target leaf was about 30 cm for both the hyperspectral imaging camera and the light source. The hyperspectral image was acquired by placing a leaf on a white reference panel to distinguish the leaf from the background (**Figure 1A**).

For the third experiment, whole plant hyperspectral images were acquired as top view with the hyperspectral camera mounted 100 cm above the target plant in the same growth room. Halogen lamp was mounted 50 cm above the target plant to irradiate light, and a separate white reference panel was added to the light source part (**Figure 1B**).

After hyperspectral image scanning, we calculated transmittance and reflectance spectra of white reference panel and target samples. All hyperspectral imaging experiments were performed under the conventional indoor light environment. In addition, a white panel was adopted to monitor indoor light conditions, including irradiating halogen lamps. To minimize the light-affected environmental condition, we applied the calibration curve obtained from response correction to hyperspectral images so that the obtained spectra are localized only to leaves.

Chemical Measurement of Anthocyanin Accumulation

Anthocyanin accumulation in leaf tissue was analyzed as described in Gitelson et al.'s study (2001). Freeze-dried tissues of 0.5 g were homogenized in 10 ml of 70% methanol and passed through a paper filter after 5 min of sonication. Distilled water was then added to equal 0.2 of the extract volume, and the diluted filtrate was centrifuged in glass test tubes for 10 min at 3,000 g to separate the water–methanol phases. The absorption spectra of the water–methanol fraction were calculated using a plant reader (Epoch, BioTek Instruments Inc., Winooski, VT, United States). Anthocyanin accumulations were normalized to dry weight (mg/g).

Selection and Optimization of Reflectance Index Model to Estimate Anthocyanin Accumulation in Bok Choy

Gitelson et al. (2009) established a non-destructive technique to estimate anthocyanin contents in leaves of various tree species using various indices, such as absorbance (abs) value for anthocyanin, anthocyanin reflectance index (ARI), modified anthocyanin reflectance index (mARI), and anthocyanin content index (ACI). Among these indices, mARI is a three-band model based on measurements of tissue reflectance (R) at the specific wavelengths obtained from hyperspectral imaging: relative anthocyanin accumulation [(R759.5–797.02) × (1/R550.14–1/R701.06)] to present the actual and precise wavelengths and anthocyanin values (Merzlyak et al., 2003; Gitelson et al., 2006). We slightly modified the wavelength of this equation to correct the wavelength mismatch and visualize the precise anthocyanin value. The modified equation is [(R759.5–797.02) × (1/R550.14–1/R706.35)]. Using these previously developed anthocyanin estimation indices with our wavelength-corrected mARI, we analyzed the linear relationship between hyperspectral imaging and destructive analysis values to select the optimum index for anthocyanin estimation in bok choy. Total 20 plant samples obtained from the first experiment were subjected to both hyperspectral imaging and destructive analysis. As the result, the highest linear relationship between the destructive and non-destructive (hyperspectral imaging) of anthocyanin content was wavelength-corrected mARI ($R^2 = 0.9958$), followed by mARI ($R^2 = 0.8795$), abs value ($R^2 = 0.8115$), ARI ($R^2 = 0.4295$), and ACI ($R^2 = 0.0224$). Therefore, our wavelength-corrected mARI was selected as the anthocyanin reflectance index in bok choy, and anthocyanin accumulation expressed on a leaf area basis (mg/cm^2) were converted to dry weight (mg/g) using the following formula obtained from the data set of the first experiment: $y = 1.8123x + 0.0962$ ($R^2 = 0.9958$) (**Supplementary Figure 1**). This anthocyanin measurement method was applied to the hyperspectral imaging-based analysis of anthocyanin variations in the second and third experiments.

Leaf Image Analysis

All images of four cultivars acquired with the reference color card were corrected and auto-segmented from the background, and their color area was analyzed using the Leaf Analysis tool (NOROO KIBAN Systems Inc., Seongnam, Republic of Korea). We analyzed the area occupied by red and purple colors (RGB ± 20%) in the leaf area as percentages, separately (**Figure 4A**).

Statistical Analysis

All data were represented as means ± SD of at least four independent experiments. Statistical analysis was carried out using the SAS 9.4 software (SAS Institute Inc., Cary, NC, United States). The ANOVA was the Fisher's Least Significant Difference (LSD) analysis based on the 0.05 probability level.

RESULTS

Comparison of Anthocyanin Accumulations Estimated by Hyperspectral Imaging With the Value Measured by Wet Chemical (Destructive) Assay in Different Bok Choy Cultivars

Between the modified hyperspectral value and destructive analysis (UV) results, no significant differences in anthocyanin accumulation were observed across the four bok choy cultivars grown under the SC at GS2. Furthermore, the anthocyanin

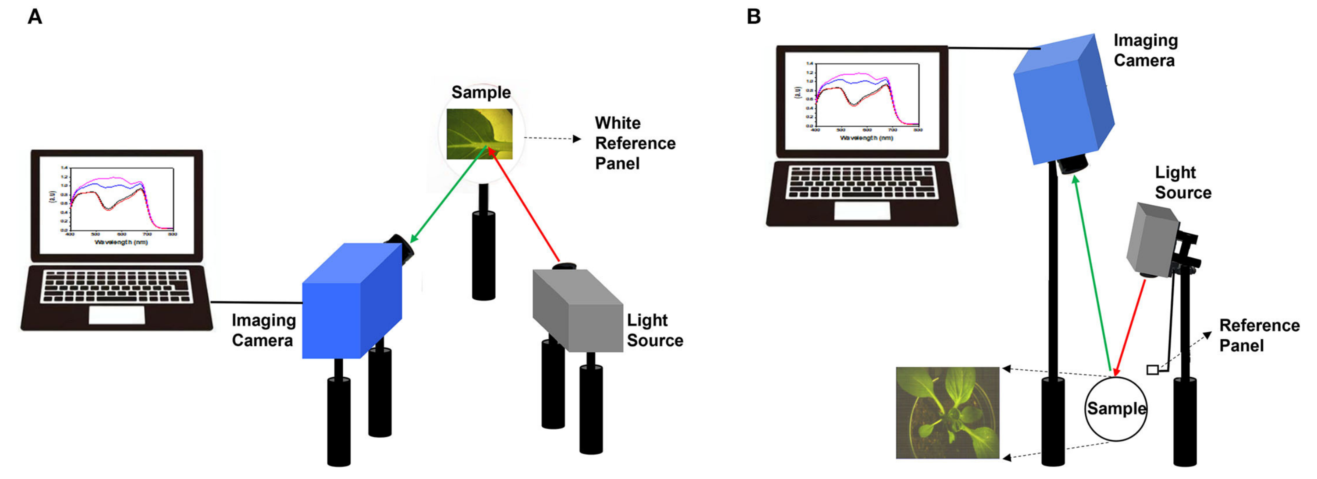


FIGURE 1 | Schematic diagram of hyperspectral imaging employed for target leaf analysis in the first and second experiment **(A)**, and whole plant analysis in the third experiment **(B)**.

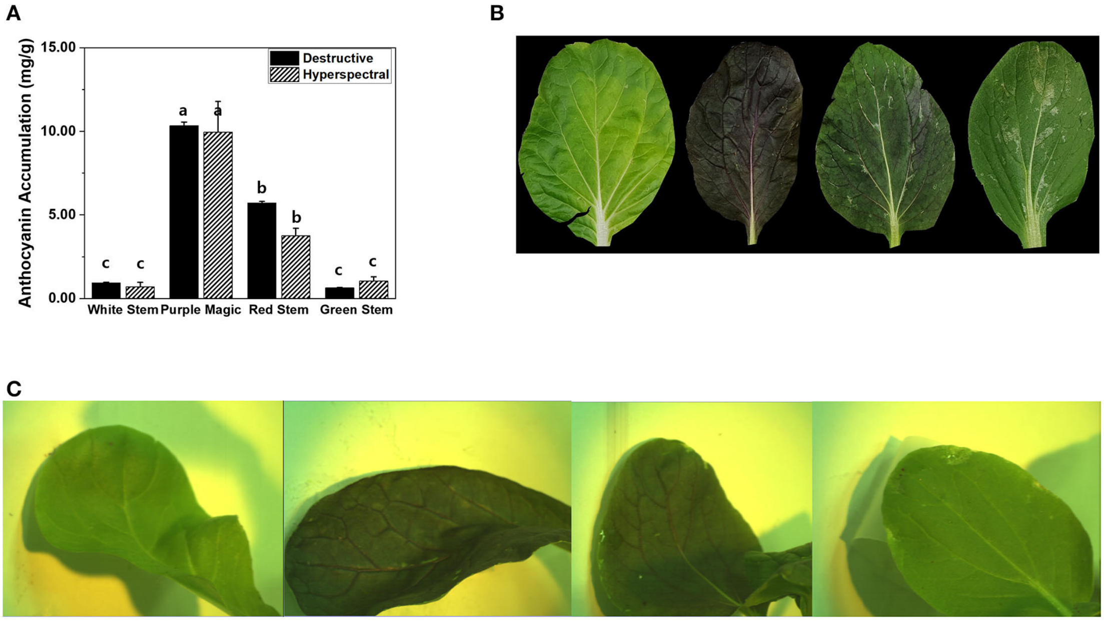


FIGURE 2 | (A) Comparison of anthocyanin accumulations (mg/g DW) obtained by destructive analysis and hyperspectral analysis among targeted leaf of four bok choy cultivars grown under the SC at GS2. These data were obtained from the second experiment described in the Materials and methods section. Means with a different letter are significantly different at $P \leqq 0.05$, using Fisher's LSD analysis. **(B)** The representative image of four different bok choy cultivars (from the left white stem, purple magic, red stem, and green stem). **(C)** The representative hyperspectral image was taken in leaf-level in four different bok choy cultivars (from the left white stem, purple magic, red stem, and green stem).

accumulations estimated by hyperspectral imaging and destructive analysis equally displayed cultivar-dependent significant differences across cultivars (**Figure 2A**). This result indicates that hyperspectral imaging can be employed to measure differences in variations of anthocyanin accumulations that occurred in the leaf tissue of different bok choy cultivars. Variations in anthocyanin accumulation in different cultivars seem to be related to the expression of color pigments in leaf tissues (**Figure 2B**). The "purple magic" showed the highest level of anthocyanin accumulation in both hyperspectral value and destructive analysis with the purple color expression in almost the whole area of the leaf, whereas "white stem" and "green stem" displayed the lowest level of anthocyanin accumulation with the no expression of purple color in a leaf. "red stem" showed a partial expression of purple color in the leaf tissue with the middle level of anthocyanin accumulation (**Figure 2**).

Variations of Anthocyanin Accumulations Caused by Different Indoor Farm Cultivation Conditions

Increased accumulation of anthocyanin under the SC was observed in "purple magic" and "red stem" by both hyperspectral imaging and destructive analysis (**Figure 3A**). In different growth stages (GS1 and GS2), no significant differences in anthocyanin accumulation between hyperspectral imaging and destructive analysis were found in each cultivar (**Figure 3B**). These results suggest that hyperspectral imaging can provide comparable analytic capability with destructive analysis to measure variations in anthocyanin accumulation under the different environmental

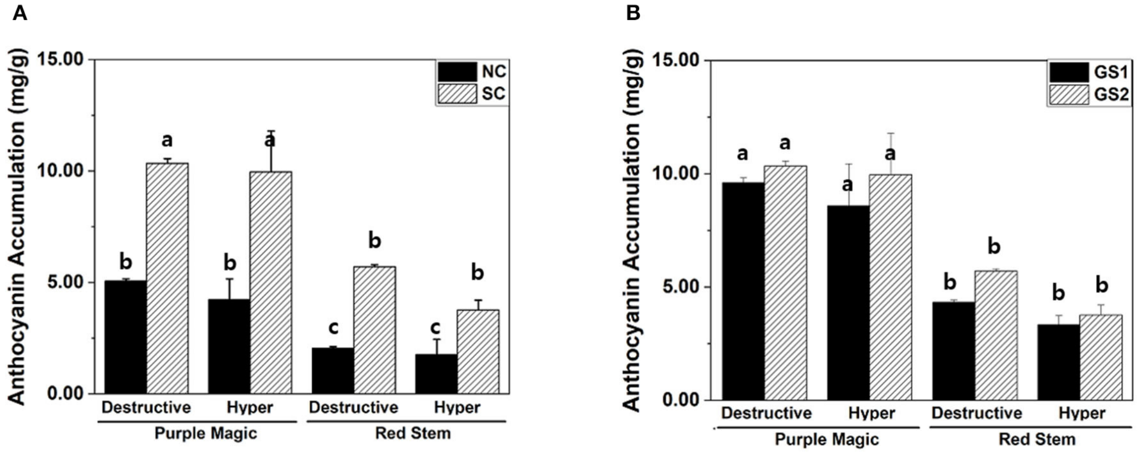


FIGURE 3 | Variation in anthocyanin accumulations (mg/g DW) of purple magic and red stem estimated by different cultivation conditions. **(A)** Destructive and hyperspectral analysis with non-stimulation and stimulation conditions. **(B)** Destructive analysis and hyperspectral analysis by cultivation period for 2 weeks (GS1) and 4 weeks (GS2). These data were obtained from the second experiment described in the Materials and methods section. Means with a different letter are significantly different at $P \leqq 0.05$, using Fisher's LSD analysis.

TABLE 1 | Variation of anthocyanin accumulation (mg/g) in leaf tissue of four bok choy cultivars grown under the different environmental conditions (NC and SC) with different growth stages (GS1 and GS2).

Genotype	Treatment[y]	TA (mg/g)[z]		TA (%)[z]
		Destructive (UV)[x]	Hyperspectral[x]	based on RGB color
"White Stem"	NC-GS1	0.880 ± 0.023^{g}	0.423 ± 0.027^{f}	0.30 ± 0.20^{f}
	SC-GS1	0.953 ± 0.043^{g}	0.529 ± 0.080^{e}	0.66 ± 0.36^{f}
	NC-GS2	0.378 ± 0.030^{h}	0.603 ± 0.051^{e}	3.01 ± 1.11^{e}
	SC-GS2	0.938 ± 0.048^{g}	0.698 ± 0.263^{e}	1.61 ± 0.79^{ef}
"Purple Stem"	NC-GS1	6.848 ± 0.158^{b}	1.982 ± 0.498^{d}	17.08 ± 10.31^{cde}
	SC-GS1	9.608 ± 0.218^{a}	8.585 ± 1.860^{b}	53.56 ± 11.62^{b}
	NC-GS2	5.060 ± 0.103^{cd}	4.232 ± 0.930^{c}	54.21 ± 11.69^{b}
	SC-GS2	10.3425 ± 0.215^{a}	9.959 ± 1.836^{a}	83.53 ± 4.20^{a}
"Red Stem"	NC-GS1	1.160 ± 0.030^{f}	0.550 ± 0.096^{e}	0.82 ± 0.31^{f}
	SC-GS1	4.325 ± 0.103^{d}	3.344 ± 0.395^{cd}	11.32 ± 4.38^{de}
	NC-GS2	2.045 ± 0.070^{e}	1.766 ± 0.678^{d}	27.53 ± 9.44^{c}
	SC-GS2	5.708 ± 0.098^{c}	3.768 ± 0.436^{c}	26.38 ± 8.53^{cd}
"Green Stem"	NC-GS1	0.863 ± 0.030^{g}	0.442 ± 0.009^{f}	0.34 ± 0.09^{f}
	SC-GS1	1.293 ± 0.093^{f}	0.573 ± 0.128^{e}	0.22 ± 0.06^{f}
	NC-GS2	0.663 ± 0.030^{gh}	0.620 ± 0.060^{e}	5.64 ± 1.88^{def}
	SC-GS2	0.633 ± 0.030^{gh}	1.046 ± 0.255^{d}	10.56 ± 2.60^{de}

Anthocyanin accumulation (mg/g) was analyzed by destructive analysis and hyperspectral imaging.

[z] *TA: total anthocyanin accumulation.*

[y] *NC-GS1: treatment with low lights and slight temperature difference for 2 weeks, SC-GS1: treatment with high lights and extreme temperature difference for 2 weeks, NC-GS2: treatment with low lights and slight temperature difference for 4 weeks, SC-GS2: treatment with high light and extreme temperature difference for 4 weeks.*

[x] *Means within each column and cultivar with different letters are significantly different at $P \leqq 0.05$, using Fisher's LSD analysis.*

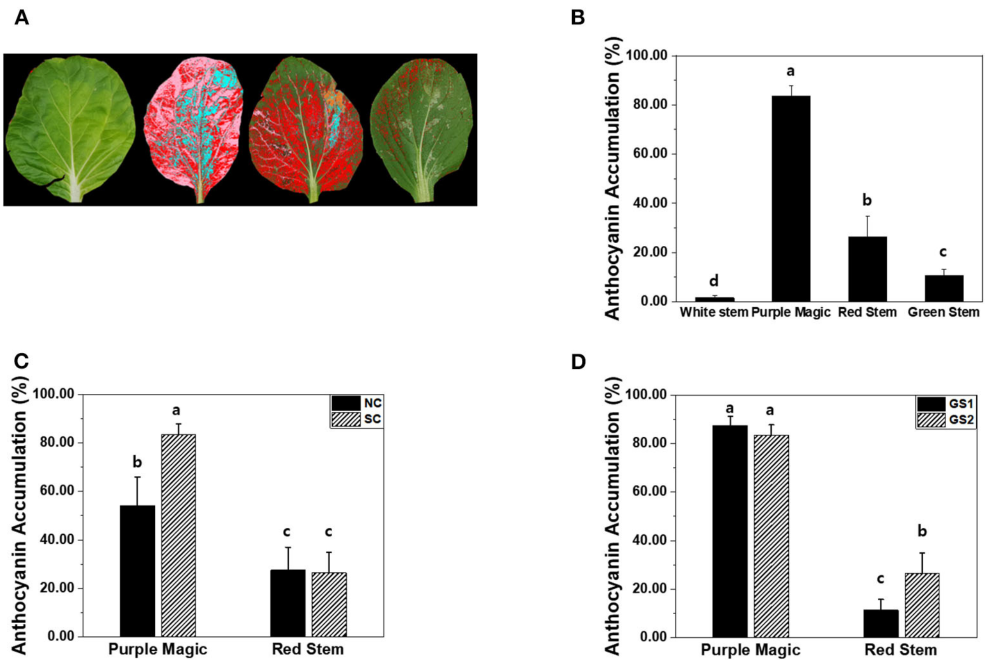


FIGURE 4 | (A) The representative image of four different bok choy cultivars (from the left white stem, purple magic, red stem, and green stem) for leaf color analysis. **(B)** Comparison of anthocyanin accumulation (%) obtained by leaf color analysis among four cultivars, bok choy. **(C)** AT accumulation (%) of purple magic and red stem by non-stimulation (NS) and stimulation (CS) conditions. **(D)** AT accumulation (%) of purple magic and red stem by cultivation period for 2 weeks (GS1) and 4 weeks (GS2). These data were obtained from the second experiment described in the Materials and methods section. Means with a different letter are significantly different at $P \leqq 0.05$, using Fisher's LSD analysis.

conditions of the indoor farm. Although significant increases of anthocyanin accumulations were also found in "white stem" and "green stem" under the SC, these increases would be caused by the limit of analytic capability of both hyperspectral imaging and destructive analysis under the relatively lower level of anthocyanin compared with its level of "purple magic" and "red stem" (**Table 1**).

Leaf Image Analysis-Based Estimation of Anthocyanin Accumulations

Leaf image analysis is one of the potential approaches to estimate anthocyanin accumulations in a non-destructive manner. Unlike hyperspectral imaging and destructive analysis, leaf image analysis significantly increased anthocyanin accumulation in GS2 of "red stem" (**Figures 3B**, **4D**). The percentages of the purple color area in the total leaf area among the four cultivars varied similar to the anthocyanin accumulation values measured by hyperspectral imaging and destructive analysis in the four cultivars (**Figure 4B**). However, leaf image analysis did not show a significant increase in anthocyanin accumulation in "red stem" under the SC compared with the results obtained by hyperspectral imaging and destructive analysis (**Figures 3A**, **4C**). The different consequence was also found in anthocyanin accumulation of "red stem" leaves in different growth stages. These results indicated that the proportion of colored leaf area may be employed to estimate the anthocyanin accumulations of different cultivars showing clear differences in anthocyanin levels with quantitative values. However, it has limitations to reflect the level of color saturation caused by pigment accumulation, such as anthocyanin accumulation.

TABLE 2 | Percentages of variations in anthocyanin accumulations (mg/g) associated with genotype, environment, and genotype by environment interaction for the leaves of four bok choy genotypes grown in different environmental conditions (NC and SC) with different growth stages (GS1 and GS2).

Source of variation	Percentage of variations in TA^z	
	Destructive (UV)	**Hyperspectral**
Genotype	77.7*	59.2*
Environment	6.7	11.1
G × E ^{y}	10.8**	26.6***
Residual	4.8***	2.9***

Anthocyanin accumulation (mg/g) was analyzed by destructive analysis and hyperspectral imaging.

^{z}TA: total anthocyanin accumulation.

^{y}G × E = genotype × environment interaction.

**, **, ***Significant at $P \leqq 0.2$, 0.05, or 0.025, respectively.*

Genotype and Environment Interaction

The ANOVA was applied to partition the variations in anthocyanin accumulation into components associated with the genotype, environment, and genotype × environment interaction (G × E) using the data obtained from hyperspectral imaging and destructive analysis, respectively (**Table 2**). Comparison of the results obtained from hyperspectral imaging and destructive analysis data demonstrated a similar proportion of genotypic and G × E effects with the same significance, indicating comparability of hyperspectral imaging with destructive analysis for the G × E study. The high proportion of anthocyanin variation was primarily contributed by the genotypic effect, indicating that selection and breeding of cultivars can be performed to optimize anthocyanin accumulation in bok choy for indoor farming. Since our data also demonstrated the significance of G × E interaction (**Table 2**), attention should be given to identifying germplasm in which anthocyanin biosynthesis is maximized in the specific environmental condition of indoor farms.

Establishment of the Hyperspectral Imaging System to Be Applied to High-Throughput Plant Phenotyping System

In our previous study, a "sensor-to-plant" type plant phenotyping system was developed to accelerate the large-scale acquisition of plant images through the moving actuators with vision sensors in real-time (Lee et al., 2018). To apply our hyperspectral imaging onto the high-throughput phenotyping system, the hyperspectral imaging system was modified to acquire a top view image of the whole plant (**Figures 1B**, **5C**). Two-week-old whole plants of four bok choy cultivars were subjected to hyperspectral imaging in the third experiment. Similarly, classified anthocyanin accumulations were observed across four cultivars in the results obtained from both hyperspectral imaging and destructive analysis (**Figure 5A**). In addition, the linear relationship between hyperspectral imaging and destructive analysis data ($R^2 = 0.9339$, RMSE $= 0.05$) suggests that the hyperspectral imaging of the whole plant is comparable with the destructive analysis and feasible to be applied onto the high-throughput plant phenotyping system (**Figure 5B**).

DISCUSSION

There were various attempts to analyze anthocyanin accumulation non-destructively, such as the hyperspectral model (Gu et al., 2018), UV-Spec. (Gitelson et al., 2009), and the fluorimetric sensor (Tuccio et al., 2011). We tested previously established hyperspectral models, such as the ACI (Gitelson et al., 2006), ARI (Gitelson et al., 2001, 2006), and mARI (Gitelson et al., 2006) to select the optimum model to analyze anthocyanin accumulations using hyperspectral imaging in bok choy. We also attempted to optimize mARI to improve the measurement

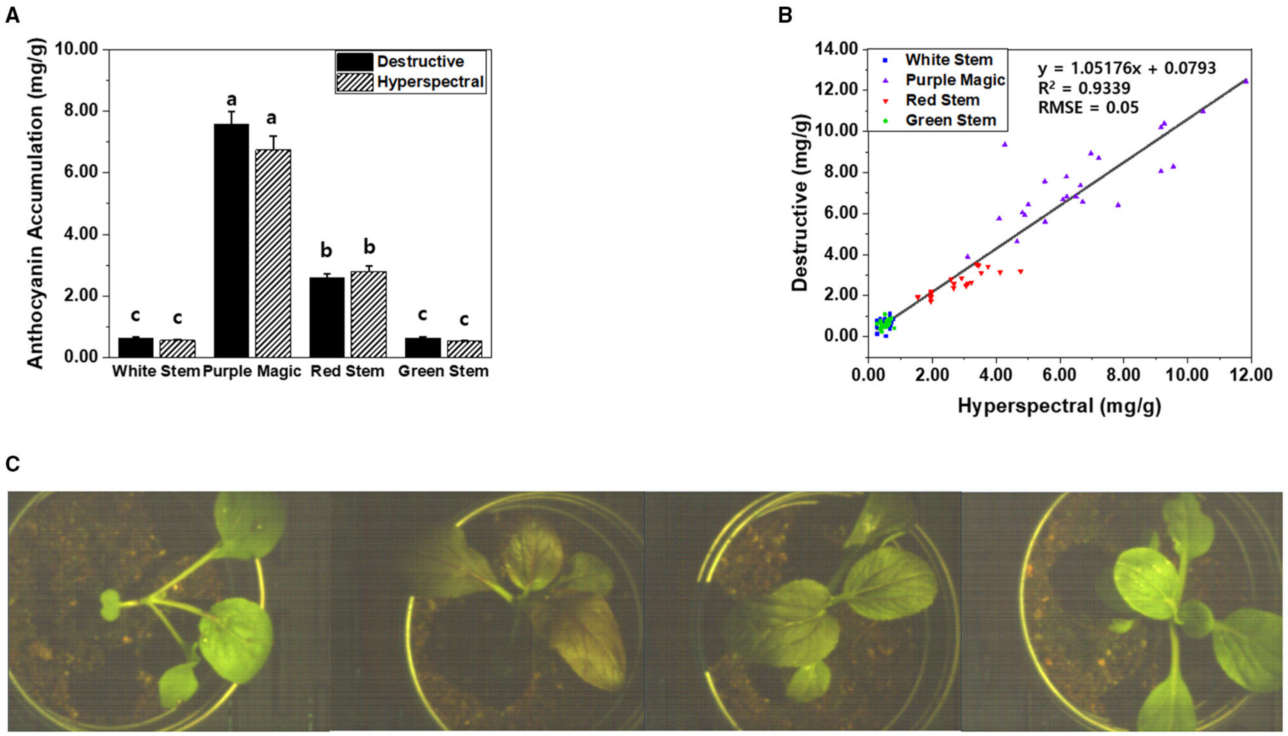


FIGURE 5 | (A) Comparison of anthocyanin accumulations (mg/g DW) obtained by destructive analysis and hyperspectral analysis from the whole plant of four bok choy cultivars grown under the SC at GS1. These data were obtained from the third experiment described in the Materials and methods section. Means with a different letter are significantly different at $P \leqq 0.05$, using Fisher's LSD analysis. **(B)** The linear relationship between hyperspectral imaging and destructive analysis data was obtained from the whole plant of four bok choy cultivars grown under the SC at GS1. **(C)** The representative hyperspectral image was taken at the whole plant level in four different bok choy cultivars (from the left white stem, purple magic, red stem, and green stem).

accuracy through the adjustment of wavelength in the first experiment, and we confirmed that our wavelength-adjusted mARI approach could improve the linear relationship ($R^2 = 0.9958$) between destructive and non-destructive analyses for anthocyanin estimation. Further studies conducted with our optimized hyperspectral imaging approach in the second and third experiments demonstrated comparable trends in changes of anthocyanin accumulations across different cultivars and different environmental conditions between hyperspectral imaging and destructive analysis, although statistical significance was limited in the cultivars and environmental conditions inducing a relatively lower level of anthocyanin accumulation.

High-throughput phenotyping has received considerable attention due to its potential for rapid identification of valuable germplasm (Matsuda et al., 2012; Banerjee et al., 2020). In our previous study, we established a high-throughput phenotyping system consisting of an image-capturing hardware module, environmental data sensors, and automated irrigation and artificial light controllers (Lee et al., 2018). In this system, a total of 28 plant trays (52.5 cm × 26.5 cm in size) are placed in a 4 × 7 matrix, and the image acquisition module moves over 28 plant trays following the pre-defined X, Y, and Z coordinates of the plant trays. In each plant tray, 8 of 10 cm pots are placed. Therefore, a total of 280 pots can be placed in our phenotyping system. Our hyperspectral imaging system was modified to be applied to our high-throughput phenotyping system (**Figure 1B**). Using our hyperspectral imaging system, it took 2 min to acquire a top view image of the whole plant in 10 cm pots, and an additional 1 min to move pots and process data. It indicated that a total of 280 plants in 10 cm pots could be scanned by hyperspectral imaging under our phenotyping system for 14 h for rapid identification of bok choy germplasms.

In the conventional indoor farm, increasing artificial light intensity and day/night temperature differences to improve anthocyanin accumulation in vegetables also may cause an increase in the energy costs. To solve this conflicting problem, not only selection of proper germplasm that can show a better performance of anthocyanin accumulation under the acceptable indoor farming environments but also identification of optimum environmental conditions through the monitoring of changes in anthocyanin accumulations under the diverse indoor farm environmental conditions need to be accomplished as a further study. According to the previous studies, lettuce grown under 5°C of night-time temperatures showed a relatively higher polyphenolic content than lettuce cultivated under 20°C for 5 days. However, phenol content showed the highest levels at 20°C in 20 days of cultivation duration (Jeong et al., 2015). In another study, anthocyanin, carotenoid, and relative chlorophyll contents were decreased when purple bok choy was exposed to low light (250 μmol/m^2/s) compared with high light (1,000 μmol/m^2/s) (Zhu et al., 2017). These reports suggest that additional studies on the change of anthocyanin accumulations in bok choy with more diverse environmental conditions for more than 4 weeks of cultivation period using hyperspectral analysis are considered to be necessary. We believe that hyperspectral imaging can be an affordable approach to perform it in a non-destructive manner.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

H-sK and JY performed the experiment, data collection, statistical data analysis, presentation, and drafted the manuscript. SP, J-SK, and YC experimented and analyzed the data. JK and HK conceived the project and experimental design and collaborated on the data interpretation and manuscript revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the Next-Generation BioGreen 21 Program (Project No. PJ012281022018), Rural Development Administration, Republic of Korea.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.693854/full#supplementary-material

Supplementary Figure 1 | Reflectance index model for estimation of anthocyanin accumulation between the destructive (UV) and non-destructive (Hyperspectral), mARI of R706.35 **(A)**, mARI of R701.06 **(B)**, abs **(C)**, ARI **(D)**, and CI **(E)**.

REFERENCES

Banerjee, B. P., Joshi, S., Thoday-Kennedy, E., Pasam, R. K., Tibbits, J., Hayden, M., et al. (2020). High-throughput phenotyping using digital and hyperspectral imaging-derived biomarkers for genotypic nitrogen response. *J. Exp. Bot.* 71, 4604–4615. doi: 10.1093/jxb/eraa143

Bian, Z. H., Yang, Q. C., and Liu, W. K. (2015). Effects of light quality on the accumulation of phytochemicals in vegetables produced in controlled environments: a review. *J. Sci. Food Agric.* 95, 869–877. doi: 10.1002/jsfa.6789

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., et al. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* 26, 1301–1308. doi: 10.1038/nbt.1506

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* 70, 1–9. doi: 10.1111/j.1751-1097.1999.tb01944.x

de Pascual-Teresa, S., Moreno, D. A., and García-Viguera, C. (2010). Flavanols and anthocyanins in cardiovascular health: a review of current evidence. *Int. J. Mol. Sci.* 11, 1679–1703. doi: 10.3390/ijms11041679

Gabrielli, M., Lancon-Verdier, V., Picouet, P., and Maury, C. (2021). Hyperspectral imaging to characterize tavle grapes. *Chemosensors* 9:71. doi: 10.3390/chemosensors9040071

Gitelson, A. A., Chivkunova, O. B., and Merzlyak, M. N. (2009). Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *Am. J. Bot.* 96, 1861–1868. doi: 10.3732/ajb.0800395

Gitelson, A. A., Keydan, G. P., and Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophys. Res. Lett.* 33:L11402. doi: 10.1029/2006GL026457

Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* 74, 38–45. doi: 10.1562/0031-8655(2001)074<0038:OPANEO>2.0.CO;2

Gu, X., Cai, W., Fan, Y., Ma, Y., Zhao, X., and Zhang, C. (2018). Estimating foliar anthocyanin content of purple corn via hyperspectral model. *Food Sci. Nutr.* 6, 572–578. doi: 10.1002/fsn3.588

Jeong, S. W., Kim, G.-S., Lee, W. S., Kim, Y.-H., Kang, N. J., Jin, J. S., et al. (2015). The effects of different night-time temperatures and cultivation durations on the polyphenolic contents of lettuce: application of principal component analysis. *J. Adv. Res.* 6, 493–499. doi: 10.1016/j.jare.2015.01.004

Lee, U., Chang, S., Putra, G. A., Kim, H., and Kim, D. H. (2018). An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. *PLoS ONE* 13:e0196615. doi: 10.1371/journal.pone.0196615

Liu,Y., Sun, Y., Xie, A., Yu, H., Yin, Y., Li, X., et al. (2017). Potential of hyperspectral imaging for rapid prediction of anthocyanin content of purple-fleshed sweet potato slices during drying process. *Food Anal. Methods* 10, 3836–3846. doi: 10.1007/s12161-017-0950-y

Matsuda, O., Tanaka, A., Fujita, T., and Iba, K. (2012). Hyperspectral imaging techniques for rapid identification of Arabidopsis mutants with altered leaf pigment status. *Plant Cell Physiol.* 53, 1154–1170. doi: 10.1093/pcp/pcs043

Merzlyak, M. N., Solovchenko, A. E., and Gitelson, A. A. (2003). Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. *Postharvest Biol. Technol.* 27, 197–211. doi: 10.1016/S0925-5214(02)00066-2

Passeri, V., Koes, R., and Quattrocchio, F. M. (2016). New challenges for the design of high value plant products: stabilization of anthocyanins in plant vacuoles. *Front. Plant Sci.* 7:153. doi: 10.3389/fpls.2016.00153

Sibley, J. L., Ruter, J. M., and Eakes, D. J. (1999). Bark anthocyanin levels differ with location in cultivars of red maple. *HortScience* 34:137. doi: 10.21273/HORTSCI.34.1.137

Simko, I., Hayes, R. J., and Furbank, R. T. (2016). Non-destructive phenotyping of lettuce plants in early stages of development with optical sensors. *Front. Plant Sci.* 7:1985. doi: 10.3389/fpls.2016.01985

Strack, D., and Wray, V. (1989). "9 - Anthocyanins," in *Methods in Plant Biochemistry Plant Phenolics*, ed. J. B. Harborne (New York, NY; London: Academic Press), 325–356.

Tuccio, L., Remorini, D., Pinelli, P., Fierini, E., Tonutti, P., Scalabrelli, G., et al. (2011). Rapid and non-destructive method to assess in the vineyard grape berry anthocyanins under different seasonal and water conditions*. *Aust. J. Grape Wine Res.* 17, 181–189. doi: 10.1111/j.1755-0238.2011.00139.x

Williams, R. J., Spencer, J. P. E., and Rice-Evans, C. (2004). Flavonoids: antioxidants or signalling molecules? *Free Radic. Biol. Med.* 36, 838–849. doi: 10.1016/j.freeradbiomed.2004.01.001

Zhu, H., Li, X., Zhai, W., Liu, Y., Gao, Q., Liu, J., et al. (2017). Effects of low light on photosynthetic properties, antioxidant enzyme activity, and anthocyanin accumulation in purple pak-choi (*Brassica campestris* ssp. Chinensis Makino). *PLoS ONE* 12:e0179305. doi: 10.1371/journal.pone.0179305

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kim, Yoo, Park, Kim, Chung, Kim and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers
in Plant Science

ORIGINAL RESEARCH
published: 19 August 2021
doi: 10.3389/fpls.2021.695102

Prediction of Drought-Induced Components and Evaluation of Drought Damage of Tea Plants Based on Hyperspectral Imaging

Sizhou Chen[1], Yuan Gao[2], Kai Fan[1], Yujie Shi[1], Danni Luo[1], Jiazhi Shen[3], Zhaotang Ding[1,3] and Yu Wang[1,3]**

[1] Tea Research Institute, Qingdao Agricultural University, Qingdao, China, [2] Jinan Agricultural Technology Promotion Service Center, Jinan, China, [3] Tea Research Institute, Shandong Academy of Agricultural Sciences, Rizhao, China

Effective evaluation of physiological and biochemical indexes and drought degree of tea plant is an important technology to determine the drought resistance ability of tea plants. At present, the traditional detection method of tea drought stress is mainly based on physiological and biochemical detection, which is not only destructive to tea plants, but also time-consuming and laborious. In this study, through simulating drought treatment of tea plant, hyperspectral camera was used to obtain spectral data of tea leaves, and three machine learning models, namely, support vector machine (SVM), random forest (RF), and partial least-squares (PLS) regression, were used to model malondialdehyde (MDA), electrolyte leakage (EL), maximum efficiency of photosystem II (*Fv/Fm*), soluble saccharide (SS), and drought damage degree (DDD) of tea leaves. The results showed that the competitive adaptive reweighted sampling (CARS)-PLS model of MDA had the best effect among the four physiological and biochemical indexes (Rcal = 0.96, Rp = 0.92, RPD = 3.51). Uninformative variable elimination (UVE)-SVM model was the best in DDD (Rcal = 0.97, Rp = 0.95, RPD = 4.28). Therefore, through the establishment of machine learning model using hyperspectral imaging technology, we can monitor the drought degree of tea seedlings under drought stress. This method is not only non-destructive, but also fast and accurate, which is expected to be widely used in tea garden water regime monitoring.

Keywords: hyperspectral imaging, machine learning, non-destructive testing, tea plants, drought assessment

OPEN ACCESS

Edited by:
Penghao Wang,
Murdoch University, Australia

Reviewed by:
Xie Jiangbo,
Zhejiang Agriculture and Forestry University, China
Xi Qiao,
Chinese Academy of Agricultural Sciences, China

****Correspondence:***
Zhaotang Ding
dzttea@163.com
Yu Wang
wangyutea@163.com

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *14 April 2021*
Accepted: *20 July 2021*
Published: *19 August 2021*

Citation:
Chen S, Gao Y, Fan K, Shi Y, Luo D, Shen J, Ding Z and Wang Y (2021) Prediction of Drought-Induced Components and Evaluation of Drought Damage of Tea Plants Based on Hyperspectral Imaging. Front. Plant Sci. 12:695102. doi: 10.3389/fpls.2021.695102

INTRODUCTION

Drought is the main factor affecting crop growth and development, which affects crop quality and yield worldwide. With climate change, especially global warming and the increase in non-agricultural water demand, drought will seriously affect the growth, yield, and quality of tea (Sharma and Kumar, 2005). According to reports, drought reduced tea production by 14–33% and caused 6–19% of plant deaths (Cheruiyot et al., 2010). At present, there are many traditional methods to detect the drought status of tea plants (Tian et al., 2019), but it is urgent to find a more timely and efficient detection method for tea drought status.

Recent studies have documented and explained the response of plant systems to drought stress. Tea plants adapt to drought stress through a series of physiological and biochemical reactions, such as osmotic pressure regulation, antioxidant activity, and plant hormone regulation (Liu and Chen, 2014). Under drought stress conditions, the content of soluble saccharide (SS) in tea plants will increase to cope with the stress. However, tea plants will cause membrane peroxidation, which will damage the membrane system and detect the increase in electrolyte leakage (EL) in plant cells. The content of malondialdehyde (MDA) as an oxidation product will increase, which will reduce the photosynthetic intensity of the cell membrane-dependent system. At this time, the maximum efficiency of the photosystem II value of plants will be lower than the normal level. In general, MDA, EL, maximum efficiency of photosystem II (*Fv/Fm*), and SS are used to evaluate the drought status of tea plants (Prieto et al., 2009; Soleimanzadeh, 2010; Guo et al., 2017). However, these traditional methods are not only time-consuming but also destructive (Tian et al., 2019).

Therefore, how to detect the physiological and biochemical components of plants under drought stress in real time is an urgent problem to be solved. Hyperspectral imaging technology, as a new phenotypic research technology, makes it possible to quickly, accurately, and non-destructively assess the water status of tea plants. Hyperspectral data have the characteristics of high spectral resolution, wide spectral range, continuous band, and rich information. Previous studies on hyperspectral imaging mainly used vegetation index or characteristic bands as input variables. The method of using vegetation index as a modeling variable has the characteristics of a small amount of data and fast calculation speed, which can be used for the large-scale data evaluation. For example, Zovko et al. (2019) found that using vegetation index to establish the prediction model can predict the drought degree of grape to a certain extent. Wang et al. (2014b) used the vegetation index (PRI, RENDVI, OSAVI, etc.) of spring wheat to build the corresponding stress prediction model, and they found that the model has a certain significance for monitoring the degree of crop stress in semi-arid stress areas. Zelazny and Lukáš (2020) found that RGI, CI, RNDVI, and GI of rape seedlings were related to drought intensity, and they took them as input variables to establish a drought stress prediction model of rape seedlings, which achieved good results. The method of characteristic bands as modeling variable has the characteristics of high accuracy and strong generalization ability. There are also related studies on this method. Kong et al. (2016) used partial least-squares (PLS) regression, LS-SVM, and ELM algorithms to extract the characteristic bands of MDA of oilseed rape leaves as the input variables of the model, and they found that the characteristic bands extracted by this method mainly concentrated in the range of 524–868 nm, and the model achieved the expected effect. Jiang et al. (2016) used competitive adaptive reweighted sampling (CARS) and GA algorithms to extract the characteristic bands of potato SS, and they found that the model had a good prediction ability in 450–470-, 520–560-, 730–810-, 860–890-, and 910–980-nm bands (Jiang et al., 2016).

Previous studies used various algorithms to analyze the correlation of different types of data and establish a robust prediction model. Shi and Cheng used multiplicative scatter correction (MSC), first derivative (1D), second derivative (2D), and Savitzky–Golay (S-G) to preprocess hyperspectral image data, and they found that these preprocessing algorithms have an excellent effect on eliminating baseline drift and multiple scattering effects (Shi et al., 2014; Cheng et al., 2019). Filho et al. used successive projections algorithm (SPA), uninformative variable elimination (UVE), CARS, and other algorithms to extract sample feature data, and they found that these algorithms can extract the most representative sample subset from the dataset (Araújo et al., 2001; Filho et al., 2004; Zhang et al., 2010; Li et al., 2019a). Qin et al. used support vector machine (SVM), random forest (RF), PLS regression, and other algorithms to model the sample set, and they found that these algorithms can adapt to different data types for modeling and analyzing and can establish stable mathematical models (Qin and He, 2005; Iverson et al., 2008; Lin et al., 2016). The above studies showed that choosing the appropriate algorithms for different types of data can save calculation time and improve the accuracy of the models. However, the comprehensive evaluation of tea drought status using hyperspectral imaging technology and the mathematical algorithm has not been reported.

In this study, hyperspectral imaging technology was used to comprehensively evaluate the drought status of tea plants. MSC, 1D, 2D, and S-G algorithms were used as preprocessing methods; SPA, UVE, and CARS algorithms were used as feature band screening methods; and SVM, RF, and PLS algorithms were used as prediction models. The principal component analysis (PCA) was used to weight the MDA, EL, and SS, which were positively correlated with the drought degree of tea plants, and a comprehensive evaluation index of drought degree of tea plants was obtained: drought damage degree (DDD), so as to more accurately reflect the drought stress suffered by tea plants.

MATERIALS AND METHODS

Experimental Design

The experiment was carried out in the greenhouse of Qingdao Agricultural University. The movable cultivation platform in the greenhouse is 3.5 m long and 1 m wide, with a total of four rows. The variety of tea plants is “Zhongcha 108,” and the age of seedlings is 2 years. The soil, substrate, and tea seedlings were disinfected, and 576 tea seedlings were cultivated in plug culture. On December 21, 2020, the tea plants will be precultured, and the tea seedlings will be irrigated quantitatively to keep the relative humidity of soil at about 50%. The air humidity in the greenhouse will be controlled at about 40% by the humidifier, and the temperature will be set at 26°C in the daytime and 20°C at night. The greenhouse was ventilated for 1–2 h every day, and the culture lasted for 2 weeks. From January 4, 2021, to January 19, 2021, the sprinkler irrigation or irrigation was stopped, the air humidifier was closed, and other conditions remained unchanged to simulate the drought stress of tea plants by high temperature and natural water loss. From 10:00 a.m. to 12:00 a.m., most of the biochemical indicators increased at noon, and we chose this time to sample and collect their data (Zhang et al., 2006; Guo et al., 2008). Each time 30 tea plants were randomly selected,

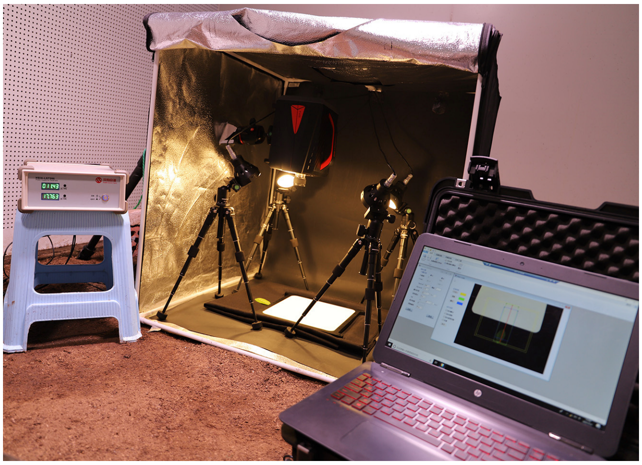


FIGURE 1 | Hyperspectral image acquisition system.

one mature leaf was taken from each tea plant, and a total of 30 samples were collected for hyperspectral data collection. The physiological and biochemical indexes of collected leaves were determined, and each sample was repeated three times. In this experiment, 180 samples were collected for six times.

Data Acquisition

Collection of Spectral Data

The hyperspectral image acquisition system device is shown in **Figure 1** (**Supplementary Figure 1** shows the detailed components of the system), which mainly includes imaging spectral camera (Gaia field pro-v10, Finland), light source (Hsia-ls-t-200w, China), displacement platform, PC, and other components. In order to get a clearer image, the exposure time is 9 ms, the field angle is 22°, and the object distance (the distance from the sample to the lens) is 38 cm. The color temperature of the light source is 3,000 K. The spectral range of the collected image is 400–1,100 nm, and the size of hyperspectral image data block is 960 pixels × 1,101 pixels × 176 bands. In order to improve the signal-to-noise ratio of hyperspectral image, the black-and-white correction method is used to remove the dark current noise caused by the internal current instability of the spectral camera (Talens et al., 2013). The formula of black and white correction is:

$$C = 65552(R - D)/(W - D)$$

where C is the corrected image, R is the original image, D and W are all black and all white images, respectively, and 65,552 is the maximum value of digital quantization value (DN).

Determination of MDA, EL, Fv/Fm, SS, and DDD

The physiological and biochemical indexes of tea leaves were measured by fresh samples, and the specific methods are as follows:

Determination of MDA and SS: the fresh leaf samples crushed by grinding machine (IKA A11, Germany) were extracted with TBA (4,6-dihydroxy-2-mercaptopyrimidine) solution at 100°C. According to the colorimetric method described by Li et al. (2019b), the absorption value of MDA and SS was read at 532 and 450 nm, respectively, by spectrophotometer (Zhou and Leul, 1999; Morales and Munné-Bosch, 2019; Tian et al., 2019).

TABLE 1 | Descriptive statistics of drought-induced components and drought damage degree of total fresh leaf samples.

Index	Maximum	Minimum	Average value	Standard deviation
MDA (mmol/kg FW)	9.61	3.26	5.95	1.76
EL (%)	49.70	18.76	33.14	7.29
Fv/Fm	0.92	0.6	0.76	0.07
SS (mmol/g FW)	13.1	5.1	8.98	1.93
DDD(Level)	9.06	3.65	6.07	1.40

Determination of EL: the leaf samples were cut and rinsed with deionized water for a short time. Under the condition of−0.1MPa, the vacuum pump (SHB-IIIA, China) was used to vacuum for 10 min. According to the method described by Tian et al., the conductivity (C_1) was measured by conductivity meter (DDSJ-308A, China). Then, the solution was boiled for 10 min, and the conductivity (C_2) was measured after cooling (Kate and Johnson, 2000; Tian et al., 2019; Takashima et al., 2021).

$$RPC\ (\%) = C_1/C_2 \times 100$$

Determination of *Fv/Fm*: After dark treatment for 20–30 min, the *Fv/Fm* value of tea leaves was determined by Fluor Pen (Fluor Pen FP110Hand held chlorophyll fluorometer, Czech Republic).

Determination of soil relative moisture: the relative moisture of the soil at the time of sample collection was determined by using a soil moisture-measuring instrument (TOP Cloud-agri TZS-I, China).

The process of obtaining DDD: three physiological data (MDA, EC, and SS) positively correlated with drought degree of tea plant were standardized, and the eigenvalues and eigenvectors of the correlation matrix were calculated, and the principal component score was calculated according to the cumulative contribution rate (the sum of the three variables is >0.85, so the three variables are available). The calculation formula of DDD can be obtained:

$$Y = 0.359X_1 + 0.341X_2 + 0.3X_3$$

where X_1 is MDA, X_2 is EL, and X_3 is SS. The contents of drought-induced components and DDD are shown in **Table 1**, mainly including maximum, minimum, average, and standard deviation. The distributions of drought-induced components and DDD of six periods under drought stress are shown in **Figure 2**; the change of soil relative humidity during drought treatment is shown in **Supplementary Figure 2**.

Extraction of Spectral Variables

In the hyperspectral image processing software Specview (Dualix spectral imaging, China), the hyperspectral image is corrected by lens correction and reflectance correction, and the

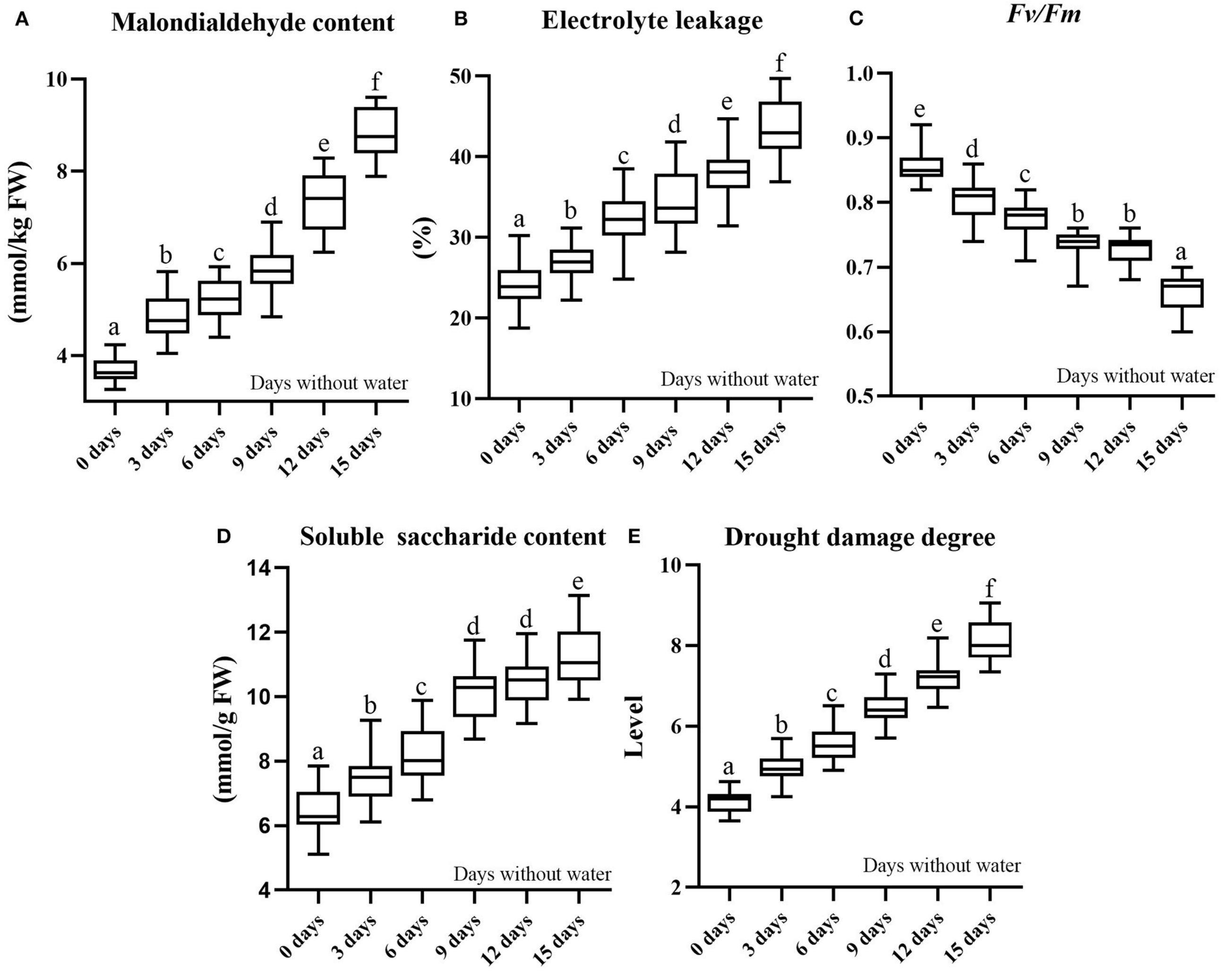


FIGURE 2 | Data distribution of drought-induced components (three repeats) and damage degree of six periods under drought stress (boxplot) **(A)**. Malondialdehyde content; **(B)**. electrolyte leakage; **(C)**. *Fv/Fm*; **(D)**. soluble saccharide content; **(E)**. drought damage degree. The data box in the figure below different letters are significantly different at $P < 0.05$ according to Duncan's test.

standardized hyperspectral image is obtained. In the remote sensing image processing software Envi5.3 (RSI, America), threshold segmentation is used to remove the background pixels of the corrected hyperspectral image, and the average spectral value of the leaf part is extracted by the combination of binarization and mask (Duan, 2016). The average spectra of all samples are extracted in turn, and the 176 × 180 (number of variables × number of samples) spectral matrix is obtained. The specific process is shown in **Figure 3**.

Spectral Data Preprocessing Method

In order to enhance the correlation between spectral parameters and tea plant indexes, the original data were preprocessed by MSC, S-G, and differential method (1D, 2D), where MSC is a common data processing method for multiwavelength modeling at present. The processed spectral data can effectively eliminate the scattering effect and enhance the quality of spectral information. The relevant formula is as follows:

$$\begin{aligned}\text{Calculate the average spectrum}:\overline{X\,(i)} &= \frac{\sum_{i=1}^{n} x\,(i)}{n}\\ \text{Linear regression}:X\,(i) &= m\,(i)*\overline{x\,(i)}+b\,(i)\\ \text{MSC correction}:X\,(i)_{(msc)} &= \frac{x\,(i)-b\,(i)}{m\,(i)}\end{aligned}$$

where X is the original spectral matrix of the sample, $X\,(i)\,, m\,(i)\,, b\,(i)\,,$ $and\ X\,(i)_{(msc)}$ are the surface original spectral mean, regression constant, regression coefficient, and MSC-corrected spectrum of the ith sample.

Savitzky–Golay (S-G estimates the ideal spectral value of the spectral data point by fitting or averaging the data points within

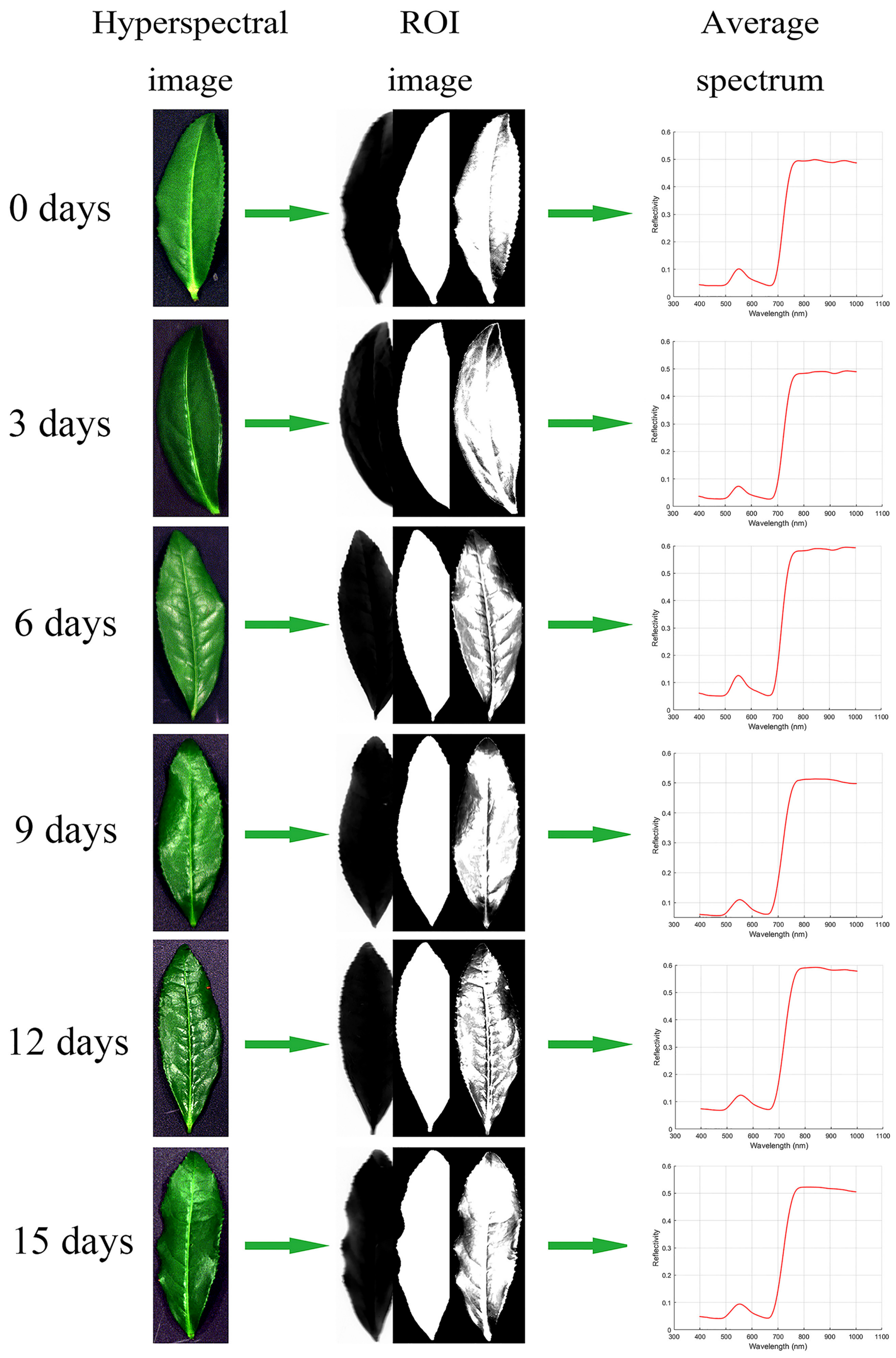


FIGURE 3 | Hyperspectral image processing flow of tea leaves: Hyperspectral image, ROI image (band math, segmentation image, masking), and average spectrum.

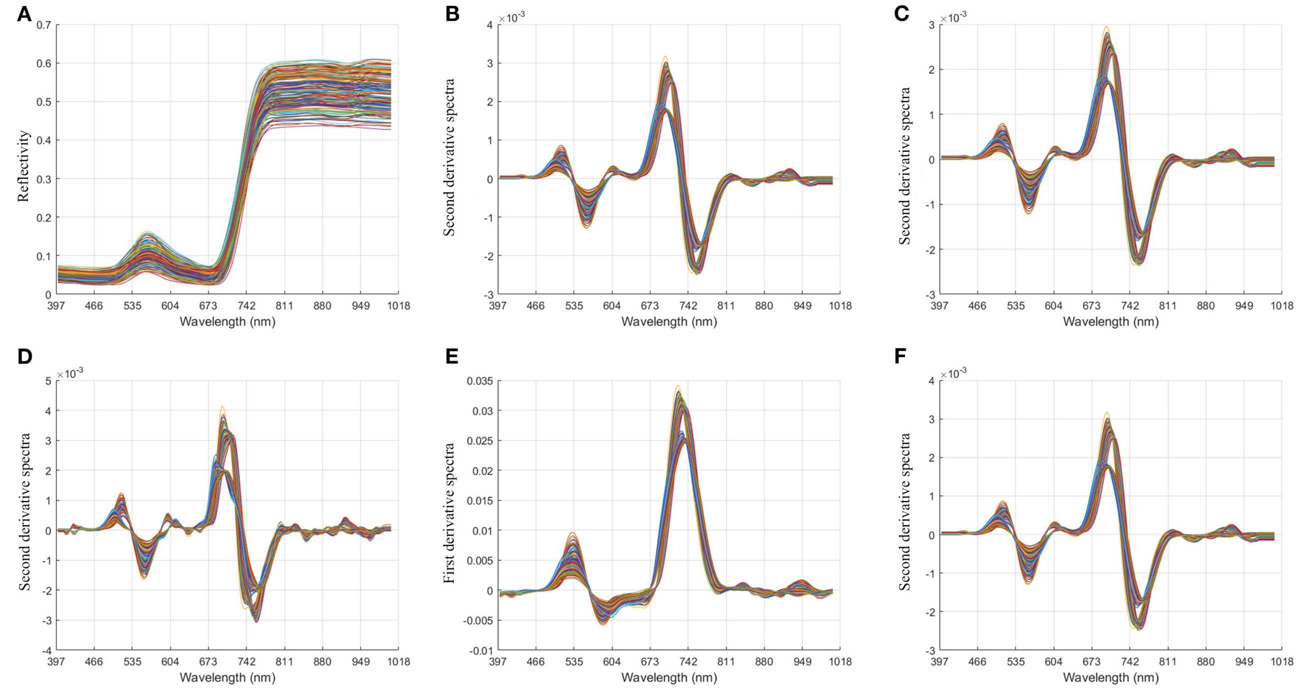


FIGURE 4 | Image comparison of unprocessed spectral data and preprocessed spectral data. **(A)** Original spectral data; **(B)**. multiplicative scatter correction + second derivative + Savitzky-Golay (17); **(C)**. multiplicative scatter correction + second derivative + Savitzky-Golay (19); **(D)**. multiplicative scatter correction+ second derivative + Savitzky-Golay (7); **(E)**. multiplicative scatter correction + first derivative + Savitzky-Golay (5); **(F)**. multiplicative scatter correction+ second derivative + Savitzky-Golay (17).

a certain size window range (the window width is generally odd) around the single-point spectral data, so as to reduce the interference of the irregular fluctuation noise signal in the spectral data to the data point and improve the signal-to-noise ratio of the spectral data. The formula of S-G smoothing algorithm is as follows:

$$X_i^* = \frac{\sum_{j=-r} X_i + W_j}{\sum_{j=-r}^{r} W_j}$$

where X_i^*, X_i is a spectral data point before and after S-G smoothing, and W_j is the weight factor obtained by smoothing the moving window with window width 2R + 1.

Derivative is mainly used for baseline correction and background interference removal of spectral data, so as to improve the resolution of spectral data. Due to the interference of different components of the sample and the experimental environment, the baseline shift (the position of the signal line changes) and the overlap of the spectral lines are directly caused. Therefore, the spectrum can be preprocessed by first derivative (1D) or second derivative (2D) to provide clearer spectral profile changes. However, when the original spectrum does not have a good signal-to-noise ratio, the derivative algorithm will also amplify the noise signal (Yan et al., 2001; Chu, 2004). The specific algorithm formula of the differential method is as follows:

$$\text{First derivative}: \frac{d_y}{d_\lambda} = \frac{y_{i+1} - y_i}{\Delta\lambda}$$

$$\text{Second derivative}: \frac{d_y^2}{d_{\lambda^2}} = \frac{y_{i+1} - 2y_i + y_{i-1}}{\Delta\lambda^2}$$

Model Accuracy Verification

The accuracy of the prediction model is measured by R^2, RMSE, and RPD. If R^2 is larger and RMSE is smaller, the accuracy of the model is higher and the model is more stable; otherwise, the accuracy of the model is lower and the model is more unstable (Cui et al., 2017). In addition, when RPD $\geq$ 2, it shows that the model has an excellent prediction ability. When $1.4 \leq \text{RPD} < 2$, it shows that the model can roughly estimate the sample, while RPD < 1.4 shows that the model cannot predict the sample (Yu et al., 2016).

RESULTS AND ANALYSIS

Significant Difference Analysis and Division of Modeling Sample Set

The drought-induced components of tea leaves were ranked according to time; the calibration set and prediction set of samples were selected according to the ratio of 3:1. The sample numbers of the calibration set and the prediction set are 135 and 45, respectively. The data distribution of the training set and

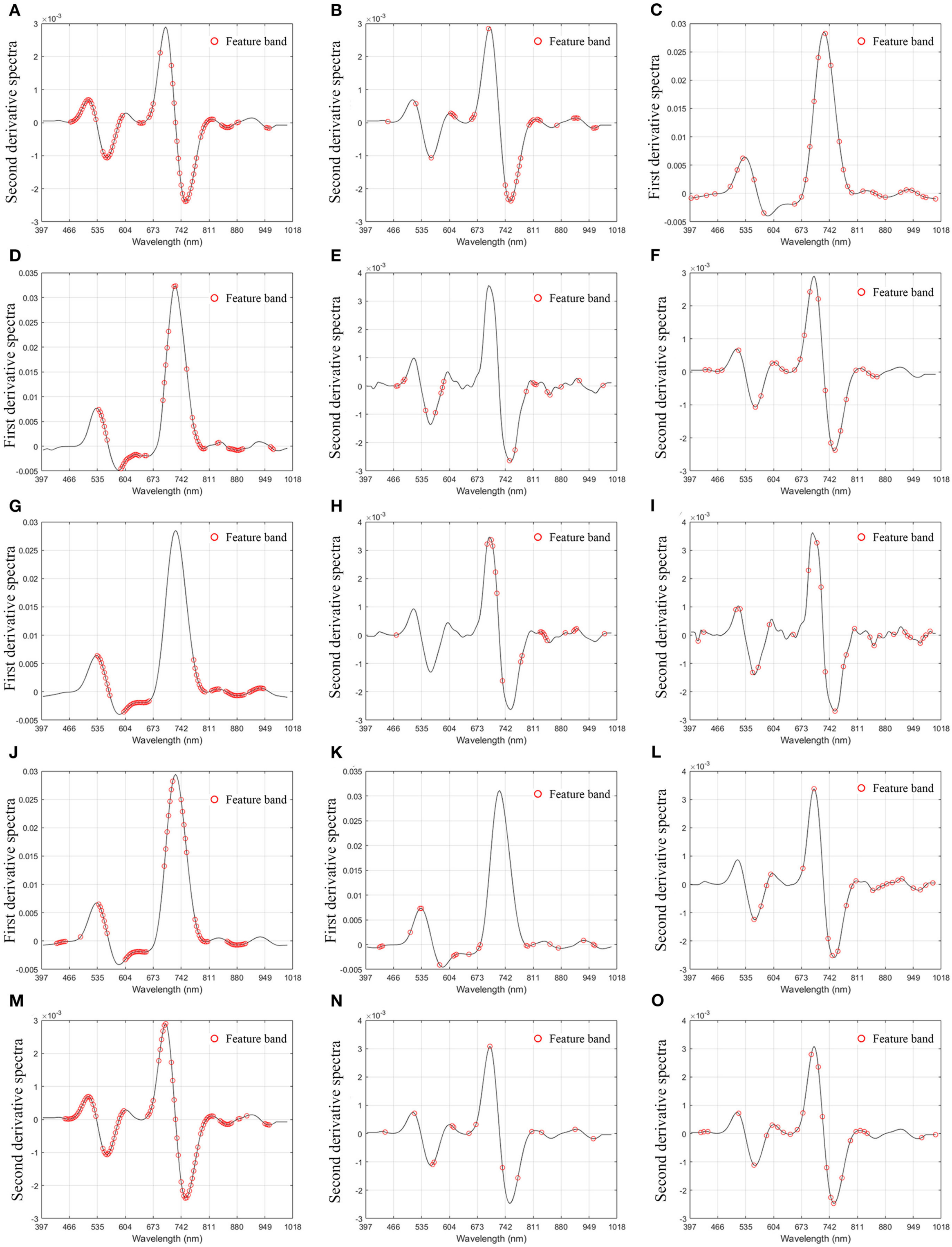


FIGURE 5 | Distribution of characteristic bands. **(A)** Malondialdehyde-uninformative variable elimination; **(B)**. Malondialdehyde-competitive adaptive reweighted sampling; **(C)**. Malondialdehyde-successive projections algorithm; **(D)**. Electrolyte leakage-uninformative variable elimination; **(E)**. Electrolyte leakage-competitive adaptive reweighted sampling; **(F)**. Electrolyte leakage-successive projections algorithm; **(G)**. *Fv/Fm*-uninformative variable elimination; **(H)**. *Fv/Fm*-competitive adaptive reweighted sampling; **(I)**. *Fv/Fm*-successive projections algorithm; **(J)**. Soluble saccharide-uninformative variable elimination; **(K)**. Soluble saccharide-competitive adaptive reweighted sampling; **(L)**. Soluble saccharide-successive projections algorithm; **(M)**. Drought damage degree-uninformative variable elimination; **(N)**. Drought damage degree-competitive adaptive reweighted sampling; **(O)**. Drought damage degree-successive projections algorithm.

TABLE 2 | Bands screening results.

Index	Screening method	Number of bands	Characteristic bands (nm)
MDA	UVE	85	466–535, 540–580, 730–760, 790–820, 830–870, 950
	CARS	36	450, 520, 600–620, 650–670, 740–780, 800, 920, 950
	SPA	33	400–460, 520, 550, 650, 670–690, 750, 810–880, 900–970
EL	UVE	57	530–550, 590–660, 690–730, 770–810, 850–910, 960
	CARS	20	460–490, 540, 560–590, 750, 790–820, 850, 880, 930
	SPA	26	430–470, 550, 600–680, 740, 750, 800–860
Fv/Fm	UVE	73	535–570, 600–670, 780–830, 840–920, 930–950
	CARS	20	460, 670, 700–740, 780, 820–850, 900–920
	SPA	27	400, 520, 540, 690, 750–810, 870–930, 960–980
SS	UVE	68	430–460, 530–570, 590–660, 690–750, 770–810, 850–910
	CARS	15	420–440, 500, 530, 580–620, 670, 810, 870, 950
	SPA	26	540–600, 670, 700, 750, 810, 850–930, 950–990
DDD	UVE	71	450–530, 540–600, 670–820, 830–870, 910, 950
	CARS	27	450, 520, 550, 600, 660, 700, 740, 810, 900, 950
	SPA	26	400–430, 520, 540, 590–670, 700–740, 810–840, 970

the prediction set is shown in **Supplementary Material**, mainly including maximum value, minimum value, average value, and standard deviation.

Preprocessing of Hyperspectral Data

In order to reduce the influence of the external environment and the dark current of the spectrometer, and reduce the baseline drift, light scattering, and other noises of the spectrum, we preprocessed the spectrum. In this paper, MSC, derivative (1D, 2D), and S-G technology are used to preprocess hyperspectral data (Tian et al., 2005; Zhao et al., 2005; Lu et al., 2019b). The spectral differences caused by different scattering levels are eliminated, and the correlation between spectra and data is enhanced. It can be seen from **Figure 4** that, through pretreatment, it is found that the peak valley of the spectral bands is obvious, avoiding the interference of overlapping peaks and improving the resolution and sensitivity of the spectrum.

Selection of Characteristic Wavelength

In order to improve the accuracy of the model and reduce the influence of noise and irrelevant bands, we screened 176 bands of spectral data. In this paper, three methods are used to select the characteristic bands: UVE, SPA, and CARS (Chen and Chen, 2005; Wu et al., 2009; Shi et al., 2018). The distribution of characteristic bands is shown in **Figure 5**. It can be seen from **Table 2** that in MDA-related characteristic bands screening method, the number of characteristic bands screened by UVE is the most, which is 85, and that by SPA was the least, 33. In the selection method of characteristic bands related to EC, the number of characteristic bands screened by UVE was the most, 57, and that by CARS was the least, 20. In the feature bands selection method related to *Fv/Fm*, the number of characteristic bands screened by UVE was the most, 73, and that by CARS was the least, 20. Among the methods for screening characteristic bands related to SS, the number of characteristic bands screened by UVE was the largest (68), and the number screened by CARS was the least (15). In the method of feature bands selection related to DG, the number of characteristic bands screened by UVE was the largest, which was 71, and that by SPA was the least, 26. It can be seen from **Table 3** that the optimal bands selection methods for MDA, EL, *Fv/Fm*, SS, and DDD models are MSC+2D+ S-G (17) +CARS, MSC+2D+S-G (19) + UVE, MSC+2D+S-G (7) +CARS, MSC+1D+S-G (5) +UVE, and MSC+2D+S–D (17) + UVE, respectively.

TABLE 3 | Optimal screening results.

Index	Optimal method	Rcal	RMSEC	Rp	RMSEP
MDA	MSC+2D+S-G (17) +CARS	0.96	0.36	0.92	0.46
EL	MSC+2D+S-G (19) + UVE	0.90	0.022	0.82	0.032
Fv/Fm	MSC+2D+S-G (7) +CARS	0.98	0.01	0.81	0.03
SS	MSC+1D+S-G (5) +UVE	0.87	0.09	0.87	0.69
DDD	MSC+2D+S-D (17) + UVE	0.98	0.28	0.95	0.32

Modeling and Analysis Based on Characteristic Bands

In order to establish the algorithm model of tea tree with different indexes, we use the feature vectors extracted by UVE, CARS, and SPA as the input variables of SVM, RF, and PLS models (Vapnik, 1998; Carrascal et al., 2010; Shao et al., 2012; Dong and Huang, 2013; Li, 2013; Zhou, 2016). **Table 4** shows the results of the validation of the model with prediction set samples; it can be seen from **Table 4** that, in MDA prediction, CARS-PLS model has the highest accuracy and SPA-RF model has the lowest accuracy. Among the models of MDA, EL, *Fv/Fm*, SS, and DDD, the models with the highest prediction accuracy are CARS-PLS, UVE-RF, CARS-SVM, UVE-PLS, and UVE-SVM respectively. The models

TABLE 4 | Modeling results.

Index	Modeling method	Rcal	RMSEC	RMSECV	Rp	RMECP	RPD
MDA	MSC+2D+S-G (17) + UVE+SVM	0.97	0.33	0.45	0.90	0.55	3.19
	MSC+1D+ S-G (15) +SPA+RF	0.96	0.34	0.36	0.91	0.54	3.01
	MSC+2D+ S-G (17) +CARS+PLS	0.96	0.36	0.38	0.92	0.46	3.51
EL	MSC+1D+S-G (21) +UVE+SVM	0.88	0.031	0.38	0.75	0.034	1.78
	MSC+2D+S-G (19) +UVE+RF	0.90	0.022	0.021	0.81	0.032	2.00
	MSC+2D+S-G (17) +SPA+PLS	0.88	0.11	0.034	0.76	0.035	1.90
Fv/Fm	MSC+2D+S-G (7) +CARS+SVM	0.98	0.01	0.02	0.81	0.03	2.29
	MSC+1D+S-G (7) +SPA+RF	0.94	0.017	0.021	0.83	0.027	2.15
	MSC+2D+S-G (5) +SPA+PLS	0.89	0.069	0.021	0.80	0.031	2.23
SS	MSC+1D+S-G (13) +UVE+ SVM	0.87	0.68	0.68	0.84	0.79	2.41
	MSC+1D+S-G (13) +SPA+RF	0.93	0.50	0.36	0.86	0.73	2.46
	MSC+1D+S-G (5) +UVE+PLS	0.87	0.09	0.71	0.87	0.69	2.72
DDD	MSC+2D+S-G (17) +UVE+SVM	0.97	0.28	0.021	0.95	0.32	4.28
	MSC+2D+S-G (15) +SPA+RF	0.96	0.29	0.29	0.92	0.40	3.27
	MSC+2D+S-G (15) +CARS+PLS	0.92	0.077	0.41	0.91	0.43	3.27

with the lowest accuracy were SPA-RF, UVE-SVM, SPA-RF, UVE-SVM, and CARS-PLS, respectively. Among the four physiological and biochemical indexes of MDA, EL, *Fv/Fm*, and SS, the CARS-PLS model of MDA had the best effect, and Rp, RMSEP, and RPD were 0.92, 0.46, and 3.51, respectively. The results showed that the UVE-SVM model of DDD index for the comprehensive evaluation of tea drought degree had the highest precision and the best effect, and Rp, RMSEP, and RPD were 0.95, 0.32, and 4.28, respectively. **Figure 6** shows the scatter distribution of the real value and the predicted value of the prediction sample set.

DISCUSSION

In this study, we found that the models of MDA, EL, *Fv/Fm*, and SS have a precise prediction ability in the inversion process of physiological/biochemical indexes and hyperspectral data of tea plants; these physiological and biochemical indexes are closely related to the drought state of tea plant, which has important physiological significance (Tian et al., 2019). Moreover, using MDA, SS, and EL to evaluate the stress degree of tea plants comprehensively can eliminate the deviation of single index evaluation to a certain extent, which is consistent with the conclusion of Liang et al. (2014). In this experiment, the estimation ability of the optimal models of MDA, EL, *Fv/Fm*, SS, and DDD all reached the expected effect ($Rp > 0.8$), which means that this method can quickly and non-destructively detect the drought state of tea plants.

The Optimization of Input Variables and Algorithms Is of Great Significance to Improve the Accuracy and Efficiency of Hyperspectral Data Inversion

First, in the selection of input datasets, a large number of previous studies used the vegetation index to evaluate stress (Wang et al., 2014a; Lu et al., 2019a). Due to the relatively small amount of information of vegetation index and the lack of stable vegetation index closely related to drought stress, the generalization ability of the final model may be reduced. Therefore, multialgorithm modeling analysis based on full bands is adopted in this experiment, which improves the accuracy of the model and makes the determination coefficients of the five models to evaluate the drought state of tea trees above 0.8, which proves the superiority of the experimental model.

In this experiment, we use a variety of feature extraction methods, including UVE, CARS, and SPA, to reduce redundant information and computing time, simplify data, and improve model accuracy. Then, the model of five indexes is established by machine learning method. The results showed that the optimization model had high precision and strong stability, which indicated that it was feasible to predict the physiological and biochemical indexes of tea and evaluate the drought status of tea by hyperspectral technology. Among them, the performance of UVE-SVM model of comprehensive index DDD ($Rcal = 0.97$, $RMSEC = 0.28$, $Rp = 0.95$, $RMSEP = 0.32$, $RPD = 4.28$) is better than that of other four physiological and biochemical indexes (MDA, EL, *Fv/Fm*, and SS), indicating that the method of combining multiple single indexes to evaluate plant drought status is better than a single index. Among the four physiological and biochemical indexes, CARS-PLS model had the highest prediction accuracy of MDA ($Rcal = 0.96$, $RMSEC = 0.36$, $Rp = 0.92$, $RMSEP = 0.46$, $RPD = 3.51$), which indicated that the relationship between MDA content and spectrum was more close than other physiological and biochemical indexes. It is expected that this model can be used to detect MDA content in tea seedlings, so as to evaluate the drought situation of tea plants. In the prediction models of EL, *Fv/Fm*, and SS, the RPD of the models was 2.72, 2.29, and 2.00 respectively, which was ≥ 2.00, indicating that the three

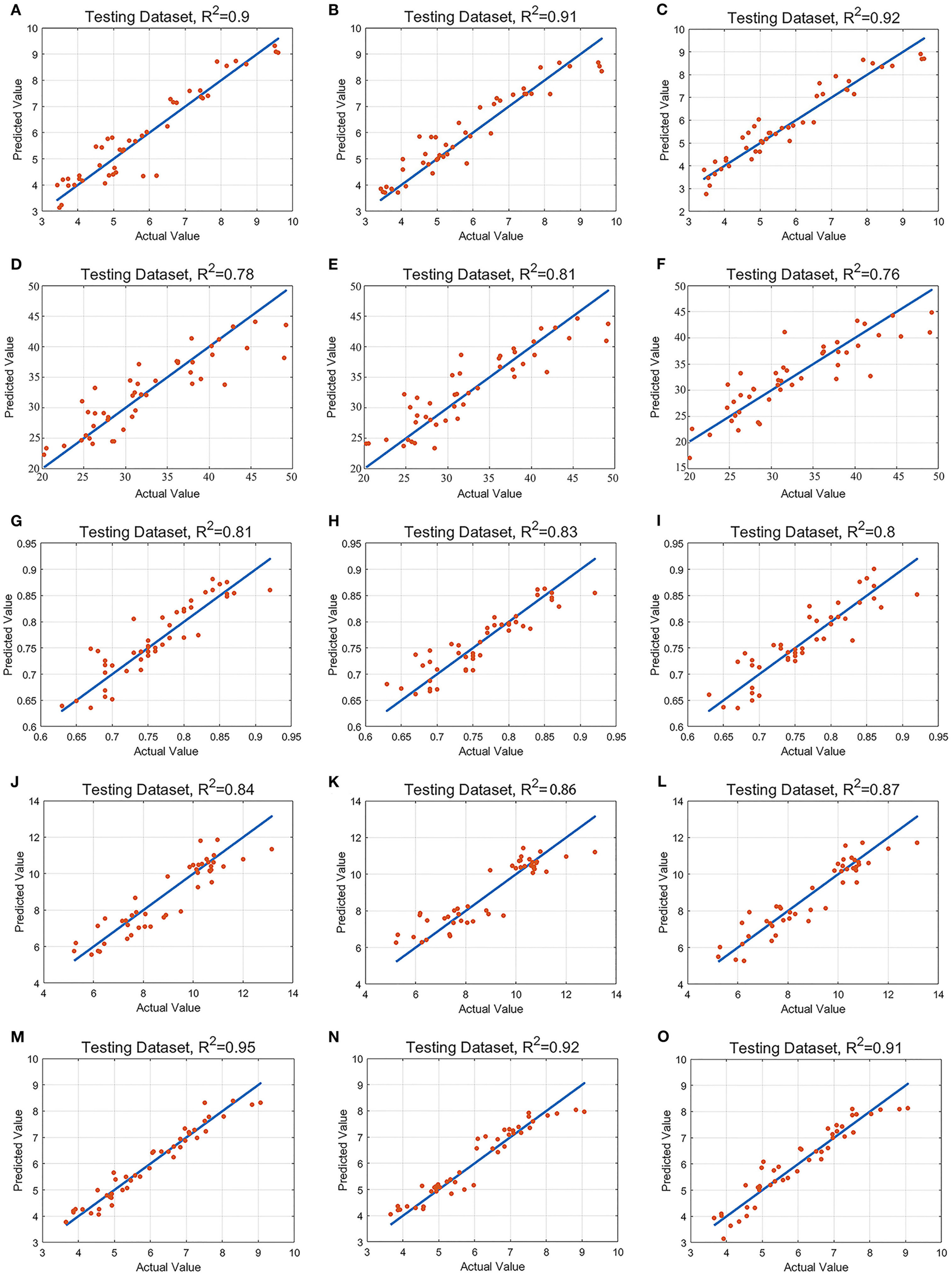


FIGURE 6 | Scatter plot of real and predicted values. **(A)**. Malondialdehyde-support vector machine; **(B)**. Malondialdehyde-random forest; **(C)**. Malondialdehyde-partial least-squares regression; **(D)**. Electrolyte leakage-support vector machine; **(E)**. Electrolyte leakage-random forest; **(F)**. Electrolyte leakage-partial least-squares regression; **(G)**. *Fv/Fm*-support vector machine; **(H)**. *Fv/Fm*-random forest; **(I)**. *Fv/Fm*-partial least-squares regression; **(J)**. Soluble saccharide-support vector machine; **(K)**. Soluble saccharide-random forest; **(L)**. Soluble saccharide-partial least-squares regression; **(M)** Drought damage degree-support vector machine; **(N)**. Drought damage degree-random forest; **(O)**. Drought damage degree-partial least-squares regression.

models had a good prediction ability and the stress state of young tea plants.

The Quantity and Quality of the Selected Characteristic Bands Have an Indirect Effect on the Model

Spectral data analysis needs to include a large number of samples, resulting in a large number of redundant data in the spectral matrix. And the original spectral data are prone to the phenomenon of spectral peak overlap, which leads to the slow speed and low efficiency of spectral analysis. In addition, the spectral matrix information unrelated to the sample detection index will have a great impact on the prediction accuracy of the model. Therefore, the performance of the prediction model can be improved by extracting characteristic wavelengths and removing redundant spectral variables from the collected spectral data.

In this experiment, we screened the variables to obtain a model with stronger generalization ability. In the screening results of MDA, EL, *Fv/Fm*, and SS, it was found that the spectral regions of 600–700, 700–780, and 800–850 nm appeared, and the positions of these peaks were closely related to the wavelength of vegetation index RENDVI and NDVI, which were also the two best indicators proposed by Kim et al. (2011) when they studied the response of plants to drought. In addition, in Kong et al.'s research (Kong et al., 2012), it is found that the characteristic bands of barley MDA are around 404 and 981 nm, and the selected characteristic bands are located at the two ends of the selected band range, in the visible and near-infrared regions, with large span and instability, and the possibility of noise interference is not excluded. In this experiment, the best characteristic band of MDA is 466–535, 540–580, 730–760, 790–820, 830–870, and 950 nm, which is different from the results of previous studies. The reason may be that with the increase of drought degree, MDA, as a product of plant peroxidation reaction, shows the increase of cell membrane permeability and respiration, which leads to the increase of reflectance in the visible region and the decrease of reflectance in the near-infrared region, thus increasing the characteristic bands (Soleimanzadeh, 2010). In research (Zhang et al., 2019), we found that the characteristic bands of conductivity of corn seeds were concentrated in the range of 400–600 and 760–1,000 nm. The screening results of UVE of EL in this experiment were 430–460, 530–570, 590–660, 690–740, 770–810, and 850–910 nm, which were similar to the previous research results; the optimal characteristic bands of *Fv/Fm* are 400, 520, 540, 690, 750–810, 870–930, and 960–980 nm. The reason may be that SPA algorithm chooses the variable combination with the least redundant information and the least collinearity, and the reflectance of near-infrared band in spectral data is quite different, which is different from previous research results. In the screening results of *Fv/Fm* characteristic bands, the corresponding bands (531 and 570 nm) of vegetation index PRI can be found, which is an effective index proposed by Wu and Niu (2008) in the study of plant photochemical vegetation index. In the visible light region of 400–700 nm, tea leaves absorbed a lot of visible light, but under drought stress, the photosynthesis of tea plants decreased, resulting in more visible light reflection and higher canopy original spectral reflectance. In the range of 700- to 1,000-nm near-infrared region, the spectral reflectance is greatly affected by the internal structure of leaves. Drought stress may lead to the disorder of internal tissue structure and rough cell wall of leaves (Mu et al., 2012), the complex leaf cavity structure scatters, and reflects near-infrared light many times, resulting in the decrease of spectral reflectance (Xu et al., 2017); in the visible light range, the utilization rate of light energy decreased and the reflectance of visible light increased, while *Fv/Fm* value and chlorophyll content could reflect the light utilization efficiency of plants. The SPA algorithm screening results of SS in this experiment were 540–600, 730, 750, 810, 850–930, and 950–990 nm. In the range of 560–719 nm, it is similar to the results of previous studies (Wang et al., 2018), but this experiment is different from previous studies in the near-infrared region. The reason may be that with the increase in SS concentration, the difference of near-infrared light reflection that leaves do not absorb becomes larger, so it is selected as the characteristic band by the algorithm.

The Algorithm Characteristics of the Model Determine the Correlation Between Hyperspectral Data and Drought Stress

Through the comparison of three modeling methods, it is found that the optimal models of different data are different, and the reason may be as follows: SVM model can make full use of the linear and non-linear information in the spectral data, but it is difficult to implement for the training set with a large amount of data. If a large part of the features of the data is lost, the RF can still maintain the accuracy, but cannot make predictions beyond the range of the training set data, which may lead to overfitting in the modeling of some specific noise data. PLS model can find the best function matching by minimizing the sum of squares of errors, but it can only use the linear information in spectral data.

In previous studies, it was found that LS-SVM was the best model for MDA content of barley under herbicide stress, and the determination coefficient of prediction set Rp = 0.84, but the RMSEC and RMSEP were 7.87 and 13.79, respectively (Kong et al., 2012), indicating that the degree of divergence of prediction results was too large. In this experiment, UVE-SVM is the best MDA model under drought stress, Rp = 0.9, RMSEP is only 0.55, which proves that this modeling method is better than LS-SVM model to some extent; In Zhang et al.'s research, MSC-GA-PLSR model was the best model for predicting the conductivity of sweet corn seeds (Zhang et al., 2019), with Rp = 0.97 and RMSEP = 0.226. In the experiment, CARS-RF model had the highest accuracy in this experiment, with Rp = 0.81 and RMSEP = 0.032. The CARS-RF model in this experiment is more stable than Zhang et al.'s GA-PLSR model; in this experiment, using a variety of algorithms and selecting the optimal model, the accuracy of *Fv/Fm* model of CARS-SVM (Rp = 0.81, RMSEP = 0.03) is higher than that of MASAVI2 model using vegetation index (Rp = 0.69, RMSEP = 8.6) (Zhao et al., 2011), and the stability is higher than that of full bands-PLS model (Rp = 0.83, RMSEP = 1.52) (Ding et al., 2015). According to Wang et al. 's research, the vegetation index DSI (D444, D455) was the best linear prediction model of SS in maize under drought stress (Wang et al., 2018). The coefficients of determination of D444 and D455 were Rp = 0.88 and Rp = 0.94, while those of RMSEP were 5.40 and 3.19, respectively. The results showed that the difference between the two models was large, which may be due to the limitations of their linear models and their limited ability to analyze complex hyperspectral data. In this experiment, the optimal SS-UVE-PLS model (Rp = 0.87, RMSEP = 0.69) is obtained through a variety of algorithms, and the anti-jamming ability and prediction accuracy are better than the former. In this experiment, the three models of comprehensive evaluation of tea drought damage have an excellent effect, among which the UVE-SVM model (Rcal = 0.97, Rp = 0.95, RPD = 4.28) is the best, which proves that the effect of the comprehensive evaluation model is better than the single physiological and biochemical index model.

CONCLUSION

In this experiment, we established the hyperspectral data models of five indexes related to drought evaluation by image segmentation, spectral preprocessing, and feature band selection. The results show that the best estimation models of the four physiological and biochemical indexes (MDA, EC, *Fv/Fm*, SS, DDD) were CARS-PLS, UVE-RF, SPA-RF, UVE-PLS, and UVE-SVM, respectively. The determination coefficients of the model prediction set were 0.92, 0.81, 0.83, 0.87, and 0.95, respectively. The models all achieve the expected results, and the prediction accuracy is very high. Among them, the model of DDD is better than the model of the four physiological and biochemical indexes, which can more comprehensively and objectively estimate the drought stress suffered by tea plants and effectively evaluate the drought resistance of tea plants.

Through the research and application of the models, the automatic irrigation of tea garden can be realized, the water-use efficiency of tea garden can be improved, and it is of great significance for water saving and consumption reduction. At the same time, this study is expected to be used to evaluate the drought resistance of different tea varieties, so as to screen out drought-resistant tea varieties.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SC carried out the experiment, collected and organized data, processed the hyperspectral image of tea leaves, and wrote the manuscript. YG and KF participated in designing the experiment and reviewed the manuscript. ZD and YW raised the hypothesis underlying this work, designed the experiment, and helped organize the manuscript structure and directed the study. YS and DL participated in designing the experiment and directed the study. All authors contributed to the article and approved the submitted version.

FUNDING

This research was subsidized by the Significant Application Projects of Agriculture Technology Innovation in Shandong Province (SD2019ZZ010), the Technology System of Modern Agricultural Industry in Shandong Province(SDAIT-19-01), and the Special Foundation for Distinguished Taishan Scholar of Shandong Province (No.ts201712057), the Livelihood Project of Qingdao City (19-6-1-64-nsh),the Project of Agricultural Science and Technology Fund in Shandong Province (2019LY002, 2019YQ010, 2019TSLH0802), Project of improved agricultural varieties in Shandong Province (2020LZGC010).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.695102/full#supplementary-material

REFERENCES

Araújo, M., Saldanha, T. C., Galvo, R. K. H., Yoneyama, T., and Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometr. Intell. Lab. Syst.* 57, 65–73. doi: 10.1016/S0169-7439(01)00119-8

Carrascal, L. M., Galván, I., and Gordo, O. (2010). Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* 118, 681–690. doi: 10.1111/j.1600-0706.2008.16881.x

Chen, B., and Chen, D. (2005). The application of uninformative variables elimination in near-infrared spectroscopy. *Spectron. Instr. Anal.* 4, 26–30.

Cheng, H., Shen, R., Chen, Y., Wan, Q., Shi, T., Wang, J., et al. (2019). Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy. *Geoderma* 336, 59–67. doi: 10.1016/j.geoderma.2018.08.010

Cheruiyot, E. K., Mumera, L. M., Ng'etich, W. K., Hassanali, A., and Wachira, F. N. (2010). High fertilizer rates increase susceptibility of tea to water stress. *J. Plant Nutr.* 33, 115–129. doi: 10.1080/01904160903392659

Chu, X. (2004). Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique. *Prog. Chem.* 16, 528–542. doi: 10.1016/j.jco.2003.08.015

Cui, X., Song, Q. J., Zhang, Y. Y., Xu, G., and Gao, J. L. (2017). Estimation of soil organic carbon content in alpine grassland using hyperspectral data. *Acta Pratacult. Sinica* 26, 20–29. doi: 10.11686/cyxb2016509

Ding, X. B., Fei, L., Chu, Z., and Yong, H. (2015). Prediction of SPAD Value in Oilseed Rape Leaves Using Hyperspectral Imaging Technique. *Guang Pu Xue Yu Guang Pu Fen Xi* 35:486. doi: 10.3964/j.issn.1000-0593(2015)02-0486-06

Dong, S. S., and Huang, Z. X. (2013). A brief theoretical overview of random forests. *J. Integr. Tech.* 2, 1–7.

Duan, H. W. (2016). *Research on Rapid Detection of Mutton pH and Total Viable Count(TVC) Based on Hyperspectral Imaging Technique*. Shihezi University.

Filho, H., Galvo, R., Araújo, M., Silva, E., and Rohwedder, J. (2004). A strategy for selecting calibration samples for multivariate modeling. *Chemometr. Intell. Lab. Syst.* 72, 83–91. doi: 10.1016/j.chemolab.2004.02.008

Guo, C. F., Sun, Y., and Zhang, M. Q. (2008). Photosynthetic characteristics and water use efficiency of tea plant under different soil moisture condition. *J. Fujian Coll. Forestr.* 28, 333–337.

Guo, Y., Zhao, S., Zhu, C., Chang, X., Yue, C., Wang, Z., et al. (2017). Identification of drought-responsive miRNAs and physiological characterization of tea plant (Camellia sinensis L.) under drought stress. *BMC Plant Biol.* 17:211. doi: 10.1186/s12870-017-1172-6

Iverson, L. R., Prasad, A. M., Matthews, S. N., and Peters, M. (2008). Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *For. Ecol. Manage.* 254, 390–406. doi: 10.1016/j.foreco.2007.07.023

Jiang, W., Fang, J., Wang, S., and Wang, R. (2016). Using CARS-SPA algorithm combined with hyperspectral to determine reducing sugars content in potatoes. *J. Northeast Agri. Univ.* 9:40. doi: 10.14257/ijhit.2016.9.9.04

Kate, M., and Johnson, G. N. (2000). Chlorophyll fluorescence—a practical guide. *J. Exp. Bot.* 51, 659–668. doi: 10.1093/jexbot/51.345.659

Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., and Lehman, B. L. (2011). Hyperspectral image analysis for water stress detection of apple trees. *Comput. Electron. Agric.* 77, 155–160. doi: 10.1016/j.compag.2011.04.008

Kong, W., Liu, F., Zhang, C., Zhang, J., and Feng, H. (2016). Non-destructive determination of Malondialdehyde (MDA) distribution in oilseed rape leaves by laboratory scale NIR hyperspectral imaging. *Sci. Rep.* 6:35393. doi: 10.1038/srep35393

Kong, W. W., Liu, F., and Fang, H. (2012). Rapid detection of malondialdehyde in herbicide-stressed barley leaves using spectroscopic techniques. *Trans. Chinese Soc. Agric. Eng.* 28, 171–175. doi: 10.3969/j.issn.1002-6819.2012.02.030

Li, P., Zhou, J., Jiang, L., Liu, X., and Du, G. (2019a). A variable selection approach of near infrared spectra based on window competitive adaptive reweighted sampling strategy. *Spectro. Spectral Anal.* 39, 106–110. doi: 10.3964/j.issn.1000-0593(2019)05-1428-05

Li, R. L., Xie, Y. D., and Tang, Y. (2019b). Effect of application accumulator plant straw on the osmotic adjustment substances and malondialdehyde content of lettuce. *IOP Conference Series: Earth Environ. Sci.* 233, 42025–42025. doi: 10.1088/1755-1315/233/4/042025

Li, X. H. (2013). Using "random forest" for classification and regression. *Chin. J. Appl. Entomol.* 50, 1190–1197. doi: 10.7679/j.issn.2095-1353.2013.163

Liang, X. D., Zeng, C. W., and Li, J. J. (2014). Evaluation and selection of drought-resistance of oat varieties. *Xinjiang Agric. Sci.* 51, 2150–2155. doi: 10.6048/j.issn.1001-4330.2014.11.031

Lin, D., Shi, S., Jian, Y., Sun, J., and Gong, W. (2016). Using different regression methods to estimate leaf nitrogen content in rice by fusing hyperspectral LiDAR Data and laser-induced chlorophyll fluorescence data. *Remote Sens. Environ.* 8, 526–526. doi: 10.3390/rs8060526

Liu, S. C., and Chen, L. (2014). Research advances on the drought-resistance mechanism and strategy of tea plant. *J. Tea Sci.* 34, 111–121. doi: 10.3969/j.issn.1000-369X.2014.02.002

Lu, B., Sun, J., Yang, N., Wu, X. H., and Zhou, X. (2019a). Prediction of Tea Diseases Based on Fluorescence Transmission Spectrum and Texture of Hyperspectral Image. *Spectro. Spect. Anal.* 39, 2515–2521. doi: 10.3964/j.issn.1000-0593(2019)08-2515-07

Lu, Y. B., Liu, W. Q., Zhang, Y. J., Zhang, K., Ying, H. E., You, K., et al. (2019b). An adaptive hierarchical savitzky-golay spectral filtering algorithm and its application. *Spectro. Spect. Anal.* 39, 2657–2663.

Morales, M., and Munné-Bosch, S. (2019). Malondialdehyde: facts and artifacts. *Plant Physiol. Commun.* 180, 1246–1250. doi: 10.1104/pp.19.00405

Mu, Y. C., Wang, R., Sun, W. T., Gong, L., Wang, Y. Q., and Li, X. W. (2012). Effect of water stress on chloroplast ultrastructure of maize. *Liaoning Agric. Sci.* 7–12. doi: 10.3969/j.issn.1002-1728.2012.05.002

Prieto, P., Llusia, A., and Estiarte A. (2009). Effects of long-term experimental night-time warming and drought on photosynthesis, Fv/Fm and stomatal conductance in the dominant species of a Mediterranean shrubland. *Acta Physiol. Plant* 31, 729–739. doi: 10.1007/s11738-009-0285-4

Qin, J., and He, Z. S. (2005). "A SVM face recognition method based on Gabor-featured key points," *inMachine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference*. doi: 10.1109/ICMLC.2005.1527850

Shao, X., Min, Z., and Cai, W. (2012). Multivariate calibration of near-infrared spectra by using influential variables. *Anal. Methods* 4, 467–473. doi: 10.1039/c2ay05609g

Sharma, P., and Kumar, S. (2005). Differential display-mediated identification of three drought-responsive expressed sequence tags in tea [Camellia sinensis (L.) O. *Kuntze]. J. Biosci.* 30, 231–235. doi: 10.1007/BF02703703

Shi, T., Liu, H., Wang, J., Chen, Y., Fei, T., and Wu, G. (2014). Monitoring arsenic contamination in agricultural soils with reflectance spectroscopy of rice plants. *Environ. Sci. Technol.* 48, 6264–6272. doi: 10.1021/es405361n

Shi, Y., Sun, D. M., Xiong, J., Wei, F., and Ma, S. C. (2018). Analysis of artificial Cow-Bezoar by near-infrared spectroscopy coupled with competitive adaptive reweighted sampling method. *Chin. Pharm. J.* 53, 1216–1221.

Soleimanzadeh (2010). Effect of Potassium Levels on Antioxidant Enzymes and Malondialdehyde Content under Drought Stress in Sunflower (Helianthus annuus L.). *Am. J. Agric. Biol. Sci.* 5:61. doi: 10.3844/ajabssp.2010.56.61

Takashima, K., Nor, A., Ando, S., Takahashi, H., and Kaneko, T. (2021). Evaluation of plant stress due to plasma-generated reactive oxygen and nitrogen species using electrolyte leakage. *Jpn. J. Appl. Phys.* 60:010504. doi: 10.35848/1347-4065/abcf5b

Talens, P., Mora, L., Morsy, N., Barbin, D. F., Elmasry, G., and Sun, D. W. (2013). Prediction of water and protein contents and quality classification of Spanish cooked ham using NIR hyperspectral imaging. *J. Food Eng.* 117, 272–280. doi: 10.1016/j.jfoodeng.2013.03.014

Tian, G. Y., Yuan, H. F., Chu, X. L., and Liu, H. Y. (2005). Near Infrared Spectra (NIR) analysis of octane number by waveletdenoising-derivative method. *Spectrosc. Spectral Anal.* 25:516. doi: 10.1016/j.saa.2004.06.052

Tian, S., Guo, R., Zou, X., Zhang, X., and Si, T. (2019). Priming with the green leaf volatile Z-3-Hexenyl Acetate enhances salinity stress tolerance in peanut (Arachis hypogaea L.) seedlings. *Front. Plant Sci.* 10:785. doi: 10.3389/fpls.2019.00785

Vapnik, V. N. (1998). Statistical learning theory. *Encyclopedia Sci. Learn.* 41, 3185–3185. doi: 10.1007/978-1-4419-1428-6_5864

Wang, X. P., Zhao, C. Y., Guo, N., Yao-Hui, L. I., Jian, S. Q., and Kai, Y. U. (2014a). Canopy hyperspectral reflectance response for spring wheat in different water stresses in semi-arid areas of the Loess Plateau. *J. Lanzhou Univ.* 50, 417–423. doi: 10.13885/j.issn.0455-2059.2014.03.019

Wang, X. P., Zhao, C. Y., Guo, N., Yao-Hui, L. I., Jian, S. Q., and Kai, Y. U. (2014b). Canopy hyperspectral reflectance response for spring wheat in different water stresses in semi-arid areas of the Loess Plateau. *J. Lanzhou Univ.* 50, 417–423.

Wang, Z. L., Chen, J. X., Cheng, Y. J., Fan, Y. F., Feng, W., Hao-Chen, L. I., et al. (2018). Assessing the soluble sugar of maize leaves in drought stress based on hyperspectral data. *J. Sichuan Agric. Univ.* 36, 436–443. doi: 10.16036/j.issn.1000-2650.2018.04.003

Wu, C. Y., and Niu, Z. (2008). Review of retrieval light use efficiency using photochemical reflectance index (PRI). *J. Plant Ecol.* 32, 734–740. doi: 10.3773/j.issn.1005-264x.2008.03.023

Wu, D., Wu, H., X., Cai, J., B., Huang, J., et al. (2009). Classifying the species of exopalaemon by using visible and near infrared spectra with uninformative variable elimination and successive projections algorithm. *J. Infra. Millimeter Waves* 28, 423–427. doi: 10.3724/SP.J.1010.2009.00423

Xu, D. Q., Liu, X. L., Wang, W., Chen, M., Kan, H. C., Li, C. F., et al. (2017). Hyper-spectral characteristics and estimation model of leaf chlorophyll content in cotton under waterlogging stress. *Chinese J. Appl. Ecol.* 28, 3289–3296. doi: 10.13287/j.1001-9332.201710.013

Yan, Y. L., Zhao, L. L., Li, J. H., Zhang, L. D., and Min, S. G. (2001). Information technology of modern NIR spectral analysis. *Guang Pu Xue Yu Guang Pu Fen Xi* 20, 777–780.

Yu, L., Hong, Y. S., Zhou, Y., Zhu, Q., Xu, L., Li, J. Y., et al. (2016). Wavelength variable selection methods for estimation of soil organic matter content using hyperspectral technique. *Trans. Chinese Soc. Agric. Eng.* 32, 95–102. doi: 10.11975/j.issn.1002-6819.2016.13.014

Zelazny, W. R., and Lukáš, J. (2020). drought stress detection in juvenile oilseed rape using hyperspectral imaging with a focus on spectra variability. *Remote Sens. Environ.* 12:3462. doi: 10.3390/rs12203462

Zhang, Q., Ming, X., and Qi, K. (2010). Region selecting methods of near infrared wavelength based on uninformative variables elimination. *J. Agric. Mechan. Res.* 11, 202–205. doi: 10.3969/j.issn.1003-188X.2010.11.049

Zhang, T. T., Zhao, B., Yang, L. M., Wang, J. H., Sun, Q., and Science, C. O. (2019). Determination of Conductivity in Sweet Corn Seeds with Algorithm of GA and SPA Based on Hyperspectral Imaging Technique. *Spectrosc. Spectral Anal.* 39, 2608–2613. doi: 10.3964/j.issn.1000-0593(2019)08-2608-06

Zhang, Z. A., Yang, F., Chen, Z. Y., and Ke-Zhang, X. U. (2006). Relationship Between Diurnal Changes of Net Photosynthetic Rate and Environmental Factors in Leaves of Zizania latifolia. *Scientia Agric. Sinica* 39, 502–509. doi: 10.3321/j.issn:0578-1752.2006.03.010

Zhao, J. W., Wang, K. L., Ouyang, Q., and Chen, Q. S. (2011). Measurement of Chlorophyll Content and Distribution in Tea Plant's Leaf Using Hyperspectral Imaging Technique. *Spectrosc Spect. Anal.* 31, 512–515. doi: 10.3964/j.issn.1000-0593(2011)02-0512-04

Zhao, Q., Zhang, G. L., and Chen, X. D. (2005). Effects of multiplicative scatter correction on a calibration model of near infrared spectral analysis. *Optics Precis. Eng.* 13, 53–58. doi: 10.1088/1009-0630/7/5/006

Zhou, W., and Leul, M. (1999). Uniconazole-induced tolerance of rape plants to heat stress in relation to changes in hormonal levels, enzyme activities and lipid peroxidation. *Plant Growth Regul.* 27, 99–104. doi: 10.1023/A:1006165603300

Zhou, Z. H. (2016). *Machine Learning*. Beijing: Tsinghua University Press.

Zovko, M., Žibrat, U., Knapič, M., Kovačić, M., and Romić, D. (2019). Hyperspectral remote sensing of grapevine drought stress. *Precis. Agric.* 12:12. doi: 10.1007/s11119-019-09640-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Chen, Gao, Fan, Shi, Luo, Shen, Ding and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership