

Article

# Assimilation of Remotely-Sensed LAI into WOFOST Model with the SUBPLEX Algorithm for Improving the Field-Scale Jujube Yield Forecasts

Tiecheng Bai <sup>1,2</sup> , Shanggui Wang <sup>1</sup>, Wenbo Meng <sup>1</sup>, Nannan Zhang <sup>1</sup>, Tao Wang <sup>1</sup>, Youqi Chen <sup>3</sup> and Benoit Mercatoris <sup>2,\*</sup> 

<sup>1</sup> Southern Xinjiang Research Center for Information Technology in Agriculture, Tarim University, Alaer 843300, China

<sup>2</sup> TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Liège University, Passage des Déportés, 2, 5030 Gembloux, Belgium

<sup>3</sup> Institute of Agricultural Resources and Regional Planning of CAAS, No.12 Zhongguancun South St., Haidian District, Beijing 100081, China

\* Correspondence: benoit.mercatoris@uliege.be

Received: 9 July 2019; Accepted: 17 August 2019; Published: 20 August 2019



**Abstract:** In order to enhance the simulated accuracy of jujube yields at the field scale, this study attempted to employ SUBPLEX algorithm to assimilate remotely sensed leaf area indices (LAI) of four key growth stages into a calibrated World Food Studies (WOFOST) model, and compare the accuracy of assimilation with the usual ensemble Kalman filter (EnKF) assimilation. Statistical regression models of LAI and Landsat 8 vegetation indices at different developmental stages were established, showing a validated  $R^2$  of 0.770, 0.841, 0.779, and 0.812, and a validated RMSE of 0.061, 0.144, 0.180, and 0.170  $m^2 m^{-2}$  for emergence, fruit filling, white maturity, and red maturity periods. The results showed that both SUBPLEX and EnKF assimilations significantly improved yield estimation performance compared with un-assimilated simulation. The SUBPLEX ( $R^2 = 0.78$  and  $RMSE = 0.64 t ha^{-1}$ ) also showed slightly better yield prediction accuracy compared with EnKF assimilation ( $R^2 = 0.73$  and  $RMSE = 0.71 t ha^{-1}$ ), especially for high-yield and low-yield jujube orchards. SUBPLEX assimilation produced a relative bias error (RBE, %) that was more concentrated near zero, being lower than 10% in 80.1%, and lower than 20% in 96.1% for SUBPLEX, 72.4% and 96.7% for EnKF, respectively. The study provided a new assimilation scheme based on SUBPLEX algorithm to employ remotely sensed data and a crop growth model to improve the field-scale fruit crops yield estimates.

**Keywords:** assimilation method; SUBPLEX; remote sensing; WOFOST model; jujube yield estimation

## 1. Introduction

Jujube (*Zizyphus jujuba*) is a significant economic tree species in China with approximately 3,250,000 hectares in 2017, and its fruit has important nutritional and medical value [1,2]. The field-scale jujube growth monitoring and yield estimation allow farmers to make management decisions, such as precision planting, irrigation optimization, fertilization, and pest management, which are also important components of precision agriculture and horticulture.

Assimilation of remote sensing information into crop models has been considered as a key technical tool for yield prediction. The goal of assimilation is to reduce the uncertainty in the spatial distribution of soil properties, crop parameters, and meteorological data of crop model applications in large areas [3]. The key input parameters or state variables, such as phenology information [4], leaf area index (LAI) [5–10], biomass [11], crop transpiration (ET) [7], and soil moisture (SM) [5,12–16], can

be observed from remote sensing data. In recent years, data assimilation methods, mainly including variational (calibration) and sequential (update) methods, have been used to integrate remote sensing data and crop models to improve the prediction accuracy of canopy state variables and yields at the field, regional, and national scale [17].

The variational method takes all available observations during the main growth season and attempts to fit the model to the observations by minimizing the cost function, thereby optimizing the initial parameters of crop models. The variational methods for remote sensing and crop model assimilation have been reported, including, shuffled complex evolution simplex algorithm (SCE-UA) [7,17–20], four-dimensional variational data assimilation (4DVAR) [9,21–24], particle swarm optimization (PSO) [25–30], Powell's conjugate direction method (Powell) [6,23,31,32], simplex search algorithm (SSA) [33,34], maximum likelihood solution (MLS) [35], golden section searching (GSS) [36], and annealing algorithm (AA) [21,24,37]. Canopy LAI is a state variable used by most studies because it directly reflects the growth of the crop. In addition, FAPAR [36,37], ET [7,38], leaf nitrogen accumulation [39], vegetation indices [21,25,31], and band reflectance [17] also demonstrate the potential as state variables for remote sensing assimilation to optimize initial parameters for crop models. The variational method attributes the model input, output, and the model's own error to the uncertainty of the initial conditions or the parameters of the model, and do not consider the state variable estimation error during the time evolution of the model parameters. Therefore, in practical applications, the assimilation accuracy of the variational method often depends on the quality and accuracy of external observation data.

Sequential methods directly update the state variables of a modeling system when observation becomes available. The magnitude of the state update then depends on the uncertainty in both the model state and the observation. Examples of sequential approaches are the ensemble Kalman filter (EnKF) [5,8,12–14,30,40–50], particle filter (PF) [51], constant gain Kalman filter (CGKF) [52,53], and ensemble square root filter (EnSRF) [15,54]. For the state variables of sequential methods, LAI is also the most focused, followed by soil moisture content (SM). The reported studies also show that the EnKF algorithm has been adopted by more research to improve assimilation accuracy.

EnKF continuously updates a new set of input parameters at each observation spot. If the state variable error statistic of remote sensing is a Gaussian distribution, the EnKF method can be considered as the preferred assimilation method because most crop growth models are nonlinear [55]. However, the EnKF approach can only work if the growth cycle (sowing date plus phenology) between the model and satellite observations in the field is correct. If the phenological information is wrong, the EnKF method is equivalent to forcing the LAI into a crop model that may belong to different growth stages and basically have a time shift between the reality and model, which may result in worse simulation accuracy [47]. If the phenology information is uncertain, the variational method is usually superior to the sequential method, which can reduce the accumulation and diffusion of remote sensing data errors in the assimilation process [3]. Of course, it also requires more computing time. The SUBPLEX method is based on the Nelder-Mead simplex algorithm (NMS), which determines an improved set of subspaces and then uses NMS to search each subspace [56]. For most applications, SUBPLEX shows higher computational efficiency for the unconstrained optimization of general multivariate functions than the simplex searching or forcing method [57]. In principle, SUBPLEX shall be one of the variational assimilation methods, which is different from EnKF assimilation process. It calculates a set of optimal input parameters based on the error of all remote sensing observations and simulated values. In addition, when the number of remote sensing observations is large, the SUBPLEX method can divide the observation spots into several lower-dimensional vectors, thereby improving computational efficiency. More importantly, for objective functions affected by remote sensing observations error, the measurement replication option of SUBPLEX can be used to avoid convergence to a false minimum [56]. Whether SUBPLEX algorithm has the potential to be applied to the assimilation of remote sensing and crop growth models also shows a valuable research exploration.

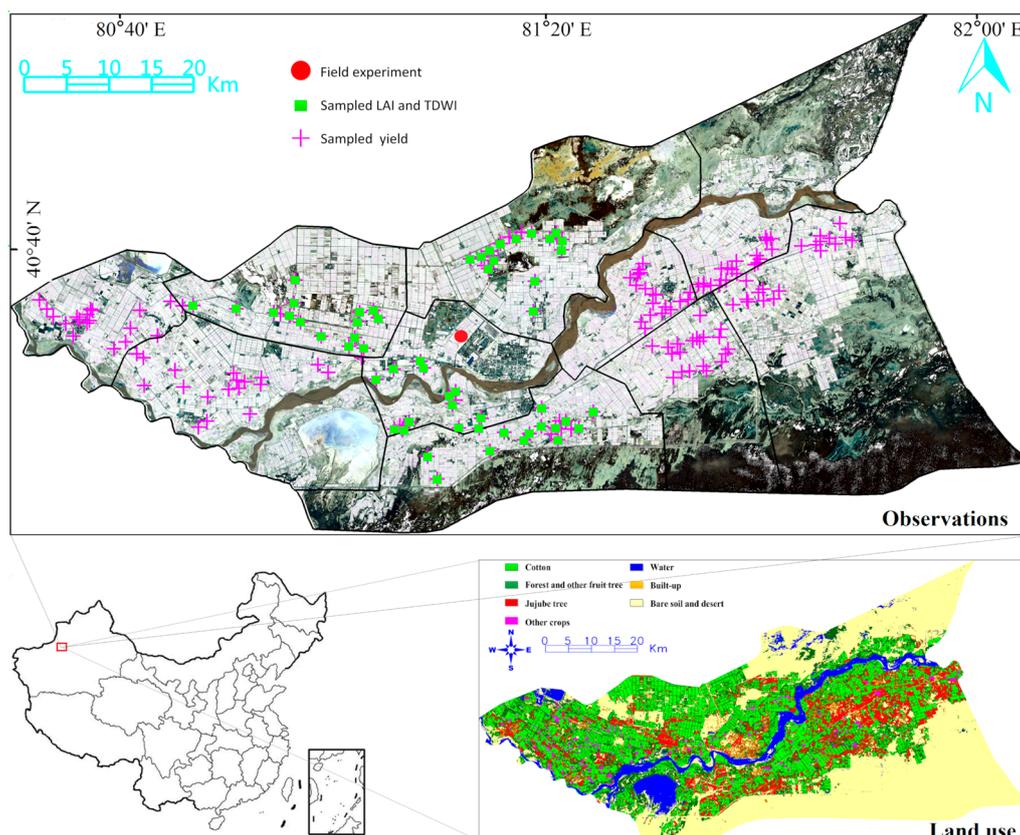
In addition, almost all remote sensing and crop growth model assimilation studies focus on annual crops, and few studies have discussed the growth simulation and assimilation of perennial fruit tree crops. Therefore, the objective of this study is to develop a data-assimilation framework based on an unconstrained optimization algorithm (SUBPLEX) that integrates remotely-sensed LAI into a calibrated WOFOST (World Food Studies) model to improve the jujube yield estimation at the field scale. To accomplish this goal, the following two specific objectives are defined:

- (1) To develop SUBPLEX algorithm to optimize the key input parameters of the WOFOST model for reducing the uncertainty.
- (2) To compare the yield prediction performance of the proposed SUBPLEX assimilation with the EnKF method.

## 2. Materials and Methods

### 2.1. Study Region

Figure 1 shows the study area, field experiment site, and the distribution of the field-scale observations for yields and LAIs. The red area in the land use map represents the jujube planting area. The study was carried out in an agricultural city with a large planting area of jujube trees (more than 50,000 ha), Alar City, Xinjiang, China. The study area covers 10 agro-ecological zones.



**Figure 1.** Study region and observed samples for this study.

The prevailing planting pattern of jujube trees in the study area is a low and dense planting pattern with a density of 2000 to 10,000 trees per hectare and a height of 1.5 to 2.5 m. The area is characterized by basins, sandy loam, and less rain, and arid warm-temperate climate. The water demand for jujube trees depends mainly on irrigation, which is managed by the government department. Generally, a reasonable irrigation strategy is carried out based on the temperature and rainfall conditions, combined with the water requirements of different development stages of jujube trees. Jujube tree

germinates in mid-April, and the fruit matures at the beginning of October and is harvested in early November.

## 2.2. Model and Data

### 2.2.1. WOFOST Model

WOFOST is a mechanistic model that explains crop growth on the basis of the underlying processes, such as phenological development, CO<sub>2</sub>-assimilation, transpiration, respiration, and how these processes are influenced by environmental conditions [58]. The potential production mode was carried out in the study because jujube in the agro-ecological zone does not usually suffer from water stress through adequate irrigation. Data collected in the two jujube orchards at the Tarim University Experiment Station were employed to calibrate WOFOST input parameters. The calibrated parameters were from five sources: Measured value, estimated value, calibrated value based on measured data, WOFOST default value, fine-tuning values around the default data. For a potential production calibration, length of the growth period and phenology were first executed, then light interception and potential biomass production, and finally assimilate distribution between crop organs. The values and details of the calibrated WOFOST model parameters for jujube growth simulation can be found in a previous study [59]. In response to the underestimation of high-yield areas, the fraction of above-ground dry matter to storage at DVS = 1.45 was increased to 0.9 (development stages, DVS).

### 2.2.2. Study Data

The observed data in this study mainly consisted of three parts, including field experiments, yield data of 181 independent orchards, and initial total crop dry weight (TDWI) and LAI data for 55 sample spots, see Table 1. In particular, the life span of leaves growing at 35 °C (SPAN) values for different subplots in the field experimental area were observed with a range of 41 to 62 days and an average of 53 days. It is very difficult to explain and verify the SPAN distribution at the regional scale [60]. Referring to the experience of existing research [9], the average SPAN of the field experiments was used to verify the simulated average SPAN of different orchards. Hundred eighty-one pieces of yield data were manually measured after harvesting in early November. Each orchard sample took up from 45 to 210 pixels in a Landsat 8 satellite image, with an average of 104. The 55 sample plots with different ages and planting densities were selected and monitored during the main growth season. LAI was monitored four times during the main growth season, which was employed to establish and validate the LAI inversion model based on remote sensing. In order to coincide with the time frequency of the Landsat satellite, LAI was measured on the same day when the satellite covered study area, using a fruit tree canopy analyzer (TOP-1300, Zhejiang Top Cloud-Agri Technology Co., Ltd., Zhejiang, China) during four major developmental stages, including emergence, flowering, white maturity, and red maturity. The TOP1300 uses the principle that canopy porosity is related to the canopy structure. It is based on Beer's law that light is weakened through the medium, and the theoretical semi-empirical formula is used to calculate the canopy structure parameters by porosity. It consists of a small fisheye camera, a measuring rod, and a microcomputer. The fisheye camera can penetrate the different horizontal and vertical heights of the canopy and quickly perform a layered measurement of the vertical distribution of the leaf area. It was corrected by manual measurement before measurement. Each sample plot consisted of four relatively homogeneous subplots (30 m × 30 m). Average LAI values from the four subplots were calculated to represent the unique LAI value in each sample plot. The value of TDWI was measured when the fifth leaf on the bud was unfolded, which was employed to evaluate the optimization performance of SUBPLEX algorithm. The definition and calculation of TDWI refer to previous research [59].

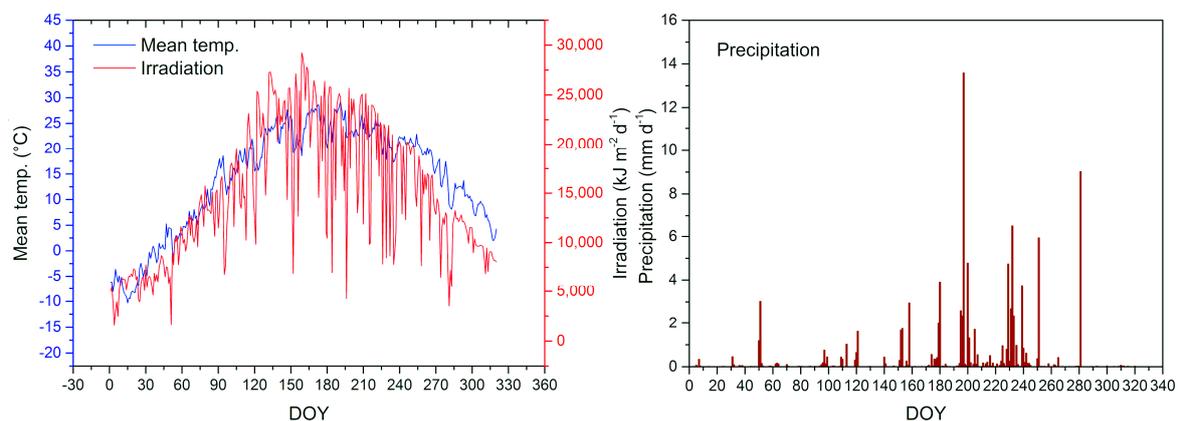
The WOFOST model was driven by meteorological data. The prepared meteorological data for the WOFOST model included irradiation, minimum temperature, maximum temperature, early morning vapor pressure, mean wind speed, and precipitation. Based on the theoretical basis of the WOFOST

model, the rate and stage of crop development are calculated by the average daily temperature [61]. Figure 2 contains daily average temperature, total radiation, and precipitation data for 2017 in our study. The average temperature data show that there were normal temperature conditions in 2017, with small fluctuations. The average radiation in the study area was strong and showed large fluctuations. The average daily temperature and total radiation during the main growing season of 2017 were 22.7 °C and 19,683 KJ m<sup>-2</sup> d<sup>-1</sup>, respectively, with a total rainfall of 79 mm.

**Table 1.** Observed field data.

Samples	Sampled Date	Max	Min	Average	STDEV.P
Yields of 181 samples	1–15 November	10.75	4.79	7.74	2.43
TDWI of 55 samples	25 April to 5 May	24.06	5.69	13.45	7.53
LAI of 55 samples	9 June	0.78	0.27	0.46	0.12
	27 July	2.51	1.03	1.76	0.37
	12 August	2.43	0.89	1.70	0.40
	28 August	2.35	0.81	1.62	0.41

Yield: t ha<sup>-1</sup>, TDWI: Total crop dry weight (kg ha<sup>-1</sup>), LAI: Leaf area index (m<sup>2</sup> m<sup>-2</sup>).



**Figure 2.** Daily mean temperature, irradiation, and precipitation of study area in 2017.

The four pieces of Landsat 8 satellite data with a spatial resolution of 30 m during the 2017 main growth season, including emergence, fruit filling, and maturity periods were acquired from the United States Geological Survey (USGS). The dates of the available remote sensing images are 9 June, 27 July, 12 August, and 28 August, respectively. Band 4 (red, 0.630–0.680 μm) and 5 (near-infrared, 0.845–0.885 μm) from the Operational Land Imager (OLI) were used in this study for vegetation index extraction and establishing LAI regression model. The geometric and atmospheric corrections were carried out for four available Landsat 8 data by reference to the Albers conical equal-area map projection using 50 field-measured ground control points and the fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) model, respectively. Their parameter settings refer to an existing study [62].

### 2.3. Remotely Sensed LAI

Since the spectral information of the crop was strongly influenced by the soil background, a soil-adjusted vegetation index (SAVI) may be a suitable choice to establish a statistical relationship during this period before the soil was covered by vegetation [63]. Referring to the research method of Huang et al. [9], the LAI inversion models at four different developmental stages were established separately. Furthermore, the accuracy of the LAI statistical regression model based on remotely sensed NDVI and SAVI was compared to determine a suitable index. Thirty-seven of the 55 samples were

used to calibrate the LAI model and the remaining 18 samples were employed to validate the model. SAVI [63] and NDVI [64], can be calculated by using Equations (1) and (2), respectively:

$$\text{SAVI} = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red} + L} \times (1 + L) \quad L = 0.5 \quad (1)$$

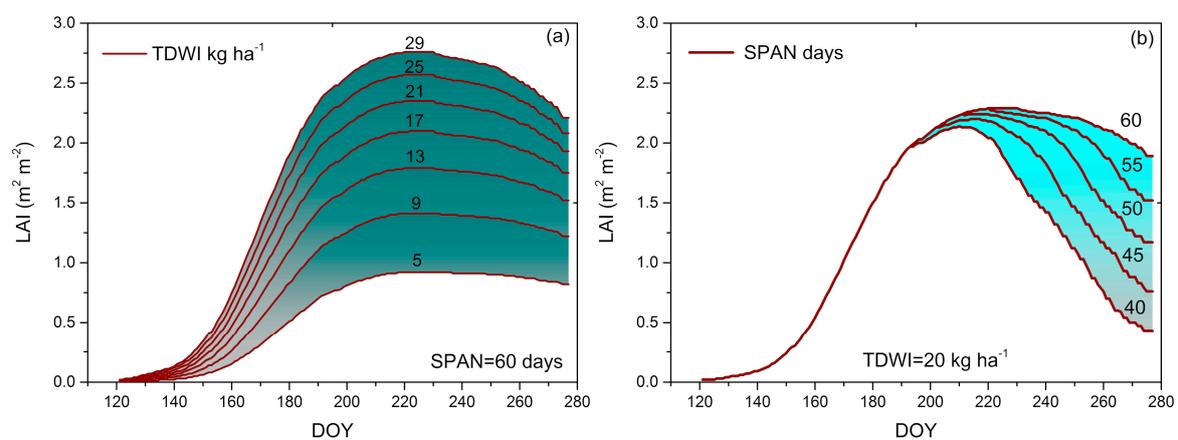
$$\text{NDVI} = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \quad (2)$$

## 2.4. Assimilation Strategy

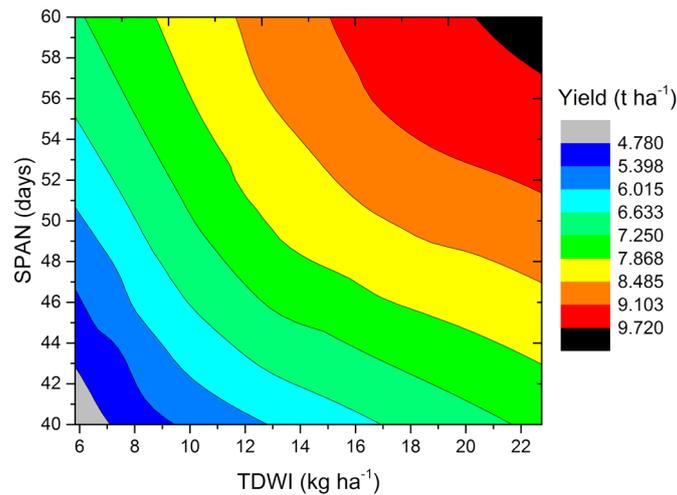
### 2.4.1. Selection of Reinitialized Parameters for WOFOST

In the WOFOST model, TDWI is equal to the dry weight sum of roots, stems, and leaves and when the crop is germinated. TDWI strongly influences the initial growth rate of the LAI and biomass, and expresses significant uncertainty in regional applications of the WOFOST model [9]. A slightly higher TDWI value will distribute more dry matter to the leaves to produce a larger LAI, which, in turn, increases photosynthesis and yield. SPAN parameter determines the rate and timing of leaf senescence, which is also affected by nutrients, pests, and diseases in practice, but WOFOST does not consider these aspects of crop growth [48]. SPAN mainly affects the downward trend of LAI in the later growing season, thereby affecting the final yield. Both two parameters can change the trajectory of the LAI during the main growth season and the maximum LAI achieved by the WOFOST model.

Previous studies have indicated that TDWI and SPAN parameters show significant uncertainty in regional applications of crop models. After optimizing these two parameters by the assimilation method, the prediction accuracy of crop yield was significantly improved [9,60]. In particular, TDWI of perennial jujube trees was strongly influenced by tree age and planting density, showing uncertainty in the same area [59]. The trends in the impact of TDWI and SPAN on the jujube LAI produced by the WOFOST model were shown in Figure 3a,b. The initial growth rate and maximum leaf area index increased with the increase of TDWI, but the growth rate was gradually decreasing. SPAN determined the rate at which green leaves will turn brown, and therefore, affected leaf senescence rate and effective green leaf index in the late growing season [60]. In addition, SPAN explained to some extent the effects of nutritional stress, insect and disease factors on crop growth and yield [9]. Thus, TDWI and SPAN input parameters were selected and recalibrated by assimilating remotely sensed LAI. The simulated jujube yield was affected by the TDWI–SPAN joint distributions (Figure 4).



**Figure 3.** Evolution of the simulated LAI profiles produced by the WOFOST model (a) LAI versus TDWI (b) LAI versus the life span of leaves growing at 35 °C (SPAN).



**Figure 4.** The effect of TDWI–SPAN joint distributions on yield.

#### 2.4.2. Assimilation Methods

The key innovation of this research was using the SUBPLEX method [56] to achieve the assimilation of the remotely sensed LAI into the WOFOST model. The objective function calculator ran the WOFOST model with the given set of input parameters, collected the simulation results, and computed the differences with the observations. Different objective functions can be selected. In this study, the objective function  $f(x)$  for SUBPLEX was constructed as the root of the mean squared error (RMSE), see Equation (3). Relative tolerance for convergence ( $\varepsilon$ ) determines the threshold at which the objective function converges. In theory, a smaller  $\varepsilon$  value will result in better assimilation accuracy. However, lower tolerance value will require more function evaluations. In this research, based on LAIs of 55 field measurements, the minimum objective function of LAIs for the four periods ranged from 0.01 to 0.27, with an average of 0.1. The test results also showed that when the  $\varepsilon$  value was reduced from 0.05 to 0.01, the coefficient of determination and accuracy of the yield prediction were only slightly improved, so in order to keep the calculation efficiency, the  $\varepsilon$  value was set to 0.05. When  $f(x)$  was less than  $\varepsilon$  (0.05), the assimilation process was ended.

$$f(x) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{s,i} - x_{o,i})^2} \quad (3)$$

where  $x_{s,i}$  and  $x_{o,i}$  represent the simulated and observed LAI values of the  $i$ th sample.  $n$  is the number of observation spots.

The key to implementing the SUBPLEX method is to set the stepsizes and subspaces, and then use a simplex searching algorithm NMS to search each subspace to perform an inner minimization. The stepsizes, stores in the vector step, determine both the scale and orientation of the initial simplex used in the inner minimizations. For the first cycle, step is equal to the initial step. Then, step is rescaled according to how much progress was made during the previous cycle. The rescaled step is shown in Equation (4):

$$\text{step} = \begin{cases} \min(\max(\|\Delta x\|_1 / (\|\text{step}\|_1, \omega), 1/\omega) \cdot \text{step} & \text{if } n_{\text{subs}} > 1 \\ \varphi \cdot \text{step} & \text{if } n_{\text{subs}} = 1 \end{cases} \quad (4)$$

where  $\Delta x$  represents the difference of observed and simulated LAI after successive iterations.  $n_{\text{subs}}$  represents the number of subspaces. When there is only one subspace, the factor  $\varphi$  that represents the simplex reduction coefficient is used to reduce the size of the simplex for the same step amount. If the  $\varphi$  value is reduced, the subspace search becomes more accurate.  $\omega$  represents the step reduction

coefficient, which is employed to control the degree to which step can be modified. A smaller  $\omega$  value can usually converge quickly to a local minimum. Conversely, a larger  $\omega$  value will reduce the rate of convergence, but a more comprehensive search can be achieved to obtain a smaller objective function value. Next, the  $i$ th component of step was reset by Equation (5):

$$\text{step}_i = \begin{cases} \text{sign}(\Delta x_i) \cdot |\text{step}_i| & \text{if } \Delta x_i \neq 0 \\ -\text{step}_i & \text{if } \Delta x_i = 0 \end{cases} \quad (5)$$

To set subspaces, the relationship between  $n_{\text{subs}}$  and the subspace dimensions  $n_{s_i}$  was shown in Equation (6):

$$\sum_{i=1}^{n_{\text{subs}}} n_{s_i} = n \quad (6)$$

And  $n_{\text{min}} \leq n_{s_i} \leq n_{\text{max}}$ , for  $i = 1, \dots, n_{\text{subs}}$ .  $n_{\text{min}}$  and  $n_{\text{max}}$  represent the minimum and maximum subspace dimensions, respectively.

And  $1 \leq n_{\text{min}} \leq n_{\text{max}} \leq n$  and  $n_{\text{min}} \lceil n/n_{\text{max}} \rceil \leq n$ .

The first step in determining the subspaces was to sort the components of the vector of progress by decreasing magnitude. The vector of progress was denoted by Equation (7):

$$\Delta x = (\Delta x_1, \dots, \Delta x_n)^T \quad (7)$$

Next, sorting  $\Delta x$ , saw Equation (8):

$$\widetilde{\Delta x} = (\Delta x_{p_1}, \dots, \Delta x_{p_n})^T \quad (8)$$

where  $|\Delta x_{p_i}| \geq |\Delta x_{p_{i+1}}|$ .

Specifically, the first subspace dimension  $n_{s_1}$  was defined by Equation (9):

$$n_{s_1} = \begin{cases} \frac{\|(\Delta x_{p_1}, \dots, \Delta x_{p_k})^T\|_1}{k} - \frac{\|(\Delta x_{p_1}, \dots, \Delta x_{p_n})^T\|_1}{n-k} & \text{if } k < n \\ \frac{\|(\Delta x_{p_1}, \dots, \Delta x_{p_n})^T\|_1}{n-k} & \text{if } k = n \end{cases} \quad (9)$$

where  $n_{\text{min}} \leq k \leq n_{\text{max}}$  and  $n_{\text{min}} \lceil (n-k)/n_{\text{max}} \rceil \leq n-k$ .

The first constraint forced  $n_{s_1}$  to be in the proper range, and the second constraint guarantees that the remaining  $(n - n_{s_1})$  vector can be partitioned. The process was repeated to determine  $n_{s_2}$ ,  $n_{s_3}$ , etc.

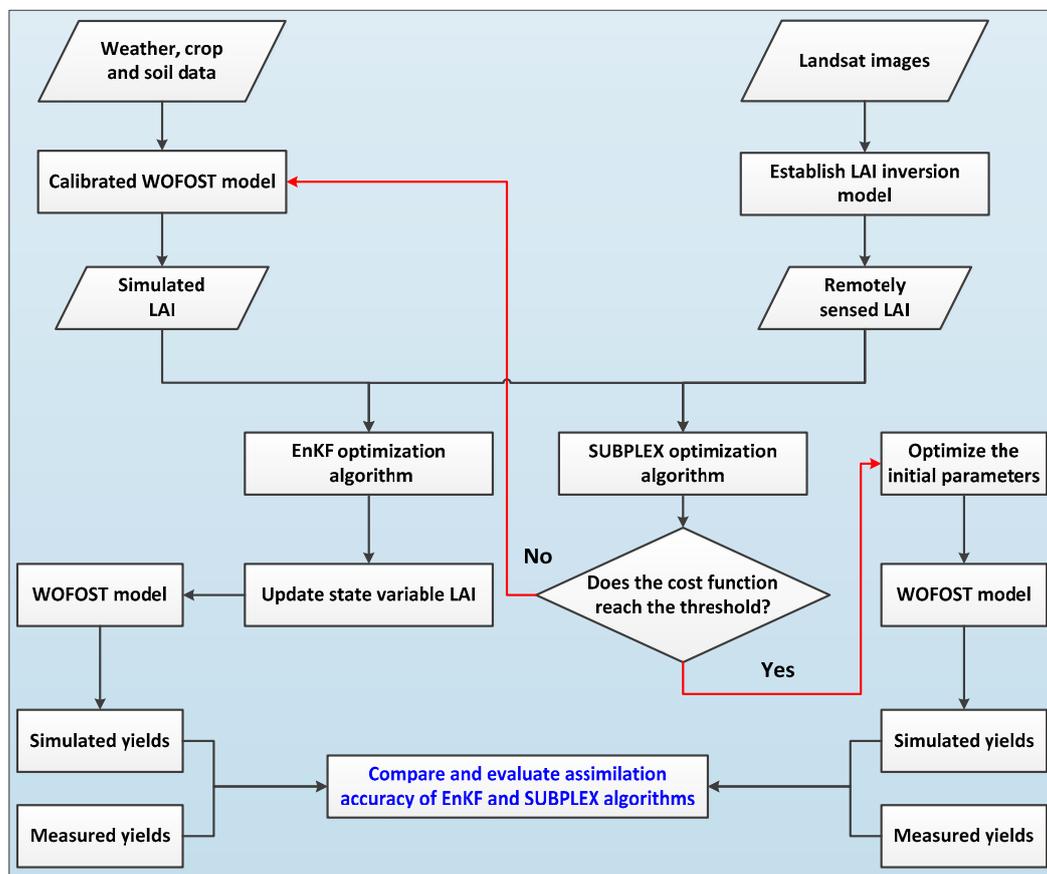
After the stepsizes and subspaces are set, the NSM algorithm is used to search each subspace to minimize the cost function. The setting and selecting parameters for data assimilation are shown in Table 2. TDWI\_range and SPAN\_range were set based on the range of TDWI and SPAN for field measurements. The  $\epsilon$  value was set with reference to the minimum value that the objective function  $f(x)$  of all observation spots can reach. The initial step was set to the expected value of an artificial setting according to the actual meaning of the parameter to be optimized, which can be further optimized by using the SUBPLEX algorithm at the iteration process. In theory, smaller  $\varphi$  and larger  $\omega$  values can produce smaller objective function values, but the computational efficiency will decrease accordingly. The actual test also showed that when  $\varphi$  was less than a usual value of 0.25 [56] and  $\omega$  was greater than a usual value of 0.1 [56], the minimized objective function value was not significantly improved, so  $\varphi$  and  $\omega$  were set to 0.25 and 0.1, respectively.

Figure 5 shows the assimilation flowchart for the SUBPLEX and EnKF methods. For open-loop simulation, orchards of the same age are set to the same TDWI values based on measurements from different orchards. The TDWI values of jujube trees aged 4 to 10 are equal to 6.27, 9.25, 13.25, 14.28, 16.29, 19.77, and 21.62 kg ha<sup>-1</sup>, respectively. SPAN is equal to 50, which is the average of calibrated SPAN based on the SUBPLEX method. The assimilation details of the EnKF method can be found in an existing study [47]. The ensemble size was set to 50 because the previous research showed that this

scale can achieve good assimilation performance [47]. EnKF updated a new set of TDWI and SPAN values when acquiring a new remote sensing observation spot to implement segmentation simulation. In this study, there were four remotely sensed LAIs, so TDWI and SPAN were updated four times by EnKF. SUBPLEX only obtained an optimal set of TDWI and SPAN values by iteratively calculating the minimum objective function values of the four remotely sensed and the simulated LAIs during the main growth stages, thereby achieving yield prediction based on the SUBPLEX assimilation method.

**Table 2.** SUBPLEX assimilation settings.

Name	Description	Set Value in This Study
$n$	the number of assimilated LAI	4
$f$	function to be minimized	Equation (3)
$m$	the number of optimized parameters for WOFOST	2 (TDWI and SPAN)
$\varphi$	simplex reduction coefficient ( $0 < \varphi < 1$ )	0.25
$\omega$	step reduction coefficient ( $0 < \omega < 1$ )	0.1
$\varepsilon$	relative tolerance for convergence	0.05
TDWI_range	the range of TDWI	5–30
SPAN_range	the range of SPAN	40–60
max-evaluation	maximum number of evaluations allowed	200
Initial step	the initial step size to compute numerical gradients	0.5 for TDWI 1 for SPAN



**Figure 5.** Assimilation flowchart.

### 3. Results

#### 3.1. Remotely Sensed LAI

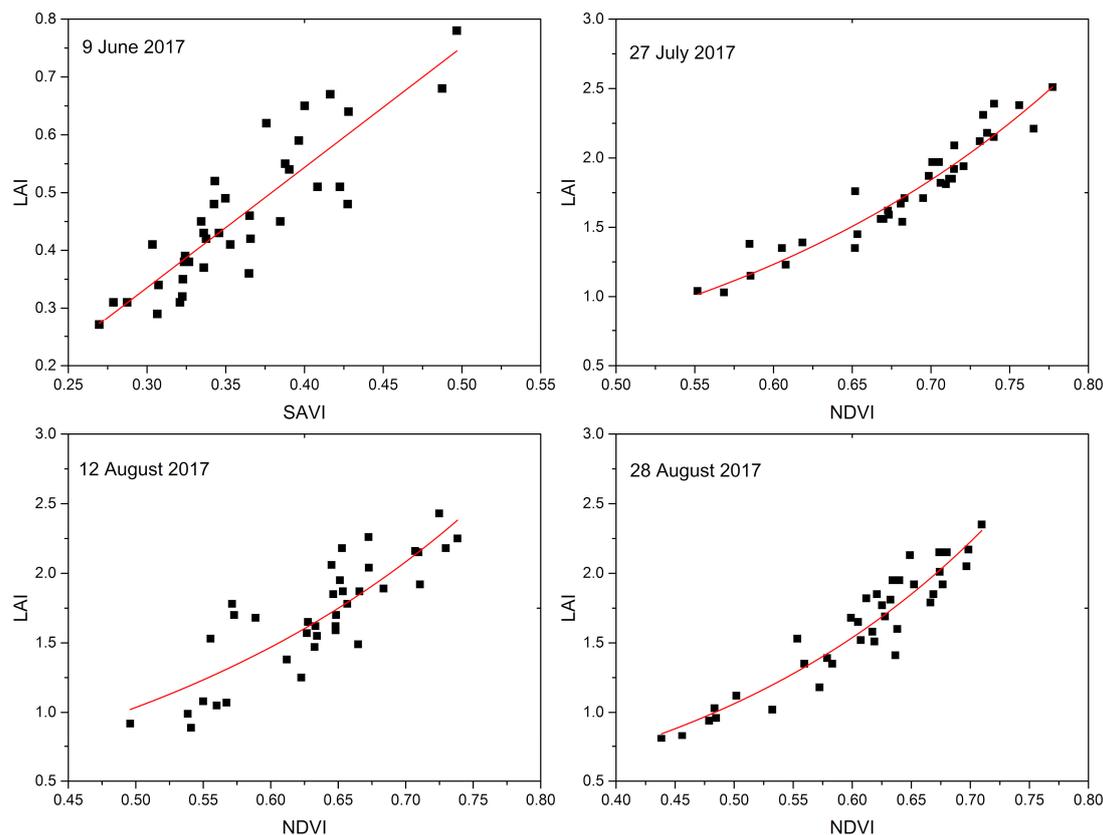
The statistical regression models between vegetation index (NDVI or SAVI) and LAI for the four jujube key growth periods are shown in Equations (10)–(13), respectively, which were obtained from regressions of 37 field-measured LAI and SAVI or NDVI (Figure 6). The difference between the 37 samples for each period was extremely significant ( $p < 0.001$ ). On 9 June between emergence and flowering, SAVI was used to establish the LAI linear regression model, while in other development stages NDVI was employed to invert LAI (exponential model) for obtaining a more accurate model. Previous studies also showed that SAVI exhibited better LAI statistical regression performance before the soil was covered by vegetation. NDVI performed better when vegetation coverage was higher [9].

$$9 \text{ June } 2017 : \text{LAI} = 2.08112 \times \text{SAVI} - 0.2889 \quad p < 0.001 \quad (10)$$

$$27 \text{ July } 2017 : \text{LAI} = 0.11071 \times e^{4.01646 \times \text{NDVI}} \quad p < 0.001 \quad (11)$$

$$12 \text{ August } 2017 : \text{LAI} = 0.18101 \times e^{3.48957 \times \text{NDVI}} \quad p < 0.001 \quad (12)$$

$$28 \text{ August } 2017 : \text{LAI} = 0.16731 \times e^{3.69701 \times \text{NDVI}} \quad p < 0.001 \quad (13)$$



**Figure 6.** LAI inversion models for four phenological development dates.

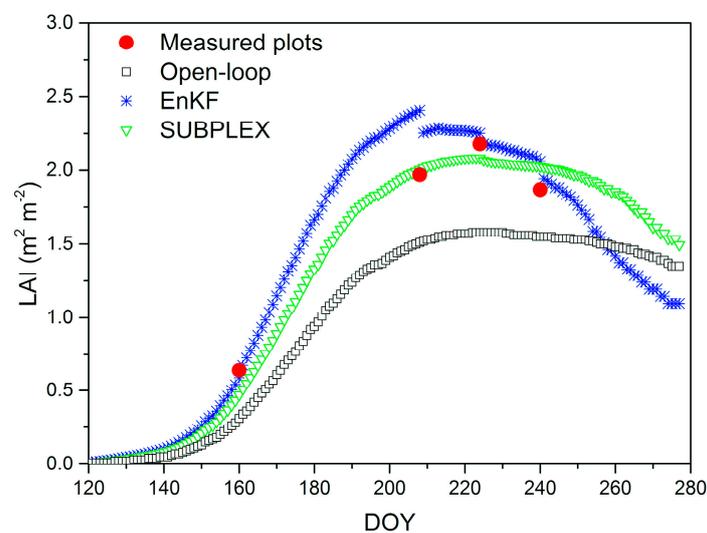
Table 3 shows the detailed results of calibration and validation LAI inversion models at four growth periods. The average  $R^2$  for calibration and validation was 0.822 and 0.801, respectively, with a range from 0.710 to 0.918. The calibrated and validated average % error was 9.9% and 10.4%, respectively. Except for 12 August 2017, the consistency and accuracy of correction sets at other days were slightly higher than the validation sets. The verification results show that the established model accurately expressed the LAI values of different phenological development stages.

**Table 3.** The comparison of LAI inversion models for four key periods.

Date	Calibrated R <sup>2</sup>	Calibrated RMSE (%) m <sup>2</sup> m <sup>-2</sup>	Validated R <sup>2</sup>	Validated RMSE (%) m <sup>2</sup> m <sup>-2</sup>
09 June 2017	0.774	0.058 (12.7)	0.770	0.061 (12.5)
27 July 2017	0.918	0.108 (6.2)	0.841	0.144 (8.1)
12 August 2017	0.710	0.229 (13.6)	0.779	0.180 (10.5)
28 August 2017	0.884	0.115 (7.1)	0.812	0.170 (10.4)

### 3.2. Assimilation Process

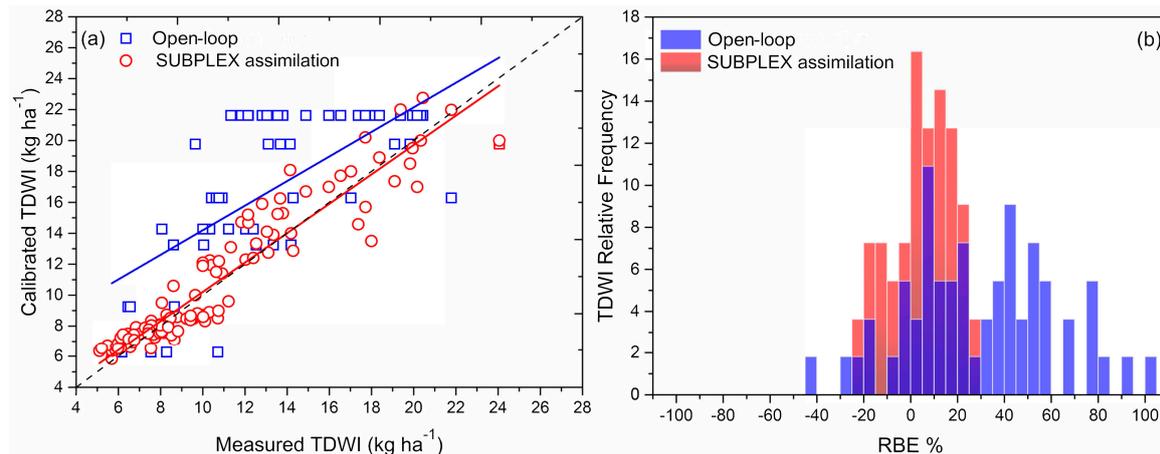
Taking a sample as an example, the LAI observed in the four key growth stages was 0.64, 1.97, 2.18, 1.87 m<sup>2</sup> m<sup>-2</sup>, respectively. For EnKF assimilation, the blue line represents the simulated state variable (Figure 7), here LAI, which was brought forward in time until an observation was available (red dot). At this point, an analysis step was performed to adjust the state of the model based on the observed LAIs, which resulted in a “jump” in the simulated state. The model was then advanced in time until the next observation was reached, and the process was repeated. The green line demonstrates the principle of a variational method of SUBPLEX. The model was first run without data assimilation (first guess). Next, all observations with an assimilation window were collected to adjust the model until it better matched the observations (analysis). In this process, model parameters or other properties were adjusted to minimize the cost function. After 47 iterations, a combination of TDWI and SPAN was screened (TDWI = 17, SPAN = 57), with a small difference against measured values (TDWI = 15.98, SPAN = 55). It is noted that Figure 7 only shows a sample to illustrate the difference in the assimilation process of the two methods for all 181 samples of which in 100 samples, the yield prediction accuracy of the SUBPLEX method was superior to the EnKF method.

**Figure 7.** SUBPLEX versus ensemble Kalman filter (EnKF) assimilation process.

### 3.3. SUBPLEX Assimilation Evaluation Based on the Field-Measured LAI

The 55 field-measured LAIs were assimilated into the WOFOST model to verify the performance of the reinitialized input parameters. For TDWI parameter, when the remotely sensed LAI was not assimilated, the average TDWI value of the same aged orchards was input to the WOFOST model. Scatter plots for re-calibrated TDWIs based on SUBPLEX assimilation versus measured 55 TDWIs are shown in Figure 8a, with an R<sup>2</sup> of 0.86. After SUBPLEX assimilation, the accuracy of the corrected TDWI with an RMSE of 1.88 kg ha<sup>-1</sup> was significantly higher than that of the unassimilated LAI (RMSE = 5.09 kg ha<sup>-1</sup>), and the calibrated average TDWI value of 13.91 kg ha<sup>-1</sup> showed an error of 3.3% compared with field-measured average value of 13.46 kg ha<sup>-1</sup>. The samples with a prediction error

between  $-10\%$  and  $10\%$ ,  $-20\%$  and  $20\%$  accounted for  $41.6\%$  and  $83.6\%$ , respectively (Figure 8b). These values showed better performance for TDWI calibration than open-loop simulation with  $21.8\%$ ,  $36.4\%$ , and  $49\%$  of  $-10\%$ – $10\%$ ,  $-20\%$ – $20\%$  and  $-40\%$ – $40\%$ , respectively. RBE values after SUBPLEX assimilation were distributed more centrally around zero compared with the values with open-loop simulation. For SPAN parameter, validating its spatial distribution was difficult due to a lack of corresponding measured data for 55 samples. However, the average value after assimilation with 49.95 days showed an error of  $5.8\%$  compared with the field experimental mean value of 53 days. The results also suggested that SUBPLEX assimilation reduced the uncertainty of TDWI and SPAN input parameters to some extent.



**Figure 8.** (a) Scatter plots for re-calibrated versus measured TDWI. (b) Frequency distribution for RBE based on recalibrated and measured TDWI. Bin size = 5.

Based on field-measured LAI, the yield prediction performance of two assimilation methods is presented in Table 4. Both SUBPLEX and EnKF significantly improved yield prediction accuracy compared with the unassimilated results, with higher  $R^2$  and lower RMSE values. The results also show that the SUBPLEX assimilation accuracy ( $R^2 = 0.86$ ,  $RMSE = 0.55$  ( $7\%$ )  $t\ ha^{-1}$ ) was slightly higher than EnKF ( $R^2 = 0.81$ ,  $RMSE = 0.65$  ( $8.2\%$ )  $t\ ha^{-1}$ ), significantly higher than open-loop simulation. The WOFOST-simulated yields with open-loop simulation overestimated the actual average yield, with an error of  $4.3\%$ . SUBPLEX and ENKF assimilations produced only  $2.7\%$  and  $2.8\%$  prediction errors for average yield, respectively, indicating better performance. For maximum yield forecast, EnKF was slightly more accurate, while for minimum production forecast, SUBPLEX was slightly better. However, both methods showed slightly higher errors in the maximum yield prediction with a value of about  $7.6\%$ . The jujube yield prediction using the field-measured LAI assimilation into the WOFSOT model indicated that SUBPLEX had the potential to be an effective assimilation method and demonstrated better accuracy than the EnKF method.

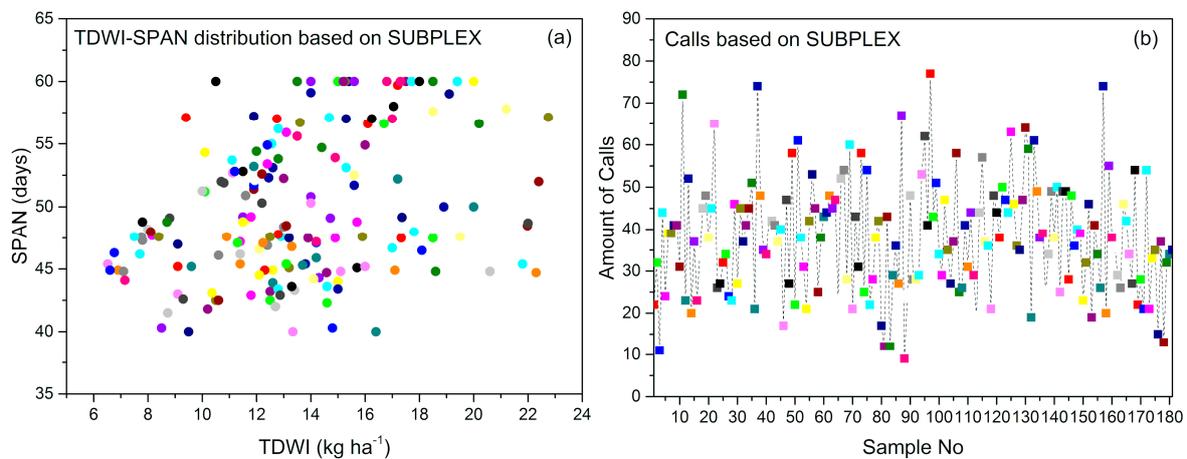
**Table 4.** Comparison of the estimated yield at the field scale based on field-measured LAI.

	Mean $t\ ha^{-1}$	Maxmium $t\ ha^{-1}$	Minmium $t\ ha^{-1}$	$R^2$	RMSE (%) $t\ ha^{-1}$
Field-measured yield at the 55 samples	7.931	10.71	4.848	–	–
Open-loop simulation	8.271	8.89	5.946	0.58	0.95 (12.1)
EnKF assimilaiton	7.717	9.887	5.303	0.81	0.65 (8.2)
SUBPLEX assimilaiton	7.707	9.713	4.791	0.86	0.55 (7.0)

### 3.4. Evaluation of Model Performance Based on Remotely Sensed LAI

After carrying out SUBPLEX assimilation, the joint distribution of TDWI and SPAN was generated for 181 samples to provide variability in parameter distribution between different samples (Figure 9a).

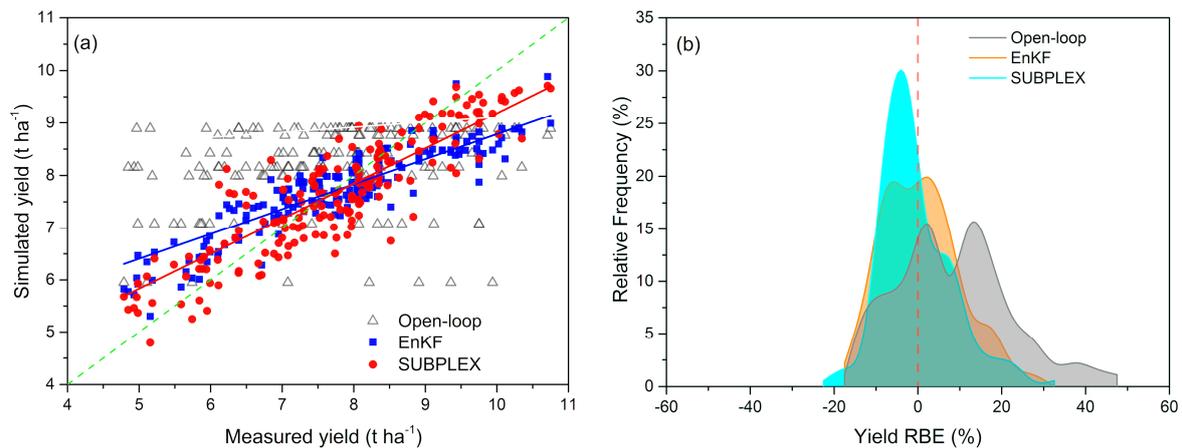
Explaining the shape of the distribution was difficult because they were the result of many interaction factors [60], such as meteorological conditions, pests and diseases, and nutritional stress. However, TDWI values recalibrated by SUBPLEX assimilation ranged from 5.84 to 22.75 kg ha<sup>-1</sup>, showing high variability between orchards. Assimilating remotely sensed LAI achieved good TDWI correction for 181 samples, with an average value of 13.46 kg ha<sup>-1</sup> compared with the field-measured mean value of 13.45 kg ha<sup>-1</sup> (55 samples). SPAN values ranged from 40 to 60 days with an average of 49.73 days, which showed an error of 6.2% versus the measured mean value at the field experiments with 53 days. Figure 9b shows that the SUBPLEX algorithm used fewer function calls (iteration times) to complete data assimilation, ranging from 9 to 77 with an average of 40, at relative tolerance  $\varepsilon = 0.05$ . The low number of iterations also indicates that the SUBPLEX algorithm showed a relatively high computational efficiency to some extent.



**Figure 9.** (a) TDWI-SPAN distribution for 181 samples. (b) Number of iterations for each sample.

The scatter plots of the simulated yields achieved from the unassimilated, EnKF, and SUBPLEX assimilations versus actual yields are shown in Figure 10a. The WOFOST-estimated yield with open-loop simulation exhibited a poor spatial distribution and prediction performance, clearly deviating from the line  $y = x$  and 71.8% of the samples were severely overestimated. Both EnKF and SUBPLEX assimilations significantly improved yield simulation performance with a scatter distribution closer to the line  $y = x$ . Although the dispersion of yield forecasts achieved by the SUBPLEX method was higher than that of ENKF, SUBPLEX showed better performance in high-yield and low-yield areas. Compared with the simulation without the assimilation and EnKF method, the SUBPLEX assimilation produced a relative bias error (RBE, %) that was more concentrated near zero (Figure 10b). The absolute RBE for yield simulation was lower than 10% in 80.1% and lower than 20% in 96.1% for SUBPLEX, 72.4% and 96.7% for EnKF.

The positive results achieved by SUBPLEX or EnKF assimilations were also confirmed by the indices of agreement and detailed prediction error (Table 5). The WOFOST simulation with open-loop simulation showed lower yield prediction accuracy, being an  $R^2$  of 0.39, RMSE of 1.06 (13.8%) t ha<sup>-1</sup>, MAE (mean absolute error, %) value of 12.5% and average RBE of 8.65%, respectively. Both SUBPLEX and EnKF methods improved the simulated accuracy, with a much higher coefficient of determination ( $R^2 \geq 0.73$ ) and lower error (RMSE  $\leq 9.2$  t ha<sup>-1</sup>). In addition, average RBE value from two assimilation methods agreed with the results shown in Figure 10 with a lower value, which shows that the error distributions were uniform to some extent. Furthermore, the SUBPLEX also indicated slightly better assimilation performance with an improvement of  $R^2$  (increased by 6.9%), RMSE (increased by 9.9%), and MAE (increased by 11.4%) than EnKF, respectively.



**Figure 10.** (a) Yield prediction scatter plots for EnKF, SBUPLEX, and open-loop simulation. (b) Frequency distribution for RBE achieved from the contrast between simulated and observed yields. Bin size = 5.

**Table 5.** Validation results of the estimated yield at the field scale using remotely sensing LAI.

	$R^2$	RMSE (%) $ha^{-1}$	MAE, %	Average RBE, %
Open-loop simulation	0.39	1.06 (13.8)	12.5	8.65
EnKF with remotely sensed LAI	0.73	0.71 (9.2)	7.73	0.77
SUBPLEX with remotely sensed LAI	0.78	0.64 (8.3)	6.85	−0.39

#### 4. Discussion

In this study, SUBPLEX demonstrates slightly better global yield prediction performance than EnKF. The detailed results also indicate the accuracy of ENKF in 100 of the 181 samples is lower than that of SUBPLEX. The reason may be that the tree age and physiological genetic factors may affect the actual emergence date and phenology time of jujube trees. The premise of ensuring the accuracy of the EnKF method is that the growth cycle between the model and the satellite observations (planting date plus phenology) is correct. If the phenological development time has a big error, the sequential approach of EnKF is also equivalent to forcing the LAI into a crop model that may belong to different growth stages, and there is essentially a time shift between the reality and the model, which may result in worse simulated accuracy [47]. The SUBPLEX as a kind of variational method may reduce the accumulation and diffusion of remote sensing data errors in the assimilation process, as well as the impact of the phenological time shift on assimilation results to some extent [3]. SUBPLEX's process is to optimize TDWI and SPAN parameters in the WOFOST model in order to ensure that the remotely-sensed LAI and the simulated WOFOST LAI match as well as possible. Then, the corrected TDWI and SPAN parameters will be used to re-drive the WOFOST model. If the WOFOST model is run from a shifted phenology time, the emergence, flowering, and maturity dates will be offset. However, the entire simulation curve is shifted overall, so it has little effect on the final predicted yield. For example, when TDWI and SPAN are set to a fixed value (TDWI = 21.6, SPAN = 53), the corrected WOFOST model is attempted to run from 115 days, 118 days, and 121 days, respectively, and the predicted yields are 9.039, 9.203, and 9.271. Obviously, when the phenological time deviation is six days, there is only a relatively small 2.5% yield prediction error. It is speculated that when the phenological development time is uncertain, the use of SUBPLEX usually results in better assimilation accuracy for yield prediction than using the EnKF method.

In theory, more state variables of remote sensing observations can produce better assimilation accuracy. In practice, satellite data for key crop growth seasons may not be available due to revisit cycles and cloud coverage limitations. Loss of remote sensing observations at critical developmental stages

may have a key impact on model performance. To further discuss the sensitivity of LAI during different phenological stages and two assimilation methods to yield prediction, only three field-measured LAIs from 55 observations were employed to establish assimilation strategies to compare the yield prediction performance of the EnKF and SUBPLEX methods (Table 6). The results show that when the observation data on 9 June, 27 July or 12 August are missing, respectively, the performance of SUBPLEX assimilation is still higher than that of the EnKF method. When the observation data of 28 August is missing, the performance of the SUBPLEX method drops sharply, and its accuracy is also significantly lower than that of the EnKF method. The main reason should be that the SPAN parameter mainly affects the LAI change in the late growing season, while the lack of LAI (28 August) in the red maturity period leads to a larger error in the corrected SPAN parameters, with a mean deviation of 2.7 days compared to the use of four LAIs. Further tests showed that when the SPAN parameter was set to a fixed value (average of field measurements, 53), the  $R^2$  and RMSE of the yield prediction were improved to 0.73 and  $0.77 \text{ t ha}^{-1}$ , respectively, which was also slightly higher than the EnKF assimilation accuracy under the same conditions. The finding also indicates that when using the SUBPLEX method to optimize SPAN parameters, the assimilated LAI shall include at least one observation spot in the middle and late stages of maturity. If this observation spot is missing, it is recommended to design an average SPAN for SUBPLEX assimilation. In addition, the combination of high or medium spatial resolution remote sensing satellite data, such as Landsat and Sentinel satellites, can be recommended to resolve the problem of remotely sensed state variables missing.

**Table 6.** Model performance of SUBPLEX versus EnKF when missing an observation.

	Available Observations	$R^2$	RMSE (%) $\text{t ha}^{-1}$	MAE, %	Average RBE, %
SUBPLEX	Four observations	0.86	0.55	5.77	−2.3
	Without 9 June	0.84	0.59	6.04	−1.75
	Without 27 July	0.85	0.58	6.01	−1.48
	Without 12 August	0.81	0.65	6.35	−3.31
	Without 28 August <sup>(a)</sup>	0.36	1.18	12.3	0.23
	Without 28 August <sup>(b)</sup>	0.73	0.77	9.02	4.49
EnKF	Four observations	0.81	0.65	6.78	−1.42
	Without 9 June	0.65	0.88	8.78	−7.69
	Without 27 July	0.75	0.75	7.93	−1.87
	Without 12 August	0.74	0.76	7.98	−1.59
	Without 28 August	0.71	0.81	8.73	−1.02

<sup>(a)</sup> represents that TDWI and SPAN parameters are optimized. <sup>(b)</sup> represents that only TDWI is optimized, SPAN is equal to a fixed value that is the average of the field measurements.

The present research also compares the computational efficiency of SUBPLEX and the brute force optimization method, with higher computational efficiency. Still taking the sample of Section 3.2 as an example, the function calls of the forced method take 1000 times, while the function calls of the SUBPLEX method are only 25. The SUBPLEX algorithm can find a solution with a similar accuracy with much fewer function calls. Of course, the computation time is closely related to the initial steps of TDWI and SPAN. When the initial step size of TDWI and SPAN is adjusted to 0.1 and 0.5, respectively, the iteration number of the cost function is increased to 43, and when they are adjusted to 0.05 and 0.1, the number of iterations is increased to 76. The running time test also shows that the SUBPLEX method runs slightly longer than the EnKF method. However, the time difference is usually less than ten seconds for different samples, which is a speed that can be fully accepted. However, the yield prediction accuracy of SUBPLEX assimilation is higher than EnKF.

The setting of key parameters of SUBPLEX will directly affect the minimization of the objective function of remote sensing inversion and model simulation, which will affect the optimization accuracy of the parameters and the final simulation yield, such as relative tolerance for convergence  $\varepsilon$ , simplex reduction coefficient  $\varphi$ , and step reduction coefficient  $\omega$ . Smaller  $\varepsilon$ ,  $\varphi$ , and larger  $\omega$  values usually result in a minimized objective function, and of course, the computation time increases. In our study,

the minimization of the objective function of all samples ranges from 0.001 to 0.26, with most samples between 0.05 and 0.1. Furthermore, when trying to decrease  $\varepsilon$  from 0.05 to 0.01, the  $R^2$  value only increases by 0.0015, and the RMSE decreases by  $0.003 \text{ t ha}^{-1}$ . Similarly, when  $\varphi$  decreases from 0.25 and  $\omega$  increases from 0.1, the final yield assimilation accuracy is not significantly improved. Therefore,  $\varepsilon = 0.05$ ,  $\varphi = 0.25$ , and  $\omega = 0.1$  are set in the study. It is recommended that when the SUBPLEX algorithm is carried out for remote sensing assimilation, the parameter setting can refer to the range of the minimum value of the objective function. When the parameter adjustment cannot significantly improve the objective function value, the parameter value at this time can be used as a suitable choice. In the future, further research on parameter optimization settings and automatic search methods has significant exploring significance for SUBPLEX assimilation. In addition, the comparison of the accuracy and computational efficiency of the SUBPLEX and the 4DVAR variational algorithm is also worth exploring.

## 5. Conclusions

In this study, the SUBPLEX algorithm was tested to assimilate remotely-sensed data into the WOFOST model to improve the modeling accuracy for jujube yield at the field scale. The state variables, here LAIs, from the four key development stages, were assimilated into a validated WOFOST model to optimize the initial input parameters, TDWI and SPAN, thereby improving yield estimation for jujube fruit trees. The results indicate that the assimilation method has the potential to enhance the yield estimation accuracy compared with unassimilated simulation, and the SUBPLEX method showed slightly better yield prediction performance versus EnKF assimilation. EnKF assimilation overestimated the actual yields in high-yield orchards and underestimated yields in low-yield orchards. The SUBPLEX method can reduce the shift error of phenological development time to some extent, thereby carrying out a more accurate yield simulation. The results indicate that the SUBPLEX algorithm can be considered as a promising assimilation method of remotely-sensed information and crop growth models. In the future, the assimilation performance of the SUBPLEX method applied to other crops at the field and regional scales needs to be further tested and compared. In addition, the determination and automatic optimization of the initial input parameters of SUBPLEX assimilation, such as initial step size, relative tolerance and initial values of the parameters to be re-calibrated, are worthy of further research and exploration to enhance its computational efficiency and assimilation accuracy.

**Author Contributions:** T.B.: methodology and writing/preparation of the original draft. S.W. and W.M.: Pretreatment for remote sensing data. N.Z.: investigation. T.W.: model analysis. Y.C. and B.M.: review and editing, and supervision.

**Funding:** This research was funded by “National Natural Science Foundation of China (41561088 and 61501314) and Science & Technology Nova Program of Xinjiang Production and Construction Corps (2018CB020).

**Acknowledgments:** Thanks to De Wit from Wageningen University for sharing the Python code for the WOFOST model and the EnKF assimilation method. The authors would like to thank the editors and anonymous reviewers for valuable comments, which are significant for improving this manuscript.

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

## References

1. Gao, Q.; Wu, C.; Wang, M. The Jujube (*Ziziphus Jujuba* Mill.) Fruit: A review of current knowledge of fruit composition and health benefits. *J. Agric. Food Chem.* **2013**, *61*, 3351–3363. [[CrossRef](#)] [[PubMed](#)]
2. Li, J.W.; Fan, L.P.; Ding, S.D.; Ding, X.L. Nutritional composition of five cultivars of chinese jujube. *Food Chem.* **2007**, *103*, 454–460. [[CrossRef](#)]
3. Jin, X.; Kumar, L.; Li, Z.; Feng, H.; Xu, X.; Yang, G.; Wang, J. A review of data assimilation of remote sensing and crop models. *Eur. J. Agron.* **2018**, *92*, 141–152. [[CrossRef](#)]
4. Zhou, G.; Liu, X.; Liu, M. Assimilating Remote Sensing Phenological Information into the WOFOST Model for Rice Growth Simulation. *Remote Sens.* **2019**, *11*, 268. [[CrossRef](#)]

5. Nearing, G.S.; Crow, W.T.; Thorp, K.R.; Moran, M.S.; Reichle, R.H.; Gupta, H.V. Assimilating remote sensing observations of leaf area index and soil moisture for wheat yield estimates: An observing system simulation experiment. *Water Resour. Res.* **2012**, *48*, W05525. [[CrossRef](#)]
6. 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](#)]
7. Huang, J.; Ma, H.; Su, W.; Zhang, X.; Huang, Y.; Fan, J.; Wu, W. Jointly Assimilating MODIS LAI and et Products into the SWAP Model for Winter Wheat Yield Estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4060–4071. [[CrossRef](#)]
8. Huang, J.; Sedano, F.; Huang, Y.; Ma, H.; Li, X.; Liang, S.; Tian, L.; Zhang, X.; Fan, J.; Wu, W. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to improve regional winter wheat yield estimation. *Agric. For. Meteorol.* **2016**, *216*, 188–202. [[CrossRef](#)]
9. Huang, J.; Tian, L.; Liang, S.; Ma, H.; Becker-Reshef, I.; Huang, Y.; Su, W.; Zhang, X.; Zhu, D.; Wu, W. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agric. For. Meteorol.* **2015**, *204*, 106–121. [[CrossRef](#)]
10. Yao, F.; Tang, Y.; Wang, P.; Zhang, J. Estimation of maize yield by using a process-based model and remote sensing data in the Northeast China Plain. *Phys. Chem. Earth* **2015**, *87*, 142–152. [[CrossRef](#)]
11. Jin, X.; Yang, G.; Xu, X.; Yang, H.; Feng, H.; Li, Z.; Shen, J.; Zhao, C.; Lan, Y. Combined multi-temporal optical and radar parameters for estimating LAI and biomass in winter wheat using HJ and RADARSAR-2 data. *Remote Sens.* **2015**, *7*, 13251–13272. [[CrossRef](#)]
12. Wang, J.; Li, X.; Lu, L.; Fang, F. Estimating near future regional corn yields by integrating multi-source observations into a crop growth model. *Eur. J. Agron.* **2013**, *49*, 126–140. [[CrossRef](#)]
13. Chakrabarti, S.; Bongiovanni, T.; Judge, J.; Zotarelli, L.; Bayer, C. Assimilation of SMOS soil moisture for quantifying drought impacts on crop yield in agricultural regions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3867–3879. [[CrossRef](#)]
14. Ines, A.V.M.; Das, N.N.; Hansen, J.W.; Njoku, E.G. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sens. Environ.* **2013**, *138*, 149–164. [[CrossRef](#)]
15. Mishra, A.K.; Ines, A.V.M.; Das, N.N.; Prakash Khedun, C.; Singh, V.P.; Sivakumar, B.; Hansen, J.W. Anatomy of a local-scale drought: Application of assimilated remote sensing products, crop model, and statistical methods to an agricultural drought study. *J. Hydrol.* **2015**, *526*, 15–29. [[CrossRef](#)]
16. Zhuo, W.; Huang, J.; Li, L.; Zhang, X.; Ma, H.; Gao, X.; Huang, H.; Xu, B.; Xiao, X. Assimilating Soil Moisture Retrieved from Sentinel-1 and Sentinel-2 Data into WOFOST Model to Improve Winter Wheat Yield Estimation. *Remote Sens.* **2019**, *11*, 1618. [[CrossRef](#)]
17. Huang, J.; Ma, H.; Sedano, F.; Lewis, P.; Liang, S.; Wu, Q.; Su, W.; Zhang, X.; Zhu, D. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST-PROSAIL model. *Eur. J. Agron.* **2019**, *102*, 1–13. [[CrossRef](#)]
18. Dong, T.; Liu, J.; Qian, B.; Zhao, T.; Jing, Q.; Geng, X.; Wang, J.; Huffman, T.; Shang, J. Estimating winter wheat biomass by assimilating leaf area index derived from fusion of Landsat-8 and MODIS data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 63–74. [[CrossRef](#)]
19. Ma, G.; Huang, J.; Wu, W.; Fan, J.; Zou, J.; Wu, S. Assimilation of MODIS-LAI into the WOFOST model for forecasting regional winter wheat yield. *Math. Comput. Model.* **2013**, *58*, 634–643. [[CrossRef](#)]
20. Ren, J.; Liu, X.; Chen, Z.; Tang, H. Extracting spatial information of harvest index for winter wheat based on modis NDVI in North China. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Honolulu, HI, USA, 25–30 July 2010; pp. 2143–2146.
21. Dong, Y.; Wang, J.; Li, C.; Yang, G.; Wang, Q.; Liu, F.; Zhao, J.; Wang, H.; Huang, W. Comparison and analysis of data assimilation algorithms for predicting the leaf area index of crop canopies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 188–201. [[CrossRef](#)]
22. Dong, Y.; Zhao, C.; Yang, G.; Chen, L.; Wang, J.; Feng, H. Integrating a very fast simulated annealing optimization algorithm for crop leaf area index variational assimilation. *Math. Comput. Model.* **2013**, *58*, 877–885. [[CrossRef](#)]
23. He, B.; Li, X.; Quan, X.; Qiu, S. Estimating the aboveground dry biomass of grass by assimilation of retrieved LAI into a crop growth model. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 550–561. [[CrossRef](#)]

24. Jin, H.; Li, A.; Wang, J.; Bo, Y. Improvement of spatially and temporally continuous crop leaf area index by integration of CERES-Maize model and MODIS data. *Eur. J. Agron.* **2016**, *78*, 1–12. [[CrossRef](#)]
25. Guo, C.; Zhang, L.; Zhou, X.; Zhu, Y.; Cao, W.; Qiu, X.; Cheng, T.; Tian, Y. Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precis. Agric.* **2017**, *19*, 55–78. [[CrossRef](#)]
26. Jin, X.; Li, Z.; Yang, G.; Yang, H.; Feng, H.; Xu, X.; Wang, J.; Li, X.; Luo, J. Winter wheat yield estimation based on multi-source medium resolution optical and radar imaging data and the AquaCrop model using the particle swarm optimization algorithm. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 24–37. [[CrossRef](#)]
27. Li, Z.; Wang, J.; Xu, X.; Zhao, C.; Jin, X.; Yang, G.; Feng, H. Assimilation of two variables derived from hyperspectral data into the DSSAT-CERES model for grain yield and quality estimation. *Remote Sens.* **2015**, *7*, 12400–12418. [[CrossRef](#)]
28. Jin, M.; Liu, X.; Wu, L.; Liu, M. An improved assimilation method with stress factors incorporated in the WOFOST model for the efficient assessment of heavy metal stress levels in rice. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *41*, 118–129. [[CrossRef](#)]
29. Liu, F.; Liu, X.; Ding, C.; Wu, L. The dynamic simulation of rice growth parameters under cadmium stress with the assimilation of multi-period spectral indices and crop model. *Field Crops Res.* **2015**, *183*, 225–234. [[CrossRef](#)]
30. Silvestro, P.C.; Pignatti, S.; Pascucci, S.; Yang, H.; Li, Z.; Yang, G.; Huang, W.; Casa, R. Estimating wheat yield in China at the field and district scale from the assimilation of satellite data into the Aquacrop and simple algorithm for yield (SAFY) models. *Remote Sens.* **2017**, *9*, 509. [[CrossRef](#)]
31. Fang, H.; Liang, S.; Hoogenboom, G. Integration of MODIS LAI and vegetation index products with the CSM-CERES-Maize model for corn yield estimation. *Int. J. Remote Sens.* **2011**, *32*, 1039–1065. [[CrossRef](#)]
32. Tian, L.; Li, Z.; Huang, J.; Wang, L.; Su, W.; Zhang, C.; Liu, J. Comparison of Two Optimization Algorithms for Estimating Regional Winter Wheat Yield by Integrating MODIS Leaf Area Index and World Food Studies Model. *Sens. Lett.* **2013**, *11*, 1261–1268. [[CrossRef](#)]
33. Claverie, M.; Demarez, V.; Duchemin, B.; Hagolle, O.; Keravec, P.; Marciel, B.; Ceschia, E.; Dejoux, J.F.; Dedieu, G. Spatialization of crop leaf area index and biomass by combining a simple crop model safy and high spatial and temporal resolutions remote sensing data. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Cape Town, South Africa, 12–17 July 2009; Volume 3.
34. Jégo, G.; Pattey, E.; Liu, J. Using Leaf Area Index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. *Field Crops Res.* **2012**, *131*, 63–74. [[CrossRef](#)]
35. Dente, L.; Satalino, G.; Mattia, F.; Rinaldi, M. Assimilation of leaf area index derived from ASAR and MERIS data into CERES-Wheat model to map wheat yield. *Remote Sens. Environ.* **2008**, *112*, 1395–1407. [[CrossRef](#)]
36. Hu, S.; Mo, X.; Lin, Z. Optimizing the photosynthetic parameter V<sub>cm</sub> by assimilating MODIS-fPAR and MODIS-NDVI with a process-based ecosystem model. *Agric. For. Meteorol.* **2014**, *198*, 320–334. [[CrossRef](#)]
37. Morel, J.; Todoroff, P.; Bégué, A.; Bury, A.; Martiné, J.F.; Petit, M. Toward a satellite-based system of sugarcane yield estimation and forecasting in smallholder farming conditions: A case study on reunion island. *Remote Sens.* **2014**, *6*, 6620–6635. [[CrossRef](#)]
38. Ines, A.V.M.; Honda, K.; Das Gupta, A.; Droogers, P.; Clemente, R.S. Combining remote sensing-simulation modeling and genetic algorithm optimization to explore water management options in irrigated agriculture. *Agric. Water Manag.* **2006**, *83*, 221–232. [[CrossRef](#)]
39. Wang, H.; Zhu, Y.; Li, W.; Cao, W.; Tian, Y. Integrating remotely sensed leaf area index and leaf nitrogen accumulation with RiceGrow model based on particle swarm optimization algorithm for rice grain yield assessment. *J. Appl. Remote Sens.* **2014**, *8*, 083674. [[CrossRef](#)]
40. Zhu, X.; Zhao, Y.; Feng, X. A methodology for estimating Leaf Area Index by assimilating remote sensing data into crop model based on temporal and spatial knowledge. *Chin. Geogr. Sci.* **2013**, *23*, 550–561. [[CrossRef](#)]
41. Zhao, Y.; Chen, S.; Shen, S. Assimilating remote sensing information with crop model using Ensemble Kalman Filter for improving LAI monitoring and yield estimation. *Ecol. Model.* **2013**, *270*, 30–42. [[CrossRef](#)]
42. Xie, Y.; Wang, P.; Bai, X.; Khan, J.; Zhang, S.; Li, L.; Wang, L. Assimilation of the leaf area index and vegetation temperature condition index for winter wheat yield estimation using Landsat imagery and the CERES-Wheat model. *Agric. For. Meteorol.* **2017**, *246*, 194–206. [[CrossRef](#)]

43. Wu, S.; Huang, J.; Liu, X.; Fan, J.; Ma, G.; Zou, J. Assimilating MODIS-LAI into crop growth model with EnKF to predict regional crop yield. In *IFIP Advances in Information and Communication Technology, Proceedings of the IFIP Advances in Information and Communication Technology, Beijing, China, 29–31 October 2011*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 370, pp. 410–418.
44. Pauwels, V.R.N.; Verhoest, N.E.C.; De Lannoy, G.J.M.; Guissard, V.; Lucau, C.; Defourny, P. Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. *Water Resour. Res.* **2007**, *43*, W04421. [[CrossRef](#)]
45. Ma, H.; Huang, J.; Zhu, D.; Liu, J.; Su, W.; Zhang, C.; Fan, J. Estimating regional winter wheat yield by assimilation of time series of HJ-1 CCD NDVI into WOFOST-ACRM model with Ensemble Kalman Filter. *Math. Comput. Model.* **2013**, *58*, 759–770. [[CrossRef](#)]
46. Li, Y.; Zhou, Q.; Zhou, J.; Zhang, G.; Chen, C.; Wang, J. Assimilating remote sensing information into a coupled hydrology-crop growth model to estimate regional maize yield in arid regions. *Ecol. Model.* **2014**, *291*, 15–27. [[CrossRef](#)]
47. De Wit, A.J.W.; van Diepen, C.A. Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. *Agric. For. Meteorol.* **2007**, *146*, 38–56. [[CrossRef](#)]
48. Curnel, Y.; de Wit, A.J.W.; Duveiller, G.; Defourny, P. Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment. *Agric. For. Meteorol.* **2011**, *151*, 1843–1855. [[CrossRef](#)]
49. Cheng, Z.; Meng, J.; Qiao, Y.; Wang, Y.; Dong, W.; Han, Y. Preliminary study of soil available nutrient simulation using a modified WOFOST model and time-series remote sensing observations. *Remote Sens.* **2018**, *10*, 64. [[CrossRef](#)]
50. Bolten, J.D.; Crow, W.T.; Jackson, T.J.; Zhan, X.; Reynolds, C.A. Evaluating the Utility of Remotely Sensed Soil Moisture Retrievals for Operational Agricultural Drought Monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 57–66. [[CrossRef](#)]
51. Jiang, Z.; Chen, Z.; Chen, J.; Liu, J.; Ren, J.; Li, Z.; Sun, L.; Li, H. Application of crop model data assimilation with a particle filter for estimating regional winter wheat yields. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4422–4431. [[CrossRef](#)]
52. Vazifedoust, M.; van Dam, J.C.; Bastiaanssen, W.G.M.; Feddes, R.A. Assimilation of satellite data into agrohydrological models to improve crop yield forecasts. *Int. J. Remote Sens.* **2009**, *30*, 2523–2545. [[CrossRef](#)]
53. Chen, Y.; Zhang, Z.; Tao, F. Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data. *Eur. J. Agron.* **2018**, *101*, 163–173. [[CrossRef](#)]
54. Huang, Y.; Zhu, Y.; Li, W.; Cao, W.; Tian, Y. Assimilating Remotely Sensed Information with the WheatGrow Model Based on the Ensemble Square Root Filter for Improving Regional Wheat Yield Forecasts. *Plant Prod. Sci.* **2013**, *16*, 352–364. [[CrossRef](#)]
55. Huang, J.; Gómez-Dans, J.; Huang, H.; Ma, H.; Wu, Q.; Lewis, P.; Liang, S.; Chen, Z.; Xue, J.; Wu, Y.; et al. Assimilation of remote sensing into crop growth models: current status and perspectives. *Agric. For. Meteorol.* **2019**, 276–277, 107609. [[CrossRef](#)]
56. Rowan, T.H. Functional Stability Analysis of Numerical Algorithms. Ph.D. Thesis, University of Texas, Austin, TX, USA, 1990.
57. Jonsén, P.; Isaksson, E.; Sundin, K.G.; Oldenburg, M. Identification of lumped parameter automotive crash models for bumper system development. *Int. J. Crashworthiness* **2009**, *14*, 533–541. [[CrossRef](#)]
58. De Wit, A.; Boogaard, H.; Fumagalli, D.; Janssen, S.; Knapen, R.; van Kraalingen, D.; Supit, I.; van der Wijngaart, R.; van Diepen, K. 25 years of the WOFOST cropping systems model. *Agric. Syst.* **2019**, *168*, 154–167. [[CrossRef](#)]
59. Bai, T.; Zhang, N.; Mercatoris, B.; Chen, Y. Improving Jujube Fruit Tree Yield Estimation at the Field Scale by Assimilating a Single Landsat Remotely-Sensed LAI into the WOFOST Model. *Remote Sens.* **2019**, *11*, 1119. [[CrossRef](#)]
60. De Wit, A.; Duveiller, G.; Defourny, P. Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations. *Agric. For. Meteorol.* **2012**, *164*, 39–52. [[CrossRef](#)]
61. Van Diepen, C.A.; Wolf, J.; van Keulen, H.; Rappoldt, C. WOFOST: A simulation model of crop production. *Soil Use Manag.* **1989**, *5*, 16–24. [[CrossRef](#)]

62. Bai, T.; Zhang, N.; Mercatoris, B.; Chen, Y. Jujube yield prediction method combining Landsat 8 Vegetation Index and the phenological length. *Comput. Electron. Agric.* **2019**, *162*, 1011–1027. [[CrossRef](#)]
63. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
64. Pettorelli, N. *The Normalized Difference Vegetation Index*; Oxford University Press: London, UK, 2013; pp. 1–208. [[CrossRef](#)]



© 2019 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/>).