Computers and Electronics in Agriculture 229 (2025) 109724

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Synergistic use of stay-green traits and UAV multispectral information in improving maize yield estimation with the random forest regression algorithm

Yuan Liu [a,b,1], Kaijian Fan [b,1], Lin Meng [a,b], Chenwei Nie [a,b], Yadong Liu [a,b], Minghan Cheng [c], Yang Song [a,b,*], Xiuliang Jin [a,b,*]

[a] National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya 572024, China
[b] Institute of Crop Sciences, Chinese Academy of Agricultural Sciences/Key Laboratory of Crop Physiology and Ecology, Ministry of Agriculture, Beijing 100081, China
[c] Jiangsu Key Laboratory of Crop Genetics and Physiology/Jiangsu Key Laboratory of Crop Cultivation and Physiology, Agricultural College, Yangzhou University, 225009 Yangzhou, China

ARTICLE INFO

Keywords:
UAV multispectral
Maize yield
Stay-Green Index (SGI)
Machine learning
Remote sensing

ABSTRACT

The timely and accurate estimation of maize yield in the field is critical for developing agricultural management strategies and ensuring food security. Most crop yield estimation models are based on absolute indices, such as spectral and textural indices, yet few studies have considered the stability of relative indices in the time dimension. To address this issue, a novel relative index, the Stay-Green Index (SGI), was extracted to characterize maize growth from the NDVI time series of Unmanned Aerial Vehicle (UAV) multispectral images. We used the Random Forest Regression (RFR) framework and a set of vegetation indices (VIs) to build the maize yield estimation model. The effects of SGI as an additional variable on yield estimation were examined at different growth stages. We finally evaluated the performance of our yield estimation model incorporating SGI. Our results showed that: (1) the correlation coefficients between VIs and the actual yield were relatively low at the silking (R1) and blister (R2) stages, close to 0, while the results were relatively high at the dough (R4), denting (R5), and maturity (R6) stages, close to 0.69; (2) the proposed SGI was significantly correlated with the actual yield; and (3) our yield estimation model with SGI can greatly improve the accuracy of yield estimation models at the R1 and R2 stages, while the performance was maintained at the R4, R5, and R6 stages. Overall, this study highlights that the model using the RFR algorithm combining SGI and VIs can improve the within-season yield estimation for maize.

1. Introduction

Maize, as one of the staple crops, is widely planted across the globe. China is the world's leading grain producer and importer, harvesting 273 million metric tons in 2021 and supplying about 20 % of the country's grain production (Gao et al., 2019; Li et al., 2021b; Yang et al., 2021). Therefore, timely and accurate maize yield estimation, especially before harvest, is of high strategic significance for China's maize import and export, agricultural decision-making, and national macroeconomic planning (Khaki et al., 2021; Ma et al., 2021). In addition, spatially continuous estimation of maize yields at early growth stages can provide guidance to farmers in applying targeted field control measures. To ensure high quality, stability, and high yields while achieving precise, efficient, and rational use of water and fertilizer is essential for the development of precision agriculture (Maimaitijiang et al., 2020; Mueller et al., 2012; Xu et al., 2021).

In recent years, with the aid of spatial information technology and to meet the needs for agricultural applications, remote sensing has now been regarded as one of the most effective tools for monitoring crop growth status and estimating grain yields (Furukawa et al., 2020; Wu et al., 2021). Remote sensing platforms, including satellites and unmanned aerial vehicles (UAVs), are primarily used to retrieve crop parameters (Guo et al., 2022; Schwalbert et al., 2020; Zhang et al., 2020). The satellite platform has made remarkable achievements in crop yield

* Corresponding authors.
E-mail addresses: songyang01@caas.cn (Y. Song), jinxiuliang@caas.cn (X. Jin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compag.2024.109724
Received 23 August 2023; Received in revised form 20 May 2024; Accepted 27 November 2024
Available online 3 December 2024

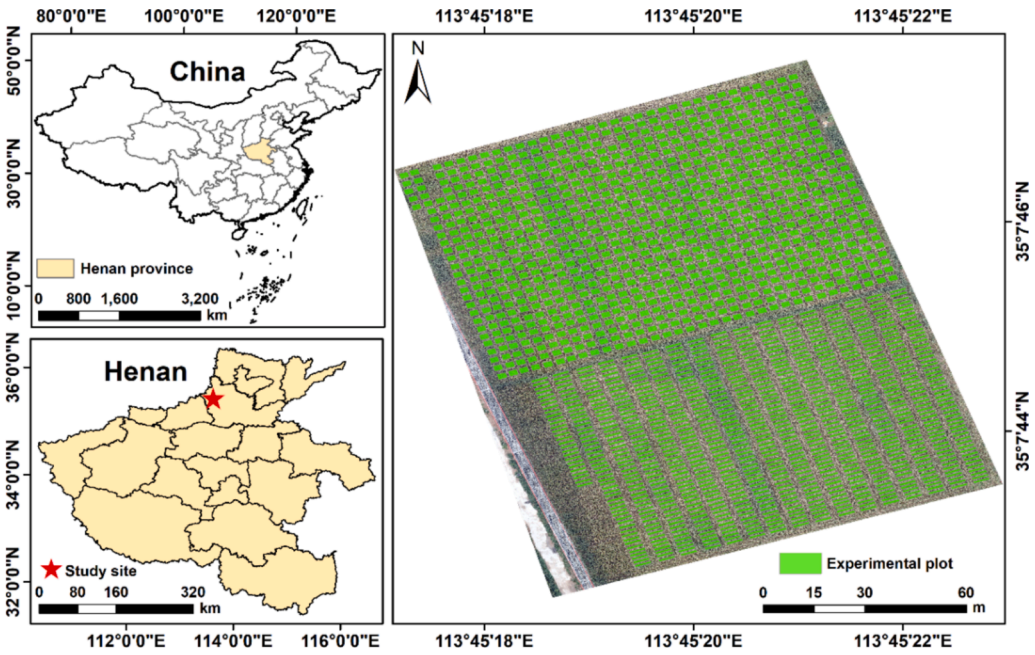

Fig. 1. Geographical location of the study area and the layout of the experimental plot.

estimation (Nagy et al., 2021; Panek and Gozdowski, 2021) with its wide observation coverage, all-day dynamic observation, and abundant information obtained (Sakamoto, 2020). When it comes to crop yield estimation at the field scale, satellite images have a long revisit period, low spatial resolution, and are greatly affected by cloudy and rainy weather, which makes obtaining images within a specified time range uncertain (Berni et al., 2009; Peng et al., 2021). Fortunately, thanks to the emergence and rapid development of UAV remote sensing technology, new solutions have been found to tackle the issues associated with crop yield estimation by satellite platforms (Xie and Yang, 2020). The emergence of UAV platforms has led to unprecedented increases in spatial, spectral, and temporal resolution for remote sensing, making it a highly effective tool for precision agriculture (Li et al., 2021c). Besides, UAVs offer high scalability and flexibility, allowing for the use of a variety of remote sensing sensors to acquire crop canopy images from different heights, all with lower operating costs and complexity than previously possible (Fu et al., 2021; Pádua et al., 2017; Wang, 2021). Consequently, the UAV platform is becoming increasingly popular for obtaining spectral and structural data related to crops (Xu et al., 2021).

Several studies have successfully used UAV images to estimate crop grain yields, including maize (Zhang et al., 2020), soybean (Maimaitijiang et al., 2020), potato (Li et al., 2021a), wheat (Yue et al., 2020), and rice yields (Wan et al., 2020). These studies mostly leverage crop canopy spectral information to model yield estimation. Canopy spectral information is typically represented by vegetation indices (VIs), which contain critical spectral information (Wang et al., 2021). The VIs can characterize canopy spectral information by using transformations of surface reflectance at two or more bands (Xu et al., 2022), with the primary purpose of enhancing sensitive information about agronomic parameters, e.g., leaf area index (LAI), nitrogen content (N), and chlorophyll (Chl), while minimizing the effects of soil, observation geometry, and other vegetation characteristics (Nie et al., 2023). At present, many kinds of VIs are employed as essential remote sensing parameters to estimate crop yields. In most studies, only a few types of VIs are selected, resulting in insufficient discussion on the correlation between VIs and crop yields. In addition, the current spectral information used to estimate yields uses transient indicators and does not consider yield estimation errors in the time dimension. According to previous studies, the influence of shaded soil on the canopy spectrum can be negligible (Simic et al., 2014). In a north–south-oriented crop canopy, the soil background is almost obscured at 9:00 and 15:00 but sunlit at 12:00. Therefore, optimizing the observation time can reduce the soil background visible to sensors (Li et al., 2020; Lu et al., 2019). Based on this, commonly used transient indicators such as VIs fluctuate within a day, resulting in the need for recalibration of the model when applied at different times. This also confirms that relative indices should be considered in yield estimation models to reduce the variation of spectral reflectance due to different times.

Stay-Green (SG) refers to the phenotypic trait of delayed leaf senescence or yellowing during post-anthesis development of crops while still maintaining photosynthesis (Christopher et al., 2014; Greco et al., 2012). Previous studies have fully confirmed that the difference in crop yields is closely related to the length of photosynthesis time, and it has been concluded that SG and crop yields are highly positively correlated (Mirosavljević et al., 2020). Therefore, the limited research on crop yield estimation has considered crop SG traits. In the family of VIs, the normalized difference vegetation index (NDVI), as an index of photosynthetic active biomass (Vannoppen and Gobin, 2021), has been used to measure conservation green characteristics (Montazeaud et al., 2016). Most of the existing studies used the values of NDVI at different growth stages to characterize crop SG traits (Christopher et al., 2016; Latif et al., 2020; Mirosavljević et al., 2020). Therefore, whether it is possible to construct a relative index of yield estimation, such as the SGI, combined with the curve of the time series NDVI, has not been verified.

Machine learning (ML) models typically attempt to establish the empirical relationship between drivers and historical crop yields (Zhang et al., 2019). The method has the capacity to analyze large amounts of data using high-performance computing and can effectively disentangle the complex relationship between numerous remote sensing information sources and crop yields through linear or nonlinear approaches (Chlingaryan et al., 2018; Rashid et al., 2021). Furthermore, ML models have the advantage of estimating crop yields independent of specific parameters (Wang et al., 2020). Among the ML models for crop yield estimation, the random forest regression (RFR) algorithm is one of the most widely used non-parametric ensemble machine learning methods and is composed of many decision trees (Yang et al., 2022). Li et al. (2021b) used the RFR algorithm to estimate the yields of three major grain crops in China (wheat, maize, and rice) and then obtained good results with a correlation coefficient (r) higher than 0.75 and normalized root mean square errors (nRMSE) lower than 18.0 %.

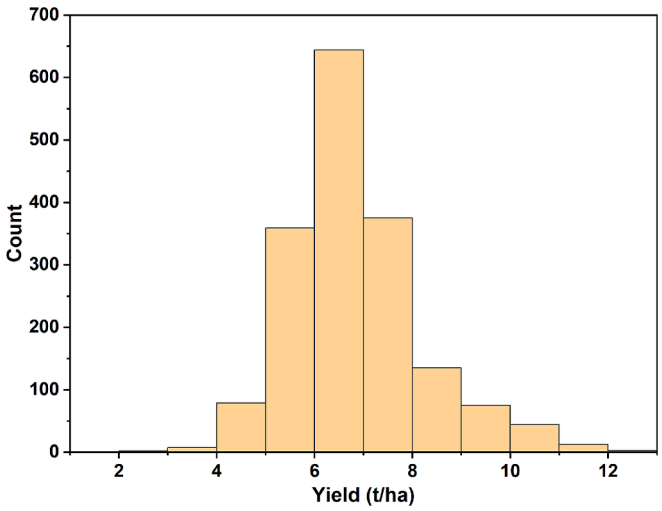

Fig. 2. The histograms of the maize measured yield.

To thoroughly investigate the correlation between various VIs and maize yields across distinct growth stages, this study was conducted to improve the accuracy of the yield estimation model by introducing a relative variable, the Stay-Green Index (SGI), which considered the time factor. The SGI were extracted from the NDVI time series of UAV multispectral images. The maize yield estimation models based on VIs and SGI were established for different growth stages using the RFR algorithm. Our objectives are (1) to explore the relationships between various VIs extracted from multispectral images at different growth stages and the actual maize yield; (2) to characterize maize growth by constructing the SGI based on the NDVI time series; and (3) to explore the importance of SGI in the yield estimation model and then to develop a method to accurately estimate maize yields by combining SGI and VIs. Our study highlights the significance of integrating the SGI into the development of yield estimation models based on VIs, which can greatly amplify the performance of crop yield estimation models.

2. Materials and methods

2.1. Study area and experimental design

The field experiment was conducted at the Xinxiang Experimental Station of the Chinese Academy of Agricultural Sciences in Henan Province (35°10′N, 113°47′E), which mainly grows maize in summer and belongs to the North China Plain (Fig. 1). The station has a typical warm continental monsoon climate with four distinct seasons: spring (characterized by dry and windy conditions), summer (marked by high temperatures and rainfall), autumn (cool weather), and winter (cold but snowless conditions). The annual average temperature is 14℃, with the coldest temperature being −0.5℃ in January and the hottest temperature being 27℃ in July. The annual average precipitation is 548.3 mm, with most of the rainfall occurring between July and August. The annual evaporation, sunshine, and frost-free periods are 1748.4 mm, 2323.9 h, and 205 days, respectively. This makes the station suitable for growing crops.

The experiment of this study was conducted from June to October 2021, and the experimental field was divided into two areas: the upper field of 908 plots with each 2.4 × 2.5 m^2; and the lower field of 828 plots with each 1.2 × 5 m^2. All plots were planted with hybrid maize of different varieties, ensuring the generalizability of the proposed method and avoiding overfitting. The planting density was 67,500 plants per hectare, with a planting row spacing of 0.6 m and a plant spacing of about 0.25 m. The field management practices, including fertilization, weeding, and insecticide usage, adhered to local production standards.

2.2. Ground measurement of maize yield

During maize maturity, the number of actual maize plants, empty rods, and double ears in each plot was counted before harvesting. When measuring the yield, the total number of maize ears in each plot was counted, and all the ears in the plot were harvested, weighed, threshed, and weighed again to calculate the seed rate. The moisture content of the grain was assessed through an average of 3–5 repetitions. The final yield of each plot was calculated based on 14 % grain water content. In total, 1736 samples of yield data were gathered for this study. Fig. 2 shows the distribution of measured maize yields. The conversion formula for grain yield is as follows:

$$GY = (W \div N) \times SR \times (1 - C) \div 0.86 \div S \quad (1)$$

where GY represents the grain yield of maize at 14 % water content, SR stands for seed rate of maize, and W, N, C, and S are the total ear weight of maize in the plot, the total ear number of maize in the plot, the grain water content, and the area of the plot, respectively.

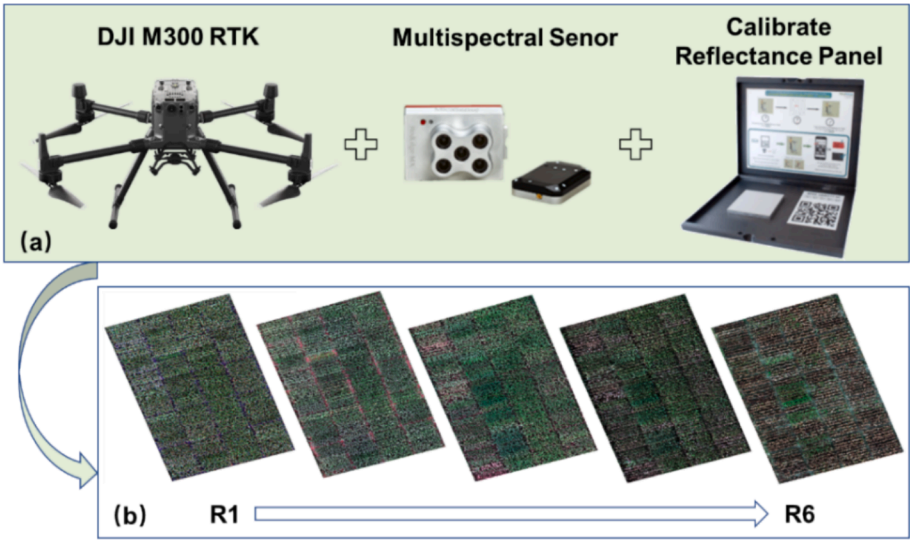

Fig. 3. The UAV platform used in this paper (a) and the multispectral images obtained by UAV at different growth stages (b).

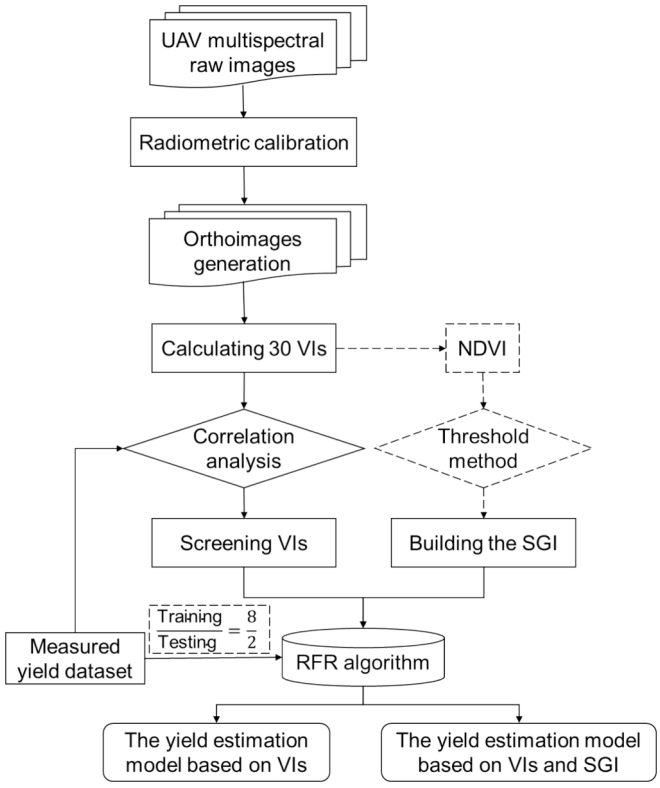

Fig. 4. The overall process of this study.

2.3. UAV data acquisition and processing

2.3.1. UAV flight campaign

The UAV used in this study is a DJI M300 RTK (DJI Innovation Co., Ltd., Shenzhen, China), equipped with a RedEdge-MX sensor (MicaSense Inc., Seattle, WA, USA), which constitutes the platform of a low-altitude UAV camera system (Fig. 3). The DJI M300 RTK comes with a six-way positioning obstacle avoidance and visual flight assistance interface, meeting the IP45 protection level, supporting up to 55 min of endurance, a transmission range of up to 15 km, and the ability to simultaneously support three payloads. With a maximum takeoff weight of 9 kg, the RTK system comes with centimeter-level position accuracy. The RedEdge-MX sensor has a dimension of 87 × 59 × 45.4 mm, a maximum resolution of 1280 × 960 pixel^2, a focal length of 5.4 mm, a mass of 231.9 g, and a field of view angle of 47.2°. The sensor comes with narrow spectral bands: blue (with a central wavelength of 475 nm and a bandwidth of 20 nm), green (with a central wavelength of 560 nm and a bandwidth of 20 nm), red (with a central wavelength of 668 nm and a bandwidth of 10 nm), red edge (with a central wavelength of 717 nm and a bandwidth of 10 nm), and near-infrared (with a central wavelength of 840 nm and a bandwidth of 40 nm). The sensor comes equipped with GPS and is capable of capturing up to five spectral images simultaneously (Lu and He, 2017).

The UAV remote sensing observations were carried out during the maize growing season of 2021, including the silking (R1), blister (R2), dough (R4), denting (R5), and maturity (R6) stages. All flight campaigns were conducted under clear and cloudless skies between 10:00 and 14:00 Beijing time to minimize the impact of cloud cover and changes in solar zenith angle on image quality. During the acquisition of multispectral images, all flight missions were undertaken at a height of 30 m and a speed of 2.1 m/s. Utilizing up to 80 % lateral and forward overlaps during flight ensured the generation of high-quality orthoimages, crucial for crafting a seamlessly mosaicked orthoimage (Lu and He, 2017). Furthermore, post-takeoff, it's essential to ascend the UAV to a specific altitude to capture images of the calibration panel, thus obtaining the

Table 1
Summary of vegetation indices for the grain yield estimation.

Vegetation index	Formulation	References
NDVI	(NIR-R)/(NIR + R)	(Rouse et al., 1973)
Ratio vegetation index (RVI)	NIR/R	(Tucker, 1979)
Enhanced Vegetation Index (EVI)	2.5(NIR-R)/(NIR + 6R-7.5B + 1)	(Huete et al., 2002)
Soil Adjusted Vegetation Index (SAVI)	1.5(NIR-R)/(NIR + R + 0.5)	(Huete, 1988)
Red-edge chlorophyll index (RECI)	(NIR/RE)-1	(Gitelson et al., 2005)
Optimized soil adjusted vegetation index (OSAVI)	1.16(NIR-R)/(NIR + R + 0.16)	(Rondeaux et al., 1996)
Green chlorophyll index (GCI)	(NIR/G)-1	(Gitelson et al., 2005)
Renormalized Difference Vegetation Index (RDVI)	$(NIR-R)/\sqrt{NIR+R}$	(Roujean and Breon, 1995)
Green Re-normalized different vegetation index (GRDVI)	$(NIR-G)/\sqrt{NIR+G}$	(Cao et al., 2013)
Modified Chlorophyll Absorption in Reflectance Index (MCARI)	[(NIR-RE)-0.2(NIR-R)] × (NIR/RE)	(Lu et al., 2017)
Normalized difference red-edge (NDRE)	(NIR-RE)/(NIR + RE)	(Gitelson and Merzlyak, 1997)
Modified Simple Ratio (MSR)	$(NIR/R-1)/\sqrt{NIR/R+1}$	(Chen, 2014)
Transformed chlorophyll absorption in reflectance index (TCARI)	3[(RE-R)-0.2(RE-G) × (RE/R)]	(Haboudane et al., 2002)
Modified Triangular Vegetation Index (MTVI)	1.2[1.2(NIR-G)-2.5(R-G)]	(Huete et al., 2002)
Green Normalized Difference Vegetation Index (GNDVI)	(NIR-G)/(NIR + G)	(Gitelson et al., 2003)
Wide dynamic range vegetation index (WDRVI)	(0.1NIR-R)/(0.1NIR + R)	(García-Martínez et al., 2020)
Two-band enhanced vegetation index (EVI2)	2.5(NIR-R)/(NIR + 2.4R + 1)	(Jiang et al., 2008)
Canopy Chlorophyll Content Index (CCCI)	NDRE/NDVI	(Fitzgerald et al., 2010)
Structure insensitive pigment index (SIPI)	(NIR-B)/(NIR + B)	(Penuelas et al., 1995)
MERIS terrestrial chlorophyll index (MTCI)	(NIR-RE)/(RE + R)	(Dash and Curran, 2010)
Triangular greenness index (TGI)	G-0.39R-0.61B	(Hunt et al., 2013)
Simple ratio pigment index (SRPI)	B/R	(Penuelas et al., 1995)
Normalized pigment chlorophyll index (NPCI)	(R-B)/(R + B)	(Ranjan et al., 2012)
Visible atmospherically resistant index (VARI)	(G-R)/(G + R-B)	(Schneider et al., 2008)
Plant senescence reflectance index (PSRI)	(B-R)/G	(Sims and Gamon, 2002)
Ratio vegetation index 2 (RVI2)	NIR/G	(Xue et al., 2004)
Triangular vegetation index (TVI)	60(NIR-G)-100(R-G)	(Broge and Leblanc, 2001)
Difference Vegetation Index (DVI)	NIR-R	(Tucker, 1979)
Green, Red Vegetation Index (GRVI)	(G-R)/(G + R)	(Tucker, 1979)
Color index of vegetation (CIVE)	0.441R-0.881G + 0.385B + 18.78745	(Kataoka et al., 2003)

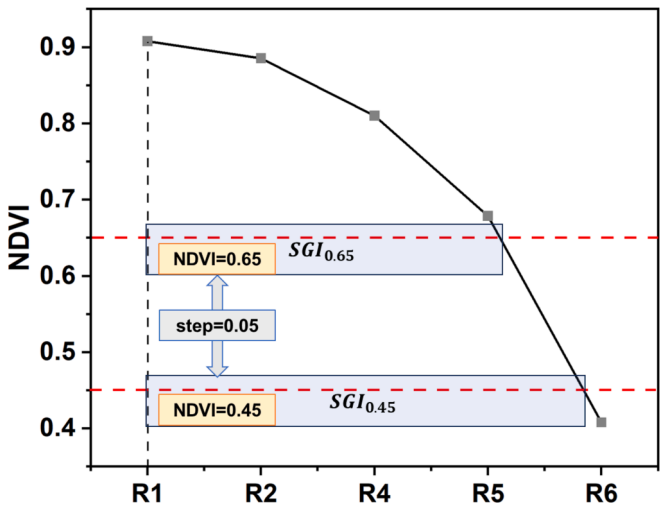

Fig. 5. The diagram of SGI construction.

necessary data for radiation calibration.

2.3.2. UAV images pre-processing

The UAV-based multispectral remote sensing images were pre-processed through the Agisoft Metashape Professional (Agisoft LLC, St. Petersburg, Russia) software, including radiometric calibration and orthoimage generation. To address challenges like atmospheric non-uniformity and sensor instability, it becomes imperative to correct the spectral reflectance of the captured multispectral images. This correction involves converting the dimensionless DN (digital number) value to the outer surface reflectivity or radiant brightness value of the atmosphere, which holds practical significance (Chen et al., 2012). During executing multispectral data acquisition tasks with the UAV, radiometric calibration with a gray calibration plate was employed. Prior to image mosaicking, the Metashape software automatically located the images containing the gray calibration plate, extracted the reflectance data from the calibration plate, and automatically carried out radiometric correction of each spectral band in the multispectral images.

To generate the orthoimages, the aerial photos needed to be aligned by using the coordinates, elevation data, and similarities. Similarly, the multispectral images obtained at different growth stages also underwent alignment. Subsequently, a systematic process of generating dense point clouds, meshes, and textures ensued. Once those were done, the multispectral images were transformed into orthoimages using a Digital Elevation Model (DEM) created with high-density points. The geographic coordinate system of the multispectral orthoimage was determined when the DEM was generated, which was GCS_WGS_1984. Each flight output results in a single high-resolution TIFF image. The pre-processed multispectral images were utilized to extract spectral information and build a relative index related to SG that was used to construct the yield estimation model. The overall process is illustrated in Fig. 4.

2.3.3. Spectral variable calculation

To extract spectral information, the UAV multispectral remote sensing orthoimages were processed in ArcGIS 10.8 (Environmental Systems Research Institute, Inc., Redlands, CA, USA) software. The extracted spectral information takes into consideration the vegetation index (VI), which transforms two or more reflectance bands to enhance vegetation features (Peng et al., 2021). For the current study, a set of previously established VIs, which were found to be sensitive to crop yield in earlier research, were selected as spectral variables (Table 1).

2.4. Construction of the stay-green index

Drawing on previous research findings and experimental results with various VIs, this paper utilized the NDVI to construct SGI. SGI is defined as the cumulative growth days when the NDVI of maize plants exceeds a certain threshold. Therefore, the focus was on establishing NDVI time series curves and determining the threshold. Initially, NDVI images were extracted successively from five stages of UAV images obtained, and NDVI time series curves were established for each plot. The first phase of UAV multispectral image capture occurred on August 5, 2021, identified as the start date for SGI calculation.

The determination of the SGI threshold in this study involved two aspects. First, based on the field conditions, when maize plants in some plots turned yellow, the corresponding NDVI value dropped below 0.7. Conversely, areas with NDVI values exceeding 0.45 were deemed relatively healthy with substantial vegetation coverage. Hence, the NDVI threshold was set between 0.45 and 0.7, with a step size of 0.05. The cumulative days of NDVI values surpassing different thresholds were then calculated. Subsequently, correlation analysis was performed between the SGI of each threshold value and the measured yield. The SGI exhibiting the highest correlation was selected for the study.

The specific SGI calculation formula is shown as follows:

$$SGI_n = (Day_{\max} - Day_{\min})_{NDVI_n} \quad (2)$$

where SGI_n is the stay-green index when NDVI is greater than the threshold n, Day is the corresponding data when NDVI is greater than the threshold n, Day_{max} is the date when the plant approaches the maturity stage, and Day_{min} is away from the maturity stage.

Fig. 5 illustrates the process of constructing the SGI. It's noteworthy that since SGI is a relative indicator based on all available images, the SGI for each plot, similar to yield data, consistently corresponds to a single value during its application.

2.5. Modeling methodology

2.5.1. Regression modeling framework

The RFR algorithm is an ensemble-based machine learning algorithm proposed by Breiman (2001), which combines multiple weak-base models to generate the optimal ensemble model (Barbosa et al., 2021). The RFR runs effectively on large and high-dimensional datasets, is less sensitive to overfitting, is robust to outliers, and exhibits great accuracy in dealing with nonlinear regression (Cheng et al., 2022; Singh et al., 2017; Yu et al., 2016). At present, the RFR algorithm has been widely used in the establishment of remote sensing crop yield estimation models and has achieved excellent results (Barbosa et al., 2021; Li et al., 2021a; Rashid et al., 2021). In order to simulate the nonlinear relationship between extracted remote sensing features and maize yield in this study, given the set of training input–output pairs (UAV image-derived features-yield), the RFR algorithm framework was used to build the maize estimation model, which was realized in Python 3.6 software (Google Inc., Mountain View, California, USA) and the scikit-learn library (version 0.24.2, Google Inc.). The RFR models were developed by optimizing two parameters: n_estimators (number of trees in the forest) and random_state (which controls both the randomness of the bootstrapping of the samples used when building trees and the sampling of the features to consider when looking for the best split at each node). Moreover, to ensure impartial and uniform comparison of RFR methods with their given input variables, 80 % of the input–output pairs (1389 samples) in the experiment were randomly selected as training datasets to establish the RFR models, and the remaining 20 % (347 samples) were used as testing datasets to validate the models.

2.5.2. Model performance assessment

Three statistical indicators, determination coefficient (R^2), RMSE, and relative root mean square error (rRMSE) were used to evaluate the

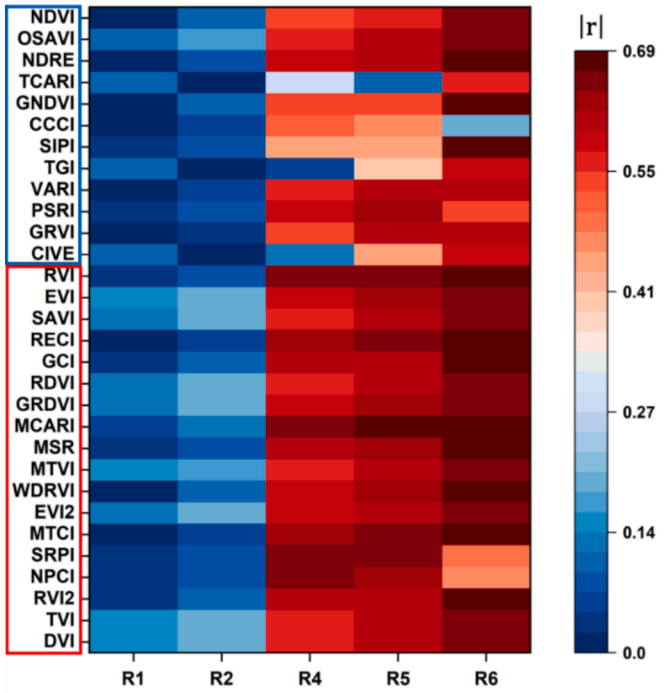

Fig. 6. The correlation between VIs and the measured yield at different growth stages was represented by the absolute correlation coefficient (|r|).

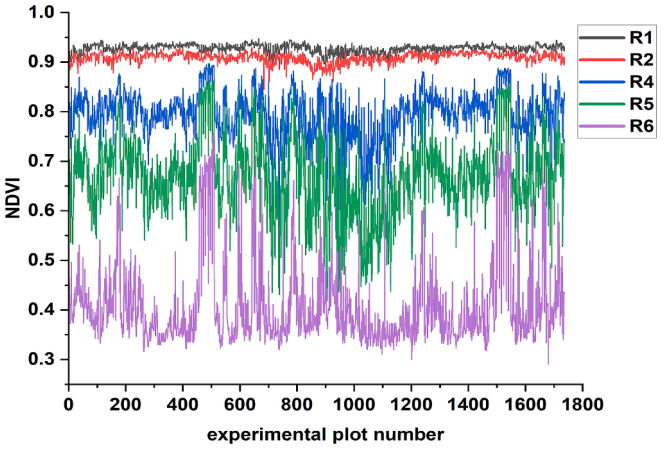

Fig. 7. Variation trend of NDVI in different experimental plots and different growth stages.

Table 2
The correlation analysis between SGI and measured yield under different thresholds.

Threshold	Correlation coefficient
$NDVI_{0.45}$	0.538**
$NDVI_{0.50}$	0.560**
$NDVI_{0.55}$	0.585**
$NDVI_{0.60}$	0.604**
$NDVI_{0.65}$	**0.617****
$NDVI_{0.70}$	0.601**

Note: ** means a significant level of 0.01.

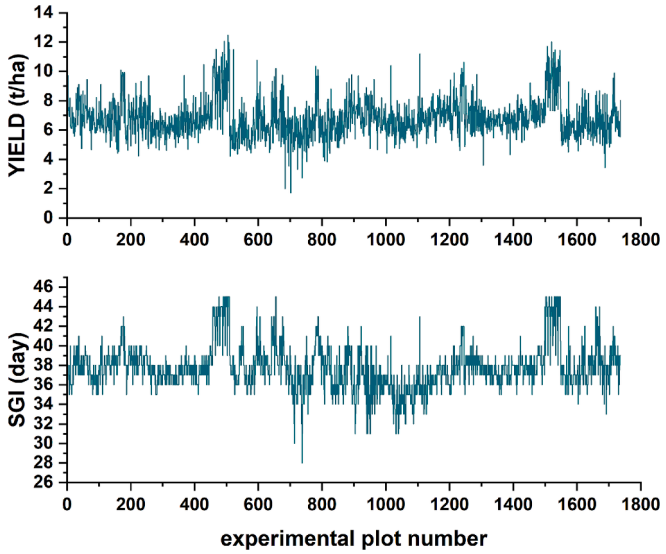

Fig. 8. Line plot of measured yield and SGI with the change of experimental plot.

Table 3
The results of yield estimation models with or without SGI at different growth stages.

Model	Training dataset			Testing dataset		
	R^2	RMSE (t/ha)	rRMSE (%)	R^2	RMSE (t/ha)	rRMSE (%)
M_{R1}	0.539	1.007	14.8	0.064	1.388	20.1
M_{R2}	0.547	0.946	13.9	0.069	1.391	20.1
M_{R4}	0.653	0.766	11.3	0.569	0.944	13.7
M_{R5}	0.618	0.802	11.8	0.561	0.949	13.7
M_{R6}	0.650	0.771	11.3	0.560	0.952	13.8
M_{R1_SGI}	0.702	0.718	10.5	0.541	0.971	14.0
M_{R2_SGI}	0.710	0.708	10.4	0.547	0.971	14.0
M_{R4_SGI}	0.656	0.763	11.2	**0.574**	**0.939**	**13.6**
M_{R5_SGI}	0.619	0.802	11.8	0.562	0.948	13.7
M_{R6_SGI}	0.654	0768	11.3	0.563	0.949	13.7

Note: M_{R1} represents the yield estimation model without SGI at the R1 stage, M_{R1_SGI} represents the yield estimation model with SGI at the R1 stage, and the rest of the same.

accuracy of each yield estimation model. The calculations of R^2, RMSE, and rRMSE are as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2(y_i - \overline{y})^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \quad (4)$$

$$rRMSE = \frac{RMSE}{\overline{x}} \times 100\% \quad (5)$$

where x_i and y_i are the measured and estimated yield of sample i, respectively, $\overline{x}$ and $\overline{y}$ are the average measured and estimated yield of sample i, respectively, n is the total number of samples.

3. Results

3.1. Correlation analysis between VIs and measured yield

The goodness of correlation was examined for various VIs with the measured yield at five different growth stages (Fig. 6). Generally, the

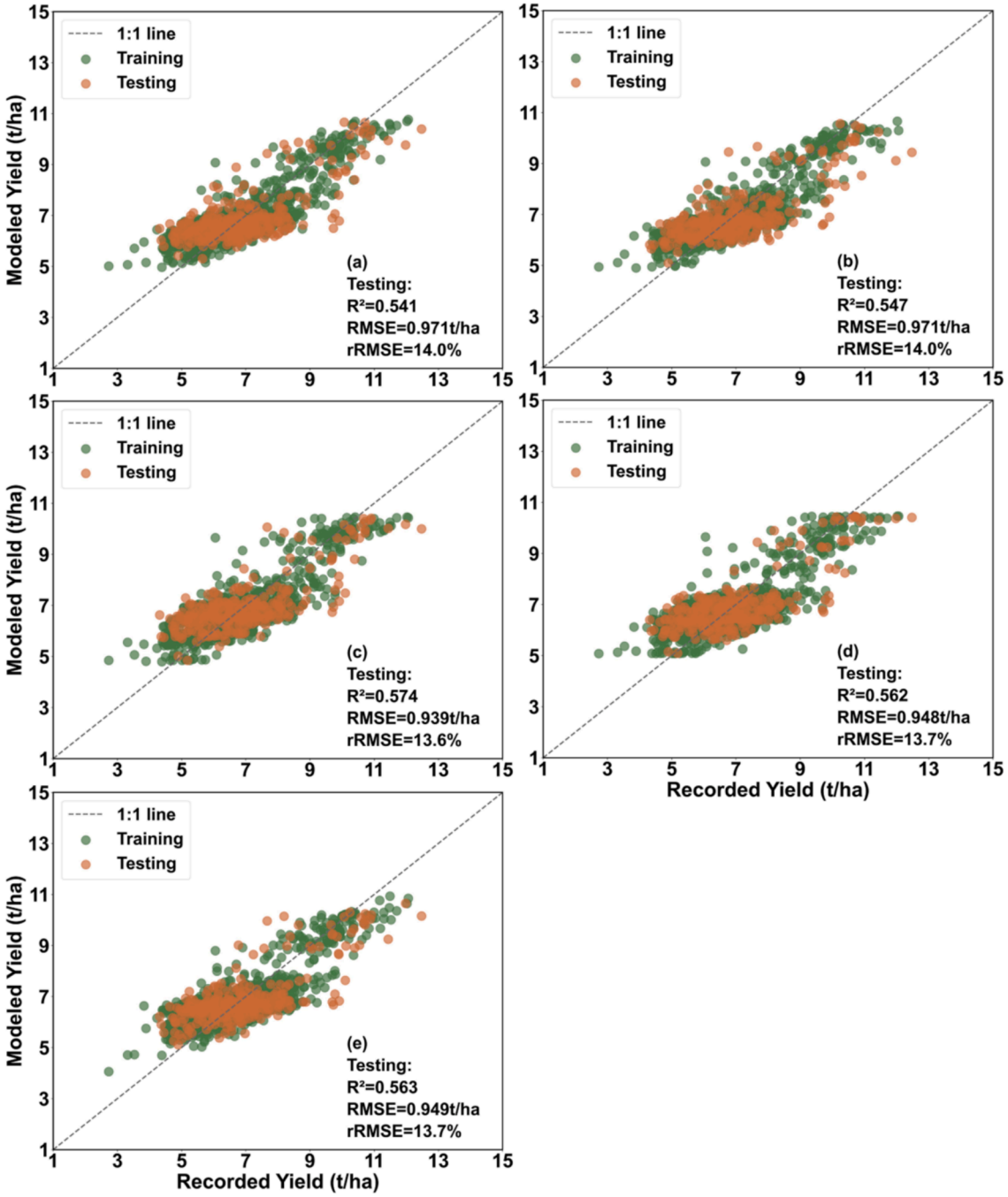

Fig. 9. The scatter plots of each yield estimation model with SGI at different growth stages, (a) the R1 stage, (b) the R2 stage, (c) the R4 stage, (d) the R5 stage, and (e) the R6 stage.

performance of VIs had a higher correlation with the measured yield at the R4, R5, and R6 stages compared to the R1 and R2 stages. Specifically, VIs had the highest correlation with the measured yield at the R6 stage. At the R1 stage, only three VIs (MTVI, TVI, and DVI) had correlation coefficients greater than 0.14, and the absolute correlation coefficients (|r|) between the selected 30 VIs and the measured yield were no more than 0.15. The correlation results in the R2 stage were slightly better than those in the R1 stage, with a maximum correlation value of 0.20 for GRDVI. Starting from the R4 stage, there was a rapid increase in the |r| between VIs and the measured yield, as shown by the transition in the figure from blue to darker red. However, TCARI, TGI, and CIVE had relatively low |r| at the R4 stage, and TCARI and CCCI had the lowest |r| at the R5 and R6 stages, respectively. Eventually, to reduce the redundancy of the yield estimation model and improve the operation's efficiency, some of the VIs were filtered according to the correlation results. The specific basis of feature filtering was that the probability of a VI in the bottom 20 VIs sets at each growth stage was more than 80 %. In Fig. 5, the filtered-out VIs are shown in the blue box, while the retained VIs are shown in the red box.

3.2. Analysis of the stay-green index

To derive the stay-green index (SGI) from the NDVI time series dataset, the variation trends of NDVI extracted from UAV multispectral images captured at different growth stages and experimental plots were analyzed (Fig. 7). The results showed that as maize approached

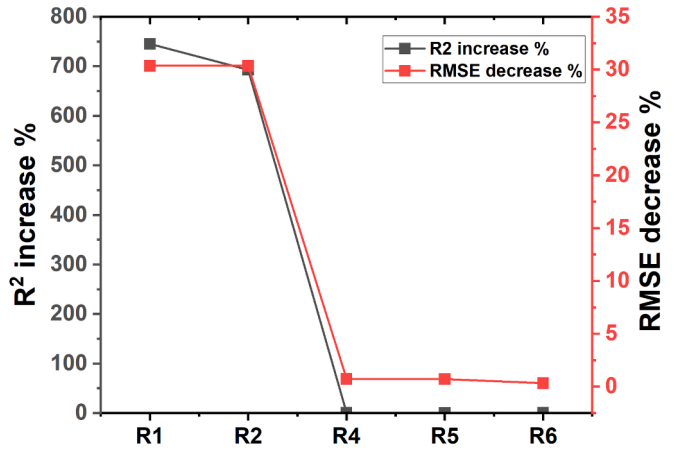

Fig. 10. The accuracy improvement after the addition of SGI to the yield estimation model.

maturity, the NDVI values gradually decreased in all experimental plots. The NDVI values at the R1 and R2 stages remained stable, around 0.9, while those at the R4, R5, and R6 stages showed increasing fluctuations between 0.3 and 0.9. This is illustrated in Fig. 7, where the black and red broken lines were concentrated in a small range, while the blue, green, and purple broken lines showed increasing fluctuations. Moreover, the differences in NDVI values between the R2 and R1 stages were not significant, but the differences between the R4, R5, and R6 stages and the R1 stage were more noticeable across different experimental plots.

The correlation analysis results of SGI and the measured yield under each threshold are shown in Table 2, and the results showed that the SGI under the selected thresholds was significantly correlated with the measured yield. The highest correlation coefficient was observed when NDVI was greater than 0.65, so this value was ultimately adopted as the threshold for NDVI in the study.

The SGI extraction threshold was determined, and the SGI calculation results of each experimental plot are shown in Fig. 8. To better describe the relationship between SGI and the measured yield, a line graph of the measured yield in each experimental plot was added to Fig. 8, which shows that the differences between the measured yield and the constructed SGI at different experimental plots were consistent. The experimental plots with high yields tended to have high SGI, while the experimental plots with low yields tended to have low SGI. Therefore, the SGI proposed in this study can well measure the yield change in different experimental plots.

3.3. Performance of the yield estimation model incorporating SGI

A model for estimating the yield of maize at various growth stages was developed using the selected 18 VIs and the extracted SGI. Two estimation models were created for each growth stage, and the only difference between them was whether SGI was included as part of the input parameters. Table 3 presents the accuracy of each model. For the different yield estimation models without SGI at different growth stages, the results showed that the performance of the M_{R1} (referred to as the yield estimation model without SGI at the R1 stage) performed poorly with R^2, RMSE, and rRMSE values of 0.064, 1.388 t/ha, and 20.1 %, respectively. Similar to the R1 stage, the model was unusable at the R2 stage. However, the models at the R4, R5, and R6 stages showed a quantitative improvement in accuracy indicators, with R^2 as high as 0.5, RMSE approaching 0.9 t/ha, and rRMSE approaching 13 %. Among all models that solely utilized VIs as parameters, M_{R6} exhibited the highest accuracy, while M_{R1} provided the lowest accuracy, which indicated the precision of yield estimation models improved as the maturity stage of maize approached. The inclusion of SGI at various stages resulted in an analogous performance for all models. Nevertheless, the M_{R4_SGI} model, which incorporated SGI at the R4 stage, demonstrated the most superior accuracy ($R^2 = 0.574$, RMSE = 0.939 t/ha, rRMSE = 13.6 %), while the other four models showed slightly inferior performance. These outcomes indicated that SGI played a crucial role in the accurate estimation of yield during the early stages. Moreover, Fig. 9 illustrates the scatter plot of each model incorporating SGI at different growth stages. To ensure better comparability between the performance of models at different growth stages, both their training and testing datasets were collated. The scatter in each graph was primarily clustered around the 1:1 line, with only a few scattered points found in the testing dataset that were considerably distanced from the 1:1 line.

To more intuitively illustrate the significance of SGI as an input parameter in the yield estimation model, the improvement in model validation dataset accuracy with the addition of SGI is presented in Fig. 10, which shows the percentage increase of R^2 and the percentage decrease of RMSE. The results indicated that SGI contributed significantly to the yield estimation models at the R1 and R2 stages, as seen by the increase of R^2 (745 % and 693 %) and decrease of RMSE (30 % and 30 %). However, at the R4, R5, and R6 stages, adding SGI to the corresponding models only slightly improved their performance, which suggested that SGI's impact on these models was not as vital. In addition, the accuracy of the yield estimation was assessed by comparing the residuals between the estimated and measured yields of the total dataset. Fig. 11 shows the residuals from yield estimation models at each growth stage, with and without SGI. The distribution of the residuals for the yield estimation models including SGI displayed a higher σ (standard deviation) than those without SGI, and μ (mean) showed a similar trend, except for the model corresponding to the R2 stage. Comprehensively, the models involving SGI were able to reduce extreme overestimation and underestimation.

4. Discussion

4.1. Relationships between VIs and the measured yield

We used thirty VIs to test their correlation with the measured yield over different growing stages (Fig. 5). As the growth stage progressed towards maturity (i.e., the vegetation coverage decreased), the correlation between VIs and measured yield increased, which aligned with the finding of Qiao et al. (2022). In their study, they discussed the correlation between 36 VIs and maize Chl content under varying vegetation coverage levels, resulting in the observation that the correlation between low vegetation coverage and maize Chl content was the strongest. The correlation analysis method employed in this study was linear; however, some of the VIs were derived from the nonlinear combination of reflectance between different bands (Xue and Su, 2017). Apart from linear relationships, there may be nonlinear response relationships between VIs and measured yield. In addition, unlike other studies that only selected 1 ~ 3 VIs for yield estimation (Ma et al., 2021; Nagy et al., 2021; Yang et al., 2022), this study retained 18 VIs after filtering out 12 VIs according to the correlation analysis. This was because, despite a high level of autocorrelation among VIs, Zeng et al. (2022) illustrated that each VI contributes differently to the research objective. Thus, combining multiple VIs can improve the accuracy of yield estimation.

4.2. The construction of SGI

The SGI constructed in this paper was determined by the NDVI and the corresponding threshold value. First, Christopher et al. (2016) demonstrated that the green parameter obtained by NDVI was a powerful predictor of grain yield under different environmental types; Rebetzke et al. (2016) concluded that NDVI is a typical surrogate for SG; and Latif et al. (2020) divided SG into four types based on NDVI. Thus, it is reasonable to utilize NDVI for calculating SGI. Secondly, by taking into account both theoretical knowledge and actual field conditions, we

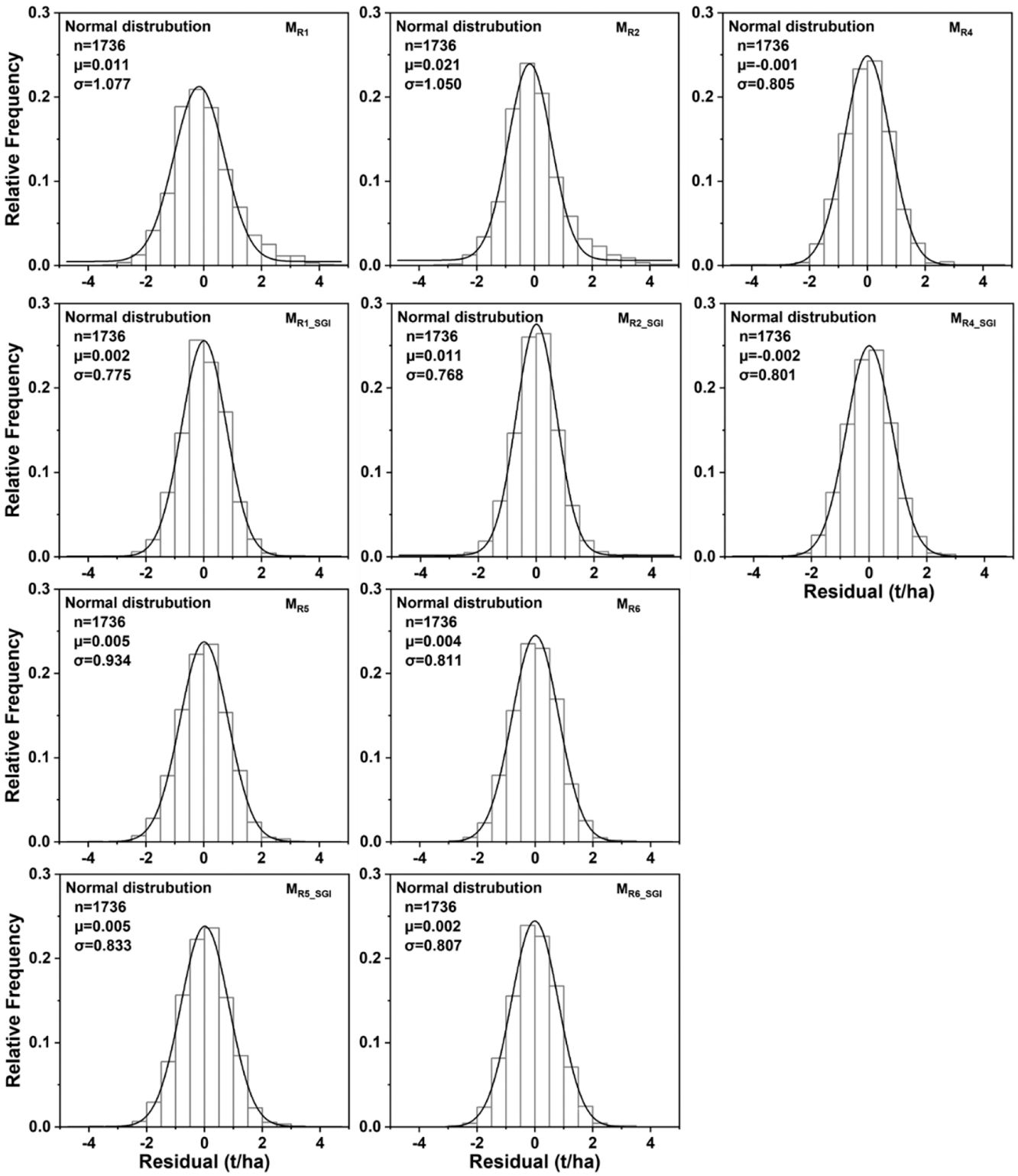

Fig. 11. The frequency histogram of residual values from the yield estimation model with or without SGI at different growth stages.

have narrowed down the range of the NDVI threshold to 0.45–0.70. We then took stepwise measurements with an interval of 0.05 across this range, and calculated the correlation between each threshold and the measured yield. As a result, we selected the threshold of NDVI = 0.65, which showed the highest correlation, as the basis for this study. The method of determining the threshold had been carefully considered and deemed appropriate for the purpose of this study.

The stability of NDVI in different field conditions was crucial for accurately monitoring maize growth and calculating SGI (Yao et al., 2017). However, in this study, the flight mission was initiated at the R1 stage, when canopy coverage had reached its peak and the NDVI signal had become saturated (Tan et al., 2020), resulting in little change in NDVI values over a certain period of time. As shown in Fig. 6, NDVI values remained high and consistent across different plots corresponding to R1 and R2 stages. Additionally, as the threshold value for the study is determined based on actual field conditions and the degree of

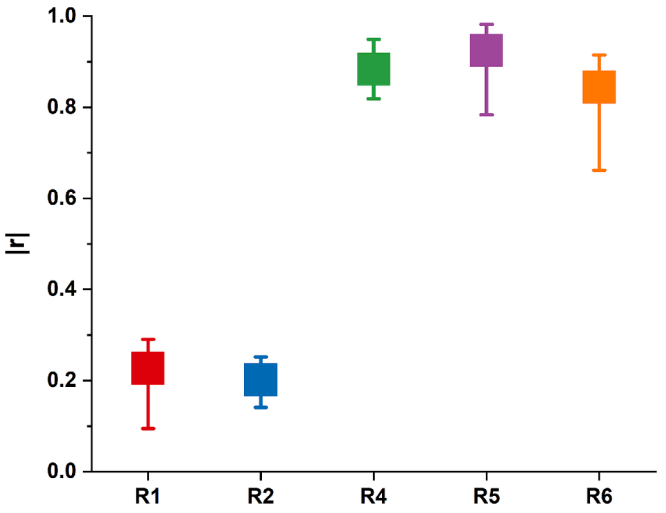

Fig. 12. The interval plot of the absolute correlation coefficient (|r|) between the SGI and VIs. The upper and lower lines represented the maximum and minimum values of |r|.

correlation with actual yield, there is no consistency in threshold values across different datasets, which reduces the universality of the yield estimation model. Moreover, the selected threshold represents the detailed situation of the entire field but not that of each experimental plot. Therefore, we proposed combining the stay-green of crops with phenology for future studies, in order to create a more organic and rational index. This approach can take into account the variability in phenology among different varieties and fertilization treatments and can potentially improve accuracy in crop yield estimation. (Schlund and Erasmi, 2020; Yang, 2023).

According to the results of this paper, the introduction of SGI demonstrates significant performance improvements in estimating yield models based on spectral indices during the early stages of grain filling. However, its effectiveness decreases in later stages. As a result, SGI needs to observe the whole corn filling process in the construction process (at least before NDVI decreases to a certain threshold), which limits the application of SGI in the early stage. However, with the accumulation of future test data, we are also trying to estimate SGI before NDVI reaches the aging threshold through HANTS or SG time series reconstruction, and even the transformer time series prediction model.

4.3. Maize yield estimation model

The yield estimation model based on VIs in Table 3 shows an improvement in accuracy as the growth process progresses, reaching the best performance at the R4 stage, with the variation range of the model's accuracy remaining small since the beginning of the R4 stage. This suggested that the model with only VIs could accurately estimate maize yield at the R4 stage, which was about one month before harvest, with an rRMSE of 13.7 %. The optimal time for estimating maize yield from Li et al. (2021b) was similar to ours but with slightly lower accuracy (rRMSE = 17.9 %).

The inclusion of SGI in the model significantly improved the yield estimation accuracy at R1 and R2 stages, with improvements of R^2 up to 745 % and 693 %, and a decrease of RMSE by 30 % and 30 %, respectively, while the contribution to the yield estimation at the R4, R5, and R6 stages was not significant (Fig. 10). This improvement was mainly attributed to the increasing correlation between SGI and VIs used in the model construction as the growth stage progressed (Fig. 12). The changes in the correlation between SGI and VIs throughout the various growth stages in the interval plot were comparable to the changes in the estimated yield accuracy at the different growth stages when SGI was incorporated into the models (as shown in Fig. 12). This meant that the correlation results at the R1 and R2 stages were quite low, with the highest |r| being merely 0.3. However, the accuracy of the model at these two stages was significantly increased due to the presence of SGI. The possible explanation for this outcome could be that SGI can provide additional information regarding maize yield from diverse perspectives and there was no significant overlap with VIs in the model. The correlation results at R4, R5, and R6 stages were highly consistent, with the highest |r| almost at par with 1, and the improvement in accuracy of the corresponding model was the least significant. This could be attributed to the collinearity of SGI and VIs being too strong, thereby diluting the influence of SGI on the RF algorithm. This also indirectly affirms that future studies should attempt to construct SGI from various perspectives as its combination with VI could lead to certain drawbacks.

In any case, the study found that the SGI proposed in the study significantly improved the ability to estimate yield at the R1 and R2 stages. Additionally, the accuracy of yield estimation remained consistent throughout the five stages of growth, leading to the optimal yield estimation at the R1 stage. Accurately estimating crop yield prior to harvesting is essential for farmers (Panek and Gozdowski, 2021; Rashid et al., 2021), and fortunately, the yield estimation model framework proposed in this paper can estimate yield accurately up to two months in advance. However, further research is needed to optimize the SGI construction method in order for it to play a more significant role in predicting crop yields at various growth stages.

5. Conclusion

Advances in UAV technology offer new opportunities for improving the ability to accurately estimate maize yields at the field level. This study utilized VIs that were extracted from UAV multispectral images to construct the maize yield estimation model. To address errors arising from instantaneous indices, we focused on developing a new relative index, the SGI, which was calculated from the time series NDVI fitting curve. The results lead to the following main conclusions: (1) The correlation between the 30 commonly used VIs and the actual maize yield varied in each growth stage, but as the growth stage approached maturity, the correlation gradually increased and was maintained at a high level. (2) The SGI proposed in this study had good consistency with the actual maize yield variation and was significantly correlated with the yield. (3) The yield estimation model that solely utilized VIs as parameters at different growth stages showed that M_{R4} exhibited the highest accuracy, with an R^2 of 0.569, an RMSE of 0.944 t/ha, and an rRMSE of 13.7 %. The inclusion of SGI in the yield estimation model improved the yield estimation ability of maize at different growth stages to varying degrees, resulting in the yield estimation accuracy of maize at different growth stages being comparable, in which the R^2 varied from 0.557 to 0.584 and the rRMSE varied from 13.3 % to 12.9 % at various maize growth stages. These findings provide a reference for the rapid and non-destructive estimation of the maize yield by UAV remote sensing and support for precise agriculture management. This study highlights the potential of multi-spectral remote sensing applications, as exemplified by the SGI, for estimating maize yield in future agricultural production.

CRediT authorship contribution statement

Yuan Liu: Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing. **Kaijian Fan:** Data curation, Formal analysis, Software, Validation, Writing – review & editing. **Lin Meng:** Formal analysis, Investigation. **Chenwei Nie:** Conceptualization, Methodology. **Yadong Liu:** Resources, Validation. **Minghan Cheng:** Resources, Supervision. **Yang Song:** Conceptualization, Methodology, Supervision, Validation. **Xiuliang Jin:** Conceptualization, Funding acquisition, Project administration, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by Nanfan special project, CAAS (ZDXM2310, YJTCY01, YBXM01), Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Agricultural Sciences (CAAS-ZDRW202107,Y2022XK22), National Natural Science Foundation of China (32301395, 42071426, 51922072, 51779161, 51009101), Key Cultivation Program of Xinjiang Academy of Agricultural Sciences (xjkcpy-2020003), Research and application of key technologies of smart brain for farm decision-making platform (2021ZXJ05A03), the Agricultural Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences, The State Key Laboratory of Water Resources and Hydropower Engineering Science (2021NSG01), and Special Fund for Independent Innovation of Agricultural Science and Technology in Jiangsu, China (CX(21)3065).

Data availability

The authors are unable or have chosen not to specify which data has been used.

References

Barbosa, B.D.S., et al., 2021. Uav-based coffee yield prediction utilizing feature selection and deep learning. Smart Agric. Technol. 1, 100010.
Berni, J., Zarco-Tejada, P.J., Suarez, L., Fereres, E., 2009. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Trans. Geosci. Remote Sens. 47 (3), 722–738.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sens. Environ. 76 (2), 156–172.
Cao, Q., et al., 2013. Non-destructive estimation of rice plant nitrogen status with crop circle multispectral active canopy sensor. Field Crop Res 154, 133–144.
Chen, W., et al., 2012. In-flight absolute radiometric calibration of uav multispectral sensor. Spectrosc. Spectr. Anal. 32 (12), 3169–3174.
Chen, J.M., 2014. Evaluation of vegetation indices and a modified simple ratio for boreal applications. Can. J. Remote. Sens. 22 (3), 229–242.
Cheng, M., et al., 2022. Using multimodal remote sensing data to estimate regional-scale soil moisture content: A case study of beijing, china. Agric Water Manag 260, 107298.
Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput. Electron. Agric. 151, 61–69.
Christopher, J.T., et al., 2014. Phenotyping novel stay-green traits to capture genetic variation in senescence dynamics. Funct. Plant Biol. 41 (11), 1035–1048.
Christopher, J.T., Christopher, M.J., Borrell, A.K., Fletcher, S., Chenu, K., 2016. Stay-green traits to improve wheat adaptation in well-watered and water-limited environments. J. Exp. Bot. 67 (17), 5159–5172.
Dash, J., Curran, P.J., 2010. The meris terrestrial chlorophyll index. Int. J. Remote Sens. 25 (23), 5403–5413.
Fitzgerald, G., Rodriguez, D., O'Leary, G., 2010. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index—the canopy chlorophyll content index (ccci). Field Crop Res 116 (3), 318–324.
Fu, H., Wang, C., Cui, G., She, W., Zhao, L., 2021. Ramie yield estimation based on uav rgb images. Sensors (basel) 21 (2), 669.
Furukawa, F., Maruyama, K., Saito, Y.K., Kaneko, M., 2020. Corn height estimation using uav for yield prediction and crop monitoring. In: Avtar, R., Watanabe, T. (Eds.), Unmanned Aerial Vehicle: Applications in Agriculture and Environment. Springer International Publishing, Cham, pp. 51–69.
Gao, J., et al., 2019. Effects of climate change on the extension of the potential double cropping region and crop water requirements in northern china. Agric. For. Meteorol. 268, 146–155.
García-Martínez, H., et al., 2020. Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and rgb images acquired with unmanned aerial vehicles. Agriculture 10 (7), 277.
Gitelson, A.A., Gritz †, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. Journal of Plant Physiology, 160(3): 271-282.
Gitelson, A.A., Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. Int. J. Remote Sens. 18 (12), 2691–2697.
Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys Res Lett 32 (8), L08403.
Greco, M., Chiappetta, A., Bruno, L., Bitonti, M.B., 2012. In posidonia oceanica cadmium induces changes in DNA methylation and chromatin patterning. J. Exp. Bot. 63 (2), 695–709.
Guo, Y., et al., 2022. Machine learning-based approaches for predicting spad values of maize using multi-spectral images. Remote Sens. (Basel) 14 (6), 1337.
Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 81 (2), 416–426.
Huete, A.R., 1988. A soil-adjusted vegetation index (savi). Remote Sens. Environ. 25 (3), 295–309.
Huete, A., et al., 2002. Overview of the radiometric and biophysical performance of the modis vegetation indices. Remote Sens. Environ. 83 (1), 195–213.
Hunt, E.R., et al., 2013. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. Int. J. Appl. Earth Obs. Geoinf. 21, 103–112.
Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112 (10), 3833–3845.
Kataoka, T., Kaneko, T., Okamoto, H. and Hata, S., 2003. Crop growth estimation system using machine vision, Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), pp. b1079-b1083 vol.2.
Khaki, S., Pham, H., Wang, L., 2021. Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning. Sci. Rep. 11 (1), 11132.
Latif, S., et al., 2020. Deciphering the role of stay-green trait to mitigate terminal heat stress in bread wheat. Agronomy 10 (7), 1001.
Li, D., et al., 2020. Improved estimation of leaf chlorophyll content of row crops from canopy reflectance spectra through minimizing canopy structural effects and optimizing off-noon observation time. Remote Sens. Environ. 248, 111985.
Li, D., et al., 2021a. Improving potato yield prediction by combining cultivar information and uav remote sensing data using machine learning. Remote Sens. (Basel) 13 (16), 3322.
Li, L., et al., 2021b. Crop yield forecasting and associated optimum lead time analysis based on multi-source environmental data across china. Agric. For. Meteorol. 308–309, 108558.
Li, X., Li, X., Liu, W., Wei, B., Xu, X., 2021c. A uav-based framework for crop lodging assessment. Eur. J. Agron. 123, 126201.
Lu, J., et al., 2017. Field detection of anthracnose crown rot in strawberry using spectroscopy technology. Comput. Electron. Agric. 135, 289–299.
Lu, N., et al., 2019. Estimation of nitrogen nutrition status in winter wheat from unmanned aerial vehicle based multi-angular multispectral imagery. Front. Plant Sci. 10, 1601.
Lu, B., He, Y., 2017. Optimal spatial resolution of unmanned aerial vehicle (uav)-acquired imagery for species classification in a heterogeneous grassland ecosystem. Giscience & Remote Sensing 55 (2), 205–220.
Ma, Y., Zhang, Z., Kang, Y., Özdoğan, M., 2021. Corn yield prediction and uncertainty analysis based on remotely sensed variables using a bayesian neural network approach. Remote Sens. Environ. 259, 112408.
Maimaitijiang, M., et al., 2020. Soybean yield prediction from uav using multimodal data fusion and deep learning. Remote Sens. Environ. 237, 111599.
Mirosavljević, M., et al., 2020. Changes in stay-green and nitrogen use efficiency traits in historical set of winter barley cultivars. Field Crop Res 249, 107740.
Montazeaud, G., et al., 2016. Predicting wheat maturity and stay–green parameters by modeling spectral reflectance measurements and their contribution to grain yield under rainfed conditions. Field Crop Res 196, 191–198.
Mueller, N.D., et al., 2012. Closing yield gaps through nutrient and water management. Nature 490 (7419), 254–257.
Nagy, A., Szabó, A., Adeniyi, O.D., Tamás, J., 2021. Wheat yield forecasting for the tisza river catchment using landsat 8 ndvi and savi time series and reported crop statistics. Agronomy 11 (4), 652.
Nie, C., et al., 2023. A comparison of methods to estimate leaf area index using either crop-specific or generic proximal hyperspectral datasets. Eur. J. Agron. 142, 126664.
Pádua, L., et al., 2017. Uas, sensors, and data processing in agroforestry: A review towards practical applications. Int. J. Remote Sens. 38 (8–10), 2349–2391.
Panek, E., Gozdowski, D., 2021. Relationship between modis derived ndvi and yield of cereals for selected european countries. Agronomy 11 (2), 340.
Peng, X., Han, W., Ao, J., Wang, Y., 2021. Assimilation of lai derived from uav multispectral data into the safy model to estimate maize yield. Remote Sens. (Basel) 13 (6), 1094.
Penuelas, J., Baret, F., Filella, I., 1995. Semiempirical indexes to assess carotenoids chlorophyll-a ratio from leaf spectral reflectance. Photosynthetica 31 (2), 221–230.
Qiao, L., et al., 2022. Uav-based chlorophyll content estimation by evaluating vegetation index responses under different crop coverages. Comput. Electron. Agric. 196, 106775.
Ranjan, R., Chopra, U.K., Sahoo, R.N., Singh, A.K., Pradhan, S., 2012. Assessment of plant nitrogen stress in wheat (triticum aestivum l.) through hyperspectral indices. Int. J. Remote Sens. 33 (20), 6342–6360.
Rashid, M., Bari, B.S., Yusup, Y., Kamaruddin, M.A., Khan, N., 2021. A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction. IEEE Access 9, 63406–63439.
Rebetzke, G.J., Jimenez-Berni, J.A., Bovill, W.D., Deery, D.M., James, R.A., 2016. High-throughput phenotyping technologies allow accurate selection of stay-green. J. Exp. Bot. 67 (17), 4919–4924.
Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55 (2), 95–107.
Roujean, J.-L., Breon, F.-M., 1995. Estimating par absorbed by vegetation from bidirectional reflectance measurements. Remote Sens. Environ. 51 (3), 375–384.

Rouse, J., Haas, R.H., Schell, J.A. and Deering, D., 1973. Monitoring vegetation systems in the great plains with erts, pp. 309-317.
Sakamoto, T., 2020. Incorporating environmental variables into a modis-based crop yield estimation method for united states corn and soybeans through the use of a random forest regression algorithm. ISPRS J. Photogramm. Remote Sens. 160, 208–228.
Schlund, M., Erasmi, S., 2020. Sentinel-1 time series data for monitoring the phenology of winter wheat. Remote Sens. Environ. 246, 111814.
Schneider, P., Roberts, D.A., Kyriakidis, P.C., 2008. A vari-based relative greenness from modis data for computing the fire potential index. Remote Sens. Environ. 112 (3), 1151–1167.
Schwalbert, R.A., et al., 2020. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern brazil. Agric. For. Meteorol. 284, 107886.
Simic, A., et al., 2014. Testing the top-down model inversion method of estimating leaf reflectance used to retrieve vegetation biochemical content within empirical approaches. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (1), 92–104.
Sims, D.A., 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 (2), 337–354.
Singh, B., Sihag, P., Singh, K., 2017. Modelling of impact of water quality on infiltration rate of soil by random forest regression. Model. Earth Syst. Environ. 3 (3), 999–1004.
Tan, C.W., et al., 2020. Quantitative monitoring of leaf area index in wheat of different plant types by integrating ndvi and beer-lambert law. Sci. Rep. 10 (1), 929.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150.
Vannoppen, A., Gobin, A., 2021. Estimating farm wheat yields from ndvi and meteorological data. Agronomy 11 (5), 946.
Wan, L., 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. For. Meteorol. 291, 108096.
Wang, F., et al., 2021. Combining spectral and textural information in uav hyperspectral images to estimate rice grain yield. Int. J. Appl. Earth Obs. Geoinf. 102, 102397.
Wang, Z., 2021. Detection and analysis of degree of maize lodging using uav-rgb image multi-feature factors and v arious classification methods. ISPRS Int. J. Geo Inf. 10, 309.
Wang, Y., Zhang, Z., Feng, L., Du, Q., Runge, T., 2020. Combining multi-source data and machine learning approaches to predict winter wheat yield in the conterminous united states. Remote Sens. (Basel) 12 (8), 1232.
Wu, S., Yang, P., Ren, J., Chen, Z., Li, H., 2021. Regional winter wheat yield estimation based on the wofost model and a novel vw-4densrf assimilation algorithm. Remote Sens. Environ. 255, 112276.
Xie, C., Yang, C., 2020. A review on plant high-throughput phenotyping traits using uav-based sensors. Comput. Electron. Agric. 178, 105731.
Xu, W., et al., 2021. Cotton yield estimation model based on machine learning using time series uav remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 104, 102511.
Xu, L., et al., 2022. An improved approach to estimate ratoon rice aboveground biomass by integrating uav-based spectral, textural and structural features. Precis. Agric. 23, 1276–1301.
Xue, L., Cao, W., Luo, W., Dai, T., Zhu, Y., 2004. Monitoring leaf nitrogen status in rice with canopy spectral reflectance. Agron. J. 96 (1), 135–142.
Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: A review of developments and applications. Journal of Sensors 2017, 1–17.
Yang, W., et al., 2021. Estimation of corn yield based on hyperspectral imagery and convolutional neural network. Comput. Electron. Agric. 184, 106092.
Yang, B., et al., 2022. The optimal phenological phase of maize for yield prediction with high-frequency uav remote sensing. Remote Sens. (Basel) 14 (7), 1559.
Yang, Q., et al., 2023. Regulating the time of the crop model clock: A data assimilation framework for regions with high phenological heterogeneity. Field Crop Res 293, 108847.
Yao, X., et al., 2017. Estimation of wheat lai at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. Remote Sens. (Basel) 9 (12), 1304.
Yu, N., et al., 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.
Yue, J., Feng, H., Li, Z., Zhou, C., Xu, K., 2020. Mapping winter-wheat biomass and grain yield based on a crop model and uav remote sensing. Int. J. Remote Sens. 42 (5), 1577–1601.
Zeng, Y., et al., 2022. Optical vegetation indices for monitoring terrestrial ecosystems globally. Nat. Rev. Earth Environ. 3 (7), 477–493.
Zhang, Z., Jin, Y., Chen, B., Brown, P., 2019. California almond yield prediction at the orchard level with a machine learning approach. Front. Plant Sci. 10, 809.
Zhang, M., Zhou, J., Sudduth, K.A., Kitchen, N.R., 2020. Estimation of maize yield and effects of variable-rate nitrogen application using uav-based rgb imagery. Biosyst. Eng. 189, 24–35.