

## Hyperspectral Imaging Inversion Method for Chlorophyll and Water Content in Maize Leaves Under Drought Stress (Postprint)

**Authors:** Wang Jingyong, Zhang Mingzhen, Ling Huarong, Wang Ziting, Gai Jingyao

**Date:** 2023-12-04T00:00:00+00:00

### Abstract

[Objective/Significance] To achieve non-destructive monitoring of crop growth status such as drought stress in corn, this study explores a non-destructive detection method for chlorophyll content and water content in corn leaves under drought stress based on hyperspectral technology. [Methods] First, a hyperspectral camera was used to collect images of seedling-stage corn leaves under different drought stress levels, and image processing techniques were employed to extract the average spectrum of the mesophyll portion. By systematically analyzing the effects of different characteristic wavelength extraction methods and machine learning regression models on the prediction performance of chlorophyll content and water content, optimal inversion models for chlorophyll content and water content were established respectively, and the construction of vegetation coefficients applicable to the inversion of chlorophyll content and water content was investigated and their inversion capability was evaluated. [Results and Discussion] Combining Stepwise Regression (SR) feature extraction with Stacking regression achieved the optimal prediction effect for chlorophyll content ( $R^2$  of 0.878, root mean square error of 0.317 mg/g); combining Successive Projections Algorithm (SPA) feature extraction with Stacking regression achieved the optimal prediction effect for water content ( $R^2$  of 0.859, RMSE of 3.75%); the newly constructed Normalized Difference Vegetation Index  $[(R410-R559)/(R410+R559)]$  and ratio coefficient  $(R400/R1171)$  achieved the highest inversion accuracy for chlorophyll content and water content respectively, significantly higher than traditional vegetation coefficients, with  $R^2$  values of 0.803 and 0.827 respectively, and root mean square errors of 0.403 mg/g and 3.28% respectively. [Conclusion] The inversion models and vegetation coefficients constructed in this study based on hyperspectral information can achieve accurate and non-destructive detection of chlorophyll content and water content in corn

leaves, providing theoretical basis and technical support for real-time monitoring of corn growth status.

## Full Text

# A Hyperspectral Image-Based Method for Estimating Water and Chlorophyll Contents in Maize Leaves under Drought Stress

WANG Jingyong<sup>1</sup>, ZHANG Mingzhen<sup>1</sup>, LING Huarong<sup>2</sup>, WANG Ziting<sup>2,3</sup>, GAI Jingyao<sup>1\*</sup>

<sup>1</sup>School of Mechanical Engineering, Guangxi University, Nanning 530004, China

<sup>2</sup>College of Agriculture, Guangxi University, Nanning 530004, China

<sup>3</sup>Guangxi Key Laboratory of Sugarcane Biology, Guangxi University, Nanning 530004, China

---

## Abstract

**[Objectives]** To achieve non-destructive monitoring of crop growth status such as drought stress in maize, this study explores a non-destructive detection method for maize leaf chlorophyll content and water content under drought stress based on hyperspectral technology. **[Methods]** First, hyperspectral cameras were used to collect images of maize leaves at the seedling stage under different drought stress levels, and image processing techniques were employed to extract the average spectrum of the leaf mesophyll region. By systematically analyzing the effects of different characteristic wavelength extraction methods and machine learning regression models on the prediction performance of chlorophyll content and water content, optimal inversion models were established for each parameter. Additionally, this study investigated the construction of vegetation indices for chlorophyll and water content inversion and evaluated their performance. **[Results and Discussion]** Combining stepwise regression (SR) feature extraction with Stacking regression yielded optimal chlorophyll content prediction performance ( $R^2$  of 0.878, RMSE of 0.317 mg/g). Combining successive projections algorithm (SPA) feature extraction with Stacking regression achieved optimal water content prediction performance ( $R^2$  of 0.859, RMSE of 3.75%). The newly constructed normalized difference vegetation index  $[(R_{410}-R_{559})/(R_{410}+R_{559})]$  and ratio index  $(R_{400}/R_{1171})$  achieved the highest inversion accuracy for chlorophyll and water content, respectively, significantly outperforming traditional vegetation indices with  $R^2$  values of 0.803 and 0.827, and RMSE values of 0.403 mg/g and 3.28%. **[Conclusions]** The inversion models and vegetation indices constructed in this study enable accurate, non-destructive detection of maize leaf chlorophyll and water content, providing theoretical basis and technical support for real-time monitoring of maize growth status.

**Keywords:** drought stress; hyperspectral technology; chlorophyll content inversion; water content inversion; machine learning

---

## 1. Introduction

Drought stress during the seedling stage of maize (*Zea mays* L.) severely affects crop quality and yield [?]. Rapid diagnosis of drought stress severity is a prerequisite for timely implementation of mitigation measures. Although traditional methods such as field surveys of crop appearance and detection of biological indicators can accurately assess crop growth status, they require substantial labor and material resources and are unsuitable for large-scale online monitoring of drought stress [?, ?]. Chlorophyll content and water content in maize leaves are important physiological indicators that reflect the degree of drought stress. Therefore, rapid and non-destructive detection of these parameters promises to enable large-scale online monitoring of drought stress in maize fields for precision agriculture, thereby safeguarding maize quality and yield.

Hyperspectral imaging technology emerged in the 1990s, combining spectral and imaging techniques to offer advantages of broad band coverage, complete spatial information, and integrated spectral-spatial data. This technology enables rapid, non-destructive detection of both external characteristics and internal composition of objects. Currently, extensive research on crop physiological information inversion based on hyperspectral imaging technology has been conducted both domestically and internationally. Some studies have collected hyperspectral images of leaves in the laboratory, extracted average spectra from leaf regions, and established physiological information inversion models after preprocessing and feature extraction. For example, Sun et al. [?] compared different preprocessing methods and established a detection model for nitrogen content in lettuce leaves using Partial Least Squares Regression (PLSR). Gao et al. [?] used correlation analysis and the random frog algorithm to extract characteristic bands, then employed PLSR to build a maize chlorophyll content estimation model. Other studies have used portable, vehicle-mounted, or UAV-based hyperspectral platforms to collect images at the plant or field scale for establishing population nutrient prediction models. For instance, Tan et al. [?] used a UAV platform to capture images of wheat under nitrogen stress at different growth stages, extracted spectral features by fusing Support Vector Regression with swarm optimization algorithms, and constructed a nitrogen content prediction model based on PLSR. Shu et al. [?] constructed and compared multiple vegetation indices for assessing water status in maize plants. However, most existing methods employ specific preprocessing and modeling approaches to build crop physiological information inversion models, while systematic and comprehensive comparative analyses of different preprocessing, feature extraction, and modeling methods and their combinations are relatively scarce, particularly for systematic analysis of physiological information inversion methods for maize seedlings under drought stress.

This study focuses on maize plants under drought stress, systematically analyzing the effects of different preprocessing methods, characteristic band extraction methods, and regression models on the prediction performance of water content and chlorophyll content. The objectives are to establish optimal inversion models for leaf chlorophyll and water content based on hyperspectral imaging technology, construct vegetation indices applicable to chlorophyll and water content inversion, and evaluate their performance. This research can provide technical support for rapid acquisition of maize growth information, aiming to achieve refined management of maize cultivation and improve yield and quality.

---

## 2. Materials and Methods

**2.1 Experimental Design and Sample Collection** The experiment was conducted in September 2022 in a greenhouse at the College of Agriculture, Guangxi University. The maize variety was “Zaoshengnuoyu 808.” Drought stress treatment was applied to maize plants at the seedling stage (four-leaf stage). Four drought gradients were established: normal water treatment  $((75 \pm 2) \pm 2) \pm 2 \times 4 \times 4 \times 3 \times 2 = 288$  maize leaf samples were collected.

**2.2 Hyperspectral Image Acquisition** The hyperspectral imaging system used in this study consisted of two push-broom hyperspectral cameras (Specim FX10 and Specim FX17, Spectral Imaging Ltd., Oulu, Finland) with parallel light sources, a camera scanning motion control system, and a computer (Fig. 1 [Figure 1: see original paper]). The main technical parameters of the hyperspectral cameras are shown in Table 1. Before image acquisition, standard white and black reference panels were used for calibration. Image acquisition was performed in a dark environment (ambient illuminance  $< 5$  lux) with consistent lighting conditions maintained throughout the process. Each leaf sample was placed in a sample holder with a black background, and the camera was driven by a motor to scan at constant speed to collect hyperspectral images. The 400-1000 nm band from the FX10 camera and the 1000-1700 nm band from the FX17 camera were extracted and co-registered through post-processing to form a single hyperspectral image, generating a total of 288 maize leaf spectral images.

**2.3 Measurement of Chlorophyll and Water Content** After hyperspectral image acquisition, chlorophyll *a* and *b* contents in the leaves were measured using standard methods [?]. For each leaf, 0.1 g of tissue was quickly cut from the area avoiding leaf veins, crushed, and placed in a test tube containing 25 mL of extraction solution (acetone and anhydrous ethanol at a 2:1 volume ratio). The tubes were placed in a dark cabinet for 24 hours to allow chlorophyll extraction and separation. After the leaves turned completely white, the extraction solution was transferred to a cuvette and measured using a spectrophotometer

(UV-1800, Shimadzu Corporation, Kyoto, Japan) to determine absorbance at 663 nm and 645 nm. Chlorophyll content was calculated using equations (1)-(3):

$$C_a = 12.72A_{663} - 2.59A_{645}$$

$$C_b = 22.88A_{645} - 4.67A_{663}$$

$$C = C_a + C_b = 20.29A_{645} + 8.05A_{663}$$

where  $A_{663}$  and  $A_{645}$  are the absorbance values of the extraction solution at 663 nm and 645 nm, respectively;  $C_a$  and  $C_b$  are the contents of chlorophyll *a* and *b* in the solution (mg/L); and  $C$  is the total chlorophyll content in the solution (mg/L).

The chlorophyll content in the solution was converted to leaf chlorophyll content *Chl* (mg/g) using equation (4):

$$Chl = \frac{V \times C}{1000 \times m}$$

where  $V$  is the volume of extraction solution (L) and  $m$  is the leaf mass (mg).

For water content measurement, leaf veins were removed and the fresh weight ( $W_F$ ) of the leaf was measured using an electronic balance. The leaf was then placed in distilled water for 2 hours to reach full turgor, removed, blotted dry, and weighed again to obtain the saturated weight ( $W_S$ ). Finally, the leaf was placed in an oven at 105°C for 0.5 hours to deactivate enzymes, then dried at 80°C to constant weight to obtain the dry weight ( $W_D$ ). Leaf (relative) water content (LWC) was calculated using equation (5):

$$LWC = \frac{W_F - W_D}{W_S - W_D} \times 100\%$$

**2.4 Data Processing Methods** The spectra of leaf mesophyll, veins, and background regions in the collected hyperspectral images showed significant differences (Fig. 2 [Figure 2: see original paper]). The background spectra had overall reflectance below 0.1. The mesophyll spectra exhibited typical plant characteristics with green peaks, red valleys, and red-edge features. Vein reflectance was higher than mesophyll in the 400-1300 nm range but lower in the 1300-1700 nm range. Since chlorophyll content in leaf veins is very low, this study performed segmentation of hyperspectral images to remove veins and background, retaining only the mesophyll region for spectral information extraction.

Significant spectral reflectance differences among background, mesophyll, and veins were concentrated at 550 nm and 1100 nm bands. Therefore, when extracting reflectance images at these two bands, the Otsu method was used to segment veins, mesophyll, and background, preserving the mesophyll region (Fig. 2(c)), from which pixels were extracted and the average mesophyll spectrum was calculated.

This study adopted an algorithm framework of “preprocessing - feature extraction - machine learning inversion” (Fig. 3 [Figure 3: see original paper]). First, spectral preprocessing was applied to eliminate noise caused by environmental and instrumental factors and highlight useful spectral information. The preprocessing methods analyzed included Savitzky-Golay smoothing (SG) [?], Multiplicative Scatter Correction (MSC) [?], Standard Normal Variate (SNV) [?], First Derivative (FD), and Second Derivative (SD).

Second, feature extraction and dimensionality reduction were performed on preprocessed spectra to reduce redundant information and prevent overfitting. Spectral dimensionality reduction and feature extraction methods included characteristic wavelength extraction and vegetation index methods. Characteristic wavelength extraction methods comprised Successive Projections Algorithm (SPA) [?], Pearson correlation coefficient method, Random Forest (RF), and Stepwise Regression (SR). Vegetation indices (VI) are linear or nonlinear combinations of reflectance at specific bands. This study analyzed the inversion capability of traditional vegetation indices (Table 2 ) for chlorophyll and water content, and re-adjusted the bands of several commonly used vegetation indices (Table 3 ) to construct new vegetation indices, including Difference Index (DI), Ratio Index (RI), and Normalized Difference Vegetation Index (NDVI), whose inversion capabilities were evaluated.

Subsequently, machine learning techniques were used to construct chlorophyll and water content inversion models. Machine learning methods included PLSR, Artificial Neural Network (ANN), Support Vector Regression (SVR), K-nearest Neighbor (KNN), and Stacking. The Stacking model used ANN, SVR, KNN, and linear regressors as base learners.

Finally, by systematically comparing the performance of different preprocessing methods, dimensionality reduction methods, and machine learning models, the optimal inversion models for chlorophyll and water content were selected. In this study, 70% of the spectral sample data (202 samples) were randomly selected as the training dataset for comparing different preprocessing and feature extraction methods and training inversion models; the remaining 30% (86 samples) served as the testing dataset for evaluating model performance. Model performance evaluation metrics included coefficient of determination ( $R^2$ ) and Root Mean Squared Error (RMSE).

### 3. Results and Discussion

**3.1 Spectral Characteristics and Trends of Maize Leaves Under Different Drought Stress Levels** Stressed and non-stressed leaves exhibited different spectral characteristics (Fig. 4 [Figure 4: see original paper]), with leaf reflectance in the 400-1700 nm range gradually increasing as drought stress severity intensified, though the overall spectral curve trend remained consistent. The plant physiological explanations for this phenomenon can be discussed by wavelength range: reflectance in the 400-750 nm range is affected by chlorophyll absorption, and drought stress causes chlorophyll degradation, reducing leaf chlorophyll content; reflectance in the 750-1350 nm range is primarily influenced by internal leaf structure, which is damaged by drought stress; reflectance in the 1350-1700 nm range is affected by water absorption, and drought stress reduces leaf water content.

**3.2 Performance Comparison of Different Spectral Preprocessing Methods** Nine preprocessing algorithms were applied to the spectral data: SG, MSC, SNV, FD, SD, SG+MSC, SG+SNV, SG+FD, and SG+SD (Fig. 5 [Figure 5: see original paper]). The SG-smoothed spectral curves showed little difference from the original spectra, as the extracted spectra were averages of the mesophyll region with minimal noise. MSC and SNV normalization made the spectral distributions more concentrated. FD and SD differential processing significantly increased the number of absorption peaks.

To evaluate preprocessing performance, the original and nine preprocessed spectral datasets were used as inputs for PLSR, with chlorophyll and water content as outputs. Five-fold cross-validation was used to determine the optimal number of principal components, and the validation set was used to evaluate prediction performance to identify the optimal preprocessing method (Table 4). SG+MSC preprocessing achieved the best performance for chlorophyll content prediction, while original spectral data performed best for water content prediction. Therefore, in subsequent model development, SG+MSC preprocessed spectra were used for chlorophyll inversion, and original spectra were used for water content inversion.

### 3.3 Construction of Inversion Models Based on Characteristic Wavelengths

**3.3.1 Analysis of Characteristic Wavelength Extraction Results** The characteristic bands selected by different extraction methods are shown in Fig. 6 [Figure 6: see original paper] and Fig. 7 [Figure 7: see original paper]. The SPA algorithm determined the number of characteristic wavelengths based on the minimum RMSE. For chlorophyll inversion, 15 characteristic wavelengths were extracted and distributed across the entire spectral range (Fig. 6(a)); for water content inversion, 16 characteristic wavelengths were extracted and concentrated in wavelengths greater than 900 nm (Fig. 7(a)).

The Pearson correlation coefficient method calculated correlation coefficients between reflectance at each wavelength and chlorophyll/water content, sorting wavelengths by absolute correlation value. Different numbers of wavelengths were sequentially selected and combined with linear regression to establish inversion models for evaluating the optimal number of characteristic wavelengths. The top 25 wavelengths with highest absolute correlation coefficients were selected for chlorophyll inversion (Fig. 6(b)), and the top 35 wavelengths were selected for water content inversion (Fig. 7(b)).

The RF algorithm calculated the Gini coefficient [?] for each wavelength to characterize variable importance for chlorophyll and water content inversion. For both parameters, wavelengths with higher Gini coefficients were concentrated in the 400-800 nm range for chlorophyll and 800-1300 nm range for water content. Different numbers of wavelengths were sequentially selected and combined with linear regression to evaluate the optimal number. The top 20 wavelengths with highest Gini coefficients were selected for chlorophyll inversion (Fig. 6(c)), and the top 25 wavelengths were selected for water content inversion (Fig. 7(c)).

The SR algorithm selected 9 characteristic wavelengths distributed across the entire spectral range for chlorophyll inversion (Fig. 6(d)), and 13 characteristic wavelengths concentrated near 1300 nm for water content inversion (Fig. 7(d)).

**3.3.2 Performance Analysis of Inversion Models Based on Characteristic Wavelengths** To compare the effects of different characteristic band extraction methods and regression methods on inversion performance, five variable sets (characteristic wavelengths extracted by SPA, correlation coefficient, RF, and SR methods, plus full wavelengths) were used as inputs for four regression methods (ANN, SVR, KNN, and Stacking), with chlorophyll and water content as outputs, resulting in 20 models each. Stacking used ANN, SVR, KNN, and linear regressors as base learners. Five-fold cross-validation determined optimal parameters, and the testing set evaluated model performance to identify the optimal inversion methods.

For chlorophyll content prediction (Table 5 ), the SR-Stacking model achieved optimal performance with testing set  $R^2$  of 0.878 and RMSE of 0.317 mg/g (Fig. 8 [Figure 8: see original paper]). Compared with the full-band Stacking model, SR-Stacking improved  $R^2$  by 2.90%, reduced RMSE by 0.036 mg/g, and decreased input variables from 1301 to 9 (a 99.31% reduction), demonstrating that SR effectively extracts characteristic wavelengths and removes redundant information to improve prediction performance.

For water content prediction (Table 6 ), the SPA-Stacking model achieved optimal performance with testing set  $R^2$  of 0.859 and RMSE of 3.75% (Fig. 9 [Figure 9: see original paper]). Compared with the full-band Stacking model, SPA-Stacking increased  $R^2$  by 0.2%, reduced RMSE by 0.03%, and decreased input variables from 1301 to 16 (a 98.77% reduction), proving that SPA effectively extracts characteristic wavelengths for water content prediction.

### 3.4 Construction of Inversion Models Based on Vegetation Indices

Traditional vegetation indices were constructed using preprocessed spectral data (Table 2), and linear regression models were established for chlorophyll and water content inversion. The prediction performance of each traditional vegetation index is shown in Table 7. All linear regression models based on traditional vegetation indices had  $R^2$  values below 0.5, indicating poor prediction performance.

Therefore, the feasibility of reselecting wavelengths from traditional vegetation indices to construct new vegetation indices was explored. For new vegetation index construction, all possible two-wavelength combinations in the 400-1700 nm range were used to calculate Difference Index (DI), Ratio Index (RI), and Normalized Difference Vegetation Index (NDVI) according to the formulas in Table 3. The correlations between these indices and chlorophyll/water content were then analyzed (Fig. 10 [Figure 10: see original paper]). For chlorophyll, the bands used to calculate vegetation indices were in the visible range, while for water content, bands from both visible and near-infrared ranges were used. Finally, the wavelength combinations with the highest absolute correlation coefficients were selected to construct new vegetation indices (Table 8).

The newly constructed vegetation indices were used to establish linear regression inversion models for chlorophyll and water content. Results showed that prediction performance improved significantly compared to traditional vegetation indices. For chlorophyll inversion, the NDVI (410, 559) model achieved the highest accuracy with  $R^2$  of 0.803 and RMSE of 0.403 mg/g. For water content inversion, the RI (400, 1171) model achieved the highest accuracy with  $R^2$  of 0.827 and RMSE of 3.28% (Table 9).

### 3.5 Visualization of Maize Leaf Physiological Information

This study used the 9 characteristic wavelengths selected by SR and the SR-Stacking model to visualize leaf chlorophyll content, and the 16 characteristic wavelengths selected by SPA and the SPA-Stacking model to visualize water content. By analyzing each pixel in the mesophyll region of hyperspectral images, the reflectance at the 9 characteristic wavelengths was input to the SR-Stacking model to output chlorophyll content for each pixel, and the reflectance at the 16 characteristic wavelengths was input to the SPA-Stacking model to output water content for each pixel. Distribution maps of leaf chlorophyll and water content were then constructed (Fig. 11 [Figure 11: see original paper]). The results demonstrated that visualization of leaf physiological parameters can reveal detailed and intuitive differences in physiological parameters between different regions of the same leaf, which can be applied to maize growth monitoring, disease detection, and plant breeding research.

#### 4. Conclusion

This study investigated rapid, non-destructive detection methods for leaf physiological information (chlorophyll and water content) under drought stress and visualization techniques for leaf physiological parameter distribution, using maize seedlings as the research object. First, a hyperspectral image segmentation algorithm based on image processing technology was developed to extract the mesophyll region and average reflectance spectra, eliminating interference from background and leaf veins. Second, by systematically analyzing the effects of different characteristic wavelength extraction methods and machine learning regression models on prediction performance, optimal inversion models for chlorophyll and water content were established. Finally, the construction of vegetation indices for chlorophyll and water content inversion was explored and their performance was evaluated.

The results demonstrate that spectral technology can achieve accurate and rapid detection of physiological parameters in drought-stressed maize leaves. For chlorophyll detection, optimal performance was achieved using SG+MSC preprocessing, SR feature extraction, and Stacking modeling. For water content detection, optimal performance was achieved using original spectra, SPA feature extraction, and Stacking modeling. This research provides methodological and technical support for early warning of drought and nutrient stress in maize. However, a limitation of this study is that spectral data were collected indoors. Considering that field environments are often more complex, future work will apply transfer learning to improve model generalization across different environments, aiming to develop a field-based online monitoring system for drought and nutrient stress in smart agriculture.

**Conflict of Interest Statement:** The authors declare no conflicts of interest.

---

#### References

- [1] LÜ M Q, SONG Y J, WENG H Y, et al. Effect of near infrared hyperspectral imaging scanning speed on prediction of water content in *Arabidopsis* [J]. *Spectroscopy and Spectral Analysis*, 2020, 40(11): 3508-3514.
- [2] ZHANG J F, TANG B, WANG J X, et al. Comparative analysis of characteristic wavelength screening of apple soluble solids based on near-infrared spectroscopy [J/OL]. *Laser & Optoelectronics Progress*. (2022-11-01) [2023-08-12]. <https://kns.cnki.net/kcms/detail/31.1690.TN.20221031.1514.004.html>
- [3] YANG J J, LU J W. Research progress of spectral technology in peanut quality detection [J]. *Crop Research*, 2018, 32(4): 345-348.
- [4] SUN J, JIN X M, MAO H P, et al. A model for predicting nitrogen of lettuce leaves based on hyperspectral imaging [J]. *Chinese Journal of Analytical Chemistry*, 2014, 42(5): 672-677.
- [5] GAO D H, LI M Z, ZHANG J Y, et al. Improvement of chlorophyll content estimation on maize leaf by vein removal in hyperspectral image [J]. *Computers*

- and *Electronics in Agriculture*, 2021, 184: 106077.
- [6] TAN K Z, WANG S W, SONG Y Z, et al. Estimating nitrogen status of rice canopy using hyperspectral reflectance combined with BPSO-SVR in cold region [J]. *Chemometrics and Intelligent Laboratory Systems*, 2018, 172: 9-19.
- [7] SHU M Y, DONG Q Z, FEI S P, et al. Improved estimation of canopy water status in maize using UAV-based digital and hyperspectral images [J]. *Computers and Electronics in Agriculture*, 2022, 197: 106982.
- [8] LICHTENTHALER H K, WELLBURN A R. Determinations of total carotenoids and chlorophylls a and b of leaf extracts in different solvents [J]. *Biochemical Society Transactions*, 1983, 11(5): 591-592.
- [9] MASSAOUDI M, REFAAT S S, ABU-RUB H, et al. PLS-CNN-BiLSTM: An end-to-end algorithm-based Savitzky-Golay smoothing and evolution strategy for load forecasting [J]. *Energies*, 2020, 13(20): 5464.
- [10] BAO Y D, KONG W W, LIU F, et al. Detection of glutamic acid in oilseed rape leaves using near infrared spectroscopy combined with SiPLS and the least squares-support vector machine [J]. *International Journal of Molecular Sciences*, 2012, 13(11): 14563-14574.
- [11] CHEN Q S, JIANG P, ZHAO J W. Measurement of total flavone content in snow lotus (*Saussurea involucrata*) using near infrared spectroscopy combined with interval PLS and genetic algorithm [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2010, 76(1): 50-55.
- [12] ARAÚJO M C U, SALDANHA T C B, GALVÃO R K H, et al. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis [J]. *Chemometrics and Intelligent Laboratory Systems*, 2001, 57(2): 65-73.
- [13] ZARCO-TEJADA P J, BERJÓN A, LÓPEZ-LOZANO R, et al. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy [J]. *Remote Sensing of Environment*, 2005, 99(3): 271-287.
- [14] MEHDAOUI R, ANANE M. Exploitation of the red-edge bands of Sentinel-2 to improve the estimation of durum wheat yield in Grombalia region (Northeastern Tunisia) [J]. *International Journal of Remote Sensing*, 2020, 41(23): 9045-9063.
- [15] GITELSON A A, KEYDAN G P, MERZLYAK M N. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves [J]. *Geophysical Research Letters*, 2006, 33(11): L11402.
- [16] FITZGERALD G, RODRIGUEZ D, O' LEARY G. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index: The canopy chlorophyll content index (CCCI) [J]. *Field Crops Research*, 2010, 116(3): 318-324.
- [17] JORDAN C F. Derivation of leaf-area index from quality of light on the forest floor [J]. *Ecology*, 1969, 50(4): 663-666.
- [18] DASH J, CURRAN P J. The MERIS terrestrial chlorophyll index [J]. *International Journal of Remote Sensing*, 2004, 25(23): 5403-5413.
- [19] GAO B C. NDWI: A normalized difference water index for remote sensing of vegetation liquid water from space [J]. *Remote Sensing of Environment*,

1996, 58(3): 257-266.

[20] HUNT E R, ROCK B N. Detection of changes in leaf water content using near- and middle-infrared reflectances [J]. *Remote Sensing of Environment*, 1989, 30(1): 43-54.

[21] ZHANG J J, ZHANG W, XIONG S P, et al. Comparison of new hyperspectral index and machine learning models for prediction of winter wheat leaf water content [J]. *Plant Methods*, 2021, 17(1): 1-14.

[22] PENUELAS J, PINOL J, OGAYA R, et al. Estimation of plant water concentration by the reflectance Water Index WI (R900/R970) [J]. *International Journal of Remote Sensing*, 1997, 18(13): 2869-2875.

[23] ZARCO-TEJADA P J, USTIN S L. Modeling canopy water content for carbon estimates from MODIS data at land EOS validation sites [C]// *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium*. Piscataway, New Jersey, USA: IEEE, 2002: 342-344.

[24] SUN H, FENG M C, XIAO L J, et al. Assessment of plant water status in winter wheat (*Triticum aestivum* L.) based on canopy spectral indices [J]. *PLoS ONE*, 2019, 14(6): e0216890.

[25] MENZE B H, KELM B M, MASUCH R, et al. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data [J]. *BMC Bioinformatics*, 2009, 10(1): 1-16.

---

*Foundation items: National Natural Science Foundation of China (31901466); Guangxi Science and Technology Base and Talent Project (AD22035915)*

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv – Machine translation. Verify with original.*