

## Article

# Improving Satellite-Based Retrieval of Maize Leaf Chlorophyll Content by Joint Observation with UAV Hyperspectral Data

Siqi Yang<sup>1,2</sup>, Ran Kang<sup>1,2,\*</sup>, Tianhe Xu<sup>1,2</sup>, Jian Guo<sup>1,2</sup>, Caiyun Deng<sup>1,2</sup>, Li Zhang<sup>1,2</sup>, Lulu Si<sup>3</sup>  
and Hermann Josef Kaufmann<sup>2,4</sup>

<sup>1</sup> School of Space Science and Technology, Shandong University, Weihai 264209, China; yangsq147@mail.sdu.edu.cn (S.Y.); thxu@sdu.edu.cn (T.X.); dengcaiyun@sdu.edu.cn (C.D.); zhangli\_2000@mail.sdu.edu.cn (L.Z.)

<sup>2</sup> Shandong Key Laboratory of Optical Astronomy and Solar-Terrestrial Environment, School of Space Science and Physics, Shandong University, Weihai 264209, China

<sup>3</sup> Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, 46022 Valencia, Spain; lsi@doctor.upv.es

<sup>4</sup> Institute of Space Sciences, Shandong University, Weihai 264209, China

\* Correspondence: ran\_kang@sdu.edu.cn

**Abstract:** While satellite-based remote sensing offers a promising avenue for large-scale LCC estimations, the accuracy of evaluations is often decreased by mixed pixels, attributable to distinct farming practices and diverse soil conditions. To overcome these challenges and to account for maize intercropping with soybeans at different growth stages combined with varying soil backgrounds, a hyperspectral database for maize was set up using a random linear mixed model applied to hyperspectral data recorded by an unmanned aerial vehicle (UAV). Four methods, namely, Euclidean distance, Minkowski distance, Manhattan distance, and Cosine similarity, were used to compare vegetation spectra from Sentinel-2A with the newly constructed database. In a next step, widely used vegetation indices such as NDVI, NAOC, and CAI were tested to find the optimum method for LCC retrieval, validated by field measurements. The results show that the NAOC had the strongest correlation with ground sampling information ( $R^2 = 0.83$ ,  $RMSE = 0.94 \mu\text{g}/\text{cm}^2$ , and  $MAE = 0.67 \mu\text{g}/\text{cm}^2$ ). Additional field measurements sampled at other farming areas were applied to validate the method's transferability and generalization. Here too, validation results showed a highly precise LCC estimation ( $R^2 = 0.93$ ,  $RMSE = 1.10 \mu\text{g}/\text{cm}^2$ , and  $MAE = 1.09 \mu\text{g}/\text{cm}^2$ ), demonstrating that integrating UAV hyperspectral data with a random linear mixed model significantly improves satellite-based LCC retrievals.

**Keywords:** UAV hyperspectral; chlorophyll content; maize intercropping; random linear mixed model; spectral matching



**Citation:** Yang, S.; Kang, R.; Xu, T.; Guo, J.; Deng, C.; Zhang, L.; Si, L.; Kaufmann, H.J. Improving Satellite-Based Retrieval of Maize Leaf Chlorophyll Content by Joint Observation with UAV Hyperspectral Data. *Drones* **2024**, *8*, 783. <https://doi.org/10.3390/drones8120783>

Academic Editor: Yangquan Chen

Received: 25 October 2024

Revised: 11 December 2024

Accepted: 21 December 2024

Published: 23 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Maize, one of the most widely cultivated crops worldwide, intercropping with soybean has become a common practice for farmers to enhance soil fertility and increase yield and profitability. However, satellite monitoring of leaf chlorophyll content (LCC) in intercropping systems poses severe challenges. In particular, the similar growth characteristics of maize and soybean result in recording mixed pixels, leading to a less precise LCC estimation. But accuracy is crucial for a timely assessment of crop growth stages in global food security and the economy [1]. Crop growth is usually described by various indicators, among which LCC is crucial as it is the essential base for photosynthesis in plants, significantly influencing the photosynthetic rate and vegetation productivity [2–4]. High-precision measurements of LCC have typically been made through plot-based field sample collection and chemical analysis methods, which are disruptive, time-consuming, and challenging to apply in large-scale precision mapping [3]. In contrast, remote sensing

techniques and methods offer an efficient solution for a nondestructive monitoring of biochemical parameters due to their ability to record unique spectral characteristics in the visible and near-infrared (VIS/NIR) wavelength regions [1,4]. Besides ground-based measurements, remote sensing platforms can be broadly categorized into satellite, aircraft, and unmanned aerial vehicle (UAV) systems [5]. Satellite remote sensing provides comprehensive, large-scale, and continuous monitoring of the Earth's surface. It can promptly provide crop distribution information, making it an ideal tool for monitoring the growth conditions of large areas [6,7]. However, the technical trade-off between high spatial and spectral resolution limits the development of LCC satellite monitoring accuracy. Satellite-derived multispectral data with a moderate spatial resolution such as Sentinel-2 already reduce these accuracies significantly [8,9]. Moreover, intercropping of different species and varying fractions and types of soils result in mixed pixels and thus further complicate LCC estimations.

Research has suggested that enhancing the spatiotemporal resolution is necessary for the accurate quantification of LCC [10,11]. To overcome this technical constraint of satellite observations at moderate spatial resolutions, UAV platforms equipped with hyperspectral sensors can be used. They provide benefits such as a closer alignment with satellite monitoring as compared to laboratory measurements; a more flexible data acquisition; and, most importantly, a higher spectral and spatiotemporal resolution [11–14]. On the other hand, the main limitation of UAV deployment is the coverage, which is constrained by feasible flight altitudes, their battery capacity, and the weight they can carry [15]. Additionally, there are costs associated with UAV hyperspectral data acquisition and processing, and all procedures are very time-consuming. Operational challenges, such as weather conditions, regulatory restrictions, and terrain limitations also restrict UAV monitoring, reducing the flexibility and scalability of widespread agricultural applications. As a result, most accurately estimating maize LCCs considering intercrop situations based on open-access satellite images remains an ongoing challenge.

Vegetation exhibits unique spectral characteristics that are useable for LCC detection. The blue and red wavelength regions are associated with chlorophyll-a,b and carotenoid absorptions needed for photosynthesis. In contrast, NIR wavelengths are strongly reflected by the mesophyll (sponge parenchyma in the leaves). When vegetation experiences stress from various disturbances, the reflectance values in the visible and NIR bands change, and the red edge shifts slightly to a shorter wavelength (blue-shift). Additionally, two distinct bands at 950 and 1180 nm can be used to detect water stress.

Existing approaches to LCC estimations can be divided into physically based, empirical, and hybrid methods [7]. Physically based approaches rely on detailed radiative transfer models that simulate light propagation within single leaves as well as plant canopies. Thereby, models such as PROSPECT, PROSAIL, and WOFOST provide a comprehensive understanding of biophysical parameters that influence the spectral properties of vegetation. But the complexity and computational demands of these models are challenging for large-scale and real-time applications [16–18]. Empirical methods are generally based on statistical relationships between measured and observed characteristics and in the case of chlorophyll content are often used to develop vegetation indices. These methods are relatively simple to implement and can provide quite accurate estimations. Some researchers developed hybrid methods that combine physically based approaches with computationally efficient algorithms (e.g., machine learning) [11,19]. Despite the advantages, hybrid methods are more complex than empirical methods due to the integration of various datasets and models. The sensitivity to varying model parameters makes it difficult to interpret the results [20,21].

Numerous studies have attempted to create various vegetation indices. According to their functions, they can be categorized into three types: general, improved, and functional vegetation indices (FVI). General vegetation indices, such as the normalized difference vegetation index (NDVI), were first proposed by Kriegler et al. [22] and are still widely used in assessing vegetation health and biomass [23]. Their ability to correctly estimate LCCs is

limited as they tend to saturate at high levels of chlorophyll content, are less sensitive to chlorophyll variations, and prone to soil/rock mineralogy [24]. In contrast, the difference vegetation index (DVI) [25] is more responsive to changes in vegetation density but also susceptible to chlorophyll saturation and soil brightness [26,27]. Improved vegetation indices such as the enhanced vegetation index (EVI) [28,29], the chlorophyll vegetation index (CVI) [30], and the red-edge chlorophyll index ( $CI_{red-edge}$ ) [31] have been proposed accordingly. These reduce atmospheric and soil disturbances to some extent in correlation with LCCs and can be applied to a broader range of environmental conditions [32,33], but they remain sensitive to the effects of soil type and coverage, and they are influenced by the leaf structure [34–36]. This issue is partly resolved by FVIs, such as the soil-adjusted vegetation index (SAVI) [37] and the modified chlorophyll absorption in the reflectance index (MCARI) [38], which are sensitive to chlorophyll content but less influenced by the respective structure of different canopies. However, these indices are limited in reducing the influence of environmental backgrounds [39,40] as compared to indices such as the normalized area over the reflectance curve (NAOC) and the chlorophyll absorption integral (CAI). Also, methods based on integrals are more sensitive to changes in chlorophyll content. Their performance through included red-edge bands demonstrate a superior ability to effectively mitigate the issue of vegetation indices (VIs) being susceptible to saturation in LCC estimation [41,42]. Hereby, the normalized area calculation relies on the integral between the red and the red-edge bands. Hence, the NAOC or CAI approaches are better suited for hyperspectral data than for multispectral satellite data. LCC has shown a strong linear correlation with FVIs, but the FVI chlorophyll models exhibited significant inversion errors throughout the growing season [43]. Consequently, despite the numerous FVIs already developed, there are further future challenges for improving satellite-based LCC monitoring.

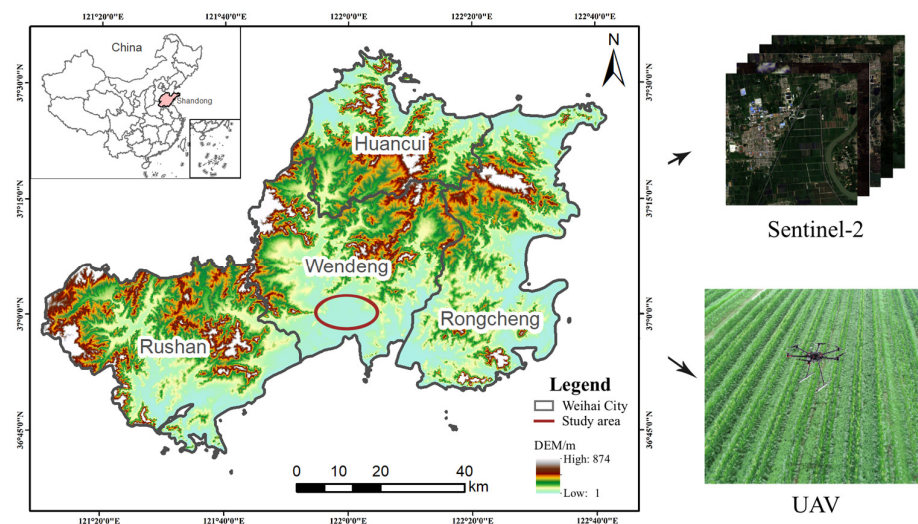
Spectral matching, a method commonly used in remote sensing to compare and identify the similarity between spectra from different sources, presents a promising approach to enhance the precision of LCC estimation. It enables the identification of an unknown target material by selecting pure pixels in remote sensing images by searching spectral libraries, based on automated processes of end-member extraction and material abundance estimation [44]. Also, various similarity measurements have been successfully applied for image retrieval. Gupta and Rajan (2010) employed matching algorithms such as the Euclidean distance (ED) to study the temporal shifts in vegetation and crop cross pixels based on different geographical locations [45]. Similarly, Li et al. (2021) [46] achieved a high precision of spatial distribution of wheat by utilizing ED and Manhattan distance (MD) due to their ability to capture global similarity trends. For applications of high dimensional target detection, particularly when targets are unpredictable in size and shape, a weighted Chebyshev distance (CD) method has shown improved efficiency in hyperspectral image classification [47]. Additionally, similarity measures based on spectral shapes, such as Cosine similarity (CS), originating from vector processing, have proven relatively reliable in target detection and classification [48,49]. Although spectral matching methods are well developed, there are ongoing efforts to evaluate the effectiveness of similarity measures between remote sensing multispectral data and UAV hyperspectral data.

Our research primarily aims to considerably improve the accuracy of satellite-based LCC estimations of maize by joint measurements accomplished by a hyperspectral UAV device. Thereby, we focus on the following objectives: (1) Exploring the impact of overlapping spectral characteristics from intercropping and soil backgrounds during key growth stages. (2) Creating an airborne (UAV) hyperspectral database of maize using a random linear model to mitigate the impact of the above mentioned interfering facts on satellite-based recordings. (3) Identifying a suitable spectral matching method and an appropriate FVI to increase the accuracy of satellite-based LCC estimation across different regions.

## 2. Materials and Methods

### 2.1. Study Area and Farming Background

The study area is located in the south of the Shandong Province (Figure 1). It is characterized by gently rolling hills and a moderate climate, and it lies at an average altitude of 60 m above sea level. The terrain is free from obstacles in the air, and thus it is an ideal experimental site for UAV data collection.



**Figure 1.** Sketch map of the study area.

In 2019, China issued ‘The National High-Standard Farmland Construction Plan’ (2021–2030), which proposes that China will create 80 million hectares of high-standard agricultural land by 2030. This initiative aims to ensure an annual stable grain production capacity of over 600 billion kilograms, thereby strengthening the foundation of the national food security [50]. Shandong Province is one of the 13 major grain production areas in China, primarily cultivating maize and wheat. Our study area, Zetou, is a high-yield field demonstration area in this province, with strip intercropping of nearly 100 hectares of maize and soybeans. Its primary cultivation practice is a rotation of winter wheat with summer maize, providing an optimal window for maize observation.

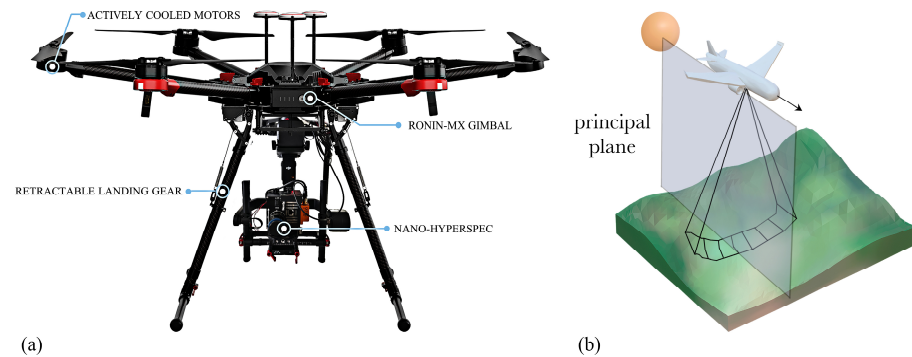
In this context, a meta-analysis exploited that leveraging species complementarities through intercropping maize and soybean significantly enhances land productivity while reducing the need for nitrogen fertilizers [51]. In an intercropping system, maize and soybean exhibit slightly distinct growth cycles that complement each other. Maize typically has a longer growth cycle, beginning with planting in early spring and reaching maturity in late summer or early fall, around 100 to 130 days after planting. Soybeans have a slightly shorter growth cycle, typically maturing in 90 to 120 days. Despite having distinct growth cycles, maize and soybean exhibit several similarities in their growth stages. In our test site, maize is seeded in late April or early May, while soybeans are usually sown mid of May. Both crops enter a rapid vegetative growth phase by mid-June or early July, reaching their peak canopy cover and biomass accumulation in August.

### 2.2. Data Sources and Processing

#### 2.2.1. UAV Based Hyperspectral Data Collection

The hyperspectral data collection was carried out with a ‘Matrice 600 Pro’, a six-rotor unmanned aerial vehicle (UAV) developed by ‘DJI Innovations’. It can lift payloads of up to 4.5 kg. For data acquisition, we operated a ‘Nano-Hyperspec’ camera (Headwall Photonics, Fitchburg, MA, USA). It is an integrated hyperspectral imaging sensor designed for visible and near infrared (VNIR) spectral recordings from 400 to 1000 nm and configured with a Ronin-MX gimbal that can electrically provide off-nadir recordings at different angles (Figure 2a). DJI’s ‘Ground Station Pro’ app was used for drone navigation, including flight

planning, setting up flight parameters, and managing teams and projects. All flights were directed along the principal plane to avoid hot spot issues [52] (Figure 2b). Considering optimum image coverage and high spatial resolution, the flight height was set at 50 m, achieving a pixel size of  $3.1 \times 3.1$  cm with the sensor aligned to the nadir.



**Figure 2.** (a) Headwall Nano-Hyperspec VNIR imaging sensor mounted on the Matrice 600 Pro. (b) Schematic diagram of the principal plane guiding flight passes of aircraft and UAV during recordings.

From 22 July to 23 September, five UAV hyperspectral experiments were conducted during the critical maize growth stages of jointing, tasseling, and milking (Figure 3). The flights took place at around 10:30 a.m., synchronized with the Sentinel-2 overpasses. The instrument was calibrated, and the recorded hyperspectral data were radiometrically processed to bottom of atmosphere (BOA) reflectance and orthorectified to be ready for further processing and evaluation.



**Figure 3.** UAV data collection at five dates during the three growth stages of maize. V6~V10 are the sixth-leaf and tenth-leaf stages of maize. VT marks the tasseling stage, and R3 represents the milking stage.

### 2.2.2. Satellite Data

Satellite data used in this study were recorded by a multispectral instrument (MSI) aboard the Sentinel 2A and 2B satellites that belong to the ESA's (European Space Agency, Paris, France) optical Earth observation fleet integrated into the Copernicus Program. MSI provides data of 13 spectral bands located in the VNIR and SWIR range at a spatial resolution from 10 m to 60 m at a swath width of 290 km. The revisit period is five days based on two identical satellites. ESA provides surface reflectance images (BOA, Level-1C) that have undergone detailed geometric corrections and are available through the Copernicus Explorer website [53]. We further applied tools such as the Sen2Cor to obtain the high-precision level '2A', as well as the Sen2Res plugin to resample the 20 m and 60 m bands to a uniform spatial 10 m pixel size while preserving their original reflectance values. We

selected Sentinel bands B1 to B8 that span from the VIS to the NIR wavelengths and are suitable and for LCC estimations [54].

### 2.2.3. Soil Library

To simulate variations of spectral characteristics of chlorophyll combined with different soil backgrounds, we used the ENVI ASTER spectral library with 2443 endmember spectra (EM) provided by the Jet Propulsion Laboratory to obtain the respective soil spectra [55]. The wavelength range of these spectra reached from 0.4 to 15.4  $\mu\text{m}$ . We extracted all bands from 443 to 833 nm to ensure consistency with the Sentinel VIS-NIR wavelengths. The primary soil type found in the research area is a brown soil (Cambisol). Therefore, we used nine brown soil spectra, including fine sandy loam, gravelly fine sandy loam, micaceous loam, gray silty clay, and silty loam to simulate different soil backgrounds for maize.

### 2.2.4. The MuSyQ LCC Product

The ‘Multi-source data Synergized Quantitative remote sensing production system of Leaf Chlorophyll Content’ (MuSyQ LCC) provides a 30 m/10-day resolution coverage of Sentinel-2 MSI data for China. It was generated based on a chlorophyll-sensitive index and processed with time series reconstruction and gap filling. A total of 308 ground measurements including deciduous broadleaf forests, evergreen needle leaf forests, and grasslands showed an overall accuracy with an  $R^2$  of 0.49 and an RMSE of 9.39, as well as an accuracy of croplands with an  $R^2$  of 0.40 and an RMSE of 9.51 [56,57]. To evaluate the accuracy of the MuSyQ LCC product at the sampling points within our study area, we applied cubic convolution interpolation to resample the 30 m resolution data to a finer 10 m resolution, aligning it with the resolution of the Sentinel-2 data used.

### 2.2.5. Field Measurements of LCC

Almost simultaneously during the Sentinel overflights and parallel to the UAV recordings, we conducted chlorophyll measurements at 56 field locations on both sides of a road, spaced a few rows apart. LCCs were recorded with a ‘SPAD-502 Plus’ chlorophyll meter (Konica Minolta Inc., Marunouchi, Chiyoda, Tokyo, Japan). The instrument was calibrated with a reading checker to ensure it provides accurate readings. Equation (1) represents the relationship between the SPAD and the chlorophyll values [58]. The statistical summary of the ground LCC measurements within our study area is presented in Table 1.

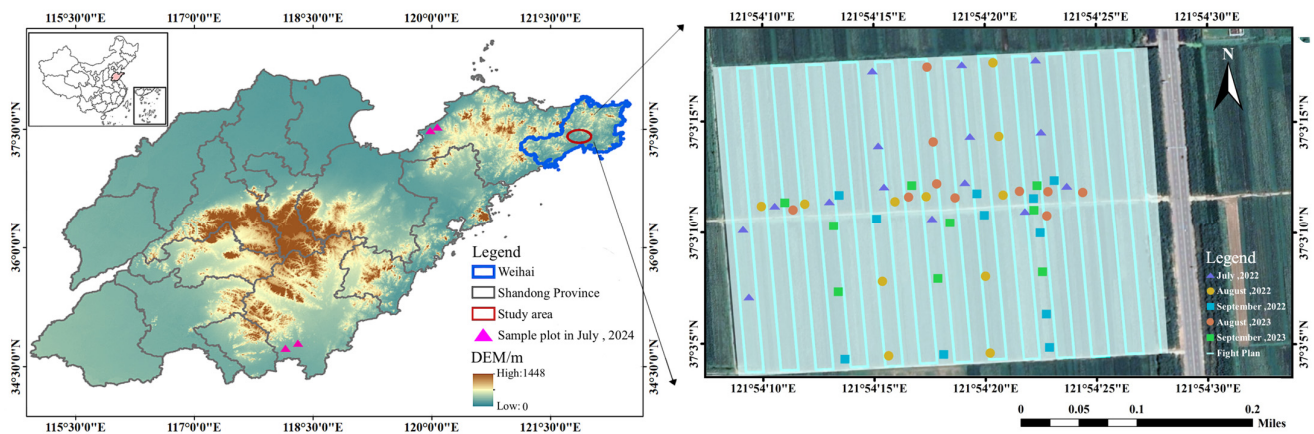
$$y = 0.996x + 1.52 \quad (1)$$

where  $x$  represents the SPAD values and  $y$  is the chlorophyll content with units of  $\mu\text{g}/\text{cm}^2$ .

**Table 1.** In situ LCC measurements of maize at different growing stages.

Dataset	Sample	Min	Mean	Max	SD
Jointing	13	47.00	50.60	55.71	3.05
Tasseling	24	44.90	54.10	60.59	4.17
Milking	19	47.90	53.43	56.87	2.68
Total	56	44.90	52.71	60.59	3.30
Extra samples	4	52.50	54.00	55.50	1.17

In July 2024, four additional measurements were taken from two remote and different farming areas in Shandong Province covering pure spring maize fields and areas intercropping with soybeans on cinnamon soil (Ferralsol) and Shajiang black soils (Vertisols). Each sample site was measured three times, and an average value was calculated and set as the final chlorophyll content. The location of the sample sites and the UAV flight plan are shown in Figure 4.



**Figure 4.** Distribution of field sampling locations and the respective UAV flight plans.

### 2.3. Methodology

To achieve the goal of improving the accuracy of LCC analyses based on multispectral satellite data by hyperspectral UAV recordings, we firstly developed a hyperspectral database of maize. Thereby, we used a random linear mixed model (RLMM) based on UAV data, accounting for different soil backgrounds and intercropping with soybeans. To keep the spatial resolution in accordance with Sentinel-2 imagery, a precision maize coverage was computed by resampling the collected UAV data to a 10 m GSD. Next, we quantified the similarity between the selected Sentinel-2 multispectral data and the maize hyperspectral database. Thereby, four spectral matching methods were tested to identify the spectrum with the highest similarity of the UAV-derived maize database as compared to the Sentinel-2 data and to find the most suitable algorithm for our purpose. The derived spectrum was then used as an input variable for the RLMM model to calculate the maize coverage and compare it with the previously obtained one. In a next step, the optimally matching spectrum was returned to the RLMM model to derive the corresponding maize hyperspectral EM. Finally, three vegetation indices were applied to estimate the LCCs of Sentinel-2 and to verify the results with the collected ground truth measurements.

#### 2.3.1. Linear Mixed Model and Endmember Extraction

Single pixel information of satellite data, even recorded at a 10 m GSD, usually contain the spectral information of more than one material due to the consistency of natural surfaces. In our case, the Sentinel pixels reflected a mixture of maize and soybeans combined with varying soils and their fractions. Using a mixed model can separate various substances depicted in one pixel area by their individual spectral characteristics. These individual substances, also called EMs, have different fractional abundances. Thus, their spectral attributes can be combined to recreate the spectral variations detected by a sensor. A well-established respective standard is the linear mixed model (LMM) [59], which assumes that the observed reflectance spectrum of a given pixel is a linear combination of a small number of EMs. The LMM can be expressed as follows:

$$M = \sum_{i=1}^n a_i R_i + \epsilon \quad (2)$$

where  $M$  is the hybrid reflectance spectra, and  $a_i$  is the abundance of the  $i^{\text{th}}$  EM, where  $0 \leq a_i \leq 1$ . The sum of the abundances is equal to 1, which ensures the total contribution of all materials within a pixel adds up to 100%.  $R_i$  is the reflectance spectra of the  $i^{\text{th}}$  EM, and  $\epsilon$  is the error.

EMs of maize and soybean were extracted from the UAV hyperspectral data obtained by setting up appropriate thresholds for the NDVI [60], which were confirmed by field observations. Accurate NDVI thresholds allow for identifying pure pixels/EMs and subse-

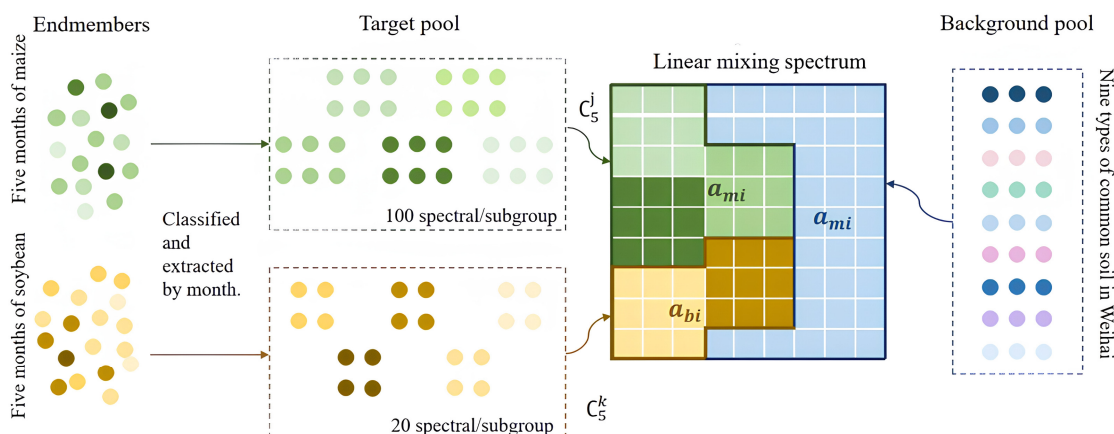
quently calculating the precise percentage of maize coverage within each 10 m pixel. Each UAV NDVI image was classified into 16 clusters to establish the thresholds for soybean and maize. Table 2 presents the NDVI thresholds determined by comparing the UAV images.

**Table 2.** NDVI thresholds of UAV hyperspectral data recorded during different growth stages.

Stage	Date	Threshold	
		Maize	Soybean
V6~V10	22 July 2022	0.657	0.806
VT	22 August 2022	0.699	0.728
R3	21 September 2022	0.701	0.744
VT	8 August 2023	0.713	0.749
R3	9 September 2023	0.674	0.721

### 2.3.2. Synthetically Hybrid Spectral Database Created by an RLMM

We used a random sampling method (RLMM) to create the database in order to avoid an excessively large dataset [61]. First, spectra were divided into a crop and a soil group, and the crop group was further subdivided into five-month-based subgroups of maize and soybean, respectively. Hence, the soil background comprised up to nine subgroups. The maize hyperspectral database was created with these subgroups in the following steps (Figure 5).



**Figure 5.** Flowchart illustrating the creation of the hyperspectral database for maize. In the target pool, different green dots are the EMs of maize, and the yellow dots are the EMs of soybeans. In the background pool, different colors represent different soil types. Linear mixing spectrum comprise data from both target pool and background pool.

To calculate the different portions of maize in linearly mixed spectra under various maize intercropping scenarios, 100 maize EMs and 20 soybean EMs were extracted from each of the five UAV hyperspectral images. Thereby, we produced a target pool that comprised 500 maize EMs and 100 soybean EMs. Concurrently, a soil background pool was established, composed of nine soil subgroups, to generate different soil portions of the linear mixed spectra. By recognizing the significant differences in EMs between various growth stages of maize and soybean, at least one EM was chosen randomly from the target pool during each sampling process, ensuring the bare soil situation was excluded. Accordingly, the EMs were linearly combined with soil spectra from the background pool to construct mixed pixels of all possible combinations. In this context, considering the maize abundance was pertinent since previous studies have shown that in areas where vegetation coverage was less than 0.39 at a 10 m GSD, estimated LCC values were lower than that of the ground measurements and showed a lower accuracy in the estimation [37].

Given that the crops in our study area are relatively lush during July, August, and September, setting the crop abundance to a value greater than zero helped reduce the

computation time. Hence, we set the minimum abundance of the crop group to 0.4 up to 1 with a step size of 0.01. The formula for the creation of random synthetically hybrid spectra is as follows:

$$\begin{aligned}
 M_i &= a_{si}R_{si} + a_{mi}\frac{\sum_{x=0}^j R_{mx}}{j} + a_{bi}\frac{\sum_{y=0}^k R_{by}}{k} & j, k \in [0, 5] \cap j + k \neq 0 \\
 a_{si} + a_{mi} + a_{bi} &= 1 & a_{si}, a_{mi}, a_{bi} \geq 0
 \end{aligned}
 \tag{3}$$

where  $M_i$  is the reflectance of the linear mixed pixels,  $a_{si}$  is the  $i^{th}$  soil background abundance,  $R_{si}$  relates to its corresponding soil reflectivity, and  $a_{mi}$  is the abundance of the  $i^{th}$  maize component spectrum.  $R_{mx}$  denotes the reflectance of the selected maize EMs from the target pool. The corresponding  $a_{bi}$  is the abundance of the soybean, and  $R_{by}$  represents the reflectance of each selected soybean EMs. The number of subgroups  $j$  and  $k$  in the target pool are both set from 0 to 5 with a step size of 1, and the total number of combinations in the target pool is determined by  $C_5^j * C_5^k$ , where  $j$  and  $k$  represent the number of maize and soybean subgroups. Note that if either  $j$  or  $k$  equals 0, it represents none of the EMs selected from the subgroup. Ultimately, along with the nine subgroups in the background pool, we generated a maize hyperspectral database composed of 552,183 spectra.

### 2.3.3. Spectral Matching of Sentinel-2 Data with the Simulated Hyperspectral Maize Data

Spectral matching identifies similarities by calculating the distance or angles between different spectral data and vectors, allowing for the assignment of the data to the most appropriate group. Four commonly applied methods, Euclidean distance (ED), Manhattan distance (MD), Chebyshev distance (CD), and Cosine similarity (CS), were used to match Sentinel-2 data with the simulated hyperspectral data based on their complementary mathematical properties and widespread application. These methods capture both global similarity trends (ED, MD) and localized variations (CD), while addressing spectral angular relationships (CS). To ensure consistent wavelength alignment, we applied linear interpolation to adjust the Sentinel-2 wavelengths with the UAV hyperspectral data. The CS values ranged between 1 and  $-1$ . When the value was 1, it signified complete identity, denoting the highest degree of matching. Conversely, when the value was  $-1$ , it indicated the lowest degree of matching. Definitions and formulae are provided in Table 3.

**Table 3.** Methods used for evaluating the similarity between UAV and Sentinel-2 data.

Methods	Pros	Definitions	Equations	Reference
Euclidean distance	<ul style="list-style-type: none"> <li>• Captures overall magnitude differences across bands.</li> <li>• More sensitive to large spectral differences.</li> </ul>	It uses the square root of the difference in magnitude and direction between two vectors to measure the distance.	$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - y_{jk})^2}$	[62]
Manhattan distance	<ul style="list-style-type: none"> <li>• Robust to outliers, less sensitive to large variations.</li> <li>• Useful for hyperspectral data with subtle variations across bands.</li> </ul>	It is calculated by taking the sum of distances between the x and y coordinates.	$d_{ij} = \sum_{k=1}^n  x_{ik} - y_{jk} $	[62]
Chebyshev distance	<ul style="list-style-type: none"> <li>• Emphasizes the largest spectral difference.</li> <li>• Effective in identifying significant band-to-band discrepancies.</li> </ul>	It is defined as the maximum absolute difference along any coordinate dimension	$d_{ij} = \max( x_{ik} - y_{jk} )$	[63]

Table 3. Cont.

Methods	Pros	Definitions	Equations	Reference
Cosine similarity	<ul style="list-style-type: none"> <li>Focuses on spectral shape rather than magnitude.</li> <li>Robust to scaling effects (e.g., illumination variations).</li> </ul>	It assesses the degree of similarity between two vectors in a vector space by calculating the cosine value of their angle.	$similarity = \frac{x_{ik} \cdot y_{jk}}{\ x_{ik}\  \cdot \ y_{jk}\ }$	[64]

Note that  $d_{ij}$  represents the linear distance between two vectors.  $x_{ik}$  is the input vector and also the reference vector, and  $y_{jk}$  is the comparison vector.  $x_{ik} \cdot y_{jk}$  denotes the dot product of vectors, and  $\|x_{ik}\|$  and  $\|y_{jk}\|$  are the magnitudes of vectors  $x_{ik}$  and  $y_{jk}$ , respectively.

The accuracy of the spectral matching methods was evaluated by calculating and comparing the maize coverage between Sentinel-2 and UAV images. In generating the maize hyperspectral database, the abundance of maize in Equation (4) represents its coverage. Consequently, the maize coverage of each 10 m pixel in the Sentinel-2 images corresponded to the 60 sampling points derived through the RLMM after spectral matching with the constructed hyperspectral database. In order to evaluate the accuracy of the four matching methods, the precise maize area was calculated using NDVI thresholds from UAV hyperspectral images. The accuracy was then assessed by counting the number of pixels that derived the similar maize coverage compared to those obtained via NDVI thresholds, representing the actual maize coverage for each 10 m pixel.

#### 2.3.4. Vegetation Indices Tested for LCC Estimations and Accuracy Assessment Normalized Area over Reflectance Curve

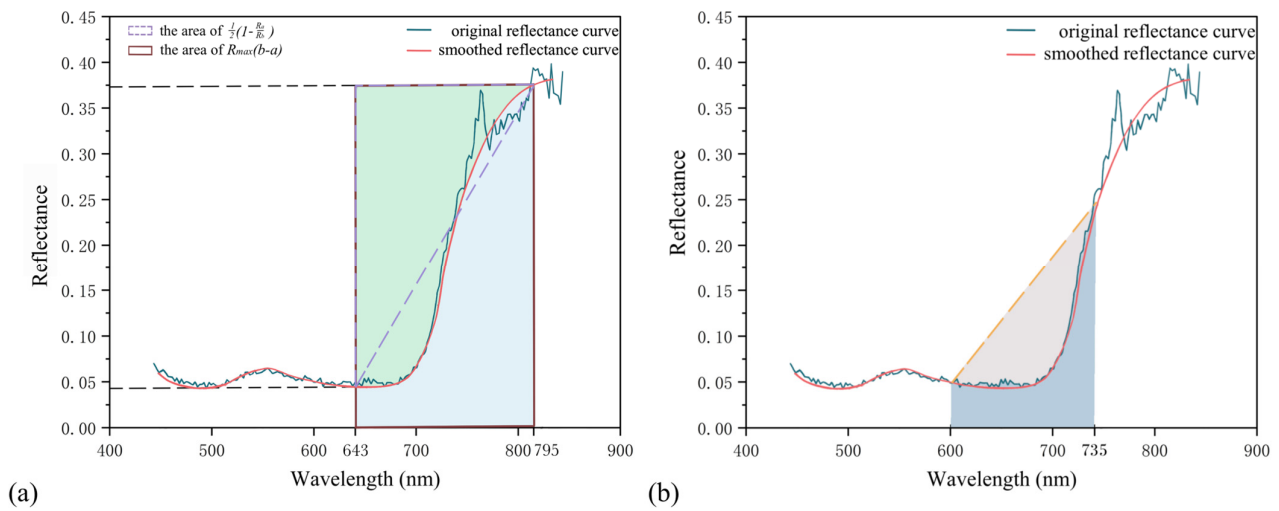
The normalized area over reflectance curve (NAOC) index is a hyperspectral index for estimating chlorophyll content, which has been validated with a variety of vegetation types and shows good accuracy for chlorophyll estimation [42,65,66]. However, it performs poorly in processing multispectral remote sensing data due to the limited number of bands available for calculating integrals. The equation of the NAOC index is expressed as follows:

$$NAOC = 1 - \frac{\int_{643 \text{ nm}}^{795 \text{ nm}} R d\lambda}{R_{max}(795 \text{ nm} - 643 \text{ nm})} \quad (4)$$

where  $R$  is the reflectance,  $\lambda$  is the wavelength, and  $R_{max}$  is the maximum reflectance in near-infrared wavelength region from 643 nm to 795 nm. In this context,  $R_{max}$  equals the reflectance at 795 nm.  $R_{max}(795 \text{ nm} - 643 \text{ nm})$  represents the area of the rectangle within the brown line (Figure 6a). Dividing the integral  $\int_{643 \text{ nm}}^{795 \text{ nm}} R d\lambda$  in blue by the area of the rectangle and subtracting by one will obtain the green area over the reflectance curve (AOC) in Figure 6a.

For multispectral satellite image processing, Facundo et al. [65] proposed the trapezoidal rule to simplify the computation process by considering the ratio of the remaining triangle to the rectangle after subtracting the trapezoidal area. The area of the dotted triangle corresponds to the NAOC approximation as described in Equation (5), which entirely depends on the integration limits (643 nm, 795 nm). Therefore, the optimized NAOC index can be regarded as a straightforward arithmetic spectral index based on two bands. This method is particularly suitable for calculating NAOC values with a few numbers of bands, such as that of Sentinel-2. The formula is defined as follows:

$$NAOC \approx \frac{1}{2} \left( 1 - \frac{R_{643 \text{ nm}}}{R_{795 \text{ nm}}} \right) \quad (5)$$



**Figure 6.** (a) Illustration of the NAOC with an original (green) and a smoothed (orange) hyperspectral reflectance curve of vegetation. The green area corresponds to the AOC, and the blue area is the integral from 643 nm to 795 nm (b) Diagram of the CAI spectral index. The dark blue area is the integral from 600 nm to 735 nm, and the brown gray area is the integral of spectral envelope.

### Chlorophyll Absorption Integral

The chlorophyll absorption integral (CAI) is a further metric utilized to estimate plant chlorophyll content. It is computed by quantifying the area between two specific points between the red and the NIR region. This integral, depicted in Figure 6b, provides a quantitative measure of the chlorophyll concentration, facilitating the comprehension of plant health, photosynthetic activity, and the dynamic variations of the entire ecosystem [67]. The calculation of the CAI is provided as follows:

$$CAI = \int_{R_{600}}^{R_{735}} \frac{Rs_i}{Re_i} \quad (6)$$

where  $R_{735}$  is the reflectance at 735 nm, and  $R_{600}$  is the reflectance at 600 nm.  $Rs_i$  represents the reflectance of the spectrum at band  $i$ , and  $Re_i$  represents the reflectance of the envelope at band  $i$ . As such, the CAI is quite reliant upon spectral envelope measurements.

### Accuracy Assessment of Vegetation Indices

The performance of several vegetation indices concerning the accuracy of LCC retrievals was evaluated by determining the coefficient ( $R^2$ ), which indicates how well the model's prediction ability performs [68]. The root mean square error (RMSE) provides a measure of the magnitude of prediction errors, and the mean absolute error (MAE) is a straightforward measure of the average magnitude of errors in a set of predictions without considering their direction, which helps in understanding the overall accuracy of the model [69]. The formulas and meanings of these indicators are as follows:

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (7)$$

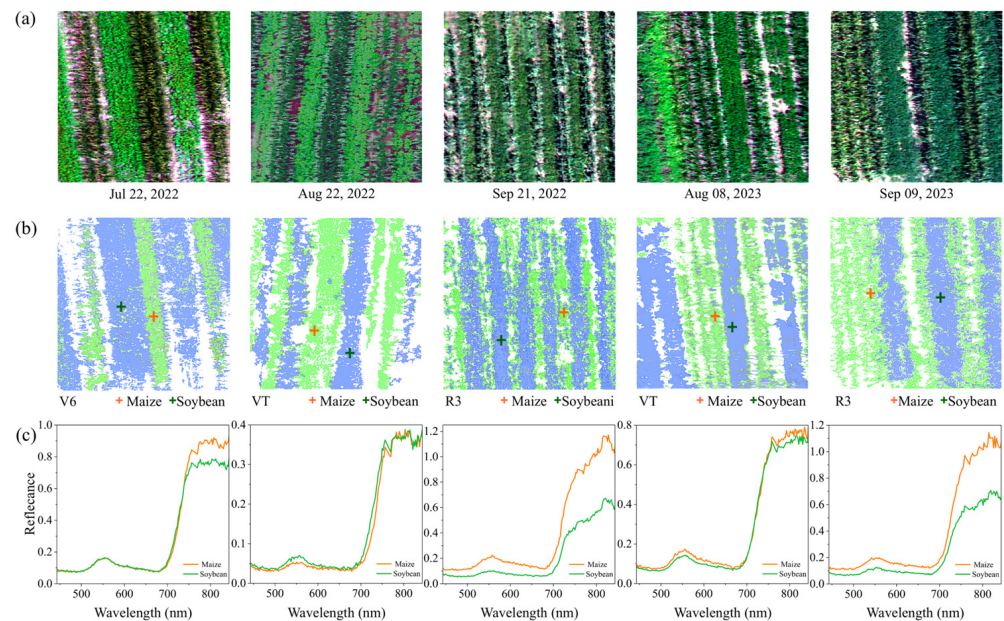
$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where  $n$  represents the sample size,  $\hat{y}_i$  denotes the actual value,  $y_i$  refers to the predicted value, and  $\bar{y}_i$  indicates the mean value.

### 3. Results

#### 3.1. Issues of Intercropping and Varying Soil Fractions

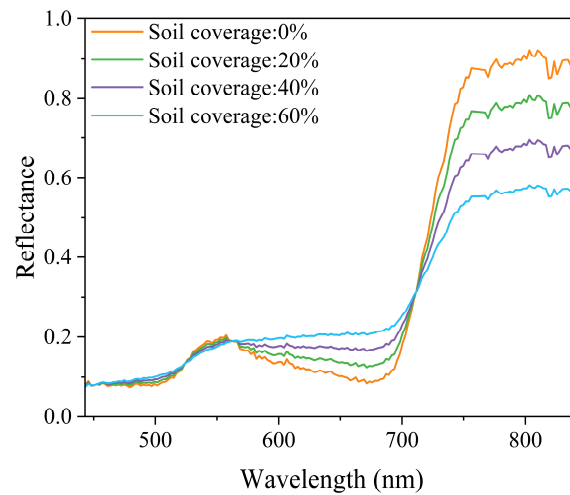
Figure 7 shows the five UAV recordings of our test site taken at different seasons of the year in combination with the respective classified results based on NDVI calculations. It clearly illustrates the problematic of intercropping when recorded with a medium to lower spatial resolution as provided by most Earth observing satellite scanners. The additionally depicted reflectance curves of maize and soybeans exhibit quite similar spectral characteristics, especially in August, which makes their separation and clear identification challenging.



**Figure 7.** (a) Randomly selected intercropping areas of the five UAV hyperspectral recordings. (b) Corresponding results of the NDVI threshold classification, identifying different crop areas: blue = maize; green = soybeans. (c) Spectral characteristics of maize and soybeans measured at the cursor position.

Besides the intercropping issue, the influence of different soil backgrounds arises from the fact that the reflectance values captured by air and space sensors are often a mixture of signals from the canopy and the soil, provided the vegetation cover is not closed. Because of that, nine types of brown soil were incorporated into the construction of the hyperspectral database. Figure 8 shows the changes in spectral reflectance of maize mixed with varying proportions of soil background. With the increasing fractions of soil background, crop reflectance decreased in the red-edge and NIR regions, as the lower reflectance of the soil background diluted the stronger reflectance from the vegetation. In contrast, reflectance increased in the VIS region due to the presence of the soil. It suggests that soil backgrounds generally reduce the spectral reflectance of crop canopies, resulting in an underestimation of LCC measurements.

With an understanding of the above two key issues, we obtained approximately 700,000 EMs from the five UAV hyperspectral images by employing NDVI thresholds (Table 2). Given the limitations of computer computational capabilities, 200 and 100 EMs were randomly selected from each subgroup of maize and soybean, respectively. Moreover, nine EMs from the soil group served as one of the primary data sources for establishing the hyperspectral database containing mixed spectra.

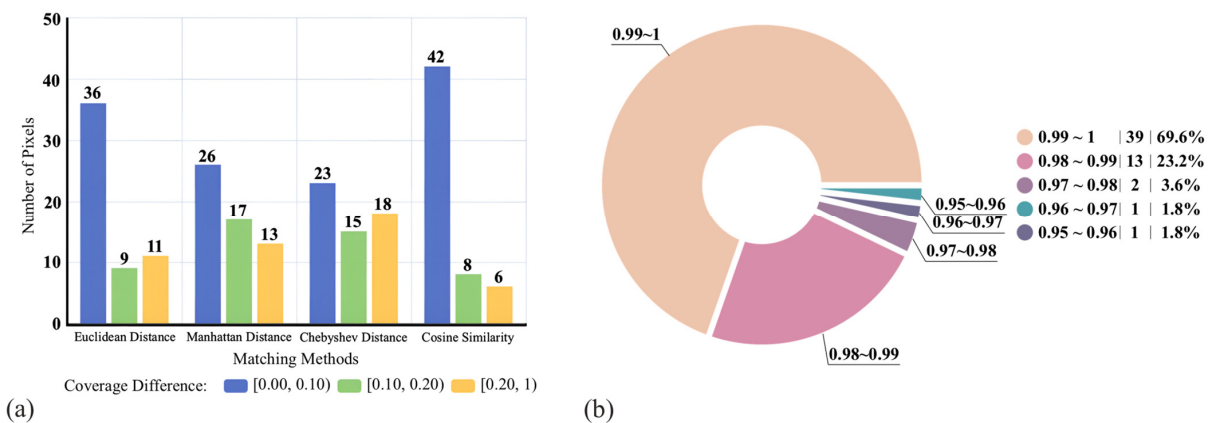


**Figure 8.** Variations of maize spectra with increasing soil background fractions.

### 3.2. Selection of the Optimum Method for Spectral Matching

We extracted 56 pixels from the Sentinel-2 images, corresponding to ground sampling points collected in 2022 and 2023. Spectra of these pixels were matched with the constructed maize hyperspectral database using the ED, MD, CD, and CS methods. Each method generated 56 hyperspectral data that matched the extracted Sentinel-2 spectra, allowing maize abundance to be derived from these hyperspectral data using the RLMM model. Then, a unified evaluation standard had to be established to evaluate the accuracy of the four methods. Therefore, comparing maize abundance with the precise coverage acquired from UAV data, we divided the coverage differences into three classes with an error margin of  $\pm 0.10$  accounting for potential discrepancies, including [0.00–0.10), [0.10–0.20), and [0.20–1). The method that yielded the highest number of matches was identified as the most suitable spectral matching method. This approach also evaluated the feasibility of matching the constructed spectral library and satellite spectra.

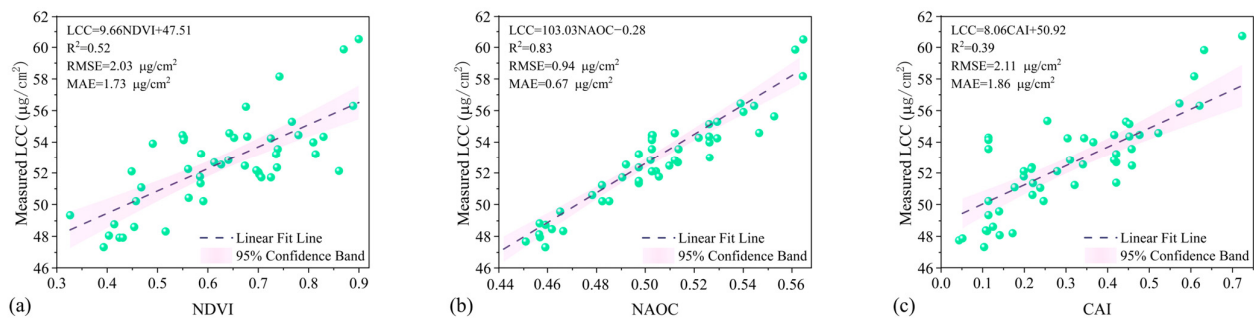
Figure 9a illustrates that the Cosine similarity method achieved the highest accuracy, with 42 pixels having maize coverage differences less than 0.10 (10%), 8 pixels between 0.1 and 0.2, and 6 pixels larger than 0.2 as compared to UAV data, thus outperforming the distance-based methods. This superiority was due to the advantages in handling variations in the overall reflectance while comparing the shapes of the spectral curves. Figure 9b shows the Cosine similarity values computed between Sentinel-2 and the constructed hyperspectral database. For all 56 pixels from Sentinel-2, the matched hyperspectral data had a similarity greater than 0.95, with 69.9% of the pixels showing a matching degree close to 1. Hence, the CS method was the selected method for matching the spectra.



**Figure 9.** (a) Number of pixels corresponding to distinct coverage differences based on four spectral matching methods. (b) Distribution of calculated Cosine similarity values.

### 3.3. Comparative Analysis of Different Vegetation Indices for LCC Retrieval

After performing spectral matching, the derived optimally matched hyperspectral data still contained interference from intercropping or soil background, resulting in mixed pixels. This spectrally mixed data would be returned to the RLMM, which computes the corresponding maize EMs as the maize hyperspectral reflectance information for the Sentinel-2 pixels. Thereby, the NDVI, CAI, and NAOC indices were calculated based on the 60 derived hyperspectral spectra. A linear fitting model was employed to perceive the relationship between the three indices and in situ measured LCCs. To ensure that the model had sufficient learning capacity, we allocated 80% of the data to the training set and used the remaining 20% for validation. The correlation from the training dataset shown in Figure 10 reveals that the NDVI proved less effective than the NAOC in LCC estimation due to its sensitivity to structural characteristics and soil background effects. Changes of structural attributes can impact the NDVI values negatively, independently of chlorophyll levels, leading to a weaker correlation ( $R^2 = 0.52$ ). Although, chlorophyll primarily absorbs light in the blue and red wavebands [70], these often exhibit lower correlations. That is because strong chlorophyll absorptions lead to a rather low reflectance, resulting in saturation levels that reduce the sensitivity of these bands to variations in chlorophyll.



**Figure 10.** Correlations between the training datasets of measured and retrieved LCCs by the following functional indices: (a) NDVI, (b) CAI, and (c) NAOC.

In contrast, the red edge and NIR regions generally exhibited better correlations with chlorophyll levels and were less prone to saturation. Therefore, the NAOC captured the integrated spectral information across a broader wavelength range, incorporating the red edge and NIR regions, making it more sensitive and results in higher correlations to chlorophyll ( $R^2 = 0.83$ ;  $RMSE = 0.94 \mu\text{g}/\text{cm}^2$ ; and  $MAE = 0.67 \mu\text{g}/\text{cm}^2$ ).

The CAI method also integrates spectral data, typically centered on the absorption feature near 670 nm, which may not precisely focus on the wavelengths most indicative of the chlorophyll content. Similar to the NDVI, CAI can be influenced by factors such as soil background and saturation at high chlorophyll levels. It leads to a less linear response and potentially lower correlation with measured chlorophyll content.

### 3.4. Accuracy Assessment of LCC Retrieval

We compared the results in various ways to investigate the feasibility of the proposed method in this study (Table 4). The Sentinel-2 data, using the simplified NAOC index, produced relatively weak results in correlation and accuracy for LCC estimation. The  $R^2$  value of 0.44 indicates that the model can only explain 44% of the variance, suggesting a low-to-moderate correlation. The relatively high RMSE ( $1.96 \mu\text{g}/\text{cm}^2$ ) and MAE ( $1.63 \mu\text{g}/\text{cm}^2$ ) values reflect the presence of considerable errors between the estimated and measured LCCs. This suggests that factors such as soil background and the canopy structure may contribute to a reduced model performance. Since the MuSyQ LCC product is only available for 2022, a total of 25 sample sites were used to validate its accuracy. The MuSyQ LCC product exhibited poor results for LCC estimations, with an  $R^2$  value of 0.43, consistent with the validation of Li, J. et al. [56]. Furthermore, the RMSE and MAE values were significantly higher than those from other validation data sources, at  $14.39 \mu\text{g}/\text{cm}^2$  and  $13.95 \mu\text{g}/\text{cm}^2$ ,

respectively. Such high error values suggest that the lower spatial resolution of the satellite data, coupled with interference from mixed pixels or environmental conditions, contributes to an inaccurate LCC estimation. Our method produced the best overall results among the different validation data sources based on the joint observations with UAV hyperspectral data. With an  $R^2$  value of 0.90, the method indicates a strong linear correlation with the measured LCC values. Moreover, the RMSE ( $1.10 \mu\text{g}/\text{cm}^2$ ) and MAE ( $0.89 \mu\text{g}/\text{cm}^2$ ) were also relatively low, demonstrating that the method provides accurate and reliable results. These also suggest that constructing the hyperspectral database offers a comprehensive and representative spectral dataset, capturing key characteristics that are critical for accurate satellite-based LCC estimations.

**Table 4.** Accuracy of LCC estimation through different validation data sources.

Validation Data Sources	Number of Samples	$R^2$	RMSE/ $(\mu\text{g}/\text{cm}^2)$	MAE/ $(\mu\text{g}/\text{cm}^2)$
Sentinel-2 (NAOC)	12	0.44	1.96	1.63
MuSyQ-LCC	25	0.43	14.39	13.95
Constructed hyperspectral database	12	0.90	1.10	0.89
Constructed hyperspectral database	8	0.89	1.12	0.90
Additional sample points in 2024	4	0.93	1.10	1.09

We used 80% of the 56 field measurements of LCC to calculate the linear correlation with the NAOC index, and 20% of data was selected for model validation. To test if the amount of data was sufficient for our method, we randomly choose only 40 field measurements to examine our model's reliability with the same procedure mentioned above. The results show that with the reduced number of samples, our method still showed a high accuracy of LCC estimation with an  $R^2$  value of 0.89, an RMSE of 1.12, and a MAE of 0.90. To further verify the generalization of the method across different areas, four additional points were evaluated in 2024, including intercropping and sole maize scenarios. The result showed the highest degree of correlation with the measured LCCs. With an  $R^2$  value of 0.93, this was the best performance based on the correlation coefficient. The RMSE ( $1.10 \mu\text{g}/\text{cm}^2$ ) and MAE ( $1.09 \mu\text{g}/\text{cm}^2$ ) values were also relatively low, further demonstrating the model's reliability also at areas with different soil backgrounds.

In conclusion, based on its high accuracy, the constructed hyperspectral database is a reliable data source for LCC estimations. The data gathered during four additional ground sampling campaigns at two further locations also performed well, demonstrating the generalization and transferability of our proposed method. In contrast, the Sentinel-2 NAOC and MuSyQ exhibited lower accuracy and higher errors, suggesting they need to be improved to be suitable for satellite-based precise LCC estimations. The determined coefficients, RMSE, and MAE obtained are presented in the Table 4:

#### 4. Discussion

Retrieving LCC from space-based observations is challenging, as atmospheric effects, canopy characteristics, and background reflectance can affect the detection of subtle differences in canopy reflectance caused by variations in LCC [38]. As maize–soybean intercropping has become a more common agricultural practice, its influence has increased and therefore cannot be overlooked. Studies demonstrated that intercropped fields often exhibit spectral responses similar to those of sole crop fields, leading to frequent misclassifications [71–73]. Therefore, to our knowledge, incorporating UAV hyperspectral data addressing different scenarios of maize and soybean during key growth stages and accounting for the influence of soil background reflectance based on a RLMM approach is reported for the first time to improve the accuracy of satellite-based LCC estimation.

#### 4.1. How Can the Constructed Hyperspectral Database Improve Satellite LCC Monitoring Accuracy?

Satellite-based estimations of LCCs are often associated with relatively severe uncertainties. Investigated by various empirical and physically based approaches as well as different sensor data, accuracies ranged between 7 and 10  $\mu\text{g}/\text{cm}^2$  RMSE (RMSD) (Table 5).

**Table 5.** Related studies of LCC estimation based on operational satellite data.

Data Source	Spatial Resolution	Method	RMSE/RMSD	Reference
SPOT-5	10 m	physically based	7.1 $\mu\text{g}/\text{cm}^2$	[74]
Landsat-5	30 m	physically based	8.73 $\mu\text{g}/\text{cm}^2$	[75]
Sentinel-2	30 m	empirical	9.51 $\mu\text{g}/\text{cm}^2$	[76]
Sentinel-2	20 m	hybrid	10.4 $\mu\text{g}/\text{cm}^2$	[77]

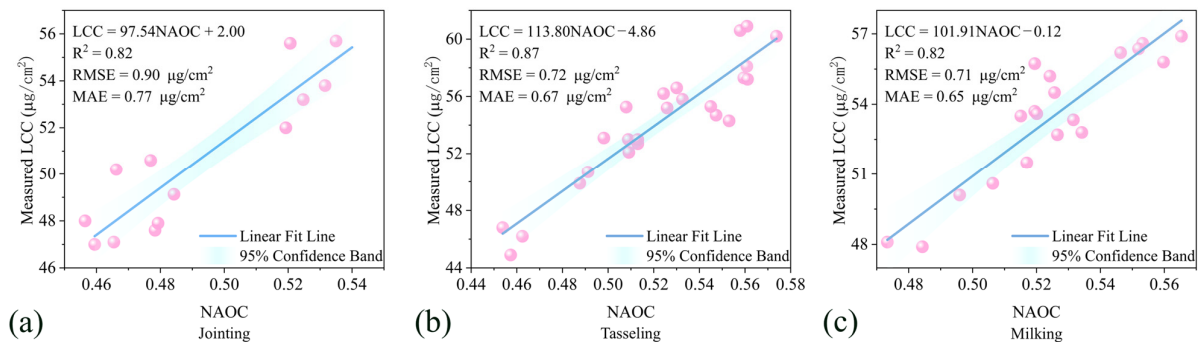
With the development of close-range sensing, UAV hyperspectral image data have been verified to have an excellent correlation ( $>0.96$ ) with field spectra, making this a reliable tool for estimating maize growth with high accuracy and stability, with RMSE (RMSD) values approximately ranging between 1 and 5  $\mu\text{g}/\text{cm}^2$  [78–80]. However, UAV hyperspectral imaging remains limited by its cost and coverage. The challenge, therefore, is how to leverage the advantages of UAV hyperspectral data to enhance satellite-based LCC estimations. The primary issue affecting the accuracy of LCC estimation in satellite remote sensing is the inevitable occurrence of mixed pixels. Although studies focus on mitigating the canopy structure and soil backgrounds, they often neglect circumstances such as intercropping. The spectral reflectance for maize and soybeans is similar and difficult to distinguish during the key growing phase (Figure 7). This supports the earlier findings of Mahlayeye et al. (2024) [71], who demonstrated that the overall distribution and median of field-averaged spectral reflectance for maize and intercropped maize were similarly comparable at the emergence-seedling phase.

After performing spectral matching, the RLMM enables the determination of the proportions of maize, soybean intercropping, and soil background. Subsequently, the canopy chlorophyll content (CCC) can be calculated by multiplying the LCCs with the maize proportion and the Leaf Area Index (LAI) under the assumption that all leaves in the plant have the same chlorophyll content [81]. This approach offers a valuable way for converting multispectral data to hyperspectral data, thereby reducing the impact of mixed pixels and addressing challenges in satellite-based LCC monitoring, particularly in intercropping scenarios. In our case, the LCC was calculated instead of the CCC because, in reality, chlorophyll is not uniformly distributed throughout the plant, and optical remote sensing systems are mostly sensitive to the upper portion of the canopy. From a vertical distribution perspective, the top eight leaf layers account for only 55–65% of the total chlorophyll content; at the canopy level, this can be less than 55% [82]. Therefore, further profound investigations into both methods and the canopy structure are required to achieve precise CCC estimates.

#### 4.2. Comparison of LCC Estimation Across Different Growth Stages

The LCC increased from approximately 40 to 60  $\mu\text{g}/\text{cm}^2$  as maize developed toward the milking stage (Table 1), which aligns with the findings reported by Argenta et al. (2004) [83]. Analysis of the relationships across different growth stages within the 400–1000 nm spectral wavelength revealed that the most sensitive stages for detecting chlorophyll content were the shooting and trumpet stages, where the absolute correlation coefficient exceeded 0.6 [84]. We evaluated the performance of our method separately according to different growth stages of maize, and the results indicated minimal estimation differences across them (Figure 11). In August, during the tasseling (shooting) stage, the method demonstrated a strong ability to capture the overall variation ( $R^2 = 0.87$ ), which can be attributed to the strong correlation ( $>0.6$ ). The jointing and milking stages showed similar results ( $R^2 = 0.82$ ), but the RMSE

for the milking stage was the lowest, indicating smaller prediction. This variation might be due to differences in sample sizes across stages, with some subsets of the data being predicted quite well. Overall, our method proved effective across the various key growth stages (Figure 11), with the best LCC estimation occurring at the tasseling (shooting) stage, which confirms the findings from Sun et al. [84]. We did not focus on the entire growth season because during these key growing stages, the LCC is a critical determinant of photosynthetic efficiency, biomass accumulation, and grain filling, making it a robust indicator of yield potential. Second, previous studies have shown that index based chlorophyll models tend to exhibit significant inversion errors throughout the entire growing season, particularly during the early vegetative and late grain filling stages, where RMSE values ranged from 36% to 87.4% [43]. Therefore, focusing on the key growth stages provides a more reliable approach for chlorophyll estimation and yield prediction.



**Figure 11.** Accuracy of LCC estimation for different growth stages of maize based on our proposed method. (a) Jointing stage. (b) Tasseling stage. (c) Milking stage.

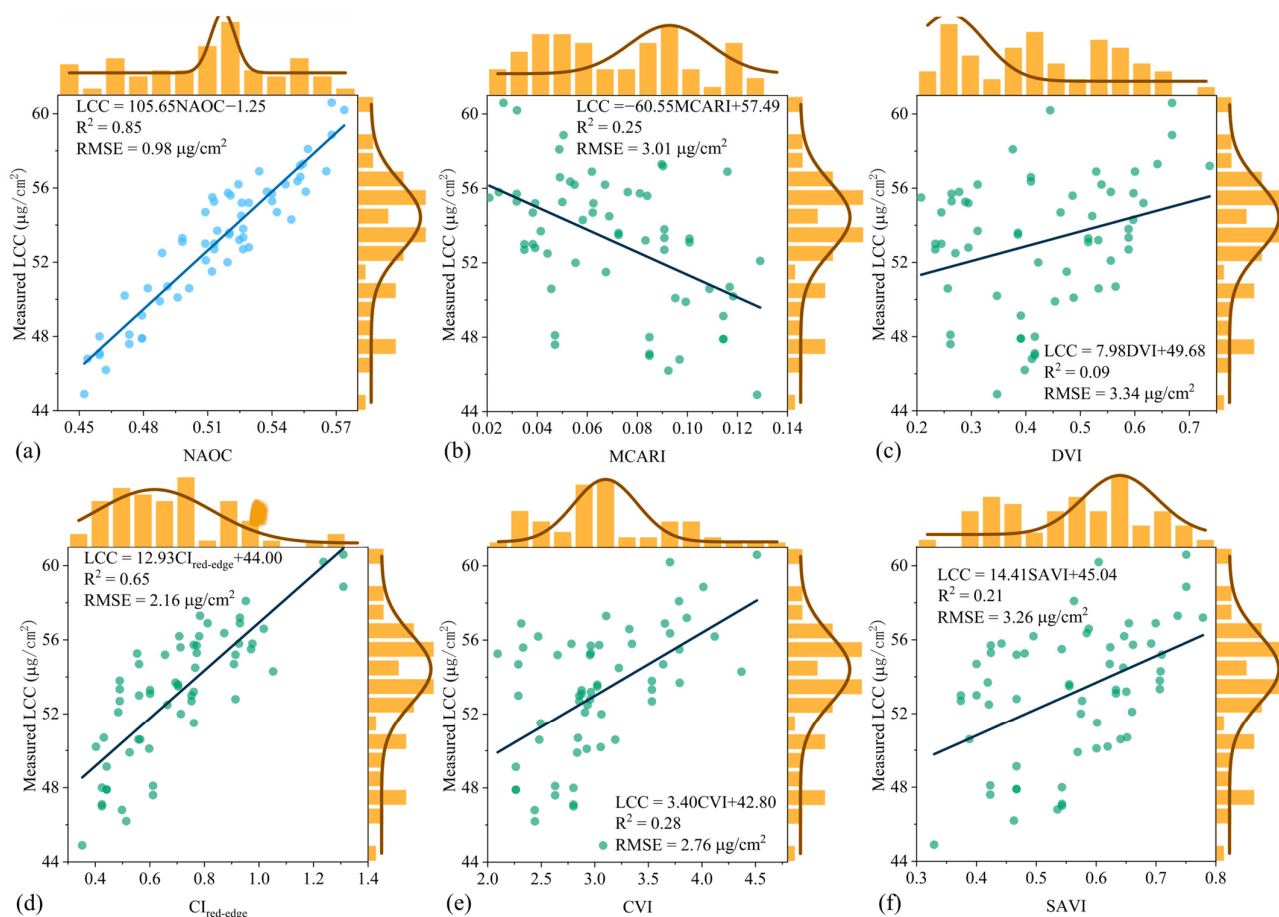
To verify the accuracy of our method across different key growth stages, all 60 samples and 6 commonly used vegetation indices (NAOC, MACARI,  $CI_{red-edge}$ , CVI, SAVI, and DVI) were calculated with the in situ measured LCCs. Among these indices, the NAOC (Figure 12a) demonstrated the best estimation capability, with an  $R^2$  of 0.85 and RMSE of  $0.98 \mu\text{g}/\text{cm}^2$ . As the sample size increased, the NAOC index showed an improved LCC estimation accuracy (Figure 10), which indicated that NAOC provides a robust linear fit with the LCCs without requiring additional ground measurements [42]. The  $CI_{red-edge}$  also depicts a strong LCC estimation ability in very different canopy architectures and leaf structures [42]. Although its accuracy was slightly lower than that of the NAOC, the  $CI_{red-edge}$  (Figure 12d) still demonstrates an acceptable level of accuracy when validated with our dataset ( $R^2 = 0.65$ ,  $\text{RMSE} = 2.16 \mu\text{g}/\text{cm}^2$ ). The low accuracy of the MCARI (Figure 12b), DVI (Figure 12c), CVI (Figure 12e), and SAVI (Figure 12f) in predicting chlorophyll content is likely due to their limited sensitivity to high chlorophyll concentrations and the relatively weak response of their selected bands to changes in chlorophyll spectral behavior.

#### 4.3. Transferability of the Processing Concept and Implications for Further Work

To examine the transferability and thus the generalization of our method, we made further measurements in different locations in Shandong Province. Four samples were collected from two other geographical areas, each of them characterized by varying soil types and different farming practices. These locations were selected to provide a representative set of conditions, capturing variations in agricultural environments and to enhance the robustness of the proposed method.

However, there are quite some constraints that have to be addressed. Although our validation process produced strong results, demonstrating high accurate LCC estimations across different cropping scenarios, further validation with a higher number and a more diverse set of field measurements is planned for future works. Also, the inclusion of sites characterized by varying climatic conditions is foreseen, as these can significantly influence the monitoring of crop LCCs due to their impact on plant phenology. Additionally, we

plan to randomly extract more maize and soybean EMs. Applying them to the RLMM can further help to reduce the canopy effect to some extent. The necessary synchronization of UAV flights and field measurements with satellite overflights imposes further temporal constraints, due to varying illumination conditions. This limits the size of the target area to be recorded by the UAV, the sampling time and consequently the number of samples. Finally, soil moisture and vegetation uniformity need to be addressed in the future, as these may introduce higher variability when applied to more heterogeneous landscapes. All these limitations emphasize the importance of cautious extrapolation and highlight the need for future validation efforts across diverse regions and conditions.



**Figure 12.** Relationship between LCC and six widely used FVIs: (a) Normalized area over reflectance curve (NAOC); (b) Modified Chlorophyll Absorption Ratio Index (MCARI); (c) Difference Vegetation Index (DVI); (d) Red Edge Chlorophyll Index ( $CI_{\text{red-edge}}$ ); (e) Chlorophyll Vegetation Index (CVI); (f) Soil-adjusted Vegetation Index (SAVI). Linear models were constructed using the full dataset of 60 ground samples from our study area. The correlation coefficients and RMSE values are displayed in each scatterplot, with the distributions of the FVI and LCC values presented as histograms along the top and at the side of each graph.

In our field investigations, for example, we observed instances of maize intercropped with garlic, a significant upcoming intercropping practice. Due to sulfur-containing compounds, garlic possesses natural pest-repellent properties, which can help deter pests such as aphids and beetles, potentially reducing the need for chemical pesticides. Typically, garlic is planted in the fall (around October to November) before the first hard frost and is harvested in late spring to early summer (May to June), coinciding with the early growth stages of maize. Consequently, monitoring LCCs during these early stages is inevitably affected by the presence of garlic leaves, potentially leading to interference in remotely sensed data. Currently, there is a limited understanding of how garlic affects LCCs in such

systems, and further investigations are needed to assess the impact of this combination on the LCC estimation. Therefore, expanding field experiments by including different regions would help to account for variations in crop management practices and soil characteristics, thereby ensuring a more robust generalizability of our method.

## 5. Conclusions

Accurate satellite-based LCC estimates provide a large-scale and comprehensive overview of crop growth conditions. By conducting joint observations with five UAV hyperspectral datasets covering key maize growth stages, we constructed a hyperspectral database of maize with 552,183 spectra based on a random linear mixed model (RLMM) to mitigate the effects from intercropping, different growth stages, varying canopy structure, and soil backgrounds. Four spectral matching methods were tested to combine Sentinel-2 reflectance data with the constructed hyperspectral database. The results demonstrated that the Cosine similarity was the most effective approach, achieving high accuracy by accounting for variations in spectral reflectance. Furthermore, comparing different vegetation indices confirmed that the NAOC outperformed traditional commonly used indices such as NDVI and  $CI_{\text{red-edge}}$  in capturing LCC-related spectral information. Our developed approach shows higher accuracy than the LCC product (MuSyQ) and has been successfully validated across different areas, reinforcing its potential for a broader application in agricultural remote sensing.

**Author Contributions:** Conceptualization, R.K.; methodology, S.Y., R.K. and L.S.; validation, J.G.; investigation, R.K., J.G., L.Z. and L.S.; data curation, T.X.; writing—original draft preparation, S.Y. and R.K.; writing—review and editing, S.Y., R.K., T.X., L.Z., C.D. and H.J.K.; visualization, S.Y. and J.G.; supervision, T.X., C.D. and H.J.K.; funding acquisition, T.X. and C.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Key Research and Development Program of Shandong (2021ZDSYS01), the National Natural Science Foundation of China (No. 42301327), the New Liberal Arts Research and Reform Project of the Ministry of Education (2021140084), and the Humanities and Social Sciences Youth Team Project of Shandong University (Weihai).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

**Acknowledgments:** We thank the anonymous reviewers for helping to improve the manuscript with their conscientious and constructive comments and suggestions.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Zhang, X.; He, Y.; Wang, C.; Xu, F.; Li, X.; Tan, C.; Chen, D.; Wang, G.; Shi, L. Estimation of corn canopy chlorophyll content using derivative spectra in the  $O_2$ —A absorption band. *Front. Plant Sci.* **2019**, *10*, 1047. [[CrossRef](#)] [[PubMed](#)]
2. Chen, L.; Chang, Q.-r.; Gao, Y.-f.; Luo, L.-l.; Zheng, Y. Hyperspectral estimation model of chlorophyll content in kiwifruit leaves. *J. Northwest A F Univ. (Nat. Sci. Ed.)* **2020**, *48*, 79–89. [[CrossRef](#)]
3. Yang, J.; Liao, G.-p.; Liu, F.; Guan, C.-y. Prediction of chlorophyll content of rape leaves with hyperspectral imaging technology. *J. Agric. Sci. Technol.* **2020**, *22*, 86–96. [[CrossRef](#)]
4. Flynn, K.C.; Frazier, A.E.; Admas, S. Performance of chlorophyll prediction indices for *Eragrostis tef* at Sentinel-2 MSI and Landsat-8 OLI spectral resolutions. *Precis. Agric.* **2020**, *21*, 1057–1071. [[CrossRef](#)]
5. Muruganatham, P.; Wibowo, S.; Grandhi, S.; Samrat, N.H.; Islam, N. A systematic literature review on crop yield prediction with deep learning and remote sensing. *Remote Sens.* **2022**, *14*, 1990. [[CrossRef](#)]
6. Qiao, L.; Gao, D.; Zhang, J.; Li, M.; Sun, H.; Ma, J. Dynamic influence elimination and chlorophyll content diagnosis of maize using UAV spectral imagery. *Remote Sens.* **2020**, *12*, 2650. [[CrossRef](#)]
7. Verrelst, J.; Camps-Valls, G.; Muñoz-Mari, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.; Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—A review. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 273–290. [[CrossRef](#)]
8. Alvarez-Vanhard, E.; Corpetti, T.; Houet, T. UAV & satellite synergies for optical remote sensing applications: A literature review. *Sci. Remote Sens.* **2021**, *3*, 100019. [[CrossRef](#)]

9. Ma, Y.; Qiu, C.; Zhang, J.; Pan, D.; Zheng, C.; Sun, H.; Feng, H.; Song, X. Potato Leaf Chlorophyll Content Estimation through Radiative Transfer Modeling and Active Learning. *Agron. J.* **2023**, *13*, 3071. [[CrossRef](#)]
10. Sadeh, Y.; Zhu, X.; Dunkerley, D.; Walker, J.P.; Zhang, Y.; Rozenstein, O.; Manivasagam, V.; Chenu, K. Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *96*, 102260. [[CrossRef](#)]
11. Brewer, K.; Clulow, A.; Sibanda, M.; Gokool, S.; Naiken, V.; Mabhaudhi, T. Predicting the chlorophyll content of maize over phenotyping as a proxy for crop health in smallholder farming systems. *Remote Sens.* **2022**, *14*, 518. [[CrossRef](#)]
12. Xie, C.; Yang, C. A review on plant high-throughput phenotyping traits using UAV-based sensors. *Comput. Electron. Agric.* **2020**, *178*, 105731. [[CrossRef](#)]
13. Nhamo, L.; Magidi, J.; Nyamugama, A.; Clulow, A.D.; Sibanda, M.; Chimonyo, V.G.; Mabhaudhi, T. Prospects of improving agricultural and water productivity through unmanned aerial vehicles. *Agriculture* **2020**, *10*, 256. [[CrossRef](#)]
14. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Hartling, S.; Esposito, F.; Fritschi, F.B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* **2020**, *237*, 111599. [[CrossRef](#)]
15. Cracknell, A.P. UAVs: Regulations and law enforcement. *Int. J. Remote Sens.* **2017**, *38*, 3054–3067. [[CrossRef](#)]
16. Féret, J.-B.; de Boissieu, F. prospect: An R package to link leaf optical properties with their chemical and structural properties with the leaf model PROSPECT. *J. Open Source Softw.* **2024**, *9*, 6027. [[CrossRef](#)]
17. Zhuo, W.; Huang, J.; Xiao, X.; Huang, H.; Bajgain, R.; Wu, X.; Gao, X.; Wang, J.; Li, X.; Wagle, P. Assimilating remote sensing-based VPM GPP into the WOFOST model for improving regional winter wheat yield estimation. *Eur. J. Agron.* **2022**, *139*, 126556. [[CrossRef](#)]
18. Zhang, C.; Chen, Z.; Yang, G.; Xu, B.; Feng, H.; Chen, R.; Qi, N.; Zhang, W.; Zhao, D.; Cheng, J. Removal of canopy shadows improved retrieval accuracy of individual apple tree crowns LAI and chlorophyll content using UAV multispectral imagery and PROSAIL model. *Comput. Electron. Agric.* **2024**, *221*, 108959. [[CrossRef](#)]
19. Wan, L.; Ryu, Y.; Dechant, B.; Lee, J.; Zhong, Z.; Feng, H. Improving retrieval of leaf chlorophyll content from Sentinel-2 and Landsat-7/8 imagery by correcting for canopy structural effects. *Remote Sens. Environ.* **2024**, *304*, 114048. [[CrossRef](#)]
20. Wang, W.; Cheng, Y.; Ren, Y.; Zhang, Z.; Geng, H. Prediction of chlorophyll content in multi-temporal winter wheat based on multispectral and machine learning. *Front. Plant Sci.* **2022**, *13*, 896408. [[CrossRef](#)]
21. Gao, C.; Li, H.; Wang, J.; Zhang, X.; Huang, K.; Song, X.; Yang, W.; Feng, M.; Xiao, L.; Zhao, Y. Combined use of spectral resampling and machine learning algorithms to estimate soybean leaf chlorophyll. *Comput. Electron. Agric.* **2024**, *218*, 108675. [[CrossRef](#)]
22. Krieglner, F.; Malila, W.; Nalepka, R.; Richardson, W. Preprocessing Transforms Their Effects on Multispectral Recognition. In Proceedings of the Sixth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, USA, 13–16 October 1969; pp. 97–121.
23. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.
24. Zhengxing, W.; Chuang, L.; Alfredo, H. From AVHRR-NDVI to MODIS-EVI: Advances in vegetation index research. *Acta Ecol. Sin.* **2003**, *23*, 979–987. [[CrossRef](#)]
25. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
26. Naji, T.A. Study of vegetation cover distribution using DVI, PVI, WdVI indices with 2D-space plot. In Proceedings of the IBN Al-Haitham First International Scientific Conference, Baghdad, Iraq, 13–14 December 2017.
27. Gao, S.; Yan, K.; Liu, J.; Pu, J.; Zou, D.; Qi, J.; Mu, X.; Yan, G. Assessment of remote-sensed vegetation indices for estimating forest chlorophyll concentration. *Ecol. Indic.* **2024**, *162*, 112001. [[CrossRef](#)]
28. Wang, C.; Wu, Y.; Hu, Q.; Hu, J.; Chen, Y.; Lin, S.; Xie, Q. Comparison of vegetation phenology derived from solar-induced chlorophyll fluorescence and enhanced vegetation index, and their relationship with climatic limitations. *Remote Sens.* **2022**, *14*, 3018. [[CrossRef](#)]
29. Gond, V.; Fayolle, A.; Pennec, A.; Cornu, G.; Mayaux, P.; Camberlin, P.; Doumenge, C.; Fauvet, N.; Gourlet-Fleury, S. Vegetation structure and greenness in Central Africa from Modis multi-temporal data. *Philos. Trans. R. Soc. B Biol. Sci.* **2013**, *368*, 20120309. [[CrossRef](#)]
30. Vincini, M.; Frazzi, E. Comparing narrow and broad-band vegetation indices to estimate leaf chlorophyll content in planophile crop canopies. *Precis. Agric.* **2011**, *12*, 334–344. [[CrossRef](#)]
31. Madonsela, S.; Cho, M.A.; Naidoo, L.; Main, R.; Majozi, N. Exploring the utility of Sentinel-2 for estimating maize chlorophyll content and leaf area index across different growth stages. *J. Spat. Sci.* **2023**, *68*, 339–351. [[CrossRef](#)]
32. Cheng, J.; Yang, H.; Qi, J.; Sun, Z.; Han, S.; Feng, H.; Jiang, J.; Xu, W.; Li, Z.; Yang, G. Estimating canopy-scale chlorophyll content in apple orchards using a 3D radiative transfer model and UAV multispectral imagery. *Comput. Electron. Agric.* **2022**, *202*, 107401. [[CrossRef](#)]
33. Gurung, R.B.; Breidt, F.J.; Duttin, A.; Ogle, S.M. Predicting Enhanced Vegetation Index (EVI) curves for ecosystem modeling applications. *Remote Sens. Environ.* **2009**, *113*, 2186–2193. [[CrossRef](#)]
34. Xu, Z.-Y.; Sun, B.; Zhang, W.-F.; Li, Y.-F.; Yan, Z.-Y.; Yue, W.; Teng, S.-H. An evaluation of a remote sensing method based on optimized triangular vegetation index (TVI) for aboveground shrub biomass estimation in shrub-encroached grassland. *Cao Ye Xue Bao* **2023**, *32*, 1. [[CrossRef](#)]

35. Li, D.; Chen, J.M.; Yu, W.; Zheng, H.; Yao, X.; Cao, W.; Wei, D.; Xiao, C.; Zhu, Y.; Cheng, T. Assessing a soil-removed semi-empirical model for estimating leaf chlorophyll content. *Remote Sens. Environ.* **2022**, *282*, 113284. [[CrossRef](#)]
36. Venancio, L.P.; Mantovani, E.C.; do Amaral, C.H.; Neale, C.M.U.; Gonçalves, I.Z.; Filgueiras, R.; Eugenio, F.C. Potential of using spectral vegetation indices for corn green biomass estimation based on their relationship with the photosynthetic vegetation sub-pixel fraction. *Agric. Water Manag.* **2020**, *236*, 106155. [[CrossRef](#)]
37. Liu, L.; Xie, Y.; Zhu, B.; Song, K. Rice leaf chlorophyll content estimation with different crop coverages based on Sentinel-2. *Ecol. Inform.* **2024**, *81*, 102622. [[CrossRef](#)]
38. Daughtry, C.S.; Walthall, C.; Kim, M.; De Colstoun, E.B.; McMurtrey Iii, J. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
39. Han, D.; Cai, H.; Zhang, L.; Wen, Y. Multi-sensor high spatial resolution leaf area index estimation by combining surface reflectance with vegetation indices for highly heterogeneous regions: A case study of the Chishui River Basin in southwest China. *Ecol. Inform.* **2024**, *80*, 102489. [[CrossRef](#)]
40. Gupta, S.K.; Pandey, A.C. Spectral aspects for monitoring forest health in extreme season using multispectral imagery. *Egypt. J. Remote Sens. Space Sci.* **2021**, *24*, 579–586. [[CrossRef](#)]
41. Padalia, H.; Sinha, S.K.; Bhave, V.; Trivedi, N.K.; Kumar, A.S. Estimating canopy LAI and chlorophyll of tropical forest plantation (North India) using Sentinel-2 data. *Adv. Space Res.* **2020**, *65*, 458–469. [[CrossRef](#)]
42. Delegido, J.; Alonso, L.; Gonzalez, G.; Moreno, J. Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 165–174. [[CrossRef](#)]
43. Yang, H.; Ming, B.; Nie, C.; Xue, B.; Xin, J.; Lu, X.; Xue, J.; Hou, P.; Xie, R.; Wang, K. Maize canopy and leaf chlorophyll content assessment from leaf spectral reflectance: Estimation and uncertainty analysis across growth stages and vertical distribution. *Remote Sens.* **2022**, *14*, 2115. [[CrossRef](#)]
44. Shwetank, S.; Jain, K.; Bhatia, K. Development of digital spectral library and supervised classification of rice crop varieties using hyperspectral image processing. *AJG* **2012**, *11*. [[CrossRef](#)]
45. Gupta, S.; Rajan, K. Temporal signature matching for land cover classification. In Proceedings of the International Society for Photogrammetry and Remote Sensing-Technical Commission VIII Symposium, Kyoto, Japan, 9–12 August 2010.
46. Li, F.; Ren, J.; Wu, S.; Zhao, H.; Zhang, N. Comparison of regional winter wheat mapping results from different similarity measurement indicators of NDVI time series and their optimized thresholds. *Remote Sens.* **2021**, *13*, 1162. [[CrossRef](#)]
47. Demirci, S.; Erer, I.; Ersoy, O. Weighted Chebyshev Distance Algorithms for Hyperspectral Target Detection and Classification Applications. Available online: <https://docs.lib.purdue.edu/ecetr/463/> (accessed on 28 November 2024).
48. Lv, X.; Yuan, L.; Rao, C.; Wu, X.; Qing, X.; Weng, X. Structure and near-infrared spectral properties of mesoporous silica for hyperspectral camouflage materials. *Infrared Phys. Technol.* **2023**, *129*, 104558. [[CrossRef](#)]
49. Lin, Z.; Jiang, Y.; Wu, C. Fast spectral clustering with local cosine similarity graphs for hyperspectral images. *J. Appl. Remote Sens.* **2024**, *18*, 024502. [[CrossRef](#)]
50. FAOLEX Database. Available online: <https://www.fao.org/faolex/results/details/en/c/LEX-FAOC205794/> (accessed on 1 March 2024).
51. Xu, Z.; Li, C.; Zhang, C.; Yu, Y.; van der Werf, W.; Zhang, F. Intercropping maize and soybean increases efficiency of land and fertilizer nitrogen use; A meta-analysis. *Field Crops Res.* **2020**, *246*, 107661. [[CrossRef](#)]
52. Ortega-Terol, D.; Hernandez-Lopez, D.; Ballesteros, R.; Gonzalez-Aguilera, D. Automatic hotspot and sun glint detection in UAV multispectral images. *Sensors* **2017**, *17*, 2352. [[CrossRef](#)]
53. Copernicus Open Access Hub. Available online: <https://browser.dataspace.copernicus.eu/> (accessed on 1 September 2023).
54. Agency, E.S. SentiWiki/Sentinel-2 S2 Mission. Available online: <https://sentiwiki.copernicus.eu/web/s2-mission> (accessed on 17 July 2024).
55. ENVI ASTER Spectral Library. Available online: <http://speclib.jpl.nasa.gov> (accessed on 23 September 2023).
56. Li, J.; Zhang, H.; Wang, X.; Zhang, Z.; Gu, C.; Wen, Y.; Chu, T.; Liu, Q. A dataset of 30m/10-day leaf chlorophyll content MuSyQ GF-series (2019–2020, China, version 01). *China Sci. Data* **2022**, *7*, 1. [[CrossRef](#)]
57. MuSyQ LCC Product. Available online: <https://www.scidb.cn/en/detail?dataSetId=846695127865884672> (accessed on 22 August 2022).
58. Chlorophyll Meter SPAD-502Plus. Available online: <https://sensing.konicaminolta.asia/product/chlorophyll-meter-spad-502plus/> (accessed on 23 September 2023).
59. Wei, J.; Wang, X. An overview on linear unmixing of hyperspectral data. *Math. Probl. Eng.* **2020**, *2020*, 3735403. [[CrossRef](#)]
60. Du, B.; Zhang, J.; Wang, Z.; Mao, D.; Zhang, M.; Wu, B. Crop mapping based on Sentinel-2A NDVI time series using object-oriented classification and decision tree model. *J. Geo-Inf. Sci.* **2019**, *21*, 740–751. [[CrossRef](#)]
61. Zhou, S.; Kaufmann, H.; Bohn, N.; Bochow, M.; Kuester, T.; Segl, K. Identifying distinct plastics in hyperspectral experimental lab-, aircraft-, and satellite data using machine/deep learning methods trained with synthetically mixed spectral data. *Remote Sens. Environ.* **2022**, *281*, 113263. [[CrossRef](#)]
62. Suwanda, R.; Syahputra, Z.; Zamzami, E.M. Analysis of euclidean distance and manhattan distance in the K-means algorithm for variations number of centroid K. In Proceedings of the 4th International Conference on Computing and Applied Informatics 2019 (ICCAI 2019), Medan, Indonesia, 26–27 November 2019.

63. Yang, W.; Long, H.; Shao, Y.; Du, Q. Chebyshev Distance-based Density Calculation and K-means Clustering Algorithm. *Commun. Technol.* **2019**, *52*, 833–838. [[CrossRef](#)]
64. Aronsson, F. Large Scale Cluster Analysis with Hadoop and Mahout. Available online: <https://lup.lub.lu.se/luur/download?func=downloadFile&recordId=5148140&fileId=5148143> (accessed on 4 July 2024).
65. Carmona, F.; Rivas, R.; Fonnegra, D.C. Vegetation Index to estimate chlorophyll content from multispectral remote sensing data. *Eur. J. Remote Sens.* **2015**, *48*, 319–326. [[CrossRef](#)]
66. Parida, B.R.; Kumari, A. Mapping and modeling mangrove biophysical and biochemical parameters using Sentinel-2A satellite data in Bhitarkanika National Park, Odisha. *Model. Earth Syst. Environ.* **2021**, *7*, 2463–2474. [[CrossRef](#)]
67. Oppelt, N.; Mauser, W. Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *Int. J. Remote Sens.* **2004**, *25*, 145–159. [[CrossRef](#)]
68. Taneja, P.; Vasava, H.B.; Fatholouloumi, S.; Daggupati, P.; Biswas, A. Predicting soil organic matter and soil moisture content from digital camera images: Comparison of regression and machine learning approaches. *Can. J. Soil Sci.* **2022**, *102*, 767–784. [[CrossRef](#)]
69. Rong, C.; Ding, J.; Li, Y. An interdisciplinary survey on origin-destination flows modeling: Theory and techniques. *ACM Comput. Surv.* **2023**, *4*, 1–49. [[CrossRef](#)]
70. Tomo, T.; Allakhverdiev, S.I. *Chlorophyll Species and Their Functions in the Photosynthetic Energy Conversion*; Springer: Cham, Switzerland, 2021; pp. 133–161, ISBN 978-3-030-67406-9.
71. Mahlayeye, M.; Darvishzadeh, R.; Nelson, A. Characterising maize and intercropped maize spectral signatures for cropping pattern classification. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *128*, 103699. [[CrossRef](#)]
72. Hegarty-Craver, M.; Polly, J.; O’Neil, M.; Ujeneza, N.; Rineer, J.; Beach, R.H.; Lapidus, D.; Temple, D.S. Remote crop mapping at scale: Using satellite imagery and UAV-acquired data as ground truth. *Remote Sens.* **2020**, *12*, 1984. [[CrossRef](#)]
73. Gumma, M.K.; Tummala, K.; Dixit, S.; Collivignarelli, F.; Holecz, F.; Kolli, R.N.; Whitbread, A.M. Crop type identification and spatial mapping using Sentinel-2 satellite data with focus on field-level information. *Geocarto Int.* **2022**, *37*, 1833–1849. [[CrossRef](#)]
74. Houborg, R.; Anderson, M.; Daughtry, C. Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at the field scale. *Remote Sens. Environ.* **2009**, *113*, 259–274. [[CrossRef](#)]
75. Croft, H.; Chen, J.; Zhang, Y.; Simic, A.; Noland, T.; Nesbitt, N.; Arabian, J. Evaluating leaf chlorophyll content prediction from multispectral remote sensing data within a physically-based modelling framework. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 85–95. [[CrossRef](#)]
76. Zhang, H.; Li, J.; Liu, Q.; Lin, S.; Huete, A.; Liu, L.; Croft, H.; Clevers, J.G.; Zeng, Y.; Wang, X. A novel red-edge spectral index for retrieving the leaf chlorophyll content. *Methods Ecol. Evol.* **2022**, *13*, 2771–2787. [[CrossRef](#)]
77. Xu, M.; Liu, R.; Chen, J.M.; Liu, Y.; Shang, R.; Ju, W.; Wu, C.; Huang, W. Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach. *Remote Sens. Environ.* **2019**, *224*, 60–73. [[CrossRef](#)]
78. Sudu, B.; Rong, G.; Guga, S.; Li, K.; Zhi, F.; Guo, Y.; Zhang, J.; Bao, Y. Retrieving SPAD values of summer maize using UAV hyperspectral data based on multiple machine learning algorithm. *Remote Sens.* **2022**, *14*, 5407. [[CrossRef](#)]
79. Zhang, Y.; Chen, J.M.; Miller, J.R.; Noland, T.L. Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sens. Environ.* **2008**, *112*, 3234–3247. [[CrossRef](#)]
80. Gao, D.; Li, M.; Zhang, J.; Song, D.; Sun, H.; Qiao, L.; Zhao, R. Improvement of chlorophyll content estimation on maize leaf by vein removal in hyperspectral image. *Comput. Electron. Agric.* **2021**, *184*, 106077. [[CrossRef](#)]
81. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, 08. [[CrossRef](#)]
82. Ciganda, V.S.; Gitelson, A.A.; Schepers, J. How deep does a remote sensor sense? Expression of chlorophyll content in a maize canopy. *Remote Sens. Environ.* **2012**, *126*, 240–247. [[CrossRef](#)]
83. Argenta, G.; Silva, P.R.F.d.; Sangoi, L. Leaf relative chlorophyll content as an indicator parameter to predict nitrogen fertilization in maize. *Cienc. Rural* **2004**, *34*, 1379–1387. [[CrossRef](#)]
84. Hong, S.; Minzan, L.; Yane, Z.; Yong, Z.; Haihua, W. Detection of corn chlorophyll content using canopy spectral reflectance. *Sensor Lett.* **2010**, *8*, 134–139. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.