remote sensing

MDPI

Article

A Robust Algorithm for Estimating Surface Fractional Vegetation Cover from Landsat Data

Linqing Yang [1], Kun Jia [1,*], Shunlin Liang [1,2], Xiangqin Wei [3], Yunjun Yao [1] and Xiaotong Zhang [1]

1 State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; linqingyang@mail.bnu.edu.cn (L.Y.); sliang@bnu.edu.cn (S.L.); boyyunjun@163.com (Y.Y.); xtngzhang@bnu.edu.cn (X.Z.)

2 Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA

3 Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; weixq@radi.ac.cn

* Correspondence: jiakun@bnu.edu.cn; Tel.: +86-10-5880-0152

Received: 7 July 2017; Accepted: 10 August 2017; Published: 19 August 2017

Abstract: Fractional vegetation cover (FVC) is an essential land surface parameter for Earth surface process simulations and global change studies. The currently existing FVC products are mostly obtained from low or medium resolution remotely sensed data, while many applications require the fine spatial resolution FVC product. The availability of well-calibrated coverage of Landsat imagery over large areas offers an opportunity for the production of FVC at fine spatial resolution. Therefore, the objective of this study is to develop a general and reliable land surface FVC estimation algorithm for Landsat surface reflectance data under various land surface conditions. Two machine learning methods multivariate adaptive regression splines (MARS) model and back-propagation neural networks (BPNNs) were trained using samples from PROSPECT leaf optical properties model and the scattering by arbitrarily inclined leaves (SAIL) model simulations, which included Landsat reflectance and corresponding FVC values, and evaluated to choose the method which had better performance. Thereafter, the MARS model, which had better performance in the independent validation, was evaluated using ground FVC measurements from two case study areas. The direct validation of the FVC estimated using the proposed algorithm (Heihe: R^2 = 0.8825, RMSE = 0.097; Chengde using Landsat 7 ETM+: R^2 = 0.8571, RMSE = 0.078, Chengde using Landsat 8 OLI: R^2 = 0.8598, RMSE = 0.078) showed the proposed method had good performance. Spatial-temporal assessment of the estimated FVC from Landsat 7 ETM+ and Landsat 8 OLI data confirmed the robustness and consistency of the proposed method. All these results indicated that the proposed algorithm could obtain satisfactory accuracy and had the potential for the production of high-quality FVC estimates from Landsat surface reflectance data.

Keywords: fractional vegetation cover (FVC); PROSAIL; multivariate adaptive regression splines (MARS); Landsat; machine learning

1. Introduction

Fractional vegetation cover (FVC), which is defined as the percentage of the vertical projected area of green vegetation to the total ground area [1,2], is an important parameter for many environmental and climate-related modeling applications [3–5], such as dynamic global vegetation models [6], soil erosion models [7], and weather prediction models [3]. FVC also plays a vital role in the exchange of carbon, water and energy at the land surface [8]. Remote sensing techniques that have large spatial coverage and temporal continuity provide potential avenues for accurately estimating FVC.

Therefore, the estimation of FVC from remote sensing data at the regional, even global scales is of great significance.

Generally, from the methodological point of view, there are three main FVC estimation methods using remotely sensed data: empirical methods, pixel un-mixing modeling, and physical model-based methods [9–11]. The empirical methods are based on the statistical relationships between FVC and spectral band reflectance or vegetation indices from airborne or satellite spectra [8,12]. Generally, the normalized difference vegetation index (NDVI) derived from the reflectance of the red and near-infrared (NIR) bands is the most frequently used index for regression models development of FVC estimation [11]. Moreover, some other vegetation indices, such as enhanced vegetation index (EVI), visible atmospherically resistant index (VARI) and modified three-band maximal gradient difference vegetation index (MTGDVI), are also used for FVC estimation [13,14]. For example, Graetz et al. (1988) estimated the FVC of the sparse grassland in the semi-acidic soil area using the linear regression model based on the Landsat TM band 5 and the measured data of the FVC [15]. Purevdor et al. (1998) built four nonlinear models by applying the empirical model, to assess the FVC in Mongolia and Japan's grassland areas. Dymond et al. (1992) estimated the FVC of the degraded grassland in New Zealand utilizing the SPOT data based on building the nonlinear empirical relation between the surface FVC and the normalized difference vegetation index (NDVI). Empirical methods are simple to implement and thus widely used for estimating FVC at the regional scale. However, Empirical methods rely on in situ measurement data in specific regions, and the measured results are fairly accurate only if the study area is small. The accuracy will be substantially reduced in large scale application as there will be many constraints. Therefore, the empirical model was only suitable for specific vegetation types and regions, and, if expanded to a large scale, these models may be invalid.

The pixel un-mixing model assumes that each pixel is composed of several components and considers the fraction of vegetation compositions as the FVC of the pixel [11,16,17]. The dimidiate pixel model, which assumes that the pixels are only composed of vegetation and non-vegetation components and that the spectral information being just a linear synthesis of the two parts, is the simplest and most widely used pixel un-mixing model and has achieved many reliable results at the regional scales [3,11,18]. For example, GF-1 WFV data were used to estimate FVC using dimidiate pixel model in the Beijing-Tianjin-Hebei region [19]. Landsat data were used to estimate FVC with the dimidiate pixel model in the San Pedro River basin area [20]. However, it is difficult to determine the endmembers and the spectral of the endmembers over large areas for FVC estimation because the land surface is complex and the spectral characteristics of objects are varied.

The physical model-based methods for FVC estimation are based on the inversion of canopy radiative transfer models that simulate the physical relationships between vegetation canopy spectral reflectance and FVC. Such physical models have clear physical mechanism and can be adapted to a wide range of scenarios [21]. However, because of the complexity of the physical models, direct inversion is generally complex. Usually, machine learning methods, such as artificial neural networks (ANNs) and lookup table method (LUT), are used for indirect inversion of the physical models by training with a pre-computed reflectance database from the physical models [22]. Machine learning methods have the advantages of computational efficiency and robustness to noisy data and can approximate multivariate nonlinear relationships, which make them popular choices for large-area FVC estimation from remote-sensing data [1,23–25]. Several FVC products such as Global LAnd Surface Satellite (GLASS) FVC product that is one of the products extended from the GLASS product suite [26], in which multivariate adaptive regression splines (MARS) was used to estimate FVC from MODIS data; GEOV1 FVC product, in which Baret et al. used back-propagation neural networks (BPNNs); and the MERIS and CYCLOPES FVC products, which were generated using NNs and the PROSPECT+SAIL model [27], have been developed based on machine-learning methods.

However, in terms of FVC products, the current FVC products were obtained mainly from low- or medium-resolution remote sensing data such as SPOT-VGT, SEAWIFS, MERIS, MODIS and AVHRR data [1,22,28–30], which limits the FVC applications to the regional and local scales [31].

The development of FVC products from decametric spatial resolution sensors will be better for addressing these applications closely related to agriculture, ecosystem and environmental management. Long time series and the availability of well-calibrated coverage of Landsat imagery over large areas offer an opportunity for the production of FVC estimates at fine spatial resolution. Therefore, a general and effective land surface FVC estimation algorithm for Landsat surface reflectance data is of great significance for the temporal continuous obtaining of FVC estimates with fine spatial resolution on the Earth's surface.

Therefore, the objective of this study is to develop a robust FVC estimation algorithm for Landsat reflectance data under various land-surface conditions. In this study, FVC is estimated using hybrid methods through the generation of a reflectance and FVC database from the PROSAIL model simulations and powerful machine learning methods, such as BPNNs and multivariate adaptive regression splines (MARS) algorithms. The spatial and temporal consistencies of the FVC estimates from Landsat 7 ETM+ and Landsat 8 OLI in 2014 were compared to explore the potential of large-area and long-term FVC estimates as well as the consistency of estimating FVC from different Landsat sensors. Finally, the accuracy of the proposed algorithm was validated over two case study areas.

The remainder of the paper is organized as follows. Section 2 presents the datasets and pre-processing, and the method developed to estimate FVC as well. Section 3 evaluates the spatiotemporal consistency of the obtained FVC estimates and describes the experimental area and the field FVC measurements as well as the results of the validation. Finally, Sections 4 and 5 outline the discussion and conclusions, respectively.

2. Data and Methods

The proposed algorithm is shown in Figure 1. The basic ingredients of the algorithm include the Landsat data pre-processing, PROSAIL model simulations and machine learning methods. In the following subsections, these components are described in detail.

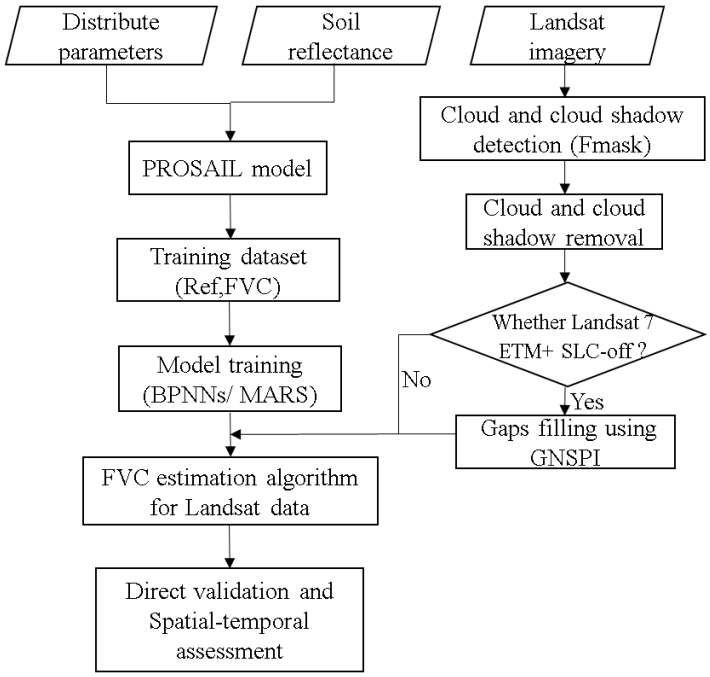

Figure 1. Flowchart of the fractional vegetation cover (FVC) estimation method for Landsat data. Notes: PROSAIL: PROSPECT leaf optical properties model and the scattering by arbitrarily inclined leaves (SAIL) model; Ref: Reflectance; BPNNs: back-propagation neural networks; MARS: multivariate adaptive regression splines; Fmask: Function of mask; ETM+: Enhanced Thematic Mapper Plus; SLC: Scan Line Corrector; GNSPI: Geostatistical Neighborhood Similar Pixel Interpolator.

2.1. Data and Pre-Processing

With data acquisition beginning in 1972 and continuing today [32], Landsat data are ideal for long-term land-surface mapping and monitoring. The Landsat surface reflectance data can be downloaded through the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA, access at: http://earthexplorer.usgs.gov/). The radiometric calibration and atmospheric correction of Landsat images were preprocessed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software. However, since 2003, all Landsat 7 ETM+ images have data gaps due to a malfunction in the scan line corrector (SLC) [33,34], resulting in about 22% image loss [35]. Despite these data gaps, Landsat 7 imagery is still widely used in land cover and land use change studies through a variety of gap-filling and dense time stack approaches [36–39]. Furthermore, many of the Landsat images are inevitably covered by clouds and cloud shadows, which limits their utility. Therefore, it is crucial to recover the cloud-contaminated pixels and fill the gaps so that temporally and spatially continuous FVC estimation can be obtained.

2.1.1. Cloud and cloud shadow detection

To remove cloud-contaminated pixels, Function of mask (Fmask) algorithm [40,41], which was originally developed for masking clouds, cloud shadows, and snow for Landsat 4–7 and had been expanded to Landsat 8, was used to detect clouds and cloud shadows by using an object-based cloud and cloud shadow matching algorithm [40]. The Fmask results had been widely used and integrated into the Landsat surface reflectance data provided by USGS [42,43] and were downloaded together with the Landsat surface reflectance data in this study.

2.1.2. Cloud and cloud shadow removal

In this study, the cloud-contaminated pixels were constructed in the proposed algorithm using modified Neighborhood Similar Pixel Interpolator (NSPI) algorithm [44]. Modified NSPI requires an auxiliary image, which is cloud free for the cloudy parts of the cloudy image. After detection of the cloud, the cloud free image was classified to search for similar pixels using unsupervised classifier, ISODATA, which can automatically merge and split classes according to the spectral similarity between pixels to get the optimal classification results. Finally, the cloud-contaminated pixels were reconstructed using the weighted spectro-spatial and spectro-temporal information of cloud-free images.

2.1.3. Gap filling for Landsat 7 ETM+ data

For the SLC-off gaps of ETM+ data, after removing cloud and cloud shadow, the Geostatistical Neighborhood Similar Pixel Interpolator (GNSPI) algorithm [45], which uses neighboring pixels with similar spectral characteristics, was used to predict the value of missing pixels using spatial and temporal information of gap-free images [46]. To be specific, GNSPI needs Landsat images acquired at other dates (labeled as the reference image), which are similar to the target image in seasonality, and have smallest spectral changes compared with the target image, as input images to fill the target SLC-off image. Sample pixels with similar spectral characteristics and similar temporal change pattern to the gap pixels are selected based on their spatial auto-covariance and other geostatistical criteria. The GNSPI algorithm uses a weighted average interpolator to predict the un-scanned pixels. The weights are calculated using the geostatistical theory (semi-variograms) and a time series as the gap-filling data. The predetermined parameters in GNSPI are defined as follows: (a) the sample size of sample pixels to 20; (b) the maximum window size to 12; (c) the estimated number of classes to 4; (d) the number of images in the time-series to 1; (e) the range of reflectance value of the image between 0 and 1; and (f) the size of block to 500.

2.2. Training Dataset Generated from PROSAIL Radiative Transfer Model

In this study, the PROSAIL model, which combined the PROSPECT leaf optical properties model and the scattering by arbitrarily inclined leaves (SAIL) canopy reflectance model [47], was used to simulate the reflectance of vegetation canopies due to its ease of use and general robustness. The SAIL model is a canopy bidirectional reflectance distribution function model that can simulate canopy reflectance from a spectral range 400 to 2500 nm [48]. The input parameters needed in the SAIL model include leaf reflectance, leaf transmittance, LAI, soil reflectance (SR), average leaf inclination angle (ALA) assuming an ellipsoidal distribution of canopy structure, solar zenith angle (SZA), viewing zenith angle (VZA), hot spot parameter (Hot), and relative azimuth angle (RAZ). The PROSPECT model is used to describe leaf optical properties at the leaf level in order to simulate directional-hemispherical reflectance and transmittance [49,50]. It is primarily based on the representation of a leaf as one or several absorbing thin plates with rough surfaces that produce isotropic scattering [51]. The inputs of the model are the leaf structure parameter (N_i, unitless), leaf chlorophyll a + b concentration (C_{ab}, $\mu g/cm^2$), equivalent water thickness (C_w, g/cm^2), dry matter content (C_m, g/cm^2), carotenoid content (C_{ar}, g/cm^2) and brown pigment content (C_{brown}).

Since there was a transform relationship between FVC and LAI, the FVC values were converted to LAI as an input variable of the PROSAIL model. In this study, the gap function $P_0(\theta)$ for a given canopy structure can be related to LAI through a theoretical model [52]:

$$P_0(\theta) = e^{-\lambda_0 \frac{G(\theta,\theta_1)}{\cos\theta} * \mathrm{LAI}} \quad (1)$$

where $P_0(\theta)$ is the gap fraction, θ is the direction where the gap fraction is computed (for nadir, $\theta = 0$), and $G(\theta, \theta_1)$ is the orthogonal projection of a unit leaf area along direction θ. It depends on the leaf inclination distribution, which is characterized by average leaf inclination angle θ_1. The parameter λ_0 is the leaf dispersion or clumping. In this study, leaf canopies are considered to be homogeneous and randomly distributed. FVC can then be calculated through $P_0(0^\circ)$:

$$FVC = 1 - P_0\left(0^\circ\right) \quad (2)$$

The main input variables of the PROSAIL model are listed in Table 1.

Table 1. Input parameters in the PROSAIL model.

	Parameters	Units	Range (or Value)	Step
PROSPECT	C_{ab}	$\mu g/cm^2$	30−60	10
	C_m	g/cm^2	0.005−0.015	0.005
	C_{ar}	g/cm^2	0	-
	C_w	cm	0.005−0.015	0.005
	C_{brown}	-	0−0.5	0.5
SAIL	N	-	1−1.5	0.5
	FVC	-	0−0.95	0.05
	ALA	°	30−70	10
	Hot	-	0.1	-
	SZA	°	25−55	10

Notes: C_{ab}: chlorophyll a + b concentration; C_m: dry matter content; C_{ar}: carotenoid content; C_w: equivalent water thickness; C_{brown}: brown pigment content; SAIL: scattering by arbitrarily inclined leaves; N: the leaf structure parameter; ALA: average leaf inclination angle; Hot: hot spot parameter; SZA: solar zenith angle.

Reflectance values of soils were also used as input variables for the PROSAIL model. In this study, the soil reflectances were selected from a globally distributed soil spectral library released by the International Soil Reference and Information Centre (ISRIC, access at: http://www.isric.org). The original soil reflectances had 785 sample locations and 4437 profiles containing various soil types

with different properties [53]. The sample locations were distributed across 58 countries, spanning Africa, Asia, Europe, North America, and South America. The original soil reflectances were first resampled from an interval of 10 nm to 1 nm. Then, the resampled soil reflectances were simulated to the corresponding Landsat spectra through the following formula:

$$\rho = \sum_{i=1}^{N} \frac{\beta(\lambda) \times \rho(\lambda)}{\sum_{i=1}^{N} \rho(\lambda)} \tag{3}$$

where ρ and $\rho(\lambda)$ are the corresponding simulated Landsat soil reflectances and the resampled soil reflectances, respectively. $\beta(\lambda)$ represents the weight of the band's spectral response function of Landsat sensors.

To remove data redundancy produced by similar soil reflectances and to avoid a large amount of calculation in the PROSAIL simulation, representative soil reflectances were determined from the original data. The simulated Landsat soil reflectances were divided into 40 categories through K-means clustering with spectral angle mapper as the similarity statistics for clustering analysis [54]. Considering two spectral vectors with n wavebands, where $X = (x_1, x_2 \ldots, x_n)$ and $Y = (y_1, y_2 \ldots, y_n)$, the spectral angle could be defined as the following:

$$\alpha_{XY} = \cos^{-1}\left[\frac{\sum_{i=1}^{N} x_i y_i}{\left[\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} y_i^2\right)^{\frac{1}{2}}\right]}\right] \tag{4}$$

where X and Y represent two different soil spectral reflectance vectors, α_{XY} is the spectral angle between the two spectral vectors X and Y, and the value range of α is between 0 and $\pi/2$. The two spectral vectors are completely similar when $\alpha = 0$ and completely different when $\alpha = \pi/2$. When the value of α is between 0 and $\pi/2$, larger α values indicate a greater difference between the two spectral vectors. In this study, if the spectral angle of soil reflectances and their belonging category's center was smaller than 0.05, it was considered as similar reflectance, and these similar reflectances were averaged as a representative soil reflectance. Finally, 40 soil reflectances were determined to represent the possible range of soil spectral shapes (Figure 2).

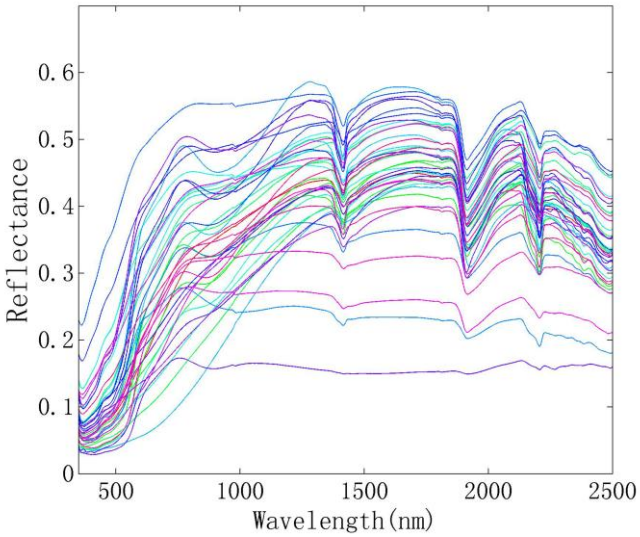

Figure 2. Reflectance of the 40 soils used to represent the possible range of spectral shapes.

For any combination of the input variables, the top of the canopy reflectance was computed for each wavelength and then resampled to simulate the Landsat observations using the spectral response

profiles. It is essential to build a dedicated model for each Landsat sensor's data because of their different spectral response profiles for each band. For one specific Landsat sensor, 288,000 cases of matched reflectances and FVC values covering various land surface conditions were obtained, of which 99% were selected randomly as the training dataset and the other 1 percent were used as a test dataset.

2.3. Machine Learning Methods

Machine learning methods essentially establish a flexible fitting model by learning the relationship between the input variables and the output variables. The hyper-parameters of the model are typically adjusted to minimize the prediction error in an independent validation dataset. Two popular machine learning methods, back-propagation neural networks (BPNNs) and multivariate adaptive regression splines (MARS), were considered because of their relatively fast training, good performance, and robustness to the overfitting problem. The related details of the two methods are presented in the following subsections.

2.3.1. The BPNNs Model

BPNNs, a popular type of neural network, have been proven as an effective algorithm for estimating land surface vegetation variables, such as LAI and FAPAR [1,27,55]. Therefore, BPNNs were selected for the comparison of the performance among machine learning methods. The BPNNs training procedure is divided into two parts: a forward propagation of information and a backward propagation of error. The back-propagation algorithm network adjusts the weights in each successive layer to reduce the errors at each level. In the linkage of the layers, the transmission of information procedure is unidirectional transmission to the input layer, the information be treated in input layer, hidden layer, and be transmitted to output layer. The status of each layer can only be affected by the next layer. If the anticipated outcome is not generated in the output layer, the algorithm switches to back-propagation, and the error between the outcome and the expected value is returned along the original path. In the present study, the BPNNs first learned from the training dataset and built relationships between reflectance and FVC, then the trained BPNNs could produce optimal FVC estimates based on the actual reflectance of the remotely sensed data. The inputs of the BPNNs included the reflectance of the red and NIR bands, and the output was the corresponding FVC. The number of nodes in the hidden layer was set to four. The BPNNs activation function in the hidden layer was set to "tansig", the transfer function for the output layer was set to "purelin", and the training function was set to "trainlm". Because of its efficient convergence capacity, the Levenberg–Marquardt minimization algorithm was used to calibrate the synaptic coefficients [36].

2.3.2. The MARS Model

MARS is a nonparametric and multivariate regression analysis model and has been demonstrated to obtain satisfactory FVC results from MODIS reflectance data [29,56]. Without strong assumptions, MARS is capable of modeling complex nonlinear relationships among variables by fitting piecewise linear regressions. Furthermore, MARS can search for the relative importance of independent variables to the dependent variable when many potential independent variables are involved. MARS builds models of the form [56]:

$$f(x) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} [S_{k,m}(x(k,m) - t_{k,m})] \quad (5)$$

where a_0 is the coefficient of the constant basis function; a_m are the coefficients of the model, which are estimated to yield the best fit to the training data; M is the number of basis functions; K_m is the number of splits that generate the mth basis function; $s_{k,m}$ takes values of either 1 or -1 and indicates the right/left sense of the associated step function; $x(k,m)$ is the label of the independent variable; and $t_{k,m}$ indicates the knot locations. In this study, the inputs of the MARS included the reflectance of the red and NIR bands, and the output was the corresponding FVC.

The optimal MARS model is determined by a two-stage process: first, MARS constructs a very large number of basis functions to fit the data, where variables can interact with each other to fit the data. Then, the basis functions are deleted in the order of least importance using the generalized cross-validation (GCV) criterion [57]. The importance of a variable can be assessed by observing the decrease in the calculated GCV when it is removed from the model. MARS is capable of reliably tracking the extremely complex data structures hidden in high-dimensional data. Additional details regarding the MARS model building process can be found in the work of Friedman et al. [57].

2.4. Direct Validation and Spatial-Temporal Assessment

With the goal of assessing the accuracy of the FVC estimates, RMSE and R^2 between the FVC estimates and the FVC values calculated from the near surface digital images over the Heihe and Chengde test areas were computed. Due to the time inconsistency between the FVC estimates and field measured FVC, the estimated FVC were interpolated through the linear interpolation method to the acquisition dates of the corresponding field measured FVC.

2.4.1. Accuracy Assessment in Heihe Region

The first test area is located in the Yingke oasis region (Figure 3) of the Heihe River Basin (HRB), southwest of Zhangye City, in the Gansu province of China. The HRB is a typical inland river basin (endorheic basin), which is located in the northwest of China, belonging to the arid region. The annual precipitation of the area is approximately 140 mm, and the annual mean temperature is approximately 7 °C–10 °C. The landscape of the test area is relatively flat and homogeneous, and most of the area is agricultural land in which corn is the dominant vegetation type, as well as some small patches of pepper, melon, and red bean.

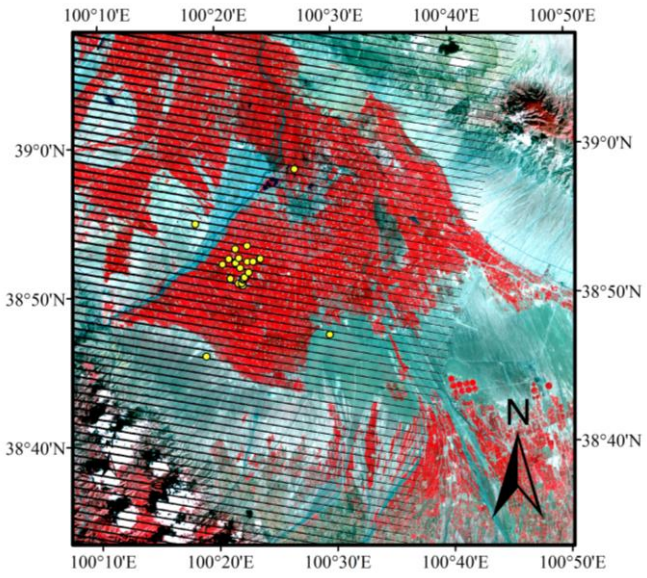

Figure 3. The geographic location of the Heihe test area. The yellow points represent the location of sample plots.

In total, 16 homogeneous sampling sites were used, and the size of each sampling site was 10 m × 10 m. The field survey was performed via digital photography using a Nikon D3000 camera from 24 May to 28 August 2012, which covered the whole growing season of corn. The observations were conducted every five days before 23 July and every ten days after 23 July. A long stick equipped with the camera at the end was used to control shooting height. The footprint of the sensor at the nadir direction required a coverage that ranged 1.5–2 times the total width of the row canopy. The photographic approach guarantees that the field of view comprises two or more rows in one

photograph. In each sample site, nine photographs were taken, along with the two diagonals of the squared sampling site, and the FVC estimates from the photographs of the nine locations were averaged to estimate the FVC for the sampling site [58].

The FVC of a digital photograph was defined as the percentage of the number of vegetation pixels to the total number of pixels in the digital photograph. First, the image edges were cut off before FVC extraction [59] to remove distortion and perspective effects. Then, the thresholding method was used to extract FVC from the digital images [60]. This automatic classification method transfers images from the RGB color space to the Commission Internationale d'Eclairage LAB color space, which can easily distinguish green vegetation from the background. Green vegetation and non-vegetation pixels were separated through histogram clustering, which supposes that green vegetation and the background distribution of greenness in the color space are Gaussian. Finally, reasonable FVC was extracted from the digital photographs and used to validate the performance of the proposed method.

2.4.2. Accuracy Assessment in Chengde Region

The second study area is located in Chengde, Hebei Province, China (Figure 4), which belongs to the semi-humid and semi-dry climate region of the temperate zone. The annual average precipitation is about 500 mm, and the annual average temperature ranges from −1.4 °C to 4.7 °C [61]. The test area is covered with various vegetation types, such as broad-leaf and coniferous forest, grassland, wetland, and cropland. The abundant land cover types make the test area suitable to validate the effectiveness of the FVC estimation algorithm for Landsat data.

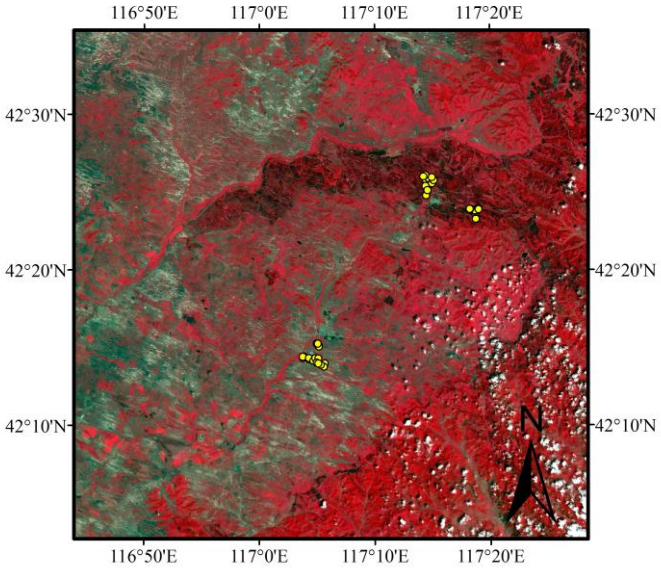

Figure 4. The geographic location of the Chengde test area and locations of the ground measurements are displayed as yellow dots.

To synchronize remote sensing observations with ground observations, the field measurements were conducted via digital photography using a Nikon D90 camera from 24 to 27 July 2014. The designed experimental plot was homogeneous and the designed experimental plot size was 30 m × 30 m. Five survey points were measured with one point at the center of the site and the other four points located at the corners of the square. The vegetation types of the sampling sites contained wheat, wetland, mixed forest, corn, pine, grass and shrub, potato, and grass. Within each sample site, usually five survey points were measured with one point at the center of the site and the other four points located at the diagonal of the square. For low vegetation types, the digital images were acquired from the nadir at approximately two meters above the ground at each survey point. The field FVC in

the forest regions was determined using the following equation that accounts for the FVC of trees and understory vegetation viewed at the nadir direction between tree gaps [58]:

$$FVC = f_{up} + (1 - f_{up}) \times f_{down} \quad (6)$$

The original digital images were saved in JPEG format with a size of 4288 pixels × 2848 pixels. To remove distortion from center projection of the digital camera, the edges, about 40% of the length and width of the image, were cut off. Then, FVC values of the cropped images were extracted using the improved Gaussian simulation and segmentation method in CIE L*a*b* color space [60]. Finally, the FVC calculated from digital image were used to directly evaluate the performance of the proposed algorithm.

Since the field survey over Chengde test area was conducted in the year 2014, there are two sensors (Landsat 7 ETM+ and Landsat 8 OLI) providing reflectance data. Therefore, the validation over Chengde area was conducted using FVC estimates generated from data of the two sensors.

2.4.3. Spatial-Temporal Analysis

In order to explore the potential of FVC retrieval with Landsat data over large areas and to assess spatial and temporal consistencies between the FVC values derived from different Landsat sensors. Multitemporal FVC maps during the 2014 growth season derived from the proposed algorithm over the oasis of the middle Heihe River, which is located in the northwest China, were used to investigate the spatial consistency and spatial continuity of the proposed algorithm. To observe the consistency of the FVC estimates from the two sensors, several representative FVC temporal profiles such as wheat, corn, wetland, Gobi and desert estimated from Landsat 7 ETM+ and Landsat 8 OLI data were extracted and compared.

3. Results

This section is designed to show the theoretical performance and experimental evidence of the performance of the proposed FVC estimation algorithm for Landsat data. The spatial-temporal FVC estimates from different sensors were compared to explore the potential of large area and long-term FVC data generation as well as the consistency of FVC estimates from different Landsat sensors.

3.1. Accuracy Assessment Over the Simulated Dataset

It is worth noting that different sensors have different spectral response functions. This leads to slight differences in surface reflectance for each sensor. The performances of the MARS and BPNNs were evaluated over the test dataset (2880 samples), corresponding to Landsat 7 ETM+ and Landsat 8 OLI sensors. For brevity, only the results of the Landsat 7 ETM+ are shown. Results revealed good accuracy and low bias in Landsat 7 ETM+ simulated reflectances. MARS (R^2 = 0.9414, RMSE = 0.07) was slightly superior to BPNNs (R^2 = 0.9388, RMSE = 0.072) in all statistical quality measures (Figure 5). It could be seen that the scatters was discontinuous, that was because the step of FVC was set as 0.05 when the training dataset was generated using PROSAIL model.

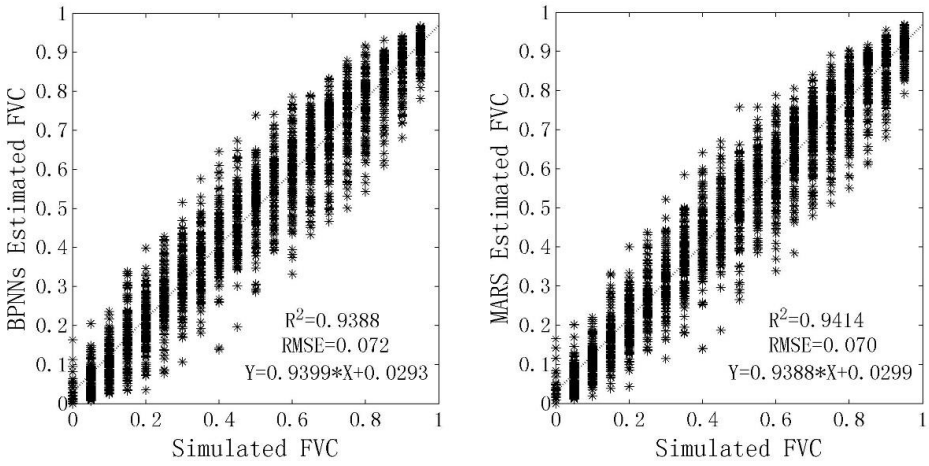

Figure 5. Theoretical performances of: back-propagation neural networks (BPNNs) (**left**); and multivariate adaptive regression splines (MARS) (**right**) using Landsat 7 ETM+ simulated data.

To analyze models' robustness to local consistency and reliability of the estimated error [62,63], Figure 6 exhibits the evolution of the RMSE as a function of the number of predictions. The shapes of the two curves in Figure 6 are similar for two methods, which shows a consistent better performance of MARS over BPNNs. When the number of predictions reaches 2880, the scatter plots of FVC estimates between simulated FVC is just as presented in Figure 5. Furthermore, the accuracy of MARS consistently shows a slightly better result than that of BPNNs. These results indicated that MARS was the more robust and accurate machine learning method in FVC estimation for Landsat data. Therefore, MARS was suitable to generate FVC estimates from Landsat reflectance data.

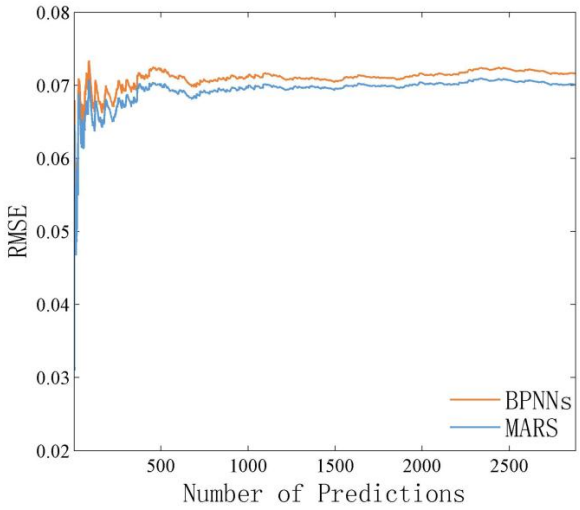

Figure 6. Performance in the test set averaged over the Landsat 7 ETM+ training-test data splitting as a function of the number of used predictions.

3.2. Accuracy Assessment Using Field Survey FVC

3.2.1. Accuracy Assessment in Heihe Region

The scatter plot of the FVC estimates from the proposed algorithm and the FVC calculated from digital image in Heihe region are shown in Figure 7. Most of the scatter points lie around the 1:1 line, and the overall performance (R^2 = 0.8825, RMSE = 0.097) presents a satisfactory FVC estimation result, which confirms the reasonability and reliability of the proposed algorithm. Furthermore, the validation

results of the temporal representative field survey measurements over the Heihe region indicate the effectiveness of the proposed algorithm further.

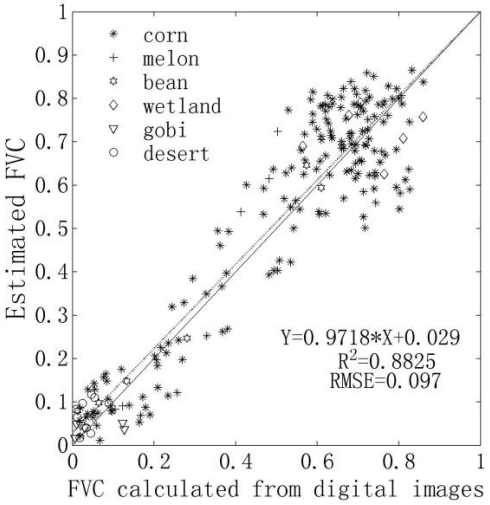

Figure 7. Comparison of the FVC estimates from Landsat 7 ETM+ reflectance data using the proposed algorithm and those extracted from the field photos in Heihe region.

3.2.2. Accuracy Assessment in Chengde Region

A comparison of the estimated FVC from Landsat 7 ETM+ and Landsat 8 OLI reflectance data using the proposed algorithm and the FVC extracted from the field photos in Chengde area was conducted (Figure 8). Both of the scatter plots show good performance of FVC estimation, with RMSE of 0.078 and R^2 of 0.8571 for ETM+ data, and RMSE of 0.078 and R^2 of 0.8598 for OLI data, which further confirm the reliability and effectiveness of the proposed algorithm. Meanwhile, in terms of each vegetation type, the estimated FVC values have high consistency with the FVC values calculated from digital images. Thus, the proposed algorithm is reliable for various vegetation types. Although no prior knowledge was used, the proposed algorithm achieved satisfactory accuracy under various vegetation types.

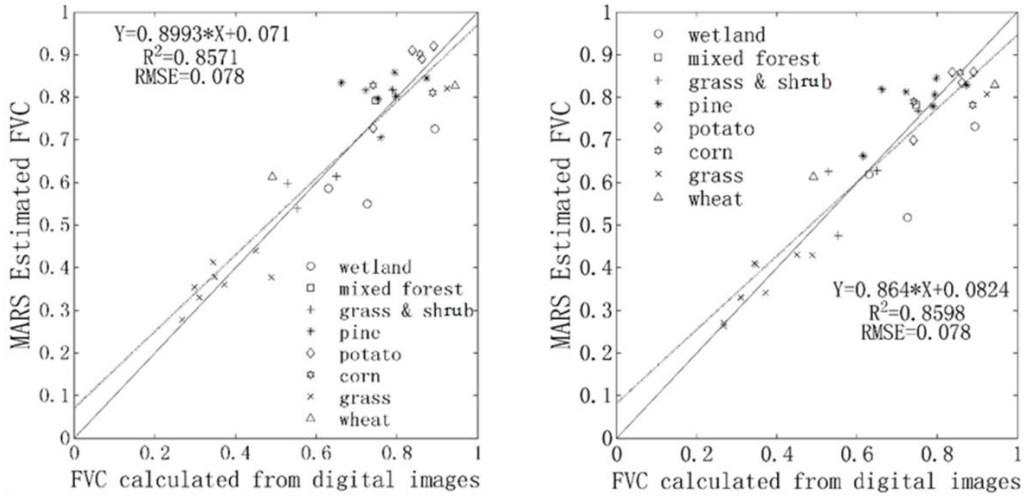

Figure 8. Scatter plots of the estimated FVC from: Landsat 7 ETM+ (**left**); and Landsat 8 OLI (**right**) reflectance data using the proposed algorithm and FVC extracted from the field photos in Chengde area.

3.3. Spatial-Temporal Analysis

3.3.1. Spatial Analysis

The MARS-Landsat 7 ETM+ and MARS-Landsat 8 OLI FVC maps for the very beginning of the season (late-April), the fast growing season (late-June), and the maximum FVC (late-August) in the Heihe area (Path: 133; Row: 33) were generated and the closest acquisition dates between Landsat 7 ETM+ and Landsat 8 OLI were taken into account for the comparison.

The standard false color composited images with cloud and cloud shadow contaminations, as well as gaps for Landsat 7 ETM+ data, are shown in Figure 9a. Then, the standard false color composited images after processing by the proposed procedure of the method are shown in Figure 9b. Finally, the corresponding FVC maps generated using the proposed algorithm are shown in Figure 9c. In Figure 9a,b, it can be seen that clouds and cloud shadows are removed well. Meanwhile, it is clear that the filled images have no strips and are similar to the closest date's corresponding Landsat 8 OLI images. These results indicate the effectiveness of the data processing procedures for cloud removal and gap filling in the proposed FVC estimation algorithm.

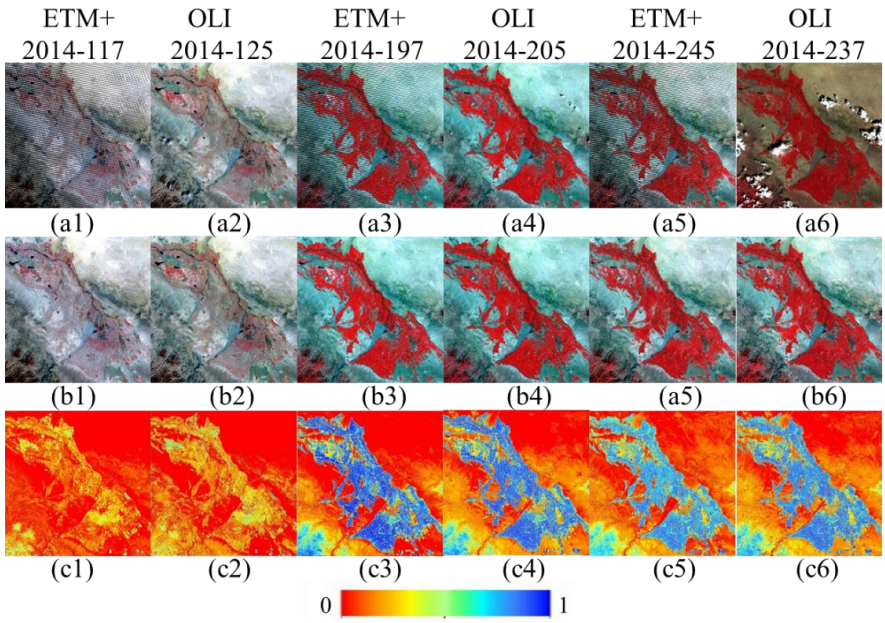

Figure 9. The FVC maps derived from Landsat 7 ETM+ and Landsat 8 OLI reflectance data using the proposed algorithm (**c1**–**c6**). (**a1**–**a6**) The original standard false color composited Landsat images; and (**b1**–**b6**) the corresponding standard false color composited Landsat images with cloud and cloud shadow removed as well as gaps filled.

First inspection of the maps shows the occurrence of very low FVC values corresponding to late-April (Landsat 7 ETM+, DOY: 117; Landsat 8 OLI, DOY: 125), because the sowing dates are around mid or late April in northwest China. The late-July (Landsat 7 ETM+, DOY: 197; Landsat 8 OLI, DOY: 205) maps show the highest FVC estimates because corn, which is the dominant vegetation type, reaches the growth peak season. The late-August (Landsat 7 ETM+, DOY: 245; Landsat 8 OLI, DOY: 237) FVC maps show that FVC values begin to decrease. Generally, the distributions of the FVC values are consistent with the actual conditions of land cover distributions and seasonal vegetation variations. The farmland regions present high FVC values that are consistent with the actual seasonal variations as well, whereas the desert regions have FVC values close to zero for the whole year. Furthermore, the FVC maps generated using the proposed algorithm achieved good performance of spatial continuity, which further confirm that preprocessing Landsat data could obtain continuous reflectance values and thus generate spatially continuous FVC data.

In order to verify the performance of cloud removal and gap filling, the relationship between FVC values estimated from Landsat 7 ETM+ which include the actual pixels and reconstructed pixels, respectively, and FVC values estimated from Landsat 8 OLI were compared. Figure 10 shows the location of the pixels as well as the FVC estimation results (Landsat 7 ETM+, DOY: 197; Landsat 8 OLI, DOY: 205). To better assess the similarity of results of actual image and filled image, Figure 11 shows the scatter plot of FVC results estimated from actual Landsat 7 ETM+ image and FVC values estimated from actual Landsat 8 OLI image, and the scatter plot of FVC results estimated from filled Landsat 7 ETM+ image and FVC values estimated from actual Landsat 8 OLI image. In Figure 11, it can be seen that data points in the scatter plots fall close to the 1:1 line. The R^2 values are 0.9832 and 0.9856, respectively, which indicate the FVC results have good consistency and the filled data are close to the actual data. Generally, the method used in this study to remove cloud and fill gaps is reliable and can generate consistent FVC estimation results compared with actual images.

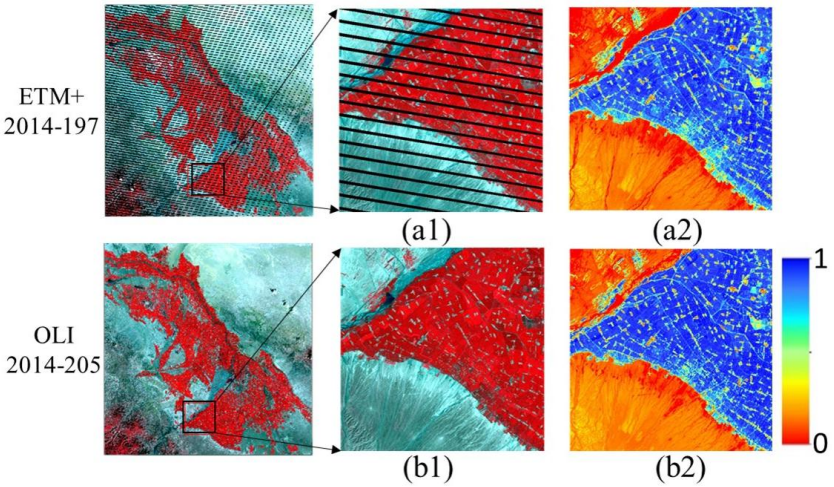

Figure 10. The FVC maps (**a2**) and (**b2**) derived from Landsat 7 ETM+ and Landsat 8 OLI reflectance data using the proposed algorithm. (**a1**) and (**b1**) The corresponding standard false color composited Landsat images with cloud and cloud shadow removed as well as gaps filled.

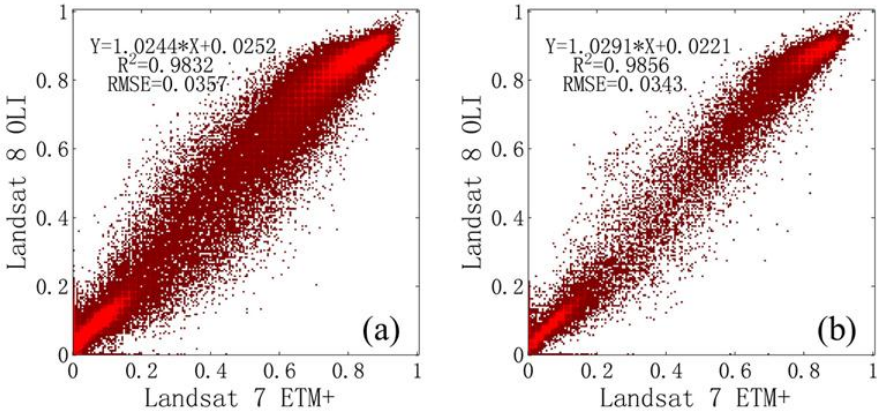

Figure 11. Comparison of relationship between FVC values estimated from Landsat 7 ETM+ and FVC values estimated from Landsat 8 OLI. The density map presents the densities of the points. The FVC values estimated from Landsat 7 ETM+: (**a**) from actual scanned pixels; and (**b**) estimated from predicted values of un-scanned pixels.

3.3.2. Time Series Analysis

Several representative FVC temporal profiles estimated from Landsat 7 ETM+ and Landsat 8 OLI data for wheat, corn, wetland, Gobi and desert are shown in Figure 12. In general, FVC estimates derived from MARS-Landsat 7 ETM+ and MARS-Landsat 8 OLI data are consistent with the seasonal phenological cycle in the oasis of the Heihe River Basin. It is interesting to note that the differences between Landsat 7 ETM+ and Landsat 8 OLI data do not induce large differences and show good temporal consistency in FVC estimates, which indicates the potential of the consistency of FVC estimates from different Landsat sensors' data. Meanwhile, the two datasets also have similar magnitudes, dynamic ranges and seasonal variations, which could reflect actual vegetation growth characteristics. In summary, the FVC results from MARS-Landsat 7 ETM+ and MARS-Landsat 8 OLI data indicate that the proposed algorithm is capable of producing FVC with good temporal consistency that also can reflect the seasonal phenological cycle.

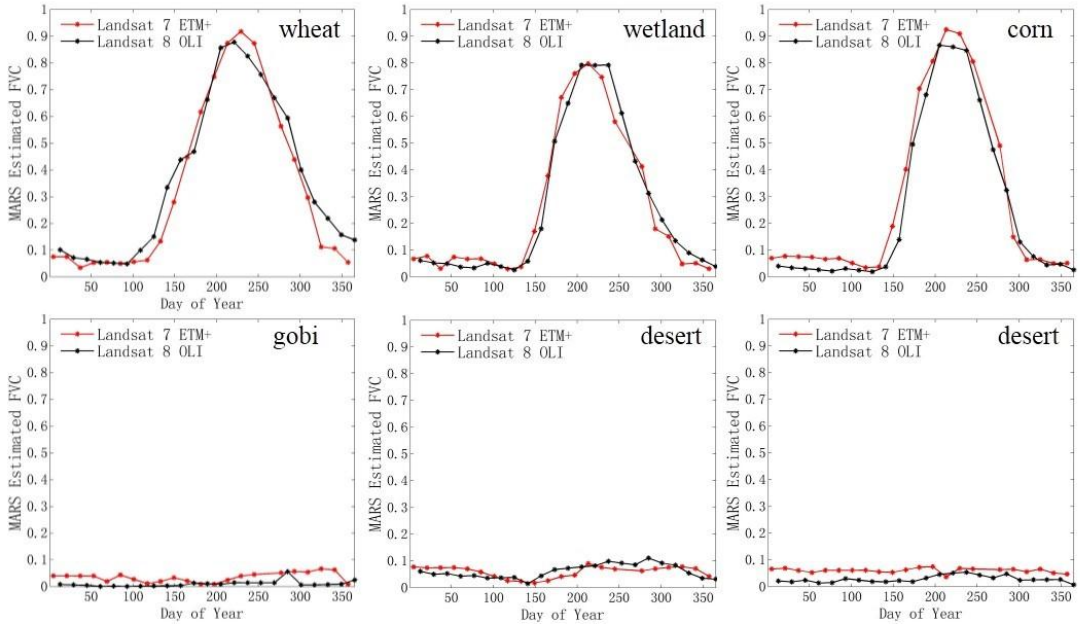

Figure 12. Temporal dynamics of the Landsat 7 ETM+ and Landsat 8 OLI FVC estimates within six representative vegetation types in the Heihe River area.

4. Discussion

In this study, a general and effective FVC estimation algorithm based on training MARS with training samples generated from radiative transfer model simulations for Landsat surface reflectance data was proposed. MARS are powerful for the estimation of FVC estimation because of its ability to perform adaptive, nonlinear data fitting. Furthermore, the radiative transfer model, which can simulate a broad range of land cover situations, can simulate training dataset more globally applicable. The algorithm achieved satisfactory FVC estimation accuracy based on the field survey values and had the ability to produce consistent FVC estimates from different Landsat sensors. Therefore, the proposed algorithm has the potential for generation high-quality FVC at a fine resolution over large areas.

In the prior studies, FVC estimation using Landsat data mainly employed on empirical methods and pixel unmixing models which was limited to a small test area. The parameters in the empirical methods and pixel unmixing models are difficult to determine and might fail when expanded to regional scales. The proposed algorithm does not need any prior knowledge, such as vegetation types or empirically obtained parameters, which overcomes the problems in empirical methods and pixel unmixing models. Furthermore, the currently existing FVC products are mostly obtained from low- or medium-resolution remotely sensed data such as SPOT-VET, MERIS, MODIS and AVHRR

data. This study testified that the proposed algorithm has the potential for FVC estimation at a fine spatial resolution, which would be better for applications related to agriculture, ecosystem and environment management.

In this study, the performance of the proposed algorithm was evaluated over two test areas. Most of the Heihe test area could be regarded as homogeneous landscapes and the landscapes of Chengde test area belong to heterogeneous landscapes containing various vegetation types. The performance of the proposed algorithm achieved satisfactory accuracy over temporal representative field survey data and FVC observations of various vegetation types, which indicated the effectiveness of the algorithm over both homogeneous and heterogeneous landscapes. Furthermore, the algorithm has the potential to estimate FVC in larger area. More validation should be conducted to confirm the general applicability over larger regions, even globally. In the future, more field experiments for testing the proposed algorithm are needed to further confirm the accuracy of the proposed algorithm. Given it is difficult to perform actual FVC field surveys, taking digital photos from the vertical direction may be the preferred choice for FVC extraction in field surveys.

ESA's satellites Sentinel-2 (S2) and Sentinel-3 (S3) aim to ensure continuity for Landsat 5/7, SPOT-5, SPOT-Vegetation and Envisat MERIS observations by providing Earth images of high spatial, spectral and temporal resolution. S2 and S3 will deliver near real-time operational products with a high accuracy land surface parameter for land monitoring, which leads to an urgent need for developing robust and accurate retrieval methods. MARS is powerful to estimate FVC and other biophysical parameters because of its ability to perform adaptive, nonlinear data fitting. Furthermore, radiative transfer model can simulate various land cover situations and then generating representative training data. Therefore, the proposed algorithm also has the potential to be applied to develop FVC estimation using S2 and S3 data.

5. Conclusions

The availability of free Landsat data makes the quantitative mapping of the spatial and temporal FVC possible. In this study, a general and reliable FVC estimation method for Landsat surface reflectance data was proposed, which was based on the radiative transfer model and machine learning method. Two machine learning methods, MARS and BPNNs, were trained using samples from PROSAIL radiative transfer model simulations, which included Landsat reflectance and corresponding FVC values, and evaluated to choose the method which had better performance. Then, the MARS model, which had better performance in the independent validation, was evaluated using spatial-temporal assessment and ground FVC measurements from two case study areas. The direct validation of the FVC estimates using the proposed algorithm had good performance. Meanwhile, spatial-temporal assessment of the estimated FVC from Landsat 7 ETM+ and Landsat 8 OLI data confirmed the robustness and consistency of the proposed method. All these results indicated that the proposed algorithm could achieve satisfactory FVC estimation accuracy and had the potential for the production of high-quality FVC estimates from Landsat surface reflectance data. Furthermore, the algorithm had the potential to estimate FVC in larger area. Therefore, future work will focus on an extensive assessment of the performance of the proposed algorithm using additional ground measurement data.

Acknowledgments: This study was partially supported by the National Natural Science Foundation of China (Nos. 41671332 and 41331173) and the National Key Research and Development Program of China (No. 2016YFA0600103). The authors would also like to thank Mu, X. from Beijing Normal University for providing the ground reference data in Heihe region.

Author Contributions: K.J., L.Y. and S.L. conceived and designed the experiments; L.Y. performed the experiments; K.J. and L.Y. conducted the analysis of the results; and all the authors contributed towards writing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Baret, F.; Weiss, M.; Lacaze, R.; Camacho, F.; Makhmara, H.; Pacholcyzk, P.; Smets, B. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. *Remote Sens. Environ.* **2013**, *137*, 299–309. [CrossRef]
2. Zhang, X.; Liao, C.; Li, J.; Sun, Q. Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 506–512. [CrossRef]
3. Gutman, G.; Ignatov, A. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.* **1998**, *19*, 1533–1543. [CrossRef]
4. Matsui, T.; Lakshmi, V.; Small, E.E. The Effects of satellite-derived vegetation cover variability on simulated land–atmosphere interactions in the NAMS. *J. Clim.* **2005**, *18*, 21–40. [CrossRef]
5. Zhang, X.; Wu, B.; Ling, F.; Zeng, Y.; Yan, N.; Yuan, C. Identification of priority areas for controlling soil erosion. *Catena* **2010**, *83*, 76–86. [CrossRef]
6. Sitch, S.; Smith, B.; Prentice, I.C.; Arneth, A.; Bondeau, A.; Cramer, W.; Kaplan, J.O.; Levis, S.; Lucht, W.; Sykes, M.T.; et al. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Glob. Chang. Biol.* **2003**, *9*, 161–185. [CrossRef]
7. Morgan, R.P.C.; Quinton, J.N.; Smith, R.E.; Govers, G.; Poesen, J.W.A.; Auerswald, K.; Chisci, G.; Torri, D.; Styczen, M.E. The European Soil Erosion Model (EUROSEM): A dynamic approach for predicting sediment transport from fields and small catchments. *Earth Surf. Process. Landf.* **1998**, *23*, 527–544. [CrossRef]
8. Xiao, J.; Moody, A. A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA. *Remote Sens. Environ.* **2005**, *98*, 237–250. [CrossRef]
9. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379. [CrossRef]
10. Guerschman, J.P.; Hill, M.J.; Renzullo, L.J.; Barrett, D.J.; Marks, A.S.; Botha, E.J. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sens. Environ.* **2009**, *113*, 928–945. [CrossRef]
11. Jiapaer, G.; Chen, X.; Bao, A. A comparison of methods for estimating fractional vegetation cover in arid regions. *Agric. For. Meteorol.* **2011**, *151*, 1698–1710. [CrossRef]
12. Carlson, T.N.; Ripley, D.A. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* **1997**, *62*, 241–252. [CrossRef]
13. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [CrossRef]
14. Jiang, Z.; Huete, A.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [CrossRef]
15. Graetz, R.D.; Pech, R.P.; Davis, A.W. The assessment and monitoring of sparsely vegetated rangelands using calibrated Landsat data. *Int. J. Remote Sens.* **1988**, *9*, 1201–1222. [CrossRef]
16. Jimenez-Munoz, J.C.; Sobrino, J.A.; Plaza, A.; Guanter, L.; Moreno, J.; Martinez, P. Comparison between fractional vegetation cover retrievals from vegetation indices and spectral mixture analysis: Case study of PROBA/CHRIS data over an agricultural area. *Sensors* **2009**, *9*, 768–793. [CrossRef] [PubMed]
17. Johnson, B.; Tateishi, R.; Kobayashi, T. Remote sensing of fractional green vegetation cover using spatially-interpolated endmembers. *Remote Sens.* **2012**, *4*, 2619–2634. [CrossRef]
18. Wu, B.; Li, M.; Yan, C.; Zhou, W.; Yan, C. Developing method of vegetation fraction estimation by remote sensing for soil loss equation: A case in the upper basin of miyun reservoir. In Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; Volume 6, pp. 4352–4355.
19. Jackson, T.J.; Chen, J.M.; Gong, P.; Liang, S.; Zhan, Y.; Meng, Q.; Wang, C.; Li, J.; Zhou, K.; Li, D. Fractional vegetation cover estimation over large regions using GF-1 satellite data. In Proceedings of the Land Surface Remote Sensing II, Beijing, China, 13 October 2014; Volume 9260, p. 92604B.

20. Qi, J.; Marsett, R.C.; Moran, M.S.; Goodrich, D.C.; Heilman, P.; Kerr, Y.H.; Dedieu, G.; Chehbouni, A.; Zhang, X.X. Spatial and temporal dynamics of vegetation in the San Pedro River basin area. *Agric. For. Meteorol.* **2000**, *105*, 55–68. [CrossRef]
21. 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]
22. Roujean, J.-L. Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation. *J. Geophys. Res.* **2002**, *107*. [CrossRef]
23. Ahmad, S.; Kalra, A.; Stephen, H. Estimating soil moisture using remote sensing data: A machine learning approach. *Adv. Water Resour.* **2010**, *33*, 69–80. [CrossRef]
24. Verger, A.; Baret, F.; Camacho, F. Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations. *Remote Sens. Environ.* **2011**, *115*, 415–426. [CrossRef]
25. Jia, K.; Liang, S.; Gu, X.; Baret, F.; Wei, X.; Wang, X.; Yao, Y.; Yang, L.; Li, Y. Fractional vegetation cover estimation algorithm for Chinese GF-1 wide field view data. *Remote Sens. Environ.* **2016**, *177*, 184–191. [CrossRef]
26. Liang, S.; Zhao, X.; Liu, S.; Yuan, W.; Cheng, X.; Xiao, Z.; Zhang, X.; Liu, Q.; Cheng, J.; Tang, H.; et al. A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. *Int. J. Digit. Earth* **2013**, *6*, 5–33. [CrossRef]
27. Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Niño, F.; Weiss, M.; Samain, O.; et al. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. *Remote Sens. Environ.* **2007**, *110*, 275–286. [CrossRef]
28. Jia, K.; Liang, S.; Liu, S.; Li, Y.; Xiao, Z.; Yao, Y.; Jiang, B.; Zhao, X.; Wang, X.; Xu, S.; et al. Global land surface fractional vegetation cover estimation using general regression neural networks from MODIS surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4787–4796. [CrossRef]
29. Yang, L.; Jia, K.; Liang, S.; Liu, J.; Wang, X. Comparison of four machine learning methods for generating the GLASS fractional vegetation cover product from MODIS Data. *Remote Sens.* **2016**, *8*, 682. [CrossRef]
30. García-Haro, F.J.; Camacho-de Coca, F.; Miralles, J.M. Inter-comparison of SEVIRI/MSG and MERIS/ENVISAT biophysical products over Europe and Africa. In Proceedings of the 2nd MERIS/(A)ATSR User Workshop, Frascati, Italy, 22–26 September 2008; p. 8.
31. Li, W.; Weiss, M.; Waldner, F.; Defourny, P.; Demarez, V.; Morin, D.; Hagolle, O.; Baret, F. A Generic algorithm to estimate LAI, FAPAR and FCOVER variables from SPOT4_HRVIR and Landsat sensors: Evaluation of the consistency and comparison with ground measurements. *Remote Sens.* **2015**, *7*, 15494–15516. [CrossRef]
32. Williams, D.L.; Goward, S.; Arvidson, T. Landsat. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1171–1178. [CrossRef]
33. Arvidson, T.; Goward, S.; Gasch, J.; Williams, D. Landsat-7 long-term acquisition plan. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1137–1146. [CrossRef]
34. Ju, J.; Roy, D.P. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sens. Environ.* **2008**, *112*, 1196–1211. [CrossRef]
35. Kontgis, C.; Schneider, A.; Ozdogan, M. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. *Remote Sens. Environ.* **2015**, *169*, 255–269. [CrossRef]
36. Castrence, M.; Nong, D.; Tran, C.; Young, L.; Fox, J. Mapping urban transitions using multi-temporal Landsat and DMSP-OLS night-time lights imagery of the Red River Delta in Vietnam. *Land* **2014**, *3*, 148–166. [CrossRef]
37. Hilker, T.; Wulder, M.A.; Coops, N.C.; Seitz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999. [CrossRef]
38. Huang, C.; Goward, S.N.; Masek, J.G.; Thomas, N.; Zhu, Z.; Vogelmann, J.E. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens. Environ.* **2010**, *114*, 183–198. [CrossRef]
39. Schneider, A. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sens. Environ.* **2012**, *124*, 689–704. [CrossRef]

40. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]
41. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]
42. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.K. A Landsat surface reflectance dataset for north America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [CrossRef]
43. Maiersperger, T.K.; Scaramuzza, P.L.; Leigh, L.; Shrestha, S.; Gallo, K.P.; Jenkerson, C.B.; Dwyer, J.L. Characterizing LEDAPS surface reflectance products by comparisons with AERONET, field spectrometer, and MODIS data. *Remote Sens. Environ.* **2013**, *136*, 1–13. [CrossRef]
44. Zhu, X.; Gao, F.; Liu, D.; Chen, J. A Modified neighborhood similar pixel interpolator approach for removing thick clouds in Landsat images. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 521–525. [CrossRef]
45. Chen, J.; Zhu, X.; Vogelmann, J.E.; Gao, F.; Jin, S. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* **2011**, *115*, 1053–1064. [CrossRef]
46. Campos-Taberner, M.; Garcia-Haro, F.J.; Moreno, A.; Gilabert, M.A.; Martinez, B.; Sanchez-Ruiz, S.; Camps-Valls, G. Development of an earth observation processing chain for crop bio-physical parameters at local scale. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium, Milan, Italy, 26–31 July 2015; pp. 29–32.
47. 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]
48. Verhoef, W. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sens. Environ.* **1984**, *16*, 125–141. [CrossRef]
49. Jacquemoud, S.; Baret, F. PROSPECT: A model of leaf optical properties spectra. *Remote Sens. Environ.* **1990**, *34*, 75–91. [CrossRef]
50. Jacquemoud, S. Comparison of Four Radiative Transfer models to simulate plant canopies reflectance direct and inverse mode. *Remote Sens. Environ.* **2000**, *74*, 471–481. [CrossRef]
51. Allen, W.A.; Gausman, H.W.; Richardson, A.J.; Thomas, J.R. Interaction of isotropic light with a compact plant leaf. *J. Opt. Soc. Am.* **1969**, *59*, 1376–1379. [CrossRef]
52. Nilson, T. A theoretical analysis of the frequency of gaps in plant stands. *Agric. Meteorol.* **1971**, *8*, 25–38. [CrossRef]
53. Shepherd, K.D.; Palm, C.A.; Gachengo, C.N.; Vanlauwe, B. Rapid Characterization of organic resource quality for soil and livestock management in tropical agroecosystems using near-infrared spectroscopy. *Agron. J.* **2003**, *95*. [CrossRef]
54. Dennison, P.E.; Halligan, K.Q.; Roberts, D.A. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sens. Environ.* **2004**, *93*, 359–367. [CrossRef]
55. Jia, K.; Liang, S.; Wei, X.; Li, Q.; Du, X.; Jiang, B.; Yao, Y.; Zhao, X.; Li, Y. Fractional forest cover changes in northeast China from 1982 to 2011 and its relationship with climatic variations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 775–783. [CrossRef]
56. Barron, A.R.; Xiao, X. Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]
57. Friedman, J.H. Multivariate adaptive regression splines (with discussion). *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]
58. Mu, X.; Huang, S.; Ren, H.; Yan, G.; Song, W.; Ruan, G. Validating GEOV1 fractional vegetation cover derived from coarse-resolution remote sensing images over croplands. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 439–446. [CrossRef]
59. Zhao, J.; Xie, D.; Mu, X.; Liu, Y.; Yan, G. Accuracy evaluation of the ground-based fractional vegetation cover measurement by using simulated images. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 3347–3350.
60. Liu, Y.; Mu, X.; Wang, H.; Yan, G.; Henebry, G. A novel method for extracting green fractional vegetation cover from digital images. *J. Veg. Sci.* **2012**, *23*, 406–418. [CrossRef]

61. Jia, K.; Liang, S.; Zhang, L.; Wei, X.; Yao, Y.; Xie, X. Forest cover classification using Landsat ETM+ data and time series MODIS NDVI data. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 32–38. [CrossRef]
62. Montavon, G.; Braun, M.L.; Krueger, T.; Muller, K.-R. Analyzing local structure in kernel-based learning: Explanation, complexity, and reliability assessment. *IEEE Signal Process. Mag.* **2013**, *30*, 62–74. [CrossRef]
63. Campos-Taberner, M.; García-Haro, F.J.; Camps-Valls, G.; Grau-Muedra, G.; Nutini, F.; Crema, A.; Boschetti, M. Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring. *Remote Sens. Environ.* **2016**, *187*, 102–118. [CrossRef]