Article

Retrieval of Crop Canopy Chlorophyll: Machine Learning vs. Radiative Transfer Model

Mir Md Tasnim Alam [1,*], **Anita Simic Milas** [1], **Mateo Gašparović** [2] **and Henry Poku Osei** [1]

1   School of Earth, Environment and Society, Bowling Green State University, Bowling Green, OH 43403, USA
2   Faculty of Geodesy, University of Zagreb, 10000 Zagreb, Croatia
*   Correspondence: alamm@bgsu.edu

Abstract: In recent years, the utilization of machine learning algorithms and advancements in unmanned aerial vehicle (UAV) technology have caused significant shifts in remote sensing practices. In particular, the integration of machine learning with physical models and their application in UAV–satellite data fusion have emerged as two prominent approaches for the estimation of vegetation biochemistry. This study evaluates the performance of five machine learning regression algorithms (MLRAs) for the mapping of crop canopy chlorophyll at the Kellogg Biological Station (KBS) in Michigan, USA, across three scenarios: (1) application to Landsat 7, RapidEye, and PlanetScope satellite images; (2) application to UAV–satellite data fusion; and (3) integration with the PROSAIL radiative transfer model (hybrid methods PROSAIL + MLRAs). The results indicate that the majority of the five MLRAs utilized in UAV–satellite data fusion perform better than the five PROSAIL + MLRAs. The general trend suggests that the integration of satellite data with UAV-derived information, including the normalized difference red-edge index (NDRE), canopy height model, and leaf area index (LAI), significantly enhances the performance of MLRAs. The UAV–RapidEye dataset exhibits the highest coefficient of determination (R^2) and the lowest root mean square errors (RMSE) when employing kernel ridge regression (KRR) and Gaussian process regression (GPR) (R^2 = 0.89 and 0.89 and RMSE = 8.99 µg/cm^2 and 9.65 µg/cm^2, respectively). Similar performance is observed for the UAV–Landsat and UAV–PlanetScope datasets (R^2 = 0.86 and 0.87 for KRR, respectively). For the hybrid models, the maximum performance is attained with the Landsat data using KRR and GPR (R^2 = 0.77 and 0.51 and RMSE = 33.10 µg/cm^2 and 42.91 µg/cm^2, respectively), followed by R^2 = 0.75 and RMSE = 39.78 µg/cm^2 for the PlanetScope data upon integrating partial least squares regression (PLSR) into the hybrid model. Across all hybrid models, the RapidEye data yield the most stable performance, with the R^2 ranging from 0.45 to 0.71 and RMSE ranging from 19.16 µg/cm^2 to 33.07 µg/cm^2. The study highlights the importance of synergizing UAV and satellite data, which enables the effective monitoring of canopy chlorophyll in small agricultural lands.

Keywords: hybrid model; radiative transfer model; machine learning; data fusion; multispectral sensors; unmanned aerial vehicle

1. Introduction

Leaf chlorophyll, an essential element of the photosynthetic system in crops, has long been recognized as one of the leaf pigments most responsive to external conditions. It serves as a primary indicator of plant health and productivity, influencing photosynthesis, nutrient uptake, and stress responses [1]. Two instrumental approaches facilitate the estimation of canopy chlorophyll derived from remote sensing data: (1) the integration and inversion of radiative transfer models (RTMs) and (2) statistical (machine learning) regression methods. In recent years, the integration of RTMs and machine learning statistical methods has emerged as a promising pathway in capturing the complexities of canopy chlorophyll estimation while leveraging both physical principles and model flexibility [2,3].

The integration of leaf and canopy RTMs has been prevalent for several decades [4–8]. While the leaf RTMs simulate the optical properties of leaves based on their biochemical constituents, such as chlorophyll, carotenoids, and leaf structure parameters [9,10], the canopy RTMs simulate the interactions between electromagnetic radiation and vegetation canopies, considering canopy architecture factors like the leaf angle distribution, leaf area index, and soil background [11,12]. This integration allows for the more comprehensive representation of vegetation reflectance by accounting for both the biochemical compositions of individual leaves and the structural characteristics of the canopy. The PROSAIL model is one of the most widely used integrated RTMs; it combines the leaf-level optical properties simulated by the PROSPECT model [13–15] with the canopy-level interactions modeled by the Scattering by Arbitrarily Inclined Leaves (SAIL) model [16–18]. By combining these two models, the PROSAIL model can simulate the spectral reflectance of vegetation canopies with greater accuracy and realism. Over the years, advancements in understanding the optical properties of leaves and canopies have led to refinements in the model's algorithms and parameterizations. These improvements include incorporating more realistic representations of the canopy architecture, leaf angle distribution, and soil background characteristics for SAIL [4,16,19–21] and enhancements in handling pigments and incorporating additional biochemical components for PROSPECT [22,23]. Developed by Verhoef and Bach in the early 2000s [21], the 4SAIL (4-scale SAIL) model introduced additional complexity by incorporating the horizontal and vertical heterogeneities of the canopy into the simulation. The latest leaf model version, the PROSPECT-PRO model, was developed by Feret et al. [24]; it can separate the nitrogen-based constituents (proteins) from the carbon-based constituents and estimate the leaf mass per area as the sum of the proteins and carbon-based constituents. This increased complexity allows for the more detailed and realistic representation of the canopy structure and improves the accuracy of forward (reflectance) simulations, particularly in dense and heterogeneous canopies [25,26].

The inversion of RTMs plays a crucial role in extracting meaningful information from remote sensing data and deriving the canopy and leaf chlorophyll content [27,28]. The inversion techniques of RTMs commonly utilize look-up table (LUT) approaches [7,29–32] and machine learning methods such as artificial neural networks [33,34]. The efficacy of LUT and neural network methods hinges on the training process and a database containing canopy biophysical properties and corresponding canopy reflectance spectra, including parameters like the leaf chlorophyll content and leaf area index [35]. In contrast, the iterative optimization approach allows for the direct retrieval of biophysical parameters from the observed reflectance, bypassing the need for training data, albeit with the drawback of high computational requirements [36]. The ill-posed nature of problems within RTMs underscores the importance of carefully selecting the initial parameter values and incorporating regularization techniques in the inverse process [37,38].

The statistical modeling of the chlorophyll content can involve simple techniques such as a linear regression equation or more sophisticated approaches utilizing machine learning regression algorithms (MLRAs), which excel in identifying patterns in remote sensing data and are particularly effective when dealing with complex, non-linear relationships. MLRAs such as decision trees, support vector machines, or neural networks are trained on the extracted features from remote sensing data and corresponding chlorophyll measurements. In recent times, data fusion techniques have gained popularity in improving the accuracy of estimating the chlorophyll content. By merging data from multiple sources, data fusion techniques help to overcome the limitations of individual data sources and provide a more comprehensive and accurate estimation of the chlorophyll content [39,40]. By integrating data from multiple sources and leveraging the power of machine learning, regression models applied to data fusion offer a robust and efficient approach to retrieving the chlorophyll content in vegetation [41,42]. Once trained, the models can be applied to new remote sensing data, enabling the comprehensive mapping and monitoring of the chlorophyll distribution in vegetation across various ecosystems. The advancements of unmanned aerial vehicle (UAV) technology have led to a new era in remote sensing, enhancing the

modeling process through the synergy between very-high-spatial-resolution UAV data and satellite images. However, the integration of satellite data with UAV-derived information still requires formalization and evaluation across various research studies and at different levels, such as 'data comparison', 'multiscale explanation', 'model calibration', and 'data fusion' [43]. According to Alvarez-Vanhard et al. [43], two thirds of all UAV–satellite data fusion publications are related to enhancing the satellite spatial resolution by incorporating very-high-spatial-resolution information from UAVs. While studies related to the retrieval of the canopy chlorophyl content commonly utilize multispectral or hyperspectral satellite or UAV imagery [44,45], UAV–satellite data fusion has been emerging slowly [40,46,47], and the scientific literature must demonstrate how and to what extent UAV information can enhance different satellite data. It is worth noting that while RTMs offer a more detailed understanding but are often computationally intensive, empirical models are site-specific and lack a physical interpretation of the spectral interactions with vegetation [48,49].

In this study, the aim is to assess and contrast the effectiveness of multiple MLRAs when utilized with UAV–satellite fused data or integrated within hybrid RTMs (PRO-SAIL + MLRAs), as outlined by Verrelst et al. [38,50]. The performance of five MLRAs in mapping the crop canopy chlorophyl is quantified over the Kellogg Biological Station in Michigan, USA, in three scenarios: (1) applied to Landsat 7, RapidEye, and PlanetScope images; (2) applied to UAV–satellite fused data; and (3) integrated within the PROSAIL radiative transfer model (hybrid methods PROSAIL + MLRAs). The five MLRAs are kernel ridge regression (KRR), least squares linear regression (LSLR), partial least squares regression (PLSR), Gaussian process regression (GPR), and neural networks (NN). It is anticipated that this study will attract increasing research interest in the future development of the integration of data-driven machine learning methods and RTMs, with an emphasis on data fusion.

2. Data and Methods

2.1. Study Area

The study site is located within the Kellogg Biological Station (KBS) in Southwestern Michigan, USA (Figure 1). The KBS includes a diverse range of ecosystems, such as agricultural lands, wetlands, forests, and lakes. Since the 1980s, the area has undergone continuous monitoring, offering invaluable long-term data on the agricultural, hydrological, and meteorological processes in the region. The study site (42°24′32.3″N, 85°22′23.8″W) covers a farmland area of approximately 800 m × 1000 m (Figure 1). The site is maintained through the Main Cropping System Experiment (MCSE) of the Long-Term Ecological Research (LTER) program, which is part of a nationwide network of LTER sites established by the National Science Foundation [51,52]. The KBS MCSE site is well known for its rotational cultivation of corn, soybean, and wheat using different chemical and soil management treatments. In 2017, when the current study was conducted, the main crop was corn. Twenty-four parcels of corn were randomly distributed and placed among the parcels of alfalfa, early successional communities, and young poplar trees.

The corn parcels, each with an approximate size of one hectare (87 × 105 m), have been grown under four agricultural treatments (T1–T4), each with 6 replicates (R1–R6) (Figure 1). The conventional tillage (T1) and no tillage (T2) treatments receive conventional levels of chemical inputs. The reduced input (T3) treatment is a biologically based treatment with a low chemical input. The certified organic (T4) treatment receives no chemical inputs, and it is rotary-hoed for weed control.

The MCSE soils are well-drained Alfisol loams of the Kalamazoo series (fine–loamy, mixed, mesic Typic Hapludalfs) mixed with well-drained loams of the Oshtemo series (coarse–loamy, mixed, mesic Typic Hapludalfs), as described by Robertson et al. [51]. The study area has a humid continental climate with warm summers and cold winters.
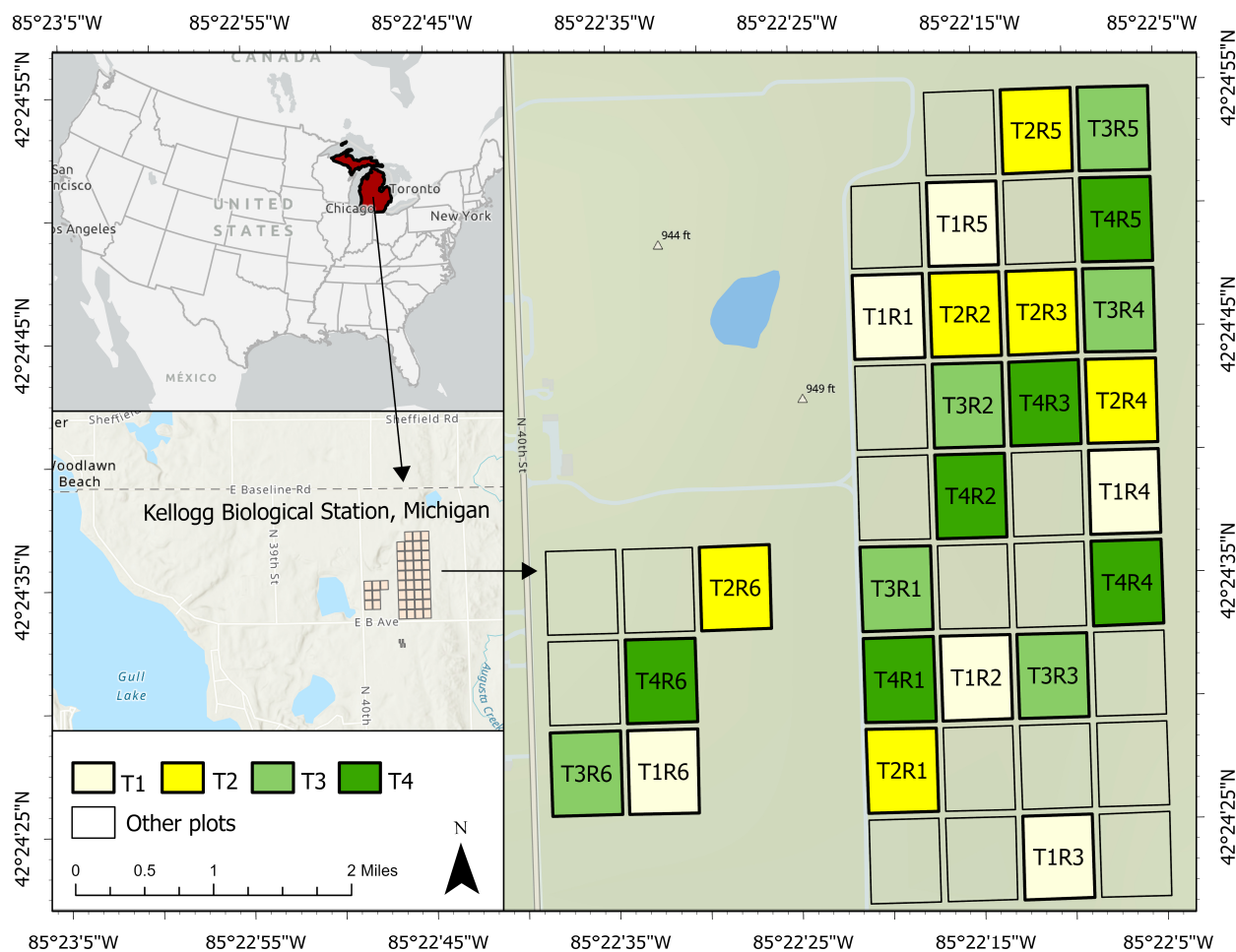
Figure 1. Study site (42°24′32.3′′N, 85°22′23.8′′W) with 24 parcels of corn grown under four chemical and management treatments.

2.2. Field and Satellite Data

Field data and leaf chlorophyll content and leaf area measurements were collected on 11 August 2017. Mid-August 2017 was the peak of the growing season for corn at the study site [53], when the LAI of the crop canopy reached its maximum and the effect of the soil was minimal. Leaf chlorophyl measurements were collected at three randomly chosen locations at each parcel with the Konica Minolta Chlorophyll Meter SPAD-502Plus. The meter has accuracy of ±1 SPAD unit [54]. The measurements were taken over several leaves at each location, at various positions within each leaf, and from the top and middle parts of plants, as expected to be seen by the sensors. The sampling locations were approximately 10–15 m apart and positioned as a triangle around the center of each parcel, to avoid any negative impact from the edge effect and mini-plots placed at some corners of the parcels. Digital hemispherical photographs were taken using a Canon EOS Rebel T5 digital SLR 18.7-megapixel camera with the Sigma 8 mm F3.5 EX DG Circular Fisheye lens at each location and used to derive the leaf area index (LAI). At each location, the set of measurements was collected within a radius of 5 m. The photographs were processed using the Can-Eye (V6.49) software package [55].

The eBee AG Sensefly UAV and the Sequoia camera were used to collect UAV spectral information over the KBS on 11 August 2017 (Figure 2) [56]. The lateral and longitudinal overlap was set to 75%, and the spatial resolution/ground sampling distance was ~13 cm. The reflectance images were created for each green (530–570 nm), red (640–680 nm), red-edge (RE) (730–740 nm), and near-infrared (NIR) (770–810 nm) band (for more information on UAV data collection and processing, see [56]) (Figure 2).
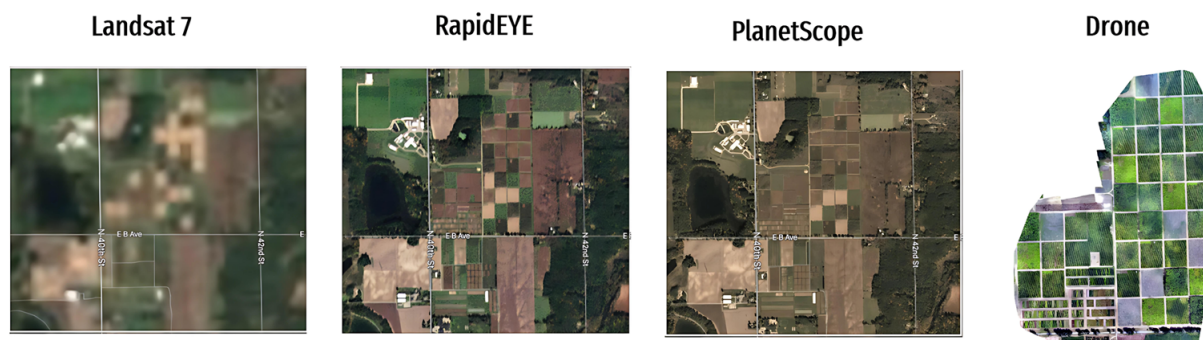
Figure 2. Images acquired in 2017 for the study site by Landsat 7 (8 August), RapidEye (9 August), PlanetScope (8 August), and UAV (11 August).

The Landsat 7 and PlanetScope images were acquired on 8 August 2017, while the RapidEye imagery was obtained on 9 August 2017 (Figure 2 and Table 1). Besides the availability of the imagery, the intention was to explore the performance of imagery with different spatial and spectral resolutions to capture the overall patterns of the findings, i.e., to compare the chlorophyll retrieval capabilities of the MLRAs and hybrid models using medium-, high-, and very-high-spatial-resolution imagery, including data fusion. Landsat 7 was selected for its greater number of spectral bands, aiming to evaluate whether a freely available satellite image with a lower spatial resolution could yield satisfactory results. RapidEye was chosen to examine the impact of including the red-edge spectral band and a fine spatial resolution on the retrieval process. PlanetScope was incorporated into the study due to its exceptionally fine spatial resolution of 3.125 m, despite the absence of the red-edge band. Unfortunately, cloud-free Landsat 8/9 or Sentinel-2 datasets for this period were not available.

Table 1. Sensors and bands used in the current study. Note: GSD—ground sampling distance.

	Landsat 7		RapidEye		PlanetScope		UAV	
	Band Center (nm)	GSD (m)	Band Center (nm)	GSD (m)	Band Center (nm)	GSD (m)	Band Center (nm)	GSD (m)
Blue	485	30	440	5	455	3.125	-	-
Green	560	30	520	5	545	3.125	550	0.13
Red	665	30	670	5	660	3.125	650	0.13
Red Edge	-	-	690	5	-	-	720	0.13
NIR	835	30	760	5	865	3.125	800	0.13
SWIR1	1650	30	-	-	-	-	-	-
SWIR2	2200	30	-	-	-	-	-	-

2.3. Methods and Models

Five machine learning regression algorithms (MLRAs) were employed across the three scenarios to assess their performance in estimating the canopy chlorophyll content (CCC): (1) application to Landsat 7, RapidEye, and PlanetScope images; (2) application to UAV–satellite fused data; and (3) integration within the hybrid radiative transfer model PROSAIL + MLRAs (Figure 3). The five MLRAs utilized in this study were kernel ridge regression (KRR), least squares linear regression (LSLR), partial least squares regression (PLSR), Gaussian process regression (GPR), and neural networks (NN).

Figure 3. The flowchart illustrates the methods employed to derive the canopy chlorophyll content (CCC) across the study area. Note that MLRA stands for 'machine learning regression algorithm', which includes kernel ridge regression (KRR), least squares linear regression (LSLR), partial least squares regression (PLSR), Gaussian process regression (GPR), and neural networks (NN); NV signifies non-vegetated spectra.
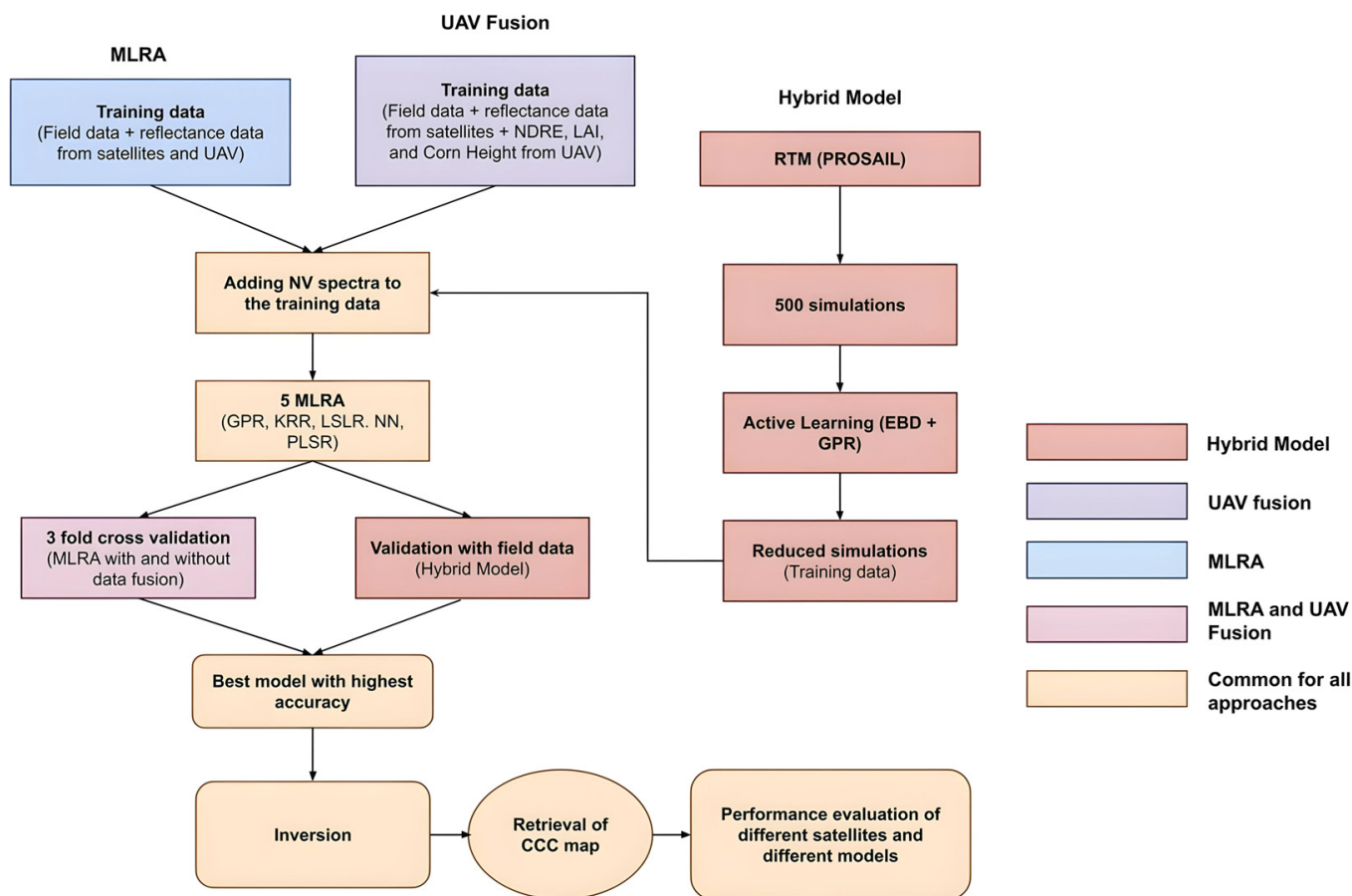
Five diverse sets of MLRAs were implemented to ensure robust and accurate predictions while exploring their relative performance. LSLR, due to its simplicity and interpretability, provides a baseline for performance and is computationally efficient for large datasets, which is paramount for satellite data processing [57]. PLSR was chosen for its ability to handle collinear data and its proven success in estimating the LAI and CCC, outperforming other linear regression methods [31,58,59]. NN offer adaptability and have demonstrated superiority in vegetation mapping, making them a compelling choice for our CCC estimation [60,61]. Moreover, NN have shown enhanced efficiency in mapping vegetation attributes compared to vegetation indices, as demonstrated in studies by Uno et al. [62] and Wang et al. [63]. KRR was included for its competitive performance, speed, and ease in handling high-dimensional spectroscopic data without the need for dimensionality reduction [64,65]. For instance, Wang et al. [66] assessed the high accuracy of KRR in LAI estimation, comparing KRR with multiple linear regression and PLSR. Similarly, Peng et al. [67] demonstrated the effectiveness of KRR in identifying the chlorophyll content, confirming its suitability for the retrieval of the CCC from satellite images. Lastly, GPR was used for its Bayesian framework, which provides probabilistic outputs, offering not only accurate predictions but also quantifying the associated uncertainties, which is particularly valuable in remote sensing applications [68–70]. A comparative analysis conducted by Caicedo et al. [64] revealed that GPR excelled over most other MLRAs in predicting the CCC and LAI from spectroscopic data. This finding was verified by Ashourloo et al. [71], who reported that GPR provided the most accurate detection of leaf rust

disease, surpassing the performance of PLSR and support vector machine (SVM). KRR and GPR are the most efficient methods for the hybrid technique, performing better than other algorithms [3]. Each MLRA brought unique strengths to the study, facilitating a comprehensive approach to accurate CCC retrieval. The fundamental advantages and disadvantages of these MLRAs are outlined in Table 2.

Table 2. Advantages and disadvantages of the MLRAs used in the study.

Algorithm Name	Advantages	Disadvantages	Source
Kernel ridge regression (KRR)	Handles non-linear relationships with kernel functions	The memory requirement for the storage of the kernel matrix can be quite high for large datasets, which can be a limitation for systems with limited memory resources	[72]
Least squares linear regression (LSLR)	Simple, interpretable, computationally efficient	Prone to overfitting with high-dimensional data	[57]
Partial least squares regression (PLSR)	Reduces dimensionality and handles correlated features	Interpretation of coefficients can be challenging	[73]
Gaussian process regression (GPR)	Provides uncertainty estimates for predictions, simple to train, and works well with comparatively smaller datasets	Computationally expensive for large datasets	[69,74]
Neural networks (NN)	Highly flexible, learns complex patterns in data	Can be prone to overfitting and requires careful configuration	[75]

In the first scenario, reflectance values were extracted from each satellite image in the ENVI 5.6 software [76] and paired with the measured CCC values, calculated as the product of the leaf chlorophyll content and LAI, to form the training dataset for each MLRA. In the second scenario, the application to UAV–satellite fused data, the training data included the extracted values from the satellite imagery and from three UAV-derived products, the NDRE index, LAI, and crop height model [56], which were spatially resampled and geo-registered before they were integrated with the satellite information. Additionally, nine non-vegetated spectra (pixels from bare soil, roads, water bodies) were added to the training dataset to enhance the accuracy of the MLRAs. The 3-fold cross-validation method was employed to determine their accuracy.

In the third scenario, the hybrid methods integrated PROSAIL RTM and each MLRA, all implemented within the ARTMO toolbox [77,78]. PROSAIL combines the PROSPECT-PRO RTM, which operates at the leaf level [24], incorporated into 4SAIL, a canopy RTM [16]. The model can simulate the bidirectional reflectance of the canopy across the spectrum of 400 to 2500 nm. The RTM incorporates various biochemical input parameters, such as pigments, proteins, and water content, along with biophysical input parameters like the LAI, average leaf inclination angle, spectral soil background, and viewing geometries [2,20,79].

PROSAIL creates training databases of the vegetation properties and simulates their spectral signals via LUT, a reliable, physically based inversion technique [32]. In Table 3, the input parameters of the PROSAIL RTM are listed. Among the parameters, the LAI and leaf chlorophyll content were determined based on in situ measurements, while the diffuse/direct radiation was set to the default value provided by the ARTMO toolbox [78]. The remaining parameters were established using references from the literature (Table 3). The Latin hypercube sampling method [80] was employed and 500 simulations were selected to train each MLRA to facilitate the avoidance of data redundancy while increasing the computational speed for further analysis [81]. The simulations were resampled to match the 6 bands of Landsat 7, 5 bands of RapidEye, 4 bands of PlanetScope, and 4 bands of the UAV using the spectral bandwidth and response function unique to each sensor. The active learning (AL) strategy was employed to optimize the model by selecting the best

samples from the simulations, which used the Euclidean distance-based diversity (EBD) sampling approach [74] and GPR machine learning regression algorithm [74]. AL is based on intelligent sampling, which optimizes the selection of training datasets from LUTs and increases the accuracy of regression algorithms [74,82]. The hybrid model was executed with each MLRA using satellite imagery: Landsat 7, RapidEye, and PlanetScope, as well as the UAV images.

Table 3. Descriptions and ranges of parameters used as input data in PROSPECT-PRO and 4SAIL models.

Model	Parameter	Description	Unit	Distribution	Range	Source
PROSPECT-PRO	N	Leaf structure	unitless	Uniform	1–2	[83]
	Cab	Leaf chlorophyll content	$\mu g/cm^2$	Uniform	0–80	-
	Ccx	Leaf carotenoid content	$\mu g/cm^2$	Uniform	2–20	[3]
	Canth	Leaf anthocyanin content	$\mu g/cm^2$	Uniform	0–2	[83]
	EWT	Leaf water content	cm	Uniform	0.001–0.02	[83]
	Cp	Leaf protein content	g/cm^2	Uniform	0.001–0.0015	[3]
	Cbrown	Brown pigment content	$\mu g/cm^2$	-	0	[83]
	CBC	Carbon-based constituents	g/cm^2	Uniform	0.001–0.01	[83]
4SAIL	ALA	Average leaf inclination angle	deg	Uniform	20–70	[1]
	LAI	Leaf area index	m^2/m^2	Uniform	0–6	-
	HOT	Hot spot parameter	m/m	Uniform	0.01–0.5	[1]
	SZA	Solar zenith angle	deg	Uniform	20–35	[82]
	OZA	Observer azimuth angle	deg	-	0	[82]
	RAA	Relative azimuth angle	deg	-	0	[82]
	BG	Soil brightness	unitless	-	0.8	[9]
	DR	Diffuse/direct radiation	unitless	-	80	-

CCC maps were generated as outputs for each of the three scenarios for every MLRA. Subsequently, the estimated and measured CCC values were correlated and compared for each model [84]. The effectiveness of each model in predicting the CCC was reported and compared.

3. Results

The five MLRAs applied to satellite images ('without fusion' scenario) demonstrate relatively strong performance for RapidEye and PlanetScope (maximum coefficient of determination $R^2 = 0.62$ and $R^2 = 0.53$, respectively, for GPR), but considerably poorer performance for Landsat 7 (maximum $R^2 = 0.30$ for GPR) (Table 4), although the root mean square error and normalized root mean square error (RMSE = 22.28 $\mu g/cm^2$ and NRMSE = 24.96%, respectively) are relatively low for GPR and similar to other datasets. Other models applied to Landsat 7, such as LSLR (RMSE = 158.53 $\mu g/cm2$, NMRSE = 177.56%) and PLSR (RMSE = 153.76 $\mu g/cm2$, NMRSE = 172.22%), demonstrate exceptionally high error values. Overall, GPR emerges as the most effective method among the five algorithms for all three satellite images. The finer spatial resolution of the RapidEye and PlanetScope images positively influences the outcomes of the regression modeling for most of the algorithms, particularly for GPR. The superior performance observed for RapidEye can be attributed to the presence of the red-edge band, which is known to be a reliable predictor of the chlorophyll content. This underscores the significance of both the spectral and spatial properties of satellite data within this method, such as those provided by the RapidEye and PlanetScope images.

Table 4. Performance of MLRAs with and without data fusion and performance of hybrid PROSAIL + MLRA models using five MLRAs: kernel ridge regression (KRR); Gaussian process regression (GPR); neural network (NN); partial least squares regression (PLSR); least squares linear regression (LSLR).

	Performance of Five MLRAs Applied to Satellite Images								
	Landsat 7			RapidEye			PlanetScope		
	RMSE (μg/cm^2)	NRMSE (%)	R^2	RMSE (μg/cm^2)	NRMSE (%)	R^2	RMSE (μg/cm^2)	NRMSE (%)	R^2
GPR	**22.28**	**24.96**	**0.30**	**16.51**	**18.49**	**0.62**	**19.61**	**21.96**	**0.53**
KRR	24.65	27.61	0.23	20.91	23.42	0.45	22.34	25.03	0.37
LSLR	158.53	177.56	0.07	22.62	25.34	0.34	22.27	24.95	0.40
NN	29.74	33.31	0.25	23.96	26.83	0.37	24.24	27.15	0.33
PLSR	153.76	172.22	0.07	22.62	25.34	0.34	21.11	23.65	0.43
	Performance of five MLRAs applied to fused satellite and UAV images								
GPR	10.61	11.88	0.85	9.65	10.81	0.89	11.69	13.09	0.83
KRR	**10.22**	**11.45**	**0.86**	**8.99**	**10.07**	**0.89**	**9.64**	**10.79**	**0.87**
LSLR	19.06	21.34	0.67	48.50	54.30	0.36	13.37	14.97	0.76
NN	12.83	14.37	0.78	14.41	16.15	0.75	14.66	16.42	0.75
PLSR	24.73	27.70	0.49	79.50	89.90	0.26	24.30	27.21	0.36
	Performance of hybrid PROSAIL + MLRA models applied to satellite images								
GPR	42.91	85.96	0.51	**19.16**	**21.46**	**0.66**	76.33	152.76	0.47
KRR	**33.10**	**66.24**	**0.77**	26.13	29.27	0.69	148.12	296.45	0.57
LSLR	71.83	143.76	0.02	28.54	31.97	0.71	40.66	81.37	0.75
NN	148.72	297.63	0.34	**25.60**	**28.68**	**0.71**	67.73	135.55	0.48
PLSR	73.64	147.37	0.02	27.53	32.99	0.71	**39.78**	**79.60**	**0.75**

Note: best results are in bold.

Combining UAV information (NDRE, LAI, and canopy height model) with satellite data yields considerably improved performance across almost all five MLRAs across all satellite images. Particularly noteworthy is the significant enhancement observed for UAV–Landsat 7, where the R^2 values increase substantially for all five models, with the best performance for KRR, which achieves its maximum value of R^2 = 0.86 (up from R^2 = 0.23 in the 'without fusion' scenario). Moreover, the GPR model demonstrates a significant improvement, with its R^2 increasing from 0.30 to 0.85 when UAV information is integrated with the Landsat 7 dataset. Similarly, for PlanetScope, the data fusion significantly enhances the performance of the MLRAs, with KRR and GPR reaching an R^2 of 0.87 and 0.83, respectively (up from R^2 = 0.37 and 0.53, in the 'without fusion' scenario, respectively). The UAV–RapidEye dataset reaches the best results, with the maximum values of R^2 = 0.89 for both KRR and GPR, compared to its previous R^2 of 0.45 and 0.62, respectively. KRR and GPR consistently emerge as the most effective predictors across images, demonstrating comparatively low RMSE and NRMSE values. While the integration of UAV information significantly enhances the performance of the NN for all three satellites, LSLR and PLSR do not demonstrate any trends. Overall, the results for the data fusion scenario are consistent regardless of the spatial and spectral resolution of the imagery. Although the red-edge band may contribute the most to the optimal results in the UAV–RapidEye scenario, the observed differences across the models are not significant or uniformly improved.

The hybrid models (PROSAIL + MLRAs) exhibit generally higher R^2 values for all three satellite images compared to the empirical MLRA methods for the 'without fusion' scenario. However, the performance is lower for most of the models across all datasets for the 'with fusion' scenario. For instance, GPR shows only a slight improvement for the hybrid models employing RapidEye (R^2 = 0.66) when compared to the 'without fusion' scenario (R^2 = 0.62), while it performs worse for the 'with fusion' scenario (R^2 = 0.89). KRR and GPR are the best models for Landsat data (R^2 = 0.77 and 0.51, respectively; RMSE = 33.10 μg/cm^2 and 42.92 μg/cm^2, respectively). LSLR and PLSR reach their maximum performance for the

data with a finer spatial resolution (RapidEye and PlanetScope) within the hybrid model (i.e., PROSAIL + PLSR: R^2 = 0.71 and 0.75, for RapidEye and PlanetScope, respectively).

Despite the careful parametrization of the PROSAIL model, the RMSE and NRMSE values for most models are exceedingly high across all runs using the Landsat 7 (RMSE = 33.10–148.72 µg/cm^2) and PlanetScope (RMSE = 37.36–148.72 µg/cm^2) imagery. Notably, only the RapidEye-based hybrid models exhibit insignificantly different errors, with the maximum RMSE value of 33.07 µg/cm^2. While the Landsat 7- and PlanetScope-based hybrid models achieve higher R^2 values for KRR and PLSR, respectively, the RapidEye-based hybrid models emerge as the most stable option, demonstrating high R^2 values and the lowest error values across all PROSAIL + MLRAs.

The Landsat-based CCC maps over the study site appear fuzzier compared to the maps produced with the RapidEye and PlanetScope data (Figure 4) due to their coarser spatial resolutions. Notably, the UAV–RapidEye data fusion approach exhibits the largest range of estimated CCC values but has a correlation coefficient R = 0.94 when compared with the field-measured CCC (Figure 5). Among the three models, the 'with fusion' scenario demonstrates the highest correlation coefficient between the measured and estimated CCC values (R = 0.93 from Landsat and PlanetScope and R = 0.94 for RapdiEye) (Figure 5).
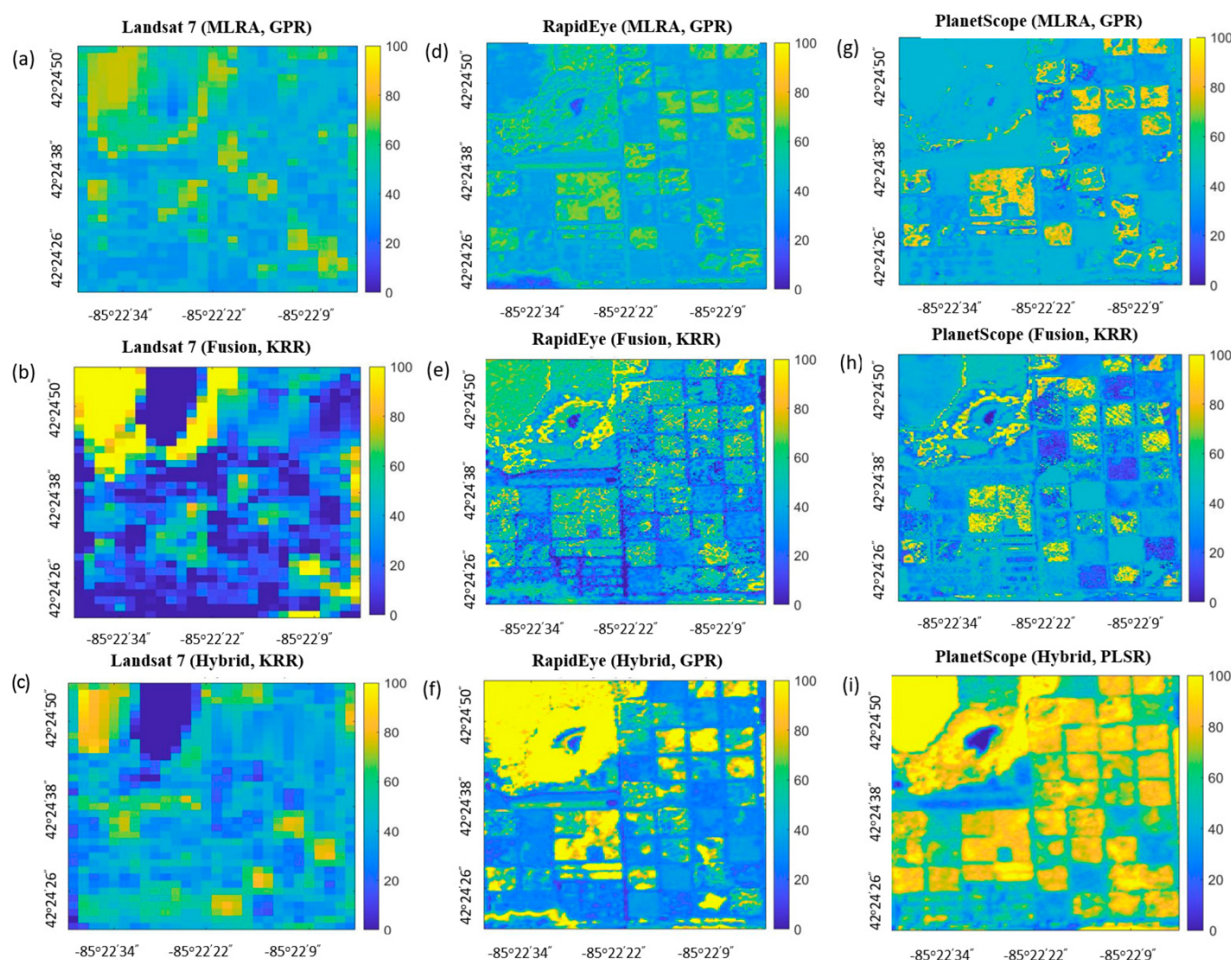

Figure 4. Canopy chlorophyl content (CCC) maps generated from (**a–c**) Landsat 7, (**d–f**) RapidEye, and (**g–i**) PlanetScope satellite images using the best-performing MLRA applied to satellite images, applied to fused UAV–satellite imagery, and integrated within the hybrid model (PROSAIL + MLRA) for each satellite dataset, respectively.
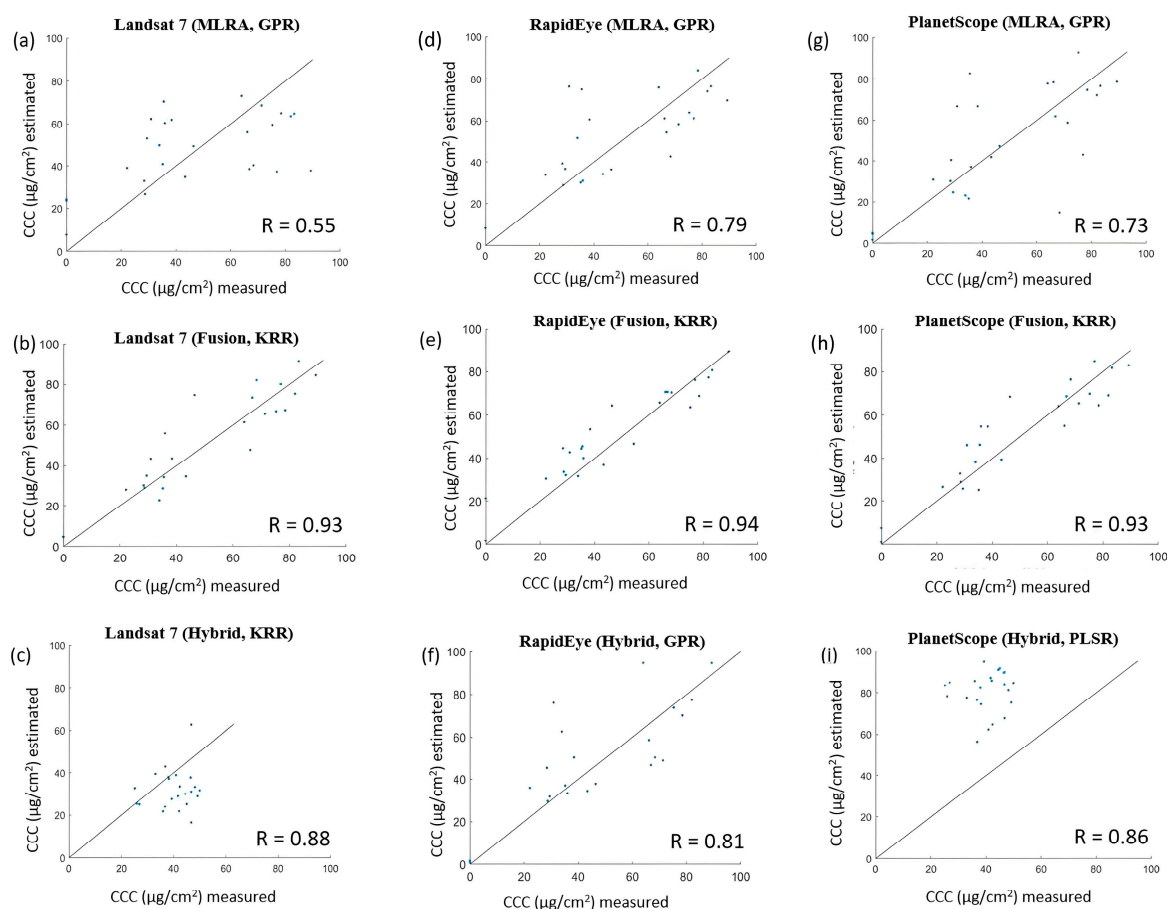
Figure 5. Relationship between measured and estimated canopy chlorophyll content (CCC) values for maps generated from (**a**–**c**) Landsat 7, (**d**–**f**) RapidEye, and (**g**–**i**) PlanetScope satellite images using the best-performing MLRA applied to satellite images, applied to fused UAV–satellite imagery, and integrated within the hybrid model (PROSAIL + MLRA) for each satellite dataset, respectively.

There is a considerable improvement between the estimated and measured CCC values from the 'without fusion' to 'with fusion' scenario, increasing from R = 0.55 to R = 0.93 using the best-performing MLRAs for the Landsat data. The RapidEye-based empirical models demonstrate the highest R values when compared with other satellites (R = 0.79 and R = 0.94 for the 'without fusion' and 'with fusion' scenario, respectively) but slightly lower values within the hybrid modeling than the two other satellites (R = 0.81 for RapidEye vs. R = 0.88 for Landsat and R = 0.86 for PlanetScope). The Landsat-based hybrid model achieves the highest R value of 0.88; however, the estimated CCC values are slightly underestimated. In contrast, the estimated CCC values are highly overestimated for the PlanetScope imagery (Figure 5).

The models were also run on a single UAV image to explore the performance of the models on a very fine-resolution image, potentially serving as a reference point (Table 5). The MLRA methods applied to the UAV bands achieved similar results across all models, with the best performance observed in LSLR, PLSR, and KRR (R^2 = 0.74, 0.73, and 0.72, respectively). The addition of the NDRE, LAI, and height model to the UAV bands considerably improved the performance of all MLRAs, reaching R^2 = 0.92 for KRR. The hybrid model did not perform well for most of the MLRAs. The highest R^2 was reached with KRR (R^2 = 0.25). Overall, all empirical MLRAs generally performed well in both the with and without fusion scenarios. A similar trend is observed when the estimated and measured CCC values are compared for the UAV image, while reaching its best R value for the 'with fusion' scenario (R = 0.96) (Figure 6).

Table 5. Performance of proposed MLRAs applied to UAV image with and without UAV-generated NDRE, LAI, and canopy height information and integrated into the hybrid PROSAIL + MLRA models.

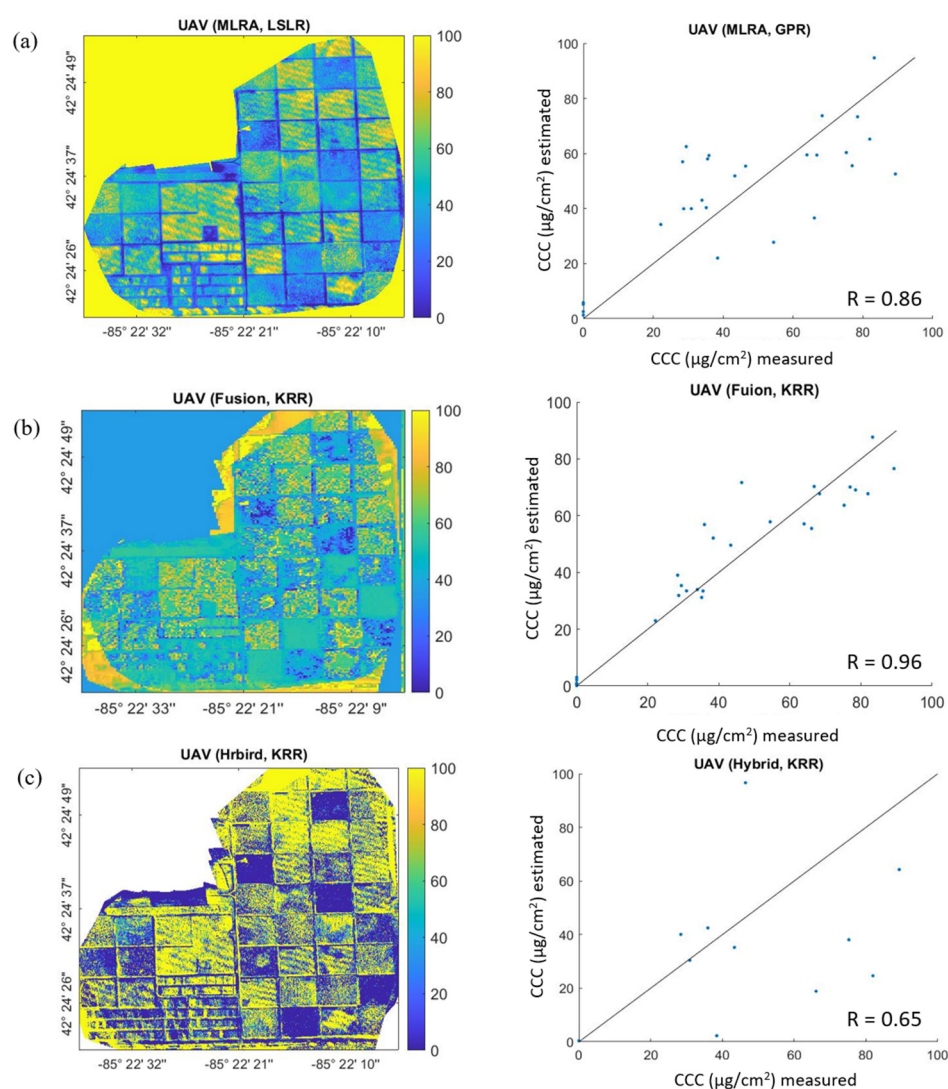|	MLRAs Applied to UAV Image			MLRAs Applied to UAV Image Including UAV-Derived NDRE, LAI, and Canopy Height Model			Hybrid (PROSAIL + MLRA) Applied to UAV Image		
	RMSE $\mu g/cm^2$	NRMSE %	R^2	RMSE $\mu g/cm^2$	NRMSE %	R^2	RMSE $\mu g/cm^2$	NRMSE %	R^2
GPR	17.60	19.72	0.67	9.27	10.38	0.91	38.89	43.56	0.06
KRR	16.11	18.05	0.72	**8.31**	**9.31**	**0.92**	**83.21**	**93.20**	**0.25**
LSLR	**15.57**	**17.44**	**0.74**	9.77	10.94	0.90	92.46	103.57	0.02
NN	18.49	20.70	0.66	13.34	14.94	0.81	35.80	40.10	0.02
PLSR	16.59	18.58	0.73	9.26	10.37	0.91	92.46	103.57	0.02

Note: best results are in bold.

Figure 6. Canopy chlorophyl content (CCC) maps generated from (**a**) UAV image; (**b**) UAV image including UAV-derived NDRE, LAI, and canopy height model; (**c**) hybrid (PROSAIL + MLRA) model applied to UAV image—all using the best-performing MLRA for each scenario—followed by graphs showing the relationship between the measured and estimated CCC.

4. Discussion

4.1. Empirical Modeling Using MRLAs

The primary focus of this study has been to contrast data fusion versus hybrid approaches for small agricultural land. In this context, data fusion involves integrating commonly used UAV-derived products, while hybrid modeling exclusively utilizes satellite bands. The findings affirm the significance of fusing UAV–satellite data for the retrieval of the canopy chlorophyll content in crops. The 'with fusion' empirical MLRAs exhibit notably superior performance compared to other modeling scenarios, including the 'without fusion' empirical MLRA approach and hybrid models (PROSAIL + MLRAs). With the highest R^2 values (max R^2 = 0.89) and the lowest RMSE (RMSE = 9.65 µg/cm^2), this approach demonstrates stability, with minimal performance variability among the five MLRAs. Two MLRAs, GPR and KRR, stand out for their strong and consistent performance: GPR excels as the best-performing model for the 'without fusion' scenario across all three satellites, while KRR leads as the best-performing model for the 'with fusion' empirical scenarios across all three UAV–satellite combinations. Another noticeable trend is the significant improvement in the performance of NN within the data fusion scenario for all satellites. The incorporation of the NDRE bands likely contributed to the NN's enhanced ability to estimate the chlorophyll values regardless of the spatial resolution of the satellite images.

Although the concept is underexplored, several studies in the agriculture or precision agriculture fields demonstrate the benefits of using machine learning techniques in combination with UAVs or the synergies between UAVs and satellites to effectively explore the impact of multiscale phenomena [43]. Various agriculture-related studies show the advantages of using UAV imagery and satellite data, such as Seninel-2 [85,86] and Worldview-2/3 [41]. Singhal et al. [85], who utilized UAV imagery to estimate leaf chlorophyll, praised KRR as the top-performing MLRA among those tested, surpassing GPR. Similarly, Zhou et al. [87] compared multiple MLRAs, including KRR and GPR, and concluded that KRR excelled, particularly when working with high-resolution UAV sensor data. Wang et al. [88] used UAV hyperspectral data to retrieve the CCC and found that the backward propagation neutral network worked better than the support vector machine and PLSR, praising the incorporation of the RE-related parameters in the modeling. On the other hand, Maimaitijiang et al. [41] suggested that random forest regression outperformed PLSR, the support vector machine, and extreme learning regression with a newly proposed activation function.

Based on our observations and a quick sensitivity analysis, two major properties of UAV images significantly contribute to the robust results of the 'with fusion' empirical modeling. Firstly, the very high spatial resolution of UAV imagery enables the detailed and precise capture of vegetation characteristics. Secondly, the inclusion of NDRE proves valuable in detecting subtle changes in chlorophyll content, leveraging the specific absorption properties of chlorophyll in the red-edge region [89,90]. Additionally, it is noted that the UAV-derived canopy height model effectively distinguishes corn parcels from other crops and vegetation types, such as poplar trees, grass, and alfalfa. Although the primary goal of this study was not to map the differences between the four treatments (T1–T4, as described in the Study Site section), most of the models using RapidEye and PlanetScope imagery generally distinguish parcels with different treatments. This is especially evident in the case of T4 (organic) parcels with minimal or no chemical inputs (see Figures 1, 2 and 4). In both 'with fusion' modeling and modeling using solely UAV imagery, the LAI emerges as a pivotal input parameter. Its inclusion significantly boosts the regression power of all models and bolsters the correlation between the estimated and measured CCC. This observation is consistent with the findings of Simic Milas et al. [56], who underscored the importance of the LAI, particularly when combined with UAV-derived biochemical parameters, in CCC estimation. Consequently, incorporating canopy structural parameters into the process of upscaling UAV measurements to the satellite level for CCC monitoring is imperative, especially across heterogeneous fields, as demonstrated in our study. A

similar beneficial approach of integrating UAV-derived canopy structural information with satellite reflectance data was highlighted by [41].

Interestingly, while the 'without fusion' scenario suggests generally more sensitivity to the spatial and spectral resolutions of satellite data—with Landsat-based models exhibiting the least powerful and RapidEye showing the most powerful predictive capabilities for most of the models—this trend is mitigated within the 'with fusion' scenario. It is surmised that very-high-spatial-resolution UAV imagery helps to overcome, or at least minimizes, the gap between in situ and satellite reflectance measurements. Overall, the integration of UAV information with satellite imagery has been shown to enhance the accuracy of analyses [41,43]. Consequently, the ongoing advancement of UAV technology is expected to further bolster the synergistic approach between UAVs and satellites, bridging the gap between sensor capacities and potentially eliminating the need for in situ measurements.

4.2. Hybrid Models (PROSAIL + MRLAs)

The PROSAIL model has been a cornerstone in remote sensing research for more than three decades, owing to its ability to account for both the spectral and structural characteristics of canopies based on leaf-level biochemistry [20,79]. However, the findings of this study suggest that the performance of the hybrid models is strongly influenced by the choice of MLRA used in conjunction with PROSAIL, as well as their capacity to handle the spatial and spectral resolutions of the input satellite data.

Although PROSAIL + KRR demonstrates the highest R^2 for the Landsat data ($R^2 = 0.77$), followed by PROSAIL + LSLR/PLSR for the PlanetScope images ($R^2 = 0.75$), the hybrid models exhibit their best overall performance with the RapidEye data. Across all MLRAs, the RapidEye data display a relatively high R^2 and relatively low errors (max $R^2 = 0.71$, RMSE = 19.16 $\mu g/cm^2$). RapidEye's spatial resolution of 5 m is finer than Landsat's 30 m resolution and coarser than PlanetScope's 3.125 m resolution. However, RapidEye includes the RE band, which is not available in the PlanetScope and Landsat datasets. The combination of spectral and spatial information from RapidEye is anticipated to enhance the stability of the hybrid models, as evidenced by both the R^2 and RMSE/NRMSE values. Similar to the 'with fusion' empirical modeling, the hybrid model with RapidEye demonstrates minimal performance variability among the five MLRAs.

As briefly mentioned above, the RE band is widely recognized for its sensitivity to changes in chlorophyll content [85,91,92]. Over time, numerous RE chlorophyll indices have been developed and documented in the literature, enhancing spectral data for both the canopy chlorophyll content, such as the Canopy Chlorophyll Content Index [93,94], and the leaf chlorophyll content, including the Chlorophyll Index—Red Edge (CLRE) [95], NDRE [96], and Chlorophyll-Sensitive Index (CSI) [97].

When considering both the R^2 and error values as criteria for comparison, KRR demonstrates strong performance for both the Landsat 7 and RapidEye data, generally followed by GPR. Both KRR and GPR excel particularly with datasets containing more abundant spectral information or bands, such as Landsat 7 and RapidEye. Despite their differing mathematical formulations, both KRR and GPR are kernel-based methods. They rely on kernel functions to compute the similarities between pairs of data points in a high-dimensional feature space. This characteristic allows them to capture non-linear relationships between input features and output variables, enabling them to model complex data relationships effectively. Furthermore, both methods incorporate regularization techniques to prevent overfitting, thereby enhancing their generalization performance on unseen data [69,72].

Several studies have suggested that integrating PROSAIL with GPR yields robust performance in hybrid models. For instance, Guo et al. [2] found that integrating PROSAIL with GRP alongside an AL strategy enhanced the model performance (from $R^2 = 0.57$ to $R^2 = 0.74$ and RMSE from 5.60 to 3.96 when applied to Sentinel-2 imagery). GPR is a potent nonparametric probabilistic algorithm [69] that requires a relatively small amount of training data to establish relationships between the spectra and parameters, corresponding to the number of simulations generated by the PROSAIL model. Active learning (AL)-

based optimized training datasets are better suited to real-world scenarios as they are queried against in situ data. GPR trained with AL-based optimized datasets results in higher retrieval certainties compared to the training method with a full dataset [83]. Furthermore, GPR does not necessitate large training datasets, and it generates uncertainties in chlorophyll content estimates [50].

LSLR and PLSR demonstrated superior results in this study for the RapidEye and PlanetScope data compared to the Landsat data, where both methods performed poorly. It is hypothesized that these two statistical approaches work significantly better for fine-spatial-resolution data, a phenomenon also observed in the 'without fusion' empirical modeling. When dealing with fine-spatial-resolution data, both LSLR and PLSR can offer advantages, although they operate differently and may be better suited for different aspects of analysis [98,99]. They are both linear regression techniques that are simple to implement; however, PLSR addresses multicollinearity in a better way and aims to achieve both prediction accuracy and dimensionality reduction simultaneously, while LSLR primarily focuses on predicting the response variable accurately [57,73]. LSLR works well when there is little collinearity among the predictors and the relationship between the predictors and the response can be accurately modeled with a linear equation. In the current study, there was no significant difference in the results yielded by LSLR and PLSR for each satellite. Similarly to LSLR/PLSR, the NN performed weakly for the Landsat data but considerably better for the RapidEye data, suggesting the possible importance of the RE band for this model. However, unlike PLSR and LSLR, the NN is commonly used in remote sensing for the prediction of vegetation parameters and crop yields [100,101]. Nevertheless, it inevitably faces overfitting issues [99]. Additionally, the design parameters and implementation of the NN involve complex and time-consuming processes, and its performance can be compromised when dealing with low-dimensional datasets [102].

Regarding the spectral and spatial resolutions of the input data, the study of Guo et al. [2] revealed that the PROSAIL + GPR model performed better with Planet data compared to four-band Sentinel-2 data but not as well as with 10-band Sentinel-2 imagery. In this study, the same model performed better for Landsat than for PlanetScope. However, the fact that the best performance was obtained with RapidEye underscores the importance of red-edge information in the hybrid models' performance, as explained above. This coincides with the findings of Guo et al. [2].

The hybrid models PROSAIL + MLRAs performed better than the empirical machine learning methods applied to satellite data in our study area, but they did not outperform the UAV data fusion technique. The hybrid model is generally used for a larger study area where there is much heterogeneity in the surface available [38]. The wide range of simulated training data for hybrid models helps the regression algorithms to predict various surfaces apart from only crops. However, for small agricultural fields, simple machine learning and data fusion techniques are preferable for higher overall accuracy and lower uncertainty.

4.3. Uncertainties

One of the main irregularities that may impact the results of this study is related to the process of UAV–satellite data fusion. As explained earlier, the UAV-generated products were properly resampled, geo-registered, and combined within the satellite data. However, a better approach must be utilized, especially for the UAV–Landsat data integration, as the difference between the two spatial resolutions is large (30 m vs. 0.13 m) [86]. The process of geo-registration between satellite and UAV data is a challenging task and may add to the uncertainties.

Although physically based models, as opposed to empirical statistical models, can offer more general and in-depth estimates of biochemical variables, their proper parameterization is critical to correctly simulate remote sensing information used as a reference for model calibration. Similar to other physically based models, the hybrid models used in this study are highly sensitive to some input parameters. However, to minimize the

uncertainties and optimize model performance, the Matlab (R2021a) version of Gaussian process regression (GPR) within the advanced sampling technique AL was applied in the current study to match the simulated reflectance to our field data. Specifically, the Euclidean distance-based diversity (EBD) method was used for sample selection (see the Method section).

Some predictive uncertainties of the hybrid models, attributed to uncertainties in the model parameterization, could be investigated by employing prior networks (PNs), which utilize a unique approach to generate uncertainties in model performance. Unlike traditional methods that focus on the model itself or the training data, PNs concentrate on understanding the uncertainty arising from disparities between the data that the model was trained on and the new data that it encounters. This methodology is advantageous because it assists the model in distinguishing between uncertainty stemming from unfamiliar data and uncertainty regarding its own settings [103]. For this study, Landsat 7 data were used to build the model and compare the performance with satellites that had a finer spatial resolution. Landsat 8/9 or Sentinel-2 could have been used for this purpose, but the higher cloud coverage of these images at the study time hindered us from choosing these satellites.

The very fine spatial resolution of the UAV image with the RE and NIR bands, which are commonly highly correlated with canopy chlorophyll [104,105], demonstrated superior results for MLRAs compared to the hybrid method. One reason could be related to the parameterization. The input parameters (i.e., the solar zenith angle, observer azimuth angle, relative azimuth angle) for the PROSAIL model were adapted from previous studies that were developed using satellite data. As a result, PROSAIL could not precisely simulate the reflectance of UAV imagery, leading to suboptimal results with the hybrid technique. However, future research could address this limitation by refining the parameterization to better accommodate UAV data. Moreover, in the current study, the focus was on corn; however, the input parameters may vary across different types of vegetation and should be modified accordingly for different crop types.

Due to their somewhat coarse spatial resolution compared to the size of the parcels in the study area, Landsat data could introduce some minor uncertainties. However, a visual evaluation of the extracted values from the central area of the parcels minimized the negative impact of the parcel-edge effect in the current study. The parcels were generally homogenous and flat. Thus, Landsat data are a viable option for the agricultural settings of the Kellog Biological Station.

In summary, integrating UAV data with data fusion techniques and machine learning algorithms provides a potent method to improve the chlorophyll retrieval accuracy. This involves merging high-resolution UAV imagery with complementary satellite data and utilizing advanced machine learning techniques. Such integration enables researchers to gain more precise insights into the chlorophyll distribution within vegetation, thereby enhancing our understanding of plant health, productivity, and environmental dynamics. The increasing adoption of UAVs in remote sensing applications is due to their outstanding spatial resolution capabilities, versatility, and adaptability in data acquisition. This was reaffirmed by employing solely UAV data in all three scenarios in this study. While the results of empirical modeling using UAV data alone demonstrated strong performance when based on UAV bands (including the red-edge band), the incorporation of additional information (such as NDRE, LAI, and canopy height model) significantly enhanced the performance of all MLRAs for the UAV image. Another advantage of UAV–satellite data fusion was observed in its ability to neutralize the impact of the spatial and spectral resolutions of the satellite data on the performance of the models.

The primary message of this study is not to advocate for MLRAs over UAV–satellite data, but rather to emphasize the importance of enhancing the integration of UAV–satellite data fusion within hybrid models. RTMs indeed play a crucial role in estimating the chlorophyll content using remote sensing data, including both UAV and satellite imagery. RTMs allow sensitivity analyses to understand the influence of various factors, such as the leaf biochemical content, canopy structure, and atmospheric conditions, on the observed

spectral signals and chlorophyll estimation accuracy. Overall, RTMs are indispensable tools in understanding the physical principles underlying remote sensing observations and in developing accurate and robust methods for the estimation of the chlorophyll content in vegetation using UAV and satellite data [26]. In the current study, the comparison of the two modeling approaches was based on a field campaign and satellite data acquired in mid-August 2017, during the peak of the growing season [53], to minimize the impact of the crop phenology, varying LAI values, and soil effects. In future studies, the influence of the crop growth stages on chlorophyl retrieval using standalone MLRAs and hybrid models should be considered. Due to phenological changes, enhanced by agricultural treatments, the structural and biochemical properties of crops can significantly alter their spectral responses, affecting the relationship between the crop parameters and vegetation spectral information. For instance, vegetation indices often suffer from problems related to spectral saturation, vegetation senescence, the soil background, and the canopy structure. Moreover, the acquisition of ground data becomes more uncertain due to variations in the leaf area and leaf chlorophyl content distribution. Regardless of the approach used to retrieve the CCC, the early and late growing stages of crops result in unavoidable uncertainties [106,107].

5. Conclusions

This study evaluated the performance of five machine learning regression algorithms (MLRAs) for the mapping of the crop canopy chlorophyll content (CCC) at the Kellogg Biological Station (KBS) in Michigan, USA, across three scenarios: (1) application to Landsat 7, RapidEye, and PlanetScope images; (2) application to UAV–satellite data fusion; and (3) integration within the hybrid radiative transfer model (PROSAIL + MLRAs). The five MLRAs were kernel ridge regression (KRR), least squares linear regression (LSLR), partial least squares regression (PLSR), Gaussian process regression (GPR), and a neural network (NN). The research also investigated the impact of the different spatial and spectral resolutions of the satellite data on the performance of the five MLRAs. Based on the results obtained, the following overall conclusions were drawn.

- The five MLRAs applied to UAV–satellite data fusion outperformed their application to satellite bands or integration within hybrid models (PROSAIL + MLRAs) in small agricultural areas such as the KBS.
- UAV–satellite data fusion neutralized and mitigated the impact of the spatial and spectral resolution of the satellite imagery on the MLRAs' performance.
- The red-edge-related information of RapidEye proved advantageous for all models across all three study scenarios, contributing to the stability of the models with minimal performance variability.
- The leaf area index (LAI) emerged as a critical parameter, necessitating incorporation with UAV-derived products in estimating biochemical parameters.
- The choice of MLRAs significantly influenced the performance of the hybrid models (PROSAIL + MLRAs).
- GPR and KRR emerged as standout models, demonstrating strong performance across various scenarios.

This study emphasizes the crucial role of integrating both UAV and satellite data to optimize the utilization of MLRAs for the mapping of the canopy chlorophyll content in small agricultural areas. It provides valuable insights for further advancement in hybrid model development.

Author Contributions: Conceptualization, M.M.T.A. and A.S.M.; methodology, M.M.T.A.; software, M.M.T.A.; validation, M.M.T.A. and A.S.M.; formal analysis, M.M.T.A.; resources, M.M.T.A. and M.G.; data curation, M.M.T.A.; writing—original draft preparation, M.M.T.A. and A.S.M.; writing—review and editing, M.M.T.A., M.G., H.P.O. and A.S.M.; visualization, M.M.T.A. and H.P.O.; supervision, A.S.M. and M.G.; project administration, A.S.M. All authors have read and agreed to the published version of the manuscript.

References

1. Chakhvashvili, E.; Siegmann, B.; Muller, O.; Verrelst, J.; Bendig, J.; Kraska, T.; Rascher, U. Retrieval of Crop Variables from Proximal Multispectral UAV Image Data Using PROSAIL in Maize Canopy. *Remote Sens.* **2022**, *14*, 1247. [CrossRef] [PubMed]
2. Guo, A.; Ye, H.; Li, G.; Zhang, B.; Huang, W.; Jiao, Q.; Qian, B.; Luo, P. Evaluation of Hybrid Models for Maize Chlorophyll Retrieval Using Medium-and High-Spatial-Resolution Satellite Images. *Remote Sens.* **2023**, *15*, 1784. [CrossRef]
3. Tagliabue, G.; Boschetti, M.; Bramati, G.; Candiani, G.; Colombo, R.; Nutini, F.; Pompilio, L.; Rivera-Caicedo, J.P.; Rossi, M.; Rossini, M.; et al. Hybrid Retrieval of Crop Traits from Multi-Temporal PRISMA Hyperspectral Imagery. *ISPRS J. Photogramm. Remote Sens.* **2022**, *187*, 362–377. [CrossRef] [PubMed]
4. Jacquemoud, S.; Baret, F.; Andrieu, B.; Danson, F.M.; Jaggard, K. Extraction of Vegetation Biophysical Parameters by Inversion of the PROSPECT+ SAIL Models on Sugar Beet Canopy Reflectance Data. *Appl. TM AVIRIS Sensors. Remote Sens. Environ.* **1995**, *52*, 163–172. [CrossRef]
5. Kuusk, A. A Fast, Invertible Canopy Reflectance Model. *Remote Sens. Environ.* **1995**, *51*, 342–350. [CrossRef]
6. Bicheron, P.; Leroy, M. A Method of Biophysical Parameter Retrieval at Global Scale by Inversion of a Vegetation Reflectance Model. *Remote Sens. Environ.* **1999**, *67*, 251–266. [CrossRef]
7. Simic, A.; Chen, J.M.; Noland, T.L. Retrieval of Forest Chlorophyll Content Using Canopy Structure Parameters Derived from Multi-Angle Data: The Measurement Concept of Combining Nadir Hyperspectral and off-Nadir Multispectral Data. *Int. J. Remote Sens.* **2011**, *32*, 5621–5644. [CrossRef]
8. Croft, H.; Chen, J.M.; Zhang, Y.; Simic, A. Modelling Leaf Chlorophyll Content in Broadleaf and Needle Leaf Canopies from Ground, CASI, Landsat TM 5 and MERIS Reflectance Data. *Remote Sens. Environ.* **2013**, *133*, 128–140. [CrossRef]
9. Sun, J.; Shi, S.; Wang, L.; Li, H.; Wang, S.; Gong, W.; Tagesson, T. Optimizing LUT-Based Inversion of Leaf Chlorophyll from Hyperspectral Lidar Data: Role of Cost Functions and Regulation Strategies. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102602. [CrossRef]
10. Miraglio, T.; Adeline, K.; Huesca, M.; Ustin, S.; Briottet, X. Monitoring LAI, Chlorophylls, and Carotenoids Content of a Woodland Savanna Using Hyperspectral Imagery and 3D Radiative Transfer Modeling. *Remote Sens.* **2019**, *12*, 28. [CrossRef]
11. Koetz, B.; Sun, G.; Morsdorf, F.; Ranson, K.J.; Kneubühler, M.; Itten, K.; Allgöwer, B. Fusion of Imaging Spectrometer and LIDAR Data over Combined Radiative Transfer Models for Forest Canopy Characterization. *Remote Sens. Environ.* **2007**, *106*, 449–459. [CrossRef]
12. Schiefer, F.; Schmidtlein, S.; Kattenborn, T. The Retrieval of Plant Functional Traits from Canopy Spectra through RTM-Inversions and Statistical Models Are Both Critically Affected by Plant Phenology. *Ecol. Indic.* **2021**, *121*, 107062. [CrossRef]
13. Jacquemoud, S.; Baret, F. PROSPECT: A Model of Leaf Optical Properties Spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91. [CrossRef]
14. Romero, A.; Aguado, I.; Yebra, M. Estimation of Dry Matter Content in Leaves Using Normalized Indexes and PROSPECT Model Inversion. *Int. J. Remote Sens.* **2012**, *33*, 396–414. [CrossRef]
15. Shiklomanov, A.N.; Dietze, M.C.; Viskari, T.; Townsend, P.A.; Serbin, S.P. Quantifying the Influences of Spectral Resolution on Uncertainty in Leaf Trait Estimates through a Bayesian Approach to RTM Inversion. *Remote Sens. Environ.* **2016**, *183*, 226–238. [CrossRef]
16. Verhoef, W. Light Scattering by Leaf Layers with Application to Canopy Reflectance Modeling: The SAIL Model. *Remote Sens. Environ.* **1984**, *16*, 125–141. [CrossRef]
17. Han, D.; Liu, J.; Zhang, R.; Liu, Z.; Guo, T.; Jiang, H.; Wang, J.; Zhao, H.; Ren, S.; Yang, P. Evaluation of the SAIL Radiative Transfer Model for Simulating Canopy Reflectance of Row Crop Canopies. *Remote Sens.* **2023**, *15*, 5433. [CrossRef]
18. Andrieu, B.; Baret, F.; Jacquemoud, S.; Malthus, T.; Steven, M. Evaluation of an Improved Version of SAIL Model for Simulating Bidirectional Reflectance of Sugar Beet Canopies. *Remote Sens. Environ.* **1997**, *60*, 247–257. [CrossRef]
19. Verhoef, W.; Jia, L.; Su, Z. *Optical-Thermal Canopy Radiance Directionality Modelling by Unified 4SAIL Model*; National Aerospace Laboratory NLR: Amsterdam, The Netherlands, 2007.
20. Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P.J.; Asner, G.P.; François, C.; Ustin, S.L. PROSPECT+ SAIL Models: A Review of Use for Vegetation Characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. [CrossRef]

21. Verhoef, W.; Bach, H. Coupled Soil–Leaf-Canopy and Atmosphere Radiative Transfer Modeling to Simulate Hyperspectral Multi-Angular Surface Reflectance and TOA Radiance Data. *Remote Sens. Environ.* **2007**, *109*, 166–182. [CrossRef]

22. Schaepman, M.E.; Wamelink, G.W.W.; van Dobben, H.F.; Gloor, M.; Schaepman-Strub, G.; Kooistra, L.; Clevers, J.G.P.W.; Schmidt, A.; Berendse, F. River Floodplain Vegetation Scenario Development Using Imaging Spectroscopy Derived Products as Input Variables in a Dynamic Vegetation Model. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 1179–1188. [CrossRef]

23. Malenovský, Z.; Albrechtová, J.; Lhotáková, Z.; Zurita-Milla, R.; Clevers, J.; Schaepman, M.E.; Cudlín, P. Applicability of the PROSPECT Model for Norway Spruce Needles. *Int. J. Remote Sens.* **2006**, *27*, 5315–5340. [CrossRef]

24. Féret, J.-B.; Berger, K.; de Boissieu, F.; Malenovský, Z. PROSPECT-PRO for Estimating Content of Nitrogen-Containing Leaf Proteins and Other Carbon-Based Constituents. *Remote Sens. Environ.* **2021**, *252*, 112173. [CrossRef]

25. Parry, C.K.; Nieto, H.; Guillevic, P.; Agam, N.; Kustas, W.P.; Alfieri, J.; McKee, L.; McElrone, A.J. An Intercomparison of Radiation Partitioning Models in Vineyard Canopies. *Irrig. Sci.* **2019**, *37*, 239–252. [CrossRef]

26. Cao, B.; Guo, M.; Fan, W.; Xu, X.; Peng, J.; Ren, H.; Du, Y.; Li, H.; Bian, Z.; Hu, T. A New Directional Canopy Emissivity Model Based on Spectral Invariants. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6911–6926. [CrossRef]

27. Chaabouni, S.; Kallel, A.; Houborg, R. Improving Retrieval of Crop Biophysical Properties in Dryland Areas Using a Multi-Scale Variational RTM Inversion Approach. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *94*, 102220. [CrossRef]

28. Verrelst, J.; Rivera, J.P.; Leonenko, G.; Alonso, L.; Moreno, J. Optimizing LUT-Based RTM Inversion for Semiautomatic Mapping of Crop Biophysical Parameters from Sentinel-2 and-3 Data: Role of Cost Functions. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 257–269. [CrossRef]

29. Vicent, J.; Verrelst, J.; Sabater, N.; Alonso, L.; Rivera-Caicedo, J.P.; Martino, L.; Muñoz-Marí, J.; Moreno, J. Comparative Analysis of Atmospheric Radiative Transfer Models Using the Atmospheric Look-up Table Generator (ALG) Toolbox (Version 2.0). *Geosci. Model Dev.* **2020**, *13*, 1945–1957. [CrossRef]

30. Vicent, J.; Sabater, N.; Alonso, L.; Verrelst, J.; Moreno, J. Alg: A Toolbox for the Generation of Look-Up Tables Based on Atmospheric Radiative Transfer Models. In Proceedings of the 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 23–26 September 2018; pp. 1–5.

31. Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C. Inversion of a Radiative Transfer Model for Estimating Vegetation LAI and Chlorophyll in a Heterogeneous Grassland. *Remote Sens. Environ.* **2008**, *112*, 2592–2604. [CrossRef]

32. Weiss, M.; Baret, F.; Myneni, R.; Pragnère, A.; Knyazikhin, Y. Investigation of a Model Inversion Technique to Estimate Canopy Biophysical Variables from Spectral and Directional Reflectance Data. *Agronomie* **2000**, *20*, 3–22. [CrossRef]

33. He, Y.; Gong, Z.; Zheng, Y.; Zhang, Y. Inland Reservoir Water Quality Inversion and Eutrophication Evaluation Using BP Neural Network and Remote Sensing Imagery: A Case Study of Dashahe Reservoir. *Water* **2021**, *13*, 2844. [CrossRef]

34. Ai, B.; Wen, Z.; Jiang, Y.; Gao, S.; Lv, G. Sea Surface Temperature Inversion Model for Infrared Remote Sensing Images Based on Deep Neural Network. *Infrared Phys. Technol.* **2019**, *99*, 231–239. [CrossRef]

35. Houborg, R.; Boegh, E. Mapping Leaf Chlorophyll and Leaf Area Index Using Inverse and Forward Canopy Reflectance Modeling and SPOT Reflectance Data. *Remote Sens. Environ.* **2008**, *112*, 186–202. [CrossRef]

36. Jacquemoud, S.; Bacour, C.; Poilvé, H.; Frangi, J.-P. Comparison of Four Radiative Transfer Models to Simulate Plant Canopies Reflectance: Direct and Inverse Mode. *Remote Sens. Environ.* **2000**, *74*, 471–481. [CrossRef]

37. Kimes, D.S.; Knyazikhin, Y.; Privette, J.L.; Abuelgasim, A.A.; Gao, F. Inversion Methods for Physically-based Models. *Remote Sens. Rev.* **2000**, *18*, 381–439. [CrossRef]

38. Verrelst, J.; Malenovský, Z.; Van der Tol, C.; Camps-Valls, G.; Gastellu-Etchegorry, J.-P.; Lewis, P.; North, P.; Moreno, J. Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surv. Geophys.* **2019**, *40*, 589–629. [CrossRef] [PubMed]

39. Boiarskii, B.; Hasegawa, H. Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content. *J. Mech. Contin. Math. Sci.* **2019**, *spl1*, 20–29. [CrossRef]

40. Chen, C.; Chen, Q.; Li, G.; He, M.; Dong, J.; Yan, H.; Wang, Z.; Duan, Z. A Novel Multi-Source Data Fusion Method Based on Bayesian Inference for Accurate Estimation of Chlorophyll-a Concentration over Eutrophic Lakes. *Environ. Model. Softw.* **2021**, *141*, 105057. [CrossRef]

41. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Daloye, A.M.; Erkbol, H.; Fritschi, F.B. Crop Monitoring Using Satellite/UAV Data Fusion and Machine Learning. *Remote Sens.* **2020**, *12*, 1357. [CrossRef]

42. Chusnah, W.N.; Chu, H.-J.; Tatas; Jaelani, L.M. Machine-Learning-Estimation of High-Spatiotemporal-Resolution Chlorophyll-a Concentration Using Multi-Satellite Imagery. *Sustain. Environ. Res.* **2023**, *33*, 11. [CrossRef]

43. 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]

44. Shu, M.; Shen, M.; Zuo, J.; Yin, P.; Wang, M.; Xie, Z.; Tang, J.; Wang, R.; Li, B.; Yang, X.; et al. The Application of UAV-Based Hyperspectral Imaging to Estimate Crop Traits in Maize Inbred Lines. *Plant Phenomics* **2021**, *2021*, 9890745. [CrossRef] [PubMed]

45. Lelong, C.C.; Burger, P.; Jubelin, G.; Roux, B.; Labbé, S.; Baret, F. Assessment of Unmanned Aerial Vehicles Imagery for Quantitative Monitoring of Wheat Crop in Small Plots. *Sensors* **2008**, *8*, 3557–3585. [CrossRef] [PubMed]

46. Gunia, M.; Laine, M.; Malve, O.; Kallio, K.; Kervinen, M.; Anttila, S.; Kotamäki, N.; Siivola, E.; Kettunen, J.; Kauranne, T. Data Fusion System for Monitoring Water Quality: Application to Chlorophyll-a in Baltic Sea Coast. *Environ. Model. Softw.* **2022**, *155*, 105465. [CrossRef]

47. Maimaitijiang, M.; Ghulam, A.; Sidike, P.; Hartling, S.; Maimaitiyiming, M.; Peterson, K.; Shavers, E.; Fishman, J.; Peterson, J.; Kadam, S. Unmanned Aerial System (UAS)-Based Phenotyping of Soybean Using Multi-Sensor Data Fusion and Extreme Learning Machine. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 43–58. [CrossRef]

48. Houborg, R.; Soegaard, H.; Boegh, E. Combining Vegetation Index and Model Inversion Methods for the Extraction of Key Vegetation Biophysical Parameters Using Terra and Aqua MODIS Reflectance Data. *Remote Sens. Environ.* **2007**, *106*, 39–58. [CrossRef]

49. Sefer, A.; Yapar, A.; Yelkenci, T. Imaging of Rough Surfaces by RTM Method. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–12. [CrossRef]

50. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine Learning Regression Algorithms for Biophysical Parameter Retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [CrossRef]

51. Robertson, G.P.; Collins, S.L.; Foster, D.R.; Brokaw, N.; Ducklow, H.W.; Gragson, T.L.; Gries, C.; Hamilton, S.K.; McGuire, A.D.; Moore, J.C.; et al. Long-Term Ecological Research in a Human-Dominated World. *BioScience* **2012**, *62*, 342–353. [CrossRef]

52. Robertson, G.P. *The Ecology of Agricultural Landscapes: Long-Term Research on the Path to Sustainability*; Oxford University Press: Oxford, UK, 2015.

53. Simic Milas, A.; Vincent, R.K. Monitoring Landsat Vegetation Indices for Different Crop Treatments and Soil Chemistry. *Int. J. Remote Sens.* **2017**, *38*, 141–160. [CrossRef]

54. Uddling, J.; Gelang-Alfredsson, J.; Piikki, K.; Pleijel, H. Evaluating the Relationship between Leaf Chlorophyll Concentration and SPAD-502 Chlorophyll Meter Readings. *Photosynth. Res.* **2007**, *91*, 37–46. [CrossRef]

55. CAN_EYE_User_Manual.Pdf. Available online: https://jecam.org/wp-content/uploads/2018/07/CAN_EYE_User_Manual.pdf (accessed on 9 May 2024).

56. Simic Milas, A.; Romanko, M.; Reil, P.; Abeysinghe, T.; Marambe, A. The Importance of Leaf Area Index in Mapping Chlorophyll Content of Corn under Different Agricultural Treatments Using UAV Images. *Int. J. Remote Sens.* **2018**, *39*, 5415–5431. [CrossRef]

57. Wolberg, J.; Wolberg, E.J. The Method of Least Squares. In *Designing Quantitative Experiments: Prediction Analysis*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 47–89.

58. Atzberger, C.; Guérif, M.; Baret, F.; Werner, W. Comparative Analysis of Three Chemometric Techniques for the Spectroradiometric Assessment of Canopy Chlorophyll Content in Winter Wheat. *Comput. Electron. Agric.* **2010**, *73*, 165–173. [CrossRef]

59. Yi, Q.; Jiapaer, G.; Chen, J.; Bao, A.; Wang, F. Different Units of Measurement of Carotenoids Estimation in Cotton Using Hyperspectral Indices and Partial Least Square Regression. *ISPRS J. Photogramm. Remote Sens.* **2014**, *91*, 72–84. [CrossRef]

60. Kalacska, M.; Lalonde, M.; Moore, T.R. Estimation of Foliar Chlorophyll and Nitrogen Content in an Ombrotrophic Bog from Hyperspectral Data: Scaling from Leaf to Image. *Remote Sens. Environ.* **2015**, *169*, 270–279. [CrossRef]

61. Malenovský, Z.; Homolová, L.; Zurita-Milla, R.; Lukeš, P.; Kaplan, V.; Hanuš, J.; Gastellu-Etchegorry, J.-P.; Schaepman, M.E. Retrieval of Spruce Leaf Chlorophyll Content from Airborne Image Data Using Continuum Removal and Radiative Transfer. *Remote Sens. Environ.* **2013**, *131*, 85–102. [CrossRef]

62. Uno, Y.; Prasher, S.O.; Lacroix, R.; Goel, P.K.; Karimi, Y.; Viau, A.; Patel, R.M. Artificial Neural Networks to Predict Corn Yield from Compact Airborne Spectrographic Imager Data. *Comput. Electron. Agric.* **2005**, *47*, 149–161. [CrossRef]

63. Wang, F.; Huang, J.; Wang, Y.; Liu, Z.; Peng, D.; Cao, F. Monitoring Nitrogen Concentration of Oilseed Rape from Hyperspectral Data Using Radial Basis Function. *Int. J. Digit. Earth* **2013**, *6*, 550–562. [CrossRef]

64. Caicedo, J.P.R.; Verrelst, J.; Muñoz-Marí, J.; Moreno, J.; Camps-Valls, G. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1249–1259. [CrossRef]

65. Rivera-Caicedo, J.P.; Verrelst, J.; Muñoz-Marí, J.; Camps-Valls, G.; Moreno, J. Hyperspectral Dimensionality Reduction for Biophysical Variable Statistical Retrieval. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 88–101. [CrossRef]

66. Wang, F.; Huang, J.; Lou, Z. A Comparison of Three Methods for Estimating Leaf Area Index of Paddy Rice from Optimal Hyperspectral Bands. *Precis. Agric.* **2011**, *12*, 439–447. [CrossRef]

67. Peng, Y.; Huang, H.; Wang, W.; Wu, J.; Wang, X. Rapid Detection of Chlorophyll Content in Corn Leaves by Using Least Squares-Support Vector Machines and Hyperspectral Images. *J. Jiangsu Univ. -Nat. Sci. Ed.* **2011**, *32*, 125–174.

68. Camps-Valls, G.; Verrelst, J.; Munoz-Mari, J.; Laparra, V.; Mateo-Jimenez, F.; Gomez-Dans, J. A Survey on Gaussian Processes for Earth-Observation Data Analysis: A Comprehensive Investigation. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 58–78. [CrossRef]

69. Williams, C.K.; Rasmussen, C.E. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006; Volume 2.

70. Verrelst, J.; Alonso, L.; Caicedo, J.P.R.; Moreno, J.; Camps-Valls, G. Gaussian Process Retrieval of Chlorophyll Content from Imaging Spectroscopy Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *6*, 867–874. [CrossRef]

71. Ashourloo, D.; Aghighi, H.; Matkan, A.A.; Mobasheri, M.R.; Rad, A.M. An Investigation into Machine Learning Regression Techniques for the Leaf Rust Disease Detection Using Hyperspectral Measurement. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 4344–4351. [CrossRef]

72. Suykens, J.a.K.; Vandewalle, J. Chaos Control Using Least-Squares Support Vector Machines. *Int. J. Circuit Theory Appl.* **1999**, *27*, 605–615. [CrossRef]

73. Geladi, P.; Kowalski, B.R. Partial Least-Squares Regression: A Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [CrossRef]

74. Verrelst, J.; Rivera-Caicedo, J.P.; Reyes-Muñoz, P.; Morata, M.; Amin, E.; Tagliabue, G.; Panigada, C.; Hank, T.; Berger, K. Mapping Landscape Canopy Nitrogen Content from Space Using PRISMA Data. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 382–395. [CrossRef]

75. Hagan, M.T.; Menhaj, M.B. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [CrossRef]

76. NV5 Geospatial Solutions & Services Expertise. Available online: https://www.nv5.com/geospatial/ (accessed on 9 May 2024).

77. Verrelst, J.; Rivera, J.; Alonso, L.; Moreno, J. ARTMO: An Automated Radiative Transfer Models Operator Toolbox for Automated Retrieval of Biophysical Parameters through Model Inversion. In Proceedings of the EARSeL 7th SIG-Imaging Spectroscopy Workshop, Edinburgh, UK, 11–13 April 2011; pp. 11–13.

78. ARTMO Toolbox. Available online: https://artmotoolbox.com/ (accessed on 9 May 2024).

79. Berger, K.; Atzberger, C.; Danner, M.; D'Urso, G.; Mauser, W.; Vuolo, F.; Hank, T. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. *Remote Sens.* **2018**, *10*, 85. [CrossRef]

80. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [CrossRef]

81. Tuia, D.; Volpi, M.; Copa, L.; Kanevski, M.; Munoz-Mari, J. A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 606–617. [CrossRef]

82. Candiani, G.; Tagliabue, G.; Panigada, C.; Verrelst, J.; Picchi, V.; Rivera Caicedo, J.P.; Boschetti, M. Evaluation of Hybrid Models to Estimate Chlorophyll and Nitrogen Content of Maize Crops in the Framework of the Future CHIME Mission. *Remote Sens.* **2022**, *14*, 1792. [CrossRef] [PubMed]

83. Berger, K.; Verrelst, J.; Féret, J.-B.; Hank, T.; Wocher, M.; Mauser, W.; Camps-Valls, G. Retrieval of Aboveground Crop Nitrogen Content with a Hybrid Machine Learning Method. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *92*, 102174. [CrossRef] [PubMed]

84. Brown, L.A.; Ogutu, B.O.; Dash, J. Estimating Forest Leaf Area Index and Canopy Chlorophyll Content with Sentinel-2: An Evaluation of Two Hybrid Retrieval Algorithms. *Remote Sens.* **2019**, *11*, 1752. [CrossRef]

85. Singhal, G.; Bansod, B.; Mathew, L.; Goswami, J.; Choudhury, B.U.; Raju, P.L.N. Chlorophyll Estimation Using Multi-Spectral Unmanned Aerial System Based on Machine Learning Techniques. *Remote Sens. Appl. Soc. Environ.* **2019**, *15*, 100235. [CrossRef]

86. Priyanka; Srivastava, P.K.; Rawat, R. Retrieval of Leaf Chlorophyll Content Using Drone Imagery and Fusion with Sentinel-2 Data. *Smart Agric. Technol.* **2023**, *6*, 100353. [CrossRef]

87. Zhou, X.; Zhang, J.; Chen, D.; Huang, Y.; Kong, W.; Yuan, L.; Ye, H.; Huang, W. Assessment of Leaf Chlorophyll Content Models for Winter Wheat Using Landsat-8 Multispectral Remote Sensing Data. *Remote Sens.* **2020**, *12*, 2574. [CrossRef]

88. Wang, Q.; Chen, X.; Meng, H.; Miao, H.; Jiang, S.; Chang, Q. UAV Hyperspectral Data Combined with Machine Learning for Winter Wheat Canopy SPAD Values Estimation. *Remote Sens.* **2023**, *15*, 4658. [CrossRef]

89. Chang-Hua, J.U.; Yong-Chao, T.; Xia, Y.A.O.; Wei-Xing, C.A.O.; Yan, Z.H.U.; Hannaway, D. Estimating Leaf Chlorophyll Content Using Red Edge Parameters. *Pedosphere* **2010**, *20*, 633–644.

90. Horler, D.N.H.; Dockray, M.; Barber, J.; Barringer, A.R. Red Edge Measurements for Remotely Sensing Plant Chlorophyll Content. *Adv. Space Res.* **1983**, *3*, 273–277. [CrossRef]

91. 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]

92. Alam, M.M.T.; Milas, A. *Machine Learning-Based Estimation of Canopy Chlorophyll Content in Crops from Multiple Satellite Images with Various Spatial Resolutions*; The Geological Society of America (GSA): Pittsburgh, PA, USA, 2023; Volume 55, No. 6. [CrossRef]

93. Fitzgerald, G.; Rodriguez, D.; O'Leary, G. Measuring and Predicting Canopy Nitrogen Nutrition in Wheat Using a Spectral Index—The Canopy Chlorophyll Content Index (CCCI). *Field Crops Res.* **2010**, *116*, 318–324. [CrossRef]

94. El-Shikha, D.M.; Barnes, E.M.; Clarke, T.R.; Hunsaker, D.J.; Haberland, J.A.; Pinter Jr, P.J.; Waller, P.M.; Thompson, T.L. Remote Sensing of Cotton Nitrogen Status Using the Canopy Chlorophyll Content Index (CCCI). *Trans. ASABE* **2008**, *51*, 73–82. [CrossRef]

95. Macedo, L.S.; Kawakubo, F.S. Temporal Analysis of Vegetation Indices Related to Biophysical Parameters Using Sentinel 2A Images to Estimate Maize Production. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIX*; SPIE: Bellingham, WA, USA, 2017; Volume 10421, pp. 213–220.

96. Barnes, E.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.L. Coincident Detection of Crop Water Stress, Nitrogen Status, and Canopy Density Using Ground Based Multispectral Data. In Proceedings of the Fifth International Conference on Precision Agriculture and Other Resource Management, Bloomington, MN, USA, 16–19 July 2000.

97. Zhang, H.; Li, J.; Liu, Q.; Zhao, J.; Dong, Y. A Highly Chlorophyll-Sensitive and LAI-Insensitive Index Based on the Red-Edge Band: CSI. In Proceedings of the IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 5014–5017.

98. Bi, K.; Gao, S.; Niu, Z.; Zhang, C.; Huang, N. Estimating Leaf Chlorophyll and Nitrogen Contents Using Active Hyperspectral LiDAR and Partial Least Square Regression Method. *J. Appl. Remote Sens.* **2019**, *13*, 034513. [CrossRef]

99. Peng, Z.; Guan, L.; Liao, Y.; Lian, S. Estimating Total Leaf Chlorophyll Content of Gannan Navel Orange Leaves Using Hyperspectral Data Based on Partial Least Squares Regression. *IEEE Access* **2019**, *7*, 155540–155551. [CrossRef]

100. Yu, K.; Li, F.; Gnyp, M.L.; Miao, Y.; Bareth, G.; Chen, X. Remotely Detecting Canopy Nitrogen Concentration and Uptake of Paddy Rice in the Northeast China Plain. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 102–115. [CrossRef]

101. Farifteh, J.; Van der Meer, F.; Atzberger, C.; Carranza, E.J.M. Quantitative Analysis of Salt-Affected Soil Reflectance Spectra: A Comparison of Two Adaptive Methods (PLSR and ANN). *Remote Sens. Environ.* **2007**, *110*, 59–78. [CrossRef]

102. Song, S.; Gong, W.; Zhu, B.; Huang, X. Wavelength Selection and Spectral Discrimination for Paddy Rice, with Laboratory Measurements of Hyperspectral Leaf Reflectance. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 672–682. [CrossRef]

103. Malinin, A.; Gales, M. Predictive Uncertainty Estimation via Prior Networks. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2018; Volume 31.

104. Datt, B. Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll A+ b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sens. Environ.* **1998**, *66*, 111–121. [CrossRef]

105. Curran, P.J.; Dungan, J.L.; Gholz, H.L. Exploring the Relationship between Reflectance Red Edge and Chlorophyll Content in Slash Pine. *Tree Physiol.* **1990**, *7*, 33–48. [CrossRef]

106. Sun, Q.; Jiao, Q.; Qian, X.; Liu, L.; Liu, X.; Dai, H. Improving the Retrieval of Crop Canopy Chlorophyll Content Using Vegetation Index Combinations. *Remote Sens.* **2021**, *13*, 470. [CrossRef]

107. Peng, Y.; Nguy-Robertson, A.; Arkebauer, T.; Gitelson, A.A. Assessment of Canopy Chlorophyll Content Retrieval in Maize and Soybean: Implications of Hysteresis on the Development of Generic Algorithms. *Remote Sens.* **2017**, *9*, 226. [CrossRef]