Article

Oil Palm (*Elaeis guineensis*) Mapping with Details: Smallholder versus Industrial Plantations and their Extent in Riau, Sumatra

Adrià Descals [1] , **Zoltan Szantoi** [2,3,*] , **Erik Meijaard** [4,5,6] , **Harsono Sutikno** [7] , **Guruh Rindanata** [7] **and Serge Wich** [8,9]

[1] CREAF, Cerdanyola del Vallès, 08193 Barcelona, Catalonia, Spain; a.descals@creaf.uab.cat
[2] European Commission, Joint Research Centre, 20127 Ispra, Italy
[3] Stellenbosch University, Stellenbosch 7602, South Africa
[4] Borneo Futures, Bandar Seri Begawan BA 2711, Brunei Darussalam; emeijaard@gmail.com
[5] Durrell Institute of Conservation and Ecology, University of Kent, Canterbury CT2 7NR, UK
[6] School of Biological Sciences, University of Queensland, St Lucia, QLD 4072, Australia
[7] Austindo Nusantara Jaya Tbk, Kawasan Mega Kuningan, Jakarta 12950, Indonesia; harsono.sutikno@anj-group.com (H.S.); guruh.rindanata@anj-group.com (G.R.)
[8] School of Biological and Environmental Sciences, Liverpool John Moores University, James Parsons Building, Byrom, Street, Liverpool L3 3AF, UK; S.A.Wich@ljmu.ac.uk
[9] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park, 904, 1098 XH Amsterdam, The Netherlands
* Correspondence: zoltan.szantoi@ec.europa.eu; Tel.: +39-0332-783656

check for updates

Abstract: Oil palm is rapidly expanding in Southeast Asia and represents one of the major drivers of deforestation in the region. This includes both industrial-scale and smallholder plantations, the management of which entails specific challenges, with either operational scale having its own particular social and environmental challenges. Although, past studies addressed the mapping of oil palm with remote sensing data, none of these studies considered the discrimination between industrial and smallholder plantations and, furthermore, between young and mature oil palm stands. This study assesses the feasibility of mapping oil palm plantations, by typology (industrial versus smallholder) and age (young versus mature), in the largest palm oil producing region of Indonesia (Riau province). The impact of using optical images (Sentinel-2) and radar scenes (Sentinel-1) in a Random Forest classification model was investigated. The classification model was implemented in a cloud computing system to map the oil palm plantations of Riau province. Our results show that the mapping of oil palm plantations by typology and age requires a set of optimal features, derived from optical and radar data, to obtain the best model performance (OA = 90.2% and kappa = 87.2%). These features are texture images that capture contextual information, such as the dense harvesting trail network in industrial plantations. The study also shows that the mapping of mature oil palm trees, without distinction between smallholder and industrial plantations, can be done with high accuracy using only Sentinel-1 data (OA = 93.5% and kappa = 86.9%) because of the characteristic backscatter response of palm-like trees in radar scenes. This means that researchers, certification bodies, and stakeholders can adequately detect mature oil palm stands over large regions without training complex classification models and with Sentinel-1 features as the only predictive variables. The results over Riau province show that smallholders represent 49.9% of total oil palm plantations, which is higher than reported in previous studies. This study is an important step towards a global map of oil palm plantations at different production scales and stand ages that can frequently be updated. Resulting insights would facilitate a more informed debate about optimizing land use for meeting global vegetable oil demands from oil palm and other oil crops.

1. Introduction

Southeast Asia is one of the most rapidly deforesting regions in the world [1]. As such, deforestation has been well-documented in the region, including its drivers, which are mostly due to various tree cash crops [2]. It is estimated that, by 2012, Indonesia lost up to 83% more primary forest than the Amazon region [3]. A major driver in the region is oil palm, which is still expanding at a high rate, due to the favorable climate and policy conditions [4]. While the majority of these expansions are truly large scale (i.e., industrial) driven by companies, local farmers are also expanding their plantations (i.e., smallholders) [5]. The global extent of industrial-scale plantations was recently estimated at 18.7 million ha (mha) [5], but the extent of smaller-scale plantations remains largely unknown. Mosnier et al. [6] report that, as large scale oil palm plantation developers increasingly comply with sustainability standards, the area cultivated by small-scale producers will likely increase. Smallholders have limited financial means to clear and replant old plantations, which are generally replanted after 25 years of production. For such reasons, there is a risk that smallholders abandon their plantations after one production cycle and, eventually, establish new plantations in other areas, thereby producing a new wave of deforestation. Although this is not yet happening [7].

The extent of industrial-scale plantations in Southeast Asia, particularly in Indonesia and Malaysia, has been the focus of several studies that have relied on medium-resolution (10–30 m/pixel) satellite images. There has been a multitude of methods used to map industrial plantations that have various levels of automation [8–12]. These studies use radar satellite data (L-band in ALOS PALSAR and C-band in Sentinel-1) as the main source for the classification of oil palm plantations. Closed oil palm stands present a characteristic radar backscatter, which entails a high separability from other tropical plantations [8]. Concretely, oil palm plantations show a higher backscatter than other vegetation types in the dual cross-polarization bands and, secondly, the single co-polarization bands and the dual cross-polarization bands present a large difference. More recently, such analyses have moved to cloud processing platforms, holding various freely available satellite datasets, in an attempt to fully automate mapping [9,10].

To date, several studies have attempted to map oil palm plantations without discriminating plantations based on their type (industrial versus smallholder) and age. One study has attempted to discriminate industrial and smallholder oil palm plantations [11], but its study area had a small coverage (<50 km^2) and considered only oil palm plantations in peatlands, which in consequence, did not allow the generalization of the results at a regional scale for the entire Sumatra island. Due to the difficulty in detecting young oil palm plantations, most of these studies have mapped mature (>3 years, but potentially >8 years in some studies [12]) oil palm stands, apart from a recent work that discriminated both classes [13]. As a result of these studies, the global extent of industrial plantations is relatively well known, although it remains based on a patchwork of a large number of different studies with different methodologies. Consequently, there is no standardized global map of industrial-scale mature oil palm plantations [5]. The extent of smallholder plantations is, furthermore, poorly quantified and estimates the proportion of smallholder plantations vary extensively between countries, with Nigeria at the high end (~94%) and areas in Indonesia and Malaysia containing ~40% [5].

Hence, it is apparent that there are several challenges to overcome before attaining a globally automated map of industrial and smallholder plantations, on which young and mature stands are mapped at frequent intervals (e.g., one year). In this study, we aim to show the suitability of radar (Sentinel-1) and optical (Sentinel-2) satellite data for the automated detection of oil palms and for discriminating industrial and smallholder plantations. Thus, we collected a sample dataset in Riau province (Indonesia) and assessed the performance of Sentinel-1 and Sentinel-2 for the mapping of

oil palm plantations. The classification model was implemented in a large scale cloud processing system [14]. Such an automated method of detecting industrial and smallholder plantations has the benefit that it can potentially lead to a global oil-palm plantation map. It can be adopted by certifying bodies, such as the Roundtable on Sustainable Oil Palm (RSPO) and additionally help to create a transparent and free mapping tool for all of the involved stakeholders. It would also help in ongoing debate about how to best meet future vegetable oil demands in the world, and which crops to use, given their respective yields and social and environmental impacts [15].

## 2. Materials and Methods

2.1. Data and Study Area

2.1.1. Sentinel Missions

This study uses Synthetic Aperture Radar (SAR) data, collected by Copernicus' Sentinel-1 satellites (A and B), and optical multispectral images from Copernicus' Sentinel-2 satellites (A and B); both missions were launched by the European Space Agency. The Sentinel-1 mission provides SAR scenes in the C-band with a global revisit time of 3 days. Sentinel-1 SAR Ground Range Detected (GRD) with the highest spatial resolution (10 meters) was used in this study. The scenes were taken with the interferometric wide swath mode in both ascending and descending orbits. The first scenes of Sentinel-1 for the study area were taken in October 2014.

The Sentinel-2 mission is composed of two multispectral satellites with various spectral resolutions (10, 20, and 60-meter resolution). We used the 10- and 20-meter resolution bands of level-1C orthorectified Top-Of-Atmosphere reflectance. The first images of Sentinel-2 for the study area were taken in October 2016. Due to the timing and the constellation of the four satellites, this study focuses on the second half of the year 2018 when both missions operated at their full capabilities.

2.1.2. Training and Validation Data Collection

The training and validation dataset was based on 4000 points that were collected in Riau province. We discarded the points that were located within pixels that presented missing values in the Sentinel-1 or Sentinel-2 composite images. Thus, only 3448 points were eventually used for the training and validation of the algorithm (Figure 1). The points were subdivided for a 10-fold cross-validation (10% of the points for training and 90% of the points for validation). We collected the points with a visual interpretation over the study area using the high-resolution (sub-meter) Google Earth orthophotos, Sentinel-1, and Sentinel-2 images. The points were collected with the geometry editing tools, that are available on the Google Earth Engine (GEE) code editor. Another feature of the GEE code editor, which helped in visual interpretation, is the fast visualization of satellite imagery. The classes that we visually distinguished were: 1. Industrial mature oil palm plantations (621 points—18.01%), 2. Industrial young oil palm plantations (727 points—21.08%), 3. Smallholder mature oil palm plantations (1016 points—29.47%), 4. Smallholder young oil palm plantations (244 points—7.08%), and 5. Other land uses that are not oil palm plantations (840 points—24.36%).
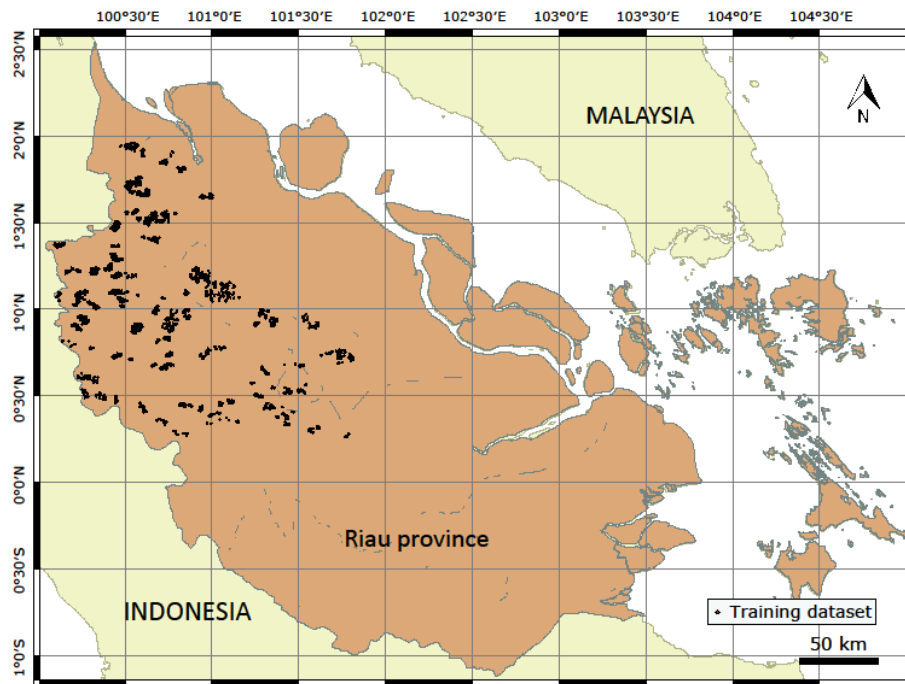
**Figure 1.** Riau province, and the locations of the training and validation points used in this study.

The visual interpretation was supported with geotagged photos taken around the village of Binanga (North Sumatra), adjacent to Riau province, on 25 and 26 January 2019. The photos were taken in 23 spots and depicted both industrial and smallholder oil palm plantations at different growth stages (Figure 2). Such geotagged photos were useful for identifying the spatial patterns observed from satellite data that are characteristic of the different types of oil palm plantations. Here, the canopy closure is the feature that makes possible the discrimination between mature and young plantations: Young plantations present an open canopy until the age of 3 to 8 years (Figure 2b,c), while the canopy in mature plantations cover most of the surface (Figure 2d). Therefore, it was assumed that an oil palm plantation gets to its maturity when it reaches the full canopy coverage, at a varying threshold of 3 to 8 years.
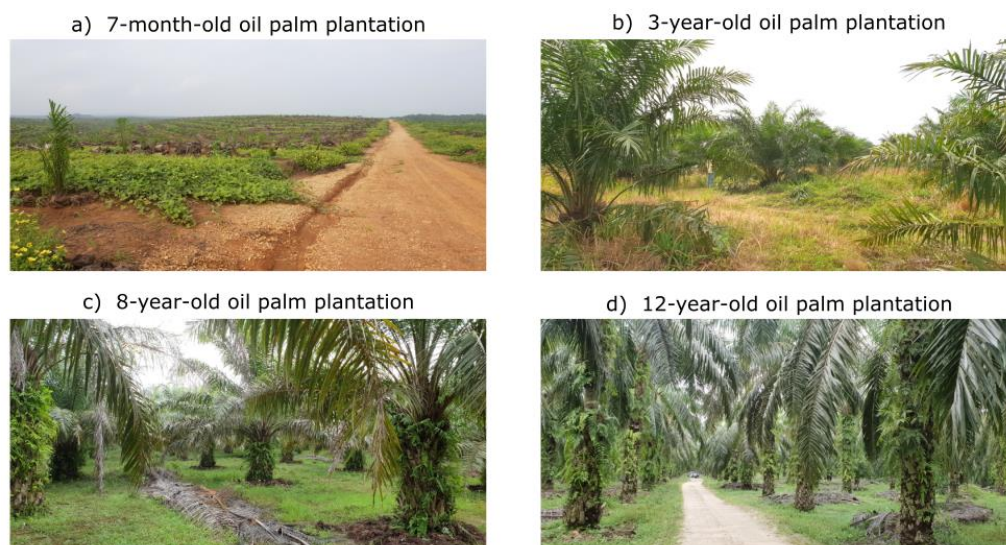
Figure 2. (**a**–**d**) Field photos around the village of Binanga (North Sumatra). The photos depict oil palm plantations of different ages. Note that oil palm stands reach the full canopy coverage between the ages of 3 and 8 years (photos © Erik Meijaard).

Industrial oil palm plantations were visually recognized through the presence of dense trail and road networks. In flat surface plantations, the harvesting trails are usually built in straight lines and, thus, form a rectilinear grid (Figure 3a). In contrast, the industrial plantations that are constructed over steep terrain usually present curvy trails (Figure 3b,c). Similar trail networks can be found in other plantations, such as pulpwood (Figure 3d). However, oil palm can be discriminated from other plantations due to its characteristic backscatter response in radar images [8]. Smallholders, unlike industrial plantations, can be distinguished with satellite images because they often form a landscape mosaic, composed of different small plantations such as rubber and pulpwood (Figure 3e). Moreover, smallholder plantations, when presented in clusters, show a less dense trail network than industrial plantations (Figure 3f). One of the reasons for this is that, commonly, smallholder plantations are often planted and, afterwards around the time of the first harvest, the harvesting trails are developed from existing rural non-straight roads. Roads in industrial-scale plantations are instead developed at the start of plantation and, therefore, equidistantly placed for optimal harvesting.
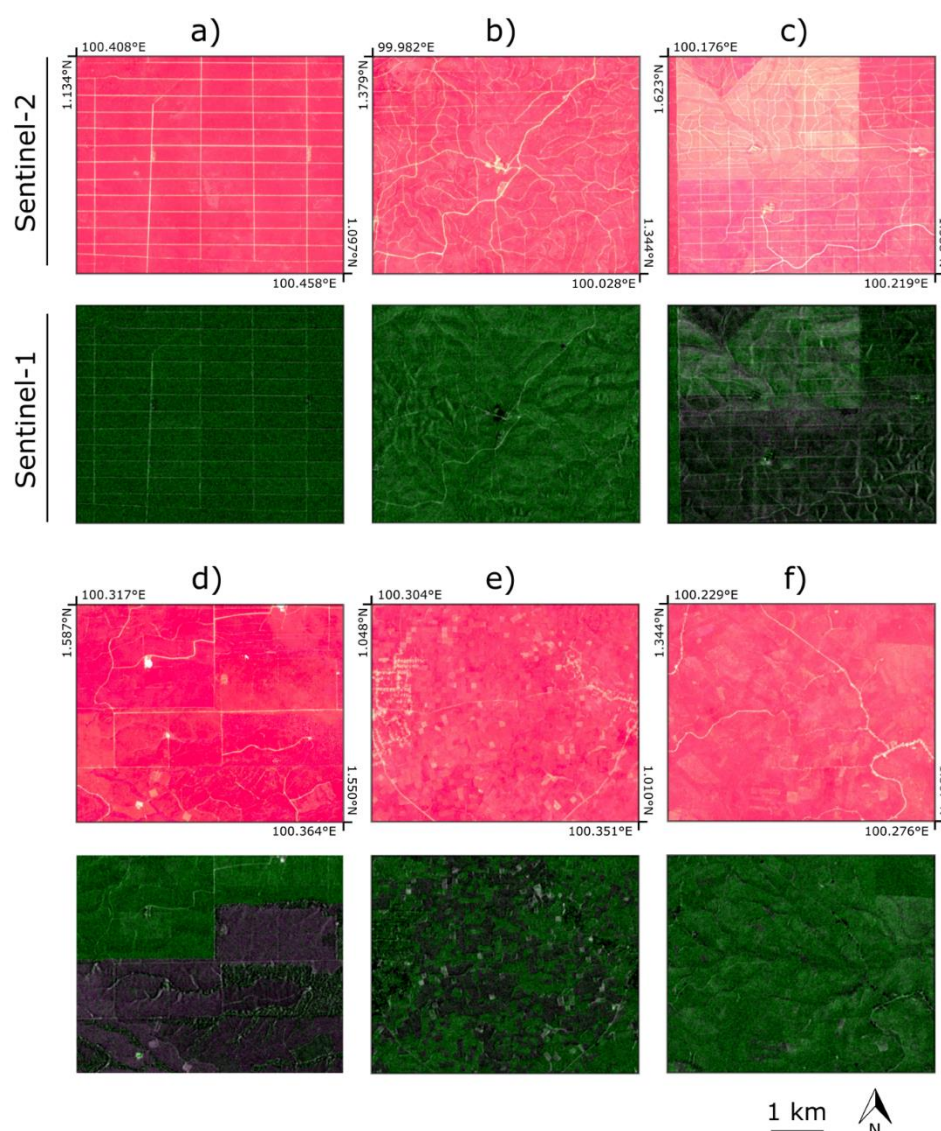
Figure 3. Six areas in Riau province that show (**a**) an industrial oil palm plantation in a flat terrain, (**b**) an industrial oil palm plantation in a hilly terrain, (**c**) industrial young plantation, (**d**) a pulpwood plantation, (**e**) a mosaic of different types of smallholder plantations (mostly oil palm and rubber), and (**f**) a homogeneous area of smallholder oil palm plantations. Images in the upper panel show Sentinel-2 composites (NIR-R-G), while the images in the lower panel are Sentinel-1 composites (VV-VH-VV).

2.1.3. Study Area

While, oil palm plantations can be found in many parts of the tropics, they are mainly grown in Southeast Asia, particularly in Malaysia and Indonesia. Riau province was selected as our area of focus because the region is the largest producer of palm oil in Indonesia, and contributes up to 24% of the national production [1]. Riau province covers the central area of the island of Sumatra and the Riau Archipelago. In this study, we considered the area of the province that covers the island of Sumatra, which represents an area of 84,360 km^2. Riau province has also a particular interest for oil palm mapping because the growth of smallholder plantations is thought to have contributed up to the 21% increase in the overall oil palm plantation area from 2004 to 2009 [1].

2.2. Algorithm Overview

The processing chain for oil palm classification starts with the compositing of daily Sentinel-1 and Sentinel-2 images hosted in Google Earth Engine [14]. In addition to the backscatter and spectral bands, a feature extraction was applied to generate additional images that might improve the classification model. A Random Forest algorithm was trained with four sets of input variables: (1) Sentinel-1 backscatter bands and features, (2) Sentinel-2 spectral bands and features, (3) Sentinel-1 and Sentinel-2 bands without features, and (4) an optimal set of Sentinel-1 and Sentinel-2 features.

Three classification setups were considered depending on the number and typology of classes. Setup I considers the 5 classes explained in the 'Study area' section: 1. Industrial mature plantation, 2. Industrial young plantation, 3. Smallholder mature plantation, 4. Smallholder young plantation, and 5. Other land uses. The classification setup II considers three classes: Mature industrial plantations, mature smallholder plantations, and other land uses including young oil palm plantations. Lastly, the setup III is a binary classification that distinguishes between mature oil palm plantations, without differentiating typology, and the rest of land uses. Figure 4 shows a representation of the algorithm.
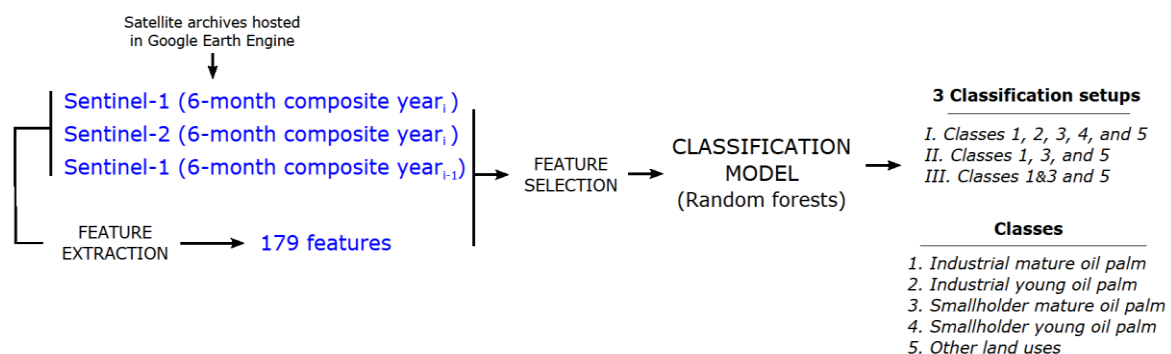
Figure 4. Diagram of the algorithm for oil palm classification.

2.2.1. Sentinel-1 and Sentinel-2 Compositing

The composite images are based on the median value of daily images during a six-month period in Sentinel-2 (year$_1$), and two six-month periods in Sentinel-1 (year$_1$ and year$_{1-i}$). For our case study, the second half of the year 2018 was used for both Sentinel-1 and Sentinel-2, and the second half of the year 2017 for Sentinel-1. The size of the compositing window was set to six months because it was the minimum window size that could produce a cloud-free composite in Riau province.

For Sentinel-1, two backscatter bands were used: Dual cross-polarization band VH (vertical transmit/horizontal receive) and the single co-polarization band VV (vertical transmit/vertical receive). These bands were previously corrected with the local incident angle and using the 30-meter Shuttle Radar Topography Mission (SRTM) dataset as the elevation data. For the Sentinel-2 composite, we firstly masked the daily images with the quality band QA60, available in the product level-1C. Additionally, the clouds were masked with a cloud-score algorithm adapted for Sentinel-2 [16]. Since GEE does all its computations at a given scale, regardless of the spatial resolution of the original image,

we computed all the composite band images at 10 meters, although the Sentinel-2 bands 6, 7, 8A, 11, and 12 have 20 meters of spatial resolution.

2.2.2. Feature Extraction

Feature extraction is a process that aims to improve the accuracy in classification models by generating a set of informative variables (features) from an original dataset. We derived several additional predictive features out of the Sentinel-1 backscatter and Sentinel-2 spectral bands using the median filter and texture analysis based on the Gray-Level Co-Occurrence Matrix (GLCM) [17,18]. The feature extraction resulted in 179 features. Table 1 shows the different convolutional operations, the input bands, and the kernel size used in the convolution.

Table 1. Convolutional operations applied to Sentinel-1 and Sentinel-2 to extract additional predictive features for the classification models. Vertical transmit/horizontal receive (VH) and vertical transmit/vertical receive (VV) are the cross-polarization bands of Sentinel-1. B_i represents the spectral bands of Sentinel-2, where i is the number of the band. The kernel size is the radius in pixels of a squared kernel.

Operation	Abbreviation	Input Bands	Kernel Size (Pixels)	Ref.
Median	smooth	VH, VV, B2, B3, B4, B6, B7, B8, B8A, B11, B12, evi	2, 5, and 10	
GLCM asymmetry	asm			
GLCM contrast	contrast			
GLCM correlation	corr			
GLCM variance	var			
GLCM inv. dif. moment	idm			
GLCM sum average	savg			
GLCM sum variance	svar	VV_smooth_ksize5, VH_smooth_ksize5	10 and 30	[17]
GLCM sum entropy	sent	VV, B4, B7, B8, B11, evi		
GLCM entropy	ent			
GLCM difference variance	dvar			
GLCM difference entropy	dent			
GLCM correlation 1	imcorr1			
GLCM correlation 2	imcorr2			
GLCM max. correlation coef.	maxcorr			
GLCM dissimilarity	diss			
GLCM inertia	inertia	VV_smooth_ksize5, VH_smooth_ksize5	10 and 30	[18]
GLCM cluster shade	shade	VV, B4, B7, B8, B11, evi		
GLCM cluster prominence	prom			

2.2.3. Feature Selection

The 179 extracted features plus the original bands were analyzed with three different feature selection methods in order to rank the most relevant features in the classification model and filter redundant and non-informative data. The feature selection was only applied to the classification setup that distinguishes among the 5 classes; industrial young and industrial mature plantations, smallholder young and smallholder mature plantations, and other land uses. The three feature selection methods included: (1) Sequential feature selection [19] or wrapper method, which is a greedy search algorithm that evaluates the accuracy of a subset of features. We used the forward feature selection, which first evaluates the model accuracy obtained with each single feature and then selects the feature f_1 with the highest accuracy. The kappa coefficient was used as the accuracy metric. Afterwards, the model accuracy of each combination of feature f_1 with the rest of the features is evaluated. The method then selects the combination of features f_1 and f_2 with the highest accuracy. This sequential step is repeated until the feature f_{n-1}. (2) Gini importance [20,21], which is an implicit method in Random Forest classification. We estimated the Gini coefficient with the implementation of Random Forest in the Scikit-learn library. The Gini coefficient is calculated as the total decrease in node impurity averaged over all decision trees of the Random Forest. The impurity measure is estimated with the Gini index [20]. (3). Permutation analysis [21], which is a feature selection technique that assesses the relevance of the features by testing the accuracy decrease when permuting the values of a single

variable in each decision tree. The features are ranked by their accuracy decrease (kappa coefficient) compared to the accuracy without permutation. The *n* features that show the highest accuracy decrease are selected for the training stage of the model.

2.2.4. Random Forest Classification

The Random Forest classification algorithm [20] was selected for its fast computing times in model training and prediction of samples. In this study, a fast implementation is particularly important for the huge amount of data that is processed. Another particular advantage of Random Forest is the low requirement for pre-processing the training data and the ability to predict data when an observation presents missing values. Lastly, Random Forest presents a low tendency for overfitting due to the random procedure during the training of the decision trees [20].

Random Forest is categorized as an ensemble classifier since it involves the training of several decision trees $h = \{h_1(x), \ldots, h_k(x)\}$. A decision tree itself is a classification algorithm, which builds a model with a tree-like structure. The decision tree algorithm first finds the feature f_i that best splits a training set (X, Y) in two groups. The selection of feature f_i is based on a measure of homogeneity in the split groups. The Gini impurity was used as the measure of homogeneity. The algorithm iteratively splits the training set until it reaches a stopping criterion, which is defined by the user with a set of parameters. The parameters include the maximum depth of the tree, the minimum number of training samples to split a node, and the number of features to consider a split [22]. In random forest, given a training set (X, Y), each decision tree h_i is trained with a random sampling of the training data $D = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ in each decision tree. Furthermore, the subset of the training data D considers a random selection of features $F = \{f_1, \ldots, f_j\}$. The final prediction class of a given sample (x_t, y_t) is the mode (majority vote) of the resulting class predictions in the k decision trees.

Other supervised classification models were tested to further justify the choice of Random Forest. We compared the performance of Random Forests (RF), k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Classification and Regression Tree (CART), Naive Bayes (NB), and Minimum Distance (MD) [23]. The comparison was done with a forward feature selection, in which the kappa coefficient was evaluated for the 26 most relevant features of each model. For this analysis, the models were trained without cross-validation using 30% of the samples for training and the remaining 70% for validation. Although, the evaluation of the models was done with the Scikit-learn library, we chose these classification models because these are implemented in GEE, except k-NN, and thus the model comparison may serve for GEE users in future studies.

2.3. Accuracy Assessment and Model Selection

Each classification setup (I, II, and III) was evaluated with the kappa coefficient and the overall accuracy. We also assessed the impact of the four sets of input variables. The models were evaluated with a 10-fold cross-validation, using one fold for training and nine folds for validation. The cross-validation generated a mean accuracy metric and its standard deviation, which gave us guidance on model performance. However, the cross-validation technique can be misleading in model comparison because it overestimates the true positive error. Thus, we used McNemar's test [24] to test pairwise if the performance of the best model in a classification setup was significantly higher than the rest of the models.

2.4. Post-Processing

A post-classification procedure was applied to improve the appearance of the classification setup I, with Sentinel-1 and Sentinel-2 optimal features as input variables of the model. The purpose of the post-processing is only to enhance the visual appeal of the product. Thus, the accuracy assessment and model comparison reported is based on the classification image before post-processing. The post-processing consisted of the following steps:

1. Morphology operations (closing) to remove misclassifications of mature smallholder plantations, particularly inside the mature industrial plantations.
2. Reclassification of smallholder young plantations misclassified as industrial young plantations. We reclassified the disconnected patches with a size lower than 150 pixels.
3. Mode filter to remove salt-and-pepper effects with a squared-kernel of 3 pixels of diameter.
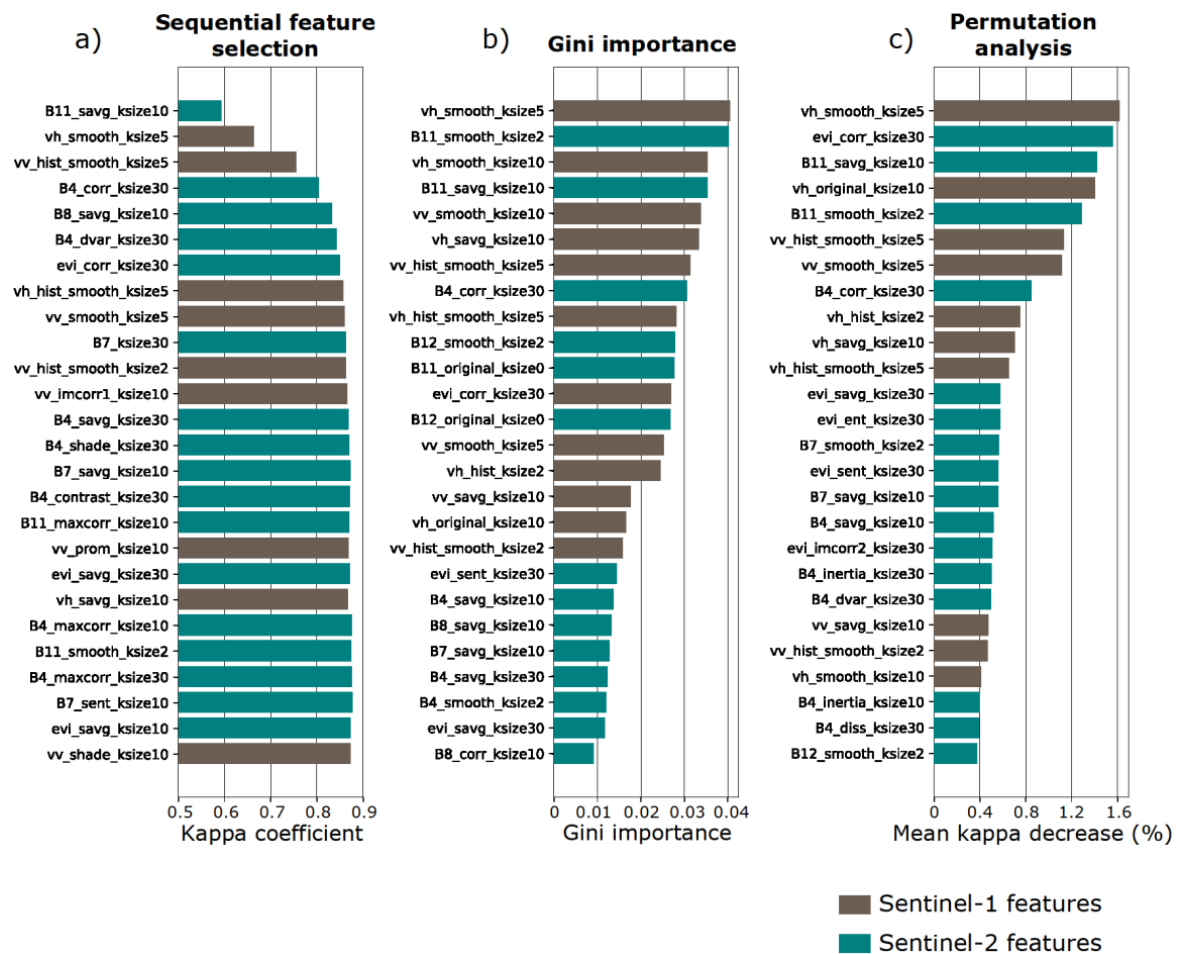
3. Results

The performance of the models for the different classification setups is shown in Figure 5. Random Forest was the model with the best accuracy in the three setups (kappa = 92.8%, 93.1%, and 95.8%), followed by Support Vector Machine in setup I and II (kappa = 89.8% and 89.5%) and k-NN in setup III (kappa = 95.6%). Except Naive Bayes and Minimum Distance in setups I and II, the models show a similar performance with a kappa above 80% in the three setups. The number of features in which the kappa accuracy saturated lies around 15 features in setup I, 10 features in setup II, and 5 features in setup III.
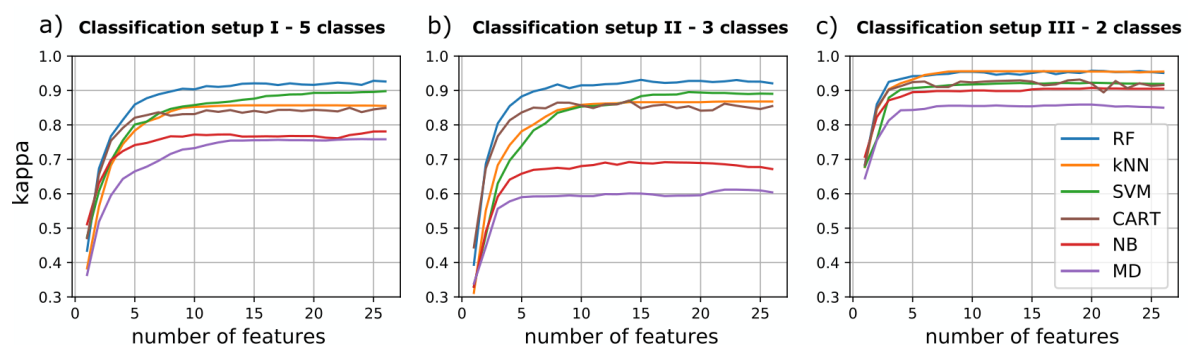
Figure 5. Results of the forward feature selection for six supervised classification models: Random Forest (RF), k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Classification and Regression Tree (CART), Naive Bayes (NB), and Minimum Distance (MD). Each plot is the results of a different classification setup: (**a**) 1. Industrial mature oil palm, 2. Industrial young oil palm, 3. Smallholder mature oil palm, 4. Smallholder young oil palm, and 5. Other land uses, (**b**) 1. Industrial mature oil palm, 2. Smallholder mature oil palm, and 3. Other land uses, and (**c**) Mature oil palm, and 2. Other land uses. The models were trained with 30% of the samples and tested with the remaining 70%.

Figure 6 shows the 26 most relevant features picked by the sequential feature selection, the permutation analysis, and the Gini coefficient. The three methods selected features derived from both Sentinel-1 and Sentinel-2 (Sentinel-1/Sentinel-2 features: 9/17 in sequential feature selection, 12/14 in Gini coefficient, and 10/16 in permutation analysis). The three methods tended to select only extracted features, although the median features Bi_smooth_ksized are highly correlated with the original spectral bands.

Figure 6. Ranking of features that resulted from the three feature selection methods: (**a**) Sequential feature selection, (**b**) Gini coefficient, and (**c**) permutation analysis. Features with a grey bar are Sentinel-1 features, while features with a blue bar are derived from Sentinel-2 bands. The feature selection method that shows the best accuracy is the sequential feature selection, which saturates from the 15th feature with a kappa coefficient of 87.2 ± 2.2%. The accuracies are averaged kappa coefficients, obtained with a 10-fold cross-validation with 10% of the samples for training and the remaining 90% for validation.

The set of features that showed the highest kappa with the lowest number of features was selected. The kappa coefficients obtained with the 26 most relevant features in each method are 87.3 ± 1.7% for the sequential feature selection, 84.5 ± 1.3% for Gini coefficient, and 84.8 ± 2.0% for the permutation analysis. The sequential feature selection shows the highest kappa coefficient, which already saturates when adding more than 15 features. For this reason, we selected the first 15 features of the sequential feature selection due to its high performance (kappa = 87.2 ± 2.2%) with a lower number of features than the other methods.

The accuracy assessment shows the added value of combining Sentinel-1 and Sentinel-2 data for the classification of oil palm plantations. Table 2 shows the overall accuracy and kappa coefficient for the classification models trained with Sentinel-1 features, Sentinel-2 features, the 15 selected features of the sequential feature selection (Figure 6), and with Sentinel-1 and Sentinel-2 original bands. The classification models trained with the selected features show the best performance in all classification setups. The models trained with the selection of optimal features excel particularly for the 3-class (OA = 92.6% and kappa = 88.6%) and the 5-class models (OA = 90.2% and kappa = 87.2%), while the performance of the 2-class models is similar to the results obtained with the other configurations of input variables.

Table 2. Overall accuracy (OA) and kappa coefficient of the three classification setups. Each setup considers a different number of output classes. We trained four models for each classification setup with different sets of input variables: (a) Models trained with the backscatter bands and all derived features of Sentinel-1, (b) models trained with the spectral bands and all derived features of Sentinel-2, (c) models trained with an optimal set of Sentinel-1 and Sentinel-2 features, and (d) models trained only with the Sentinel-1 backscatter bands and Sentinel-2 spectral bands. The accuracies are averaged OA and kappa coefficients, obtained with a 10-fold cross-validation with 10% of the samples for training and the remaining 90% for validation.

Classification Setup	(a) Sentinel-1 (All Bands and Features)		(b) Sentinel-2 (All Bands and Features)		(c) Sentinel-1 & Sentinel-2 (Selected Bands and Features)		(d) Sentinel-1 & Sentinel-2 (All Bands— without Features)	
	OA (%)	Kappa (%)	OA (%)	Kappa (%)	OA (%)	Kappa (%)	OA (%)	Kappa (%)
5 classes [1]	78.2	71.5	79.4	73.1	90.2	87.2	78.2	71.5
3 classes [2]	82.6	73.4	81.6	71.5	92.6	88.6	80.2	69.7
2 classes [3]	93.5	86.9	90.3	80.5	95.3	90.5	94.3	88.6

[1] Classes: 1. Industrial mature oil palm, 2. Industrial young oil palm, 3. Smallholder mature oil palm, 4. Smallholder young oil palm, and 5. Other land uses. [2] Classes: 1. Industrial mature oil palm, 2. Smallholder mature oil palm, and 3. Other land uses. [3] Classes: 1. Mature oil palm and 2. Other land uses.

The confusion matrices of the best models (Figure 7), along the user's (UA) and producer's (PA) accuracies, show a high thematic accuracy for all the classes (UA and PA > 85%) in the three classification setups, apart of class 4 Smallholder young oil palm in setup I, which shows a low producer's accuracy (PA = 64.5%). Although, the partial accuracies are high, the industrial and smallholder mature plantations present higher confusion compared to the other classes in setups I and II. We also observe that the models could distinguish the oil palm classes, regardless of the age or typology, from other land uses; Other land uses presents a UA and PA above 93% in the three setups.
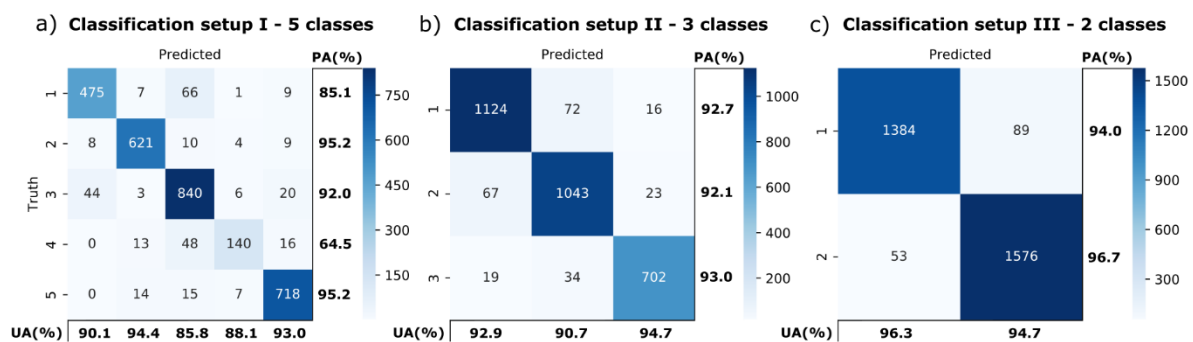
Figure 7. Confusion matrices of the models that were trained with the selection of relevant Sentinel-1 and Sentinel-2 features. Each confusion matrix is the result of a different classification setup, which considers a different number of output classes: (**a**) 1. Industrial mature oil palm, 2. Industrial young oil palm, 3. Smallholder mature oil palm, 4. Smallholder young oil palm, and 5. Other land uses, (**b**) 1. Industrial mature oil palm, 2. Smallholder mature oil palm, and 3. Other land uses, and (**c**) 1. Mature oil palm, and 2. Other land uses.

McNemar's test [24] was applied to assess whether the accuracy of the best model (the model trained with the selected features) is significantly higher than the rest of the models in each classification setup with a significance level lower than 1%. The null hypothesis was rejected for all cases except the classification setup III trained with Sentinel-1 and Sentinel-2 without added features (p-value = 0.747). This means that the proportion of errors in the model trained with all the features of Sentinel-1 and Sentinel-2 is not significantly higher than the model trained without the features when detecting oil palm trees.

The post-classification step improved the appearance of the classification. Based on the visual comparison with Sentinel-1 and Sentinel-2 composites (Figure 8a,b), we confirmed that the step corrected the major issues of the classification. Figure 8 exemplifies the improvement before (Figure 8c) and after (Figure 8d) the post-classification in setup I model trained with the optimal features. Despite these improvements, the accuracy did not increase remarkably after the post-classification, with an improvement of 1.2%.
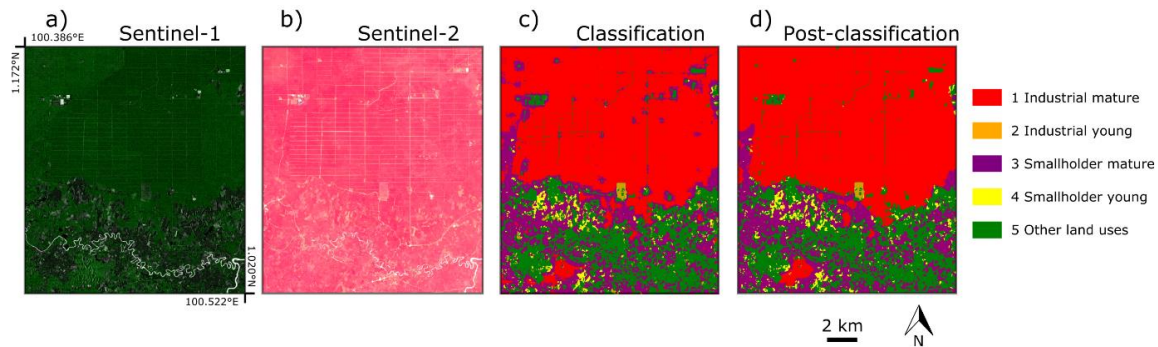
Figure 8. Example of the improvements of the post-classification step. Images (**a**) and (**b**) show the Sentinel-1 and Sentinel-2 composites respectively. Image (**c**) shows the resulting raw classification image, and image (**d**) shows the classification image after the post-classification step.

Figure 9 shows the map of oil palm plantations in Riau province for classification setup I, using the set of optimal features from Sentinel-1 and Sentinel-2. The total amount of oil palm is 31,020 km^2, which represents 36.8% of the land surface of Riau on the Sumatran mainland. Within the total surface of oil palm plantations, 37.8% is industrial mature, 12.3% is industrial young, 42.0% is smallholder mature, and 7.9% is smallholder young. Thus, the ratio of smallholders is 49.9% over all oil palm plantations.
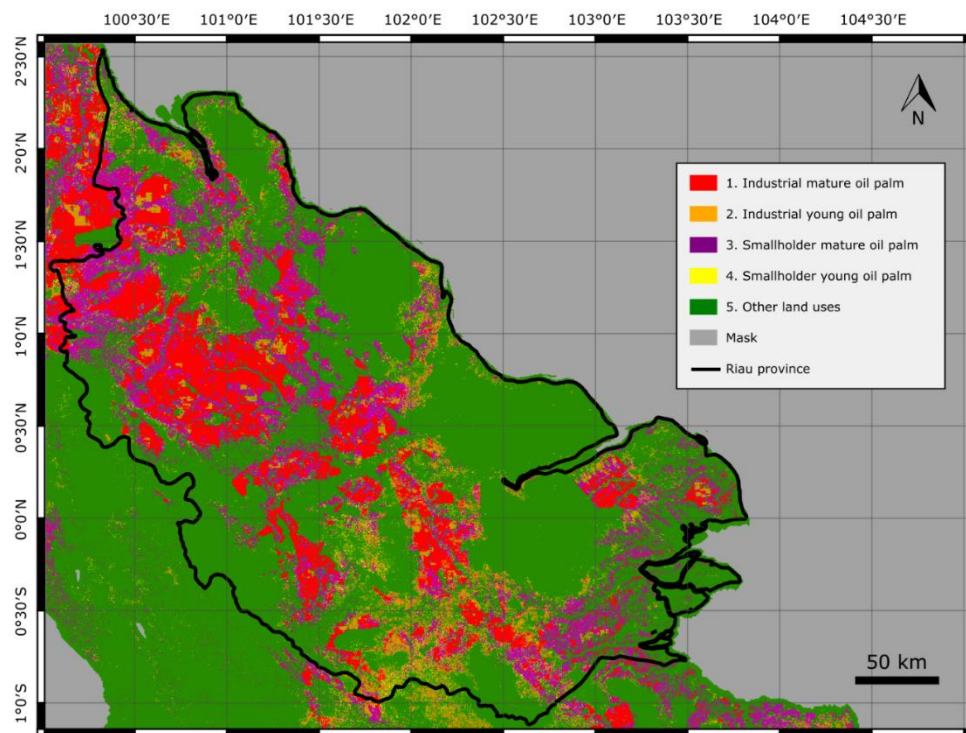
Figure 9. Map of oil palm plantations in Riau province derived from a set of optimal Sentinel-1 and Sentinel-2 features for the second half of 2018.

4. Discussion

The study aimed to show the feasibility of distinguishing smallholder and industrial oil palm plantations with satellite images. The importance of optical and SAR data in the classification of oil palm plantations was analyzed. The results showed that a combined use of Sentinel-1 and Sentinel-2, with a set of optimal additional features, led to the highest accuracy when classifying smallholders and large plantations. The code that generates the results of this study is available at [25]. To our knowledge, this is the first study that aimed to distinguish smallholder and industrial oil palm plantations using state-of-the-art satellite remote sensing data. The fusion of the employed Sentinel-1 and Sentinel-2 bands is also unprecedented in this topic and presents a further step, compared to previous studies that used ALOS PALSAR or only Sentinel-1 scenes.

The high accuracy obtained using only Sentinel-1 (OA = 93.5% and kappa = 86.9%), for oil palm classifications, without distinguishing between smallholders and industrial plantations, are comparable to the user's accuracy of 95.6% obtained in a recent study [26] for an oil palm class. The results of the Riau case study confirmed the usefulness of SAR data for mature oil palm mapping [8]. The detection of oil palms, without distinction of typology, can still be improved with optical imagery (Sentinel-2). The results also show that feature extraction is not necessary when detecting mature oil palm trees. This is particularly relevant for further studies that aim to detect roughly the oil palm plantations at a regional scale, without distinction of typology and computationally expensive algorithms.

The characteristic canopy of oil palm plantations might explain the high relevance of Sentinel-1 in the models. The shape of palm-like trees produces a characteristically high backscatter response in the dual-band VH. Evidence of the importance of Sentinel-1 is the high relevance of the VH band and its features in the three feature selection methods. Despite good results for Sentinel-1 in mature oil palm mapping, Sentinel-1 solely cannot distinguish smallholders and industrial plantations. A distinction between smallholders and industrial plantations necessarily requires additional features derived from Sentinel-1 and Sentinel-2 images, which are effective in capturing the shape and density of the harvesting trail networks in industrial oil palm plantations.

The results do not corroborate the study of Oon and colleagues [11], which concluded that the sole use of Sentinel-1 dual bands can distinguish smallholder and industrial plantations in peatlands. The findings of the Riau case study suggest that Sentinel-1 bands show a low accuracy (OA = 82.6% and kappa = 73.4%) for such classification problems. One possible explanation for this is that the conclusions by Oon et al. are based on the results obtained from an insufficient number of training and testing points (98 points) that may overfit to its small study area (<50 km^2). In contrast, the present study uses a larger training dataset (3,448 points) that covers an extensive heterogeneous area across Riau province. The training dataset used in the present study serves as evidence that the distinction between smallholders and industrial plantations is a challenging problem that requires additional features, based on textural analysis, that capture distinctive contextual information of the oil palm plantations.

The study also highlights the usefulness of cloud computing for regional and global environmental studies that make use of large satellite datasets and require high processing speed. In our study, the algorithm uses 217 daily Sentinel-1 images and 827 daily Sentinel-2 images. The processing of the 10-meter resolution oil palm map of Riau province, which covers an area of 84,360 km^2, was processed in the cloud platform and took about 11 hours. This processing includes the image compositing of Sentinel-1 and Sentinel-2, the extraction of the selected features, the training of the Random Forest, and the image classification. The code and algorithms written in Google Earth Engine can be easily shared and run among different users. For instance, a demo code is available for oil palm mapping [27] and for the visualization of the results of setup I [28] in GEE. The shareability of code can be useful, not only for code development, but also for obtaining quick results for environmental and land use monitoring. For instance, in our case study, the classification in Riau province revealed an unexpectedly high ratio of smallholders, which represent the 49.9% of all palm oil plantations in Riau province in 2018.

The model comparison emphasized the usefulness of Random Forests for fast modeling and image classification, even in cloud-based platforms, such as GEE. Random Forest delivered higher accuracy numbers, compared to other supervised classification models, that are implemented in GEE. Although, SVM and other kernel-based classification methods may excel at drawing the decision boundary between classes, their use requires expert knowledge of data pre-processing and hyperparameter tuning in the training stage. Instead, RF is a fast and easy-to-use classification model that requires less parameterization to control overfitting and, thus, can be used more broadly by the scientific community or industry.

The algorithm we used showed good performance, particularly for distinguishing smallholders and industrial plantations in Riau province. However, the level of performance might differ in other parts of the world, particularly for such classification setup, that considers the typology and age of the plantations. The industrial oil palm plantation in other regions might present a different construction design and trail network that would lead to a different set of relevant features for the classification model. Moreover, other palm-like plantations, such as sago plantations in Papua, might get misclassified as oil palm. The similar canopy of oil palm and other palm-like species entails a similar backscatter response, which in turn leads to a high confusion between oil palm and other palm-like plantations. Therefore, our recommendation is that the classification model should be trained again with a different sample dataset collected around the new study area. The last shortcoming we observed is the high confusion between young palm trees and bare soil, which is also reported in [29]. The classification of young (<3 year-old) plantations is very challenging due to the low canopy coverage of young palm trees. Young plantations seem to be unplanted from space even with 10-meter optical imagery. Recent use of Convolutional Neural Networks has proven its usefulness in similar remote sensing studies [30,31], and the more extensive use of contextual information in deep learning may lead to a more accurate classification of young and mature oil palm plantations. Future studies should address these limitations and opportunities to accurately map oil palm plantations at global scale.

5. Conclusions

The present study aimed to show the suitability of optical and SAR satellite data to classify oil palm plantations. The results showed that the combined use of Sentinel-1 co-polarization bands and Sentinel-2 spectral bands allows the detection of oil palm stands. The detection of mature oil palm stands is possible thanks to the characteristic backscatter response of the palm canopy. However, the Sentinel-1 and Sentinel-2 raw images are not enough for a classification problem that aims to distinguish between smallholders and industrial oil palm plantations, or a distinction between young and mature plantations. Such classification setup requires a set of additional relevant features, such as texture and convolutional operations, that capture the spatial patterns that are characteristic in industrial plantations. The most significant pattern in industrial plantations is the dense trail network, which tends to be non-existent in smallholder plantations.

The study was carried out with Google Earth Engine, a cloud-based platform that allows the rapid classification of Sentinel-1 and Sentinel-2 data for a given study area. The shareability of our algorithm and the possibility to run a trained classification model everywhere in the world makes GEE (or other cloud-based processing systems) a suitable tool for environmental monitoring. Researchers and environmentalists can easily reproduce the classification model for mapping oil palm trees over large regions and detect new plantations in sensitive and protected areas. It is our aim that this study will eventually lead to an automated and standardized global map of oil palm and its various categories that can be used by the various parties involved in palm oil, such as the Round Table for Sustainable Palm Oil (RSPO), Governments, NGOs, companies, and other stakeholders.

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

References

1. Susanti, A. *Oil Palm Expansion in Indonesia: Serving People, Planet and Profit?* Eburon Academic Publishers: Utrecht, The Netherlands, 2016.

2. Austin, K.G.; Schwantes, A.; Gu, Y.; Kasibhatla, P.S. What causes deforestation in Indonesia? *Environ. Res. Lett.* **2019**, *14*, 024007. [CrossRef]

3. Margono, B.A.; Potapov, P.V.; Turubanova, S.; Stolle, F.; Hansen, M.C. Primary forest cover loss in Indonesia over 2000–2012. *Nat. Clim. Chang.* **2014**, *4*, 730. [CrossRef]

4. Abood, S.A.; Lee, J.S.H.; Burivalova, Z.; Garcia-Ulloa, J.; Koh, L.P. Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. *Conserv. Lett.* **2015**, *8*, 58–67. [CrossRef]

5. Meijaard, E.; Garcia-Ulloa, J.; Sheil, D.; Wich, S.; Carlson, K.; Juffe-Bignoli, D.; Brooks, T. *Oil Palm and Biodiversity: A Situation Analysis by the IUCN Oil Palm Task Force*; International Union for Conservation of Nature and Natural Resources (IUCN): Gland, Switzerland, 2018.

6. Mosnier, A.; Boere, E.; Reumann, A.; Yowargana, P.; Pirker, J.; Havlík, P.; Pacheco, P. *Palm Oil and Likely Futures: Assessing the Potential Impacts of Zero Deforestation Commitments and a Moratorium on Large-Scale Oil Palm Plantations in Indonesia*; CIFOR: Bogor, Indonesia, 2017; Volume 177.

7. Austin, K.; Mosnier, A.; Pirker, J.; McCallum, I.; Fritz, S.; Kasibhatla, P. Shifting patterns of oil palm driven deforestation in Indonesia and implications for zero-deforestation commitments. *Land Use Policy* **2017**, *69*, 41–48. [CrossRef]

8. Miettinen, J.; Liew, S.C. Separability of insular Southeast Asian woody plantation species in the 50 m resolution ALOS PALSAR mosaic product. *Remote Sens. Lett.* **2011**, *2*, 299–307. [CrossRef]

9. Lee, J.S.H.; Wich, S.; Widayati, A.; Koh, L.P. Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. *Remote Sens. Appl. Soc. Environ.* **2016**, *4*, 219–224. [CrossRef]

10. Nomura, K.; Mitchard, E.T.; Patenaude, G.; Bastide, J.; Oswald, P.; Nwe, T. Oil palm concessions in southern Myanmar consist mostly of unconverted forest. *Sci. Rep.* **2019**, *9*, 1–9. [CrossRef]

11. Oon, A.; Ngo, K.D.; Azhar, R.; Ashton-Butt, A.; Lechner, A.M.; Azhar, B. Assessment of ALOS-2 PALSAR-2L-band and Sentinel-1 C-band SAR backscatter for discriminating between large-scale oil palm plantations and smallholdings on tropical peatlands. *Remote Sens. Appl. Soc. Environ.* **2019**, *13*, 183–190. [CrossRef]

12. Koh, L.P.; Miettinen, J.; Liew, S.C.; Ghazoul, J. Remotely sensed evidence of tropical peatland conversion to oil palm. *PNAS* **2011**, *108*, 5127–5132. [CrossRef]

13. Szantoi, Z.; Smith, S.E.; Strona, G.; Koh, L.P.; Wich, S.A. Mapping orangutan habitat and agricultural areas using Landsat OLI imagery augmented with unmanned aircraft system aerial photography. *Int. J. Remote Sens.* **2017**, *38*, 2231–2245. [CrossRef]

14. 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]

15. Meijaard, E.; Sheil, D. The Moral Minefield of Ethical Oil Palm and Sustainable Development. *Front. For. Glob. Chang.* **2019**, *2*. [CrossRef]

16. Chastain, R.; Housman, I.; Goldstein, J.; Finco, M. Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States. *Remote Sens. Environ.* **2019**, *221*, 274–285. [CrossRef]

17. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, 610–621. [CrossRef]

18. Conners, R.W.; Trivedi, M.M.; Harlow, C.A. Segmentation of a high-resolution urban scene using texture operators. *Comput. Vision Graph. Image Process.* **1984**, *25*, 273–310. [CrossRef]

19. Langley, P.; Sage, S. Induction of selective Bayesian classifiers. In *Uncertainty Proceedings 1994*; Elsevier: Amsterdam, The Netherlands, 1994; pp. 399–406.

20. Breiman, L. Some properties of splitting criteria. *Mach. Learn.* **1996**, *24*, 41–47. [CrossRef]

21. Louppe, G.; Wehenkel, L.; Sutera, A.; Geurts, P. Understanding variable importances in forests of randomized trees. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–10 December 2013; pp. 431–439.

22. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. *Classification and Regression Trees*; Chapman and Hall: Wadsworth, NY, USA, 1984.

23. Hastie, T.; Tibshirani, R.; Friedman, J.; Franklin, J. The elements of statistical learning: Data mining, inference and prediction. *Math. Intell.* **2005**, *27*, 83–85.

24. Dietterich, T.G. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput.* **1998**, *10*, 1895–1923. [CrossRef]

25. GitHub Repository. Available online: https://github.com/adriadescals/oil_palm_riau (accessed on 25 October 2019).

26. Miettinen, J.; Shi, C.; Liew, S.C. Land cover distribution in the peatlands of Peninsular Malaysia, Sumatra and Borneo in 2015 with changes since 1990. *Glob. Ecol. Conserv.* **2016**, *6*, 67–78. [CrossRef]

27. Google Earth Engine Code. Available online: https://code.earthengine.google.com/70a61b2ea4d2e93562397f1acbe6b337 (accessed on 25 October 2019).

28. Google Earth Engine Code. Available online: https://code.earthengine.google.com/9be26b9317e8d72eea6f7bcb42f2dc19 (accessed on 25 October 2019).

29. Miettinen, J.; Gaveau, D.L.; Liew, S.C. Comparison of visual and automated oil palm mapping in Borneo. *Int. J. Remote Sens.* **2019**, *40*, 8174–8185. [CrossRef]

30. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [CrossRef]

31. Padarian, J.; Minasny, B.; McBratney, A.B. Using deep learning for digital soil mapping. *Soil* **2019**, *5*, 79–89. [CrossRef]