





Article

# Spatial Rice Yield Estimation Based on MODIS and Sentinel-1 SAR Data and ORYZA Crop Growth Model

Tri D. Setiyono <sup>1,\*</sup>, Emma D. Quicho <sup>1</sup>, Luca Gatti <sup>2</sup>, Manuel Campos-Taberner <sup>3</sup> , Lorenzo Busetto <sup>4</sup> , Francesco Collivignarelli <sup>2</sup>, Francisco Javier García-Haro <sup>3</sup> , Mirco Boschetti <sup>4</sup> , Nasreen Islam Khan <sup>1</sup> and Francesco Holecz <sup>2</sup>

<sup>1</sup> International Rice Research Institute, DAPO Box 7777, Metro Manila 1301, Philippines; e.quicho@irri.org (E.D.Q.); n.khan@irri.org (N.I.K.)

<sup>2</sup> Sarmap, Cascine di Barico 10, Purasca 6989, Switzerland; lgatti@sarmap.ch (L.G.); fcolli@sarmap.ch (F.C.); fholecz@sarmap.ch (F.H.)

<sup>3</sup> Department of Earth Physics and Thermodynamics, Faculty of Physics, Universitat de València, València 46100, Spain; Manuel.Campos@uv.es (M.C.-T.); J.Garcia.Haro@uv.es (F.J.G.-H.)

<sup>4</sup> Institute for Electromagnetic Sensing of the Environment, Italian National Research Council, Via Bassini 15, Milan 20133, Italy; busetto.l@irea.cnr.it (L.B.); boschetti.m@irea.cnr.it (M.B.)

\* Correspondence: t.setiyono@irri.org; Tel.: +63-2-580-5600

Received: 7 December 2017; Accepted: 23 January 2018; Published: 14 February 2018

**Abstract:** Crop insurance is a viable solution to reduce the vulnerability of smallholder farmers to risks from pest and disease outbreaks, extreme weather events, and market shocks that threaten their household food and income security. In developing and emerging countries, the implementation of area yield-based insurance, the form of crop insurance preferred by clients and industry, is constrained by the limited availability of detailed historical yield records. Remote-sensing technology can help to fill this gap by providing an unbiased and replicable source of the needed data. This study is dedicated to demonstrating and validating the methodology of remote sensing and crop growth model-based rice yield estimation with the intention of historical yield data generation for application in crop insurance. The developed system combines MODIS and SAR-based remote-sensing data to generate spatially explicit inputs for rice using a crop growth model. MODIS reflectance data were used to generate multitemporal *LAI* maps using the inverted Radiative Transfer Model (RTM). SAR data were used to generate rice area maps using MAPScape-RICE to mask *LAI* map products for further processing, including smoothing with logistic function and running yield simulation using the ORYZA crop growth model facilitated by the Rice Yield Estimation System (Rice-YES). Results from this study indicate that the approach of assimilating MODIS and SAR data into a crop growth model can generate well-adjusted yield estimates that adequately describe spatial yield distribution in the study area while reliably replicating official yield data with root mean square error, RMSE, of 0.30 and 0.46 t ha<sup>-1</sup> (normalized root mean square error, NRMSE of 5% and 8%) for the 2016 spring and summer seasons, respectively, in the Red River Delta of Vietnam, as evaluated at district level aggregation. The information from remote-sensing technology was also useful for identifying geographic locations with peculiarity in the timing of rice establishment, leaf growth, and yield level, and thus contributing to the spatial targeting of further investigation and interventions needed to reduce yield gaps.

**Keywords:** rice yield; leaf area index; *LAI*; remote sensing; MODIS; synthetic aperture radar; SAR; Sentinel-1; ORYZA crop growth model

## 1. Introduction

Smallholder farmers, representing 85% of the world's farms [1], face numerous risks to their agricultural production from pest and disease outbreaks, extreme weather events, and market shocks that threaten their household food and income security [2,3]. These risks are further exacerbated by the unfavorable impact of climate change on rainfall distribution, the rise in sea level, and the increased frequency and duration of extreme weather events such as droughts, floods, cyclones, and storm surges [4], with tropical countries as the most vulnerable geographies [5]. Crop insurance is therefore one viable strategy to mitigate the impact of this global environmental challenge through its contribution to transferring and reducing the risk of crop loss [6–8]. Area-yield index insurance is a promising form of crop insurance and it is more attractive to farmers and the insurance industry because of the lower associated basis risk [9], and farmers' willingness to pay was found to be higher with this form of insurance than with weather-based insurance [10]. The implementation of area yield-based insurance in developing and emerging economy countries, however, is constrained by the limited availability of detailed historical yield records of more than 10 years [9].

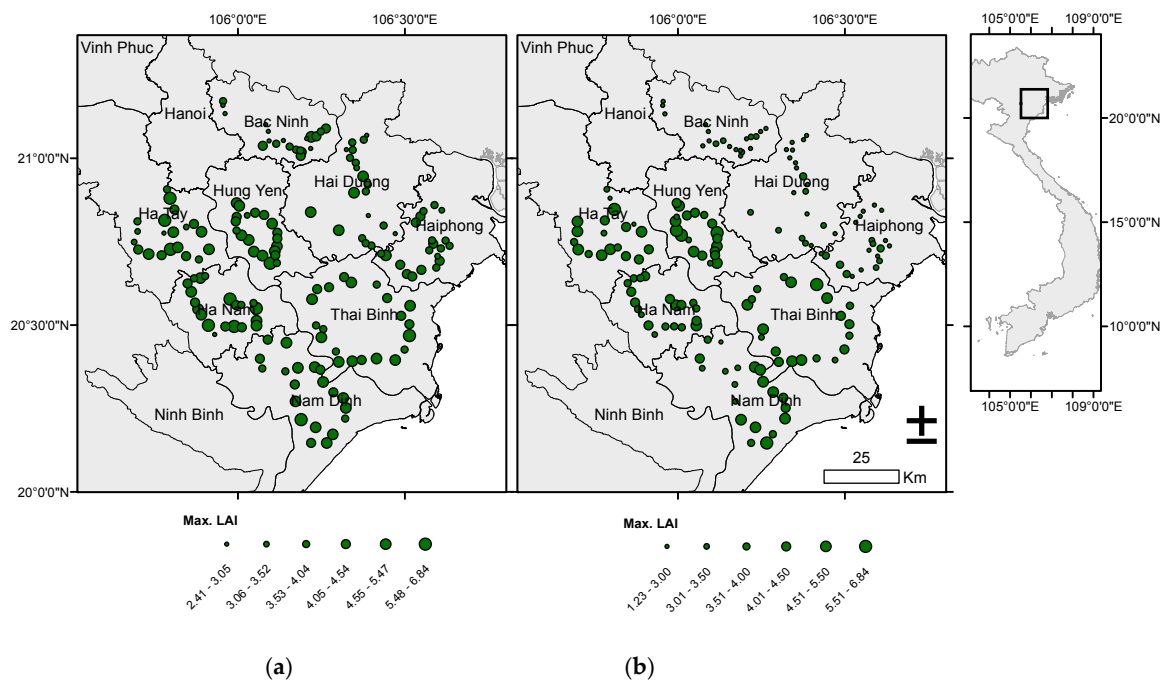
The provision of crop yield data using a biophysical modeling approach is desirable due to the unbiased and replicable nature [11,12] and it can be carried out by exploiting remote-sensing data [13–18], using rainfall data in a statistical framework [19], using a crop growth model [20], or by integrating remote-sensing data and a crop growth model [21,22]. The latter approach is more promising than the empirical approach of translating remotely sensed vegetation indices directly into crop yield and production values [23–25]. This is because the integration approach exploits the synergies between: (i) remote-sensing technology strength in capturing spatial and temporal variation related to agro-practices (e.g., crop establishment dates) and seasonal crop development (i.e., phenology) and vegetation status (e.g., leaf area index); and (ii) process-based crop growth model strength in reliably simulating yield by capturing biophysical growth drivers (microclimate, water, and nutrient) once key parameters are properly assigned [22].

The present study demonstrates a methodology of remote-sensing and crop growth model-based rice yield estimation at district level aggregation and evaluates its performance in a specific geography (Red River Delta, Vietnam) and rice growing seasons (summer and spring 2016). The methodology is intended to be applied to other geographies and capture more growing seasons and years in the past. This study is based on the successful framework of the Remote Sensing-based Information and Insurance from Crops in Emerging economies (RIICE) project, in which rice yield is estimated using crop models in conjunction with Earth Observation (EO)-derived crop-related information. In particular, it was demonstrated that, when Synthetic Aperture Radar (SAR) data with an appropriate temporal revisiting time (e.g., Sentinel-1 data) are available, it is possible to accurately identify rice cultivated areas and estimate crop start of season (SOS) [26] information, which can be used as spatialized inputs for the ORYZA crop growth simulation model [27] to produce spatially explicit yield estimations [28]. These yield estimates are of paramount importance for developing innovative area yield index insurance schemes [29]. To fully perform this task, however, the availability of "historic" information on yield spatial variability is required for a reasonably long period (e.g., 10 years) [30] and therefore it is not possible to rely on the SAR Sentinel-1 approach for the provision of such spatial yield data given that SAR data from Sentinel-1 have been made available only in recent years (Sentinel-1A and Sentinel-1B from 2014 and 2016, respectively). Consequently, to provide the required historical spatially explicit information, the only consistent source of remotely sensed data is represented by optical moderate resolution imagery with quasi-daily revisiting cycle such as SPOT (satellite pour l'observation de la terre)-Vegetation [31] and TERRA/AQUA-MODIS (moderate-resolution imaging spectroradiometer) [32]. The main objective of this article is therefore to develop a processing chain based on the integration of added-value products derived from Sentinel-1 (provided by the European Space Agency, ESA) and MODIS (provided by the National Aeronautics and Space Administration, NASA) data into a crop growth model for spatialized yield estimation in the Red River Delta of Vietnam.

This article is structured as follows. Section 2 describes the study area and datasets used. Section 3 describes the algorithms for retrieving phenological stages and biophysical parameters from MODIS and SAR, as well as the yield modeling approach based on a crop growth model. Section 4 describes product validation and inter-comparison whereas results and discussion are presented in Sections 5 and 6, respectively. One relevant aspect is to assess the reliability of moderate resolution optical data (i.e., MODIS data archive) in the yield estimation system by comparing results with default RIICE solution based on Sentinel-1 (S1) data as well as with reported yield data from the official source.

## 2. Study Site and Data

This study focuses on eight rice-producing provinces in the Red River Delta in Vietnam: Bach Ninh, Ha Nam, Ha Noi, Hai Duong, Hai Phong, Hung Yen, Nam Dinh, and Thai Binh (Figure 1). Two other provinces, Vinh Phuc and Ninh Binh, were not considered in this study because of their relatively small contribution to total rice production in the Red River Delta. The study area is dominated by irrigated rice, which is cultivated in the spring and summer seasons. The spring season spans from February to June and typically involves rice varieties of about 120-day maturity duration. The summer season usually starts in June and ends in early October, involving rice varieties of 90–120-day maturity duration. The rice crop establishment method used in the area is mostly transplanting.



**Figure 1.** Maps of the study site in the Red River Delta, Vietnam. Graduated symbols indicate the magnitude of maximum *LAI* measured in situ during the 2016: spring (a); and summer (b) seasons.

The remote-sensing data used in this study were obtained from MODIS and Sentinel-1 sensors for the year 2016 over the study area. Concerning MODIS, a combined time series of the TERRA and AQUA 16-day Composite Vegetation Indices at 250-m resolution distributed by the National Aeronautics and Space Administration (NASA) were used (MODIS products MOD13Q1 and MYD13Q1 [33]). These products include both the NDVI (Normalized Difference Vegetation Index [34]) and *EVI* (Enhanced Vegetation Index [35]) spectral indices, and surface reflectance in four spectral bands (blue, red, near infrared, and middle infrared (band 7—2105–2155 nm). Since MOD13Q1 and MYD13Q1 are based on compositing periods shifted by eight days between each other, the time series has a theoretical eight-day frequency. MODIS reflectance data from the entire 2016 were used to capture the 2016 spring and summer seasons. Concerning SAR data, Sentinel-1A data distributed by

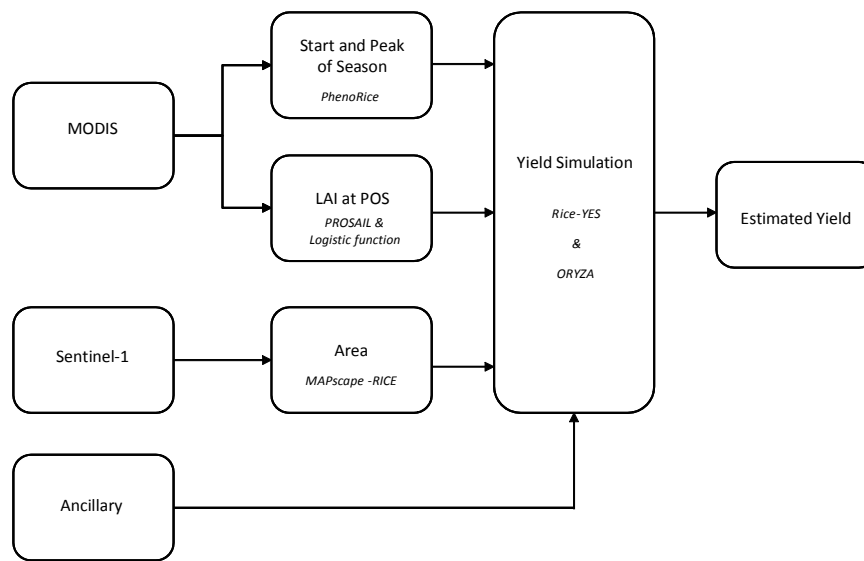
the European Space Agency (ESA) (C-band, VV and VH polarization) with a 12-day regular repeat cycle and 20-m spatial resolution [36] were used. Sentinel-1 SAR images from 17 December 2015 to 21 May 2016 were used for the 2016 spring season whereas images from 2 June to 12 October 2016 were used for the 2016 summer season.

Non-remote-sensing data used in the study include meteorological, soil, and agronomic management data, used as spatialized inputs for the ORYZA model [37]. In particular, daily solar radiation ( $\text{Mj m}^{-2} \text{ day}^{-1}$ ) data were obtained from the NASA POWER (Prediction Of Worldwide Energy Resource) website (<http://power.larc.nasa.gov>), minimum and maximum daily temperature ( $^{\circ}\text{C}$ ) data were obtained from the NCDC Global Summary of the Day (GSOD, <http://www.ncdc.noaa.gov>), whereas rainfall data ( $\text{mm day}^{-1}$ ) were obtained from Climate Hazards Group InfraRed precipitation with station data (CHIRPS, <http://chg.geog.ucsb.edu/data/chirps/>). All meteorological data were resampled to 15 arc minutes resolution. Soil data were extracted from the World Inventory of Soil Emission potential (WISE) dataset (<http://www.icasa.net/toolkit/wise.htm>) and Harmonized World Soil Database (HWSD) (<http://webarchive.iiasa.ac.at>). Agronomic management information such as rice variety duration, crop establishment method, water management, and amount of inorganic N fertilizer was inferred for the study area based on interviews with farmers as part of ground data collection conducted in eight provinces and involving 20 rice fields in each province. The ground observations were made at a regular interval following Sentinel-1A satellite acquisition date. Data collection includes the geographic coordinates of the monitored field, actual weather and field conditions, rice phenological stages, and leaf area index values during each field visit. These data were collected using smartphones with pre-structured questionnaires in Geo-Open Data Kit (Geo-ODK) [38] and the PocketLAI mobile application [39]. The accuracy of GPS coordinates data from Geo-ODK is 4–7 m. Toward the end of the season, rice and non-rice land type surveys were conducted to validate the rice area map for each season. A total of 428 and 246 data points were randomly collected across the study site during the spring and summer season, respectively, for the rice area map validation. Official rice yield and area data for each province in the study area were obtained from the Vietnam General Statistics Office (GSO) available at district-level granularity and therefore this is the spatial aggregation unit used for comparing yield results in this study. Sixty-six districts were included in the yield comparison exercise. In addition, yield comparisons were conducted at provincial-level aggregation. Provincial-level aggregated yield data were calculated by summing the total rice production across districts (area  $\times$  yield) in the same province and dividing the resulting output by the sum of rice area for the given province. In Vietnam, district is one level of administrative unit below province. The average physical areas of provinces and districts in the Red River Delta are approximately 1400 and 100  $\text{km}^2$ , respectively.

### 3. System Description

The developed system combining MODIS- and SAR-based remote-sensing data to generate spatially explicit rice yield using a crop growth model consists of the following components (Figure 2):

- i Seasonality information: The selected seasonal drivers of crop growth are: (i) start and peak of season (SOS and POS) derived from MODIS using the rule-based PhenoRice algorithm [40]; and (ii) multitemporal leaf area index (*LAI*) estimates derived from MODIS reflectance data using a hybrid approach exploiting machine learning algorithms fed with the synthetic database generated from the PROSAIL Radiative Transfer Model (RTM) [41]. These estimations are then fitted into a logistic function for *LAI* [42].
- ii Rice area maps are derived from Sentinel-1 SAR data using MAPScape-Rice [26].
- iii Yield simulation using the Rice Yield Estimation System (Rice-YES) [28] are performed to coordinate both remote-sensing (output from MODIS and Sentinel-1) and non-remote-sensing (meteorological, soil, and agronomic management) data, and mainstream them into the ORYZA crop growth model [27].



**Figure 2.** Schematic diagram of the components involved in the generation of spatially explicit rice yield using remote sensing (optical and SAR) and a crop growth model. Ancillary data include weather, soil, and agronomic management information. Rice-YES, Rice Yield Estimation System.

### 3.1. MODIS Rice Information Retrieval

#### 3.1.1. Estimation of Rice Start and Peak of Season

Retrieval of phenological indicators from MODIS data is based on the use of the PhenoRice method [40], which allows identifying rice cultivated pixels and analyzing their phenological signal in a consistent and flexible way. The method works in two steps: (1) rice crop detection; and (2) phenological dates estimation. Rice crop detection is based on the analysis of time series of *EVI* and *NDFI* (Normalized Difference Flood Index [43]) spectral indices.

$$EVI = G * \frac{\rho_{B2} - \rho_{B1}}{\rho_{B2} + C1 * \rho_{B1} - C2 * \rho_{B3} + L} \quad (1)$$

$$NDFI = \frac{\rho_{B1} - \rho_{B7}}{\rho_{B1} + \rho_{B7}} \quad (2)$$

where  $\rho_{B1}$  is the RED band (central wavelength, 645 nm),  $\rho_{B2}$  is NIR (near infra-red) band (central wavelength, 858.5 nm),  $\rho_{B3}$  is BLUE band (central wavelength, 469 nm),  $L = 1$ ,  $C1 = 6$ ,  $C2 = 7.5$ ,  $G$  (gain factor) = 2.5, and  $\rho_{B7}$  is the SWIR (short wave infra-red) band (central wavelength, 2130 nm).

In particular, PhenoRice uses a rule-based method that classifies a MODIS pixel as a rice crop if: (1) a clear and unambiguous flood condition is detected based on *NDFI* values; and (2) the *EVI* signal following the identified flooding is consistent with what is expected for a rice crop (e.g., rapid growth, followed by a relatively stationary peak period and a rapid decrease related to yellowing and harvesting). If a pixel is considered to belong to rice, phenological dates estimation is also conducted. In particular, the date corresponding to the strongest minima of the *EVI* curve identified in the proximity of the detected agronomic flooding is considered as a proxy of crop establishment (e.g., rice transplanting period), and the absolute maxima of the curve in a phenologically sound period as a proxy of crop heading (peak of season, POS). The start of season date is finally computed as the date at which the *EVI* signal increases 10% of the min/max range from the minimum value. A more detailed description of the PhenoRice algorithm and its performance can be found in Boschetti et al. [40]. The values of SOS and POS (for both the spring and summer rice seasons) obtained from PhenoRice were used as inputs in the ORYZA crop growth model (Section 3.3).

### 3.1.2. LAI Retrieval

Retrieval of LAI from MODIS surface reflectance was conducted by inverting the PROSAIL Radiative Transfer Model [44]. PROSAIL simulates surface reflectance of the vegetated surface of interest (in our case, rice crops in the tropics) in the range of 400 to 2500 nm. For this purpose, PROSAIL uses a set of bio-chemical and structural parameters at canopy and leaf levels, namely, leaf chlorophyll, dry matter, and relative water contents ( $C_{ab}$ ,  $C_{dm}$ , and  $C_{WREL}$ , respectively); the leaf mesophyll parameter ( $N$ ); average leaf angle ( $ALA$ ); LAI; and the hot-spot parameter ( $H_s$ ). In addition, a soil brightness parameter ( $\beta_s$ ) was used and applied as a multiplicative factor over a site-specific spectral background library to represent diverse rice background reflectances. With the objective of considering non-homogeneous pixels (i.e., rice field boundaries), a parameter ( $vCover$ ) was used for representing pixels formed by a fraction of pure vegetation and pure flooded/bare soil [45].

PROSAIL was run 2000 times to obtain a database composed of surface reflectance corresponding to the MODIS bands available in the MOD13Q1/MYD13Q1 products such as blue, red, near infrared (NIR), and middle-infrared (MIR) and the associated LAI values. In particular, PROSAIL variables were parameterized according to the distributions in Table 1. The values were similar to those of other studies with the aim of simulating rice crop surface reflectance [41,46].

**Table 1.** Distribution of the PROSAIL parameters used for the simulation of rice reflectance.

	Parameter	Min	Max	Mean	Std	Type
Canopy	LAI ( $m^2 m^{-2}$ )	0	6	3	2	Gaussian (*)
	ALA ( $^\circ$ )	30	80	60	20	Gaussian
	$H_s$	0.1	0.5	0.2	0.2	Gaussian
	$vCover$	0.5	1.0	1.0	0.2	Truncated Gaussian (*)
Leaf	$N$	1.2	2.2	1.5	0.3	Gaussian
	$C_{ab}$ ( $\mu g cm^{-2}$ )	20	90	45	30	Gaussian
	$C_{dm}$ ( $g cm^{-2}$ )	0.005	0.03	0.01	0.008	Gaussian
	$C_{WREL}$	0.6	0.8	-	-	Uniform
Soil	$\beta_s$	0.30	1.20	0.90	0.25	Gaussian

LAI, leaf area index; ALA, average leaf angle;  $H_s$ , hotspot;  $vCover$ , fraction of vegetation covering a pixel;  $N$ , mesophyll parameter;  $C_{ab}$ , leaf chlorophyll content;  $C_{dm}$ , dry mater content;  $C_{WREL}$ , relative water content;  $\beta_s$ , soil brightness parameter; Std, standard deviation. (\*) A 5% of pure background spectra ( $vCover = 0$ ) was added to represent bare areas.

The simulated database was then used for calibrating a Gaussian process regression (GPR) [46] model, which has proven to be an efficient and robust machine learning non-linear regression tool for bio-physical parameter retrieval [47]. In our case, the GPR lays down a relationship between the four MOD13Q1 surface reflectances ( $x = [x_1, x_2, x_3, x_4] \in \mathbb{R}^4$ ) and the LAI ( $y \in \mathbb{R}$ ) as  $y_* = \sum_{i=1}^N \alpha_i K(x_i, x_*) + \alpha_0$ , where  $\alpha_i$  is the weight assigned to each training spectrum  $x_i$ ,  $K$  is a kernel function that measures the similarity between the real MODIS spectrum and all the simulated ( $n = 2000$ ) training spectra, and  $\alpha_0$  is the bias in the regression function. In this study, we used the automatic relevance determination (ARD) kernel:

$$K(x_i, x_j) = \nu \exp \left( - \sum_{b=1}^4 \frac{(x_i^{(b)} - x_j^{(b)})^2}{2\sigma_b^2} \right) + \sigma_n^2 \delta_{ij} \quad (3)$$

where  $\nu$ ,  $\sigma_n$ , and  $\sigma_b$  are the so-called kernel hyperparameters accounting for a scaling factor, the noise standard deviation, and the spread of the relations for each particular spectral band, respectively.

Given our training dataset  $D \equiv \{x_n, y_n | n = 1, \dots, 2000\}_{i=1}^{2000}$  and a real MOD13Q1 spectrum test, the GPR provides a predictive distribution of its corresponding output  $p(y_* | x_*, D) = N(y_* | \mu_*, \sigma_*^2)$ ,

from which we can compute mean predictions,  $\mu_*$  (e.g., *LAI* estimates), and also the predictive variance  $\sigma_*^2$  as:

$$\mu_* = k_*^\top (K + \sigma_n^2 I_n)^{-1} y = k_*^\top \alpha \quad (4)$$

$$\sigma_*^2 = \sigma^2 + k_{**} - k_*^\top (K + \sigma_n^2 I_n)^{-1} k_* \quad (5)$$

where  $k_*$ ,  $k_{**}$ , and  $K$  are the covariance between the training and the real spectra, the prior covariance, and the training samples covariance, respectively, and  $k_*^\top$  denotes vector  $k_*(x_i, x_j)$ .

The resulting *LAI* time series data were then smoothed using the *LAI* phenology approach based on partitioning of leaf expansion and senescence processes [42,48]. To account for variability in rice maturity duration, the first step of the *LAI* smoothing was to adjust the time variable from days after emergence to maturity progress index (*MPI*), which is calculated as follows:

$$MPI = 100(DAE/MTD) \quad (6)$$

where *DAE* is the number of days after emergence at each time step (daily) and *MTD* is rice maturity duration in days. In the subsequent step, for each day after emergence, both expanded and senesced *LAI* components were calculated using the following logistic function:

$$LAI_j = \frac{a_j}{1 + b_j \exp(-MPIc_j)} \quad (7)$$

where  $j$  in the *LAI* term denotes either an expanded (*E*) or senesced (*S*) component of the *LAI*, while  $a$ ,  $b$ , and  $c$  are logistic model parameters. Expanded and senesced *LAI* components were computed independently from each other. Finally, *LAI* is calculated by subtracting the senesced components ( $LAI_S$ ) from the expanded components ( $LAI_E$ ) for each day:

$$LAI = LAI_E - LAI_S \quad (8)$$

The logistic function parameters were derived by fitting the *LAI* time series derived from MODIS. Inferred rice *LAI* values at early expansion stage were defined as *MPI* equals 33. These single-date *LAI* values become the inputs for the subsequent processing in Rice Yield Estimation System (Rice-YES) and ORYZA crop growth model (Section 3.3).

### 3.2. SAR Sentinel-1 Data Processing

#### 3.2.1. Identification of Rice Cultivated Areas

A fully automated processing chain was followed to convert the Sentinel-1 single-look complex (SLC) time series into terrain-geocoded  $\sigma^\circ$  values. The processing chain is a module within MAPscape-RICE, which included the following steps: strip mosaicking, image co-registration, time series speckle filtering, terrain geocoding, radiometric calibration and normalization, anisotropic non-linear diffusion (ANLD) filtering, and the removal of atmospheric attenuation. The ANLD filter performs a smoothing and a high level of regularization in homogeneous areas, while preserving inhomogeneous signal variations (discontinuities), which are typical of features such as field edges, roads, rivers, and other image information that we want to preserve. On the other hand, although microwave signals can penetrate clouds, it is possible to observe sudden changes in the dielectric constant at shorter wavelengths (X- and C-band) due to atmospheric moisture changes, dew, rain, and severe storms. These events affect the temporal signature of  $\sigma^\circ$  in two ways: (i) the thick layer of water vapor generates a strong decrease in  $\sigma^\circ$  during the event, followed by a strong increase after the event; and (ii) the intense rainfall generates a strong increase in  $\sigma^\circ$  during the event, followed by a strong decrease after the event. These effects were removed by identifying anomalous peaks or troughs and correcting the  $\sigma^\circ$  values by means of an interpolator.

The multi-temporal rule-based rice detection algorithm, originally proposed by [26] for X-band HH SAR data and adapted to C-band VV/VH data [41], was then applied on the obtained  $\sigma^\circ$  time series to identify rice area within the study area (for both the spring and summer seasons). Processing has been carried out in MAPscape-RICE. The rice identification algorithm is based on a set of rules defined based on a priori knowledge on the rice calendar, crop practices, and agro-ecological conditions of the study areas. The main rules are as follows:

1. Rice exclusion condition: The average  $\sigma^\circ$  is lower/higher than expected, or average  $\sigma^\circ$  is below a minimum value for longer than expected or variation of  $\sigma^\circ$  larger than expected.
2. Presence of flooding conditions at the start of season (SOS): The temporal series is analyzed starting from the first image acquisition to identify when  $\sigma^\circ$  drops below a maximum value for SOS flooding; this time is set as SOS.
3. Confirmation of rice presence: After flooding detection, rice presence is confirmed if  $\sigma^\circ$  increases after SOS, which is consistent with the expected value for rice crops (rapid growth of rice biomass after flooding).
4. Late rice condition: When the length of the S-1A time series after detected SOS is shorter than expected, rice is classified as being sown late in the season.
5. New rice/crop season: An unexpected drop in  $\sigma^\circ$  after SOS, which suggests a new flood or a new crop season. Note that, if Rule 2 is not met during the entire temporal series, another rule allows one to determine whether flooding occurred before the first Sentinel-1A acquisition. If  $\sigma^\circ$  variation is above a pre-defined value, which is expected for rice, and  $\sigma^\circ$  decreases after the maximum value, the pixel is classified as early rice.

The rice areas derived from Sentinel-1 time series were used as a masking layer for *LAI* retrieval based on both MODIS (Section 3.1.2) and SAR (Section 3.2.2) data. A 250-m MODIS pixel was determined as verified rice pixel and used for further processing when it overlaid with at least 75% of the 30-m Sentinel 1 rice area pixels. Such threshold was used to avoid strong contamination of MODIS reflectance data from non-rice surface.

### 3.2.2. *LAI* Retrieval

Since the objective of this study is to assess the reliability of rice yield estimation based on optical MODIS data by comparing the results with the already proven solution based on SAR data [26,28], *LAI* maps also had to be derived from the Sentinel-1A time series to be used as exogenous input for crop model estimates. Here, we describe the method used for *LAI* data retrieval from Sentinel-1A. Overall structure of the SAR-based yield estimation is very similar to the schematic diagram in Figure 1, except, in this case, Sentinel-1 is the only source of remote-sensing-derived products going toward yield simulation.

Leaf area index values at early leaf expansion stage, EES (at 33% of rice maturity progress, roughly 40 days after SOS for 110-day rice), were inferred from SAR backscatter using the modified version of the water cloud vegetation model [49] based on empirical assumption of plant height and moisture dynamic described in [50]. According to that method, *LAI* can be derived from  $\sigma^\circ$  using the following formula:

$$LAI = \frac{B + A \left( \ln \left[ \frac{10^{0.1\sigma^\circ} - \alpha \cos(\theta)}{\sigma_G - \alpha \cos(\theta)} \right] \cos(\theta) / (-2\beta) \right)}{\ln \left[ \frac{10^{0.1\sigma^\circ} - \alpha \cos(\theta)}{\sigma_G} \right] \cos(\theta) / (-2\beta) + C} \quad (9)$$

where  $\sigma^\circ$  is the radar backscattering in decibels (dB) at C-band (5.3 GHz) derived from Sentinel-1, as described in Section 3.2.1;  $\alpha$  is the backscattering coefficient at full canopy closure ( $\text{m}^2 \text{m}^{-2}$ );  $\beta$  is the coefficient of attenuation per unit canopy water ( $\text{m}^2 \text{kg}^{-1}$ );  $\sigma_G$  is the backscattering coefficient from canopy background ( $\text{m}^2 \text{m}^{-2}$ );  $\theta$  is the incident angle of the radar beam ( $^\circ$ ); and  $A$ ,  $B$ , and  $C$  are the coefficients of a non-linear regression relating rice *LAI* to canopy water ( $\text{kg m}^{-3}$ ) and height (m). The values of 10.19, 0.1534, and 1.511 were used for  $A$ ,  $B$ , and  $C$ , respectively, based on [50].



The values of 0.209, 0.336, and 0.0340 were used for  $\alpha$ ,  $\beta$ , and  $\sigma^{\circ}_G$ , respectively. The processing chain from multi-temporal backscatter to *LAI* at EES also involves normalization and filtering of the raw backscatter inputs. These steps are important requirements prior to the use of the backscatter inputs in the water cloud vegetation model to ensure that the model performs within the acceptable range of backscatter inputs between  $-14$  (minimum  $\sigma^{\circ}$ ) and  $-8$  (maximum  $\sigma^{\circ}$ ).

### 3.3. Yield Simulation from Crop Model and EO Product Integration

On the downstream side of the system, the ORYZA crop growth model was used to simulate rice yield using remote-sensing and non-remote-sensing inputs described above. ORYZA is a process-based crop simulation model underpinned with largely mechanistic components such as photosynthesis, respiration, and water-nitrogen balance [27]. The interface between SOS and POS dates and *LAI* values obtained from remote sensing and the ORYZA crop growth model is the Rice Yield Estimation (Rice-YES) system, a software written in Object Pascal, Python, and R [28,51]. Rice-YES facilitates the use of inferred *LAI* values derived from either optical MODIS data or SAR data to calibrate relative leaf growth rate parameters in ORYZA.

Rather than running a yield simulation in each pixel of the *LAI* maps, *LAI* values from either MODIS or SAR were first grouped based on specific intervals. Spatial allocations of these *LAI* groups were determined by also considering other spatially varying inputs, including weather and SOS data. The resulting combinations of *LAI* groups, SOS, and weather pixels were assigned unique spatial IDs and the number of simulation runs for ORYZA was determined. After running the ORYZA model, the yield results were remapped based on their corresponding spatial IDs. With this approach, the rice yield outputs have the same resolution as the remote-sensing data inputs while processing efficiency is maintained at a high level by eliminating redundant simulation runs with similar *LAI* inputs and identical SOS and weather information. In this case, the yield outputs with MODIS-based approach were at 250 m, whereas the reference SAR-derived yield outputs were at 30 m in spatial resolution.

ORYZA model runs on daily time step. The assimilation of MODIS and SAR-derived *LAI* products begins after the rice crop reached the early expansion stage (roughly 40 days after SOS for 110-day rice). During this early part of the rice-growing cycle, leaf expansion parameters in ORYZA can be effectively calibrated against real ground conditions inferred from satellite observations. *LAI* values derived from MODIS or SAR were used to recalibrate the early leaf growth parameters in ORYZA, in particular the relative leaf growth rate (RGRL) variable. The RGRL variable was an ideal target for calibration given the known sensitivity of biomass accumulation and yield output of ORYZA to changes in this coefficient [52].

## 4. Product Validation and Inter-Comparison

Given that SAR-derived rice area maps are used to mask *LAI* products for yield estimation in this study, it is essential to demonstrate the accuracy of these maps. This is done by comparing rice area derived from SAR against ground data that were collected as part of the ground observations in this study (Section 2). The accuracy assessments consider producers' and users' accuracy, overall accuracy, and Kappa Index [53].

Yield estimates obtained with the Optical + SAR + ORYZA (MODIS-tested approach) and SAR+ORYZA (Sentinel-1, RIICE default approach) configurations were aggregated from the corresponding sensor's resolution to the district and provincial administrative level and then compared between each other and against official government-reported yield using analysis of variance least significant difference, ANOVA LSD, statistics tests and root mean square error (RMSE) and normalized root mean square error (NRMSE) [54]:

$$\text{RMSE} = \sqrt{\sum (S_i - O_i)^2 / n} \quad (10)$$

$$\text{NRMSE} = \left( \sqrt{\sum (S_i - O_i)^2 / n} \right) / (\sum O_i / n) \quad (11)$$

where  $S$  accounts for yield estimates from the tested approach for the  $i$ th province or district in a given geography of interest,  $O$  is the yield from the reference source (other model or official government-reported yield) to be compared against the tested approach for the same province or district in the geography of interest, and  $n$  is the number of pairs of yield estimates from both sources (tested approach vs. other model and tested approach vs. official yield). Yield comparisons at provincial-level aggregation are provided in tabular form in the Results Section, whereas comparisons at district-level aggregation are presented as box plots and scatter plots to emphasize the results of ANNOVA LSD and RMSE analyses, respectively. Equation (10) was also used for evaluating the intermediate  $LAI$  output derived from MODIS and logistic function against the observed in situ  $LAI$  and in such case the  $S$  and  $O$  refer to  $LAI$  variable instead of yield.

## 5. Results

The rice area map derived from Sentinel SAR data in this study has a high accuracy of 90.7% and 94.7% for the 2016 spring and summer season, respectively (Table 2). The Kappa Index values greater than 0.80 indicate a strong agreement between the SAR-derived rice area and the land survey information from the ground observations [53].

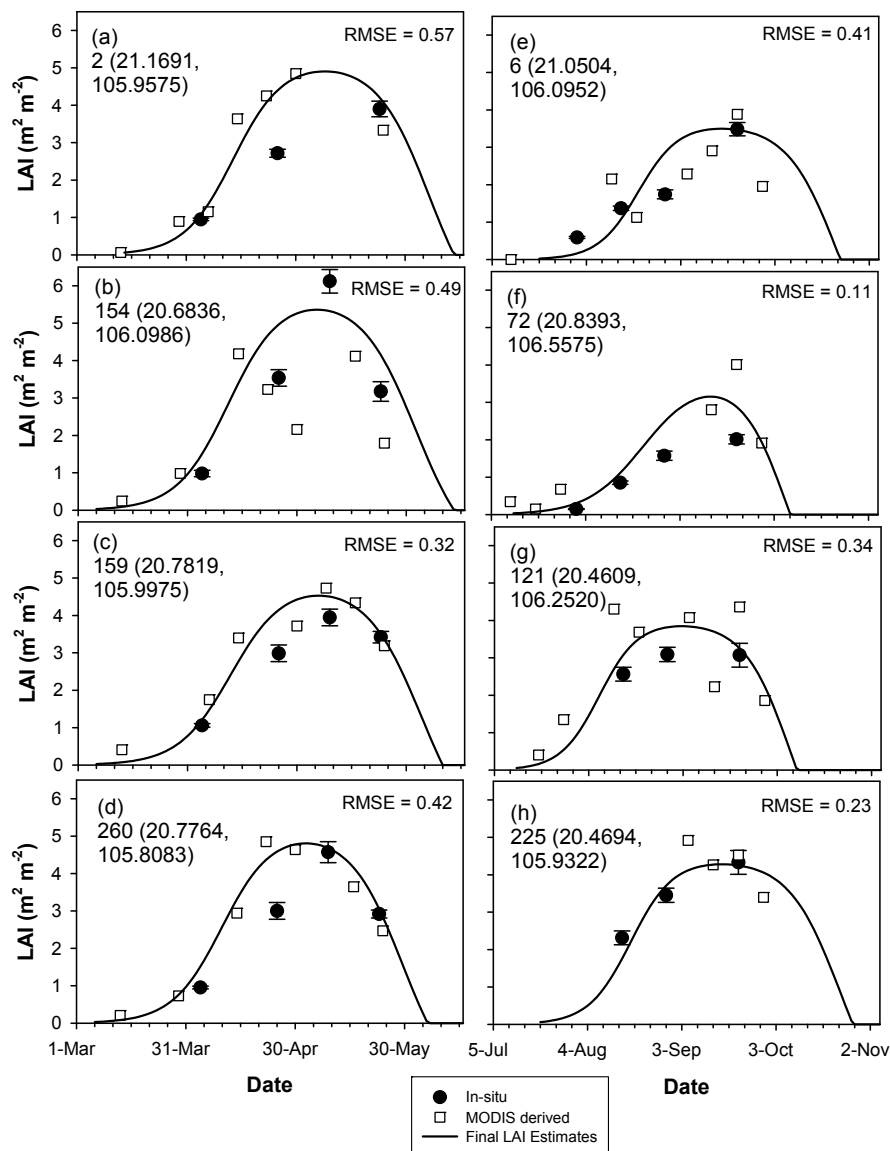
**Table 2.** Accuracy assessment of rice area derived from SAR Sentinel-1 used in this study for the 2016 spring and summer seasons in the Red River Delta of Vietnam. Survey and classified data represent the number of points identified as rice (R) or non-rice (NR).

	2016 Spring Season				2016 Summer Season			
	Survey Data				Survey Data			
		R	NR	Accuracy	R	NR	Accuracy	
<b>Classified data</b>	R	302	9	97.1%	R	216	0	100%
	NR	31	86	73.5%	NR	13	17	56.7%
Accuracy		90.7%	90.5%	90.7%		94.3%	100%	94.7%
Average producers' accuracy				90.6%				97.2%
Average users' accuracy				85.3%				78.3%
Overall accuracy				90.7%				94.7%
Kappa Index				0.81				0.89

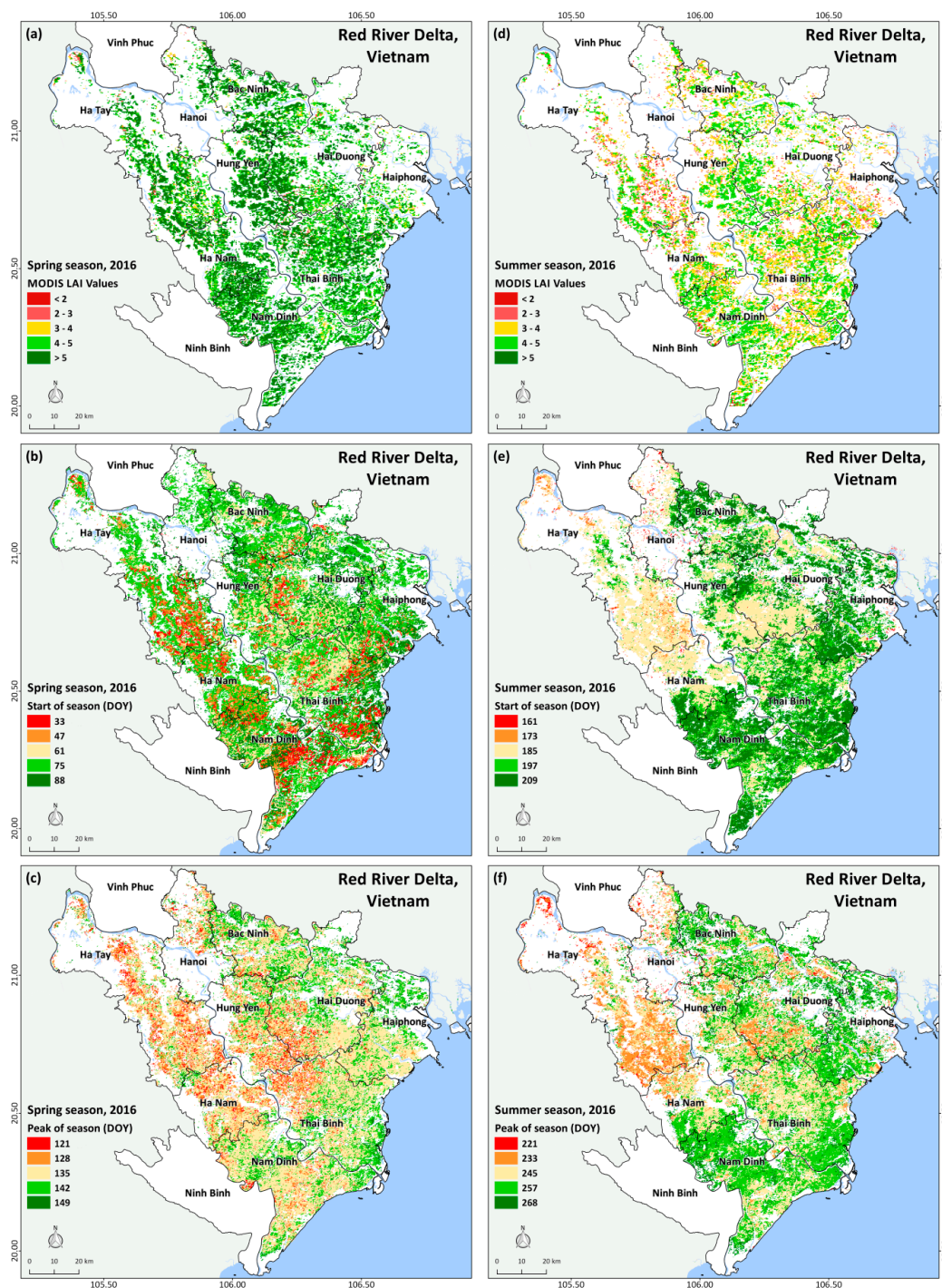
Regarding crop status, the in situ  $LAI$  observations collected at the study sites indicated variation in maximum rice  $LAI$  across seasons (spring vs. summer). In general, higher maximum  $LAI$  values were observed during the spring season with the exception of some areas in the western part of Hay Tay and Hung Yen provinces where the reverse was the case (Figure 1).

Final  $LAI$  estimates derived by smoothing the MODIS-derived  $LAI$  with the  $LAI$  phenology approach (Section 3.1.2) show reasonable agreement against in situ  $LAI$ , with root mean square error (RMSE) ranging from 0.11 to 0.57 (Figure 3). Spatial distribution of estimated maximum  $LAI$  at the study site is shown in Figure 4a,d. Consistent with in situ data, the map of estimated  $LAI$  from MODIS suggests in general higher values for the spring (Figure 4a) than for the summer (Figure 4d) season. In terms of spatial pattern, the maps agree with in situ observations of some areas with relatively high maximum  $LAI$  during the summer season (e.g., in Hung Yen Province) and low  $LAI$  during both seasons (e.g., in Hai Phong Province). Figure 4b–f shows the outputs of the MODIS data processing for estimates of phenological dates. It is evident from these maps that, in the spring season, crop establishment generally spanned about 50 days (from DOY 30 to 80, January to March), although some patches of delayed transplanting (about DOY 80, late March) were detected in the eastern part of the area (Hai Phong and Thai Binh provinces). During the summer season, crop establishment occurred mostly around DOY 190 (July), but with a clear pattern of anticipated

transplanting around DOY 160 (June) in the southern part of Ha Tay Province. Peak of season (POS) during the spring season was detected around DOY 120 and 130 (April/May), with a few exceptions of maximum development at about DOY 140 (end of May), whereas, in the summer season, a larger spatial variability was observed in POS from DOY 220 (August) in the central and northwestern parts of the study area and around DOY 250 (September) in the southeastern part. These estimates are in accordance with ground reference information as recorded during field interviews, indicating that rice transplanting period in the spring season was concentrated in February (DOY 32–60) and harvest from the end of May to early June (DOY 143 to 163); hence, full flowering (i.e., POS) generally occurred around the end of April to early May (DOY 115–128). Regarding the summer season, field survey observations recorded a majority of rice area planted from the end of June to the third week of July 2016 (DOY 178–198).



**Figure 3.** Estimated LAI using MODIS and the Radiative Transfer Model and as well combined with the double logistic model against ground-based LAI for selected study locations (location ID and geo-coordinates are shown in the upper right corner of each panel) during the 2016: spring (a–d); and summer (e–h) seasons in the Red River Delta of Vietnam.

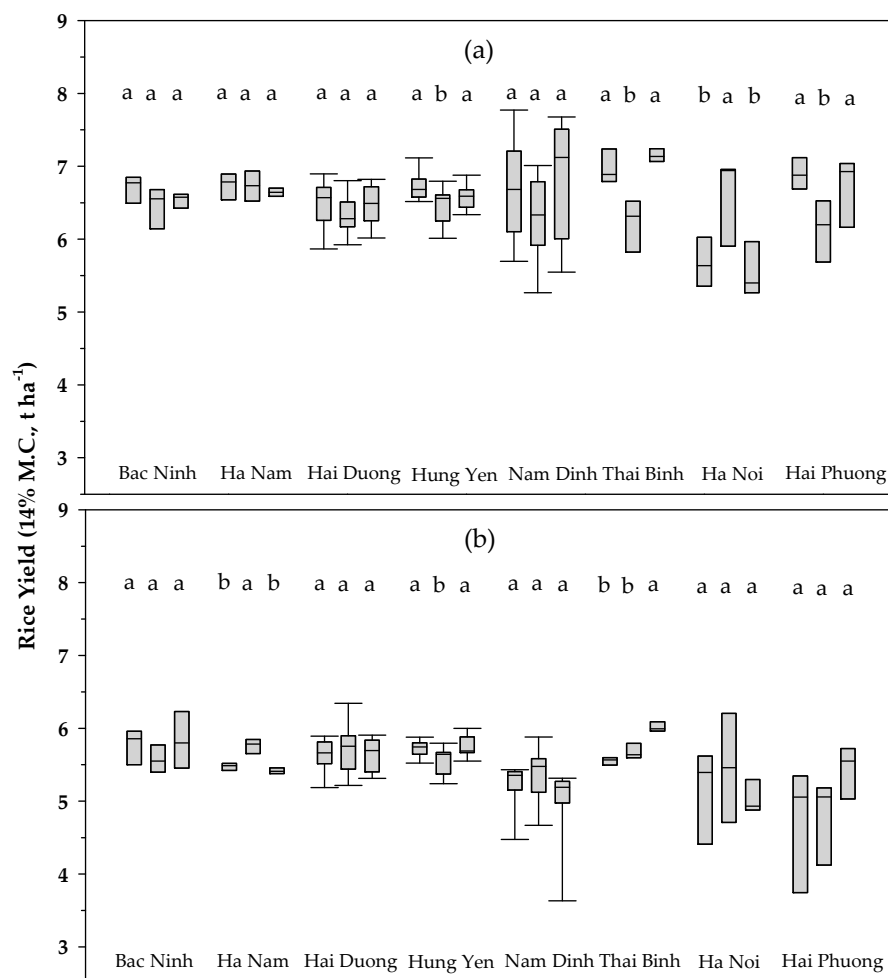


**Figure 4.** Estimated rice LAI ( $\text{m}^2 \text{m}^{-2}$ ) at peak of season (POS) for the 2016: spring (a); and summer (d) seasons in the Red River Delta of Vietnam; and the corresponding: start of season (SOS) (b,e); and POS (c,f) in day of year (DOY), as estimated by PhenoRice.

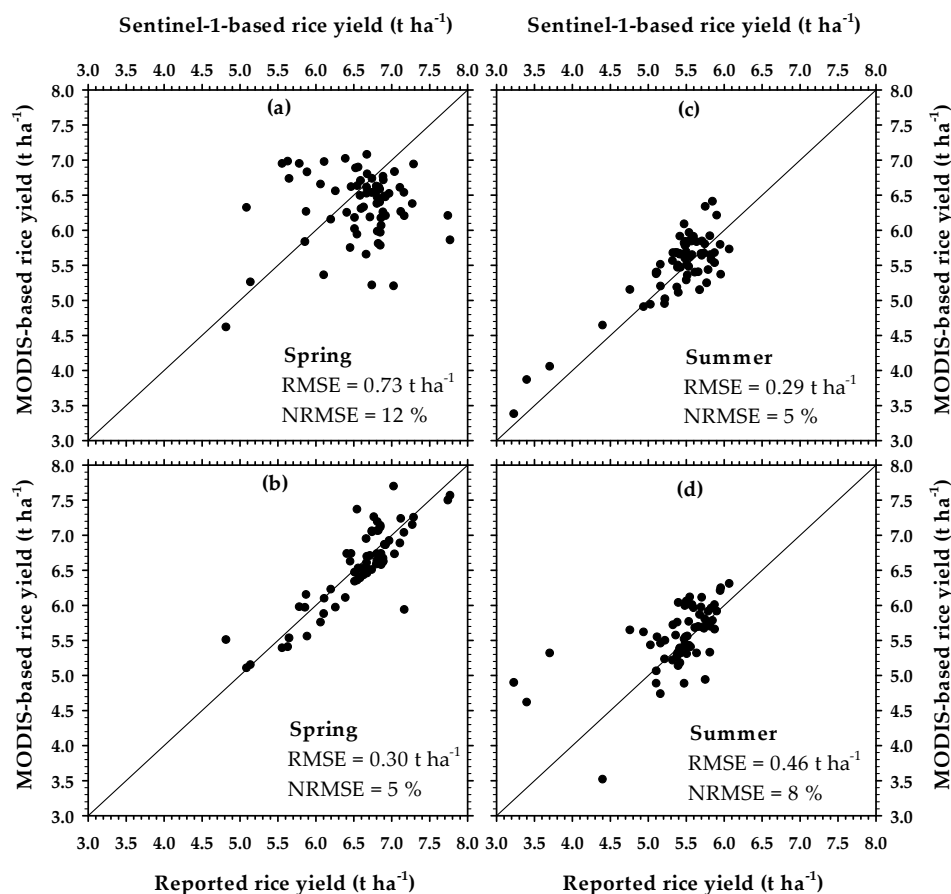
Yield simulation across provinces indicates good agreement between MODIS-derived yield and SAR-based yield and against official government-reported yield with ANOVA LSD test, suggesting no significant difference between the datasets (Table 3). At provincial-level aggregation, the corresponding RMSE (and NRMSE) for comparisons of MODIS-derived yield vs. SAR-derived yield is  $0.59 \text{ t ha}^{-1}$  (9%) and  $0.17 \text{ t ha}^{-1}$  (3%), for the spring and summer season, respectively, whereas RMSE for comparisons

of MODIS-derived yield vs. reported yield is  $0.20 \text{ t ha}^{-1}$  (3%) and  $0.20 \text{ t ha}^{-1}$  (4%) for the spring and summer season, respectively (Table 3).

Similarly, comparison of yield datasets at district-level aggregation shows that MODIS-derived yield was in agreement with reported yield from the government as shown by no significant differences in two datasets based on ANOVA LSD test results in 15 out of 16 cases (Figure 5). Comparisons between MODIS-derived yields and SAR-derived yields, however, show some discrepancies, especially in the spring season. Scatter plots comparing yield results at district-level aggregation from MODIS against SAR and reported official yield are shown in Figure 6. MODIS-derived yield against SAR-based yield has RMSE of 0.73 and  $0.29 \text{ t ha}^{-1}$  (NRMSE of 12% and 5%) for the spring and summer season, respectively. MODIS-derived rice yield data against reported official yield have RMSE of 0.30 and  $0.46 \text{ t ha}^{-1}$  (NRMSE of 5% and 8%) for the spring and summer season, respectively (Figure 6). The higher yield values found in the spring than in the summer season are consistent with both observations and simulation of maximum *LAI* (Table 3 and Figures 1, 3 and 6).



**Figure 5.** Box plots of rice yield aggregated at the district level summarized for eight provinces in the Red River Delta of Vietnam (study sites) for the 2016: spring (a); and summer (b) seasons. In each province, the three box plots represent yield data derived from MODIS, Sentinel-1, and an official source, sequentially, with the corresponding group of three letters near the top of the graphs representing analysis of variance least significant difference, ANOVA LSD, assignment comparing the three datasets.



**Figure 6.** Scatter plot comparing rice yield aggregated at the district level as derived from: MODIS vs. Sentinel-1 (a,c); and MODIS vs. official source (reported) (b,d), for the 2016: spring (a,b); and summer (c,d) seasons in the Red River Delta of Vietnam.  $n = 66$  districts.

**Table 3.** Comparison between average rice yield (14% moisture content, MC) at provincial level derived from MODIS, SAR, and official source, for the 2016 summer and spring seasons in the Red River Delta of Vietnam.

Province	Rice Yield (14% MC) t ha <sup>-1</sup>					
	2016 Spring Season			2016 Summer Season		
	MODIS	Sentinel-1	Reported	MODIS	Sentinel-1	Reported
Bac Ninh	6.71	6.39	6.50	5.85	5.60	5.71
Ha Nam	6.74	6.78	6.63	5.49	5.80	5.53
Ha Noi	5.66	6.81	6.07	5.56	5.51	5.52
Hai Duong	6.43	6.28	6.44	5.68	5.77	5.54
Hai Phong	6.84	6.17	6.95	5.25	5.11	5.63
Hung Yen	6.72	6.44	6.58	5.75	5.58	5.80
Nam Dinh	6.77	6.35	6.92	5.31	5.43	5.16
Thai Binh	6.94	6.14	7.14	5.58	5.66	5.24
Mean <sup>1</sup>	6.60 <sup>a</sup>	6.42 <sup>a</sup>	6.65 <sup>a</sup>	5.56 <sup>a</sup>	5.56 <sup>a</sup>	5.52 <sup>a</sup>
RMSE <sup>2</sup> (t ha <sup>-1</sup> )		0.59	0.20		0.17	0.20
NRMSE <sup>2</sup> (%)		9	3		3	4

<sup>1</sup> In each season, mean values across provinces are followed by letters indicating ANOVA LSD ( $p = 0.05$ ) ranking, with the same letters indicating no significant difference between the different sources of yield data. <sup>2</sup> In each season, RMSE and NRMSE evaluating MODIS-based yield against SAR-based yield and official government-reported yield are shown in the corresponding reference yield columns.

## 6. Discussion

The final *LAI* estimates for the Red River Delta from MODIS obtained by smoothing with a logistic function in this study have a comparable range of RMSE (0.11–0.57) as previous studies involving higher resolution reflectance data from Sentinel-2 and the same inverted PROSAIL Radiative Transfer Model in Spain (0.56), Italy (0.82), and Greece (0.77) [41]. However, this previous approach did not involve smoothing using a logistic function. Time series observation indicates that *LAI* estimates from MODIS in this present study showed occasional unexpected pronounced drops in *LAI* values (Figure 3) similar to ones previously found in Spain [41], which was likely due to cloud obstruction. Time series analysis also reveals some degree of uncertainty in the in situ *LAI* data and it also shows, although not as pronounced as the MODIS-derived *LAI*, some anomaly in the *LAI* dynamics during the growing season. *LAI* data obtained from PocketLAI have been known to have limitations vis-à-vis the conventional method [55]. Given such occasional unexpected fluctuation of *LAI* time series from MODIS and in situ data, smoothing using a logistic function in this study is strongly justifiable to allow reliable *LAI* evolution data during the growing season. Further investigation is needed to evaluate whether, in some cases, the drop in *LAI* values during the growing season may be caused by environmental stresses. In such case, however, the *LAI* data are expected to remain low for the remainder of the season following the drop, something that was not observed in this present study in the Red River Delta during the 2016 growing seasons, inferring that rice in this area during these particular seasons did not suffer major stresses. This inference is supported by the good rice growing conditions reported by farmers based on the field interviews and expert opinions. Another possible reason for the discrepancy between MODIS derived *LAI* versus ground data was the heterogeneity of the surface conditions captured within MODIS 250-m pixels. Such phenomena had been prevented and minimized, however, by requiring in-situ field observations taken at locations that were relatively homogenous, conditions that are commonly found in the study site.

Even with some degree of uncertainty as observed based on some deviation between MODIS-based and SAR-based yield results at district-level aggregation as indicated by ANOVA LSD test and relatively high RMSE of  $0.73 \text{ t ha}^{-1}$  (NRMSE of 12%), overall, MODIS-based yield estimations performed satisfactorily given the good agreement with official yield data at provincial and district aggregation levels (less than 8% bias). The high accuracy obtained for yield estimates at district level aggregation in this study suggests that a successful integration of various approaches to properly exploit both optical and SAR data has been achieved. For example, a proper characterization of the crop growth simulation model has been possible because of a biological function describing *LAI* time series [42]. Optical data from MODIS were used for detecting phenology [40] (input for model simulation initialization) and *LAI* dynamic [56] (drivers for forcing model simulation according to real crop development) over the growing season, whereas high-resolution SAR data were used to identify precise rice area to compensate for the low resolution of MODIS and given the previous success of rice area detection using SAR and rule-based algorithm [26]. In this study, the deviation between modeled yield (configured with MODIS and SAR input data) and reference yield from an official source aggregated at the district level was 5% and 8% for the spring and summer season, respectively. This error level is comparable to, or better than, the reported performance of a similar approach of integrating a remote-sensing and rice crop growth model in Japan with error level ranging from 3.5% to 121.4% [22]. District-level yield data from the official source from the present study suggest that yield values vary from  $-23\%$  to  $15\%$  of the mean value of  $6.65 \text{ t ha}^{-1}$  for the spring season and from  $-36\%$  to  $14\%$  of the mean value of  $5.52 \text{ t ha}^{-1}$  for the summer season (Figure 5). Against such spatial yield variation of the reference yield (official source), the bias of MODIS-based yield against the official yield of 5% to 8% is reasonably low. MODIS-based yield estimates also comparably show similar variation in yield from  $-41\%$  to  $9\%$  of the mean value of  $5.56 \text{ t ha}^{-1}$  and  $-27\%$  to  $18\%$  of the mean value of  $6.6 \text{ t ha}^{-1}$  for the summer and spring season, respectively. Therefore, the yield estimation approach presented here is a promising method for reconstructing credible historical yield data to support an area yield-based crop insurance program

for rice growing areas in the tropics by exploiting the consistently available archives of MODIS data. Further investigation is needed to elucidate the why the discrepancy of MODIS-based yield results against SAR-based results was greater in the spring than in the summer season.

In this study, we use SAR to mask rice area in MODIS-based *LAI* maps. The question remains: how to derive rice area information to generate archive yield data, given that SAR data from Sentinel-1 were not available before 2014. This could be a potential source of uncertainty for the onward effort of generating historical yield data using the approach presented. Although the ultimate use of the information would be for aggregation to relatively coarser resolution (e.g., district level), the likely overall impact of this uncertainty in rice area data on the ultimate yield estimate products would be rather minimal. The latest rice area data, given that they are sufficiently demonstrated to be highly accurate and reliable, should be used in this case. Over time, when SAR data continue to accumulate, theoretically, the historical yield data will improve because more of the simulated yields will be based on the actual SAR-based rice area to mask MODIS-based *LAI* data from the same year. Another possible approach to generate “past rice cultivated area” to be considered representative for a given period (some years) is to fully exploit archive SAR data (e.g., ERS/ASAR); this approach was already demonstrated in tropical Asian conditions [57].

Although the synergetic method presented in this study is robust in generating spatial yield information, as indicated by good agreement with the high-resolution SAR-based yield and official yield data at district level aggregation, it is important to note the level of heterogeneity in the Red River Delta. Larger uncertainty values are expected in other geographies associated with sensor resolution relative to the size of homogeneous rice patches. This study can be extended to other rice growing regions given the global availability of both MODIS and Sentinel-1 SAR data. However, the first priority for implementing this study should be in geographies where Sentinel-1 provides regular SAR data acquisition with at least a 12-day temporal frequency in order to ensure reliable rice area estimates. Other datasets such as meteorological and soil data for the crop model are relatively straightforward to prepare with the source data available globally. Extra effort will be needed for local information gathering to collect agronomic management information and some limited ground data (rice area, *LAI*, and yield) for validation purposes. In areas where it is known to have large heterogeneity such as diverse planting dates across small spaces, higher resolution optical remote-sensing data such as Landsat and Sentinel-2 can be exploited.

Using spatial *LAI* and yield estimate results from this study supported by evidence from the in situ data, it will be useful to further investigate the possible reasons why some areas in one season (e.g., Hung Yen Province) or in both seasons (Hai Phong Province) showed relatively higher or lower maximum *LAI* and yield than other areas. Such an endeavor can reveal the underlying factors conforming to the outstanding performance of rice in certain areas or factors that limit attaining higher yield in other areas. This understanding can contribute to the effort of minimizing yield gaps and maximizing resource use efficiency to support sustainable rice production and improve our understanding on how biophysical and socioeconomic drivers affect spatial yield distribution.

## 7. Conclusions

This study highlights the importance of incorporating algorithms dedicated for Earth Observation (EO) in processing data with a biologically based function for *LAI* time series and local and expert knowledge of the area for successful integration of information for yield simulation using a crop growth model. In this study, we demonstrated a synergetic method of exploiting optical and SAR remote-sensing data in the context of assimilation of critical information such as rice area, start of season, and *LAI* for rice yield simulation. In the study areas, MODIS-based yield estimates performed relatively well in both rice growing seasons, showing a good agreement with both the SAR-based high-resolution approach and the official reported yield. Hence, the approach shows itself to be well suited for reconstructing historical yield data in the context of crop insurance programs. In addition to supporting risk reduction efforts by supporting crop insurance, the information from remote-sensing



technology can be used to identify and target hot-spot areas for interventions to reduce yield gaps and optimize resource use to ensure sustainable rice production.

**Acknowledgments:** Funding for this research was provided by the Swiss Agency for Development and Cooperation (SDC) under the Remote Sensing-based Information and Insurance for Crops in Emerging Economies (RIICE) project. The authors express their appreciation to Michael Anthony (sarmap) for management support for the project, Do Minh Phuong (NIAPP, Vietnam) for the provision of ground data from the Red River Delta of Vietnam, Gene Christopher Romuga (IRRI) for assistance in weather data preparation, Cornelia Garcia (IRRI) for GIS data preparation, and Rosendo Gutierrez (IRRI) for administrative support. SAR data were provided by the European Space Agency (ESA) from Sentinel-1A satellite. MODIS reflectance data were provided by the National Aeronautics and Space Administration (NASA) from Terra and Aqua satellites. Some maps in this manuscript are overlaid on Google Maps layers, © Google, 2016. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of the authors or their institutes concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

**Author Contributions:** T.D.S., E.Q., L.G., M.C.-T., L.B., J.G.-H., M.B., N.I.K., and F.H. wrote and edited the manuscript. T.D.S., M.C.-T., J.G.-H., L.G., and F.C. prepared and processed *LAI* data from remote-sensing data. L.B. processed MODIS time series to estimate rice phenology. T.D.S. and E.Q. ran yield simulations using the ORYZA crop growth model and *LAI* derived from remote-sensing data, and analyzed the outputs.

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

## References

1. Nagayet, O. Small Farms: Current Status and Key Trends. In *The Future of Small Farms, Proceedings of a Research Workshop, Wye, UK, 26–29 June 2005*; International Food Policy Research Institute: Washington, DC, USA, 2005; pp. 355–367.
2. O'Brien, K.; Leichenko, R.; Kelkar, U.; Venema, H.; Aandahl, G.; Tompkins, H.; Javed, A.; Bhadwal, S.; Barg, S.; Nygaard, L.; et al. Mapping vulnerability to multiple stressors: Climate change and globalization in India. *Glob. Environ. Chang.* **2004**, *14*, 303–313. [[CrossRef](#)]
3. Morton, J.F. The impacts of climate change on smallholder and subsistence agriculture. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19680–19685. [[CrossRef](#)] [[PubMed](#)]
4. Intergovernmental Panel on Climate Change. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E., Eds.; Cambridge University Press: Cambridge, UK, 2007.
5. Mendelsohn, R. The impact of climate change on agriculture in developing countries. *J. Nat. Resour. Policy Res.* **2009**, *1*, 5–19. [[CrossRef](#)]
6. Miranda, M.J.; Gonzalez-Vega, C. Systemic risk, index insurance, and optimal management of agricultural loan portfolios in developing countries. *Am. J. Agric. Econ.* **2011**, *93*, 399–406. [[CrossRef](#)]
7. Shannon, H.D.; Motha, R.P. Managing weather and climate risks to agriculture in North America, Central America and the Caribbean. *Weather Clim. Extremes* **2015**, *10*, 50–56. [[CrossRef](#)]
8. Swain, M. *Crop Insurance for Adaptation to Climate Change in India*; Asia Research Centre Working Paper; London School of Economics and Political Science: London, UK, 2014.
9. Miranda, M.J.; Farrin, K. Index insurance for developing countries. *Appl. Econ. Perspect. Policy* **2012**, *34*, 391–427. [[CrossRef](#)]
10. Carter, M.R.; Galarza, F.; Boucher, S. *Underwriting Area-Based Yield Insurance to Crowd-In Credit Supply and Demand*; Working Paper No. 07-003; Department of Agricultural and Resource Economics, University of California: Davis, CA, USA, 2007.
11. Dillon, C.R.; Mjelde, J.W.; McCarl, B.A. Biophysical simulation in support of crop production decisions: A case study in the Blacklands regions of Texas. *South. J. Agric. Econ.* **1989**, *21*, 73–86. [[CrossRef](#)]
12. Löw, F.; Biradar, C.; Fliemann, E.; Lamers, J.P.A.; Conrad, C. Assessing gaps in irrigated agricultural productivity through satellite earth observations—A case study of the Fergana Valley, Central Asia. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *57*, 118–134. [[CrossRef](#)]
13. Hamar, D.; Ferencz, C.S.; Lichtenberger, J.; Tarcsai, G.Y.; Ferencz Árkos, I. Yield estimation for corn and wheat in the Hungarian Great Plain using Landsat MSS data. *Int. J. Remote Sens.* **1996**, *17*, 1689–1699. [[CrossRef](#)]
14. Clevers, J.G.P.W. A simplified approach for yield prediction of sugar beet based on optical remote sensing data. *Remote Sens. Environ.* **1997**, *61*, 221–2280. [[CrossRef](#)]

15. Ferencz, C.; Bognár, P.; Lichtenberger, J.; Hamar, D.; Tarcsai, G.; Timár, G.; Ferencz Árkos, I. Crop yield estimation by satellite remote sensing. *Int. J. Remote Sens.* **2004**, *25*, 4113–4149. [[CrossRef](#)]
16. Doraiswamy, P.C.; Sinclair, T.R.; Hollinger, S.; Akhmedov, B.; Stern, A.; Prueger, J. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sens. Environ.* **2005**, *97*, 192–202. [[CrossRef](#)]
17. Prasad, A.K.; Chai, L.; Singh, R.P.; Kafatos, M. Crop yield estimation model for Iowa using remote sensing and surface parameters. *Int. J. Appl. Earth Obs. Geoinf.* **2006**, *8*, 26–33. [[CrossRef](#)]
18. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [[CrossRef](#)]
19. Koide, N.; Robertson, A.; Ines, A.; Qian, J.H.; DeWitt, D.; Lucero, A. Prediction of rice production in the Philippines using seasonal climate forecasts. *J. Appl. Meteorol. Clim.* **2013**, 552–569. [[CrossRef](#)]
20. Lansigan, F. Evaluating the effects of anticipated climate change on rice production in the Philippines. *J. Agric. Met.* **1993**, *48*, 779–782. [[CrossRef](#)]
21. Fang, H.; Liang, S.; Hoogenboom, G.; Teasdale, J.; Cavigelli, M. Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. *Int. J. Remote Sens.* **2008**, *29*, 3011–3032. [[CrossRef](#)]
22. Maki, M.; Sekiguchi, K.; Homma, K.; Hirooka, Y.; Oki, K. Estimation of rice yield by SIMRIW-RS, a model that integrates remote sensing data into a crop growth model. *J. Agric. Met.* **2017**, *73*, 2–8. [[CrossRef](#)]
23. Benedetti, R.; Rossini, P. On the use of NDVI profiles as a tool for agricultural statistics: The case study of wheat yield estimate and forecast in Emilia Romagna. *Remote Sens. Environ.* **2003**, *45*, 311–326. [[CrossRef](#)]
24. Fuller, D.O. Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987–1993. *Int. J. Remote Sens.* **1998**, *19*, 2013–2018. [[CrossRef](#)]
25. Zhang, X.; Zhang, Q. Monitoring interannual variation in global crop yield using long-term AVHRR and MODIS observations. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 191–205. [[CrossRef](#)]
26. Nelson, A.; Setiyono, T.; Rala, A.B.; Quicho, E.D.; Raviz, J.V.; Abonete, P.J.; Maunahan, A.A.; Garcia, C.A.; Bhatti, H.Z.M.; Villano, L.S.; et al. Toward an operational SAR-based rice monitoring system in Asia: Examples from 13 demonstration sites across Asia in the RIICE project. *Remote Sens.* **2014**, *6*, 10773–10812. [[CrossRef](#)]
27. Li, T.; Angeles, O.; Marcaida, M., III; Manalo, E.; Manalili, M.P.; Radanielson, A.; Mohanty, S. From ORYZA2000 to ORYZA (v3): An improved simulation model for rice in drought and nitrogen-deficient environments. *Agric. For. Meteorol.* **2017**, *237*, 246–256. [[CrossRef](#)] [[PubMed](#)]
28. Setiyono, T.D.; Holecz, F.; Khan, N.I.; Barbieri, M.; Quicho, E.; Collivignarelli, F.; Maunahan, A.; Gatti, L.; Romuga, G.C. Synthetic Aperture Radar (SAR)-based paddy rice monitoring system: Development and application in key rice producing areas in Tropical Asia. *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *54*, 1–15. [[CrossRef](#)]
29. Kang, M.G. *Innovative Agricultural Insurance Products and Schemes*; Agricultural Management, Marketing and Finance Occasional Paper 12; Food and Agriculture Organization of the United Nations: Rome, Italy, 2007.
30. Barnett, B.J.; Black, J.R.; Hu, Y.; Skees, J.R. Is area yield insurance competitive with farm yield insurance? *J. Agric. Resour. Econ.* **2005**, *30*, 285–301.
31. Saint, G. SPOT 4 VEGETATION system: Association with high resolution data for multiscale studies. *Adv. Space Res.* **1996**, *17*, 107–110. [[CrossRef](#)]
32. Justice, C.O.; Vermote, E.; Townshend, J.R.G.; Defries, R.; Roy, D.P.; Hall, D.K.; Salomonson, V.V.; Privette, J.L.; Riggs, G.; Strahler, A.; et al. The moderate resolution imaging spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1228–1249. [[CrossRef](#)]
33. Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC, 2015. Available online: <https://lpdaac.usgs.gov/node/844> (accessed on 20 March 2017).
34. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Progress Report RSC 1978-1; Remote Sensing Center, Texas A&M University: College Station, TX, USA, 1973.
35. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
36. Torres, R.; Snoeji, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.; Rommen, B.; Floury, N.; Brown, M.; et al. GMES Sentinel-1 mission. *Remote Sens. Environ.* **2012**, *120*, 9–24. [[CrossRef](#)]

37. Bouman, B.A.M.; Kropff, M.J.; Tuong, T.P.; Wopereis, M.C.S.; Ten Berge, H.F.M.; van Laar, H.H. *ORYZA2000: Modeling Lowland Rice*; International Rice Research Institute, Los Baños, Philippines, and Wageningen University and Research Centre: Wageningen, The Netherlands, 2001.
38. Anokwa, Y.; Hartung, C.; Brunette, W.; Borriello, G.; Lerer, A. Open source data collection in the developing world. *Computer* **2009**, *42*, 97–99. [[CrossRef](#)]
39. Confalonieri, R.; Foi, M.; Casa, R.; Aquaro, S.; Tona, E.; Peterle, M.; Boldini, A.; De Carli, G.; Ferrari, A.; Finotto, G.; et al. Development of an app for estimating leaf area index using a smartphone: Trueness and precision determination and comparison with other indirect methods. *Comput. Electron. Agric.* **2013**, *96*, 67–74. [[CrossRef](#)]
40. Boschetti, M.; Busetto, L.; Manfron, G.; Laborte, A.G.; Asilo, S.; Pazhanivelan, S.; Nelson, A. PhenoRice: A method for automatic extraction of spatio-temporal information on rice crops using satellite data time series. *Remote Sens. Environ.* **2017**, *194*, 347–365. [[CrossRef](#)]
41. Campos-Taberner, M.; García-Haro, F.; Camps-Valls, G.; Grau-Muedra, G.; Nutini, F.; Busetto, L.; Katsantonis, D.; Stavrakoudis, D.; Minakou, C.; Gatti, L.; et al. Exploitation of SAR and optical Sentinel data to detect rice crop and estimate seasonal dynamics of leaf area index. *Remote Sens.* **2017**, *9*, 248. [[CrossRef](#)]
42. Setiyono, T.D.; Weiss, A.; Specht, J.E.; Cassman, K.G.; Dobermann, A. Leaf area index simulation in soybean grown under near-optimal conditions. *Field Crop. Res.* **2008**, *108*, 82–92. [[CrossRef](#)]
43. Boschetti, M.; Nutini, F.; Manfron, G.; Brivio, P.A.; Nelson, A. Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping Systems. *PLoS ONE* **2014**, *9*, e88741. [[CrossRef](#)] [[PubMed](#)]
44. 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](#)]
45. 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 Part 1: Principles of the algorithm. *Remote Sens. Environ.* **2007**, *110*, 275–286. [[CrossRef](#)]
46. Rasmussen, C.E.; Williams, C.K. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, USA, 2006.
47. Campos-Taberner, M.; García-Haro, F.J.; Moreno, A.; Gilabert, M.A.; Sanchez-Ruiz, S.; Martinez, B.; Camps-Valls, G. Mapping leaf area index with a smartphone and Gaussian processes. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2501–2505. [[CrossRef](#)]
48. Setiyono, T.D.; Bastidas, A.M.; Cassman, K.G.; Weiss, A.; Dobermann, A. Nodal leaf area distribution in soybean plants grown in high yield environments. *Agron. J.* **2011**, *103*, 1–8. [[CrossRef](#)]
49. Attema, E.P.W.; Ulaby, F.T. Vegetation modeled as a water cloud. *Radio Sci.* **1978**, *13*, 357–364. [[CrossRef](#)]
50. Shen, S.H.; Yang, S.B.; Li, B.B.; Zhao, X.Y.; Tan, B.X.; Li, Z.Y.; Le Toan, T. Study on ENVISAT ASAR data assimilation in rice growth model for yield estimation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2009**, *37*, 951–956.
51. Setiyono, T.D.; Quicho, E.D.; Romuga, G.C. *Rice Yield Estimation (Rice-YES) User Manual*; International Rice Research Institute: Los Baños, Philippines, 2017.
52. Tan, J.; Cui, Y.; Luo, Y. Global sensitivity analysis of outputs over rice-growth process in ORYZA model. *Environ. Model. Softw.* **2016**, *83*, 36–46. [[CrossRef](#)]
53. Congalton, R.G. Accuracy assessment and validation of remotely sensed and other spatial information. *Int. J. Wildl. Fire* **2001**, *10*, 321–328. [[CrossRef](#)]
54. Janssen, P.H.M.; Heuberger, P.S.C. Calibration of process-oriented models. *Ecol. Model.* **1995**, *83*, 55–66. [[CrossRef](#)]
55. Campos-Taberner, M.; García-Haro, F.J.; Confalonieri, R.; Martínez, B.; Moreno, A.; Sánchez-Ruiz, S.; Gilabert, M.A.; Camacho, F.; Boschetti, M.; Busetto, L. Multitemporal monitoring of plant area index in the Valencia rice district with PocketLAI. *Remote Sens.* **2016**, *8*, 202. [[CrossRef](#)]

56. 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](#)]
57. Nelson, A.D.; Boschetti, M.; Manfron, G.; Holecz, F.; Collivignarelli, F.; Gatti, L.; Barbieri, M.; Villano, L.; Chandna, P.; Setiyono, T. Combining Moderate-Resolution Time-Series RS Data from SAR and Optical Sources for Rice Crop Characterisation: Examples from Bangladesh. In *Land Applications of Radar Remote Sensing*; Holecz, F., Pasquali, P., Milisavljevic, N., Closson, D., Eds.; InTech: Rijeka, Croatia, 2014.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).