Mapping of Cotton Fields Within-Season Using Phenology-Based Metrics Derived from a Time Series of Landsat Imagery

Dhahi Al-Shammari *, Ignacio Fuentes®, Brett M. Whelan, Patrick Filippi® and Thomas F. A. Bishop

Sydney Institute of Agriculture, School of Life & Environmental Science, The University of Sydney, Central Ave, Eveleigh, Sydney, NSW 2015, Australia; ignacio.fuentes@sydney.edu.au (I.F.); brett.whelan@sydney.edu.au (B.M.W.); patrick.filippi@sydney.edu.au (P.F.); thomas.bishop@sydney.edu.au (T.F.A.B.)

* Correspondence: dhahi.al-shammari@sydney.edu.au

Received: 23 July 2020; Accepted: 16 September 2020; Published: 17 September 2020

Abstract: A phenology-based crop type mapping approach was carried out to map cotton fields throughout the cotton-growing areas of eastern Australia. The workflow was implemented in the Google Earth Engine (GEE) platform, as it is time efficient and does not require processing in multiple platforms to complete the classification steps. A time series of Normalised Difference Vegetation Index (NDVI) imagery were generated from Landsat 8 Surface Reflectance Tier 1 (L8SR) and processed using Fourier transformation. This was used to produce the harmonised-NDVI (H-NDVI) from the original NDVI, and then phase and amplitude values were generated from the H-NDVI to visualise active cotton in the targeted fields. Random Forest (RF) models were built to classify cotton at early, mid and late growth stages to assess the ability of the model to classify cotton as the season progresses, with phase, amplitude and other individual bands as predictors. Results obtained from leave-one-season-out cross validation (LOSOCV) indicated that Overall Accuracy (OA), Kappa, Producer’s Accuracies (PA) and User’s Accuracy (UA), increased significantly when adding amplitude and phase as predictor variables to the model, than prediction using H-NDVI or raw bands only. Commission and omission errors were reduced significantly as the season progressed and more in-season imagery was available. The methodology proposed in this study can map cotton crops accurately based on the reconstruction of the unique cotton reflectance trajectory through time. This study confirms the importance of phenological metrics in improving in-season cotton fields mapping across eastern Australia. This model can be used in conjunction with other datasets to forecast yield based on the mapped crop type for improved decision making related to supply chain logistics and seasonal outlooks for production.

Keywords: remote sensing; land use classification; crop type mapping; commodity mapping

1. Introduction

Remote identification of crop type is widely studied for many agricultural applications. In most cases, this is done retrospectively after the growing season [1], however, in-season mapping is required for better decisions for policymakers and others in the agribusiness sector [2]. For example, knowing the type of crop and the hectares of a particular crop can help in making crucial management, marketing, and logistical decisions [3]. Cotton is an economically important crop in Australia and is worth $2 billion dollars annually on average. The number of cotton fields grown in Australia considerably fluctuates from season to season. For example, in 2011, 599,630 hectares were grown, whereas in 2019,
only 292,380 hectares were grown [4]. It is therefore important to develop a reliable method to map cotton fields in-season for regional irrigation allocations, especially in areas where water scarcity is dominant [5]. The conventional way of collecting land-use information is mostly achieved by surveying farmers or conducting land surveys. This is an accurate way of mapping crop type, but it is relatively expensive and time-consuming [6]. Remote sensing technologies offer great potential for collecting temporal and spatial data to support decisions required for crop monitoring and management [7]. This is usually cost and time-efficient at the farm-scale, but also for large areas [8].

There are many methods for crop type mapping using remote sensing. For example, conventional classification is widely used where one value per pixel is used to classify crop type. This approach is often performed relying on images from a single date [9]. Single date images were used to map citrus groves for economic assessment purposes [10]. However, the spectral similarities of some citrus grove plants with other vegetation covers caused misclassification in some cases. The drawback of this type of classification is that it does not track the growth trajectory over a period of time, which can result in misclassification of crop types [11].

Recently, phenology-based metrics of crops have been employed for crop type mapping tasks to overcome issues related to conventional crop mapping methods. Phenology is one of the characteristics used to monitor periodic plant life-cycle events and how climate influences these events, and this can be used for crop type identification [6]. Crops in a particular environment have specific phenological stages at defined time intervals in the season [12]. Phenological-based methods aim to track the change of phenological trajectories which vary from one crop to another in terms of the start of phenomena (start of crop growing), duration of phenomena (length of growing season), and shape of trajectories (shape of time series) [7]. Most crops have life cycles which are characterised by specific lengths and timings. This can create a unique signal for each individual crop type within a region. The unique signals are a result of the change of a plant’s vigour and/or intensity, mainly due to the time of planting, emergence, growing, maturing and harvesting. In some cases, these changes can be due to temporal constraints, such as plant diseases, or permanent constraints, such as soil constraints, so we may observe different trajectories in the same field which can introduce errors into modelling based on phenology.

There has been a noticeable number of studies in agriculture that focus on the extraction of phenological features using remotely sensed data. For example, it was found that incorporation of crop phenology for crop classification can increase its accuracy. Serra and Pons (2008) [13] built a multitemporal model from 36 Landsat images (2002–2005) to classify different Mediterranean crops (fruit, alfalfa, rice, winter cereals, maize, and other crops) [13]. Their model achieved over 93% accuracy when including multi-temporal signals. However, the efficiency of phenology-based crop type mapping can be improved by adding important features that lead to better discrimination between classes [14]. For example, the Fourier transformation (FT) has added a new representation of time series images by extracting new features to the phenology analysis [15]. These are the amplitude and phase, that provide a description of vegetation status over time [16]. Phase and amplitude bands can be constructed using sine and cosine functions where the amplitude represents the size of change from a base threshold, and the phase is a measure of the period over which the change occurred [17]. Jakubauskas et al. (2002) [15] used Advanced Very High-Resolution Radiometer (AVHRR) Normalised Difference Vegetation Index (NDVI) time series to build phase and amplitude images for crop identification in Kansas, the United States of America. They concluded that the first harmonic term was adequate for identifying crops that have a distinct growing season in the study region, such as corn and alfalfa, and that the second harmonic term was adequate for other crops, such as wheat. In another study, identification of sugarcane was achieved using the amplitude feature, which aided in distinguishing between sugarcane and pasture [18]. This is because sugarcane has a higher amplitude (more above-ground biomass) than pasture during growth stages, and then drops after harvest.

Mapping large areas of crops can be very time consuming and very expensive in terms of computational effort. Therefore, platforms such as the Google Earth Engine (GEE) interface can
be used to perform this in relatively short time periods [19]. The classifier package in the Google Earth Engine provides both generative and discriminative models including Random Forest (RF) [20], Classification and Regression Trees (CART) [21], Support Vector Machine (SVM) [22] and Naïve Bayes [23]. A classifier can be selected depending on the purpose and volume of data being used for a study. Generally, RF requires less time than the SVM for training and can provide comparable accuracies to the SVM classifier [24,25]. RF can give better results when datasets are large, as compared to the Naïve Bays, which can give good results with small datasets. Moreover, compared to Decision Trees (DT), RF overcomes overfitting by constructing an ensemble of DTs. Agricultural applications of the GEE platform such as those in [26–28] have all shown benefits from the flexibility and efficiency of the platform in solving big data problems.

In machine learning models, training and validation of a model requires representative samples to achieve reliable and accurate results. In many studies, machine learning models use subsets of data for training and validation. There are many methods used to train and evaluate models. For instance, some studies use samples that are drawn from the same datasets based on assigning a specific percent of samples for training and percent of samples for evaluation. For example, in a study conducted by Deschamps et al. (2012) [29], data from 2009 season were used to classify different types of beans, cereals and oilseeds. The ground data (fields) were split randomly into equal numbers of training and validation subsets where each class in the training and validation sets were assigned equally. Accuracies obtained from this study were above 80%. In another study, Waldner et al. (2015) [30] classified land cover and crop type using data from the 2013 season. The model was trained on 50% of data which was split randomly. In this study, crop type classification reached an accuracy of 89%. Another study conducted by Hao et al. (2018) [31] showed data obtained from two campaigns (2011 and 2013) that were split randomly for crop type classification. Howard and Wylie (2014) [32] created annual crop type maps for the 2000–2011 seasons. Approximately 16.2 million observations were used to train and validate their model where 90% of data were used for training and 10% for validation. Their model achieved a classification accuracy of 87%.

Other studies rely on using cross validation (CV) techniques [33] to evaluate models where the dataset is partitioned into a number of folds, where a specific number of folds are used for training and the remaining folds are used for validation. There are a limited number of studies that have used the CV technique for crop classification purposes. For example, Pena and Brenning (2015) [34] classified four fruit-tree crops in central Chile using Landsat-8 time series for the 2013/2014 season. They evaluated three models’ performances (linear discriminant analysis, random forest and support vector machine) using the cross validation technique. CV was performed at the field level where 10 folds were used to partition 2291 fields for training and validation subsets. They repeated the CV with different training sample sizes (100, 200, 400 and 800, and all 2291). They achieved a low misclassification error rate <0.21. However, as mentioned earlier, data used in their study were taken from only one season.

However, one limitation of all these studies is that they use observations from the same season to calibrate and validate the models. This means to use this data operationally for mapping crop types, they are testing the interpolation quality of their models rather than their ability to predict in-season without in-season training data. The real challenge is to evaluate a model using a subset that has never been seen by a model.

Therefore, this study aims to map the cotton crop in-season across the major cotton fields of eastern Australia based on the phenology information derived from a time series of Landsat 8 imagery in the GEE platform. All models will be tested using the leave-one-season-out cross validation (LOSOCV) technique where an entire season will be left out to test the ability of the model to map cotton fields that have never been modelled. Furthermore, we will test the quality of the model at different stages of the growing season. Successful development of an approach will enable near real-time in-season mapping of cotton that can be used for managing supply chain logistics and seasonal outlooks for the agricultural sector.
2. Materials and Methods

2.1. Study Area

This study covered major cotton production areas (about 90% of the cotton production region in eastern Australia) stretching from the Macintyre River south of Queensland (QLD), to the Murrumbidgee River in New South Wales (NSW), Australia, for 2013/2014, 2014/2015 and 2019/2020 seasons. This area is known for the dominance of cotton crops in the summer season. Figure 1 shows the study region in NSW and QLD. Actual cotton field boundaries were retrieved from the SataCrop (https://satacrop.com.au/) tool which provides in-season boundaries. This data is collected for the primary purpose of mitigating the risk of spray drift, which allows growers to understand where sensitive crops are located in proximity to their spray operation.

Those field boundaries are uploaded or drawn by farmers on a voluntary basis in the SataCrop tool. Therefore, some filtering of the polygons was based on the NDVI threshold. Polygons were assessed based on the mean NDVI for the entire growing season where polygons with NDVI values that exceeded the threshold of 0.3 were kept and polygons that showed lower values than this threshold were discarded and assumed as bare soil or a mix between bare soil and weeds. Additionally, some polygons drawn in Victoria and Northern Queensland were not used, as these were well outside where cotton is typically grown. The retrieved field boundaries covered ~1,195,439 hectares (ha) with 2742 polygon boundaries (fields). Field boundaries for three seasons were extracted from this tool and a new subset of 286, 211 and 231 field boundaries with a total area of 8426, 6482 and 4643 ha for 2013/2014, 2014/2015 and 2019/2020 seasons, respectively, were created. The new subset was used to train and validate the classification models.

These sites are located along paths 91 to 93 and row 80 to 84 (Figure 1) of the Landsat 8 satellite mission and are covered by a total number of 408 remotely sensed images during this period.

Figure 1. Study area shows the paths and rows used for this study in New South Wales and Queensland, Australia.
2.2. Analytical Process

In Australia, the variation in cotton planting/harvesting dates are primarily due to differences in climate and weather conditions between the different growing regions. However, the cotton season generally starts from October (planting) and ends in March/April (picking). For this reason, the start and end dates of images were used to take into account the variation in sowing dates across the study area.

Figure 2 explains the methodology used in this study. The first step included the acquisition of the surface reflectance bands and removal of pixels with clouds and cloud shadows using the function of mask (FMASK) described by [35]. After that, the NDVI was calculated and subsequently used for the generation of amplitude and phase bands. Three-time windows (Figure 3) for early-, mid- and late-season were selected based on the cotton crop calendar for the NSW and QLD regions. Images were gathered from the beginning of the growing season (which usually starts in September/October according to the cotton growing calendar in Australia) to the end of the season (April/May). The three-time windows were assigned to enable the assessment of the models with increasing amounts of in-season information. For example, for window 1, five images were available at each single path and row. Those five images were used to create metrics to map cotton at early stages. For window 2, the images from the early season (window 1) were supplemented with an additional four to five (depending on revisit time at each path and row) images gathered after the end of window 1. For window 3, a further four to five images were added to the nine to 10 images used in window 1 and 2. Window 3 also includes images from window 1 and 2, in total meaning that approximately 13 to 14 images were used for mapping cotton at the late stage.

The second step was splitting the datasets into training and validation sets. The third step was fitting the classification models and evaluating them accordingly.

Figure 2. Flowchart of analytical process, the steps of which will be detailed in the following sections. (1) Data acquisition, processing and metric variable extraction. (2) Data preparation for analysis. (3) Data analysis and validation. Window 1 includes images from the 1st of October to the 1st of January for each season. Window 2 includes images from the 1st of October to the 1st of February (this includes images from window 1 and window 2). Window 3 includes images from 1st of October to the 1st of May, including also images from windows 1 and 2.
October 1st was set for the start of imagery acquisition and the 1st of May for the end of acquisition to cover the entire study area in terms of the revisit period of the Landsat 8 satellite and also the difference in sowing dates. In this way, it was ensured that all fields in NSW and QLD were covered for the period of study. The analysis of time series data within each of the time-windows tests if the cotton crop can be spectrally mapped at early-, mid- or late-season as well as identifying the reduction/increase in accuracies at each time-window as more in-season data becomes available [36].

![Figure 3](image_url)

**Figure 3.** Example of windows selected for each classification model in the 2013/2014 season. Window 1 indicates the first model time-period from (1 October 2013 to 1 January 2014), Window 2 indicates the second model time-period from (1 October 2013 to 1 February 2014). Window 3 indicates the third model time-period from (1 October 2013 to 1 May 2014). The dashed line set to 0.3 is the Normalised Difference Vegetation Index (NDVI) threshold to detect cotton. Phase and amplitude were calculated within each window (one phase and one amplitude per window).

2.3. Data Acquisition and Preparation

The Google Earth Engine (GEE) platform is a tool that helps to acquire, process, analyse and visualise remote sensing (RS) observations that are available freely in different temporal and spatial scales [37,38]. In this study, GEE was used to carry out all steps starting from surface reflectance data acquisition to the accuracy assessment stage. All surface reflectance bands were acquired from the Landsat 8 Tier 1 surface reflectance with a spatial resolution of 30 m and temporal resolution of 16 days and preprocessed in GEE using available functions including masking clouds/cloud shadows using the FMASK algorithm, which is capable of identifying and removing clouds/cloud shadows from pixels in all images in the collection [35]. The FMASK algorithm labels potential clouds and cloud shadows of each pixel using the ‘pixel_qa’ band.

Images were acquired for the entire growing season to extract the reflectance band information and the NDVI [39] was calculated for each pixel (Equation (1)):

\[
NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}
\] (1)
where $\rho_{\text{NIR}}$ is reflectance in the Near Infrared wavelength range (0.636–0.673 $\mu$m), and $\rho_{\text{red}}$ is the red wavelength range (0.851–0.879 $\mu$m). All acquired images were stacked together to build a time series of images to be used in the classification process.

Phenological metrics derived from time series data can provide very useful information for temporal monitoring of vegetation. However, due to noise in the data and missing data (caused by removing pixels affected by clouds and cloud shadows), temporal monitoring applications can be hindered. Therefore, it is usual to reconstruct or harmonise the times series data using a continuous model to improve the signal quality and deal with data gaps [40].

2.3.1. Harmonised NDVI (H-NDVI)

A Fourier Transformation method (FT) was used to harmonise the original NDVI (H-NDVI), and the coefficients were added to the linear model to get the amplitude ($A$) and phase ($\varphi$) (Equation (2)):

$$H-\text{NDVI}(t)_i = \beta_0 + \beta_1 t + \beta_2 \cos(2\pi \omega t) + \beta_3 \sin(2\pi \omega t) + e_t$$

where $H-\text{NDVI}(t)_i$ is the value of H-NDVI at time $t$ within a pixel ($i$), $\beta_0$ is a constant component of the regression which represents the start of greenness at each pixel, $\beta_1$ is a linear coefficient (linear trend), $\beta_2 = A \cos(\varphi)$ and $\beta_3 = A \sin(\varphi)$ are the harmonic terms that represent the cyclical period of 32 weeks for cotton crops (for this study) and $\omega$ the angular frequency which was set to 1 to limit the time series by one cycle per time unit (year), and $e_t$ represents the error term.

A specific H-NDVI threshold was necessary to confirm active cotton growth in the target fields, delineate the start of the cotton growing season and designate the beginning of the time series data for further analysis [41]. Visual assessment was performed relying on the H-NDVI threshold which was set to 0.3 (after testing different H-NDVI thresholds) to ensure an accurate identification of active growth of cotton within field boundaries. Following this process, time series data sets were constructed for the three time windows from which amplitude and phase data were extracted.

2.3.2. Amplitude

The amplitude (Equation (3)) was recovered from the H-NDVI to indicate the size of H-NDVI change during the growing season:

$$\text{Amplitude} (A) = \sqrt{\beta_2^2 + \beta_3^2}$$

where $\beta_2$ and $\beta_3$ are the fitted coefficients used to calculate the magnitude of H-NDVI change. The accumulation of magnitudes in the time series shows the amount of change through time.

2.3.3. Phase

Phase images were constructed using (Equation (4)) to indicate when the season H-NDVI peak occurs (timing of growth stage):

$$\text{Phase} (\varphi) = \text{atan2} \left( \frac{\beta_3}{\beta_2} \right)$$

where $\text{atan2}$ calculates the arctangent at two dimensions to obtain the phase of H-NDVI [42].

These metrics for each pixel were then collated with the reflectance band information and the H-NDVI for all images in the data sets for the three time periods.

2.4. The Model

2.4.1. Random Forest (RF) Classification Model

An RF classification model was used due to its ability to indicate the best input variables based on their contribution in the model. These models are also more computationally efficient compared to
other classifiers [43]. Since this study seeks to evaluate some phenological-based metrics that might help for cotton detection from surface reflectance data, only two land cover classes were generated (cotton versus others that includes bare soil, water, other crops and pasture and native vegetation).

A pixel-based supervised classification was performed, and training of the RF model was conducted in space and time where temporal predictors were used as input for each pixel location. Time series metric variables were assigned to each model based on allocated time windows. Each data point in the training and validation datasets included more metric values by including new scenes as the growing season progresses. For example, mapping cotton at early stages included a time series constructed from five observations for each pixel. For window 2, the number of observations increased to 10 observations for each pixel (including observations from window 1). For window 3, the number of observations increased to 15 observations (including observations from window 1 and 2).

The RF classifier was trained using the training subset with one Rifle decision tree per class, a size of terminal node set to 1, a fraction of input to bag per tree of 0.5 and the square root of the number of variables per split. A total number of 728 field boundaries with a total area of 19,551 ha for 2013/2014, 2014/2015 and 2019/2020 were used to train and validate the model. Each sample unit consisted of a single polygon (covering a single field) containing mean values of the metrics (each model has a different combination of metrics) which were constructed (Table 1). A comparison between random sampling and Leave-One-Season-Out Cross Validation (LOSOCV) was performed. With random sampling approach, all samples from all years were used to train and validate the models where the dataset was split to 80% for training and 20% for validation and this was repeated 20 times and the mean of model quality metrics were calculated. For the LOSOCV, two years were used to train each model and one season left out for validation. This was repeated so that each season was included in the validation set once. The RF model was tested with five different combinations of metric variables. The first model was based on only the H-NDVI to test the reliability of using the constructed H-NDVI in identifying cotton fields from other classes. The second model was based on using surface reflectance bands only. The third model was based on using only the phase and amplitude images. The fourth model was based on using H-NDVI, phase and amplitude. The fifth model was based on using all metrics as variables.

### Table 1. Metric variable selection for each classification model.

| Selected Metric Variables | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---------------------------|---------|---------|---------|---------|---------|
| Bands 1 to 11             | ✓       | ×       | ×       | ×       | ✓       |
| H-NDVI                    | ×       | ✓       | ×       | ✓       | ✓       |
| Amplitude                 | ×       | ×       | ✓       | ✓       | ✓       |
| Phase                     | ×       | ×       | ✓       | ✓       | ✓       |

#### 2.4.2. Accuracy Assessment

Many accuracy assessment methods are available to validate the model predictions. However, the most common method to assess classification accuracy in remote sensing is based on the error matrix [44]. Classification accuracy was computed by establishing a confusion/error matrix [45] on the training and validation samples [46]. Once the confusion matrix was built, overall accuracies, kappa values, producers’ accuracies, users’ accuracies, omission and commission errors were derived from this matrix to report the performance of each model.

The overall accuracy (OA), which is the simplest description for the model performance, is calculated by dividing the total number of correctly classified samples by the total number of samples (Equation (5)).

Overall accuracy ranges between 0 when there is no agreement and 1 when there is complete agreement:

$$OA = \frac{C_s}{n_s}$$  \hspace{1cm} (5)
where OA stands for the overall accuracy, $C_s$ is the number of Correct samples ($C_s$) classified, and $N_s$ is the total number of samples.

The Kappa statistic is considered a more robust accuracy measure because it considers the random probability of agreement. It estimates the percentage of correctly classified samples compared to chance [47]. Kappa coefficient values range between 0 (when there is no agreement) and 1 (when there is strong agreement). Kappa coefficient is calculated using (Equation (6)):

$$K = \frac{P_o - P_e}{1 - P_e}$$

(6)

where $K$ is the kappa coefficient, $P_o$ is the correctly allocated samples (the proportion of cases in agreement), and $P_e$ is the hypothetical random agreement.

Producer accuracy (PA) indicates the percentage of correctly classified samples of an individual class in the “map maker” point of view [48]. PA was also derived from the error matrix and it ranges from 0 (when the samples of a class are totally misclassified, to 1 when all samples of a class are all correctly classified) (Equation (7)):

$$PA = \frac{C_c}{C_t}$$

(7)

where $PA$ is the producer accuracy, $C_c$ is the accurately classified samples of a specific class divided by the total number of samples of the same class ($C_t$).

User’s accuracy (UA) is a measure of how well the classified pixels belong to the actual class (vegetation cover). It measures the reliability of the map according to the “user’s” point of view [49], and it can be calculated by dividing the total number of correct classifications of a particular class by the total number of reference data for that class (Equation (8)):

$$UA = \frac{C_c}{C_r}$$

(8)

where $UA$ is the user’s accuracy, $C_c$ is the total of correctly classified pixels of a specific class, and $C_r$ is the total number of a reference data of that particular class.

RF variable importance is a very useful feature that ranks the importance of individual predictors regarding their contribution in predicting response and their interaction with the other predictor variables [50–52]. For this study, mean decrease in accuracy (MDA) [53] was used to measure the importance of the metric variables for the three time periods.

Commission Error (CE) and Omission Error (OE) (which are supplementary to UA and PA) were also considered to assess the models’ performances. CE can have occurred when pixels from a specific class are wrongly classified as other classes in the validation data. CE was simply calculated as (CE = 1 − UA). OE occurs when pixels of a specific class are excluded from that class in the validation data. OE can be calculated as (OE = 1 − PA).

3. Results

The entire study area is shown in the left side of Figure 4, but a small region (representing cotton fields near Moree, NSW) is highlighted to show the different model mapping results due to the large extent of the study area. The selected time windows allowed the capture of crop variation at the start, middle and towards the end of the season effectively.

Tables 2 and 3 show the effect of using different combinations of metrics variables on the models’ performances. The accuracy assessment includes the overall accuracy, kappa coefficient, User’s accuracy, Producer’s accuracy, omission and commission.

According to the results achieved from random sampling technique (Table 2), high accuracies can be achieved when using all data to train and validate the models, and this is in contrast with the models trained on LOSOCV which scored lower accuracies, especially in early season. In Table 2, the only low results when using random sampling were when using H-NDVI only, which resulted in
an OA of 0.70 to 0.79 and kappa coefficient of 0.57 to 0.70. UA, PA were also decreased, and the CE and OE were increased significantly. This is expected because using data to train and validate a model from the same season is likely to guarantee similar signatures of metric variables in both training and validation sets, unlike the LOSOCV, where a metric variable of a new season might not have been modelled before, which makes it hard to decide which class a specific signal might belong to.

Conversely, LOSOCV accuracies increased and errors decreased with the progress of the growing season. Moreover, the overall accuracy, kappa coefficient, UA and PA increased with the addition of amplitude and phase in all three time periods as well as with the allocation of more information (with the progression of the season). It can be seen in Table 3 that the best results were obtained from the models trained on a combination of surface reflectance bands, amplitude, phase and H-NDVI, followed by models using amplitude, phase and H-NDVI and models that used only phase and amplitude. The poorest results were obtained when using only H-NDVI. Overall accuracies increased significantly as the season progressed from a mean of ~0.57 (early season) to a mean of ~0.97 when using all metrics as variables. Kappa coefficients were very low for all three-time windows with a mean of ~0.12 when using only H-NDVI and a mean of ~0.46 when using all metrics (late season classification). The obtained kappa values are generally low even though classification is performed at late stages of plant growth, which means that relying on only H-NDVI to map cotton can be very difficult and not reliable. UA was high in the three time periods and for all four models except when using the H-NDVI, which produced relatively lower UA. PA recorded lower accuracies in the early season models. Commission and omission errors were also high at the early stages of classification. However, these errors were reduced significantly with the progress of the growing season and this is not surprising, since more classes were classified correctly at the mid and late seasons.

**Figure 4.** Zoomed-in maps (2013/2014 season) of some cotton fields using the leave-one-season-out cross validation (LOSOCV) technique. (A) the results of using harmonised Normalised Difference Vegetation Index (H-NDVI) only, (B) amplitude and phase only, (C) H-NDVI, amplitude and phase (D) H-NDVI, amplitude, phase and reflectance bands (1–11).
Table 2. Lists of overall Accuracy (OA), kappa, User’s Accuracy (UA), Producer’s Accuracy (PA) accuracies, and Omission Error (OE) and Commission errors (CE) achieved from random sampling based Random Forest (RF) classification applied to the validation datasets.

| Model                                      | Period        | OA    | Kappa  | UA   | PA   | CE  | OE  |
|---------------------------------------------|---------------|-------|--------|------|------|-----|-----|
| Surface reflectance bands                   | Early         | 0.97  | 0.96   | 0.99 | 0.99 | 0.01| 0.01|
| H-NDVI only                                 | Early         | 0.59  | 0.43   | 0.71 | 0.71 | 0.29| 0.29|
| Amplitude and phase                         | Mid           | 0.85  | 0.79   | 0.98 | 0.98 | 0.02| 0.02|
| HNDVI, amplitude and phase                  | Late          | 0.89  | 0.85   | 0.99 | 0.99 | 0.001| 0.001|
| H-NDVI, amplitude, phase and surface reflectance bands | Early | 0.97  | 0.96   | 0.99 | 0.99 | 0.01| 0.01|

Table 3. Lists of overall Accuracy (OA), kappa, User’s Accuracy (UA), Producer’s Accuracy (PA), Omission Error (OE) and Commission Errors (CE) achieved from Leave-One-Season-Out Cross Validation (LOSOCV) based Random Forest (RF) classification applied to the validation datasets.

| Model                                      | Period        | OA    | Kappa  | UA   | PA   | CE  | OE  |
|---------------------------------------------|---------------|-------|--------|------|------|-----|-----|
| Surface reflectance bands only              | Early         | 0.56  | 0.01   | 0.99 | 0.56 | 0.01| 0.44|
| Mid                                         | 0.90          | 0.22  | 0.90   | 0.92 | 0.1  | 0.8 |
| Late                                        | 0.92          | 0.48  | 0.90   | 0.90 | 0.1  | 0.1 |
| H-NDVI only                                 | Early         | 0.43  | 0.22   | 0.51 | 0.18 | 0.48| 0.81|
| Mid                                         | 0.53          | 0.31  | 0.90   | 0.34 | 0.09 | 0.65|
| Late                                        | 0.55          | 0.42  | 0.75   | 0.32 | 0.24 | 0.67|
| Amplitude and phase                         | Early         | 0.63  | 0.45   | 0.90 | 0.56 | 0.1 | 0.43|
| Mid                                         | 0.81          | 0.71  | 0.95   | 0.97 | 0.15 | 0.06|
| Late                                        | 0.95          | 0.91  | 0.99   | 0.99 | 0.01 | 0.01|
| H-NDVI, amplitude and phase                 | Early         | 0.59  | 0.34   | 0.95 | 0.40 | 0.04| 0.59|
| Mid                                         | 0.97          | 0.95  | 0.99   | 0.97 | 0.01 | 0.02|
| Late                                        | 0.99          | 0.99  | 0.99   | 0.99 | 0.01 | 0.01|
| H-NDVI, amplitude, phase and surface reflectance bands | Early | 0.60  | 0.45   | 0.91 | 0.41 | 0.08| 0.69|
| Mid                                         | 0.95          | 0.92  | 0.98   | 0.98 | 0.01 | 0.01|
| Late                                        | 0.98          | 0.96  | 0.97   | 0.99 | 0.02 | 0.01|

| Model                                      | Period        | OA    | Kappa  | UA   | PA   | CE  | OE  |
|---------------------------------------------|---------------|-------|--------|------|------|-----|-----|
| Surface reflectance bands only              | Early         | 0.55  | 0.24   | 0.99 | 0.23 | 0.01| 0.77|
| Mid                                         | 0.59          | 0.27  | 0.98   | 0.31 | 0.02 | 0.69|
| Late                                        | 0.82          | 0.60  | 0.99   | 0.71 | 0.01 | 0.29|
| H-NDVI only                                 | Early         | 0.47  | 0.10   | 0.53 | 0.71 | 0.46| 0.28|
| Mid                                         | 0.50          | 0.12  | 0.55   | 0.78 | 0.44 | 0.21|
| Late                                        | 0.56          | 0.20  | 0.54   | 0.73 | 0.45 | 0.26|
### Table 3. Cont.

| Model                                      | Period | OA    | Kappa | UA  | PA  | CE  | OE  |
|--------------------------------------------|--------|-------|-------|-----|-----|-----|-----|
| Amplitude and phase                        | Early  | 0.68  | 0.42  | 0.80| 0.92| 0.19| 0.07|
|                                            | Mid    | 0.90  | 0.77  | 0.96| 0.85| 0.03| 0.15|
|                                            | Late   | 0.91  | 0.91  | 0.95| 0.96| 0.05| 0.04|
| H-NDVI, amplitude and phase                | Early  | 0.73  | 0.52  | 0.82| 0.70| 0.17| 0.32|
|                                            | Mid    | 0.90  | 0.81  | 0.86| 0.99| 0.13| 0.01|
|                                            | Late   | 0.97  | 0.94  | 0.99| 0.97| 0.009| 0.021|
| H-NDVI, amplitude, phase and reflectance bands | Early | 0.84  | 0.60  | 0.90| 0.89| 0.01| 0.03|
|                                            | Mid    | 0.95  | 0.90  | 0.95| 0.96| 0.01| 0.04|
|                                            | Late   | 0.98  | 0.96  | 0.99| 0.99| 0.002| 0.01|
| 2019/2020                                  |        |       |       |     |     |     |     |
| Surface reflectance bands only             | Early  | 0.59  | 0.21  | 0.75| 0.18| 0.25| 0.82|
|                                            | Mid    | 0.67  | 0.31  | 0.99| 0.26| 0.01| 0.74|
|                                            | Late   | 0.70  | 0.43  | 0.93| 0.36| 0.07| 0.64|
| H-NDVI only                                | Early  | 0.50  | 0.13  | 0.73| 0.75| 0.26| 0.24|
|                                            | Mid    | 0.52  | 0.21  | 0.83| 0.51| 0.16| 0.55|
|                                            | Late   | 0.51  | 0.24  | 0.66| 0.57| 0.33| 0.42|
| Amplitude and phase                        | Early  | 0.54  | 0.35  | 0.78| 0.12| 0.21| 0.80|
|                                            | Mid    | 0.70  | 0.45  | 0.98| 0.50| 0.01| 0.58|
|                                            | Late   | 0.93  | 0.89  | 0.96| 0.94| 0.02| 0.03|
| H-NDVI, amplitude and phase                | Early  | 0.73  | 0.48  | 0.92| 0.67| 0.7| 0.32|
|                                            | Mid    | 0.80  | 0.55  | 0.78| 0.88| 0.11| 0.12|
|                                            | Late   | 0.93  | 0.85  | 0.98| 0.95| 0.02| 0.04|
| H-NDVI, amplitude, phase and reflectance bands | Early | 0.60  | 0.35  | 0.66| 0.70| 0.44| 0.37|
|                                            | Mid    | 0.86  | 0.74  | 0.95| 0.76| 0.04| 0.26|
|                                            | Late   | 0.96  | 0.93  | 0.97| 0.96| 0.02| 0.03|

In this study, the gap-filling technique using FT (Figure 5 shows the mean original NDVI time series and H-NDVI time series values of cotton fields for the three growing seasons; Figure 6 shows the mean error derived from the H-NDVI time series for the cotton fields for the three growing seasons) helped to compensate masked pixels with interpolated values in the time series and building of the unique signal. The FT was used to decompose the NDVI signal into sine and cosine waves, whereas other studies (e.g., Inglada et al. 2015 [54]) used the linear interpolation to compensate for missing information. However, the linear interpolation might be good in some cases where the gap and variance are small. In contrast, if the gap and variance are large (which is usually a problem of clouds) then more sophisticated techniques such as FT are required [55]. Furthermore, the FT technique also preserves the seasonality of the NDVI which helped to create a distinct trajectory for the cotton and other classes and made them separable especially at the mid and late seasons [56]. This also can be supported by Figure 6, showing a very low error produced by the harmonic model. A study conducted by Hendrik and Cyrus (2006) [57] reported that FT-based NDVI attributes have limitations for crop-type classification. However, their study was based on coarse resolution (1.1 km) data derived from AVHRR, unlike this study which was based on moderate resolution data from L8SR (30 m).

Variable importance plots of model 4 are shown in Figure 7 (for early-season, mid-season, and late-season) which were derived from each RF model. Based on the MDA results, the constructed metric variables showed better contribution to the models than the other reflectance bands. More specifically, the H-NDVI was very important for the model at early stages of cotton growth. However, the contribution of this metric variable decreased with the progress of the growing season. The decrease of H-NDVI importance was influenced by the increased contribution of amplitude with the progress of the growing season (mid- and late-season). Similarly, phase contributed more in the
late-season model than in the early- and mid-season models. This is due to the fact that at the beginning of the season, there was insufficient time series data to build a unique phase for the cotton crop.

**Figure 5.** Mean NDVI and the transformed NDVI (H-NDVI) of the study area for the three time periods. (A) Original and H-NDVI time series for 2013/2014 season. (B) Original and H-NDVI time series for 2014/2015 season. (C) Original and H-NDVI time series for 2019/2020 season.

**Figure 6.** Mean error derived from the H-NDVI for the study area using the three time windows. (A) Mean H-NDVI error for the 2013/2014 season. (B) Mean H-NDVI error for the 2014/2015 season. (C) Mean H-NDVI error for the 2019/2020 season.
4. Discussion

Mapping cotton fields using only H-NDVI reduced the accuracy, which could be due to the fact that there are many areas that can reflect similar greenness during the growing season, so relying on only changes in greenness might lead to a confusion of cotton pixels with other classes. Therefore, the delineation of cotton fields was more reliable using the amplitude and phase images only. This also can be seen in Table 3 when using H-NDVI with the amplitude and phase bands, which led to a slight reduction in all accuracies, yet higher than models 1 and 3 (Table 3). The amplitude and phase outperformed the use H-NDVI only and this is because amplitude and phase are the basis to determine the amount change and time of change in the greenness as mentioned previously. But H-NDVI is still very important and provides complementary information to the models. Conversely, even though the H-NDVI, amplitude and phase were the dominant metric variables, the use of the surface reflectance bands (1–11) was still important to the models because they carry additional information that the three metric variables might not have. This is especially the case for band 5 (Near Infrared) which has been used in some crop-type studies and it is very important in crop type classification applications [58,59].

Random sampling technique outperformed the LOSOCV and this was expected because the dataset was split randomly into training and validation. This means data from the same season can be found in both training and validation sets. The only poor results obtained from the random sampling technique was when using H-NDVI only. In the three time windows, H-NDVI yielded very low results and this might be because most classes showed similar greenness, as summer is the time of high vegetative growth, and without capturing the time and amount of H-NDVI change, the separation of cotton from other classes was not effective.

Conversely, the models’ performances improved with the progress of the season using the LOSOCV (Table 3), which was because, at the early stages of cotton, the phenological trajectory was not
yet clear. This led to confusing pixels between cotton and other classes which reflect similar amplitude at the same acquired time window, especially at early stages of growth [1]. However, the trajectory becomes unique mid- and late- season. The confusion between classes might be avoided by using higher temporal resolution imagery with shorter revisit periods such as Sentinel 2, which can improve capturing distinct trajectories to differentiate different classes. For instance, Vaudour et al. (2015) [46] found that confusion between spring barley and bare soil at early stages of crop growth occurred because of low plant cover at early stages [60]. However, that study was not based on analysis of time series information, which might be an additional cause of confusion. Therefore, incorporating high spatiotemporal imagery might be beneficial for distinguishing between classes based on phenological information. The only concern about higher resolution imagery is that it is very expensive in terms of computation time and may be useful for small scale studies only. For example, it is reported that fine resolution imagery is useful for overcoming over/underestimation of small-sized fields due to the sub-pixel proportion that can cause mixed pixels with other classes [61]. However, in this present study, field sizes are relatively large with an average of 394 hectares, so the L8SR spatial resolution of 30-m was sufficient to avoid this issue. However, cotton mapping at early stages might be improved by using higher temporal resolution images such as from Sentinel 2, which unfortunately are only available from mid-2015 onwards, so could not be applied here.

The accuracies shown in Table 3 reflect the usefulness of amplitude and phase and their contribution to the model (Figure 6) when using them without other inputs. This is because temporal amplitude and phase images delineate cotton fields based on the size of NDVI changes and the time of that change throughout the growing season, which was obviously different for other vegetation in the study area. This can be seen in Figure 3 where the reflectance from cotton fields started from bare soil (NDVI ~0.2) in October to February (NDVI ~0.9) when the maximum changes (peak) occurred and then decreased to the bare soil status after harvesting (~April and May). The shape of trajectory was the main driver of the high accuracy and this indicates the change of temporal profile [62,63] which gave a description of the cycle [64].

The unique amplitude could be used to identify cotton, and this was similar to that found by Jakubauskas et al. (2001) [49]. However, they produced seven additive terms which correlated to crop identification. They found that the first term of amplitude can be correlated to a single season, whereas the second term of amplitude (cosine waves) can be correlated to fallow rotation. Cotton fields in the study area were classified using the first term which yielded high accuracy for mapping cotton fields.

Up to now, there are no other studies that have mapped cotton over large areas in Australia. Also, the use of amplitude and phase for in-season mapping of cotton fields is not studied widely and no other studies have handled this aspect. The LOSOCV in this study allowed to test the ability of the models to classify classes on data that were not used to train the models. Therefore, this technique provided results that are reliable and not biased.

5. Conclusions

The aim of this study was to demonstrate the potential of mapping cotton fields in NSW and QLD within season, by applying a simple method using phenology-based indices derived from a time series of imagery. GEE helped to map large areas of cotton fields without the need of multiple platforms for processing and analysing data. The results obtained from this study show that mapping cotton fields at early-, mid- and late-season can be done with high accuracy. Including a signal transformation to create amplitude and phase to feed the random forest classifier increased the accuracy. This approach can be extended by including more crop types to the model to test its ability to discriminate between a wide range of crops. Furthermore, different vegetation indices can be included in this model to test their sensitivity to discriminate different classes.

More work needs to be done on summer crops such as sorghum, which have a similar reflectance signal to cotton. Also, more work needs to be done in a more complicated system where many crops are
grown in the same area to examine the validity of the model on other crops. There is also an opportunity to further test the transferability of the model by including more data from different seasons.

While the present study addressed the cotton detection in NSW and QLD, Australia, it might be applied in other regions. However, it relies on “ground truth” data, and therefore it constrains its applicability on areas with sparse data. A global crop classifier might be envisioned by including the latitudinal/longitudinal gradient. However, this is a matter for future research.

**Author Contributions:** Conceptualization, D.A.-S.; Formal analysis, D.A.-S. and I.F.; Methodology, D.A.-S. and T.F.A.B; Supervision, B.M.W. and T.F.A.B; Writing—original draft, D.A.-S.; Writing—review and editing, D.A.-S., I.F., B.M.W., P.F. and T.F.A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The corresponding author would like to gratefully acknowledge the financial support (CSIRO/Data61 Postgraduate Research Stipend and Supplementary Scholarship in Digital Agriculture) from the Commonwealth Scientific and Industrial Research Organisation (CSIRO). This research was also funded by the University of Sydney, and partly funded by the Cotton Research and Development Corporation (CRDC).

**Acknowledgments:** The authors would like to acknowledge Kristian Maras (Senior Research Informatics Technical Officer) for technical assistance provided as well as the Sydney Informatics Hub, a Core Research Facility of the University of Sydney.

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

**References**

1. Azar, R.; Villa, P.; Stroppiana, D.; Crema, A.; Boschetti, M.; Brivio, P.A. Assessing in-season crop classification performance using satellite data: A test case in Northern Italy. *Eur. J. Remote Sens.* 2016, 49, 361–380. [CrossRef]

2. Bozbek, T.; Sezener, V.; Unay, A. The effect of sowing date and plant density on cotton yield. *J. Agron.* 2006, 5, 122–125.

3. Williams, A.; Mushtaq, S.; Kouadio, L.; Power, B.; Marcussen, T.; McRae, D.; Cockfield, G. An investigation of farm-scale adaptation options for cotton production in the face of future climate change and water allocation policies in southern Queensland, Australia. *Agric. Water Manag.* 2018, 196, 124–132. [CrossRef]

4. Cotton Australia. Statistics. Available online: https://cottonaustralia.com.au/statistics (accessed on 15 April 2020).

5. French, A.N.; Hunsaker, D.J.; Thorp, K.R. Remote sensing of evapotranspiration over cotton using the TSEB and METRIC energy balance models. *Remote Sens. Environ.* 2015, 158, 281–294. [CrossRef]

6. Zhong, L.; Gong, P.; Biging, G.S. Phenology-based crop classification algorithm and its implications on agricultural water use assessments in California’s Central Valley. *Photogramm. Eng. Remote Sens.* 2012, 78, 799–813. [CrossRef]

7. Chen, J.; Chen, J.; Liu, H.; Peng, S. Detection of Cropland Change Using Multi-Harmonic Based Phenological Trajectory Similarity. *Remote Sens.* 2018, 10, 1020. [CrossRef]

8. Chua, R.; Qingbin, X.; Bo, Y. Crop Monitoring Using Multispectral Optical Satellite Imagery; Twenty First Century Aerospace Technology (Asia) Pte. Ltd.: Singapore, 2017.

9. Dadhwal, V.; Singh, R.; Dutta, S.; Parihar, J. Remote sensing based crop inventory: A review of Indian experience. *Trop. Ecol.* 2002, 43, 107–122.

10. Shrivastava, R.J.; Gebelein, J.L. Land cover classification and economic assessment of citrus groves using remote sensing. *ISPRS J. Photogramm. Remote Sens.* 2007, 61, 341–353. [CrossRef]

11. Peña-Barraquín, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* 2011, 115, 1301–1316. [CrossRef]

12. Rodriguez-Galiano, V.F.; Dash, J.; Atkinson, P.M. Characterising the land surface phenology of Europe using decadal MERIS data. *Remote Sens.* 2015, 7, 9390–9409. [CrossRef]

13. Serra, P.; Pons, X. Monitoring farmers’ decisions on Mediterranean irrigated crops using satellite image time series. *Int. J. Remote Sens.* 2008, 29, 2293–2316. [CrossRef]

14. Blaes, X.; Vanhalle, L.; Defourny, P. Efficiency of crop identification based on optical and SAR image time series. *Remote Sens. Environ.* 2005, 96, 352–365. [CrossRef]
15. Mingwei, Z.; Qingbo, Z.; Zhongxin, C.; Jia, L.; Yong, Z.; Chongfa, C. Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. *Int. J. Appl. Earth Obs. Geoinf.* 2008, 10, 476–485. [CrossRef]

16. Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Comput. Electron. Agric.* 2002, 37, 127–139. [CrossRef]

17. Immerzeel, W.; Quiroz, R.; De Jong, S. Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series. *Int. J. Remote Sens.* 2005, 26, 2281–2296. [CrossRef]

18. Rudorff, B.F.T.; Adami, M.; De Aguiar, D.A.; Gusso, A.; Da Silva, W.F.; De Freitas, R.M. Temporal series of EVI/MODIS to identify land converted to sugarcane. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; pp. IV-252–IV-255.

19. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* 2018, 10, 1509. [CrossRef]

20. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32. [CrossRef]

21. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 1985; Volume 19, p. 144. [CrossRef]

22. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* 1995, 20, 273–297. [CrossRef]

23. Webb, G.I. Naïve Bayes. In *Encyclopedia of Machine Learning*; Sammut, C., Webb, G.I., Eds.; Springer: Boston, MA, USA, 2010; pp. 713–714. [CrossRef]

24. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* 2005, 26, 217–222. [CrossRef]

25. Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sens. Environ.* 2017, 198, 369–383. [CrossRef]

26. Li, H.; Wan, W.; Fang, Y.; Zhu, S.Y.; Chen, X.; Liu, B.J.; Hong, Y. A Google Earth Engine-enabled software for efficiently generating high-quality user-ready Landsat mosaic images. *Environ. Model. Softw.* 2019, 112, 16–22. [CrossRef]

27. Yalew, S.G.; van Griensven, A.; van der Zaag, P. AgriSuit: A web-based GIS-MCDA framework for agricultural land suitability assessment. *Comput. Electron. Agric.* 2016, 128, 1–8. [CrossRef]

28. Shelestov, A.; Lavreniuk, M.; Kussul, N.; Novikov, A.; Skakun, S. Large Scale Crop Classification Using Google Earth Engine Platform. *Int. Geosci. Remote Sens. Symp.* 2017, 3696–3699. [CrossRef]

29. Deschamps, B.; McNairn, H.; Shang, J.; Jiao, X. Towards operational radar-only crop type classification: Comparison of a traditional decision tree with a random forest classifier. *Can. J. Remote Sens.* 2012, 38, 60–68. [CrossRef]

30. Waldner, F.; Lambert, M.-J.; Li, W.; Weiss, M.; Demarez, V.; Morin, D.; Marais-Sicre, C.; Hagolle, O.; Baret, F.; Defourny, P. Land cover and crop type classification along the season based on biophysical variables retrieved from multi-sensor high-resolution time series. *Remote Sens.* 2015, 7, 10400–10424. [CrossRef]

31. Hao, P.; Wu, M.; Niu, Z.; Wang, L.; Zhan, Y. Estimation of different data compositions for early-season crop type classification. *PeerJ* 2018, 6, e4834. [CrossRef]

32. Howard, D.M.; Wylie, B.K. Annual crop type classification of the US Great Plains for 2000 to 2011. *Photogramm. Eng. Remote Sens.* 2014, 80, 537–549. [CrossRef]

33. Browne, M.W. Cross-validation methods. *J. Math. Psychol.* 2000, 44, 108–132. [CrossRef]

34. Peña, M.; Brenning, A. Assessing fruit-tree crop classification from Landsat-8 time series for the Maipo Valley, Chile. *Remote Sens. Environ.* 2015, 171, 234–244. [CrossRef]

35. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* 2012, 118, 83–94. [CrossRef]

36. Lira Melo de Oliveira Santos, C.; Augusto Camargo Lamparelli, R.; Kelly Dantas Araujo Figueiredo, G.; Dupuy, S.; Boury, J.; Luciano, A.C.; Torres, R.D.; Le Maire, G. Classification of crops, pastures, and tree plantations along the season with multi-sensor image time series in a subtropical agricultural region. *Remote Sens.* 2019, 11, 334. [CrossRef]

37. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 2017, 202, 18–27. [CrossRef]
38. Xiong, J.; Thenkabail, P.S.; Gumma, M.K.; Teluguntla, P.; Poehnelt, J.; Congalton, R.G.; Yadav, K.; Thau, D. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS J. Photogramm. Remote Sens.* 2017, 126, 224–244. [CrossRef]

39. Rouse, J., Jr. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; NASA: Washington, DC, USA, 1972.

40. Cai, Z.; Jönsson, P.; Jin, H.; Eklundh, L. Performance of smoothing methods for reconstructing NDVI time-series and estimating vegetation phenology from MODIS data. *Remote Sens.* 2017, 9, 1271. [CrossRef]

41. Høgda, K.A.; Tømmervik, H.; Karlsen, S.R. Trends in the start of the growing season in Fennoscandia 1982–2011. *Remote Sens.* 2013, 5, 4304–4318. [CrossRef]

42. Slabaugh, G.G. Computing Euler angles from a rotation matrix. *Retrieved August* 1999, 6, 39–63.

43. Rodriguez-Galiano, V.E.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* 2012, 67, 93–104. [CrossRef]

44. Foody, G.M. Thematic map comparison. *Photogramm. Eng. Remote Sens.* 2004, 70, 627–633. [CrossRef]

45. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* 2002, 80, 185–201. [CrossRef]

46. Lucas, I.; Janssen, F.; van der Wel, F.J. Accuracy assessment of satellite derived landcover data: A review. *Photogramm. Eng. Remote Sens.* 1994, 60, 426–479.

47. Brennan, R.L.; Prediger, D.J. Coefficient kappa: Some uses, misuses, and alternatives. *Educ. Psychol. Meas.* 1981, 41, 687–699. [CrossRef]

48. Shao, G.; Wu, J. On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecol.* 2008, 23, 505–511. [CrossRef]

49. Rossiter, D. *Technical Note: Statistical Methods for Accuracy Assessment of Classified Thematic Maps*; International Institute for Geo-information Science & Earth Observation (ITC): Enschede, The Netherlands, 2004.

50. Strobl, C.; Boulesteix, A.-L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinform.* 2008, 9, 307. [CrossRef]

51. Breiman, L. *Manual on Setting Up, Using, and Understanding Random Forests v3.1*; Statistics Department, University of California Berkeley: Berkeley, CA, USA, 2002; Volume 1, p. 58.

52. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* 2002, 2, 18–22.

53. Han, H.; Guo, X.; Yu, H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 26–28 August 2016; pp. 219–224.

54. Ingilada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedeu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* 2015, 7, 12356–12379. [CrossRef]

55. Pascual-Granado, J.; Garrido, R.; Suarez, J.C. MIARMA: A minimal-loss information method for filling gaps in time series Application to CoRoT light curves. *Astron. Astrophys.* 2015, 575, A78. [CrossRef]

56. Wilson, B.T.; Knight, J.F.; McRoberts, R.E. Harmonic regression of Landsat time series for modeling attributes from national forest inventory data. *ISPRS J. Photogramm. Remote Sens.* 2018, 137, 29–46. [CrossRef]

57. Wagenseil, H.; Samimi, C. Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: Results from a dry savannah environment in Namibia. *Int. J. Remote Sens.* 2006, 27, 3455–3471. [CrossRef]

58. Picoli, M.C.A.; Camara, G.; Sanches, I.; Simoes, R.; Carvalho, A.; Maciel, A.; Coutinho, A.; Esquerdo, J.; Antunes, J.; Begotti, R.A.; et al. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS J. Photogramm. Remote Sens.* 2018, 145, 328–339. [CrossRef]

59. Sun, C.L.; Bian, Y.; Zhou, T.; Pan, J.L. Using of Multi-Source and Multi-Temporal Remote Sensing Data Improves Crop-Type Mapping in the Subtropical Agriculture Region. *Sensors* 2019, 19, 2401. [CrossRef]

60. Vaudour, E.; Noirot-Cosson, P.; Membrive, O. Early-season mapping of crops and cultural operations using very high spatial resolution Pleiades images. *Int. J. Appl. Earth Obs. Geoinf.* 2015, 42, 128–141. [CrossRef]
62. Lhermitte, S.; Verbesselt, J.; Verstraeten, W.W.; Coppin, P. A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sens. Environ.* **2011**, *115*, 3129–3152. [CrossRef]

63. Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Harmonic analysis of time-series AVHRR NDVI data. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 461–470.

64. Negrón Juárez, R.I.; T. Liu, W. FFT analysis on NDVI annual cycle and climatic regionality in Northeast Brazil. *Int. J. Climatol. J. R. Meteorol. Soc.* **2001**, *21*, 1803–1820. [CrossRef]

© 2020 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/).