




Article

Crop Intensity Mapping Using Dynamic Time Warping and Machine Learning from Multi-Temporal PlanetScope Data

Raihan Rafif , Sandiaga Swahyu Kusuma, Siti Saringatin, Giara Iman Nanda, Pramaditya Wicaksono 
and Sanjiwana Arjasakusuma * 

Department of Geographic Information Science, Faculty of Geography, Universitas Gadjah Mada, Yogyakarta 55223, Indonesia; raihanrafif@mail.ugm.ac.id (R.R.); sandiaga.kusuma@mail.ugm.ac.id (S.S.K.); Siti.saringatin@mail.ugm.ac.id (S.S.); giaraiman94@mail.ugm.ac.id (G.I.N.); prama.wicaksono@ugm.ac.id (P.W.)
* Correspondence: sanjiwana.arjasakusuma@ugm.ac.id

Abstract: Crop intensity information describes the productivity and the sustainability of agricultural land. This information can be used to determine which agricultural lands should be prioritized for intensification or protection. Time-series data from remote sensing can be used to derive the crop intensity information; however, this application is limited when using medium to coarse resolution data. This study aims to use 3.7 m-PlanetScope™ Dove constellation data, which provides daily observations, to map crop intensity information for agricultural land in Magelang District, Indonesia. Two-stage histogram matching, before and after the monthly median composites, is used to normalize the PlanetScope data and to generate monthly data to map crop intensity information. Several methods including Time-Weighted Dynamic Time Warping (TWDTW) and the machine-learning algorithms: Random Forest (RF), Extremely Randomized Trees (ET), and Extreme Gradient Boosting (XGB) are employed in this study, and the results are validated using field survey data. Our results show that XGB generated the highest overall accuracy (OA) ($95 \pm 4\%$), followed by RF ($92 \pm 5\%$), ET ($87 \pm 6\%$), and TWDTW ($81 \pm 8\%$), for mapping four-classes of cropping intensity, with the near-infrared (NIR) band being the most important variable for identifying cropping intensity. This study demonstrates the potential of PlanetScope data for the production of cropping intensity maps at detailed resolutions.

Keywords: crop intensity; classification; extreme gradient boosting; random forests; extremely randomized trees



Citation: Rafif, R.; Kusuma, S.S.; Saringatin, S.; Nanda, G.I.; Wicaksono, P.; Arjasakusuma, S. Crop Intensity Mapping Using Dynamic Time Warping and Machine Learning from Multi-Temporal PlanetScope Data. *Land* **2021**, *10*, 1384. <https://doi.org/10.3390/land10121384>

Academic Editors: Bin Chen, Yufang Jin and Le Yu

Received: 24 November 2021
Accepted: 8 December 2021
Published: 14 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The need for intensive production on agricultural land is increasing due to massive climate change, human activities, and urbanization [1]. It is estimated that the world population will increase from 7.2 billion to 9.6–11.2 billion by 2100 [2,3]. However, this increasing need for food supplies is also accompanied by an increasing need for space for settlements [4]. In Sleman, Yogyakarta Province, Indonesia, a study by Harini et al. [5] has shown that land conversion significantly affects the ability to provide food (in this case, rice), where some areas are projected to have insufficient food availability by 2025. Even in Romania, conversion of agricultural land into built-up land does not only occur in peri-urban areas, but has also penetrated into rural areas [6]. Therefore, the increasing agricultural production is necessary to meet increasing demand for food [7].

Although agricultural areas are declining, the demand for food can be met by increasing land productivity, for example, by using fertilizers or adopting new farming systems to practice continuous planting throughout the year or increasing crop intensity [8]. However, such practices can also add more stress to agricultural land, which can result in reduced yields or environmental problems such as decreased groundwater or increased soil salinity [8,9]. Therefore, it is necessary to obtain cropping intensity information to

assess sustainability and to avoid the conversion of highly productive agricultural land to other land uses.

The monitoring of agricultural land can be performed by using remote sensing data, which can provide time-series data over the same areas so that the temporal dynamics of agricultural land can be better observed [2,10,11]. Many studies monitoring cropping dynamics have been conducted; for instance, a study by Maus et al. [12] has succeeded in identifying single and multiple cropping in agricultural land, forests, and grasslands using a transformation of the time-series Enhanced Vegetation Index (EVI) with an overall accuracy of 91.21%, while Foerster et al. [10] were able to apply spectral-temporal profile analysis combined with phenological information to mapping agricultural land types in Germany using a Normalized Vegetation Index (NDVI) transformation applied to Landsat Thematic Mapper (TM)/Enhanced Thematic Mapper (ETM) images with an overall accuracy gain of 65.7%. In addition to these studies, the integration of Landsat 8 Operational Land Imager (OLI) time-series data with Sentinel-2 data succeeded in obtaining an overall accuracy of 93% in mapping levels of agricultural intensity in China [1]. However, these studies are still limited to the use of medium to coarse spatial resolution data, so the dynamics of arable lands managed by smallholders will be generalized.

Several obstacles will be encountered when mapping of agricultural land is carried out in a relatively small area with high heterogeneity [13]. This heterogeneity of agricultural types can always be found on non-industrial agricultural land (privately-owned agricultural land), which makes it difficult to carry out mapping using remote sensing images with low spatial resolution—such as those derived from the Moderate-resolution Imaging Spectroradiometer (MODIS) and Landsat—due to the large number of mixed pixels [13–15]. In addition, the data are often obstructed by cloud-cover, and cloud-free composites are difficult to generate due to low temporal resolution, especially for data with high and medium spatial resolution.

The availability of imagery from the PlanetScope microsat constellation provides a relatively high spatial resolution (3.7 m) with daily observations. This high spatial and temporal resolution has been used for near real-time monitoring of forests [16], vegetation phenology [17], and for yield prediction for cover and cash crops [18]. However, this has not yet been applied to map cropping intensity. In addition, the PlanetScope data suffers from radiometric and geometric inconsistencies due to sensor upgrades in the constellation [19]. Therefore, a study to explore the capability of time-series PlanetScope data for mapping cropping intensity remains necessary.

Our study aims to explore the accuracy of cropping intensity mapping with PlanetScope time-series data using Magelang District, Indonesia, as a test area. In addition, since various methods have been developed to map cropping intensity using remote sensing data, we also aim to compare the accuracy of using machine learning and distance-based calculation by using a Time-Weighted Dynamic Time Warping (TWDTW) algorithm [20] for mapping crop intensity.

2. Materials and Methods

2.1. Study Area and Datasets

2.1.1. Study Area

The study area of this research covers 34.5 km² and is located in the southern part of the administrative area of Magelang Regency, Central Java Province (Figure 1). In general, Magelang Regency is composed of highlands with an average height of 360 m above sea level (masl), in the form of a basin, and is surrounded by the Merapi, Merbabu, Andong, Telomoyo, and Sumbing volcanoes and the Menoreh Mountains. This area has a fairly diverse set of land uses, but is dominated by agricultural land. These agricultural lands are mostly found on alluvial plains [21], where two major rivers (the Progo and Elo) flow across the region.

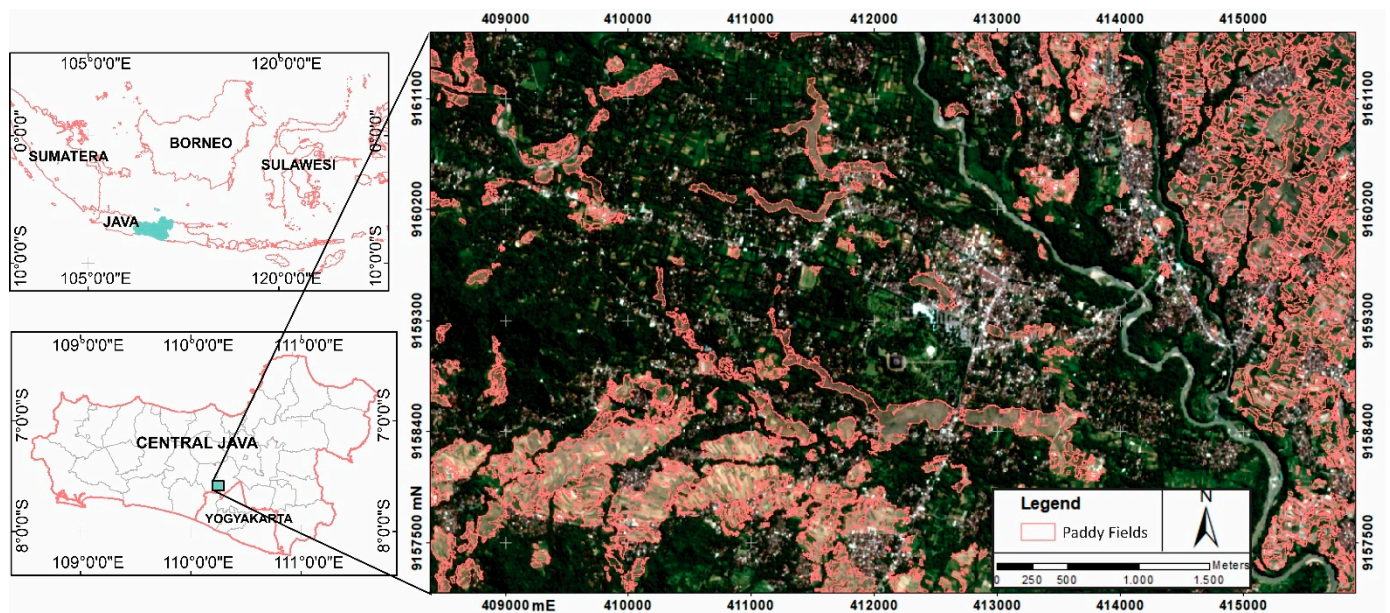


Figure 1. Study area observed in this research, the red polygon represents the paddy fields derived from another study by Kusuma, et al. [22].

The local government recorded an average temperature of 25.62 °C and an average rainfall of 2589 mm/year in Magelang Regency in 2014. The soil in the study area, especially in the area surrounding the Borobudur Temple, is composed of breccia, andesite, tuff, lapilli-tuff, agglomerates, and andesite lava, which formed in old andesite formations. These rocks are the main reason for the development of fertile alluvial sedimentary soils, which has resulted in the extensive use of the region as agricultural land. In 2013, the local government recorded that there were around 78,890 hectares of agricultural land in Magelang Regency.

2.1.2. Datasets

Multispectral PlanetScope time series data, recorded throughout 2019 and obtained from Planet Labs, was used in this study. The data have 4 spectral bands (blue, green, red, and NIR bands) and a spatial resolution of 3.7 m. In this study, the PlanetScope data was corrected to the level of surface reflectance (SR). The study uses 190 scenes of PlanetScope imagery that were pre-processed to generate 12 months of data.

Image preprocessing was performed on the PlanetScope imagery using SR correction. Image preprocessing was carried out in four main stages: (1) cloud masking; (2) relative calibration (the first histogram matching); (3) monthly median composites; and (4) monthly relative calibration (the second monthly histogram matching). Cloud masking was achieved by performing multispectral classification using a Random Forest (RF) algorithm to create a classification model of cloud cover, cloud shadow, and non-cloud, which was applied to all images.

Relative calibration was carried out to calibrate/normalize the pixel values of the PlanetScope image scenes, so that a more representative contrast to the actual appearance of objects was achieved and to enable comparison between images. In this study, we used two stages of histogram matching to normalize inconsistencies in PlanetScope pixel values. The first stage was performed after the cloud masking step, prior to making monthly median composites, by using one image in each one-month time recording period as a reference for all images in the same one-month time recording period. The second histogram matching was then performed after producing the monthly median composite images (which numbered 12 images), by taking a single one-month image as a reference for all one-month images, so that the pixel values of all monthly median composite images were comparable. The histogram matching process was conducted in RStudio using the 'raster' and 'RStoolbox' packages [23,24].

In addition, the paddy field data which were obtained from Kusuma et al. [24] Kusuma et al. [22], a study which used Worldview-3 data and Object-Based Image Analysis, was used to mask non-paddy fields from the analysis. The accuracy of the paddy field map reached 95% [22].

2.2. Classification Methods

The methodological workflow was composed of the following steps: (1) data pre-processing to remove cloud cover and shadows, and also to create consistent pixel values for all the PlanetScope imagery using a histogram matching process; (2) data processing, which consisted of producing the NDVI time-series, the normalized difference water index (NDWI), the k-means clustering model, and training and validation samples; (3) classifying data, where TWDTW, Extreme Gradient Boosting (XGB), RF, and Extremely Randomized Trees (ET) were used to classify objects; and (4) accuracy assessment, carried out to assess the classification accuracies obtained by TWDTW, XGB, RF, and ET (Figure 2).

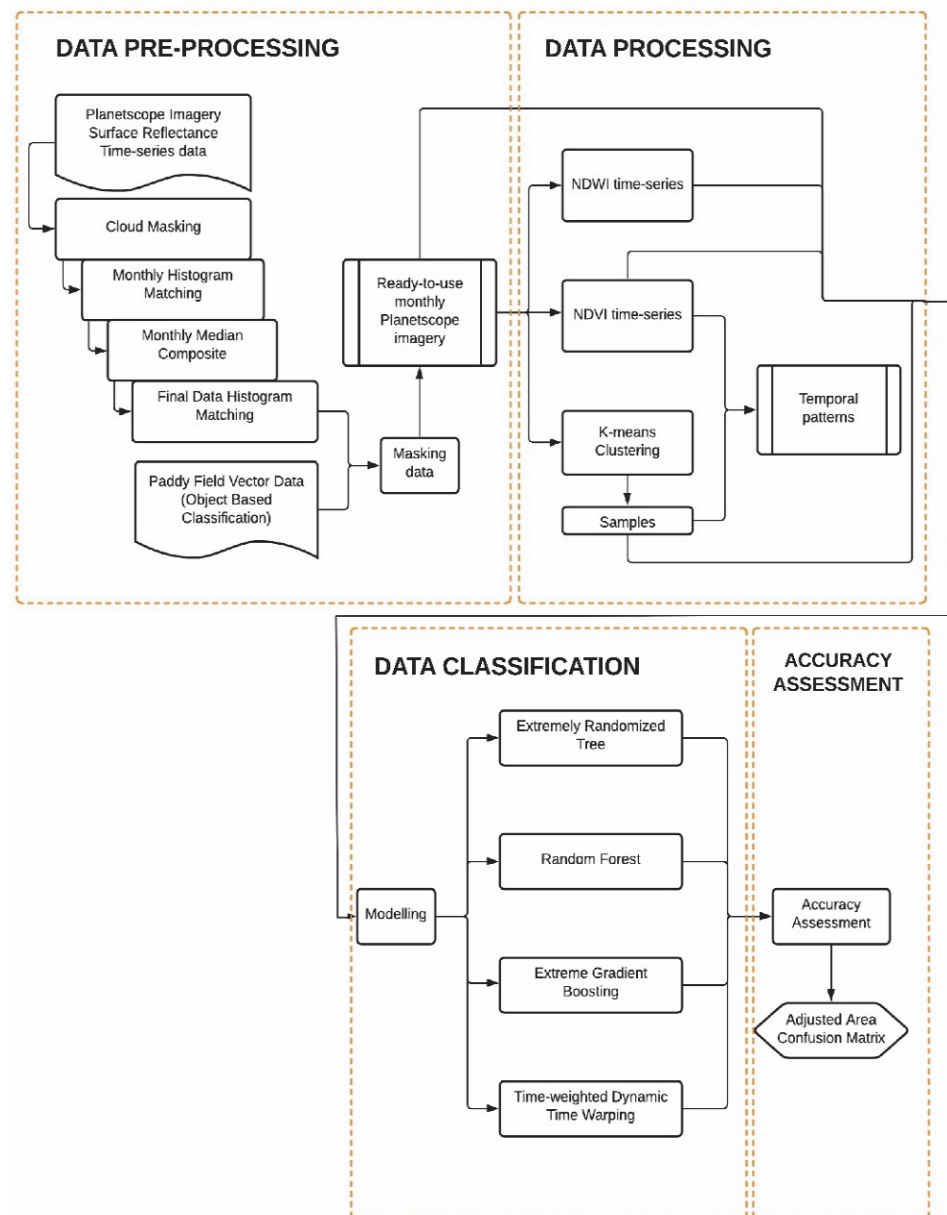


Figure 2. The workflow of the method used.

2.2.1. K-Means Clustering

K-means clustering is a method that produces conclusions from data sets using only vector data input without producing a defined output class. The purpose of k-means clustering is to group similar pixels into one class by finding the underlying pattern [25]. Therefore, a fixed number of classes (k) is needed in the dataset used [25]. K-means clustering was used to indicate the cropping intensity classes on the time-series PlanetScope imagery. The number of classes (k) used in this method was four. The resulting classes in Figure 3 were used as a reference for determining the training samples that would be carried out when making the classification model using machine learning.

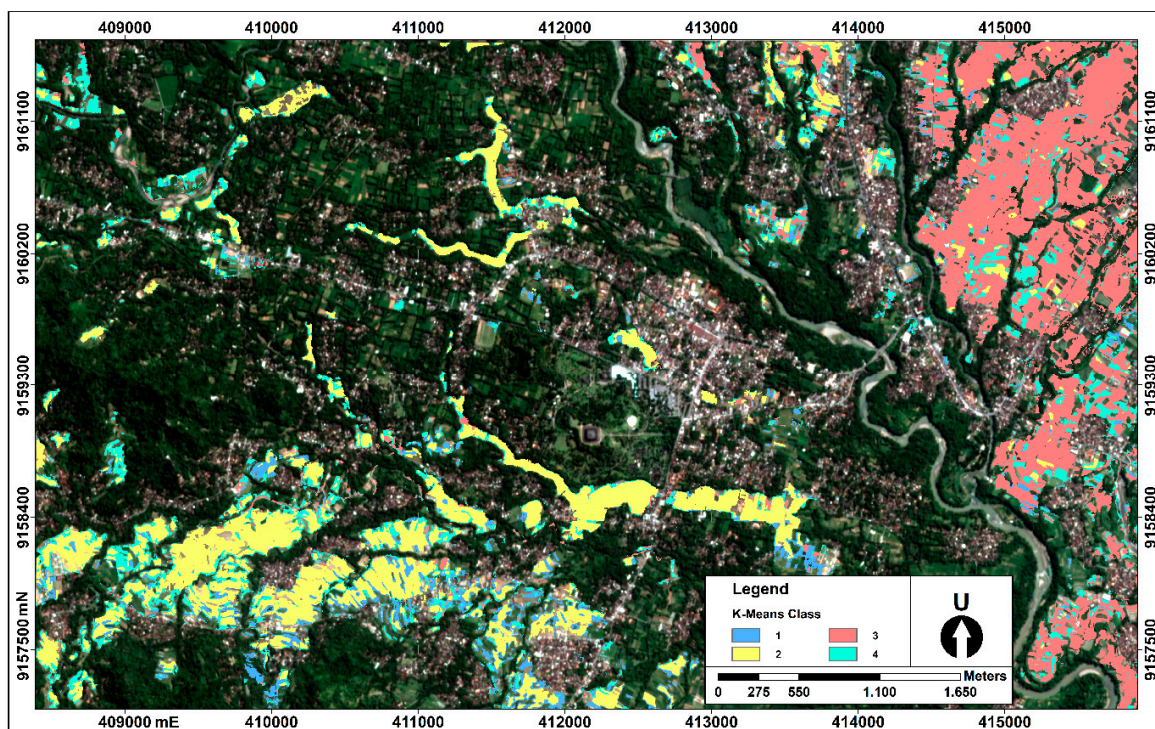


Figure 3. K-means clustering in the study area used to define the fieldwork campaign.

2.2.2. Training Data Collection

Training samples are used as input when making a classification model. Training samples were obtained after the k-means model was created to determine variations in the phenology of the paddy fields. The determination of the training sample was carried out by visual observation of the fluctuations of pixels' reflectance in the NDVI time-series index, which was considered to represent the reflection of the paddy field from the beginning to the end of the year. Training samples were grouped into four classes according to the k-means model, with a total number of 321 points. The training sample was then used as input for creating the TWDTW, XGB, RF, and ET classification models.

2.2.3. Time-Weighted Dynamic Time Warping (TWDTW)

Time-Weighted Dynamic Time Warping (TWDTW) [20] is a method based on Dynamic Time Warping (DTW) that utilizes time weights in the DTW distance calculation formula [26]. The utilization of the DTW method for time-series analysis using remote sensing imagery was initially carried out by Maus, Appel and Giorgino [20], due to several factors that cause non-uniformity in time-series data. The function in this process runs by looking at the time lag between points in one sequence and points in another sequence. If the time lag between points is large, then the weight of the time given is also large [26].

The first step of the TWDTW method was determination of the temporal pattern of agricultural plant phenology using a vegetation index transformation of the Normalized Difference Vegetation Index (NDVI) with the 'dtwsat create-patterns' function available in the 'dtwsat' RStudio package [20]. The next step was TWDTW classification using the logistic function. The logistic function was chosen because it proved to have a low penalty for a small-time curve and a high penalty for a large-time curve. This allows the logistic function to provide better performance, in terms of accuracy, compared to the TWDTW linear function [20]. The parameter of weight function was applied to all classes with the same value [20]. The values of α and β were used to construct the time curve. α is the amount of time-weighting given, while β is the amount of time used as a reference when assigning weight. The value used for β in this study was 50, while the value used for α was -0.1 . These values were chosen based on the best results from the research conducted by Belgiu and Csillik [2].

The computational time taken by TWDTW classification became a concern, especially in the pixel-based classification method where each pixel was treated independently. Therefore, a scheme was needed to reduce the computational time taken. The inputs used in this modeling were 12 temporal images with four spectral bands (blue, green, red, and NIR), NDVI, and NDWI. Each variable was divided into 10 tiles and the classification process was run using parallel computing. The TWDTW classification was run on a configuration using 10 out of 12 cores, with 2666 MHz processing speed and 16-GB of memory. The total computational time for producing the TWDTW classification model was 7.5 h.

2.2.4. Machine Learning Classification

Machine learning is a field of study that uses computational algorithms to transform empirical data into usable models. Machine learning takes datasets, processes them, and tries to find causal variables [27]. There are several methods in machine learning, and these can be divided into three categories according to the learning processed used: supervised, semi-supervised, or unsupervised [27]. The methods can also be divided into four categories: classification, clustering, regression, and anomaly detection. One of the machine learning methods that is most often used is tree-based ensemble machine learning [27].

Some of the methods included in tree-based ensemble machine learning include Extreme Gradient Boosting (XGB), Random Forest (RF), and Extremely Randomized Tree (ET). XGB is performed by generating multiple models by assigning each base classifier to a random subsample of the original dataset, and then combining the results from all models to determine the final prediction [28]. XGB is thought to perform well even when compared to the deep learning method [29]. RF is performed by creating many classification trees with a bootstrap; this is used as a sampling technique to train a classification tree that aims to find a random subset of variables to separate each node [30]. There are two important parameters in classification processes which use the RF method. The two parameters are the variable (Mtry) and the number of sample data (Ntree). ET is performed by creating clusters of unpruned decision trees based on the traditional top-down method [31]. The ET algorithm tries to increase the randomization of the trees, so that the trees' diversity can be increased—this can substantially increase the accuracy of the prediction [32].

The samples collected for constructing the machine learning model in Figure 4, were divided into a training dataset (80%) and a test dataset (20%). In addition, prior to conducting the analysis, a variable selection step was conducted to filter out the insignificant variables by using Boruta [33], which was able to filter out the high-dimensional variables to obtain only the important variables [34]. In addition, hyperparameter tuning was conducted in this study by using grid search for RF and random search for ET and XGB, since more parameters are needed for those algorithms. The machine learning process was conducted using the 'caret' package in RStudio [35].

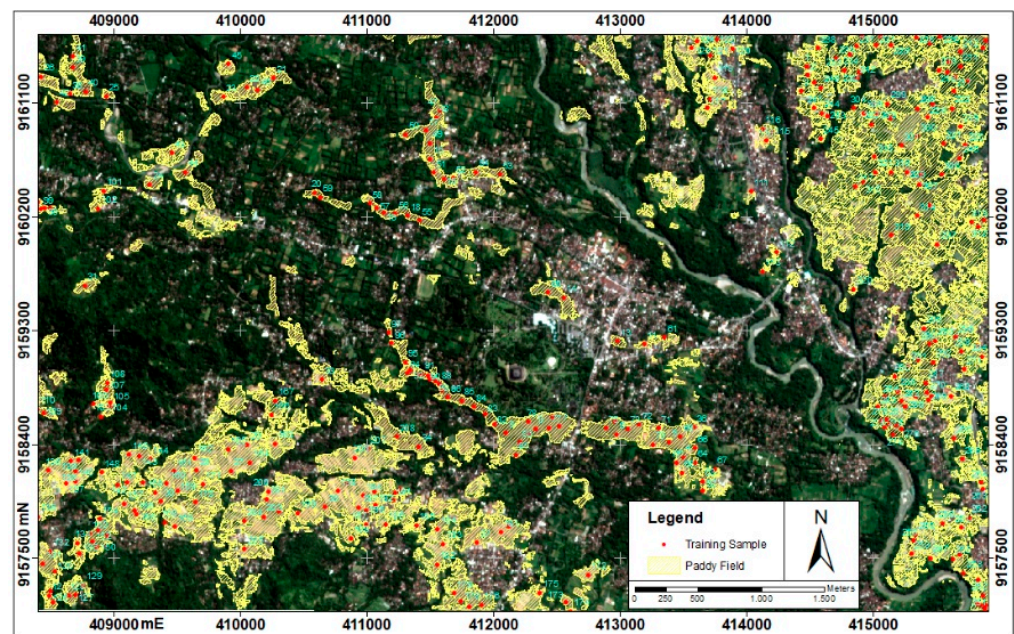


Figure 4. Training samples distribution map. The yellow polygons represent the paddy fields; this information was derived from a study by Kusuma et al. [22].

2.2.5. Accuracy Assessment

The modeling of the validation sample was carried out based on the previously performed k-means classification, which had resulted in four cropping intensity classes. Stratified random sampling [36] was used to determine the number of samples. The minimum number of samples required when using the 90% confidence level was 70. Fieldwork was carried out by conducting interviews with farmers for five days from 13 to 18 September 2021, and a total of 139 samples were obtained as shown in Figure 5. The validation samples were selected after fieldwork, because the PlanetScope image data used were obtained in 2019 while fieldwork could only be carried out in 2021; 94 samples were to be used in the accuracy-test process.

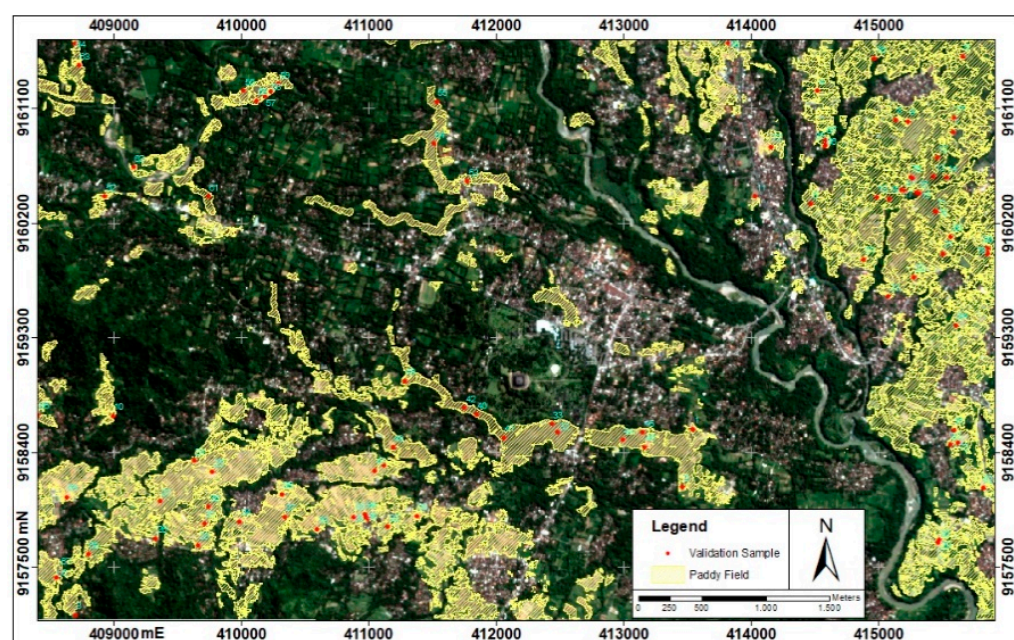


Figure 5. Distribution map of validation samples. The yellow polygons represent paddy fields, and were derived from a study by Kusuma, Arjasakusuma, Rafif, Saringatin, Wicaksono, and Aziz [24].

The accuracy assessment was performed by using an error matrix calculation. An error matrix is a simple tabulation calculation based on classes produced by classifying remote sensing data against reference data for the sample tested [37]. The main components generated in an error matrix are overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA). These components are used as a reference when assessing the level of accuracy of the resulting model. OA, UA, and PA can be formulated as follows:

$$P_j = \frac{P_{jj}}{P_j} \quad (1)$$

$$U_i = \frac{P_{ii}}{P_i} \quad (2)$$

$$OA = \sum_{j=1}^q P_{jj} \quad (3)$$

P_j is the PA obtained from the proportion of samples in a particular class in the field that are also mapped as that class (P_{jj}), with the total sample of that class in the field ($P_{.j}$). Meanwhile, U_i is the UA obtained from the proportion of samples mapped to a certain class and in that class in the field (P_{ii}) with the total sample being mapped to that class ($P_{i.}$), while OA is the OA obtained from the accumulation of P_{jj} in all mapped classes.

Calculation of the uncertainty from the accuracy obtained is also important since accuracy assessment is a product of statistical analysis which possess an uncertainty [38]. One method for doing so is the use of a confidence interval, by calculating the standard error of the estimates made. The use of only the error matrix is not considered to describe the design of the sampling technique that was carried out, the area or proportion of the area of the mapped class, or the uncertainty in the accuracy results obtained [39]. Therefore, in addition to the error matrix used in this study (which included the OA, UA, and PA), calculation of the uncertainty in the accuracy results that were obtained also needed to be performed. In addition, the accuracy calculated for the mapping result classes using a highly accurate classification method can also be biased [37]. Therefore, it is important to calculate the adjusted area of the sample from the field identification results.

The adjusted area can be determined with the following formula:

$$A_j = A_{tot} \sum_i W_i \frac{n_{ij}}{n_i} \quad (4)$$

A_{tot} is the total area of the mapped class. W_i is the proportion of the area that is classified as class i ($A_{m,i}$) with A_{tot} . n_{ij} represents the proportion of samples in a particular class in the field that is also mapped as that class to the total number of samples that is classified as that class on the map (n_i).

The determination of the adjusted area has a standard error that is used to calculate the variance of the adjusted area. The determination of standard error was carried out using a standard error calculation (Cochran, 1977) as follows:

$$S(P_j) = \sqrt{\sum_{i=1}^q W_i^2 \frac{n_{ij}}{n_i} \left(1 - \frac{n_{ij}}{n_i}\right) \frac{1}{n_{ij} - 1}} \quad (5)$$

Then, using a 90% confidence interval, the variance of the adjusted area was obtained using the following formula:

$$A_j \pm 2 \times S(A_j) \quad (6)$$

where

$$S(A_j) = A_{tot} \times S(P_j) \quad (7)$$

The margin of error was defined by a z-score (percentile of the standard normal distribution), where the value of the z-score depends on the confidence interval used (for a 90% confidence level, $z = 1.645$). The final step in the accuracy assessment was measuring the uncertainty in the accuracy results using the variance calculation. This was carried out using the estimated variance for OA, UA, and PA using the following formula:

Overall accuracy (OA)

$$V(OA) = \sum_{i=1}^q W_i^2 U_i (1 - U_i) / (n_i - 1) \tag{8}$$

User’s accuracy (UA)

$$V(U_i) = U_i (1 - U_i) / (n_i - 1) \tag{9}$$

Producer’s accuracy (PA)

$$V(P_j) = \frac{1}{N_j^2} \left[\frac{N_j^2 (1 - P_j)^2 U_j (1 - U_j)}{N_j - 1} + P_j^2 \sum_{i \neq j}^q N_i^2 \frac{n_{ij}}{n_i} \left(1 - \frac{n_{ij}}{n_i} \right) / (n_i - 1) \right] \tag{10}$$

where

$$N_j = \sum_{i=1}^q \frac{N_i n_{ij}}{n_i} \tag{11}$$

N_j is the estimated total number of pixels of class j in the reference data, N_j is the marginal total of class j on the map, while n_j is the total number of samples that were produced for class j on the map.

3. Results

3.1. Relative Calibration of PlanetScope Results

The spectral inconsistency in PlanetScope data requires the data to be calibrated prior to analysis, especially when time-series data is the input for analysis. In this study, we applied a two-stage histogram matching process to obtain spectral values that were comparable across time. The time-series values from the monthly median datasets, both prior to the second histogram matching and after the second histogram matching, using the three samples of unchanged objects, which represented water, built-up areas, and vegetation, are presented in Figure 6. The results show more comparable values for water, built-up and vegetation objects after the second histogram matching in different months. Example of the improvement can be seen from the water object (Figure 6a,b), where the value dip in months 5, 6, and 7 is now having similar values as compared to the other values over year, after the two-stage histogram matching process.

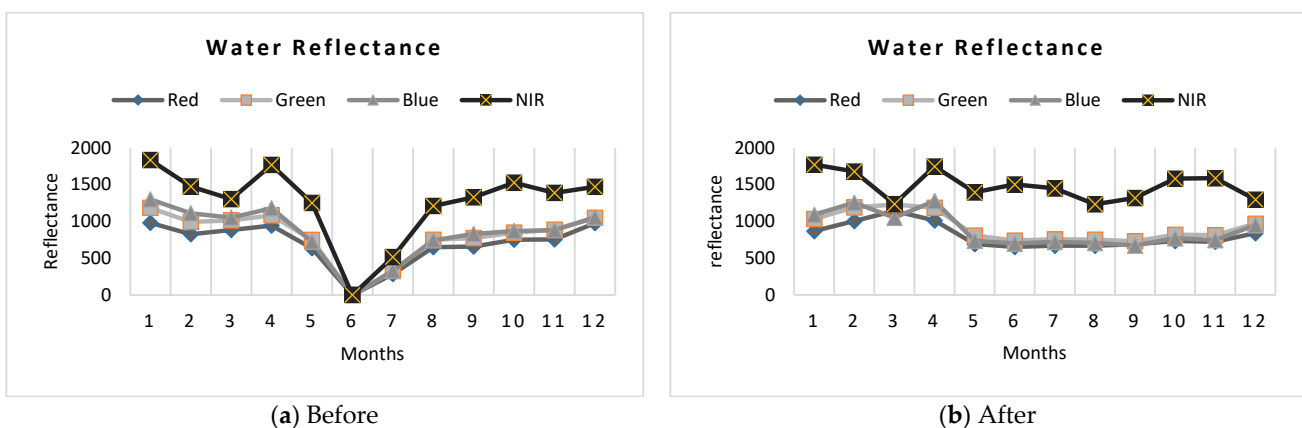


Figure 6. Cont.

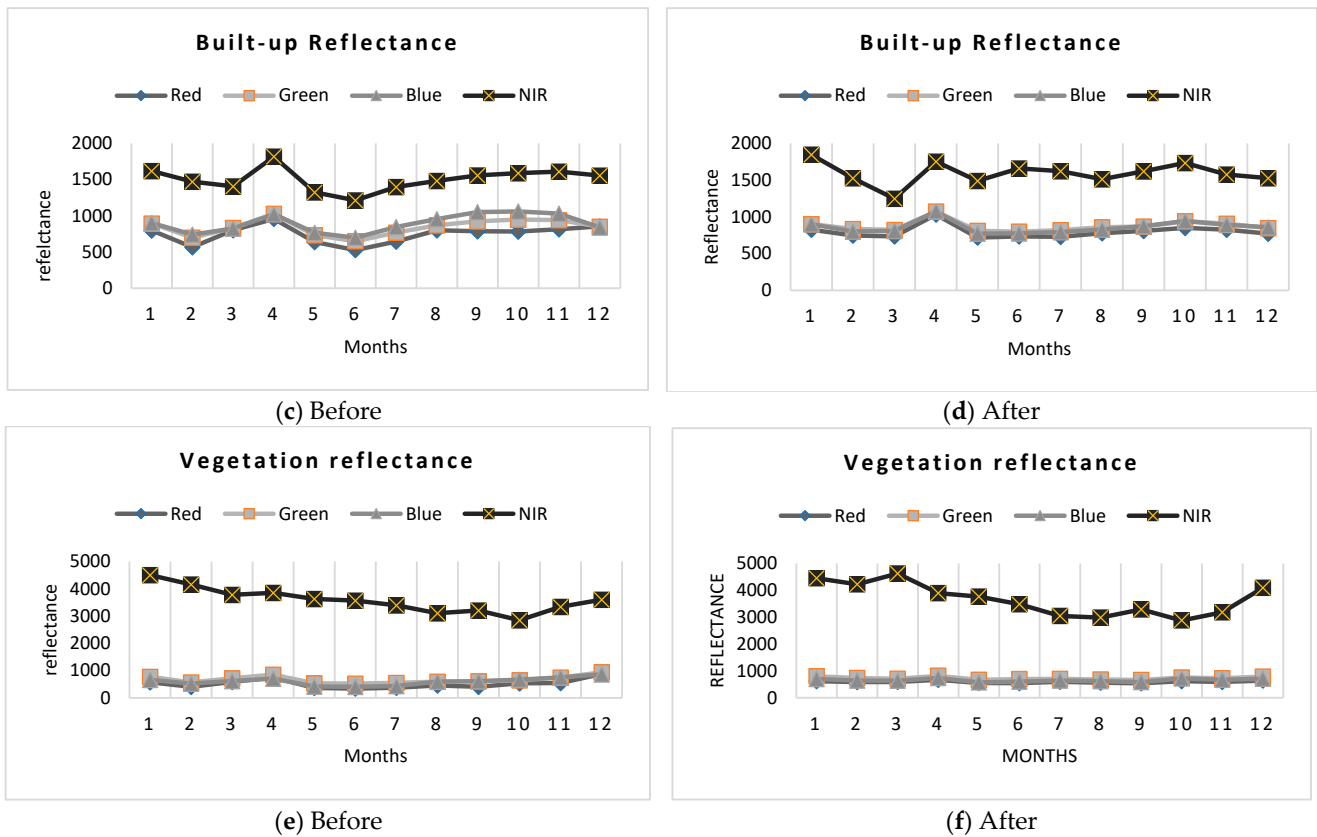


Figure 6. The samples of time-series values from unchanged objects using median composite before the second histogram matching (a,c,e), and after the second histogram matching (b,d,f).

3.2. Temporal Patterns from Different Cropping Intensities

We identified four types of cropping intensities in the study area: once per year, twice per year, thrice per year, and four times per year. To map these cropping intensities, 321 training samples were taken and used as the input for the TWDTW and machine learning classifications. Furthermore, we also identified the average temporal patterns from the spectral bands of PlanetScope, which can be seen in Figure 7a–d. It can be seen from the figures that the temporal patterns in NIR were influenced by the greenness and density level of vegetation [40], displaying the most distinctive patterns between different cropping intensities; this is beneficial for identifying different cropping intensities in PlanetScope imagery. The other bands (red, green, and blue) showed a relatively similar temporal pattern between different cropping intensities.

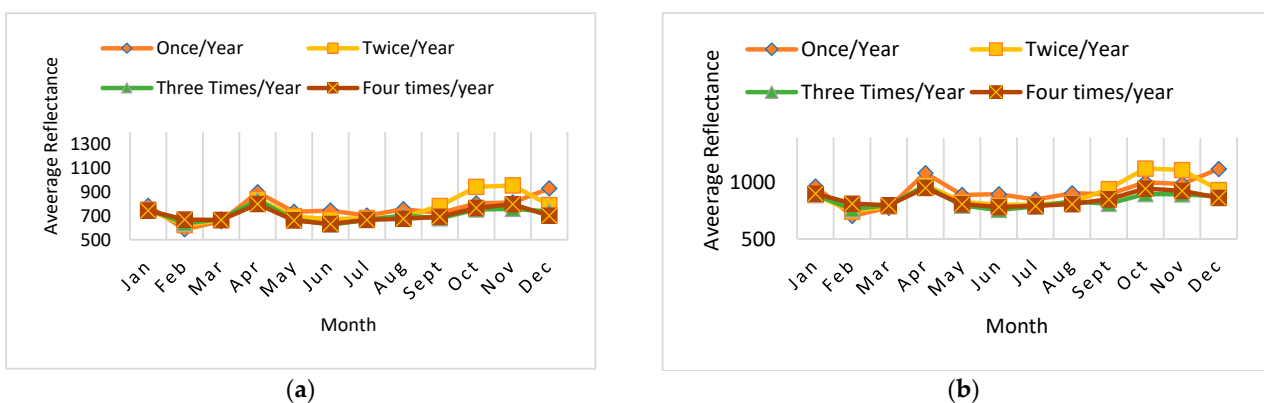


Figure 7. Cont.

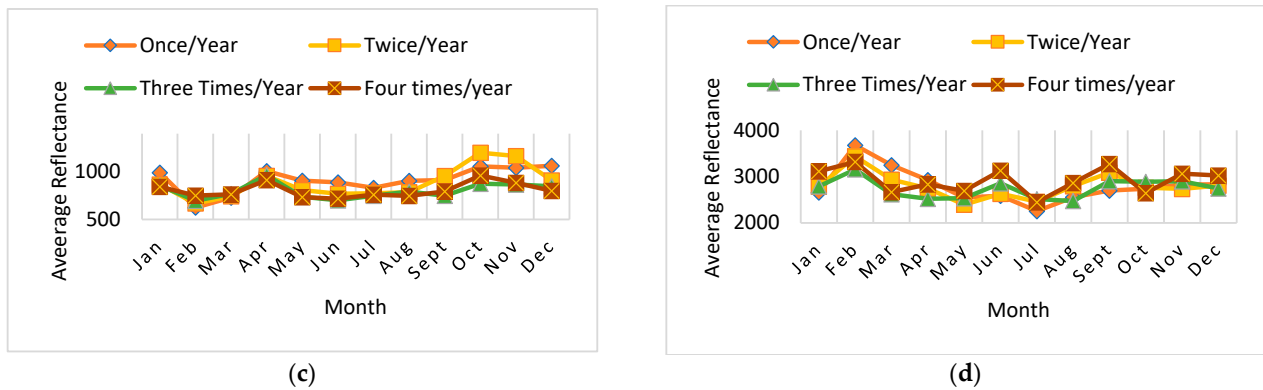


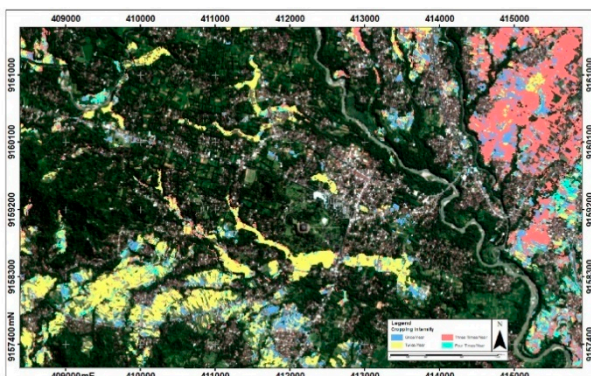
Figure 7. The average reflectance temporal patterns from different cropping intensities identified in the study areas. (a) Average Reflectance in Blue Band; (b) Average Reflectance in Green Band; (c) Average Reflectance in Green Band; (d) Average Reflectance in NIR Band.

3.3. Classification Results

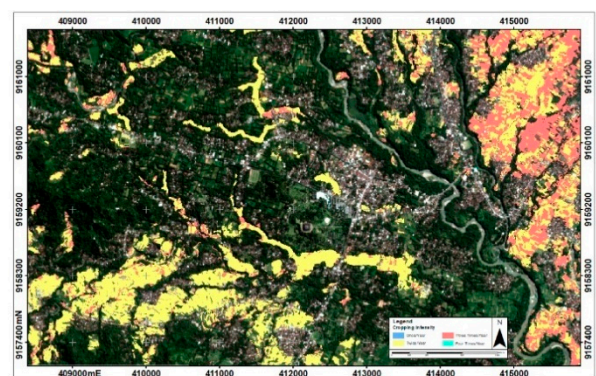
The classification results from using TWDTW and the RF, XGB, and ET machine learning models for mapping cropping intensity can be seen in Figure 8a–d. From the results, we can see that TWDTW classified once per year and thrice per year cropping intensity areas as significantly larger than the other classification results did, while the machine learning algorithms produced similar results for cropping intensity maps. Table 1 shows that all the machine learning results predicted that the study area was dominated by twice to thrice per year cropping intensities, in contrast to the results from TWDTW, which predicted that the areas were dominated by thrice per year and once per year cropping intensities. This difference indicates the need to validate the maps using different methods to obtain the accuracy of each map.

Table 1. The predicted areas based on different algorithms.

| Methods | Once/Year | Twice/Year | Three Times/Year | Four Times/Year |
|---------|--------------------|------------|------------------|-----------------|
| | (km ²) | | | |
| TWDTW | 3705.94 | 2446.64 | 7716.7 | 2157.55 |
| RF | 1.44 | 6531.57 | 5164.37 | 2.61 |
| XGB | 37.01 | 6540.59 | 5363.76 | 75.41 |
| ET | 0.95 | 8251.99 | 3509.56 | 1.42 |



(a)



(b)

Figure 8. Cont.

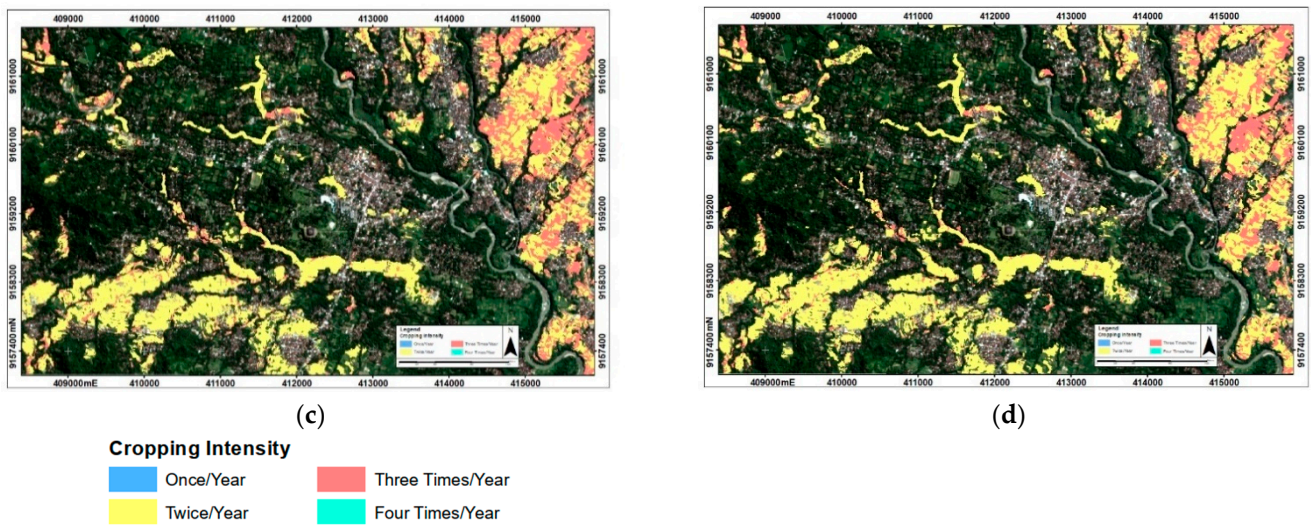


Figure 8. The classification results using tested algorithms in this study, for instance, (a) TWDTW, (b) XGB, (c) RF, and (d) ET.

3.4. Model Accuracy and Variable Importance from Machine Learning

For the machine learning models, model accuracy could also be calculated using the test datasets. The RF model produced a model accuracy of 89.13% with the best parameter of mtry at a value of 10; the ET model produced a model accuracy of 90.16% with mtry of 28 and random cuts at a value of 1; and XGB produced a model accuracy of 86.71%. It is evident that ET produced the best accuracy for classifying the training dataset. However, additional accuracy assessment using an independent dataset is still needed to explore the robustness of the machine learning models.

The machine learning models also provided a list of important variables which were used to generate the models. For the variable importance analysis, all variables were used to construct the model, since feature selection using Boruta determined that all variables were useful for constructing the machine learning model. This information was useful for inferring the ability of different variables when being used in the classification process using machine learning. We can also infer which variables from spectral bands and spectral indices performed better than other variables, and the time information (month) where the cropping intensity mapping was useful. The 15 most important variables from a total of 72 variables are presented in Figure 9a–c.

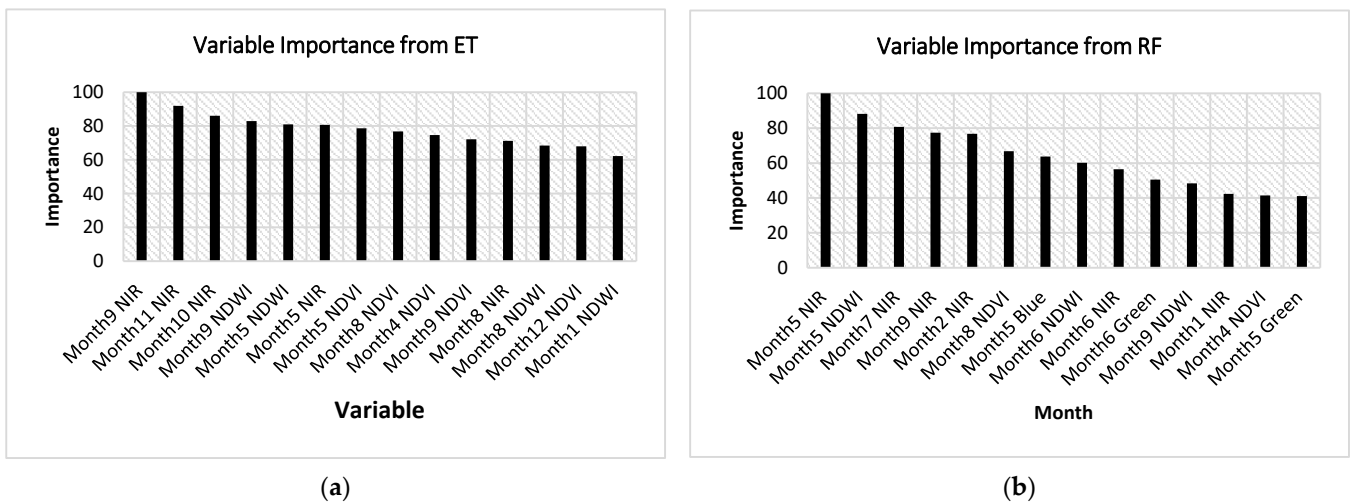


Figure 9. Cont.

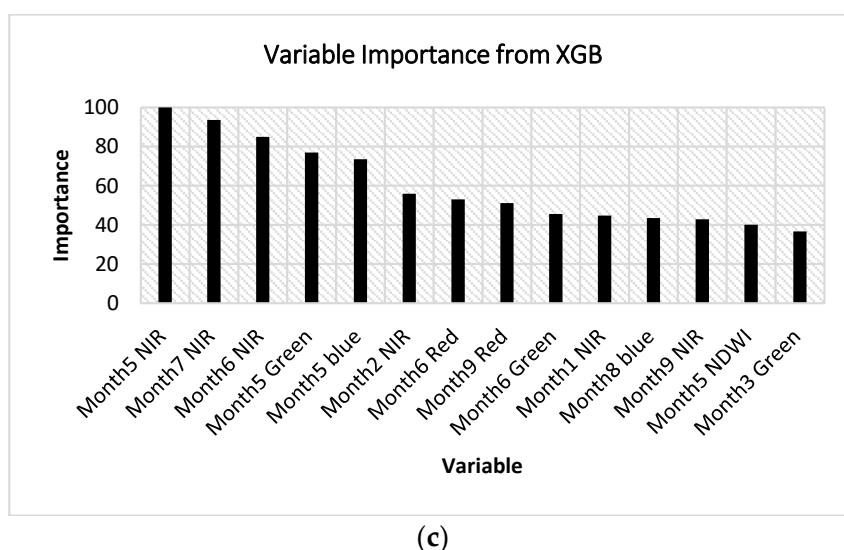


Figure 9. The identified important variables (top 15) in the machine learning models of (a) ET, (b) RF, and (c) XGB.

The results of the variable importance analysis show that the NIR band and NDVI are beneficial for mapping cropping intensity using time-series PlanetScope imagery. This is in line with the ability of the temporal patterns of NIR to distinguish different cropping intensities, as shown in Section 3.2, while NDVI time-series data also has been deemed beneficial to describe the dynamic in the paddy fields, and for crop identification [41,42]. NIR band data captured in September and May are the most important variables used in the ET, RF, and XGB models. In addition, by looking at the timing of the important variables, we can also see that the variables representing the middle of the year (May, June, and July) and end of the year (September, October, and November) are useful for mapping different cropping intensities. This reflects the beginning of planting at the start of the rainy season (end of year) and the beginning of second planting (around the months of April and May).

3.5. Classification Accuracies

From the accuracy assessment using the area-adjusted error matrix, we found that XGB achieved the highest overall accuracy at $95 \pm 4\%$, followed by RF ($92 \pm 5\%$), ET ($87 \pm 6\%$), and TWDTW ($81 \pm 8\%$) (see Table 2). UA generated for the once/year class obtained a value of 70% for all four methods (XGB, RF, TWDTW, and ET). Meanwhile, for the twice/year class, there was a slight difference with the highest value obtained by the RF method (100%) followed by XGB (97.56%), ET (95.12%), and TWDTW (87.8%). UA in the three times/year class was significantly different, with the highest value obtained by the XGB method (92.86%), followed by TWDTW (89.29%), RF (82.14%), and ET, which obtained the lowest value (67.86%). The four times/year class had UA of the same value (60%) for each of the four methods used. The PA generated for the once/year class by all four methods (XGB, RF, TWDTW, and ET) was 100%. Meanwhile, for the twice/year class, there was a difference between the methods, with the highest value obtained by the XGB method (94.23%) followed by RF (87.1%), ET (87%), and finally TWDTW with 53%. The RF method produced the highest PA value for the three times/year class (99.32%), followed by XGB (96.26%), ET (86%), and TWDTW (86%). The four times/year class had the same value of PA (100%) for each of the four methods used. The complete calculation of accuracy can be seen in the Supplementary Materials Section.

Table 2. The producer's, users, and overall accuracies of TWDTW, XGB, RF, and ET methods.

| Methods | Once/Year | | Twice/Year | | Three Times/Year | | Four Times/Year | | OA |
|---------|-----------|----------|------------|---------|------------------|----------|-----------------|----------|---------|
| | UA | PA | UA | PA | UA | PA | UA | PA | |
| XGB | 70 ± 25% | 100 ± 0% | 98 ± 4% | 94 ± 6% | 93 ± 8% | 96 ± 13% | 60 ± 22% | 100 ± 0% | 95 ± 4% |
| RF | 70 ± 25% | 100 ± 0% | 100 ± 0% | 87 ± 8% | 82 ± 12% | 99 ± 1% | 60 ± 22% | 100 ± 0% | 92 ± 5% |
| ET | 70 ± 25% | 100 ± 0% | 95 ± 6% | 87 ± 5% | 68 ± 15% | 86 ± 3% | 60 ± 22% | 100 ± 0% | 87 ± 6% |
| TWDTW | 70 ± 25% | 100 ± 0% | 88 ± 9% | 53 ± 7% | 89 ± 10% | 86 ± 16% | 60 ± 22% | 100 ± 0% | 81 ± 8% |

4. Discussion

The PlanetScope constellation provides daily temporal resolution data with a relatively high spatial resolution of 3.7 m. Daily temporal resolution is needed to increase the possibility of obtaining a cloud-free observation during the rainy season. The time-series dataset can then be employed for various applications; one is to map cropping intensity on agricultural land, which can be used to increase protection of productive agricultural lands from land-use conversion. The problem with PlanetScope is geometric and radiometric inconsistency [19], which can hamper mapping accuracy.

Our study performs a relative radiometric correction by performing two stages of histogram matching to achieve consistent spectral values across the period of observation. Histogram matching has been previously used to calibrate areas under the influence of cloud and cloud-shadows so that they match with surrounding pixels captured in clear conditions [43]. In our study, the two stages of histogram matching calibration were conducted so that pixel values from the same month of observation have similar/comparable values to the consistent median composite values produced, by selecting one image with low cloud cover as the reference for histogram matching, and conducting the second histogram matching by achieving consistent values between the median composite values, and also by using one set of monthly median composite data as the reference.

The calibrated PlanetScope data were then used to map cropping intensity information using the RF, ET, XGB, and TWDTW algorithms, with XGB being the most accurate and robust method, and TWDTW the least accurate method. Our results are different from the accuracy of TWDTW found in the studies of Belgiu and Csillik [2] and Cheng and Wang [44], where TWDTW produced better accuracy than the RF and SVM methods; although, accuracy results from our TWDTW analysis is still within the range of the accuracy reported in both these studies. However, our study is similar to the study conducted by Dadi [45], which reported higher accuracy from using RF than TWDTW for cropland mapping. The lower accuracy of TWDTW in our study may have been caused by the relatively similar temporal patterns of different cropping intensities in the study areas, which may have resulted in misclassification, similar to problem reported by Dadi [45]. TWDTW process also generated longer computation time due to the $O(n^2)$ complexity problem from calculating the per-pixel distance to the training areas, which requires parallel/multi-core process to fasten the analysis [46]. In addition, the classification of TWDTW does not consider the phenology of crops (seasonality and duration) [47]; instead, it relies on the ability from the temporal spectral values from the training areas to classify objects. Therefore, improvement for TWDTW can be made by incorporating the crop phenology information to the weight [48], or by using the modified fuzzy TWDTW classification to account for the uncertainty in the classification process [49].

Machine learning methods perform well, yet faster in this study because they can generate non-linear models by looking at the optimal combinations of variables that are able to distinguish between different cropping intensities. For the next study, the development of deep learning using remote sensing time-series dataset can be tested for mapping crop intensity. Deep neural network architectures, such as one-dimensional Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Temporal

Convolutional Network (TCN), have been used for obtaining phenological information and crop classification [50,51]. Although, a study by Rußwurm and Körner [52] suggested that conventional machine learning algorithms such as RF can give competitive accuracy for classification as long as proper pre-processing (atmospheric correction and data-selection) to the satellite imagery is conducted.

The accuracy reported in our study still emphasizes the potential of high-spatial and high-temporal PlanetScope imagery to map cropping intensity in agricultural land. In the future, high-spatial and temporal satellite imagery data, such as the constellation of Capella Synthetic Aperture Radar (SAR) system [53], may also be tested separately or in combination with PlanetScope for mapping the cropping intensity dynamics. The crop mapping using TWDTW from time-series SAR data using Sentinel-1 has been utilized and studied where the classification accuracy can reach 80.6% [48]. Our study also focused on the cropping intensity information where future studies could target to characterize and map the crop commodities planting cycle in the paddy using temporal signatures from time-series remote sensing data.

5. Conclusions

Our study demonstrates the usefulness of monthly time-series PlanetScope imagery for mapping cropping intensity information at 3.7 m spatial resolution. The two-stage relative calibration using histogram matching was able to increase the consistency of values derived from the time-series PlanetScope imagery. In addition, our study also highlights the higher accuracy of tree-based ensemble machine learning methods such as XGB, RF, and ET, as compared to the TWDTW algorithm. XGB produced the best overall accuracy with $95 \pm 4\%$, followed by RF ($92 \pm 5\%$), ET ($87 \pm 6\%$), and finally TWDTW ($81 \pm 8\%$). The lower overall accuracy from TWDTW was produced due to the similarity in the temporal signatures between different cropping intensities, especially in the visible bands of PlanetScope. XGB also produced a map with higher user's and producer's accuracy; this means that it had a low rate of omission and commission errors, showing the robustness of the generated model. NIR variables and variables representing the months during the end of the year and middle of the year were shown to be beneficial for assisting machine learning model construction for mapping cropping intensity, as can be seen in the variable importance results. The TWDTW algorithm resulted in the lowest accuracy in this study due to the lack of ability of bands other than NIR to differentiate between different cropping intensities. Our study highlights the potential for the high-spatial and high-temporal PlanetScope constellation system to generate cropping intensity maps at higher spatial resolution with high accuracy. Therefore, similar constellation systems developed in the future can be utilized to produce an accurate and details cropping intensity map, to indicate the productive agricultural lands for protection or to intensify the production over the non-productive agricultural areas, especially over the smallholder farming areas.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/land10121384/s1>, Table S1: Confusion Matrix.

Author Contributions: Conceptualization, R.R., S.A. and P.W.; methodology, R.R., S.A. and S.S.K.; validation, R.R.; formal analysis, R.R.; investigation, R.R.; resources, S.A.; data curation, R.R.; writing—original draft preparation, R.R. and S.A.; writing—review and editing, R.R. and S.A.; visualization, S.S.; supervision, P.W.; project administration, G.I.N.; funding acquisition, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Direktorat Sumber Daya, Direktorat Jenderal Pendidikan Tinggi, Riset dan Teknologi, Kementerian Pendidikan, Kebudayaan, and Riset dan Teknologi, grant number 6/E1/KP.PTNBH/2021 and 006/E4/AK.04.PTNBH/2021 with 1637/UN1/DITLIT/DIT-LIT/PT/2021 and 6479/UN1/DITLIT/DIT-LIT/PT/2021.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Data generated in this study are available upon request.

Acknowledgments: The authors are grateful to the PlanetScope data provided under the scheme of PlanetScope for Education and Research Program. In addition, the authors would like to send the gratitude towards the Direktorat Sumber Daya, Direktorat Jenderal Pendidikan Tinggi, Riset dan Teknologi, Kementerian Pendidikan, Kebudayaan, and Riset dan Teknologi for the funding provided to this study. The author would like express our gratitude to the remote sensing laboratory and Geo-AI (Artificial Intelligence) research group of in the Faculty of Geography, Universitas Gadjah Mada for facilitating this research.

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

References

- Liu, L.; Xiao, X.; Qin, Y.; Wang, J.; Xu, X.; Hu, Y.; Qiao, Z. Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine. *Remote Sens. Environ.* **2020**, *239*, 111624. [CrossRef]
- Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [CrossRef]
- Gerland, P.; Raftery, A.E.; Ševčíková, H.; Li, N.; Gu, D.; Spoorenberg, T.; Alkema, L.; Fosdick, B.K.; Chunn, J.; Lalic, N. World population stabilization unlikely this century. *Science* **2014**, *346*, 234–237. [CrossRef] [PubMed]
- Liu, Y.; Yang, Y.; Li, Y.; Li, J. Conversion from rural settlements and arable land under rapid urbanization in Beijing during 1985–2010. *J. Rural Stud.* **2017**, *51*, 141–150. [CrossRef]
- Harini, R.; Yunus, H.S.; Hartono, S. Agricultural land conversion: Determinants and impact for food sufficiency in Sleman Regency. *Indones. J. Geogr.* **2012**, *44*, 120–133.
- Ioja, I.; Onose, D.; Nita, M.; Vanau, G.; Patroescu, M.; Gavrilidis, A.; Saghin, I.; Zarea, R. The conversion of agricultural lands into built surfaces in Romania. *Recent Res. Urban Sustain. Green Dev.* **2011**, *6*, 115–120.
- Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 1–13. [CrossRef]
- Vavorita, B. Decentralization and Rice Production in Bali Province. *J. Public Adm. Stud.* **2018**, *3*, 44–50. [CrossRef]
- Hao, P.-Y.; Tang, H.-J.; Chen, Z.-X.; Le, Y.; Wu, M.-Q. High resolution crop intensity mapping using harmonized Landsat-8 and Sentinel-2 data. *J. Integr. Agric.* **2019**, *18*, 2883–2897. [CrossRef]
- Foerster, S.; Kaden, K.; Foerster, M.; Itzerott, S. Crop type mapping using spectral-temporal profiles and phenological information. *Comput. Electron. Agric.* **2012**, *89*, 30–40. [CrossRef]
- Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogram. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]
- Maus, V.; Câmara, G.; Cartaxo, R.; Sanchez, A.; Ramos, F.M.; De Queiroz, G.R. A time-weighted dynamic time warping method for land-use and land-cover mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3729–3739. [CrossRef]
- Smith, H.W. Evaluating Multiple Sensors for Mapping Cropped Area of Smallholder Farms in the Eastern Indo-Gangetic Plains. Master's Thesis, University of Michigan, Ann Arbor, MI, USA, 2019.
- Cordero-Sancho, S.; Bergen, K.M. Relationships of agricultural land use to an expanded road network within tropical forest landscapes of Cameroon and Republic of the Congo. *Prof. Geogr.* **2018**, *70*, 60–72. [CrossRef]
- Xiao, X.; Boles, S.; Froking, S.; Li, C.; Babu, J.Y.; Salas, W.; Moore, B., III. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sens. Environ.* **2006**, *100*, 95–113. [CrossRef]
- Francini, S.; McRoberts, R.E.; Giannetti, F.; Mencucci, M.; Marchetti, M.; Scarascia Mugnozza, G.; Chirici, G. Near-real time forest change detection using PlanetScope imagery. *Eur. J. Remote Sens.* **2020**, *53*, 233–244. [CrossRef]
- Cheng, Y.; Vrieling, A.; Fava, F.; Meroni, M.; Marshall, M.; Gachoki, S. Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sens. Environ.* **2020**, *248*, 112004. [CrossRef]
- Breunig, F.M.; Galvão, L.S.; Dalagnol, R.; Dauve, C.E.; Parraga, A.; Santi, A.L.; Della Flora, D.P.; Chen, S. Delineation of management zones in agricultural fields using cover-crop biomass estimates from PlanetScope data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *85*, 102004. [CrossRef]
- Frazier, A.E.; Hemingway, B.L. A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery. *Remote Sens.* **2021**, *13*, 3930. [CrossRef]
- Maus, V.; Appel, M.; Giorgino, T. Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis (Software). R Package Version 0.2.6. 2017. Available online: <https://cran.r-project.org/web/packages/dtwSat/index.html> (accessed on 7 December 2021).
- Arjasakusuma, S.; Kusuma, S.; Mahendra, W.; Astriviany, N. Mapping Paddy Field Extent and Temporal Pattern Variation in a Complex Terrain Area using Sentinel 1-Time Series Data: Case Study of Magelang District, Indonesia. *Int. J. Geoinform.* **2021**, *17*, 79–88. [CrossRef]

22. Kusuma, S.S.; Arjasakusuma, S.; Rafif, R.; Saringatin, S.; Wicaksono, P.; Aziz, A.A. Assessment of Image Segmentation and Deep Learning for Mapping Paddy Fields Using Worldview-3 in Magelang, Central Java Provinces, Indonesia. In Proceedings of the 7th Geoinformation Science Symposium, Yogyakarta, Indonesia, 25–28 October 2021.
23. Hijmans, R.J.; Van Etten, J.; Cheng, J.; Mattiuzzi, M.; Sumner, M.; Greenberg, J.A.; Lamigueiro, O.P.; Bevan, A.; Racine, E.B.; Shortridge, A. Package ‘Raster’; R Package Version 3.5.9. 2021. Available online: <https://cran.r-project.org/web/packages/raster/index.html> (accessed on 7 December 2021).
24. Leutner, B.; Horning, N.; Leutner, M.B. Package ‘RStoolbox’; Version 0.1; R Foundation for Statistical Computing: Vienna, Austria, 2017.
25. Duda, T.; Canty, M. Unsupervised classification of satellite imagery: Choosing a good algorithm. *Int. J. Remote Sens.* **2002**, *23*, 2193–2212. [[CrossRef](#)]
26. Zhou, Y. Reduction of Computation Time of Dynamic Time Warping Based Methods Used for Cropland Mapping. Master’s Thesis, University of Twente, Enschede, The Netherlands, 2019.
27. Edgar, T.; Manz, D. *Research Methods for Cyber Security*; Syngress: Cambridge, MA, USA, 2017.
28. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
29. Shwartz-Ziv, R.; Armon, A. Tabular Data: Deep Learning is Not All You Need. *arXiv* **2021**, arXiv:2106.03253. [[CrossRef](#)]
30. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
31. Ampomah, E.K.; Qin, Z.; Nyame, G. Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. *Information* **2020**, *11*, 332. [[CrossRef](#)]
32. Rainforth, T.; Wood, F. Canonical correlation forests. *arXiv* **2015**, arXiv:1507.05444.
33. Kursu, M.B.; Rudnicki, W.R. Feature selection with the Boruta package. *J. Stat. Softw.* **2010**, *36*, 1–13. [[CrossRef](#)]
34. Arjasakusuma, S.; Swahyu Kusuma, S.; Phinn, S. Evaluating variable selection and machine learning algorithms for estimating forest heights by combining lidar and hyperspectral data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 507. [[CrossRef](#)]
35. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B. *The Caret Package*; Vienna, Austria, 2012. Available online: <https://cran.r-project.org/package=caret> (accessed on 22 August 2020).
36. Cochran, W.G. *Sampling Techniques*; Wiley: New York, NY, USA, 1977.
37. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [[CrossRef](#)]
38. Arjasakusuma, S.; Kamal, M.; Hafizt, M.; Forestriko, H.F. Local-scale accuracy assessment of vegetation cover change maps derived from Global Forest Change data, ClasLite, and supervised classifications: Case study at part of Riau Province, Indonesia. *Appl. Geomat.* **2018**, *10*, 205–217. [[CrossRef](#)]
39. Olofsson, P.; Foody, G.M.; Stehman, S.V.; Woodcock, C.E. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **2013**, *129*, 122–131. [[CrossRef](#)]
40. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007.
41. Arjasakusuma, S.; Swahyu Kusuma, S.; Rafif, R.; Saringatin, S.; Wicaksono, P. Combination of Landsat 8 OLI and Sentinel-1 SAR Time-Series Data for Mapping Paddy Fields in Parts of West and Central Java Provinces, Indonesia. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 663. [[CrossRef](#)]
42. Qiu, P.; Wang, X.; Cha, M.; Li, Y. Crop identification based on TWDTW method and time series GF-1 WFV. *Sci. Agric. Sin.* **2019**, *52*, 2951–2961.
43. Helmer, E.H.; Ruefenacht, B. Cloud-free satellite image mosaics with regression trees and histogram matching. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 1079–1089. [[CrossRef](#)]
44. Cheng, K.; Wang, J. Forest-Type Classification Using Time-Weighted Dynamic Time Warping Analysis in Mountain Areas: A Case Study in Southern China. *Forests* **2019**, *10*, 1040. [[CrossRef](#)]
45. Dadi, M.M. Assessing the Transferability of Random Forest and Time-Weighted Dynamic Time Warping for Agriculture Mapping. Master’s Thesis, University of Twente, Enschede, The Netherlands, 2019.
46. De Oliveira, S.S.T.; Rodrigues, V.J.D.S.; Ferreira, L.G.; Martins, W.S. P-twtdtw: Parallel processing of time series remote sensing images using manycore architectures. In Proceedings of the 2018 Symposium on High Performance Computing Systems (WSCAD), São Paulo, Brazil, 1–3 October 2018; pp. 252–258.
47. Belgiu, M.; Zhou, Y.; Marshall, M.; Stein, A. Dynamic time warping for crops mapping. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 947–951. [[CrossRef](#)]
48. Gella, G.W.; Bijker, W.; Belgiu, M. Mapping crop types in complex farming areas using SAR imagery with dynamic time warping. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 171–183. [[CrossRef](#)]
49. Moola, W.S.; Bijker, W.; Belgiu, M.; Li, M. Vegetable mapping using fuzzy classification of Dynamic Time Warping distances from time series of Sentinel-1A images. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102405. [[CrossRef](#)]
50. Tang, P.; Du, P.; Xia, J.; Zhang, P.; Zhang, W. Channel Attention-Based Temporal Convolutional Network for Satellite Image Time Series Classification. *IEEE Geosci. Remote Sens. Lett.* **2021**. [[CrossRef](#)]
51. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of three deep learning models for early crop classification using sentinel-1A imagery time series—A case study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [[CrossRef](#)]

-
52. Rußwurm, M.; Körner, M. Self-attention for raw optical satellite time series classification. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 421–435. [[CrossRef](#)]
 53. Farquharson, G.; Woods, W.; Stringham, C.; Sankarambadi, N.; Riggi, L. The capella synthetic aperture radar constellation. In *Proceedings of the EUSAR 2018—12th European Conference on Synthetic Aperture Radar*, Aachen, Germany, 4–7 June 2018; pp. 1–5.